THE UNIVERSITY OF CHICAGO

FULLY NONLINEAR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

BY
BENJAMIN SEEGER

CHICAGO, ILLINOIS
JUNE 2019

To my family

TABLE OF CONTENTS

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Takis Souganidis. He has been an extremely dedicated mentor who heavily invested in my development as a mathematician. I have learned a great deal from him about what it means to conduct meaningful research and how to communicate it to others. It is safe to say that I would not be the mathematician, or person, I am today without his guidance, advice, and care.

I would also like to thank Pierre-Louis Lions for many useful conversations throughout my Ph.D. studies. I have gained many insights from his visits and lectures at UChicago, and I hope to continue to do so in the future as I work with him in the coming years.

More thanks are due to the other analysis faculty at UChicago, particularly Carlos Kenig, Wilhelm Schlag, Luis Silvestre, and Charlie Smart. The strong variety of analysis was one of the main reasons I came to UChicago, and indeed, I have been gifted with a well-rounded education in modern analysis from the many classes, seminars, and working groups organized by these professors, as well as much advice and support from each of them.

My early days at UChicago would have been extremely difficult without my friendships with the other first year students, which were formed as we struggled and helped each other through the shared workload of the first-year program. Later, my older academic siblings, Ben Fehrman, Jessica Lin, and Olga Turanova, showed me kindness and encouragement, and they continue to be sources of support and fruitful discussion. The other student of Takis from my year, Lisa Kava, became one of my closest friends at UChicago as we helped each other navigate a long year of learning new and difficult mathematics.

Finally, to my family - my parents Andreas and Susanne, my sister Lena, my brother David, and my wife Ami - you are my most important role models. Ami, you are my biggest fan, and I am yours. Thank you for all the love, respect, and encouragement.

ABSTRACT

This thesis is concerned with the theory and applications of certain fully nonlinear stochastic partial differential equations. First, we present several new results regarding the well-posedness of the equations. Among these are proofs of the comparison principle for equations with nontrivial spatial dependence. We also prove some new path-stability estimates, and we give a very general proof of existence using Perron's method, which characterizes the unique solution as the maximal sub-solution.

We also discuss a general framework for approximating solutions numerically. A variety of convergent approximation schemes are considered, including finite difference schemes and Trotter-Kato splitting formulas, and the results are general enough to allow for many more examples. For first-order equations, we derive explicit error estimates.

Finally, we introduce a family of homogenization problems that arise from scaling limits of fully nonlinear equations with highly oscillatory spatio-temporal dependence. We prove, under suitable assumptions on the nonlinearities and the random dependence, that the limiting behavior is governed by a spatially homogenous, stochastic Hamilton-Jacobi equation.

NOTATION

Let $N \geq 1$ and $U \subset \mathbb{R}^N$ be open.

- $C_b^k(U)$ is the space of k-times differentiable functions with bounded derivatives. The bounds of the derivatives of $f \in C_b^k(U)$ are denoted $\left\|D^j f\right\|_{\infty,U}$ for $j = 1, 2, \ldots, k$. If $U = [0,T] \subset \mathbb{R}$, we also write $\left\|f^{(j)}\right\|_{\infty,[0,T]} = \left\|f^{(j)}\right\|_{\infty,T}$.

- $C^{0,\alpha}(U)$ is the space of α-Hölder continuous functions. The semi-norm of $f \in C^{0,\alpha}(U)$ is denoted by $[f]_{\alpha,U}$. If $U = [0,T] \subset \mathbb{R}$, we also write $[f]_{\alpha,[0,T]} = [f]_{\alpha,T}$. If $\alpha = 1$, we also write $[f]_1 = \|Df\|_\infty = \mathrm{Lip}(f)$ and $C^{0,1}(U) = \mathrm{Lip}(U)$.

- $LSC(U)$ = the space of lower-semicontinuous functions on U

- $USC(U)$ = the space of upper-semicontinuous functions on U

- $(B)UC(U)$ = the space of (bounded) uniformly continuous functions on U

- For a locally finite function $u : U \to \mathbb{R}$, u^* and u_* denote the upper- and lower-semicontinuous envelopes of u.

- $\mathbb{S}^N$ denotes the space of symmetric N-by-N matrices. I denotes the identity matrix. For $A \in \mathbb{S}^N$, $\|A\| := \max_{|v|=1} Av \cdot v$. For $B \in \mathbb{R}^N \times \mathbb{R}^N$, $\mathrm{Sym}(B) \in \mathbb{S}^N$ is defined by $(BB^t + B^t B)/2$.

- For $W \in C([0,\infty))$ and $0 \leq s \leq t < \infty$, $\mathrm{osc}(W,s,t) := \max_{r_1,r_2 \in [s,t]} |W(r_1) - W(r_2)|$.

- $B_r(x) := \{y \in \mathbb{R}^N : |x-y| < r\}$, $B_r := B_r(0)$, $N_{r,s}(x,t) := \{(y,q) : |x-y| < r,\ |q-t| < s\}$, $N_r(x,t) := N_{r,r}(x,t)$, $S^{N-1} := \partial B_1$.

- For $a, b \in \mathbb{R}$, $a \wedge b = \min\{a,b\}$ and $a \vee b = \max\{a,b\}$.

- For a function $H : \mathbb{R}^N \to \mathbb{R}$, $H^*(q) := \sup_{p \in \mathbb{R}^N} (p \cdot q - H(p))$.

CHAPTER 1
INTRODUCTION

1.1 Background

This thesis is concerned with recent developments in the theory and applications of fully nonlinear rough, or stochastic, partial differential equations.

Many complex phenomena in the real world are modeled with partial differential equations that, due to statistical uncertainty, have random, possibly singular, spatio-temporal dependence. This dependence can be mathematically described using white noise, which leads to stochastic partial differential equations. In recent years, the study of such equations has attracted much attention in view of the many applications to physics, economics, and beyond, as well as the interesting mathematics that arise from the interplay between probability and analysis.

As a consequence of the rough nature of the randomness, the standard methods for defining solutions of partial differential equations are often not sufficient in the stochastic setting. It is therefore necessary to come up with a new, well-posed notion of weak solutions. One can then investigate the effect that the stochasticity has on the structure and behavior of the solutions. This question can be studied in many contexts, including, but not limited to, regularity, long-time behavior, homogenization effects, scaling limits, numerical approximations and simulations, and large deviations.

In this thesis, we are interested in fully nonlinear and degenerately parabolic equations. For some unknown function $u : \mathbb{R}^d \times [0,\infty) \to \mathbb{R}$, these take the form

$$du = F(D^2u, Du, x, t)\, dt + \sum_{i=1}^{m} H^i(Du, x) \cdot dW^i \quad \text{in } \mathbb{R}^d \times (0,\infty). \tag{1.1.1}$$

The initial data $u(\cdot, 0) := u_0$, as well as the functions $F : \mathbb{S}^d \times \mathbb{R}^d \to \mathbb{R}$ and $H = (H^1, \dots, H^m) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^m$ are continuous, although more regularity is required for the latter in order to make sense of the equation. The nonlinearity F is degenerate elliptic, which means that it is increasing in the $\mathbb{S}^d$ variable, where $\mathbb{S}^d$ is the space of symmetric d-by-d matrices. The terms dW^i denote the time-differential of the continuous path $W = (W^1, W^2, \ldots, W^m)$. Of interest to us are random paths that are not regular enough to put (1.1.1) in the classical viscosity solution framework, for example, if W is an m-dimensional Brownian motion, or a more general semimartingale. As we discuss later on, (1.1.1) will then be understood, at least formally, through the Stratonovich formulation.

The study of parabolic stochastic partial differential equations goes back to the work of Krylov and Rozovskiĭ [44] in the quasilinear, uniformly parabolic setting (see also Pardoux and Peng [69]), which is based on using the martingale approach to solve stochastic differential equations in an appropriate Banach space. On the other hand, a global theory for weak solutions of fully-nonlinear, degenerately parabolic equations like (1.1.1) was not available until the works of Lions and Souganidis [53, 54, 55, 56, 76] that introduced pathwise (stochastic, rough) viscosity solutions.

Degenerate parabolic stochastic partial differential equations arise in a number of settings. One example is the theory of stochastic control with partial information, or, more generally, pathwise stochastic control. As is well known, (see, for instance, the books of Bardi and Capuzzo-Dolcetta [8] or Fleming and Soner [30]), the so-called value function from stochastic control is the solution of a fully nonlinear, Hamilton-Jacobi-Bellman equation. However, in applications of finance, only partial information is known for the stochastic dynamical system being controlled. This uncertainty comes up in the partial differential equation as rough, stochastic dependence, leading to an equation like (1.1.1) where the Hamiltonians H^i are linear in the gradient variable. This setting, going back to Pardoux [67, 68], has also been studied extensively from the point of view of rough path theory by many authors, including, but not limited to, Caruana, Friz, and Oberhauser [21] and Buckdahn and Ma [17, 18].

An example for which H is fully nonlinear arises in the study of moving interfaces, a subject whose mathematical analysis is relevant to models for phase transitions, population dynamics, turbulent combustion, and more. A moving interface can be described by a family of hypersurfaces $\{\Gamma_t\}_{t\geq 0} \subset \mathbb{R}^d$ evolving according to a given normal velocity $V(Dn, n, x, t)$, where $n = n(x,t)$ is the outward normal vector to the surface Γ_t at the point x, and Dn denotes its spatial derivative, which encodes the principle curvatures of the surface. A well-studied example is

$$V(Dn, n, x, t) := -\operatorname{tr}(Dn) + a(n, x, t), \tag{1.1.2}$$

where $\kappa := \operatorname{tr}(Dn)$ is the mean curvature of the surface and $a : S^{d-1} \times \mathbb{R}^d \times [0, \infty)$ is a first-order perturbation.

An interface evolving in this manner is known to develop singularities or change topological type in finite time, in which case the normal direction and curvature cease to be well-defined. In many applications, however, it is of interest to describe the motion of the interface after singularities occur. A great number of strategies for doing so have been developed. One notable example in the case of mean-curvature flow (that is, when $a \equiv 0$ in (1.1.4)) is Brakke's theory of varifolds [16]. For very general interface motions, the most successful tool is the level-set approach introduced by Osher and Sethian [64] and generalized by many authors, for example Barles and Souganidis [13]. For each $t \geq 0$, the hypersurface Γ_t is identified with the level-set of a function u:

$$\Gamma_t := \{x \in \mathbb{R}^d : u(x,t) = 0\}. \tag{1.1.3}$$

One can then formally derive a nonlinear, degenerately parabolic partial differential equation to describe the evolution of the function u. In the case of (1.1.2), for example, this takes the form

$$u_t = \Delta u - D^2 u \frac{Du}{|Du|} \cdot \frac{Du}{|Du|} + a\left(\frac{Du}{|Du|}, x, t\right)|Du| \quad \text{in } \mathbb{R}^d \times (0, \infty). \tag{1.1.4}$$

The theory of viscosity solutions yields the existence of a unique, continuous, weak solution of (1.1.4), as long as a is continuous in all variables. As a consequence, the singular behavior of moving interfaces can be understood through an analysis of the equation.

In many physical applications, the perturbation a in (1.1.2) is a singular, space-time noise term, an example being

$$a(n,x,t) := \sum_{i=1}^{m} a^i(n,x)dB^i(t) \tag{1.1.5}$$

for some coefficients $a^i : S^{d-1} \times \mathbb{R}^d \to \mathbb{R}$ and independent Brownian motions $\{B^i\}_{i=1}^m$. In this case, the equation (1.1.4) is a special case of (1.1.1):

$$du = \left(\Delta u - D^2 u \frac{Du}{|Du|} \cdot \frac{Du}{|Du|} \right) dt + \sum_{i=1}^{m} a^i \left(\frac{Du}{|Du|}, x \right) |Du| \circ dB^i \quad \text{in } \mathbb{R}^d \times (0,\infty). \tag{1.1.6}$$

1.2 The main difficulties

We briefly summarize some of the difficulties in generalizing the viscosity solution theory to solutions of (1.1.1) with rough time dependence. A more detailed discussion can be found in [53, 76].

The first important difficulty in the study of equations such as (1.1.1) is that global smooth solutions do not exist in general. This is due to the nonlinearity of the equation, which implies that shocks in the solutions form in finite time. In the non-rough setting, that is, when the path W is smooth, this was overcome with the Crandall-Lions notion of viscosity solutions, a well-developed theory for which powerful existence, uniqueness, and stability results have been obtained. For a detailed description of the theory and a list of references, see the User's Guide of Crandall, Ishii, and Lions [25].

Extending the theory of [44] to (1.1.1) is not possible in general. For example, establishing the existence of solutions, which is usually done by proving tightness of probability laws of suitable approximations in the right function space, is made difficult by the full nonlinearity and degenerate parabolicity. A related problem arises when trying to characterize the law of (1.1.1) as the viscosity solution of a partial differential equation in infinite dimensions, since the resulting equation does not appear to be amenable to the existing theory of Crandall and Lions [26, 27].

Another major difficulty is that the standard viscosity solution theory breaks down for (1.1.1), since it is not possible to make sense of the equation in a pointwise sense. If W is smooth, then a sub-solution of (1.1.1) is a function u such that, whenever $\phi \in C^2(\mathbb{R}^d \times (0,\infty))$ and $u(x,t) - \phi(x,t)$ achieves a local maximum at (x_0, t_0), then

$$\phi_t(x_0,t_0) \le F(D^2\phi(x_0,t_0), D\phi(x_0,t_0), x_0, t_0) + \sum_{i=1}^m H^i(D\phi(x_0,t_0), x_0)\dot{W}^i(t_0),$$

with a similar definition for super-solutions. However, this inequality makes no sense if $\dot{W}$ is not defined pointwise, as is the case if W is a Brownian motion and, therefore, nowhere differentiable.

In view of the evolution structure of the equation, the definition may be rephrased in an integral sense. This has been done for equations with integrable time-dependence; see the works of Ishii [36] and Lions and Perthame [49] when $F \equiv 0$, or of Nunziante [62, 63] for second-order equations. In particular, if $\phi \in C^2(\mathbb{R}^d \times [0,\infty))$, then, for all $s < t$,

$$\begin{aligned}\sup_{x\in\mathbb{R}^d} \{u(x,t) - \phi(x,t)\} &\le \sup_{x\in\mathbb{R}^d} \{u(x,s) - \phi(x)\} \\ &+ \int_s^t \sup_{x \in A(r)} \left\{ -\phi_r(x,r) + F(D^2\phi(x), D\phi(x), x, r)dr + \sum_{i=1}^m H^i(D\phi(x),x)dW^i(r) \right\},\end{aligned}$$

where $A(r)$ is the set of points for which $u(\cdot,r) - \phi(\cdot,r)$ attains a maximum. This strategy allows the viscosity solution theory for (1.1.1) to be applied to the case of paths with finite total variation, but still does not cover the most general, rough case.

It turns out that it is possible to find, on sufficiently small time intervals $I \subset (0,\infty)$, smooth-in-space solutions Φ of the rough, Hamilton-Jacobi part of (1.1.1)

$$d\Phi = \sum_{i=1}^m H^i(D\Phi, x) \cdot dW^i \quad \text{in } \mathbb{R}^d \times I,$$

or, equivalently, for all $s, t \in I$,

$$\Phi(x,t) - \Phi(x,s) = \sum_{i=1}^m \int_s^t H^i(D\Phi(x,r), x) \cdot dW_r^i.$$

This was originally accomplished by Kunita [45] with the method of characteristics, which here become a system of stochastic differential equations. The same can be accomplished if W is a geometric rough path using the flow properties of rough differential equations proved by Lyons and Qian [59].

The key idea behind the pathwise viscosity theory is to use these local, spatially smooth solutions as test functions to absorb the "rough" part of (1.1.1). That is, the definition involves a local version of the so-called method of flow transformations, a standard tool by which a stochastic partial differential equation is converted into a classical one with random coefficients. The trade-off is that, being exact solutions, the test functions are very inflexible. Hence, it is not obvious how to use them as barriers, and, as a result, it is necessary to essentially "reinvent" the arguments and results from the standard theory of viscosity solutions.

1.3 A summary of the results

We briefly describe the results outlined in this thesis. For more detailed explanations, see the introductory material for each respective chapter. The results are also contained in the author's works [73, 74, 75].

In Chapter 2, we recall the theory of stochastic and rough flows, present the Lions-Souganidis definition of pathwise viscosity solutions, and prove some important properties.

Many specialized test functions are also constructed that will be used in the various quantitative arguments throughout the thesis.

The comparison principle is the focus of Chapter 3. Its proof is given in many settings for which the Hamiltonians have nontrivial spatial dependence, which, in view of the rough time-dependence, creates a number of difficulties.

We study the existence of solutions in Chapter 4. First, we prove precise path stability estimates for equations with convex, spatially-dependent Hamiltonians, which extends the class of first-order equations for which existence is known. We then give a very general proof of existence of solutions for second-order equations using Perron's method, which characterizes the unique solution as the maximal sub-solution.

We present a general framework for approximating solutions of (1.1.1) numerically in Chapter 5. A variety of convergent approximation schemes are considered, including finite difference schemes and Trotter-Kato splitting formulas, and the results are general enough to allow for many more examples. For first-order equations, we derive explicit error estimates.

Finally, in Chapter 6, we introduce a family of homogenization problems that arise from scaling limits of fully nonlinear equations with highly oscillatory spatio-temporal dependence. We prove, under suitable assumptions on the nonlinearities and the random dependence, that the limiting behavior is governed by a spatially homogenous, stochastic Hamilton-Jacobi equation.

CHAPTER 2

DEFINITIONS AND PRELIMINARY RESULTS

2.1 Introduction

In this chapter, we introduce the definition of pathwise viscosity solutions of (1.1.1). Central to this discussion is the ability to obtain local-in-time, smooth-in-space solutions of the "rough," Hamilton-Jacobi part of the equation. We first explain which assumptions on

$$\begin{cases} H = (H^1, H^2, \dots, H^m) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^m, \\ \phi : \mathbb{R}^d \to \mathbb{R}, \quad \text{and} \\ W = (W^1, W^2, \dots, W^m) : [0, \infty) \to \mathbb{R}^m \end{cases}$$

guarantee the existence, for a sufficiently small $h > 0$, of a smooth-in-space solution of

$$\begin{cases} d\Phi = \sum_{i=1}^{m} H^i(D\Phi(x,t), x) \cdot dW^i & \text{in } \mathbb{R}^d \times (t_0 - h, t_0 + h) \\ \Phi(\cdot, t_0) = \phi & \text{in } \mathbb{R}^d. \end{cases} \tag{2.1.1}$$

Once we have done so, we give some equivalent definitions of pathwise viscosity sub- and super-solutions, and prove some of their properties. Finally, we construct particular solutions of (2.1.1) (or rather, of certain variants that arise in the method of doubling variables) that will be used in comparison and stability proofs in various settings.

2.2 The method of characteristics

The main approach to constructing solutions of (2.1.1) is through the study of the system of characteristic equations given, for fixed $x, p \in \mathbb{R}^d$ and $t_0 \in (0, \infty)$, by

$$\begin{cases} dX = -\sum_{i=1}^m D_p H^i(P, X) \cdot dW^i, & X(x,p,t_0) = x, \\ dP = \sum_{i=1}^m D_x H^i(P, X) \cdot dW^i, & P(x,p,t_0) = p. \end{cases} \tag{2.2.1}$$

If (2.2.1) has a global solution, then, assuming these objects are well-defined, we will set

$$\begin{cases} \mathbf{X}(x,t) = X(x, D\phi(x), t), \quad \mathbf{P}(x,t) = P(x, D\phi(x), t), \quad \text{and} \\ \mathbf{Z}(x,t) = \phi(x) + \sum_{i=1}^m \int_{t_0}^t \Big(H^i(\mathbf{P}, \mathbf{X}) - \mathbf{P} \cdot D_p H^i(\mathbf{P}, \mathbf{X}) \Big) \cdot dW^i. \end{cases}$$

If H^i and ϕ are regular enough, then $x \mapsto \mathbf{X}(x,t)$ will be a diffeomorphism for t sufficiently close to t_0. Then, formally, the solution Φ of (2.1.1) is given by

$$\Phi(x,t) := \mathbf{Z}(\mathbf{X}^{-1}(x,t), t). \tag{2.2.2}$$

We make these calculations rigorous in what follows.

2.2.1 *Stochastic and rough flows*

We first discuss the case where $W : [0, \infty) \to \mathbb{R}^m$ is stochastic process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For simplicity of presentation, we focus here only on the case where W is a Brownian motion, although the discussion follows similarly if W is a semi-martingale. Without loss of generality, we may assume that $\Omega := C([0,\infty), \mathbb{R}^m)$ with the topology of local uniform convergence, $\mathcal{F}$ is the Borel σ-algebra on Ω, and $\mathbb{P}$ is the Wiener measure. With this viewpoint, the statements and results below are understood to hold with

$\mathbb{P}$-probability one, that is, for $\mathbb{P}$-almost every path $W \in \Omega$.

In this setting, (2.1.1) and (2.2.1) are interpreted with the Stratonovich formulation, and the symbol "$\cdot$" is replaced with a "$\circ$." In particular, the maps X and P solve the equations

$$\begin{cases} dX = -\sum_{i=1}^{m} D_p H^i(P, X) \circ dW^i, & X(x, p, t_0) = x, \\ dP = \sum_{i=1}^{m} D_x H^i(P, X) \circ dW^i, & P(x, p, t_0) = p. \end{cases} \tag{2.2.3}$$

This has to do with the way that the standard Itô stochastic calculus is applied in the construction of smooth flows below, as well as the stability of these systems with respect to a regularization of the path.

The following lemma is a special case from a large family of results about stochastic flows of diffeomorphisms studied by Kunita [45].

Lemma 2.2.1. *Assume that, for some $n \geq 0$, $H \in C_b^{4+n}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$. Then, with $\mathbb{P}$-probability one, $(x, p) \mapsto (X(x, p, t), P(x, p, t))$ is a C^{1+n}-diffeomorphism for all $t \in \mathbb{R}$.*

The theory of rough paths, which was introduced by Lyons [60] and subsequently developed by many authors, including Gubinelli [34] and Friz and Victoir [33] (see also the book of Friz and Hairer [32]) provides a general analytic and deterministic framework for constructing rough integrals and solutions of differential equations with rough time dependence. The focus is on constructing solution maps for systems like (2.2.1) that are continuous with respect to a suitably augmented rough-path metric, which measures, in addition to the Hölder distance between paths, the distance between certain iterated integrals. In analogy with the results on stochastic flows, one can construct flows of rough diffeomorphisms with the theory of rough paths, as we now demonstrate. The presentation below follows most closely that of [32].

Let $\alpha \in (1/3, 1/2]$ and $m \geq 1$ be fixed.

Definition 2.2.1. *The map* $\mathbf{W} = (W, \mathbb{W}) \in C([0,T], \mathbb{R}^m) \times C([0,T]^2, \mathbb{R}^m \otimes \mathbb{R}^m)$ *is said to be an* α*-Hölder continuous rough path, and we write* $\mathbf{W} \in \mathscr{C}^\alpha([0,T], \mathbb{R}^m)$*, if*

$$
\begin{cases} \|\mathbf{W}\|_{\mathscr{C}^\alpha} := \sup_{s \neq t} \dfrac{|W_s - W_t|}{|s-t|^\alpha} + \sup_{s \neq t} \dfrac{|\mathbb{W}_{st}|}{|s-t|^{2\alpha}} < \infty, \text{ and} \\ \mathbb{W}_{st} - \mathbb{W}_{su} - \mathbb{W}_{ut} = (W_u - W_s) \otimes (W_t - W_u) \quad \text{for any } s, u, t \in [0,T]. \end{cases} \tag{2.2.4}
$$

If $\mathbf{W}$ *also satisfies*

$$
\mathrm{Sym}(\mathbb{W}_{st}) = \frac{1}{2}(W_t - W_s) \otimes (W_t - W_s) \quad \text{for any } s, t \in [0,T], \tag{2.2.5}
$$

then $\mathbf{W}$ *is said to be a geometric rough path, and we write* $\mathbf{W} \in \mathscr{C}_g^\alpha([0,T], \mathbb{R}^m)$*.*

The quantity $\|\mathbf{W}\|_{\mathscr{C}^\alpha}$ is called the rough-path norm of $\mathbf{W}$, although $\mathscr{C}^\alpha$ is not a linear space, due to the nonlinear nature of the second two constraints in (2.2.4).

If W is smooth, then $\mathbb{W}$ can be given by the Riemann-Stieltjes integrals

$$
\mathbb{W}_{s,t}^{ij} := \int_s^t (W_r^i - W_s^i)\, dW_r^j \quad \text{for } 1 \leq i, j, \leq m.
$$

However, given $W \in C^{0,\alpha}([0,T], \mathbb{R}^m)$, $\mathbb{W}$ is not in general uniquely determined by the algebraic constraints of Definition 2.2.1, precisely because $\alpha \leq 1/2$. Indeed, if $(W, \mathbb{W}) \in \mathscr{C}^\alpha$, then so is $(W, \tilde{\mathbb{W}})$ if

$$
\tilde{\mathbb{W}}_{st} := \mathbb{W}_{st} + F(t) - F(s) \quad \text{for some } F \in C^{0,2\alpha}([0,T], \mathbb{R}^m \otimes \mathbb{R}^m). \tag{2.2.6}
$$

The same remarks about the nonuniqueness of $\mathbb{W}$ are true for paths in $\mathscr{C}_g^\alpha$, in which case (2.2.6) defines another geometric rough-path lift if F takes values in the antisymmetric tensors.

In view of the fact that Brownian paths are α-Hölder continuous for any $\alpha \in \left(0, \frac{1}{2}\right)$, a Brownian motion can be viewed as a random rough path, if $\mathbb{W}$ is properly defined. For instance, both the Itô and Stratonovich iterated integrals

$$\mathbb{W}^{\mathrm{Ito},ij}_{s,t} = \int_s^t \left(W^i_r - W^i_s\right) dW^j_r \quad \text{and} \quad \mathbb{W}^{\mathrm{Strat},ij}_{s,t} = \int_s^t \left(W^i_r - W^i_s\right) \circ dW^j_r$$

satisfy the first algebraic constraint in Definition 2.2.1. Moreover,

$$\mathbb{W}^{\mathrm{Strat}}_{s,t} = \mathbb{W}^{\mathrm{Ito}}_{s,t} + \frac{t-s}{2} I,$$

which is a special case of (2.2.6). As is well-known, $(W, \mathbb{W}^{\mathrm{Strat}})$ belongs to the space of geometric rough paths $\mathscr{C}^\alpha_g$, while $(W, \mathbb{W}^{\mathrm{Ito}})$ does not.

In the sequel, we focus only on rough paths which are geometric. One reason is that (2.2.5) ensures that geometric rough paths satisfy, formally, the standard chain and product rules from differential calculus – in the stochastic setting, this means that there are no Itô correction terms. Furthermore, for any $\mathbf{W} \in \mathscr{C}^\alpha_g([0,T], \mathbb{R}^m)$, there exists a sequence of smooth paths $W^n : [0,T] \to \mathbb{R}^m$ such that, as $n \to \infty$, W^n and $\mathbb{W}^n$ converge uniformly to respectively W and $\mathbb{W}$, and, furthermore, $\sup_{n \in \mathbb{N}} \|\mathbf{W}^n\|_{\mathscr{C}^\alpha} < \infty$. As a particular example, the Stratonovich Brownian rough path can be approximated in the rough path metric by piecewise-affine interpolations, or by convolving with a standard mollfier. These classical facts belong to a family of results about so-called Wong-Zakai approximations [83, 84].

The analogue of Lemma 2.2.1 is the following.

Lemma 2.2.2. *Let $W \in \mathscr{C}^\alpha_g([0,T], \mathbb{R}^m)$, and assume that, for some $n \geq 0$, $H \in C^{4+n}_b(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$. Then $(x,p) \mapsto (X(x,p,t), P(x,p,t))$ is a C^{1+n}-diffeomorphism for all $t \in \mathbb{R}$, and its derivatives in x and p solve the rough differential equations obtained by differentiating the system* (2.2.1). *Moreover, the solution maps $\mathbf{W} \mapsto D^k_{x,p}(X,P)$ for $1 \leq k \leq 1+n$ are continuous from $\mathscr{C}^\alpha$ into C.*

The constructions and definitions above could be adapted to treat rough paths with less Hölder regularity, in which case more iterated integrals are involved in the definition (2.2.4), and more regularity is required for H. For ease of presentation, we do not pursue this further here, especially since the case of $\alpha \in (1/3, 1/2]$ is already enough to cover many interesting examples.

2.2.2 The case of vanishing Poisson brackets

Suppose now that the Poisson brackets of the Hamiltonians vanish:

$$\left\{H^i, H^j\right\} := \sum_{k=1}^{d} \left(\frac{\partial H^i}{\partial p_k} \frac{\partial H^j}{\partial x_k} - \frac{\partial H^i}{\partial x_k} \frac{\partial H^j}{\partial p_k} \right) = 0 \quad \text{for all } i, j = 1, 2, \ldots, m. \tag{2.2.7}$$

If (2.2.7) holds, then the Hamiltonian flows commute (see for instance Arnold [7]). In particular, for each $i = 1, 2, \ldots, m$ and for $\tau \in \mathbb{R}$, define the flow $\psi^i_\tau : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ by $\psi^i_\tau(x, p) = (X^i(x, p, \tau), P^i(x, p, \tau)$, where

$$\begin{cases} \dot{X}^i = -D_p H^i(P^i, X^i), & X^i(x, p, 0) = x, \\ \dot{P}^i = D_x H^i(P^i, X^i), & P^i(x, p, 0) = p. \end{cases} \tag{2.2.8}$$

Lemma 2.2.3. *Assume that the collection $\{H^i\}_{i=1}^m \subset C^2(\mathbb{R}^d \times \mathbb{R}^d)$ satisfies* (2.2.7)*. Then the solution (X, P) of* (2.2.1) *is given by the formula*

$$(X, P)(p, x, t) := \psi^m_{W^m(t) - W^m(t_0)} \circ \cdots \circ \psi^2_{W^2(t) - W^2(t_0)} \circ \psi^1_{W^1(t) - W^1(t_0)}(p, x). \tag{2.2.9}$$

This formula holds if W is smooth. It also holds if W is a geometric rough path, for example if W is a Brownian motion and (2.2.1) is interpreted with the Stratonovich differential. In fact, (2.2.9) makes sense whenever W is an arbitrary continuous path, and so this gives a way, via a density argument, to make sense of (2.2.1) for continuous paths when (2.2.7) holds.

The condition (2.2.7) holds trivially when $m = 1$, that is, when there is only a single path. It also holds automatically if each H^i is independent of x. Indeed, in this case, (2.2.1) reduces to

$$
\begin{cases} dX = -\displaystyle\sum_{i=1}^m D_p H^i(P) \cdot dW^i, & X(x,p,t_0) = x, \\ dP = 0, & P(x,p,t_0) = p, \end{cases} \tag{2.2.10}
$$

which is solved uniquely by $P(x,p,t) = p$ and

$$
X(x,p,t) := x - \sum_{i=1}^m D_p H^i(p) \left(W^i(t) - W^i(t_0) \right).
$$

2.2.3 Local-in-time, smooth-in-space solutions

We now demonstrate that the method of characteristics, discussed in the previous subsection, can be used to construct solutions of (2.1.1) in any of the three settings listed above, provided that the Hamiltonians are sufficiently regular. We focus mainly on the rough-path setting, since it contains the stochastic case, and then conclude with some remarks for continuous paths when (2.2.7) is satisfied.

Fix $\phi \in C_b^2(\mathbb{R}^d)$ and $t_0 \in [0,T]$, and set

$$
\begin{cases} \mathbf{X}(x,t) = X(x, D\phi(x), t), \quad \mathbf{P}(x,t) = P(x, D\phi(x), t), \quad \text{and} \\ \mathbf{Z}(x,t) = \phi(x) + \displaystyle\sum_{i=1}^m \int_{t_0}^t \left(H^i(\mathbf{P}, \mathbf{X}) - \mathbf{P} \cdot D_p H^i(\mathbf{P}, \mathbf{X}) \right) \cdot dW^i. \end{cases} \tag{2.2.11}
$$

The expression defining $\mathbf{Z}$ is interpreted as a rough integral if W is a geometric rough path, and $(\mathbf{X}, \mathbf{P}, \mathbf{Z})$ solves the rough characteristic system

$$\text{(2.2.12)} \qquad \begin{cases} d\mathbf{X} = -\sum_{i=1}^{m} D_p H^i(\mathbf{P}, \mathbf{X}) \cdot dW^i, & \mathbf{X}(x, t_0) = x, \\ d\mathbf{P} = \sum_{i=1}^{m} D_x H^i(\mathbf{P}, \mathbf{X}) \cdot dW^i, & \mathbf{P}(x, t_0) = D\phi(x), \\ d\mathbf{Z} = \sum_{i=1}^{m} \Big(H^i(\mathbf{P}, \mathbf{X}) - \mathbf{P} \cdot D_p \mathbf{H}^i(\mathbf{P}, \mathbf{X}) \Big) \cdot dW^i, & \mathbf{Z}(x, t_0) = \phi(x). \end{cases}$$

The boundedness of $D^2\phi$ and the flow properties in Lemmas 2.2.1 or 2.2.2 yield the following:

Lemma 2.2.4. *Assume that, for some $n \geq 0$, $H \in C^{4+n}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$ and $\phi \in C_b^{2+n}(\mathbb{R}^d)$. Then there exists $h > 0$ depending only on $\left\|D^2\phi\right\|_\infty$, the derivatives of H, and $\|\mathbf{W}\|_{\mathscr{C}^\alpha}$ such that, for all $t \in (t_0 - h, t_0 + h)$, $x \mapsto \mathbf{X}(x, t)$ is invertible on $\mathbb{R}^d$, and both $t \mapsto \mathbf{X}(\cdot, t)$ and $t \mapsto \mathbf{X}^{-1}(\cdot, t)$ belong to $C((t_0 - h, t_0 + h); C_b^{1+n}(\mathbb{R}^d))$.*

The fact that $\mathbf{X}$ is invertible allows us to define

$$\text{(2.2.13)} \qquad \Phi(x, t) = S(t, t_0)\phi(x) := \mathbf{Z}(\mathbf{X}^{-1}(x, t), t) \quad \text{for } (x, t) \in \mathbb{R}^d \times (t_0 - h, t_0 + h).$$

Lemma 2.2.5. *Assume that, for some $n \geq 0$, $H \in C^{4+n}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$ and $\phi \in C_b^{2+n}(\mathbb{R}^d)$, and let $h > 0$ be as in Lemma 2.2.4. Then the function Φ defined by* (2.2.13) *belongs to $C((t_0 - h, t_0 + h); C_b^{2+n}(\mathbb{R}^d))$, and is a solution of the pathwise Hamilton-Jacobi equation* (2.1.1).

Proof. From the equations satisfied by $D_x\mathbf{X}$ and $D_x\mathbf{P}$, a straightforward calculation yields

$$d\left(D_x\mathbf{Z} - \mathbf{P} \cdot D_x\mathbf{X}\right) = 0,$$

and therefore

$$D\Phi(\mathbf{X}, t) \cdot D_x\mathbf{X} = D_x\mathbf{Z} = P \cdot D_x\mathbf{X}.$$

It follows that $\mathbf{P}(x,t) = D\Phi(\mathbf{X}(x,t),t)$ for all $(x,t) \in \mathbb{R}^d \times (t_0 - h, t_0 + h)$, and so Lemma 2.2.4 implies that Φ maps $(t_0 - h, t_0 + h)$ continuously into $C_b^{2+n}(\mathbb{R}^d)$. It is then standard to verify that Φ solves (2.1.1). □

The next lemma summarizes some properties of the solution operators $S(t, t_0)$. The proofs are immediate or follow from the classical case by approximating W with appropriate smooth paths and passing to the limit. Indeed, if $\{W^n\}_{n=1}^{\infty}$ is a sequence of smooth paths that converge, as $n \to \infty$, to $\mathbf{W}$ in the rough path topology, as explained in Section 2.2, then the stability of the system (2.2.12) with respect to the rough path norm yields $h > 0$ independent of n such that the classical solution Φ^n to (2.1.1) driven by W^n is smooth on $\mathbb{R}^d \times (t_0 - h, t_0 + h)$, and, as $n \to \infty$, Φ^n converges uniformly to $S(t, t_0)\phi$.

Lemma 2.2.6. *Let $t_0 \in [0, T]$ and $\phi_1, \phi_2 \in C_b^2(\mathbb{R}^d)$, and choose $h > 0$ such that $S(\cdot, t_0)\phi_1$ and $S(\cdot, t_0)\phi_2$ belong to $C((t_0 - h, t_0 + h), C_b^2(\mathbb{R}^d))$.*

(a) For any $t \in (t_0 - h, t_0 + h)$ and $k \in \mathbb{R}$, $S(t, t_0)(\phi_1 + k) = S(t, t_0)\phi_1 + k$.

(b) For any $t \in (t_0 - h, t_0 + h)$, $\sup_{\mathbb{R}^d} (S(t, t_0)\phi_1 - S(t, t_0)\phi_2) \leq \sup_{\mathbb{R}^d}(\phi_1 - \phi_2)$.

(c) For any $r, s, t \in (t_0 - h, t_0 + h)$, $S(r, s)S(s, t)\phi_1 = S(r, t)\phi_1$.

Property (b) is simply the comparison principle for smooth solutions of (2.1.1). Note that (b) actually holds with equality, because of property (c), which follows from the uniqueness for (2.1.1) and the fact that the equation is reversible in the interval $(t_0 - h, t_0 + h)$.

By estimating the deviation of the characteristic $\mathbf{X}(x, t)$ from its starting point x, we obtain the following domain of dependence property for $S(t, t_0)$. For a compact set $K \subset \mathbb{R}^d$ and $r > 0$, define

$$K_r := \{x \in K : \operatorname{dist}(x, K^c) > \delta\}.$$

Lemma 2.2.7. *For every $R > 0$, there exists a nondecreasing, continuous function $\rho_R : [0,\infty) \to [0,\infty)$ with $\rho_R(0) = 0$ such that, if $K \subset \mathbb{R}^d$ is compact, $\phi_1, \phi_2 \in C_b^2(\mathbb{R}^d)$ satisfy $\|D\phi_1\|_\infty, \|D\phi_2\|_\infty \le R$, and $h > 0$ is such that $S(\cdot, t_0)\phi_1, S(\cdot, t_0)\phi_2 \in C(t_0 - h, t_0 + h), C_b^2(\mathbb{R}^d))$ and $K_{\rho_R(h)}$ is nonempty, then, for all $t \in (t_0 - h, t_0 + h)$,*

$$\sup_{K_{\rho_R(|t-t_0|)}} (S(t,t_0)\phi_1 - S(t,t_0)\phi_2) \le \sup_K (\phi_1 - \phi_2).$$

Proof. Set

$$\rho_R(\sigma) := \sup_{|p|\le R} \sup_{|t-t_0|\le\sigma} \sup_{x\in\mathbb{R}^d} |X(x,p,t) - x|.$$

The modulus of continuity of X is uniform for bounded p, and otherwise depends only on $\mathbf{W}$ and the derivatives of H. Therefore, ρ_R is finite, nondecreasing, continuous, and satisfies $\rho_R(0) = 0$.

For $i = 1, 2$, let $(\mathbf{X}_i, \mathbf{P}_i, \mathbf{Z}_i)$ be as in (2.2.11) for ϕ_i, and notice that, for any $t \in (t_0 - h, t_0 + h)$,

$$\left|\mathbf{X}_i^{-1}(x,t) - x\right| = \left|\mathbf{X}_i^{-1}(x,t) - X(\mathbf{X}_i^{-1}(x,t), D\phi_i(\mathbf{X}_i^{-1}(x,t)), t)\right| \le \rho_R(|t-t_0|). \tag{2.2.14}$$

Suppose first that $\phi_1 = \phi_2$ in K, and let x be in the interior of $K_{\rho_R(|t-t_0|)}$; that is, $\mathrm{dist}(x, K^c) > \rho_R(|t-t_0|)$. In view of (2.2.14), $y := \mathbf{X}_1^{-1}(x,t)$ lies in the interior of K. This implies that $\phi_1(y) = \phi_2(y)$ and $D\phi_1(y) = D\phi_2(y)$, so that $(\mathbf{X}_1, \mathbf{P}_1, \mathbf{Z}_1)(y,t) = (\mathbf{X}_2, \mathbf{P}_2, \mathbf{Z}_2)(y,t)$. Therefore $y = \mathbf{X}_2^{-1}(x,t)$, and

$$S(t,t_0)\phi_1(x) = \mathbf{Z}_1(\mathbf{X}_1^{-1}(x,t),t) = \mathbf{Z}_1(y,t) = \mathbf{Z}_2(y,t) = \mathbf{Z}_2(\mathbf{X}_2^{-1}(x,t),t) = S(t,t_0)\phi_2(x).$$

By continuity, the equality is true for any $x \in K_{\rho_R(|t-t_0|)}$.

Now assume $\phi_1 \le \phi_2$ in K, fix $\varepsilon > 0$, and let $\tilde{\phi}_2 \in C_b^2(\mathbb{R}^d)$ be such that $\phi_2 = \tilde{\phi}_2$ in K,

$\phi_1 \le \tilde{\phi}_2 + \varepsilon$ in $\mathbb{R}^d$, and $\left\| D\tilde{\phi}_2 \right\|_\infty \le R$. Then Lemma 2.2.6(a) yields, for all $x \in K_{\rho_R(|t-t_0|)}$,

$$S(t,t_0)\phi_1(x) \le S(t,t_0)(\tilde{\phi}_2+\varepsilon)(x) = S(t,t_0)(\phi_2+\varepsilon)(x) = S(t,t_0)\phi_2(x)+\varepsilon.$$

Letting $\varepsilon \to 0$ finishes the proof in this case. For general ϕ_1 and ϕ_2, the result follows from Lemma 2.2.6(a) and the fact that $\phi_1 \le \phi_2 + \sup_K(\phi_1 - \phi_2)$ in K. □

We conclude this section by explaining how the previous results improve in the case of vanishing Poisson brackets (2.2.7); namely, the less regularity is required for the Hamiltonians, and W is allowed to be merely continuous. In particular, for each fixed $i = 1, 2, \ldots, m$, the system (2.2.8) provides a classical way to construct, for some sufficiently small $\tau > 0$, smooth solutions of the individual Hamilton-Jacobi equations

$$\begin{cases} U_t = H^i(DU^i, x) & \text{in } \mathbb{R}^d \times (-\tau,\tau), \\ U(\cdot,0) = \phi & \text{in } \mathbb{R}^d. \end{cases} \tag{2.2.15}$$

If (2.2.7) holds and, for some $n \ge 0$, $H \in C^{2+n}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$ and $\phi \in C_b^{2+n}(\mathbb{R}^d)$, then the solution U will be in $C_b^{2+n}(\mathbb{R}^d \times (-\tau,\tau))$ as well.

Let $S^i(t) : C_b^2(\mathbb{R}^d) \to C_b^2(\mathbb{R}^d)$ denote the solution operator for (2.2.15). That is, for each $\phi \in C_b^2(\mathbb{R}^d)$ and sufficiently small $\tau > 0$, $S^i(t)\phi(x) = U(x,t)$ for all $(x,t) \in \mathbb{R}^d \times (-\tau,\tau)$, where U is the solution of (2.2.15).

Lemma 2.2.8. *Assume that the collection* $\{H^i\}_{i=1}^m \subset C^{2+n}(\mathbb{R}^d \times \mathbb{R}^d)$ *satisfies* (2.2.7). *Let* $\phi \in C_b^{2+n}(\mathbb{R}^d)$. *Then there exists* $h > 0$ *depending only on* $\left\| D^2\phi \right\|_\infty$ *such that* (2.1.1) *has a unique solution* $\Phi \in C((t_0-h, t_0+h), C_b^{2+n}(\mathbb{R}^d))$, *which is given by*

$$\Phi(x,t) := \prod_{i=1}^m S^i\left(W^i(t) - W^i(t_0)\right)\phi(x).$$

The quantity $h > 0$ in Lemma 2.2.8 is chosen so that the continuous paths $\{W^i\}_{i=1}^m$

satisfy

$$\sup_{|t-t_0|<h} \max_{i=1,2,\ldots,m} \left| W^i(t) - W^i(t_0) \right| < \tau,$$

where $\tau > 0$ is small enough to ensure that smooth solutions exist for the individual equation (2.2.15) for each $i = 1, 2, \ldots, m$.

Finally, note that, when the $\{H^i\}_{i=1}^m$ satisfy (2.2.7), then the propagation speed in Lemma 2.2.7 is equal to

$$\rho_R(\sigma) = \sup_{|p|\leq R} \sup_{|t-t_0|\leq\sigma} \sup_{x\in\mathbb{R}^d} \max_{i=1,2,\ldots,m} \left| D_p H^i(p,x) \right| \left| W_t^i - W_{t_0}^i \right|,$$

in accordance with the classical result on finite speed of propagation for Hamilton-Jacobi equations.

2.3 The definition of pathwise viscosity solutions

The local-in-time spatially-smooth solutions constructed in the previous section are used to define global sub- and super-solutions for the equation

$$du = F(D^2u, Du, u, x, t)\, dt + \sum_{i=1}^m H^i(Du, x) \cdot dW^i \quad \text{in } \mathbb{R}^d \times (0, \infty). \tag{2.3.1}$$

In analogy with the classical viscosity solution theory, test functions of the form $S(t, t_0)\phi$ are used to cancel out the "rough part" of (2.3.1) (the term involving dW^i), thereby dealing with the difficulties discussed in the Introduction.

Definition 2.3.1. *A function $u \in USC(\mathbb{R}^d \times [0,T])$ (resp. $u \in LSC(\mathbb{R}^d \times [0,T])$) is called a pathwise viscosity sub-solution (resp. super-solution) of (2.3.1) if u is bounded from above (resp. from below) and, whenever $\phi \in C_b^2(\mathbb{R}^d)$, $\psi \in C^1([0,T])$, $h > 0$, $S(t,t_0)\phi \in C_b^2(\mathbb{R}^d)$ for $t \in (t_0 - h, t_0 + h)$, and*

$$u(x,t) - S(t,t_0)\phi(x) - \psi(t)$$

attains a local maximum (resp. minimum) at $(x_0, t_0) \in \mathbb{R}^d \times (t_0 - h, t_0 + h)$*, then*

$$\begin{aligned} &\psi'(t_0) \le F(D^2\phi(x_0,t_0), D\phi(x_0,t_0), u(x_0,t_0), x_0, t_0) \\ &\left(\textit{resp. } \psi'(t_0) \ge F(D^2\phi(x_0,t_0), D\phi(x_0,t_0), u(x_0,t_0), x_0, t_0)\right). \end{aligned} \tag{2.3.2}$$

A solution of (2.3.1) *is both a sub- and super-solution.*

The following remarks regarding Definition 2.3.1 are useful in many arguments, and are analogous to observations from the classical viscosity theory.

Lemma 2.3.1. *(a) Assume that* u *satisfies the hypotheses of Definition 2.3.1, except that* (2.3.2) *only holds when* $u(x,t) - S(t,t_0)\phi(x) - \psi(t)$ *attains a strict maximum (resp. minimum) at* (x_0, t_0)*, that is, when*

$$u(x,t) - S(t,t_0)\phi(x) - \psi(t) \le u(x_0,t_0) - \phi(x_0) - \psi(t_0) \quad (\textit{resp. } \ge)$$

for all $(x,t) \in \mathbb{R}^d \times (t_0 + h, t_0 + h)$*, with equality if and only if* $(x,t) = (x_0,t_0)$*. Then* u *is a pathwise viscosity sub- (resp. super-) solution in the sense of Definition 2.3.1.*

(b) If $0 < t_0 \le T$ *and* u *is a sub- (resp. super-) solution in* $\mathbb{R}^d \times (0, t_0)$*, then it is a sub- (resp. super-) solution in* $\mathbb{R}^d \times (0, t_0]$*.*

It follows that it is sufficient to consider strict maxima or minima, as well as maxima or minima over half open neighborhoods like $B_r(x_0) \times (t_0 - r, t_0]$ instead of $N_r(x_0, t_0)$.

Proof of Lemma 2.3.1. Since the proofs for sub- and super-solutions are similar, we only present the sub-solution case.

(a) Assume that u is upper-semicontinuous and bounded from above, and $u(x,t) - S(t,t_0)\phi(x) - \psi(t)$ attains a local maximum at $(x_0,t_0) \in \mathbb{R}^d \times (t_0 - h, t_0 + h)$. In view of Lemma 2.2.6(a), we may assume, without loss of generality, that $u(x_0,t_0) = \phi(x_0)$. In particular, for some $r > 0$,

$$u(x, t_0) - \phi(x) \leq 0 \quad \text{for all } x \in B_r(x_0).$$

Choose $\tilde{\phi} \in C_b^2(\mathbb{R}^d)$ such that $\tilde{\phi}(x) = \phi(x) + |x - x_0|^4$ for $x \in B_r(x_0)$ and $u(\cdot, t_0) < \tilde{\phi}$ on $\mathbb{R}^d \backslash B_r(x_0)$, and set $\tilde{\psi}(t) := \psi(t) + |t - t_0|^2$. Then, by Lemma 2.2.6(b),

$$u(x, t) - S(t, t_0)\tilde{\phi}(x) - \tilde{\psi}(t)$$

attains a strict maximum at (x_0, t_0). The result now follows from the fact that $D\tilde{\phi}(x_0) = D\phi(x_0)$, $D^2\tilde{\phi}(x_0) = D^2\phi(x_0)$, and $\tilde{\psi}'(t_0) = \psi'(t_0)$.

(b) Assume that u is upper-semicontinuous and bounded from above, and, for some $r > 0$, $u(x, t) - S(t, t_0)\phi(x) - \psi(t)$ attains a maximum in $B_r(x_0) \times (t_0 - r, t_0]$ at (x_0, t_0). By replacing ϕ and ψ with respectively $\tilde{\phi}$ and $\tilde{\psi}$ as in part (a), the maximum may be assumed to be strict over $\mathbb{R}^d \times (t_0 - r, t_0]$.

Fix $\nu > 0$ and assume that (x_ν, t_ν) is a maximum point for

$$u(x, t) - S(t, t_0)\phi(x) - \psi(t) - \frac{\nu}{t_0 - t}$$

over $\overline{B_1(x_0)} \times [t_0 - r, t_0]$. Then $t_\nu \in [t_0 - r, t_0)$ for all $\nu > 0$, because u is bounded. Let $(y, s) \in \overline{B_1(x_0)} \times [t_0 - r, t_0]$ be an accumulation point of the sequence $\{(x_\nu, t_\nu)\}_{\nu > 0}$ as $\nu \to 0$, and assume that $s \neq t_0$. For fixed $(x, t) \in \overline{B_1(x_0)} \times [t_0 - r, t_0)$,

$$u(x, t) - S(t, t_0)\phi(x) - \psi(t) - \frac{\nu}{t_0 - t} \leq u(x_\nu, t_\nu) - S(t_\nu, t_0)\phi(x_\nu) - \psi(t_\nu) - \frac{\nu}{t_0 - t_\nu}.$$

Letting $\nu \to 0$ along a subsequence such that $(x_\nu, t_\nu) \to (y, s)$ yields

$$u(x, t) - S(t, t_0)\phi(x) - \psi(t) \leq u(y, s) - S(s, t_0)\phi(y) - \psi(s),$$

and, in view of the semicontinuity of u, the same inequality holds for $(x,t) = (x_0, t_0)$, contradicting the strictness of the maximum point (x_0, t_0). It follows that the whole sequence (x_ν, t_ν) converges to (x_0, t_0), and in particular, for sufficiently small ν, $(x_\nu, t_\nu) \in B_1(x_0) \times (t_0 - r, t_0)$. Therefore, Definition 2.3.1 yields

$$\psi'(t_\nu) \leq \psi'(t_\nu) + \frac{\nu}{(t_0 - t_\nu)^2} \leq F(D^2 S(t_\nu, t_0)\phi(x_\nu), DS(t_\nu, t_0)\phi(x_\nu), u(x_\nu, t_\nu), x_\nu, t_\nu),$$

and the proof is finished upon letting $\nu \to 0$. □

It is a simple consequence of Lemma 2.3.1 that the following is equivalent to Definition 2.3.1 in the first-order setting, that is, when $F \equiv 0$.

Definition 2.3.2. *A function $u \in USC(\mathbb{R}^d \times (0,\infty))$ (resp. $LSC(\mathbb{R}^d \times (0,\infty))$ is a pathwise viscosity sub- (resp. super-) solution of* (2.3.1) *if u is bounded from above (resp. from below) and, whenever $t_0 > h > 0$, $\Phi \in C((t_0 - h, t_0 + h), C_b^1(\mathbb{R}^d))$ is a solution of* (2.1.1)*, and $u(\cdot, t) - \Phi(\cdot, t)$ achieves a global maximum (resp. minimum) for all $t \in (t_0 - h, t_0 + h)$, then*

$$t \mapsto \max_{x \in \mathbb{R}^d} (u(x,t) - \Phi(x,t)) \quad \textit{is nonincreasing}$$

(resp.

$$t \mapsto \min_{x \in \mathbb{R}^d} (u(x,t) - \Phi(x,t)) \quad \textit{is nondecreasing).}$$

Finally, we remark that, in certain cases, the class of test functions used in Definitions 2.3.1 and 2.3.2 can be further restricted to those with even more regularity, as is in the case in the classical viscosity solution theory. For example, if, for some $n \geq 0$, $H \in C^{4+n}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$ (or $H \in C^{2+n}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$ and (2.2.7) holds), then one only needs to use test functions $\Phi \in C((t_0 - h, t_0 + h), C_b^{2+n}(\mathbb{R}^d))$. This follows from standard arguments in the theory of viscosity solutions, as well as the contraction property of the solution operator implied by Lemma 2.2.5.

2.4 Some special test functions

As in the classical viscosity theory, many quantitative arguments involve doubling variables, and it is therefore important to have objects that behave like the penalizing "distance function"

$$(x,y) \mapsto \frac{|x-y|^2}{2\delta}. \tag{2.4.1}$$

Due to the nature of the test functions in the definition, these objects will need to be solutions of appropriate "doubled" versions of (2.1.1).

