Lecture notes for Math 272, Winter 2021

Lenya Ryzhik

February 4, 2021
These notes will be continuously updated during the course.

The plan for this class is to cover the following topics:

I. Basic theory of Hamilton-Jacobi equations:
   1. Existence and long time behavior for the viscous Hamilton-Jacobi equations.
   2. Basic viscosity solutions theory for the first order Hamilton-Jacobi equations.
   3. The Lions-Papanicolaou-Varadhan theorem and applications to periodic homogenization.
   4. Long time behavior for the Lax-Oleinik semigroup, and very rudimentary aspects of the Fathi theory.

II. Hamilton-Jacobi equations with a constraint and applications to the biological modeling.

III. An introduction to mean-field games, based on the lecture notes by P. Cardaliaguet and A. Porretta.

Part I of these lecture notes is a draft of a chapter in a book in preparation with Sasha Kiselev and Jean-Michel Roquejoffre. The preliminary version of the draft of this chapter was written mostly by Jean-Michel. All mistakes are, obviously, mine.

The draft will be updated as we go, potentially with major re-writes back and forth. Because of that, I plan to update the lecture notes after each lecture, to reflect what was actually presented in class, and not upload the full draft of Chapter 2 of these notes from the start.

In addition, I include Chapter 1 (which is actual Chapter 2 of the book draft) in the lecture notes because some of the results of that chapter will be used in class, and it is easy to refer to them in this way. However, this content is included solely for your convenience, the class will not cover that chapter and will start with Chapter 2 of these notes (which is Chapter 3 of the book draft).

The texts of Chapter 1 and 2 have not been finalized so all comments are extremely welcome!
# Contents

1 **Diffusion equations**  
1.1 Introduction to the chapter ........................................... 5  
1.2 A probabilistic introduction to the evolution equations ............ 7  
1.3 The maximum principle interlude: the basic statements .......... 13  
1.4 The forced linear heat equation .................................... 16  
1.4.1 The Duhamel formula ........................................... 17  
1.4.2 Regularity gain: from a bounded $g(t, x)$ to a Hölder $J(t, x)$ 20  
1.4.3 Regularity gain: from Hölder $g(t, x)$ to differentiable $J(t, x)$ 23  
1.4.4 Regularity gain: from a Hölder $g$ to Hölder derivatives of $J$ 26  
1.5 Regularity for the nonlinear heat equations ........................ 32  
1.5.1 Existence and regularity for a semi-linear diffusion equation 32  
1.5.2 The regularity of the solutions to a quasi-linear heat equation 38  
1.6 A survival kit in the jungle of regularity .......................... 47  
1.7 The principal eigenvalue for elliptic operators and the Krein-Rutman theorem 56  
1.7.1 The periodic principal eigenvalue ................................ 56  
1.7.2 The Krein-Rutman theorem: the periodic parabolic case ...... 58  
1.7.3 Back to the principal periodic elliptic eigenvalue ................ 60  
1.7.4 The principal eigenvalue and the comparison principle ........ 68  
1.8 Reaction-diffusion waves ......................................... 71  
1.8.1 The long time behavior for the Allen-Cahn equation ........ 71  
1.8.2 Spreading in unbalanced Allen-Cahn equations, and related models 82  
1.8.3 When the medium is inhomogeneous: pulsating waves ...... 83  

2 **Inviscid Hamilton-Jacobi equations**  
2.1 Introduction to the chapter ......................................... 93  
2.2 An informal derivation of the Hamilton-Jacobi equations ....... 95  
2.3 The simple world of viscous Hamilton-Jacobi equations .......... 98  
2.3.1 The wave solutions ........................................... 99  
2.3.2 Long time convergence and uniqueness of the wave solutions 107  
2.4 A glimpse of the classical solutions to the Hamilton-Jacobi equations 110  
2.4.1 Smooth solutions and their limitations ........................ 110  
2.4.2 An example of classical global solutions ...................... 112  
2.5 Viscosity solutions ............................................... 116  
2.5.1 The definition and the basic properties of the viscosity solutions 117
2.5.2 Uniqueness of the viscosity solutions .................................................. 123
2.5.3 Finite speed of propagation ................................................................. 125
2.6 Construction of solutions ........................................................................ 127
  2.6.1 Existence of waves, and the Lions-Papanicolaou-Varadhan theorem ...... 128
  2.6.2 Existence of the solution to the Cauchy problem ............................... 138
2.7 When the Hamiltonian is strictly convex: the Lagrangian theory .............. 143
  2.7.1 The Lax-Oleinik formula and viscosity solutions .............................. 144
  2.7.2 Semi-concavity and $C^{1,1}$ regularity ............................................. 152
2.8 Large time behavior in a particular case ............................................... 157
  2.8.1 Counting the waves ............................................................................ 158
  2.8.2 The large time behavior ..................................................................... 164
2.9 Convergence of the Lax-Oleinik semigroup ............................................ 167
Chapter 1

Diffusion equations

1.1 Introduction to the chapter

Parabolic equations of the form

$$\frac{\partial u}{\partial t} - \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{j=1}^{n} b_j(x) \frac{\partial u}{\partial x_j} = f(x,u,\nabla u),$$

are ubiquitous in mathematics and various applications in physics, biology, economics and other fields. While there are many textbooks on the subject, ranging from the most elementary to extremely advanced, most of them concentrate on the highly non-trivial questions of the existence and regularity of the solutions. We have chosen instead to focus on some striking qualitative properties of the solutions that, nevertheless, can be proved with almost no background in analysis, using only the very basic regularity results. The unifying link in this chapter will be the parabolic maximum principle and the Harnack inequality. Together with the parabolic regularity, they will be responsible for the seemingly very special behavior that we will observe in the solutions of these equations.

The chapter starts with an informal probabilistic introduction. While we do not try to motivate the basic diffusion equations by models in the applied sciences here, an interested reader would have no difficulty finding the connections between such equations and models in physics, biology, chemistry and ecology in many basic textbooks. On the other hand, the parabolic equations have a deep connection with probability. Indeed, some of the most famous results in the parabolic regularity theory were proved by probabilistic tools. It is, therefore, quite natural to start the chapter by explaining how the basic linear models arise, in a very simple manner, from limits of a random walk. We reassure the reader that the motivation from the physical or life sciences will not be absent from this book, as some of the later chapters will precisely be motivated by problems in fluid mechanics or biology. We also keep the probabilistic considerations to an elementary level, without any use of stochastic analysis.

The probabilistic section is followed by a brief interlude on the maximum principle. There is nothing original in the exposition, and we do not even present the proofs, as they can be found in many textbooks on PDE. We simply recall the statements that we will need.

We then proceed to the section on the existence and regularity theory for the nonlinear heat equations: the reaction-diffusion equations and viscous Hamilton-Jacobi equations. They
arise in many models in physical and biological sciences, and our "true" interest is in the qualitative behavior of their solutions, as these reflect the corresponding natural phenomena. However, an unfortunate feature of the nonlinear partial differential equations is that, before talking knowledgeably about their solutions or their behavior, one first has to prove that they exist. This will, as a matter of fact, be a non-trivial problem in the last two chapters of this book, where we look at the fluid mechanics models, for which the existence of the solutions can be quite subtle. As the reaction-diffusion equations that we have in mind here and in Chapter ?? both belong to a very well studied class and are much simpler, it would not be inconceivable to brush their existence theory under the rug, invoking other books. This would not be completely right, for several reasons. The first is that we do not want to give the impression that the theory is inaccessible: it is quite simple and can be explained very easily. The second reason is that we wish to explain both the power and the limitation of the parabolic regularity theory, so that the difficulty of the existence issues for the fluid mechanics models in the latter chapters would be clearer to the reader. The third reason is more practical: even for the qualitative properties that we aim for, we still need to estimate derivatives. So, it is better to say how this is done.

The next section contains a rather informal guide to the regularity theory for the parabolic equations with inhomogeneous coefficients. We state the results we will need later, and outline the details of some of the main ideas needed for the proofs without presenting them in full – they can be found in the classical texts we mention below. We hope that by this point the reader will be able to study the proofs in these more advanced textbooks without losing sight of the main ideas. This section also contains the Harnack inequality. What is slightly different here is the statement of a (non-optimal) version of the Harnack inequality that will be of an immediate use to us in the first main application of this chapter, the convergence to the steady solutions in the one-dimensional Allen-Cahn equations on the line. The reason we have chosen this example is that it really depends on nothing else than the maximum principle and the Harnack inequality, illustrating how far reaching this property is. It is also a perfect example of how a technical information, such as bounds on the derivatives, has a qualitative implication – the long time behavior of the solutions.

The next section concerns the principal eigenvalue of the second order elliptic operators, a well-treated subject in its own right. We state the Krein-Rutman theorem and, just to show the reader that we are not using any machinery heavier than the results we want to prove, we provide a proof in the context of the second order elliptic and parabolic operators. It shares many features with the convergence proof of the next section, without its sometimes technically involved details. We hope the reader will realize the ubiquitous character of the ideas presented.

We end the chapter with the study of reaction-diffusion fronts. While it is, in its own right, a huge subject that is still advancing at the time of the writing of this chapter, we have decided that talking about them was a good way to follow the main pledge of this book: show the reader results that are interesting and representative of the theory, while not being the most advanced or up-to-date. With nothing else than the tools displayed in this chapter, we will see that we can say a lot about the large time organization of this class of models, a striking example being the convergence to pulsating waves: periodicity in space will generate a sort of time periodicity for the solutions.

This chapter is quite long so we ask the reader to be prepared to persevere through the
more technical places, but we feel that it is worth showing how far one may go with the sole aid of the maximum principle and a few estimates. We hope that in the end the reader will find the effort rewarding.

A note on notation. We will follow throughout the book the summation convention: the repeated indices are always summed over, unless specified otherwise. In particular, we will usually write equations such as (1.1.1) in the form

\[
\frac{\partial u}{\partial t} - a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(x) \frac{\partial u}{\partial x_j} = f(x, u, \nabla u),
\]

(1.1.2)
or

\[
\frac{\partial u}{\partial t} - a_{ij}(x) \partial_{x_i} \partial_{x_j} u + b_j(x) \partial_{x_j} u = f(x, u, \nabla u).
\]

(1.1.3)

We hope the reader will get accustomed to this convention sufficiently fast so that it causes no confusion or inconvenience.

1.2 A probabilistic introduction to the evolution equations

Let us explain informally how the linear equations of the form (1.1.2), with \( f \equiv 0 \) arise from random walks, in a very simple way. One should emphasize that many of the qualitative properties of the solutions to the parabolic and integral equations, such as the maximum principle and regularity, on a very informal level, are an "obvious" consequence of the microscopic random walk model. For simplicity, we will mostly consider the one-dimensional case, the reader can, and should, generalize this approach to higher dimensions – this is quite straightforward.

Discrete equations and random walks

The starting point in our derivation of the evolution equations is a discrete time Markov jump process \( X_{n\tau} \), with a time step \( \tau > 0 \), defined on a lattice with mesh size \( h \):

\[
h \mathbb{Z} = \{0, \pm h, \pm 2h, \ldots\}.
\]

The particle position evolves as follows: if the particle is located at a position \( x \in h \mathbb{Z} \) at the time \( t = n\tau \) then at the time \( t = (n + 1)\tau \) it jumps to a random position \( y \in h \mathbb{Z} \), with the transition probability

\[
P(X_{(n+1)\tau} = y \mid X_{n\tau} = x) = k(x - y), \quad x, y \in h \mathbb{Z}.
\]

(1.2.1)

Here, \( k(x) \) is a prescribed non-negative kernel such that

\[
\sum_{y \in h \mathbb{Z}} k(y) = 1.
\]

(1.2.2)

The classical symmetric random walk with a spatial step \( h \) and a time step \( \tau \) corresponds to the choice \( k(\pm h) = 1/2 \), and \( k(y) = 0 \) otherwise – the particle may only jump to the nearest neighbor on the left and on the right, with equal probabilities.
In order to connect this process to an evolution equation, let us take a function $f : h\mathbb{Z} \rightarrow \mathbb{R}$, defined on our lattice, and introduce

$$u(t, x) = \mathbb{E}(f(X_t(x))).$$  \hspace{1cm} (1.2.3)

Here, $X_t(x)$, $t \in \tau \mathbb{N}$, is the above Markov process starting at a position $X_0(x) = x \in h\mathbb{Z}$ at the time $t = 0$. If $f \geq 0$ then one may think of $u(t, x)$ as the expected value of a “prize” to be collected at the time $t$ at a (random) location of $X_t(x)$ given that the process starts at the point $x$ at the time $t = 0$. An important special case is when $f$ is the characteristic function of a set $A$. Then, $u(t, x)$ is the probability that the jump process $X_t(x)$ that starts at the position $X_0 = x$ is inside the set $A$ at the time $t$.

As the process $X_t(x)$ is Markov, the function $u(t, x)$ satisfies the following relation

$$u(t + \tau, x) = \mathbb{E}(f(X_{t+\tau}(x))) = \sum_{y \in h\mathbb{Z}} P(X_\tau = y | X_0 = x) \mathbb{E}(f(X_t(y))) = \sum_{y \in h\mathbb{Z}} k(x - y) u(t, y).$$  \hspace{1cm} (1.2.4)

This is because after the initial step when the particle jumps at the time $\tau$ from the starting position $x$ to a random position $y$, the process “starts anew”, and runs for time $t$ between the times $\tau$ and $t + \tau$. Equation (1.2.4) can be re-written, using (1.2.2) as

$$u(t + \tau, x) - u(t, x) = \sum_{y \in h\mathbb{Z}} k(x - y) [u(t, y) - u(t, x)].$$  \hspace{1cm} (1.2.5)

The key point of this section is that the discrete equation (1.2.5) leads to various interesting continuum limits as $h \downarrow 0$ and $\tau \downarrow 0$, depending on the choice of the transition kernel $k(y)$, and on the relative size of the spatial mesh size $h$ and the time step $\tau$. In other words, depending on the microscopic model – the particular properties of the random walk – we will end up with different macroscopic continuous models.

The heat equation and random walks

Before showing how a general parabolic equation with non-constant coefficients can be obtained via a limiting procedure from a random walk on a lattice, let us show how this can be done for the heat equation

$$\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2},$$  \hspace{1cm} (1.2.6)

with a constant diffusivity constant $a > 0$. We will assume that the transition probability kernel has the form

$$k(x) = \phi \left( \frac{x}{h} \right), \quad x \in h\mathbb{Z},$$  \hspace{1cm} (1.2.7)

with a non-negative function $\phi(m) \geq 0$ defined on $\mathbb{Z}$, such that

$$\sum_m \phi(m) = 1.$$  \hspace{1cm} (1.2.8)

This form of $k(x)$ allows us to re-write (1.2.5) as

$$u(t + \tau, x) - u(t, x) = \sum_{y \in h\mathbb{Z}} \phi \left( \frac{x - y}{h} \right) [u(t, y) - u(t, x)],$$  \hspace{1cm} (1.2.9)
or, equivalently,

\[ u(t + \tau, x) - u(t, x) = \sum_{m \in \mathbb{Z}} \phi(m)[u(t, x - mh) - u(t, x)]. \] (1.2.10)

In order to arrive to the heat equation in the limit, we will make the assumption that jumps are symmetric on average:

\[ \sum_{m \in \mathbb{Z}} m \phi(m) = 0. \] (1.2.11)

Then, expanding the right side of (1.2.10) in \( h \) and the left side in \( \tau \), we obtain

\[ \tau \frac{\partial u(t, x)}{\partial t} = \frac{\alpha h^2}{2} \frac{\partial^2 u}{\partial x^2}(t, x) + \text{lower order terms}, \] (1.2.12)

with

\[ \alpha = \sum_m |m|^2 \phi(m). \] (1.2.13)

To balance the left and the right sides of (1.2.12), we need to take the time step \( \tau = h^2 \) – note that the scaling \( \tau = O(h^2) \) is essentially forced on us if we want to balance the two sides of this equation. Then, in the limit \( \tau = h^2 \downarrow 0 \), we obtain the heat equation

\[ \frac{\partial u(t, x)}{\partial t} = \frac{\alpha}{2} \frac{\partial^2 u(t, x)}{\partial x^2}. \] (1.2.14)

The diffusion coefficient \( \alpha \) given by (1.2.13) is the second moment of the jump size – in other words, it measures the “overall jumpiness” of the particles. This is a very simple example of how the microscopic information, the kernel \( \phi(m) \), translates into a macroscopic quantity – the overall diffusion coefficient \( \alpha \) in the macroscopic equation (1.2.14).

**Exercise 1.2.1** Show that if (1.2.11) is violated and

\[ b = \sum_{m \in \mathbb{Z}} m \phi(m) \neq 0, \] (1.2.15)

then one needs to take \( \tau = h \), and the (formal limit) is the advection equation

\[ \frac{\partial u(t, x)}{\partial t} + b \frac{\partial u(t, x)}{\partial x} = 0, \] (1.2.16)

without any diffusion.

**Exercise 1.2.2** A reader familiar with the basic probability theory should relate the limit in (1.2.16) to the law of large numbers and explain the relation \( \tau = h \) in these terms. How can (1.2.14) and the relation \( \tau = h^2 \) between the temporal and spatial steps be explained in terms of the central limit theorem?
Parabolic equations with variable coefficients and drifts and random walks

In order to connect a linear parabolic equation with inhomogeneous coefficients, such as (1.1.2) with the right side $f \equiv 0$:

$$\frac{\partial u}{\partial t} - a(x) \frac{\partial^2 u}{\partial x^2} + b(x) \frac{\partial u}{\partial x} = 0,$$

(1.2.17)

to a continuum limit of random walks, we consider a slight modification of the microscopic dynamics that led to the heat equation in the macroscopic limit. We go back to (1.2.4):

$$u(t + \tau, x) = E(f(X_{t+\tau}(x))) = \sum_{y \in h\mathbb{Z}} P(X_{\tau} = y | X_{0} = x) E(f(X_{t}(y))) = \sum_{y \in h\mathbb{Z}} k(x, y) u(t, y).$$

(1.2.18)

Here, $k(x, y)$ is the probability to jump to the position $y$ from a position $x$. Note that we no longer assume that the law of the jump process is spatially homogeneous: the transition probabilities depend not only on the difference $x - y$ but both on $x$ and $y$. However, we will assume that $k(x, y)$ is "locally homogeneous". This condition translates into considering

$$k(x, y; h) = \phi \left( x, \frac{x - y}{h}, h \right).$$

(1.2.19)

The "slow" spatial dependence of the transition probability density is encoded in the dependence of the function $\phi(x, z; h)$ on the "macroscopic" variable $x$, while its "fast" spatial variations are described by the dependence of $\phi(x, z; h)$ on the variable $z$.

**Exercise 1.2.3** Make sure you can interpret this point. Think of "freezing" the variable $x$ and only varying the $z$-variable.

We will soon see why we introduce the additional dependence of the transition density on the mesh size $h$ – this will lead to a non-trivial first order term in the parabolic equation we will obtain in the limit. We assume that the function $\phi(x, m; h)$, with $x \in \mathbb{R}$, $m \in \mathbb{Z}$ and $h \in (0, 1)$, satisfies

$$\sum_{m \in \mathbb{Z}} \phi(x, m; h) = 1 \text{ for all } x \in \mathbb{R} \text{ and } h \in (0, 1),$$

(1.2.20)

which leads to the analog of the normalization (1.2.2):

$$\sum_{y \in h\mathbb{Z}} k(x, y) = 1 \text{ for all } x \in h\mathbb{Z}. \quad (1.2.21)$$

This allows us to re-write (1.2.18) in the familiar form

$$u(t + \tau, x) - u(t, x) = \sum_{y \in h\mathbb{Z}} \phi \left( x, \frac{x - y}{h}, h \right) [u(t, y) - u(t, x)],$$

(1.2.22)

or, equivalently,

$$u(t + \tau, x) - u(t, x) = \sum_{m \in \mathbb{Z}} \phi(x, m; h)[u(t, x - mh) - u(t, x)],$$

(1.2.23)
We will make the assumption that the average asymmetry of the jumps is of the size \( h \). In other words, we suppose that
\[
\sum_{m \in \mathbb{Z}} m \phi(x, m; h) = b(x) h + O(h^2),
\] (1.2.24)
that is,
\[
\sum_{m \in \mathbb{Z}} m \phi(x, m; 0) = 0 \text{ for all } x \in \mathbb{R},
\]
and
\[
b(x) = \sum_{m \in \mathbb{Z}} m \frac{\partial \phi(x, m; h = 0)}{\partial h}
\] (1.2.25)
is a given smooth function. The last assumption we will make is that the time step is \( \tau = h^2 \), as before. Expanding the left and the right side of (1.2.23) in \( h \) now leads to the parabolic equation
\[
\frac{\partial u}{\partial t} = -b(x) \frac{\partial u(t, x)}{\partial x} + a(x) \frac{\partial^2 u(t, x)}{\partial x^2},
\] (1.2.26)
with
\[
a(x) = \frac{1}{2} \sum_{m \in \mathbb{Z}} |m|^2 \phi(x, m; h = 0).
\] (1.2.27)
This is a parabolic equation of the form (1.1.2) in one dimension. We automatically satisfy the condition \( a(x) > 0 \) (known as the ellipticity condition) unless \( \phi(x, m; h = 0) = 0 \) for all \( m \in \mathbb{Z} \setminus \{0\} \). That is, \( a(x) = 0 \) only at the positions where the particles are completely stuck and can not jump at all. Note that the asymmetry in (1.2.24), that is, the mismatch in the typical jump sizes to the left and right, leads to the first order term in the limit equation (1.2.26) – because of that the first-order coefficient \( b(x) \) is known as the drift, while the second-order coefficient \( a(x) \) (known as the diffusivity) measures "the overall jumpiness" of the particles, as seen from (1.2.27).

**Exercise 1.2.4** Relate the above considerations to the method of characteristics for the first order linear equation
\[
\frac{\partial u}{\partial t} + b(x) \frac{\partial u}{\partial x} = 0.
\]
How does it arise from similar considerations?

**Exercise 1.2.5** It is straightforward to generalize this construction to higher dimensions leading to general parabolic equations of the form (1.1.2). Verify that the diffusion matrices \( a_{ij}(x) \) in (1.1.2) that arise in this fashion, will always be nonnegative, in the sense that for any \( \xi \in \mathbb{R}^n \) and all \( x \), we have (once again, as the repeated indices are summed over):
\[
a_{ij}(x) \xi_i \xi_j \geq 0.
\] (1.2.28)
This is very close to the lower bound in the ellipticity condition on the matrix \( a_{ij}(x) \) which says that there exists a constant \( c > 0 \) so that for any \( \xi \in \mathbb{R}^n \) and \( x \in \mathbb{R}^n \) we have
\[
c|\xi|^2 \leq a_{ij}(x) \xi_i \xi_j \leq c^{-1}|\xi|^2.
\] (1.2.29)
We see that the ellipticity condition appears very naturally in the probabilistic setting.
Summarizing, we see that parabolic equations of the form (1.1.2) arise as limits of random walks that make jumps of the size $O(h)$, with a time step $\tau = O(h^2)$. Thus, the overall number of jumps by a time $t = O(1)$ is very large, and each individual jump is very small. The drift vector $b_j(x)$ appears from the local non-zero mean of the jump direction and size, and the diffusivity matrix $a_{ij}(x)$ measures the typical jump size. In addition, the diffusivity matrix is nonegative-definite: condition (1.2.28) is satisfied.

**Parabolic equations and branching random walks**

Let us now explain how random walks can lead to parabolic equations with a zero-order term:

$$\frac{\partial u}{\partial t} - a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(x) \frac{\partial u}{\partial x_j} + c(x)u = 0. \tag{1.2.30}$$

This will help us understand qualitatively the role of the coefficient $c(x)$. Once again, we will consider the one-dimensional case for simplicity, and will only give the details for the case

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} + c(x)u = 0, \tag{1.2.31}$$

as the non-constant diffusion matrix $a_{ij}(x)$ and drift $b_j(x)$ can be treated exactly as in the case $c = 0$.

In order to incorporate the zero order term we need to allow the particles not only jump but also branch – this is the reason why the zero-order term will appear in (1.2.30). As before, our particles make jumps on the lattice $h\mathbb{Z}$, at the discrete times $t \in \tau \mathbb{N}$. We start at $t = 0$ with one particle at a position $x \in h\mathbb{Z}$. Let us assume that at the time $t = n\tau$ we have a collection of $N_t$ particles $X_1(t, x), \ldots, X_{N_t}(t, x)$ (the number $N_t$ is random, as will immediately see). At the time $t$, each particle $X_m(t, x)$ behaves independently from the other particles. With the probability

$$p_0 = 1 - |c(X_m(t))|\tau,$$

it simply jumps to a new location $y \in h\mathbb{Z}$, chosen with the transition probability $k(X_m(t) - y)$, as in the process with no branching. If the particle at $X_m(t, x)$ does not jump – this happens with the probability $p_1 = 1 - p_0$, there are two possibilities. If $c(X_m(t)) < 0$, then it is replaced by two particles at the same location $X_m(t, x)$ that remain at this position until the time $t + \tau$. If $c(X_m(t)) > 0$ and the particle does not jump, then it is removed. This process is repeated independently for all particles $X_1(t, x), \ldots, X_{N_t}(t, x)$, giving a new collection of particles at the locations $X_1(t + \tau, x), \ldots, X_{N_t}(t + \tau, x)$ at the time $t + \tau$. If $c(x) > 0$ at some positions, then the process can terminate when there are no particles left. If $c(x) \leq 0$ everywhere, then the process continues forever.

To connect this particle system to an evolution equation, given a function $f$, we define, for $t \in \tau \mathbb{N}$, and $x \in h\mathbb{Z},$

$$u(t, x) = \mathbb{E}[f(X_1(t, x)) + f(X_2(t, x)) + \cdots + f(X_{N_t}(t, x))].$$

The convention is that $f = 0$ inside the expectation if there are no particles left. This is similar to what we have done for particles with no branching. If $f$ is the characteristic function of a set $A$, then $u(t, x)$ is the expected number of particles inside $A$ at the time $t > 0$. 

12
In order to get an evolution equation for \( u(t,x) \), we look at the initial time when we have just one particle at the position \( x \): if \( c(x) \leq 0 \), then this particle either jumps or branches, leading to the balance

\[
u(t + \tau, x) = (1 + c(x)\tau) \sum_{y \in h\mathbb{Z}} k(x - y)u(t, y) - 2c(x)\tau u(t, x), \quad \text{if } c(x) \leq 0, \tag{1.2.32}
\]

which is the analog of (1.2.4). If \( c(x) > 0 \) the particle either jumps or is removed, leading to

\[
u(t + \tau, x) = (1 - |c(x)|\tau) \sum_{y \in h\mathbb{Z}} k(x - y)u(t, y). \tag{1.2.33}
\]

In both cases, we can re-write the balances similarly to (1.2.5):

\[
u(t + \tau, x) - u(t, x) = (1 - |c(x)|\tau) \sum_{y \in h\mathbb{Z}} k(x - y)(u(t, y) - u(t, x)) - c(x)\tau u(t, x). \tag{1.2.34}
\]

We may now take the transition probability kernel of the familiar form

\[
k(x) = \phi\left(\frac{x}{h}\right),
\]

with a function \( \phi(m) \) as in (1.2.7)-(1.2.8). Taking \( \tau = h^2 \) leads, as in (1.2.12), to the diffusion equation but now with a zero-order term:

\[
\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} - c(x)u. \tag{1.2.35}
\]

Thus, the zero-order coefficient \( c(x) \) can be interpreted as the branching (or killing, depending on the sign of \( c(x) \)) rate of the random walk. The parabolic maximum principle for \( c(x) \geq 0 \) that we will discuss in the next section simply means, on this informal level, that if the particles never branch, and can only be removed, their expected number can not grow in time.

**Exercise 1.2.6** Add branching to the random walk we have discussed in Section 1.2 of this chapter, and obtain a more general parabolic equation, in higher dimensions:

\[
\frac{\partial u}{\partial t} - a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(x) \frac{\partial u}{\partial x_j} + c(x)u = 0. \tag{1.2.36}
\]

### 1.3 The maximum principle interlude: the basic statements

As the parabolic maximum principle underlies most of the parabolic existence and regularity theory, we first recall some basics on the maximum principle for parabolic equations. They are very similar in spirit to what we have described in the previous chapter for the Laplace and Poisson equations. This material can, once again, be found in many standard textbooks, such as [56], so we will not present most of the proofs but just recall the statements we will need.
We consider a (more general than the Laplacian) elliptic operator of the form

\[ Lu(x) = -a_{ij}(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(t, x) \frac{\partial u}{\partial x_j}, \quad (1.3.1) \]

in a bounded domain \( x \in \Omega \subseteq \mathbb{R}^n \) and for \( 0 \leq t \leq T \). Note that the zero-order coefficient is set to be zero for the moment. The ellipticity of \( L \) means that the matrix \( a_{ij}(t, x) \) is uniformly positive-definite and bounded. That is, there exist two positive constants \( \lambda > 0 \) and \( \Lambda > 0 \) so that, for any \( \xi \in \mathbb{R}^n \), and \( 0 \leq t \leq T \), and any \( x \in \Omega \), we have

\[ \lambda |\xi|^2 \leq a_{ij}(t, x) \xi_i \xi_j \leq \Lambda |\xi|^2. \quad (1.3.2) \]

We also assume that all coefficients \( a_{ij}(t, x) \) and \( b_j(t, x) \) are continuous and uniformly bounded.

Given a time \( T > 0 \), define the parabolic cylinder \( \Omega_T = \{0, T\} \times \Omega \) and its parabolic boundary as

\[ \Gamma_T = \{ x \in \Omega, \ 0 \leq t \leq T : \text{either} \ x \in \partial \Omega \text{ or } t = 0 \}. \]

In other words, \( \Gamma_T \) is the part of the boundary of \( \Omega_T \) without “the top” \( \{(t, x) : t = T, x \in \Omega\} \).

**Theorem 1.3.1** (The weak maximum principle) Let a function \( u(t, x) \) satisfy

\[ \frac{\partial u}{\partial t} + Lu \leq 0, \quad x \in \Omega, \ 0 \leq t \leq T, \quad (1.3.3) \]

and assume that \( \Omega \) is a smooth bounded domain. Then \( u(t, x) \) attains its maximum over \( \Omega_T \) on the parabolic boundary \( \Gamma_T \), that is,

\[ \sup_{\Omega_T} u(t, x) = \sup_{\Gamma_T} u(t, x). \quad (1.3.4) \]

As in the elliptic case, we also have the strong maximum principle.

**Theorem 1.3.2** (The strong maximum principle) Let a smooth function \( u(t, x) \) satisfy

\[ \frac{\partial u}{\partial t} + Lu \leq 0, \quad x \in \Omega, \ 0 \leq t \leq T, \quad (1.3.5) \]

in a smooth bounded domain \( \Omega \). Then if \( u(t, x) \) attains its maximum over \( \Omega_T \) at an interior point \( (t_0, x_0) \notin \Gamma_T \) then \( u(t, x) \) is equal to a constant in \( \Omega_T \).

We will not prove these results here, the reader may consult [56] or other standard textbooks on PDEs for a proof. One standard generalization of the maximum principle is to include the lower order term with a sign, as in the elliptic case – compare to Theorem ?? in Chapter ??.

Namely, it is quite straightforward to show that if \( c(x) \geq 0 \) then the maximum principle still holds for parabolic equations (1.3.5) with an operator \( L \) of the form

\[ Lu(x) = -a_{ij}(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(t, x) \frac{\partial u}{\partial x_j} + c(t, x)u. \quad (1.3.6) \]
The proof can, once again, be found in [56]. However, as we have seen in the elliptic case, in the maximum principles for narrow domains (Theorem ?? in Chapter ??) and domains of a small volume (Theorem ?? in the same chapter), the sign condition on the coefficient $c(t, x)$ is not necessary for the maximum principle to hold. Later in this chapter, we will discuss a more general condition that quantifies the necessary assumptions on the operator $L$ for the maximum principle to hold in a unified way.

A consequence of the maximum principle is the comparison principle, a result that holds also for operators with zero order coefficients and in unbounded domains. In general, the comparison principle in unbounded domains holds under a proper restriction on the growth of the solutions at infinity. Here, for simplicity we assume that the solutions are uniformly bounded.

**Theorem 1.3.3** Let the smooth uniformly bounded functions $u(t, x)$ and $v(t, x)$ satisfy

$$
\frac{\partial u}{\partial t} + Lu + c(t, x)u \geq 0, \quad 0 \leq t \leq T, \quad x \in \Omega
$$

and

$$
\frac{\partial v}{\partial t} + Lv + c(t, x)v \leq 0, \quad 0 \leq t \leq T, \quad x \in \Omega,
$$

in a smooth (and possibly unbounded) domain $\Omega$, with a bounded function $c(t, x)$. Assume that $u(0, x) \geq v(0, x)$ and

$$
u(t, x) \geq v(t, x) \text{ for all } 0 \leq t \leq T \text{ and } x \in \partial \Omega.
$$

Then, we have

$$u(t, x) \geq v(t, x) \text{ for all } 0 \leq t \leq T \text{ and all } x \in \Omega.
$$

Moreover, if in addition, $u(0, x) > v(0, x)$ on an open subset of $\Omega$ then $u(t, x) > v(t, x)$ for all $0 < t < T$ and all $x \in \Omega$.

The assumption that both $u(t, x)$ and $v(t, x)$ are uniformly bounded is important if the domain $\Omega$ is unbounded – without this condition even the Cauchy problem for the standard heat equation in $\mathbb{R}^n$ may have more than one solution, while the comparison principle implies uniqueness trivially. An example of non-uniqueness is discussed in detail in [79] – such solutions grow very fast as $|x| \to +\infty$ for any $t > 0$, while satisfying the initial condition $u(0, x) \equiv 0$. The extra assumption that $u(t, x)$ is bounded allows to rule out this non-uniqueness issue. Note that the special case $\Omega = \mathbb{R}^n$ is included in Theorem 1.3.3, and in that case only the comparison at the initial time $t = 0$ is needed for the conclusion to hold for bounded solutions. Once again, a reader not interested in treating the proof as an exercise should consult [56], or another of his favorite basic PDE textbooks. We should stress that in the rest of this book we will only consider solutions for which the uniqueness holds.

A standard corollary of the parabolic maximum principle is the following estimate.

**Exercise 1.3.4** Let $\Omega$ be a (possibly unbounded) smooth domain, and $u(t, x)$ be the solution of the initial boundary value problem

$$
u_t + Lu + c(t, x)u = 0, \quad \text{in } \Omega,
$$

$$u(t, x) = 0 \text{ for } x \in \partial \Omega,
$$

$$u(0, x) = u_0(x).
$$
Assume (to ensure the uniqueness of the solution) that $u$ is locally in time bounded: for all $T > 0$ there exists $C_T > 0$ such that $|u(t,x)| \leq C_T$ for all $t \in [0,T]$ and $x \in \Omega$. Assume that the function $c(t,x)$ is bounded, with $c(t,x) \geq -M$ for all $x \in \Omega$, and show that then $u(t,x)$ satisfies

$$|u(t,x)| \leq \|u_0\|_{L^\infty} e^{Mt}, \text{ for all } t > 0 \text{ and } x \in \Omega.$$  \hspace{1cm} (1.3.10)

The estimate (1.3.10) on the possible growth (or decay) of the solution of (1.3.9) is by no means optimal, and we will soon see how it can be improved.

We also have the parabolic Hopf Lemma, of which we will only need the following version.

**Lemma 1.3.5** (The parabolic Hopf Lemma) Let $u(t,x) \geq 0$ be a solution of

$$u_t + Lu + c(t,x)u = 0, \quad 0 \leq t \leq T,$$

in a ball $B(z,R)$. Assume that there exists $t_0 > 0$ and $x_0 \in \partial B(z,R)$ such that $u(t_0,x_0) = 0$, then we have

$$\frac{\partial u(t_0,x_0)}{\partial \nu} < 0.$$ \hspace{1cm} (1.3.11)

The proof is very similar to that of the elliptic Hopf Lemma, and can be found, for instance, in [75].

### 1.4 The forced linear heat equation

The regularity theory for the parabolic equations is an extremely rich and fascinating subject that is often misunderstood as "technical". To keep things relatively simple, we are not going to delve into it head first. Rather, we focus in this section on the regularity results for the forced linear heat equation in the whole space:

$$u_t - \Delta u = g(t,x), \quad t > 0, \quad x \in \mathbb{R}^n,$$ \hspace{1cm} (1.4.1)

with an initial condition $u(0,x) = u_0(x)$. As we will see almost immediately, in Proposition 1.4.3, the contribution of the initial condition can be treated in a very simple way, and the main question is what can we say about the regularity of $u(t,x)$ in terms of the prescribed regularity of $g(t,x)$. In Section 1.5, the answers to these seemingly technical and "boring" issues will allow us to address the question of existence and regularity of solutions to "much more interesting" nonlinear equations, in a very large class. The completely explicit results for the heat equation we describe in this section also explain quite well how one can approach general inhomogeneous parabolic equations – we explain this at the qualitative level in Section 1.6.

This section is both longer and more technical than what the reader has encountered so far in the book. This techniques are mostly elementary but still require us to get our hands dirty and the computations reveal some of the very important cancellations that underline the regularity theory in the general case discussed in Section 1.6. We proceed in several steps. First we show that if $g(t,x)$ is bounded, without any other assumptions on its regularity,
then the function $u(t, x)$ is Hölder continuous both in $t$ and in $x$, and the corresponding Hölder norms of $u$ are bounded by the $L^\infty$-norm of $g$. This is done in Section 1.4.2, and the main result there is Proposition 1.4.7. Next, in Section 1.4.3 we assume that $g(t, x)$ is Hölder continuous and show that then $u(t, x)$ is once differentiable in $t$ and twice in $x$, with the corresponding bounds on the derivatives in terms of the Hölder norm of $g$. This is made precise in Proposition 1.4.13, and a generalization to higher order derivatives is explained in Proposition 1.4.20. Finally, in Section 1.4.4 we show that if $g(t, x)$ is Hölder continuous then the first derivative in time and second derivatives in space are not just bounded but actually themselves Hölder continuous – this is stated in Proposition 1.4.18. Of course, in the first place, the reader may wonder what we mean by a solution to the heat equation that is not necessarily differentiable. This is explained in Section 1.4.1 in terms of the Duhamel formula.

The proofs of all these results are painfully computational but they open the gates to beautiful results in the theory of nonlinear diffusion equations, so the payoff for the hard work in this section is quite high. The hope is that the reader will emerge at the end of this section with the understanding that the parabolic regularity theory does require some calculations but is by no means mysterious or inaccessible. As we will see later, the results it provides are not light but worth their weight in gold.

Recommendaion. This section contains many exercises that are computational in nature and may at the first look appear somewhat unappealing to the reader. We strongly encourage you to do them as they show the machinery and details behind the beautiful theory.

1.4.1 The Duhamel formula

We consider the forced linear heat equation

$$u_t = \Delta u + g(t, x), \quad (1.4.2)$$

posed in the whole space $x \in \mathbb{R}^n$, and with an initial condition

$$u(0, x) = u_0(x). \quad (1.4.3)$$

The basic question for us in this section is how regular the solution of (1.4.2)-(1.4.3) is, in terms of the regularity of the initial condition $u_0(x)$ and the forcing term $g(t, x)$. The function $u(t, x)$ is given explicitly by the Duhamel formula

$$u(t, x) = v(t, x) + \int_0^t w(t, x; s)ds. \quad (1.4.4)$$

Here, $v(t, x)$ is the solution to the homogeneous heat equation

$$v_t = \Delta v, \quad x \in \mathbb{R}^n, \quad t > 0, \quad (1.4.5)$$

with the initial condition $v(0, x) = u_0(x)$, and $w(t, x; s)$ is the solution of the Cauchy problem

$$w_t(t, x; s) = \Delta w(t, x; s), \quad x \in \mathbb{R}^n, \quad t > s, \quad (1.4.6)$$

that runs starting at the time $s$, and is supplemented by the initial condition at $t = s$:

$$w(t = s, x; s) = g(s, x). \quad (1.4.7)$$
Exercise 1.4.1 If the reader has not previously encountered the Duhamel formula, you should consider it in a more general setting of a forced problem on a Banach space \( X \):

\[ \frac{du}{dt} = Lu + g, \quad t > 0, \]  

(1.4.8)

with an initial condition \( u(0) = u_0 \in X \), for some general linear operator \( L : X \to X \) and forcing \( g \in C([0, T]; X) \). The basic assumption on the operator \( L \) is that for any \( v_0 \in X \) the initial value problem

\[ \frac{dv}{dt} = Lv, \quad t > 0, \]  

(1.4.9)

with the initial condition \( v(0) = v_0 \in X \) has a unique bounded solution \( u(t) \in X \) for all \( t \geq 0 \).

Show that for all \( 0 \leq t \leq T \) the function \( u(t) \) can be written as

\[ u(t) = v(t) + \int_0^t w(t; s)ds. \]

Here, \( v(t) \) is the solution to the initial value problem (1.4.9) with \( v(0) = u_0 \), and \( w(t; s) \) solves the initial value problem starting at a time \( s < t \):

\[ \frac{dw}{dt} = Lw, \quad t > s, \]  

(1.4.10)

with the initial condition \( w(s) = g(s) \).

Let us denote the solution of the Cauchy problem (1.4.5) as

\[ v(t, x) = e^{t\Delta}u_0. \]  

(1.4.11)

This defines the operator \( e^{t\Delta} \). It maps the initial condition of the heat equation to its solution at the time \( t \), and is given explicitly as

\[ e^{t\Delta}f(x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x-y|^2/(4t)} f(y)dy. \]  

(1.4.12)

With this notation, another way to write the Duhamel formula (1.4.2) is

\[ u(t, x) = e^{t\Delta}u_0(x) + \int_0^t e^{(t-s)\Delta}g(s, x)ds, \]  

(1.4.13)

or, more explicitly:

\[ u(t, x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x-y|^2/(4t)}u_0(y)dy + \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-|x-y|^2/(4(t-s))}g(s, y)dyds. \]  

(1.4.14)

Of course, we can make these expressions much shorter and more elegant if we introduce the heat kernel

\[ G(t, x) = \frac{1}{(4\pi t)^{n/2}}e^{-|x|^2/(4t)}, \]  

(1.4.15)

and rewrite them in terms of convolutions with \( G(t, x) \). We keep the formulas for \( u(t, x) \) as explicit as feasible on purpose, to keep the potential singularities as visible as possible, so that the reader would be alert of the potential dangers in the estimates.

Here is an exercise on the Duhamel formula for a different partial differential equation.
Exercise 1.4.2 Consider the one-dimensional wave equation

\[ u_{tt} - u_{xx} = g(t, x), \quad t > 0, \quad x \in \mathbb{R}, \tag{1.4.16} \]

with zero initial condition \( u(0, x) = u_t(0, x) = 0 \). Show that its solution is given by the Duhamel formula

\[ J_{\text{wave}}(t, x) = \frac{1}{2} \int_0^t \int_{x+(t-s)}^{x-(t-s)} g(s, y)dyds. \tag{1.4.17} \]

The first term in (1.4.14) is rather benign as far as regularity is concerned. We use the notation

\[ |k| = k_1 + \cdots + k_n, \]

for a multi-index \( k = (k_1, \ldots, k_n) \), and

\[ D^{k}_{x}u = \frac{\partial^{|k|}u}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}. \]

Proposition 1.4.3 Let \( u_0 \) be a bounded and continuous function and set

\[ \begin{align*}
  v(t, x) &= e^{t\Delta}u_0 - \frac{1}{(4\pi t)^{n/2}} \int e^{-(x-y)^2/(4t)}u_0(y)dy. 
\end{align*} \tag{1.4.18} \]

Show that for any \( t > 0 \) and for any multi-index \( k \) with \( |k| = m \) there exists \( C_m > 0 \) that depends only on \( m \) so that

\[ |D^{k}_{x}v(t, x)| \leq \frac{C_m}{t^{m/2}}\|u_0\|_{L^\infty}, \quad |\partial^{m}_{t}v(t, x)| \leq \frac{C_m}{t^{m}}\|u_0\|_{L^\infty}, \quad \text{for all } t > 0 \text{ and } x \in \mathbb{R}^n. \tag{1.4.19} \]

The reader should note the following simple observations. First, the estimates on the derivatives of \( v(t, x) \) in (1.4.19) blow-up as \( t \downarrow 0 \). This is expected – we only assume that \( u_0(x) \) is continuous. More importantly, the estimates on the derivatives at a positive time \( t > 0 \) depend only on the \( L^\infty \)-norm of \( u_0 \) – this is the instant regularization effect of the heat equation.

Exercise 1.4.4 Prove Proposition 1.4.3. The proof involves nothing but calculus and remembering when an integral with an integrand that depends on a parameter can be differentiated in this parameter.

For the rest of this section we focus on the second term in (1.4.14),

\[ J(t, x) = \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}}e^{-(x-y)^2/(4(t-s))}g(s, y)dyds. \tag{1.4.20} \]

Note that the function \( J(t, x) \), if it is sufficiently regular, is the solution to the forced linear heat equation

\[ J_t = \Delta J + g(t, x), \tag{1.4.21} \]

posed in the whole space \( x \in \mathbb{R}^n \), with the initial condition \( J(0, x) \equiv 0 \). However, as a priori we do not know that \( J(t, x) \) is differentiable, for now, we can not be quite sure that (1.4.21) makes classical sense as stated. It is potentially problematic because of the term \( (t-s)^{-n/2} \) in (1.4.20) that blows up as \( s \uparrow t \). In particular, a naive attempt to differentiate the integrand in \( t \) or \( x \) would lead to expressions that are too singular to be absolutely integrable without some cancellations.
Exercise 1.4.5 Differentiate the integrand in $J(t,x)$ in $t$ blindly, observe the singularity as $s \to t$ and get stuck.

However, to see that the singularity is not as dangerous as it may naively seem, observe that a simple change of variables shows that if $g(t,x)$ is bounded then so is $J(t,x)$:

$$|J(t,x)| \leq \|g\|_{L^\infty} \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} dy ds = \frac{t\|g\|_{L^\infty}}{\pi^{n/2}} \int e^{-z^2} dz = t\|g\|_{L^\infty}. \quad (1.4.22)$$

We used above the simple change of variables

$$z = \frac{x-y}{2\sqrt{t-s}}, \quad (1.4.23)$$

that will be a recurring refrain throughout this section and beyond. In other words, the integral (1.4.20) defines an $L^\infty$ function $J(t,x)$ if $g(t,x)$ itself is an $L^\infty$ function. The reader should informally think of $J(t,x)$ as the solution to (1.4.21) even if it does not have sufficient regularity to be a classical solution. In the remainder of this section we focus exactly to the question of regularity of $J(t,x)$.

Exercise 1.4.6 Deduce the upper bound (1.4.22) for $J(t,x)$ directly from the parabolic maximum principle, without any explicit computations.

1.4.2 Regularity gain: from a bounded $g(t,x)$ to a Hölder $J(t,x)$

The estimate (1.4.22) can be restated as an $L^\infty - L^\infty$ bound:

$$\|J\|_{L^\infty} \leq t\|g\|_{L^\infty}. \quad (1.4.24)$$

Such bounds are useful but they do not give any better regularity for the function $J(t,x)$ than for $g(t,x)$: it says that if $g$ is bounded then so is $J$. On the other hand, the following proposition gives a quantifiable way to say that if $g(t,x)$ is bounded, and without any assumptions on the continuity of $g$, then the function $J$ is Hölder continuous in $t$ and differentiable in $x$. Hence, it is more regular than the assumed regularity of $g$. This is a very simple example of the general phenomenon of parabolic regularity: solution is better than the input data, such as the initial condition or forcing.

Recall the notion of the Hölder norm of a function $g(t,x)$ defined for $(t,x) \in [0,T] \times \mathbb{R}^n$:

$$\|g\|_{C_t^\alpha C_x^\beta} = \|g\|_{L^\infty} + \sup_{0 \leq t \leq t', x,x' \in \mathbb{R}^n} \frac{|g(t,x) - g(t',x')|}{|t-t'|^\alpha + |x-x'|^\beta}, \quad (1.4.25)$$

with the supremum taken over all $0 \leq t, t' \leq T$ and $x,x' \in \mathbb{R}^n$ such that $(t,x) \neq (t',x')$. We will use the notation $C_t^\alpha C_x^\beta([0,T] \times \mathbb{R}^n)$, or $C_t^\alpha C_x^\beta$ for short, for the space of Hölder continuous functions on $[0,T] \times \mathbb{R}^n$ with a finite Hölder norm. We apologize to the reader for the use of this cumbersome notation but it allows us to distinguish between the regularity in time and space and avoid various other notational pitfalls.
**Proposition 1.4.7** Let \( g(t, x) \) be a measurable bounded function, so that \( g \in L^\infty([0, T] \times \mathbb{R}^n) \), and \( J(t, x) \) be given by (1.4.20). Then,

(i) the function \( J(t, x) \) is once differentiable in \( x \) for all \( t > 0 \) and \( x \in \mathbb{R}^n \),

(ii) for any \( \alpha \in (0, 1) \), the function \( J(t, x) \) is \( C^\alpha \)-Hölder continuous in \( t \) for all \( t > 0 \) and all \( x \in \mathbb{R}^n \), and

(iii) for all \( 1 \leq k \leq n \), the derivatives \( \partial_{x_k} J(t, x) \) are \( C^\alpha \)-Hölder continuous in \( x \), for all \( t > 0 \) and \( x \in \mathbb{R}^n \), for all \( \alpha \in (0, 1) \).

Moreover, there exist \( C > 0 \) and \( C_\alpha > 0 \) so that for all \( t, t' \in [0, T] \) and \( x, x' \in \mathbb{R}^n \) we have

\[
\begin{align*}
|\partial_{x_k} J(t, x)| &\leq C \sqrt{t} \| g \|_{L^\infty([0, T] \times \mathbb{R}^n)}, \\
|\partial_{x_k} J(t, x) - \partial_{x_k} J(t, x')| &\leq C_\alpha \| g \|_{L^\infty([0, T] \times \mathbb{R}^n)} |x - x'|^\alpha, \\
|J(t, x) - J(t', x)| &\leq C_\alpha \| g \|_{L^\infty([0, T] \times \mathbb{R}^n)} |t - t'|^\alpha.
\end{align*}
\]

The difference in regularity of \( J(t, x) \) in \( t \) and \( x \) is not an artifact of the proof. It is easy to see that \( J(t, x) \) need not be differentiable in \( t \) if all we know about \( g(t, x) \) is that it is bounded. Indeed, the reader can simply think of \( g(t, x) = \text{sgn}(t - 1) \), in which case \( J(t, x) = |t - 1| - 1 \) and is not differentiable at \( t = 1 \) but is Hölder continuous for all \( t \geq 0 \). The next exercise shows that neither can one expect the function \( J(t, x) \) to be twice continuously differentiable in \( x \) under the assumption that \( g(t, x) \) is bounded and not necessarily continuous. Hence, the claimed regularity of \( J(t, x) \) in Proposition 1.4.7 is ”reasonably optimal”.