To understand this, suppose that u and v are a sub- and super-solution of the first order equation

$$du = \sum_{i=1}^m H^i(Du,x)\cdot dW^i \quad \text{in } \mathbb{R}^d \times (0,\infty).$$

Then the doubled quantity $z(x,y,t) := u(x,t) - v(y,t)$ is a sub-solution of the equation

$$dz = \sum_{i=1}^m \Big(H^i(D_x z, x) - H^i(-D_y z, y) \Big) \cdot dW^i \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (0,\infty). \tag{2.4.2}$$

Our focus will be on constructing local-in-time, smooth-in-x solutions of (2.4.2) that resemble the classical penalizing quantity (2.4.1), as well as proving some precise quantitative statements.

2.4.1 Nonsmooth Hamiltonians

If H fails to be C^2, then it may only be possible to construct solutions of (2.1.1) for certain initial data. We will concentrate here on the x-independent case

$$\begin{cases} d\Phi = \displaystyle\sum_{i=1}^m H^i(D\Phi)\cdot dW^i & \text{in } \mathbb{R}^d \times (t_0 - h, t_0 + h) \\ \Phi(\cdot,0) = \phi & \text{in } \mathbb{R}^d, \end{cases} \tag{2.4.3}$$

where W is continuous and H satisfies

$$H^i = H_1^i - H_2^i \text{ for convex } H_1^i, H_2^i : \mathbb{R}^d \to \mathbb{R}. \tag{2.4.4}$$

It turns out that (2.4.4) is enough to ensure that (2.4.3) has a unique, global viscosity solution (see [55, 76]), and, moreover, there are precise path-stability estimates, as is explained in more detail in Chapter 4.

We will make some further assumptions here to simplify the presentation, namely that

$$H^i = H_1^i - H_2^i \text{ for convex } H_1^i, H_2^i : \mathbb{R}^d \to \mathbb{R} \text{ nonnegative} \tag{2.4.5}$$

and

$$\|DH\|_\infty < \infty. \tag{2.4.6}$$

The non-negativity in (2.4.5) is imposed only to simplify some arguments in what follows, and the setting can be reduced to the general case by transforming the equation appropriately. Meanwhile, if one works with Lipschitz solutions, one can always redefine H outside of a bounded set so that (2.4.6) holds. Note also that (2.4.6) implies that H grows at most linearly as $|p| \to +\infty$.

Assume $\eta : \mathbb{R}^d \to \mathbb{R}$ is strictly convex, and, for $\delta > 0$, define

$$\begin{aligned} \Phi(x,t) := \sup_{p \in \mathbb{R}^d} \Bigg\{ p \cdot x - \eta(p) - \delta \sum_{i=1}^m \left(H_1^i(p) + H_2^i(p) \right) \\ + \sum_{i=1}^m H^i(p)(W^i(t) - W^i(t_0)) \Bigg\}. \end{aligned} \tag{2.4.7}$$

Lemma 2.4.1. *If the open interval $I \ni t_0$ is such that*

$$\sup_{t \in I} \max_{i=1,2,\ldots,m} \left| W^i(t) - W^i(t_0) \right| < \delta,$$

then the function Φ *defined by* (2.4.7) *belongs to* $C(I, C^{1,1}(\mathbb{R}^d))$*, and is a solution of* (2.4.3) *with*

$$(2.4.8) \qquad \Phi(\cdot, t_0) = \phi(x) := \sup_{p \in \mathbb{R}^d} \left\{ p \cdot x - \eta(p) - \delta \sum_{i=1}^{m} \left(H_1^i(p) + H_2^i(p) \right) \right\}.$$

Proof. For all $x \in \mathbb{R}^d$ and $t \in I$, the function inside the brackets in (2.4.7) is strictly concave as a function of p, and therefore attains a unique global maximum. The smoothness of Φ in x then follows from the implicit function theorem.

Now, for $t \in \mathbb{R}$, let $S^i(t) : UC(\mathbb{R}^d) \to UC(\mathbb{R}^d)$ be the solution operator for the equation $u_t = H^i(Du)$. If $\psi \in UC(\mathbb{R}^d)$ is convex, then the Hopf formula proved by Lions and Rochet [50] gives

$$S^i(t)\psi(x) = \sup_{p \in \mathbb{R}^d} \left\{ p \cdot x - \psi^*(p) + tH^i(p) \right\},$$

and so (2.4.7) can be rewritten as

$$\Phi(x, t) = \prod_{i=1}^{m} S^i(W^i(t) - W^i(t_0))\phi(x)$$

with ϕ as in (2.4.8). If W is smooth, then the fact that Φ is a solution of (2.1.1) is justified by the regularity of Φ and a simple calculation. The result holds for continuous W by a density argument. □

We now construct a function $\Phi_\delta : \mathbb{R}^d \times [0, T]^2 \times C([0, T], \mathbb{R}^m) \to \mathbb{R}$ that is equal to a particular choice of (2.4.7) near the diagonal $\{(t, t) \in [0, T]^2\}$, and such that $\Phi_\delta(x-y, s, t; W)$ exhibits similar growth as (2.4.1) when $|x - y|$ is large.

Define the neighborhood $U_\delta(W)$ by

$$U_\delta(W) := \left\{ (s, t) \in [0, T]^2 : \operatorname{osc}(W, s, t) < \delta \right\},$$

let the projection $\pi_\delta(W) : [0, T]^2 \to \overline{U_\delta(W)}$ be such that $\pi_\delta(W)(s, t)$ is the element $(\tilde{s}, \tilde{t}) \in$

$\overline{U_\delta(W)}$ closest to (s,t) on the line $\tilde{s}+\tilde{t} = s+t$, and set

(2.4.9)
$$\Phi_\delta(x,s,t;W) := \begin{cases} \sup_{p\in\mathbb{R}^d} \Big\{ p\cdot x - \frac{\delta}{2}|p|^2 - \delta\sum_{i=1}^m (H_1^i(p)+H_2^i(p)) \\ \qquad + \sum_{i=1}^m H^i(p)\left(W^i(s)-W^i(t)\right)\Big\} & \text{if } (s,t)\in\overline{U_\delta(W)}, \\ \Phi_\delta\left(x,\pi_\delta(W)(s,t);W\right) & \text{if } (s,t)\notin\overline{U_\delta(W)}. \end{cases}$$

Lemma 2.4.2. *For some $C = C(\|DH\|_\infty) > 0$ and for all $\delta > 0$ and $W \in C([0,T];\mathbb{R}^m)$, the following hold:*

(a) For all $x\in\mathbb{R}^d$ and $(s,t),(\tilde{s},\tilde{t})\in U_\delta(W)$,

$$\left|\Phi_\delta(x,s,t;W)-\Phi_\delta(x,\tilde{s},\tilde{t};W)\right| \le C\left(1+\frac{|x|}{\delta}\right)\left(|W(s)-W(\tilde{s})|+\left|W(t)-W(\tilde{t})\right|\right).$$

(b) For all $(s,t)\in[0,T]^2$, $x\mapsto\Phi_\delta(x,s,t;W)$ is convex and semiconcave with constant $\frac{1}{\delta}$. That is,

$$0 \le D^2\Phi_\delta(x,s,t;W) \le \frac{1}{\delta}I_d \quad \text{in the sense of distributions.}$$

(c) For all $x\in\mathbb{R}^d$ and $s,t\in[0,T]$,

$$\frac{1}{2(C+1)\delta}|x|^2 - C\delta \le \Phi_\delta(x,s,t;W) \le \frac{1}{2\delta}|x|^2.$$

(d) For any fixed $y\in\mathbb{R}^d$ and $t\in[0,T]$, the functions

$$(x,s)\mapsto\Phi_\delta(x-y,s,t;W) \quad \text{and} \quad (x,s)\mapsto -\Phi_\delta(y-x,t,s;W)$$

are $C(I,C^{1,1}(\mathbb{R}^d))$-solutions of (2.4.3), where $I := \{s\in[0,T] : \operatorname{osc}(W,s,t)<\delta\}$.

Proof. (a) It is enough to prove the time-regularity for the s-variable alone, in view of the identity $\Phi_\delta(x,s,t;W) = \Phi_\delta(x,t,s;-W)$.

We show that there exists $C = C(\|DH\|_{oo}) > 0$ such that, for any $x \in \mathbb{R}^d$ and $(s,t) \in U_\delta(W)$, the unique maximum p^* achieved in the definition of Φ_δ satisfies $\delta\,|p^*| \le C\delta + |x|$. The result then follows from the formula for Φ_δ and the linear growth of H implied by (2.4.6).

If

$$J(p) := p\cdot x - \frac{\delta}{2}|p|^2 - \delta\sum_{i=1}^m (H_1^i(p) + H_2^i(p)) + \sum_{i=1}^m H^i(p)\left(W^i(s) - W^i(t)\right),$$

then, for any $q \in \mathbb{R}^d$, (2.4.6) and the inequality $J(p^*) \ge J(p^*+q)$ imply that

$$\delta p^* \cdot \frac{q}{|q|} - \frac{\delta}{2}|q| \le |x| + C\delta.$$

Setting $q = t\frac{p^*}{|p^*|}$ and sending $t \to 0^+$ yields the claim.

(b) As a pointwise supremum of affine functions, Φ_δ is clearly convex, while the semiconcavity follows from elementary convex analysis and the convexity of

$$p \mapsto \delta\sum_{i=1}^m (H_1^i(p) + H_2^i(p)) - \sum_{i=1}^m H^i(p)\left(W^i(s) - W^i(t)\right).$$

(c) This can be deduced from Young's inequality and the fact that, for some $C = C(\|DH\|_\infty) > 0$ and for all $p \in \mathbb{R}^d$ and $(s,t) \in U_\delta(W)$,

$$0 \le \delta\sum_{i=1}^m (H_1^i(p) + H_2^i(p)) - \sum_{i=1}^m H^i(p)\left(W^i(s) - W^i(t)\right) \le C\delta(1+|p|).$$

(d) This is a direct consequence of Lemma 2.4.1. $\square$

2.4.2 Convex Hamiltonians

We now turn to Hamiltonians with nontrivial spatial dependence. Here, we consider a single Hamiltonian which is smooth, strictly convex, and satisfies certain growth bounds. More precisely,

$$\begin{cases} H \in C_b^2(B_R \times \mathbb{R}^d) \text{ for all } R > 0,\ D_p^2 H \text{ is strictly positive, and} \\ \text{there exist convex, increasing functions } \underline{\nu}, \overline{\nu} : [0,\infty) \to \mathbb{R} \text{ such that} \\ \underline{\nu}(|p|) \le H(p,x) \le \overline{\nu}(|p|) \quad \text{for all } (p,x) \in \mathbb{R}^d \times \mathbb{R}^d. \end{cases} \tag{2.4.10}$$

For $x, y \in \mathbb{R}^d$ and $\tau > 0$, define

$$\begin{cases} \mathcal{A}(x,y,\tau) := \left\{ \gamma \in W^{1,\infty}([0,\tau],\mathbb{R}^d) : \gamma_0 = x,\ \gamma_\tau = y \right\} \quad \text{and} \\ \ell(x,y,\tau)_s = \ell_s := x + \dfrac{s}{\tau}(y-x). \end{cases}$$

Note that $\ell(x,y,\tau) \in \mathcal{A}(x,y,\tau)$ and $\mathcal{A}(x,y,\tau) = W_0^{1,\infty}([0,\tau],\mathbb{R}^d) + \ell(x,y,\tau)$.

The distance function associated to H is

$$L(x,y,\tau) := \inf \left\{ \int_0^\tau H^*\left(-\dot{\gamma}_s, \gamma_s\right)\, ds : \gamma \in \mathcal{A}(x,y,\tau) \right\}. \tag{2.4.11}$$

We summarize its main properties in the next lemma. For $R > 0$, define

$$\Delta_R := \left\{ (x,y) \in \mathbb{R}^d \times \mathbb{R}^d : |x-y| \le R \right\}.$$

Lemma 2.4.3. *Assume that H satisfies* (2.4.10)*. Then the following hold:*

(a) L is a viscosity solution of

$$\frac{\partial L}{\partial \tau} = H(D_x L, x) \quad \textit{and} \quad \frac{\partial L}{\partial \tau} = H(-D_y L, y) \quad \textit{in } \mathbb{R}^d \times \mathbb{R}^d \times (0,\infty).$$

In particular, $H(D_xL(x,y,\tau),x) = H(-D_yL(x,y,\tau),y)$ *whenever* L *is differentiable at* (x,y,τ).

(b) For all $x, y \in \mathbb{R}^d$ *and* $\tau > 0$,

$$\tau\overline{\nu}^*\left(\frac{|x-y|}{\tau}\right) \leq L(x,y,\tau) \leq \tau\underline{\nu}^*\left(\frac{|x-y|}{\tau}\right).$$

Furthermore, there exists $\gamma \in \mathcal{A}(x,y,\tau)$ *such that* $L(x,y,\tau) = \int_0^\tau H^*(-\dot{\gamma}_s, \gamma_s)\, ds$, *and, for some* $c \geq 1$ *and almost every* $s \in [0,\tau]$,

$$\frac{|x-y|}{c\tau} \leq |\dot{\gamma}_s| \leq \frac{c|x-y|}{\tau}.$$

(c) For all $R > 0$, *there exists a constant* $C = C_R > 0$ *such that*

$$|D_xL| + |D_yL| \leq C \quad \textit{and} \quad D^2L \leq CI_{2d} \quad \textit{on } \Delta_R \times \left[\frac{1}{R}, R\right].$$

Proof. (a) This follows from well known variational formulae for solutions of Hamilton-Jacobi equations. See, for instance, Lions [51].

(b) The bounds for L are immediate from (2.4.10) and (2.4.11). In view of the C^2-regularity and uniform convexity, a classical variational argument yields the existence of a minimizer γ. The bounds for $\dot{\gamma}$ can then be inferred from the Euler-Lagrange equation.

(c) Pick $(x,y) \in \Delta_R$, $\tau > 1/R$, and $h \in B_1$, and let $\gamma \in \mathcal{A}(x,y,\tau)$ be a minimizer for $L(x,y,\tau)$. Then $\{s \mapsto \gamma_s + s/\tau h : s \in [0,\tau]\} \in \mathcal{A}(x, y+h, \tau)$, and it follows from part (b) and (2.4.10) that

$$L(x,y+h,\tau) - L(x,y,\tau) \leq \int_0^\tau \left(H^*\left(-\dot{\gamma}_s - h, \gamma_s + sh\right) - H^*(-\dot{\gamma}_s, \gamma_s)\right)\, ds \leq C|h|.$$

The opposite inequality is obtained by switching the roles of y and $y+h$. This yields the bound for D_yL, and the argument for D_xL is similar.

Now let $h, k \in B_1$ and set $\eta_s := h + s/\tau(k-h)$. Then $\gamma \pm \eta \in \mathcal{A}(x \pm h, y \pm k, \tau)$, and, for all $s \in [0, \tau]$,

$$|\eta_s| + |\dot{\eta}_s| \leq C\left(|h| + |k|\right). \tag{2.4.12}$$

The proof is finished upon applying (2.4.10) to obtain

$$\begin{aligned}\int_0^\tau H^*(-\dot{\gamma}_s - \dot{\eta}_s, \gamma_s + \eta_s)\, ds + \int_0^\tau H^*(-\dot{\gamma}_s + \dot{\eta}_s, \gamma_s - \eta_s)\, ds - 2\int_0^\tau H^*(-\dot{\gamma}_s, \gamma_s)\, ds \\ \leq C(|h|^2 + |k|^2).\end{aligned}$$

□

As the next lemma demonstrates, the semiconcavity estimate allows L to be used as a test function in the definition of viscosity solutions, because it implies the existence of many points of differentiability.

Lemma 2.4.4. *Under the same assumptions as Lemma 2.4.3, assume that $\phi \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$ and $L(\cdot, \cdot, \tau_0) - \phi$ attains a local minimum at (x_0, y_0). Then L is differentiable at (x_0, y_0, τ_0) with*

$$\begin{cases} (D_x L(x_0, y_0, \tau_0), D_y L(x_0, y_0, \tau_0)) = (D_x \phi(x_0, y_0), D_y \phi(x_0, y_0)) \quad \text{and} \\ \dfrac{\partial L}{\partial \tau}(x_0, y_0, \tau_0) = H(D_x L(x_0, y_0, \tau_0), x_0) = H(-D_y L(x_0, y_0, \tau_0), y_0). \end{cases}$$

Proof. In view of the semiconcavity of $L(\cdot, \cdot, \tau_0)$ on $\mathbb{R}^d \times \mathbb{R}^d$, the super-differential of $L(\cdot, \cdot, \tau_0)$ is nonempty at every point. Meanwhile, $(p_0, q_0) := D\phi(x_0, y_0)$ belongs to the sub-differential of $L(\cdot, \cdot, \tau_0)$ at (x_0, y_0). This implies that $L(\cdot, \cdot, \tau_0)$ is differentiable at (x_0, y_0), and the first line above holds.

Choose $\psi^+, \psi^- \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$ such that

$$\begin{cases} \psi^- \leq L(\cdot,\cdot,\tau_0) \leq \psi^+, \\ \psi^-(x_0,y_0) = L(x_0,y_0,\tau_0) = \psi^+(x_0,y_0), \quad \text{and} \\ D\psi^-(x_0,y_0) = D\psi^+(x_0,y_0) = (p_0,q_0). \end{cases}$$

The method of characteristics can then be used to construct, for sufficiently small $\mu > 0$, solutions $\Psi^\pm \in C^2(\mathbb{R}^d \times \mathbb{R}^d \times (\tau_0 - \mu, \tau_0 + \mu))$ of the equations

$$\frac{\partial \Psi^\pm}{\partial \tau}(x,y,\tau) = H(D_x\Psi^\pm(x,y,\tau),x) \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (\tau_0 - \mu, \tau_0 + \mu).$$

The comparison principle and Lemma 2.4.3(a) then yield

$$\text{(2.4.13)} \quad \Psi^-(x,y,\tau) \leq L(x,y,\tau) \leq \Psi^+(x,y,\tau) \quad \text{for all } (x,y,\tau) \in \mathbb{R}^d \times \mathbb{R}^d \times (\tau_0 - \mu, \tau_0 + \mu).$$

Finally, the regularity of H and the equations for $\Psi^\pm$ allow for the Taylor expansion

$$\begin{aligned} \Psi^\pm(x,y,\tau) &= L(x_0,y_0,\tau_0) + p \cdot (x - x_0) + q \cdot (y - y_0) \\ &\quad + H(p_0,x_0)(\tau - \tau_0) + O(|x - x_0|^2 + |y - y_0|^2 + |\tau - \tau_0|^2). \end{aligned}$$

Together with (2.4.13), this shows that L is differentiable at (x_0, y_0, τ_0) and

$$\frac{\partial L}{\partial \tau}(x_0,y_0,\tau_0) = H(D_xL(x_0,y_0,\tau_0),x_0).$$

A similar argument using the equation $\Psi_\tau = H(-D_y\Psi, y)$ gives the final desired equality

$$\frac{\partial L}{\partial \tau}(x_0,y_0,\tau_0) = H(-D_yL(x_0,y_0,\tau_0),y_0).$$

□

The semiconcavity estimate sometimes does not provide enough regularity for L to be useful as a test function. This is the case in the proof of the comparison principle in Section 3.2. Under some more conditions on H, however, it is possible to prove that L is actually smooth in a small neighborhood of the diagonal. We present an example here where H has some extra homogeneity. Assume that

$$(2.4.14) \qquad \begin{cases} p \mapsto H^*(p,\cdot) \text{ and } p \mapsto H(p,\cdot) \text{ are strictly convex in } \mathbb{R}^d \backslash \{0\}, \\ H^* \in C_b^2(B_R \backslash \overline{B_{1/R}} \times \mathbb{R}^d) \text{ for all } R > 1, \\ H^*(p,\cdot) > 0 \text{ for all } p \neq 0, \text{ and} \\ H^*(tp,y) = t^q H^*(p,y) \text{ for some } q > 1 \text{ and for all } (p,y) \in \mathbb{R}^d \times \mathbb{R}^d \text{ and } t > 0. \end{cases}$$

The homogeneity and the regularity given in (2.4.14) yield, for some $0 < c_0 \leq C_0$ and $C > 0$, and for all $(p,y) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$(2.4.15) \qquad \begin{cases} c_0|p|^q \leq H^*(p,y) \leq C_0|p|^q, \quad |D_y H^*(p,y)| \leq C|p|^q, \\ c_0|p|^{q-1} \leq |D_p H^*(p,y)| \leq C_0|p|^{q-1}, \quad |D^2_{py} H^*(p,y)| \leq C|p|^{q-1}, \\ c_0|p|^{q-2}\, I_d \leq D^2_p H^*(p,y) \leq C_0|p|^{q-2}\, I_d, \quad |D^2_y H^*(p,y)| \leq C|p|^q. \end{cases}$$

Standard convex analysis implies that H satisfies (2.4.14) and (2.4.15) with the exponent $q' = \frac{q}{q-1}$.

A general class of examples satisfying (2.4.15) is given, for some smooth, positive definite matrix g, by $H(p,x) := (g(x)p \cdot p)^{q'/2}$. This particular Hamiltonian is studied by Lions and Souganidis in [31] for $q = 2$. Generalizations of the above theory appear in [53].

The homogeneity of H^* implies that

$$(2.4.16) \qquad L(x,y,\tau) = \frac{1}{\tau^{q-1}} L(x,y) \quad \text{where} \quad L(x,y) := L(x,y,1).$$

We then have the following properties.

Lemma 2.4.5. *Assume* (2.4.14).

(a) $L(\cdot, 1)$ *is a viscosity solution of*

$$-(q-1)L + H(D_x L, x) = 0 \quad \textit{and} \quad -(q-1)L + H(-D_y L, y) = 0 \quad \textit{in } \mathbb{R}^d \times \mathbb{R}^d.$$

(b) There exists $r_0 > 0$ *such that* $L \in C^1(\Delta_{r_0})$.

Proof. (a) This is a corollary of Lemma 2.4.3(a) and the identity (2.4.16).

(b) Define $I : W_0^{1,q}([0,1], \mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ by

$$I(\gamma, x, y) := \int_0^1 H^*(-\dot{\gamma}_s + x - y, \gamma_s + x + s(y-x))\, ds.$$

Henceforth, the arguments of H^* and all of its derivatives are $(-\dot{\gamma}_s + x - y, \gamma_s + x + s(y-x))$.

Observe that $L(x,y) = \min_{\gamma \in W_0^{1,q}} I(\gamma, x, y)$, and, in view of Lemma 2.4.3(b), the minimum is attained for some $\gamma \in W_0^{1,\infty}$ satisfying, for some $c, C > 0$ and almost every s,

$$c|x-y| \le |\dot{\gamma}_s + y - x| \le C|x-y|.$$

As a result, we can assume that H^* grows at most quadratically by redefining H^* outside of $B_R \times \mathbb{R}^d$ for some large $R > 0$. It follows that the map I can be defined as before on $W_0^{1,2} \times \mathbb{R}^d \times \mathbb{R}^d$, and that I has the same minimizers.

For $x, y \in \mathbb{R}^d$, fix a minimizer γ. Then $D_\gamma I(\gamma, x, y) = 0$, and, for all $\xi, \eta \in W_0^{1,2}$,

$$\begin{aligned} &D_\gamma^2 I(\gamma, x, y)[\xi, \eta] \\ &\quad = \int_0^1 \left(D_p^2 H^* \dot{\xi}_s \cdot \dot{\eta}_s - D_{px}^2 H^* \dot{\xi}_s \cdot \eta_s - D_{xp}^2 H^* \xi_s \cdot \dot{\eta}_s + D_x^2 H^* \xi_s \cdot \eta_s \right) ds. \end{aligned}$$

In view of (2.4.15), there exist $c, C > 0$ such that

$$D^2_\gamma I(\gamma, x, y)[\eta, \eta] \geq c|x-y|^{q-2} \int_0^1 \left(|\dot{\eta}_s|^2 - C|x-y||\dot{\eta}_s||\eta_s| - C|x-y|^2|\eta_s|^2 \right) \, ds.$$

Young's and Poincaré's inequalities give, for a larger C and smaller c,

$$D^2_\gamma I(\gamma, x, y)[\eta, \eta] \geq c|x-y|^{q-2} \left(1 - C|x-y|^2\right) \int_0^1 |\dot{\eta}_s|^2 \, ds.$$

Set $r_0 := \frac{1}{\sqrt{2C}}$, where C is the constant from the previous line. Then, if $(x, y) \in \Delta_{r_0}$,

$$D^2_\gamma I(\gamma, x, y)[\eta, \eta] \geq \frac{c}{2}|x-y|^{q-2} \, \|\eta\|^2_{W^{1,2}_0}.$$

As a consequence, for $(x, y) \in \Delta_{r_0}$ with $x \neq y$, $I(\cdot, x, y)$ has a unique minimizer $\gamma = \gamma(x, y)$. Also, for any $\eta \in W^{1,2}_0$,

$$\left\| D^2_\gamma I(\gamma(x, y), x, y)[\cdot, \eta] \right\|_{\left(W^{1,2}_0\right)^*} \geq \frac{c}{2}|x-y|^{q-2} \, \|\eta\|_{W^{1,2}_0}.$$

It follows from the implicit function theorem that $(x, y) \mapsto \gamma(x, y)$ is C^1, and, therefore, $L(x, y) = I(\gamma(x, y), x, y)$ is C^1 in $\Delta_{r_0} \backslash \Delta_0$.

Lemma 2.4.3(b) implies that

$$c_0|x-y|^q \leq L(x, y) \leq C_0|x-y|^q,$$

and therefore L is differentiable on Δ_0 with $D_x L = D_y L = 0$. The bounds in (2.4.15) and a similar argument as in the proof of Lemma 2.4.3(c) yield

$$\lim_{R \to 0} \sup_{\Delta_R} \left(|D_x L| + |D_y L| \right) = 0,$$

and we conclude that L is C^1 in all of Δ_{r_0}. □

2.4.3 Level-set equations

For Hamiltonians depending on space that are not convex, there is no natural analogue to the "distance"-type functions constructed in the previous subsections. Therefore, in order to analyze local-in-time, smooth-in-space solutions of (1.1.1), it is necessary to study the characteristic equations.

As an example, we consider a Hamiltonian that arises in the study of level-set equations describing certain interface-motion problems. For some $a \in C_b^2(S^{d-1} \times \mathbb{R}^d)$, set

$$a(p,x) := \begin{cases} a\left(\frac{p}{|p|}, x\right)|p| & \text{if } p \neq 0, \text{ and} \\ 0 & \text{if } p = 0, \end{cases}$$

so that a is positively 1-homogenous in the gradient variable.

The standard method of characteristics breaks down here because of the singularity of $D_p a$ at $p = 0$, and thus, as before, we construct local smooth solutions only for particular initial data. For $\delta > 0$ and $x, y \in \mathbb{R}^d$, set

$$\phi(x,y) := \frac{1}{2}\left[(|x-y|-\delta)_+\right]^2. \tag{2.4.17}$$

Lemma 2.4.6. *There exists $\tau = \tau_\delta > 0$ and a function $U \in C^2(\mathbb{R}^d \times \mathbb{R}^d \times (-\tau,\tau)^2)$ with bounded second-order derivatives in space such that*

(a) $U = U(x,y,s,t)$ is a solution of

$$\begin{cases} U_s = a(D_x U, x) \quad \text{and} \quad U_t = -a(-D_y U, y) \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (-\tau,\tau)^2 \quad \text{and} \\ U(x,y,0,0) = \phi(x,y) \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d, \end{cases}$$

and the function $\tilde{U}(x,y,t) := U(x,y,t,t)$ is a solution of

$$\tilde{U}_t = a(D_x\tilde{U}, x) - a(-D_y\tilde{U}, y) \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (-\tau,\tau). \tag{2.4.18}$$

(b) There exists a constant $C > 0$ depending only on the derivatives of a such that, for $(x, y, s, t) \in \mathbb{R}^d \times \mathbb{R}^d \times (-\tau, \tau)^2$,

$$|U(x,y,s,t) - \phi(x,y)| \le C\left(|t||x-y|^2 + |s-t|\right). \tag{2.4.19}$$

Proof. Throughout the proof, we use the following facts, which follow from the positive homogeneity of a: for all $t > 0$,

$$\begin{cases} a(tp,x) = ta(p,x), \quad D_p a(tp,x) = D_p a(p,x), \\ D_p a(p,x)\cdot p = a(p,x), \quad \text{and} \quad D^2_{pp} a(p,x)\cdot D_p a(p,x) = 0. \end{cases} \tag{2.4.20}$$

The quantity $\tau > 0$ below, which depends only on δ and the derivatives of a, may change from line to line.

(a) Consider the system of equations

$$\begin{cases} \dot{X} = -D_p a(M,X) & X(x,y,0) = x, \\ \dot{M} = (I - M\otimes M) D_x a(M,X) & M(x,y,0) = \dfrac{x-y}{|x-y|}, \\ \dot{Y} = -D_p a(N,Y) & Y(x,y,0) = y, \\ \dot{N} = (1 - N\otimes N) D_y a(N,Y) & N(x,y,0) = \dfrac{x-y}{|x-y|}. \end{cases} \tag{2.4.21}$$

The quantity $\sigma := |M|$ solves

$$\dot{\sigma} = (D_x a(M,X)\cdot M)\cdot\left(\frac{1}{\sigma} - \sigma\right), \quad \sigma(0) = 1.$$

The unique solution to this equation is $\sigma \equiv 1$, and therefore $|M(x,t)| = |N(x,t)| = 1$ for all t. In particular, the system (2.4.21) has a global-in-time solution for all $x \neq y$.

There exists a sufficiently small $\tau = \tau_\delta > 0$ such that

$$(x, y) \mapsto (X(x, y, s), Y(x, y, t))$$

is a diffeomorphism of $\{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d : x \neq y\}$ for $s, t \in (-\tau, \tau)$. For such s, t, define $U(\cdot, s, t)$ in the image of this diffeomorphism implicitly by

$$U(X(x, s), Y(y, t), s, t) := \phi(x, y).$$

For $|x - y| \leq \delta/2$, and $s, t \in (-\tau, \tau)$, define also

$$U(x, y, s, t) := 0.$$

In order to show that U is well-defined in $\mathbb{R}^d \times \mathbb{R}^d \times (-\tau, \tau)^2$, we prove that, if τ is sufficiently small, then the two, possibly different, definitions of $U(x, y, s, t)$ actually coincide. In particular, it is necessary to demonstrate that the two definitions agree at (x, y, s, t) for which

$$|X(x, y, s) - Y(x, y, t)| < \delta/2. \tag{2.4.22}$$

The derivatives $\dot{X}$ and $\dot{Y}$ are uniformly bounded in view of (2.4.21) and the boundedness of $D_p a$, so there exists a sufficiently small τ depending on δ such that (2.4.22) holds only if $|x-y| < \delta$. For such x and y, we have $U(X(x, y, s), Y(x, y, t), s, t) = \phi(x, y) = 0$. Therefore, U is well-defined, and, moreover, belongs to $C^2(\mathbb{R}^d)$ with bounded second-order derivatives.

For $t \in \mathbb{R}$ and $|x - y| \leq \delta$, define $P(x, y, t) = Q(x, y, t) = 0$, and, for $|x - y| > \delta$, let $P(x, y, t)$ and $Q(x, y, t)$ be the solutions of

$$\begin{cases} \dot{P} = D_x a(P, X) & P(x, y, 0) = D_x \phi(x, y), \\ \dot{Q} = D_y a(Q, Y) & Q(x, y, 0) = -D_y \phi(x, y). \end{cases} \tag{2.4.23}$$

A calculation using (2.4.20) and (2.4.21) shows that

$$P(x,y,t) = |P(x,y,t)|M(x,y,t) \quad \text{and} \quad Q(x,y,t) = |Q(x,y,t)|N(x,y,t) \quad \text{for all } x \neq y,$$

and, if $|x-y| > \delta$, $X(x,y,\cdot)$ and $Y(x,y,\cdot)$ also solve the equations

$$\dot{X} = -D_p a(P,X) \quad \text{and} \quad \dot{Y} = -D_p a(Q,Y).$$

Routine calculations from the method of characteristics reveal that U and $\tilde{U}$ solve the desired equations.

(b) For sufficiently large $C > 0$ depending only on δ and the derivatives of a, the functions

$$\phi(x,y) \pm Ct(|x-y|^2 + 1)$$

are a sub- and super-solution of the equations in part (a). The desired bounds then follow from the comparison principle. □

2.4.4 Separated dependence

Let $H, f \in C^3(\mathbb{R}^d)$ be such that

$$\|Df\|_\infty + \left\|D^2 f\right\|_\infty + \left\|D^3 f\right\|_\infty + \left\|D^2 H\right\|_\infty + \left\|D^3 H\right\|_\infty < \infty. \tag{2.4.24}$$

For a particular smooth function φ_λ defined below, we use the method of characteristics to construct, for a sufficiently small $\tau_\lambda > 0$, a smooth solution of

$$\begin{cases} U_t = H(D_x U) + f(x) - H(-D_y U) - f(y) & \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (-\tau_\lambda, \tau_\lambda) \quad \text{and} \\ U(x,y,0) = \varphi_\lambda(x-y) & \text{in } \mathbb{R}^d \times \mathbb{R}^d. \end{cases} \tag{2.4.25}$$

A large part of this subsection is devoted to achieving sharp lower bounds for τ_λ.

The construction below, and the proofs of its properties, can be achieved for more general Hamiltonians, for example, those that satisfy, for some constant $C > 0$ and all $(p,x) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$
(2.4.26) \qquad \begin{cases} |D_x H(p,x)| \le C, \ |D_x^2 H(p,x)| \le C, \ |D_{xp}^2 H(p,x)| \le C, \\ |D_p^2 H(p,x)| \le C, \ |D_x^3 H(p,x)| \le C, \ |D_{xxp}^3 H(p,x)| \le C, \\ |D_{ppx}^3 H(p,x)| \le C(1+|p|)^{-1}, \ |D_p^3 H(p,x)| \le C. \end{cases}
$$

To simplify the presentation, we focus here on the Hamiltonian as in (2.4.25).

Fix $R > 0$, choose $g \in C^2((0,\infty))$ such that

$$
\begin{cases} g' > 0, \quad \|g'\|_\infty + \|g''\|_\infty \le C \quad \text{for some } C > 0 \text{ independent of } R, \text{ and} \\ g(s) := \begin{cases} s & \text{if } 0 \le s \le R, \\ R+1 & \text{if } s \le R+1, \end{cases} \end{cases}
$$

and set

$$
\varphi_{R,\lambda}(z) := \varphi_\lambda(z) = g\left(\frac{\lambda}{2}|z|^2\right).
$$

Observe that, for all $\lambda > 0$ and some $C = C_R > 0$,

$$
|\varphi_\lambda(z)| \le C, \quad |D\varphi_\lambda(z)| \le C\lambda^{1/2}, \quad \text{and} \quad |D^2\varphi_\lambda(z)| \le C\lambda.
$$

The dependence on R will be suppressed in what follows. In Section 3.4, where $\varphi_{R,\lambda}$ is used in the proof of the comparison principle for certain second-order equations, the quantity R depends on the given sub- and super-solution and remains fixed throughout the proof.

The system of characteristics for (2.4.25) is given by

$$\begin{cases} \dot{X} = -DH(P), & X(x,y,0) = x, \\ \dot{P} = Df(X), & P(x,y,0) = D\varphi_\lambda(x-y), \\ \dot{Y} = -DH(Q), & Y(x,y,0) = y, \\ \dot{Q} = Df(Y), & Q(x,y,0) = D\varphi_\lambda(x-y). \end{cases} \tag{2.4.27}$$

The relation to (2.4.25) is made via the identities

$$\begin{cases} P(x,y,t) = D_x U(X(x,y,t), Y(x,y,t), t) \quad \text{and} \\ Q(x,y,t) = -D_y U(X(x,y,t), Y(x,y,t), t). \end{cases} \tag{2.4.28}$$

In order to analyze the length of the interval of existence for solutions of (2.4.25), we directly study the Jacobian

$$J(x,y,t) := \det \begin{pmatrix} D_x X(x,y,t) & D_y X(x,y,t) \\ D_x Y(x,y,t) & D_y Y(x,y,t) \end{pmatrix}.$$

For definiteness, we define

$$\tau_\lambda := \inf \left\{ |t| \le 1 : J(x,y,t) = \frac{1}{2} \quad \text{for some} \quad x, y \in \mathbb{R}^d \right\}.$$

The method of characteristics then yields a smooth solution of (2.4.25) in $\mathbb{R}^d \times \mathbb{R}^d \times (-\tau_\lambda, \tau_\lambda)$.

To prove the comparison principle for the equation

$$du = F(D^2u, Du, x, t)dt + (H(Du) + f(x)) \cdot dW \quad \text{in } \mathbb{R}^d \times (0, \infty),$$

it is important to have a precise lower bound for τ_λ in terms of λ. From the bound $|D^2\varphi_\lambda| \le C\lambda$, it is not difficulty to prove, for some uniform constant $c > 0$, that $c\lambda \le \tau_\lambda \le 1$. Following the strategy in Section 3.4, it is then possible to prove the comparison principle whenever

$W \in C^{0,2/3}([0,\infty))$. This does not cover many other interesting cases, such as when W is a Brownian motion or other general semimartingale.

On the other hand, observe that, if $f \equiv 0$, then $\varphi_\lambda(x-y)$ is a stationary solution of (2.4.25). That is, in that case, $\tau_\lambda \equiv +\infty$. Treating the function $f \in C^3(\mathbb{R}^d)$ as a perturbation, it is possible to prove a much sharper lower bound for τ_λ.

Lemma 2.4.7. *There exist $c, C > 0$ depending only on R and the bounds in* (2.4.24) *such that, for all $\lambda \geq 1$, $\tau_\lambda \geq c\lambda^{-1/4}$, and, for all $\varepsilon > 0$, $(x,y,t) \in \mathbb{R}^d \times \mathbb{R}^d \times (-\tau_\lambda, \tau_\lambda)$,*

$$\begin{cases} |D_x U(x,y,t)| + |D_y U(x,y,t)| \leq C\lambda^{1/2}, \\ |D_x U(x,y,t) + D_y U(x,y,t)| \leq C\left(|x-y| \wedge 1\right)|t|, \end{cases} \tag{2.4.29}$$

$$\begin{cases} \left\| D^2 U(x,y,t) \right\| \leq C\lambda, \\ D^2 U(x,y,t) + \varepsilon D^2 U(x,y,t)^2 \leq C\lambda(1+\lambda\varepsilon) \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + C|t| \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \end{cases} \tag{2.4.30}$$

and

$$|U(x,y,t) - \varphi_\lambda(x,y)| \leq C|x-y||t|. \tag{2.4.31}$$

Proof. Throughout the proof, the positive constants c and C depend only on R and the bounds in (2.4.24), and may change from line to line.

To simplify the presentation, we consider only the one-dimensional case $d = 1$. However, all of the arguments carry through in the general case. We also adapt the shorthand

$$X_\alpha = D_\alpha X,\ Y_\alpha = D_\alpha Y,\ P_\alpha = D_\alpha P, \text{ and } Q_\alpha = D_\alpha Q \quad \text{for } \alpha = x, y.$$

Finally, we may assume below that $0 < t \leq 1$. The proofs for $t \in [-1, 0)$ are similar.

Define

$$\mathcal{A}_x := \begin{pmatrix} 0 & -H''(P) \\ f''(X) & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{A}_y := \begin{pmatrix} 0 & -H''(Q) \\ f''(Y) & 0 \end{pmatrix}.$$

Then the equations for X_α, Y_α, P_α, and Q_α with $\alpha = x, y$ are given by

$$\begin{cases} \dfrac{d}{dt}\begin{pmatrix} X_\alpha \\ P_\alpha \end{pmatrix} = \mathcal{A}_x \begin{pmatrix} X_\alpha \\ P_\alpha \end{pmatrix}, & \begin{pmatrix} X_x & X_y \\ P_x & P_y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \varphi''_\lambda & -\varphi''_\lambda \end{pmatrix}, \\ \dfrac{d}{dt}\begin{pmatrix} Y_\alpha \\ Q_\alpha \end{pmatrix} = \mathcal{A}_y \begin{pmatrix} Y_\alpha \\ Q_\alpha \end{pmatrix}, & \begin{pmatrix} Y_x & Y_y \\ Q_x & Q_y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \varphi''_\lambda & -\varphi''_\lambda \end{pmatrix}. \end{cases} \tag{2.4.32}$$

Through a change of variables, we relate (2.4.32) to the system

$$\begin{cases} \dfrac{d}{dt}\begin{pmatrix} \xi_i \\ \phi_i \end{pmatrix} = \mathcal{A}_x \begin{pmatrix} \xi_i \\ \phi_i \end{pmatrix}, & \begin{pmatrix} \xi_1 & \xi_2 \\ \phi_1 & \phi_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \dfrac{d}{dt}\begin{pmatrix} \eta_i \\ \psi_i \end{pmatrix} = \mathcal{A}_y \begin{pmatrix} \eta_i \\ \psi_i \end{pmatrix}, & \begin{pmatrix} \eta_1 & \eta_2 \\ \psi_1 & \psi_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{cases} \tag{2.4.33}$$

via the formulas

$$\begin{cases} \begin{pmatrix} X_x \\ P_x \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \phi_1 \end{pmatrix} + \varphi''_\lambda \begin{pmatrix} \xi_2 \\ \phi_2 \end{pmatrix}, & \begin{pmatrix} X_y \\ P_y \end{pmatrix} = -\varphi''_\lambda \begin{pmatrix} \xi_2 \\ \phi_2 \end{pmatrix}, \\ \begin{pmatrix} Y_x \\ Q_x \end{pmatrix} = \varphi''_\lambda \begin{pmatrix} \eta_2 \\ \psi_2 \end{pmatrix}, \quad \text{and} & \begin{pmatrix} Y_y \\ Q_y \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \psi_1 \end{pmatrix} - \varphi''_\lambda \begin{pmatrix} \eta_2 \\ \psi_2 \end{pmatrix}, \end{cases} \tag{2.4.34}$$

and therefore

$$J = \xi_1\eta_1 + \varphi''_\lambda(\xi_2\eta_1 - \xi_1\eta_2).$$

The bound for f' immediately gives the first line of (2.4.29) after inverting the characteristics in (2.4.28). We also obtain $\left|\frac{d}{dt}(P-Q)\right| \le C$, and so, because $(P-Q)(0)=0$,

$$|P-Q| \le Ct. \tag{2.4.35}$$

The bounds for H'' and f'' then yield $\left|\frac{d}{dt}(X-Y)\right| \le C(|X-Y|+1)$, and so an application of Gronwall's inequality gives

$$c|x-y| \le |X-Y| \le C|x-y|. \tag{2.4.36}$$

This implies that $\left|\frac{d}{dt}(P-Q)\right| \le C(|P-Q|+|x-y|)$, so that, with another application of Gronwall's inequality, we obtain

$$|P-Q| \le C|x-y|t. \tag{2.4.37}$$

In view of (2.4.28), the bounds (2.4.35) and (2.4.37) together give the second line of (2.4.29) after inverting the characteristics.

The matrices $\mathcal{A}_x$ and $\mathcal{A}_y$ are uniformly bounded in λ, which yields

$$\begin{cases} |\xi_i|+|\eta_i|+|\phi_i|+|\psi_i| \le C \quad \text{for } i=1,2, \text{ and} \\ |\xi_2|+|\eta_2|+|\phi_1|+|\psi_1| \le Ct. \end{cases} \tag{2.4.38}$$

Furthermore, the uniform bounds for $H^{(3)}$ and $f^{(3)}$, as well as the bound $|X-Y|+|P-Q| \le C|x-y|$, give $|\mathcal{A}_x-\mathcal{A}_y| \le C\left(|x-y|\wedge 1\right)$. The Duhamel formula applied to

$$\frac{d}{dt}\begin{pmatrix}\xi_i-\eta_i\\ \phi_i-\psi_i\end{pmatrix} = \mathcal{A}_x\begin{pmatrix}\xi_i-\eta_i\\ \phi_i-\psi_i\end{pmatrix} + \left(\mathcal{A}_x-\mathcal{A}_y\right)\begin{pmatrix}\eta_i\\ \psi_i\end{pmatrix}$$

can then be used to obtain the bound

$$|\xi_i - \eta_i| + |\phi_i - \psi_i| \le C\,(|x-y| \wedge 1)\,t. \tag{2.4.39}$$

Now define $J_1 := \xi_1\eta_1$ and $\Delta := \xi_2\eta_1 - \xi_1\eta_2$, so that

$$J = J_1 + \varphi_\lambda''(x-y)\Delta.$$

In view of (2.4.38) and the bound for H'',

$$\left|\dot{J}_1\right| = \left|\dot{\xi}_1\eta_1 + \xi_1\dot{\eta}_1\right| = \left|-h''(P)\phi_1\eta_1 - h''(Q)\eta_1\psi_1\right| \le C.$$

Next, observe that

$$|\Delta| \le |\xi_2 - \eta_2||\eta_1| + |\eta_1 - \xi_1||\eta_2| \le C|x-y|t,$$

while

$$\begin{aligned}\dot{\Delta} &= \dot{\xi}_2\eta_1 + \xi_2\dot{\eta}_1 - \dot{\xi}_1\eta_2 - \xi_1\dot{\eta}_2 = -H''(P)(\phi_2\eta_1 - \phi_1\eta_2) - H''(Q)(\xi_2\psi_1 - \xi_1\psi_2)\\ &= -H''(P)\left[(\phi_2\eta_1 - \xi_1\psi_2) - (\phi_1\eta_2 - \xi_2\psi_1)\right] + (H''(P) - H''(Q))(\xi_1\psi_2 - \xi_2\psi_1).\end{aligned}$$

The first term can be bounded by

$$\left|-H''(P)\left[(\phi_2\eta_1 - \xi_1\psi_2) - (\phi_1\eta_2 - \xi_2\psi_1)\right]\right| \le C|x-y|t,$$

and the second term by

$$\left|(H''(P) - H''(Q))(\xi_1\psi_2 - \xi_2\psi_1)\right| \le C|P-Q|t \le C|x-y|t.$$

Therefore $\left|\dot{\Delta}\right| \le C|x-y|t$, and we conclude that

$$\left|\dot{J}\right| \le \left|\dot{J}_1\right| + |\varphi_\lambda''(x-y)| \left|\dot{\Delta}\right| \le C + C\lambda|x-y|t \cdot \mathbf{1}_{\varphi_\lambda''(x-y)\neq 0} \le C\lambda^{1/2}t.$$

By the mean value theorem, there exists $s_\lambda \in [0,\tau_\lambda]$ such that $\dot{J}(s_\lambda)\tau_\lambda = -1/2$, and so

$$\frac{1}{2} = |\dot{J}(s_\lambda)|\tau_\lambda \le C\lambda^{1/2}\tau_\lambda^2,$$

which is the desired lower bound for τ_λ.

Differentiating (2.4.28) yields

$$\begin{pmatrix} P_x & -Q_x \\ P_y & -Q_y \end{pmatrix} = \begin{pmatrix} X_x & Y_x \\ X_y & Y_y \end{pmatrix} \begin{pmatrix} U_{xx} & U_{yx} \\ U_{xy} & U_{yy} \end{pmatrix}.$$

Because $|t| \le \tau_\lambda$, the first matrix on the right is invertible, so that

$$\begin{pmatrix} U_{xx} & U_{yx} \\ U_{xy} & U_{yy} \end{pmatrix} = \begin{pmatrix} X_x & Y_x \\ X_y & Y_y \end{pmatrix}^{-1} \begin{pmatrix} P_x & -Q_x \\ P_y & -Q_y \end{pmatrix} = J^{-1} \begin{pmatrix} Y_y & -Y_x \\ -X_y & X_x \end{pmatrix} \begin{pmatrix} P_x & -Q_x \\ P_y & -Q_y \end{pmatrix}.$$

Therefore, in view of (2.4.34),

$$\begin{aligned} J \begin{pmatrix} U_{xx} & U_{yx} \\ U_{xy} & U_{yy} \end{pmatrix} &= \begin{pmatrix} \eta_1 - \varphi_\lambda''\eta_2 & -\varphi_\lambda''\eta_2 \\ \varphi_\lambda''\xi_2 & \xi_1 + \varphi_\lambda''\xi_2 \end{pmatrix} \begin{pmatrix} \phi_1 + \varphi_\lambda''\phi_2 & -\varphi_\lambda''\psi_2 \\ -\varphi_\lambda''\phi_2 & -\psi_1 + \varphi_\lambda''\psi_2 \end{pmatrix} \\ &= \left(\begin{pmatrix} \eta_1 & 0 \\ 0 & \xi_1 \end{pmatrix} + \varphi_\lambda'' \begin{pmatrix} -\eta_2 & -\eta_2 \\ \xi_2 & \xi_2 \end{pmatrix} \right) \left(\begin{pmatrix} \phi_1 & 0 \\ 0 & -\psi_1 \end{pmatrix} + \varphi_\lambda'' \begin{pmatrix} \phi_2 & -\psi_2 \\ -\phi_2 & \psi_2 \end{pmatrix} \right) \\ &= \begin{pmatrix} \eta_1\phi_1 & 0 \\ 0 & -\xi_1\psi_1 \end{pmatrix} + \varphi_\lambda'' \left(\begin{pmatrix} -\phi_1\eta_2 & \psi_1\eta_2 \\ \phi_1\xi_2 & -\psi_1\xi_2 \end{pmatrix} + \begin{pmatrix} \phi_2\eta_1 & -\psi_2\eta_1 \\ -\xi_1\phi_2 & \xi_1\psi_2 \end{pmatrix} \right). \end{aligned}$$

Define

$$A := J \begin{pmatrix} U_{xx} & U_{yx} \\ U_{xy} & U_{yy} \end{pmatrix}.$$

For any fixed $v, w \in \mathbb{R}$, by writing

$$\begin{pmatrix} v \\ w \end{pmatrix} = \frac{1}{2} \begin{pmatrix} v+w \\ v+w \end{pmatrix} + \frac{1}{2} \begin{pmatrix} v-w \\ w-v \end{pmatrix},$$

we estimate, in view of (2.4.38), (2.4.39), and the Cauchy-Schwarz inequality,

$$\begin{aligned} A(v,w)\cdot(v,w) &\leq C\lambda|v-w|^2 + C(|v|^2+|w|^2)\phi_\lambda''(x-y)|x-y|^2 t \\ &\leq C\lambda|v-w|^2 + C(|v|^2+|w|^2)t, \end{aligned}$$

or, in other words,

$$A \leq C\lambda \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + Ct \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A similar computation yields

$$A^2 \leq C\lambda^2 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + Ct \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and this gives (2.4.30), since $t \in (-\tau_\lambda, \tau_\lambda)$ and $|J^{-1}| \leq 2$.

Finally, define

$$U^{\pm}(x,y,t) := (\lambda/2)|x-y|^2 \pm \|Df\|_\infty (|x-y|^2+\delta)^{1/2} t.$$

Then U^+ and U^- are respectively a smooth super- and sub-solution of (2.4.25). The comparison principle then gives

$$|U(x,y,t) - \varphi_\lambda(x-y)| \le \|Df\|_\infty (|x-y|^2 + \delta)^{1/2} t,$$

and (2.4.31) follows upon taking $\delta \to 0$. $\square$

CHAPTER 3

THE COMPARISON PRINCIPLE

3.1 Introduction

In this chapter, we prove the comparison principle for a variety of pathwise equations when the Hamiltonian depends on the spatial variable. More precisely, we prove that, if $u \in USC(\mathbb{R}^d \times (0,\infty))$ and $v \in LSC(\mathbb{R}^d \times (0,\infty))$ are respectively a sub- and super-solution bounded for bounded times, then, for all $t > 0$,

$$\sup_{x \in \mathbb{R}^d} (u(x,t) - v(x,t)) \le \sup_{x \in \mathbb{R}^d} (u(x,0) - v(x,0)) . \tag{3.1.1}$$

The comparison principle is a central part of the well-posedness theory for viscosity solutions. It immediately implies the uniqueness of solutions for initial value problems, and, with a variant of the proof, one can obtain several estimates regarding the stability of solutions with respect to perturbations in the data. This is done in Chapter 4 to prove certain pathwise stability estimates, and also in Chapter 5 to prove a pathwise error estimate for the approximation schemes studied there. Finally, the comparison principle is a necessary ingredient in the Perron method for constructing solutions, as is explained in Section 4.3.

The comparison principle was first proved for pathwise equations by Lions and Souganidis [54, 55, 56] when the Hamiltonian does not depend on space, that is, for equations of the form

$$du = F(D^2u, Du, u, x, t)\; dt + \sum_{i=1}^m H^i(Du) \cdot dW^i \quad \text{in } \mathbb{R}^d \times (0,\infty).$$

The key observation in this setting is that the standard penalizing quantity $(\lambda/2)|x-y|^2$ is a smooth, stationary solution of the doubled equation

$$d\Phi = \sum_{i=1}^m \Big(H^i(D_x\Phi) - H^i(-D_y\Phi) \Big) \cdot dW^i \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (0,\infty).$$

If H has nontrivial spatial dependence, this is no longer true, because the doubled equation becomes

$$d\Phi = \sum_{i=1}^{m} \Big(H^i(D_x\Phi, x) - H^i(-D_y\Phi, y) \Big) \cdot dW^i \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (0, \infty).$$

If H and W are sufficiently regular, then, as is standard in the literature, one can prove the comparison principle with the aid of smooth test functions of the form

$$\Phi(x, y, t) := \frac{\lambda}{2}|x - y|^2 \pm C \int_0^t \Big| \dot{W}(s) \Big| \, ds.$$

However, this strategy breaks down as soon as one considers paths with infinite variation.

In the first two sections of this chapter, we prove the comparison principle for certain spatially inhomogenous first-order equations. We introduce a new idea in the proof that deals with the difficulties that arise for equations set on unbounded domains, and is based on the equivalent formulation (Definition 2.3.2) for the definition of pathwise viscosity solutions. First, we consider Hamiltonians that are uniformly convex in the gradient variable, and we make use of the stationary solution of the doubled equation constructed in Section 2.4. We then move to a level-set equation that arises in the study of front propagation, and thus prove the uniqueness for the moving interface problem when the normal velocity is inhomogenous, anisotropic, and rough in time.

In the last section, we study second-order equations. First, we present a general parabolic maximum principle along the lines of [56], the ideas for which originate in works of Jensen [41] and Crandall and Ishii [24]. This is then used to prove the comparison principle for second order equations when the Hamiltonians have separated gradient and spatial dependence.

3.2 Spatially-dependent, uniformly convex Hamiltonians

We first prove the comparison principle for the equation

$$du = H(Du, x) \cdot dW \quad \text{in } \mathbb{R}^d \times (0, \infty), \qquad u(\cdot, 0) = u_0 \quad \text{on } \mathbb{R}^d \tag{3.2.1}$$

under the assumption that H satisfies the convexity and homogeneity assumptions of (2.4.14), which we reproduce here for convenience:

$$\begin{cases} p \mapsto H^*(p, \cdot) \text{ and } p \mapsto H(p, \cdot) \text{ are strictly convex in } \mathbb{R}^d \backslash \{0\}, \\ H^* \in C_b^2(B_R \backslash \overline{B_{1/R}} \times \mathbb{R}^d) \text{ for all } R > 1, \\ H^*(p, \cdot) > 0 \text{ for all } p \neq 0, \text{ and} \\ H^*(tp, y) = t^q H^*(p, y) \text{ for some } q > 1 \text{ and for all } (p, y) \in \mathbb{R}^d \times \mathbb{R}^d \text{ and } t > 0. \end{cases} \tag{3.2.2}$$

Recall also the definition of the distance function, whose properties are proved in Lemmas 2.4.3 and 2.4.5, defined by

$$L(x, y, \tau) := \inf \left\{ \int_0^\tau H^* \left(-\dot{\gamma}_s, \gamma_s \right) \, ds : \gamma \in \mathcal{A}(x, y, \tau) \right\}$$

where, for $x, y \in \mathbb{R}^d$ and $\tau > 0$,

$$\mathcal{A}(x, y, \tau) := \left\{ \gamma \in W^{1,\infty}([0, \tau], \mathbb{R}^d) : \gamma_0 = x, \; \gamma_\tau = y \right\}.$$

Define also $L(x, y) := L(x, y, 1)$ and recall that the homogeneity of H^* implies that

$$L(x, y, \tau) = \frac{1}{\tau^{q-1}} L(x, y).$$

Throughout the proof below, let $\omega : [0, \infty) \to [0, \infty)$ be the modulus of continuity for W on $[0, T]$.

Proof of the comparison principle for (3.2.1). We argue by contradiction and assume that there exist $T > 0$, $t_0 \in (0, T]$, sufficiently small $\mu > 0$, and sufficiently large $\lambda > 0$ such that the function

$$t \mapsto \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} \Big(u(x,t) - v(y,t) - \lambda^{q-1} L(x,y)\Big) - \mu t$$

achieves its maximum on $[0, T]$ at t_0. In view of Lemma 2.3.1(b), we may assume that $t_0 < T$.

For $\lambda > 0$, define

$$\Phi_\lambda(x, y, s, t) := L\left(x, y, \lambda^{-1} + W_s - W_t\right) = \left(\frac{\lambda}{1 + \lambda(W_s - W_t)}\right)^{q-1} L(x, y).$$

Then Lemma 2.4.5(a) and (b) yield that, whenever $(x, y) \in \Delta_{r_0}$ and $\omega(|s - t|) < \frac{1}{2\lambda}$,

$$(x, s) \mapsto \Phi_\lambda(x, y, s, t) \quad \text{and} \quad (y, t) \mapsto -\Phi_\lambda(x, y, s, t)$$

are C^1 in respectively x and y, and solve (3.2.1).

Let

$$M_0 := \sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \max\left(|u(x, t)|, |v(x, t|\right),$$

choose

$$\lambda > \frac{3}{2}\left(\frac{2M_0}{c_0 r_0^q}\right)^{\frac{1}{q-1}}, \tag{3.2.3}$$

where r_0 is as in Lemma 2.4.5(b), and, for $\theta > 0$, consider the auxiliary function

$$(s, t) \mapsto \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} \left(u(x, s) - v(y, t) - \Phi_\lambda(x, y, s, t)\right) - \frac{|s - t|^2}{2\theta} - \mu\frac{s + t}{2}, \tag{3.2.4}$$

which attains a maximum at some $(s_\theta, t_\theta) \in S_\lambda := \left\{(s, t) \in [0, T]^2 : \omega(|s - t|) \le (2\lambda)^{-1}\right\}$.

Then $\lim_{\theta\to 0}\frac{|s_\theta - t_\theta|^2}{2\theta} = 0$ and

$$\lim_{\theta\to 0}\sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}(u(x,s_\theta) - v(y,t_\theta) - \Phi_\lambda(x,y,s_\theta,t_\theta)) - \frac{|s_\theta - t_\theta|^2}{2\theta} - \mu\frac{s_\theta + t_\theta}{2}$$
$$= \max_{t\in[0,T]}\sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}\Big(u(x,t) - v(y,t) - \lambda^{q-1}L(x,y)\Big) - \mu t$$
$$= \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}\Big(u(x,t_0) - v(y,t_0) - \lambda^{q-1}L(x,y)\Big) - \mu t_0.$$

Therefore, as $\theta\to 0$, $(s_\theta, t_\theta)\to(t_0,t_0)$, and so, for sufficiently small θ, we have $s_\theta > 0$, $t_\theta > 0$, and $\omega(|s_\theta - t_\theta|) < \frac{1}{2\lambda}$.

We show that, for each fixed $y\in\mathbb{R}^d$,

$$s\mapsto\sup_{x\in\mathbb{R}^d}(u(x,s) - v(y,t_\theta) - \Phi_\lambda(x,y,s,t_\theta))$$

is nonincreasing in the interval $[a,b] := \big\{s\in[0,T] : \omega(|s-t_\theta|) < (2\lambda)^{-1}\big\}$. The same will then also be true for

$$s\mapsto\sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}(u(x,s) - v(y,t_\theta) - \Phi_\lambda(x,y,s,t_\theta))\,. \tag{3.2.5}$$

Indeed, for each fixed $s\in(a,b]$, we have $\omega(|\hat{s} - t_\theta|) < \frac{1}{2\lambda}$, and so, for all $x,y\in\mathbb{R}^d$,

$$c_0\left(\frac{2}{3}\lambda\right)^{q-1}|x-y|^q \le \Phi_\lambda(x,y,s,t_\theta),$$

which implies that $x\mapsto u(x,s) - \Phi_\lambda(x,y,s,t_\theta)$ attains a global maximum over $\mathbb{R}^d$. Moreover, for any such maximum point x, we have $|x-y| < r_0$, which is a consequence of (3.2.3) and

$$c_0\left(\frac{2}{3}\lambda\right)^{q-1}|x-y|^q \le u(x,s) - u(y,s) \le 2M_0.$$

Thus, in view of Lemma 2.4.5(b), Φ_λ is C^1 in the x-variable in a neighborhood of the maximum point. Definition 2.3.2 then yields the claim, and a similar argument for the super-solution v yields that

$$t \mapsto \inf_{x,y\in\mathbb{R}^d} \left(v(y,t) - u(x,s_\theta) + \Phi_\lambda(x,y,s_\theta,t)\right) \tag{3.2.6}$$

is nondecreasing on $[c,d] := \left\{t \in [0,T] : \omega(|s_\theta - t|) < (2\lambda)^{-1}\right\}$.

We now return to the maximum point (s_θ, t_θ) of (3.2.4). The map

$$s \mapsto \sup_{(x,y)\in\mathbb{R}^d} \left(u(x,s) - v(y,t_\theta) - \Phi_\lambda(x,y,s,t_\theta)\right) - \frac{|s-t_\theta|^2}{2\theta} - \mu\frac{s}{2}$$

attains a maximum at s_θ, and, since $\omega(|s_\theta - t_\theta|) < \frac{1}{2\lambda}$, we have $s_\theta > a$. Because (3.2.5) is nonincreasing, we have $\frac{\mu}{2} + \frac{s_\theta - t_\theta}{\theta} \leq 0$. Similarly, the map

$$t \mapsto \inf_{(x,y)\in\mathbb{R}^d} \left(v(y,t) - u(x,s_\theta) + \Phi_\lambda(x,y,s_\theta,t)\right) + \frac{|s_\theta - t|^2}{2\theta} + \mu\frac{t}{2}$$

attains a minimum at t_θ with $t_\theta > c$, so, because (3.2.6) is nondecreasing, we have $-\frac{\mu}{2} + \frac{s_\theta - t_\theta}{\theta} \geq 0$. We conclude that $\mu \leq 0$, a contradiction, and the result follows. □

3.3 Level-set equations

We now prove the comparison principle for the equation

$$du = a\left(\frac{Du}{|Du|}, x\right)|Du| \cdot dW \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d \tag{3.3.1}$$

when W is continuous and $a \in C_b^2(S^{d-1}, \mathbb{R}^d)$. Recall that this is the level-set equation corresponding to the motion of an interface evolving according to the normal velocity $a(n,x) \cdot dW$.

For ease of notation, we extend a to be 1-homogenous as follows:

$$a(p,x) := \begin{cases} a\left(\frac{p}{|p|}, x\right)|p| & \text{if } p \neq 0 \text{ and} \\ a(0,x) = 0. \end{cases}$$

The proof of the comparison principle for (3.3.1) shares many similarities with the argument in the previous section. However, since we do not assume convexity for a in the gradient variable, there is no analogue of the "distance" quantity used in that setting. Instead, we use the smooth solution U which was constructed in Lemma 2.4.6 with the method of characteristics.

Proof of the comparison principle for (3.3.1). Fix $\lambda, \delta, \mu, T > 0$ and let ϕ be as in (2.4.17):

$$\phi(x,y) := \frac{1}{2}\left[(|x-y|-\delta)_+\right]^2.$$

Consider the function

$$t \mapsto \sup_{x,y\in\mathbb{R}^d} (u(x,t) - v(y,t) - \lambda\phi(x,y)) - \mu t, \tag{3.3.2}$$

which attains a maximum at some $\hat{t} \in [0,T]$. Suppose for the sake of contradiction that $\hat{t} > 0$. We also assume that $\hat{t} < T$, which is justified by Lemma 2.3.1(b). Then let $0 < h < \hat{t}$ be such that

$$\max_{t\in[\hat{t}-h,\hat{t}]} \left|W(t) - W(\hat{t}-h)\right| < \tau,$$

where τ is the short time of existence for the solution U constructed in Lemma 2.4.6.

It is standard to check that, for δ and μ fixed, there exists $\omega_\lambda > 0$ satisfying $\lim_{\lambda\to\infty}\omega_\lambda = 0$ such that the supremum in (3.3.2) may be restricted to $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ satisfying

$$\lambda|x-y|^2 \leq \omega_\lambda.$$

Now define

$$\Phi_\lambda(x,y,s,t) := \lambda U(x,y,W(s)-W(\hat{t}-h),W(t)-W(\hat{t}-h)),$$

where U is the function constructed in Lemma 2.4.6.

The claim is that

$$t \mapsto \sup_{x,y\in\mathbb{R}^d} (u(x,t)-v(y,t)-\Phi_\lambda(x,y,t,t)) \quad \text{is nonincreasing on } [\hat{t}-h,\hat{t}]. \tag{3.3.3}$$

Step 1 of (3.3.3). In view of the estimate (2.4.19), for each fixed $s,t \in [\hat{t}-h,\hat{t}]$, the function $x \mapsto u(x,s)-\Phi(x,y,s,t)$ attains a maximum in $\mathbb{R}^d$. Therefore, Lemma 2.4.6 and Definition 2.3.2 imply that

$$s \mapsto \sup_{x\in\mathbb{R}^d} (u(x,s)-\Phi(x,y,s,t))$$

is nonincreasing. Similarly,

$$t \mapsto \sup_{y\in\mathbb{R}^d} (v(y,t)+\Phi(x,y,s,t))$$

is nondecreasing. Therefore,

$$(s,t) \mapsto \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} (u(x,s)-v(y,t)-\Phi(x,y,s,t))$$

is nonincreasing in both s and t.

Step 2 of (3.3.3). Suppose now that, for some smooth ψ,

$$t \mapsto \sup_{x,y\in\mathbb{R}^d} (u(x,t)-v(y,t)-\Phi_\lambda(x,y,t,t))-\psi(t)$$

attains a maximum at some $t^* \in (\hat{t} - h, \hat{t}]$. Then, for fixed $\beta > 0$,

$$(s,t) \mapsto \sup_{x,y\in\mathbb{R}^d} (u(x,s) - v(y,t) - \Phi_\lambda(x,y,s,t)) - \frac{|s-t|^2}{\beta} - \psi(s)$$

attains a maximum at some (s_β, t_β) such that $\lim_{\beta\to 0} s_\beta = \lim_{\beta\to 0} t_\beta = t^*$. In view of Step 1, this implies that $\psi'(s_\beta) \le 0$. Taking $\beta \to 0$ implies that $\psi'(s^*) \le 0$, which finishes the proof of the claim.

As a result of (3.3.3),

$$\begin{aligned} u(\hat{x},\hat{t}) - v(\hat{y},\hat{t}) - \Phi_\lambda(\hat{x},\hat{y},\hat{t},\hat{t}) &\le \max_{x,y\in\mathbb{R}^d} \big(u(x,\hat{t}-h) - v(y,\hat{t}-h) - \lambda\phi(x,y)\big) \\ &\le u(\hat{x},\hat{t}) - v(\hat{y},\hat{t}) - \lambda\phi(\hat{x},\hat{y}) - \mu h, \end{aligned}$$

so that rearranging terms and once more appealing to (2.4.19) yields

$$\mu h \le \lambda \left(\tilde{U}(\hat{x},\hat{y},W(\hat{t}) - W(\hat{t}-h)) - \phi(\hat{x},\hat{y})\right) \le C\lambda|\hat{x}-\hat{y}|^2\tau,$$

where $\tilde{U}$ is as in Lemma 2.4.6. For sufficiently large λ, this is a contradiction.

It follows that, for all large λ,

$$\sup_{x\in\mathbb{R}^d} (u(x,t) - v(x,t)) \le \sup_{x,y\in\mathbb{R}^d} (u(x,0) - v(y,0) - \lambda\phi(x,y)) + \mu t.$$

Sending first $\lambda \to +\infty$ and then $\mu \to 0$ shows that

$$\sup_{x\in\mathbb{R}^d} (u(x,t) - v(x,t)) \le \sup_{|x-y|\le\delta} (u(x,0) - v(y,0)),$$

and the proof is finished upon sending $\delta \to 0$.

□

3.4 Second order equations

In this last section, we present some tools for proving the comparison principle for second-order pathwise equations. First we prove a pathwise analogue of the so-called "Theorem of Sums" from the classical viscosity solution theory. We then use this to prove the comparison principle for equations where the Hamiltonians have a special separated form.

3.4.1 A pathwise "Theorem of Sums"

The next theorem is an analogue of the result in [24] (see also [25]) known as the "Theorem of Sums." Its ideas go back to [41], where the comparison principle for viscosity solutions of general, fully nonlinear elliptic partial differential equations was proved for the first time. In particular, we indirectly make use of Alexandroff's result [1] that convex functions are almost-everywhere twice-differentiable. In addition, the regularizing "sup- and inf-convolutions" known to experts who study viscosity solutions appear here in a modified form, due to the particular nature of the test functions specified in Definition 2.3.1. This modification is also used in the proof of the comparison principle in [56].

We consider the rough partial differential equation

$$du = F(D^2u, Du, u, x, t)dt + \sum_{i=1}^{m} H^i(Du, x) \cdot dW^i \quad \text{in } \mathbb{R}^d \times (0, T] \tag{3.4.1}$$

where $F \in C(\mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times [0, T])$ is degenerate elliptic, increasing in the u variable, and uniformly continuous for bounded (X, p, r); $H \in C^4(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^m)$; and $W \in \mathscr{C}_g^{\alpha}([0, T], \mathbb{R}^m)$ for some $\alpha \in (1/3, 1/2]$.

Theorem 3.4.1. *Assume that $u \in USC(\mathbb{R}^d \times [0, T])$ and $v \in LSC(\mathbb{R}^d \times [0, T])$ are respectively a sub- and supersolution of* (3.4.1)*, $(x_0, y_0, t_0) \in \mathbb{R}^d \times \mathbb{R}^d \times (0, T]$, $h \in (0, t_0)$,*

$\psi \in C^1((t_0 - h, t_0 + h))$, $\Phi \in C((t_0 - h, t_0 + h), C_b^2(\mathbb{R}^d))$ is a solution of

$$
(3.4.2) \qquad d\Phi = \sum_{i=1}^{m} \Big(H^i(D_x\Phi, x) - H^i(-D_y\Phi, y) \Big) \cdot dW^i \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (t_0 - h, t_0 + h),
$$

and

$$
u(x,t) - v(y,t) - \Phi(x,y,t) - \psi(t)
$$

achieves a local maximum at $(x_0, y_0, t_0) \in \mathbb{R}^d \times \mathbb{R}^d \times (t_0 - h, t_0 + h)$. Set $\phi := \Phi(\cdot, t_0)$. Then, for every $\varepsilon > 0$, there exist $X_\varepsilon, Y_\varepsilon \in \mathbb{S}^d$ such that

$$
(3.4.3) \qquad \begin{aligned} -\left(\left\| D^2\phi(x_0,y_0) \right\| + \frac{1}{\varepsilon} \right) \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} &\le \begin{pmatrix} X_\varepsilon & 0 \\ 0 & -Y_\varepsilon \end{pmatrix} \\ &\le D^2\phi(x_0,y_0) + \varepsilon (D^2\phi(x_0,y_0))^2 \end{aligned}
$$

and

$$
\psi'(t_0) \le F(X_\varepsilon, D_x\phi(x_0,y_0), u(x_0,t_0), x_0, t_0) - F(Y_\varepsilon, -D_y\phi(x_0,y_0), v(y_0,t_0), y_0, t_0).
$$

Notice that, if $\phi(x,y) = (\lambda/2)|x-y|^2$ and $\varepsilon = \lambda^{-1}$, then (3.4.3) becomes

$$
(3.4.4) \qquad -3\lambda \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \le \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le 3\lambda \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}.
$$

Proof of Theorem 3.4.1. Assume without loss of generality that $(x_0, y_0) = (0,0)$ and $\phi(0,0) = 0$. We may also assume, in view of Lemma 2.3.1, that $t_0 < T$ and the maximum is strict. Set $p := D_x\phi(0,0)$, $q := D_y\phi(0,0)$, $A := D^2\phi(0,0)$, and $a := \psi'(t_0)$, and fix $\delta > 0$. Then, for some $r > 0$ and for all $(x,y) \in B_r(0) \times B_r(0)$,

$$
\phi(x,y) \le p \cdot x + q \cdot y + \frac{1}{2}(A + \delta I)(x,y) \cdot (x,y).
$$

Step 1: Construction of the sup- *and* inf-*convolutions.* If $n \geq 1$, $\mathcal{A} \in \mathbb{S}^n$, and $X, \Xi \in \mathbb{R}^n$, then the Cauchy-Schwarz inequality implies, for every $\varepsilon > 0$, that

$$\begin{aligned}\frac{1}{2}\mathcal{A}X \cdot X - \frac{1}{2}\mathcal{A}\Xi \cdot \Xi &= \mathcal{A}\Xi \cdot (X - \Xi) + \frac{1}{2}\mathcal{A}(X - \Xi) \cdot (X - \Xi) \\ &\leq \frac{1}{2}\left(\varepsilon \left\|\mathcal{A}\Xi\right\|^2 + \left(\|\mathcal{A}\| + \frac{1}{\varepsilon}\right)|X - \Xi|^2\right),\end{aligned}$$

so that

$$\frac{1}{2}\mathcal{A}X \cdot X \leq \frac{1}{2}(\mathcal{A} + \varepsilon\mathcal{A}^2)\Xi \cdot \Xi + \frac{1}{2}\left(\|\mathcal{A}\| + \frac{1}{\varepsilon}\right)|X - \Xi|^2.$$

Letting $n = 2d$, $\mathcal{A} = A + \delta I := A_\delta$, $X = (x, y)$, and $\Xi = (\xi, \eta)$ above, and setting $A_{\delta,\varepsilon} = A_\delta + \varepsilon A_\delta^2$ yields

$$\begin{aligned}\phi(x, y) \leq p \cdot (x - \xi) + q \cdot (y - \eta) + \frac{1}{2}\left(\|A_\delta\| + \frac{1}{\varepsilon}\right)(|x - \xi|^2 + |y - \eta|^2) \\ + p \cdot \xi + q \cdot \eta + \frac{1}{2}A_{\delta,\varepsilon}(\xi, \eta) \cdot (\xi, \eta)\end{aligned}$$

for all $(x, y, \xi, \eta) \in B_r(0) \times B_r(0) \times \mathbb{R}^d \times \mathbb{R}^d$.

Now let $S_+(\cdot, t_0)$, $S_-(\cdot, t_0)$, and $S(\cdot, t_0)$ be the solution operators for respectively

$$dU = H(DU, x) \cdot dW, \quad dU = -H(-DU, x) \cdot dW, \quad \text{and} \quad dU = H(DU, x) - H(-DU, y).$$

If $f, g \in C^2(\mathbb{R}^d)$, $\tilde{\phi}(x, y) := f(x) + g(y)$, and

$$\begin{cases} S_+(\cdot, t_0)f,\ S_-(\cdot, t_0)g \in C((t_0 - h, t_0 + h), C_b^2(\mathbb{R}^d)) \quad \text{and} \\ S(\cdot, t_0)\tilde{\phi} \in C((t_0 - h, t_0 + h), C_b^2(\mathbb{R}^d \times \mathbb{R}^d)), \end{cases}$$

then, in view of the uniqueness of the spatially-smooth solutions constructed in Chapter 2, for all $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d \times (t_0 - h, t_0 + h)$,

$$S(t, t_0)\tilde{\phi}(x, y) = S_+(t, t_0)f(x) + S_-(t, t_0)g(y).$$

Shrinking h if necessary, define, for $(x, y, \xi, \eta, t) \in B_r(0) \times B_r(0) \times \mathbb{R}^d \times \mathbb{R}^d \times (t_0 - h, t_0 + h)$,

$$\begin{cases} \Phi_+(x, \xi, t) := S_+(t, t_0) \left(p \cdot (\cdot - \xi) + \frac{1}{2} \left(\|A_\delta\| + \frac{1}{\varepsilon} \right) |\cdot - \xi|^2 \right)(x), \\ \Phi_-(y, \eta, t) := S_-(t, t_0) \left(q \cdot (\cdot - \eta) + \frac{1}{2} \left(\|A_\delta\| + \frac{1}{\varepsilon} \right) |\cdot - \eta|^2 \right)(y), \\ \overline{u}(\xi, t) := \sup_{x \in B_r(0)} \left(u(x, t) - \Phi_+(x, \xi, t) \right), \quad \text{and} \\ \underline{v}(\eta, t) := \inf_{y \in B_r(0)} \left(v(y, t) + \Phi_-(y, \eta, t) \right). \end{cases}$$

It then follows that

$$\overline{u}(\xi, t) - \underline{v}(\eta, t) - p \cdot \xi - q \cdot \eta - \frac{1}{2} A_{\delta, \varepsilon}(\xi, \eta) \cdot (\xi, \eta) - a(t - t_0) - \frac{\delta}{2} |t - t_0|^2$$

attains a local maximum at $(\xi, \eta, t) = (0, 0, t_0)$ in $\mathbb{R}^d \times \mathbb{R}^d \times (t_0 - h, t_0 + h)$.

Step 2: Regularity of $\overline{u}$ *and* $\underline{v}$. Lemmas 2.2.2 and 2.2.5 yield that $D^2 \Phi_+$ and $D^2 \Phi_-$ are continuous on $\mathbb{R}^d \times \mathbb{R}^d \times (t_0 - h, t_0 + h)$, and, therefore,

$$\begin{cases} K_+(t) := \sup_{(x, \xi) \in B_r(0) \times \mathbb{R}^d} \left\| D^2_\xi \Phi_+(x, \xi, t) \right\| \quad \text{and} \\ K_-(t) := \sup_{(y, \eta) \in B_r(0) \times \mathbb{R}^d} \left\| D^2_\eta \Phi_-(y, \eta, t) \right\| \end{cases}$$

are continuous on $(t_0 - h, t_0 + h)$ with

$$K_+(0) = K_-(0) = \|A_\delta\| + \frac{1}{\varepsilon}.$$

Then $\overline{u}$ and $\underline{v}$ are respectively semiconvex and semiconcave in the spatial variable, with the following inequalities holding in the distributional sense:

$$D^2_\xi \overline{u}(\xi, t) \geq -K_+(t) I \quad \text{and} \quad D^2_\eta \underline{v}(\eta, t) \leq K_-(t) I.$$

Next, observe that, if h is sufficiently small, then, for all $t \in (t_0 - h, t_0 + h)$, the supremum and infimum in the definitions of respectively $\overline{u}$ and $\underline{v}$ are achieved for some $x(t)$ and $y(t)$ in $B_r(0)$. This is because, for $t = t_0$, the extrema are attained at respectively $x = 0$ and $y = 0$, and are strict by assumption.

Set

$$\begin{cases} K := \sup_{x \in B_r} \sup_{\xi \in \mathbb{R}^d} \sup_{t \in (t_0 - h, t_0 + h)} \left\{ \left| D_x^2 \Phi_+(x, \xi, t) \right| + \left| D_x \Phi_+(x, \xi, t) \right| \right\}, \\ M := \sup_{x \in \mathbb{R}^d} \sup_{t \in (t_0 - h, t_0 + h)} u(x, t), \quad \text{and} \\ \alpha_0 := \sup_{|X| + |p| \leq K} \sup_{r \leq M} \sup_{x \in \mathbb{R}^d} \sup_{t \in (t_0 - h, t_0 + h)} F(X, p, r, x, t), \end{cases}$$

fix $s \in (t_0 - h, t_0 + h)$ and $\xi \in \mathbb{R}^d$, and assume, for some $\mu > 0$, that $\overline{u}(\xi, t) - \alpha_0 t - \mu t$ attains a maximum at some $\bar{t} \in (s, t_0 + h]$. Then, for some $\bar{x} \in B_r(0)$,

$$u(x, t) - \Phi(x, \xi, t) - \alpha_0 t - \mu t$$

attains a local maximum at $(\bar{x}, \bar{t})$. Therefore, Definition 2.3.1 yields

$$\alpha_0 + \mu \leq F(D_x^2 \Phi(\bar{x}, \xi, \bar{t}), D_x \Phi(\bar{x}, \xi, \bar{t}), u(\bar{x}, \bar{t}), \bar{x}, \bar{t}) \leq \alpha_0.$$

This is a contradiction, and therefore, for all $\xi \in \mathbb{R}^d$, $s, t \in (t_0 - h, t_0 + h)$, and $\mu > 0$,

$$\overline{u}(\xi, t) - \overline{u}(\xi, s) \leq (\alpha_0 + \mu)(t - s).$$

It follows that, in the distributional sense,

$$\frac{\partial}{\partial t} \overline{u}(\xi, t) \leq \alpha_0.$$

Similarly, there exists $\beta_0 > 0$ such that

$$\frac{\partial}{\partial t}\underline{v}(\eta, t) \geq -\beta_0.$$

Step 3: A parabolic maximum principle. In view of the regularity proved in Step 2 for $\overline{u}$ and $\underline{v}$, we now have the following: there exist $\{(p_n, q_n, a_n)\}_{n=1}^{\infty} \subset \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$ and $\{(\xi_n, \eta_n, t_n)\}_{n=1}^{\infty} \subset \mathbb{R}^d \times \mathbb{R}^d \times (t_0 - h, t_0 + h)$ such that

$$\lim_{n\to\infty}(p_n, q_n, a_n, \xi_n, \eta_n, t_n) = (0, 0, 0, 0, 0, t_0),$$

$$\overline{u}(\xi, t) - \underline{v}(\eta, t) - (p + p_n) \cdot \xi - (q + q_n) \cdot \eta - \frac{1}{2}A_{\delta,\varepsilon}(\xi, \eta) \cdot (\xi, \eta) - (a_n + a)t - \frac{\delta}{2}|t - t_0|^2$$

attains a local maximum at (ξ_n, η_n, t_n), $\overline{u}$ and $\underline{v}$ are twice-differentiable in the spatial variable at respectively ξ_n and η_n, there exist α_n and β_n in respectively the super- and sub-differential of $\overline{u}(\xi_n, t_n)$ and $\underline{v}(\eta_n, t_n)$ such that

$$\alpha_n - \beta_n = a_n + a + \delta(t_n - t_0),$$

$$\begin{pmatrix} D_\xi \overline{u}(\xi_n, t_n) \\ -D_\eta \underline{v}(\eta_n, t_n) \end{pmatrix} = \begin{pmatrix} p + p_n \\ q + q_n \end{pmatrix} + A_{\delta,\varepsilon} \begin{pmatrix} \xi_n \\ \eta_n \end{pmatrix},$$

and

$$\text{(3.4.5)} \qquad \begin{pmatrix} -K_+(t_n)I & 0 \\ 0 & K_-(t_n)I \end{pmatrix} \leq \begin{pmatrix} D^2_\xi \overline{u}(\xi_n, t_n) & 0 \\ 0 & -D^2_\eta \underline{v}(\eta_n, t_n) \end{pmatrix} \leq A_{\delta,\varepsilon}.$$

These facts follow from the general parabolic maximum principle in [24] (see also Krylov [42] and Tso [80]).

Step 4: Construction of local-in-time, smooth-in-space solutions. Fix $\theta > 0$. Then, for some $r_n > 0$ and all $(\xi, \eta, t) \in B_{r_n}(\xi_n) \times B_{r_n}(\eta_n) \times (t_n - r_n, t_n]$,

$$\begin{cases} \overline{u}(\xi, t) \le \overline{u}(\xi_n, t_n) + (\alpha_n - \theta)(t - t_n) + D\overline{u}(\xi_n, t_n) \cdot (\xi - \xi_n) \\ \qquad + \frac{1}{2}(D^2\overline{u}(\xi_n, t_n) + \theta I)(\xi - \xi_n) \cdot (\xi - \xi_n) \quad \text{and} \\ \underline{v}(\eta, t) \ge \underline{v}(\eta_n, t_n) + (\beta_n + \theta)(t - t_n) + D\underline{v}(\eta_n, t_n) \cdot (\eta - \eta_n) \\ \qquad + \frac{1}{2}(D^2\underline{v}(\eta_n, t_n) - \theta I)(\eta - \eta_n) \cdot (\eta - \eta_n). \end{cases}$$

Now, if

$$\begin{cases} J_n^+(\xi) := D\overline{u}(\xi_n, t_n) \cdot (\xi - \xi_n) + \frac{1}{2}(D^2\overline{u}(\xi_n, t_n) + \theta I)(\xi - \xi_n) \cdot (\xi - \xi_n), \\ J_n^-(\eta) := D\underline{v}(\eta_n, t_n) \cdot (\eta - \eta_n) + \frac{1}{2}(D^2\underline{v}(\eta_n, t_n) - \theta I)(\eta - \eta_n) \cdot (\eta - \eta_n), \\ g^+(t) := (\alpha_n - \theta)t, \quad \text{and} \\ g^-(t) := (\beta_n + \theta)t, \end{cases}$$

it follows that, for some $x_n, y_n \in \overline{B_r(0)}$,

$$u(x, t) - \Phi_+(x, \xi, t) - J_n^+(\xi) - g^+(t) \tag{3.4.6}$$

attains a maximum in $\overline{B_r(0)} \times B_{r_n}(\xi_n) \times (t_n - r_n, t_n]$ at (x_n, ξ_n, t_n), and

$$v(y, t) + \Phi_-(y, \eta, t) - J_n^-(\eta) - g^-(t) \tag{3.4.7}$$

attains a minimum in $\overline{B_r(0)} \times B_{r_n}(\eta_n) \times (t_n - r_n, t_n]$ at (y_n, η_n, t_n).

Observe that $(x_n, y_n) \to (0, 0)$ as $n \to \infty$. Indeed, if $(\overline{x}, \overline{y})$ is an accumulation point of (x_n, y_n), then

$$u(x, t_0) - v(y, t_0) - p \cdot x - q \cdot y - \frac{1}{2}(\|A_\delta\| + \varepsilon^{-1})(|x|^2 + |y|^2)$$

attains a maximum in $\overline{B_r(0)} \times \overline{B_r(0)}$ at $(\overline{x}, \overline{y})$, and, therefore, $(\overline{x}, \overline{y}) = (0, 0)$. Thus, if n is large enough, x_n and y_n belong to $B_r(0)$.

Define

$$\begin{cases} \Phi_n^+(x,t) := \inf_{\xi \in B_{r_n}(\xi_n)} \left\{ J_n^+(\xi) + \Phi_+(x,\xi,t) \right\} \quad \text{and} \\ \Phi_n^-(y,t) := \sup_{\eta \in B_{r_n}(\eta_n)} \left\{ J_n^-(\eta) - \Phi_-(y,\eta,t) \right\}. \end{cases}$$

The claim is that, for (x,t) and (y,t) sufficiently close to respectively (x_n,t_n) and (y_n,t_n), the infimum and supremum above are attained uniquely at some $\xi_n(x,t) \in B_{r_n}(\xi_n)$ and $\eta_n(y,t) \in B_{r_n}(\eta_n)$. Indeed, these are attained if

$$\begin{cases} D_\xi J_n^+(\xi_n(x,t)) + D\Phi_+(x,\xi_n(x,t),t) = 0 \quad \text{and} \\ D_\eta J_n^-(\eta_n(x,t)) + D\Phi_-(y,\eta_n(x,t),t) = 0. \end{cases} \tag{3.4.8}$$

The claim then easily follows from the implicit function theorem, since

$$\begin{cases} D_\xi J_n^+(\xi_n) + D\Phi_+(x_n,\xi_n,t_n) = 0 \quad \text{and} \\ D_\xi^2 J_n^+(\xi_n) + D^2\Phi_+(x_n,\xi_n,t_n) \geq D^2\overline{u}(\xi_n,t_n) + \theta - K_+(t_n) \geq \theta > 0, \end{cases}$$

with a similar justification for Φ_-. Furthermore, $\xi_n(x_n,t_n) = \xi_n$, $\eta_n(y_n,t_n) = \eta_n$, and both functions are smooth in x and continuous in t. Since

$$\begin{cases} \Phi_n^+(x,t) = J_n^+(\xi_n(x,t)) + \Phi_+(x,\xi_n(x,t),t) \quad \text{and} \\ \Phi_n^-(y,t) = J_n^-(\eta_n(y,t)) - \Phi_-(y,\eta_n(y,t),t), \end{cases}$$

another computation shows that, shrinking r_n if necessary, Φ_n^+ and Φ_n^- and smooth-in-x-and-y solutions of

$$d\Phi = \sum_{i=1}^m H^i(D\Phi,x) \cdot dW^i$$

in respectively $B_{r_n}(x_n) \times (t_n - r_n, t_n]$ and $B_{r_n}(y_n) \times (t_n - r_n, t_n]$.

Step 5: Derivation of (3.4.3). Definition 2.3.1 yields

$$\begin{cases} \alpha_n - \theta \leq F(D^2\Phi_n^+(x_n,t_n), D\Phi_n^+(x_n,t_n)) \quad \text{and} \\ \beta_n + \theta \geq F(D^2\Phi_n^-(y_n,t_n), D\Phi_n^-(y_n,t_n)), \end{cases}$$

so that

$$\begin{aligned} a + a_n + \delta(t_n - t_0) - 2\theta \leq F(D^2\Phi_n^+(x_n,t_n), D\Phi_n^+(x_n,t_n), u(x_n,t_n), x_n, t_n) \\ - F(D^2\Phi_n^-(y_n,t_n), D\Phi_n^-(y_n,t_n), v(y_n,t_n), y_n, t_n). \end{aligned}$$

Differentiating Φ_n^+ gives

$$\begin{cases} D_x\Phi_n^+(x_n,t_n) = D_x\Phi_+(x_n,\xi_n,t_n) \quad \text{and} \\ D_x^2\Phi_n^+(x_n,t_n) = D_x^2\Phi_+(x_n,\xi_n,t) + D_{x\xi}^2\Phi_+(x_n,\xi_n,t_n)D_x\xi_n(x_n,t_n), \end{cases}$$

while differentiating (3.4.8) in x and evaluating at $(x,t) = (x_n,t_n)$ yields

$$\left(D^2J_n^+(\xi_n) + D_\xi^2\Phi_+(x_n,\xi_n,t_n)\right) D_x\xi_n(x_n,t_n) + D_{x\xi}^2\Phi_+(x_n,\xi_n,t_n) = 0.$$

Set $X_n := D_\xi^2\overline{u}(\xi_n,t_n)$ and $Y_n := D_\eta^2\overline{v}(\eta_n,t_n)$, and note that, in view of (3.4.5), as $n \to \infty$ along a subsequence, X_n and Y_n converge to some $X_\delta, Y_\delta \in \mathbb{S}^d$ satisfying

$$-\left(\|A_\delta\| + \frac{1}{\varepsilon}\right)\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X_\delta & 0 \\ 0 & -Y_\delta \end{pmatrix} \leq A_{\delta,\varepsilon}. \tag{3.4.9}$$

Along the same subsequence, as $n \to \infty$, it follows that $D_x\xi_n(x_n,t_n) \to \sigma$, where σ is given by

$$\left(X_\delta + \theta I + (\|A_\delta\| + \varepsilon^{-1})I\right)\sigma - (\|A_\delta\| + \varepsilon^{-1})I = 0.$$

Thus

$$\lim_{n\to\infty} D_x^2\Phi_n^+(x_n,t_n) = (\|A_\delta\| + \varepsilon^{-1})(X_\delta + \theta I)\left(X_\delta + \theta I + (\|A_\delta\| + \varepsilon^{-1})I\right)^{-1} \le X_\delta + \theta I,$$

and similarly,

$$\lim_{n\to\infty} D_x^2\Phi_n^-(y_n,t_n) \ge Y_\delta - \theta I.$$

This implies that

$$a - 2\theta \le F(X_\delta + \theta I, p, u(x_0,t_0), x_0, t_0) - F(Y_\delta - \theta I, -q, v(y_0,t_0)y_0, t_0),$$

so that letting $\theta \to 0$ yields

$$a \le F(X_\delta, p, u(x_0,t_0), x_0, t_0) - F(Y_\delta, -q, v(y_0,t_0), y_0, t_0).$$

Finally, from (3.4.9), it follows that, for some sequence δ_n decreasing to 0, X_{δ_n} and Y_{δ_n} converge, as $n \to \infty$, to respectively X and Y in $\mathbb{S}^d$, where X and Y satisfy (3.4.3). Taking the limit above for $\delta = \delta_n$ yields the desired inequality. □

3.4.2 Hamiltonians with separated dependence

We now use the Theorem of Sums from the previous subsection to prove the comparison principle for $\mathbb{Z}^d$-periodic sub- and super-solutions of the equation

$$du = F(D^2u, Du, u, x, t)\, dt + (H(Du) + f(x)) \cdot dW \quad \text{in } \mathbb{R}^d \times (0,\infty). \tag{3.4.10}$$

We assume that $H, f \in C^3(\mathbb{R}^d)$ satisfy (2.4.24) with f $\mathbb{Z}^d$-periodic, and $W \in C^{0,\alpha}$ for some $\alpha \ge 1/3$. We work on the torus to avoid some of the difficulties that come up in the setting of unbounded domains.

The assumptions for the nonlinearity F are more complicated. Matrix inequalities of the form below appear in the standard viscosity solution theory, as in the work of Ishii [38] (see also Chapter 2 of [25]). We assume that

$$\begin{cases} F \in C(\mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{T}^d \times [0,\infty)), \\ X \mapsto F(X,p,r,x,t) \quad \text{is continuous uniformly in } (p,r,x,t), \end{cases} \tag{3.4.11}$$

$$r \mapsto F(X,p,r,x,t) \quad \text{is nonincreasing,} \tag{3.4.12}$$

and

$$\begin{cases} \text{for any } C>0, \text{ there exists } \omega : [0,\infty) \to [0,\infty) \text{ satisfying } \lim_{s\to 0+} \omega(s) = 0 \\ \text{such that, whenever } \lambda > 0,\ X, Y \in \mathbb{S}^d, \text{ and} \\ -C\lambda \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \le \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le C\lambda \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}, \\ \text{then, for all } x,y \in \mathbb{T}^d,\ p,q \in \mathbb{R}^d,\ r \in \mathbb{R}, \text{ and } t \in [0,\infty), \\ F(Y,p,r,y,t) - F(X,q,r,x,t) \\ \qquad \le \omega\left(\lambda(|x-y|^2 + |p-q|^2) + (1+|p|+|q|)(|x-y|+|p-q|)\right). \end{cases} \tag{3.4.13}$$

Notice that (3.4.13) implies that F is degenerate elliptic. Indeed, (3.4.13) holds whenever F is independent of (p,x) and degenerate elliptic, or, more generally, if F is of the form

$$F(X,p,r,x,t) = F_1(X,r,t) + F_0(p,x,r,t)$$

where $F_1 \in C(\mathbb{S}^d \times \mathbb{R} \times [0,\infty))$, $F_0 \in C(\mathbb{R}^d \times \mathbb{T}^d \times \mathbb{R} \times [0,\infty)$, F_1 is degenerate elliptic and uniformly continuous in the $\mathbb{S}^d$-variable, F_0 and F_1 are nonincreasing in the $\mathbb{R}$-variable, and, for some $\omega : [0,\infty) \to [0,\infty)$ satisfying $\lim_{s\to 0^+} \omega(s)$,

$$|F_0(p,x,r,t) - F_0(q,y,r,t)| \le \omega((1+|p|+|q|)(|x-y|+|p-q|))$$
$$\text{for all} \quad (p,q,x,y,r,t) \in \mathbb{R}^{2d} \times \mathbb{T}^{2d} \times \mathbb{R} \times [0,T].$$

The requirement in (3.4.13) is general enough to cover quasilinear equations with more involved spatial and gradient dependence. For instance, $F : \mathbb{S}^d \times \mathbb{R}^d \times \mathbb{T}^d \to \mathbb{R}$ satisfies (3.4.11), (3.4.12), and (3.4.13) if F is given by

$$F(X,p,x) := \sum_{i,j=1}^{d} \sum_{k=1}^{m} \sigma_{ik}(p,x)\sigma_{jk}(p,x)X_{ij} \quad \text{for some } \sigma \in C^{0,1}(\mathbb{R}^d \times \mathbb{T}^d, \mathbb{R}^{d\times m}).$$

Many more examples can be constructed by observing that, if $\{F_{\alpha\beta}\}_{\alpha\in A,\beta\in B}$ is a family of functions satisfying (3.4.11), (3.4.12), and (3.4.13) uniformly in α and β, then the same is true for

$$\inf_{\alpha\in A}\sup_{\beta\in B} F_{\alpha\beta} \quad \text{and} \quad \sup_{\beta\in B}\inf_{\alpha\in A} F_{\alpha\beta}.$$

The assumption that W is at least 1/3-Hölder continuous is not related to rough-path theory, and indeed, it is not required that W have a geometric rough-path lift. It turns out that there is a precise interplay between the Hölder regularity of the path and the time τ_λ of existence for the smooth solution constructed in Lemma 2.4.7.

Proof of the comparison principle for (3.4.10). Fix $T > 0$ and assume without loss of generality that $u(\cdot,0) \le v(\cdot,0)$. Suppose, for the sake of contradiction that there exists $\mu > 0$ and $(x_0,t_0) \in \mathbb{T}^d \times (0,T]$ such that

$$u(x_0,t_0) - v(x_0,t_0) - \mu t_0 = \max_{(x,t)\in\mathbb{T}^d\times[0,T]} (u(x,t) - v(x,t) - \mu t) > 0.$$

As before, we may assume that $t_0 < T$ without loss of generality.

For $\lambda > 0$, the function

$$u(x,t) - v(y,t) - \frac{\lambda}{2}|x-y|^2 - \mu t$$

attains a maximum at $(x_\lambda, y_\lambda, t_\lambda) \in \mathbb{T}^d \times \mathbb{T}^d \times [0,T]$. Standard arguments from the viscosity solution theory imply that, as $\lambda \to 0$,

$$(3.4.14) \quad \begin{cases} (x_\lambda, y_\lambda, t_\lambda) \to (x_0, x_0, t_0), \\ u(x_\lambda, t_\lambda) - v(y_\lambda, t_\lambda) - \frac{\lambda}{2}|x_\lambda - y_\lambda|^2 - \mu t_\lambda \to u(x_0,t_0) - v(x_0,t_0) - \mu t_0, \quad \text{and} \\ \frac{\lambda}{2}|x_\lambda - y_\lambda|^2 \to 0. \end{cases}$$

In particular, for all large λ, $t_\lambda \in (0,T)$ and $u(x_\lambda, t_\lambda) > v(y_\lambda, t_\lambda)$, and, for some $R > 0$, $(\lambda/2)|x_\lambda - y_\lambda|^2 \leq R$ and

$$\Psi(x,y,t) := u(x,t) - v(y,t) - \varphi_{\lambda,R}(x-y) - \mu t$$

attains a global maximum in $\mathbb{R}^d \times \mathbb{R}^d \times [0,T]$ at $(x_\lambda, y_\lambda, t_\lambda)$, where

$$\varphi_{\lambda,R}(z) = \varphi_\lambda = g((\lambda/2)|x-y|^2)$$

is the cutoff function from subsection 2.4.4 (the R in that section agrees with the R here).

For all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $h > 0$, $\Psi(x,y,t_\lambda - h) \leq \Psi(x_\lambda, y_\lambda, t_\lambda)$, and therefore

$$(3.4.15) \quad \begin{aligned} &\sup_{x,y\in\mathbb{R}^d} \{u(x, t_\lambda - h) - v(y, t_\lambda - h) - \varphi_\lambda(x-y)\} \\ &\qquad \leq u(x_\lambda, t_\lambda) - v(y_\lambda, t_\lambda) - \varphi_\lambda(x-y) - \mu h. \end{aligned}$$

Recall that Lemma 2.4.7 implies that, for some $c > 0$ depending on the bounds in (2.4.24) and τ_λ satisfying $\tau_\lambda \geq c\lambda^{-1/4}$, there exists a C^2-solution U of

$$\begin{cases} U_t = H(D_x U) + f(x) - H(-D_y U) - f(y) & \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (-\tau_\lambda, \tau_\lambda) \quad \text{and} \\ U(x,y,0) = \varphi_\lambda(x-y) & \text{in } \mathbb{R}^d \times \mathbb{R}^d. \end{cases} \tag{3.4.16}$$

It then follows that, for some constant $C > 0$ depending only on α, the bounds in (2.4.24), and the α-seminorm of W on $[0,T]$, if $h_\lambda = C\lambda^{-\frac{1}{4\alpha}}$, then

$$|W(s) - W(t)| < \tau_\lambda \quad \text{for any } s,t \in [0,T] \text{ with } |s-t| \leq h_\lambda.$$

For sufficiently large λ, $t_\lambda - h_\lambda > 0$, and so $\Phi(x,y,t) := U(x,y,W(t) - W(t_\lambda - h_\lambda))$ belongs to $C([t_\lambda - h_\lambda, t_\lambda], C_b^2(\mathbb{R}^d \times \mathbb{R}^d))$ and is a solution of

$$d\Phi = \big(H(D_x\Phi) + f(x) - H(-D_y\Phi) - f(y)\big) \cdot dW \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times [t_\lambda - h_\lambda, t_\lambda].$$

Define

$$U_\lambda = U_{\lambda,R} := \left\{ (x,y) \in \mathbb{R}^d \times \mathbb{R}^d : \frac{\lambda}{2}|x-y|^2 \leq R \right\}.$$

We claim there exists $\rho : [0,\infty) \to [0,\infty)$ satisfying $\lim_{\lambda \to \infty} \rho(\lambda) = 0$ such that

$$\begin{aligned} &\sup_{(x,y) \in U_\lambda} (u(x,t) - v(y,t) - \Phi(x,y,t)) \\ &\qquad \leq \sup_{x,y \in \mathbb{R}^d} \left(u(x, t_\lambda - h_\lambda) - v(y, t_\lambda - h_\lambda) - \frac{\lambda}{2}|x-y|^2 \right) + \rho(\lambda) h_\lambda. \end{aligned} \tag{3.4.17}$$

Fix $\nu > 0$ and let $(x^\lambda, y^\lambda, t^\lambda)$ be a maximum point in $U_\lambda \times [t_\lambda - h_\lambda, t_\lambda]$ for

$$u(x,t) - v(y,t) - \Phi(x,y,t) - \nu t.$$

Then (3.4.17) will follow if we can show that $t^\lambda > t_\lambda - h_\lambda$ implies that $\nu \leq \rho(\lambda)$ for some ρ as in the claim.

The estimate (2.4.31) gives

$$|\Phi(x,y,t)-\varphi_\lambda(x-y)| \le C|x-y|\,|W(t)-W(t_\lambda-h_\lambda)| \le C\lambda^{-1/4}|x-y|,$$

and so

$$\begin{aligned} u(x,t)-v(y,t)-\varphi_\lambda(x-y)-C\lambda^{-1/4}|x-y| \\ &\le u(x,t)-v(y,t)-\Phi(x,y,t) \\ &\le u(x,t)-v(y,t)-(\lambda/2)|x-y|^2+C\lambda^{-1/4}|x-y|. \end{aligned}$$

Similar arguments as the ones for (3.4.14) imply that $\lim_{\lambda\to\infty}(\lambda/2)|x^\lambda-y^\lambda|=0$ uniformly in $\nu>0$. In particular, for large enough λ, (x^λ,y^λ) is in the interior of the set U_λ, and $u(x^\lambda,t^\lambda)>v(y^\lambda,t^\lambda)$.

For fixed $\varepsilon>0$, Theorem 3.4.1 gives $X_\lambda, Y_\lambda \in \mathbb{S}^d$ satisfying

$$(3.4.18)\qquad \begin{aligned} -\left(\left\|D^2\Phi(x^\lambda,y^\lambda,t^\lambda)\right\|+\frac{1}{\varepsilon}\right)\begin{pmatrix} I & 0\\ 0 & I\end{pmatrix} \le \begin{pmatrix} X_\lambda & 0\\ 0 & -Y_\lambda\end{pmatrix} \\ \le D^2\Phi(x^\lambda,y^\lambda,t^\lambda)+\varepsilon(D^2\Phi(x^\lambda,y^\lambda,t^\lambda))^2 \end{aligned}$$

and

$$(3.4.19)\qquad \begin{aligned} \nu &\le F(X_\lambda, D_x\Phi(x^\lambda,y^\lambda,t^\lambda), u(x^\lambda,t^\lambda), x^\lambda, t^\lambda) \\ &\quad - F(Y_\lambda, -D_y\Phi(x^\lambda,y^\lambda,t^\lambda), v(y^\lambda,t^\lambda), y^\lambda, t^\lambda)) \\ &\le F(X_\lambda, D_x\Phi(x^\lambda,y^\lambda,t^\lambda), v(y^\lambda,t^\lambda), x^\lambda, t^\lambda) \\ &\quad - F(Y_\lambda, -D_y\Phi(x^\lambda,y^\lambda,t^\lambda), v(y^\lambda,t^\lambda), y^\lambda, t^\lambda)), \end{aligned}$$

where we have used (3.4.12). Choosing $\varepsilon=\lambda^{-1}$ above and applying the bounds in (2.4.30) from Lemma 2.4.7, (3.4.18) becomes, for some constant $C>0$ depending only on the bounds in (2.4.24),

$$-C\lambda \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X_\lambda & 0 \\ 0 & -Y_\lambda \end{pmatrix} \leq C\lambda \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + C\lambda^{-1/4} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

We now use the bound (2.4.29) from Lemma 2.4.7, which implies

$$|D_x\Phi(x^\lambda, y^\lambda, t^\lambda) + D_y\Phi(x^\lambda, y^\lambda, t^\lambda)| \leq Ch_\lambda^\alpha \leq C\lambda^{-1/4}, \tag{3.4.20}$$

as well as the assumptions (3.4.11) and (3.4.13), to conclude from (3.4.19) that, for some $\omega : [0,\infty) \to [0,\infty)$ satisfying $\lim_{s\to 0^+} \omega(s)$,

$$\nu \leq \omega\left(\lambda|x^\lambda - y^\lambda|^2 + |x^\lambda - y^\lambda| + \lambda^{-1/4}\right) := \rho(\lambda),$$

and thus (3.4.17) is established.

Finally, combining (2.4.31) and (3.4.17) gives

$$\begin{aligned} \mu \leq \frac{\Phi(x_\lambda, y_\lambda, t_\lambda) - (\lambda/2)|x_\lambda - y_\lambda|^2}{h_\lambda} &\leq C|x_\lambda - y_\lambda| h_\lambda^{\alpha-1} + \rho(\lambda) \\ &\leq C\lambda^{1/2}|x_\lambda - y_\lambda| \cdot \lambda^{\frac{1-3\alpha}{4\alpha}} + \rho(\lambda), \end{aligned}$$

and so we obtain a contradiction for sufficiently large λ in view of (3.4.14). □

CHAPTER 4

EXISTENCE

4.1 Introduction

In this section, we address the issue of existence of solutions for the initial value problem (1.1.1).

The strategy for proving existence in this setting has been to regularize the paths and extract limits, as described in [54, 55, 53, 76]. More precisely, the program has been as follows:

1. For a fixed path $W \in \mathscr{C}_g^\alpha([0,T], \mathbb{R}^m)$, choose any two smooth families of paths $\{\zeta^n\}_{n=1}^\infty$ and $\{\eta^n\}_{n=1}^\infty$ that both converge, as $n \to \infty$, to W in the rough path metric.

2. Let u^n and v^n be the classical viscosity solutions of (1.1.1) with respectively the paths ζ^n and η^n and prove that

$$\lim_{n\to\infty} \sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u^n(x,t) - v^n(x,t)| = 0.$$

This shows that the solution operator for (1.1.1) has a unique, continuous extension to the space of geometric rough paths.

Accomplishing step 2 above is done through comparison-principle type arguments, and depends on the stability of the rough, Hamilton-Jacobi part (2.1.1) of the equation under perturbations in the rough-path metric. Note that, when (2.2.7) holds, which is the case if, for example, $m = 1$ or the Hamiltonians do not depend on space, then the smooth approximations ζ^n and η^n need only converge to W locally uniformly.