**Exercise 1.4.8** Give an example of a bounded function \( g(t, x), t \geq 0, x \in \mathbb{R} \), such that \( J(t, x) \) is not twice continuously differentiable in \( x \) even though the derivative \( \partial_x J(t, x) \) is \( \alpha \)-Hölder continuous in \( x \) for any \( \alpha \in (0, 1) \).

The next exercise asks you to compare the gain of regularity for the heat equation and for the wave equation.

**Exercise 1.4.9** Does the result of Proposition 1.4.7 apply to the solution to the wave equation given in Exercise 1.4.2?

**Proof of Proposition 1.4.7.** Let us freeze \( t > 0 \), fix some \( 1 \leq i \leq n \), and prove that \( J(t, x) \) is differentiable in \( x_i \). The first inclination may be to simply differentiate the integrand in (1.4.20), as suggested, albeit with a warning, in Exercise 1.4.4. This cannot be done in the \( t \)-variable, simply because we have seen that \( J(t, x) \) need not be differentiable in \( t \). Such differentiation in the \( x \)-variable can be justified, but it is also instructive to work from scratch with the finite differences, as we will need to do that with the time increments anyway. Let \( e_i \) be the unit vector in the \( x_i \)-direction and write

\[
\frac{J(t, x + he_i) - J(t, x)}{h} = \frac{1}{h} \int_0^t \int_{\mathbb{R}^n} \left[ e^{-(x + he_i - y)^2/4(t-s)} - e^{-(x - y)^2/4(t-s)} \right] g(s, y) \frac{dyds}{(4\pi(t-s))^{n/2}}.
\]

The familiar change of variables (1.4.23)

\[
z = \frac{x - y}{2\sqrt{t-s}},
\]

\[ (1.4.27) \]
leads to

\[
\frac{J(t, x + he_i) - J(t, x)}{h} = \frac{1}{h} \int_0^t \int_{\mathbb{R}^n} \left[ e^{-(z+he_i/(2\sqrt{t-s}))^2} - e^{-z^2} \right] g(s, x - 2z\sqrt{t-s}) \frac{dz ds}{(4\pi)^{n/2}}
\]

\[
= \int_0^t Q_{h,i}(t, s) \frac{ds}{\sqrt{t-s}},
\]

with

\[
Q_{h,i}(t, s) = \frac{\sqrt{t-s}}{h} \int_{\mathbb{R}^n} \left[ e^{-(z+he_i/(2\sqrt{t-s}))^2} - e^{-z^2} \right] g(s, x - 2z\sqrt{t-s}) \frac{dz}{(4\pi)^{n/2}}.
\]

Exercise 1.4.10 Show that if \( g \in L^\infty([0, T] \times \mathbb{R}^n) \), then for almost every \( 0 < s \leq t \) fixed we have

\[
\lim_{h \to 0} Q_{h,i}(t, s) = \tilde{Q}_i(t, s) := -\int_{\mathbb{R}^n} z_i e^{-z^2} g(s, x - 2z\sqrt{t-s}) \frac{dz}{(4\pi)^{n/2}},
\]

and that there exists \( C > 0 \) so that for almost every \( 0 < s \leq t \) and all \( h \in (0, 1) \) we have

\[
|Q_{h,i}(t, s)| \leq C\|g\|_{L^\infty}.
\]

The result of Exercise 1.4.10 allows us to use the Lebesgue dominated convergence theorem and pass to the limit \( h \to 0 \) in (1.4.29) and conclude that

\[
\frac{\partial J(t, x)}{\partial x_i} = -\int_0^t \int_{\mathbb{R}^n} z_i e^{-z^2} g(s, x - 2z\sqrt{t-s}) \frac{dz}{(4\pi)^{n/2}} \frac{ds}{\sqrt{t-s}}.
\]

This shows both that \( J(t, x) \) is differentiable in \( x \) and that the first bound in (1.4.26) holds:

\[
\left| \frac{\partial J(t, x)}{\partial x_i} \right| \leq C\sqrt{t}\|g\|_{L^\infty}.
\]

Exercise 1.4.11 Verify that differentiating the integrand twice in \( x \) leads to the same kind of (seemingly non-integrable) singularity in \( t-s \) as differentiating once in \( t \).

Nevertheless, the Hölder continuity of \( J(t, x) \) in time is proved by a very similar, except slightly longer, argument. We again compute a partial difference. Assume, for convenience, that \( t' \geq t \), and write

\[
J(t, x) - J(t', x) = \int_0^t \int_{\mathbb{R}^n} \left( \frac{e^{-|x-y|^2/4(t-s)}}{(4\pi(t-s))^{n/2}} - \frac{e^{-|x-y|^2/4(t'-s)}}{(4\pi(t'-s))^{n/2}} \right) g(s, y) dy ds
\]

\[
- \int_t^{t'} \int_{\mathbb{R}^n} \frac{e^{-|x-y|^2/4(t'-s)}}{(4\pi(t'-s))^{n/2}} g(s, y) dy ds = I_1(t, t', x) + I_2(t, t', x).
\]
The second term above satisfies the simple estimate
\[ |I_2(t, t', x)| \leq \|g\|_{L^\infty} |t' - t|, \tag{1.4.36} \]
obtained via the by now automatic change of variables as in (1.4.28). As for \(I_1\), we write, using the Newton-Leibniz formula in the \(t\) variable, for a fixed \(s \in [0, t]\)
\[ e^{-|x-y|^2/4(t'-s)} - e^{-|x-y|^2/4(t-s)} = \int_s^{t'} \frac{h(z)}{(4\pi(t-s))^{n/2+1}} d\tau, \quad z = \frac{x-y}{\sqrt{\tau-s}}, \tag{1.4.37} \]
with an integrable function
\[ h(z) = \left( -\frac{n}{2} + \frac{|z|^2}{4} \right) e^{-|z|^2}. \]
Thus, we have, changing the variables \(y \to z\) in the integral over \(\mathbb{R}^n\), and integrating \(z\) out, using integrability of \(h(z)\):
\[ |I_1(t, t', x)| \leq C\|g\|_{L^\infty} \int_0^t \int_s^{t'} \frac{d\tau}{\tau-s} ds = C\|g\|_{L^\infty} \int_0^t \log \left( \frac{t'-s}{t-s} \right) ds \]
\[ = C\|g\|_{L^\infty} (t' \log t' - t \log t - (t' - t) \log(t' - t)). \tag{1.4.38} \]
This proves that
\[ |I_1(t, t', x)| \leq C\|g\|_{L^\infty} |t' - t|^{\alpha}, \tag{1.4.39} \]
for all \(\alpha \in (0, 1)\).

**Exercise 1.4.12** Consider the partial differences
\[ \frac{\partial J(t, x + he_j)}{\partial x_i} - \frac{\partial J(t, x)}{\partial x_i} \]
using expression (1.4.33) for \(\partial_x J(t, x)\) and use a trick similar to (1.4.37) to show that \(\partial_x J(t, x)\) is Hölder continuous and the last estimate in (1.4.26) holds.

This exercise finishes the proof of Proposition 1.4.7. \(\square\)

Let us stress again that the logarithmic term \(\log(t - t')\) that appears in (1.4.38) is not a fluke of the proof: it represents a genuine obstacle to differentiability of \(J(t, x)\) in time if \(g(t, x)\) is just bounded and not Hölder continuous in \(t\). This fact is absolutely crucial in the parabolic regularity theory, and not just in the present reasonably simple context.

### 1.4.3 Regularity gain: from Hölder \(g(t, x)\) to differentiable \(J(t, x)\)

Proposition 1.4.7 shows that if we assume that \(g(t, x)\) is bounded then \(J(t, x)\) is differentiable in space and Hölder continuous in time, and, as we have seen, one can not expect a better regularity for \(J(t, x)\) without further assumptions on the function \(g(t, x)\). We now assume that \(g(t, x)\) itself is Hölder continuous in time and space, and show that then \(J(t, x)\) is differentiable in \(t\) and twice differentiable in \(x\).
Proposition 1.4.13 Assume that \( g(t, x) \in C^{\alpha/2}_t C^\alpha_x ([0, T] \times \mathbb{R}^n) \), so that there exists \( K > 0 \) such that for all \( 0 \leq t, t' \leq T \) and \( x, x' \in \mathbb{R}^n \) we have

\[
|g(t, x) - g(t', x')| \leq K (|t - t'|^{\alpha/2} + |x - x'|^\alpha)
\]  
(1.4.40)

for some \( \alpha \in (0, 1) \). Then \( J(t, x) \) given by (1.4.20) is twice continuously differentiable in \( x \), and once continuously differentiable in \( t \) over \((0, T) \times \mathbb{R}^n \). Moreover, there exists \( C > 0 \) so that

\[
\|D^2_x J\|_{L^\infty([0, T] \times \mathbb{R}^n)} + \|\partial_t J\|_{L^\infty([0, T] \times \mathbb{R}^n)} \leq C \|g\|_{C^{\alpha/2, \alpha}([0, T] \times \mathbb{R}^n)}.
\]  
(1.4.41)

Note the difference in the Hölder exponents in \( t \) and \( x \) in the assumption (1.4.40) on the function \( g(t, x) \). It is of course not necessary as any Hölder assumptions in \( x \) and \( t \) would lead to result. But it is very natural, as will be clear from the argument below. In a similar fashion, the gain of regularity in time and space for \( J \) is different: one derivative in time and two derivatives in space. Both instances are related to the different scaling of the heat equation and other parabolic problems in time and space.

Continuing the theme started in Exercise 1.4.9, we ask the reader to consider the following question.

Exercise 1.4.14 Does the result of Proposition 1.4.13 apply to the solution \( J_{wave}(t, x) \) to the wave equation given in Exercise 1.4.2?

Proof. One could look again at the partial differences, as in the proof of Proposition 1.4.7. However, we will use a different strategy, to illustrate another method. We will take \( \delta \in (0, t) \) small, and consider an approximation

\[
J_\delta(t, x) = \int_0^{t-\delta} e^{(t-s)\Delta} g(s, \cdot)(x) ds = \int_0^{t-\delta} \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} g(s, y) dy ds.
\]  
(1.4.42)

Note that \( J_\delta(s, x) \) is the solution to the Cauchy problem (in the variable \( s \), with \( t \) fixed)

\[
\frac{\partial J_\delta}{\partial s} = \Delta J_\delta + H(t - s - \delta) g(s, x),
\]  
(1.4.43)

with the initial condition \( J_\delta(0, x) = 0 \). Here, we have introduced the cut-off \( H(s) = 1 \) for \( s < 0 \) and \( H(s) = 0 \) for \( s > 0 \).

The function \( J_\delta(t, x) \) is smooth both in \( t \) and \( x \) for all \( \delta > 0 \) – this is easy to check simply by differentiating the integrand in (1.4.42) in \( t \) and \( x \), since that does not produce any singularity due to \( t - s > \delta \). Moreover, \( J_\delta(t, x) \) converges uniformly to \( J(t, x) \) as \( \delta \downarrow 0 \) – this follows from the estimate

\[
|J(t, x) - J_\delta(t, x)| \leq \delta \|g\|_{L^\infty},
\]  
(1.4.44)

that can be checked as in (1.4.22).

Exercise 1.4.15 Check that (1.4.44) holds.

As a consequence of (1.4.44), the derivatives of \( J_\delta(t, x) \) converge weakly, in the sense of distributions, to the corresponding weak derivatives of \( J(t, x) \). Thus, to show that, say, the second derivatives (understood in the sense of distributions) \( \partial_{x_1 x_2} J(t, x) \) are actually continuous functions, it suffices to prove that the partial derivatives \( \partial_{x_1 x_2} J_\delta(t, x) \) converge uniformly to a continuous function, and that is what we will do. In other words, we are relying on the following real analysis exercise.
Exercise 1.4.16 Assume that \( f_n(x), x \in \mathbb{R} \), is a sequence of infinitely differentiable functions that converges uniformly on \( \mathbb{R} \) to a limit \( g \in C(\mathbb{R}) \). Show that then \( f'_n \to g' \) in the sense of distributions. Suppose, in addition, that there is a function \( p \in C(\mathbb{R}) \) such that \( f'_n \to p \), also uniformly on \( \mathbb{R} \). Show that then \( g(x) \) is continuously differentiable and \( p(x) = g'(x) \) for all \( x \in \mathbb{R} \).

We will look in detail at \( \partial_x J_\delta \), with \( i \neq j \). As the integrand for \( J_\delta \) has no singularity at \( s = t \), we may simply differentiate under the integral sign

\[
\frac{\partial^2 J_\delta(t, x)}{\partial x_i \partial x_j} = \int_0^{t-\delta} \int_{\mathbb{R}^n} \frac{(x_i - y_i)(x_j - y_j)}{4(t-s)^2(4\pi(t-s))^{n/2}} e^{-|x-y|^2/4(t-s)} g(s, y) ds dy.
\]

The extra factor \((t-s)^2\) in the denominator can not be removed simply by the change of variable (1.4.28) – as the reader can immediately check, this would still leave a non-integrable extra factor of \((t-s)^{-1}\) that would cause an obvious problem in passing to the limit \( \delta \downarrow 0 \).

A very simple but absolutely crucial observation that will come to our rescue here is that, as \( i \neq j \), we have

\[
\int_{\mathbb{R}^n} (x_i - y_i)(x_j - y_j) e^{-|x-y|^2/4(t-s)} dy = 0. \tag{1.4.45}
\]

This allows us to write

\[
\frac{\partial^2 J_\delta(t, x)}{\partial x_i \partial x_j} = \int_0^{t-\delta} \int_{\mathbb{R}^n} \frac{(x_i - y_i)(x_j - y_j)}{4(t-s)^2(4\pi(t-s))^{n/2}} e^{-|x-y|^2/4(t-s)} (g(s, y) - g(t, x)) ds dy. \tag{1.4.46}
\]

Note that we use here crucially the fact that \( \delta > 0 \) and all integrals are finite because of that. Now, we can use the regularity of \( g(s, y) \) to help us. In particular, the H"older continuity assumption (1.4.40) gives

\[
\left| \frac{(x_i - y_i)(x_j - y_j)}{4(t-s)^2(4\pi(t-s))^{n/2}} e^{-|x-y|^2/4(t-s)} (g(s, y) - g(t, x)) \right| \leq C|z|^2 e^{-|z|^2/4}\left[(t-s)^{\alpha/2} + |x-y|^\alpha\right] ||g||_{C^{\alpha/2}_z} \leq \frac{C}{(t-s)^{1-\alpha/2}} \frac{k(z)}{(4\pi(t-s))^{n/2}} ||g||_{C^{\alpha/2}_z}, \tag{1.4.47}
\]

still with \( z = (x - y)/\sqrt{t-s} \), as in (1.4.28), and

\[
k(z) = |z|^2 e^{-|z|^2/4}(1 + |z|^\alpha).
\]

As before, the factor of \((t-s)^{n/2}\) in the right side of (1.4.47) goes into the volume element

\[
dz = \frac{dy}{(t-s)^{n/2}},
\]

and we only have the factor \((t-s)^{1-\alpha/2}\) left in the denominator in (1.4.47), which is integrable in \( s \), unlike the factor \((t-s)^{-1}\) one would get without using the cancellation in (1.4.45) and the H"older regularity of \( g(t, x) \). Thus, after accounting for the Jacobian factor, the integrand
in the expression for $\partial_{x_i x_j} J_\delta$ is dominated by an integrable function in $z$. This has two consequences. First, the Lebesgue dominated convergence theorem implies that

$$\frac{\partial^2 J_\delta(t, x)}{\partial x_i \partial x_j} \to Z_{ij}(t, x) := \int_0^t \int_{\mathbb{R}^n} \frac{(x_i - y_i)(x_j - y_j)}{4(t-s)^2(4\pi(t-s))^{n/2}} e^{-|x-y|^2/4(t-s)}(g(s, y) - g(t, x)) \, ds \, dy,$$

as $\delta \to 0$, pointwise in $t$ and $x$. In addition, the bound on the integrand in (1.4.47) implies that the convergence in (1.4.48) is uniform in $x \in \mathbb{R}^n$. In particular, the continuity of the limit $Z_{ij}(t, x)$ follows as well. Invoking the claim of Exercise 1.4.16, we now deduce that

$$\frac{\partial^2 J(t, x)}{\partial x_i \partial x_j} = Z_{ij}(t, x) \text{ for } i \neq j \text{ and all } t > 0 \text{ and } x \in \mathbb{R}^n, \quad (1.4.49)$$

and that these mixed second derivatives are continuous. In addition, we also see from (1.4.47) that

$$\left| \frac{\partial^2 J_\delta(t, x)}{\partial x_i \partial x_j} \right| \leq C \int_0^t \frac{ds}{(t-s)^{1-\alpha/2}} \|g\|_{C_t^{\alpha/2}C_x^{\alpha}} \leq C t^{\alpha/2} \|g\|_{C_t^{\alpha/2}C_x^{\alpha}}, \quad (1.4.50)$$

and thus the derivatives of $J(t, x)$ obey the same bound:

$$\left| \frac{\partial^2 J(t, x)}{\partial x_i \partial x_j} \right| \leq C t^{\alpha/2} \|g\|_{C_t^{\alpha/2}C_x^{\alpha}}. \quad (1.4.51)$$

**Exercise 1.4.17** Complete the argument by looking at the remaining derivatives $\partial_t J(t, x)$ and $\partial_{x_i x_i} J(t, x)$. In both cases, one would start with $J_\delta$, find a cancellation such as in (1.4.45), leading to a version of (1.4.46), and then pass to the limit $\delta \downarrow 0$ using the Hölder regularity of $g(t, x)$.

### 1.4.4 Regularity gain: from a Hölder $g$ to Hölder derivatives of $J$

Proposition 1.4.13 is slightly sub-optimal: it says that if $g$ is Hölder continuous then $J$ is twice differentiable in $x$ and once in $t$ but says nothing about the continuity or regularity of these derivatives. We now show that they are actually themselves Hölder continuous then the Hölder continuity passes on to the corresponding derivatives of the solution.

**Proposition 1.4.18** Assume that $g(t, x) \in C_t^{\alpha/2}C_x^{\alpha}([0, T] \times \mathbb{R}^n)$, so that there exists $K > 0$ such that for all $0 \leq t, t' \leq T$ and $x, x' \in \mathbb{R}^n$ we have

$$|g(t, x) - g(t', x')| \leq K \left( |t - t'|^{\alpha/2} + |x - x'|^{\alpha} \right) \quad (1.4.52)$$

for some $\alpha \in (0, 1)$. Then, there exists $C > 0$ such that

$$\|D_{x}^2 J\|_{C_t^{\alpha/2,\alpha}([0,T] \times \mathbb{R}^n)} + \|\partial_t J\|_{C_t^{\alpha/2}C_x^{\alpha}([0, T] \times \mathbb{R}^n)} \leq C \|g\|_{C_t^{\alpha/2}C_x^{\alpha}([0, T] \times \mathbb{R}^n)}. \quad (1.4.53)$$

There is a subtle but important point here. Our first result, Proposition 1.4.7 said that if $g$ is in $L^\infty$ then you nearly gain one derivative in time and two derivatives in space, but only nearly: the true result is that $J$ and $D_x J$ are $\alpha$-Hölder continuous in $t$ and $x$ for any $\alpha \in (0, 1)$ but it is not true that $\partial_t J$ and $D_x^2 J$ necessarily exist. Proposition 1.4.18 says, on the other
hand, that if \( g \) is Hölder continuous and not just bounded, so that \( g \in C_t^{\alpha/2}C_x^\alpha \), then you fully gain one derivative in \( t \) and two in \( x \): \( \partial_t J \) and \( D_x^2 J \) are bounded in the same space \( C_t^{\alpha/2}C_x^\alpha \) as \( g \). This result is optimal, one can not expect anything better, as can be seen simply from the form of the heat equation

\[
J_t - \Delta J = g.
\]

**Warning.** In the proof below we denote by the constants \( C, C' \) etc. various universal constants that do not depend on anything but elementary calculus, and, in particular, not on the function \( g, t \) or \( x \). We make no attempt to optimize them. We also set, for some brevity

\[
M_g = \| g \|_{C_t^{\alpha/2}C_x^\alpha}.
\]

**Proof.** Our analysis follows what we did in Section 1.4.3 except we have to look at the Hölder differences for the second derivatives. The function \( J(t, x) \) is given by the Duhamel formula

\[
J(t, x) = \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} g(s, y) dyds.
\]

As in the proof of Proposition 1.4.13, we are going to examine only \( \partial_{x_i x_j} J \), with \( i \neq j \), leaving the other derivatives to the reader as a lengthy but straightforward exercise. Let us set

\[
h_{ij}(z) = \frac{z_i z_j}{(4\pi)^{n/2}} e^{-|z|^2}, \quad D(s, t, x, y) = h_{ij} \left( \frac{x - y}{2\sqrt{t-s}} \right) \frac{g(s, y) - g(t, x)}{t-s}.
\]

so that we may write (1.4.48)-(1.4.49) as

\[
\frac{\partial^2 J(t, x)}{\partial x_i \partial x_j} = \int_t^t \int_{\mathbb{R}^n} D(s, t, x, y) dsdy.
\]

Now, for \( 0 < t < t' \leq T \) and \( x, x' \) in \( \mathbb{R}^n \), we have

\[
\frac{\partial^2 J(t', x')}{\partial x_i \partial x_j} - \frac{\partial^2 J(t, x)}{\partial x_i \partial x_j} = \int_t^{t'} \int_{\mathbb{R}^n} D(s, t', x', y) dyds + \int_0^t \int_{\mathbb{R}^n} (D(s, t', x', y) - D(s, t, x, y)) dyds = J_1(t, t', x') + J_2(t, t', x').
\]

**Exercise 1.4.19** Verify that no additional ideas other than what has already been developed in the proof of Propositions 1.4.7 and 1.4.13 are required to prove that the integral \( J_1 \) satisfies an inequality of the form

\[
|J_1(t, t', x')| \leq C M_g |t-t'|^{\alpha/2}.
\]

As for the integral \( J_2 \), we need to look at it a little deeper. The change of variables

\[
z = \frac{x - y}{2\sqrt{t-s}}
\]

transforms (1.4.57) into

\[
\frac{\partial^2 J(t, x)}{\partial x_i \partial x_j} = \int_0^t \int_{\mathbb{R}^n} h_{ij}(z) \left( \frac{g(s, x - 2\sqrt{t-s}z) - g(t, x)}{t-s} \right) dz.
\]
and $J_2$ becomes

\[
J_2(t, t', x, x') = \int_0^t \int_{\mathbb{R}^n} h_{ij}(z) \left[ \frac{g(s, x' - 2\sqrt{t' - s} z) - g(t', x')}{t' - s} - \frac{g(s, x - 2\sqrt{t - s} z) - g(t, x)}{t - s} \right] \frac{dsdz}{\pi^{n/2}}
\]

\[
= \int_0^t \int_{\mathbb{R}^n} h_{ij}(z) \left[ \frac{g(s, x' - 2\sqrt{t' - s} z) - g(t', x')}{t' - s} - \frac{g(s, x - 2\sqrt{t' - s} z) - g(t, x)}{t - s} \right] \frac{dsdz}{\pi^{n/2}}
\]

\[
+ \int_0^t \int_{\mathbb{R}^n} h_{ij}(z) \left[ g(s, x - 2\sqrt{t' - s} z) - g(t', x) - g(s, x - 2\sqrt{t - s} z) - g(t, x) \right] \frac{dsdz}{\pi^{n/2}}
\]

\[
= J_{21}(t, t', x, x') + J_{22}(t, t', x, x').
\]

(1.4.60)

We estimate each term separately.

**The estimate of $J_{22}(t, t', x, x')$.** We split the time integration domain $0 \leq s \leq t$ into the intervals

\[
A = \{ s : t - (t' - t) \leq s \leq t \}, \quad B = \{ 0 \leq s \leq t - (t' - t) \}.
\]

Note that if $t' - t \geq t$, then $A = [0, t]$ and $B$ is empty. The Hölder regularity of $g(t, x)$ in (1.4.52) implies that

\[
|g(s, x - 2\sqrt{t' - s} z) - g(t', x)| \leq 2M_g(t' - s)^{\alpha/2}(1 + |z|^\alpha),
\]

and

\[
|g(s, x - 2\sqrt{t - s} z) - g(t, x)| \leq 2M_g(t - s)^{\alpha/2}(1 + |z|^\alpha).
\]

Note that for $s \in A$ we have

\[
t' - s \leq 2(t' - t), \quad t - s \leq (t' - t).
\]

Hence, the contribution to $J_{22}$ by the integral over the interval $A$ can be bounded as

\[
J_{22A}(t, t', x, x') \leq 2M_g \int_{t-(t'-t)}^t \int_{\mathbb{R}^n} |h_{ij}(z)|(1 + |z|^{\alpha}) \left[ \frac{1}{(t' - s)^{1-\alpha/2}} + \frac{1}{(t - s)^{1-\alpha/2}} \right] \frac{dsdz}{\pi^{n/2}}
\]

\[
\leq C_{t'}(t' - t)^{\alpha/2} M_g,
\]

(1.4.63)

with a constant $C_{t'}$ that depends only on $\alpha \in (0, 1)$. We used (1.4.61) and (1.4.62) above.

To estimate the contribution to $J_{22}$ by the integral over the interval $B$, note that for $s \in B$ both increments $t - s$ and $t' - s$ are strictly positive, so that the integrand is not singular. Let us also recall that $h_{ij}$ has zero integral. Thus, we may remove both $g(t, x)$ and $g(t', x')$ from the integral. This allows us to rewrite $J_{22B}$ as

\[
J_{22B}(t, t', x, x') = \int_0^{t-(t'-t)} \int_{\mathbb{R}^n} \left( \frac{g(s, x - 2\sqrt{t' - s} z) - g(s, x - 2\sqrt{t - s} z)}{t' - s} \right) h_{ij}(z) \frac{dsdz}{\pi^{n/2}}
\]

\[
= \int_0^{t-(t'-t)} \int_{\mathbb{R}^n} \left( \frac{g(s, x' - 2\sqrt{t' - s} z) - g(s, x - 2\sqrt{t' - s} z)}{t' - s} \right) h_{ij}(z) \frac{dsdz}{\pi^{n/2}}
\]

\[
+ \int_0^{t-(t'-t)} \int_{\mathbb{R}^n} \frac{g(s, x - 2\sqrt{t' - s} z) - g(s, x - 2\sqrt{t - s} z)}{t - s} h_{ij}(z) \frac{dsdz}{\pi^{n/2}}
\]

\[
= J_{221} + J_{222}.
\]

(1.4.64)
Note that the integrand in the term $J_{21}^B$ can be bounded from above by

$$CM_g |z|^2 e^{-|z|^2} \frac{(|\sqrt{t' - s} - \sqrt{t - s})^\alpha}{t' - s},$$

(1.4.65)

with a constant $C > 0$ that only depends on $\alpha \in (0,1)$. Integrating out the $z$-variable then gives

$$J_{21}^B(t, t', x, x') \leq CM_g \int_0^{t - (t' - t)} \left(\frac{t' - s}{t - s}\right)^\alpha ds \leq CM_g \int_0^{t - (t' - t)} \frac{1}{t - s} \frac{(t' - t)^\alpha}{(t - s)^\alpha/2} ds \leq CM_g (t' - t)^{\alpha/2}. \quad (1.4.66)$$

To estimate $J_{22}^B$ we again use the zero integral property of $h_{ij}(z)$ to write this term as

$$J_{22}^B(t, t', x, x') = \int_0^{t - (t' - t)} \int_{\mathbb{R}^n} (g(s, x - 2\sqrt{t - sz}) - g(s, x)) \left(\frac{1}{t' - s} - \frac{1}{t - s}\right) h_{ij}(z) dsdz \pi^{n/2}. \quad (1.4.67)$$

The integrand in (1.4.69) can be bounded by

$$CM_g |t - s|^{\alpha/2} |z| |z|^2 e^{-|z|^2} \frac{t' - t}{(t - s)^2}. \quad (1.4.68)$$

Integrating out the $z$-variable and then the $s$ variable, we obtain

$$|J_{22}^B(t, t', x, x')| \leq CM_g |t - t'|^{\alpha/2}. \quad (1.4.69)$$

We conclude that

$$J_{22}(t, t', x, x') \leq CM_g (t' - t)^{\alpha/2}, \quad 0 < t \leq t'. \quad (1.4.70)$$

**The estimate of $J_{21}(t, t', x, x')$.** Now, we estimate

$$J_{21}(t, t', x, x') = \int_0^t \int_{\mathbb{R}^n} h_{ij}(z) \left[ \frac{g(s, x' - 2\sqrt{t' - sz}) - g(s, x')}{t' - s} - \frac{g(s, x - 2\sqrt{t - sz}) - g(t', x)}{t' - s} \right] dsdz \pi^{n/2}$$

$$= J_{21}^A + J_{21}^B. \quad (1.4.71)$$

The two terms above refer to the integration over the time interval $A = \{t - |x - x'|^2 \leq s \leq t\}$ and its complement $B$. As before, if $t \leq |x - x'|^2$, then we only have $A = \{0 \leq s \leq t\}$. In the first domain, we just use the bounds

$$|g(s, x' - 2\sqrt{t' - sz}) - g(s, x')| \leq CM_g (t' - s)^{\alpha/2}(1 + |z|^\alpha) \quad (1.4.72)$$

and

$$|g(s, x - 2\sqrt{t - sz}) - g(t', x)| \leq CM_g (t' - s)^{\alpha/2}(1 + |z|^\alpha). \quad (1.4.73)$$

After integrating out the $z$-variable, this leads to

$$|J_{21}^A(t, t', x, x')| \leq CM_g \int_0^t (t' - s)^{-1 + \alpha/2} ds \leq CM_g (t' - t)^{\alpha/2} + |x - x'|^\alpha. \quad (1.4.74)$$
Next, as \( h_{ij} \) has zero mass and \( t' - s \) is strictly positive when \( s \in B \), we can drop the terms involving \( g(t', x') \) and \( g(t', x) \) leading to

\[
J_{21}^B(t, t', x, x') = \int_0^{t-|x-x'|^2} \int_{\mathbb{R}^n} h_{ij}(z) \frac{g(s, x' - 2\sqrt{t' - s}z) - g(s, x - 2\sqrt{t' - s}z)}{t' - s} ds dz \frac{1}{\pi^{n/2}}
\]

\[
= \int_0^{t-|x-x'|^2} \int_{\mathbb{R}^n} \left( h_{ij} \left( \frac{x' - y}{2\sqrt{t' - s}} \right) - h_{ij} \left( \frac{x - y}{2\sqrt{t' - s}} \right) \right) g(s, y) ds dy \frac{1}{t' - s} \frac{1}{(4\pi(t' - s))^{n/2}}.
\]  

(1.4.75)

Once again, because \( h_{ij} \) has zero mass we have

\[
J_{21}^B(t, t', x, x') = \int_0^{t-|x-x'|^2} \int_{\mathbb{R}^n} \left( h_{ij} \left( \frac{x' - y}{2\sqrt{t' - s}} \right) - h_{ij} \left( \frac{x - y}{2\sqrt{t' - s}} \right) \right) g(s, y) ds dy \frac{1}{(4\pi(t' - s))^{n/2}}.
\]

The integrand above can be re-written as

\[
\left( h_{ij} \left( \frac{x' - y}{2\sqrt{t' - s}} \right) - h_{ij} \left( \frac{x - y}{2\sqrt{t' - s}} \right) \right) g(s, y) - g(t', x') \frac{ds dy}{(t' - s)^{3/2}}
\]

\[
= \frac{1}{2} \int_0^1 g(s, y) - g(s, x_\sigma) + g(s, x_\sigma) - g(t', x') (x' - x) \cdot \nabla h_{ij} \left( \frac{x_\sigma - y}{2\sqrt{t' - s}} \right) d\sigma,
\]

with \( x_\sigma = \sigma x + (1 - \sigma)x' \). It follows that

\[
|J_{21}^B(t, t', x, x')| \leq CM_g |x - x'| \int_0^{t-|x-x'|^2} \int_{\mathbb{R}^n} \left| \nabla h_{ij} \left( \frac{x_\sigma - y}{2\sqrt{t' - s}} \right) \right| ds
dy (t' - s)^{3/2}.
\]

(1.4.77)

Using the estimates

\[
|\nabla h(z)| \leq C|z|^3 e^{-|z|^2},
\]

and \( |x' - x_\sigma| \leq |x - x'| \), and making the usual change of variable

\[
z = \frac{x_\sigma - y}{2\sqrt{t' - s}},
\]

and integrating out the \( z \) and \( \sigma \) variables, we arrive at

\[
|J_{21}^B(t, t', x, x')| \leq CM_g |x - x'| \int_0^{t-|x-x'|^2} \frac{1}{(t - s)^{(3-\alpha)/2}} + \frac{|x - x'|^\alpha}{(t - s)^{3/2}} ds.
\]

(1.4.78)

Integrating out the \( s \) variable, we obtain

\[
|J_{21}^B(t, t', x, x')| \leq CM_g |x - x'| (|x - x'|^{\alpha - 1} + |x - x'|^\alpha |x - x'|^{\alpha - 1}) \leq CM_g |x - x'|^\alpha,
\]

(1.4.79)

thus \( J_{21} \) is also Hölder continuous, finishing the proof. \( \square \)
Higher order derivatives

The previous results can be generalized to the higher order derivatives of $J(t, x)$ assuming that the corresponding derivatives of $g(t, x)$ exist and are Hölder continuous. Recall that we use the notation

$$|k| = k_1 + \cdots + k_n,$$

for a multi-index $k = (k_1, \ldots, k_n)$, and

$$D_x^k u = \frac{\partial^{k_1}(x_1) \cdots \partial^{k_n}(x_n)}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}.$$

The following result will be indispensable in the analysis of nonlinear equations, despite its seemingly technical nature.

**Proposition 1.4.20** Assume that the function $g(t, x)$ is $M$ times continuously differentiable in $t$ and $K$ times continuously differentiable in $x$, and $\partial^M_t D_x^k g(t, x) \in C^{\alpha/2}_x C^\alpha_t ([0, T] \times \mathbb{R}^n)$ for all multi-indices $k$ with $|k| = K$, for some $\alpha \in (0, 1)$. Then $J(t, x)$ given by (1.4.20) is $M + 1$ times continuously differentiable in $t$, and $K + 2$ times continuously differentiable in $x$ for all $0 \leq t \leq T$ and $x \in \mathbb{R}^n$. Moreover, there exists $C > 0$ that depends on $M$ and $K$ so that for any multi-indices $k$ and $k'$ with $|k| = K$ and $|k'| = K + 2$ we have

$$\|\partial^{M+1} D_x^k J\|_{C^{\alpha/2}_x C^\alpha_t ([0, T] \times \mathbb{R}^n)} + \|\partial^M D_x^k J\|_{C^{\alpha/2}_x C^\alpha_t ([0, T] \times \mathbb{R}^n)} \leq C \sup_{0 \leq |r| \leq K, 0 \leq m \leq M} \|D_t^m D_x^r g\|_{C^{\alpha/2}_x C^\alpha_t ([0, T] \times \mathbb{R}^n)}.$$  

(1.4.80)

In particular, if $g(t, x)$ is infinitely differentiable with each derivative uniformly bounded in $t$ and $x$ then so is $J(t, x)$.

**Exercise 1.4.21** Provide the proof of Proposition 1.4.20. One way to run the argument is to solve $\partial_t v - \Delta v = D_t^m D_x^k g$, $v(0) = 0$, apply the results we proved above to $v$ and then show that $D_t^m D_x^k u = v$.

A remark on the constant coefficients case

To finish this section, consider solutions to general constant coefficients equations of the form

$$u_t - a_{ij} \partial_{x_i, x_j} u + b_j \partial_{x_j} u + cu = f(t, x).$$

We assume that $a_{ij}$, $b_i$, and $c$ are constants, and the matrix $A := (a_{ij})$ is positive definite: there exists a constant $\lambda > 0$ so that for any vector $\xi \in \mathbb{R}^n$ we have

$$a_{ij} \xi_i \xi_j \geq \lambda |\xi|^2.$$  

(1.4.82)

Assume also that $f$ is an $\alpha$-Hölder function over $[0, T] \times \mathbb{R}^n$, and take the initial condition $v(0, x) \equiv 0$. The function $v(t, x) = u(t, x + B t) \exp(\beta t)$, with $B = (b_1, \ldots, b_n)$, solves

$$v_t - a_{ij} \partial_{x_i, x_j} v = f(t, x + B t).$$

(1.4.83)

The change of variable $w(t, x) = v(t, \sqrt{A} x)$ brings us back to the forced heat equation:

$$w_t - \Delta w = f(t, \sqrt{A} (x + B t)).$$

(1.4.84)

We see that the conclusion of Proposition 1.4.18 also applies to other parabolic equations with constant coefficients, as long as the ellipticity condition (1.4.82) holds.
Exercise 1.4.22 Consider the solutions of the equation
\[ u_t - u_{xx} + u_y = f(t, x, y), \]  
(1.4.85) in \( \mathbb{R}^2 \) and use this example to convince yourself that the ellipticity condition is necessary for the Hölder regularity as in Proposition 1.4.18 to hold.

**Congratulations.** We congratulate the reader who managed to follow the lengthy computations in this section!

### 1.5 Regularity for the nonlinear heat equations

In this section, we reap the fruit of our labour in the previous section and prove global in time existence of solutions to some nonlinear parabolic equations. We will not strive to achieve the sharpest results. Rather, we have in mind two particular classes of nonlinear parabolic equations, for which eventually we would like to understand the large time behavior: the semi-linear and quasi-linear equations of the simplest form. The truth is that the two examples we consider here contain some of the main features under which the more general global existence and regularity results hold: the Lipschitz behavior of the nonlinearity, and the smooth spatial dependence of the coefficients in the equation. Thus, after reading this section the reader should be well prepared to digest the more general results described in other, more specialized books.

#### 1.5.1 Existence and regularity for a semi-linear diffusion equation

First, we consider semi-linear parabolic equations of the form
\[ u_t = \Delta u + f(t, x, u). \]  
(1.5.1)

Such equations are generally known as the reaction-diffusion equations, and are very common in biological and physical sciences. We will discuss the origins of such equations, and the behavior of the solutions to a class of such equations in great detail in Chapter ??.

We will consider (1.5.1) posed in \( \mathbb{R}^n \), and equipped with a bounded and continuous initial condition
\[ u(0, x) = u_0(x). \]  
(1.5.2)

As in the theory of nonlinear ordinary differential equations, we need to assume some Lipschitz property of the function \( f(t, x, u) \) in the \( u \)-variable. Otherwise we may run into blow-up issues, familiar from the solutions to the ordinary differential equation
\[ \frac{du}{dt} = u^2, \quad u(0) = u_0. \]  
(1.5.3)

Recall that if \( u_0 > 0 \) then solution to (1.5.3) exists only until the time \( T_0 = 1/u_0 \) and
\[ \lim_{t \to T_0} u(t) = +\infty. \]  
(1.5.4)

This is something we would like to avoid in this expository section.
There are two possible assumptions that will ensure that solutions to (1.5.1) exist and do not blow up in a finite time. First, we may simply assume that the function $f$ is smooth in all its variables and globally Lipschitz in $u$: there exists a constant $C_f > 0$ so that

$$|f(t, x, u) - f(t, x, u')| \leq C_f|u - u'|,$$

for all $t \geq 0$, $x \in \mathbb{R}^n$ and $u, u' \in \mathbb{R}$. (1.5.5)

Alternatively, we may assume that $f(t, x, u)$ is smooth in all its variables, and locally Lipschitz in $u$: for every $K > 0$ there exists $C_K > 0$ such that

$$|f(t, x, u) - f(t, x, u')| \leq C_K|u - u'|,$$

for all $t \geq 0$, $x \in \mathbb{R}^n$ and $|u|, |u'| \leq K$, (1.5.6)

and, in addition, there exist $M_1 < M_2$ so that

$$f(t, x, M_1) = f(t, x, M_2) = 0 \quad \text{for all } t \geq 0 \text{ and } x \in \mathbb{R}^n.$$

(1.5.7)

Under this assumption we will show that solutions corresponding to initial data $u_0$ such that $M_1 \leq u_0 \leq M_2$ will be globally regular. One reason why (1.5.6)-(1.5.7) is a useful alternative to the global Lipschitz assumption in (1.5.5) is the Fisher-KPP equation

$$u_t = \Delta u + u(1-u),$$

with the predator-prey nonlinearity $f(u) = u(1-u)$ that does not satisfy (1.5.5) but which does obey (1.5.6)-(1.5.7). We refer the reader to Chapter ?? for the discussion of how this equation arises in the biological modeling and other applications, as well as to the explanation of its name.

Another important example is the time-dependent version of the Allen-Cahn equation we have encountered in Chapter ??:

$$u_t = \Delta u + u - u^3.$$ (1.5.9)

Here, once again, the nonlinearity $f(u) = u - u^3$ satisfies (1.5.6)-(1.5.7) but not (1.5.5). We will prove the following existence result under assumptions (1.5.6)-(1.5.7).

**Theorem 1.5.1** Assume that assumptions (1.5.6)-(1.5.7) hold with some $M_1 < M_2$, and the initial condition $u_0(x)$ is bounded and smooth, and

$$M_1 \leq u_0(x) \leq M_2 \text{ for all } x \in \mathbb{R}^n.$$ (1.5.10)

Then, there exists a unique bounded smooth solution $u(t,x)$ to (1.5.1)-(1.5.2), which, in addition, satisfies

$$M_1 < u(t,x) < M_2 \text{ for all } t > 0 \text{ and } x \in \mathbb{R}^n.$$ (1.5.11)

Moreover, for all $T > 0$ each derivative of $u$ is uniformly bounded over $[T, +\infty) \times \mathbb{R}^n$.

Assumption (1.5.7) may seem too stringent to the reader. Its role is to ensure that $u(t,x)$ satisfies the uniform bounds in (1.5.11). We explain in Exercise 1.5.4 how this assumption can be relaxed, while still ensuring that (1.5.11) holds. We also do not need to assume that the initial condition $u_0(x)$ is smooth – it suffices to assume that it is bounded and continuous. This is the subject of Exercise 1.5.5.
Let us first explain the two simpler claims in Theorem 1.5.1: the bounds in (1.5.11) and uniqueness. Let \( u(t, x) \) be a bounded smooth solution to (1.5.1)-(1.5.2) with an initial condition \( u_0(x) \) that satisfies (1.5.10). Consider the function \( v(t, x) = u(t, x) - M_1 \). This function satisfies

\[
v_t = \Delta v + c(t, x)v, \tag{1.5.12}
\]

with

\[
c(t, x) = \frac{f(t, x, u(t, x)) - f(t, x, M_1)}{u(t, x) - M_1}. \tag{1.5.13}
\]

As the function \( u(t, x) \) is bounded, assumptions (1.5.7) and (1.5.6) imply that \( c(t, x) \) is also bounded. Hence, the comparison principle in Theorem 1.3.3 can be applied to (1.5.12). As \( v(0, x) = 0 \) for all \( x \in \mathbb{R}^n \), it follows that \( v(t, x) > 0 \) for all \( t > 0 \), so that \( u(t, x) > M_1 \) for all \( x \in \mathbb{R}^n \). The other inequality in (1.5.11) can be proved similarly.

Uniqueness of bounded solutions is proved in an analogous fashion. Assume that \( u_1(x) \) and \( u_2(x) \) are two smooth bounded solutions to the Cauchy problem (1.5.1)-(1.5.2). Then \( w = u_1 - u_2 \) satisfies

\[
w_t = \Delta w + c(t, x)w, \tag{1.5.14}
\]

with the initial condition \( w(0, x) = 0 \) and a bounded function

\[
c(t, x) = \frac{f(t, x, u_1(t, x)) - f(t, x, u_2(t, x))}{u_1(t, x) - u_2(t, x)}. \tag{1.5.15}
\]

The comparison principle then implies that both \( w(t, x) \leq 0 \) and \( w(t, x) \geq 0 \), thus \( w(t, x) \equiv 0 \), proving the uniqueness.

Thus, the main issue in the proof of Theorem 1.5.1 is to prove the existence of a bounded solution to (1.5.1)-(1.5.2). As the function \( f(t, x, u) \) is not necessarily globally Lipschitz, we are going to use the following trick based on the fact that \( f \) satisfies (1.5.7). Consider a function \( \tilde{f}(t, x, u) \) such that

\[
f(t, x, u) = \tilde{f}(t, x, u) \text{ for all } x \in \mathbb{R}^n \text{ and } M_1 \leq u \leq M_2, \tag{1.5.15}
\]

and there exists \( K > 0 \) so that

\[
|\tilde{f}(t, x, u)| \leq K \text{ for all } x \in \mathbb{R}^n \text{ and } u \in \mathbb{R}. \tag{1.5.16}
\]

We may also ensure that \( \tilde{f}(t, x, u) \) is globally Lipschitz: there exists \( C_f > 0 \) so that

\[
|\tilde{f}(t, x, u_1) - \tilde{f}(t, x, u_2)| \leq C_f|u_1 - u_2|, \text{ for all } t \geq 0, x \in \mathbb{R}^n \text{ and all } u_1, u_2 \in \mathbb{R}, \tag{1.5.17}
\]

as compared to \( f(t, x, u) \) that is only locally Lipschitz in \( u \). Note that (1.5.7) holds automatically for \( \tilde{f}(t, x, u) \) because of (1.5.15). Hence, by what we have just shown, any smooth bounded solution to the initial value problem

\[
u_t = \Delta u + \tilde{f}(t, x, u),
u(0, x) = u_0(x), \tag{1.5.18}
\]
with an initial condition $u_0(x)$ that satisfies assumption (1.5.10) will obey (1.5.11):

$$M_1 < u(t,x) < M_2 \text{ for all } t > 0 \text{ and } x \in \mathbb{R}^n.$$  

(1.5.19)

It follows that $\tilde{f}(t,x,u(x,t)) \equiv f(t,x,u(x,t))$, and thus any bounded solution to (1.5.18) is a solution to (1.5.1) with the same initial condition. Therefore, it suffices to construct a bounded solution to (1.5.18).

A typical approach to the existence proofs in nonlinear problems is to use a fixed point argument. To this end, it is useful, and standard, to rephrase the parabolic initial value problem (1.5.1)-(1.5.2) as an integral equation, using the Duhamel formula. This is done as follows. Given a fixed $T > 0$ and initial condition $u_0(x)$, we define an operator $T$ as a mapping of the space $C([0,T] \times \mathbb{R}^n)$ to itself via

$$[Tu](t,x) = e^{t\Delta}u_0(x) + \int_0^t e^{(t-s)\Delta} \tilde{f}(s,\cdot,u(s,\cdot))(x)ds$$

(1.5.20)

$$= e^{t\Delta}u_0(x) + \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} \tilde{f}(s,y,u(s,y))dyds,$$

with the operator $e^{t\Delta}$ defined in (1.4.12):

$$e^{t\Delta} \eta(x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-(x-y)^2/(4t)} \eta(y)dy.$$  

(1.5.21)

The Duhamel formula for the solution to the Cauchy problem (1.5.18) can be now succinctly restated as

$$u(t,x) = [Tu](t,x).$$  

(1.5.22)

In other words, any smooth bounded solution to the initial value problem is a fixed point of the operator $T$.

**Exercise 1.5.2** Show that if a function $u(t,x)$ that is continuously differentiable in $t$ and twice continuously differentiable in $x$ satisfies (1.5.22), then $u(t,x)$ is a solution to the initial value problem (1.5.18).

Thus, to prove the existence part of Theorem 1.5.1 we need to show that a fixed point of the operator $T$ exists and is sufficiently regular to differentiate it once in $t$ and twice in $x$.

**Existence of a fixed point: the Picard iteration argument on a short time interval**

The first step is to prove the existence of a fixed point of $T$ in $C([0,T] \times \mathbb{R}^n)$ for $T > 0$ sufficiently small. In the second step, we will extend the existence result to all $T > 0$.

We will use the standard Picard iteration approach: set $u^{(0)} = 0$ and define

$$u^{(n+1)}(t,x) = Tu^{(n)}(t,x).$$  

(1.5.23)

In particular, we have

$$u^{(1)}(t,x) = e^{t\Delta}u_0.$$  

(1.5.24)
As the initial condition \( u_0(x) \) is continuous and bounded, the function \( u^{(1)}(t, x) \) is infinitely differentiable in \( t \) and \( x \). Proposition 1.4.20, combined with a simple induction argument, shows that \( u^{(n)}(t, x) \) are smooth for \( t > 0 \) and \( x \in \mathbb{R}^n \), for all \( n \geq 1 \).

The global Lipschitz property (1.5.17) of \( \tilde{f}(t, x, u) \) allows us to write

\[
|\mathcal{T}u(t, x) - \mathcal{T}v(t, x)| \leq \int_0^t \int_{\mathbb{R}^n} \frac{e^{-(x-y)^2/(4(t-s))}}{(4\pi(t-s))^{n/2}} |\tilde{f}(s, y, u(s, y)) - \tilde{f}(s, y, v(s, y))| dy ds
\]

\[
\leq C_f \int_0^t \int_{\mathbb{R}^n} \frac{e^{-(x-y)^2/(4(t-s))}}{(4\pi(t-s))^{n/2}} |u(s, y) - v(s, y)| dy ds
\]

\[
\leq C_f T \sup_{0 \leq s \leq T, y \in \mathbb{R}^n} |u(s, y) - v(s, y)|.
\]

This shows that if \( T < C_f^{-1} \), then the mapping \( \mathcal{T} \) is a contraction on \( C([0, T] \times \mathbb{R}^n) \) and thus has a unique fixed point in \( C([0, T] \times \mathbb{R}^n) \). Before we extend this result to all \( T > 0 \) we first show that the fixed point is a smooth function hence a classical solution to (1.5.18) on \([0, T]\).

**The bootstrap argument**

Smoothness of the fixed point \( u(t, x) \) is proved using what is commonly called a boot-strap argument. The key observation is the following.