When this argument is made quantitative, one can obtain precise stability estimates for the solution operator. One setting in which this has been done is for first-order equations when the Hamiltonians do not depend on space, as was proved in [55]:

Theorem 4.1.1. *Assume that, for each* $i = 1, 2, \ldots, m$, $H^i : \mathbb{R}^d \to \mathbb{R}$ *satisfies* (2.4.4), $u_0 \in \mathrm{Lip}(\mathbb{R}^d)$ *with* $\mathrm{Lip}(u_0) \le L$, $W^1, W^2 \in C([0,\infty), \mathbb{R}^m)$, *and* u^1 *and* u^2 *solve, for* $j = 1, 2$,

$$\begin{cases} du^j = \displaystyle\sum_{i=1}^m H^i(Du^j) \cdot dW^i & \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \\ u^j(\cdot, 0) = u_0 & \text{in } \mathbb{R}^d. \end{cases}$$

Then, for all $T > 0$ *and some* $C = C_L > 0$,

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \left| u^1(x,t) - u^2(x,t) \right| \le C \max_{t\in[0,T]} \left| W^1(t) - W^2(t) \right|.$$

The first goal of this section is to extend Theorem 4.1.1 to the setting where the Hamiltonian has nontrivial spatial dependence. When $m = 1$ and H is convex in the gradient variable, we obtain uniform Lipschitz bounds and path-stability estimates that depend only on the growth of the Hamiltonian. This gives a unique extension of the solution operator to continuous paths. It also allows for the study of other asymptotic problems for the equation, and we use the estimates to investigate certain homogenization problems in Chapter 6.

As has already been discussed in the introduction to Chapter 3, additional difficulties arise in the pathwise viscosity solution theory for spatially dependent Hamiltonians. The main tool for dealing with these is the "distance function" constructed in Section 2.4.

We then present a general existence proof for equations of the form

$$\begin{cases} du = F(D^2u, Du, u, x, t)\, dt + \displaystyle\sum_{i=1}^m H^i(Du, x) \cdot dW^i & \text{in } \mathbb{R}^d \times (0, T], \\ u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d. \end{cases} \tag{4.1.1}$$

Our objective is to use the comparison principle to prove directly that the maximal subsolution defines the unique solution of (4.1.1). This method of establishing the existence of solutions goes back to Perron [71], who used it to solve the Dirichlet problem for the Laplace equation. Ishii [37] then generalized the strategy to the setting of viscosity solutions of Hamilton-Jacobi equations (see also Section 4 of [25]). This gave an elegant alternative to proving the existence of solutions for singular, first-order equations, as opposed to the method of vanishing viscosity, which was used before by, among others, Lions [51]. Similarly, the Perron construction for (4.1.1) presented here indicates that the solution theory put forth by Definition 2.3.1 is powerful enough to avoid the consideration of auxiliary equations driven by smooth paths.

The Perron construction consists of two steps. First, we prove a different kind of stability for (4.1.1), namely, the stability of sub-solutions under pointwise suprema. We also construct suitable barriers to ensure that the maximal sub-solution is actually a super-solution, and that it achieves the desired initial data. The analysis is more involved than for the classical viscosity solution setting, due to the rough nature of the test functions.

4.2 Path stability estimates for convex Hamiltonians

Assume that $H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies

$$(4.2.1) \qquad \begin{cases} H \in C(\mathbb{R}^d \times \mathbb{R}^d),\ p \mapsto H(p,x) \text{ is convex for all } x \in \mathbb{R}^d, \text{ and} \\ \text{there exist convex, increasing functions } \underline{\nu}, \overline{\nu} : [0,\infty) \to \mathbb{R} \text{ such that} \\ \underline{\nu}(|p|) \le H(p,x) \le \overline{\nu}(|p|) \quad \text{for all } (p,x) \in \mathbb{R}^d \times \mathbb{R}^d, \end{cases}$$

and, for two smooth (or piecewise smooth) paths $W^1, W^2 : [0,\infty) \to \mathbb{R}$ and $u_0^1, u_0^2 \in \mathrm{Lip}(\mathbb{R}^d)$, consider the viscosity solutions u^1 and u^2 of

$$u_t^j = H(Du^j, x)\dot{W}^j \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u^j(\cdot,0) = u_0^j \quad \text{in } \mathbb{R}^d.$$

Theorem 4.2.1. *Set* $L := \max\left(\mathrm{Lip}(u_0^1), \mathrm{Lip}(u_0^2)\right)$. *Then, for all* $t > 0$ *and for* $j = 1, 2$,

$$\mathrm{Lip}(u^j(\cdot,t)) \le \underline{\nu}^{-1}\left(\overline{\nu}(L)\right),$$

and, for all $T > 0$*,*

$$\max_{(x,t)\in\mathbb{R}^d\times[0,T]} \left|u^1(x,t) - u^2(x,t)\right| \le \max_{x\in\mathbb{R}^d} \left|u_0^1(x) - u_0^2(x)\right| + \overline{\nu}(L) \max_{t\in[0,T]} \left|W^1(t) - W^2(t)\right|$$
$$+ \underline{\nu}(0)_- \left(\max_{t\in[0,T]} \left|W^1(t) - W^2(t)\right| - (W^1(T) - W^2(T)) \right).$$

As a consequence, the solution operator for the equation

$$du = H(Du, x) \cdot dW \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u(\cdot, 0) = u_0 \tag{4.2.2}$$

extends continuously to the space of continuous paths. That is, if $\{W^n\}_{n=1}^{\infty}$ is a family of smooth paths approximating W uniformly as $n \to \infty$, and if u^n is the solution of (4.2.2) with the path W^n, then u^n converges uniformly in $\mathbb{R}^d \times [0,T]$ to a unique limit $u \in BUC(\mathbb{R}^d \times [0,T])$ as $n \to \infty$, where u is independent of the choice of approximations. Moreover, if the comparison principle holds for (4.2.2) with W continuous, then the limiting function u is the unique pathwise viscosity solution of (4.2.2). As we proved in Chapter 3, this is the case if H has the strict convexity and homogeneity specified by (2.4.14), or if

$$H(p,x) := a\left(\frac{p}{|p|}, x\right)|p|$$

with $a \in C_b^2(S^{d-1}, \mathbb{R}^d)$ and such that H is convex.

The notable feature of Theorem 4.2.1 is that the estimates do not depend on the regularity of H, but rather, only on the rate of growth of H in the gradient variable. We exploit this fact in Chapter 6 to study some equations where the dependence of the Hamiltonian on the spatial variable is highly oscillatory.

Both results in Theorem 4.2.1 follow from the next proposition, which is proved with the distance function from Section 2.4 that was already used for the comparison principle in Section 3.2. Its hypotheses require more regularity for the Hamiltonian than is specified by (4.2.1). The proof of Theorem 4.2.1 then involves a further regularization of H, and the result will follow upon obtaining estimates that do not depend on the regularization parameter.

Proposition 4.2.1. *Assume that* H *satisfies* (4.2.1),

$$\begin{cases} H \in C_b^2(B_R \times \mathbb{R}^d) \quad \textit{for all } R > 0, \textit{ and} \\ H \textit{ is uniformly convex.} \end{cases} \tag{4.2.3}$$

For $u_0, v_0 \in BUC(\mathbb{R}^d)$ *and* $\zeta, \eta \in C^1([0,\infty))$ *with* $\zeta_0 = \eta_0$, *let* u *be a sub-solution of*

$$u_t = H(Du, x)\dot{\zeta}(t) \quad \textit{in } \mathbb{R}^d \times (0,\infty), \qquad u(\cdot, 0) = u_0 \quad \textit{in } \mathbb{R}^d,$$

and v *a super-solution of*

$$v_t = H(Dv, x)\dot{\eta}(t) \quad \textit{in } \mathbb{R}^d \times (0,\infty), \qquad v(\cdot, 0) = v_0 \quad \textit{in } \mathbb{R}^d.$$

Then, for all $T > 0$ *and* $0 < \lambda < \left(\max_{0 \le t \le T}(\zeta_t - \eta_t)_-\right)^{-1}$,

$$\sup_{(x,y,t) \in \mathbb{R}^d \times \mathbb{R}^d \times [0,T]} \left(u(x,t) - v(y,t) - \left(\frac{1}{\lambda} + \zeta_t - \eta_t\right) \underline{\nu}^* \left(\frac{\lambda|x-y|}{1 + \lambda(\zeta_t - \eta_t)} \right) \right) \le \sup_{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d} \left(u_0(x) - v_0(y) - \frac{1}{\lambda} \overline{\nu}^*(\lambda|x-y|) \right).$$

Equipped with Proposition 4.2.1, we proceed with the

Proof of Theorem 4.2.1. Step 1. Assume first that H satisfies (4.2.3) in addition to (4.2.1). Applying Proposition 4.2.1 to the case $u = v = u^1$ and $\zeta = \eta = W^1$ yields, for all $(x,y,t) \in \mathbb{R}^d \times \mathbb{R}^d \times (0,\infty)$,

$$\begin{aligned} u^1(x,t) - u^1(y,t) &\le \inf_{\lambda > 0} \left\{ \frac{1}{\lambda} \underline{\nu}^*(\lambda|x-y|) + \sup_{s \ge 0} \left\{ Ls - \frac{1}{\lambda} \overline{\nu}^*(\lambda s) \right\} \right\} \\ &= \inf_{\lambda > 0} \left\{ \frac{\underline{\nu}^*(\lambda|x-y|) + \overline{\nu}(L)}{\lambda} \right\} = \underline{\nu}^{-1}\left(\overline{\nu}(L)\right)|x-y|. \end{aligned}$$

Thus $\mathrm{Lip}(u^1(\cdot,t)) \le \underline{\nu}^{-1}\left(\overline{\nu}(L)\right)$, and similarly for u^2.

Now setting $(u,v,\zeta,\eta) := (u^1,u^2,W^1,W^2)$ in Proposition 4.2.1 gives

$$u^1(x,t) - u^2(x,t) \le \left(\frac{1}{\lambda} - (W_t^1 - W_t^2)\right)\underline{\nu}^*(0) + \max_{x\in\mathbb{R}^d}\left|u_0^1(x) - u_0^2(x)\right| + \frac{1}{\lambda}\overline{\nu}(L).$$

The claim follows upon choosing $\lambda = (\max_{s\in[0,t]}\left|W_s^1 - W_s^2\right|)^{-1}$ and using the fact that

$$\underline{\nu}^*(0) = -\min_{r\ge 0}\underline{\nu}(r) = -\underline{\nu}(0) \le \underline{\nu}(0)_-.$$

Step 2. We now return to the general case, where H satisfies only (4.2.1). Let $\phi \in C^2(\mathbb{R}^d)$ be nonnegative and supported in $B_1(0)$ with $\int\phi = 1$, and, for $\rho > 0$, define

$$\phi_\rho(z) := \frac{1}{\rho^d}\phi\left(\frac{z}{\rho}\right)$$

and

$$H_\rho(p,x) := \rho|p|^2 + \iint_{\mathbb{R}^d\times\mathbb{R}^d} H(q,y)\phi_\rho(p-q)\phi_\rho(x-y)\,dq\,dy.$$

It is straightforward to verify that $\lim_{\rho\to 0} H_\rho = H$ locally uniformly, and H_ρ satisfies both (4.2.1) and (4.2.3) with

$$\begin{cases} \overline{\nu}_\rho(s) := \rho s^2 + \overline{\nu}(s+\rho) \quad \text{and} \\ \underline{\nu}_\rho(s) := \rho s^2 + \underline{\nu}\left((s-\rho)_+\right). \end{cases}$$

Let u_ρ^1 and u_ρ^2 be as in the statement of Theorem 4.2.1 for the Hamiltonian H_ρ. As proved above, u_ρ^1 and u_ρ^2 satisfy the Lipschitz bound and stability estimate for $\overline{\nu}_\rho$ and $\underline{\nu}_\rho$. Classical arguments from the theory of viscosity solutions yield the local uniform convergence, as $\rho \to 0$, of u_ρ^j to u^j for $j = 1,2$, where u^j are as in the statement of Theorem 4.2.1 for the Hamiltonian H. Since $\overline{\nu}_\rho$ and $\underline{\nu}_\rho$ converge, as $\rho\to 0$, to $\overline{\nu}$ and $\underline{\nu}$, the proof is complete. □

We now present the proof of Proposition 4.2.1. As in Section 3.2, for $R > 0$, define

$$\Delta_R := \left\{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d : |x - y| \le R\right\}.$$

Recall also the definition of the distance function, whose properties are proved in Lemmas 2.4.3 and 2.4.5, defined by

$$L(x, y, \tau) := \inf\left\{\int_0^\tau H^*\left(-\dot\gamma_s, \gamma_s\right)\, ds : \gamma \in \mathcal{A}(x, y, \tau)\right\}$$

where, for $x, y \in \mathbb{R}^d$ and $\tau > 0$,

$$\mathcal{A}(x, y, \tau) := \left\{\gamma \in W^{1,\infty}([0, \tau], \mathbb{R}^d) : \gamma_0 = x,\ \gamma_\tau = y\right\}.$$

Proof of Proposition 4.2.1. Classical viscosity solution arguments show that $z(x, y, t) := u(x, t) - v(y, t)$ is a sub-solution of

$$z_t = H(D_x z, x)\dot\zeta - H(-D_y z, y)\dot\eta \quad \text{in } \mathbb{R}^d \times \mathbb{R}^d \times (0, \infty). \tag{4.2.4}$$

For $0 < \lambda < \left(\max_{0\le t\le T}(\zeta_t - \eta_t)_-\right)^{-1}$, define

$$\Phi_\lambda(x, y, t) := L\left(x, y, \frac{1}{\lambda} + \zeta_t - \eta_t\right).$$

A simple computation and Lemma 2.4.3(a) reveal that Φ satisfies (4.2.4) at any point (x, y, t) of differentiability.

Next, for $0 < \beta < 1$ and $\mu > 0$, define

$$\Psi(x, y, t) := u(x, t) - v(y, t) - \Phi_\lambda(x, y, t) - \frac{\beta}{2}(|x|^2 + |y|^2) - \mu t.$$

The comparison principle yields that $|u(x,t)| \leq M$ and $|v(x,t)| \leq M$ on $\mathbb{R}^d \times [0,T]$, where

$$M = \max\left\{\|u_0\|_\infty + |H(0)| \max_{0\leq t\leq T} |\zeta(t)|, \|v_0\|_\infty + |H(0)| \max_{0\leq t\leq T} |\eta(t)|\right\}.$$

Therefore, Ψ attains a maximum on $\mathbb{R}^d \times \mathbb{R}^d \times [0,T]$ at some $(\hat{x},\hat{y},\hat{t})$ that depends on β, λ, and μ. Assume for the sake of contradiction that $\hat{t} > 0$. In view of Lemma 2.3.1(b), we may also assume $\hat{t} < T$.

Rearranging terms in the inequality $\Psi(0,0,\hat{t}) \leq \Psi(\hat{x},\hat{y},\hat{t})$ gives

$$\frac{\beta}{2}(|\hat{x}|^2 + |\hat{y}|^2) \leq u(\hat{x},\hat{t}) - v(\hat{y},\hat{t}) - (u(\hat{0},\hat{t}) - v(\hat{0},\hat{t})) \leq 4M. \tag{4.2.5}$$

The inequality $\Psi(\hat{y},\hat{y},\hat{t}) \leq \Psi(\hat{x},\hat{y},\hat{t})$ and Lemma 2.4.3(b) yield

$$\left(\frac{1}{\lambda} + \zeta_{\hat{t}} - \eta_{\hat{t}}\right) \overline{\nu}^* \left(\frac{\lambda|x-y|}{1+\lambda(\zeta_{\hat{t}} - \eta_{\hat{t}})}\right) \leq u(\hat{x},\hat{t}) - u(\hat{y},\hat{t}) + \frac{\beta}{2}(|\hat{y}|^2 - |\hat{x}|^2) \leq 6M. \tag{4.2.6}$$

Then (4.2.5) and (4.2.6) together imply that, for some $R > 0$ depending on λ, M, $\|\zeta\|_{\infty,T}$, and $\|\eta\|_{\infty,T}$, but independent of β, $(\hat{x},\hat{y}) \in \Omega_{R,\beta}$, where

$$\begin{aligned}\Omega_{R,\beta} &:= \Delta_R \cap B_{R\beta^{-1/2}} \\ &= \left\{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d : \left(|x|^2 + |y|^2\right)^{1/2} \leq R\beta^{-1/2} \quad \text{and} \quad |x-y| \leq R\right\}.\end{aligned}$$

In the arguments that follow, the constant $C > 0$ depends only on R, and may change from line to line.

For $0 < \delta < 1$, set

$$\begin{aligned}\Psi_\delta(x,y,z,w,t) := u(x,t) - v(y,t) - \frac{1}{2\delta}(|x-z|^2 + |y-w|^2) - \Phi_\lambda(z,w,t) \\ - \frac{\beta}{2}(|z|^2 + |w|^2) - \mu t - \frac{1}{2}\left(|x-\hat{x}|^2 + |y-\hat{y}|^2 + |t-\hat{t}|^2\right)\end{aligned}$$

and assume that the maximum of Ψ_δ on $\Omega_{R,\beta}\times\Omega_{R,\beta}\times[0,T]$ is attained at $(x_\delta, y_\delta, z_\delta, w_\delta, t_\delta)$.

Lemma 2.4.3(c) gives $|D_z\Phi_\lambda| + |D_w\Phi_\lambda| + \beta(|z|+|w|) \le C$ on $\Omega_{R,\beta}\times\Omega_{R,\beta}\times[0,T]$. Rearranging terms in the inequality $\Psi_\delta(x_\delta, y_\delta, x_\delta, y_\delta, t_\delta) \le \Psi_\delta(x_\delta, y_\delta, z_\delta, w_\delta, t_\delta)$ yields

$$\begin{aligned}\frac{1}{2\delta}\left(|x_\delta - z_\delta|^2 + |y_\delta - w_\delta|^2\right) &\le \Phi_\lambda(x_\delta, y_\delta, t_\delta) - \Phi_\lambda(z_\delta, w_\delta, t_\delta)\\ &\quad + \frac{\beta}{2}(|x_\delta|^2 + |y_\delta|^2 - |z_\delta|^2 - |w_\delta|^2) + C\left(|x_\delta - z_\delta| + |y_\delta - w_\delta|\right),\end{aligned}$$

and, hence, $|x_\delta - z_\delta| + |y_\delta - w_\delta| \le C\delta$.

Since $(\hat{x}, \hat{y}, \hat{x}, \hat{y}, \hat{t}) \in \Omega_{R,\beta}\times\Omega_{R,\beta}\times[0,T]$ and $\Psi_\delta(\hat{x}, \hat{y}, \hat{x}, \hat{y}, \hat{t}) = \Psi(\hat{x}, \hat{y}, \hat{t})$,

$$\begin{aligned}&\Psi(\hat{x}, \hat{y}, \hat{t}) = u(\hat{x}, \hat{t}) - v(\hat{y}, \hat{t}) - \Phi_\lambda(\hat{x}, \hat{y}, \hat{t}) - \frac{\beta}{2}(|\hat{x}|^2 + |\hat{y}|^2) - \mu\hat{t}\\ &\le u(x_\delta, t_\delta) - v(y_\delta, t_\delta) - \frac{1}{2}|x_\delta - \hat{x}|^2 - \frac{1}{2}|y_\delta - \hat{y}|^2 - \frac{1}{2\delta}(|x_\delta - z_\delta|^2 + |y_\delta - w_\delta|^2)\\ &- \Phi_\lambda(z_\delta, w_\delta, t_\delta) - \frac{\beta}{2}(|z_\delta|^2 + |w_\delta|^2) - \mu t_\delta - \frac{1}{2}|t_\delta - \hat{t}|^2\\ &\le \Psi(x_\delta, y_\delta, t_\delta) + \Phi_\lambda(x_\delta, y_\delta, t_\delta) - \Phi_\lambda(z_\delta, w_\delta, t_\delta) + \frac{\beta}{2}(|x_\delta|^2 + |y_\delta|^2 - |z_\delta|^2 - |w_\delta|^2)\\ &- \frac{1}{2}\left(|x_\delta - \hat{x}|^2 - |y_\delta - \hat{y}|^2 - |t_\delta - \hat{t}|^2\right).\end{aligned}$$

Rearranging terms and using $\Psi(x_\delta, y_\delta, t_\delta) \le \Psi(\hat{x}, \hat{y}, \hat{t})$, we see that

$$|x_\delta - \hat{x}|^2 + |y_\delta - \hat{y}|^2 + |t_\delta - \hat{t}|^2 \le C\delta.$$

Therefore, for sufficiently small δ, $(x_\delta, y_\delta, z_\delta, w_\delta, t_\delta)$ is a local interior maximum point of Ψ_δ in $\Omega_{R,\beta}\times\Omega_{R,\beta}\times(0,T)$.

Since

$$\begin{aligned}(x, y, t) \mapsto u(x,t) - v(y,t) &- \frac{1}{2\delta}\left(|x - z_\delta|^2 + |y - w_\delta|^2\right)\\ &- \Phi_\lambda(z_\delta, w_\delta, t) - \mu t - \frac{1}{2}\left(|x - \hat{x}|^2 - |y - \hat{y}|^2 - |t - \hat{t}|^2\right)\end{aligned}$$

attains an interior maximum at $(x_\delta, y_\delta, t_\delta)$, the definition of viscosity solutions yields

$$\begin{aligned}\mu + t_\delta - \hat{t} + \Phi_{\lambda,t}(z_\delta, w_\delta, t_\delta) \leq H\left(\frac{x_\delta - z_\delta}{\delta} + x_\delta - \hat{x}, x_\delta\right)\dot{\zeta}_{t_\delta} \\ - H\left(-\frac{y_\delta - w_\delta}{\delta} - (y_\delta - \hat{y}), y_\delta\right)\dot{\eta}_{t_\delta}.\end{aligned}$$

Next, (z_δ, w_δ) is a minimum point of

$$(z, w) \mapsto \Phi_\lambda(z, w, t_\delta) + \frac{1}{2\delta}(|x_\delta - z|^2 + |y_\delta - w|^2) + \frac{\beta}{2}(|z|^2 + |w|^2).$$

In view of Lemma 2.4.4, Φ_λ is differentiable at $(z_\delta, w_\delta, t_\delta)$, and so

$$\begin{cases} D_x\Phi_\lambda(z_\delta, w_\delta, t_\delta) = \dfrac{x_\delta - z_\delta}{\delta} - \beta z_\delta, \\ D_y\Phi_\lambda(z_\delta, w_\delta, t_\delta) = \dfrac{y_\delta - w_\delta}{\delta} - \beta w_\delta, \quad \text{and} \\ \Phi_{\lambda,t}(z_\delta, w_\delta, t_\delta) = H(D_x\Phi_\lambda(z_\delta, w_\delta, t_\delta), z_\delta)\dot{\zeta}_{t_\delta} - H(-D_y\Phi_\lambda(z_\delta, w_\delta, t_\delta), w_\delta)\dot{\eta}_{t_\delta}. \end{cases}$$

It follows that

$$\begin{aligned}\mu + t_\delta - \hat{t} + \Phi_{\lambda,t}(z_\delta, w_\delta, t_\delta) \leq H\left(D_x\Phi_\lambda(z_\delta, w_\delta, t_\delta) + \beta z_\delta + x_\delta - \hat{x}, x_\delta\right)\dot{\xi}_{t_\delta} \\ - H\left(-D_y\Phi_\lambda(z_\delta, w_\delta, t_\delta) - \beta w_\delta - (y_\delta - \hat{y}), y_\delta\right)\dot{\zeta}_{t_\delta}.\end{aligned}$$

The bounds for $(\hat{x}, \hat{y}, \hat{t})$ and $(x_\delta, y_\delta, z_\delta, w_\delta, t_\delta)$ and the local Lipschitz regularity of H yield

$$\mu \leq C(\beta^{1/2} + \delta^{1/2} + \delta)\left(\left\|\dot{\xi}\right\|_{\infty,T} + \left\|\dot{\zeta}\right\|_{\infty,T}\right).$$

We obtain a contradiction for sufficiently small enough δ and β.

Therefore, for all $\mu > 0$ and $t \in [0,T]$,

$$\begin{aligned}
&\lim_{\beta\to 0} \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} \left(u(x,t) - v(y,t) - \Phi_\lambda(x,y,t) - \frac{\beta}{2}(|x|^2 + |y|^2) \right) \\
&\quad = \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} \left(u(x,t) - v(y,t) - \Phi_\lambda(x,y,t) \right) \\
&\quad \le \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} \left(u_0(x) - v_0(y) - L(x,y,1/\lambda) \right) + \mu t.
\end{aligned}$$

The desired inequality is established upon letting $\mu \to 0$ and using the bounds in Lemma 2.4.3(b). $\square$

4.3 Perron's method

We consider the initial value problem

$$\begin{cases} du = F(D^2u, Du, u, x, t)\, dt + \displaystyle\sum_{i=1}^m H^i(Du, x) \cdot dW^i \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \\ u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d, \end{cases} \tag{4.3.1}$$

where $u_0 \in BUC(\mathbb{R}^d)$, $T > 0$, and $W = (W^1, W^2, \dots, W^m) \in \mathscr{C}_g^\alpha$ for some $\alpha \in (1/3, 1/2]$. As explained before, our goal is to use the comparison principle to prove that the unique solution of (4.3.1) is given by the maximal sub-solution.

We assume that $F : \mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times [0,T] \to \mathbb{R}$ is continuous, bounded for bounded $(X,p,r) \in \mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R}$, degenerate elliptic, and nonincreasing in $r \in \mathbb{R}$; that is,

$$\begin{cases} F \in C(\mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times [0,T]) \cap C_b(B_R(0) \times \mathbb{R}^d \times [0,T]) \quad \text{for all } R > 0, \\ X \mapsto F(X, \cdot, \cdot, \cdot, \cdot) \text{ is nondecreasing, and } r \mapsto F(\cdot, \cdot, r, \cdot, \cdot) \text{ is nonincreasing,} \end{cases} \tag{4.3.2}$$

and the Hamiltonians are regular enough so that (2.1.1) has local-in-time, C^2-in-space solutions:

$$H \in C_b^4(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^m). \tag{4.3.3}$$

Recall that, if the $\{H^i\}$ satisfy (2.2.7) (for example, if $m = 1$ or there is no spatial dependence), then it suffices to have $H \in C_b^2(B_R \times \mathbb{R}^d; \mathbb{R}^m)$ for all $R > 0$.

Most importantly, we assume that the comparison principle holds for (4.3.1). That is,

$$\begin{cases} \text{whenever } u \in USC(\mathbb{R}^d \times (0,\infty)) \text{ and } v \in LSC(\mathbb{R}^d \times (0,\infty)) \text{ are} \\ \text{respectively a bounded sub- and super-solution of (4.3.1), then, for all } t > 0, \\ \sup_{x\in\mathbb{R}^d} (u(x,t) - v(x,t)) \le \sup_{x\in\mathbb{R}^d} (u(x,0) - v(x,0)). \end{cases} \tag{4.3.4}$$

Theorem 4.3.1. *If* (4.3.2), (4.3.3), *and* (4.3.4) *hold, then there exists a unique solution of* (4.3.1) *in* $BUC(\mathbb{R}^d \times [0,T])$, *which is given by*

$$u(x,t) := \sup\{v(x,t) : v(\cdot,0) \le u_0 \textit{ and } v \textit{ is a sub-solution of } (4.3.1)\}. \tag{4.3.5}$$

More assumptions are generally required for F, H, and W in order for the comparison principle to hold. This is especially the case when H has nontrivial spatial dependence, even when $m = 1$, as discussed in Chapter 3. The only assumptions we directly use for the Perron construction, besides the comparison principle, are (4.3.2) and (4.3.3).

4.3.1 Construction of sub- and super-solutions

We build some particular global sub- and super-solutions of (4.3.1) achieving arbitrary initial data in $C_b^2(\mathbb{R}^d)$. Throughout the proofs below, we repeatedly use the properties of the smooth-solution operator $S(t,s) : C_b^2(\mathbb{R}^d) \to C_b^2(\mathbb{R}^d)$ used in Section 2.2.3 to construct local-in-time, smooth-in-space solutions of (2.1.1).

Lemma 4.3.1. *For every $\phi \in C_b^2(\mathbb{R}^d)$, there exist a locally bounded sub- and super-solution $\underline{u}$ and $\overline{u}$ of* (4.3.1) *with $\underline{u}(\cdot,0) = \overline{u}(\cdot,0) = \phi$. Moreover, $\underline{u}$ and $\overline{u}$ are both continuous in a neighborhood of $\mathbb{R}^d \times \{0\}$.*

Proof. Only the sub-solution is constructed, since the argument for the super-solution is similar.

For some $h > 0$, there exists a solution $\Phi \in C([0,h], C_b^2(\mathbb{R}^d))$ of (2.1.1) in $\mathbb{R}^d \times [0,h]$ with $\Phi(\cdot,0) = \phi$, which is given by $\Phi(x,t) = S(t,0)\phi(x)$ for $(x,t) \in \mathbb{R}^d \times [0,h]$. Set

$$\begin{cases} R := \sup\limits_{0\le t\le h} \|\Phi(\cdot,t)\|_{C_b^2(\mathbb{R}^d)} \quad \text{and} \\ C := \inf\limits_{|X|+|p|+|u|\le R} \inf\limits_{(x,t)\in\mathbb{R}^d\times[0,T]} F(X,p,u,x,t), \end{cases}$$

and define $\underline{u}(\cdot,t) := \Phi(\cdot,t) - Ct$ for $t \in [0,h]$.

For some fixed $h_0 > 0$, the function Φ_0 defined by $\Phi_0(x,t,s) := S(t,s)(0)(x)$ satisfies $\Phi_0(\cdot,t,s) \in C_b^2(\mathbb{R}^d)$ when $s \le t \le s + h_0$. Note that Lemma 2.2.6(a) implies that $S(t,s)(M)(x) = \Phi_0(x,t,s) + M$ for all $M \in \mathbb{R}$.

Define

$$\begin{cases} R_0 := \sup\limits_{0\le t-s\le h_0} \|\Phi_0(\cdot,t,s)\|_{C_b^2(\mathbb{R}^d)}, \\ C_0 := \inf\limits_{|X|+|p|+|u|\le R_0,(x,t)\in\mathbb{R}^d\times[0,T]} F(X,p,u,x,t), \end{cases}$$

and, for $k = 0,1,2,3,\ldots, \left\lceil \frac{T-h}{h_0} \right\rceil - 1$ and $t \in (h + kh_0, h + (k+1)h_0]$,

$$\begin{cases} M_k := \sup\limits_{x\in\mathbb{R}^d} [\underline{u}(x, h+kh_0)]_- \quad \text{and} \\ \underline{u}(\cdot,t) := \Phi_0(\cdot,t,h+kh_0) - M_k - C_0(t-h-kh_0). \end{cases}$$

By construction, $\underline{u}$ is upper-semicontinuous, bounded, and continuous on $\mathbb{R}^d \times [0,h]$.

Now choose $\eta \in C_b^2(\mathbb{R}^d)$ and $\psi \in C^1([0,T])$, let $h_1 > 0$ and $t_0 \in [0,T]$ be such that

$S(\cdot, t_0)\eta \in C((t_0 - h_1, t_0 + h_1), C_b^2(\mathbb{R}^d))$, and assume that

$$\underline{u}(x,t) - S(t,t_0)\eta(x) - \psi(t)$$

attains a strict maximum in $\mathbb{R}^d \times (t_0 - h_1, t_0 + h_1)$ at (x_0, t_0). In view of Lemma 2.3.1(b), it suffices to consider $t_0 \neq h + kh_0$ for any $k \in \mathbb{N} \cup \{0\}$. Assume also that $t_0 \in (h + kh_0, h + (k+1)h_0)$ for some $k \geq 0$, as the proof for $t_0 \in (0, h)$ is similar. Then

$$D\Phi_0(x_0, t_0) = D\eta(x_0), \quad D^2\Phi_0(x_0, t_0) \leq D^2\eta(x_0), \quad \text{and} \quad \psi'(t_0) = -C_0,$$

where the last equality follows from Lemma 2.2.6(c). Therefore,

$$\begin{aligned} &\psi'(t_0) - F(D^2\eta(x_0), D\eta(x_0), \underline{u}(x_0, t_0), x_0, t_0) \\ &\leq -C_0 - F(D^2\Phi_0(x_0, t_0), D\Phi_0(x_0, t_0), \Phi_0(x_0, t_0, h + kh_0), x_0, t_0) \leq 0, \end{aligned}$$

and so $\underline{u}$ is a sub-solution of (4.3.1). □

4.3.2 The proof of Theorem 4.3.1

The proof that the Perron construction yields a solution involves two main steps. First, it is clear from Definition 2.3.1 that the maximum of a finite number of sub-solutions is also a sub-solution, with a corresponding statement holding true for the minimum of a finite number of super-solutions. We generalize this observation to infinite families.

Lemma 4.3.2. *Let $\mathcal{F}$ be a family of sub- (resp. super-) solutions of* (4.3.1)*. Define*

$$U(x,t) := \sup_{v \in \mathcal{F}} v(x,t) \quad \left(\textit{resp.} \ \inf_{v \in \mathcal{F}} v(x,t) \right).$$

Assume that $U^* < \infty$ (resp. $U_* > -\infty$). Then U^* (resp. U_*) is a sub- (resp. super-) solution of (4.3.1).

Proof. We give only the proof for sub-solutions, since it is almost identical for super-solutions.

Let $\phi \in C_b^2(\mathbb{R}^d)$, $\psi \in C^1([0,T])$, $t_0 > 0$, and $h > 0$ be such that $S(\cdot, t_0)\phi \in C((t_0 - ht_0 + h), C_b^2(\mathbb{R}^d))$, assume that

$$U^*(x,t) - S(t,t_0)\phi(x) - \psi(t)$$

attains a local maximum at $(x_0, t_0) \in \mathbb{R}^d \times (t_0 - h, t_0 + h)$, and, without loss of generality, assume $x_0 = 0$, $\phi(0) = 0$, and $\psi(t_0) = 0$. Set $p := D\phi(0)$, $X := D^2\phi(0)$, and $a := \psi'(t_0)$.

For fixed $\delta > 0$, shrink $h > 0$, if necessary, and let $r > 0$ be such that

$$\begin{cases} \phi(x) \le p \cdot x + \dfrac{1}{2} Xx \cdot x + \delta |x|^2 & \text{for all } x \in B_r(0) \quad \text{and} \\ \psi(t) \le a(t - t_0) + \delta |t - t_0| & \text{for all } t \in (t_0 - h, t_0 + h). \end{cases}$$

Choose ϕ_1 and ϕ_2 in $C_b^2(\mathbb{R}^d)$ such that

$$\begin{cases} \phi_1(x) = p \cdot x + \dfrac{1}{2} Xx \cdot x + \delta |x|^2 & \text{for } x \in B_r(x_0), \\ \phi_2(x) = p \cdot x + \dfrac{1}{2} Xx \cdot x + 2\delta |x|^2 & \text{for } x \in B_r(x_0), \text{ and} \\ \phi \le \phi_1 \le \phi_2 & \text{on } \mathbb{R}^d. \end{cases}$$

If h is sufficiently small, then $S(\cdot, t_0)\phi_1$ and $S(\cdot, t_0)\phi_2$ belong to $C((t_0 - h, t_0 + h), C_b^2(\mathbb{R}^d))$, and Lemma 2.2.6(b) implies that

$$S(t,t_0)\phi \le S(t,t_0)\phi_1 \le S(t,t_0)\phi_2 \quad \text{for all } t \in (t_0 - h, t_0 + h).$$

In particular, $U^*(x,t) - S(t,t_0)\phi_1(x) - \psi(t)$ attains a local maximum at $(0, t_0)$.

Let $(x_n, t_n) \in \mathbb{R}^d \times (t_0 - h, t_0 + h)$ and $v_n \in \mathcal{F}$ be such that, as $n \to \infty$, $(x_n, t_n) \to (0, t_0)$ and $v_n(x_n, t_n) \to U^*(0, t_0)$, and let (x_n', t_n') be the maximum point attained over $\overline{B_r(0)} \times [t_0 - h, t_0 + h]$ by the function

$$v_n(x,t) - S(t,t_0)\phi_2(x) - a(t - t_0) - 2\delta(|t - t_0|^2 + n^{-1})^{1/2}. \tag{4.3.6}$$

Then

$$
\begin{aligned}
v_n(x_n, t_n) &\leq v_n(x'_n, t'_n) + S(t_n, t_0)\phi_2(x_n) - S(t'_n, t_0)\phi_2(x'_n) + a(t_n - t'_n) \\
&\quad + 2\delta\left((|t_n - t_0|^2 + n^{-1})^{1/2} - (|t'_n - t_0|^2 + n^{-1})^{1/2}\right).
\end{aligned} \tag{4.3.7}
$$

Let $(y, s) \in \overline{B_r(0)} \times [t_0 - h, t_0 + h]$ be an accumulation point of the sequence $\{(x'_n, t'_n)\}_{n\in\mathbb{N}}$. Passing to the limit in (4.3.7) yields

$$
\begin{aligned}
U^*(0, t_0) &\leq U^*(y, s) - S(s, t_0)\phi_2(y) - a(s - t_0) - 2\delta|s - t_0| \\
&\leq U^*(0, t_0) + S(s, t_0)\phi_1(y) - S(s, t_0)\phi_2(y) - \delta|s - t_0| \\
&\leq U^*(0, t_0) - \delta|s - t_0|,
\end{aligned} \tag{4.3.8}
$$

and, therefore, $s = t_0$. Inserting this fact into (4.3.8) gives $\phi_2(y) \leq \phi_1(y)$, which implies that

$$
y = 0, \quad \lim_{n\to\infty}(x'_n, t'_n) = (0, t_0), \quad \text{and} \quad \lim_{n\to\infty} v_n(x'_n, t'_n) = U^*(0, t_0).
$$

In particular, for sufficiently large n, $(x'_n, t'_n) \in B_r(0) \times (t_0 - h, t_0 + h)$.

Finally, set $\Phi(x, t) := S(t, t_0)\phi_2(x)$. (The definition of solutions) gives

$$
a + 2\delta\frac{t'_n - t_0}{(|t'_n - t_0|^2 + n^{-1})^{-1}} \leq F(D^2\Phi(x'_n, t'_n), D\Phi(x'_n, t'_n), v_n(x'_n, t'_n), x'_n, t'_n).
$$

Upon letting $n \to \infty$ and $\delta \to 0$, this becomes $a \leq F(X, p, U^*(0, t_0), 0, t_0)$, as desired. □

The second step is to show that if a "strict" sub-solution has its values increased in a certain way in a sufficiently small open cylinder, then the resulting function is another sub-solution. This "bump" construction is less straightforward than in the classical viscosity solution, due to the limited flexibility in the choice of test functions. The domain of dependence result Lemma 2.2.7 will play an important role in the proof.

Lemma 4.3.3. *Suppose that w is a sub-solution of* (4.3.1), *and that w_* fails to be a super-solution. Then there exists $(x_0, t_0) \in \mathbb{R}^d \times (0, T]$ such that, for all $\kappa > 0$,* (4.3.1) *admits a* *sub-solution* w_κ *satisfying*

$$w_\kappa \geq w, \quad \sup(w_\kappa - w) > 0, \quad \textit{and} \quad w_\kappa = w \quad \textit{in} \quad \Big(\mathbb{R}^d \times [0,T]\Big) \setminus N_\kappa(x_0,t_0).$$

Proof. By assumption, there exist $\phi \in C_b^2(\mathbb{R}^d)$, $\psi \in C^1([0,T])$, $(x_0,t_0) \in \mathbb{R}^d \times (0,T]$, and $h \in (0,\kappa)$ such that $S(\cdot,t_0)\phi \in C((t_0-h,t_0+h), C_b^2(\mathbb{R}^d))$,

$$w_*(x,t) - S(t,t_0)\phi(x) - \psi(t)$$

attains a local minimum at (x_0,t_0), and

$$\psi'(t_0) - F(D^2\phi(x_0), D\phi(x_0), w_*(x_0,t_0), x_0, t_0) < 0. \tag{4.3.9}$$

Assume again $x_0 = 0$, $\phi(0) = 0$, and $\psi(t_0) = 0$, set $X := D^2\phi(0)$, $p := D\phi(0)$, and $a := \psi'(t_0)$, and define the nondecreasing functions $\omega_1, \omega_2 : [0,\infty) \to [0,\infty)$ by

$$\begin{cases} \omega_1(\sigma) := \sup\limits_{|x|\leq\sigma} \dfrac{|\phi(x) - p\cdot x - \frac{1}{2}Xx\cdot x|}{|x|^2} & \text{and} \\ \omega_2(\sigma) := \sup\limits_{|t-t_0|\leq\sigma} \dfrac{|\psi(t) - a(t-t_0)|}{|t-t_0|}. \end{cases}$$

Let $\gamma \in (0,1)$, $r \in (0,\kappa)$, and $s \in (0,h)$ be such that

$$\omega_1(r), \omega_2(s) \leq \frac{\gamma}{2}, \tag{4.3.10}$$

and set

$$\delta := \gamma \min\left(\frac{r^2}{16}, \frac{s}{8}\right). \tag{4.3.11}$$

Choose $\hat{\eta} \in C_b^2(\mathbb{R}^d)$ satisfying

$$\begin{cases} \hat{\eta}(x) = p \cdot x + \dfrac{1}{2} X x \cdot x - \gamma |x|^2 \text{ in } B_r(x_0) \quad \text{and} \\ \hat{\eta} \le \phi \text{ in } \mathbb{R}^d, \end{cases}$$

and, if necessary, shrink $h > 0$ so that $S(\cdot, t_0)\hat{\eta} \in C((t_0 - h, t_0 + h), C_b^2(\mathbb{R}^d))$. Observe that this may also result in s, and therefore δ, becoming smaller.

For $(x,t) \in \mathbb{R}^d \times (t_0 - h, t_0 + h)$, define

$$\hat{w}(x,t) := w_*(0,t_0) + \delta + S(t,t_0)\hat{\eta}(x) + a(t - t_0) - \gamma(|t - t_0|^2 + \delta^2)^{1/2}.$$

Then, if γ, r, and s (and therefore δ) are sufficiently small, $\hat{w}$ satisfies the sub-solution property of Definition 2.3.1 in $N_{r,s}(0,t_0)$. Indeed, choose $\zeta \in C_b^2(\mathbb{R}^d)$, $\alpha \in C^1([0,T])$, and $(\hat{x},\hat{t}) \in N_{r,s}(0,t_0)$; let $\hat{h} > 0$ be such that $S(\cdot,\hat{t})\zeta \in C(\hat{t} - \hat{h}, \hat{t} + \hat{h}, C_b^2(\mathbb{R}^d))$; and assume that

$$\hat{w}(x,t) - S(t,\hat{t})\zeta(x) - \alpha(t)$$

attains a strict maximum at $(\hat{x},\hat{t})$. This implies that

$$DS(\hat{t},t_0)\hat{\eta}(\hat{x}) = D\zeta(\hat{x}) \quad \text{and} \quad D^2 S(\hat{t},t_0)\hat{\eta}(\hat{x}) \le D^2\zeta(\hat{x}),$$

and, because

$$t \mapsto \sup_{\mathbb{R}^d} \big(S(t,t_0)\hat{\eta} - S(t,\hat{t})\zeta\big) + at - \gamma(|t - t_0|^2 + \delta^2)^{1/2} - \alpha(t)$$

attains a maximum at $\hat{t}$, Lemma 2.2.6(c) yields

$$a - \gamma \frac{\hat{t} - t_0}{\big(|\hat{t} - t_0|^2 + \delta^2\big)^{1/2}} = \alpha'(\hat{t}).$$

Therefore, from the strict inequality in (4.3.9), the continuity of the solution map $S(t,t_0)$ on $C_b^2(\mathbb{R}^d)$, and the continuity of F, it follows that $\alpha'(\hat{t}) \leq F(D^2\zeta(\hat{x}), D\zeta(\hat{x}), \hat{w}(\hat{x},\hat{t}), \hat{x}, \hat{t})$ if γ, r, and s are small enough.

Define

$$R := \max_{|t-t_0|\leq h} \max\left\{\|DS(t,t_0)\phi\|_\infty, \|DS(t,t_0)\hat{\eta}\|_\infty\right\},$$

and shrink s further so that

$$\rho_R(s) \leq \frac{r}{8}, \tag{4.3.12}$$

where ρ_R is the modulus from the domain-of-dependence result Lemma 2.2.7. The claim is that

$$w(x,t) > \hat{w}(x,t) \quad \text{in } N_{7r/8,s}(0,t_0)\backslash\overline{N_{5r/8,s/2}(0,t_0)}. \tag{4.3.13}$$

As a first step, observe that

$$w(x,t) - \hat{w}(x,t) \geq w_*(x,t) - \hat{w}(x,t) \geq -\delta + S(t,t_0)\phi(x) - S(t,t_0)\hat{\eta}(x)$$
$$+ (\gamma - \omega_2(s))|t-t_0|.$$

Suppose that $(x,t) \in N_{7r/8,s}(0,t_0)\backslash\overline{N_{7r/8,s/2}}$, that is, $s/2 < |t-t_0| < s$ and $|x| < 7r/8$. Then (4.3.10) and (4.3.11) give

$$w(x,t) - \hat{w}(x,t) \geq -\delta + (\gamma - \omega_2(s)) \cdot \frac{s}{2} \geq -\delta + \frac{\gamma s}{4} \geq \frac{\gamma s}{8} > 0. \tag{4.3.14}$$

To prove (4.3.13) for $(x,t) \in N_{7r/8,s}\backslash\overline{N_{5r/8,s}}$, we apply Lemma 2.2.7 to the annulus $K =$

$\overline{B_r(0)}\backslash B_{r/2}(0)$ and obtain

$$\inf_{r/2\le|x|\le r}(\phi(x)-\hat{\eta}(x)) \le \inf\left\{S(t,t_0)\phi(x)-S(t,t_0)\hat{\eta}(x) : \operatorname{dist}(x,K^c)\ge \rho_R(|t-t_0|)\right\}$$
$$\le \inf\left\{S(t,t_0)\phi(x)-S(t,t_0)\hat{\eta}(x) : \min\left(r-|x|,|x|-\frac{r}{2}\right)\ge \rho_R(s)\right\}.$$

Combining this with (4.3.10), (4.3.11), and (4.3.12), it follows that, whenever $5r/8<|x|<7r/8$ and $|t-t_0|<s$,

$$w(x,t)-\hat{w}(x,t)\ge -\delta+\inf_{r/2\le|x|\le r}(\phi(x)-\hat{\eta}(x))\ge -\delta+(\gamma-\omega_1(r))\cdot\left(\frac{r}{2}\right)^2$$
$$\ge -\delta+\frac{\gamma r^2}{8}\ge \frac{\gamma r^2}{16}>0.$$

This finishes the proof of (4.3.13).

Finally, define

$$w_\kappa(x,t):=\begin{cases}\max(\hat{w}(x,t),w(x,t)) & \text{for } (x,t)\in N_{7r/8,s}(0,t_0), \text{ and}\\ w(x,t) & \text{for } (x,t)\notin N_{7r/8,s}(0,t_0).\end{cases}$$

Then $w_\kappa\ge w$, and $w_\kappa=w$ outside of $N_\kappa(0,t_0)$. If (x_n,t_n) is such that $\lim_{n\to\infty}(x_n,t_n)=(0,t_0)$ and $\lim_{n\to\infty}w(x_n,t_n)=w_*(0,t_0)$, then

$$\lim_{n\to\infty}(w(x_n,t_n)-\hat{w}(x_n,t_n))=-(1-\gamma)\delta<0,$$

so that $\sup_{N_\kappa(0,t_0)}(w_\kappa-w)>0$. Finally, w_κ is a sub-solution. This is evident on $(\mathbb{R}^d\times[0,T])\backslash\overline{N_{7r/8,s}(0,t_0)}$, as well as in the interior of $N_{7r/8,s}(0,t_0)$, because there w_κ is equal to the pointwise maximum of two sub-solutions. It remains to verify the sub-solution property on the boundary of $N_{7r/8,s}(0,t_0)$, and this follows because, in view of (4.3.13), $w_\kappa=w$ in a neighborhood of the boundary of $N_{7r/8,s}(0,t_0)$. □

Finally, we present the

Proof of Theorem 4.3.1. Observe first that, in view of Lemma 4.3.1 and the comparison principle (4.3.4), u is well-defined and bounded.

Fix $\varepsilon > 0$, let $\phi^\varepsilon \in C_b^2(\mathbb{R}^d)$ be such that

$$\phi^\varepsilon - \varepsilon \leq u_0 \leq \phi^\varepsilon + \varepsilon \quad \text{on } \mathbb{R}^d,$$

and let $\underline{u}^\varepsilon$ and $\overline{u}^\varepsilon$ be a sub- and super-solution, continuous in a neighborhood of $\mathbb{R}^d \times \{0\}$, and achieving respectively the initial data $\phi^\varepsilon - \varepsilon$ and $\phi^\varepsilon + \varepsilon$. The existence of the sub- and super-solution are guaranteed by Lemma 4.3.1.

The comparison principle yields

$$\underline{u}^\varepsilon \leq u_* \leq u \leq u^* \leq \overline{u}^\varepsilon \quad \text{on } \mathbb{R}^d \times [0, T],$$

and, in view of the continuity of $\underline{u}^\varepsilon$ and $\overline{u}^\varepsilon$ near $\mathbb{R}^d \times \{0\}$,

$$\phi^\varepsilon - \varepsilon \leq u_*(\cdot, 0) \leq u^*(\cdot, 0) \leq \phi^\varepsilon + \varepsilon.$$

Since ε is arbitrary, it follows that $u(\cdot, 0) = u_0$ and $\lim_{(x,t)\to(x_0,0)} u(x,t) = u_0(x_0)$ for all $x_0 \in \mathbb{R}^d$.

Lemma 4.3.2 now implies that u^* is a sub-solution of (4.3.1) with $u^*(x, 0) \leq u_0(x)$. The formula (4.3.5) for u then yields $u^* \leq u$, and therefore $u^* = u$. That is, u is itself upper-semicontinuous and a sub-solution.

On the other hand, u_* is a super-solution. If this were not the case, then Lemma 4.3.3 would imply the existence of a sub-solution $\tilde{u} \geq u$ and a neighborhood $N \subset \mathbb{R}^d \times (0, T]$ such that $\tilde{u} = u$ in $(\mathbb{R}^d \times [0, T]) \backslash N$ and $\sup_N(\tilde{u} - u) > 0$, contradicting the maximality of u.

The comparison principle (4.3.4) gives $u^* \leq u_*$, and, as a consequence of the definition of semicontinuous envelopes, $u_* \leq u^*$. Therefore, $u = u_* = u^*$ is a solution of (4.3.1) with $u = u_0$ on $\mathbb{R}^d \times \{0\}$. The uniqueness of u follows from yet another application of the comparison principle. $\square$

CHAPTER 5

APPROXIMATION SCHEMES

5.1 Introduction

In this chapter, we construct numerical schemes to approximate solutions, and prove that they converge under quite general assumptions. Among the approximations that we study are finite-difference schemes and Trotter-Kato splitting formulas. The former raise the possibility of numerical implementation, which we justify with precise error estimates in the first-order setting.

For a fixed, finite horizon $T > 0$, we consider the initial value problem

$$\begin{cases} du = F(D^2u, Du)\, dt + \sum_{i=1}^{m} H^i(Du) \cdot dW^i \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \\ u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d, \end{cases} \tag{5.1.1}$$

where

$$F \in C^{0,1}(\mathbb{S}^d \times \mathbb{R}^d) \quad \text{is degenerate elliptic,} \tag{5.1.2}$$

$W = (W^1, W^2, \dots, W^m) \in C([0,T], \mathbb{R}^m)$, and $u_0 \in BUC(\mathbb{R}^d)$. The precise assumptions on $H = (H^1, H^2, \dots, H^m) : \mathbb{R}^d \to \mathbb{R}^m$ are specified later, and depend on the setting. As is the case earlier in the thesis, H should be regular enough to allow for the construction, for sufficiently small intervals I and smooth ϕ, of local-in-time, spatially-smooth solutions of the equation

$$d\Phi = \sum_{i=1}^{m} H^i(D\Phi) \cdot dW^i \quad \text{in } \mathbb{R}^d \times I \quad \text{and} \quad \Phi(\cdot, t_0) = \phi \quad \text{in } \mathbb{R}^d. \tag{5.1.3}$$

5.1.1 The main results

We first summarize the main results in the context of certain finite-difference schemes. To simplify the presentation, assume $d = m = 1$, F and H are both smooth, and F depends only on u_{xx}, so that (5.1.1) becomes

$$du = F(u_{xx})\,dt + H(u_x) \cdot dW \quad \text{in } \mathbb{R} \times (0,T] \quad \text{and} \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}, \tag{5.1.4}$$

or, in the first order case, when $F \equiv 0$,

$$du = H(u_x) \cdot dW \quad \text{in } \mathbb{R} \times (0,T] \quad \text{and} \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}. \tag{5.1.5}$$

The approximations are constructed through the use of a scheme operator, which, for $h > 0$, $0 \le s \le t \le T$, and $\zeta \in C([0,T];\mathbb{R})$, is a map $S_h(t,s;\zeta) : BUC(\mathbb{R}) \to BUC(\mathbb{R})$, whose properties will be made more precise in Section 5.2.

Throughout the chapter, the symbol $\mathcal{P}$ denotes a partition of $[0,T]$ and $|\mathcal{P}|$ its mesh size, that is,

$$\begin{cases} \mathcal{P} := \{0 = t_0 < t_1 < \cdots, t_N = T\} \quad \text{and} \\ |\mathcal{P}| := \max\limits_{n=0,1,\ldots,N-1} (t_{n+1} - t_n). \end{cases}$$

Given such a partition $\mathcal{P}$ and a path $\zeta \in C([0,T];\mathbb{R})$, usually a piecewise linear approximation of W, we first define the function $\tilde{u}_h(\cdot;\zeta,\mathcal{P})$ by

$$\begin{cases} \tilde{u}_h(\cdot,0;\zeta,\mathcal{P}) := u_0 \quad \text{and} \\ \tilde{u}_h(\cdot,t;\zeta,\mathcal{P}) := S_h(t,t_n;\zeta)\tilde{u}_h(\cdot,t_n;\zeta,\mathcal{P}) \quad \text{for } n = 0,1,\ldots,N-1,\ t \in (t_n,t_{n+1}]. \end{cases} \tag{5.1.6}$$

The strategy is to choose families of approximating paths $\{W_h\}_{h>0}$ and partitions $\{\mathcal{P}_h\}_{h>0}$ satisfying

$$\lim_{h\to 0^+} \|W_h - W\|_\infty = 0 = \lim_{h \to 0^+} |\mathcal{P}_h|, \tag{5.1.7}$$

in such a way that the function

$$u_h(x,t) := \tilde{u}_h(x,t;W_h,\mathcal{P}_h) \tag{5.1.8}$$

is an efficient approximation of the solution of (5.1.4). We give several different specifications for $\mathcal{P}_h$ and W_h below. While technical, these are all made with the same idea in mind, namely, to ensure that the approximation W_h is "mild" enough with respect to the partition. In particular, for any consecutive points t_n and t_{n+1} of the partition $\mathcal{P}_h$, and for sufficiently small h, the ratio

$$\frac{|W_h(t_{n+1}) - W_h(t_n)|}{h}$$

should be less than some fixed constant. This is a special case of the kind of Courant-Lewy-Friedrichs (CFL) conditions required for the schemes in this chapter, which are discussed in more detail in the forthcoming sections.