**Lemma 1.5.3** Let \( u(t, x) \) be a fixed point of the operator \( \mathcal{T} \) in \( C([0, T]; \mathbb{R}^n) \), so that it satisfies (1.5.22), and is bounded and continuous on \([0, T] \times \mathbb{R}^n \). Then \( u(t, x) \) is infinitely differentiable in \( t \) and \( x \) for all \( t > 0 \) and \( x \in \mathbb{R}^n \).

**Proof.** Let us write (1.5.22) as

\[
u(t, x) = u^{(1)}(t, x) + D[u](t, x),
\]

with

\[
u^{(1)}(t, x) = e^{t\Delta}u_0,
\]

and

\[
D[u](t, x) = \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} g(s, y) dy ds,
\]

where

\[
g(t, x) = \tilde{f}(t, x, u(t, x)).
\]

As we have noted, the function \( u^{(1)}(t, x) \) is infinitely differentiable for any \( t > 0 \) and \( x \in \mathbb{R}^n \) simply because it is a solution to the heat equation with a bounded and continuous initial condition \( u_0 \). Thus, we only need to deal with the Duhamel term \( D[u](t, x) \). To treat this term, we will use Propositions 1.4.7 and 1.4.18. The function \( g(t, x) \) defined in (1.5.29) is bounded on \([0, T] \times \mathbb{R}^n \) because of (1.5.16). Hence, we may apply Proposition 1.4.7 and deduce that \( D[u](t, x) \) is actually Hölder continuous on \([0, T] \times \mathbb{R}^n \), and we have a priori bounds

\[
\|D[u]\|_{L^\infty} \leq C\|g\|_{L^\infty},
\]

\[
\|\partial_x D[u]\|_{L^\infty} \leq C\|g\|_{L^\infty}
\]

(1.5.30)
and for any $\alpha \in (0, 1)$ there is $C_\alpha$ so that
\[
|D[u](t, x) - D[u](t', x)| \leq C_\alpha \|g\|_{L^\infty} |t - t'|^\alpha. \tag{1.5.31}
\]
From this and (1.5.26), we conclude that $u$ itself satisfies the same bounds:
\[
\|u\|_{L^\infty} \leq \|u_0\|_{L^\infty} + C\|g\|_{L^\infty},
\|\partial_x u\|_{L^\infty} \leq \|D_x u_0\|_{L^\infty} + C\|g\|_{L^\infty} \tag{1.5.32}
\]
\[|u(t, x) - u(t', x)| \leq \|D^2_x u_0\|_{L^\infty} |t - t'| + C_\alpha \|g\|_{L^\infty} |t - t'|^\alpha.
\]
Therefore, $u(t, x)$ is not just continuous and bounded, but also Hölder continuous in $t$ and $x$, with explicit bounds above. Then, so is $g(t, x) = \tilde{f}(t, x, u(t, x))$, and then Proposition 1.4.18 tells us that $D[u](t, x)$ is differentiable once in $t$ and twice in $x$ and the derivatives $\partial_t u$ and $D^2_x u$ are themselves Hölder continuous. Then, (1.5.26), in turn, implies that $u(t, x)$ is differentiable in $t$ and twice differentiable in $x$, with Hölder continuous derivatives, and thus $g(t, x)$ possesses the same regularity. We may iterate this argument, using Proposition 1.4.18, each time gaining derivatives in $t$ and $x$, and conclude that, actually, $u(t, x)$ is infinitely differentiable in $t$ and $x$. This is known as a boot-strap argument.

**The global in time existence**

To show that existence of a solution can be extended to all $T > 0$, note that, as we have shown that the fixed point $u(t, x)$ of the mapping $T$ is smooth, we know that $u(t, x)$ is a classical solution to the initial value problem (1.5.18) on the time interval $0 \leq t \leq T$, hence it satisfies
\[
M_1 \leq u(T, x) \leq M_2, \quad \text{for all } x \in \mathbb{R}^n. \tag{1.5.33}
\]
Moreover, the existence time $T$ does not depend on $u_0$. Therefore, we can repeat the Picard iteration argument on the time intervals $[T, 2T]$, $[2T, 3T]$, and so on, eventually constructing a global in time solution to the Cauchy problem. This finishes the proof of Theorem 1.5.1.

**Exercise 1.5.4** Assumption (1.5.7) is more stringent than necessary. Show that the claim of Theorem 1.5.1 holds also if instead of (1.5.7) we assume that there exist $M_1$ and $M_2$ such that
\[
f(t, x, M_1) \text{sgn}(M_1) \leq 0 \text{ and } f(t, x, M_2) \text{sgn}(M_2) \leq 0 \quad \text{for all } t \geq 0 \text{ and } x \in \mathbb{R}^n, \tag{1.5.34}
\]
and the initial condition $u_0(x)$ satisfies (1.5.10).

**Exercise 1.5.5** As the reader may have noticed, the only place where we have used the assumption that the initial condition $u_0(x)$ is smooth is in estimates (1.5.32). It was used there to bound the derivatives of the contribution $u^{(1)}(t, x) = e^{t\Delta}u_0$ to $u(t, x)$ in (1.5.26). However, this term is smooth even if $u_0(x)$ is just continuous and not necessarily smooth. Use this to show that the conclusion of Theorem 1.5.1 holds if we only assume that $u_0$ is bounded and continuous.
1.5.2 The regularity of the solutions to a quasi-linear heat equation

One may wonder if the treatment that we have given to the semi-linear heat equation (1.5.1) is too specialized. To dispel this concern, we show how the above approach can be extended to equations with a drift and quasi-linear heat equations of the form

\[ u_t - \Delta u = f(t, x, \nabla u), \quad (1.5.35) \]

posed for \( t > 0 \) and \( x \in \mathbb{R}^n \). The nonlinearity is now stronger: it depends not on \( u \) itself but on its gradient \( \nabla u \). We ask that the nonlinear term \( f(t, x, p) \) satisfies the following two hypotheses: first, there exists \( C_1 > 0 \) so that

\[ |f(t, x, 0)| \leq C_1 \text{ for all } t \geq 0 \text{ and } x \in \mathbb{R}^n, \quad (1.5.36) \]

and, second, \( f \) is uniformly Lipschitz in the \( p \)-variable: there exists \( C_2 > 0 \) so that

\[ |f(t, x, p_1) - f(t, x, p_2)| \leq C_2|p_1 - p_2|, \text{ for all } t \geq 0 \text{ and } x, p_1, p_2 \in \mathbb{R}^n. \quad (1.5.37) \]

One consequence of (1.5.36) and (1.5.37) is a uniform bound

\[ |f(t, x, p)| \leq C_3(1 + |p|), \text{ for all } t \geq 0, \text{ } x, p \in \mathbb{R}^n, \quad (1.5.38) \]

showing that \( f(t, x, p) \) grows at most linearly in \( p \). We also require that \( f(t, x, p) \) is smooth in \( t, x \) and \( p \), and obeys the estimates

\[ |\partial^m_t f(t, x, p)| + |D^k_x f(t, x, p)| \leq C_{m,k}(1 + |p|), \text{ for all } t \geq 0, \text{ } x, p \in \mathbb{R}^n, \quad (1.5.39) \]

for any \( m \geq 1 \) and multi-index \( k \in \mathbb{Z}^n \). This smoothness assumption can be greatly relaxed but we are not concerned with the optimal results here.

Two standard examples of equations of the form (1.5.35) are parabolic equations with constant diffusion and nonuniform drifts, such as

\[ u_t = \Delta u + b_j(t, x) \frac{"}{dx_j}, \quad (1.5.40) \]

with a prescribed drift \( b(t, x) = (b_1(t, x), \ldots, b_n(t, x)) \), and viscous regularizations of the Hamilton-Jacobi equations, such as

\[ u_t = \Delta u + f(|\nabla u|). \quad (1.5.41) \]

We will encounter both of them in the sequel. Our goal is to prove the following.

**Theorem 1.5.6** Under the above assumptions, equation (1.5.35), equipped with a bounded continuous initial condition \( u_0 \), has a unique smooth solution \( u(t, x) \) over \((0, +\infty) \times \mathbb{R}^n\), which is bounded with all its derivatives over every set of the form \((\varepsilon, T) \times \mathbb{R}^n\), with \( 0 < \varepsilon < T \).

We will use the ideas displayed in the proof of Theorem 1.5.1. However, a serious additional difficulty for a quasi-linear equation (1.5.35) compared to a semi-linear equation such as (1.5.1) is that it involves a nonlinear function of the gradient of the function \( u \), which, a priori, may not be smooth at all. That is, if \( u \) is not smooth, and its gradient is only a distribution,
giving the meaning to a nonlinear function \( f(x, \nabla u) \) becomes problematic. Note that there is no problem of that sort with the Laplacian \( \Delta u \) in (1.5.35), as we may interpret it in the sense of distributions. In addition, if we try to write down the Duhamel formula for (1.5.35), an analog to (1.5.20)-(1.5.18), it would take the form

\[
 u(t, x) = [\mathcal{T}u](t, x),
\]

(1.5.42)

with the operator \( \mathcal{T} \) given now by

\[
[\mathcal{T}u](t, x) = e^{t\Delta}u_0(x) + \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-\frac{(x-y)^2}{4(t-s)}} f(s, y, \nabla u(s, y)) dy ds.
\]

(1.5.43)

This operator can not be considered as a mapping on \( C([0, T] \times \mathbb{R}^n) \) because of the term involving \( \nabla u \). Hence, the strategy in the proof of Theorem 1.5.1 needs to be modified.

A natural and standard idea is to regularize the nonlinear term, and then pass to the limit, removing the regularization. We will consider the following nonlocal approximation to (1.5.35):

\[
 u^\varepsilon_t - \Delta u^\varepsilon = f(t, x, \nabla v^\varepsilon), \quad v^\varepsilon = e^{\varepsilon\Delta}u^\varepsilon.
\]

(1.5.44)

When \( \varepsilon > 0 \) is small, one expects the solutions to (1.5.35) and (1.5.44) to be close as

\[
 e^{\varepsilon \Delta} \psi \to \psi, \text{ as } \varepsilon \to 0.
\]

(1.5.45)

**Exercise 1.5.7** For \( \psi \) in which function spaces does the convergence in (1.5.45) hold? For instance, does it hold in \( L^2 \) or \( L^\infty \)? How about \( C^1(\mathbb{R}) \)?

A damper on our expectations is that the convergence in (1.5.45) does not automatically translate into the convergence of the corresponding gradients, unless we already know that \( \psi \) is differentiable. In other words, there is no reason to expect that

\[
 \nabla(e^{\varepsilon \Delta} \psi) \to \nabla \psi,
\]

simply because the right side may not exist. Unfortunately, a result of this kind is exactly what we need in order to understand the convergence of the term \( f(x, \nabla v^\varepsilon) \) in (1.5.44).

Nevertheless, a huge advantage of (1.5.44) over (1.5.35) is that the function \( v^\varepsilon \) that appears inside the nonlinearity is smooth if \( u^\varepsilon \) is merely continuous, as long as \( \varepsilon > 0 \). This can be used to show that the Cauchy problem for (1.5.44) has a unique smooth solution.

**Exercise 1.5.8** Show that, for every \( \varepsilon > 0 \) and every bounded function \( u(x) \), we have

\[
 \| \nabla(e^{\varepsilon \Delta} u) \|_{L^\infty} \leq \frac{C}{\sqrt{\varepsilon}} \| u \|_{L^\infty}.
\]

(1.5.46)

Use this fact, and the strategy in the proof of Theorem 1.5.1, to prove that (1.5.44), equipped with a bounded continuous initial condition \( u_0 \), has a unique smooth solution \( u^\varepsilon \) over a set of the form \( (0, T_\varepsilon] \times \mathbb{R}^n \), with a time \( T_\varepsilon > 0 \) that depends on \( \varepsilon > 0 \) but not on the initial condition \( u_0 \).
**Recommendation.** The reader should take this exercise very seriously. You do not need any tools beyond what has been already done in this chapter, and it presents a good opportunity to check your understanding so far.

Having constructed solutions to (1.5.44) on a finite time interval \([0, T_\varepsilon]\), in order to obtain a global in time solution to the original equation (1.5.35), we need to do two things: (1) extend the existence of the solutions to the approximate equation (1.5.44) to all \(t > 0\), and (2) pass to the limit \(\varepsilon \to 0\) and show that the limit of \(u^\varepsilon\) exists (possibly along a sub-sequence) and satisfies “the true equation” (1.5.35). The latter step will require uniform bounds on \(\nabla u^\varepsilon\) that do not depend on \(\varepsilon\) – something much better than what is required in Exercise 1.5.8. The last step will be to prove uniqueness of such global in time smooth solution to (1.5.35) but that is much simpler.

**Global in time existence of the approximate solution**

To show that the solution to (1.5.44) exists for all \(t > 0\), and not just on the interval \([0, T_\varepsilon]\), we use the Duhamel formula

\[
 u^\varepsilon(t, x) = e^{t\Delta}u_0(x) + \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} f(s, y, \nabla u^\varepsilon(s, y)) dy ds. \tag{1.5.47}
\]

Assumption (1.5.38), together with the gradient bound (1.5.46), implies an estimate

\[
 |f(t, x, \nabla u^\varepsilon(t, x))| \leq C(1 + |\nabla u^\varepsilon(t, x)|) \leq C\left(1 + \frac{\|u^\varepsilon(t, \cdot)\|_{L^\infty}}{\varepsilon}\right), \tag{1.5.48}
\]

that can be used in (1.5.47) to yield a Gronwall inequality

\[
 \|u^\varepsilon(t, \cdot)\|_{L^\infty} \leq \|u_0\|_{L^\infty} + Ct + \frac{C}{\sqrt{\varepsilon}} \int_0^t \|u^\varepsilon(s, \cdot)\|_{L^\infty} ds. \tag{1.5.49}
\]

We used the maximum principle to bound the first term in the right side of (1.5.47), and (1.5.48) together with the standard change of variables (1.4.28):

\[
 z = \frac{x - y}{2\sqrt{t-s}}, \tag{1.5.50}
\]

to estimate the integral in the right side of (1.5.47).

We set

\[
 Z_\varepsilon(t) = \int_0^t \|u^\varepsilon(s, \cdot)\|_{L^\infty} ds,
\]

and write (1.5.49) as

\[
 \frac{dZ_\varepsilon}{dt} \leq \|u_0\|_{L^\infty} + Ct + \frac{C}{\sqrt{\varepsilon}} Z_\varepsilon. \tag{1.5.51}
\]

Multiplying (1.5.51) by \(\exp(-Ct/\sqrt{\varepsilon})\) and integrating, keeping in mind that \(Z_\varepsilon(0) = 0\), gives

\[
 Z_\varepsilon(t) \leq \frac{\sqrt{\varepsilon}}{C} e^{Ct/\sqrt{\varepsilon}} (\|u_0\|_{L^\infty} + Ct). \tag{1.5.52}
\]
Using this bound in (1.5.49) gives the estimate
\[ \|u^\varepsilon(t, \cdot)\|_{L^\infty} \leq \left(\|u_0\|_{L^\infty} + Ct\right)\left(1 + e^{Ct/\sqrt{\varepsilon}}\right). \] (1.5.53)

Therefore, the $L^\infty$-norm of the solution can grow by at most a fixed factor over the time interval $[0, T\varepsilon]$. This estimate, together with the result of Exercise 1.5.8 allows us to restart the Cauchy problem on the time interval $[T\varepsilon, 2T\varepsilon]$, and then on $[2T\varepsilon, 3T\varepsilon]$, and so on, showing that the regularized problem (1.5.44) admits a global in time solution.

**Passing to the limit $\varepsilon \downarrow 0$**

A much more serious challenge than proving the global in time existence of $u^\varepsilon$ is to send $\varepsilon \downarrow 0$, and recover a smooth solution of the original equation (1.5.35) in the limit. Note that the upper bound (1.5.53) deteriorates very badly as $\varepsilon \downarrow 0$. Hence, we need to come up with much better bounds than that in order to pass to the limit $\varepsilon \downarrow 0$. To do this, we will obtain the Hölder estimates for $u^\varepsilon$ and its derivatives up to the second order in space and the first order in time, that will be independent of $\varepsilon$. The Ascoli-Arzelà theorem will then provide us with the compactness of the family $u^\varepsilon$, and allow us to pass to the limit along a subsequence and obtain a solution to (1.5.35).

**Exercise 1.5.9** Assume that there exists $\alpha \in (0, 1)$ such that, for all $\delta > 0$ and $T > \delta$, there is $C_\delta(T) > 0$, that is independent of $\varepsilon \in (0, 1)$, for which we have the following Hölder regularity estimates:
\[ \left| \frac{\partial}{\partial t} \left( u^\varepsilon(t, x) - u^\varepsilon(t', x') \right) \right| + \left| D_x^2 \left( u^\varepsilon(t, x) - u^\varepsilon(t', x') \right) \right| \leq C_\delta(T) \left( |t - t'|^{\alpha/2} + |x - x'|^\alpha \right), \] (1.5.54)
for all $t, t' \in [\delta, T]$ and $x, x' \in \mathbb{R}^n$, together with a uniform bound
\[ |u^\varepsilon(t, x)| \leq C(T), \text{ for all } 0 \leq t \leq T \text{ and all } x \in \mathbb{R}^n. \] (1.5.55)

Write down a complete proof that then there exists a subsequence $u^{\varepsilon_k}(t, x)$ that converges to a limit $u(t, x)$ as $k \to +\infty$, and, moreover, that $\nabla u^\varepsilon \to \nabla u$. In which space does the convergence take place? Show that the limit $u(t, x)$ is twice continuously differentiable in space, and once continuously differentiable in time, and is a solution to (1.5.35). For now, we leave open the question of why the limit satisfies the initial conditions as well.

This exercise gives us the road map to the construction of a solution to (1.5.35): we “only” need to establish the Hölder estimates (1.5.54) for the solutions to the approximate equation (1.5.44). We will use the following lemma, that is a slight generalization of the Gronwall lemma, and which is very useful in estimating the derivatives for the solutions of the parabolic equations.

**Lemma 1.5.10** Let $\varphi(t)$ be a nonnegative bounded function that satisfies, for all $0 \leq t \leq T$:
\[ \varphi(t) \leq \frac{a}{\sqrt{t}} + b \int_0^t \frac{\varphi(s)}{\sqrt{t - s}} ds. \] (1.5.56)
Then, for all $T > 0$, there is $C(T, b) > 0$ that depends on $T$ and $b$ but not on $\|\varphi\|_{L^\infty}$, such that
\[ \varphi(t) \leq \frac{C(T, b)a}{\sqrt{t}}, \text{ for all } 0 < t \leq T. \] (1.5.57)
Proof. First, note that we can write \( \varphi(t) = a\psi(t) \), leading to

\[
\psi(t) \leq \frac{1}{\sqrt{t}} + b \int_0^t \frac{\psi(s)}{\sqrt{t - s}} ds.
\] (1.5.58)

Then, iterating (1.5.58) we obtain

\[
\psi(t) \leq \sum_{k=0}^n I_n(t) + R_{n+1}(t),
\] (1.5.59)

for any \( n \geq 0 \), with

\[
I_{n+1}(t) = b \int_0^t \frac{I_n(s)}{\sqrt{t - s}} ds, \quad I_0(t) = \frac{1}{\sqrt{t}},
\] (1.5.60)

and

\[
R_{n+1}(t) = b \int_0^t \frac{R_n(s)}{\sqrt{t - s}}, \quad R_0(t) = \psi(t).
\] (1.5.61)

We claim that there exist a constant \( c > 0 \), and \( p > 1 \) so that

\[
I_n(t) \leq \frac{(ct)^{n/2}}{(n!)^{1/p}}.
\] (1.5.62)

Indeed, this bound holds for \( n = 0 \), and if it holds for \( I_n(t) \), then we have

\[
I_{n+1}(t) = b \int_0^t \frac{I_n(s)}{\sqrt{t - s}} ds \leq \frac{bc^{n/2}}{(n!)^{1/p}} \int_0^t \frac{s^{(n-1)/2} ds}{\sqrt{t - s}} = \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \int_0^1 \frac{\tau^{(n-1)/2} d\tau}{\sqrt{1 - \tau}}
\]

\[
\leq \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \left( \int_0^1 \tau^{3(n-1)/2} d\tau \right)^{1/3} \left( \int_0^1 \frac{d\tau}{(1 - \tau)^{3/4}} \right)^{2/3}
\]

\[
= \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \left( 3n/2 - 1/2 \right)^{1/3} \leq \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \left( n + 1 \right)^{1/3}.
\] (1.5.63)

We used above the Hölder inequality with exponents 3 and 3/2. Thus, the bound (1.5.62) holds with \( p = 3 \) and \( c = 16b^2 \).

As we assume that \( \varphi(t) \) is bounded, so is \( R_0(t) = \psi(t) \). This leads to a better bound for \( R_n(t) \) than for \( I_n(t) \): we claim that there exist a constant \( c > 0 \), and \( p > 1 \) so that

\[
R_n(t) \leq \frac{(ct)^{n/2}}{(n!)^{1/p}} \| \psi \|_{L^\infty}.
\] (1.5.64)

The computation is very similar to (1.5.63): we know that (1.5.64) holds for \( n = 0 \), and if it holds for some \( n \), then we have

\[
R_{n+1}(t) = b \int_0^t \frac{R_n(s)}{\sqrt{t - s}} ds \leq \frac{bc^{n/2}}{(n!)^{1/p}} \int_0^t \frac{s^{n/2} ds}{\sqrt{t - s}} = \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \int_0^1 \frac{\tau^{(n-1)/2} d\tau}{\sqrt{1 - \tau}}
\]

\[
\leq \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \left( \int_0^1 \tau^{3(n-1)/2} d\tau \right)^{1/3} \left( \int_0^1 \frac{d\tau}{(1 - \tau)^{3/4}} \right)^{2/3}
\]

\[
= \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \left( 3n/2 - 1/2 \right)^{1/3} \leq \frac{bc^{n/2}t^{(n+1)/2}}{(n!)^{1/p}} \left( n + 1 \right)^{1/3}.
\] (1.5.65)
Once again, we can take $p = 3$ and $c = 16b^2$. We conclude that

$$R_n(t) \to 0 \text{ as } n \to +\infty, \text{ uniformly on } [0, T].$$

(1.5.66)

Going back to (1.5.59), we see that

$$\varphi(t) \leq \frac{a}{\sqrt{t}} + a \sum_{n=1}^{\infty} I_n(t).$$

(1.5.67)

Now, the desired estimate (1.5.57) follows from (1.5.67) and (1.5.62). \[ \square \]

With the claim of Lemma 1.5.10 in hand, let us go back to the Duhamel formula (1.5.47)

$$u^\varepsilon(t, x) = e^{t\Delta} u_0(x) + \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} f(s, y, \nabla v^\varepsilon(s, y))dyds.$$

(1.5.68)

We first get a Hölder bound on $\nabla u^\varepsilon$. The maximum principle implies that

$$\|e^{t\Delta} u_0\|_{L^\infty} \leq \|u_0\|_{L^\infty},$$

(1.5.69)

and also that the gradient

$$\nabla v^\varepsilon = e^{t\Delta} \nabla u^\varepsilon,$$

satisfies the bound

$$\|\nabla v^\varepsilon(t, \cdot)\|_{L^\infty} \leq \|\nabla u^\varepsilon(t, \cdot)\|_{L^\infty}.$$  \hspace{1cm} (1.5.70)

We use these estimates, together with assumption (1.5.38) on the function $f(t, x, p)$, and the change of variables (1.5.50), in the Duhamel formula (1.5.68), leading to

$$\|u^\varepsilon(t, \cdot)\|_{L^\infty} \leq \|u_0\|_{L^\infty} + Ct + C \int_0^t \|\nabla u^\varepsilon(s, \cdot)\|_{L^\infty} ds.$$

(1.5.71)

The next step is to take the gradient of the Duhamel formula:

$$\nabla u^\varepsilon(t, x) = \nabla(e^{t\Delta} u_0(x)) + \nabla \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} f(s, y, \nabla v^\varepsilon(s, y))dyds$$

$$= \nabla(e^{t\Delta} u_0(x)) + \int_0^t \nabla \left[e^{(t-s)\Delta} f(s, \cdot, \nabla v^\varepsilon(s, \cdot))\right](x)ds.$$ \hspace{1cm} (1.5.72)

The first term in the right side is estimated as in (1.5.46):

$$\|\nabla(e^{t\Delta} u_0)\|_{L^\infty} \leq \frac{C}{\sqrt{t}} \|u_0\|_{L^\infty}.$$  \hspace{1cm} (1.5.73)

To bound the gradient of the integral term in (1.5.72), we note that (1.5.73), together with assumption (1.5.38) give

$$\|\nabla e^{(t-s)\Delta} f(s, \cdot, \nabla v^\varepsilon(s, \cdot))\|_{L^\infty} \leq \frac{C}{\sqrt{t-s}} \|f(s, \cdot, \nabla v^\varepsilon(s, \cdot))\|_{L^\infty} \leq \frac{C}{\sqrt{t-s}} (1 + \|\nabla v^\varepsilon(s, \cdot)\|_{L^\infty}).$$ \hspace{1cm} (1.5.74)
Using (1.5.70) once again and putting together (1.5.72), (1.5.73) and (1.5.74), leads to
\[ \| \nabla u^\varepsilon(t, \cdot) \|_{L^\infty} \leq \frac{C}{\sqrt{t}} \| u_0 \|_{L^\infty} + C \sqrt{t} + C \int_0^t \frac{\| \nabla u^\varepsilon(s, \cdot) \|_{L^\infty}}{\sqrt{t - s}} ds. \] (1.5.75)

Writing
\[ \frac{C}{\sqrt{t}} \| u_0 \|_{L^\infty} + C \sqrt{t} \leq \frac{C \| u_0 \|_{L^\infty} + CT}{\sqrt{t}}, \quad 0 \leq t \leq T, \]
we can put (1.5.75) into the form of (1.5.56). Lemma 1.5.10 implies then that there exists a constant \( C(T) > 0 \), independent of \( \varepsilon \), such that
\[ \| \nabla u^\varepsilon(t, \cdot) \|_{L^\infty} \leq \frac{C(T)}{\sqrt{t}}, \quad 0 < t \leq T. \] (1.5.76)

This bound, which is uniform in \( \varepsilon \in (0, 1) \), is absolutely crucial and allows us to proceed relatively effortlessly. Note that even though the right side of (1.5.76) blows up as \( t \downarrow 0 \), we cannot expect any better bound than (1.5.76) as we only assume that the initial condition \( u_0(x) \) is continuous and not necessarily differentiable.

The first simple observation is that using the estimate (1.5.76) in (1.5.71) gives a uniform bound on \( u^\varepsilon \) itself:
\[ \| u^\varepsilon(t, \cdot) \|_{L^\infty} \leq C(T), \quad 0 < t \leq T. \] (1.5.77)
This is the uniform bound (1.5.55) in Exercise 1.5.9. In other words, for \( t \in (\delta, \infty) \) for any \( \delta > 0 \), the family \( u^\varepsilon(t, \cdot) \) is uniformly bounded in the Sobolev space \( W^{1,\infty}(\mathbb{R}^n) \) – the space of \( L^\infty \) functions with gradients (in the sense of distributions) that are also \( L^\infty \) functions:
\[ \| u^\varepsilon(t, \cdot) \|_{W^{1,\infty}} \leq \frac{C(T)}{\sqrt{t}}, \quad 0 < t \leq T. \] (1.5.78)

The constant \( C(T) \) depends only on \( T \), the constant \( C_3 \) in (1.5.38) and \( \| u_0 \|_{L^\infty} \).

The uniform bound on the gradient in (1.5.76) seems a far cry from what we need in Exercise 1.5.9 – there, we require a Hölder estimate on the second derivatives in \( x \), and so far we only have a uniform bound on the first derivative. We do not even know yet that the first derivatives are Hölder continuous. Surprisingly, the end of the proof is actually not far off. Take some \( 1 \leq i \leq n \), and set
\[ z_i^\varepsilon(t, x) = \frac{\partial u^\varepsilon(t, x)}{\partial x_i}. \]
Note that such differentiation is perfectly legal since the functions \( u^\varepsilon \) are smooth. The equation for \( z_i^\varepsilon \) is (using, as usual, the summation convention for repeated indices)
\[ \partial_i z_i^\varepsilon - \Delta z_i^\varepsilon = \partial_i f(t, x, \nabla u^\varepsilon) + \partial_{ij} f(t, x, \nabla u^\varepsilon) \partial_{x_j} q_i^\varepsilon, \quad q_i^\varepsilon = \varepsilon^\Delta z_i^\varepsilon . \] (1.5.79)
We look at (1.5.79) as an equation for \( z_i^\varepsilon \), with a given function \( \nabla u^\varepsilon(t, x) \) that satisfies the already proved uniform bound
\[ \| \nabla u^\varepsilon(t, \cdot) \|_{L^\infty} \leq \frac{C(T)}{\sqrt{t}}, \quad 0 < t \leq T, \] (1.5.80)
that follows immediately from (1.5.76). Thus, (1.5.79) is of the form
\[ \partial_t z^\varepsilon_i - \Delta z^\varepsilon_i = G(t, x, \nabla q^\varepsilon_i), \quad q^\varepsilon_i = e^\varepsilon \Delta z^\varepsilon_i, \] (1.5.81)
with
\[ G(t, x, p) = \partial_x f(t, x, \nabla v^\varepsilon(t, x)) + \partial_{p_j} f(t, x, \nabla v^\varepsilon(t, x)) p_j. \] (1.5.82)
The function \( G(t, x, p) \) satisfies the assumptions on the nonlinearity \( f(t, x, p) \) stated at the beginning of this section – it is simply a linear function in the variable \( p \), and the gradient bound (1.5.80), together with the smoothness assumptions on \( f(t, x, p) \) in (1.5.39), and the Lipschitz estimate (1.5.37), implies that
\[ |G(t, x, p)| \leq C(T) \sqrt{1 + |p|}. \] (1.5.83)
Hence, on any time interval \([\delta, T]\) with \( \delta > 0 \), the function \( z^\varepsilon_i \) satisfies an equation of the type we have just analyzed for \( u^\varepsilon \), and our previous analysis shows that
\[ \|\nabla z^\varepsilon_i(t, \cdot)\|_{L^\infty} \leq \frac{C(T, \delta)}{\sqrt{t - \delta}}, \quad \delta < t \leq T. \] (1.5.84)

The constant \( C(T, \delta) \) depends on \( \delta > 0 \) because it depends on \( \|z^\varepsilon_i(\delta, \cdot)\|_{L^\infty} \) and because (1.5.83) produces an upper bound on \( G(t, x, p) \) for \( t > \delta \) that depends on \( \delta > 0 \). Rephrasing (1.5.84), we have the bound
\[ \|D^2 u^\varepsilon(t, \cdot)\|_{L^\infty} \leq \frac{C(T, \delta)}{\sqrt{t - \delta}}, \quad \delta < t \leq T, \] (1.5.85)
with a constant \( C(T, \delta) \) that depends on \( \delta > 0, T > 0 \) and \( \|u_0\|_{L^\infty} \).

This is almost what we need in (1.5.54) – we also need to show that \( D^2 u^\varepsilon \) are Hölder continuous, and deal with the time derivative and Hölder continuity in \( t \). With the information we have already obtained, we know that the right side of (1.5.81) is a uniformly bounded function, on any time interval \([\delta, T]\), with \( \delta > 0 \). Proposition 1.4.7 implies then immediately that \( \nabla z_i(t, x) \) is Hölder continuous in \( x \) and \( z_i(t, x) \) itself is Hölder continuous in \( t \) on the time interval \([2\delta, T]\), with bounds that do not depend on \( \varepsilon > 0 \). In addition, the uniform bound on \( \|D^2 u^\varepsilon\|_{L^\infty} \) and equation (1.5.44) itself imply a uniform bound on the time derivative:
\[ \|\partial_t u^\varepsilon(t, \cdot)\|_{L^\infty} \leq \frac{C(T, \delta)}{\sqrt{t - \delta}}, \quad \delta < t \leq T. \] (1.5.86)

To get a Hölder bound on the time derivative
\[ \zeta^\varepsilon(t, x) = \frac{\partial u^\varepsilon(t, x)}{\partial t}, \] (1.5.87)
we differentiate (1.5.44) in time to get the following equation
\[ \partial_t \zeta^\varepsilon - \Delta \zeta^\varepsilon = \partial_t f(t, x, \nabla v^\varepsilon) + \partial_{p_j} f(t, x, \nabla v^\varepsilon) \partial_x \eta^\varepsilon, \quad \eta^\varepsilon = e^\varepsilon \Delta \zeta^\varepsilon. \] (1.5.88)
This equation has the same form as equation (1.5.81) for \( z^\varepsilon_i(t, x) \). In addition, (1.5.86) gives an a priori bound for \( \zeta^\varepsilon(\delta, \cdot) \). Hence, arguing as above gives an analog of (1.5.84):
\[ \|\nabla \zeta^\varepsilon(t, \cdot)\|_{L^\infty} \leq \frac{C(T, \delta)}{\sqrt{t - \delta}}, \quad \delta < t \leq T. \] (1.5.89)
This allows us to bound $\nabla \eta^\varepsilon$, so that we can view (1.5.88) as an equation of the form

$$\partial_t \zeta^\varepsilon - \Delta \zeta^\varepsilon = F^\varepsilon(t,x),$$  \hspace{1cm} (1.5.90)

with a function $F(t,x)$ that satisfies a uniform bound

$$\|F^\varepsilon(t,\cdot)\|_{L^\infty} \leq \frac{C(T,\delta)}{\sqrt{t-\delta}}, \quad \delta < t \leq T.$$  \hspace{1cm} (1.5.91)

As we know that $\zeta^\varepsilon(t,x)$ is uniformly bounded by (1.5.86), Proposition 1.4.7 can be used again, this time to deduce that $\zeta^\varepsilon(t,x)$ is Hölder continuous in $t$ on any time interval $[\delta,T]$, with a bound that does not depend on $\varepsilon > 0$. Thus, (1.5.54) is finally proved.

Now, given any $\delta > 0$, Exercise 1.5.9 allows us to find a sequence $\varepsilon_k \to 0$ so that $u^{\varepsilon_k}$ converges to a limit $u(t,x)$ locally uniformly, and $\nabla u^{\varepsilon_k}$ converge to $\nabla u(t,x)$ on any time interval $[\delta,T]$. A standard diagonal argument allows us to pass to the limit on $(0,T)$. The limit is also twice continuously differentiable in $x$ and once continuously differentiable in $t$, and these derivatives themselves are Hölder continuous. Passing to the limit in (1.5.44)

$$u_t^\varepsilon - \Delta u^\varepsilon = f(t,x,\nabla v^\varepsilon), \quad v^\varepsilon = e^{\varepsilon \Delta} u^\varepsilon,$$  \hspace{1cm} (1.5.92)

leads to

$$u_t - \Delta u = f(t,x,\nabla u),$$  \hspace{1cm} (1.5.93)

as desired. In order to prove that $u(t,x)$ satisfies the initial condition, we go back to the Duhamel formula (1.5.47) to obtain, with the help of (1.5.80):

$$|u^\varepsilon(t,x) - e^{t\Delta} u_0(x)| \leq \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} |f(s,y,\nabla v^\varepsilon(s,y))| dy ds$$

$$\leq C(T) \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))} dy ds \leq C(T) \sqrt{t}.$$  \hspace{1cm} (1.5.94)

Passing to the limit $\varepsilon_k \to 0$ implies that $u(0,x) = u_0(x)$, so that the initial condition is satisfied.

**Exercise 1.5.11** Differentiate the equation for $u$ and iterate the above argument, showing that the solution is actually infinitely differentiable.

All that is left in the proof of Theorem 1.5.6 is to prove the uniqueness of a smooth solution. We will invoke the maximum principle again. Recall that we are looking for smooth solutions, so the difference $w = u_1 - u_2$ between any two solutions $u_1$ and $u_2$ simply satisfies an equation with a drift:

$$w_t - \Delta w = b(t,x) \cdot \nabla w,$$  \hspace{1cm} (1.5.95)

with a smooth drift $b(t,x)$ such that

$$f(x,\nabla u_1(t,x)) - f(x,\nabla u_2(t,x)) = b(t,x) \cdot [\nabla u_1(t,x) - \nabla u_2(t,x)].$$

As $w(0,x) \equiv 0$, the comparison principle of Theorem 1.3.3 implies that $w(t,x) \equiv 0$ and $u_1 \equiv u_2$. This completes the proof of Theorem 1.5.6. □

**Exercise 1.5.12** Prove that, if $u_0$ is smooth, then smoothness holds up to $t = 0$. Prove that equation (1.5.35) holds up to $t = 0$, that is:

$$u_t(0,x) = \Delta u_0(x) + f(x,\nabla u_0(x)).$$
1.5.3 Applications to linear equations with a drift

Let us now discuss how the above results can be made more quantitative for linear equations of the form

\[ u_t = \Delta u + b_j(t, x) \frac{\partial u}{\partial x_j} + c(t, x)u, \quad t > 0, \quad x \in \mathbb{R}^n, \]  

(1.5.96)

with smooth coefficients \( b_j(t, x) \) and \( c(t, x) \). We recall, once again, that the repeated indices are summed. When \( c(t, x) = 0 \), this equation has the form of the non-linear equation (1.5.35) considered in the previous section. In particular, Theorem 1.5.6 implies immediately that given any initial condition \( u_0(x) \) that is a bounded continuous function, the equation (1.5.96) has a unique solution \( u(t, x) \) that is infinitely differentiable for all \( t > 0 \) and \( x \in \mathbb{R}^n \) such that \( u(0, x) = u_0(x) \). The same result holds, with essentially an identical proof when \( c(t, x) \) is smooth.

Exercise 1.5.13 Extend the result of Theorem 1.5.6 to equations of the form (1.5.96) with smooth coefficients \( b_j(t, x) \) and \( c(t, x) \).

A more important claim is that the quantitative regularity results formulated in Proposition 1.4.3 for the linear heat equation in the whole space also hold essentially verbatim for (1.5.96).

1.6 A survival kit in the jungle of regularity

In our noble endeavor to carry out as explicit computations as possible, we have not touched the question of regularity of solutions to inhomogeneous equations where the diffusivity can be not constant. An inhomogeneous drift has been treated in Section 1.5.2, We address here the following question: given a linear inhomogeneous equation of the form

\[ u_t - a_{ij}(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(t, x) \frac{\partial u}{\partial x_j} + c(t, x) = f(t, x), \]  

(1.6.1)

the coefficients \( a_{ij} \), \( b_j \) and the right side \( f \) having a certain given degree of smoothness, what is the best regularity that one may expect from \( u \)? The question is a little different from what we did for the nonlinear equations, where one would first prove a certain, possibly small, amount of regularity, in the hope that this would be sufficient for a bootstrap argument leading to a much better regularity than in one iteration step. The answer to the question of maximal regularity is, in a nutshell: if the coefficients have a little bit of continuity, such as the Hölder continuity, then the derivatives \( u_t \) and \( D^2 u \) have the same regularity as \( f \). This, however, is true up to some painful exceptions: continuity for \( f \) will not entail, in general, the continuity of \( u_t \) and \( D^2 u \). This is exactly what we have seen in Section 1.4 for the forced heat equation, so this should not come as a surprise to the reader.

The question of the maximal regularity for linear parabolic equations has a certain degree of maturity, an interested reader should consult [89] to admire the breadth, beauty and technical complexity of the available results. Our goal here is much more modest: we will explain why the Hölder continuity of \( f \) will entail the Hölder continuity of \( u_t \) and \( D^2 u \) – the result we have already seen for the heat equation using the explicit computations with the Duhamel term.
When $a_{ij}(t, x) = \delta_{ij}$ (the Kronecker symbol), the second order term in (1.6.1) is the Laplacian, and our work was already almost done in Theorem 1.5.6 even though we have not formulated the precise Hölder estimates on the solution in the case when equation (1.5.35) happens to be linear. Nevertheless, the reader should be able to extract them from the proof of that theorem and discover a version of Proposition 1.4.20 for an equation of the form

$$u_t - \Delta u + b_j(t, x) \frac{\partial u}{\partial x_j} + c(t, x)u = f(t, x),$$

(1.6.2)

with smooth coefficients $b_j(t, x)$, $j = 1, \ldots, n$ and $c(t, x)$, and $f(t, x) \in C_t^{\alpha/2} C_x^\alpha$. We will try to convince the reader, without giving the full details of all the proofs, that this carries over to variable diffusion coefficients, and, importantly, to problems with boundary conditions. Our main message here is that all the ideas necessary for the various proofs have already been displayed, and that ”only” technical complexity and dexterity are involved. Our discussion follows Chapter 4 of [89], which presents various results with much more details. Let us emphasize again that in this section, we will only give a sketch of the proofs, and sometimes we will not state the results in a formal way.

When the diffusion coefficients are not continuous, but merely bounded, the methods described in this chapter break down. Chapter ??, based on the Nash inequality, explains to some extent how to deal with such problems by a very different approach.

The Cauchy problem for the inhomogeneous coefficients

We have all the ideas to understand the first big piece of this section, the Cauchy problem for the parabolic equations with variable coefficients in the whole space, without any forcing:

$$u_t - a_{ij}(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(t, x) \frac{\partial u}{\partial x_j} + c(t, x)u = 0, \quad t > 0, \quad x \in \mathbb{R}^n,$$

$$u(0, x) = u_0(x), \quad x \in \mathbb{R}^n.$$

(1.6.3)

We make the following assumptions on the coefficients: first, they are sufficiently regular – the functions $(a_{ij}(t, x))_{1 \leq i, j \leq N}$, $(b_j(t, x))_{1 \leq j \leq N}$ and $c(t, x)$, all $\alpha$-Hölder continuous over $[0, T] \times \mathbb{R}^n$. Second, we make the ellipticity assumption, generalizing (1.4.82): there exist $\lambda > 0$ and $\Lambda > 0$ so that for any vector $\xi \in \mathbb{R}^n$ and any $x \in \mathbb{R}^n$ we have

$$\lambda|\xi|^2 \leq a_{ij}(t, x)\xi_i\xi_j \leq \Lambda|\xi|^2.$$

(1.6.4)

We assume that the initial condition $u_0(x)$ is a continuous function – this assumption can be very much weakened but we do not focus on it right now.

**Theorem 1.6.1** The Cauchy problem (1.6.3) has a unique solution $u(t, x)$, whose Hölder norm on the sets of the form $[\varepsilon, T] \times \mathbb{R}^n$ is controlled by the $L^\infty$ norm of $u_0$.

The statement of this theorem is deliberately vague – the correct statement should become clear to the reader after we outline the ideas of the proof.

**Exercise 1.6.2** Show that the uniqueness of the solution is an immediate consequence of the maximum principle.
Thus, the main issue is the construction of a solution with the desired regularity. The idea is to construct the fundamental solution of (1.6.3), that is, the solution \( E(t,s,x,y) \) of (1.6.3) on the time interval \( s \leq t \leq T \), instead of \( 0 \leq t \leq T \):

\[
\partial_t E - a_{ij}(t,x) \frac{\partial E}{\partial x_i} x_j + b_j(t,x) \frac{\partial E}{\partial x_j} + c(t,x) E = 0, \quad t > s, \ x \in \mathbb{R}^n, \tag{1.6.5}
\]

with the initial condition

\[
E(t = s, s, x, y) = \delta(x - y), \tag{1.6.6}
\]

the Dirac mass at \( x = y \). The solution of (1.6.3) can then be written as

\[
u(t,x) = \int_{\mathbb{R}^n} E(t,0,x,y)u_0(y)dy. \tag{1.6.7}\]

If can show that \( E(t,s,x,y) \) is smooth enough (at least away from \( t = s \)), \( u(t,x) \) will satisfy the desired estimates as well – they can be obtained by differentiating or taking partial differences in (1.6.7). Note that regularity of \( E \) for \( t > s \) is a very strong property: the initial condition in (1.6.6) at \( t = s \) is a measure – and we need to show that for all \( t > s \) the solution is actually a smooth function. On the other hand, this is exactly what happens for the heat equation

\[ u_t = \Delta u, \]

where the fundamental solution is

\[ E(t,s,x,y) = \frac{1}{(4\pi(t-s))^{n/2}} e^{-(x-y)^2/(4(t-s))}, \]

and is smooth for all \( t > s \).

**Exercise 1.6.3** Go back to the equation

\[ u_t - u_{xx} + u_y = 0. \]

considered in Exercise 1.4.22. Show that its fundamental solution is not a smooth function in the \( y \)-variable. Thus, the ellipticity condition is important for this property.

The understanding of the regularity of the solutions of the Cauchy problem is also a key to the inhomogeneous problem because of the Duhamel principle.

**Exercise 1.6.4** Let \( f(t,x) \) be a Hölder-continuous function over \([0,T] \times \mathbb{R}^n\). Use the Duhamel principle to write down the solution of

\[
u_t - a_{ij}(t,x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(t,x) \frac{\partial u}{\partial x_j} + c(t,x) u = f(t,x), \quad t > 0, \ x \in \mathbb{R}^n, \]

\[ u(0,x) = u_0(x), \quad x \in \mathbb{R}^n, \tag{1.6.8} \]

in terms of \( E(t,s,x,y) \).
Thus, everything boils down to constructing the fundamental solution \( E(t, s, x, y) \), and a way to do it is via the parametrix method. Let us set \( b_j = c = 0 \) – this does not affect the essence of the arguments but simplifies the notation. The philosophy is that the possible singularities of \( E(t, s, x, y) \) are localized at \( t = s \) and \( x = y \) (as for the heat equation). Therefore, in order to capture the singularities of \( E(t, s, x, y) \) we may try to simply freeze the coefficients in the equation at \( t = s \) and \( x = y \), and compare \( E(t, s, x, y) \) to the fundamental solution \( E_0(t, s, x, y) \) of the resulting equation:

\[
\begin{align*}
\partial_t E_0 - a_{ij}(s, y) \frac{\partial^2 E_0}{\partial x_i \partial x_j} &= 0, \quad t > s, \quad x \in \mathbb{R}^n, \\
E_0(t = s, x) &= \delta(x - y), \quad x \in \mathbb{R}^n.
\end{align*}
\]

There is no reason to expect the two fundamental solutions to be close – they satisfy different equations. Rather, the expectation is that that \( E \) will be a smooth perturbation of \( E_0 \) – and, since \( E_0 \) solves an equation with constant coefficients (remember that \( s \) and \( y \) are fixed here), we may compute it exactly.

To this end, let us write the equation for \( E(t, s, x, y) \) as

\[
\begin{align*}
\partial_t E - a_{ij}(s, y) \frac{\partial^2 E}{\partial x_i \partial x_j} &= F(t, x), \quad t > s, \quad x \in \mathbb{R}^n, \\
E(t = s, x) &= \delta(x - y), \quad x \in \mathbb{R}^n,
\end{align*}
\]

with the right side

\[
F(t, x, y) = (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 E}{\partial x_i \partial x_j}.
\]

The difference

\[
R_0 = E - E_0
\]

satisfies

\[
\begin{align*}
\partial_t R_0 - a_{ij}(s, y) \frac{\partial^2 R_0}{\partial x_i \partial x_j} &= (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 E_0}{\partial x_i \partial x_j} + F_0(t, x), \quad t > s, \\
F_0(t, x) &= (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 R_0}{\partial x_i \partial x_j}.
\end{align*}
\]

Let us further decompose

\[
R_0 = E_1 + R_1.
\]

Here, \( E_1 \) is the solution of

\[
\begin{align*}
\partial_t E_1 - a_{ij}(s, y) \frac{\partial^2 E_1}{\partial x_i \partial x_j} &= (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 E_0}{\partial x_i \partial x_j}, \quad t > s, \\
E_1(t = s, x) &= 0.
\end{align*}
\]

The remainder \( R_1 \) solves

\[
\begin{align*}
\partial_t R_1 - a_{ij}(s, y) \frac{\partial^2 R_1}{\partial x_i \partial x_j} &= (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 E_1}{\partial x_i \partial x_j} + F_1(t, x), \quad t > s,
\end{align*}
\]

50
with \( R_1(t = s, x) = 0 \), and

\[
F_1(t, x) = (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 R_1}{\partial x_i \partial x_j}.
\] (1.6.16)

Equation (1.6.14) for \( E_1 \) is a forced parabolic equation with constant coefficients – as we have seen, its solutions behave exactly like those of the standard heat equation with a forcing, except for rotations and dilations. We may assume without loss of generality that \( a_{ij}(s, y) = \delta_{ij} \), so that the reference fundamental solution is

\[
E_0(t, s, x, y) = \frac{1}{(4\pi(t - s))^{n/2}} e^{-(x-y)^2/(4(t-s))},
\] (1.6.17)

and (1.6.14) is simply a forced heat equation:

\[
\partial_t E_1 - \Delta E_1 = [a_{ij}(t, x) - \delta_{ij}] \frac{\partial^2 E_0(t, s, x, y)}{\partial x_i \partial x_j}, \quad t > s, \quad x \in \mathbb{R}^n.
\] (1.6.18)

The functions \( a_{ij}(t, x) \) Hölder continuous, with \( a_{ij}(s, y) = \delta_{ij} \). The regularity of \( E_1 \) can be approached by the tools of the previous sections – after all, (1.6.14) is just another forced heat equation! The next exercise may be useful for understanding what is going on.

**Exercise 1.6.5** Consider, instead of (1.6.14) the solution of

\[
\partial_t z - \Delta z = \frac{\partial^2 E_0(t, s, x, y)}{\partial x_i \partial x_j}, \quad t > s, \quad x \in \mathbb{R}^n,
\] (1.6.19)

with \( z(t = s, x) = 0 \). How does its regularity compare to that of \( E_0 \)? Now, what can you say about the regularity of the solution to (1.6.18), how does the factor \( [a_{ij}(t, x) - \delta_{ij}] \) help to make \( E_1 \) more regular than \( z \)? In which sense is \( E_1 \) more regular than \( E_0 \)?

With this understanding in hand, one may consider the iterative process: write

\[ R_1 = E_2 + R_2, \]

with \( E_2 \) the solution of

\[
\partial_t E_2 - a_{ij}(s, y) \frac{\partial^2 E_2}{\partial x_i \partial x_j} = (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 E_1}{\partial x_i \partial x_j}, \quad t > s,
\] (1.6.20)

with \( E_2(t = s, x) = 0 \), and \( R_2 \) the solution of

\[
\partial_t R_2 - a_{ij}(s, y) \frac{\partial^2 R_2}{\partial x_i \partial x_j} = (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 E_2}{\partial x_i \partial x_j} + F_2(t, x), \quad t > s,
\] (1.6.21)

with \( R_2(t = s, x) = 0 \), and

\[
F_2(t, x) = (a_{ij}(t, x) - a_{ij}(s, y)) \frac{\partial^2 R_2}{\partial x_i \partial x_j}.
\] (1.6.22)
Continuing this process, we have a representation for $E(t, s, x, y)$ as
\[
E = E_0 + E_1 + \cdots + E_n + R_n,
\] (1.6.23)
with each next term $E_j$ more regular than $E_0, \ldots, E_{j-1}$. Regularity of all $E_j$ can be inferred as in Exercise 1.6.5. One needs, of course, also to estimate the remainder $R_n$ to obtain a "true theorem" but we leave this out of this chapter, to keep the presentation short. An interested reader should consult the aforementioned references for full details. We do, however, offer the reader another (non-trivial) exercise.

**Exercise 1.6.6** Prove that $E(s, t, x, y)$ obeys Gaussian estimates of the form:
\[
m e^{-|x-y|^2/dt} \leq E(s, t, x, y) \leq M e^{-|x-y|^2/dt},
\]
for all $0 < s < t, T$ and $x, y \in \mathbb{R}^n$. The constants $m$ and $M$, unfortunately, depend very much on $T$; however the constants $d$ and $D$ do not.

**Interior regularity**

So far, we have considered parabolic problems in the whole space $\mathbb{R}^n$, without any boundaries. One of the miracles of the second order diffusion equations is that the regularity properties are **local**. That is, the regularity of the solutions in a given region only depends on how regular the coefficients are in a slightly larger region. Consider, again, the inhomogeneous parabolic equation
\[
u_t - a_{ij}(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(t, x) \frac{\partial u}{\partial x_j} + c(t, x)u = f(t, x), \quad t > 0,
\] (1.6.24)
and assume that the coefficients $a_{ij}(t, x), b_j(t, x)$ and $c(t, x)$, and forcing $f(t, x)$, are $\alpha$-Hölder in $S = [0, T] \times B_R(x_0)$. It turns out that the derivatives $D^2u(t, x)$ and $\partial_t u(t, x)$ are $\alpha$-Hölder in a smaller set of the form $S = [\varepsilon, T] \times B_{(1-\varepsilon)R}(x_0)$, for any $\varepsilon > 0$. The most important point is that the Hölder norm of $u$ in $S$ is controlled only by $\varepsilon, R$, and the Hölder norms of the coefficients and the $L^\infty$ bound of $u$, both inside the larger set $S$. Note that the Hölder estimates on $u$ in terms of the $L^\infty$-norm of $u$ over $S$ do not hold in the original set $S$, we need a small margin, going down to the smaller set $S_\varepsilon$. This is very similar to what happens for the heat equation: the bounded solution to
\[
u_t = \Delta u, \quad t > 0, \quad x \in \mathbb{R}^n,
\] (1.6.25)
with an initial condition $u(0, x) = \nabla u$ satisfies a bound
\[
\|\nabla u(t, \cdot)\|_{L^\infty} \leq \frac{C}{\sqrt{t}} \|u_0\|_{L^\infty},
\] (1.6.26)
that gives information only for $t > 0$ – this is exactly the margin we have discussed above.
Exercise 1.6.7 Prove this fact. One standard way to do it is to pick a nonnegative and smooth function \( \gamma(x) \), equal to 1 in \( B_{R/2}(x_0) \) and 0 outside of \( B_R(x) \), and to write down an equation for \( v(t, x) = \gamma(x)u(t, x) \). Note that this equation is now posed on \((0, T] \times \mathbb{R}^n\), and that the whole space theory can be applied. The computations should be, at times cumbersome. If in doubt, consult [56]. Looking ahead, we will use this strategy in the proof of Proposition 1.7.10 in Section 1.7 below, so the reader may find it helpful to read this proof now.

Regularity up to the boundary

Specifying the Dirichlet boundary conditions allows to get rid of this small margin, and this is the last issue that we are going to discuss in this section. Let us consider equation (1.6.24), posed this time in \((0, T] \times \Omega\), where \( \Omega \) is bounded smooth open subset of \( \mathbb{R}^n \). As a side remark, it is not crucial that \( \Omega \) be bounded. However, if \( \Omega \) is unbounded, we should ask its boundary to oscillate not too much at infinity. Let us supplement (1.6.24) by an initial condition 
\[
 u(0, x) = u_0(x) \quad \text{in } \Omega,
\]
with a continuous function \( u_0 \), and the Dirichlet boundary condition
\[
 u(t, x) = 0 \quad \text{for } 0 \leq t \leq T \quad \text{and } x \in \partial\Omega. \tag{1.6.27}
\]

Theorem 1.6.8 Assume \( a_{ij}(t, x), b_j(t, x), c(t, x), \) and \( f(t, x) \) are \( \alpha \)-Hölder in \((0, T] \times \overline{\Omega} \) – note that, here, we do need the closure of \( \Omega \). The above initial-boundary value problem has a unique solution \( u(t, x) \) such that \( D^2u(t, x) \) and \( \partial_t u(t, x) \) are \( \alpha \)-Hölder in \([\varepsilon, T] \times \overline{\Omega} \), with their Hölder norms controlled by the \( L^\infty \) bound of \( u_0 \), and the Hölder norms of the coefficients and \( f \).

The way to prove this result parallels the way we followed to establish Theorem 1.6.1. First, we write down an explicit solution on a model situation. Then, we prove the regularity in the presence of a Hölder forcing in the model problem. Once this is done, we turn to general constant coefficients. Then, we do the parametrix method on the model situation. Finally, we localize the problem and reduce it to the model situation.

Let us be more explicit. The model situation is the heat equation in a half space
\[
 \Omega_n = \mathbb{R}^n_+ := \{ x = (x_1, \ldots, x_n) \in \mathbb{R}^n : x_n > 0 \}.
\]
Setting \( x' = (x_1, \ldots, x_{n-1}) \), we easily obtain the solution of the initial boundary value problem
\[
 u_t - \Delta u = 0, \quad t > 0, \quad x \in \Omega_n, \tag{1.6.28}
\]
\[
 u(t, x', 0) = 0, \quad u(0, x) = u_0(x),
\]
as
\[
 u(t, x) = \int_{\mathbb{R}^n} E_0(t, x, y)u_0(y)dy, \tag{1.6.29}
\]
with the fundamental solution
\[
 E_0(t, x, y) = \frac{e^{-(x' - y')^2/4t}}{(4\pi t)^{n/2}} \left( e^{-(x_n - y_n)^2/4t} - e^{-(x_n + y_n)^2/4t} \right). \tag{1.6.30}
\]
Let us now generalize step by step: for an equation with a constant drift

\[ u_t - \Delta u + b_j \partial_{x_j} u = 0, \quad t > 0, \ x \in \Omega_n, \quad (1.6.31) \]

the change of unknowns \( u(t, x) = e^{\frac{x \cdot b_n}{2}} v(t, x) \) transforms the equation into

\[ v_t - \Delta v + b_j \partial_{x_j} v - \frac{b_n^2}{4} v = 0, \quad t > 0, \ x \in \Omega_n. \quad (1.6.32) \]

Thus, the fundamental solution, for (1.6.31) is

\[ E(t, x, y) = e^{\frac{t \partial_{B'}^2}{4} - \frac{B_{n-1}^2}{2}} E_0(t, x - tB', y), \quad B' = (b_1, \ldots, B_{n-1}, 0). \quad (1.6.33) \]

For an equation of the form

\[ u_t - a_{ij} \partial_{x_i} \partial_{x_j} u = 0, \quad t > 0, \ x \in \Omega_n, \quad (1.6.34) \]

with a constant positive-definite diffusivity matrix \( a_{ij} \), we use the fact that the function

\[ u(t, x) = v(t, \sqrt{A^{-1}}x), \]

with \( v(t, x) \) a solution of the heat equation

\[ v_t = \Delta v, \]

solves (1.6.34). For an equation mixing the two sets of coefficients, one only has to compose the transformations. At that point, one can, with a nontrivial amount of computations, prove the desired regularity for the solutions of

\[ u_t - a_{ij} \partial_{x_i} \partial_{x_j} u + b_j \partial_{x_j} u + cu = f(t, x) \quad (1.6.35) \]

with constant coefficients, and the Dirichlet boundary conditions on \( \partial \Omega_n \). Then, one can use the parametrix method to obtain the result for general inhomogeneous coefficients. This is how one proves Theorem 1.6.8 for \( \Omega_n = \mathbb{R}^n \).

How does one pass to a general \( \Omega \)? Unfortunately, the work is not at all finished yet. One still has to prove a local version of the already proved theorem in \( \Omega_n \), in the spirit of the local regularity in \( \mathbb{R}^n \), up to the fact that we must not avoid the boundary. Once this is done, consider a general \( \Omega \). Cover its boundary \( \partial \Omega \) with balls such that, in each of them, \( \partial \Omega \) is a graph in a suitable coordinate system. By using this new coordinate system, one retrieves an equation of the form (1.6.8), and one has to prove that the diffusion coefficients satisfy a coercivity inequality. At this point, maximal regularity for the Dirichlet problem is proved.

Of course, all kinds of local versions (that is, versions of Theorem 1.6.8 where the coefficients are \( \alpha \)-Hölder only in a part of \( \Omega_n \)) are available. Also, most of the above material is valid for the Neumann boundary conditions

\[ \partial_\nu u = 0 \text{ on } \partial \Omega, \]

or Robin boundary conditions

\[ \partial_\nu u + \gamma(t, x)u = 0 \text{ on } \partial \Omega. \]

We encourage the reader who might still be interested in the subject to try to produce the full proofs, with an occasional help from the books we have mentioned.
The Harnack inequalities

We will only touch here on the Harnack inequalities, a very deep and involved topic of parabolic equations. In a nutshell, the Harnack inequalities allow to control the infimum of a positive solution of a parabolic equation by a universal fraction of its maximum, modulo a time shift. They provide one possible, and very beautiful, path to prove regularity, but we will ignore this aspect here. They are also mostly responsible for the behaviors that are very specific to the diffusion equations, as will be seen in the next section.

We are going to prove what is, in a sense, a “poor man’s” version. It is not as scale invariant as one would wish, and uses the regularity theory instead of proving it. It is, however, suited to what we wish to do, and already gives a good account of what is going on. Consider our favorite equation

$$u_t - \Delta u + b_j(t, x) \frac{\partial u}{\partial x_j} + c(t, x) u = 0, \quad (1.6.36)$$

with smooth coefficients $b_j$ and $c$, posed for $t \in (0, T)$, and $x \in B_{R+1}(0)$. We stress that the variable smooth diffusion coefficients could be put in the picture.

**Theorem 1.6.9** Let $u(t, x) \geq 0$ be a non-negative bounded solution of (1.6.36) for $0 \leq t \leq T$ and $x \in B_{R+1}(0)$, and assume that for all $t \in [0, T]$:

$$\sup_{|x| \leq R} u(t, x) \leq k_2, \quad \sup_{|x| \leq R} u(t, x) \geq k_1. \quad (1.6.37)$$

There is a constant $h_R > 0$, that does not depend on $T$, but that depends on $k_1$ and $k_2$, such that, for all $t \in [1, T]$:

$$h_R \leq \inf_{|x| \leq R} u(t, x). \quad (1.6.38)$$

**Proof.** The proof is by contradiction. Assume that there exists a sequence $u_n$ of the solutions of (1.6.36) satisfying (1.6.37), and $t_n \in [1, T]$, and $x_n \in B_R(0)$, such that

$$\lim_{n \to +\infty} u_n(t_n, x_n) = 0. \quad (1.6.39)$$

Up to a possible extraction of a subsequence, we may assume that

$$t_n \to t_\infty \in [1, T] \text{ and } x_n \to x_\infty \in B_R(0).$$

The Hölder estimates on $u_n$ and its derivatives in Theorem 1.6.8 together with the Ascoli-Arzela theorem, imply that the sequence $u_n$ is relatively compact in $C^2([t_\infty - 1/2] \times B_{R+1/2}(0))$. Hence, again, after a possible extraction of a subsequence, we may assume that $u_n$ converges to $u_\infty \in C^2([t_\infty - 1/2, T] \times B_{R+1/2}(0))$, together with its first two derivatives in $x$ and the first derivatives in $t$. Thus, the limit $u_\infty(t, x)$ satisfies (1.6.36) for $t_\infty - 1/2 \leq t \leq T$, and $x \in B_{R+1/2}(0)$, and is non-negative. It also satisfies the bounds in (1.6.37), hence it is not identically equal to zero. Moreover it satisfies $u_\infty(t_\infty, x_\infty) = 0$. This contradicts the strong maximum principle. $\square$
1.7 The principal eigenvalue for elliptic operators and the Krein-Rutman theorem

One consequence of the strong maximum principle is the existence of a positive eigenfunction for an elliptic operator in a bounded domain with the Dirichlet or Neumann boundary conditions. Such eigenfunction necessarily corresponds to the eigenvalue with the smallest real part. A slightly different way to put it is that the strong maximum principle makes the Krein-Rutman Theorem applicable, which in turn, implies the existence of such eigenfunction. In this section, we will prove this theorem in the context of parabolic operators with time periodic coefficients. We then deduce, in an easy way, some standard properties of the principal elliptic eigenvalue.

1.7.1 The periodic principal eigenvalue

The maximum principle for elliptic and parabolic problems has a beautiful connection to the eigenvalue problems, which also allows to extend it to operators with a zero-order term. We will first consider the periodic eigenvalue problems, that is, elliptic equations where the coefficients are 1-periodic in every direction in $\mathbb{R}^n$, and the sought for solutions are all 1-periodic in $\mathbb{R}^n$. It would, of course, be easy to deduce, by dilating the coordinates, the same results for coefficients with general periods $T_1, \ldots, T_n$ in the directions $e_1, \ldots, e_n$. We will consider operators of the form

$$\mathcal{L}u(x) = -\Delta u + b_j(x) \frac{\partial u}{\partial x_j} + c(x)u, \quad (1.7.1)$$

with bounded, smooth and 1-periodic coefficients $b_j(x)$ and $c(x)$. We could also consider more general operators of the form

$$\mathcal{L}u(x) = -a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(x) \frac{\partial u}{\partial x_j} + c(x)u,$$

with uniformly elliptic (and 1-periodic), and regular coefficients $a_{ij}$, with the help of the elliptic regularity theory. This will not, however, be needed for our purposes. In order to avoid repeating that the coefficients and the solutions are 1-periodic, we will just say that $x \in \mathbb{T}^n$, the $n$-dimensional unit torus.

The key spectral property of the operator $\mathcal{L}$ comes from the comparison principle. To this end, let us recall the Krein-Rutman theorem. It says that if $M$ is a compact operator in a strongly ordered Banach space $X$ (that is, there is a solid cone $K$ which serves for defining an order relation: $u \leq v$ iff $v - u \in K$), that preserves $K$: $Mu \in K$ for all $u \in K$, and maps the boundary of $K$ into its interior, then $M$ has an eigenfunction $\phi$ that lies in this cone:

$$M\phi = \lambda \phi. \quad (1.7.2)$$

Moreover, the corresponding eigenvalue $\lambda$ has the largest real part of all eigenvalues of the operator $M$. The classical reference [46] has a nice and clear presentation of this theorem but one can find it in other textbooks, as well.
How can we apply this theorem to the elliptic operators? The operator \( L \) given by (1.7.1) is not compact, nor does it preserve any interesting cone. However, let us assume momentarily that \( c(x) \) is continuous and \( c(x) > 0 \) for all \( x \in \mathbb{T}^n \). Then the problem

\[
Lu = f, \quad x \in \mathbb{T}^n
\]

has a unique solution, and, in addition, if \( f(x) \geq 0 \) and \( f \not\equiv 0 \), then \( u(x) > 0 \) for all \( x \in \mathbb{T}^n \). Indeed, let \( v(t, x) \) be the solution of the initial value problem

\[
v_t + Lv = 0, \quad t > 0, \quad x \in \mathbb{T}^n,
\]

with \( v(0, x) = f(x) \). The comparison principle implies a uniform upper bound

\[
|v(t, x)| \leq e^{-\bar{c}t} \|f\|_{L^\infty},
\]

with

\[
\bar{c} = \inf_{x \in \mathbb{T}^n} c(x) > 0.
\]

This allows us to define

\[
u(x) = \int_0^\infty v(t, x) x.
\]

**Exercise 1.7.1** Verify that if \( c(x) > 0 \) for all \( x \in \mathbb{T}^n \), then \( u(x) \) given by (1.7.7) is a solution to (1.7.3). Use the maximum principle to show that (1.7.3) has a unique solution.

This means that we may define the inverse operator \( M = L^{-1} \). This operator preserves the cone of the positive functions, and maps its boundary (non-negative functions that vanish somewhere in \( \Omega \)) into its interior – this is a consequence of the strong maximum principle that holds if \( c(x) > 0 \). In addition, \( M \) is a compact operator from \( C(\mathbb{T}^n) \) to itself. Hence, the inverse operator satisfies the assumptions of the Krein-Rutman theorem.

**Exercise 1.7.2** Compactness of the inverse \( M \) follows from the elliptic regularity estimates. One way to convince yourself of this fact is to consult Evans [56]. Another way is to go back to Theorem 1.5.6, use it to obtain the Hölder regularity estimates on \( v(t, x) \), and translate them in terms of \( u(x) \) to show that, if \( f \) is continuous, then \( \nabla u \) is \( \alpha \)-Hölder continuous, for all \( \alpha \in (0, 1) \). The Arzela-Ascoli theorem implies then compactness of \( M \). Hint: be careful about the regularity of \( v(t, x) \) as \( t \downarrow 0 \).

Thus, there exists a positive function \( f \) and \( \mu \in \mathbb{R} \) so that the function \( u = \mu f \) satisfies (1.7.3). Positivity of \( f \) implies that the solution of (1.7.3) is also positive, hence \( \mu > 0 \). As \( \mu \) is the eigenvalue of \( L^{-1} \) with the largest real part, \( \lambda = \mu^{-1} \) is the eigenvalue of \( L \) with the smallest real part. In particular, it follows that all eigenvalues \( \lambda_k \) of the operator \( L \) have a positive real part.

If the assumption \( c(x) \geq 0 \) does not hold, we may take \( K > \|c\|_{L^\infty} \), and consider the operator

\[
L'u = Lu + Ku.
\]
The zero-order coefficient of $L'$ is

$$c'(x) = c(x) + K \geq 0.$$  

Hence, we may apply the previous argument to the operator $L'$ and conclude that $L'$ has an eigenvalue $\mu_1$ that corresponds to a positive eigenfunction, and has the smallest real part among all eigenvalues of $L'$. The same is true for the operator $L$, with the eigenvalue

$$\lambda_1 = \mu_1 - K.$$  

We say that $\lambda_1$ is the principal periodic eigenvalue of the operator $L$.

1.7.2 The Krein-Rutman theorem: the periodic parabolic case

As promised, we will prove the Krein-Rutman Theorem in the context of the periodic eigenvalue problems. Our starting point will be a slightly more general problem with time-periodic coefficients:

$$u_t - \Delta u + b_j(t,x) \frac{\partial u}{\partial x_j} + c(t,x)u = 0, \quad x \in \mathbb{T}^n. \quad (1.7.8)$$  

Here, the coefficients $b_j(t,x)$ and $c(t,x)$ are smooth, 1-periodic in $x$ and $T$-periodic in $t$. Let $u(t,x)$ be the solution of the Cauchy problem for (1.7.8), with a 1-periodic, continuous initial condition

$$u(t,x) = u_0(x). \quad (1.7.9)$$  

We define the "time $T$" operator $S_T$ as

$$[S_Tu_0](x) = u(T,x). \quad (1.7.10)$$

**Exercise 1.7.3** Use the results of Section 1.5 to show that $S_T$ is compact operator on $C(\mathbb{T}^n)$ that preserves the cone of positive functions.

We are going to prove the Krein-Rutman Theorem for $S_T$ first.

**Theorem 1.7.4** The operator $S_T$ has an eigenvalue $\bar{\mu} > 0$ that corresponds to a positive eigenfunction $\phi_1(x) > 0$. The eigenvalue $\bar{\mu}$ is simple: the only solutions of

$$(S_T - \bar{\mu})u = 0, \quad x \in \mathbb{T}^n$$

are multiples of $\phi_1$. If $\mu$ is another (possibly non-real) eigenvalue of $S_T$, then $|\mu| < \bar{\mu}$.

**Proof.** Let us pick any positive function $\phi_0 \in C(\mathbb{T}^n)$, set $\psi_0 = \phi_0/\|\phi_0\|_{L^\infty}$, and consider the iterative sequence $(\phi_n, \psi_n)$:

$$\phi_{n+1} = S_T\psi_n, \quad \psi_{n+1} = \frac{\phi_{n+1}}{\|\phi_{n+1}\|_{L^\infty}}.$$
Note that, because $\phi_0$ is positive, both $\phi_n$ and $\psi_n$ are positive for all $n$, by the strong maximum principle. For every $n$, let $\mu_n$ be the smallest $\mu$ such that

$$\phi_{n+1}(x) \leq \mu \phi_n(x), \quad \text{for all } x \in \mathbb{T}^n. \tag{1.7.11}$$

Note that (1.7.11) holds for large $\mu$, because each of the $\phi_n$ is positive, hence the smallest such $\mu$ exists. It is also clear that $\mu_n \geq 0$. We claim that the sequence $\mu_n$ is non-increasing. To see that, we apply the operator $S_T$ to both sides of the inequality (1.7.11) with $\mu = \mu_n$, written as

$$S_T \psi_n(x) \leq \mu_n \psi_n(x), \quad \text{for all } x \in \mathbb{T}^n. \tag{1.7.12}$$

and use the fact that $S_T$ preserves positivity, to get

$$(S_T \circ S_T) \psi_n(x) \leq \mu_n S_T \psi_n(x), \quad \text{for all } x \in \mathbb{T}^n, \tag{1.7.13}$$

which is

$$S_T \phi_{n+1}(x) \leq \mu_n \phi_{n+1}(x), \quad \text{for all } x \in \mathbb{T}^n. \tag{1.7.14}$$

Dividing both sides by $\|\phi_{n+1}\|_{L^\infty}$, we see that

$$S_T \psi_{n+1}(x) \leq \mu_n \psi_{n+1}(x), \quad \text{for all } x \in \mathbb{T}^n, \tag{1.7.15}$$

hence

$$\phi_{n+2}(x) \leq \mu_n \phi_{n+1}(x), \quad \text{for all } x \in \mathbb{T}^n. \tag{1.7.16}$$

It follows that $\mu_{n+1} \leq \mu_n$.

Thus, $\mu_n$ converges to a limit $\bar{\mu}$.

**Exercise 1.7.5** Show that, up to an extraction of a subsequence, the sequence $\psi_n$ converges to a limit $\psi_\infty$, with $\|\psi_\infty\|_{L^\infty} = 1$.

The corresponding subsequence $\phi_{n_k}$ converges to $\phi_\infty = S_T \psi_\infty$, by the continuity of $S_T$. And we have, by (1.7.11):

$$S_T \psi_\infty \leq \bar{\mu} \psi_\infty. \tag{1.7.17}$$

If we have the equality in (1.7.17):

$$S_T \psi_\infty(x) = \bar{\mu} \psi_\infty(x) \text{ for all } x \in \mathbb{T}^n, \tag{1.7.18}$$

then $\psi_\infty$ is a positive eigenfunction of $S_T$ corresponding to the eigenvalue $\bar{\mu}$. If, on the other hand, we have

$$S_T \psi_\infty(x) < \bar{\mu} \psi_\infty(x), \text{ on an open set } U \subset \mathbb{T}^n, \tag{1.7.19}$$

they we may use the fact that $S_T$ maps the boundary of the cone of non-negative functions into its interior. In other words, we use the strong maximum principle here. Applying $S_T$ to both sides of (1.7.17) gives then:

$$S_T \phi_\infty < \bar{\mu} \phi_\infty \text{ for all } x \in \mathbb{T}^n. \tag{1.7.20}$$

This contradicts, for large $n$, the minimality of $\mu_n$. Thus, (1.7.19) is impossible, and $\bar{\mu}$ is the sought for eigenvalue. We set, from now on, $\phi_1 = \psi_\infty$:

$$S_T \phi_1 = \bar{\mu} \phi_1, \quad \phi_1(x) > 0 \text{ for all } x \in \mathbb{T}^n. \tag{1.7.21}$$
Exercise 1.7.6 So far, we have shown that $\bar{\mu} \geq 0$. Why do we know that, actually, $\bar{\mu} > 0$?

Let $\phi$ be an eigenfunction of $S_T$ that is not a multiple of $\phi_1$, corresponding to an eigenvalue $\mu$:

$$S_T \phi = \mu \phi.$$ 

Let us first assume that $\mu$ is real, and so is the eigenfunction $\phi$. If $\mu \geq 0$, after multiplying $\phi$ by an appropriate factor, we may assume without loss of generality that $\phi_1(x) \geq \phi(x)$ for all $x \in \mathbb{T}^n$, $\phi_1 \not\equiv \phi$, and there exists $x_0 \in \mathbb{T}^n$ such that $\phi_1(x_0) = \phi(x_0)$. The strong comparison principle implies that then

$$S_T \phi_1(x) > S_T \phi(x) \text{ for all } x \in \mathbb{T}^n.$$ 

It follows, in particular, that

$$\bar{\mu} \phi_1(x_0) > \mu \phi(x_0),$$

hence $\bar{\mu} > \mu \geq 0$, as $\phi_1(x_0) = \phi(x_0) > 0$. This argument also shows that $\bar{\mu}$ is a simple eigenvalue.

If $\mu < 0$, then we can choose $\phi$ (after multiplying it by a, possibly negative, constant) so that, first,

$$\phi_1(x) \geq \phi(x), \quad \phi(x) \geq -\phi_1(x), \quad \text{for all } x \in \mathbb{T}^n,$$ 

(1.7.22)

and there exists $x_0 \in \mathbb{T}^n$ such that

$$\phi(x_0) = \phi_1(x_0).$$

Applying $S_T$ to the second inequality in (1.7.22) gives, in particular,

$$\mu \phi(x_0) > -\bar{\mu} \phi_1(x_0),$$

(1.7.23)

thus $\bar{\mu} > -\mu$. In both cases, we see that $|\mu| < \bar{\mu}$.

Exercise 1.7.7 Use a similar consideration for the case when $\mu$ is complex. In that case, it helps to write the corresponding eigenfunction as

$$\phi = u + iv,$$

and consider the action of $S_T$ on the span of $u$ and $v$, using the same comparison idea. Show that $|\mu| < \bar{\mu}$. If in doubt, consult [46].

This completes the proof of Theorem 1.7.4. □

1.7.3 Back to the principal periodic elliptic eigenvalue

Consider now again the operator $L$ given by (1.7.1):

$$Lu(x) = -\Delta u + b_j(x) \frac{\partial u}{\partial x_j} + c(x) u,$$ 

(1.7.24)

with bounded, smooth and 1-periodic coefficients $b_j(x)$ and $c(x)$. One consequence of Theorem 1.7.4 is the analogous result for the principal periodic eigenvalue for $L$. We will also refer to the following as the Krein-Rutman theorem.
Theorem 1.7.8 The operator $L$ has a unique eigenvalue $\lambda_1$ associated to a positive function $\phi_1$. Moreover, each eigenvalue of $L$ has a real part strictly larger than $\lambda_1$.

Proof. The operator $L$ falls, of course, in the realm of Theorem 1.7.4, since its time-independent coefficients are $T$-periodic for all $T > 0$. We are also going to use the formula

$$L\phi = -\lim_{t \downarrow 0} \frac{S_t\phi - \phi}{t}, \quad (1.7.25)$$

for smooth $\phi(x)$, with the limit in the sense of uniform convergence. This is nothing but an expression of the fact that the function $u(t,x) = [S_t\phi](x)$ is the solution of

$$u_t + Lu = 0, \quad (1.7.26)$$

with the initial condition $u(0,x) = \phi(x)$, and if $\phi$ is smooth, then (1.7.26) holds also at $t = 0$.

Given $n \in \mathbb{N}$, let $\bar{\mu}_n$ be the principal eigenvalue of the operator $S_{1/n}$, with the principal eigenfunction $\phi_n > 0$:

$$S_{1/n}\phi_n = \bar{\mu}_n\phi_n,$$

normalized so that $\|\phi_n\|_\infty = 1$.

Exercise 1.7.9 Show that

$$\lim_{n \to \infty} \bar{\mu}_n = 1$$

directly, without using (1.7.27) below.

As $(S_{1/n})^n = S_1$ for all $n$, we conclude that $\phi_n$ is a positive eigenfunction of $S_1$ with the eigenvalue $(\bar{\mu}_n)^n$. By the uniqueness of the positive eigenfunction, we have

$$\bar{\mu}_n = (\bar{\mu}_1)^{1/n}, \quad \phi_n = \phi_1. \quad (1.7.27)$$

Note that, by the parabolic regularity, $\phi_1$ is infinitely smooth, simply because it is a multiple of $S_1\phi_1$, which is infinitely smooth. Hence, (1.7.25) applies to $\phi_1$, and

$$L\phi_1 = -\lim_{n \to +\infty} n(S_{1/n} - I)\phi_1 = -\lim_{n \to +\infty} n(\bar{\mu}_1^{1/n} - 1)\phi_1 = -(\log \bar{\mu}_1)\phi_1.$$

We have thus proved the existence of an eigenvalue $\lambda_1 = -\log \bar{\mu}_1$ of $L$ that corresponds to a positive eigenfunction. It is easy to see that if

$$L\phi = \lambda\phi,$$

then

$$S_1\phi = e^{-\lambda} \phi.$$

It follows that $L$ can have only one eigenvalue corresponding to a positive eigenfunction. As we know that all eigenvalues $\mu$ of $S_1$ satisfy $|\mu| < \bar{\mu}_1$, we conclude that $\lambda_1$ is the eigenvalue of $L$ with the smallest real part. □

If $L$ is symmetric – that is, it has the form

$$Lu = -\frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + c(x)u, \quad (1.7.28)$$

61
with \( a_{ij} = a_{ji} \), then the first eigenvalue is given by the minimization over \( H^1(\mathbb{T}^n) \) of the Rayleigh quotient

\[
\lambda_1 = \inf_{u \in H^1(\mathbb{T}^n)} \frac{\int_{\mathbb{T}^n} (a_{ij}(x) (\partial_i u)(\partial_j u) + c(x)u^2(x))dx}{\int_{\mathbb{T}^n} u^2(x)dx}.
\]  

(1.7.29)

The existence and uniqueness (up to a factor) of the minimizer is a classical exercise that we do not reproduce here. As for the positivity of the minimizer, we notice that, if \( \phi \) is a minimizer of the Rayleigh quotient, then \( |\phi_1| \) is also a minimizer, thus the unique minimizer is a positive function.

We end this section with a proposition that may look slightly academic, because it has to do with lowering the smoothness of the coefficients - something that we have not been so much interested in so far. A first reason to state it here is that it involves a nice juggling of estimates. Another reason is that it will have a true application in the next chapter. For an \( \mathbb{R}^n \)-valued function \( v(x) \) we denote the divergence of \( v(x) \) by

\[
\nabla \cdot v = \sum_{j=1}^{n} \frac{\partial v_j}{\partial x_j}.
\]

Proposition 1.7.10 Let \( b(x) \) be a smooth vector field over \( \mathbb{T}^n \). The linear equation

\[
-\Delta e + \nabla \cdot (eb) = 0, \quad x \in \mathbb{T}^n,
\]

(1.7.30)

has a unique solution \( e^*_1(x) \) normalized so that

\[
||e^*_1||_{L^\infty} = 1,
\]

(1.7.31)

and such that \( e^*_1 > 0 \) on \( \mathbb{T}^n \). Moreover, for all \( \alpha \in (0, 1) \), the function \( e^*_1 \) is \( \alpha \)-Hölder continuous, with the \( \alpha \)-Hölder norm bounded by a universal constant depending only on \( ||b||_{L^\infty(\mathbb{T}^n)} \).

A key point here is that the Hölder regularity of the solution only depends on the \( L^\infty \)-norm of \( b(x) \) but not on its smoothness or any of its derivatives – this is typical for equations in the divergence form, and we will see much more of this in Chapter ???. This is very different from what we have seen so far in this chapter: we have always relied on the assumption that the coefficients are smooth, and the Hölder bounds for the solutions depended on the regularity of the coefficients. A very remarkable fact is that for equations in the divergence form, such as (1.7.30), one may often obtain bounds on the regularity of the solutions that depend only on the \( L^\infty \)-norm of the coefficients but not on their smoothness. Such bounds are much harder to get for equations in the non-divergence form.

Proof of Proposition 1.7.10

Let us denote

\[
L\phi = -\Delta \phi - b_j(x) \frac{\partial \phi}{\partial x_j}.
\]

(1.7.32)
The constant functions are the principal periodic eigenfunctions of $L$ and zero is the principal eigenvalue:

$$L1 = 0.$$  \hfill (1.7.33)

Thus, by Theorem 1.7.8, the operator $L$ has no other eigenvalue with a non-positive real part, which entails the same result for the operator

$$L^* \phi = -\Delta \phi + \nabla \cdot (b(x)\phi).$$

In particular, zero is the principal eigenvalue of $L^*$, associated to a positive eigenfunction $e_1^*(x) > 0$:

$$L^* e_1^* = 0, \quad \text{for all } x \in \mathbb{T}^n,$$

and we can normalize $e_1^*$ so that that (1.7.31) holds. Thus, existence of $e_1^*(x)$ is the easy part of the proof.

The challenge is, of course, to bound the Hölder norms of $e_1^*$ in terms of $\|b\|_{L^\infty(\mathbb{T}^n)}$ only. We would like to use a representation formula, as we already did many times in this chapter. In other words, we would like to treat the term $\nabla \cdot (e_1^* b)$ as a force, and convolve it with the fundamental solution of the Laplace equation in $\mathbb{R}^n$. For that, we need various quantities to be sufficiently integrable, so we first localize the equation, and then write a representation formula. This is very similar to the proof of the interior regularity estimates that we have mentioned very briefly in Section 1.6 – see Exercise 1.6.7. We recommend the reader to go back to this Section after finishing the current proof, and attempt this exercise again, setting $a_{ij}(t,x) = \delta_{ij}$ in (1.6.24) for simplicity.

Let $\Gamma(x)$ be a nonnegative smooth cut-off function such that $\Gamma(x) \equiv 1$ for $x \in [-2,2]^n$ and $\Gamma(x) \equiv 0$ outside $(-3,3)^n$. The function $v(x) = \Gamma(x)e_1^*(x)$ satisfies

$$-\Delta v = -2\nabla \Gamma \cdot \nabla e_1^* - e_1^* \Delta \Gamma - \Gamma \nabla \cdot (e_1^* b), \quad x \in \mathbb{R}^n.$$  \hfill (1.7.34)

Remember that $e_1^*$ is bounded in $L^\infty$, thus so is $v$. As we will see, nothing should be feared from the cumbersome quantities like $\Delta \Gamma$ or $\nabla \Gamma$. We concentrate on the space dimensions $n \geq 2$, leaving $n = 1$ as an exercise. Let $E(x)$ be the fundamental solution of the Laplace equation in $\mathbb{R}^n$: the solution of

$$-\Delta u = f, \quad x \in \mathbb{R}^n,$$  \hfill (1.7.35)

is given by

$$u(x) = \int_{\mathbb{R}^n} E(x - y)u(y)dy.$$  \hfill (1.7.36)

Then we have

$$v(x) = \int_{\mathbb{R}^n} E(x - y)\left[-2\Gamma(x) \cdot \nabla e_1^*(y) - e_1^*(y)\Delta \Gamma(y) - \Gamma(y) \nabla \cdot (e_1^*(y)b(y)) \right]dy.$$  \hfill (1.7.37)

After an integration by parts, we obtain

$$v(x) = \int_{\mathbb{R}^n} \left((\nabla E(x - y) \cdot \nabla \Gamma(y))e_1^*(y) + E(x - y)e_1^*(y)\Delta \Gamma(y) + \nabla(E(x - y)\Gamma(y)) \cdot b(y)e_1^*(y) \right)dy.$$  \hfill (1.7.38)
The key point is that no derivatives of \( b(x) \) or \( e^*_i(x) \) appear in the right side of (1.7.38) – this is important as the only a priori information that we have on these functions is that they are bounded in \( L^\infty \). Thus, the main point is to prove that integrals of the form

\[
P(x) = \int_{\mathbb{R}^n} E(x-y) G(y) dy,
\]

with a bounded and compactly supported function \( G(x) \), and

\[
I(x) = \int_{\mathbb{R}^n} \nabla E(x-y) \cdot F(y) dy,
\]

with a bounded and compactly supported vector-valued function \( F : \mathbb{R}^n \to \mathbb{R}^n \), are \( \alpha \)-Hölder continuous for all \( \alpha \in (0, 1) \), with the Hölder constants depending only on \( \alpha \) and the \( L^\infty \)-norms of \( F \) and \( G \). Both \( F \) and \( G \) are supported inside the cube \([-3, 3]^n\). We will only consider the integral \( I(x) \), as this would also show that \( \nabla P(x) \) is \( \alpha \)-Hölder. Using the expression

\[
\nabla E(z) = c_n \frac{z}{|z|^n},
\]

we see that

\[
|I(x) - I(x')| \leq c_n \|F\|_{L^\infty} K(x, x'),
\]

with

\[
K(x, x') = \int_{(-3,3)^n} \left| \frac{x-y}{|x-y|^n} - \frac{x'-y}{|x'-y|^n} \right| dy.
\]

Pick now \( \alpha \in (0, 1) \). We estimate \( K \) by splitting the integration domain into two disjoint pieces:

\[
A_x = \{ y \in (-3,3)^n : |x-y| \leq |x-x'|^\alpha \}, \quad B_x = \{ y \in (-3,3)^n : |x-y| > |x-x'|^\alpha \},
\]

and denote by \( K_A(x, x') \) and \( K_B(x, x') \) the contribution to \( K(x, x') \) by the integration over each of these two regions. To avoid some unnecessary trouble, we assume that \( |x-x'| \leq l_\alpha \), with \( l_\alpha \) such that

\[
3l \leq l_\alpha \quad \text{for all } l \in [0, l_\alpha].
\]

With this choice, we have

\[
|x'-y| \leq |x'-x| + |x-y| \leq 2|x-x'|^\alpha \quad \text{if } y \in A_x,
\]

and

\[
|x'-y| \geq |x-y| - |x'-x| \geq 2|x-x'| \quad \text{if } y \in B_x.
\]

It follows that

\[
K_A(x, x') \leq \int_{|x-y| \leq |x-x'|^\alpha} \frac{dy}{|x-y|^{n-1}} + \int_{|x'-y| \leq 2|x-x'|^\alpha} \frac{dy}{|x'-y|^{n-1}} \leq C|x-x'|^\alpha.
\]

To estimate \( K_B \), we write

\[
\left| \frac{x-y}{|x-y|^n} - \frac{x'-y}{|x'-y|^n} \right| \leq C|x-x'| \int_0^1 \frac{d\sigma}{|x_\sigma - y|^n}, \quad x_\sigma = \sigma x + (1-\sigma)x'.
\]
Note that for all $y \in B_x$ we have

$$|x_\sigma - y| \geq |x - y| - |x - x_\sigma| \geq |x - x'|^\alpha - |x - x'| \geq 2|x' - x|,$$

and $|y| \leq 3\sqrt{n}$, hence

$$K_B(x, x') \leq |x - x'| \int_0^1 d\sigma \int_{B_x} \frac{dy}{|x_\sigma - y|^n} \leq |x - x'| \int_0^1 d\sigma \int_{|x_\sigma - y| \geq |x' - x'|} \frac{\chi(|y| \leq 3\sqrt{n})dy}{|x_\sigma - y|^n},$$

(1.7.48)

which implies the uniform $\alpha$-Hölder bound for $I(x)$, for all $\alpha \in (0, 1)$.

**The Dirichlet principal eigenvalue, related issues**

We have so far talked about the principal eigenvalue for spatially periodic elliptic problems. This discussion applies equally well to problems in bounded domains, with the Dirichlet or Neumann boundary conditions. In the rest of this book, we will often encounter the Dirichlet problems, so let us explain this situation. Let $\Omega$ be a smooth bounded open subset of $\mathbb{R}^n$, and consider our favorite elliptic operator

$$Lu = -\Delta u + b_j(x) \frac{\partial u}{\partial x_j} + c(x)u,$$

(1.7.49)

with smooth coefficients $b_j(x)$ and $c(x)$. One could easily look at the more general problem

$$Lu = -a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_j(x) \frac{\partial u}{\partial x_j} + c(x)u,$$

(1.7.50)

with essentially identical results, as long as the matrix $a_{ij}(x)$ is uniformly elliptic – we will avoid this just to keep the notation simpler. We are interested in the eigenvalue problem

$$Lu = \lambda u \quad \text{in } \Omega,$$

$$u = 0 \quad \text{on } \partial \Omega,$$

(1.7.51)

and, in particular, in the existence of a positive eigenfunction $\phi > 0$ in $\Omega$. The strategy will be as in the periodic case, to look at the initial value problem

$$u_t - \Delta u + b_j(x) \frac{\partial u}{\partial x_j} + c(x)u = 0, \quad t > 0, \quad x \in \Omega,$$

$$u = 0, \quad t > 0, \quad x \in \partial \Omega,$$

$$u(0, x) = u_0(x).$$

(1.7.52)

The coefficients $b_j$ and $c$ are smooth in $(t, x)$ and $T$-periodic in $t$. Again, we set

$$(S_T u_0)(x) = u(T, x).$$

The main difference with the periodic case is that, here, the cone of continuous functions which are positive in $\Omega$ and vanish on $\partial \Omega$ has an empty interior, so we can not repeat verbatim the proof of the Krein-Rutman theorem for the operators on $\mathbb{T}^n$. 65
Exercise 1.7.11 Revisit the proof of the Krein-Rutman theorem in that case and identify the place where the proof would fail for the Dirichlet boundary conditions.

What will save the day is the strong maximum principle and the Hopf Lemma. We are not going to fully repeat the proof of Theorems 1.7.4 and 1.7.8, but we are going to prove a key proposition that an interested reader can use to prove the Krein-Rutman theorem for the Dirichlet problem.

Proposition 1.7.12 Assume $u_0 \in C^1(\Omega)$ - that is, $u_0$ has derivatives that are continuous up to $\partial\Omega$, and that $u_0 > 0$ in $\Omega$, and both $u_0 = 0$ and $\partial\nu u_0 < 0$ on $\partial\Omega$. Then there is $\mu_1 > 0$ defined by the formula

$$\mu_1 = \inf\{\mu > 0 : S_T u_0 \leq \mu u_0\}. \tag{1.7.53}$$

Moreover, if $\mu_2 > 0$ is defined as

$$\mu_2 = \inf\{\mu > 0 : (S_T \circ S_T)u_0 \leq \mu S_T u_0\}, \tag{1.7.54}$$

then either $\mu_1 > \mu_2$, or $\mu_1 = \mu_2$, and in the latter case $(S_T \circ S_T)u_0 \equiv \mu_2 S_T u_0$.

Proof. For the first claim, the existence of the infimum in (1.7.53), we simply note that

$$\mu u_0 \geq S_T u_0,$$

as soon as $\mu > 0$ is large enough, because $\partial\nu u_0 < 0$ on $\partial\Omega$, $u_0 > 0$ in $\Omega$, and $S_T u_0$ is a smooth function up to the boundary. As for the second item, let us first observe that

$$u(t + T, x) \leq \mu_1 u(t, x), \tag{1.7.55}$$

for any $t > 0$, by the maximum principle. Let us assume that

$$u(2T, x) \neq \mu_1 u(T, x). \tag{1.7.56}$$

Then the maximum principle implies that

$$u(2T, x) < \mu_1 u(T, x) \text{ for all } x \in \Omega. \tag{1.7.57}$$

As

$$\max_{x \in \Omega}|u(2T, x) - \mu_1 u(T, x)| = 0,$$

the parabolic Hopf lemma, together with (1.7.55) and (1.7.56), implies the existence of $\delta > 0$ such that

$$\partial\nu(u(2T, x) - \mu_1 u(T, x)) \geq \delta > 0, \text{ for all } x \in \partial\Omega. \tag{1.7.58}$$

It follows that for $\varepsilon > 0$ sufficiently small, we have

$$u(2T, x) - \mu_1 u(T, x) \leq -\frac{\delta}{2}d(x, \partial\Omega) \text{ for } x \in \Omega \text{ such that } d(x, \partial\Omega) < \varepsilon.$$

On the other hand, once again, the strong maximum principle precludes a touching point between $u(2T, x)$ and $\mu_1 u(T, x)$ inside

$$\overline{\Omega}_\varepsilon = \{x \in \Omega : d(x, \partial\Omega) \geq \varepsilon\}.$$
Therefore, there exists $\delta_1$ such that

$$u(2T, x) - \mu_1 u(T, x) \leq -\delta_1,$$

for all $x \in \overline{\Omega}$. We deduce that there is a – possibly very small – constant $c > 0$ such that

$$u(2T, x) - \mu_1 u(T, x) \leq -cd(x, \partial \Omega) \quad \text{in } \Omega.$$ 

However, $u(T, x)$ is controlled from above by $Cd(x, \partial \Omega)$, for a possibly large constant $C > 0$. All in all, we have

$$u(2T, x) \leq (\mu_1 - \frac{c}{C})u(T, x),$$

hence (1.7.56) implies that $\mu_2 < \mu_1$, which proves the second claim of the proposition. □

**Exercise 1.7.13** Deduce from Proposition 1.7.12 the versions of Theorems 1.7.4 and 1.7.8 for operators $S_T$ and $L$, this time with the Dirichlet boundary conditions.

Thus, the eigenvalue problem (1.7.51), has a principal eigenvalue that enjoys all the properties we have proved in the periodic one: it has the least real part among all eigenvalues, and is the only eigenvalue associated to a positive eigenfunction.

**Exercise 1.7.14** Assume that $L$ is symmetric; it has the form

$$Lu = -\frac{\partial}{\partial x_j} \left(a_{ij}(x) \frac{\partial u}{\partial x_i} \right) + c(x)u \quad \text{(1.7.59)}$$

Then, the principal eigenvalue is given by the minimization of the Rayleigh quotient over the Sobolev space $H^1_0(\Omega)$:

$$\lambda_1 = \inf_{u \in H^1_0(\Omega), \|u\|_{L^2} = 1} \int_\Omega \left(a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} + c(x)u^2(x) \right) dx. \quad \text{(1.7.60)}$$

**Exercise 1.7.15** Adapt the preceding discussion to prove the existence of a principal eigenvalue to the Neumann eigenvalue problem

$$Lu = \lambda u, \quad x \in \Omega,$$

$$\partial_n u = 0, \quad x \in \partial \Omega. \quad \text{(1.7.61)}$$

**Exercise 1.7.16** *(The principal eigenvalue in an unbounded domain)* Consider the differential operator

$$L = -\Delta + c(x), \quad x \in \mathbb{R}^n.$$ 

Assume that $c$ is bounded and uniformly continuous. Assume the existence of $c_\infty > 0$ such that

$$\lim_{|x| \to +\infty} c(x) = c_\infty. \quad \text{(1.7.62)}$$

Also assume that $c(x) < c_\infty$ for all $x \in \mathbb{R}^n$. The goal of the exercise is to prove a Krein-Rutman type theorem for $L$. For $n \geq 1$, let $\mu_n$ be the principal eigenvalue of $L$ in $B_n(0)$, with Dirichlet conditions.
1. Show that \((\mu_n)_n\) is a decreasing sequence, bounded by \(-\|c\|_\infty\). Let \(\mu_\infty\) be its limit.

2. Let \(\psi_{\varepsilon,n}(x)\) solve
\[
-\Delta \psi + \varepsilon \psi = 0 \quad (B_n(0) \setminus B_R(0)) \\
\psi = 1 \quad (\partial B_R(0)) \\
\psi = 0 \quad (\partial B_n(0)).
\]
Show that \(|\partial_r \psi_{\varepsilon,n}| = O(\sqrt{\varepsilon})\) on \(\partial B_R(0)\), as soon as \(n\) is large enough.

3. Let \(\phi_{2R}\) be the first Dirichlet eigenfunction in \(B_{2R}(0)\), that is equal to 1 on \(\partial B_R(0)\) (why is \(\phi_{2R}\) radial?). Let \(\phi_n\) be equal to \(\phi_{2R}\) in \(B_R(0)\), and \(\psi_{\varepsilon,n}\) in \(B_n(0) \setminus B_R(0)\). Show that, if \(\varepsilon > 0\) is small enough, \(R\) large and \(n\) very large, then we have
\[
L \phi_n \leq (c_\infty - \varepsilon) \phi_n.
\]

4. Deduce that \(\mu_\infty \leq c_\infty - \varepsilon\).

5. Conclude that \(L\) has the Krein-Rutman property.

6. Show that the first eigenfunction decays exponentially fast at infinity.

**Exercise 1.7.17** Set
\[
L = -\Delta + c(x), \quad x \in \mathbb{R}^n,
\]
the function \(c\) satisfying (1.7.62) for some positive \(c_\infty\). We do not, however, assume \(c(x) < c_\infty\) anymore. Find as many properties of the preceding exercise as possible that would fail without this assumption.

**Exercise 1.7.18** Redo the existence part of Exercise 1.7.16 with the aid of the Rayleigh quotients, without any approximation on a finite domain.

These three exercises give just a glimpse at what happens to the principal eigenvalue in unbounded domains – an interested reader should investigate further, starting with the variational formulations of [1] and [110], and continuing with the more recent papers [19, 24].

### 1.7.4 The principal eigenvalue and the comparison principle

Let us now connect the principal eigenvalue and the comparison principle. Since we are at the moment dealing with the Dirichlet problems, let us remain in this context. There would be nothing significantly different about the periodic problems.