As an example of the types of schemes studied in this chapter, we consider here the following adaptation of the Lax-Friedrichs finite difference approximation for scalar conservation laws. A formulation for approximating viscosity solutions of Hamilton-Jacobi equations with no space-time dependence was studied by Crandall and Lions [23].

For some $\varepsilon_h > 0$ to be determined, define

$$\begin{aligned} S_h(t,s;\zeta)u(x) := u(x) &+ H\left(\frac{u(x+h)-u(x-h)}{2h}\right)(\zeta(t)-\zeta(s)) \\ &+ \left[F\left(\frac{u(x+h)+u(x-h)-2u(x)}{h^2}\right)\right. \\ &+ \left.\varepsilon_h\left(\frac{u(x+h)+u(x-h)-2u(x)}{h^2}\right)\right](t-s). \end{aligned} \tag{5.1.9}$$

The first result, which is qualitative in nature, applies to the simple setting above as follows:

Theorem 5.1.1. *Assume that, in addition to* (5.1.7), W_h *and* $\mathcal{P}_h$ *satisfy*

$$|\mathcal{P}_h| \le \frac{h^2}{\|F'\|_\infty} \quad \text{and} \quad \varepsilon_h := h\left\|\dot{W}_h\right\|_\infty \xrightarrow{h\to 0} 0.$$

Then, as $h \to 0$, *the function* u_h *defined by* (5.1.8) *using the scheme operator* (5.1.9) *converges locally uniformly to the solution* u *of* (5.1.4).

We obtain explicit error estimates for finite difference approximations of the pathwise Hamilton-Jacobi equation (5.1.5). The results below are stated for the following scheme, which is defined, for some $\theta \in (0, 1]$, by

$$\begin{aligned} S_h(t,s;\zeta)u(x) := u(x) + H\left(\frac{u(x+h)-u(x-h)}{2h}\right)(\zeta(t)-\zeta(s)) \\ + \frac{\theta}{2}\left(u(x+h)+u(x-h)-2u(x)\right). \end{aligned} \tag{5.1.10}$$

Note that this corresponds to choosing $\varepsilon_h := \frac{\theta h^2}{2(t-s)}$ in (5.1.9).

The main tool for proving rates of convergence is the following pathwise error estimate. For the remaining results in the introduction, it is assumed that $\mathrm{Lip}(u_0) \le L$ for some $L > 0$.

Theorem 5.1.2. *There exists* $C > 0$ *depending only on the Lipschitz constant* L *such that, if* $\zeta \in C([0,T],\mathbb{R})$ *is piecewise linear over the partition* $\mathcal{P}$ *such that*

$$\max_{n=0,1,\ldots,N-1} |\zeta(t_{n+1}) - \zeta(t_n)| \le \frac{\theta}{\|H'\|_\infty} h,$$

and $\tilde{u}$ *is the solution of*

$$\tilde{u}_t = H(\tilde{u}_x)\dot{\zeta} \quad \text{in } \mathbb{R}\times(0,T] \quad \text{and} \quad \tilde{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R},$$

then, for all $\varepsilon, h > 0$,

$$\sup_{(x,t)\in\mathbb{R}\times[0,T]} |\tilde{u}_h(x,t;\zeta,\mathcal{P}) - \tilde{u}(x,t)| \\ \leq \frac{1}{\varepsilon}\sum_{n=0}^{N-1}(t_{n+1}-t_n)^2 + C\sqrt{N}h + \max_{s,t\in[0,T]}\left\{C\,|\zeta(s)-\zeta(t)| - \frac{|s-t|^2}{2\varepsilon}\right\}.$$

The rates of convergence are then established by choosing families of paths $\{W_h\}_{h>0}$ and partitions $\{\mathcal{P}_h\}_{h>0}$ in order to optimize the estimate from Theorem 5.1.2.

We do so first for an arbitrary, fixed continuous path W, whose modulus of continuity on $[0,T]$ is denoted by $\omega : [0,\infty) \to [0,\infty)$. For $h > 0$, define ρ_h implicitly by

$$\lambda := \frac{(\rho_h)^{1/2}\omega((\rho_h)^{1/2})}{h} < \frac{\theta}{\|H'\|_\infty}, \tag{5.1.11}$$

and let the partition $\mathcal{P}_h$ and path W_h satisfy

$$\begin{cases} \mathcal{P}_h := \{n\rho_h \wedge T\}_{n\in\mathbb{N}_0},\ M_h := \lfloor(\rho_h)^{-1/2}\rfloor, \\ \text{and, for } k\in\mathbb{N}_0 \text{ and } t\in[kM_h\rho_h,(k+1)M_h\rho_h), \\ W_h(t) := W(kM_h\rho_h) + \left(\dfrac{W((k+1)M_h\rho_h) - W(kM_h\rho_h)}{M_h\rho_h}\right)(t - kM_h\rho_h)\,. \end{cases} \tag{5.1.12}$$

Theorem 5.1.3. *There exists* $C > 0$ *depending only on* L *such that, if* u_h *is constructed using* (5.1.8) *and* (5.1.10) *with* $\mathcal{P}_h$ *and* W_h *as in* (5.1.11) *and* (5.1.12)*, and* u *is the pathwise viscosity solution of* (5.1.5)*, then*

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - u(x,t)| \leq C(1+T)\omega((\rho_h)^{1/2}).$$

As an example, if $W \in C^{0,\alpha}([0,T])$, then (5.1.11) means that $\rho_h = O(h^{2/(1+\alpha)})$, and the rate of convergence in Theorem 5.1.3 is $O(h^{\alpha/(1+\alpha)})$.

We now describe some constructions in the case that W is a Brownian motion, for which

(5.1.5) becomes the stochastic Hamilton-Jacobi equation

$$du = H(u_x) \circ dW \quad \text{in } \mathbb{R} \times (0,T] \quad \text{and} \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}. \tag{5.1.13}$$

As a special case of Theorem 5.1.3, the approximating paths and partitions may be taken to satisfy (5.1.12) with ρ_h given by

$$\lambda := \frac{(\rho_h)^{3/4} |\log \rho_h|^{1/2}}{h} < \frac{\theta}{\|H'\|_\infty}. \tag{5.1.14}$$

A second way of defining the partitions and paths may be achieved through the use of certain stopping times:

$$\begin{cases} T_0 := 0, \quad T_{k+1} := \inf \left\{ t > T_k : \max_{r,s \in [T_k,t]} |W(r) - W(s)| > \dfrac{h^{1/3}}{|\log h|^{2/3}} \right\}, \\ W_h(t) := W(T_k) + \dfrac{W(T_{k+1}) - W(T_k)}{T_{k+1} - T_k}(t - T_k) \quad \text{for } t \in [T_k, T_{k+1}), \\ M_h := \left\lceil \dfrac{\|H'\|_\infty}{(h|\log h|)^{2/3}} \right\rceil, \quad \text{and} \\ \mathcal{P}_h := \left\{ t_n := T_k + (n - kM_h)\dfrac{T_{k+1} - T_k}{M_h} : kM_h \le n < (k+1)M_h,\ k \in \mathbb{N}_0 \right\}. \end{cases} \tag{5.1.15}$$

Theorem 5.1.4. *Suppose that W is a Brownian motion, and assume that either*

$$\begin{cases} \mathcal{P}_h \text{ and } W_h \text{ are as in (5.1.12) with } \rho_h \text{ defined by (5.1.14), or} \\ \mathcal{P}_h \text{ and } W_h \text{ are as in (5.1.15).} \end{cases}$$

If u_h is constructed using (5.1.8) and (5.1.10), and u is the solution of (5.1.13), then there exists a deterministic constant $C > 0$ depending only on L and λ such that, with probability one,

$$\limsup_{h \to 0} \sup_{(x,t) \in \mathbb{R}^d \times [0,T]} \frac{|u_h(x,t) - u(x,t)|}{h^{1/3} |\log h|^{1/3}} \le C(1+T).$$

The final type of result involves convergence in distribution in the space $BUC(\mathbb{R}^d \times [0,T])$, with the topology of local uniform convergence.

For random variables $\{X_\delta\}_{\delta>0}$ and X taking values in some topological space $\mathcal{X}$, recall that X_δ is said to converge in distribution (or in law) to X and $\delta \to 0$ if the law ν_δ of X_δ on $\mathcal{X}$ converges weakly to the law ν of X. That is, for any bounded continuous function $\phi : \mathcal{X} \to \mathbb{R}$,

$$\lim_{\delta\to 0} \int_{\mathcal{X}} \phi \, d\nu_\delta = \int_{\mathcal{X}} \phi \, d\nu.$$

Below, the paths W_h are taken to be appropriately scaled simple random walks, and, as a consequence, W_h converges in distribution to a Brownian motion W (see for instance Billingsley [14]). This corresponds above to $\mathcal{X} = C([0,T], \mathbb{R}^m)$ and ν the Wiener measure on $\mathcal{X}$.

Let λ, ρ_h, W_h, and $\mathcal{P}_h$ be given, for some probability space $(\mathcal{A}, \mathcal{G}, \mathbf{P})$, by

$$(5.1.16) \qquad \begin{cases} \lambda := \dfrac{(\rho_h)^{3/4}}{h} \le \dfrac{\theta}{\|H'\|_\infty}, \quad M_h := \lfloor (\rho_h)^{-1/2} \rfloor, \\ \mathcal{P}_h := \{t_n\}_{n=0}^N = \{n\rho_h \wedge T\}_{n\in\mathbb{N}_0}, \\ \{\xi_n\}_{n=1}^\infty : \mathcal{A} \to \{-1, 1\} \text{ are independent}, \\ \mathbf{P}(\xi_n = 1) = \mathbf{P}(\xi_n = -1) = \dfrac{1}{2}, \quad W(0) = 0, \quad \text{and} \\ W_h(t) := W_h(kM_h\rho_h) + \dfrac{\xi_k}{\sqrt{M_h\rho_h}}(t - kM_h\rho_h) \\ \text{for } k \in \mathbb{N}_0, \ t \in [kM_h\rho_h, (k+1)M_h\rho_h). \end{cases}$$

Theorem 5.1.5. *If u_h is constructed using* (5.1.8) *and* (5.1.10) *with W_h and $\mathcal{P}_h$ as in* (5.1.16)*, and u is the solution of* (5.1.13)*, then, as $h \to 0$, u_h converges to u in distribution.*

5.1.2 Overview of the theory in the non-rough setting

The comparison principle for the equation

$$(5.1.17) \qquad u_t = F(D^2u, Du) \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d$$

gives monotonicity of the solution with respect to the initial condition. That is, if $u(\cdot,0) \leq v(\cdot,0)$, then $u(\cdot,t) \leq v(\cdot,t)$ for all future times $t > 0$.

In addition, (5.1.17) is stable under local uniform convergence: if, for $n \geq 0$, $u_{0,n}, u_0 \in BUC(\mathbb{R}^d)$, $F_n, F \in C(\mathbb{R}^d)$, $u_n \in BUC(\mathbb{R}^d \times [0,T])$ solves

$$u_{n,t} = F_n(D^2 u_n, Du_n) \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \quad u_n(\cdot,0) = u_{0,n} \quad \text{in } \mathbb{R}^d, \tag{5.1.18}$$

and, as $n \to \infty$,

$$u_{0,n} \to u_0 \quad \text{and} \quad F_n \to F \quad \text{locally uniformly,} \tag{5.1.19}$$

then, as $n \to \infty$, u_n converges locally uniformly to u, the viscosity solution of (5.1.17).

These properties can be summarized in terms of the solutions operators for (5.1.17), which are, for $t \geq 0$, the maps $S(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ for which the solution u of (5.1.17) is given by $u(x,t) = S(t)u_0(x)$. For all $s, t \geq 0$, $\phi, \psi \in BUC(\mathbb{R}^d)$, and $k \in \mathbb{R}$, these satisfy

$$\begin{cases} (a) \quad S(0)\phi = \phi, \\ (b) \quad S(t+s) = S(t)S(s), \\ (c) \quad S(t)(\phi + k) = S(t)\phi + k, \quad \text{and} \\ (d) \quad \sup_{\mathbb{R}^d} (S(t)\phi - S(t)\psi) \leq \sup_{\mathbb{R}^d} (\phi - \psi). \end{cases} \tag{5.1.20}$$

Property (5.1.20)(c) implies that (5.1.20)(d) is equivalent to the monotonicity of $S(t)$. That is, if $\phi \leq \psi$, then $S(t)\phi \leq S(t)\psi$ for all $t \geq 0$.

The stability property above can be rephrased as saying that, if (5.1.19) holds and if $S_n(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ is the family of solution operators corresponding to (5.1.18), then, as $n \to \infty$, $S_n(t)u_{0,n}(x) \to S(t)u_0(x)$ locally uniformly. The philosophy behind the creation of approximation schemes is to generalize this result, by constructing, for $h > 0$ and

$\rho > 0$, suitable operators $S_h(\rho) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ that satisfy properties similar to those in (5.1.20). In particular, for all $\phi \in BUC(\mathbb{R}^d)$ and $k \in \mathbb{R}$, and for some increasing function $h \mapsto \rho_h$ satisfying $\lim_{h\to 0} \rho_h = 0$,

$$
\text{(5.1.21)} \qquad \begin{cases} (a) \quad S_h(t)(\phi + k) = S_h(t)\phi + k, \\ (b) \quad \sup_{\mathbb{R}^d} (S_h(\rho)\phi - S_h(\rho)\psi) \le \sup_{\mathbb{R}^d} (\phi - \psi) \quad \text{whenever } 0 < \rho \le \rho_h, \text{ and} \\ (c) \quad \lim_{h\to 0} \sup_{0<\rho\le\rho_h} \left| \dfrac{S_h(\rho)\phi - \phi}{\rho} - F(D^2\phi, D\phi) \right| = 0 \quad \text{for all } \phi \in C^2(\mathbb{R}^d). \end{cases}
$$

Given a partition $\mathcal{P}_h$ satisfying $|\mathcal{P}_h| \le \rho_h$, the approximate solution $u_h : BUC(\mathbb{R}^d \times [0,T])$ is assembled by first setting $u_h(\cdot, 0) := u_0$ and then iteratively defining

$$
\text{(5.1.22)} \qquad u_h(\cdot, t) := S_h(t - t_n)u_h(\cdot, t_n) \quad \text{for } n = 0, 1, 2, \dots, N-1 \text{ and } t \in (t_n, t_{n+1}].
$$

One example of particular interest is the class of finite difference approximations, for which $S_h(\rho)u$ depends on the function u only through its values on the discrete lattice $h\mathbb{Z}^d$. A major consideration for such schemes is to establish a relationship between the resolutions of the discrete grids in time and space, that is, to choose the map $h \mapsto \rho_h$ in such a way that the properties in (5.1.21) are satisfied. Such a relationship is known as a Courant-Friedrichs-Lewy (CFL) condition [22], and various examples will be studied throughout the chapter.

As is well-known, solutions of (5.1.17) are generally not C^2 on all of $\mathbb{R}^d \times [0,T]$, even if F, H, and u_0 are all smooth, and so (5.1.21)(c) is not enough to prove the convergence of u_h to u as $h \to 0$. It is here that the monotonicity of $S_h(\rho)$, which is implied by (5.1.21)(a) and (b), is vital, since it allows the scheme operator to be applied to the smooth test functions coming from the definition of viscosity solutions.

A finite difference scheme operator S_h, in its simplest form, when $d = 1$ (the last assumption here made only to simplify the presentation), is given, for some $F_h \in C^{0,1}(\mathbb{R} \times \mathbb{R} \times \mathbb{R})$, by

$$
\text{(5.1.23)} \qquad \begin{aligned} &S_h(\rho)u(x) := u(x) \\ &+ \rho F_h\left(\frac{u(x+h)+u(x-h)-2u(x)}{h^2}, \frac{u(x+h)-u(x)}{h}, \frac{u(x)-u(x-h)}{h}\right). \end{aligned}
$$

The scheme (5.1.23) automatically satisfies (5.1.21)(a), while (5.1.21)(b) holds if the function

$$
\mathbb{R}\times\mathbb{R}\times\mathbb{R} \ni (u, u_-, u_+) \mapsto u + \rho F_h\left(\frac{u_+ + u_- - 2u}{h^2}, \frac{u_+ - u}{h}, \frac{u - u_-}{h}\right)
$$

is nondecreasing in each argument when $0 < \rho \leq \rho_h$. This will be the case if

$$
\text{(5.1.24)} \qquad \rho_h := \lambda h^2
$$

for some sufficiently small constant $\lambda > 0$. For the first-order equation

$$
\text{(5.1.25)} \qquad u_t = H(Du) \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{on } \mathbb{R}^d,
$$

the CFL condition becomes instead

$$
\text{(5.1.26)} \qquad \rho_h = \lambda h.
$$

The function F_h is related to F through a consistency requirement, which here means that, for all $X \in \mathbb{R}$ and $p \in \mathbb{R}$,

$$
\text{(5.1.27)} \qquad \lim_{h\to 0} F_h(X,p,p) = F(X,p) \quad \text{and} \quad \sup_{h>0} \|DF_h\|_\infty < \infty.
$$

Property (5.1.21)(c) can then be readily verified by using Taylor approximations to estimate the finite differences of functions $\phi \in C^2(\mathbb{R}^d)$.

An instructive example in the first-order setting is the following analogue of the Lax-

Friedrichs scheme for scalar conservation laws. Let $\varepsilon_h > 0$ and define, for $x \in \mathbb{R}$,

$$
\begin{aligned}
S_h(\rho)u(x) := u(x) + \rho \Bigg[H\left(\frac{u(x+h) - u(x-h)}{2h} \right) \\
+ \varepsilon_h \left(\frac{u(x+h) + u(x-h) - 2u(x)}{h^2} \right) \Bigg].
\end{aligned}
\tag{5.1.28}
$$

Here, H_h is given by

$$
H_h(p,q) = H\left(\frac{p+q}{2}\right) + \frac{\varepsilon_h}{h}(p-q).
$$

The final term in (5.1.28) is a discrete analogue of the small viscosity used in the method of vanishing viscosity to obtain weak solutions of first-order equations. It is used here to inject monotonicity into the scheme, because, if, for some fixed $\theta > 0$ and $\lambda > 0$, the small parameter ε_h is defined by

$$
\varepsilon_h := \frac{\theta h}{2\lambda}, \tag{5.1.29}
$$

then (5.1.21)(b) is satisfied as long as (5.1.26) holds with $\theta \le 1$ and $\lambda \le \frac{\theta}{\|H'\|_\infty}$.

In [23], Crandall and Lions found error estimates for this and and other explicit finite difference schemes for homogenous Hamilton-Jacobi equations. More precisely, it was proved for the above example that there exists a constant $C > 0$ depending only on $\|DH\|_\infty$, $\|Du_0\|_\infty$, and λ such that, if u_h is defined as in (5.1.22) and (5.1.28), and if u solves (5.1.25), then

$$
\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - u(x,t)| \le C(1+T)h^{1/2}.
$$

This same rate was later established by Souganidis [77] for both explicit and implicit finite difference schemes for equations with Lipschitz spatial and time dependence, and the same method was applied to study other approximations such as max-min representations and Trotter-Kato splitting formulas [78].

Barles and Souganidis [9] considered schemes for second order equations, using a shorter, qualitative proof of convergence relying on the method of half-relaxed limits from the classical viscosity solution theory. Kuo and Trudinger [46, 47] also investigated such schemes in great detail and constructed several examples. The question of estimating the rates of convergence for such approximations of second order equations was analyzed from many points of view. Barles and Jakobsen [10, 11, 12] achieved algebraic convergence rates for stochastic control problems, taking advantage of the fact that F is convex in that setting. Jakobsen [39, 40] and Krylov [43] also established rates of convergence for nonconvex problems under some restrictions on F. If F is uniformly elliptic, then rates of convergence can be found under very general assumptions using techniques from the regularity theory for fully nonlinear, uniformly elliptic equations, as exhibited by Caffarelli and Souganidis [19], and later by Turanova [80] for inhomogenous equations.

5.1.3 *Difficulties in the pathwise setting*

The lack of regularity for W complicates the task of constructing scheme operators for (5.1.1) that are both monotone and consistent. Consider, for example, modifying the Lax-Friedrichs scheme (5.1.28) for the equation

$$du = H(u_x) \cdot dW \quad \text{in } \mathbb{R} \times (0,T] \quad \text{and} \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}. \tag{5.1.30}$$

If W is sufficiently regular, then it is reasonable to define a time-inhomogenous scheme operator by

$$\begin{aligned} S_h(t,s)u(x) &:= u(x) + H\left(\frac{u(x+h)-u(x-h)}{2h}\right)(W(t)-W(s)) \\ &\quad + \varepsilon_h\left(\frac{u(x+h)+u(x-h)-2u(x)}{h^2}\right)(t-s). \end{aligned} \tag{5.1.31}$$

Proceeding as in the previous subsection, a simple calculation reveals that $S_h(t,s)$ is monotone for $0 \le t-s \le \rho_h$, if ρ_h and ε_h are such that, for some $\theta \le 1$,

$$\varepsilon_h := \frac{\theta h^2}{2(t-s)}$$

and

$$\lambda := \max_{|t-s|\le\rho_h} \frac{\operatorname{osc}(W,s,t)}{h} \le \lambda_0 := \frac{\theta}{\|H'\|_\infty}. \tag{5.1.32}$$

On the other hand, for any $s,t \in [0,T]$ with $|s-t|$ sufficiently small, spatially smooth solutions Φ of (5.1.30) have the expansion,

$$\begin{aligned}\Phi(x,t) &= \Phi(x,s) + H(\Phi_x(x,s))(W(t)-W(s)) \\ &\quad + H'(\Phi_x(x,s))^2\Phi_{xx}(x,s)(W(t)-W(s))^2 + O(|W(t)-W(s)|^3),\end{aligned} \tag{5.1.33}$$

so that, if $0 \le t-s \le \rho_h$, for some $C > 0$ depending only on H,

$$\begin{aligned}\sup_{\mathbb{R}} |S_h(t,s)\Phi(\cdot,s) - \Phi(\cdot,t)| &\le C \sup_{r\in[s,t]} \left\| D^2\Phi(\cdot,r)\right\|_\infty \left(|W(t)-W(s)|^2 + h^2\right) \\ &\le C \sup_{r\in[s,t]} \left\| D^2\Phi(\cdot,r)\right\|_\infty (1+\lambda_0^2)h^2.\end{aligned} \tag{5.1.34}$$

Therefore, in order for the scheme to have a chance of converging, ρ_h should satisfy

$$\lim_{h\to 0} \frac{h^2}{\rho_h} = 0. \tag{5.1.35}$$

Both (5.1.32) and (5.1.35) can be achieved when W is continuously differentiable, or merely Lipschitz, by setting

$$\rho_h := \frac{\lambda h}{\left\|\dot{W}\right\|_\infty}.$$

More generally, if $W \in C^{0,\alpha}([0,T])$ with $\alpha > \frac{1}{2}$, and if

$$(\rho_h)^\alpha := \frac{\lambda h}{[W]_{\alpha,T}}, \tag{5.1.36}$$

then both (5.1.32) and (5.1.35) are satisfied, since

$$\frac{h^2}{\rho_h} = \left(\frac{[W]_{\alpha,T} h^{2\alpha-1}}{\lambda} \right)^{1/\alpha} \xrightarrow{h \to 0} 0.$$

However, this approach fails as soon as the quadratic variation path

$$\langle W \rangle_T := \lim_{|\mathcal{P}| \to 0} \sum_{n=0}^{N-1} |W(t_{n+1}) - W(t_n)|^2$$

is non-zero, as (5.1.32) and (5.1.35) together imply that $\langle W \rangle_T = 0$. This rules out, for instance, the case where W is the sample path of a Brownian motion, or, more generally, any nontrivial semimartingale.

Motivated by the theory of rough differential equations, it is natural to explore whether the scheme operator (5.1.31) can be altered in some way to refine the estimate in (5.1.34), potentially allowing (5.1.35) to be relaxed and ρ_h to converge more quickly to zero as $h \to 0^+$. More precisely, the next term in the expansion (5.1.33) suggests taking $W \in C^{0,\alpha}([0,T], \mathbb{R}^m)$ with $\alpha > \frac{1}{3}$ (or more generally, W with p-variation for $p < 3$) and defining

$$\begin{aligned} S_h(t,s)u(x) &:= u(x) + H\left(\frac{u(x+h) - u(x-h)}{2h} \right) (W(t) - W(s)) \\ &+ \frac{1}{2} H' \left(\frac{u(x+h) - u(x-h)}{2h} \right)^2 \\ &\quad \cdot \left(\frac{u(x+h) + u(x-h) - 2u(x)}{h^2} \right) (W(t) - W(s))^2 \\ &+ \frac{\theta}{2} \left(u(x+h) + u(x-h) - 2u(x) \right). \end{aligned} \tag{5.1.37}$$

As can easily be checked, (5.1.37) is monotone as long as (5.1.32) holds,

$$\operatorname{Lip}(u) \le L, \quad \theta + \|H'\|_\infty \lambda^2 \le 1, \quad \text{and} \quad \lambda \le \frac{\theta}{\|H'\|_\infty (1 + 2L \|H''\|_\infty)}.$$

On the other hand, the error in (5.1.34) would then be of order $h^2 + |W(t) - W(s)|^3$, which again leads to a requirement like (5.1.35). This seems to indicate that we should also incorporate higher order corrections in (5.1.37) to deal with the second-order spatial derivatives of u. However, this will disrupt the monotonicity of the scheme in general, since it will no longer be possible to use discrete maximum principle techniques.

For this reason, we develop a more effective strategy that works for any continuous path. Namely, rather than modifying the scheme itself, we regularize the path W. If $\{W_h\}_{h>0}$ is a family of smooth paths converging uniformly, as $h \to 0$, to W, then $\langle W_h \rangle_T = 0$ for each fixed $h > 0$, and therefore, W_h and ρ_h can be chosen so that (5.1.32) and (5.1.35) hold for W_h rather than W. Various methods for implementing this procedure, both qualitative and quantitative, are explored throughout the chapter.

5.2 The general convergence result and applications

The constructions in this chapter rely on the properties of a family of scheme operators, indexed by $h > 0$, $s, t \in [0,T]$ with $s \le t$, and a path $\zeta \in C([0,T], \mathbb{R}^m)$:

$$S_h(t,s;\zeta) : (B)UC(\mathbb{R}^d) \to (B)UC(\mathbb{R}^d).$$

We assume throughout that S_h commutes with translations in both the independent and dependent variables, in order to reflect the corresponding translation invariance of (5.1.1). That is,

$$(5.2.1) \qquad S_h(t,s;\zeta)(u+k) = S_h(t,s;\zeta)u + k \quad \text{for all} \quad k \in \mathbb{R} \quad \text{and} \quad u \in (B)UC(\mathbb{R}^d),$$

and

$$S_h(t,s;\zeta)\circ\tau_v = \tau_v\circ S_h(t,s;\zeta) \quad \text{for all} \quad v\in\mathbb{R}^d,\ \text{where } \tau_v u(x) := u(x+v). \tag{5.2.2}$$

For a Hamiltonian H satisfying

$$H\in C^k(\mathbb{R}^d,\mathbb{R}^m) \quad \text{for some } k\geq 2 \tag{5.2.3}$$

and a fixed continuous path $W\in C([0,T],\mathbb{R}^m)$, we consider a family of paths $\{W_h\}_{h>0}\subset C([0,T],\mathbb{R}^m)$ and a partition width $\rho_h>0$ satisfying

$$h\mapsto\rho_h \text{ is increasing,} \quad \lim_{h\to 0}\|W_h-W\|_\infty = 0 = \lim_{h\to 0}\rho_h; \tag{5.2.4}$$

$$\begin{cases} \text{if } u_1\leq u_2 \text{ and } s,t\in[0,T] \text{ satisfy } 0\leq t-s\leq\rho_h, \\ \text{then } S_h(t,s;W_h)u_1\leq S_h(t,s;W_h)u_2; \end{cases} \tag{5.2.5}$$

and

$$\begin{cases} \text{if } I\subset\mathbb{R},\ \Phi_h\in C(I,C^k(\mathbb{R}^d)) \text{ is a solution of} \\ d\Phi_h = \displaystyle\sum_{i=1}^m H^i(D\Phi_h)\cdot dW_h^i \text{ in } \mathbb{R}^d\times I, \\ s_h,t_h\in I,\ 0\leq t_h-s_h\leq\rho_h,\ \phi\in C^k(\mathbb{R}^d),\ R>0, \\ \text{and } \displaystyle\lim_{h\to 0}\|\Phi_h(\cdot,s_h)-\phi\|_{C^k(\mathbb{R}^d)} = 0, \text{ then} \\ \displaystyle\lim_{h\to 0}\frac{S_h(t_h,s_h;W_h)\Phi_h(\cdot,s_h)(x)-\Phi_h(x,s_h)}{t_h-s_h} = F(D^2\phi(x),D\phi(x)) \\ \text{uniformly for } x\in\mathbb{R}^d \text{ and } \displaystyle\max_{j=2,3,\ldots,k}\left\|D^j\phi\right\|_\infty\leq R. \end{cases} \tag{5.2.6}$$

The scheme operator is used to build approximate solutions as follows. For a fixed path $\zeta\in C([0,T];\mathbb{R}^m)$, partition $\mathcal{P}=\{0=t_0<t_1<\cdots<t_N=T\}$ of $[0,T]$, and initial datum

$u_0 \in BUC(\mathbb{R}^d)$, define

$$
(5.2.7) \quad \begin{cases} \tilde{u}_h(\cdot, 0; \zeta, \mathcal{P}) := u_0 \quad \text{and} \\ \tilde{u}_h(\cdot, t; \zeta, \mathcal{P}) := S_h(t, t_n; \zeta) u_h(\cdot, t_n; \zeta, \mathcal{P}) \quad \text{for } n = 0, 1, \ldots, N-1, \ t \in (t_n, t_{n+1}]. \end{cases}
$$

Theorem 5.2.1. *Assume that $H \in C^k(\mathbb{R}^d, \mathbb{R}^m)$ and S_h, W_h, and ρ_h satisfy (5.2.1) - (5.2.6). Let $\{\mathcal{P}_h\}_{h>0}$ be a family of partitions of $[0,T]$ such that $|\mathcal{P}_h| \leq \rho_h$ for all $h > 0$, and define $u_h := \tilde{u}_h(\cdot; W_h, \mathcal{P}_h)$. Then, as $h \to 0$, u_h converges locally uniformly to the pathwise viscosity solution u of (5.1.1).*

The proof of Theorem 5.2.1, which, as in [9], makes use of the method of half-relaxed limits, will be postponed until the end of this section. In the following sub-sections, we demonstrate its utility in a variety of contexts.

5.2.1 Finite difference schemes

Define, for $x \in \mathbb{R}^d$ and $y \in \mathbb{Z}^d \backslash \{0\}$, the discrete derivatives

$$
(5.2.8) \quad \begin{aligned} &D^+_{h,y} u(x) := \frac{u(x+hy) - u(x)}{h|y|}, \quad D^-_{h,y} u(x) := \frac{u(x) - u(x-hy)}{h|y|}, \\ &\text{and} \quad D^2_{h,y} u(x) := D^+_{h,y} D^-_{h,y} u(x) = \frac{u(x+hy) + u(x-hy) - 2u(x)}{h^2 |y|^2}. \end{aligned}
$$

Observe that there exists a universal constant $C > 0$ such that, if $u \in C^{1,1}(\mathbb{R}^d)$, $h > 0$, and $y \in \mathbb{Z}^d \backslash \{0\}$, then

$$
(5.2.9) \quad \left\| D^{\pm}_{h,y} u - Du \cdot \frac{y}{|y|} \right\|_\infty \leq C \left\| D^2 u \right\|_\infty h,
$$

and, if $u \in C^2(\mathbb{R}^d)$,

$$
(5.2.10) \quad \left\| D^2_{h,y} u - D^2 u \frac{y}{|y|} \cdot \frac{y}{|y|} \right\|_\infty \leq C \sup_{|x_1 - x_2| \leq h} \left| D^2 u(x_1) - D^2 u(x_2) \right|.
$$

For some fixed $N \in \mathbb{N}$, define

$$\begin{cases} \mathbb{Z}^d_N := \left\{ y \in \mathbb{Z}^d : \max_{i=1,2,\dots,d} |y_i| \le N \right\}, \\ D^{\pm}_{h,N} := \{D^{\pm}_{h,y}\}_{y \in \mathbb{Z}^d_N \setminus \{0\}}, \quad D_{h,N} := \left(D^+_{h,N} \; D^-_{h,N} \right), \quad \text{and} \\ D^2_{h,N} := \{D^2_{h,y}\}_{y \in \mathbb{Z}^d_N \setminus \{0\}}. \end{cases}$$

Then, for some given functions

$$\begin{cases} H_h \in C^{0,1}(\mathbb{R}^{(2N+1)^d - 1} \times \mathbb{R}^{(2N+1)^d - 1} \times \mathbb{R}) \quad \text{and} \\ F_h \in C^{0,1}(\mathbb{R}^{2N+1)^d - 1} \times \mathbb{R}^{(2N+1)^d - 1} \times \mathbb{R}^{(2N+1)^d - 1}), \end{cases}$$

the scheme operators for finite difference approximations take the form

$$S_h(t,s;\zeta)u(x) := u(x) + F_h\left(D^2_h u(x), D_h u(x)\right)(t-s) + H_h\left(D_h u(x), \zeta(t) - \zeta(s)\right). \tag{5.2.11}$$

Properties (5.2.1) and (5.2.2) are immediate, while the question of whether (5.2.11) satisfies (5.2.5) or (5.2.6) is reduced to routine calculations involving F_h and H_h.

Hamilton-Jacobi equations

We first study the first-order setting, for which $F = F_h = 0$. We assume that

$$H^i = H^i_1 - H^i_2 \text{ for convex } H^i_1, H^i_2 : \mathbb{R}^d \to \mathbb{R} \text{ nonnegative,} \tag{5.2.12}$$

$$\|DH\|_\infty < \infty, \tag{5.2.13}$$

and, for some $C = C_L > 0$,

$$
(5.2.14) \quad \begin{cases} |DH_h(\cdot, \Delta\zeta)| \le C\left(|\Delta\zeta| + h\right) & \text{for all } h > 0 \text{ and } \Delta\zeta \in \mathbb{R}^m, \text{ and} \\ H_h\left(p, p, \Delta\zeta\right) = \displaystyle\sum_{i=1}^m H^i(p)(\Delta\zeta)^i & \text{for all } h > 0,\ p \in \mathbb{R}^{(2N+1)^d - 1}, \text{ and } \Delta\zeta \in \mathbb{R}^m. \end{cases}
$$

In order for monotonicity to hold, more precise bounds for the derivatives of H_h are required. Let elements of $\mathbb{R}^{(2N+1)^d-1}$ be labeled by $\{p_y\}_{y \in \mathbb{Z}_N^d \setminus \{0\}}$, and assume that, for some $C = C_L > 0$, $\theta \in [0,1]$, and $\lambda_0 > 0$,

$$
(5.2.15) \qquad \begin{cases} \displaystyle\sum_{y \in \mathbb{Z}_N^d \setminus \{0\}} \frac{1}{|y|}\left(\frac{\partial H_h}{\partial q_y} - \frac{\partial H_h}{\partial p_y}\right) \le \frac{1-\theta}{\lambda_0}\left|\Delta\zeta\right| + \theta h \quad \text{and} \\ \displaystyle\frac{\partial H_h}{\partial q_y} - \frac{\partial H_h}{\partial p_{-y}} \ge C\left(h - \frac{|\Delta\zeta|}{\lambda_0}\right) \quad \text{for all } y \in \mathbb{Z}_N^d \setminus \{0\}. \end{cases}
$$

Lemma 5.2.1. *Suppose that H_h satisfies* (5.2.14)*. Then there exists $C = C_L > 0$ such that, whenever $\zeta \in C([0,T], \mathbb{R}^m)$, $\operatorname{osc}(\zeta, s, t) \le \lambda_0 h$ for some $s, t \in I$, and $\Phi \in C(I, C^{1,1}(\mathbb{R}^d))$ is a solution of*

$$
d\Phi = \sum_{i=1}^m H^i(D\Phi) \cdot d\zeta^i \quad \text{in } \mathbb{R}^d \times I,
$$

then

$$
\left\|S_h(t,s;\zeta)\Phi(\cdot,s) - \Phi(\cdot,t)\right\|_\infty \le C\left\|D^2\Phi\right\|_\infty h^2.
$$

If, in addition, H_h satisfies (5.2.15)*, then, whenever $u_1, u_2 \in (B)UC(\mathbb{R}^d)$ with $u_1 \le u_2$ and $\operatorname{osc}(\zeta, s, t) \le \lambda_0 h$,*

$$
S_h(t,s;\zeta)u_1 \le S_h(t,s;\zeta)u_2.
$$

Motivated by the above result, the schemes for first-order equations in Section 5.3, for which we obtain explicit error estimates, will be assumed to satisfy the conclusions of Lemma 5.2.1.

Proof of Lemma 5.2.1. Let $\Phi \in C(I, C^{1,1}(\mathbb{R}^d))$ be as in the statement of the lemma. Then there exists $C = C_L > 0$ such that, for all $s, t \in I$,

$$\left\| \Phi(\cdot, t) - \Phi(\cdot, s) - \sum_{i=1}^{m} H^i(D\Phi(\cdot, s)) \left(\zeta^i(t) - \zeta^i(s) \right) \right\|_\infty \le C \left\| D^2 \Phi \right\|_\infty |\zeta(t) - \zeta(s)|^2 .$$

Therefore,

$$\begin{aligned} |S_h(t, s; \zeta)\Phi(\cdot, s)(x) - \Phi(x, t)| &\le C \left\| D^2\Phi \right\|_\infty \left(h^2 + |\zeta(t) - \zeta(s)|\, h + |\zeta(t) - \zeta(s)|^2 \right) \\ &\le C(1 + \lambda_0 + \lambda_0^2) h^2. \end{aligned}$$

Meanwhile, if $\mathcal{S}_h : \mathbb{R}^{(2N+1)^d} \to \mathbb{R}$ is the map implicitly defined by

$$\mathcal{S}_h \left(\{u(x + y)\}_{y \in \mathbb{Z}^d_N} \right) = S_h(t, s; \zeta) u(x),$$

then (5.2.15) implies that $\mathcal{S}_h$ is increasing in each of its arguments as long as $\operatorname{osc}(\zeta, s, t) \le \lambda_0 h$. $\square$

We now mention two specific examples. The first is the analogue of the Lax-Friedrichs scheme discussed in the introduction. Here, H_h is defined, for some $\theta \in (0, 1]$, by

$$H_h(p, q, \Delta\zeta) := H \left(\frac{p + q}{2} \right) \Delta\zeta + \frac{\theta h}{2d} \sum_{k=1}^{d} (q_k - p_k) ,$$

where the vector $(p, q) \in \mathbb{R}^d \times \mathbb{R}^d$ stands for the discrete derivatives

$$\begin{aligned} &p = \left(D^+_{h,e_1}, D^+_{h,e_2}, \ldots, D^+_{h,e_d} \right), \quad q = \left(D^-_{h,e_1}, D^-_{h,e_2}, \ldots, D^-_{h,e_d} \right), \\ &e_k := (0, 0, \ldots, 0, \underbrace{1}_{k}, 0, \ldots, 0) \quad \text{for } k = 1, 2, \ldots, d. \end{aligned}$$

Then (5.2.14) and (5.2.15) are satisfied with $\lambda_0 := \frac{\theta}{d\|DH\|_\infty}$.

If $d = 1$, the different intervals of monotonicity of H may be exploited to create "upwind" schemes. As a simple example, assume that $H \geq H(0) = 0$, H is increasing for $p > 0$, and decreasing for $p < 0$, and define

$$H_h(p, q, \Delta\zeta) := [H(p_+) + H(-q_-)]\,(\Delta\zeta)_+ - [H(q_+) + H(-p_-)]\,(\Delta\zeta)_-.$$

Then (5.2.14) and (5.2.15) hold with $\theta = 0$ and $\lambda_0 := \frac{1}{2\|H'\|_\infty}$.

As far as the approximating paths W_h are concerned, Lemma 5.2.1 implies that (5.2.5) and (5.2.6) will hold, with $k = 2$, if ρ_h and W_h satisfy

(5.2.16)
$$\sup_{0 \leq t-s \leq \rho_h} |W_h(t) - W_h(s)| \leq \lambda_0 h \quad \text{and} \quad \lim_{h \to 0} \frac{h^2}{\rho_h} = 0.$$

If W_h is piecewise-smooth, then

$$\sup_{0 \leq t-s \leq \rho_h} |W_h(t) - W_h(s)| \leq \left\| \dot{W}_h \right\|_\infty \rho_h.$$

For definiteness, we now take W_h to be the piecewise linear interpolation of W with stepsize η_h, for some increasing function $h \mapsto \eta_h$ satisfying $\lim_{h \to 0^+} \eta_h = 0$. Then, if $\omega : [0, \infty) \to [0, \infty)$ is the modulus of continuity for W, for some $C > 0$, we have

(5.2.17)
$$\left\| \dot{W}_h \right\|_\infty \leq C \frac{\omega(\eta_h)}{\eta_h}.$$

Then the first part of (5.2.16) may be replaced with the slightly stronger assumption

(5.2.18)
$$\frac{C\omega(\eta_h)\rho_h}{h\eta_h} \leq \lambda_0.$$

To be more explicit, suppose that $W \in C^{0,\alpha}([0, T], \mathbb{R}^m)$ and $\eta_h = (\rho_h)^\gamma$ for some $\gamma > 0$. Then (5.2.18) will hold if ρ_h is defined by

$$\lambda := \frac{C[W]_{\alpha,T}(\rho_h)^{1-\gamma+\alpha\gamma}}{h} \leq \lambda_0.$$

This yields

$$\frac{h^2}{\rho_h} \approx (\rho_h)^{1-2\gamma+2\alpha\gamma},$$

so that (5.2.16) will be satisfied if

$$0 < \gamma < \frac{1}{2(1-\alpha)}.$$

If $\alpha > \frac{1}{2}$, then γ is allowed to be 1, and in particular, it is natural to define W_h to be the piecewise linear interpolation of W on a partition of step-size $\eta_h = \rho_h$. Notice also that such paths have trivial quadratic variation.

However, for $\alpha \leq \frac{1}{2}$, γ is forced to be less than 1, and so we must make W_h a milder approximation. The work in the subsequent sections suggests that choosing $\gamma = \frac{1}{2}$ gives the best rate of convergence regardless of the regularity of the path W.

A second-order example

Verifying (5.2.5) and (5.2.6) is more complicated for finite difference approximations of second-order equations. Rather than stating very general assumptions on F_h or H_h, we perform these calculations for a specific scheme. More examples can be formed by adapting the methods of [46, 47].

Assume for simplicity that $d = 1$, $H \in C^3(\mathbb{R}, \mathbb{R}^m)$, and that F depends only on u_{xx}, and define, for some $\varepsilon_h > 0$,

$$\begin{cases} H_h(p, q, \Delta\zeta) := H\left(\dfrac{p+q}{2}\right)\Delta\zeta \quad \text{and} \quad F_h(X) = F(X) + \varepsilon_h X \\ \text{for} \quad X = D^2_{h,1}u \text{ and } (p, q) = D_{h,1}u. \end{cases}$$

Note that F is increasing, and so S_h, W_h, and ρ_h satisfy (5.2.5) if

$$\rho_h := \lambda h^2 \quad \text{with} \quad \lambda \leq \frac{1}{2\left\|F'\right\|_\infty}$$

and

$$\left\|\dot{W}_h\right\|_\infty \leq \frac{2}{\|H'\|_\infty} \cdot \frac{\varepsilon_h}{h}. \tag{5.2.19}$$

Now let $\Phi_h \in C(I, C^3(\mathbb{R}))$ and $\phi \in C^3(\mathbb{R})$ be as in (5.2.6). Observe that it is possible to find such a solution because of the added regularity for H, and, moreover,

$$\sup_{h>0} \sup_{t\in I} \left(\left\|\Phi_{h,xx}(\cdot, t)\right\|_\infty + \left\|\Phi_{h,xxx}(\cdot, t)\right\|_\infty \right) < \infty.$$

Then, for some $C > 0$ depending only on $\left\|H'\right\|_\infty$, and for all $\rho \in (0, \lambda h^2)$,

$$\begin{aligned} &\left| \Phi_h(x, t+\rho) - \Phi_h(x,t) - \sum_{i=1}^m H^i(\Phi_{h,x}(x,t))(W_h(t+\rho) - W_h(t)) \right| \\ &\leq C \left(\left\|\Phi_{h,xx}\right\|_\infty \max_{t\leq s\leq t+\rho} |W_h(s) - W_h(t)|^2 \right) \leq C\lambda \left\|\Phi_{h,xx}\right\|_\infty (\varepsilon_h)^2 \rho. \end{aligned}$$

The estimates (5.2.9) and (5.2.10) then imply that, for s_h and t_h as in (5.2.6),

$$\begin{aligned} &|S_h(t_h, s_h; W_h)\Phi_h(\cdot, s_h)(x) - \Phi_h(x, t_h) - (t_h - s_h)F(\phi_{xx}(x,t))| \\ &\leq C\rho_h \cdot \left(\left\|\Phi_{h,xx}\right\|_\infty \varepsilon_h + \left\|\Phi_{h,xxx}\right\|_\infty h + \left\|\Phi_{h,xx}(\cdot, s_h) - \phi_{xx}\right\|_\infty \right), \end{aligned}$$

and so (5.2.6) holds if $\lim_{h\to 0} \varepsilon_h = 0$. This, in turn, requires that

$$\lim_{h\to 0} h \left\|\dot{W}_h\right\|_\infty = 0,$$

or that W_h satisfies (5.2.17) with η_h such that

$$\lim_{h\to 0} \frac{h\omega(\eta_h)}{\eta_h} = 0.$$

Taking $W \in C^{0,\alpha}([0,T], \mathbb{R}^m)$ and $\eta_h = (\rho_h)^\gamma = \lambda^\gamma h^{2\gamma}$ for some $\gamma > 0$ as a concrete example, this leads once more to the restriction

$$0 < \gamma < \frac{1}{2(1-\alpha)}.$$

5.2.2 Other approximations

Stability for (5.1.1)

Theorem 5.2.1 may be used to obtain an alternative verification of the stability properties for (5.1.1). Suppose that

$$(5.2.20) \qquad \begin{cases} u_0^\varepsilon \in C^{0,1}(\mathbb{R}^d), \quad W^\varepsilon, W \in C([0,T];\mathbb{R}^m), \quad H^\varepsilon, H \in C^2(\mathbb{R}^d;\mathbb{R}^m), \\ F^\varepsilon, F \in C(\mathbb{S}^d \times \mathbb{R}^d) \text{ are degenerate elliptic,} \\ \text{and } \lim_{\varepsilon \to 0} \left(\|u_0^\varepsilon - u_0\|_\infty, \|W^\varepsilon - W\|_\infty, \|H^\varepsilon - H\|_{C^2}, \|F^\varepsilon - F\|_\infty \right) = 0, \end{cases}$$

and let $u^\varepsilon \in BUC(\mathbb{R}^d \times [0,T])$ be the unique solution of

$$(5.2.21) \qquad \begin{cases} du^\varepsilon = F^\varepsilon(D^2u^\varepsilon, Du^\varepsilon)\, dt + \sum_{i=1}^m H^{i,\varepsilon}(Du^\varepsilon) \cdot dW^{i,\varepsilon} \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \\ u^\varepsilon(\cdot, 0) = u_0^\varepsilon \quad \text{in } \mathbb{R}^d. \end{cases}$$

Theorem 5.2.2. *Assume* (5.2.20) *and let u^ε and u solve respectively* (5.2.21) *and* (5.1.1). *Then, as $\varepsilon \to 0$, u^ε converges locally uniformly to u.*

Proof. The comparison principle implies that the solution operator for (5.2.21) is contractive, and therefore, it suffices to assume that $u_0^\varepsilon = u_0$ for all $\varepsilon > 0$.

For $s \le t$, $\zeta \in C([0,T];\mathbb{R}^m)$, and $h > 0$, let $S_\varepsilon(t,s;\zeta) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ be the solution operator for (5.2.21) driven by the path ζ instead of W^ε. Properties (5.2.1) and (5.2.2) are readily verified, and, letting $\rho_h = \rho_\varepsilon$ be arbitrary and setting $W_h = W^\varepsilon$, (5.2.5) follows immediately from the comparison principle.

Finally, in view of the uniform bound for D^2H^ε, for any interval $I \subset [0,T]$ and solution

$\Phi \in C(I, C^2(\mathbb{R}^d))$ of (5.1.3), there exists a family of solutions $\Phi^\varepsilon \in C(I, C^2(\mathbb{R}^d))$ solving (5.1.3) with the Hamiltonian H^ε and path W^ε, converging in $C(I, C^2(\mathbb{R}^d))$ to Φ as $\varepsilon \to 0$. This can be verified with the method of characteristics, as in Section 2.2 (recall that, since H is independent of x, we are in the setting where (2.2.7) holds). Therefore, (5.2.6) is a consequence of Definition 2.3.1 and the local uniform convergence of F^ε to F. Theorem 5.2.1 now gives the result. □

A splitting formula

It is also possible to derive general Trotter-Kato type splitting formulas for (5.1.1). Here, we present a specific example.

Assume that

$$F \in C^{1,1}(\mathbb{S}^d \times \mathbb{R}^d) \quad \text{and} \quad H \in C^4(\mathbb{R}^d, \mathbb{R}^m),$$

and, for $\zeta \in C([0,T], \mathbb{R}^m)$, let $S_F(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ and $S_H(t,s;\zeta) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ be the solution operators for respectively

$$u_t = F(D^2u, Du) \quad \text{and} \quad du = \sum_{i=1}^m H^i(Du) \cdot d\zeta^i.$$

Define

$$S_h(t,s;\zeta) = S_F(t-s)S_H(t,s;\zeta).$$

Theorem 5.2.3. *For any sequence of approximating paths* $\{W_h\}_{h>0}$ *and modulus* $h \to \rho_h$ *satisfying* (5.2.4)*, the triple* (S_h, W_h, ρ_h) *satisfies* (5.2.1) - (5.2.6).

Proof. Properties (5.2.1) - (5.2.5) are immediate from the definitions of the above objects. Let $I \subset [0,T]$, $s_h, t_h \in I$, $\Phi_h \in C(I; C^4(\mathbb{R}^d))$, and $\phi \in C^4(\mathbb{R}^d)$ be as in (5.2.6). Such a solution Φ exists in view of the additional regularity assumed for H.

For any $x \in \mathbb{R}^d$,

$$S_h(t_h, s_h; W_h)\Phi_h(\cdot, s_h)(x) - \Phi_h(x, t_h) = S_F(t_h - s_h)\Phi_h(\cdot, t_h)(x) - \Phi_h(x, t_h).$$

Define $\phi_h := \Phi_h(\cdot, t_h)$, which satisfies

$$R := \sup_{h>0} \|\phi_h\|_{C^4(\mathbb{R}^d)} < \infty \quad \text{and} \quad \lim_{h\to 0} \|\phi_h - \phi\|_{C^2(\mathbb{R}^d)} = 0,$$

and let

$$v_h(x,t) := \phi_h(x) + tF(D^2\phi_h(x), D\phi_h(x)).$$

Then, for some universal constant $C > 0$, v_h is a viscosity super-solution of the equation

$$v_{h,t} \geq F(D^2 v_h, Dv_h) - C\,\|F\|_{C^{1,1}(\mathbb{R}^d)}\,R\rho_h \quad \text{in } \mathbb{R}^d \times [0, \rho_h],$$

so that, for all $\rho \in (0, \rho_h)$,

$$\sup_{x\in\mathbb{R}^d} \left(S_F(\rho)\phi_h(x) - v_h(x,\rho)\right) \leq C\,\|F\|_{C^{1,1}(\mathbb{R}^d)}\,R\rho_h\rho.$$

A similar argument, using that v_h satisfies an analogous viscosity sub-solution property, gives a lower bound, whence

$$\left|S_F(t_h - s_h)\phi_h(x) - \phi_h(x) - (t_h - s_h)F(D^2\phi_h(x), D\phi_h(x))\right| \leq C\,\|F\|_{C^{1,1}(\mathbb{R}^d)}\,R\rho_h(t_h - s_h).$$

Property (5.2.6) now follows, with $k = 4$, from the fact that

$$\lim_{h\to 0} F(D^2\phi_h, D\phi_h) = F(D^2\phi, D\phi) \quad \text{uniformly.}$$

□

5.2.3 The proof of Theorem 5.2.1

Because the proof is presented for the case when H satisfies (5.2.3), solutions of equation (5.1.1) are in the sense of Definition 2.3.1. The proof can be modified to treat the case when H is not smooth and instead only satisfies (2.4.4), but we do not pursue this here.

Define

$$u^\star(x,t) = \limsup_{h\to 0,(y,s)\to(x,t)} u_h(y,s) \quad \text{and} \quad u_\star(x,t) = \liminf_{h\to 0,(y,s)\to(x,t)} u_h(y,s)$$

(not to be confused with the notation u^* and u_* for upper- and lower-semicontinuous envelopes). The functions $u^\star$ and $u_\star$, called the half-relaxed limits of u_h, are respectively upper- and lower- semicontinuous. Furthermore, $u_\star \le u^\star$ on $\mathbb{R}^d \times [0,T]$ and $u_\star(\cdot,0) \le u_0 \le u^\star(\cdot,0)$ on $\mathbb{R}^d$. The goal will be to show that $u_\star = u^\star$, which yields the local uniform convergence of u_h and the fact that the limit u solves (5.1.1).