The principal eigenfunction \(\phi_1 > 0\), solution of
\[
L \phi_1 = \lambda_1 \phi_1, \quad \text{in } \Omega, \quad (1.7.63)
\]
\[
\phi_1 = 0 \quad \text{on } \partial \Omega, \quad (1.7.64)
\]
with
\[
Lu = -\Delta u + b_j(x) \frac{\partial u}{\partial x_j} + c(x) u, \quad (1.7.65)
\]
in particular, provides a special solution
\[ \psi(t, x) = e^{-\lambda_1 t} \phi_1(x) \] (1.7.66)
for the linear parabolic problem
\[ \psi_t + L\psi = 0, \quad t > 0, x \in \Omega \] (1.7.67)
\[ \psi = 0 \text{ on } \partial \Omega. \]
Consider then the Cauchy problem
\[ v_t + Lv = 0, \quad t > 0, x \in \Omega \] (1.7.68)
\[ v = 0 \text{ on } \partial \Omega, \]
\[ v(0, x) = g(x), \quad x \in \Omega, \]
with a smooth bounded function \( g(x) \) that vanishes at the boundary \( \partial \Omega \). We can find a constant \( M > 0 \) so that
\[ -M \phi_1(x) \leq g(x) \leq M \phi_1(x), \quad \text{for all } x \in \Omega. \]
The comparison principle then implies that for all \( t > 0 \) we have a bound
\[ -M \phi_1(x) e^{-\lambda_1 t} \leq v(t, x) \leq M \phi_1(x) e^{-\lambda_1 t}, \quad \text{for all } x \in \Omega, \] (1.7.69)
which is very useful, especially if \( \lambda_1 > 0 \). The assumption that the initial condition \( g \) vanishes at the boundary \( \partial \Omega \) is not necessary but removes the technical step of having to show that even if \( g(x) \) does not vanish on the boundary, then for any positive time \( t_0 > 0 \) we can find a constant \( C_0 \) so that \( |v(t_0, x)| \leq C_0 \phi_1(x) \). This leads to the bound (1.7.69) for all \( t > t_0 \).

Let us now apply the above considerations to the solutions of the elliptic problem
\[ Lu = g(x), \quad \text{in } \Omega, \] (1.7.70)
\[ u = 0 \text{ on } \partial \Omega, \]
with a non-negative function \( g(x) \). When can we conclude that the solution \( u(x) \) is also non-negative? The solution of (1.7.70) can be formally written as
\[ u(x) = \int_0^\infty v(t, x) dt. \] (1.7.71)
Here, the function \( v(t, x) \) satisfies the Cauchy problem (1.7.68). If the principal eigenvalue \( \lambda_1 \) of the operator \( L \) is positive, then the integral (1.7.71) converges for all \( x \in \Omega \) because of the estimates (1.7.69), and the solution of (1.7.70) is, indeed, given by (1.7.71). On the other hand, if \( g(x) \geq 0 \) and \( g(x) \neq 0 \), then the parabolic comparison principle implies that \( v(t, x) > 0 \) for all \( t > 0 \) and all \( x \in \Omega \). It follows that \( u(x) > 0 \) in \( \Omega \).

Therefore, we have proved the following theorem that succinctly relates the notions of the principal eigenvalue and the comparison principle.
Theorem 1.7.19 If the principal eigenvalue of the operator $L$ is positive then solutions of the elliptic equation (1.7.70) satisfy the comparison principle: $u(x) > 0$ in $\Omega$ if $g(x) \geq 0$ in $\Omega$ and $g(x) \not\equiv 0$.

This theorem allows to look at the maximum principle in narrow domains introduced in the previous chapter from a slightly different point of view: the narrowness of the domain implies that the principal eigenvalue of $L$ is positive no matter what the sign of the free coefficient $c(x)$ is. This is because the “size” of the second order term in $L$ increases as the domain narrows, while the “size” of the zero-order term does not change. Therefore, in a sufficiently narrow domain the principal eigenvalue of $L$ will be positive (recall that the required narrowness does depend on the size of $c(x)$). A similar philosophy applies to the maximum principle for the domains of a small volume.

We conclude this topic with another characterization of the principal eigenvalue of an elliptic operator in a bounded domain, which we leave as an (important) exercise for the reader. Let us define

$$\mu_1(\Omega) = \sup \{ \lambda : \exists \phi \in C^2(\Omega) \cap C^1(\overline{\Omega}), \phi > 0 \text{ and } (L - \lambda)\phi \geq 0 \text{ in } \Omega \},$$

(1.7.72)

and

$$\mu'_1(\Omega) = \inf \{ \lambda : \exists \phi \in C^2(\Omega) \cap C^1(\overline{\Omega}), \phi = 0 \text{ on } \partial\Omega, \phi > 0 \text{ in } \Omega, \text{ and } (L - \lambda)\phi \leq 0 \text{ in } \Omega \}.$$  

(1.7.73)

Exercise 1.7.20 Let $L$ be an elliptic operator in a smooth bounded domain $\Omega$, and let $\lambda_1$ be the principal eigenvalue of the operator $L$, and $\mu_1(\Omega)$ and $\mu'_1(\Omega)$ be as above. Show that

$$\lambda_1 = \mu_1(\Omega) = \mu'_1(\Omega).$$

(1.7.74)

As a hint, say, for the equality $\lambda_1 = \mu_1(\Omega)$, we suggest, assuming existence of some $\lambda > \lambda_1$ and $\phi > 0$ such that

$$(L - \lambda)\phi \geq 0,$$

to consider the Cauchy problem

$$u_t + (L - \lambda)u = 0, \text{ in } \Omega$$

with the initial data $u(0, x) = \phi(x)$, and with the Dirichlet boundary condition $u(t, x) = 0$ for $t > 0$ and $x \in \partial\Omega$. One should prove two things: first, that $u(t, x) \leq 0$ for all $t > 0$, and, second, that there exists some constant $C > 0$ so that

$$u(t, x) \geq C\phi_1(x)e^{(\lambda - \lambda_1)t},$$

where $\phi_1$ is the principal Dirichlet eigenfunction of $L$. This will lead to a contradiction. The second equality in (1.7.74) is proved in a similar way.
1.8 Reaction-diffusion waves

As a conclusion to this chapter, we will be interested here in one-dimensional models of the form

\[ u_t - u_{xx} = f(x, u), \quad t > 0, \; x \in \mathbb{R}, \] (1.8.1)

the assumptions on \( f \) being made precise as the study develops. We will see in this section how the possibility of comparing two solutions of the same problem will imply their convergence in the long time limit, putting to work the two main characters we have seen so far in this chapter: the comparison principle and the Harnack inequality. We will also put to good use the ideas developed for the existence of the principal eigenvalues for elliptic operators, they are the same - although they will sometimes be imbedded in more or less technical considerations. We will start by the simplest possible model: the Allen-Cahn equation, with \( f(x, u) = f(u) = u(1 - u)^2 \). There is an explicit steady solution, and we will show in detail how every solution of the problem, that vaguely looks like the steady solution at both ends at time \( t = 0 \), will converge to a translate of it for large times. The rest of the chapter will be devoted to showing that the idea is universal, and helps the understanding of seemingly more complicated, or unrelated models. We will first treat nonlinearities that are less symmetric than the Allen-Cahn one, giving rise to travelling waves, that will attract the whole dynamics of the solutions. We will finally investigate the large-time behavior of (1.8.1) with an \( f(x, u) \) periodic in \( x \). This is going to give raise to pulsating waves, i.e. waves that look time-periodic in some Galilean reference frame. These waves will be shown to be globally attracting, thus giving some substance to the advertiseemnt (that we made in the introduction) about how space periodicity generates time-periodicity.

1.8.1 The long time behavior for the Allen-Cahn equation

We consider the one-dimensional Allen-Cahn equation

\[ u_t - u_{xx} = f(u), \] (1.8.2)

with

\[ f(u) = u - u^3. \] (1.8.3)

Recall that we have already considered the steady solutions of this equation in Section ?? of Chapter ??, and, in particular, the role of its explicit time-independent solutions

\[ \phi(x) = \tanh \left( \frac{x}{\sqrt{2}} \right), \] (1.8.4)

and its translates \( \phi_{x_0}(x) := \phi(x + x_0), \; x_0 \in \mathbb{R} \).

**Exercise 1.8.1** We have proved in Chapter ?? that, if \( \psi(x) \) is a steady solution to (1.8.2) that satisfies

\[ \lim_{x \to -\infty} \psi(x) = -1, \quad \lim_{x \to +\infty} \psi(x) = 1, \]

then \( \psi \) is a translate of \( \phi \). For an alternative proof, draw the phase portrait of the equation

\[ -\psi'' = f(\psi) \] (1.8.5)
in the \((\psi, \psi')\) plane. For an orbit \((\psi, \psi')\) connecting \((-1,0)\) to \((1,0)\), show that the solution tends to \((-1,0)\) exponentially fast. Multiply then (1.8.5) by \(\psi'\), integrate from \(-\infty\) to \(x\) and conclude.

Recall that the Allen-Cahn equation is a simple model for a physical situation when two phases are stable, corresponding to \(u = \pm 1\). The time dynamics of the initial value problem for (1.8.2) corresponds to a competition between these two states. The fact that

\[
\int_{-1}^{1} f(u) \, du = 0 \quad (1.8.6)
\]

means that the two states are "equally stable" – this is a necessary condition for (1.8.2) to have a time-independent solution \(\phi(x)\) such that

\[
\phi(x) \to \pm 1, \quad \text{as } x \to \pm \infty. \quad (1.8.7)
\]

In other words, such connection between +1 and −1 exists only if (1.8.6) holds.

Since the two phases \(u = \pm 1\) are equally stable, one expects that if the initial condition \(u_0(x)\) for (1.8.2) satisfies

\[
\lim_{x \to -\infty} u_0(x) = -1, \quad \lim_{x \to +\infty} u_0(x) = 1, \quad (1.8.8)
\]

then, as \(t \to +\infty\), the solution \(u(t,x)\) will converge to a steady equilibrium, that has to be a translate of \(\phi\). This is the subject of the next theorem, that shows, in addition, that the convergence rate is exponential.

**Theorem 1.8.2** There exists \(\omega > 0\) such that for any uniformly continuous and bounded initial condition \(u_0\) for (1.8.2) that satisfies (1.8.8), we can find \(x_0 \in \mathbb{R}\) and \(C_0 > 0\) such that

\[
|u(t,x) - \phi(x + x_0)| \leq C_0 e^{-\omega t}, \quad \text{for all } x \in \mathbb{R} \text{ and } t > 0. \quad (1.8.9)
\]

Since there is a one parameter family of steady solutions, naturally, one may ask how the solution of the initial value problem chooses a particular translation of \(\phi\) in the long time limit. In other words, one would like to know how the shift \(x_0\) depends on the initial condition \(u_0\). However, this dependence is quite implicit and there is no simple expression for \(x_0\).

There are at least two ways to prove Theorem 1.8.2, both of them need the forthcoming Lemma 1.8.4, that bounds the level sets of the solution. Once this is at hand, a first option is to solve the following

**Exercise 1.8.3** Assume Lemma 1.8.4 to be true.

1. Verify that the energy functional

\[
J(u) = \int_{\mathbb{R}} \left( \frac{1}{2} |u_x|^2 - F(u) \right) \, dx, \quad F(u) = \int_{-1}^{u} f(v) \, dv;
\]

decreases in time for any solution \(u(t,x)\) of (1.8.2).

2. With the aid of Lemma 1.8.4 and the preceding question, show that the solution eventually comes very close to a translate \(\phi_{x_0}(x)\), uniformly on \(\mathbb{R}\).
3. Prove a Krein-Rutman type property for the operator

\[ \mathcal{M}u = -u_{xx} - f'(\phi_{x_0})u. \]

What is the principal eigenvalue, and what is an associated eigenfunction?

4. If \( v(x) \) is close to \( \phi_{x_0} \), show that one may decompose it uniquely as

\[ v(x) = \phi_{x_0}(x + X) + w, \]

\( X \) small and \( w \) small, orthogonal to the null space of \( M \).

5. Assume that \( u_0 \) is close to \( \phi_{x_0} \).
   
   • Show the existence of \( T > 0 \) such that the decomposition
     
     \[ u(t, x) = \phi_{x_0}(x + X(t)) + w(t, x) \]
     
     \( X \) small and \( w \) small, orthogonal to the null space of \( M \), holds at least up to time \( T \).
   
   • Write a system of equations for \( (X(t), w(t, x)) \).
   
   • Deduce that \( T \) can be chosen infinite and that \( X(t) \) converges, exponentially in time, to some \( x_1 \) close to \( x_0 \).

6. Round up everything and conclude.

This is, more or less, the method devised in the beautiful paper of Fife and McLeod [64]. It has been generalized to gradient systems in a remarkable paper of Risler [122], which proves very precise spreading estimates of the leading edge of the solutions, only based on a one-dimensional set of energy functionals. Risler’s ideas were put to work on the simpler example (1.8.2) in a paper by Gallay and Risler [71].

We chose to present an alternative method, entirely based on sub and super-solutions that come closer and closer to each other. It avoids the spectral arguments, and is more flexible as there are many reaction-diffusion problems where the comparison principle and the Harnack inequality are available but the energy functionals do not exist. The reader should also be aware that there are many problems, such as many reaction-diffusion systems, where the situation is the opposite: the energy functional exists but the comparison principle is not applicable.

Before we begin, we note that the function \( f \) satisfies

\[ f'(u) \leq -1 \quad \text{for} \quad |u| \geq 5/6, \quad f'(u) \leq -3/2 \quad \text{for} \quad |u| \geq 11/12. \]  

(1.8.10)

We will also take \( R_0 > 0 \) such that

\[ |\phi(x)| \geq 11/12 \quad \text{for} \quad |x| \geq R_0. \]  

(1.8.11)
A bound on the level sets

The first ingredient is to prove that the level sets of \( u(t,x) \) do not, indeed, go to infinity, so that the region of activity, where \( u(t,x) \) is not too close to \( \pm 1 \), happens, essentially, in a compact set. This crucial step had already been identified by Fife and McLeod, and we reproduce here their argument. The idea is to squish \( u(t,x) \) between two different translates of \( \phi \), with a correction that goes to zero exponentially in fast time.

**Lemma 1.8.4** Let \( u_0 \) satisfy the assumptions of the theorem. There exist \( \xi^\pm_\infty \in \mathbb{R} \), and \( q_0 > 0 \), such that

\[
\phi(x + \xi^-_\infty) - q_0 e^{-t} \leq u(t,x) \leq \phi(x + \xi^+_\infty) + q_0 e^{-t},
\]

for all \( t \geq 0 \) and \( x \in \mathbb{R} \).

**Proof.** For the upper bound, we are going to devise two functions \( \xi^+(t) \) and \( q(t) \) such that

\[
\bar{u}(t,x) = \phi(x + \xi^+(t)) + q(t)
\]

is a super-solution to (1.8.2), with an increasing but bounded function \( \xi^+(t) \), and an exponentially decreasing function \( q(t) = q_0 \exp(-t) \). One would also construct, in a similar way, a sub-solution of the form

\[
\underline{u}(t,x) = \phi(x + \xi^-(t)) - q(t),
\]

possibly increasing \( q \) a little, with a decreasing but bounded function \( \xi^-(t) \).

Let us denote

\[
N[u] = \partial_t u - u_{xx} - f(u).
\]

Now, with \( \bar{u}(t,x) \) as in (1.8.13), we have

\[
N[\bar{u}] = \dot{\bar{q}} + \dot{\xi}^+ \phi'(\zeta) - f(\phi(\zeta) + q) + f(\phi(\zeta)),
\]

with \( \zeta = x + \xi^+(t) \). Our goal is to choose \( \xi^+(t) \) and \( q(t) \) so that

\[
N[\bar{u}] \geq 0, \quad \text{for all } t \geq 0 \text{ and } x \in \mathbb{R},
\]

so that \( \bar{u}(t,x) \) is a super-solution to (1.8.2). We will consider separately the regions \( |\zeta| \leq R_0 \) and \( |\zeta| \geq R_0 \).

**Step 1. The region** \( |\zeta| \geq R_0 \). First, we have

\[
\phi(\zeta) + q(t) \geq 11/12 \text{ for } \zeta \geq R_0,
\]

as \( q(t) \geq 0 \). If we assume that \( q(0) \leq 1/12 \) and make sure that \( q(t) \) is decreasing in time, then we also have

\[
\phi(\zeta) + \dot{q} \leq -5/6 \text{ for } \zeta \leq -R_0.
\]

We have, therefore, as long as \( \xi^+(t) \) is increasing, using (1.8.10):

\[
N[\bar{u}] \geq \dot{q} - f(\phi(\zeta) + q) + f(\phi) \geq \dot{q} + q, \text{ for } |\zeta| \geq R_0.
\]

It suffices, therefore, to choose

\[
q(t) = q(0)e^{-t},
\]

74
with \( q(0) \leq 1/12 \), and an increasing \( \xi^+(t) \), to ensure that

\[
N[\bar{u}] \geq 0, \quad \text{for all } t \geq 0 \text{ and } |\zeta| \geq R_0. \tag{1.8.20}
\]

**Step 2. The region** \( |\zeta| \leq R_0 \). This time, we have to choose \( \xi^+(t) \) properly. We write

\[
N[\bar{u}] \geq \dot{q} + \dot{\xi}^+ \phi' (\zeta) - M_f q, \quad M_f = \| f' \|_{L^\infty}, \tag{1.8.21}
\]

and choose

\[
\dot{\xi}^+ = \frac{1}{k_0} \left( - \dot{q} + M_f q \right), \quad k_0 = \inf_{|\zeta| \leq R_0} \phi' (\zeta), \tag{1.8.22}
\]

to ensure that the right side of (1.8.21) is non-negative. Using expression (1.8.19) for \( q(t) \), we obtain

\[
\xi^+(t) = \xi^+(0) + \frac{q(0)}{k_0} (1 + M_f)(1 - e^{-t}). \tag{1.8.23}
\]

To summarize, with the above choices of \( q(t) \) and \( \xi^+(t) \), we know that \( \bar{u} \) satisfies (1.8.17).

It remains to choose \( q(0) \) and \( \xi^+(0) \) so that \( \bar{u}(t,x) \) is actually above \( u(t,x) \) – as we have already established (1.8.17), the comparison principle tells us that we only need to verify that

\[
\bar{u}(0,x) \geq u_0(x), \quad \text{for all } x \in \mathbb{R}. \tag{1.8.24}
\]

Because \( u_0 \) tends to \( \pm 1 \) at \( \pm \infty \), there exists \( \xi_0^+ \) (possibly quite large), and \( q_0 \in (0, 1/12) \) such that

\[
u \leq \phi(x + \xi_0^+) + q_0. \tag{1.8.25}
\]

Thus, it is enough to choose \( q(0) = q_0, \xi^+(0) = \xi_0^+ \). □

**Exercise 1.8.5** Follow the same strategy to construct a sub-solution \( u(t,x) \) as in (1.8.14).

Lemma 1.8.4 traps nicely the level sets of \( u \). But will this imply convergence to a steady solution, or will the level sets of \( u(t,x) \) oscillate inside a bounded set? First, let us restate our findings in a more precise way. We have shown the following

**Corollary 1.8.6** Assume that we have

\[
\phi(x + \xi_0^-) - q_0 \leq u_0(x) \leq \phi(x + \xi_0^+) + q_0, \tag{1.8.26}
\]

with \( 0 \leq q_0 \leq 1/12 \). Then, we have

\[
\phi(x + \xi^-(t)) - q(t) \leq u_0(x) \leq \phi(x + \xi^+(t)) + q(t). \tag{1.8.27}
\]

with \( q(t) = q_0 e^{-t} \), and

\[
\xi^+(t) = \frac{q_0}{k_0} (1 + M_f)(1 - e^{-t}), \quad \xi^-(t) = \xi_0^- - \frac{q_0}{k_0} (1 + M_f)(1 - e^{-t}). \tag{1.8.28}
\]

One issue here is that the gap between \( \xi^+(t) \) and \( \xi^-(t) \) is not decreasing in time but rather increasing – the opposite of what we want! Our goal is to show that we can actually choose \( \xi^+(t) \) and \( \xi^-(t) \) in (1.8.27) so that the "sub-solution/super-solution gap" \( \xi^+(t) - \xi^-(t) \) would decrease to zero as \( t \to +\infty \) – this will prove convergence of the solution to a translate of \( \phi \).
The mechanism to decrease this difference will be kindly provided by the strong maximum principle. The idea is to iteratively trap the solutions, at an increasing sequence of times, between translates of \( \phi_0 \), that will come closer and closer to each other, thus implying the convergence. However, as there will be some computations, it is worth explaining beforehand what the main idea is, and which difficulties we will see.

Let us consider for the moment a slightly better situation than in Lemma 1.8.4 – assume that \( u_0(x) \) is actually trapped between \( \phi(x + \xi_0^-) \) and \( \phi(x + \xi_0^+) \), without the need for an additional term \( q(t) \):

\[
\phi(x + \xi_0^-) \leq u_0(x) \leq \phi(x + \xi_0^+). \tag{1.8.29}
\]

Then, \( u(t, x) \) is at a positive distance from one of the two translates, on compact sets, at least for \( 0 \leq t \leq 1 \), say, \( \phi(x + \xi_0^+) \). This is where the strong maximum principle strikes: at \( t = 1 \), it will make the infimum of \( \phi(x + \xi_0^+) - u(t, x) \) strictly positive, at least on a large compact set. We would like to think that then we may translate \( \phi(x + \xi_0^+) \) to the right a little, decreasing \( \xi_0^+ \), while keeping it above \( u(1, x) \). The catch is that, potentially, the tail of \( u(1, x) \) – that we do not control very well at the moment – might go over \( \phi(x + \xi) \), as soon as \( \xi \) is just a little smaller than \( \xi_0^+ \). Let us ignore this, and assume that magically we have

\[
\phi(x + \xi_0^-) \leq u(1, x) \leq \phi(x + \xi_1^+), \tag{1.8.30}
\]

with

\[
\xi_1^+ = \xi_0^+ - \delta(\xi_0^+ - \xi_0^-), \tag{1.8.31}
\]

with some \( \delta > 0 \). If we believe in this scenario, we might just as well hope that the situation may be iterated: at the time \( t = n \), we have

\[
\phi(x + \xi_n^-) \leq u(n, x) \leq \phi(x + \xi_n^+), \tag{1.8.32}
\]

with

\[
\xi_{n+1}^+ - \xi_{n+1}^- \leq (1 - \delta)(\xi_n^+ - \xi_n^-). \tag{1.8.33}
\]

This would imply a geometric decay of \( \xi_n^+ - \xi_n^- \) to zero, which, in turn, would imply the exponential convergence of \( u(t, x) \) to a translate of \( \phi \).

The gap in the previous argument is, of course, in our lack of control of the tail of \( u(t, x) \) that prevents us from being sure that (1.8.30), with \( \xi_1^+ \) as in (1.8.31), holds everywhere on \( \mathbb{R} \) rather than on a compact set. There is no way we can simply ignore it: we will see in Chapter ?? that the dynamics of many respectable equations is controlled exactly by the tail of its solutions. Such will not be the case here, but we will have to go through the pain of controlling the tail of \( u \) at every step. This leads to the somewhat heavy proof that follows, which is itself a simplified version of [100], where global exponential stability of transition waves is shown. However, there is essentially no other idea than what we have just explained, the rest are just technical embellishments. The reader should also recall that we have already encountered a tool for the tail-control in the Allen-Cahn equation: Corollary ?? in Chapter ?? served exactly that purpose in the proof of Theorem ??.

We are going to use something very similar here.
The proof of Theorem 1.8.2

As promised, the strategy is a refinement of the proof of Lemma 1.8.4. We will construct a sequence of sub-solutions \( u_n \) and super-solutions \( \overline{u}_n \) defined for \( t \geq T_n \), such that

\[
 u_n(t, x) \leq u(t, x) \leq \overline{u}_n(t, x) \quad \text{for} \quad t \geq T_n. \tag{1.8.34}
\]

Here, \( T_n \to +\infty \) is a sequence of times with

\[
 T_n + T \leq T_{n+1} \leq T_n + 2T, \tag{1.8.35}
\]

and the time step \( T > 0 \) to be specified later on. The sub- and super-solutions will be of the familiar form (1.8.27)-(1.8.28):

\[
 u_n(t, x) = \phi(x + \xi_n^+(t)) - q_n e^{-(t-T_n)}, \quad \overline{u}_n(t, x) = \phi(x + \xi_n^-(t)) + q_n e^{-(t-T_n)}, \quad t \geq T_n, \tag{1.8.36}
\]

with \( \xi_n^\pm(t) \) as in (1.8.28):

\[
 \xi_n^+(t) = \xi_n^+ + \frac{q_n}{k_0}(1 + M_f)(1 - e^{-(t-T_n)}), \quad \xi_n^-(t) = \xi_n^- - \frac{q_n}{k_0}(1 + M_f)(1 - e^{-(t-T_n)}). \tag{1.8.37}
\]

The reader has surely noticed a slight abuse of notation: we denote by \( \xi_n^\pm \) the values of \( \xi_n^\pm(t) \) at the time \( t = T_n \). This allows us to avoid introducing further notation, and we hope it does not cause too much confusion.

Our plan is to switch from one pair of sub- and super-solutions to another at the times \( T_n \), and improve the difference in the two shifts at the "switching" times, to ensure that

\[
 \xi_{n+1}^+ - \xi_{n+1}^- \leq (1 - \delta)(\xi_n^+ - \xi_n^-), \tag{1.8.38}
\]

with some small but fixed constant \( \delta > 0 \) such that

\[
 e^{-T} \leq c_T \delta \leq \frac{1}{4}. \tag{1.8.39}
\]

The constant \( c_T \) will also be chosen very small in the end – one should think of (1.8.39) as the requirement that the time step \( T \) is very large. This is natural: we can only hope to improve on the difference \( \xi_n^+ - \xi_n^- \), as in (1.8.38), after a very large time step \( T \). The shifts can be chosen so that they are uniformly bounded:

\[
 |\xi_n^\pm| \leq M, \tag{1.8.40}
\]

with a sufficiently large \( M \) – this follows from the bounds on the level sets of \( u(t, x) \) that we have already obtained. As far as \( q_n \) are concerned, we will ask that

\[
 0 \leq q_n \leq c_q \delta(\xi_n^+ - \xi_n^-), \tag{1.8.41}
\]

with another small constant \( c_q \) to be determined. Note that at \( t = 0 \) we may ensure that \( q_0 \) satisfies (1.8.41) simply by taking \( \xi_0^+ \) sufficiently positive and \( \xi_0^- \) sufficiently negative.

As we have uniform bounds on the location of the level sets of \( u(t, x) \), and the shifts \( \xi_n^\pm \) will be chosen uniformly bounded, as in (1.8.40), after possibly increasing \( R_0 \) in (1.8.11), we can ensure that

\[
 \phi(x + \xi_n^+(t)) \geq 11/12, \quad u(t, x) \geq 11/12, \quad \text{for} \quad x \geq R_0 \quad \text{and} \quad t \geq T_n, \tag{1.8.42}
\]
and

\[ -1 < \phi(x + \xi_n^+(t)) \leq 11/12, \quad -1 < u(t, x) \leq -11/12, \]  \quad \text{for } x \leq -R_0 \text{ and } t \geq T_n, \quad (1.8.43) \]

which implies

\[ f'(\phi(x + \xi_n^+(t))) \leq -1, \quad f'(u(t, x)) \leq -1, \]  \quad \text{for } |x| \geq R_0 \text{ and } t \geq T_n. \quad (1.8.44) \]

Let us now assume that at the time \( t = T_n \) we have the inequality

\[ \phi(x + \xi_n^-) - q_n \leq u(T_n, x) \leq \phi(x + \xi_n^+) + q_n, \quad (1.8.45) \]

with the shift \( q_n \) that satisfies (1.8.41). Our goal is to find a time \( T_{n+1} \in [T_n + T, T_n + 2T] \), and the new shifts \( \xi_{n+1}^\pm \) and \( q_{n+1} \), so that (1.8.45) holds with \( n \) replaced by \( n + 1 \) and the new gap \( \xi_{n+1}^+ - \xi_{n+1}^- \) satisfies (1.8.38). We will consider two different cases.

**Case 1: the solution gets close to the super-solution.** Let us first assume that there is a time \( \tau_n \in [T_n + T, T_n + 2T] \) such that the solution \( u(\tau_n, x) \) is "very close" to the super-solution \( \bar{u}_n(\tau_n, x) \) on the interval \( \{|x| \leq R_0 + 1\} \). More precisely, we assume that

\[ \sup_{|x| \leq R_0 + 1} \left( \bar{u}_n(\tau_n, x) - u(\tau_n, x) \right) \leq \delta(\xi_n^+ - \xi_n^-). \quad (1.8.46) \]

We will show that in this case we may take \( T_{n+1} = \tau_n \), and set

\[ \xi_{n+1}^+ = \xi_n^+(\tau_n), \quad \xi_{n+1}^- = \xi_n^- + (\xi_n^+(\tau_n) - \xi_n^+) + \delta(\xi_n^+ - \xi_n^-), \quad (1.8.47) \]

as long as \( \delta \) is sufficiently small, making sure that

\[ \xi_{n+1}^+ - \xi_{n+1}^- = (1 - \delta)(\xi_n^+ - \xi_n^-), \quad (1.8.48) \]

and also choose \( q_{n+1} \) so that

\[ q_{n+1} = c_q \delta(\xi_{n+1}^+ - \xi_{n+1}^-). \quad (1.8.49) \]

As far as the super-solution is concerned, we note that

\[ u(\tau_n, x) \leq \phi(x + \xi_{n}^+(\tau_n)) + q_ne^{-(t-T_n)} \leq \phi(x + \xi_{n}^+(\tau_n)) + c_q \delta(\xi_n^+ - \xi_n^-)e^{-T} \]
\[ \leq \phi(x + \xi_{n}^+(\tau_n)) + q_{n+1}, \quad (1.8.50) \]

for all \( x \in \mathbb{R} \), provided that \( T \) is sufficiently large, independent of \( n \).

For the sub-solution, we first look at what happens for \( |x| \leq R_0 + 1 \) and use (1.8.46):

\[ u(\tau_n, x) \geq \phi(x + \xi_{n}^+(\tau_n)) + q_ne^{-(\tau_n-T_n)} - \delta(\xi_n^+ - \xi_n^-), \]  \quad \text{for all } |x| \leq R_0 + 1. \quad (1.8.51) \]

Thus, for \( |x| \leq R_0 + 1 \) we have

\[ u(\tau_n, x) \geq \phi(x + \xi_{n}^+(\tau_n)) - \delta(\xi_n^+ - \xi_n^-) \geq \phi(x + \xi_{n}^+ - C_R(\xi_n^+ - \xi_n^-)) \geq \phi(x + \xi_{n+1}^-), \quad (1.8.52) \]

with the constant \( C_R \) that depends on \( R_0 \), as long as \( \delta > 0 \) is sufficiently small.
It remains to look at $|x| \geq R_0 + 1$. To this end, recall that
\[ u(\tau_n, x) \geq \phi(x + \xi_n^-(\tau_n)) - q_n e^{-(\tau_n - T_n)}, \quad \text{for all } x \in \mathbb{R}, \] (1.8.53)
so that, as follows from the definition of $\xi_n^-(t)$, we have
\[ u(\tau_n, x) \geq \phi(x + \xi_n^- - Cq_n) - q_n e^{-2T}, \quad \text{for all } x \in \mathbb{R}. \] (1.8.54)
Observe that, as $\phi(x)$ is approaching $\pm 1$ as $x \to \pm \infty$ exponentially fast, there exist $\omega > 0$ and $C > 0$ such that, taking into account (1.8.41) we can write for $|x| \geq R_0 + 1$:
\[ \phi(x + \xi_n^- - Cq_n) \geq \phi(x + \xi_n^- + (\xi_n^+(\tau_n) - \xi_n^-) + \delta(\xi_n^+ - \xi_n^-)) - C\delta e^{-\omega R_0 (\xi_n^+ - \xi_n^-)} \geq \phi(x + \xi_{n+1}^-) - q_{n+1}, \] (1.8.55)
as long as $R_0$ is large enough. Here, we have used $\xi_{n+1}^-$ and $q_{n+1}$ as in (1.8.47) and (1.8.49).
We conclude that
\[ u(\tau_n, x) \geq \phi(x + \xi_{n+1}^-) - q_{n+1}, \quad \text{for } |x| \geq R_0 + 1. \] (1.8.56)
Summarizing, if (1.8.46) holds, we set $T_{n+1} = \tau_n$, define the new shifts $\xi_{n+1}^\pm$ as in (1.8.47) and (1.8.49), which ensures that the “shift gap” is decreased by a fixed factor, so that (1.8.48) holds, and we can restart the argument at $t = T_{n+1}$, because
\[ \phi(x + \xi_{n+1}^-) - q_{n+1} \leq u(T_{n+1}, x) \leq \phi(x + \xi_{n+1}^+ + q_{n+1}), \quad \text{for all } x \in \mathbb{R}. \] (1.8.57)
Of course, if at some time $\tau_n \in [T_n + T, T_n + 2T]$ we have, instead of (1.8.46) that
\[ \sup_{|x| \leq R_0 + 1} \left( u(\tau_n, x) - u(\tau_n, x) \right) \leq \delta(\xi_n^+ - \xi_n^-), \] (1.8.58)
then we could repeat the above argument essentially verbatim, using the fact that now the solution is very close to the sub-solution on a very large interval.

**Case 2: the solution and the super-solution are never too close.** Next, let us assume that for all $t \in [T_n + T, T_n + 2T]$, we have
\[ \sup_{|x| \leq R_0 + 1} \left( \bar{u}_n(t, x) - u(t, x) \right) \geq \delta(\xi_n^+ - \xi_n^-). \] (1.8.59)
Because $\xi_n^+(t)$ is increasing, we have, for all $|x| \leq R_0 + 1$ and $t \in [T_n + T, T_n + 2T]$: \[ \bar{u}_n(t, x) \leq \phi(x + \xi_n^+(T_n + 2T)) + q_n e^{-T} \leq \phi(x + \xi_n^+(T_n + 2T) + q_n e^{-T} \rho_0), \] (1.8.60)
with
\[ \rho_0 = \left( \inf_{|x| \leq R_0 + M + 10} \phi'(x) \right)^{-1}. \] (1.8.61)
Here, $M$ is the constant in the upper bound (1.8.40) for $\xi_n^+$. Note that by choosing $T$ sufficiently large we can make sure that the argument in $\phi$ in the right side of (1.8.60) is within the range of the infimum in (1.8.61). The function
\[ w_n(t, x) = \phi(x + \xi_n^+(T_n + 2T) + q_n e^{-T} \rho_0) - u(t, x). \]
that appears in the right side of (1.8.60) solves a linear parabolic equation
\[
\partial_tw_n - \partial_{xx}w_n + a_n(t, x)w_n = 0, \tag{1.8.62}
\]
with the coefficient \(a_n\) that is bounded in \(n, t\) and \(x\):
\[
a_n(t, x) = \frac{f(\phi(x + \xi_n^+(T_n + 2T) + q_ne^{-T}\rho_0)) - f(u(t, x))}{\phi(x + \xi_n^+(T_n + 2T) + q_ne^{-T}\rho_0) - u(t, x)}.
\tag{1.8.63}
\]
It follows from assumption (1.8.59) and (1.8.60) that
\[
\sup_{|x|\leq R_0 + 1} w_n(t, x) \geq \delta(\xi_n^+ - \xi_n^-), \quad \text{for all } t \in [T_n + T, T_n + 2T],
\tag{1.8.64}
\]
but in order to improve the shift, we would like to have not the supremum but the infimum in the above inequality. And here the Harnack inequality comes to the rescue: we will use Theorem 1.6.9 for the intervals \(|x| \leq R_0 + 1\) and \(|x| \leq R_0\). For that, we need to make sure that at least a fraction of the supremum in (1.8.64) is attained on \([-R_0, R_0]\): there exists \(k_1\) so that
\[
\sup_{|x|\leq R_0} w_n(t, x) \geq k_1\delta(\xi_n^+ - \xi_n^-), \quad \text{for all } T_n + T \leq t \leq T_n + 2T.
\tag{1.8.65}
\]
However, if there is a time \(T_n + T \leq s_n \leq T_n + 2T\) such that
\[
\sup_{|x|\leq R_0} w_n(s_n, x) \leq \frac{\delta}{2}(\xi_n^+ - \xi_n^-), \tag{1.8.66}
\]
then we have
\[
\bar{u}(s_n, x) - u(s_n, x) \leq \frac{\delta}{2}(\xi_n^+ - \xi_n^-) \quad \text{for all } |x| \leq R_0.
\tag{1.8.67}
\]
This is the situation we faced in Case 1, and we can proceed as in that case. Thus, we may assume that
\[
\sup_{|x|\leq R_0} w_n(t, x) \geq \frac{\delta}{2}(\xi_n^+ - \xi_n^-) \quad \text{for all } T_n + T \leq t \leq T_n + 2T.
\tag{1.8.68}
\]
In that case, we may apply the Harnack inequality of Theorem 1.6.9 to (1.8.62) on the intervals \(|x| \leq R_0 + 1\) and \(|x| \leq R_0\): there exists a Harnack constant \(h_{R_0}\) that is independent of \(T\), such that
\[
w_n(t, x) \geq h_{R_0}\delta(\xi_n^+ - \xi_n^-), \quad \text{for all } t \in [T_n + T + 1, T_n + 2T] \text{ and } |x| \leq R_0.
\tag{1.8.69}
\]

**Exercise 1.8.7** Show that, as a consequence, we can find \(\rho_1 > 0\) that depends on \(R_0\) but not on \(n\) such that for \(|x| \leq R_0\) and \(T_n + T + 1 \leq t \leq T_n + 2T\), we have
\[
\bar{w}_n(t, x) = \phi(x + \xi_n^+(T_n + 2T) + \rho_0 e^{-T}q_n - \rho_1 h_{R_0}\delta(\xi_n^+ - \xi_n^-)) - u(t, x) \geq 0.
\tag{1.8.70}
\]
Let us now worry about what $\tilde{w}_n$ does for $|x| \geq R_0$. In this range, the function $\tilde{w}_n$ solves another linear equation of the form

$$\partial_t \tilde{w}_n - \partial_{xx} \tilde{w}_n + \tilde{a}_n(t, x) \tilde{w}_n = 0,$$

(1.8.71)

with $\tilde{a}_n(t, x) \geq 1$ that is an appropriate modification of the expression for $a_n(t, x)$ in (1.8.63). In addition, at the boundary $|x| = R_0$, we have $\tilde{w}_n(t, x) \geq 0$, and at the time $t = T_n + T$, we have an estimate of the form

$$\tilde{w}_n(T_n + T, x) \geq -K(\xi_n^+ - \xi_n^-), \quad |x| \geq R_0.$$  

(1.8.72)

**Exercise 1.8.8** What did we use to get (1.8.72)?

Therefore, the maximum principle applied to (1.8.71) implies that

$$\tilde{w}_n(T_n + 2T, x) \geq -Ke^{-T}(\xi_n^+ - \xi_n^-), \quad |x| \geq R_0.$$  

(1.8.73)

We now set $T_{n+1} = T_n + 2T$. The previous argument shows that we have

$$u(T_{n+1}, x) \leq \phi(x + \xi^n_+(T_{n+1}) + \rho_0e^{-T}q_n - hR_0\delta(\xi_n^+ - \xi_n^-)) + q_{n+1},$$

(1.8.74)

with

$$0 \leq q_{n+1} \leq Ke^{-T}(\xi_n^+ - \xi_n^-).$$

(1.8.75)

In addition, we still have the lower bound:

$$u(T_{n+2}, x) \geq \phi(x + \xi^n_-(T_{n+1})) - e^{-T}q_n.$$  

(1.8.76)

It only remains to define $\xi^{\pm}_{n+1}$ and $q_{n+1}$ properly, to convert (1.8.74) and (1.8.76) into the form required to restart the iteration process. We take

$$q_{n+1} = \max(e^{-T}q_n, Ke^{-T}(\xi_n^+ - \xi_n^-)), \quad \xi_{n+1}^+ = \xi_n^-(T_{n+1}),$$

(1.8.77)

and

$$\xi_{n+1}^- = \xi_n^+(T_{n+1}) + \rho_0e^{-T}q_n - hR_0\rho_1\delta(\xi_n^+ - \xi_n^-).$$  

(1.8.78)

It is easy to see that assumption (1.8.41) holds for $q_{n+1}$ provided we take $T$ sufficiently large, so that

$$e^{-T} \ll c_q.$$  

(1.8.79)

The main point to verify is that the contraction in (1.8.38) does happen with the above choice. We recall (1.8.37):

$$\xi_n^+(T_{n+1}) = \xi_n^+ + \frac{q_n}{k_0}(1 + M_f)(1 - e^{-2T}), \quad \xi_n^-(T_{n+1}) = \xi_n^- - \frac{q_n}{k_0}(1 + M_f)(1 - e^{-2T}).$$

(1.8.80)

Hence, in order to ensure that

$$\xi_{n+1}^+ - \xi_{n+1}^- \leq (1 - \frac{hR_0\rho_1\delta}{2})(\xi_n^+ - \xi_n^-),$$

(1.8.81)

it suffices to make sure that the term $hR_0\rho_1\delta(\xi_n^+ - \xi_n^-)$ dominates all the other multiples of $\delta(\xi_n^+ - \xi_n^-)$ in the expression for the difference $\xi_{n+1}^+ - \xi_{n+1}^-$ that come with the opposite sign. However, all such terms are multiples of $q_n$, thus it suffices to make sure that the constant $c_q$ is small, which, in turn, can be accomplished by taking $T$ sufficiently large. This completes the proof. □
1.8.2 Spreading in unbalanced Allen-Cahn equations, and related models

Let us now discuss, informally, what one would expect, from the physical considerations, to happen to the solution of the initial value problem if the balance condition (1.8.6) fails, that is,
\[ \int_{-1}^{1} f(u) du \neq 0. \] (1.8.82)

To be concrete, let us consider the nonlinearity \( f(u) \) of the form
\[ f(u) = (u + 1)(u + a)(1 - u), \] (1.8.83)
with \( a \in (0, 1) \), so that \( u = \pm 1 \) are still the two stable solutions of the ODE
\[ \dot{u} = f(u), \]
but instead of (1.8.6) we have
\[ \int_{-1}^{1} f(u) du > 0. \]

As an indication of what happens we give the reader the following exercises. They are by no means short but they can all be done with the tools of this section, and we strongly recommend them to a reader interested in understanding this material well.

**Exercise 1.8.9** To start, show that for \( f(u) \) given by (1.8.83), we can find a special solution \( u(t, x) \) of the Allen-Cahn equation (1.8.2):
\[ u_t = u_{xx} + f(u), \] (1.8.84)
of the form
\[ u(t, x) = \psi(x + ct), \] (1.8.85)
with \( c > 0 \) and a function \( \psi(x) \) that satisfies
\[ c\psi' = \psi'' + f(\psi), \] (1.8.86)
together with the boundary condition
\[ \psi(x) \to \pm 1, \text{ as } x \to \pm \infty. \] (1.8.87)

Solutions of the form (1.8.85) are known as traveling waves. Show that such \( c \) is unique, and \( \psi \) is unique up to a translation: if \( \psi_1(x) \) is another solution of (1.8.86)-(1.8.87) with \( c \) replaced by \( c_1 \), then \( c = c_1 \) and there exists \( x_1 \in \mathbb{R} \) such that \( \psi_1(x) = \psi(x + x_1) \).

**Exercise 1.8.10** Try to modify the proof of Lemma 1.8.4 to show that if \( u(t, x) \) is the solution of the Allen-Cahn equation (1.8.84) with an initial condition \( u_0(x) \) that satisfies (1.8.8):
\[ u_0(x) \to \pm 1, \text{ as } x \to \pm \infty, \] (1.8.88)
then we have
\[ u(t, x) \to 1 \text{ as } t \to +\infty, \text{ for each } x \in \mathbb{R} \text{ fixed.} \quad (1.8.89) \]

It should be helpful to use the traveling wave solution to construct a sub-solution that will force (1.8.89). Thus, in the "unbalanced" case, the "more stable" of the two states \( u = -1 \) and \( u = +1 \) wins in the long time limit. Show that the convergence in (1.8.89) is not uniform in \( x \in \mathbb{R} \).

**Exercise 1.8.11** Let \( u(t, x) \) be a solution of (1.8.84) with an initial condition \( u_0(x) \) that satisfies (1.8.88). Show that for any \( c' < c \) and \( x \in \mathbb{R} \) fixed, we have
\[ \lim_{t \to +\infty} u(t, x - c't) = 1, \quad (1.8.90) \]
and for any \( c' > c \) and \( x \in \mathbb{R} \) fixed, we have
\[ \lim_{t \to +\infty} u(t, x - c't) = -1. \quad (1.8.91) \]

**Exercise 1.8.12** Let \( u(t, x) \) be a solution of (1.8.84) with an initial condition \( u_0(x) \) that satisfies (1.8.88). Show that there exists \( x_0 \in \mathbb{R} \) (which depends on \( u_0 \)) so that for all \( x \in \mathbb{R} \) fixed we have
\[ \lim_{t \to +\infty} u(t, x - ct) = \psi(x + x_0). \quad (1.8.92) \]

### 1.8.3 When the medium is inhomogeneous: pulsating waves

We will be interested, in this final section, in equations of the form
\[ u_t - u_{xx} = f(x, u), \quad t > 0, \ x \in \mathbb{R}, \quad (1.8.93) \]
with \( f \) 1-periodic in the variable \( x \). We will assume the following form \( f \): there is \( \theta \in (0, 1) \) such that \( f(x, u) \equiv 0 \) if \( u < \theta \), and \( f(x, u) > 0 \) if \( u > \theta \). In the vicinity of \( \theta \) there holds
\[ f_u(x, u) \geq \frac{f(x, u)}{u}. \quad (1.8.94) \]

Finally we assume that \( f(x, 1) \equiv 0 \), and that there is \( \alpha > 0 \) such that \( f_u(x, 1) < -\alpha \) for all \( x \in \mathbb{R} \). Of course, the set of such functions is by no means empty. For instance, (1.8.94) is true when
\[ f(x, u) = a(x)(u - \theta)^p_+, \]
with \( a(x) > 0 \), 1-periodic, and \( p \) a sufficiently large integer.

We wish to understand the large time behavior of a solution \( u(t, x) \) which, at time \( t = 0 \), tends to 1 as \( x \to -\infty \) and to 0 as \( x \to +\infty \). Clearly, the large time asymptotics cannot be given by a traveling wave, as the function \( f \) depends explicitly on \( x \). It turns out that, for such a nonlinearity, there is a special class of solutions generalizing traveling waves, these are called pulsating waves. The reason for that is simple: these solutions will be shown to be periodic in a well chosen Galilian reference frame. We will also show that, in fact, they attract all the solutions that, initially, has the afore-mentioned behavior.
The notion of pulsating waves was introduced by Xin at the beginning of the 90’s, see for instance the review paper [135]. It was much extended and generalized by Berestycki and Hamel, especially for models posed in the whole space, where things are more subtle than in the one-dimensional model that we present. See [14] for a much detailed account of the theory - that is, by the way, still evolving. We note that, even in one space dimension, relaxing the assumptions that we have (for instance, asking \( f(x, 1) \equiv 0 \) instead of a more natural looking assumption allowing uniform boundedness of the solutions) made may not necessarily modify the nature of the results that we are about to present, but it would certainly involve a good deal of additional work, which would not be in the sprit of this chapter. Once again, we refer to [14] for an account of what happens in the most general situations.

**Exercise 1.8.13** Under the stated assumptions on \( f \), the only solutions \( \phi(x) \in [0, 1] \) of
\[
-\phi'' = f(x, \phi), \quad x \in \mathbb{R},
\]
are all the constants between 0 and \( \theta \), and 1.

**Exercise 1.8.14** Consider a nonlinearity \( f(x, u) \) that satisfies all the above assumptions, except \( f(x, 1) \equiv 0 \). Assume, though, the existence of \( f_0(x, u) \) that satisfies them all, and that is additionally close to \( f \) in the \( C^1 \) norm.

1. Show that the (1.8.95) has a unique minimal nonconstant solution \( \phi_+(x) \), that is \( C^2 \)-close to 1.

2. Show that the first periodic eigenvalue of \( -\partial_{xx} - f_u(x, \phi_+) \) is negative.

Why have we suddenly shifted from nonlinearities of the Allen-Cahn type to nonnegative nonlinearities of the afore-mentioned type? As the reader may have guessed in view of the previous exercise, we will construct waves that connect various solutions of (1.8.95), but we will not be necessarily very easy to count - a very interesting exercise of course, but once again outside the scope of this section. We may still propose the following exercise to the interested reader.

**Exercise 1.8.15** Let \( f(x, u) \) be \( C^2 \)-close to an unbalanced Allen-Cahn nonlinearity. Find all the solutions of (1.8.95), as well as, for each of them, the sign of \( -\partial_{xx} - f_u(x, \phi_+) \). Hint: there is a catch, the limiting Allen-Cahn equation has nonconstant solutions!

These informal preliminaries being dealt with, let us now define precisely the object that will be in our proccupations.

**Definition 1.8.16** A function \( \phi(t, x) \) is a pulsating wave of (1.8.93), with speed \( c > 0 \), and connecting 1 to 0 if it satisfies the following properties:

1. \( \phi \) solves (1.8.93) of \( \mathbb{R} \times \mathbb{R} \),

2. it is \( 1/c \)-periodic in a Galilean reference frame with speed \( c \),

3. we have
\[
\lim_{x \to -\infty} \phi(t, x) = 1, \quad \lim_{t \to +\infty} \phi(t, x) = 0,
\]
pointwise in \( t \).
Let us make the following simple remarks.

**Remark 1.8.17**  
- The parabolic regularity and the time periodicity imply that the limits are in fact uniform in $t$, as soon as one looks at the phenomenon in the reference frame with speed $c$.

- For a pulsating wave with speed $c$, the time-periodicity $1/c$ is the only possible one. This forced by the 1-periodicity in $x$: in the reference frame with speed $ct$, the function $f(x, \phi)$ becomes $f(x+ct, \phi)$. If $\phi$ is $T$-periodic, then $f(x+ct,.)$ should also be $T$-periodic. Another way to view it is the following: the speed being $c$, the wave takes the time $1/c$ to cover the cell of length 1. When this is achieved, it retrieves its original shape.

- If $\phi(t,x)$ is a pulsating wave with speed $c$, then any translate in time $\phi(t+t_0,x)$ is also a pulsating wave with speed $c$.

And we may state the main achievement of the section, namely the existence and uniqueness of pulsating waves.

**Theorem 1.8.18**  
Problem (1.8.93) has a one-dimensional family of pulsating waves $\phi(t,x)$ (one can be deduced from another by a translate in $t$) with speed $c$, connecting 1 to the left to 0 to the right. The speed $c$ is unique. moreover we have $\partial_t \phi > 0$.

The last statement is a striking parallel with the main property of the traveling waves, namely that they are decreasing in $x$. Of course, here, monotonicity in $x$ is not true, what replaces it is monotonicity on $t$.

Monotonicity and uniqueness are very easily proved by the sliding ideas that we have exposed at length in the first chapter, it is therefore a good time to propose a last refresh to the reader. The idea is the same in both cases: take two different waves, and prove that a sufficiently large translate of one is below the other. Then, translate back until it is not possible anymore, and derive a contradiction. Let us give some more details and prove monotonicity first. let $\phi$ be a wave with speed $c > 0$, we infer, simply by the fact that they have limits as $x \to \pm \infty$, the existence of a large $T > 0$ such that

$$\phi(t + T, x) \leq \phi(t, x) \text{ for all } (t, x) \in [0, \frac{1}{c}] \times [-M, M],$$

and we may choose $M > 0$ such that

$$\phi(t, x) \leq \frac{\theta}{2} \text{ for } t \in [0, \frac{1}{c}], x \geq M, \quad \phi(t, x) \geq 1 - \delta \text{ for } t \in [0, \frac{1}{c}], x \leq -M,$$

with $f_u(x, u) \leq \frac{\alpha}{2}$ for $u \geq 1 - \delta$. In the reference frame moving with velocity $c$, both the wave (still denoted by $\phi$) and its translate solve:

$$\partial_t \phi - c\partial_x \phi - \partial_{xx} \phi = f(x - ct, \phi), \quad (1.8.96)$$

and $\psi(t, x) : \phi(t + T, x) - \phi(t, x)$ solve

$$\partial_t \psi - c\partial_x \psi - \partial_{xx} \psi - a(t, x)\psi = 0, \quad (1.8.97)$$
where \( a(t, x) \) is, as usual, some convex combination of \( f_u(x - ct, \phi) \). Note that now, both \( \phi \) and its translate are \( 1/c \)-periodic in \( t \). Therefore, \( \psi(t, x) \geq 0 \) for \( t \geq 0 \), \( -M \leq x \leq M \). For \( x \leq -M \) we have \( a(t, x) \geq \alpha/2 \), while for \( x \geq M \) we have an inequation of the form
\[
\partial_t \psi - c \partial_x \psi - \partial_{xx} \psi \geq 0. \tag{1.8.98}
\]
Sending \( t \to +\infty \) allows us to infer, from (1.8.97) and (1.8.98):
\[
\lim inf_{t \to +\infty} \psi(t, x) \geq 0.
\]

**Exercise 1.8.19** Work out the details by placing a subsolution below \( \psi \) on \( \mathbb{R}^+ \times (-\infty, -M) \) and \( (M, +\infty) \). Nothing difficult here, the only intermediate step is to prove that the solution of the Dirichlet advection-diffusion equation
\[
\begin{align*}
v_t - v_{xx} - cv_x &= 0, & t > 0, & x \geq M \\
v(t, 0) &= 0 \\
\lim_{t \to +\infty} v(0, x) &= 0
\end{align*}
\tag{1.8.99}
\]
tends to 0 as \( t \) increases to infinity, uniformly in \( x \). One may proceed as follows.

1. Show that it is enough to assume \( v(0, x) \geq 0 \) and to prove that the lim sup is zero.

2. Construct a super-solution of the form \( e^{-\beta t - \gamma x} \), give explicit values to \( \beta \) and \( \gamma \).

3. Prove the result if \( v(0, x) \) is compactly supported.

4. Show that (1.8.99) generates a weakly contracting semigroup, that is: \( \|v_1(t, .) - v_2(t, .)\|_{\infty} \leq \|v_1(0, .) - v_2(0, .)\|_{\infty} \).

5. Conclude.

In the area \( \{ t > 0, \ x \geq -M \} \), use the fact that \( a(t, x) \leq -\alpha/2 \).

Sending \( t \) to infinity and remembering the \( 1/c \)-periodicity of \( \psi \) allows us to infer that, actually, we have \( \phi(t + T, x) \geq \psi(t, x) \) everywhere. And so, there is \( t_{\text{min}} \geq 0 \) such that \( \phi(t + t_{\text{min}}, x) \leq \phi(t, x) \) everywhere. Assume \( t_{\text{min}} > 0 \), the inequality has to be strict at at least one point, otherwise we would have \( \phi(t + t_{\text{min}}, x) = \phi(t, x) \), something that the positive speed of propagation as well as the limits as \( x \to \pm \infty \) oppose vehemently. But then, the \( \leq \) sign should be replaced by a \(< \) sign, just to appease the strong maximum principle. So, on a large compact set that we still call \([-M, M]\), we may translate a little more and obtain, for small \( \delta > 0 \):
\[
\phi(t + t_{\text{min}} - \delta, x) \leq \phi(t, x), \quad t \in \mathbb{R}, \ -M \leq x \leq M.
\]
The previous argument can be repeated, so that the inequality holds in fact on \( \mathbb{R} \times \mathbb{R} \). This contradicts the minimality of \( t_{\text{min}} \), ensuring \( \phi(t + \delta, x) \geq \phi(t, x) \) for all \( \delta > 0 \). This entails the monotonicity of the wave, and the same argument may be used to show uniqueness.

The uniqueness of the speed is hardly any more difficult. Let \( c_1 \) and \( c_2 \) be two potential wave speeds, write down (1.8.93) in the reference frame moving with speed \( c_1 \). Translate \( \phi_1 \) enough so that it is below \( \phi_2 \) in a large compact set, then use the following
Exercise 1.8.20 Construct a sub-solution to (1.8.96) under the form
\[ \phi_1(t, x) = \phi_1(t + \xi(t), x) - q e^{-\alpha t/2 - c^2 t/4} \Gamma(x) - q e^{-\alpha t/2 (1 - \Gamma(x - x_1))}, \]
where \( \Gamma \) is a smooth function that is equal to 1 on \( \mathbb{R}_- \) and 0 on \([1, +\infty)\). The constant \( x_1 \) should be large, and the constant \( q \) should be small. The inspiration should be taken from Lemma 1.8.4. The monotonicity in \( x \) of the traveling wave should now be replaced by the monotonicity in \( t \) of the pulsating wave.

Clearly, we may choose the constants to have \( \phi_1 \leq \phi_2 \). Sending \( t \to +\infty \) and using the \( 1/c \) periodicity yields \( \phi_1(t + t_0, x) \leq \phi_2(t, x) \), which implies \( c_1 \geq c_2 \). And so, \( c_1 = c_2 \) by symmetry.

It is now a good time to explain why pulsating waves exist. The construction that we are going to provide relies, once again, on rather simple comparison arguments. The idea is to solve the Cauchy Problem for (1.8.93), with a sub-solution as an initial datum. Thus we will obtain, for the resulting solution, monotonicity for free. We will then examine its behavior as \( |x| \) becomes very large, and see that it has the correct limits, and that they are taken uniformly with respect to \( t \). This can be viewed as a finite thickness property for the front, and it will provide a sufficient amount of compactness for us to prove that the limiting solution is nontrivial. A last effort, where we will again use comparison and sliding, will tell us that we have put our hand on the sought for pulsating wave.

The starting point is therefore the construction of a sub-solution. Pick \( \theta_1 \in (\theta, 1) \) close enough to 1, \( \beta > 0 \) small to be chosen later, and \( \alpha > 0 \) small so that we have
\[ f(x, u) \geq \alpha^2 (1 - u), \quad u \in [\theta_1, 1], \]
and define
\[ f(u) = \begin{cases} 
\alpha^2 (1 - u), & u \in [\theta_1, 1] \\
-\beta^2 u, & u \in [0, \theta_1). 
\end{cases} \]

Hence \( f(u) \leq f(x, u) \). A solution \( v(x) \) of
\[ -v'' = f(v), \quad \lim_{x \to -\infty} v(x) = 1 \]
is easily computed: we have
\[ v(x) = \frac{1}{2} \left( \theta_1 + \frac{1 - \theta_1}{\beta} \right) e^{-\beta x} + \frac{1}{2} \left( \theta_1 - \frac{1 - \theta_1}{\beta} \right) e^{\beta x}. \]
The sought for sub-solution is simply chosen as \( u(x) = v^+(x) \). Solve (1.8.93) with \( u(0, x) = u(x) \): we obtain a solution \( u(t, x) \) that satisfies \( \partial_t u \geq 0 \), moreover \( u(t, x) \) assumes the correct limits as \( x \to \pm \infty \); the trouble is, however, that there is no uniformity at this stage. Our main tool will be the following

Proposition 1.8.21 There is a function \( \lambda \mapsto q(\lambda) \), bounded and bounded away from 0 on every compact of \([0, 1)\), such that we have
\[ u_t(t, x) \geq q(\lambda) u(t, x) \text{ if } u(t, x) = \lambda. \]
Proof. To see that it is true on every compact set of \((\theta, 1)\) is not so difficult: pick \(\lambda \in (\theta, 1)\) and assume that there is a sequence \((t_n, x_n)\) such that \(u(t_n, x_n) = \lambda\) and \(\lim_{n \to +\infty} u(t_n, x_n) = 0\). Because \(u_t \geq 0\), the sequence of functions

\[ u_n(t, x) = u(t + t_n, x + x_n) \]

converges, uniformly on every compact, and up to the extraction of a subsequence, to a function \(u_\infty(t, x)\) that satisfies (1.8.93). Because of the Harnack inequality we have \((\partial_t u_n)\) converges to 0, so that \(u_\infty\) does not depend on \(x\), moreover \(u_\infty(0, 0) = \lambda\). The positivity of \(f\) yields a contradiction. Thus, the proposition is true for this range of \(\lambda\).

Now, pick \(\delta > 0\) small, and let us study what happens for the remaining values of \(\lambda\).

Define

\[ \Omega = \{(t, x) \in \mathbb{R} \times \mathbb{R} : u(t, x) < \theta + \delta\} \]

Obviously, (1.8.100) holds true on

\[ \{(t, x) \in \mathbb{R} \times \mathbb{R} : u(t, x) = \theta + \delta\}, \]

provided \(q(\theta + \delta)\) is chosen appropriately. The inequality also holds at \(t = 0\), by the definition of \(f\), with \(q(\lambda) = \beta^2\) for all \(\lambda \in [0, \theta]\). For short, we set

\[ q = \inf_{\lambda \in [0, \theta + \delta]} q(\lambda). \]

The function \(v(t, x) = u_t(t, x) - qu(t, x)\) solves, in \(\Omega\):

\[
\begin{align*}
v_t - v_{xx} &= f_u(x, u)v + (f_u(x, u) - \frac{f(x, u)}{u})u \\
&\geq f_u(x, u)v \text{ because of (1.8.94)}. \end{align*}
\]

Thus, \(v(t, x) \geq 0\) in \(\Omega\), what we wanted to prove. □

Proposition 1.8.21 is a very important property that will allow us to conclude almost effortlessly. Let us now denote, this time, \(\Gamma\) the set \(\{u < \theta\}\), and \(\Gamma\) its boundary within \(\mathbb{R}_+ \times \mathbb{R}\); because \(u_t \geq q\theta\) on \(\Gamma\) it is a smooth curve \(\{\tau(x), x \in \mathbb{R}\}\). We choose, once and for all:

\[ q = \inf\{q(\lambda), 0 \leq \lambda \leq \frac{1 + \theta}{2}\}. \tag{1.8.101} \]

We may always assume that \(\tau\) is defined on \([0, +\infty)\), with \(\tau(0) = 0\). The main consequence of Proposition 1.8.21 is the following statement, which says that the front has finite thickness.

Corollary 1.8.22 The level set \(\Gamma\) may be described as a smooth curve of the form \((t, X(t))\), with \(X(t) \geq \mu > 0\).

\[
\lim_{x \to X(t) + \infty} u(t, x) = 0, \quad \lim_{x \to X(t) - \infty} = 1, \tag{1.8.102}
\]

uniformly with respect to \(t \in \mathbb{R}_+\).

Proof. For small \(t > 0\), say, \(t \in [0, t_0]\), continuity implies that \(u_x(t, x) > 0\) when \((t, x) \in \Gamma\). This implies, for every \(t \in [0, t_0]\), the existence and uniqueness of the function \(X(t)\) defined
as in the proposition. And, for any \( t \) in this interval, the equation for \( u \) to the left of \( X(t) \) becomes

\[-u_{xx} + qu \leq u_t - u_{xx} = 0,\]

so that \( u(t, x) \leq \theta e^{-\sqrt{q(x - X(t))}} \), and \( u_x(t, X(t))) \leq -\theta \sqrt{q} \). What would make this beautiful estimate break down would be a time \( t_{\max} \geq t_0 \) and a point \( x_{\max} \) such that \( t_{\max} = \tau(x_{\max}) \) and \( \tau'(x_{\max}) = 0 \), in other words \( u_x(t_{\max}, x_{\max}) = 0 \). However, at time \( t_{\max} \), \( x_{\max} \) is still the rightmost point (or, at least, can be made the rightmost one if \( \tau' \) vanishes on an interval) \( x \) such that \( u(t_{\max}, x) = \theta \); this makes the previous argument work, and, in particular, \( u_x(t_{\max}, x_{\max}) \leq -\theta \sqrt{q} \). This is an impossibility. Consequently, the function \( X(t) \) is defined for all time, moreover we have

\[
\dot{X}(t) = -\frac{u_t(t, X(t))}{u_x(t, X(t))} \in \left[ \frac{q}{\left\| u_x \right\|_\infty}, \frac{\|u_t\|_\infty}{\inf_{\Gamma} u_x} \right],
\]

(1.8.103)

and each of the above quantities is a positive constant. Notice that we have also ruled out a scenario where a hole of points where \( u \) would drop below \( \theta \) would appear to the left of \( \Gamma \), simply because \( u_t > 0 \).

Our argument also shows that \( u(t, x) \) goes to 0 exponentially fast as \( x - X(t) \) becomes negative, so the only thing that should worry us now is the limit \( x - X(t) \to +\infty \). For this, it is convenient to write the equation for \( u \) in the reference frame following \( X(t) \):

\[
u_t - u_{xx} - \dot{X}(t)u_x = f(x + X(t), u),
\]

(1.8.104)

by the Implicit function Theorem there is \( \delta > 0 \) and \( \theta_1 \in (\theta, 1) \) such that

\[u(t, x) \geq \theta_1 \text{ for } x \leq -\delta.\]

The assumptions implies the existence of a small \( \alpha > 0 \) such that

\[f(x, u) \geq \alpha^2(1 - u) \text{ for } u \geq \theta_1.\]

(1.8.105)

The maximum principle implies that \( 1 - u(t, x) \leq v(t, x) \) on \( \mathbb{R}_+ \times (-\infty, \delta] \), where

\[
v_t - v_{xx} - \dot{X}(t)v_x + \alpha^2 v = 0, \quad t > 0, x < -\delta
\]

\[v(t, -\delta) = 1 - \theta_1 \]

\[v(0, x) = 1 - u(x).\]

(1.8.106)

The function \( \varpi(x) = (1 - \theta_1)e^{\varepsilon x} \) is a super-solution to (1.8.106) as soon as \( 0 < \varepsilon \leq 2\sqrt{\alpha} \), which implies the uniform exponential decay to 0 of \( 1 - u \). This proves the corollary. □

**Exercise 1.8.23** At the beginning of the proof, we merrily said "continuity implies that \( u_x(t, x) > 0 \) when \( (t, x) \in \Gamma' \)". Make this continuity argument a little more explicit.

To construct the pulsating wave, we send \( t \) to \( +\infty \); more precisely, we consider a sequence \( (t_n)_n \) going to \( +\infty \) such that the sequence

\[u_n(t, x) = u(t_n + t, [X(t_n)] + x)\]
converges, locally on every compact, to a function that we denote $\phi(t, x)$, as well as all its derivatives. We still denote by $X(t)$ the $\theta$-level set of $\phi$, we note that $\phi(t, x)$ and $X(t, x)$ enjoy all the properties listed in Proposition 1.8.21 and its corollary. Moreover, $\phi$ is now defined on the whole plane $\mathbb{R}^2$. It remains to see that it is the sought for pulsating wave, and this is where sliding will make a last appearance. In fact, our argument will be quite similar to the one we used to prove the Krein-Rutman Theorem. We claim the existence of a large $T > 0$ such that

$$\phi(t + T, x) \geq \phi(t, x - 1), \text{ for all } (t, x) \in \mathbb{R}^2. \quad (1.8.107)$$

For all $M > 0$, we set

$$\Omega_M = \{(t, x) \in \mathbb{R}^2 : -M \leq x - X(t) \leq M\}.$$ 

From Corollary 1.8.22, we may choose $M > 0$ large enough so that $u(t, x) = 0$ if $(t, x)$ is at the left of $\Omega_M$, whereas $u(t, x) \geq \theta_1$ if $(t, x)$ is at the left of $\Omega_M$. This being done, we notice that indeed, there is a large $T > 0$ such that 1.8.107 holds on $\Omega_M$. To see that the inequality holds everywhere, notice that an inequation for $v(t, x) = \phi(t + T, x) - \phi(t, x - 1)$ in the reference frame of $X(t)$ is

$$v_t - v_{xx} - \dot{X}(t)v_x + a(t, x) \geq 0,$$

with $a(t, x) = 0$ if $x \geq M$, and $a(t, x) \geq \alpha^2$ if $x \leq -M$. Notice finally the existence of $\mu > 0$ such that $\dot{X}(t) \geq \mu$. This said, pick any time $t \in \mathbb{R}$, we may assume for simplicity that $t = 0$, just by translation. At the time $-n$ ($n$ is any integer) there is $\varepsilon > 0$ independent of $n$ such that

$$u(-n, x) \leq \theta e^{-\varepsilon(x-M)} \text{ for } x \geq M,$$

and

$$1 - u(-n, x) \leq \theta_1 e^{\varepsilon(x+M)} \text{ for } x \geq -M.$$

**Exercise 1.8.24** Show the existence of a constant $K > 0$ and $\delta > 0$ such that

$$v(t, x) \geq -Ke^{-\delta(t+n)} \text{ for } t \geq -n \text{ and } x \leq -M,$$

whereas

$$v(t, x) \geq -Ke^{-\varepsilon x/2 - \delta(t+n)} \text{ for } t \geq -n \text{ and } x \geq M.$$

The second estimate is the one where $\dot{X}(t) \geq \mu$ is needed.

Sending $n$ to $+\infty$ reveals that $v(0, x) \geq 0$, so that the $T$ we have defined satisfies (1.8.107). The reader who has accepted to follow us up to this point will have no problem with this penultimate hurdle:

**Exercise 1.8.25** Call $1/c$ the smallest $T$ such that (1.8.107) holds. Then, for $T = 1/c$, the inequality becomes an equality.

Simply by uniqueness for the Cauchy Problem, we have

$$\phi(t + 1/c, x) = \phi(t, x - 1).$$

Thus, $\phi$ is 1/c-periodic in a Galilean reference frame with speed $c$, and we have found our pulsating wave.
Exercise 1.8.26 Prove that any solution of the Cauchy Problem (1.8.93) converges, exponentially fast in time, to a pulsating wave, provided that (i) $u(0, x)$ converges to 1 as $x \to \infty$, and (ii) $u(0, x)$ converges exponentially fast in $x$ to 0 as $x \to +\infty$. It will be useful to write the equation in the reference frame moving like $X(t)$, and to recycle the proof of Theorem 1.8.2 with the family of sub and super-solutions defined in Exercise 1.8.20. If you are challenged by the task, look at [100].

With this last result, we are leaving the fascinating world of invasion in reaction-diffusion models. We are not leaving the topic forever, we will meet it again in a few chapters, with equations that look very similar, but where both the mechanisms and the tools used for their investigation will be very different from those displayed here. As always, we do not claim that we have exhausted the subject, as multi-dimensional models - and also some questions in one dimensions - require arguments that are sometimes much more sophisticated than those presented here. However, the material of this chapter will be a good basis to a reader who will want to attack the numerous open questions of the field.
Chapter 2

Inviscid Hamilton-Jacobi equations

2.1 Introduction to the chapter

We will consider in this chapter the Hamilton-Jacobi equations

$$u_t + H(x, \nabla u) = 0$$ (2.1.1)

on the unit torus $\mathbb{T}^n \subset \mathbb{R}^n$, or, sometimes, in all of $\mathbb{R}^n$. Note that here, unlike in the viscous Hamilton-Jacobi equations we have considered in Chapter 1, the diffusion coefficient vanishes. There are two reasons to do that in this book, where diffusion is remarkably present everywhere. The first is to emphasize some of the difficulties and phenomena that one encounters in the absence of diffusion. Another is that, as we will see, a physically reasonable class of solutions to (2.1.1) behave very much like the solutions to a regularized problem

$$u_t^\varepsilon + H(x, \nabla u^\varepsilon) = \varepsilon \Delta u^\varepsilon,$$ (2.1.2)

with a small diffusivity $\varepsilon > 0$. Most of the techniques we have introduced so far rely on the positivity of the diffusion coefficient and will deteriorate badly when the diffusion coefficient is small. However, we will see that some of the bounds may survive even as the diffusion term vanishes, because they are helped by the nonlinear Hamiltonian $H(x, \nabla u)$. Obviously, not every nonlinearity is beneficial: for example, solutions to the linear advection equation

$$u_t + b(x) \cdot \nabla u(x) = 0,$$ (2.1.3)

are typically no better than the initial condition $u_0(x) = u(0, x)$, no matter how smooth the drift $b(x)$ is. Therefore, we will have to restrict ourselves to some class of Hamiltonians $H(x, p)$ that do help to regularize the problem. This nonlinear regularization effect is one of the main points of this chapter.

The organization of the chapter

As a warm-up, we will discuss in Section 2.3 a class of viscous Hamilton-Jacobi equations, that is, equations of the form (2.1.1) with a Laplacian added:

$$u_t + H(x, \nabla u) = \Delta u.$$ (2.1.4)
Armed with the knowledge gathered in Chapter 1, we will (not so easily, but also not at an immense cost) construct a particular type of solutions, that we will call wave solutions. Quite similarly to what happens for traveling waves for the parabolic equations, these waves will be unique up to a translation in time, and solutions to the Cauchy problem will converge exponentially fast to them. In the remainder of the chapter, we will keep these features in mind to investigate how far the behavior of solutions to the inviscid Hamilton-Jacobi equations deviates from this simple picture.

We will then go through the most naive approach, looking for the smooth solutions to (2.1.1) in Section 2.4. However, a reader familiar with the theory of conservation laws, would see immediately the connection between them and the Hamilton-Jacobi equations: in one dimension, $n = 1$, differentiating (2.1.1) in $x$, we get a conservation law for $v = u_x$:

$$v_t + (H(x,v))_x = 0.$$  (2.1.5)

The basic one-dimensional conservation laws theory tells us that it is reasonable to expect that $v(t,x)$ becomes discontinuous in $x$ at a finite time $t$, forming a shock, which means that we can not hope that solutions to the inviscid Hamilton-Jacobi equations are better than Lipschitz continuous generally. We will say a few words about the classical theory and explain why it breaks down very quickly. This is well-known, see for instance [56], where it is done very nicely. For the reader’s convenience, we recall the basics here. This somehow pessimistic message should, however, be softened: there are (perhaps, less well-known) instances where a nice classical theory can be developed, and we are going to discuss one such example here. This material will, hopefully, be a good introduction to the more abstract theory to come next.

We proceed with a discussion of the viscosity solutions of the first order Hamilton-Jacobi equations in Section 2.5. Similarly to the parabolic regularity theory in Chapter 1, it is impossible to give a reasonable overview of the state of the art in this field. Rather, we will focus on its most elementary aspect: that a viscosity solution is a solution obtained by sending a diffusion coefficient (viscosity) to zero. Our goal will be to convince the reader that, although the viscous terms will have disappeared from the equations, some nontrivial features remain, such as the large time convergence to a steady state. One may call this the Cheshire cat smile effect [29]. This is explored, once again, in stages, where we first give a relatively accurate account of the Cauchy problem without dwelling too much on technicalities.

With the solution theory of Sections 2.4 and 2.5 in hand, one may start looking for the long time behavior of the solutions we have constructed, and their convergence to plane waves. The first step in this direction, as in the viscous case, is to construct the plane waves, and consider the stationary version of (2.1.1):

$$H(x,\nabla u) = c, \quad x \in \mathbb{T}^n.$$  (2.1.6)

This we will be done in Section 2.6. After what we will have done in Section 2.3, it should be clear to the reader why (2.1.6) has a constant $c$ in the right side – solutions to (2.1.6) lead to the wave solutions for the time-dependent problem (2.1.1). As in the viscous case, we will prove that under reasonable assumptions, solutions to (2.1.6) exist only for a unique value of $c$ which has no reason to be equal to zero. Thus, the “standard” steady equation

$$H(x,\nabla u) = 0$$
typically would have no solutions.

The case of a strictly convex Hamiltonian is quite interesting, and has strong connections with dynamical systems. We are going to dwell on it in Section 2.7, and show surprising regularizing properties that are not due to diffusion anymore. After that, we will come back to the large time behavior of the solutions to the Cauchy problem, in Sections 2.8 and 2.9. We will first settle on a particular, and important class of equations in Section 2.8, for which we will prove, just as in the viscous case, the convergence to a wave. Alas, even though the speed \( c \) is unique, we will lose the uniqueness of the profile of the steady solutions – unlike in the diffusive case, (2.1.6) may have non-unique solutions, even up to a translation. We are going to investigate this phenomenon, that may be considered as the second main point of this chapter, in some detail. This major difference with the diffusive Hamilton-Jacobi equations will not be enough to prevent the large time convergence, but will force us to find a selection mechanism that will make up for the loss of diffusion.

In the last section, we use these new ideas to explain that, in fact, convergence to a wave holds for general equations of the form (2.1.1), as long as the Hamiltonian \( H \) is strictly convex in its second variable. In order to achieve this objective, we will (although we do not pretend to give a comprehensive treatment of this vast subject, that is still evolving at the time of the writing) give a reasonably comprehensive view of the issues posed by these deceptively simple models.

### 2.2 An informal derivation of the Hamilton-Jacobi equations

We begin by providing an informal derivation of the Hamilton-Jacobi equations, in the spirit of what we have done in Section 1.2 for the linear and semi-linear parabolic equations. The material of this section will reappear in Section 2.7 in the form of the Lax-Oleinik formula for the solutions to the Hamilton-Jacobi equations.

As in Section 1.2, we start with a random walk on a lattice of size \( h \) in \( \mathbb{R}^n \), and a time step \( \tau \). The walker evolves as follows. If the walker is located at a position \( X(t) \in h\mathbb{Z}^n \) at a time \( t = m\tau, m \in \mathbb{N} \), then at a time \( t + \tau \) it finds itself at a position

\[
X(t + \tau) = X(t) + v(t)\tau + h\xi(t).
\]

Here, \( \xi(t) \in \mathbb{R}^n \) is an \( \mathbb{R}^n \)-valued random variable such that each of the coordinates \( \xi_k(t) \), with \( k = 1, \ldots, n \), are independent and take the values \( \pm 1 \) with probabilities equal to \( 1/2 \), so that

\[
\mathbb{E}(\xi_k(t)) = 0, \quad \mathbb{E}(\xi_k(t)\xi_m(t')) = \delta_{km}\delta_{t,t'},
\]

for all \( 1 \leq k,m \leq n \) and all \( t,t' \). The velocity \( v(t) \) is known as a control, that the walker can choose from a set \( A \) of admissible velocities. The choice of the velocity \( v \) on the time interval \( [t, t + \tau] \) comes with a cost \( L(v)\tau \), where \( L(v) \) is a prescribed cost function. At the terminal time \( T = N\tau \) the walker finds itself at a position \( X(T) \) and pays the terminal cost \( f(X(T)) \), where \( f(x) \) is also a given function. The total cost of the trajectory that starts
at a time $t = m\tau$ at a position $x$ and continues until the time $T = N\tau$ is

$$w(t, x; V) = \sum_{k=m}^{N} L(v(k\tau))\tau + f(X(N\tau)). \quad (2.2.3)$$

Note that the total cost involves both the running cost and the terminal cost. We have denoted here by $V = (v(t), v(t + \tau), \ldots, v((N - 1)\tau))$ the whole sequence of the controls (velocities) chosen by the walker between the times $t$ and $T$. The quantity of interest is the least possible average cost, optimized over all choices of the velocities:

$$u(t, x) = \inf_{V \in \mathcal{A}_{t,T}} \mathbb{E} w(t, x; V) = \inf_{V \in \mathcal{A}_{t,T}} \mathbb{E} \sum_{k=m}^{N} L(v(k\tau))\tau + f(X(N\tau)). \quad (2.2.4)$$

Here, the expectation $\mathbb{E}$ is taken with respect to the random variables $\xi(s)$, for all $s = k\tau$ with $m \leq k < N$ that describe the random contribution at each of the time steps between $t$ and $T$. The set $\mathcal{A}_{t,T}$ is the set of all possible controls chosen between the times $t = n\tau$ and $T = N\tau$. The velocities $v$ are viewed as not random, as they can be chosen by the walker. The function $u(t, x)$ is known as the value function and is the basic object of study in the control theory.

As the velocities $v(s)$ are chosen separately by the walker at each time $s$ between $t$ and $T$, and the random variables $\xi(s)$ and $\xi(s')$ are independent for $s \neq s'$, the function $u(t, x)$ satisfies the following relation:

$$u(t, x) = \inf_{V \in \mathcal{A}} \mathbb{E} [L(v)\tau + u(t + \tau, x + v\tau + h\xi(t))]. \quad (2.2.5)$$

This is the simplest version of a dynamic programming principle, a fundamental notion of the control theory. Here, $v$ is the velocity chosen at the initial time $t$ and the expectation is taken solely with respect to the random variable $\xi(t)$.

A version of the dynamic programming principle, such as (2.2.5), is a very common starting point for the derivation of the Hamilton-Jacobi and other related types of equations. To illustrate this idea, let us assume that $u(t, x)$ is a sufficiently smooth function. Expanding the right side of (2.2.5) in $h \ll 1$ and $\tau \ll 1$ gives

$$u(t, x) = \inf_{v \in \mathcal{A}} \mathbb{E} [L(v)\tau + u(t + \tau, x + v\tau + h\xi(t))] = u(t, x) + \tau u_t + \frac{\tau^2}{2} u_{tt}(t, x)$$

$$+ \inf_{v \in \mathcal{A}} \mathbb{E} \left[ L(v)\tau + (v\tau + h\xi(t)) \cdot \nabla u(t, x) + \tau (v\tau + h\xi(t)) \cdot \nabla u_t(t, x) \right]$$

$$+ \frac{1}{2} \sum_{i,j=1}^{n} (v_i\tau + h\xi_i(t))(v_j\tau + h\xi_j(t)) \frac{\partial^2 u(t, x)}{\partial x_i \partial x_j} + l.o.t. \quad (2.2.6)$$

Note that the terms of the order $O(1)$ in the left and the right sides of (2.2.6) cancel automatically. In addition, the terms that are linear in $\xi(t)$ vanish after taking the expectation. It is easy to see then that, as in the random walk approximation of the diffusion equations we have encountered in Section 1.2, the interesting choice of the temporal and spatial steps $\tau$ and $h$ is

$$h^2 = 2D\tau, \quad (2.2.7)$$
with a diffusion coefficient $D$. Then, after taking into account the aforementioned cancellations, the leading order terms in (2.2.6) are of the order $O(\tau) = O(h^2)$. Keeping in mind (2.2.2), we see that they combine to give the following equation for $u(t, x)$:

$$u_t(t, x) + \inf_{v \in A} \left[ L(v) + v \cdot \nabla u(t, x) \right] + D\Delta u(t, x) = 0.$$  \hfill (2.2.8)

Let us introduce the function

$$H(p) = \inf_{v \in A} \left[ L(v) + v \cdot p \right],$$  \hfill (2.2.9)

defined for $p \in \mathbb{R}^n$. Then (2.2.8) can be written as

$$u_t + H(\nabla u) + D\Delta u = 0. \hfill (2.2.10)$$

This equation should be supplemented by the terminal condition $u(T, x) = f(x)$ that comes from the definition of the value function. Recall that $f(x)$ is the terminal cost function.

Equation (2.2.10) is backward in time. It is convenient to define the function

$$\bar{u}(t, x) = u(T - t, x),$$

which satisfies the forward in time Cauchy problem:

$$\bar{u}_t = H(\nabla \bar{u}) + D\Delta \bar{u}, \quad t > 0,$$

$$\bar{u}(0, x) = f(x), \hfill (2.2.11)$$

and for the sake of convenience we will focus on this forward in time Cauchy problem.

This is how the viscous Hamilton-Jacobi equations can be derived informally. Their rigorous derivation starting with a continuous in space and time stochastic control problem is not very different but requires the use of the stochastic calculus and the Ito formula. The inviscid equations of the form

$$u_t = H(\nabla u), \hfill (2.2.12)$$

are derived in a very similar way but the random walk is taken to be purely deterministic, driven solely by the control $v$, with $\xi(t) = 0$.

**Exercise 2.2.1** Generalize the above derivation to obtain a spatially inhomogeneous Hamilton-Jacobi equation of the form

$$u_t = H(x, \nabla u) + D\Delta u.$$  \hfill (2.2.13)

**Exercise 2.2.2** Show that the function $H(p)$ defined in (2.2.9) is concave.

This exercise explains why we will often consider below the Hamilton-Jacobi equations of the form

$$u_t + H(x, \nabla u) = D\Delta u,$$  \hfill (2.2.14)

with a convex Hamiltonian $H(p)$, either with $D > 0$ or $D = 0$. 

97
2.3 The simple world of viscous Hamilton-Jacobi equations

As a warm-up to the chapter, we are going to use the knowledge gathered in Chapter 1 for the study of the long time behavior of the solutions to the viscous Hamilton-Jacobi equations. This problem falls in the same class as what we did in Section 1.8.1, where we proved, essentially with the sole aid of the strong maximum principle and the Harnack inequality, the convergence of the solutions to the Cauchy problem for the Allen-Cahn equation to a translate of a stationary solution. The main difference is that now we will have to fight a little to show the existence of a steady state, while the long time convergence will be relatively effortless.

We are interested in the large time behavior of the solutions to the Cauchy problem

\[ u_t - \Delta u = H(x, \nabla u), \quad t > 0, \quad x \in \mathbb{T}^n, \quad (2.3.1) \]

with a given initial condition \( u(0, x) = u_0(x) \). This is an equation of the form (1.5.35) that we have considered in Section 1.5.2, and we make the same assumptions on the nonlinearity, that we now denote by \( H \), the standard notation in the theory of the Hamilton-Jacobi equations, as in that section. First, we assume that \( H \) is smooth and 1-periodic in \( x \). We also make the Lipschitz assumption on the function \( H(x, p) \): there exists \( C_L > 0 \) so that

\[ |H(x, p_1) - H(x, p_2)| \leq C_L|p_1 - p_2|, \quad \text{for all } x, p_1, p_2 \in \mathbb{R}^n. \quad (2.3.2) \]

In addition, we assume that \( H \) is growing linearly in \( p \) at infinity: there exist \( \alpha > 0 \) and \( \beta > 0 \) so that

\[ 0 < \alpha \leq \liminf_{|p| \to +\infty} \frac{H(x, p)}{|p|} \leq \limsup_{|p| \to +\infty} \frac{H(x, p)}{|p|} \leq \beta < +\infty, \quad \text{uniformly in } x \in \mathbb{T}^n. \quad (2.3.3) \]

One consequence of (2.3.3) is that \( H(x, p) \) is uniformly bounded from below. Note also that if \( u(t, x) \) solves (2.3.1) then \( u(t, x) + Kt \) solves (2.3.1) with the Hamiltonian \( H(x, p) \) replaced by \( H(x, p) + K \). Therefore, we may assume without loss of generality that there exist \( C_{1,2} > 0 \) so that

\[ C_1(1 + |p|) \leq H(x, p) \leq C_2(1 + |p|), \quad \text{for all } x \in \mathbb{T}^n \text{ and } p \in \mathbb{R}^n, \quad (2.3.4) \]

so that, in particular,

\[ H(x, p) > C_1 \quad \text{for all } x \in \mathbb{T}^n \text{ and } p \in \mathbb{R}^n. \quad (2.3.5) \]

As we have seen in Section 1.5.2, these assumptions ensure the existence of a unique smooth 1-periodic solution \( u(t, x) \) to (2.3.1) supplemented by a continuous, 1-periodic initial condition \( u_0(x) \). In order to discuss its long time behavior, we need to introduce a special class of solutions of (2.3.1).

**Theorem 2.3.1** Under the above assumptions, there exists a unique \( c \in \mathbb{R} \) so that (2.3.1) has solutions (that we will call the wave solutions) of the form

\[ w(t, x) = ct + \phi(x), \quad (2.3.6) \]

with a 1-periodic function \( \phi(x) \). The profile \( \phi(x) \) is unique up to an additive constant: if \( w_1(t, x) \) and \( w_2(t, x) \) are two such solutions then there exists \( k \in \mathbb{R} \) so that \( \phi_1(x) - \phi_2(x) \equiv k \) for all \( x \in \mathbb{T}^n \).
The constant $c$ is often referred to as the speed of the plane wave. The reason is that the solutions to the Hamilton-Jacobi equations, apart from the optimal control theory context that we have discussed above, also often describe the height of an interface, so that $c$ may be thought of as the speed at which the height of the interface is moving up, and $\phi(x)$ as the fixed profile of that interface as it moves up at a constant speed.

**Exercise 2.3.2** Consider the following discrete growing interface model, defined on the lattice $h\mathbb{Z}$, with a time step $\tau$, with $u(t,x)$ being the interface height at the time $t$ at the position $x$:

$$u(t+\tau,x) = \frac{1}{2}[u(t,x-h)+u(t,x+h)] + \frac{1}{2}[F(u(t,x+h)-u(t,x))+F(u(t,x)-u(t,x-h))] + V(t,x),$$

with a given function $F(p)$, and a prescribed source $V(t,x)$. The terms in the right side of (2.3.7) can be interpreted as follows: (1) the first term has an equilibrating effect, leveling the interface out, (2) the second term says that the rate of the interface growth depends on its slope – things falling from above can stick to the interface, and (3) the last term is a outside source of the interface growth (things falling from above). Find a scaling limit that relates $\tau$, $h$ and $F$ so that in the limit you get a Hamilton-Jacobi equation of the form

$$u_t = \Delta u + H(x, \nabla u) + V(t,x).$$

(2.3.8)

The large time behavior of $u(t,x)$ is summarized in the next theorem.

**Theorem 2.3.3** Let $u(t,x)$ be the solution of the Cauchy problem for (2.3.1) with a continuous 1-periodic initial condition $u_0$. There is a wave solution $w(t,x)$ of the form (2.3.6), a constant $\omega > 0$ that does not depend on $u_0$ and $C_0 > 0$ that depends on $u_0$ such that

$$|u(t,x) - w(t,x)| \leq C_0 e^{-\omega t},$$

(2.3.9)

for all $t \geq 0$ and $x \in \mathbb{T}^n$.

We will first prove the existence part of Theorem 2.3.1, and that will occupy most of the rest of this section, while its uniqueness part and the convergence claim of Theorem 2.3.3 will be proved together rather quickly in the end.

### 2.3.1 The wave solutions

**Outline of the existence proof**

We first present an outline of the existence proof, before going into the details of the argument. Plugging the ansatz (2.3.6) into (2.3.1) and integrating over $\mathbb{T}^n$ gives a relation

$$c = \int_{\mathbb{T}^n} H(x, \nabla \phi(x)) dx.$$ 

(2.3.10)

The equation for $\phi$ can, therefore, be written as

$$- \Delta \phi = H(x, \nabla \phi(x)) - \int_{\mathbb{T}^n} H(z, \nabla \phi(z)) dz,$$

(2.3.11)
and this will be the starting point of our analysis.

We are going to use a continuation method. As this strategy is typical for the existence proofs for many nonlinear PDEs, it is worth sketching out the general plan. Instead of just looking at (2.3.11) with a given Hamiltonian $H(x, p)$, we consider a family of equations

$$- \Delta \phi_{\sigma} = H_{\sigma}(x, \nabla \phi_{\sigma}) - \int_{\mathbb{T}^n} H_{\sigma}(z, \nabla \phi_{\sigma}) dz,$$

with the Hamiltonians

$$H_{\sigma}(x, p) = (1 - \sigma)H_0(x, p) + \sigma H(x, p),$$

parametrized by $\sigma \in [0, 1]$. At $\sigma = 0$, we start with a particular choice of $H_0(x, p)$ for which we know that (2.3.12) has a solution: we take

$$H_0(x, p) = \sqrt{1 + |p|^2},$$

so that $\phi_0(x) \equiv 0$ is an explicit solution to (2.3.12) with $\sigma = 0$. At $\sigma = 1$, we end with

$$H_1(x, p) = H(x, p).$$

The goal is to show that (2.3.12) has a solution for all $\sigma \in [0, 1]$ and not just for $\sigma = 0$ by showing that the set $\Sigma$ of $\sigma$ for which (2.3.12) has a solution is both open and closed in $[0, 1]$.

Showing that $\Sigma$ is closed requires a priori bounds on the solution $\phi_{\sigma}$ of (2.3.12) that would both be uniform in $\sigma \in [0, 1]$ and ensure the compactness of the sequence $\phi_{\sigma_n}$ of solutions to (2.3.12) as $\sigma_n \to \sigma$ in a suitable function space. The estimates should be strong enough to ensure that the limit $\phi_{\sigma}$ is a solution to (2.3.12).

In order to show that $\Sigma$ is open, one relies on the implicit function theorem. Let us assume that (2.3.12) has a solution $\phi_{\sigma}(x)$ for some $\sigma \in [0, 1]$. In order to show that (2.3.12) has a solution for $\sigma + \varepsilon$, with a sufficiently small $\varepsilon$, we are led to consider the linearized problem

$$- \Delta h - \frac{\partial H_{\sigma}(x, \nabla \phi_{\sigma})}{\partial p_j} \frac{\partial h}{\partial x_j} + \int_{\mathbb{T}^n} \frac{\partial H_{\sigma}(z, \nabla \phi_{\sigma})}{\partial p_j} \frac{\partial h(z)}{\partial z_j} dz = f,$$

with

$$f(x) = H(x, \nabla \phi_{\sigma}) - H_0(x, \nabla \phi_{\sigma}) - \int_{\mathbb{T}^n} H(z, \nabla \phi_{\sigma}(z)) dz + \int_{\mathbb{T}^n} H_0(z, \nabla \phi_{\sigma}(z)) dz.$$

The implicit function theorem guarantees existence of the solution $\phi_{\sigma+\varepsilon}$, provided that the linearized operator in the left side of (2.3.15) is invertible, with the norm of the inverse a priori bounded in $\sigma$. This will show that the set $\Sigma$ of $\sigma \in [0, 1]$ for which the solution to (2.3.12) exists is open.

The bounds on the operator that maps $f \to h$ in (2.3.15) also require the a priori bounds on $\phi_{\sigma}$. Thus, both proving that $\Sigma$ is open and that it is closed require us to prove the a priori uniform bounds on $\phi_{\sigma}$. Therefore, our first step will be to assume that a solution $\phi_{\sigma}(x)$ to (2.3.12) exists and obtain a priori bounds on $\phi_{\sigma}$. Note that if $\phi_{\sigma}$ is a solution to (2.3.12), then $k + \phi_{\sigma}$ is also a solution for any $k \in \mathbb{R}$. Thus, it is more natural to obtain a priori bounds
on $\nabla \phi_{\sigma}$ than on $\phi_{\sigma}$ itself, and then normalize the solution so that $\phi_{\sigma}(0) = 0$ to ensure that $\phi_{\sigma}$ is bounded.

It is important to observe that the Hamiltonians $H_{\sigma}(x, p)$ obey the Lipschitz bound (2.3.2), with a Lipschitz constant $C_L$ that does not depend on $\sigma \in [0, 1]$, and estimate (2.3.4) also holds for $H_{\sigma}$ with the same $C_{1,2} > 0$ for all $\sigma \in [0, 1]$. The key bound to prove will be to show that there exists a constant $K > 0$ that depends only on the Lipschitz constant of $H$ in (2.3.2) and the two constants in the linear growth estimate (2.3.4) such that any solution to (2.3.12) satisfies

$$\| \nabla \phi_{\sigma} \|_{L^\infty(\mathbb{T}^n)} \leq K.$$  \hfill (2.3.17)

We stress that this bound will be obtained not just for one Hamiltonian but for all Hamiltonians with the same Lipschitz constant $C_L$ in (2.3.2) that satisfy (2.3.4) with the same $C_{1,2} > 0$. The estimate (2.3.17) will turn out to be sufficient to apply the argument we have outlined above.

An a priori $L^1$-bound on the gradient

Before establishing the $L^\infty$-bound (2.3.17), let us first prove that there exists a constant $C > 0$ that only depends on $C_L$ in (2.3.2) and $C_{1,2}$ in (2.3.4) such that any solution $\phi_{\sigma}(x)$ of (2.3.12) satisfies

$$\int_{\mathbb{T}^n} H_{\sigma}(x, \nabla \phi_{\sigma}) dx \leq C.$$  \hfill (2.3.18)

Because of the lower bound in (2.3.3), this is equivalent to an a priori $L^1$-bound on $|\nabla \phi_{\sigma}|$:

$$\int_{\mathbb{T}^n} |\nabla \phi_{\sigma}(x)| dx \leq C,$$ \hfill (2.3.19)

with a possibly different $C > 0$ that still depends only on $C_L$ and $C_{1,2}$. To prove (2.3.18), we will rely on Proposition 1.7.10 in Chapter 1 that we recall here for the convenience of the reader.

**Proposition 2.3.4** Let $b(x)$ be a smooth vector field over $\mathbb{T}^n$. The linear equation

$$-\Delta e + \nabla \cdot (eb) = 0, \quad x \in \mathbb{T}^n,$$  \hfill (2.3.20)

has a unique solution $e_1^*(x)$ normalized so that

$$\|e_1^*\|_{L^\infty} = 1,$$ \hfill (2.3.21)

and such that $e_1^* > 0$ on $\mathbb{T}^n$. Moreover, for all $\alpha \in (0, 1)$, the function $e_1^*$ is $\alpha$-Hölder continuous, with the $\alpha$-Hölder norm bounded by a universal constant depending only on $\|b\|_{L^\infty(\mathbb{T}^n)}$.

Let us first see why it implies (2.3.18). An immediate consequence of the normalization (2.3.21) and the claim about the Hölder norm of $e_1^*$, together with the positivity of $e_1^*$ is that

$$\int_{\mathbb{T}^n} e_1^*(x) dx \geq K_1 > 0,$$ \hfill (2.3.22)
with a constant $K_1 > 0$ that depends only on $\|b\|_{L^\infty}$. Now, given a solution $\phi_\sigma(x)$ of (2.3.12), set
\[
b_j(x) = \int_0^1 \partial_{p_j} H_\sigma(x, r \nabla \phi_\sigma(x)) \, dr,
\]
so that
\[
b(x) \cdot \nabla \phi_\sigma(x) = \sum_{j=1}^n b_j(x) \frac{\partial \phi_\sigma}{\partial x_j} = H_\sigma(x, \nabla \phi_\sigma) - H_\sigma(x, 0),
\]
and (2.3.12) can be re-stated as
\[
- \Delta \phi_\sigma - b_j(x) \frac{\partial \phi_\sigma}{\partial x_j} = H_\sigma(x, 0) - \int_{\mathbb{T}^n} H_\sigma(z, \nabla \phi_\sigma) \, dz.
\]
Note that while $b(x)$ does depend on $\nabla \phi_\sigma$, the $L^\infty$-norm of $b(x)$ depends only on the Lipschitz constant $C_L$ of the function $H_\sigma(x, p)$ in the $p$-variable. Let now $e_1^*$ be the solution to (2.3.20) given by Proposition 2.3.4, with the above $b(x)$. Multiplying (2.3.25) by $e_1^*$ and integrating over $\mathbb{T}^n$ yields
\[
0 = \int_{\mathbb{T}^n} e_1^*(x) H_\sigma(x, 0) \, dx - \left( \int_{\mathbb{T}^n} e_1^*(x) \, dx \right) \left( \int_{\mathbb{T}^n} H_\sigma(z, \nabla \phi_\sigma) \, dz \right),
\]
hence
\[
\int_{\mathbb{T}^n} H_\sigma(x, \nabla \phi_\sigma) \, dx \leq \left( \int_{\mathbb{T}^n} e_1^*(x) \, dx \right)^{-1} \int_{\mathbb{T}^n} e_1^*(x) H_\sigma(x, 0) \, dx,
\]
and (2.3.19) follows from (2.3.22) and (2.3.4). As the constant $K_1$ in (2.3.22) depends only on the $L^\infty$-norm of $b(x)$ that, in turn, depends only on $C_L$, the constant $C$ in the right side of (2.3.18), indeed, depends only on $C_L$ and $C_{1,2}$.

**An a priori $L^\infty$ bound on the gradient**

So far, we have obtained an a priori $L^1$-bound (2.3.19) for the gradient of any solution $\phi_\sigma$ to (2.3.12). Now, we improve this estimate to an $L^\infty$ bound.

**Proposition 2.3.5** There is a constant $C > 0$ that depends only on the constants $C_L$ and $C_{1,2}$, such that any solution $\phi_\sigma$ to
\[
- \Delta \phi_\sigma = H_\sigma(x, \nabla \phi_\sigma) - \int_{\mathbb{T}^n} H_\sigma(z, \nabla \phi_\sigma) \, dz,
\]
satisfies
\[
\|\nabla \phi_\sigma\|_{L^\infty(\mathbb{T}^n)} \leq C.
\]
As a consequence, if $\phi_\sigma$ is normalized such that $\phi_\sigma(0) = 0$, then we also have $\|\phi_\sigma\|_{L^\infty(\mathbb{T}^n)} \leq C$.

**Proof.** We borrow the strategy in the proof of Proposition 1.7.10. Let $\phi_\sigma$ be a solution to (2.3.28) such that $\phi_\sigma(0) = 0$. The only estimate we have so far is the $L^1$-bound (2.3.19) for $\nabla \phi_\sigma$ – the idea is to estimate the $L^\infty$-norm $\|\nabla \phi_\sigma\|_{L^\infty(\mathbb{T}^n)}$ solely from the $L^1$-norm of $\phi_\sigma$ and the equation.
Let \( \Gamma(x) \) be as in the proof of Proposition 1.7.10: a nonnegative smooth function equal to 1 in the cube \([-2, 2]^n\) and to zero outside of the cube \((-3, 3)^n\), and set \( \psi(x) = \Gamma(x)\phi_\sigma(x) \). The function \( \psi(x) \) satisfies an equation similar to what we have seen in (1.7.34):

\[
-\Delta \psi = -2\nabla \Gamma \cdot \nabla \phi_\sigma - \phi_\sigma \Delta \Gamma + F(x), \quad x \in \mathbb{R}^n,
\]

with

\[
F(x) = \Gamma(x) \left[ H_\sigma(x, \nabla \phi_\sigma(x)) - \int_{\mathbb{T}^n} H_\sigma(z, \nabla \phi_\sigma(z)) dz \right].
\]

The only a priori information we have about \( F(x) \) and the term \( \nabla \Gamma \cdot \nabla \phi_\sigma(x) \) so far is that they are supported inside \([-3, 3]^n\) and are uniformly bounded in \( L^1(\mathbb{R}^n) \) via (2.3.18) and (2.3.19).