Step 1: Finiteness of $u^\star$ and $u_\star$. Observe that, for any constant $k \in \mathbb{R}$, the function

$$\Phi_h(x,t) = k + \sum_{i=1}^m H^i(0) W_h^i(t)$$

is a smooth solution of (5.1.3) for all $(x,t) \in \mathbb{R}^d \times [0,T]$. Therefore, in view of (5.2.5) and (5.2.6),

$$u_h(x,t) \le \|u_0\|_\infty + \sum_{i=1}^m H^i(0) W_h^i(t) + T(F(0,0)+1)$$

for all sufficiently small $h > 0$, and so $u^\star(x,t) < \infty$ for all $(x,t) \in \mathbb{R}^d \times [0,T]$. A similar argument gives $u_\star > -\infty$.

Step 2: The solution inequalities. In this step, we demonstrate that $u^\star$ and $u_\star$ satisfy respectively the sub- and super-solution properties in Definition 2.3.1 for equation (5.1.1). Only the argument for $u^\star$ is presented, since the proof for $u_\star$ is similar.

Assume that $(x_0,t_0) \in \mathbb{R}^d \times (0,T]$, $I \ni t_0$, $\psi \in C^1([0,T])$, $\Phi \in C(I, C^k(\mathbb{R}^d))$ solves

(5.1.3) with

$$\max_{j=2,3,\ldots,k} \sup_{t\in I} \left\| D^j \Phi(\cdot, t) \right\|_\infty < \infty,$$

and $u^\star(x,t) - \Phi(x,t) - \psi(t)$ attains a local maximum at (x_0, t_0). In view of Lemma 2.3.1(a), by adding an extra quadratic term, it may be assumed that this maximum is strict in $\mathbb{R}^d \times I$, and that

$$\lim_{|x|\to+\infty} \inf_{t\in I} \frac{\Phi(x,t)}{|x|} = +\infty. \tag{5.2.22}$$

The definition of $u^\star$ implies that there exist $y_h \in \mathbb{R}^d$ and $s_h \in [0,T]$ such that

$$\lim_{h\to 0} (y_h, s_h, u_h(y_h, s_h)) = (x_0, t_0, u^\star(x_0, t_0)).$$

The method of characteristics from Section 2.2, and the fact that $\lim_{h\to 0} \|W_h - W\|_\infty = 0$, yield the existence of a subinterval of I containing t_0, relabeled as I for convenience, such that, for all $h > 0$, there exists a solution $\Phi_h \in C(I, C^k(\mathbb{R}^d))$ of

$$d\Phi_h = \sum_{i=1}^m H^i(D\Phi_h) \cdot dW_h \quad \text{in } \mathbb{R}^d \times I \quad \text{and} \quad \Phi_h(\cdot, t_0) = \Phi(\cdot, t_0) \quad \text{in } \mathbb{R}^d$$

that satisfies (5.2.22) uniformly in h, and Φ_h converegs to Φ in $C(I, C^k(\mathbb{R}^d))$ as $h \to 0$. It follows that

$$u_h(x,t) - \Phi_h(x,t) - \psi(t)$$

attains a global maximum at $(\hat{y}_h, \hat{s}_h)$ over $\mathbb{R}^d \times \overline{I}$ such that $\{\hat{y}_h\}_{h>0}$ is bounded. This gives, in particular,

$$u_h(y_h, s_h) - \Phi_h(y_h, s_h) - \psi(s_h) \le u_h(\hat{y}_h, \hat{s}_h) - \Phi_h(\hat{y}_h, \hat{s}_h) - \psi(\hat{s}_h).$$

Let $(\hat{x}, \hat{t})$ be a limit point of the sequence $\{(\hat{y}_h, \hat{s}_h)\}_{h>0}$. Taking $h \to 0$ along the appropriate subsequence above results in the inequality

$$u^\star(x_0,t_0) - \Phi(x_0,t_0) - \psi(t_0) \le u^\star(\hat{x},\hat{t}) - \Phi(\hat{x},\hat{t}) - \psi(\hat{t}).$$

The strictness of the original maximum then implies that $\lim_{h\to 0}(\hat{y}_h,\hat{s}_h) = (x_0,t_0)$.

Because $|\mathcal{P}_h| \le \rho_h \xrightarrow{h\to 0} 0$, it follows that, for sufficiently small h, there exists $t_n \in \mathcal{P}_h$ such that $t_n < \hat{s}_h \le t_{n+1}$ and $t_n \in I$. Then, for all $x \in \mathbb{R}^d$,

$$u_h(x,t_n) \le u_h(\hat{y}_h,\hat{s}_h) + \Phi_h(x,t_n) - \Phi_h(\hat{y}_h,\hat{s}_h) + \psi(t_n) - \psi(\hat{s}_h). \tag{5.2.23}$$

Applying the operator $S_h(\hat{s}_h,t_n;W_h)$ to both sides of (5.2.23), using (5.2.5) and the fact that $0 < \hat{s}_h - t_n \le \rho_h$, and rearranging terms yields

$$\frac{\psi(\hat{s}_h) - \psi(t_n)}{\hat{s}_h - t_n} \le \frac{S_h(\hat{s}_h,t_n;W_h)\Phi_h(\cdot,t_n)(\hat{y}_h) - \Phi_h(\hat{y}_h,\hat{s}_h)}{\hat{s}_h - t_n}.$$

Sending $h \to 0$ and using (5.2.6) gives $\psi'(t_0) \le F(D^2\Phi(x_0,t_0), D\Phi(x_0,t_0))$, as desired.

Step 3: Initial data. We now prove that $u^\star(x,0) = u_0(x) = u_\star(x,0)$. Only the first equality is considered, and since $u^\star(x,0) \ge u_0(x)$, it suffices to show that $u^\star(x,0) \le u_0(x)$.

Let $\phi \in C^k(\mathbb{R}^d)$ be such that

$$R := \max_{j=2,3,\dots,k} \left\| D^j \phi \right\|_\infty < \infty$$

and $u_0 \le \phi$ on $\mathbb{R}^d$, and let $I \ni 0$ and $\Phi \in C(I, C^k(\mathbb{R}^d))$ be a solution of (5.1.3) with $\Phi(\cdot,0) = \phi$. Define $\phi_h \in C^{0,1}(\mathbb{R}^d \times [0,T])$ as in (5.2.7) with the initial condition $\phi_h(\cdot,0) = \phi$, path W_h, and partition $\mathcal{P}_h$. Then (5.2.5) and (5.2.6) yield, for some $C > 0$ depending only on R and $\|DF\|_\infty$, and for any $(y,s) \in \mathbb{R}^d \times I$ and sufficiently small h,

$$u_h(y,s) \le \phi_h(y,s) \le \Phi(y,s) + Cs.$$

Sending $(y,s) \to (x,0)$ and $h \to 0$, this becomes $u^\star(x,0) \le \phi(x)$, completing the argument since ϕ was arbitrary.

Step 4: The comparison principle. In view of the comparison principle, $u^\star(x,t) \le u_\star(x,t)$ for all $(x,t) \in \mathbb{R}^d \times [0,T]$. Therefore $u^\star = u_\star$, and the result is proved.

5.3 Convergence rates for first-order equations

In this last section, we focus on deriving quantitative error estimates for schemes in the first-order setting.

5.3.1 The pathwise error estimate

We first obtain an estimate for the error between the viscosity solution $\tilde{u}$ of

$$\tilde{u}_t = \sum_{i=1}^m H^i(D\tilde{u})\dot{\zeta}^i(t) \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \quad \tilde{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d \tag{5.3.1}$$

and the approximate solution $\tilde{u}_h(\cdot;\zeta,\mathcal{P})$ given by (5.2.7), which, for convenience, we define again here:

$$\begin{cases} \tilde{u}_h(\cdot,0;\zeta,\mathcal{P}) := u_0 \quad \text{and} \\ \tilde{u}_h(\cdot,t;\zeta,\mathcal{P}) := S_h(t,t_n;\zeta)u_h(\cdot,t_n;\zeta,\mathcal{P}) \quad \text{for } n = 0,1,\dots,N-1,\ t \in (t_n,t_{n+1}]. \end{cases} \tag{5.3.2}$$

We will henceforth always assume that $\mathrm{Lip}(u_0) \le L$, and that the Hamiltonians satisfy (5.2.12) and (5.2.13). Also, in addition to (5.2.1) and (5.2.2), the schemes in this part of the chapter will be required to satisfy the following quantitative versions of (5.2.5) and (5.2.6): for some $\lambda_0 > 0$,

$$\text{if } u_1 \le u_2 \text{ and } \operatorname{osc}(\zeta,s,t) \le \lambda_0 h, \quad \text{then } S_h(t,s;\zeta)u_1 \le S_h(t,s;\zeta)u_2, \tag{5.3.3}$$

and

$$
(5.3.4) \quad \begin{cases} \text{there exists } C = C_L > 0 \text{ such that, if } \zeta \in C([0,T],\mathbb{R}^m),\ \Phi \in C(I, C^{1,1}(\mathbb{R}^d)) \\ \text{is a solution of } d\Phi = \displaystyle\sum_{i=1}^m H^i(D\Phi)\cdot d\zeta^i \text{ in } \mathbb{R}^d \times I, \text{ and } \operatorname{osc}(\zeta, s, t) \le \lambda_0 h, \text{ then} \\ \|S_h(t,s;\zeta)\Phi(\cdot,s) - \Phi(\cdot,t)\|_\infty \le C \left\| D^2\Phi \right\|_\infty h^2. \end{cases}
$$

This is motivated by the properties obtained in Lemma 5.2.1 for the finite difference approximations discussed in subsection 5.2.1.

Fix a partition

$$
\mathcal{P} = \{0 = t_0 < t_1 < t_2 < \cdots < t_N = T\}
$$

of $[0,T]$, set $(\Delta t)_n := t_{n+1} - t_n$, and let $\zeta : [0,T] \to \mathbb{R}^m$ be any continuous path satisfying

$$
(5.3.5) \quad \begin{cases} \zeta(0) = 0, \quad \zeta \text{ is affine on } [t_n, t_{n+1}] \text{ for every } n = 0,1,2,\dots,N-1, \text{ and} \\ \displaystyle\max_{n=0,1,2,\dots,N-1} |\zeta(t_{n+1}) - \zeta(t_n)| \le \lambda_0 h. \end{cases}
$$

Theorem 5.3.1. *There exists $C = C_L > 0$ such that, if S_h satisfies* (5.2.1), (5.2.2), (5.3.3), *and* (5.3.4), *ζ and $\mathcal{P}$ satisfy* (5.3.5), *and $\tilde{u}$ and $\tilde{u}_h$ are as in* (5.3.1) *and* (5.3.2) *with $\|Du_0\|_\infty \le L$, then, for all $\varepsilon, h > 0$,*

$$
\begin{aligned}
\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} & |\tilde{u}_h(x,t;\zeta,\mathcal{P}) - \tilde{u}(x,t)| \\
& \le \frac{1}{\varepsilon}\sum_{n=0}^{N-1} (\Delta t_n)^2 + C\sqrt{N}h + \max_{s,t\in[0,T]} \left\{ C\,|\zeta(s) - \zeta(t)| - \frac{|s-t|^2}{2\varepsilon} \right\}.
\end{aligned}
$$

A central role of the proof will be played by the smooth solution Φ_δ constructed in subsection 2.4.1, whose properties are outlined in Lemma 2.4.2.

Before proving Theorem 5.3.1, we state some regularity estimates for $\tilde{u}$ and $\tilde{u}_h$. First, the monotonicity of the scheme operator S_h, the comparison principle for (5.3.1), and the translation invariance of the solution operators for each immediately yield the Lipschitz bounds

$$\|D\tilde{u}\|_\infty , \ \|D\tilde{u}_h\|_\infty \leq L. \tag{5.3.6}$$

The regularity of $\tilde{u}_h$ and $\tilde{u}$ in the time variable is established by the next result.

Lemma 5.3.1. *There exists $C = C_L > 0$ such that, for all $(x, s, t) \in \mathbb{R}^d \times [0, T] \times [0, T]$ with $s < t$,*

$$|\tilde{u}(x, t) - \tilde{u}(x, s)| \leq C \operatorname{osc}(\zeta, s, t) \tag{5.3.7}$$

and, for all $m, n \in \{0, 1, 2, \dots N\}$ with $m < n$,

$$|\tilde{u}_h(x, t_n; \zeta, \mathcal{P}) - \tilde{u}_h(x, t_m; \zeta, \mathcal{P})| \leq C \left(h\sqrt{n - m} + \operatorname{osc}(\zeta, t_m, t_n) \right). \tag{5.3.8}$$

Proof. For fixed $s \geq 0$, if $\tilde{v}(x, t) := \tilde{u}(x, t + s)$, then $\tilde{v}$ is the solution of (5.3.1) with the path $\eta(t) := \zeta(t + s)$. The bound (5.3.7) then follows easily from Theorem 4.1.1.

To prove (5.3.8), observe first that, in view of Lemma 2.4.2(c), there exists $C = C_L > 0$ such that, for all $z \in \mathbb{R}^d$ and $\delta > 0$,

$$L|z| \leq \Phi_\delta(z, t_m, t_m; \zeta) + C\delta.$$

Then (5.3.6) yields, for all $x, y \in \mathbb{R}^d$,

$$\begin{aligned} \tilde{u}_h(x, t_m; \zeta, \mathcal{P}) &\leq \tilde{u}_h(y, t_m; \zeta, \mathcal{P}) + L|x - y| \\ &\leq \tilde{u}_h(y, t_m; \zeta, \mathcal{P}) + \Phi_\delta(x - y, t_m, t_m; \zeta) + C\delta. \end{aligned} \tag{5.3.9}$$

Keeping y fixed, we then apply the operator $\prod_{k=m}^{n-1} S_h(t_{k+1}, t_k; \zeta, \mathcal{P})$ to the left- and right-hand sides of (5.3.9). The inequality is preserved because of the monotonicity of this operator implied by (5.3.3) and (5.3.5). According to (5.3.2), the left-hand side becomes $\tilde{u}_h(x, t_n; \zeta, \mathcal{P})$. Iteratively using (5.3.4) to compare the right-hand side to $\Phi_\delta(x-y, t_n, t_m; \zeta, \mathcal{P})$ yields, in view of Lemma 2.4.2(b),

$$\begin{aligned}\tilde{u}_h(x, t_n; \zeta, \mathcal{P}) &\leq \tilde{u}_h(y, t_m; \zeta, \mathcal{P}) + \Phi_\delta(x-y, t_n, t_m; \zeta) + C\left(\delta + (n-m)\left\|D^2\Phi_\delta\right\|_\infty h^2\right) \\ &\leq \tilde{u}_h(y, t_m; \zeta, \mathcal{P}) + \Phi_\delta(x-y, t_n, t_m; \zeta) + C\left(\delta + \frac{(n-m)h^2}{\delta}\right),\end{aligned}$$

as long as $\operatorname{osc}(\zeta, t_m, t_n) \leq \delta$. Setting $x = y$ gives

$$\tilde{u}_h(x, t_n; \zeta, \mathcal{P}) - \tilde{u}_h(x, t_m; \zeta, \mathcal{P}) \leq C \inf\left\{\delta + \frac{(n-m)h^2}{\delta} : \delta \geq \operatorname{osc}(\zeta, t_m, t_n)\right\}.$$

If $\operatorname{osc}(\zeta, t_m, t_n) \leq h\sqrt{n-m}$, then the right-hand side is optimized by choosing $\delta = h\sqrt{n-m}$. Otherwise, setting $\delta = \operatorname{osc}(\zeta, t_m, t_n)$ gives the result, since in this case,

$$\frac{(n-m)h^2}{\delta} = \frac{(n-m)h^2}{\operatorname{osc}(\zeta, t_m, t_n)} \leq h\sqrt{n-m}.$$

The lower bound for $\tilde{u}_h(\cdot, t_n; \zeta, \mathcal{P}) - \tilde{u}_h(\cdot, t_m; \zeta, \mathcal{P})$ is proved similarly. □

We now prove Theorem 5.3.1. Using the notation of Lemma 2.4.2, we define

$$U_\delta(\zeta) := \left\{(s,t) \in [0,T]^2 : \operatorname{osc}(\zeta, s, t) < \delta\right\}.$$

Proof of Theorem 5.3.1. Throughout the proof, to simplify the presentation, we set $\tilde{u}_h(x,t) := \tilde{u}_h(x, t; \zeta, \mathcal{P})$. Fix a constant $\overline{C} = \overline{C}_L > 0$ to be determined later, and let $\alpha, \mu : [0,T] \to \mathbb{R}$ be the nondecreasing, lower-semicontinuous, piecewise constant functions defined by

$$\begin{cases} \alpha(0) = \mu(0) = 0, \quad \alpha(s) - \alpha(t_n) := [(\Delta t)_n]^2, \quad \text{and} \quad \mu(s) - \mu(t_n) = \overline{C}h^2 \\ \text{for } n = 0, 1, 2, \ldots, N-1 \text{ and } s \in (t_n, t_{n+1}]. \end{cases}$$

Choose $\varepsilon > 0$ and

$$\delta > \max\left\{2\lambda_0 h, \max_{s,t\in[0,T]}\left\{\overline{C}\,|\zeta(s)-\zeta(t)| - \frac{|s-t|^2}{2\varepsilon}\right\}\right\}, \tag{5.3.10}$$

and define the auxiliary function $\Psi : [0,T]\times[0,T] \to \mathbb{R}$ by

$$\Psi(s,t) = \sup_{x,y\in\mathbb{R}^d}\{\tilde{u}_h(x,s) - \tilde{u}(y,t) - \Phi_\delta(x-y,s,t;\zeta)\} - \frac{|s-t|^2}{2\varepsilon} - \frac{\mu(s)}{\delta} - \frac{\alpha(s)}{\varepsilon}, \tag{5.3.11}$$

where Φ_δ is the "distance function" given in (2.4.9).

Step 1: We first prove that, if $\overline{C}$ is sufficiently large, then

$$\max_{[0,T]^2}\Psi = \max\left\{\max_{s\in[0,T]}\Psi(s,0), \max_{t\in[0,T]}\Psi(0,t)\right\}. \tag{5.3.12}$$

Assume for the sake of contradiction that, for some $\sigma > 0$, $\Psi(s,t) - \sigma t$ attains its maximum in $[0,T]\times[0,T]$ at $(\hat{s},\hat{t})$ with $\hat{s} > 0$ and $\hat{t} > 0$.

The first observation is that, for some $M = M_L > 0$, the supremum in (5.3.11) may be restricted to $x, y \in \mathbb{R}^d$ satisfying $|x-y| \le M\delta$. This is because, for any $s,t\in[0,T]$ and for some $C' = C'_L > 0$,

$$\sup_{x,y\in\mathbb{R}^d}\{\tilde{u}_h(x,s) - \tilde{u}(y,t) - \Phi_\delta(x-y,s,t;\zeta)\} \ge \sup_{x\in\mathbb{R}^d}\{\tilde{u}_h(x,s) - \tilde{u}(x,t)\} - C'\delta,$$

while, if $|x-y| > M\delta$, then (5.3.6) and Lemma 2.4.2(c) give, for some $C = C_L > 0$,

$$\begin{aligned}\tilde{u}_h(x,s) - \tilde{u}(y,t) - \Phi_\delta(x-y,s,t;\zeta) &\le \sup_{x\in\mathbb{R}^d}\{\tilde{u}_h(x,s) - \tilde{u}(x,t)\} + L|x-y| - \frac{|x-y|^2}{2(C+1)\delta} + C\delta \\ &\le \sup_{x\in\mathbb{R}^d}\{\tilde{u}_h(x,s) - \tilde{u}(x,t)\} - \frac{M^2\delta}{4(C+1)} + (C + (C+1)L^2)\delta \\ &< \sup_{x\in\mathbb{R}^d}\{\tilde{u}_h(x,s) - \tilde{u}(x,t)\} - C'\delta,\end{aligned}$$

where the last inequality holds if M is sufficiently large.

As a result, if $\overline{C}$ is large enough, then $(\hat{s}, \hat{t}) \in U_{\delta/2}(W)$. To verify this, we rearrange terms in the inequality $\Psi(\hat{s}, \hat{s}) \leq \Psi(\hat{s}, \hat{t})$ and use Lemmas 2.4.2(a) and 5.3.1 to obtain, for some $C = C_L > 0$,

$$\begin{aligned}\frac{|\hat{s}-\hat{t}|^2}{2\varepsilon} &\leq \sup_{|x-y|\leq M\delta} \left\{\tilde{u}_h(y,\hat{s}) - \tilde{u}(y,\hat{t}) + \Phi_\delta(x-y,\hat{s},\hat{s};\zeta) - \Phi_\delta(x-y,\hat{s},\hat{t};\zeta)\right\} \\ &\leq C\operatorname{osc}(\zeta,\hat{s},\hat{t}).\end{aligned}$$

Consequently,

$$\begin{aligned}\overline{C}\operatorname{osc}(\zeta,\hat{s},\hat{t}) &\leq \max_{s,t\in[0,T]} \left\{\overline{C}\operatorname{osc}(\zeta,s,t) - \frac{|s-t|^2}{2\varepsilon}\right\} + \frac{|\hat{s}-\hat{t}|^2}{2\varepsilon} \\ &\leq \max_{s,t\in[0,T]} \left\{\overline{C}\,|\zeta(s)-\zeta(t)| - \frac{|s-t|^2}{2\varepsilon}\right\} + C\operatorname{osc}(\zeta,\hat{s},\hat{t}) \leq \delta + C\operatorname{osc}(\zeta,\hat{s},\hat{t}),\end{aligned}$$

so that

$$\operatorname{osc}(\zeta,\hat{s},\hat{t}) \leq \frac{\delta}{\overline{C}-C} < \frac{\delta}{2} \quad \text{if} \quad \overline{C} > C+2.$$

Now, if $\hat{n} \in \{0,1,2,\ldots,N-1\}$ is the integer satisfying $t_{\hat{n}} < \hat{s} \leq t_{\hat{n}+1}$, then, because ζ is affine on $[t_{\hat{n}}, t_{\hat{n}+1}]$, we have that

$$|\zeta(\hat{s}) - \zeta(t_{\hat{n}})| \leq \lambda_0 h < \frac{\delta}{2},$$

and so the triangle inequality yields $(t_{\hat{n}}, \hat{t}) \in U_\delta(\zeta)$. This, in turn, means that $(s, \hat{t}) \in U_\delta(\zeta)$ for all $s \in [t_{\hat{n}}, \hat{s}]$.

We next use Definition 2.3.2 to establish the inequality

$$\frac{\hat{s}-\hat{t}}{\varepsilon} \geq \sigma. \tag{5.3.13}$$

In view of Lemma 2.4.2(c), for any $x \in \mathbb{R}^d$, the function

$$y \mapsto \tilde{u}(y,t) + \Phi_\delta(x-y,\hat{s},t;\zeta)$$

attains a global minimum over $\mathbb{R}^d$. Definition 2.3.2 and Lemma 2.4.2(d) then imply that

$$t \mapsto \inf_{y\in\mathbb{R}^d} \{\tilde{u}(y,t) + \Phi_\delta(x-y,\hat{s},t;\zeta)\}$$

is nondecreasing on $I := \{t \in [0,T] : (\hat{s},t) \in U_\delta(\zeta)\}$, and therefore, so is

$$\phi(t) := \inf_{x,y\in\mathbb{R}^d} \{\tilde{u}(y,t) - \tilde{u}_h(x,\hat{s}) + \Phi_\delta(x-y,\hat{s},t;\zeta)\}.$$

Since $\phi(t) + \frac{|\hat{s}-t|^2}{2\varepsilon} + \sigma t$ attains a minimum at $\hat{t} \in I$, (5.3.13) follows.

On the other hand, we obtain a contradiction by using (5.3.3) and (5.3.4) to show that

$$\frac{\hat{s}-\hat{t}}{\varepsilon} \leq 0. \tag{5.3.14}$$

The first step is to prove that, for each $y \in \mathbb{R}^d$, the function

$$a(s) := \sup_{x\in\mathbb{R}^d} \left\{\tilde{u}_h(x,s) - \Phi_\delta(x-y,s,\hat{t};\zeta)\right\} - \frac{\mu(s)}{\delta}$$

satisfies

$$\max_{[t_{\hat{n}},t_{\hat{n}+1}]} a = a(t_{\hat{n}}).$$

Indeed, if this were not the case, then, for some $s^* \in [t_{\hat{n}}, t_{\hat{n}+1}]$ and sufficiently small $\beta > 0$,

$$a(t_{\hat{n}}) \leq a(s^*) - \beta(s^* - t_{\hat{n}}).$$

Lemma 2.4.2(c) implies that the supremum in the definition of $a(s^*)$ is attained for some

$x^* \in \mathbb{R}^d$, and so it follows that, for all $x \in \mathbb{R}^d$,

$$\begin{aligned}\tilde{u}_h(x, t_{\hat{n}}) \leq \tilde{u}_h(x^*, s^*) + \Phi_\delta(x - y, t_{\hat{n}}, \hat{t}; \zeta) - \Phi_\delta(x^* - y, s^*, \hat{t}; \zeta) \\ - \frac{\mu(s^*) - \mu(t_{\hat{n}})}{\delta} - \beta(s^* - t_{\hat{n}}).\end{aligned} \tag{5.3.15}$$

In view of (5.3.3) and the fact that $\operatorname{osc}(\zeta, t_{\hat{n}}, s^*) \leq \lambda_0 h$, the operator $S_h(s^*, t_{\hat{n}}; \zeta)$ is monotone. Applying it to both sides of the inequality (5.3.15), setting $x = x^*$, rearranging terms, and using (5.3.4) and Lemma 2.4.2(b) and (d) yield

$$\frac{\overline{C} h^2}{\delta} + \beta(s^* - t_{\hat{n}}) = \frac{\mu(s^*) - \mu(t_{\hat{n}})}{\delta} + \beta(s^* - t_{\hat{n}}) \leq C \left\| D^2 \Phi \right\|_\infty h^2 \leq \frac{C h^2}{\delta}.$$

This results in a contradiction as long as $\overline{C}$ is chosen to be at least the constant C on the right-hand side.

As a consequence,

$$\psi(s) := \sup_{x, y \in \mathbb{R}^d} \left\{ \tilde{u}_h(x, s) - \tilde{u}(y, \hat{t}) - \Phi_\delta(x - y, s, \hat{t}; \zeta) \right\} - \frac{\mu(s)}{\delta}$$

attains its maximum in $[t_{\hat{n}}, t_{\hat{n}+1}]$ at $t_{\hat{n}}$, and therefore, because $\psi(s) - \frac{|s - \hat{t}|^2}{2\varepsilon} - \frac{\alpha(s)}{\varepsilon}$ attains a maximum at $\hat{s}$,

$$\psi(t_{\hat{n}}) - \frac{|t_{\hat{n}} - \hat{t}|^2}{2\varepsilon} - \frac{\alpha(t_{\hat{n}})}{\varepsilon} \leq \psi(\hat{s}) - \frac{|\hat{s} - \hat{t}|^2}{2\varepsilon} - \frac{\alpha(\hat{s})}{\varepsilon} \leq \psi(t_{\hat{n}}) - \frac{|\hat{s} - \hat{t}|^2}{2\varepsilon} - \frac{\alpha(\hat{s})}{\varepsilon}$$

which, after rearranging terms, yields (5.3.14). Together with (5.3.13), this establishes (5.3.12).

Step 2: The next claim is that, for some $C = C_L > 0$,

$$\max_{\{0\} \times [0,T] \cup [0,T] \times \{0\}} \Psi \leq C \left(\delta + \sqrt{N} h \right) + \max_{s, t \in [0,T]} \left\{ C \left| \zeta(s) - \zeta(t) \right| - \frac{|s - t|^2}{2\varepsilon} \right\}.$$

Assume that Ψ attains its maximum at $(\hat{s}, \hat{t})$, with either $\hat{s} = 0$ or $\hat{t} = 0$.

If $\hat{s} = \hat{t} = 0$, then Lemmas 2.4.2(c) and 5.3.1 yield $C = C_L > 0$ such that

$$\Psi(0,0) = \sup_{x,y\in\mathbb{R}^d} \{u_0(x) - u_0(y) - \Phi_\delta(x-y,0,0;\zeta)\}$$
$$\leq \sup_{x,y\in\mathbb{R}^d} \left\{L|x-y| - \frac{1}{2(C+1)\delta}|x-y|^2\right\} + C\delta \leq \left(C + \frac{(C+1)L^2}{2}\right)\delta.$$

Assume now that $\hat{s} = 0$. Then, in view of Lemmas 2.4.2(c) and 5.3.1,

$$\begin{aligned}\Psi(0,\hat{t}) &= \sup_{|x-y|\leq M\delta} \left\{u_0(x) - \tilde{u}(y,\hat{t}) - \Phi_\delta(x-y,0,\hat{t};\zeta)\right\} - \frac{\hat{t}^2}{2\varepsilon}\\
&\leq C\delta + \sup_{|x-y|\leq M\delta} \left\{u_0(y) - \tilde{u}(y,\hat{t}) - \Phi_\delta(x-y,0,\hat{t};\zeta)\right\} - \frac{\hat{t}^2}{2\varepsilon}\\
&\leq C\delta + \max_{t\in[0,T]} \left(C\operatorname{osc}(\zeta,0,t) - \frac{t^2}{2\varepsilon}\right) = C\delta + \max_{s,t\in[0,T]}\left(C|\zeta(s)-\zeta(t)| - \frac{|s-t|^2}{2\varepsilon}\right).\end{aligned}$$

Finally, if $\hat{t} = 0$, then Lemma 5.3.1 gives

$$\begin{aligned}\Psi(\hat{s},0) &\leq \sup_{|x-y|\leq M\delta} \{\tilde{u}_h(x,\hat{s}) - u_0(y) - \Phi_\delta(x-y,\hat{s},0;\zeta)\} - \frac{\hat{s}^2}{2\varepsilon}\\
&\leq C\delta + \sup_{x\in\mathbb{R}^d} \{\tilde{u}_h(x,\hat{s}) - u_0(x)\} - \frac{\hat{s}^2}{2\varepsilon}\\
&\leq C\left(\delta + \sqrt{N}h\right) + \max_{s,t\in[0,T]}\left\{C\left|\zeta(s)-\zeta(t)\right| - \frac{|s-t|^2}{2\varepsilon}\right\}.\end{aligned}$$

Step 3. Combining the previous two steps and rearranging terms yields, for all $(x,t) \in \mathbb{R}^d \times [0,T]$,

$$\begin{aligned}\tilde{u}_h(x,t) - \tilde{u}(x,t) &\leq \frac{1}{\varepsilon}\sum_{n=0}^{N-1}(\Delta t_n)^2 + C\left(\delta + \frac{Nh^2}{\delta} + \sqrt{N}h\right)\\
&\quad + \max_{s,t\in[0,T]}\left\{C\left|\zeta(s)-\zeta(t)\right| - \frac{|s-t|^2}{2\varepsilon}\right\}.\end{aligned}$$

The inequality is optimized by setting

$$\delta := \max \left\{ C\sqrt{N}h, \max_{s,t\in[0,T]} \left(C\left|\zeta(s) - \zeta(t)\right| - \frac{|s-t|^2}{2\varepsilon} \right) \right\}.$$

Note that, if the constant $C = C_L > 0$ in the definition of δ is sufficiently large, then (5.3.10) is satisfied. This finishes the proof of the upper bound for $\tilde{u}_h - \tilde{u}$, and the lower bound is proved similarly. □

5.3.2 Convergence rates for a fixed continuous path

We now apply the pathwise error estimate from Theorem 5.3.1 to obtain a rate of convergence for schemes approximating solutions of the Hamilton-Jacobi equation

$$du = \sum_{i=1}^{m} H^i(Du) \cdot dW^i \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d. \tag{5.3.16}$$

It will always be assumed, as above, that $\operatorname{Lip}(u_0) \le L$, (5.2.12), and (5.2.13), and that the scheme operator S_h satisfies (5.2.1), (5.2.2), (5.3.3), and (5.3.4).

We first examine the setting in which W is a fixed, deterministic path, and then some extensions are presented in the case where W is a Brownian motion. Following Section 5.2, we define $u_h := \tilde{u}_h(\cdot; W_h, \mathcal{P}_h)$, with $\tilde{u}_h$ as in (5.3.2), for an appropriate family of approximating paths $\{W_h\}_{h>0}$ and partitions $\{\mathcal{P}_h\}_{h>0}$. Let $\tilde{u}$ be the viscosity solution of

$$\tilde{u}_t = \sum_{i=1}^{m} H^i(D\tilde{u})\dot{W}_h \quad \text{in } \mathbb{R}^d \times (0,T] \quad \text{and} \quad \tilde{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d. \tag{5.3.17}$$

The error $u_h - u$ is then controlled by using Theorems 5.3.1 and 4.1.1 to estimate respectively the differences $u_h - \tilde{u}$ and $\tilde{u} - u$.

Fix $W \in C([0,T];\mathbb{R}^m)$, and let $\omega : [0,\infty) \to [0,\infty)$ be its modulus of continuity. Define

ρ_h implicitly by

$$\lambda := \frac{\omega((\rho_h)^{1/2})}{(\rho_h)^{1/2}} < \lambda_0 \tag{5.3.18}$$

and set $\mathcal{P}_h := \{n\rho_h \wedge T\}_{n=0}^N$, where N is the smallest integer for which $N\rho_h \wedge T = T$.

Recall from subsection 5.2.1 that taking W_h to be the piecewise linear interpolation of W over the partition $\mathcal{P}_h$ may not, in general, yield a convergent scheme. Instead, we set

$$M_h := \left\lfloor (\rho_h)^{-1/2} \right\rfloor$$

and define W_h as follows: for $k \in \mathbb{N}_0$ and $t \in [kM_h\rho_h, (k+1)M_h\rho_h)$,

$$W_h(t) := W(kM_h\rho_h) + \left(\frac{W((k+1)M_h\rho_h) - W(kM_h\rho_h)}{M_h\rho_h} \right)(t - kM_h\rho_h)\,. \tag{5.3.19}$$

Observe that the approximating path W_h satisfies (5.2.17) with $\eta_h = (\rho_h)^{1/2}$.

Now set $u_h := \tilde{u}_h(\cdot; W_h, \mathcal{P}_h)$, with $\tilde{u}_h$ as in (5.3.2), and let $\tilde{u}$ be the solution of (5.3.17).

Theorem 5.3.2. *There exists $C = C_{L,\lambda} > 0$ such that*

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - u(x,t)| \le C(1+T)\omega((\rho_h)^{1/2}). \tag{5.3.20}$$

As an example, assume that $W \in C^{0,\alpha}([0,T], \mathbb{R}^m)$ and set

$$\lambda := [W]_{\alpha,T} \frac{(\rho_h)^{(1+\alpha)/2}}{h}.$$

Then, as long as $\lambda < \lambda_0$, the scheme converges with a rate of order $(\rho_h)^{\alpha/2} \approx h^{\alpha/(1+\alpha)}$.

Proof of Theorem 5.3.2. First, notice that, in view of (5.3.18), W_h satisfies (5.3.5). In particular, for some $C = C_L > 0$,

$$\max_{s,t\in[0,T]} \left(C|W_h(s) - W_h(t)| - \frac{|s-t|^2}{2\varepsilon} \right) \leq C\lambda_0 h + \max_{n\in\mathbb{N}_0} \left(Cn\lambda_0 h - \frac{n^2\rho_h^2}{2\varepsilon} \right)$$
$$\leq C\lambda_0 h + \frac{(C\lambda_0 h)^2 \varepsilon}{2(\rho_h)^2}.$$

Theorem 5.3.1 then gives, for any $\varepsilon > 0$,

$$\max_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - \tilde{u}(x,t)| \leq \frac{N(\rho_h)^2}{\varepsilon} + C\sqrt{N}h + \frac{(C\lambda_0 h)^2\varepsilon}{2(\rho_h)^2}$$
$$\leq \frac{T\rho_h}{\varepsilon} + C\sqrt{T}\frac{h}{\sqrt{\rho_h}} + \frac{(C\lambda_0 h)^2\varepsilon}{2(\rho_h)^2}.$$

Upon choosing $\varepsilon = \sqrt{T}\frac{(\rho_h)^{3/2}}{h}$, this becomes

$$\max_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - \tilde{u}(x,t)| \leq C\sqrt{T}\frac{h}{\sqrt{\rho_h}} = C\sqrt{T}\omega((\rho_h)^{1/2}). \tag{5.3.21}$$

Notice that the error term takes the form $\sqrt{\frac{h^2}{\rho_h}}$, which is consistent with the discussion in subsection 5.2.1.

Combining this with Theorem 4.1.1, we obtain

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - u(x,t)| \leq C\left(\sqrt{T}\omega((\rho_h)^{1/2}) + \omega(M_h\rho_h)\right),$$

and the result is proved in view of the choice of M_h. □

5.3.3 Brownian paths

For the rest of the chapter, we investigate schemes for which W is a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which, for definiteness, we may take to be $C([0,T], \mathbb{R})$ with $\mathbb{P}$ the Wiener measure. The expectation and variance with respect to $\mathbb{P}$ are denoted by respectively $\mathbb{E}$ and Var. To simplify the presentation, it is assumed that $m = 1$, so that W is one-dimensional, although all three schemes below can be adapted to the case when $m > 1$.

Regular partitions

Theorem 5.3.2 may be applied in this situation by using the fact that oscillations of Brownian paths are controlled by the Lévy modulus of continuity. More precisely,

$$\mathbb{P}\left(\limsup_{\delta \to 0} \sup_{\delta \le t \le T-\delta} \frac{|W(t) - W(t+\delta)|}{\sqrt{2\delta \left|\log \delta\right|}} = 1 \right) = 1. \tag{5.3.22}$$

Theorem 5.3.3. *Let ρ_h be defined implicitly by*

$$\lambda := \frac{(\rho_h)^{3/4} \left|\log \rho_h\right|^{1/2}}{h} < \lambda_0, \tag{5.3.23}$$

and let u_h, $\mathcal{P}_h$, and W_h be as in the previous subsection. Then there exists a deterministic constant $C = C_{L,\lambda} > 0$ such that, if u is the solution of (5.3.16), *then*

$$\mathbb{P}\left(\limsup_{h \to 0} \sup_{(x,t) \in \mathbb{R}^d \times [0,T]} \frac{|u_h(x,t) - u(x,t)|}{(\rho_h)^{1/4} \left|\log \rho_h\right|^{1/2}} \le C(1+T) \right) = 1.$$

Proof. Define $M_h := \lfloor (\rho_h)^{-1/2} \rfloor$ and $K_h := \lfloor T/(M_h \rho_h) \rfloor$. The definitions of W_h and λ give

$$\begin{aligned}
\max_{n=0,1,2\ldots,N-1} \frac{|W_h(n\rho_h) - W_h((n+1)\rho_h)|}{h} &= \max_{k=0,1,2\ldots,K_h} \frac{|W(kM_h\rho_h) - W((k+1)M_h\rho_h|}{M_h h} \\
&= \lambda \max_{k=0,1,2\ldots,K_h} \frac{|W(kM_h\rho_h) - W((k+1)M_h\rho_h)|}{M_h(\rho_h)^{3/4} \left|\log \rho_h\right|^{1/2}} \\
&\le \lambda \frac{\max_{|s-t| \le (\rho_h)^{1/2}} |W(s) - W(t)|}{(\rho_h)^{1/4}(1 - (\rho_h)^{1/2}) \left|\log \rho_h\right|^{1/2}}.
\end{aligned}$$

Therefore, in view of (5.3.22), for any $\delta > 0$,

$$\mathbb{P}\left(\max_{n=0,1,2,\ldots,N-1} \frac{|W_h(n\rho_h) - W_h((n+1)\rho_h)|}{h} \le \lambda \frac{1+\delta}{1 - (\rho_h)^{1/2}} \text{ for small } h \right) = 1.$$

Taking $\delta \in (0, \lambda_0/\lambda - 1)$ above, this implies that

$$\mathbb{P}\left(\limsup_{h\to 0} \max_{n=0,1,2,\ldots,N-1} \frac{|W_h(n\rho_h) - W_h((n+1)\rho_h)|}{h} < \lambda_0\right) = 1,$$

so that, for some $h_0 > 0$,

$$\mathbb{P}\left(|W_h(n\rho_h) - W_h((n+1)\rho_h)| \leq \lambda_0 h \quad \text{for all } 0 < h < h_0 \text{ and } n = 0, 1, 2, \ldots, N_h - 1\right) = 1.$$

Shrinking h_0, if necessary, it may be concluded from (5.3.20) and (5.3.22) that

$$\mathbb{P}\left(\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - u(x,t)| \leq CT(\rho_h)^{1/4} |\log \rho_h|^{1/2} \text{ for all } 0 < h < h_0\right) = 1.$$

□

Observe that (5.3.23) implies that $\lim_{h\to 0} \frac{\log \rho_h}{\log h} = \frac{4}{3}$, so that the convergence rate in Theorem 5.3.3 can be rewritten as

$$\limsup_{h\to 0} \sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \frac{|u_h(x,t) - u(x,t)|}{h^{1/3} |\log h|^{1/3}} \leq C(1+T). \tag{5.3.24}$$

Random partitions

For the next scheme, the partitions $\mathcal{P}_h$ are defined using a sequence of stopping times adapted to the filtration $\mathcal{F}_t$ of the Brownian motion W. This raises the possibility of improving the rate of convergence in Theorem 5.3.3 by using techniques such as the law of large numbers to eliminate the logarithmic correction in (5.3.24). Unfortunately, this does not seem to be possible because of the presence of the term

$$\max_{s,t\in[0,T]} \left\{C|W_h(s) - W_h(t)| - \frac{|s-t|^2}{2\varepsilon}\right\} \tag{5.3.25}$$

in the estimate from Theorem 5.3.1. On the one hand,

$$\max_{s,t\in[0,T]} \mathbb{E}\left\{C|W_h(s)-W_h(t)| - \frac{|s-t|^2}{2\varepsilon}\right\} \le \max_{\rho\ge 0}\left\{C\rho^{1/2} - \frac{\rho^2}{2\varepsilon}\right\} \le \frac{3C^{4/3}}{2^{7/3}}\varepsilon^{1/3},$$

which suggests that taking $\varepsilon = h$ optimizes the error estimate. However, as $\varepsilon, h \to 0$, the term (5.3.25) is on the order of $\varepsilon^{1/3}\,|\log\varepsilon|^{2/3}$, and, as a consequence, $\varepsilon = \varepsilon_h$ must be slightly smaller than h in order to recover the error estimate (5.3.24).

For $h > 0$, define $\eta_h := h^{1/3}\,|\log h|^{-2/3}$, set $T_0 = T_0(h) := 0$, and, for $k \in \mathbb{N}_0$,

$$\begin{cases} T_{k+1} = T_{k+1}(h) := \inf\{t > T_k(h) : \operatorname{osc}(W, T_k(h), t) > \eta_h\} \quad \text{and} \\ \tau_{k+1} = \tau_{k+1}(h) := T_{k+1}(h) - T_k(h). \end{cases}$$

Observe that $\{T_k\}_{k=0}^{\infty}$ is an increasing sequence of stopping times, and, for each fixed k, $h \to T_k(h)$ decreases as $h \to 0$. Therefore, by the strong Markov property for Brownian motion, for each fixed h, $\{\tau_k(h)\}_{k=1}^{\infty}$ is a collection of independent, identically distributed random variables. As a result, for any integer $\ell > 0$, there exists a constant $c_\ell > 0$ such that, for all k,

$$\mathbb{E}[\tau_k(h)^\ell] = c_\ell(\eta_h)^{2\ell}.$$

Indeed, it is well known that the first exit time of a Brownian motion from a bounded interval has finite moments of any order. The exact formula follows from the scaling properties of Brownian motion, so that

$$c_\ell := \mathbb{E}\left[\inf\left\{t > 0 : \operatorname{osc}\left(W, 0, t^{1/\ell}\right) > 1\right\}\right].$$

Let W_h be the piecewise interpolation of W over the partition

$$\{0 = T_0(h) < T_1(h) < T_2(h) < \cdots\}.$$

That is,

$$\begin{cases} W_h(t) := W(T_k(h)) + \dfrac{W(T_{k+1}(h)) - W(T_k(h))}{\tau_{k+1}(h)}(t - T_k(h)) \\ \text{whenever} \quad T_k(h) \le t < T_{k+1}(h). \end{cases}$$

Define

$$\begin{cases} M_h := \left\lceil \dfrac{\eta_h}{\lambda_0 h} \right\rceil = \left\lceil (\lambda_0 h^{2/3} |\log h|^{2/3})^{-1} \right\rceil, \\ t_0 = t_0(h) := 0, \quad t_n = t_n(h) := T_k(h) + (n - kM_h)\dfrac{\tau_{k+1}(h)}{M_h}, \\ \Delta t_n = \Delta t_n(h) := t_{n+1}(h) - t_n(h) = \dfrac{\tau_{k+1}(h)}{M_h} \\ \text{whenever } k = 0, 1, 2, \dots \text{ and } kM_h \le n < (k+1)M_h, \\ K_h := \sup \left\{ k \in \mathbb{N}_0 : T_k(h) \le T \right\}, \quad \text{and} \\ N_h := \sup \left\{ n \in \mathbb{N}_0 : t_n(h) \le T \right\}, \end{cases}$$

and note that $h \mapsto K_h$ increases as $h \to 0$.

We have now specified the path W_h, which is piecewise affine over the partition

$$\mathcal{P}_h := \left\{ 0 = t_0(h) < t_1(h) < t_2(h) < \cdots < t_{N_h}(h) \le T \right\},$$

in such a way that (5.3.5) holds for $\zeta = W_h$. Indeed, if $n = 0, 1, 2, \dots, N - 1$ and k is such that $kM_h \le t_n < t_{n+1} \le (k+1)M_h$, then

$$|W_h(t_{n+1}) - W_h(t_n)| = \frac{|W(T_{k+1}) - W(T_k)|}{M_h} \le \lambda_0 h.$$

Finally, set $u_h := \tilde{u}_h(\cdot; W_h, \mathcal{P}_h)$ and let u be the stochastic viscosity solution of (5.3.16).

Theorem 5.3.4. *There exists a deterministic constant $C = C_L > 0$ such that*

$$\mathbb{P}\left(\limsup_{h \to 0} \max_{(x,t) \in \mathbb{R}^d \times [0,T]} \frac{|u_h(x,t) - u(x,t)|}{h^{1/3} \left|\log h\right|^{1/3}} \le C(1 + T) \right) = 1.$$

We proceed with a series of lemmas that indicate how to control the various terms appearing in the estimate from Theorem 5.3.1.

Lemma 5.3.2.

$$\mathbb{P}\left(\limsup_{h\to 0} K_h\eta_h^2 \leq \frac{T}{c_1}\right) = 1.$$

Proof. Fix α and β such that $1 < \beta^{2/3} < \alpha$, and define $h_m := \beta^{-m}$. Note that

$$\lim_{m\to\infty} \frac{\eta_{h_{m+1}}}{\eta_{h_m}} = \frac{1}{\beta^{1/3}}.$$

The monotonicity of K_h and η_h implies that

$$\mathbb{P}\left(\sup_{h_{m+1}\leq h<h_m} K_h\eta_h^2 > \frac{\alpha T}{c_1}\right) \leq \mathbb{P}\left(K_{h_{m+1}} > \frac{\alpha T}{c_1\eta_{h_m}^2}\right). \tag{5.3.26}$$

Set

$$k_m := \left\lceil \frac{\alpha T}{c_1\eta_{h_m}^2}\right\rceil,$$

so that

$$k_m c_1\eta_{h_{m+1}}^2 \geq \alpha T\left(\frac{\eta_{h_{m+1}}}{\eta_{h_m}}\right)^2 \xrightarrow{m\to\infty} \alpha\beta^{-2/3}T > T,$$

and therefore, for any fixed $\gamma > 0$ and all sufficiently large m, $k_m c_1\eta_{h_{m+1}}^2 \geq (1+\gamma)T$.

Define $\sigma^2 := c_2 - c_1^2$, so that $\mathrm{Var}(\tau_k(h)) = \sigma^2\eta_h^4$ for all k and h. Continuing (5.3.26) and applying Markov's inequality yields, for some fixed positive constant $C > 0$ and for all sufficiently large m,

$$\begin{aligned}\mathbb{P}\left(K_{h_{m+1}} > \frac{\alpha T}{c_1\eta_{h_m}^2}\right) &= \mathbb{P}\left(\sum_{k=1}^{k_m} \tau_k(h_{m+1}) \leq T\right)\\ &\leq \mathbb{P}\left(\sum_{k=1}^{k_m}\left(\tau_k(h_{m+1}) - c_1\eta_{h_{m+1}}^2\right) \leq -\gamma T\right)\\ &\leq \frac{k_m\sigma^2\eta_{h_{m+1}}^4}{\gamma^2T^2} \leq C\beta^{-2m/3}.\end{aligned}$$

The Borel-Cantelli lemma applied to the events

$$E_m := \left\{ \sup_{h_{m+1} \le h < h_m} K_h \eta_h^2 > \frac{\alpha T}{c_1} \right\}$$

gives

$$\mathbb{P}\left(\limsup_{h \to 0} K_h \eta_h^2 > \frac{\alpha T}{c_1} \right) = \mathbb{P}\left(\limsup_{m \to \infty} E_m \right) = 0,$$

and we may conclude upon sending $\alpha \to 1^+$. □

Lemma 5.3.3.

$$\mathbb{P}\left(\limsup_{h \to 0} \frac{1}{h\eta_h} \sum_{n=0}^{N-1} (\Delta t_n)^2 \le \frac{T \lambda_0 c_2}{c_1} \right) = 1.$$

Proof. Fix α and β satisfying $1 < \beta^{7/3} < \alpha$ and set $h_m := \beta^{-m}$. If, for some m, $h_{m+1} \le h < h_m$, then

$$\sum_{n=0}^{N_h - 1} (\Delta t_n(h))^2 \le \sum_{k=1}^{K_h + 1} M_h \left(\frac{\tau_k(h)}{M_h} \right)^2 \le \lambda_0 \frac{h_m}{\eta_{h_{m+1}}} \sum_{k=1}^{K_{h_{m+1}}+1} \tau_k(h_m)^2.$$

Fix $m_0 \in \mathbb{N}$ and define the event

$$E_{m_0} := \left\{ K_{h_{m+1}} + 1 \le \hat{K}_m := \left\lceil \frac{\alpha T}{c_1 \eta_{h_{m+1}}^2} \right\rceil \quad \text{for all} \quad m \ge m_0 \right\}.$$

In view of Lemma 5.3.2, $\lim_{m_0 \to \infty} \mathbb{P}\left(E_{m_0}\right) = 1$.

Now, for any $m \geq m_0$,

$$\begin{aligned}
&\mathbb{P}\left(\left\{\sup_{h_{m+1}\leq h<h_m} \frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 > \frac{\alpha^2 T\lambda_0 c_2}{c_1}\right\}\cap E_{m_0}\right)\\
&\quad\leq \mathbb{P}\left(\sum_{k=1}^{\hat{K}_m}\tau_k(h_m)^2 > \frac{\alpha^2 T c_2 h_{m+1}(\eta_{h_{m+1}})^2}{c_1 h_m}\right)\\
&\quad= \mathbb{P}\left(\sum_{k=1}^{\hat{K}_m}\left(\tau_k(h_m)^2 - c_2\eta_{h_m}^4\right) > \frac{\alpha^2 T c_2 h_{m+1}(\eta_{h_{m+1}})^2}{c_1 h_m} - \hat{K}_m c_2\eta_{h_m}^4\right)\\
&\quad\leq \mathbb{P}\left(\sum_{k=1}^{\hat{K}_m}\left(\tau_k(h_m)^2 - c_2\eta_{h_m}^4\right) > \frac{\alpha T c_2\eta_{h_m}^2}{c_1}\left(\frac{\alpha h_{m+1}(\eta_{h_{m+1}})^2}{h_m(\eta_{h_m})^2} - \frac{\eta_{h_m}^2}{\eta_{h_{m+1}}^2}\right) - c_2\eta_{h_m}^4\right).
\end{aligned}$$

Since

$$\lim_{m\to\infty}\left(\frac{\alpha h_{m+1}\eta_{h_{m+1}}^2}{h_m\eta_{h_m}^2} - \frac{\eta_{h_m}^2}{\eta_{h_{m+1}}^2}\right) = \frac{\alpha}{\beta^{5/3}} - \beta^{2/3} > 0,$$

it follows that, for some fixed $\gamma > 0$, all sufficiently large m_0, and all $m > m_0$,

$$\begin{aligned}
&\mathbb{P}\left(\left\{\sup_{h_{m+1}\leq h<h_m} \frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 > \frac{\alpha^2 T\lambda_0 c_2}{c_1}\right\}\cap E_{m_0}\right)\\
&\quad\leq \mathbb{P}\left(\sum_{k=1}^{\hat{K}_m}\left(\tau_k(h_m)^2 - c_2\eta_{h_m}^4\right) > \gamma\eta_{h_m}^2\right).
\end{aligned}$$

Set $\sigma^2 := c_4 - c_2^2 > 0$. Then Markov's inequality gives, for some constant $C > 0$ independent of m,

$$\begin{aligned}
&\mathbb{P}\left(\left\{\sup_{h_{m+1}\leq h<h_m} \frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 > \frac{\alpha^2 T\lambda_0 c_2}{c_1}\right\}\cap E_{m_0}\right)\\
&\quad\leq \frac{\hat{K}_m\sigma^2\eta_{h_m}^4}{\gamma^2} \leq C\eta_{h_m}^2 \leq C\beta^{-2m/3}.
\end{aligned}$$

An application of the Borel-Cantelli lemma for the events

$$\left\{ \sup_{h_{m+1} \le h < h_m} \frac{1}{h\eta_h} \sum_{n=0}^{N_h - 1} (\Delta t_n(h))^2 > \frac{\alpha^2 T \lambda_0 c_2}{c_1} \right\} \cap E_{m_0}$$

yields

$$\mathbb{P}\left(\left\{ \limsup_{h \to 0} \frac{1}{h\eta_h} \sum_{n=0}^{N_h - 1} (\Delta t_n(h))^2 > \frac{\alpha^2 T \lambda_0 c_2}{c_1} \right\} \cap E_{m_0} \right) = 0.$$

Sending $m_0 \to \infty$ and then $\alpha \to 1^+$ finishes the proof. □

Lemma 5.3.4. *For any deterministic constant $C > 0$,*

$$\mathbb{P}\left(\limsup_{\varepsilon \to 0} \frac{\max_{s,t \in [0,T]} \left\{ C \left| W(s) - W(t) \right| - \frac{|s-t|^2}{2\varepsilon} \right\}}{\varepsilon^{1/3} \left| \log \varepsilon \right|^{2/3}} \le \frac{4C^{4/3}}{3^{2/3}} \right) = 1.$$

Proof. Let $1 < \beta < \alpha$. If, for some $\delta > 0$,

$$\text{(5.3.27)} \qquad \operatorname{osc}(W, k\delta, (k+1)\delta) \le \sqrt{2\beta}\delta^{1/2} \left| \log \delta \right|^{1/2} \quad \text{for all} \quad k = 0, 1, 2, \dots, \left\lceil \frac{T}{\delta} \right\rceil,$$

then

$$\begin{aligned}
&\max_{s,t \in [0,T]} \left\{ C|W(s) - W(t)| - \frac{|s-t|^2}{2\varepsilon} \right\} \\
&\qquad \le \sqrt{2\beta} C \delta^{1/2} \left| \log \delta \right|^{1/2} + \max_{n \in \mathbb{N}_0} \left\{ \sqrt{2\beta} C \delta^{1/2} \left| \log \delta \right|^{1/2} n - \frac{n^2 \delta^2}{2\varepsilon} \right\} \\
&\qquad \le \sqrt{2\beta} C \delta^{1/2} \left| \log \delta \right|^{1/2} + \beta C^2 \frac{\left| \log \delta \right|}{\delta} \varepsilon.
\end{aligned}$$

Taking $\delta := C^{2/3} 3^{-1/3} \beta^{1/3} \varepsilon^{2/3} \left| \log \varepsilon \right|^{1/3}$ yields, for some deterministic function $c(\varepsilon)$ that converges to 0 as $\varepsilon \to 0$,

$$\max_{s,t \in [0,T]} \left\{ C|W(s) - W(t)| - \frac{|s-t|^2}{2\varepsilon} \right\} \le \frac{4C^{4/3}}{3^{2/3}} \beta^{2/3} \varepsilon^{1/3} |\log \varepsilon|^{2/3} (1 + c(\varepsilon)),$$

and, therefore, if ε is sufficiently small,

$$\max_{s,t\in[0,T]}\left\{C|W(s)-W(t)|-\frac{|s-t|^2}{2\varepsilon}\right\}\leq\frac{4C^{4/3}}{3^{2/3}}\alpha^{2/3}\varepsilon^{1/3}|\log\varepsilon|^{2/3}.$$

Define

$$\varepsilon_m:=\alpha^{-m}\quad\text{and}\quad\delta_m:=\frac{C^{2/3}\beta^{1/3}\varepsilon_m^{2/3}\,|\log\varepsilon_m|^{1/3}}{3^{1/3}},$$

and note that

$$\lim_{m\to\infty}\frac{\varepsilon_m^{1/3}|\log\varepsilon_m|^{2/3}}{\varepsilon_{m+1}^{1/3}|\log\varepsilon_{m+1}|^{2/3}}=\alpha^{1/3}.$$

It follows that, for sufficiently large m,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{\varepsilon_{m+1}\leq\varepsilon<\varepsilon_m}\frac{\max_{s,t\in[0,T]}\left\{C\,|W(s)-W(t)|-\frac{|s-t|^2}{2\varepsilon}\right\}}{\varepsilon^{1/3}\,|\log\varepsilon|^{2/3}}>\frac{4C^{4/3}\alpha}{3^{2/3}}\right)\\
&\leq\mathbb{P}\left(\max_{s,t\in[0,T]}\left\{C|W(s)-W(t)|-\frac{|s-t|^2}{2\varepsilon_m}\right\}>\frac{4C^{4/3}\alpha\varepsilon_{m+1}^{1/3}|\log\varepsilon_{m+1}|^{2/3}}{3^{2/3}}\right)\\
&\leq\mathbb{P}\left(\max_{s,t\in[0,T]}\left\{C|W(s)-W(t)|-\frac{|s-t|^2}{2\varepsilon_m}\right\}>\frac{4C^{4/3}\alpha^{2/3}\varepsilon_m^{1/3}|\log\varepsilon_m|^{2/3}}{3^{2/3}}\right)\\
&\leq\mathbb{P}\left(\operatorname{osc}(W,k\delta_m,(k+1)\delta_m)>\sqrt{2\beta}\delta_m^{1/2}\,|\log\delta_m|^{1/2}\ \text{ for some } k=0,1,2,\ldots,\left\lceil\frac{T}{\delta_m}\right\rceil\right)\\
&\leq\left\lceil\frac{T}{\delta_m}\right\rceil\mathbb{P}\left(\max_{[0,1]}W-\min_{[0,1]}W>\sqrt{2\beta}\,|\log\delta_m|^{1/2}\right)\\
&\leq 2\left\lceil\frac{T}{\delta_m}\right\rceil\mathbb{P}\left(\max_{[0,1]}W>\sqrt{2\beta}\,|\log\delta_m|^{1/2}\right)\\
&\leq CT\delta_m^{\beta-1}\leq CT\alpha^{-\gamma m}\quad\text{for}\quad\gamma=\frac{2}{3}(\beta-1)>0.
\end{aligned}$$

The symmetry and scaling properties of Brownian motion, as well as the reflection principle, are all used above. In particular, since the processes

$$\left\{t\mapsto\max_{s\in[0,t]}W(s)-W(t)\right\}\quad\text{and}\quad\{t\mapsto|W(t)|\}$$

are identically distributed, so are the random variables

$$\max_{[0,1]} W - \min_{[0,1]} W \quad \text{and} \quad \max_{[0,1]} |W| = \max\left\{ \max_{[0,1]} W, -\min_{[0,1]} W \right\}.$$

The Borel-Cantelli lemma implies that

$$\mathbb{P}\left(\limsup_{\varepsilon\to 0} \frac{\max_{s,t\in[0,T]} \left\{ C\,|W(s)-W(t)| - \frac{|s-t|^2}{2\varepsilon} \right\}}{\varepsilon^{1/3}\,|\log\varepsilon|^{2/3}} > \frac{4C^{4/3}\alpha}{3^{2/3}} \right) = 0,$$

and sending $\alpha \to 1^+$ gives the result. □

Proof of Theorem 5.3.4. Let $\tilde{u}$ be the solution of (5.3.17). Then Theorem 4.1.1 gives, for some $C = C_L > 0$,

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |\tilde{u}(x,t) - u(x,t)| \le C\,|W(t) - W_h(t)|$$
$$\le C \max_{k=0,1,2,\ldots,K_h} \operatorname{osc}\left(W, T_k(h), T_{k+1}(h)\right) \le C\frac{h^{1/3}}{|\log h|^{2/3}}.$$

Next, define $\varepsilon_h := \frac{h}{|\log h|}$ and recall the pathwise estimate from Theorem 5.3.1:

$$\max_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - \tilde{u}(x,t)|$$
$$\le \frac{1}{\varepsilon_h}\sum_{n=0}^{N_h-1} (\Delta t_n(h))^2 + C\sqrt{N_h}h + \max_{s,t\in[0,T]}\left\{ C\,|W_h(s) - W_h(t)| - \frac{|s-t|^2}{2\varepsilon_h} \right\}.$$

From the definitions of N_h, M_h, and K_h, and from Lemma 5.3.2, it follows that, for some $C = C_L > 0$, with probability one, for all sufficiently small h,

$$C\sqrt{N_h}h \le C\sqrt{(K_h+1)M_h}h \le CTh^{1/3}\,|\log h|^{1/3}.$$

Meanwhile, Lemma 5.3.3 yields $C = C_L > 0$ such that, with probability one, for all sufficiently small h,

$$\frac{1}{\varepsilon_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 = \frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n)^2 \cdot h^{1/3}\left|\log h\right|^{1/3} \leq CTh^{1/3}\left|\log h\right|^{1/3}.$$

In view of the definition of W_h,

$$\max_{0\leq t\leq T}|W_h(t)-W(t)| \leq \max_{k=0,1,2,\ldots,K_h}\operatorname{osc}(W,T_k(h),T_{k+1}(h)) \leq \eta_h,$$

so that, with probability one, for all h,

$$\frac{\max_{s,t\in[0,T]}\left\{C\left|W_h(s)-W_h(t)\right| - \frac{|s-t|^2}{2\varepsilon_h}\right\}}{h^{1/3}\left|\log h\right|^{1/3}}$$
$$\leq \frac{\max_{s,t\in[0,T]}\left\{C\left|W(s)-W(t)\right| - \frac{|s-t|^2}{2\varepsilon_h}\right\}}{h^{1/3}\left|\log h\right|^{1/3}} + \frac{C}{\left|\log h\right|},$$

while Lemma 5.3.4 implies that, with probability one and for all sufficiently small h,

$$\max_{s,t\in[0,T]}\left\{C\left|W(s)-W(t)\right| - \frac{|s-t|^2}{2\varepsilon_h}\right\} \leq C\varepsilon_h^{1/3}|\log\varepsilon_h|^{2/3} \leq Ch^{1/3}\left|\log h\right|^{1/3}.$$

Combining all terms in the estimate finishes the proof. □

Scaled random walks and convergence in law

Here, we use independent Rademacher random variables to build an object that converges to the solution of (5.3.16) in distribution. This construction has the advantage that it is simple to implement numerically.

Fix a probability space $(\mathcal{A},\mathcal{G},\mathbf{P})$, not necessarily related to $(\Omega,\mathcal{F},\mathbb{P})$, and let $\{\xi_n\}_{n=1}^{\infty}$:

$\mathcal{A} \to \{-1, 1\}$ be independent and identically distributed with

$$\mathbf{P}(\xi_n = 1) = \mathbf{P}(\xi_n = -1) = \frac{1}{2}.$$

Define ρ_h by

$$\lambda := \frac{(\rho_h)^{3/4}}{h} \leq \lambda_0,$$

and, as before, set $M_h = \lfloor (\rho_h)^{-1/2} \rfloor$, $\mathcal{P}_h := \{t_n\}_{n=0}^N = \{n\rho_h \wedge T\}_{n=0}^N$, $W_h(0) = 0$ and, for $k \in \mathbb{N}_0$ and $t \in [kM_h\rho_h, (k+1)M_h\rho_h)$,

$$W_h(t) := W_h(kM_h\rho_h) + \frac{\xi_k}{\sqrt{M_h\rho_h}}(t - kM_h\rho_h).$$

As has already been discussed, W_h converges to the Wiener process W in distribution.

Define $u_h := \tilde{u}_h(\cdot; W_h, \mathcal{P}_h) \in BUC(\mathbb{R}^d \times [0,T])$ and let $\tilde{u} \in BUC(\mathbb{R}^d \times [0,T])$ be the solution of (5.3.17).

Theorem 5.3.5. *As $h \to 0$, u_h converges to u in distribution.*

Proof. Observe first that

$$|W_h(t_{n+1}) - W_h(t_n)| = \sqrt{\frac{\rho_h}{M_h}} \leq (\rho_h)^{3/4} \leq \lambda_0 h,$$

so that W_h satisfies (5.3.5). Then (5.3.20) becomes, for some $C = C_{L,\lambda} > 0$,

$$\max_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - \tilde{u}(x,t)| \leq C(1+T)(\rho_h)^{1/4} = C(1+T)h^{1/3}. \tag{5.3.28}$$

Theorem 4.1.1 implies that the map

$$S : X = C([0,T], \mathbb{R}) \ni \zeta \mapsto v \in BUC(\mathbb{R}^d \times [0,T]) =: Y,$$

where v is the solution of (5.3.1), is uniformly continuous. Let $\tilde{\nu}_h$ and ν be the push-forwards by S of respectively μ_h and μ, that is, for any measurable $\psi : Y \to \mathbb{R}$,

$$\int_Y \psi \, d\nu = \int_X \psi \circ S \, d\mu,$$

with the analogous relation holding for ν_h and ν. As a consequence of the Mapping Theorem (see [14]), $\tilde{\nu}_h$ converges weakly to ν. On the other hand, if ν_h is the measure on $BUC(\mathbb{R}^d \times [0, T])$ induced by u_h, then (5.3.28) and Slutzky's theorem imply that, as $h \to 0$, ν_h converges weakly to ν. $\square$

CHAPTER 6

HOMOGENIZATION AND SCALING LIMITS

6.1 Introduction

In this chapter, we study the asymptotic properties, for small $\varepsilon > 0$, of equations of the form

$$(6.1.1) \qquad \begin{cases} u_t^\varepsilon + \sum_{i=1}^m H^i(Du^\varepsilon, x/\varepsilon)\dot{\zeta}^{i,\varepsilon} = 0 & \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \\ u^\varepsilon(\cdot, 0) = u_0 & \text{in } \mathbb{R}^d. \end{cases}$$

Here, each Hamiltonian H^i is assumed to have some averaging properties in the spatial variable $y = x/\varepsilon$. This could include a variety of types of dependence, for instance periodic, almost periodic, or stationary and ergodic. The paths $\zeta^\varepsilon = (\zeta^{1,\varepsilon}, \zeta^{2,\varepsilon}, \cdots, \zeta^{m,\varepsilon})$, which converge locally uniformly to some limiting path $\zeta \in C([0,\infty), \mathbb{R}^m)$, will generally be assumed to be piecewise-C^1, although we present some results where they are only continuous.

One motivation for considering such problems is to study general equations of the form

$$(6.1.2) \qquad \begin{cases} u_t^\varepsilon + \frac{1}{\varepsilon^\gamma} H\left(Du^\varepsilon, \frac{x}{\varepsilon}, \frac{t}{\varepsilon^{2\gamma}}\right) = 0 & \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \\ u^\varepsilon(\cdot, 0) = u_0 & \text{in } \mathbb{R}^d. \end{cases}$$

In addition to the averaging dependence on space, the Hamiltonian $H : \mathbb{R}^d \times \mathbb{R}^d \times [0,\infty) \to \mathbb{R}$ is assumed to have zero expectation, so that, on average, u^ε is close to its initial value u_0. The dependence on time, meanwhile, is assumed to be "mixing" with a certain rate, so that, with the scaling of the central limit theorem, $\frac{1}{\varepsilon^\gamma} H(\cdot, \cdot, t/\varepsilon^{2\gamma})$ will resemble white noise in time as $\varepsilon \to 0$.

With $\gamma = 1$, equation (6.1.2) arises naturally as a scaled version of the equation

$$\begin{cases} u_t + H(Du, x, t) = 0 & \text{in } \mathbb{R}^d \times (0, \infty), \\ u(\cdot, 0) = \varepsilon^{-1} u_0(\varepsilon \cdot) & \text{in } \mathbb{R}^d. \end{cases} \tag{6.1.3}$$

More precisely, u and u^ε are related by $u^\varepsilon(x,t) = \varepsilon u(x/\varepsilon, t/\varepsilon^2)$. Studying the limit as $\varepsilon \to 0$ of u^ε then amounts to understanding the average long range, long time behavior of solutions of (6.1.3) with large, slowly-varying initial data.

Scaling limits for equations with mixing dependence were studied already in the early days of the theory of stochastic differential equations. Khasminskii [35], Papanicolaou and Varadhan [66], and Papanicolaou and Kohler [65] considered ordinary differential equations with smooth coefficients that satisfy similar mixing properties, and these authors proved that, in distribution, the solutions converge to diffusion Markov processes given by solutions of certain Itô stochastic differential equations (see also [45] for more details and references). Results of this form were obtained for linear parabolic partial differential equations with mixing dependence in time by, among many others, Bouc and Pardoux [15], Kushner and Huang [48], and Watanabe [82]. There are some works studying equations that homogenize in space and become white in time; see, for instance, Pardoux and Piatnitski [70].

The results presented in this thesis do not correspond to any of these, due to the nonlinear dependence on the gradient variable. In addition, our methods are quite different than those in the above cited works. The strategy there is to first establish tightness of the probability measures, and then uniquely identify the limiting measure as the law of the unique solution of a certain equation. In our setting, this second step presents a challenge, due to the nonlinearity of the equations, and we will need to prove convergence directly without using the tightness of the laws.

It is of interest to examine (6.1.2) for different values of γ. As it will turn out, the nature of the limiting behavior does not change for different values of γ, so, from a practical point of view, the quantities ε and $\delta := \varepsilon^\gamma$ can be viewed as small, independent parameters. In addition, for technical reasons, some results can only be proved under a mildness assumption on the approximate white noise dependence, which translates to a smallness condition on γ.

The Hamiltonians we consider for (6.1.2) will generally have the form

$$H(p,y,t) := \sum_{i=1}^{m} H^i(p,y)\xi^i(t), \tag{6.1.4}$$

where the random fields $\xi^i : [0,\infty) \to \mathbb{R}$ are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $0 \le s \le t \le \infty$, consider the sigma algebras $\mathcal{F}^i_{s,t} \subset \mathcal{F}$ generated by $\{\xi^i(r)\}_{r\in[s,t]}$, and define the mixing rate

$$\rho(t) := \max_{i=1,2,\ldots,m} \sup_{s\ge 0} \sup_{A\in\mathcal{F}^i_{s+t,\infty}} \sup_{B\in\mathcal{F}^i_{0,s}} \left|\mathbb{P}(A \mid B) - \mathbb{P}(A)\right|. \tag{6.1.5}$$

The quantitative mixing assumptions for the ξ^i are that

$$\begin{cases} t \mapsto \xi^i(t) \text{ is stationary,} \\ \rho(t) \xrightarrow{t\to\infty} 0, \quad \displaystyle\int_0^\infty \rho(t)^{1/2}\,dt < \infty, \\ \mathbb{E}[\xi^i(0)] = 0, \text{ and } \mathbb{E}[\xi^i(0)^2] = 1. \end{cases} \tag{6.1.6}$$

One consequence, in view of the ergodic theorem, the stationarity, and the centering assumption, is that, if

$$\zeta^i(t) := \int_0^t \xi^i(s)\,ds,$$

then $\lim_{\delta\to 0} \delta\zeta^i(t/\delta) = 0$. Understanding the long-time fluctuations of this quantity around 0 is done by considering the scaling of the central limit theorem. Setting $\zeta^{i,\delta}(t) = \delta\zeta^i(t/\delta^2)$, it is well-known [65] that, in distribution, as $\delta \to 0$, $\zeta^{i,\delta}$ converges locally uniformly to a standard Brownian motion. Indeed, with $\delta = \varepsilon^\gamma$, the equation (6.1.2) is then a specific form of (6.1.1).

The assumption of stationarity in (6.1.6) means that

$$(\xi(s_1), \xi(s_2), \dots, \xi(s_M)) \quad \text{and} \quad (\xi(s_1+t), \xi(s_2+t), \dots, \xi(s_M+t))$$

have the same joint distribution for any choice of $s_1, s_2, \dots, s_M \in [0,\infty)$ and $t \geq -\min_j s_j$. However, the conclusions reached above are unchanged if ξ is stationary only with respect to integer shifts. Later, we focus on ξ of the form

$$\xi(t) := \sum_{k=1}^{\infty} X_k \mathbf{1}_{[k-1,k)}(t) \tag{6.1.7}$$

where the $\{X_k\}_{k=1}^{\infty}$ are mutually independent, identically distributed random variables with

$$\mathbb{E}[X_k] = 0 \quad \text{and} \quad \mathbb{E}[(X_k)^2] = 1.$$

For instance, each X_k may be a standard Rademacher random variable; that is, $\mathbb{P}(X_k = 1) = \mathbb{P}(X_k = -1) = \frac{1}{2}$.

Observe that ξ defined in (6.1.7) immediately satisfies (6.1.6), since $\rho(t) = 0$ as soon as $t > 1$. In this case, the path $\zeta(t) = \int_0^t \xi(s)\, ds$ is a linearly-interpolated random walk, and the convergence of ζ^ε to a Brownian motion is just a restatement of Donsker's invariance principle, which has already been used in Chapter 5.

6.2 The difficulties and general strategy

Here we discuss some of the difficulties in the study of (6.1.1) and the strategies we use to overcome them. For simplicity of presentation, we discuss here only the case where the H^i's are periodic in space.

We first make the formal assumption that the noise is "mild" enough to allow for averaging behavior in space, and therefore, u^ε is closely approximated by a solution $\overline{u}^\varepsilon$ of an equation of the form $\overline{u}^\varepsilon_t + \overline{H}^\varepsilon(D\overline{u}^\varepsilon, t) = 0$. More precisely, we follow a standard strategy from the homogenization literature and assume that, for some auxiliary function $v : \mathbb{T}^d \times [0,\infty) \to \mathbb{R}$, u^ε has the formal expansion

$$u^\varepsilon(x,t) \approx \overline{u}^\varepsilon(x,t) + \varepsilon v(x/\varepsilon, t).$$

An asymptotic analysis yields that, for fixed $p \in \mathbb{R}^d$ (here, $p = D\overline{u}^\varepsilon(x,t)$ and $y = \frac{x}{\varepsilon}$), v solves

$$\sum_{i=1}^m H^i(D_y v + p, y)\xi^i = \overline{H}(p,\xi) \quad \text{in } \mathbb{R}^d. \tag{6.2.1}$$

Here, the fixed parameter $\xi \in \mathbb{R}^m$ stands in place of the mild white noise $\varepsilon^{-\gamma}\xi(t/\varepsilon^{2\gamma})$. The equation (6.2.1) is known as the "cell problem," and, in the theory of periodic homogenization of Hamilton-Jacobi equations, under the right conditions, there is a unique choice of constant for the right-hand side for which (6.2.1) has periodic solutions, which are called "correctors."

Taking this fact for granted for now, this means that, formally, u^ε will be closely approximated by the solution $\overline{u}^\varepsilon$ of

$$\overline{u}^\varepsilon_t + \frac{1}{\varepsilon^\gamma}\overline{H}\left(D\overline{u}^\varepsilon, \xi\left(\frac{t}{\varepsilon^{2\gamma}}\right)\right) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad \overline{u}^\varepsilon(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d. \tag{6.2.2}$$

Note that, in deriving (6.2.2), we have used the fact that $\xi \mapsto \overline{H}(\cdot,\xi)$ is positively homogenous, which can be seen from multiplying (6.2.1) by a positive constant and using the uniqueness of the right-hand side.

If

$$\mathbb{E}\left[\overline{H}(p,\xi(t))\right] = 0 \quad \text{for all } p \in \mathbb{R}^d, \tag{6.2.3}$$

then the solution of (6.2.2) with $u_0(x) = p_0 \cdot x$, which is given by

$$\overline{u}^\varepsilon(x,t) = p_0 \cdot x - \frac{1}{\varepsilon^{2\gamma}} \int_0^t \overline{H}\left(p_0, \xi\left(\frac{s}{\varepsilon^{2\gamma}}\right)\right) ds,$$

converges in distribution, as $\varepsilon \to 0$, to $p_0 \cdot x + \sigma(p_0)B(t)$, where B is a standard Brownian motion and

$$\sigma(p_0) := \left(\mathbb{E}\left[\overline{H}(p_0, \xi(0))^2\right]\right)^{1/2}.$$

However, due to the nonlinearity of the map $\xi \mapsto \overline{H}(\cdot, \xi)$, it is not clear, for an arbitrary $u_0 \in UC(\mathbb{R}^d)$, how to study the general equation (6.2.2) with the pathwise viscosity solution theory described in Chapter 2. This question, as well as the validity of (6.2.3), are taken up in the section on multiple paths below. The answers are subtle, and depend strongly on the nature of the mixing field ξ. In particular, there is not a universal limit in general.

When $m = 1$, the characterization of $\overline{H}(p, \xi)$ reduces to the study of the two Hamiltonians

$$\overline{H}(p) := \overline{H}(p, 1) \quad \text{and} \quad \overline{(-H)}(p) := \overline{H}(p, -1).$$

The equation (6.2.2) then takes the form

$$\text{(6.2.4)} \quad \begin{cases} \overline{u}^\varepsilon_t + \dfrac{1}{\varepsilon^\gamma}\overline{H}^1(D\overline{u}^\varepsilon)\xi\left(\dfrac{t}{\varepsilon^{2\gamma}}\right) + \dfrac{1}{\varepsilon^\gamma}\overline{H}^2(D\overline{u}^\varepsilon)\left|\xi\left(\dfrac{t}{\varepsilon^{2\gamma}}\right)\right| = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \\ \overline{u}^\varepsilon(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d, \end{cases}$$

where

$$\overline{H}^1(p) := \frac{\overline{H}(p) - \overline{(-H)}(p)}{2} \quad \text{and} \quad \overline{H}^2(p) := \frac{\overline{H}(p) + \overline{(-H)}(p)}{2}.$$

The coefficient $\overline{H}^2$ is equal to zero if and only if

$$\text{(6.2.5)} \qquad \overline{(-H)} = -\overline{H}.$$

Notice that, when $m = 1$, (6.2.3) is equivalent to (6.2.5).

Because (6.2.1) is interpreted in the viscosity solution sense, it is not possible to multiply the equation by -1, and so (6.2.5) is not only non-obvious, but false in general, as we show later by example. In that case, for some $p_0 \in \mathbb{R}^d$, $\overline{(-H)}(p_0) \neq -\overline{H}(p_0)$. The solution $\overline{u}^\varepsilon$ with $u_0(x) := p_0 \cdot x$ is then equal to

$$\begin{aligned}\overline{u}^\varepsilon(x,t) = p_0 \cdot x - \varepsilon^\gamma \frac{\overline{H}(p_0) - \overline{(-H)}(p_0)}{2} \zeta\left(\frac{t}{\varepsilon^{2\gamma}}\right) \\ - \varepsilon^\gamma \frac{\overline{H}(p_0) + \overline{(-H)}(p_0)}{2} \int_0^{t/\varepsilon^{2\gamma}} \left|\xi\left(\frac{s}{\varepsilon^{2\gamma}}\right)\right| \, ds,\end{aligned}$$

and so

$$\varepsilon^\gamma \overline{u}^\varepsilon(x,t) \xrightarrow{\varepsilon \to 0} -\frac{\overline{H}(p_0) + \overline{(-H)}(p_0)}{2} \mathbb{E}\left|\xi(0)\right| t \quad \text{in distribution.}$$

On the other hand, if (6.2.5) holds, then (6.2.2) becomes

$$\text{(6.2.6)} \qquad \overline{u}^\varepsilon_t + \frac{1}{\varepsilon^\gamma}\overline{H}(D\overline{u}^\varepsilon)\xi\left(\frac{t}{\varepsilon^{2\gamma}}\right) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad \overline{u}^\varepsilon(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d,$$

and the determination of whether or not $\overline{u}^\varepsilon$ has a limit depends on the properties of the effective Hamiltonian $\overline{H}$, and, in particular, whether or not it satisfies (2.4.4).

6.3 The single-noise case

The equation of interest here is

$$\text{(6.3.1)} \qquad u^\varepsilon_t + \frac{1}{\varepsilon^\gamma} H\left(Du^\varepsilon, \frac{x}{\varepsilon}\right)\xi\left(\frac{t}{\varepsilon^{2\gamma}}\right) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u^\varepsilon(\cdot,0) = u_0.$$

As suggested by the previous section, the fact that there is only one source of noise simplifies the structure of the problem. We also make use of the stability results proved in Section 4.2. The Hamiltonian is assumed to be satisfy exactly the assumptions from that section, which we recount here for convenience:

$$\begin{cases} H \in C(\mathbb{R}^d \times \mathbb{R}^d),\ p \mapsto H(p,x) \text{ is convex for all } x \in \mathbb{R}^d, \text{ and} \\ \text{there exist convex, increasing functions } \underline{\nu}, \overline{\nu} : [0,\infty) \to \mathbb{R} \text{ such that} \\ \underline{\nu}(|p|) \le H(p,x) \le \overline{\nu}(|p|) \quad \text{for all } (p,x) \in \mathbb{R}^d \times \mathbb{R}^d. \end{cases} \tag{6.3.2}$$

In order for homogenization to occur in space, it is necessary for the dependence of H on the spatial environment to possess averaging behavior. Since the result below does not rely on the specific nature of this dependence, we give the most general assumption, and later demonstrate that it holds in many cases of interest.

Let $S^\varepsilon_\pm(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ and $S_\pm(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ be the solution operators for respectively

$$U^\varepsilon_{\pm,t} \pm H(DU^\varepsilon_\pm, x/\varepsilon) = 0 \text{ in } \mathbb{R}^d \times (0,\infty), \quad U^\varepsilon_\pm(\cdot,0) = \phi \quad \text{on } \mathbb{R}^d,$$

and, for some deterministic $\overline{H} \in C(\mathbb{R}^d)$ satisfying (6.3.2),

$$U_{\pm,t} \pm \overline{H}(DU_\pm) = 0 \text{ in } \mathbb{R}^d \times (0,\infty), \quad U_\pm(\cdot,0) = \phi \quad \text{on } \mathbb{R}^d,$$

that is, $U^\varepsilon_\pm(x,t) = S^\varepsilon_\pm(t)\phi(x)$ and $U_\pm(x,t) = S_\pm(t)\phi(x)$. The assumption is that, for all $\phi \in BUC(\mathbb{R}^d)$,

$$\lim_{\varepsilon\to 0} S^\varepsilon_\pm(t)\phi(x) = S_\pm(t)\phi(x) \quad \text{locally uniformly.} \tag{6.3.3}$$

Note that the consistency condition (6.2.5) is built directly into assumption (6.3.3). As we show later, this is justified by the convexity of $p \mapsto H(p,\cdot)$, which implies (6.2.5) in either the periodic or stationary-ergodic settings.

The first result is as follows:

Theorem 6.3.1. *Assume* (6.3.2) *and* (6.3.3)*. Then there exists a Brownian motion* B :

$[0,\infty) \to \mathbb{R}$ *such that, as* $\varepsilon \to 0$, $(u^\varepsilon, \zeta^\varepsilon)$ *converges in distribution to* $(\overline{u}, B)$ *in* $BUC(\mathbb{R}^d \times [0,\infty)) \times C([0,\infty))$, *where* $\overline{u}$ *is the pathwise viscosity solution of*

$$d\overline{u} + \overline{H}(D\overline{u}) \circ dB = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad \overline{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d. \tag{6.3.4}$$

As before, the topology in $BUC(\mathbb{R}^d \times [0,\infty))$ and $C([0,\infty))$ is that of local uniform convergence.

In the case where $\gamma = 1$, (6.3.1) arises from the scaling $u^\varepsilon(x,t) = \varepsilon u(x/\varepsilon, t/\varepsilon^2)$ of the equation

$$u_t + H(Du, y)\xi(t) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u(\cdot,0) = \varepsilon^{-1}u_0(\varepsilon\cdot) \quad \text{in } \mathbb{R}^d. \tag{6.3.5}$$

Indeed, in this case, the scaling limit result has the following interpretation: the function $u_\varepsilon(x,t) := \varepsilon u(x/\varepsilon, t/\varepsilon)$, which is a solution of

$$u_{\varepsilon,t} + H(Du_\varepsilon, x/\varepsilon)\xi(t/\varepsilon) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u_\varepsilon(\cdot,0) = u_0, \tag{6.3.6}$$

converges, with probability one and locally uniformly as $\varepsilon \to 0$, to u_0. Theorem 6.3.1 then implies that $u_\varepsilon(x, t/\varepsilon)$ converges in distribution to $\overline{u}(x,t)$; in other words, we are able to describe the long time fluctuations of u_ε about its average value u_0.

Observe that the singular random field

$$\xi(t) := dB(t),$$

is strongly mixing, and indeed,

$$\zeta(t) := \int_0^t \xi(s)\, ds = \int_0^t dB(s) = B(t),$$

so that $t \mapsto \zeta^\delta(t) = \delta B(t/\delta^2)$ is equal in distribution to B by the scaling properties of

Brownian motion. It therefore is an interesting question whether the solutions u^ε of

$$du^\varepsilon + H(Du^\varepsilon, x/\varepsilon) \circ dB = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u^\varepsilon(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d \tag{6.3.7}$$

share the same limit.

Theorem 6.3.2. *In addition to the hypotheses of Theorem 6.3.1, assume that the comparison principle holds for* (6.3.7)*. Then, with probability one, as $\varepsilon \to 0$, the solution u^ε of* (6.3.7) *converges locally uniformly to the solution of* (6.3.4)*.*

An example of a Hamiltonian satisfying (6.3.2) is given, for some $a \in C(S^{d-1} \times \mathbb{R}^d)$, by

$$H(p,x) := a\left(\frac{p}{|p|}, x\right)|p| \tag{6.3.8}$$

as long as it is convex. This will hold if, for instance, a is independent of the direction variable $p/|p| \in S^{d-1}$ and is strictly positive. If, in addition, $a \in C_b^2(S^{d-1}, \mathbb{R}^d)$, then, by the results in Section 3.3, (6.3.7) satisfies the comparison principle, and so Theorem 6.3.2 holds as well.

Recall that the Hamiltonian (6.3.8) arises in the study of first-order front propagation. Indeed, with H as in (6.3.8), the level sets of u^ε evolve according to the normal velocity $-\varepsilon^{-\gamma} a(n, x/\varepsilon)\xi(t/\varepsilon^{2\gamma})$ (or $-a(n, x/\varepsilon)dB$ in the case of (6.3.7)), where n is the outward normal vector to the level set at the point x. It is not hard to prove that the effective Hamiltonian $\overline{H}$ is 1-homogenous in the gradient variable as well, and therefore, it has the form $\overline{H}(p) := \overline{a}(p/|p|)|p|$ for some $\overline{a} \in C(S^{d-1})$. The homogenization results may then be rephrased as saying that the level-set flows converge in distribution to a front evolving according to the normal velocity $\overline{a}(n)dB$ in the space of compact subsets of $\mathbb{R}^d \times [0,\infty)$, whose topology is given by the Hausdorff distance.

Theorems 6.3.1 and 6.3.2 are a consequence of a more general convergence result that holds in the single noise case. We consider collections $\{H^\varepsilon\}_{\varepsilon > 0} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ of Hamiltonians and $\{\zeta^\varepsilon\}_{\varepsilon>0} : [0,\infty) \to \mathbb{R}$ of continuous, piecewise smooth paths, and study the convergence of solutions, as $\varepsilon \to 0$, of the equations

$$(6.3.9) \qquad u_t^\varepsilon + H^\varepsilon(Du^\varepsilon, x)\dot{\zeta}^\varepsilon(t) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u^\varepsilon(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d.$$

Theorem 6.3.3. *Assume that the Hamiltonians* $\{H^\varepsilon\}_{\varepsilon>0}$ *satisfy* (6.3.2) *uniformly in* ε, *and that there exists a Hamiltonian* $H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *satisfying* (6.3.2) *such that* (6.3.3) *holds with* $S_\pm^\varepsilon$ *and* $S_\pm$ *the solution operators for respectively* $\pm H^\varepsilon$ *and* $\pm H$.

(a) If $\zeta \in C([0,\infty),\mathbb{R})$ *is such that, as* $\varepsilon \to 0$, ζ^ε *converges locally uniformly to* ζ, *then, as* $\varepsilon \to 0$, u^ε *converges locally uniformly to a pathwise viscosity solution* u *of*

$$(6.3.10) \qquad du + H(Du,x)\cdot d\zeta = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d.$$

(b) If $\{\zeta^\varepsilon\}, \zeta$ *above are defined over a probability space* $(\Omega,\mathcal{F},\mathbb{P})$ *and, as* $\varepsilon \to 0$, ζ^ε *converges locally uniformly to* ζ *in distribution, then, as* $\varepsilon \to 0$, u^ε *converges locally uniformly in distribution to a solution* u *of* (6.3.10).

(c) The limiting function u *in parts* (a) *or* (b) *is independent of the choice of approximating family* $\{\zeta^\varepsilon\}_{\varepsilon>0}$.

(d) If the comparison principle holds for (6.3.10) *for any continuous* ζ, *then the paths* $\{\zeta^\varepsilon\}_{\varepsilon>0}$ *in part* (a) *or* (b) *are permitted to be merely continuous.*

Even if the comparison principle holds for (6.3.9) when ζ^ε is continuous, the same may not necessarily be true of the equation for H, and so solutions of (6.3.10) are not guaranteed to be unique. However, it is still the case that the limiting solution u given by Theorem 6.3.3 is unique and independent of the choice of approximating paths.

The general convergence results in Theorem 6.3.3 can be applied to a variety of other settings. For instance, it recovers the already well-known stability of viscosity solutions under local uniform limits of the data. That is, if $\{\zeta^\varepsilon\}_{\varepsilon>0}$ and $\{H^\varepsilon\}_{\varepsilon>0}$ are as above and $H^\varepsilon \xrightarrow{\varepsilon\to 0} H$ locally uniformly, then the solution u^ε converges locally uniformly as $\varepsilon \to 0$ to u. Other homogenization results can also be studied, for instance, those of the form

$$u_t^\varepsilon + H\left(Du^\varepsilon, x, \frac{x}{\varepsilon}\right)\zeta^\varepsilon(t) \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u^\varepsilon(\cdot,0) = u_0$$

with $\{\zeta^\varepsilon\}_{\varepsilon>0}$ any collection of paths converging locally uniformly and almost surely (or in distribution) to a Brownian motion or other stochastic process, and with the dependence of H on the fast variable being, for instance, periodic, quasi-periodic, or stationary-ergodic.

6.3.1 Assumptions that guarantee homogenization

Given a Hamiltonian satisfying (6.3.2), we now describe in more detail some extra assumptions for H that imply (6.3.3).

Periodicity

If

(6.3.11) $$y \mapsto H(p,y) \text{ is } \mathbb{Z}^d\text{-periodic,}$$

then there exists a Hamiltonian $\overline{H} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfying the assumptions of (6.3.2) such that (6.3.3) holds. In fact, the convergence in (6.3.3) is uniform globally, not just locally, in space, and so the statements of the theorems above can be modified to obtain global-in-space convergence results.

This qualitative homogenization has been known since the work of Lions, Papanicolaou, and Varadhan [52]. In fact, instead of (6.3.2), it is enough to have

(6.3.12) $$H \in C(\mathbb{R}^d \times \mathbb{R}^d) \quad \text{and} \quad \lim_{|p|\to+\infty} \inf_{y\in\mathbb{T}^d} H(p,y) = +\infty,$$

except this does not immediately imply (6.2.5). Indeed, this fails in general if H is not convex in p (see the example below).

Because they will be used later, we record some of the facts in the periodic homogenization setting here:

Theorem 6.3.4. *Assume that H satisfies* (6.3.11) *and* (6.3.12). *Then, for every $p \in \mathbb{R}^d$, there exist unique constants $\overline{H}(p)$ and $\overline{(-H)}(p)$ such that the equations*

$$H(p + D_y v_+, y) = \overline{H}(p) \quad \textit{and} \quad -H(p + D_y v_-, y) = \overline{(-H)}(p)$$

admit periodic viscosity solutions v_+ and v_-.

The functions $p \mapsto \overline{H}(p)$ and $p \mapsto \overline{(-H)}(p)$ are continuous. Moreover, if H is convex in p, then $\overline{(-H)} = -\overline{H}$ and $\overline{H}$ is convex.

It was proved by Capuzzo-Dolcetta and Ishii [20] that, under additional regularity, the homogenization can be quantified. We record the result here for future reference.

Theorem 6.3.5. *Assume, in addition to* (6.3.11) *and* (6.3.12), *that H is locally Lipschitz. Let u^ε and $\overline{u}$ be the solutions of respectively*

$$\begin{cases} u^\varepsilon_t + H(Du^\varepsilon, x/\varepsilon) = 0 \quad \textit{and} \\ \overline{u}_t + \overline{H}(D\overline{u}) = 0 \quad \textit{in } \mathbb{R}^d \times (0,\infty), \quad \textit{and} \\ u^\varepsilon(\cdot,0) = \overline{u}(\cdot,0) = u_0 \quad \textit{in } \mathbb{R}^d. \end{cases}$$

Then, for $L > 0$, there exists $C = C_L > 0$ such that, if $\mathrm{Lip}(u_0) \le L$, then

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u^\varepsilon(x,t) - \overline{u}(x,t)| \le C(1+T)\varepsilon^{1/3}.$$

In the full generality of the above theorem, it is still not known whether the exponent 1/3 is optimal. It can be improved in some cases: for instance, if $u_0(x) = p \cdot x$ for some fixed

$p \in \mathbb{R}^d$, then the convergence rate is $O(\varepsilon)$. Recently, using techniques from Aubrey-Mather theory, Mitake, Tran, and Yu [61] obtained the optimal convergence rate $O(\varepsilon)$ under further assumptions, in particular, if $d = 1$ and H is convex, or if $d = 2$, and H is convex and positively homogenous of some degree $q \geq 1$ in the gradient variable. The result for $d = 1$ was extended by Tu [81] to equations with Hamiltonians of the form $H(Du^\varepsilon, x, x/\varepsilon)$, periodic in the x/ε variable and convex in the gradient variable.

Stationarity and ergodicity

Another well-developed research area is the stochastic homogenization of Hamilton-Jacobi equations, in which the Hamiltonian is random and its probability law is statistically homogenous and ergodic in space. More precisely, we consider Hamiltonians $H = H(p, x, \omega)$ defined on a probability space $(\mathbf{\Omega}, \mathbf{F})$, and, for $z \in \mathbb{R}^d$, define the group of translation operators $T_z : \mathbf{\Omega} \to \mathbf{\Omega}$ by $T_z H(\cdot, y) := H(\cdot, y + z)$. We assume that there exists a probability measure $\mathbf{P}$ on $(\mathbf{\Omega}, \mathbf{F})$ such that $\{T_z\}_{z \in \mathbb{R}^d}$ is measure-preserving and ergodic, that is,

$$\begin{cases} \mathbf{P} = \mathbf{P} \circ T_z \text{ for all } z \in \mathbb{R}^d, \text{ and} \\ \text{if } E \in \mathbf{F} \text{ and } T_z E = E \text{ for all } z \in \mathbb{R}^d, \text{ then } \mathbf{P}[E] = 1 \text{ or } \mathbf{P}[E] = 0. \end{cases} \tag{6.3.13}$$

If H satisfies (6.3.12) and is convex in the gradient variable, then

$$\begin{cases} \text{there exists an event } \mathbf{\Omega}_0 \in \mathcal{F} \text{ with } \mathbf{P}(\mathbf{\Omega}_0) = 1 \text{ such that} \\ \text{for all } u_0 \in UC(\mathbb{R}^d),\ T > 0, \text{ and } \omega \in \mathbf{\Omega}_0, \\ \lim\limits_{\varepsilon \to 0} \sup\limits_{(x,t) \in B_T \times [0,T]} \left| S_\pm^\varepsilon(t) u_0(x) - S_\pm(t) u_0(x) \right| = 0. \end{cases} \tag{6.3.14}$$

This result was proved independently by Souganidis [79] and Rezakhanlou and Tarver [72]. The degenerate elliptic "viscous" case was later treated by Lions and Souganidis in [57]. Armstrong and Souganidis considered the same problem in [4] from the viewpoint of the so-called metric problem.

These results all used convexity, but recently, there has been some progress in the nonconvex case, for instance by Armstrong, Tran, and Yu [5, 6], and by Armstrong and Cardaliaguet [2]. Ziliotto [85], and later Feldman and Souganidis [29], provided examples of nonconvex Hamiltonians and random media for which homogenization fails, indicating that, in general, some form of convexity is required.

There has also been extensive work in obtaining algebraic error estimates of the type obtained in the periodic case in Theorem 6.3.5. Armstrong, Cardaliaguet, and Souganidis obtained such a result in [3] by quantifying the methods in [4]. In addition to (6.3.2), it is necessary to have a stronger mixing assumption than ergodicity, namely that the environment has a finite range of dependence, as well as a condition of the law of the Hamiltonian near its minimum.

The condition (6.2.5)

We prove here a sufficient criterion for the consistency condition (6.2.5) to hold, in both the periodic and stationary-ergodic settings. We use the fact that $\overline{H}$ and $\overline{(-H)}$ are given by

$$\overline{H}(p) = \inf_{v\in\mathcal{G}} \sup_{y\in\mathbb{R}^d} H(p + D_y v, y) \quad \text{and} \quad \overline{(-H)}(p) = \sup_{v\in\mathcal{G}} \inf_{y\in\mathbb{R}^d} \left(-H(p + D_y v, y)\right), \tag{6.3.15}$$

where the supremum and infimum over $y \in \mathbb{R}^d$ are interpreted in the viscosity sense, and $\mathcal{G}$ is the set of $v \in C(\mathbb{R}^d)$ such that

$$\begin{cases} v \text{ is periodic} & \text{in the periodic setting, or} \\ \lim_{|y|\to\infty} \frac{v(y)}{|y|} = 0 \text{ almost surely} & \text{in the random setting.} \end{cases}$$

In the periodic setting, this follows from the comparison principle and the definition of correctors, and is proved in [52] and [28]. In the random setting, (6.3.15) holds for convex H, which was first proved in [57].

Lemma 6.3.1. *In addition to* (6.3.12)*, assume that* $p \mapsto H(p, \cdot)$ *is convex. Then, in either the periodic or random setting,* $\overline{H}$ *is convex and* (6.2.5) *holds.*

Proof. The convexity of $\overline{H}$ is known in both settings, so we only prove (6.2.5) here.

Suppose that, for some continuous $v : \mathbb{R}^d \to \mathbb{R}$ and $\nu \in \mathbb{R}$, $H(Dv, y) \leq \nu$ in $\mathbb{R}^d$ in the viscosity sense. The coercivity of H implies that v is Lipschitz and, hence, satisfies the inequality almost everywhere. Since H is convex, the converse also holds (see [4]).

It follows that the supremum and infimum over $y \in \mathbb{R}^d$ in (6.3.15) may be interpreted in the almost everywhere sense, and, therefore,

$$\overline{(-H)}(p) = \sup_{v \in \mathcal{G}} \operatorname*{ess\,inf}_{y \in \mathbb{R}^d} \big(-H(p + D_y v, y)\big) = -\inf_{v \in \mathcal{G}} \operatorname*{ess\,sup}_{y \in \mathbb{R}^d} H(p + D_y v, y) = -\overline{H}(p).$$

□

The argument above fails in the nonconvex case because viscosity inequalities are sensitive to multiplication by -1. Indeed, we provide a counterexample to (6.2.5) in the periodic setting, based on an example of Luo, Tran, and Yu [58].

If v_- is a corrector for $\overline{(-H)}(p)$ and $\tilde{v}(y) := -v_-(-y)$, then $\tilde{v}$ is a viscosity solution of

$$H(D_y \tilde{v} + p, -y) = -\overline{(-H)}(p).$$

Therefore, (6.2.5) is equivalent to the invariance of $\overline{H}$ under a reflection of the periodic medium, that is, $\{(p, y) \mapsto H(p, y)\}$ and $\{(p, y) \mapsto H(p, -y)\}$ have the same effective Hamiltonian. It follows that (6.2.5) is satisfied if H is even in y, although we will not use this here.

Following [58], let $F \in C([0, \infty))$ and $0 < \theta_3 < \theta_3 < \theta_1$ be such that $F(0) = 0$, $F(\theta_2) = \frac{1}{2}$, $F(\theta_3) = F(\theta_2) = \frac{1}{3}$, F is strictly increasing on $[0, \theta_2]$ and $[\theta_1, +\infty)$ and strictly decreasing on $[\theta_2, \theta_1]$, and $\lim_{r \to +\infty} F(r) = +\infty$. For $s \in (0, 1)$, let $V_s : \mathbb{R} \to \mathbb{R}$ be the 1-periodic extension of

$$V_s(y) := \begin{cases} -\frac{x}{s} & \text{for } y \in [0,s], \\ \frac{x-1}{1-s} & \text{for } y \in [s,1]. \end{cases}$$

For $(p,y) \in \mathbb{R} \times \mathbb{R}$, set $H_s(p,y) := F(|p|) + V_s(y)$. It is shown in [58] that $\overline{H_s} = \overline{H_{s'}}$ if and only if $s = s'$. In particular, since $H_s(p,-y) = H_{1-s}(p,y)$, H_s fails to satisfy (6.2.5) whenever $s \neq \frac{1}{2}$.

6.3.2 The proof of Theorem 6.3.3

Theorem 6.3.3, and its corollaries, are essentially a consequence of the stability result Theorem 4.2.1 proved in Section 4.2.

The assumption (6.3.3), meanwhile, makes it possible to obtain a limit for solutions of (6.3.9) if the path ζ^ε is replaced with a fixed, sufficiently regular path η satisfying

(6.3.16)
$$\begin{cases} \eta : [0,\infty) \to \mathbb{R} \text{ is piecewise-}C^1 \text{ and, for any } T > 0, \\ \dot{\eta} \text{ changes sign finitely many times on } [0,T]. \end{cases}$$

Lemma 6.3.2. *Let η satisfy (6.3.16), and, for some fixed $v_0 \in BUC(\mathbb{R}^d)$, let v^ε and v solve*

$$\begin{cases} v^\varepsilon_t = H^\varepsilon(Dv^\varepsilon, x)\dot{\eta}(t) & \text{in } \mathbb{R}^d \times (0,\infty), \\ v_t = H(Dv, x)\dot{\eta}(t) & \text{in } \mathbb{R}^d \times (0,\infty), \\ v^\varepsilon(\cdot,0) = v(\cdot,0) = v_0 & \text{in } \mathbb{R}^d. \end{cases}$$

Then, for any $T > 0$,

$$\lim_{\varepsilon \to 0} \sup_{(x,t) \in B_T \times [0,T]} |v^\varepsilon(x,t) - v(x,t)| = 0.$$

It is necessary to use the following well-known domain-of-dependence result for viscosity solutions of Hamilton-Jacobi equations. For a proof, see the book of Lions [51].

Lemma 6.3.3. *Suppose that* $G : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *is continuous, let* U *and* V *be respectively a sub- and super-solution of*

$$U_t = G(DU, x) \quad \textit{and} \quad V_t = G(DV, x) \quad \textit{in } \mathbb{R}^d \times (-\infty, \infty)$$

such that $\max\{\mathrm{Lip}(U), \mathrm{Lip}(V)\} \le L$, *and suppose that*

$$\mathcal{L} := \sup_{(p,x)\in B_L\times\mathbb{R}^d} \left|D_p H(p, x)\right| < \infty.$$

Then, for all $R > 0$ *and* $-\infty < s < t < \infty$,

$$\max_{x\in B_{R-\mathcal{L}(t-s)}} |U(x, t) - V(x, t)| \le \max_{x\in B_R} |U(x, s) - V(x, s)|.$$

Proof of Lemma 6.3.2. Let the partition $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$ be such that η is monotone on each interval $[t_i, t_{i+1}]$. Fix $(x, t) \in B_T \times [0, T]$, let i be such that $t \in (t_i, t_{i+1}]$, and assume without loss of generality that η is increasing on $[t_i, t_{i+1}]$.

Set $\Delta := \eta_t - \eta_{t_i}$. It then follows that

$$v^\varepsilon(\cdot, t) = S^\varepsilon_+(\Delta)v^\varepsilon(\cdot, t_i) \quad \text{and} \quad \overline{v}(\cdot, t) = S_+(\Delta)\overline{v}(\cdot, t_i).$$

We then write

$$v^\varepsilon(x, t) - \overline{v}(x, t) = \mathrm{I} + \mathrm{II}$$

where

$$\begin{cases} \mathrm{I} := S^\varepsilon_+(\Delta)v^\varepsilon(\cdot, t_i)(x) - S^\varepsilon_+(\Delta)\overline{v}(\cdot, t_i)(x) \quad \text{and} \\ \mathrm{II} := S^\varepsilon_+(\Delta)\overline{v}(\cdot, t_i)(x) - S_+(\Delta)\overline{v}(\cdot, t_i). \end{cases}$$

In view of Theorem 4.2.1, there exists $C_1 > 0$ depending only on $\mathrm{Lip}(v_0)$ such that

$$\max(\mathrm{Lip}(v^\varepsilon), \mathrm{Lip}(\overline{v})) \le C_1.$$

The finite speed of propagation statement in Lemma 6.3.3 and the uniform growth of H in the gradient variable therefore yields a constant $C_2 > 0$ depending only on $\mathrm{Lip}(v_0)$ such that

$$|\mathrm{I}| \leq \max_{y \in B_{T+C_2\Delta}} |v^\varepsilon(y, t_i) - \overline{v}(y, t_i)| \, .$$

An inductive argument yields

$$|v^\varepsilon(x,t) - \overline{v}(x,t)| \leq \sum_{i=0}^{N-1} \max_{(y,\tau) \in B_{R_i} \times [0, \Delta_i]} \left| S^\varepsilon_\pm(\tau)\overline{v}(\cdot, t_i)(y) - S_\pm(\tau)\overline{v}(\cdot, t_i)(y) \right|$$

where

$$\Delta_i := |\eta(t_{i+1}) - \eta(t_i)| \quad \text{and} \quad R_i := T + C_2 \sum_{k=i}^{N-1} \Delta_k.$$

Sending $\varepsilon \to 0$ and using (6.3.3) yields the result. □

Proof of Theorem 6.3.3. (a) Fix a path η satisfying (6.3.16) and let v^ε and $\overline{v}$ be as in the statement of Lemma 6.3.2. Then Theorems 4.1.1 and 4.2.1 yield, for some $C = C_L > 0$,

$$\max_{(x,t) \in B_T \times [0,T]} \left(|u^\varepsilon(x,t) - v^\varepsilon(x,t)| + |\overline{u}(x,t) - \overline{v}(x,t)| \right) \leq C \max_{0 \leq t \leq T} |\zeta^\varepsilon(t) - \eta(t)| \, .$$

Lemma 6.3.2 then gives

$$\limsup_{\varepsilon \to 0} |u^\varepsilon(x,t) - \overline{u}(x,t)| \leq C \max_{0 \leq t \leq T} |\zeta(t) - \eta(t)| \, .$$

The result follows, as η was arbitrary.

(b) As a consequence of the Portmanteau Theorem (see [14]), it is enough to show that, for any $T > 0$ and open set $G \subset BUC(B_T \times [0,T])$,

$$\liminf_{\varepsilon \to 0} \mathbb{P}\left(u^\varepsilon \in G\right) \geq \mathbb{P}\left(u \in G\right),$$

where, through abuse of notation, we identify u^ε and u with the restrictions to $B_T \times [0,T]$.

Below, for $\sigma > 0$, define the open set

$$G_\sigma := \left\{ v \in G : \|v - w\|_{\infty, B_T \times [0,T]} > \sigma \text{ for all } w \in G^c \right\}.$$

Fix $\delta > 0$, and let η be a random path such that (6.3.16) holds and $\|\zeta - \eta\|_{\infty,T} < \delta$. For example, η could be a piecewise linear interpolation of ζ over an appropriately defined partition, using the fact that any fixed realization of ζ is uniformly continuous on $[0,T]$.