Here, we use the assumption (2.3.4) that the Hamiltonian \( H(x, p) \) is uniformly positive. It helps to combine these two terms:

\[
G(x) = F(x) - 2\nabla \Gamma(x) \cdot \nabla \phi_\sigma(x),
\]

with \( G(x) \) supported inside \([-3, 3]^n\), and

\[
\int_{\mathbb{R}^n} |G(x)| dx \leq C,
\]

with a constant \( C > 0 \) that depends only on \( C_L \) and \( C_{1,2} \), due to (2.3.18) and (2.3.19). We also know that

\[
|G(x)| \leq C(1 + |\nabla \phi_\sigma(x)|),
\]

because of (2.3.4).

Next, we use the fundamental solution \( E(x) \) to the Laplace equation in \( \mathbb{R}^n \) to write

\[
\psi(x) = \int_{\mathbb{R}^n} E(x - y)[G(y) - \phi_\sigma(y) \Delta \Gamma(y)] dy.
\]

Differentiating (2.3.35) in \( x \) gives

\[
\nabla \psi(x) = \int_{\mathbb{R}^n} \nabla E(x - y)[G(y) - \phi_\sigma(y) \Delta \Gamma(y)] dy.
\]

Exercise 2.3.6 Note that the function \( E(x - y) \) has a singularity at \( y = x \). Show that nevertheless one can differentiate in (2.3.35) under the integral sign to obtain (2.3.36).

The function \( \nabla E(x - y) \) has an integrable singularity at \( y = x \), of the order \( |x - y|^{-n+1} \), and is bounded everywhere else. Thus, for all \( \varepsilon > 0 \) we have, with the help of (2.3.33) and (2.3.34):

\[
\left| \int_{\mathbb{R}^n} G(y) \nabla E(x - y) dy \right| \leq \int_{|x - y| \leq \varepsilon} G(y) \nabla E(x - y) dy + \int_{|x - y| \geq \varepsilon} G(y) \nabla E(x - y) dy \leq C\varepsilon^{1-n} + C\varepsilon^{-n+1} \int_{|x - y| \geq \varepsilon} |G(y)| dy \leq C\varepsilon(1 + \|\nabla \phi_\sigma\|_{L^\infty}) + C\varepsilon^{1-n}.
\]

\[\text{(2.3.37)}\]
The integral in (2.3.36) also contains a factor of $\phi_\sigma$, whereas our bounds so far deal with $\nabla \phi_\sigma$. However, we have assumed without loss of generality that $\phi_\sigma(0) = 0$, hence for any $\delta > 0$ we may write

$$
\phi_\sigma(y) = \int_0^1 y \cdot \nabla \phi_\sigma(sy) ds = \int_0^\delta y \cdot \nabla \phi_\sigma(sy) ds + \int_\delta^1 y \cdot \nabla \phi_\sigma(sy) ds,
$$

so that both, as $|y| \leq 1$, we have

$$
|\phi_\sigma(y)| \leq \|\nabla \phi_\sigma\|_{L^\infty}, \quad (2.3.38)
$$

and

$$
\int_{\mathbb{T}^n} |\phi_\sigma(y)| dy \leq C\delta \|\nabla \phi_\sigma\|_{L^\infty} + \int_\delta^1 \int_{\mathbb{T}^n} |y| \|\nabla \phi_\sigma(sy)\| dy ds \leq C\delta \|\nabla \phi_\sigma\|_{L^\infty} + C \int_\delta^1 \int_{\mathbb{T}^n} \|\nabla \phi_\sigma(y)\| dy ds \leq C\delta \|\nabla \phi_\sigma\|_{L^\infty} + C\delta^{-n}. \quad (2.3.39)
$$

We used above the a priori bound (2.3.19) on $\|\nabla \phi\|_{L^1(\mathbb{T}^n)}$. Combining (2.3.38) and (2.3.39), we obtain, as in (2.3.37):

$$
\left| \int_{\mathbb{R}^n} \phi_\sigma(y) \Delta \Gamma(y) \nabla E(x-y) dy \right| \leq \int_{|x-y| \leq \varepsilon} |\phi_\sigma(y)||\Delta \Gamma(y)||\nabla E(x-y)||dy + \int_{|x-y| > \varepsilon} |\phi_\sigma(y)||\Delta \Gamma(y)||\nabla E(x-y)||dy \leq C\varepsilon \|\phi_\sigma\|_{L^\infty} + C\varepsilon^{1-n} \int_{\mathbb{T}^n} |\phi_\sigma(y)| dy \leq C\varepsilon \|\nabla \phi_\sigma\|_{L^\infty} + C\varepsilon^{1-n} \delta \|\nabla \phi_\sigma\|_{L^\infty} + C\varepsilon^{1-n} \delta^{-n}. \quad (2.3.40)
$$

Together, (2.3.37) and (2.40) tell us that

$$
\|\nabla \psi\|_{L^\infty} \leq C\varepsilon (1 + \|\nabla \phi_\sigma\|_{L^\infty}) + C\varepsilon^{1-n} + C\|\nabla \phi_\sigma\|_{L^\infty} + C\varepsilon^{1-n} \delta \|\nabla \phi_\sigma\|_{L^\infty} + C\varepsilon^{1-n} \delta^{-n}. \quad (2.3.41)
$$

Next, observe that, because $\alpha \equiv \Gamma \equiv 1$ in $[-2,2]^n$ and $\phi_\sigma$ is 1-periodic, we have

$$
\|\nabla \phi_\sigma\|_{L^\infty(\mathbb{T}^n)} = \|\nabla (\Gamma \phi_\sigma)\|_{L^\infty([-1,1]^n)} \leq \|\nabla (\Gamma \phi_\sigma)\|_{L^\infty([-3,3]^n)} = \|\nabla \psi\|_{L^\infty}. \quad (2.3.42)
$$

Thus, if we take $\delta = \varepsilon^n$ in (2.3.41), we would obtain

$$
\|\nabla \phi_\sigma\|_{L^\infty} \leq C\varepsilon \|\nabla \phi_\sigma\|_{L^\infty} + C\varepsilon, \quad (2.3.43)
$$

with a universal constant $C > 0$ and $C\varepsilon$ that does depend on $\varepsilon$. Now, the proof of (2.3.29) is concluded by taking $\varepsilon > 0$ small enough. □

Going back to equation (2.3.11) for $\phi$:

$$
- \Delta \phi = H(x,\nabla \phi) - \int_{\mathbb{T}^n} H(x,\nabla \phi) dx, \quad (2.3.44)
$$

the reader should do the following exercise.

**Exercise 2.3.7** Use the $L^\infty$-bound on $\nabla \phi$ in Proposition 2.3.5 to deduce from (2.3.44) that, under the assumption that $H(x,p)$ is smooth (infinitely differentiable) in both variables $x$ and $p$, the function $\phi(x)$ is, actually, infinitely differentiable, with all its derivatives of order $n$ bounded by a priori constants $C_n$ that do not depend on $\phi$. 

104
The linearized problem

We need one more ingredient to finish the proof of the existence part of Theorem 2.3.1: to set-up an application of the implicit function theorem. Let \( \phi \) be a solution to (2.3.12) and let us consider the linearized problem, with an unknown \( h \):

\[
- \Delta h - \partial_{p_j} H_\sigma(x, \nabla \phi_\sigma) \partial_{x_j} h + \int_{\mathbb{T}^n} \partial_{p_j} H_\sigma(y, \nabla \phi_\sigma) \partial_{x_j} h(y) dy = f \quad x \in \mathbb{T}^n. \tag{2.3.45}
\]

We assume that \( f \in C^{1,\alpha}(\mathbb{T}^n) \) for some \( \alpha \in (0, 1) \), and \( f \) has zero mean over \( \mathbb{T}^n \):

\[
\int_{\mathbb{T}^n} f(x) dx = 0,
\]

and require that the solution \( h \) to (2.3.45) also has zero mean:

\[
\int_{\mathbb{T}^n} h(x) dx = 0. \tag{2.3.46}
\]

Proposition 2.3.8 Equation (2.3.45) has a unique solution \( h \in C^{3,\alpha}(\mathbb{T}^n) \) with zero mean. The mapping \( f \mapsto h \) is continuous from the set of \( C^{1,\alpha} \) functions with zero mean to the set of \( C^{3,\alpha}(\mathbb{T}^n) \) functions with zero mean.

Proof. The Laplacian is a one-to-one map between the set of \( C^{m+2,\alpha} \) functions with zero mean and the set of \( C^{m,\alpha}(\mathbb{T}^n) \) functions with zero mean, for any \( m \in \mathbb{N} \). Thus, we may talk about its inverse that we denote by \((-\Delta)^{-1}\). Equation (2.3.45) is thus equivalent to

\[
(I + K)h = (-\Delta)^{-1} f, \tag{2.3.47}
\]

with the operator

\[
Kh = (-\Delta)^{-1} \left( -\partial_{p_j} H_\sigma(x, \nabla \phi_\sigma) \partial_{x_j} h + \int_{\mathbb{T}^n} \partial_{p_j} H(y, \nabla \phi_\sigma) \partial_{x_j} h(y) dy \right). \tag{2.3.48}
\]

Exercise 2.3.9 Show that \( K \) is a compact operator on the set of functions in \( C^{1,\alpha}(\mathbb{T}^n) \) with zero mean.

The problem has been now reduced to showing that the only solution of

\[
(I + K)h = 0 \tag{2.3.49}
\]

with \( h \in C^{1,\alpha}(\mathbb{T}^n) \) with zero mean is \( h \equiv 0 \). Note that (2.3.49) simply says that \( h \) is a solution of (2.3.45) with \( f \equiv 0 \). Let \( e_1^* > 0 \) be given by Proposition 2.3.4, with

\[
b_j(x) = -\partial_{p_j} H_\sigma(x, \nabla \phi_\sigma). \tag{2.3.50}
\]

That is, \( e_1^* \) is the positive solution of the equation

\[
- \Delta e_1^* + \nabla \cdot (e_1^* b) = 0, \tag{2.3.51}
\]
normalized so that \( \| e_1^* \|_{L^\infty(\mathbb{T}^n)} = 1 \). The uniform Lipschitz bound on \( H_\sigma(x, p) \) in the \( p \)-variable implies that \( b(x) \) is in \( L^\infty(\mathbb{T}^n) \), and thus Proposition 2.3.4 can be applied. Multiplying (2.3.45) with \( f = 0 \) by \( e_1^* \) and integrating gives, as \( e_1^* > 0 \):

\[
\int_{\mathbb{T}^n} \partial_p H_\sigma(y, \nabla \phi_\sigma) \partial_{x_j} h(y) dy = 0.
\]

But then, the equation for \( h \) becomes simply

\[-\Delta h + b_j(x) \partial_{x_j} h = 0, \quad x \in \mathbb{T}^n,\]

which entails that \( h \) is constant, by the Krein-Rutman theorem. Because \( h \) has zero mean, we get \( h \equiv 0 \). □

**Exercise 2.3.10** Let \( H_0(x, p) \) satisfy the assumptions of Theorem 2.3.3, and assume that equation (2.3.11), with \( H = H_0 \):

\[-\Delta \phi_0 = H_0(x, \nabla \phi_0) - \int_{\mathbb{T}^n} H_0(z, \nabla \phi_0) dz, \quad (2.3.52)\]

has a solution \( \phi_0 \in C(\mathbb{T}^n) \). Consider \( H_1(x, p) \in C^\infty(\mathbb{T} \times \mathbb{R}^n) \). Prove, with the aid of Propositions 2.3.5 and 2.3.8, and the implicit function theorem that there exist \( R_0 > 0 \) and \( \varepsilon_0 > 0 \) such that if

\[|H_1(x, p)| \leq \varepsilon_0, \quad \text{for } x \in \mathbb{T}^n \text{ and } |p| \leq R_0, \quad (2.3.53)\]

then equation (2.3.11) with \( H = H_0 + H_1 \):

\[-\Delta \phi = H(x, \nabla \phi) - \int_{\mathbb{T}^n} H(z, \nabla \phi) dz, \quad (2.3.54)\]

has a solution \( \phi \).

**Existence of the solution**

We finally prove the existence part of Theorem 2.3.1. Consider \( H(x, p) \) satisfying the assumptions of the theorem. As before, we set

\[H_0(x, p) = \sqrt{1 + |p|^2} - 1, \]

and

\[H_\sigma(x, p) = (1 - \sigma)H_0(x, p) + \sigma H(x, p), \]

so that \( H_1(x, p) = H(x, p) \), and consider existence of a solution to (2.3.12):

\[-\Delta \phi_\sigma = H_\sigma(x, \nabla \phi_\sigma) - \int_{\mathbb{T}^n} H_\sigma(z, \nabla \phi_\sigma) dz, \quad (2.3.55)\]

Consider the set

\[\Sigma = \{ \sigma \in [0, 1] : \text{ equation (2.3.55) has a solution} \}.\]
Our goal is to show that $\Sigma = [0, 1]$. We know that $\Sigma$ is non empty, because $0 \in \Sigma$: indeed, $\phi_0(x) \equiv 0$ is a solution to (2.3.55) at $\sigma = 0$. Thus, if we show that $\Sigma$ is both open and closed in $[0, 1]$, this will imply that $\Sigma = [0, 1]$, and, in particular, establish the existence of a solution to (2.3.55) for $H_1(x, p) = H(x, p)$.

Now that we know that the linearized problem is invertible, the openness of $\Sigma$ is a direct consequence of the inverse function theorem, as explained in Exercise 2.3.10. Closedness of $\Sigma$ is not too difficult to see either: consider a sequence $\sigma_n \in [0, 1]$ converging to $\bar{\sigma} \in [0, 1]$, and let $\phi_n$ be a solution to (2.3.55) with $H(x, p) = H_{\sigma_n}(x, p)$, normalized so that $\phi_n(0) = 0$. (2.3.56)

Proposition 2.3.5 implies that $\|\nabla \phi_n\|_{L^\infty(T^n)} \leq C$, and thus $\|H(x, \nabla \phi_n)\|_{L^\infty} \leq C$. However, this means that $\phi_n$ solve an equation of the form

$$-\Delta \phi_n = F_n(x), \quad x \in T^n,$$

(2.3.57)

with a uniformly bounded function

$$F_n(x) = H_{\sigma_n}(x, \nabla \phi_n) - \int_{T^n} H_{\sigma_n}(z, \nabla \phi_n(z))dz.$$  

(2.3.58)

It follows that that $\phi_n$ is bounded in $C^{1,\alpha}(T^n)$, for all $\alpha \in [0, 1]$:

$$\|\phi_n\|_{C^{1,\alpha}(T^n)} \leq C.$$  

(2.3.59)

But this implies, in turn, that the functions $F_n(x)$ in (2.3.58) are also uniformly bounded in $C^\alpha(T^n)$, hence $\phi_n$ are uniformly bounded in $C^{2,\alpha}(T^n)$:

$$\|\phi_n\|_{C^{2,\alpha}(T^n)} \leq C.$$  

(2.3.60)

Now, the Arzela-Ascoli theorem implies that a subsequence $\phi_{n_k}$ will converge in $C^2(T^n)$ to a function $\tilde{\phi}$, which is a solution to (2.3.19) with $H = H_{\bar{\sigma}}$. Thus, $\bar{\sigma} \in \Sigma$, and $\Sigma$ is closed. This finishes the proof of the existence part of the theorem.

### 2.3.2 Long time convergence and uniqueness of the wave solutions

We will now prove simultaneously the claim of the uniqueness of the speed $c$ and of the profile $\phi(x)$ in Theorem 2.3.1, and the long time convergence for the solutions to the Cauchy problem stated in Theorem 2.3.3.

Let $u(t, x)$ be the solution to (2.3.1)

$$u_t = \Delta u + H(x, \nabla u), \quad t > 0, \quad x \in T^n,$$

(2.3.61)

with $u(0, x) = u_0(x) \in C(T^n)$. We also take a speed $c \in \mathbb{R}$ and a solution $\phi(x)$ to

$$\Delta \phi + H(x, \nabla \phi) = c,$$

(2.3.62)
without assuming that either $c$ or $\phi$ is unique.

We wish to prove that there exists $\bar{k} \in \mathbb{R}$ so that $u(t, x) - ct$ is attracted exponentially fast in time to $\phi(x) + \bar{k}$:

$$|u(t, x) - ct - \bar{k} - \phi(x)| \leq C_0 e^{-\omega t},$$

with some $C_0 > 0$ and $\omega > 0$, such that $C_0$ depends on the initial condition $u_0$ but $\omega$ does not.

The idea is the same as in the proof of Theorem 1.8.2 for the Allen-Cahn equation: squeeze the solution between two different wave solutions, and show that the difference between the squeezers tends to zero as $t \to +\infty$. However, the situation here is much simpler: we do not have any tail as $|x| \to +\infty$ to control, because we are now considering the problem for $x \in \mathbb{T}^n$.

Actually, the present setting realizes what would be the dream scenario for the Allen-Cahn equation.

As a simple remark, we may assume that $c = 0$, just by setting

$$\tilde{H}(x, p) = H(x, p) - c,$$

and dropping the tilde, and this is what we will do. In other words, $\phi(x)$ is the solution to

$$\Delta \phi + H(x, \nabla \phi) = 0. \tag{2.3.64}$$

Let $\phi$ be any solution to (2.3.64), and set

$$k_0^- = \sup \{k : \ u(0, x) \geq \phi(x) + k \text{ for all } x \in \mathbb{T}^n\},$$

and

$$k_0^+ = \inf \{k : \ u(0, x) \leq \phi(x) + k \text{ for all } x \in \mathbb{T}^n\}.$$

Because $\phi(x) + k_0^\pm$ solve (2.3.64) with $c = 0$, and $u(t, x)$ solves (2.3.61), we have, by the maximum principle:

$$\phi(x) + k_0^- \leq u(t, x) \leq \phi(x) + k_0^+, \text{ for all } t \geq 0 \text{ and } x \in \mathbb{T}^n. \tag{2.3.65}$$

Now, for all $q \in \mathbb{N}$, let us set

$$k_q^- = \sup \{k : \ u(t = q, x) \geq \phi(x) + k \text{ for all } x \in \mathbb{T}^n\} = \inf_{x \in \mathbb{T}^n} [u(t = q, x) - \phi(x)], \tag{2.3.66}$$

and

$$k_q^+ = \inf \{k : \ u(t = q, x) \leq \phi(x) + k \text{ for all } x \in \mathbb{T}^n\} = \sup_{x \in \mathbb{T}^n} [u(t = q, x) - \phi(x)]. \tag{2.3.67}$$

The strong maximum principle implies that the sequence $k_q^-$ is increasing, whereas $k_q^+$ is decreasing, and that, as in (2.3.65), we have

$$\phi(x) + k_q^- \leq u(t, x) \leq \phi(x) + k_q^+, \text{ for all } t \geq q \text{ and } x \in \mathbb{T}^n. \tag{2.3.68}$$

Hence, the theorem will be proved if we manage to show that

$$0 \leq k_q^+ - k_q^- \leq C a^q, \text{ for all } q \geq 0, \tag{2.3.69}$$

108
with some $C \in \mathbb{R}$ that may depend on the initial condition $u_0$ and $a \in (0, 1)$ that does not depend on $u_0$. In order to prove (2.3.69), it suffices to show that
\begin{equation}
    k^+_{q+1} - k^-_{q+1} \leq (1 - r_0)(k^+_{q} - k^-_{q}),
\end{equation}
with some $r_0 \in (0, 1)$. This is a quantification of the strong maximum principle: by the time $t = q + 1$ $u(x)$ has to detach “by a fixed amount” from the respective lower and upper bounds $\phi(x) + k^\pm_q$ that hold at $t = q$. Such estimates typically rely on the Harnack inequality, and this is what we will use.

To bring the Harnack inequality in, note that the function
\begin{equation}
    w(t, x) = u(t, x) - \phi(x) - k^-_q
\end{equation}
is nonnegative for $t \geq q$, and solves an equation of the form
\begin{equation}
    \partial_t w - \Delta w + b_j(t, x)\partial_{x_j} w = 0, \quad t > q, \quad x \in \mathbb{T}^n,
\end{equation}
with a bounded drift $b(t, x)$ given by
\begin{equation}
    b(t, x) = \int_0^1 \nabla_p H(x, (1 - s)\nabla \phi(x) + s \nabla u(t, x)) ds,
\end{equation}
so that
\begin{equation}
    b(t, x) \cdot [\nabla u(t, x) - \nabla \phi(x)] = H(x, \nabla u(t, x)) - H(x, \nabla \phi(x)),
\end{equation}
and
\begin{equation}
    |b_j(t, x)| \leq C_L, \quad \text{for all } t \geq q \text{ and } x \in \mathbb{T}^n.
\end{equation}
The Harnack inequality in Theorem ?? and (2.3.73) imply that there exists $r_0 > 0$ that depends only on $C_L$ such that
\begin{equation}
    \inf_{x \in \mathbb{T}^n} w(q + 1, x) \geq r_0 \sup_{x \in \mathbb{T}^n} w(q, x).
\end{equation}
Using (2.3.66) and (2.3.67), together with (2.3.74), we may write
\begin{equation}
    r_0 \sup_{x \in \mathbb{T}^n} w(q, x) = r_0 \sup_{x \in \mathbb{T}^n} [u(q, x) - \phi(x) - k^-_q] = r_0 [k^+_q - k^-_q] \leq \inf_{x \in \mathbb{T}^n} w(q + 1, x)
\end{equation}
\begin{equation}
    = \inf_{x \in \mathbb{T}^n} [u(q + 1, x) - \phi(x) - k^-_q] = k^-_{q+1} - k^-_q,
\end{equation}
so that
\begin{equation}
    k^-_{q+1} \geq k^-_q + r_0 [k^+_q - k^-_q].
\end{equation}
As $k^+_q \leq k^+_q$, it follows that
\begin{equation}
    k^+_{q+1} - k^-_{q+1} \leq k^+_q - k^-_q - r_0(k^+_q - k^-_q) \leq (1 - r_0)(k^+_q - k^-_q),
\end{equation}
which is (2.3.70). This implies the geometric decay as in (2.3.69), hence the theorem, because of (2.3.68) and (2.3.69). Note that the constant
\begin{equation}
    a = 1 - r_0
\end{equation}
comes from the Harnack inequality and does not depend on the initial condition $u_0$ but only on the Lipschitz constant $C_L$ of $H(x, p)$. □
Exercise 2.3.11 (i) Why does the uniqueness of \( c \) and of the profile \( \phi(x) \) follow?
(ii) How is the constant \( \omega \) in Theorem 2.3.3 related to the constant \( a \) in the above proof?

Exercise 2.3.12 Consider a modified equation, not quite of the Hamilton-Jacobi form:

\[
    u_t - \Delta u = R(x, u)\sqrt{1 + |\nabla u|^2},
\]

where \( R(x, u) \) is a smooth, positive function, that is 1-periodic in \( x \) and 1-periodic in \( u \).

(i) Let \( u_0 \in C(T^N) \), and show that the Cauchy problem for (2.3.78) with \( u(0, x) = u_0(x) \) is well posed.

(ii) Prove the existence of a unique \( T > 0 \) such that equation (2.3.78) has solutions of the form

\[
    u(t, x) = \frac{t}{T} + \phi(t, x),
\]

where \( \phi \) is \( T \)-periodic in \( t \) and 1-periodic in \( x \). We will call such a solution a wave solution. Why is it not reasonable to expect that under the above assumptions (2.3.78) has a wave solution of the form \( u(t, x) = ct + \psi(x) \) with a 1-periodic function \( \psi(x) \)?

(iii) Show that every solution of the Cauchy problem which is initially 1-periodic in \( x \) converges, exponentially fast in time, to a wave solution of the form (2.3.79).

If in doubt, you may consult [111]. Note that the topological degree argument used in that reference can be replaced by a more elementary implicit function theorem argument we have used in the existence proof here.

2.4 A glimpse of the classical solutions to the Hamilton-Jacobi equations

2.4.1 Smooth solutions and their limitations

We now turn our attention to first order inviscid Hamilton-Jacobi equations of the form

\[
    u_t + H(x, \nabla u) = 0.
\]

The standard philosophy of the construction of a solution to a first order equation is to find its values on special curves, known as characteristics, that will, hopefully, fill the whole space. This is the strategy that is also classically used to solve (2.4.1). Consider a time \( t > 0 \) and a point \( x \in \mathbb{R}^n \). In order to assign a value to \( u(t, x) \) we consider a curve \( \gamma(s) \), with \( s \in [0, t] \), such that \( \gamma(t) = x \), and set

\[
    p(s) = \nabla u(s, \gamma(s)).
\]

Here, \( u(t, x) \) is the sought for solution to (2.4.1). Assuming that everything is smooth we have, using the dot to denote the differentiation in \( s \):

\[
    \dot{p}_k(s) = \partial_{x_k} u_t(s, \gamma(s)) + \frac{\partial^2 u(s, \gamma(s))}{\partial x_k \partial x_m} \dot{\gamma}_m(s)
    = -\frac{\partial H(\gamma(s), p(s))}{\partial x_k} - \frac{\partial H(\gamma(s), p(s))}{\partial p_m} \frac{\partial^2 u(s, \gamma(s))}{\partial x_k \partial x_m} + \frac{\partial^2 u(s, \gamma(s))}{\partial x_k \partial x_m} \dot{\gamma}_m(s).
\]

(2.4.2)
We see that it is convenient to choose \( \gamma(s) \) that satisfies the following system of ODEs:

\[
\dot{\gamma}(s) = \nabla_p H(\gamma(s), p(s)) \\
\dot{p}(s) = -\nabla_x H(\gamma(s), p(s))
\]  

for \( 0 \leq s \leq t \). This dynamical system is to be complemented by the boundary conditions at \( s = 0 \) and \( s = t \):

\[
p(0) = \nabla u_0(\gamma(0)), \quad \gamma(t) = x.
\]  

The system (2.4.3) has the form of a Hamiltonian system with the Hamiltonian \( H(x, p) \), and the curves \( (\gamma(s), p(s)) \) are called the characteristic curves. In order to solve (2.1.1), we need to find a solution to (2.4.3)-(2.4.4), and it would be excellent to prove that such solution is unique. The trouble is that there is no good reason, in general, for existence and uniqueness of a solution to this boundary value problem.

Exercise 2.4.1 Consider \( x_0 \in \mathbb{R}^n \) and \( t > 0 \) and assume that \( u(t, x) \) is smooth in a ball around \( x_0 \). Prove, for instance, with the help of the implicit function theorem, that the boundary value problem (2.4.3)-(2.4.4) has a unique solution \( (\gamma(s), p(s)) \) as soon as \( t \) is small enough and \( x \) is in the vicinity of \( x_0 \), and that this solution is smooth in \( t \) and \( x \).

Once \( \gamma(s) \) and \( p(s) \) are constructed, we may assign a value to \( u(t, x) \) as follows. The function \( \varphi(s) = u(s, \gamma(s)) \) satisfies

\[
\dot{\varphi}(s) = u_t(s, \gamma(s)) + \dot{\gamma}(s) \cdot p(s) = -H(\gamma(s), p(s)) + \dot{\gamma}(s) \cdot p(s).
\]  

Integrating (2.4.5) from \( s = 0 \) to \( s = t \) gives an expression for \( u(t, x) \) in terms of the curves \( \gamma(s) \) and \( p(s) \), \( 0 \leq s \leq t \):

\[
u(t, x) = u_0(\gamma(0)) + \int_0^t \left( -H(\gamma(s), p(s)) + \dot{\gamma}(s) \cdot p(s) \right) ds.
\]  

Exercise 2.4.2 Check that (2.4.6) indeed gives a solution to (2.4.1) such that \( u(0, x) = u_0(x) \).

To see that this strategy can not always lead to smooth solutions for all times, just consider the simplest nonlinear equation in one space dimension

\[
u_t + \frac{u_x^2}{2} = 0 \quad \text{for} \ t > 0 \text{ and } x \in \mathbb{R}, \quad u(0, x) = u_0(x).
\]  

The solution to the boundary value problem (2.4.3)-(2.4.4) amounts (this is very easily checked) to finding \( \gamma(0) \) solving the equation

\[
x = \gamma(0) + tu'_0(\gamma(0)),
\]  

for a given \( t > 0 \) and \( x \in \mathbb{R} \). The issue is that this equation may, or may not have a unique solution \( \gamma(0) \). If \( u''_0 > 0 \), solution is unique and we are on the safe side. But if \( u''_0(x_0) < 0 \) at some point \( x_0 \), uniqueness fails as soon as

\[
t \geq \frac{1}{\sup(-u''_0)}.
\]
Thus, we need a more elaborate theory. Nevertheless, in the rest of this section, we wish to show the reader one interesting situation where smooth solutions can be constructed.

Before we end this short section, let us mention, in the form of an exercise (this will be revisited in the context of viscosity solutions), a very strong form of uniqueness.

**Exercise 2.4.3 (Finite speed of propagation).** Let $H$ be uniformly Lipschitz with respect to its second variable, as well as $\nabla_p H$. Let $u_0$ and $v_0$ be two smooth, compactly supported initial conditions, and assume that each generates a smooth solution to the Cauchy problem for (2.4.1), on a common time interval $[0,T]$. Compute, in terms of $H_p$, a constant $K$ such that, if

$$\text{dist}(x, \text{supp}(u_0 - v_0)) > Kt,$$

then $u(t,x) = v(t,x)$. Hint: it may be helpful to solve, first, the following question: let $b(t,x)$ be smooth and uniformly Lipschitz in its second variable. Let $u_0$ be a smooth compactly supported function, and $u(t,x)$ the solution to

$$u_t + b(t,x) \cdot \nabla u = 0, \quad t > 0, x \in \mathbb{R}^n$$

$$u(0, x) = u_0(x) \quad (2.4.8)$$

If

$$\text{dist}(x, \text{supp}(u_0)) > t\|b\|_{\infty},$$

then $u(t,x) = 0$.

**2.4.2 An example of classical global solutions**

We now discuss a situation when classical smooth solutions do exist. Consider solutions to the equation

$$u_t + \frac{1}{2} |\nabla u|^2 - R(x) = 0, \quad (2.4.9)$$

with an initial condition $u(0, x) = u_0(x)$. We assume that both $u_0$ and $R$ are strictly convex smooth functions on $\mathbb{R}^n$, such that there is $\alpha \in (0,1)$ so that, for all $x \in \mathbb{R}^n$ and all $\xi \in \mathbb{R}^n$ we have:

$$\alpha|\xi|^2 \leq (D^2 u_0(x) \xi \cdot \xi) \leq \alpha^{-1} |\xi|^2, \quad \alpha|\xi|^2 \leq (D^2 R(x) \xi \cdot \xi) \leq \alpha^{-1} |\xi|^2. \quad (2.4.10)$$

**Exercise 2.4.4** First, consider the case $R = 0$. Argue informally, just by looking at the equation and using pictures that if $u_0(x)$ is strictly convex but its Hessian is uniformly bounded then the graph of $u(t,x)$ should not form a corner, and if $u_0(x)$ is strictly concave but its Hessian is uniformly bounded then it is plausible that the graph of $u(t,x)$ will form a corner. It may be helpful to start by looking at $u(0, x) = |x|^2$ and $u(0, x) = -|x|^2$. Explain, also informally, how the comparison principle should may come into play.

We now use the approach via the characteristic curves to show that a smooth solution exists under the above assumptions. Note that the Hamiltonian for (2.4.9) is

$$H(x,p) = \frac{1}{2} |p|^2 - R(x),$$

112
and the characteristic system \((2.4.3)-(2.4.4)\) reduces to
\[
\dot{\gamma}(s) = p(s), \quad \dot{p}(s) = \nabla R(\gamma(s)),
\]
which can be written as
\[
-\gamma'' + \nabla R(\gamma) = 0,
\]
with the boundary conditions
\[
\gamma'(0) - \nabla u_0(\gamma(0)) = 0, \quad \gamma(t) = x. \tag{2.4.11}
\]
To establish uniqueness and smoothness of the solution \(u(t, x)\) to \((2.4.9)\) with the initial condition \(u(0, x) = u_0(x)\), we need to prove that \((2.4.11)-(2.4.12)\) has a unique solution \(\gamma(s)\) that depends smoothly on \(t\) and \(x\). Then, \(u(t, x)\) will be given by (2.4.6):
\[
u(t, x) = u_0(\gamma(0)) + \int_0^t (- H(\gamma(s), p(s)) + \dot{\gamma}(s) \cdot p(s)) ds. \tag{2.4.13}
\]

**Existence of the characteristic curves**

To construct a solution to \((2.4.11)-(2.4.12)\), we observe that \((2.4.11)\) is the Euler-Lagrange equation for the energy functional
\[
J_{t,x}(\gamma) = u_0(\gamma(0)) + \int_0^t \left( \frac{|\gamma'(s)|^2}{2} + R(\gamma(s)) \right) ds, \tag{2.4.14}
\]
over \(H^1([0, t])\), with the constraint \(\gamma(t) = x\).

**Exercise 2.4.5** Verify that claim: show that if the minimizer of \(J_{t,x}(\gamma)\) over the set
\[
S = \{ \gamma \in H^1[0, t] : \gamma(t) = x \}
\]
exists and is smooth then it satisfies both \((2.4.11)\) and the boundary condition at \(s = 0\) in \((2.4.12)\). Next, define what it means for \(\gamma \in H^1[0, t]\) (without assuming \(\gamma\) is smooth) to be a weak solution to \((2.4.11)-(2.4.12)\) and show that a minimizer of \(J_{t,x}\) over \(S\) (if it exists) is a weak solution.

As both \(u_0(x)\) and \(R(x)\) are strictly convex, they are bounded from below, and it is easy to see that the functional \(J_{t,x}\) is bounded from below over \(S\). Let us set
\[
\bar{J}_{t,x} = \inf_{\gamma \in S} J_{t,x}(\gamma),
\]
and let \(\gamma_n \in S\) be a minimizing sequence, so that \(J_{t,x}(\gamma_n)\) decreases to \(\bar{J}_{t,x}\). Once again, as \(u_0\) and \(R\) are bounded from below, there exists \(C > 0\) so that
\[
\int_0^t |\gamma'_n(s)|^2 ds \leq C, \quad \text{for all } n.
\]
As, in addition, \(\gamma_n(t) = x\) for all \(n\), there is a subsequence, that we will still denote by \(\gamma_n\) that converges uniformly over \([0, t]\), and weakly in \(H^1([0, t])\) to a limit \(\bar{\gamma}_{t,x} \in S\).
To prove that $J_{t,x}(\bar{\gamma}_{t,x}) = J_{t,x}$ we simply observe that by the weak convergence we have

$$\|\bar{\gamma}'_{t,x}\|_{L^2}^2 \leq \lim\inf_{n \to +\infty} \|\gamma'_n\|_{L^2}^2,$$

which, combined with the uniform convergence of $\gamma_n$ to $\bar{\gamma}_{t,x}$ on $[0,t]$ implies that

$$J_{t,x}(\bar{\gamma}_{t,x}) \leq \lim_{n \to +\infty} J_{t,x}(\gamma_n) = \bar{J}_{t,x},$$

and thus

$$J_{t,x}(\bar{\gamma}_{t,x}) = \bar{J}_{t,x}.$$

**Uniqueness of the characteristic curve**

To prove the uniqueness of the minimizer, we will use the convexity of $u_0(x)$ and $R(x)$ and not just their boundedness from below. Let $\gamma_1$ and $\gamma_2$ be two solutions to (2.4.11)-(2.4.12). The difference

$$\tilde{\gamma} = \gamma_2 - \gamma_1.$$

satisfies

$$- \tilde{\gamma}'' + A_{kj} \tilde{\gamma}_j = 0, \quad 1 \leq k \leq n,$$

with the boundary conditions

$$\tilde{\gamma}'_k(0) - B_{kj} \tilde{\gamma}_j(0) = 0, \quad \tilde{\gamma}_k(t) = 0, \quad 1 \leq k \leq n.$$

The matrices $A$ and $B$ are given by

$$A_{kj} = \int_0^1 \frac{\partial^2 R(\gamma_1(s) + \sigma(\gamma_2(s) - \gamma_1(s)))}{\partial x_k \partial x_j} d\sigma,$$

and

$$B_{kj} = \int_0^1 \frac{\partial^2 u_0(\gamma_1(s) + \sigma(\gamma_2(s) - \gamma_1(s)))}{\partial x_k \partial x_j} d\sigma.$$

Let us take the inner product of (2.4.15) with $\tilde{\gamma}$, and integrate. This gives

$$\int_0^t \left( |\tilde{\gamma}'(s)|^2 + (A\tilde{\gamma}(s) \cdot \tilde{\gamma}(s)) \right) ds + (B\tilde{\gamma}(0) \cdot \tilde{\gamma}(0)) = 0. \quad (2.4.17)$$

Using (2.4.10), we deduce that the matrices $A$ an $B$ are strictly positive definite. Thus, (2.4.17) implies that $\tilde{\gamma}(s) \equiv 0$, so that the minimizer is unique. Hence, $u(t, x)$ is well-defined by (2.4.6):

$$u(t, x) = u_0(\gamma(0)) + \int_0^t \left( - \frac{|p(s)|^2}{2} + R(\gamma(s)) + \dot{\gamma}(s) \cdot p(s) \right) ds$$

$$= u_0(\gamma(0)) + \int_0^t \left( \frac{|\dot{\gamma}(s)|^2}{2} + R(\gamma(s)) \right) ds. \quad (2.4.18)$$

This may be rephrased as

$$u(t, x) = \inf_{\gamma(t) = x} \left( u_0(\gamma(0)) + \int_0^t \left( \frac{|\gamma'(s)|^2}{2} + R(\gamma(s)) \right) ds \right). \quad (2.4.19)$$
This formula, known as the Lax-Oleinik formula, is the starting point of the Lagrangian theory of Hamilton-Jacobi equations, and has immense implications. We will spend some time with this aspect of Hamilton-Jacobi equations later in this chapter. We will see that we can take it as a good definition of a solution to the Cauchy problem, at least when the Hamiltonian is strictly convex in $p$.

**Smoothness of the solution**

Let us quickly examine the smoothness of $u(t, x)$ in $x$ in the set-up of the present section. We see from (2.4.13) that it is equivalent to the smoothness of the minimizer $\gamma$ in $x$. If $h \in \mathbb{R}$ and $i \in \{1, \ldots, n\}$, consider the partial difference

$$\gamma^i_h(s) = \frac{\gamma_{t,x+he_i}(s) - \gamma_{t,x}(s)}{h}.$$  

It solves a system similar to (2.4.15), except for the boundary condition at $s = t$ that is now $\gamma^i_h(t) = e_i$. The exact same integration by parts argument yields the uniform boundedness of $\|\gamma^i_h\|_{H^1}$, hence the uniform boundedness of $\gamma^i_h$. Sending $h$ to $0$ and repeating the analysis shows that $\gamma^i_h$ converges to the unique solution of an equation of the type (2.4.15), with

$$A(s) = D^2 R(\gamma_{t,x}(s)), \quad B = D^2 u_0(\gamma_{t,x}(0)).$$  

This argument may be repeated over and over again, to yield the $C^\infty$ smoothness of $\gamma_{t,x}$ in $t$ and $x$, as long as $u_0$ and $R(x)$ are infinitely differentiable. Finally, using (2.4.6) we can conclude that

$$u(t, x) = \bar{J}_{t,x},$$  

is infinitely differentiable as well.

**Exercise 2.4.6** Show that $u$ is convex in $x$, for all $t > 0$, in two ways. First, fix $\xi \in \mathbb{R}^n$ and get a differential equation for $Q(t, x) = (D^2 u(t, x) \xi \cdot \xi)$. Use a maximum principle type argument to conclude that $Q(t, x) > 0$ for all $t > 0$ and $x \in \mathbb{R}^n$. An alternative and more elegant way is to proceed as follows.

(i) Assume the existence of $\kappa > 0$ such that, for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ and $\sigma \in [0, 1]$, we have:

$$u(t, \sigma x + (1 - \sigma)y) \leq \sigma u(t, x) + (1 - \sigma)u(t, y) - \kappa \sigma (1 - \sigma)|x - y|^2. \tag{2.4.20}$$

Show that then the function $u(t, x)$ is strictly convex.

(ii) Show that there exists $\lambda > 0$ such that if $\gamma_{t,x}$ and $\gamma_{t,y}$ are, respectively, the minimizing curves for $u(t, x)$ and $u(t, y)$, then

$$u(t, \sigma x + (1 - \sigma)y) \leq \sigma u(t, x) + (1 - \sigma)u(t, y)$$

$$- \lambda \sigma (1 - \sigma) \left( |\gamma_{t,x}(0) - \gamma_{t,y}(0)|^2 + \|\gamma_{t,x} - \gamma_{t,y}\|_{H^1([0, t])}^2 \right). \tag{2.4.21}$$

Hint: use the test curve $\gamma_\sigma = \sigma \gamma_{t,x} + (1 - \sigma)\gamma_{t,y}$ in the Lax-Oleinik formula (2.4.19) for $u(t, \sigma x + (1 - \sigma)y)$, together with the convexity of the functions $u_0(x)$ and $R(x)$.  

115
Finish the proof of (2.4.20), by noticing that

\[ |\gamma_{t,x}(0) - \gamma_{t,y}(0)|^2 = |x - y|^2 - \int_0^t \frac{d}{ds}|\gamma_{t,x}(s) - \gamma_{t,y}(s)|^2 ds. \]

The qualitative behavior of \( u(t, x) \) can be investigated further, implying the large time stabilization of the whole solution. We will come back to this class of questions later, when we study the large time behavior of viscosity solutions on the torus. For the time being, we leave the classical theory.

### 2.5 Viscosity solutions

We have just seen that, in order to find reasonable solutions to an inviscid Hamilton-Jacobi equation

\[ u_t + H(x, \nabla u) = 0, \quad (2.5.1) \]

we should relax the constraint that "\( u \) is continuously differentiable". The first idea would be to replace it by "\( u \) is Lipschitz", and require (2.5.1) to hold almost everywhere. Alas, there are, in general, several such solutions to the Cauchy problem for (2.5.1) with a Lipschitz (or even smooth) initial condition. This parallels the fact that the weak solutions to the conservation laws are not unique – for uniqueness, one must require that the weak solution satisfies the entropy condition. See, for instance, [93] for a discussion of these issues. A simple illustration of this phenomenon is to consider the Hamilton-Jacobi equation

\[ u_t + \frac{1}{2} u_x^2 = 0, \quad (2.5.2) \]

in one dimension, with the Lipschitz continuous initial condition

\[ u_0(x) = 0 \text{ for } x \leq 0 \text{ and } u_0(x) = x \text{ for } x > 0. \quad (2.5.3) \]

It is easy to check that one Lipschitz solution to (2.5.2) that satisfies this equation almost everywhere and obeys the initial condition (2.5.3) is

\[ u^{(1)}(t, x) = 0 \text{ for } x < t/2 \text{ and } u^{(1)}(t, x) = x - t/2 \text{ for } x > t/2. \]

However, another solution to (2.5.4)-(2.5.3) is given by

\[ u^{(2)}(t, x) = 0 \text{ for } x < 0, \quad u^{(2)}(t, x) = \frac{x^2}{2t} \text{ for } 0 < x < t \text{ and } u(t, x) = x - \frac{t}{2} \text{ for } x > t. \]

**Exercise 2.5.1** Consider the solution \( u^\varepsilon(t, x) \) to a viscous version of (2.5.4):

\[ u_t^\varepsilon + \frac{1}{2}(u_x^\varepsilon)^2 = \varepsilon u_{xx}^\varepsilon, \quad (2.5.4) \]

also with the initial condition \( u^\varepsilon(0, x) = u_0(x) \), as in (2.5.3). Use the Hopf-Cole transform

\[ v^\varepsilon(t, x) = \exp \left( - \frac{u^\varepsilon(t/\varepsilon, x)}{2\varepsilon} \right), \]
to show that $v^\varepsilon$ satisfies the standard heat equation

$$v^\varepsilon_t = v^\varepsilon_{xx}.$$ 

Find $v^\varepsilon(t, x)$ explicitly and use this to show that

$$u^\varepsilon(t, x) \to u^{(2)}(t, x) \text{ as } \varepsilon \to 0.$$ 

A natural question is, therefore, to know if an additional condition, less stringent than the $C^1$-regularity, but stronger than the mere Lipschitz regularity, enables us to select a unique solution to the Cauchy problem — as the notion of the entropy solutions does for the conservation laws. Exercise 2.5.1 suggests that regularizing the inviscid Hamilton-Jacobi equation with a small diffusion can provide one such approach, but for more general Hamilton-Jacobi equations than (2.5.4), for which the Hopf-Cole transform is not available, this procedure would be much less explicit.

The above considerations have motivated the introduction, by Crandall and Lions [44], at the beginning of the 1980’s, of the notion of a viscosity solution to (2.1.1). The idea is to select, among all the solutions of (2.1.1), “the one that has a physical meaning”, intrinsically, without directly appealing to the small diffusion regularization, — though understanding the connection to physics may require some additional thought. Being weaker than the notion of a classical solution, it introduces new difficulties to the existence, regularity and uniqueness issues, as well as into getting insight into the qualitative properties of solutions.

Finally, looking ahead, we mention that even if there is a unique viscosity solution to the Cauchy problem associated to (2.1.1), there will be no clear reason for the stationary equation (2.1.6) to have a unique steady viscosity solution — unlike what we have seen in the diffusive situation.

As a concluding remark to this introduction, we must mention that we will by no means do justice to a very rich subject in this short section and provide just a brief glance of a still developing subject. The reader who wishes to learn more may enjoy reading Barles [7], or Lions [93] as a starting point.

2.5.1 The definition and the basic properties of the viscosity solutions

The definition of a viscosity solution

Let us begin with more general equations than (2.1.1) — we will restrict the assumptions as the theory develops. Consider the Cauchy problem

$$u_t + F(x, u, \nabla u) = 0, \quad t > 0, \quad x \in \mathbb{T}^n,$$

with a continuous initial condition $u(0, x) = u_0(x)$, and $F \in C(\mathbb{T}^n \times \mathbb{R} \times \mathbb{R}^n; \mathbb{R})$.

In order to motivate the notion of a viscosity solution, one takes the point of view that the smooth solutions to the regularized problem

$$u^\varepsilon_t + F(x, u^\varepsilon, \nabla u^\varepsilon) = \varepsilon \Delta u^\varepsilon$$

117
are a good approximation to $u(t, x)$. Existence of the solution to the Cauchy problem for (2.5.6) for $\varepsilon > 0$ is not really an issue – we have already seen how it can be proved. Hence, a natural attempt would be to pass to the limit $\varepsilon \downarrow 0$ in (2.5.6). It is possible to prove that there is a unique limiting solution and that one actually ends up with a nonlinear semigroup. In particular, one may show that, if we take this notion of solution as a definition, there are uniqueness and contraction properties analogous to what we will see below – see [93] for further details. Taking this limit as a definition, however, raises an important issue: there is always the danger that the solution depends on the underlying regularization – why regularize with the Laplacian? What if we were to regularize differently? For instance, what if we would consider a dispersive regularization in one dimension

$$u_\varepsilon^t + F(x, u_\varepsilon, u_\varepsilon^x) = \varepsilon u_\varepsilon^{xxx}, \quad x \in \mathbb{R},$$

(2.5.7)

which is a generalized Korteweg-de Vries equation, and let $\varepsilon \to 0$ in (2.5.7) instead?

We now describe an alternative and more intrinsic approach, instead of using (2.5.6) in this very direct fashion of passing to the limit $\varepsilon \downarrow 0$. The idea is that the key property that should be inherited from the diffusive regularization is the maximum principle, as it is usually inherent in the origins of such models in the corresponding applications, be it physics, such as motion of interfaces, or optimal control problems. There is an interesting separate question of what happens as $\varepsilon \to 0$ to the solutions coming from regularizations that do not admit the maximum principle, such as (2.5.7). The situation is not quite trivial, especially for non-convex fluxes $F$ – we refer an interested reader to [92].

Our approach will be to use the comparison principle idea to extend the notions of a sub-solution and a super-solution to (2.5.5) and then simply say that a function $u(t, x)$ is a solution to (2.5.5) if it is both a sub-solution and a super-solution. To understand the upcoming definition of a viscosity sub-solution to (2.5.5), consider first a smooth sub-solution $u(t, x)$ to the regularized problem (2.5.6):

$$u_\varepsilon^t + F(x, u_\varepsilon, \nabla u) \leq \varepsilon \Delta u.$$  (2.5.8)

Let us take a smooth function $\phi(t, x)$ such that the difference $\phi - u$ attains its minimum at a point $(t_0, x_0)$. One may simply think of the case when $\phi(t_0, x_0) = u(t_0, x_0)$ and $\phi(t, x) \geq u(t, x)$ elsewhere. Then, at this point we have

$$u_t(t_0, x_0) = \phi_t(t_0, x_0), \quad \nabla \phi(t_0, x_0) = \nabla u(t_0, x_0),$$

and

$$D^2 \phi(t_0, x_0) \geq D^2 u(t_0, x_0),$$

in the sense of the quadratic forms. It follows that

$$\phi_t(t_0, x_0) + F(x_0, u(t_0, x_0), \nabla u(t_0, x_0)) - \varepsilon \Delta \phi(t_0, x_0)$$

$$\leq u_t(t_0, x_0) + F(x_0, u(t_0, x_0), \nabla u(t_0, x_0)) - \varepsilon \Delta u(t_0, x_0) \leq 0.$$  (2.5.9)

In other words, if $u$ is a smooth sub-solution to (2.5.6), and $\phi$ is a smooth function that touches $u$ at the point $(t_0, x_0)$ from above, then $\phi$ is also a sub-solution to (2.5.6) at this point.
In a similar vein, if \( u(t, x) \) is a smooth super-solution to the regularized problem:

\[
u_t + F(x, u, \nabla u) \geq \varepsilon \Delta u,
\]

we consider a smooth function \( \phi(t, x) \) such that the difference \( \phi - u \) attains its maximum at a point \((t_0, x_0)\). Again, we may assume without loss of generality that \( \phi(t_0, x_0) = u(t_0, x_0) \) and \( \phi(t, x) \leq u(t, x) \) elsewhere. Then, at this point we have

\[
\phi_t(t_0, x_0) + F(x_0, u(t_0, x_0), \nabla \phi(t_0, x_0)) - \varepsilon \Delta \phi(t_0, x_0) \geq 0,
\]

by a computation similar to (2.5.9). That is, if \( u \) is a smooth super-solution to (2.5.6), and \( \phi \) is a smooth function that touches \( u \) at \((t_0, x_0)\) from below, then \( \phi \) is also a super-solution to (2.5.6) at this point.