Let v^ε and v be as in the statement of Lemma 6.3.2 with the path η. Then

$$\mathbb{P}\left(\lim_{\varepsilon \to 0} v^\varepsilon = v \quad \text{uniformly on } B_T \times [0,T] \right) = 1,$$

and so, by Egoroff's Theorem, there exists a deterministic $\varepsilon_0 > 0$ such that

$$\mathbb{P}\left(\|v^\varepsilon - v\|_{\infty, B_T \times [0,T]} < \delta \quad \text{for all } \varepsilon \in (0, \varepsilon_0) \right) \geq 1 - \delta.$$

Then, for all $\varepsilon \in (0, \varepsilon_0)$,

$$\begin{aligned}
\mathbb{P}\left(u^\varepsilon \in G\right) &\geq \mathbb{P}\left(v + u^\varepsilon - v^\varepsilon \in G_\delta\right) - \delta \geq \mathbb{P}\left(u + u^\varepsilon - v^\varepsilon \in G_{2\delta}\right) - \delta \\
&\geq \mathbb{P}\left(u \in G_{3\delta} \text{ and } \|\zeta^\varepsilon - \eta\|_T < \delta\right) - \delta \\
&= \mathbb{P}\left((u, \zeta^\varepsilon - \eta) \in G_{3\delta} \times B_\delta(0)\right) - \delta.
\end{aligned}$$

Since $(u, \zeta^\varepsilon - \eta)$ converges in distribution to $(u, \zeta - \eta)$ in $BUC(B_T \times [0,T]) \times C([0,T])$, letting $\varepsilon \to 0$ yields

$$\liminf_{\varepsilon \to 0} \mathbb{P}\left(u^\varepsilon \in G\right) \geq \mathbb{P}\left(u \in G_{3\delta}\right) - \delta,$$

and the result is proved because δ was arbitrary.

(c) Assume H satisfies (6.3.2), $\zeta \in C([0,T])$, $\{\zeta^\varepsilon\}_{\varepsilon>0}$ and $\{\tilde{\zeta}^\varepsilon\}_{\varepsilon>0}$ are two families of piecewise-C^1 paths on $[0,T]$ that converge uniformly to ζ as $\varepsilon \to 0$, and w^ε and $\tilde{w}^\varepsilon$ are the corresponding viscosity solutions of (6.3.10). Then Theorem 4.2.1 implies that the sequences

$\{w^\varepsilon\}_{\varepsilon>0}$ and $\{\tilde{w}^\varepsilon\}_{\varepsilon>0}$ are Cauchy, and, for some $C = C_L > 0$,

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |w^\varepsilon(x,t) - \tilde{w}^\varepsilon(x,t)| \le C \max_{t\in[0,T]} \left|\zeta^\varepsilon(t) - \tilde{\zeta}^\varepsilon(t)\right| \xrightarrow{\varepsilon\to 0} 0.$$

It follows that the solution operator for (6.3.10) extends uniquely to continuous paths. That is, as $\varepsilon \to 0$, w^ε and $\tilde{w}^\varepsilon$ converge uniformly in $\mathbb{R}^d \times [0,T]$ to a unique limit, which is a pathwise viscosity solution of (6.3.10).

In parts (a) and (b), the limiting solution is obtained exactly as the extension of this solution operator. Therefore, the limit is unique and does not depend on the approximating paths.

(d) If the comparison principle holds for (6.3.10), then this together with Theorem 4.2.1 implies that (6.3.10) has a unique pathwise viscosity solution for any choice of continuous path ζ. The results in the previous parts now follow easily for continuous paths. $\square$

6.4 The multiple-noise case

We discuss next the case when there are multiple driving signals. More precisely, we are interested in the behavior of equations of the form

$$(6.4.1)\quad u^\varepsilon_t + \frac{1}{\varepsilon^\gamma}\sum_{i=0}^m H^i(Du^\varepsilon, x/\varepsilon)\xi^i(t/\varepsilon^{2\gamma}) \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u^\varepsilon(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d.$$

Here, for each $i = 0, 1, 2, \ldots, m$, ξ^i is a mixing field satisfying (6.1.6), although we will be making more specific assumptions about these and the Hamiltonians.

To simplify the presentation, in this section we will only consider the periodic setting (6.3.11). Then, under sufficient conditions on the H^i's which are made more specific below, for every $p \in \mathbb{R}^d$ and $\xi \in \mathbb{R}^m$, there exists a unique constant $\overline{H}(p,\xi)$ such that the cell problem

$$\sum_{i=1}^{m} H^i(p + D_y v, y)\xi^i = \overline{H}(p, \xi) \tag{6.4.2}$$

admits periodic solutions $v : \mathbb{T}^d \to \mathbb{R}$. Furthermore, $\xi \mapsto \overline{H}(p, \xi)$ is positively homogenous, and

$$\mathbb{E}[\overline{H}(p, \xi(0)] = 0 \quad \text{for all } p \in \mathbb{R}^d. \tag{6.4.3}$$

As discussed in Section 6.2 above, we will attempt to show that u^ε is closely approximated by the solution $\overline{u}^\varepsilon$ of

$$\overline{u}^\varepsilon_t + \frac{1}{\varepsilon^\gamma}\overline{H}\left(D\overline{u}^\varepsilon, \xi\left(\frac{t}{\varepsilon^{2\gamma}}\right)\right) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad \overline{u}^\varepsilon(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d. \tag{6.4.4}$$

The main tool for doing so is the homogenization error estimate obtained by Capuzzo-Dolcetta and Ishii [20], or Theorem 6.3.5 above.

The limiting behavior for (6.4.4) is well understood if $u_0(x) = p_0 \cdot x$ for some fixed $p_0 \in \mathbb{R}^d$. Indeed, in view of the mixing properties of ξ and the centering property (6.4.3), as $\varepsilon \to 0$, $\overline{u}^\varepsilon$ converges locally uniformly in distribution to

$$p_0 \cdot x + \mathbb{E}\left[\overline{H}(p_0, \xi(0))^2\right]^{1/2} B(t),$$

where B is a Brownian motion. We will then be concerned with describing the limit of $\overline{u}^\varepsilon$ for arbitrary initial data u_0. Under certain assumptions on the Hamiltonians and mixing fields, the goal will be to explain that there exists $M \geq 1$ and, for each $j = 1, 2, \ldots, M$, an effective Hamiltonian $\overline{H}^j : \mathbb{R}^d \to \mathbb{R}$ satisfying (2.4.4) and a Brownian motion B^j such that, as $\varepsilon \to 0$ and in distribution, $\overline{u}^\varepsilon$ (and therefore u^ε) converges in $BUC(\mathbb{R}^d \times (0, \infty))$ to the pathwise viscosity solution $\overline{u}$ of

$$d\overline{u} + \sum_{j=1}^{M} \overline{H}^j(D\overline{u}) \circ dB^j = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad \overline{u} = u_0 \quad \text{in } \mathbb{R}^d. \tag{6.4.5}$$

Although at first glance, the nature of the problem is similar to the single-noise case, there will nevertheless be some fundamental differences. Most importantly, the deterministic effective Hamiltonians $\{\overline{H}^j\}_{j=1}^M$, and even their number M, depend on the particular law of the mixing field ξ.

We emphasize that this will imply a lack of universality for the limit of u^ε, even in probability law. This is related to the following situation: suppose that u and $\tilde{u}$ are the stochastic viscosity solutions of respectively

$$\begin{cases} du + \sum_{i=1}^{M} H^i(Du) \circ dB^i = 0 \quad \text{and} \\ d\tilde{u} + \sum_{j=1}^{\tilde{M}} \tilde{H}^j(D\tilde{u}) \circ d\tilde{B}^j = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty), \quad \text{and} \\ u(\cdot,0) = \tilde{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d, \end{cases}$$

where $\{B^i\}_{i=1}^M$, $\{\tilde{B}^j\}_{j=1}^{\tilde{M}}$ are independent Brownian motions, and

$$\sum_{i=1}^{M} \left[H^i(p) \right]^2 = \sum_{j=1}^{\tilde{M}} \left[\tilde{H}^j(p) \right]^2 \quad \text{for all } p \in \mathbb{R}^d. \tag{6.4.6}$$

The equality (6.4.6) implies that u and $\tilde{u}$ have the same distribution whenever $u_0(x) = p_0 \cdot x$ for some fixed $p_0 \in \mathbb{R}^d$. However, this fails for arbitrary $u_0 \in UC(\mathbb{R}^d)$, even if $M = \tilde{M} = 1$, as can be seen by the following example:

$$\begin{cases} du - u_x \circ dB = 0 \quad \text{and} \\ d\tilde{u} - |u_x| \circ d\tilde{B} = 0 \quad \text{in } \mathbb{R} \times (0,\infty), \quad \text{and} \\ u(x,0) = \tilde{u}(x,0) = |x| \quad \text{in } \mathbb{R}. \end{cases}$$

The Hamiltonians $H(p) = p$ and $\tilde{H}(p) = |p|$ clearly satisfy (6.4.6). However, a simple calculation yields $u(x,t) = |x + B(t)|$, while it is shown in [55, 53, 76] that

$$\tilde{u}(x,t) = \max\left\{ |x| + \tilde{B}(t), \max_{0 \le s \le t} \tilde{B}(s) \right\}.$$

The solutions u and $\tilde{u}$ then fail to have the same law. For instance, $u(0,t) = |B(t)|$, while $\tilde{u}(0,t) = \max_{0 \le s \le t} \tilde{B}(t)$ is an increasing function.

6.4.1 A general class of examples

We now make some further assumptions that give rise to a rich class of examples and results.

For the Hamiltonians $\{H^i\}_{i=1}^m$, we assume that

$$\begin{cases} H^i \in C^{0,1}(\mathbb{R}^d \times \mathbb{T}^d), \\ p \mapsto H^1(p,\cdot) + \sum_{i=2}^m H^i(p,\cdot)\xi^i \text{ is convex for all } \xi^2, \xi^3, \dots, \xi^m \in \{-1,1\}, \text{ and} \\ \lim_{|p| \to +\infty} \inf_{y \in \mathbb{T}^d} \left(H^1(p,y) - \sum_{i=2}^m \left| H^i(p,y) \right| \right) = +\infty. \end{cases} \tag{6.4.7}$$

As a consequence of Theorem 6.3.4, the cell problem (6.4.2) is solvable for all $p \in \mathbb{R}^d$ and $\xi \in \{-1,1\}^m$, and furthermore, $p \mapsto \overline{H}(p,1,\xi)$ is convex and $\xi \mapsto \overline{H}(p,\xi)$ is homogenous. That is, for all $\lambda \in \mathbb{R}$ and $\xi \in \{-1,1\}^m$,

$$\overline{H}(\cdot, \lambda\xi) = \lambda \overline{H}(\cdot, \xi). \tag{6.4.8}$$

That (6.4.8) holds for positive λ was already observed in Section 6.2. The full assertion is a consequence of Theorem 6.3.4 and the convexity assumption in (6.4.7).

The mixing fields are assumed to satisfy, for $i = 1, \ldots, m$,

$$\text{(6.4.9)} \qquad \begin{cases} \xi^i = \sum_{k=0}^{\infty} X_k^i \mathbf{1}_{(k,k+1)} & \text{where} \\ \left\{X_k^i\right\}_{i=1,2,\ldots,m,\ k=0,1,\ldots} & \text{are independent Rademacher random variables.} \end{cases}$$

In particular, if

$$\text{(6.4.10)} \qquad \xi^{i,\varepsilon}(t) := \frac{1}{\varepsilon^\gamma}\xi^i(t/\varepsilon^{2\gamma}) \quad \text{and} \quad \zeta^{i,\varepsilon}(t) := \int_0^t \xi^{i,\varepsilon}(s)\, ds,$$

then each $\zeta^{i,\varepsilon}$ is a scaled, linearly-interpolated, simple random walk on $\mathbb{Z}$, and, in distribution,

$$(\zeta^{1,\varepsilon}, \zeta^{2,\varepsilon}, \ldots, \zeta^{m,\varepsilon}) \xrightarrow{\varepsilon \to 0} (B^1, B^2, \ldots, B^m)$$

in $C([0,\infty), \mathbb{R}^m)$, where $(B^1, B^2, \ldots, B^m)$ is an m-dimensional Brownian motion.

Define

$$\begin{cases} \mathcal{A}^m := \{\mathbf{j} = (j_1, j_2, \ldots, j_l) : j_i \in \{1, 2, \ldots, m\},\ j_1 < j_2 < \cdots < j_l\}, \\ |\mathbf{j}| = |(j_1, j_2, \ldots, j_l)| := l, \quad \text{and} \\ \mathcal{A}_o^m := \{\mathbf{j} \in \mathcal{A}^m : |\mathbf{j}| \text{ is odd.}\}. \end{cases}$$

Note that $\#\mathcal{A}^m = 2^m - 1$ and $\#\mathcal{A}_0^m = 2^{m-1}$.

For any $\mathbf{j} = (j_1, j_2, \ldots, j_l) \in \mathcal{A}^m$, define

$$(6.4.11)\qquad \begin{cases} \xi^{\mathbf{j}} := \xi^{j_1}\xi^{j_2}\cdots\xi^{j_l} \quad \text{for } \xi = (\xi^1, \xi^2, \ldots, \xi^m) \in \{-1,1\}^m, \\ \overline{H}^{\mathbf{j}}(p) := \displaystyle\sum_{\xi\in\{-1,1\}^m} 2^{-m}\overline{H}(p,\xi)\xi^{\mathbf{j}}, \\ X_k^{\mathbf{j}} := X_k^{j_1}X_k^{j_2}\cdots X_k^{j_l}, \\ \zeta^{\mathbf{j}}(0) := 0, \quad \dot{\zeta}^{\mathbf{j}} := \displaystyle\sum_{k=0}^{\infty} X_k^{\mathbf{j}}\mathbf{1}_{(k,k+1)}, \quad \text{and} \quad \zeta^{\mathbf{j},\varepsilon}(t) := \varepsilon^{\gamma}\zeta^{\mathbf{j}}(t/\varepsilon^{2\gamma}). \end{cases}$$

Observe that, for each $\mathbf{j} \in \mathcal{A}_0^m$, $\overline{H}^{\mathbf{j}}$ is a difference of convex functions. Note also that if $|\mathbf{j}|$ is even, then the homogeneity property (6.4.8) implies that $\overline{H}^{\mathbf{j}} = 0$.

We then have the following result:

Theorem 6.4.1. *Assume that* $0 < \gamma < 1/6$, $u_0 \in \mathrm{Lip}(\mathbb{R}^d)$, (6.4.7), *and* (6.4.9), *and let* u^ε *be the solution of* (6.4.1). *Then there exist* 2^{m-1} *independent Brownian motions* $\{B^{\mathbf{j}}\}_{\mathbf{j}\in\mathcal{A}_o^m}$, *such that, in distribution,*

$$\left(u^\varepsilon, \{\zeta^{\mathbf{j},\varepsilon}\}_{\mathbf{j}\in\mathcal{A}_o^m}\right) \xrightarrow{\varepsilon\to 0} \left(\overline{u}, \{B^{\mathbf{j}}\}_{\mathbf{j}\in\mathcal{A}_o^m}\right) \quad \text{in } BUC(\mathbb{R}^d\times[0,T])\times C\left([0,T],\mathbb{R}^{2^{m-1}}\right),$$

where $\overline{u}$ *is the stochastic viscosity solution of*

$$(6.4.12)\qquad d\overline{u} + \sum_{\mathbf{j}\in\mathcal{A}_0^m} \overline{H}^{\mathbf{j}}(D\overline{u})\circ dB^{\mathbf{j}} = 0 \quad \text{in } \mathbb{R}^d\times(0,\infty) \quad \text{and} \quad \overline{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d.$$

The result relies on the fact that, due to the assumptions on the mixing fields ξ^i, which take their values only in $\{-1,1\}$, the general effective Hamiltonian $\overline{H}(p,\xi)$ can be decomposed using a combinatorial argument. This is described by the next lemmas.

Lemma 6.4.1. *Let* $\{X^j\}_{j=1}^m$ *be mutually independent Rademacher random variables. Then* *the random variables defined by*

$$X^{\mathbf{j}} := X^{j_1} X^{j_2} \cdots X^{j_l} \quad \textit{for } \mathbf{j} = (j_1, j_2, \ldots, j_l) \in \mathcal{A}^m$$

are pairwise independent and Rademacher.

Proof. The Rademacher property follows from the mutual independence of the $\{X^j\}_{j=1}^m$.

To prove the pairwise independence, it is enough to establish the following two facts: if X, Y, and Z are pairwise independent Rademacher random variables, then

$$\begin{cases} X \text{ and } XZ \text{ are independent, and} \\ XZ \text{ and } YZ \text{ are independent.} \end{cases}$$

Both facts are easily checked. □

Lemma 6.4.2. *Let* $f : \{-1, 1\}^m \to \mathbb{R}$. *Then*

$$f(\xi) = f_0 + \sum_{\mathbf{j} \in \mathcal{A}^m} f_{\mathbf{j}} \xi^{\mathbf{j}}, \tag{6.4.13}$$

where

$$f_0 := \frac{1}{2^m} \sum_{\xi \in \{-1,1\}^m} f(\xi) \quad \textit{and} \quad f_{\mathbf{j}} := \frac{1}{2^m} \sum_{\xi \in \{-1,1\}^m} f(\xi) \xi^{\mathbf{j}}.$$

If f is odd, then $f_0 = 0$ and the sum in (6.4.13) *is taken over* $\mathbf{j} \in \mathcal{A}_o^m$.

Proof. The last statement about odd f follows easily from the formulas. To prove these, let $\mathcal{F}^m$ be the space of all real-valued functions on $\{-1, 1\}^m$. This is a 2^m-dimensional vector space. On the other hand, the 2^m functions in the collection $\mathcal{P}^m := \{1, \{\xi^{\mathbf{j}}\}_{\mathbf{j} \in \mathcal{A}^m}\}$ are linearly independent elements of $\mathcal{F}^m$, and therefore, their span is equal to $\mathcal{F}^m$.

For $f, g \in \mathcal{F}^m$, define the inner product

$$\langle f, g \rangle_{\mathcal{F}^m} := \frac{1}{2^m} \sum_{\xi \in \{-1,1\}^m} f(\xi) g(\xi).$$

With respect to $\langle \cdot, \cdot \rangle_{\mathcal{F}^m}$, $\mathcal{P}^m$ becomes an orthonormal basis, so that, for any $f \in \mathcal{F}^m$,

$$f = \sum_{q \in \mathcal{P}^m} \langle f, q \rangle_{\mathcal{F}^m} q,$$

which is the desired conclusion. □

As a corollary, we relate the effective Hamiltonian $\overline{H} : \mathbb{R}^d \times \mathbb{R}^m$ from (6.4.2) to the definitions in (6.4.11):

Corollary 6.4.1. *For all $p \in \mathbb{R}^d$ and $\xi \in \{-1, 1\}^m$,*

$$\overline{H}(p, \xi) := \sum_{\mathbf{j} \in \mathcal{A}_o^m} \overline{H}^{\mathbf{j}}(p) \xi^{\mathbf{j}}.$$

Now let $\overline{u}^\varepsilon$ be the solution of the equation

$$\overline{u}_t^\varepsilon + \sum_{\mathbf{j} \in \mathcal{A}_o^m} \overline{H}^{\mathbf{j}}(D\overline{u}^\varepsilon) \dot{\zeta}^{\mathbf{j},\varepsilon}(t) = 0 \quad \text{in } \mathbb{R}^d \times (0, \infty) \quad \text{and} \quad \overline{u}^\varepsilon(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d, \tag{6.4.14}$$

where the $\overline{H}^{\mathbf{j}}$'s and $\zeta^{\mathbf{j}}$'s are as in (6.4.11).

Lemma 6.4.3. *There exists $C = C_L > 0$ such that, whenever $\operatorname{Lip}(u_0) \le L$, $\varepsilon > 0$ and $T > 1$,*

$$\sup_{(x,t) \in \mathbb{R}^d \times [0,T]} |u^\varepsilon(x, t) - \overline{u}^\varepsilon(x, t)| \le CT\varepsilon^{1/3 - 2\gamma}.$$

Proof. We do not give the full details of the proof, as it is very similar to that of Lemma 6.3.2, although quantitative, and made simpler by the fact that periodic homogenization is global in space, so that the finite speed of propagation property from Lemma 6.3.3 does not need to be used.

The argument follows by applying the $O(\varepsilon^{1/3})$ rate of convergence from the periodic homogenization of Hamilton-Jacobi equations on each of the approximately $1/\varepsilon^{2\gamma}$ intervals on which $\xi^\varepsilon(t)$ is constant. The effective equation on each of those intervals is given by

$$\overline{u}_t^\varepsilon + \overline{H}(D\overline{u}^\varepsilon, \varepsilon^{-\gamma}\xi(t/\varepsilon^{2\gamma})) = 0,$$

which is exactly equation (6.4.14) in view of Corollary 6.4.1. $\square$

We finally present the

Proof of Theorem 6.4.1. Because $\gamma < \frac{1}{6}$, Lemma 6.4.3 implies that, with probability one,

$$\lim_{\varepsilon \to 0} \sup_{(x,t) \in \mathbb{R}^d \times [0,T]} |u^\varepsilon(x,t) - \overline{u}^\varepsilon(x,t)| = 0.$$

In view of Lemma 6.4.1, the path

$$\{\zeta^{\mathbf{j},\varepsilon}\}_{\mathbf{j} \in \mathcal{A}_o^m} \in C\left([0,\infty), \mathbb{R}^{2^{m-1}}\right)$$

is a random walk which, as $\varepsilon \to 0$, converges in distribution to a 2^{m-1}-dimensional Brownian motion $\{B^{\mathbf{j}}\}_{\mathbf{j} \in \mathcal{A}_o^m}$.

The stability result Theorem 4.1.1 implies that the solution operator for the equation (6.4.12) given by

$$S : C\left([0,T], \mathbb{R}^{2^{m-1}}\right) \ni B \mapsto \overline{u} \in BUC(\mathbb{R}^d \times [0,T])$$

is continuous, and, therefore, so is the graph map

$$(\mathrm{Id}, S) : C\left([0,T], \mathbb{R}^{2^{m-1}}\right) \ni B \mapsto (B, \overline{u}) \in C\left([0,T], \mathbb{R}^{2^{m-1}}\right) \times BUC(\mathbb{R}^d \times [0,T]).$$

It follows from the Mapping Theorem that, if $\overline{u}^\varepsilon$ is the solution of (6.4.14), then, as $\varepsilon \to 0$, $(\overline{u}^\varepsilon, \zeta^\varepsilon)$ converges in distribution to $(\overline{u}, B)$ in $BUC(\mathbb{R}^d \times [0,T]) \times C([0,T], \mathbb{R})^{2^{m-1}}$. The result now follows from Slutzky's theorem. $\square$

6.4.2 A one-dimensional example

We begin with some examples in the case $d = 1$.

We consider the following cell problem

$$|p + v'(y)| + F(y) = \overline{H}(p) \quad \text{on } \mathbb{T}. \tag{6.4.15}$$

It what follows, define $\langle V\rangle := \int_0^1 V(y)\,dy$ for any $V \in C(\mathbb{T})$.

Lemma 6.4.4. *The effective Hamiltonian* $\overline{H}$ *in* (6.4.15) *is given by*

$$\overline{H}(p) = \max\left\{\max_{y\in\mathbb{T}} F(y), |p| + \langle F\rangle\right\}.$$

Proof. Assume first that $|p| \le \max F - \langle F\rangle$. Let $y_m \in [0,1]$ be such that $\max F = F(y_m)$. Then there exists $y^* \in [y_m, 1+y_m]$ such that

$$\int_{y_m}^{y^*} (\max F - F(t))\,dt - \int_{y^*}^{1+y_m} (\max F - F(t))\,dt = p.$$

We then define a function v on $[y_m, 1+y_m]$ as below and extend it periodically to the rest of $\mathbb{R}$:

$$v(y) := \begin{cases} \displaystyle\int_{y_m}^{y} (\max F - F(t) - p)\,dt & \text{if } y_m \le y \le y^*, \text{ and} \\ \displaystyle v(y^*) - \int_{y^*}^{y} (\max F - F(t) + p)\,dt & \text{if } y^* < y \le 1 + y_m. \end{cases}$$

It is standard to check that v is a viscosity solution of $|v' + p| + F(y) = \max F$ in $\mathbb{R}$.

Now assume that $p > \max F - \langle F\rangle$. Then

$$v(y) := \int_0^y (\langle F\rangle - F(t))\,dt$$

is a C^1 solution of the equation $|v' + p| + F(y) = p + \langle F\rangle$. A similar construction works for $p < -(\max F - \langle F\rangle)$.

$\square$

Now, for $u_0 \in \mathrm{Lip}(\mathbb{R})$, $\zeta^{1,\varepsilon}$ and $\zeta^{2,\varepsilon}$ as in (6.4.10), and $f \in C^{0,1}(\mathbb{T})$, consider the equation

$$\begin{cases} u^\varepsilon_t + |\partial_x u^\varepsilon| \dot{\zeta}^{1,\varepsilon} + f\left(\frac{x}{\varepsilon}\right)\dot{\zeta}^{2,\varepsilon} = 0 \quad \text{in } \mathbb{R}\times(0,\infty) \quad \text{and} \\ u^\varepsilon(\cdot,0) = u_0 \quad \text{in } \mathbb{R}. \end{cases} \tag{6.4.16}$$

Theorem 6.4.1 implies that, as $\varepsilon \to 0$, $(u^\varepsilon, \zeta^{1,\varepsilon}, \zeta^{2,\varepsilon})$ converges in distribution to $(\overline{u}, B^1, B^2)$, where B^1 and B^2 are independent Brownian motions and $\overline{u}$ is the solution of an equation of the form

$$\begin{cases} d\overline{u} + \overline{H}^1(\partial_x \overline{u}) \circ dB^1 + \overline{H}^2(\partial_x \overline{u}) \circ dB^2 = 0 \quad \text{in } \mathbb{R}\times(0,\infty) \quad \text{and} \\ \overline{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R}, \end{cases} \tag{6.4.17}$$

provided that $\gamma < \frac{1}{6}$. In fact, the convergence actually holds if $\gamma < 1$, as a result of the recent homogenization error estimates proved in [61].

Using the formulas in (6.4.11) and Lemma 6.4.4, we explicitly compute $\overline{H}^1$ and $\overline{H}^2$. In particular, for all $p \in \mathbb{R}^d$ and $\xi^1, \xi^2 \in \{-1, 1\}$,

$$\xi^1 \max\left\{ \max_{y\in\mathbb{T}} \left(\frac{\xi^2}{\xi^1} f(y) \right), |p| + \langle f \rangle \frac{\xi^2}{\xi^1} \right\} = H^1(p)\xi^1 + H^2(p)\xi^2.$$

It turns out that these depend on the relationship between the average of f and its amplitude. More precisely, we need to split into two cases, depending on whether

$$\langle f \rangle > \frac{\max f - \max f}{2} \quad \text{or} \quad \langle f \rangle < \frac{\max f - \min f}{2}.$$

In the first case, we say that f skews upwards, and in the second, we say that f skews downwards.

If f skews upwards, then $0 \leq \max f - \langle f \rangle < \langle f \rangle - \min f$, and

$$\overline{H}^1(p) = \begin{cases} \dfrac{\max f - \min f}{2} & \text{if } |p| \leq \max f - \langle f \rangle, \\ \dfrac{1}{2}|p| + \dfrac{1}{2}\left(\langle f \rangle - \min f\right) & \text{if } \max f - \langle f \rangle < |p| \leq \langle f \rangle - \min f, \\ |p| & \text{if } |p| > \langle f \rangle - \min f, \end{cases}$$

and

$$\overline{H}^2(p) = \begin{cases} \dfrac{\max f + \min f}{2} & \text{if } |p| \leq \max f - \langle f \rangle, \\ \dfrac{1}{2}|p| + \dfrac{1}{2}\left(\langle f \rangle + \min f\right) & \text{if } \max f - \langle f \rangle < |p| \leq \langle f \rangle - \min f, \\ \langle f \rangle & \text{if } |p| > \langle f \rangle - \min f. \end{cases}$$

If f skews downwards, then $0 \leq \langle f \rangle - \min f < \max f - \langle f \rangle$, and

$$\overline{H}^1(p) = \begin{cases} \dfrac{\max f - \min f}{2} & \text{if } |p| \leq \langle f \rangle - \min f, \\ \dfrac{1}{2}|p| + \dfrac{1}{2}\left(\max f - \langle f \rangle\right) & \text{if } \langle f \rangle - \min f < |p| \leq \max f - \langle f \rangle, \\ |p| & \text{if } |p| > \max f - \langle f \rangle, \end{cases}$$

and

$$\overline{H}^2(p) = \begin{cases} \dfrac{\max f + \min f}{2} & \text{if } |p| \leq \langle f \rangle - \min f, \\ \dfrac{1}{2}|p| + \dfrac{1}{2}\left(\max f - \langle f \rangle\right) & \text{if } \langle f \rangle - \min f < |p| \leq \max f - \langle f \rangle, \\ \langle f \rangle & \text{if } |p| > \max f - \langle f \rangle. \end{cases}$$

6.4.3 Dependence of the limit on the noise approximation

We consider equation (6.4.16) once more, but for different approximating paths.

We will still define, for $i = 1, 2$, $\zeta^{i,\varepsilon}(t) = \varepsilon^{\gamma}\zeta^i(t/\varepsilon^{2\gamma})$, where

$$\dot{\zeta}^i(t) = \sum_{k=0}^{\infty} X_k^i \mathbf{1}_{(k,k+1)}(t)$$

for some collection of independent random variables $\{X_k^i\}_{i=1,2,\ k=0,1,2,\ldots}$. We will still assume that

$$\{X_k^1\}_{k=0,1,2,\ldots} \quad \text{are independent Rademacher random variables.}$$

However, for the other collection, we will assume

$$\begin{cases} X_k^2 = \dfrac{a+b}{2} Y_k + \dfrac{a-b}{2} Z_k, \\ 0 < b < a, \quad a^2 + b^2 = 2, \quad a(\max f - \langle f \rangle) < b(\langle f \rangle - \min f), \\ \{Y_k, Z_k\}_{k=0}^{\infty} \quad \text{are independent Rademacher random variables.} \end{cases}$$

Note that

$$\begin{cases} \mathbb{P}\left(X_k^2 = a\right) = \mathbb{P}\left(X_k^2 = b\right) = \mathbb{P}\left(X_k^2 = -a\right) = \mathbb{P}\left(X_k^2 = -b\right) = \dfrac{1}{4}, \\ \mathbb{E} X_k^2 = 0, \quad \text{and} \quad \mathbb{E}|X_k^2|^2 = 1. \end{cases}$$

In particular, there exists a standard Brownian motion (B^1, B^2) such that

$$\left(\zeta^{1,\varepsilon}, \zeta^{2,\varepsilon}\right) \xrightarrow{\varepsilon \to 0} (B^1, B^2) \quad \text{in distribution.}$$

However, the limiting equation for u^ε is no longer (6.4.17). Define the approximating paths $\zeta^{\mathbf{j},\varepsilon}(t) := \varepsilon^\gamma \zeta^{\mathbf{j}}(t/\varepsilon^{2\gamma})$ for $\mathbf{j} \in \{\{1\}, \{2\}, \{3\}, \{1,2,3\}\}$, where

$$\begin{cases} \zeta^{\{1\},\varepsilon} := \zeta^{1,\varepsilon}, \quad \zeta^{\{2\}}(0) = \zeta^{\{3\}}(0) = \zeta^{\{1,2,3\}}(0) := 0, \\ \dot{\zeta}^{\{2\}}(t) := \displaystyle\sum_{k=0}^{\infty} Y_k \mathbf{1}_{(k,k+1)}(t), \quad \dot{\zeta}^{\{3\}}(t) := \displaystyle\sum_{k=0}^{\infty} Z_k \mathbf{1}_{(k,k+1)}(t), \quad \text{and} \\ \dot{\zeta}^{\{1,2,3\}}(t) := \displaystyle\sum_{k=0}^{\infty} X_k^1 Y_k Z_k \mathbf{1}_{(k,k+1)}(t). \end{cases}$$

Equation (6.4.16) then becomes

$$(6.4.18)\qquad \begin{cases} u_t^\varepsilon + |\partial_x u^\varepsilon| \dot{\zeta}^{\{1\},\varepsilon} + \dfrac{a+b}{2} f\left(\dfrac{x}{\varepsilon}\right) \dot{\zeta}^{\{2\},\varepsilon} \\ \qquad + \dfrac{a-b}{2} f\left(\dfrac{x}{\varepsilon}\right) \dot{\zeta}^{\{3\},\varepsilon} = 0 \quad \text{in } \mathbb{R} \times (0,\infty) \quad \text{and} \\ u^\varepsilon(\cdot,0) = u_0 \quad \text{in } \mathbb{R}. \end{cases}$$

This falls into the class of equations studied in Theorem 6.4.1. Applying this result shows that, if $\gamma < 1$, then, for some independent Brownian motions $B^{\mathbf{j}}$ with $\mathbf{j}$ equal to $\{1\}$, $\{2\}$, $\{3\}$, or $\{1,2,3\}$,

$$\left(u^\varepsilon, \zeta^{\{1\},\varepsilon}, \zeta^{\{2\},\varepsilon}, \zeta^{\{3\},\varepsilon}, \zeta^{\{1,2,3\},\varepsilon}\right) \xrightarrow{\varepsilon\to 0} \left(\overline{u}, B^{\{1\}}, B^{\{2\}}, B^{\{3\}}, B^{\{1,2,3\}}\right) \quad \text{in distribution,}$$

where $\overline{u}$ is the stochastic viscosity solution of

$$(6.4.19)\qquad \begin{cases} d\overline{u} + \overline{H}^{\{1\}}(\partial_x \overline{u}) \circ dB^{\{1\}} + \overline{H}^{\{2\}}(\partial_x \overline{u}) \circ dB^{\{2\}} + \overline{H}^{\{3\}}(\partial_x \overline{u}) \circ dB^{\{3\}} \\ \qquad + \overline{H}^{\{1,2,3\}}(\partial_x \overline{u}) \circ dB^{\{1,2,3\}} = 0 \quad \text{in } \mathbb{R} \times (0,\infty) \quad \text{and} \\ \overline{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R} \end{cases}$$

with

$$\overline{H}^{\{1\}}(p) :=$$
$$\begin{cases} \dfrac{a+b}{4}(\max f - \min f) & \text{if } 0 \le |p| \le b(\max f - \langle f \rangle), \\ \dfrac{1}{4}|p| + \dfrac{a}{4}(\max f - \min f) + \dfrac{b}{4}(\langle f \rangle - \min f) & \text{if } b(\max f - \langle f \rangle) \le |p| \le a(\max f - \langle f \rangle), \\ \dfrac{1}{2}|p| + \dfrac{a+b}{4}(\langle f \rangle - \min f) & \text{if } a(\max f - \langle f \rangle) \le |p| \le b(\langle f \rangle - \min f), \\ \dfrac{3}{4}|p| + \dfrac{a}{4}(\langle f \rangle - \min f) & \text{if } b(\langle f \rangle - \min f) \le |p| \le a(\langle f \rangle - \min f), \\ |p| & \text{if } |p| \ge a(\langle f \rangle - \min f), \end{cases}$$

$$\overline{H}^{\{2\}}(p) :=$$

$$\begin{cases} \dfrac{a+b}{4}(\max f + \min f) & \text{if } 0 \le |p| \le b(\max f - \langle f\rangle), \\ \dfrac{1}{4}|p| + \dfrac{a}{4}(\max f + \min f) + \dfrac{b}{4}(\langle f\rangle + \min f) & \text{if } b(\max f - \langle f\rangle) \le |p| \le a(\max f - \langle f\rangle), \\ \dfrac{1}{2}|p| + \dfrac{a+b}{4}(\langle f\rangle + \min f) & \text{if } a(\max f - \langle f\rangle) \le |p| \le b(\langle f\rangle - \min f), \\ \dfrac{1}{4}|p| + \dfrac{a}{4}(\langle f\rangle + \min f) + \dfrac{b}{2}\langle f\rangle & \text{if } b(\langle f\rangle - \min f) \le |p| \le a(\langle f\rangle - \min f), \\ \dfrac{a+b}{2}\langle f\rangle & \text{if } |p| \ge a(\langle f\rangle - \min f), \end{cases}$$

$$\overline{H}^{\{3\}}(p) :=$$

$$\begin{cases} \dfrac{a-b}{4}(\max f + \min f) & \text{if } 0 \le |p| \le b(\max f - \langle f\rangle), \\ -\dfrac{1}{4}|p| + \dfrac{a}{4}(\max f + \min f) - \dfrac{b}{4}(\langle f\rangle + \min f) & \text{if } b(\max f - \langle f\rangle) \le |p| \le a(\max f - \langle f\rangle), \\ \dfrac{a-b}{4}(\langle f\rangle + \min f) & \text{if } a(\max f - \langle f\rangle) \le |p| \le b(\langle f\rangle - \min f), \\ \dfrac{1}{4}|p| + \dfrac{a}{4}(\langle f\rangle + \min f) - \dfrac{b}{2}\langle f\rangle & \text{if } b(\langle f\rangle - \min f) \le |p| \le a(\langle f\rangle - \min f), \\ \dfrac{a-b}{2}\langle f\rangle & \text{if } |p| \ge a(\langle f\rangle - \min f), \end{cases}$$

and

$$\overline{H}^{\{1,2,3\}}(p) :=$$

$$\begin{cases} \dfrac{a-b}{4}(\max f - \min f) & \text{if } 0 \le |p| \le b(\max f - \langle f\rangle), \\ -\dfrac{1}{4}|p| + \dfrac{a}{4}(\max f - \min f) - \dfrac{b}{4}(\langle f\rangle - \min f) & \text{if } b(\max f - \langle f\rangle) \le |p| \le a(\max f - \langle f\rangle), \\ \dfrac{a-b}{4}(\langle f\rangle - \min f) & \text{if } a(\max f - \langle f\rangle) \le |p| \le b(\langle f\rangle - \min f), \\ -\dfrac{1}{4}|p| + \dfrac{a}{4}(\langle f\rangle + \min f) & \text{if } b(\langle f\rangle - \min f) \le |p| \le a(\langle f\rangle - \min f), \\ 0 & \text{if } |p| \ge a(\langle f\rangle - \min f). \end{cases}$$

6.4.4 Some front propagation problems

We now consider the following general first-order, level-set equation

$$u_t^\varepsilon + \frac{1}{\varepsilon^\gamma} A\left(\frac{x}{\varepsilon}, \frac{t}{\varepsilon^{2\gamma}}\right)|Du^\varepsilon| = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad u^\varepsilon(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d, \tag{6.4.20}$$

where

$$\begin{cases} A(y,t) := \displaystyle\sum_{i=1}^m a^i(y)\xi^i(t), \\ \xi^i \text{ satisfies (6.4.9) and } a^i \in \mathrm{Lip}(\mathbb{T}^d) \quad \text{for all } i = 1,2,\ldots,m, \quad \text{and} \\ a^1 > \displaystyle\sum_{k=2}^m |a^k|. \end{cases} \tag{6.4.21}$$

Equation (6.4.20) is the level set equation for a hypersurface evolving according to the normal velocity $-\varepsilon^{-\gamma}A(x/\varepsilon, t/\varepsilon^{2\gamma})$.

For $i = 1,2,\ldots,m$, define $H^i(p,x) := a^i(x)|p|$. The convexity and coercivity assumptions in (6.4.7) are satisfied in view of the last line of (6.4.21), and therefore Theorem 6.4.1 can used to study (6.4.20) when $\gamma < 1/6$. In this case, the effective Hamiltonian $\overline{H}$ given by (6.4.2) is positively homogenous in the gradient variable, and, from the formula in (6.4.11), so are each of the $\overline{H}^{\mathbf{j}}$ for $\mathbf{j} \in \mathcal{A}_o^m$. Therefore, each $\overline{H}^{\mathbf{j}}$ has the form

$$\overline{H}^{\mathbf{j}}(p) := \overline{a}^{\mathbf{j}}\left(\frac{p}{|p|}\right)|p| \quad \text{for some } \overline{a}^{\mathbf{j}} : S^{d-1} \to \mathbb{R}.$$

For some independent Brownian motions $\{B^{\mathbf{j}}\}_{\mathbf{j}\in\mathcal{A}_o^m}$, the limiting equation is then of the form

$$d\overline{u} + \sum_{\mathbf{j}\in\mathcal{A}_o^m} \overline{a}^{\mathbf{j}}\left(\frac{D\overline{u}^\varepsilon}{|D\overline{u}^\varepsilon|}\right)|D\overline{u}^\varepsilon| \circ dB^{\mathbf{j}} \quad \text{in } \mathbb{R}^d \times (0,\infty) \quad \text{and} \quad \overline{u}(\cdot,0) = u_0 \quad \text{in } \mathbb{R}^d.$$

In other words, as $\varepsilon \to 0$ and in distribution, the level-set flow corresponding to the normal velocity $-\varepsilon^{\gamma} A(x/\varepsilon, t/\varepsilon^{2\gamma})$ converges in the Hausdorff metric to the level-set flow with the normal velocity $dB(n,t)$, where

$$B(n,t) := \sum_{\mathbf{j}\in\mathcal{A}_o^m} \overline{a}^{\mathbf{j}}(n) B^{\mathbf{j}}(t).$$

REFERENCES

[1] A. D. Alexandroff. Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. *Leningrad State Univ. Annals [Uchenye Zapiski] Math. Ser.*, 6:3–35, 1939.

[2] Scott Armstrong and Pierre Cardaliaguet. Stochastic homogenization of quasilinear Hamilton-Jacobi equations and geometric motions. *J. Eur. Math. Soc. (JEMS)*, 20(4):797–864, 2018.

[3] Scott N. Armstrong, Pierre Cardaliaguet, and Panagiotis E. Souganidis. Error estimates and convergence rates for the stochastic homogenization of Hamilton-Jacobi equations. *J. Amer. Math. Soc.*, 27(2):479–540, 2014.

[4] Scott N. Armstrong and Panagiotis E. Souganidis. Stochastic homogenization of level-set convex Hamilton-Jacobi equations. *Int. Math. Res. Not. IMRN*, (15):3420–3449, 2013.

[5] Scott N. Armstrong, Hung V. Tran, and Yifeng Yu. Stochastic homogenization of a nonconvex Hamilton-Jacobi equation. *Calc. Var. Partial Differential Equations*, 54(2):1507–1524, 2015.

[6] Scott N. Armstrong, Hung V. Tran, and Yifeng Yu. Stochastic homogenization of nonconvex Hamilton-Jacobi equations in one space dimension. *J. Differential Equations*, 261(5):2702–2737, 2016.

[7] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, [1989?]. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.

[8] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[9] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[10] Guy Barles and Espen R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 43(2):540–558, 2005.

[11] Guy Barles and Espen R. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Math. Comp.*, 76(260):1861–1893, 2007.

[12] Guy Barles and Espen Robstad Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN Math. Model. Numer. Anal.*, 36(1):33–54, 2002.

[13] Guy Barles and Panagiotis E. Souganidis. A new approach to front propagation problems: theory and applications. *Arch. Rational Mech. Anal.*, 141(3):237–296, 1998.

[14] Patrick Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[15] R. Bouc and É. Pardoux. Asymptotic analysis of PDEs with wide-band noise disturbances, and expansion of the moments. *Stochastic Anal. Appl.*, 2(4):369–422, 1984.

[16] Kenneth A. Brakke. *The motion of a surface by its mean curvature*, volume 20 of *Mathematical Notes.* Princeton University Press, Princeton, N.J., 1978.

[17] Rainer Buckdahn and Jin Ma. Stochastic viscosity solutions for nonlinear stochastic partial differential equations. I. *Stochastic Process. Appl.*, 93(2):181–204, 2001.

[18] Rainer Buckdahn and Jin Ma. Stochastic viscosity solutions for nonlinear stochastic partial differential equations. II. *Stochastic Process. Appl.*, 93(2):205–228, 2001.

[19] Luis A. Caffarelli and Panagiotis E. Souganidis. A rate of convergence for monotone finite difference approximations to fully nonlinear, uniformly elliptic PDEs. *Comm. Pure Appl. Math.*, 61(1):1–17, 2008.

[20] I. Capuzzo-Dolcetta and H. Ishii. On the rate of convergence in homogenization of Hamilton-Jacobi equations. *Indiana Univ. Math. J.*, 50(3):1113–1129, 2001.

[21] Michael Caruana, Peter K. Friz, and Harald Oberhauser. A (rough) pathwise approach to a class of non-linear stochastic partial differential equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(1):27–46, 2011.

[22] R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.*, 100(1):32–74, 1928.

[23] M. G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comp.*, 43(167):1–19, 1984.

[24] Michael G. Crandall and Hitoshi Ishii. The maximum principle for semicontinuous functions. *Differential Integral Equations*, 3(6):1001–1014, 1990.

[25] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[26] Michael G. Crandall and Pierre-Louis Lions. Hamilton-Jacobi equations in infinite dimensions. I. Uniqueness of viscosity solutions. *J. Funct. Anal.*, 62(3):379–396, 1985.

[27] Michael G. Crandall and Pierre-Louis Lions. Hamilton-Jacobi equations in infinite dimensions. II. Existence of viscosity solutions. *J. Funct. Anal.*, 65(3):368–405, 1986.

[28] Lawrence C. Evans. Periodic homogenisation of certain fully nonlinear partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(3-4):245–265, 1992.

[29] William M. Feldman and Panagiotis E. Souganidis. Homogenization and non-homogenization of certain non-convex Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 108(5):751–782, 2017.

[30] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1993.

[31] Peter K. Friz, Paul Gassiat, Pierre-Louis Lions, and Panagiotis E. Souganidis. Eikonal equations and pathwise solutions to fully non-linear SPDEs. *Stoch. Partial Differ. Equ. Anal. Comput.*, 5(2):256–277, 2017.

[32] Peter K. Friz and Martin Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.

[33] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.

[34] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.*, 216(1):86–140, 2004.

[35] R. Z. Hasminskiĭ. A limit theorem for solutions of differential equations with a random right hand part. *Teor. Verojatnost. i Primenen*, 11:444–462, 1966.

[36] Hitoshi Ishii. Hamilton-Jacobi equations with discontinuous Hamiltonians on arbitrary open sets. *Bull. Fac. Sci. Engrg. Chuo Univ.*, 28:33–77, 1985.

[37] Hitoshi Ishii. Perron's method for Hamilton-Jacobi equations. *Duke Math. J.*, 55(2):369–384, 1987.

[38] Hitoshi Ishii. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs. *Comm. Pure Appl. Math.*, 42(1):15–45, 1989.

[39] E. R. Jakobsen. On error bounds for approximation schemes for non-convex degenerate elliptic equations. *BIT*, 44(2):269–285, 2004.

[40] Espen R. Jakobsen. On error bounds for monotone approximation schemes for multi-dimensional Isaacs equations. *Asymptot. Anal.*, 49(3-4):249–273, 2006.

[41] Robert Jensen. The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. *Arch. Rational Mech. Anal.*, 101(1):1–27, 1988.

[42] N. V. Krylov. *Nonlinear elliptic and parabolic equations of the second order*, volume 7 of *Mathematics and its Applications (Soviet Series)*. D. Reidel Publishing Co., Dordrecht, 1987. Translated from the Russian by P. L. Buzytsky [P. L. Buzytskiĭ].

[43] N. V. Krylov. On the rate of convergence of finite-difference approximations for elliptic Isaacs equations in smooth domains. *Comm. Partial Differential Equations*, 40(8):1393–1407, 2015.

[44] N. V. Krylov and B. L. Rozovskiĭ. Stochastic evolution equations. In *Current problems in mathematics, Vol. 14 (Russian)*, pages 71–147, 256. Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979.

[45] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[46] Hung Ju Kuo and Neil S. Trudinger. Discrete methods for fully nonlinear elliptic equations. *SIAM J. Numer. Anal.*, 29(1):123–135, 1992.

[47] Hung-Ju Kuo and Neil S. Trudinger. Positive difference operators on general meshes. *Duke Math. J.*, 83(2):415–433, 1996.

[48] Harold J. Kushner and Hai Huang. Limits for parabolic partial differential equations with wide band stochastic coefficients and an application to filtering theory. *Stochastics*, 14(2):115–148, 1985.

[49] P.-L. Lions and B. Perthame. Remarks on Hamilton-Jacobi equations with measurable time-dependent Hamiltonians. *Nonlinear Anal.*, 11(5):613–621, 1987.

[50] P.-L. Lions and J.-C. Rochet. Hopf formula and multitime Hamilton-Jacobi equations. *Proc. Amer. Math. Soc.*, 96(1):79–84, 1986.

[51] Pierre-Louis Lions. *Generalized solutions of Hamilton-Jacobi equations*, volume 69 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass.-London, 1982.

[52] Pierre-Louis Lions, G. C. Papanicolaou, and S. R. Srinivasa Varadhan. Homogenization of hamilton-jacobi equations. Unpublished manuscript.

[53] Pierre-Louis Lions and Panagiotis E. Souganidis. Fully nonlinear first- and second-order stochastic partial differential equations. to appear.

[54] Pierre-Louis Lions and Panagiotis E. Souganidis. Fully nonlinear stochastic partial differential equations. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(9):1085–1092, 1998.

[55] Pierre-Louis Lions and Panagiotis E. Souganidis. Fully nonlinear stochastic partial differential equations: non-smooth equations and applications. *C. R. Acad. Sci. Paris Sér. I Math.*, 327(8):735–741, 1998.

[56] Pierre-Louis Lions and Panagiotis E. Souganidis. Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations. *C. R. Acad. Sci. Paris Sér. I Math.*, 331(10):783–790, 2000.

[57] Pierre-Louis Lions and Panagiotis E. Souganidis. Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media. *Comm. Partial Differential Equations*, 30(1-3):335–375, 2005.

[58] Songting Luo, Hung V. Tran, and Yifeng Yu. Some inverse problems in periodic homogenization of Hamilton-Jacobi equations. *Arch. Ration. Mech. Anal.*, 221(3):1585–1617, 2016.

[59] Terry Lyons and Zhongmin Qian. Flow equations on spaces of rough paths. *J. Funct. Anal.*, 149(1):135–159, 1997.

[60] Terry J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.

[61] Hiroyoshi Mitake, Hung V. Tran, and Yifeng Yu. Rate of convergence in periodic homogenization of Hamilton-Jacobi equations: the convex setting. arXiv:1801.00391 [math.AP].

[62] Diana Nunziante. Uniqueness of viscosity solutions of fully nonlinear second order parabolic equations with discontinuous time-dependence. *Differential Integral Equations*, 3(1):77–91, 1990.

[63] Diana Nunziante. Existence and uniqueness of unbounded viscosity solutions of parabolic equations with discontinuous time-dependence. *Nonlinear Anal.*, 18(11):1033–1062, 1992.

[64] Stanley Osher and James A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79(1):12–49, 1988.

[65] G. C. Papanicolaou and W. Kohler. Asymptotic theory of mixing stochastic ordinary differential equations. *Comm. Pure Appl. Math.*, 27:641–668, 1974.

[66] G. C. Papanicolaou and S. R. S. Varadhan. A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations. *Comm. Pure Appl. Math.*, 26:497–524, 1973.

[67] E. Pardoux. Stochastic partial differential equations and filtering of diffusion processes. *Stochastics*, 3(2):127–167, 1979.

[68] É. Pardoux. Equations of nonlinear filtering and application to stochastic control with partial observation. In *Nonlinear filtering and stochastic control (Cortona, 1981)*, volume 972 of *Lecture Notes in Math.*, pages 208–248. Springer, Berlin, 1982.

[69] É. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lect. Notes Control Inf. Sci.*, pages 200–217. Springer, Berlin, 1992.

[70] Étienne Pardoux and Andrey Piatnitski. Homogenization of a singular random one-dimensional PDE with time-varying coefficients. *Ann. Probab.*, 40(3):1316–1356, 2012.

[71] Oskar Perron. Eine neue Behandlung der ersten Randwertaufgabe für $\Delta u = 0$. *Math. Z.*, 18(1):42–54, 1923.

[72] Fraydoun Rezakhanlou and James E. Tarver. Homogenization for stochastic Hamilton-Jacobi equations. *Arch. Ration. Mech. Anal.*, 151(4):277–309, 2000.

[73] Benjamin Seeger. Approximation schemes for viscosity solutions of fully nonlinear stochastic partial differential equations. arXiv:1802.04740 [math.AP].

[74] Benjamin Seeger. Homogenization of pathwise Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 110:1–31, 2018.

[75] Benjamin Seeger. Perron's method for pathwise viscosity solutions. *Comm. Partial Differential Equations*, 43(6):998–1018, 2018.

[76] Panagiotis E. Souganidis. Fully nonlinear first- and second-order stochastic partial differential equations. to appear.

[77] Panagiotis E. Souganidis. Approximation schemes for viscosity solutions of Hamilton-Jacobi equations. *J. Differential Equations*, 59(1):1–43, 1985.

[78] Panagiotis E. Souganidis. Max-min representations and product formulas for the viscosity solutions of Hamilton-Jacobi equations with applications to differential games. *Nonlinear Anal.*, 9(3):217–257, 1985.

[79] Panagiotis E. Souganidis. Stochastic homogenization of Hamilton-Jacobi equations and some applications. *Asymptot. Anal.*, 20(1):1–11, 1999.

[80] Kaising Tso. On an Aleksandrov-Bakel'man type maximum principle for second-order parabolic equations. *Comm. Partial Differential Equations*, 10(5):543–553, 1985.

[81] Son N. T. Tu. Rate of convergence for periodic homogenization of convex Hamilton-Jacobi equations in one dimension. arXiv:1808.06129 [math.AP].

[82] Hisao Watanabe. Averaging and fluctuations for parabolic equations with rapidly oscillating random coefficients. *Probab. Theory Related Fields*, 77(3):359–378, 1988.

[83] Eugene Wong and Moshe Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.*, 36:1560–1564, 1965.

[84] Eugene Wong and Moshe Zakai. On the relation between ordinary and stochastic differential equations. *Internat. J. Engrg. Sci.*, 3:213–229, 1965.

[85] Bruno Ziliotto. Stochastic homogenization of nonconvex Hamilton-Jacobi equations: a counterexample. *Comm. Pure Appl. Math.*, 70(9):1798–1809, 2017.