These two observations lead to the following definition, where we simply drop the requirement that \( u \) is smooth, only use the regularity of the test function that touches it from above or below, and send \( \varepsilon \to 0 \) in (2.5.9) and (2.5.11).

**Definition 2.5.2** A continuous function \( u(t, x) \) is a viscosity sub-solution to

\[
u_t + F(x, u, \nabla u) = 0,
\]

if, for all test functions \( \phi \in C^1([0, +\infty) \times \mathbb{T}^n) \) and all \((t_0, x_0) \in (0, +\infty) \times \mathbb{T}^n \) such that \((t_0, x_0) \) is a local minimum for \( \phi - u \), we have:

\[
\phi_t(t_0, x_0) + F(x_0, u(t_0, x_0), \nabla \phi(t_0, x_0)) \leq 0.
\]

Furthermore, a continuous function \( u(t, x) \) is a viscosity super-solution to (2.5.12) if, for all test functions \( \phi \in C^1((0, +\infty) \times \mathbb{T}^n) \) and all \((t_0, x_0) \in (0, +\infty) \times \mathbb{T}^n \) such that the point \((t_0, x_0) \) is a local maximum for the difference \( \phi - u \), we have:

\[
\phi_t(t_0, x_0) + F(x_0, u(t_0, x_0), \nabla \phi(t_0, x_0)) \geq 0.
\]

Finally, \( u(t, x) \) is a viscosity solution to (2.5.12) if it is both a viscosity sub-solution and a viscosity super-solution to (2.5.12).

Definition 2.5.2 extends to steady equations of the type

\[
F(x, u, \nabla u) = 0 \text{ on } \mathbb{T}^n,
\]

by requiring that \( u(x) \) is a viscosity sub-solution (respectively, super-solution) to

\[
u_t + F(x, u, \nabla u) = 0,
\]

that happens to be time-independent.

This definition was introduced by Crandall and Lions in their seminal paper [44]. The name “viscosity solution” comes out of the diffusive regularization we have discussed above. Definition 2.5.2 is intrinsic and bypasses the philosophical question we have mentioned above: "Why regularize with the Laplacian?" much like the notion of an entropy solution does this for the conservation laws. We stress, however, that it does make the assumption that the
underlying model must respect the comparison principle. Let us also note that the notion of a viscosity solution has turned out to be also very much relevant to the second order elliptic and parabolic equations—especially those fully nonlinear with respect to the Hessian of the solution. There have been spectacular developments, which are out of the scope of this chapter.

The main issue we will need to face soon is whether such a seemingly weak definition has any selective power—can it possibly ensure uniqueness of the solution? The expectation is that it should, due to the general principle that "the comparison principle implies uniqueness".

First, the following exercises may help the reader gain some intuition.

**Exercise 2.5.3** Show that a $C^1$ solution to
\[ u_t + F(x, u, \nabla u) = 0, \quad t > 0, \quad x \in \mathbb{T}^n, \]  
(2.5.15)
is a viscosity solution.

**Exercise 2.5.4** Consider the Hamilton-Jacobi equation
\[ u_t + u_x^2 = 0, \quad x \in \mathbb{R}. \]  
(2.5.16)
(i) Which of the following two functions is a viscosity solution to (2.5.16): $v(t, x) = |x| - t$ or $w(t, x) = -t - |x|$? Hint: pay attention to the fact that at the point $x = 0$ a smooth function $\phi(t, x)$ can only touch $v(t, x)$ from the bottom, and $w(t, x)$ from the top. This will tell you something about $|\phi_x(t, 0)|$ and determine the answer to this question.

(ii) Consider (2.5.16) with a zigzag initial condition $u_0(x) = u(0, x)$:
\[ u_0(x) = \begin{cases} 
  x, & 0 \leq x \leq 1/2, \\
  1 - x, & 1/2 \leq x \leq 1,
\end{cases} \]  
(2.5.17)
extended periodically to $\mathbb{R}$. How will the viscosity solution $u(t, x)$ to the Cauchy problem look like? Where will it be smooth, and where will it be just Lipschitz? Hint: it may help to do this in at least two ways: (1) use the definition of the viscosity solution, (2) use the notion of the entropy solution for the Burgers’ equation for $v(t, x) = u_x(t, x)$ if you are familiar with the basic theory of one-dimensional conservation laws.

**Exercise 2.5.5** (*Intermezzo: a Laplace asymptotics of integrals*). Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a real-valued smooth function such that there are two positive constants $\alpha$ and $\beta$ such that
\[ \varphi(x) \geq \alpha |x|^2 - \beta. \]

For $\varepsilon > 0$, consider the integral
\[ I_\varepsilon = \int_{\mathbb{R}^n} e^{-\varphi(x)/\varepsilon} \, dx. \]

The goal of this exercise is to show that
\[ \lim_{\varepsilon \to 0} \left( - \varepsilon \log I_\varepsilon \right) = \min_{x \in \mathbb{R}^n} \varphi(x). \]  
(2.5.18)

Note that it suffices to assume that
\[ \min_{x \in \mathbb{R}^n} \varphi(x) = 0, \]  
(2.5.19)
and show that
\[ \lim_{\varepsilon \to 0} \left( - \varepsilon \log I_\varepsilon \right) = 0. \]  
(2.5.20)
Exercise 2.5.6 Let us add the term $\varepsilon u_{xx}$ to the right side of (2.5.16), which produces a solution $u_\varepsilon(t, x)$. Use the Hopf-Cole transformation $z_\varepsilon(t, x) = \exp(u_\varepsilon(t, x)/\varepsilon)$, solve the linear problem for $z(t, x)$ and then pass to the limit $\varepsilon \to 0$ using Exercise 2.5.5. Study what happens when $u'_0(x)$ has limits at $\pm \infty$.

Basic properties of the viscosity solutions

We now describe some basic corollaries of the definition of a viscosity solution.

Exercise 2.5.7 Show that the maximum of two viscosity subsolutions to (2.5.15) is a viscosity subsolution, and the minimum of two viscosity supersolutions is a viscosity supersolution.

Exercise 2.5.8 (Stability) Let $F_j(x, u, p)$ be a sequence of functions in $C(T^n \times \mathbb{R} \times \mathbb{R}^n)$, which converges locally uniformly to $F \in C(T^n \times \mathbb{R} \times \mathbb{R}^n)$. Let $u_j(t, x)$ be a sequence of viscosity solutions to (2.5.5) with $F = F_j$:

$$\partial_t u_j + F_j(x, u_j, \nabla u_j) = 0, \quad t > 0, \ x \in T^n, \quad (2.5.21)$$

and assume that $u_j$ converges locally uniformly to $u \in C([0, +\infty), T^n)$. Show that then $u$ is a viscosity solution to the limiting problem

$$u_t + F(x, u, \nabla u) = 0, \quad t > 0, \ x \in T^n, \quad (2.5.22)$$

Hint: if $(t_0, x_0)$ is, for instance, a local minimum of the difference $\phi - u$, one can turn it into a strict minimum by changing $\phi(t, x)$ into $\phi(x) + M((t - t_0)^2 + |x - x_0|^2)$, without changing $\phi_t(t_0, x_0)$ and $\nabla \phi(t_0, x_0)$. In this situation, show that there is a sequence $(t_j, x_j)$ of minima of $\phi - j$ converging to $(t_0, x_0)$, and use the fact that each $u_j$ is a viscosity solution to (2.5.21) to conclude.

The above exercise is extremely important: it shows that, somewhat similar to the weak solutions, we do not need to check the convergence of the derivatives of $u_j$ to the derivatives of $u$ to know that $u$ is a viscosity solution – this is an extremely useful property to have. Exercise 2.5.8 asserts that one may safely “pass to the limit” in equation (2.5.5), as soon as estimates on the moduli of continuity of the solutions are available rather than on the derivatives.

The next proposition says that viscosity solutions that are Lipschitz continuous do satisfy the equation in the classical sense almost everywhere.

Proposition 2.5.9 Let $u$ be a locally Lipschitz viscosity solution to

$$u_t + F(x, u, \nabla u) = 0, \quad t > 0, \ x \in T^n. \quad (2.5.23)$$

Then it satisfies (2.5.23) almost everywhere.

This property relies on the following lemma [44].

Lemma 2.5.10 At a point of differentiability $(t_0, x_0)$ of $u$, one may construct a $C^1$ test function $\phi(t, x)$ such that $(t_0, x_0)$ is a local maximum (respectively, a local minimum) of $\phi - u$. 

121
**Proof.** For simplicity, we do not take the $t$-dependence into account, leaving this to the reader, so we assume that $u(x)$ is a function of $x$ that is differentiable at $x_0$. Without loss of generality, we assume that $x_0 = 0$, $u(0) = 0$, and that $\nabla u(0) = 0$, so that $u(x)$ satisfies, in the vicinity of $x = 0$:

$$u(x) = |x|\varepsilon(x), \quad \lim_{|x| \to 0} \varepsilon(x) = 0. \quad (2.5.24)$$

Our goal is to construct a $C^1$ function $\phi(x)$ such that $\phi(x) \leq u(x)$ and $\phi(0) = 0$. Note that this could be impossible for $u(x)$ that is merely Lipschitz and not differentiable – the simple counterexample is $u(x) = -|x|$. We look for a radially symmetric function $\phi(x)$ in the form $\phi(x) = |x|\zeta(|x|)$ with a $C^1$-function $\zeta(r)$ such that

$$\zeta(|x|) \leq \varepsilon(x), \quad \lim_{r \to 0} \zeta(r) = 0. \quad (2.5.25)$$

To this end, consider the decreasing sequence

$$\varepsilon_n = \inf_{2^{-n-1} \leq |r| < 0} \varepsilon(r),$$

and produce the function $\zeta(r) \leq \varepsilon(r)$ by smoothing the piecewise constant function

$$\sum_{n=0}^{+\infty} \varepsilon_n 1_{2^{-n-1} \leq r < 2^{-n}}.$$

As the sequence $\varepsilon_n \to 0$ as $n \to +\infty$, and we have chosen the dyadic intervals in the above sum, we may ensure that

$$|\zeta'(r)| \leq \frac{\alpha(r)}{r},$$

with $\alpha(r) \to 0$ as $r \to 0$. It follows that $\phi(x) = |x|\zeta(|x|)$ is the sought $C^1$-function. □

**Proof of Proposition 2.5.9.** The conclusion of this proposition follows essentially immediately from Lemma 2.5.10 and the Rademacher theorem. The latter says that a Lipschitz function is differentiable a.e., see for instance [57]. At any differentiability point we can construct a $C^1$-function $\phi(t, x)$ such that the difference $\phi - u$ attains its minimum at $(t_0, x_0)$, so that

$$\phi_t(t_0, x_0) = u_t(t_0, x_0) \text{ and } \nabla \phi(t_0, x_0) = \nabla u(t_0, x_0). \quad (2.5.26)$$

The definition of a viscosity sub-solution together with (2.5.26) implies that

$$u_t(t_0, x_0) + H(x, u(t_0, x_0), \nabla u(t_0, x_0)) = \phi_t(t_0, x_0) + H(x, u(t_0, x_0), \nabla \phi(t_0, x_0)) \leq 0.$$

Similarly, we can show that

$$u_t(t_0, x_0) + H(x, u(t_0, x_0), \nabla u(t_0, x_0)) \geq 0,$$

using the fact that $u(t, x)$ is a viscosity super-solution. This finishes the proof. □

**Warning.** For the rest of this section, a solution of (2.1.1) or (2.1.6) will always be meant in the viscosity sense.
2.5.2 Uniqueness of the viscosity solutions

Let us first give the simplest uniqueness result, that we will prove by the method of doubling of variables. This argument appears in almost all uniqueness proofs, in more or less elaborate forms.

**Proposition 2.5.11** Assume that the Hamiltonian \( H(x, p) \) is continuous in all its variables, and satisfies the coercivity assumption

\[
\lim_{|p| \to +\infty} H(x, p) = +\infty, \quad \text{uniformly in } x \in \mathbb{T}^n. \tag{2.5.27}
\]

Consider the equation

\[
H(x, \nabla u) + u = 0, \quad x \in \mathbb{T}^n. \tag{2.5.28}
\]

Let \( u \) and \( \overline{u} \) be, respectively, a viscosity sub- and a super-solution to (2.5.28), then \( u \leq \overline{u} \).

**Proof.** Assume for a moment that both \( u \) and \( \overline{u} \) are \( C^1 \)-functions, so that we can use each of them as a test function in the definition of the viscosity sub- and super-solutions. First, we use the fact that \( \overline{u} \) is a super-solution to (2.5.28) and take \( u \) as a test function. Let \( x_0 \) be a maximum of \( u - \overline{u} \), then we deduce from the definition of a viscosity super-solution to (2.5.28) that

\[
H(x_0, \nabla u(x_0)) + u(x_0) \geq 0. \tag{2.5.29}
\]

On the other hand, \( \overline{u} - u \) attains its minimum at the same point \( x_0 \), and, as \( u \) is a viscosity sub-solution to (2.5.28), and \( \overline{u} \) can serve as a test function, we have

\[
H(x_0, \nabla \overline{u}(x_0)) + \overline{u}(x_0) \leq 0. \tag{2.5.30}
\]

As \( x_0 \) is a minimum of \( \overline{u} - u \), and \( u \) and \( \overline{u} \) are differentiable, we have \( \nabla \overline{u}(x_0) = \nabla u(x_0) \), whence (2.5.29) and (2.5.30) imply

\[
u(x_0) \leq \overline{u}(x_0).
\]

Once again, as \( \overline{u} - u \) attains its minimum at \( x_0 \), we conclude that \( \overline{u}(x) \geq u(x) \) for all \( x \in \mathbb{T}^n \) if both of these functions are in \( C^1(\mathbb{T}^n) \). Unfortunately, we only know that \( u \) and \( \overline{u} \) are continuous, so we can not use the elegant argument above unless we know, in addition, that they are both \( C^1 \)-functions.

In the general case, the method of doubling the variables gives a way around the technical difficulty that \( u(x) \) and \( \overline{u}(x) \) are merely continuous and not necessarily differentiable. Let us define, for \( \varepsilon > 0 \), the penalization

\[
u_{\varepsilon}(x, y) = \overline{u}(x) - u(y) + \frac{|x - y|^2}{2\varepsilon^2}
\]

and let \((x_{\varepsilon}, y_{\varepsilon}) \in \mathbb{T}^{2n}\) be a minimum for \( u_{\varepsilon}(x, y) \).

**Exercise 2.5.12** Show that

\[
\lim_{\varepsilon \to 0} |x_{\varepsilon} - y_{\varepsilon}| = 0. \tag{2.5.31}
\]

and that the family \((x_{\varepsilon}, y_{\varepsilon})\) converges, as \( \varepsilon \to 0 \), up to a subsequence, to a point \((x_0, x_0)\), where \( x_0 \) is a minimum for \( \overline{u}(x) - u(x) \).
Consider the function
\[ \phi(x) = u(y_\varepsilon) - \frac{|x - y_\varepsilon|^2}{2\varepsilon^2}, \]
as a (smooth) quadratic function of the variable \( x \). The difference
\[ \phi(x) - \overline{u}(x) = -u_\varepsilon(x, y_\varepsilon) \]
attains its maximum, as a function of \( x \), at the point \( x = x_\varepsilon \). As \( \overline{u}(x) \) is a viscosity supersolution to (2.5.28), we have
\[ H(x_\varepsilon, \frac{y_\varepsilon - x_\varepsilon}{\varepsilon^2}) + \overline{u}(x_\varepsilon) \geq 0. \] (2.5.32)

Next, we apply the viscosity sub-solution part of Definition 2.5.13 to the quadratic test function
\[ \psi(y) = \overline{u}(x_\varepsilon) + \frac{|x_\varepsilon - y|^2}{2\varepsilon^2}. \]
The difference
\[ \psi(y) - u(y) = \overline{u}(x_\varepsilon) + \frac{|x_\varepsilon - y|^2}{2\varepsilon^2} - u(y) = u_\varepsilon(x_\varepsilon, y) \]
attains its minimum at \( y = y_\varepsilon \). As \( u(y) \) is a viscosity sub-solution to (2.5.28), we obtain
\[ H(y_\varepsilon, \frac{y_\varepsilon - x_\varepsilon}{\varepsilon^2}) + u_\varepsilon(x_\varepsilon, y) \leq 0. \] (2.5.33)
The coercivity of the Hamiltonian and (2.5.33), together with the boundedness of \( u_\varepsilon \), imply that \(|x_\varepsilon - y_\varepsilon|/\varepsilon^2\) is bounded, uniformly in \( \varepsilon \): there exists \( R \) so that
\[ \frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon^2} \leq R. \]

The uniform continuity of \( H(x, p) \) on the set \( \{(x, p) : x \in \mathbb{T}^n, p \in B(0, R)\} \) implies that, as consequence, we have
\[ H(y_\varepsilon, \frac{y_\varepsilon - x_\varepsilon}{\varepsilon^2}) = H(x_\varepsilon, \frac{y_\varepsilon - x_\varepsilon}{\varepsilon^2}) + o(1), \text{ as } \varepsilon \to 0. \]
Subtracting (2.5.33) from (2.5.32), we obtain
\[ \overline{u}(x_\varepsilon) - u(y_\varepsilon) \geq o(1), \text{ as } \varepsilon \to 0. \]

Sending \( \varepsilon \to 0 \) with the help of the result of Exercise 2.5.12 implies
\[ \overline{u}(x_0) - u(x_0) \geq 0, \]
and, as \( x_0 \) is the minimum of \( \overline{u} - u \), the proof is complete. \( \square \)

An immediate consequence of Proposition 2.5.11 is that (2.5.28) has at most one solution.
The comparison principle and weak contraction

The proof of Proposition 2.5.11 can be adapted to establish two fundamental properties for the viscosity solutions to the Cauchy problem: the comparison principle and the weak contraction property.

Exercise 2.5.13 (The comparison principle) Assume that $H(x,p)$, is a continuous function that satisfies the coercivity property (2.5.27). Let $u_1(t,x)$ and $u_2(t,x)$ be, respectively, a viscosity sub-solution, and a viscosity super-solution to

$$ u_t + H(x, \nabla u) = 0, \quad t > 0, \quad x \in \mathbb{T}^n, $$

(2.5.34)

with the initial conditions $u_{10}$ and $u_{20}$ such that $u_{10}(x) \leq u_{20}(x)$ for all $x \in \mathbb{T}^n$. Modify the proof of Proposition 2.5.11 to show that then $u_1(t,x) \leq u_2(t,x)$ for all $t \geq 0$ and $x \in \mathbb{T}^n$. This proves the uniqueness of the viscosity solutions.

Exercise 2.5.14 (Weak contraction) Let $H(x,p)$ be a continuous function that satisfies the coercivity property (2.5.27), and $u_1$ and $u_2$ be two solutions to (2.5.34) with continuous initial conditions $u_{10}$ and $u_{20}$, respectively. Show that then we have

$$ \|u_1(t,\cdot) - u_2(t,\cdot)\|_{L^\infty} \leq \|u_{10} - u_{20}\|_{L^\infty}. $$

Hint: notice that if $u(t,x)$ solves (2.5.34) then so does $u(t,x) + k$ for any $k \in \mathbb{R}$, and use Exercise 2.5.13.

2.5.3 Finite speed of propagation

We are now going to prove a finite speed of propagation property, partly to acquire some further familiarity with the notion of a viscosity solution, and partly to emphasize the sharp contrast with viscous models: if the equation carried a Laplacian, an initially nonnegative solution would instantly become positive everywhere. As this property makes better sense in $\mathbb{R}^n$ and not on the torus, this is the case we will consider.

Proposition 2.5.15 Let $H$ be uniformly Lipschitz with respect to its second variable:

$$ |H(x,p_1) - H(x,p_2)| \leq C_L|p_1 - p_2| \quad \text{for all } x \in \mathbb{R}^n \text{ and } p_1, p_2 \in \mathbb{R}^n. $$

(2.5.35)

Let $u_0$ and $v_0$ be two continuous, compactly supported initial conditions, and assume that each generates a globally Lipschitz solution, respectively denoted by $u(t,x)$ and $v(t,x)$ to the Cauchy problem

$$ u_t + H(x, \nabla u) = 0, \quad v_t + H(x, \nabla v) = 0, \quad 0 < t \leq T, \quad x \in \mathbb{R}^n, $$

(2.5.36)

with $u(0,x) = u_0(x)$ and $v(0,x) = v_0(x)$ for all $x \in \mathbb{R}^n$. Then, if $x_0 \in \mathbb{R}^n$ and $t_0 \in [0,T]$ satisfy

$$ \text{dist}(x_0, \text{supp}(u_0 - v_0)) > t_0 C_L, $$

then $u(t_0,x_0) = v(t_0,x_0)$. 

125
\textbf{Proof.} The idea is simple: assuming that everything is smooth, the function \(w = u - v\) satisfies the inequalities
\[
 w_t \leq C_L |\nabla w|, \quad (2.5.37)
\]
and
\[
 w_t \geq -C_L |\nabla w|. \quad (2.5.38)
\]

\textbf{Exercise 2.5.16} Use the method of characteristics to show that if \(w\) is a smooth function that satisfies (2.5.37) and
\[
 \text{dist}(x_0, \text{supp}(w(0, \cdot))) > C_L t_0, \quad (2.5.39)
\]
then \(w(t_0, x_0) \leq 0\), and if a smooth function \(w\) satisfies (2.5.38)-(2.5.39), then \(w(t_0, x_0) \geq 0\).

Thus, the conclusion of this proposition follows from Exercise 2.5.16 if \(u\) and \(v\) are smooth. Unfortunately, if \(u\) and \(v\) are not smooth, then we can not use the characteristics but only the definition of a viscosity solution. Let us fix a point \(x_0 \in \mathbb{R}^n\) and \(t_0 > 0\) so that
\[
 \text{dist}(x_0, \text{supp}(u_0 - v_0)) > C_L t_0, \quad (2.5.40)
\]
and let \(\phi_0(r)\) be a \(C^1\)-function equal to 1 outside of the the ball \(B_{C_L t_0 + \varepsilon}(0)\), and to 0 in the ball \(B_{C_L t_0}(0)\). The function
\[
 \bar{w}(t, x) = \|u_0 - v_0\|_{L^\infty} \phi_0(|x - x_0| + C_L t)
\]
is a smooth solution to
\[
 \partial_t \bar{w} - C_L |
abla \bar{w}| = 0, \quad t > 0, \quad x \in \mathbb{R}^n, \quad (2.5.43)
\]
such that \(\phi(t, x_0) = 0\) for \(t \leq t_0\). Moreover, because of (2.5.41), at \(t = 0\) we have
\[
 \bar{w}(0, x) = \|u_0 - v_0\|_{L^\infty} \phi_0(|x - x_0|) \geq |u_0(x) - v_0(x)| \quad \text{for all} \ x \in \mathbb{R}^n.
\]

Our goal is to show this inequality persists until the time \(t_0\):
\[
 |u(t, x) - v(t, x)| \leq \bar{w}(t, x) \quad \text{for all} \ 0 \leq t \leq t_0 \quad \text{and} \ x \in \mathbb{R}^n. \quad (2.5.45)
\]

Indeed, using (2.5.45) at \(x = x_0\) and \(t = t_0\) would give
\[
 |u(t_0, x_0) - v(t_0, x_0)| \leq \|u_0 - v_0\|_{L^\infty} \phi_0(C_L t_0) = 0,
\]
which is what we need.

The comparison principle in Exercise 2.5.13 together with (2.5.44) implies that (2.5.45) would follow if we show that \(\bar{v}(t, x) = u(t, x) + \bar{w}(t, x)\) is a viscosity super-solution to (2.5.36):
\[
 \partial_t \bar{v} + H(x, \nabla \bar{v}) \geq 0. \quad (2.5.47)
\]
Observe that (2.5.47) and (2.5.44) together would imply
\[ v(t, x) \leq \overline{v}(t, x) = u(t, x) + \overline{w}(t, x) \] for all \( 0 \leq t \leq t_0 \) and \( x \in \mathbb{R}^n \). (2.5.48)
As the roles of \( u \) and \( v \) can be reversed, we would deduce (2.5.45).

Thus, we need to prove the viscosity super-solution property for \( \overline{v}(t, x) \). Let \( \varphi(t, x) \) be a smooth test function, and \((t_1, x_1)\) be a minimum point for
\[ \overline{v}(t, x) - \varphi(t, x) = u(t, x) + \overline{w}(t, x) - \varphi(t, x) = u(t, x) - \psi(t, x), \] with a \( C^1 \)-function
\[ \psi(t, x) = \varphi(t, x) - \overline{w}(t, x). \]
In other words, \((t_1, x_1)\) is a minimum point for \( u(t, x) - \psi(t, x) \). As \( u \) is a viscosity solution to (2.5.36), it follows that
\[ \partial_t \psi(t_1, x_1) + H(x_1, \nabla \psi(t_1, x_1)) \geq 0, \] (2.5.50)
which is nothing but
\[ \partial_t \varphi(t_1, x_1) - \partial_t \overline{w}(t_1, x_1) + H(x_1, \nabla \varphi(t_1, x_1) - \nabla \overline{w}(t_1, x_1)) \geq 0, \] (2.5.51)
Using the inequality
\[ H(\bar{x}, \nabla \varphi - \nabla \overline{w}) \leq H(\bar{x}, \nabla \varphi) + C_L |\nabla \overline{w}|. \]
in (2.5.51) gives
\[ \partial_t \varphi(t_1, x_1) - \partial_t \overline{w}(t_1, x_1) + H(x_1, \nabla \varphi(t_1, x_1)) + C_L |\nabla \overline{w}(t_1, x_1)| \geq 0. \] (2.5.52)
Recalling (2.5.43), we obtain
\[ \partial_t \varphi(t_1, x_1) + H(x_1, \nabla \varphi(t_1, x_1)) \geq 0. \] (2.5.53)
We conclude that \( \overline{v}(t, x) \) is a viscosity super-solution to (2.5.36), finishing the proof.  

**Exercise 2.5.17** *(Hole filling property)*. Let \( u(t, x) \) be a viscosity solution to
\[ u_t = R(t, x)|\nabla u|, \quad t > 0, \ x \in \mathbb{R}^n, \]
with \( R(t, x) \geq R_0 > 0 \). Assume that (i) \( u(0, x) = u_0(x) \geq \delta_0 > 0 \) outside a ball \( B(0, R) \), and (ii) the set \( \mathbb{R}^n \backslash \text{(supp}(u_0)) \) is compact. Prove that there exists \( T_0 > 0 \) such that \( u(t, x) > 0 \) for all \( t \geq T_0 \), and all \( x \in \mathbb{R}^n \).

### 2.6 Construction of solutions

So far, we have set up a beautiful notion of viscosity solutions, and we have also proved that this works well in settling our worries about uniqueness, distinguishing them from the Lipschitz solutions. Now, we have to prove that, as far as existence is concerned, this new notion does better than the classical solutions, in the sense that solutions to the Cauchy
problem exist for all $t > 0$ under reasonable assumptions. In this section, we will show how these solutions can be constructed. First, we will produce wave solutions to the time-dependent problem

$$\partial_t u + H(x, \nabla u) = 0, \quad x \in \mathbb{T}^n. \quad (2.6.1)$$

Next, we are going to prove that the Cauchy problem for (2.6.1) is well-posed as soon as a continuous initial condition is specified. Eventually, we will show that the wave solutions describe the long time behavior of the solutions to the Cauchy problem.

### 2.6.1 Existence of waves, and the Lions-Papanicolaou-Varadhan theorem

Wave solutions for (2.6.1) will be sought in the same form as viscous waves, that is

$$v(t, x) = -ct + u(x), \quad (2.6.2)$$

with a constant $c \in \mathbb{R}$. A function $v(t, x)$ of this form is a solution to (2.6.1) if $u(x)$ solves a time-independent problem

$$H(x, \nabla u) = c, \quad x \in \mathbb{T}^n. \quad (2.6.3)$$

**Exercise 2.6.1** Show that a function $v(t, x)$ of the form (2.6.2) is a viscosity solution to (2.6.1) if and only if $u(x)$ is a viscosity solution to (2.6.3).

As before, we will think of $v(t, x)$ as the height of an interface, and refer to the constant $c$ as the speed of the wave, and to $u(x)$ as its shape. Let us point out that the speed is unique: (2.6.3) may have viscosity solutions for at most one $c$. Indeed, assume there exist $c_1 \neq c_2$, such that (2.6.3) has a viscosity solution $u_1$ for $c = c_1$ and another viscosity solution $u_2$ for $c = c_2$. Let $K > 0$ be such that

$$u_1(x) - K \leq u_2(x) \leq u_1(x) + K, \quad \text{for all } x \in \mathbb{T}^n.$$

By Exercise 2.6.1 the functions

$$v_{1, \pm}(t, x) = -c_1 t + u_1(x) \pm K$$

and

$$v_2(t, x) = -c_2 t + u_2(x)$$

are the viscosity solutions to the Cauchy problem (2.1.1) with the respective initial conditions

$$v_{1, \pm}(0, x) = u_1(x) \pm K, \quad v_2(0, x) = u_2(x).$$

By the comparison principle (Exercise 2.5.13), we have

$$-c_1 t + u_1(x) - K \leq -c_2 t + u_2(x) \leq -c_1 t + u_1(x) + K, \quad \text{for all } t \geq 0 \text{ and } x \in \mathbb{T}^n.$$

This is a contradiction since $c_1 \neq c_2$, and the functions $u_1$ and $u_2$ are bounded. Therefore, the wave speed $c$ is unique. Note that this does not address the question of uniqueness of the shape $u(x)$ – we leave this question for later.

The main result of this section is the following theorem, due to Lions, Papanicolaou, Varadhan [94], that asserts the existence of a constant $c$ for which (2.6.3) has a solution.
Theorem 2.6.2 Assume that $H(x,p)$ is continuous, uniformly Lipschitz:

$$|H(x,p_1) - H(x,p_2)| \leq C_L|p_1 - p_2|, \text{ for all } x \in \mathbb{T}^n, \text{ and } p_1, p_2 \in \mathbb{R}^n,$$

(2.6.4)

the coercivity condition

$$\lim_{|p| \to +\infty} H(x,p) = +\infty, \text{ uniformly in } x \in \mathbb{T}^n.$$

(2.6.5)

holds, and

$$|\nabla_x H(x,p)| \leq K_0(1 + |p|), \text{ for all } x \in \mathbb{T}^n, \text{ and } p \in \mathbb{R}^n.$$

(2.6.6)

Then there is a unique $c \in \mathbb{R}$ for which

$$H(x,\nabla) = c, \text{ } x \in \mathbb{T}^n.$$

(2.6.7)

has a solution.

It is important to point out that the periodicity assumption in $x$ on the Hamiltonian is indispensable – for instance, when $H(x,p)$ is a random function (in $x$) on $\mathbb{R}^n \times \mathbb{R}^n$, the situation is much more complicated – an interested reader should consult the literature on stochastic homogenization of the Hamilton-Jacobi equations, a research area that is active and evolving at the moment of this writing. We also mention that the only assumption made in [94] is that $H(x,p)$ is continuous and coercive. The Lipschitz condition (2.6.4) in $p$ and (2.6.6) in $x$ have been added here for convenience.

**The homogenization connection**

Before proceeding with the proof of the Lions-Papanicolaou-Varadhan theorem, let us explain how the steady equation (2.6.7) appears in the context of periodic homogenization, which was probably the main motivation behind this theorem. We can not possibly do justice to the area of homogenization here – an interested reader should explore the huge literature on the subject, with the book [115] by G. Pavliotis and A. Stuart providing a good starting point. Let us just briefly illustrate the general setting on the example of the periodic Hamilton-Jacobi equations. Consider the Cauchy problem

$$u_t^\varepsilon + H(x,\nabla^\varepsilon) = 0, \text{ } t > 0, \text{ } x \in \mathbb{R}^n,$$

(2.6.8)

in the whole space $\mathbb{R}^n$ (and not on the torus), with the Hamiltonian $H(x,p)$ that is 1-periodic in all coordinates $x_j, j = 1, \ldots, n$. We are interested in the regime where the initial condition is slowly varying and large:

$$u^\varepsilon(0, x) = \varepsilon^{-1} u_0(\varepsilon x).$$

(2.6.9)

Let us note that if one thinks of the solution to (2.6.8) as the height of some interface at the position $x \in \mathbb{R}^n$ at a time $t > 0$, then it is very natural that if $u^\varepsilon(0, x)$ varies on a scale $\varepsilon^{-1}$ in the $x$-variable, then its height should also be of the order $\varepsilon^{-1}$, which is exactly what we see in (2.6.9).

The general question of homogenization is how the "microscopic" variations in the Hamiltonian that varies on the scale $O(1)$ affect the evolution of the initial condition that varies on
the ”macroscopic” scale $O(\varepsilon^{-1})$. The goal is to describe the evolution in purely ”macroscopic” terms, and extract an effective macroscopic problem that approximates the full microscopic problem well. This allows to avoid, say, in numerical simulations, modeling the microscopic variations of the Hamiltonian, and do the simulations on the macroscopic scale – a huge advantage in engineering problems. It also happens that from the purely mathematical viewpoint, homogenization is also an extremely rich subject.

This general philosophy translates into the following strategy. As the initial condition in (2.6.9) is slowly varying, one should observe the solution on a macroscopic spatial scale, in the ”slow” variable $y = \varepsilon x$. Since $u^\varepsilon(0, x)$ is also very large itself, of the size $O(\varepsilon^{-1})$, it is appropriate to rescale it down. In other words, instead of looking at $u^\varepsilon(t, x)$ directly, we would represent it as

$$
u^\varepsilon(t, x) = \varepsilon^{-1}w^\varepsilon(t, \varepsilon x),$$

and consider the evolution of $w^\varepsilon(t, y)$, which satisfies

$$w_t^\varepsilon + \varepsilon H\left(\frac{y}{\varepsilon}, \nabla w^\varepsilon\right) = 0,$$

with the initial condition $w^\varepsilon(0, y) = u_0(y)$ that is now independent of $\varepsilon$. However, we see that $w^\varepsilon$ evolves very slowly in $t$ – its time derivative is of the size $O(\varepsilon)$. Hence, we need to wait a long time until it changes. To remedy this, we introduce a long time scale of the size $t = O(\varepsilon^{-1})$. In other words, we write

$$w^\varepsilon(t, y) = \nu^\varepsilon(\varepsilon t, y).$$

In the new variables the problem takes the form

$$\nu_s^\varepsilon + H\left(\frac{y}{\varepsilon}, \nabla \nu^\varepsilon\right) = 0, \quad y \in \mathbb{R}^n, \quad s > 0,$$

with the initial condition $\nu^\varepsilon(0, y) = u_0(y)$.

It seems that we have merely shifted the difficulty – we used to have $\varepsilon$ in the initial condition in (2.6.9) while now we have it appear in the equation itself – the Hamiltonian depends on $y/\varepsilon$. However, it turns out that we may now find an $\varepsilon$-independent problem that has a spatially uniform Hamiltonian that provides a good approximation to (2.6.11). The reason this is possible is that we have chosen the correct temporal and spatial scales to track the evolution of the solution.

Here is an informal way to find the approximating problem. Let us seek the solution to (2.6.11) in the form of an asymptotic expansion

$$\nu^\varepsilon(s, y) = \bar{\nu}(s, y) + \varepsilon v_1(s, y, \frac{y}{\varepsilon}) + \varepsilon^2 v_2(s, y, \frac{y}{\varepsilon}) + \ldots$$

The functions $v_j(s, y, z)$ are assumed to be periodic in the ”fast” variable $z$ but not in the ”slow” variables $s$ and $y$. Inserting this expansion into (2.6.11), and collecting the terms with various powers of $\varepsilon$, we obtain in the leading order

$$\bar{\nu}_s(s, y) + H\left(\frac{y}{\varepsilon}, \nabla_y \bar{\nu}(s, y) + \nabla_z v_1(s, y, \frac{y}{\varepsilon})\right) = 0.$$
As is standard in such multiple scale expansions, we consider (2.6.13) as
\[ \ddot{v}_s(s, y) + H(z, \nabla_y \ddot{v}(s, y) + \nabla_z v_1(s, y, z)) = 0, \quad z \in \mathbb{T}^n, \] an equation for \( v_1 \) as a function of the fast variable \( z \in \mathbb{T}^n \), for each \( s > 0 \) and \( y \in \mathbb{R}^n \) fixed.

In other words, for each pair of the "macroscopic" variables \( s \) and \( y \) we consider a microscopic problem in the \( z \)-variable. In the area of numerical analysis, one would call this "sub-grid modeling": the variables \( s \) and \( y \) live on the macroscopic grid, and the \( z \)-variable lives on the microscopic sub-grid.

The function \( \ddot{v}(s, y) \) will then be found from the solvability condition for (2.6.13). Indeed, the terms \( \ddot{v}_s(s, y) \) and \( \nabla_y \ddot{v}(s, y) \) in (2.6.14) do not depend on the fast variable \( z \) and should be treated as constants – we solve (2.6.14) independently for each \( s \) and \( y \). Let us then, for each fixed \( p \in \mathbb{R}^n \), consider the problem
\[ H(z, p + \nabla_z w) = c, \quad z \in \mathbb{T}^n. \] (2.6.15)
The case of interest is \( p = \nabla_y \ddot{v}(s, y) \) and \( c = -\ddot{v}_s(s, y) \) but one needs to momentarily look at (2.6.15) for an arbitrary choice of \( p \in \mathbb{R}^n \) and \( c \in \mathbb{R} \). The Lions-Papanicolaou-Varadhan theorem says that for each \( p \in \mathbb{R}^n \) there is a unique \( c \) that we will denote by \( \dddot{H}(p) \) such that (2.6.15) has a solution. We then write (2.6.15) as
\[ H(z, p + \nabla_z w) = \dddot{H}(p), \quad z \in \mathbb{T}^n. \] (2.6.16)

Hence, the solvability condition for (2.6.14) is that the function \( \ddot{v}(s, y) \) satisfies the homogenized (also known as "effective") equation
\[ \ddot{v}_s + \dddot{H}(\nabla_y \ddot{v}) = 0, \quad \ddot{v}(0, y) = u_0(y), \quad s > 0, \quad y \in \mathbb{R}^n, \] (2.6.17)
and the function \( \dddot{H}(p) \) is called the effective, or homogenized Hamiltonian. Note that the effective Hamiltonian does not depend on the spatial variable – the "small scale" variations are averaged out via the above homogenization procedure. The point is that the solution \( \ddot{v}^\varepsilon(s, y) \) of (2.6.11), an equation with highly oscillatory coefficients is well approximated by \( \ddot{v}(s, y) \), the solution of (2.6.17), an equation with spatially uniform coefficients, that is much simpler to study analytically or solve numerically.

Thus, the existence and uniqueness of the constant \( c \) for which solution of the steady equation (2.6.15) exists, is directly related to the homogenization (long time behavior) of the solutions to the Cauchy problem (2.6.8) with slowly varying initial conditions, as it provides the corresponding effective Hamiltonian. Unfortunately, there is a catch: not so much is known in general on how the effective Hamiltonian \( \dddot{H}(p) \) depends on the original Hamiltonian \( H(x, p) \), except for some very generic properties. Estimating and computing numerically the effective Hamiltonian \( \dddot{H}(p) \) is a separate interesting line of research.

**Exercise 2.6.3 (The one-dimensional case)** Compute the effective Hamiltonian \( \dddot{H}(p) \) for
\[ H(x, p) = R(x) \sqrt{1 + p^2}, \quad x \in \mathbb{T}^1, \quad p \in \mathbb{R}, \] where \( R(x) \) is a smooth 1-periodic function.
**Exercise 2.6.4** Show that for every \( p \in \mathbb{R}^n \) one can find a periodic in \( x \) function \( u(x; p) \), \( x \in \mathbb{T}^n, p \in \mathbb{R}^n \) such that the function

\[
v(t, x; p) = p \cdot x + u(x; p) - t \bar{H}(p)
\]

is a solution to

\[
v_t + H(x; \nabla v) = 0.
\]

What is the function \( u(x; p) \) in terms of the approximate expansion (2.6.12)? Explain why it is natural that the function \( u(x; p) \) appears when we try to approximate the solution to

\[
u_t^\varepsilon + H(x, \nabla u^\varepsilon) = 0,
\]

with an initial condition of the form \( u^\varepsilon(0, x) = \varepsilon^{-1}u_0(\varepsilon x) \).

**The proof of the Lions-Papanicolaou-Varadhan theorem**

Recall that our goal is to construct a solution to (2.6.7):

\[
H(x, \nabla u) = c, \quad x \in \mathbb{T}^n.
\]  

(2.6.18)

As we have already proved uniqueness of the constant \( c \), we only need to prove its existence, and, of course, construct the solution \( u(x) \). We will make use of the viscosity solution to the auxiliary problem

\[
H(x, \nabla u^\varepsilon) + \varepsilon u^\varepsilon = 0, \quad x \in \mathbb{T}^n,
\]  

(2.6.19)

with \( \varepsilon > 0 \). Note that the regularization parameter \( \varepsilon > 0 \) in (2.6.19) has nothing to do with the small parameter \( \varepsilon > 0 \) that we have used in the discussion of the periodic homogenization theory, where it referred to the separation of scales between the scale of variation of the initial condition and that of the periodic Hamiltonian. Unfortunately, it is common to use the notation \( \varepsilon \) in both of these settings. We hope that the reader will find it not too confusing.

We have already shown that (2.6.19) has at most one solution. Let us for the moment accept that the solution to the regularized problem (2.6.19) exists and show how one can finish the proof of Theorem 2.6.2 from here. Then, we will come back to the construction of a solution to (2.6.19). Our task is to pass to the limit \( \varepsilon \downarrow 0 \) in (2.6.19).

**Exercise 2.6.5** Show that for all \( \varepsilon > 0 \), the solution \( u^\varepsilon(x) \) of (2.6.19) satisfies

\[
- \frac{\|H(\cdot, 0)\|_{L^\infty}}{\varepsilon} \leq u^\varepsilon(x) \leq \frac{\|H(\cdot, 0)\|_{L^\infty}}{\varepsilon},
\]  

(2.6.20)

for all \( x \in \mathbb{T}^n \). Hint: use the comparison principle.

Note that the fact that \( u^\varepsilon(x) \) is of the size \( \varepsilon^{-1} \) is not a fluke of the estimate. For instance, if the function \( H(x, p) \) is bounded from below by a positive constant \( c_0 \), then the solution to (2.6.19) will clearly satisfy \( |u^\varepsilon(x)| \geq c_0/\varepsilon \) for all \( x \in \mathbb{T}^n \). Therefore, one can not expect that the solution to (2.6.19) converges as \( \varepsilon \to 0 \) to a solution to (2.6.18). One can, however, hope that the solution becomes large but its gradient stays bounded, so if we subtract the
large mean the difference will be bounded. Accordingly, we will decompose \( u^\varepsilon \) into its mean and oscillation:

\[
    u^\varepsilon(x) = \langle u^\varepsilon \rangle + v^\varepsilon(x),
\]

where

\[
    \langle u^\varepsilon \rangle = \int_{T^n} u^\varepsilon(y) dy. \tag{2.6.22}
\]

Recall that the torus \( T^n \) is normalized so that \( \text{Vol}(T^n) = 1 \). We will then show that there is a sequence \( \varepsilon_k \to 0 \) so that the limit

\[
    c = \lim_{\varepsilon_k \to 0} \varepsilon_k \langle u^{\varepsilon_k} \rangle \tag{2.6.23}
\]

exists, and \( v^{\varepsilon_k}(x) \) also converge uniformly on \( T^n \) to a limit \( u \) that satisfies (2.6.18) with \( c \) given by (2.6.23).

In order to pass to the limit \( \varepsilon \downarrow 0 \) in (2.6.19), we need a modulus of continuity estimate on \( u^\varepsilon \) (and hence \( v^\varepsilon \)) that does not depend on \( \varepsilon \in (0,1) \).

**Lemma 2.6.6** There is \( C > 0 \) independent of \( \varepsilon \) such that \( |\text{Lip } u^\varepsilon| \leq C \).

**Proof.** Again, we use the doubling of the independent variables. Fix \( x \in T^n \) and, for \( K > 0 \), consider the function

\[
    \zeta(y) = u^\varepsilon(y) - u^\varepsilon(x) - K|y - x|. \tag{2.6.24}
\]

Let \( \hat{x} \) be a maximum of \( \zeta(y) \) (the point \( \hat{x} \) depends on \( x \)). If \( \hat{x} = x \) for all \( x \in T^n \), then, as \( \zeta(x) = 0 \), we obtain

\[
    u^\varepsilon(y) - u^\varepsilon(x) \leq K|y - x|, \tag{2.6.25}
\]

for all \( x, y \in T^n \), which implies that \( u^\varepsilon \) is Lipschitz with the constant \( K \). If there exists some \( x \) such that \( \hat{x} \neq x \), then the function

\[
    \psi(y) = u^\varepsilon(x) + K|y - x|
\]

is, in a vicinity of the point \( y = \hat{x} \), an admissible test function, as a function of \( y \). Moreover, the difference

\[
    \psi(y) - u^\varepsilon(y) = -\zeta(y)
\]

attains its minimum at \( y = \hat{x} \). As \( u^\varepsilon(y) \) is a viscosity solution to (2.6.19), and

\[
    \nabla \psi(\hat{x}) = K \frac{\hat{x} - x}{|\hat{x} - x|},
\]

it follows that

\[
    H\left( \hat{x}, K \frac{\hat{x} - x}{|\hat{x} - x|} \right) + \varepsilon u^\varepsilon(\hat{x}) \leq 0. \tag{2.6.26}
\]

Since \( \varepsilon u^\varepsilon(x) \) is bounded by \( \|H(\cdot,0)\|_{L^\infty} \), as in (2.6.20), we deduce that

\[
    H\left( \hat{x}, K \frac{\hat{x} - x}{|\hat{x} - x|} \right) \leq \|H(\cdot,0)\|_{L^\infty}. \tag{2.6.27}
\]

133
On the other hand, the coercivity condition (2.6.5) implies that we can take $K$ sufficiently large, so that
\[ \|H(\cdot,0)\|_{L^\infty} < \inf_{x \in \mathbb{T}^n, |p|=K} H(x,p). \] (2.6.28)
Hence, if we take $K$ as in (2.6.28), then (2.6.27) can not hold. As a consequence, for such $K$ we must have $\hat{x} = x$ for all $x \in \mathbb{T}^n$. It follows that for such $K$ the inequality (2.6.25) holds for all $x, y \in \mathbb{T}^n$. This finishes the proof. □

To finish the proof of Theorem 2.6.2, we go back to the decomposition (2.6.21)-(2.6.22). The function $v^\varepsilon = u^\varepsilon - \langle u^\varepsilon \rangle$ satisfies
\[ H(x, \nabla v^\varepsilon) + \varepsilon \langle u^\varepsilon \rangle + \varepsilon v^\varepsilon = 0. \] (2.6.29)
As
\[ \int_{\mathbb{T}^n} v^\varepsilon(x) dx = 0, \]
and because of Lemma 2.6.6, the family $v^\varepsilon$ is both uniformly bounded in $L^\infty$ and is uniformly Lipschitz. As a consequence, it converges uniformly, up to extraction of a subsequence, to a function $v \in C(\mathbb{T}^n)$, and $\varepsilon v^\varepsilon \to 0$. The bound (2.6.20) implies that the family $\varepsilon \langle u^\varepsilon \rangle$ is bounded. We may, therefore, assume its convergence (along a subsequence) to a constant denoted by $-c$, as in (2.6.23). By the stability result in Exercise 2.5.8, we deduce that $v$ is a viscosity solution of
\[ H(x, \nabla v) = c. \] (2.6.30)
This finishes the proof of Theorem 2.6.2 except for the construction of a solution to (2.6.19).

**Existence of the solution to the auxiliary problem**

Let us now construct a solution to (2.6.19).

**Proposition 2.6.7** If $H(x,p)$ satisfies the assumptions of Theorem 2.6.2, then for all $\varepsilon > 0$ the problem
\[ H(x, \nabla u) + \varepsilon u = 0, \quad x \in \mathbb{T}^n, \] (2.6.31)
has a viscosity solution.

We will treat a solution to (2.6.31) as a fixed point of the map $S[u] = u$ defined via
\[ H(x, \nabla u) + Mu = (M - \varepsilon)v, \quad x \in \mathbb{T}^n, \] (2.6.32)
with $M > 0$ to be chosen appropriately. The point is that if $M$ is sufficiently large, we will be able to prove that this map is a contraction on $C(\mathbb{T}^n)$, hence has a fixed point. Any such fixed point is a solution to (2.6.31). Our first task is to prove the following lemma.

**Lemma 2.6.8** There exists $M_0 > 0$ so that for all $M > M_0$ and all $f \in C(\mathbb{T}^n)$ there exists a solution to
\[ H(x, \nabla u) + Mu = f, \quad x \in \mathbb{T}^n. \] (2.6.33)
This lemma shows that the map $S$ is well-defined for $M > M_0$. Its proof will use an explicit construction of the solutions via a limiting procedure that will give us sufficiently strong a priori bounds that will allow us to deduce that $S$ is a contraction.
The proof of Lemma 2.6.8

We take a function \( f \in C(\mathbb{T}^n) \), and consider a regularized problem
\[
-\delta \Delta u^{\gamma,\delta} + H(x, \nabla u^{\gamma,\delta}) + Mu^{\gamma,\delta} = f_\gamma(x), \quad x \in \mathbb{T}^n,
\] (2.6.34)
with \( \delta > 0 \) and \( \gamma > 0 \), and
\[
f_\gamma = G_\gamma * f.
\] (2.6.35)
Here, \( G_\gamma \) is a compactly supported smooth approximation of identity:
\[
G_\gamma(x) = \gamma^{-n}G\left(\frac{x}{\gamma}\right), \quad G(x) \geq 0, \quad \int_{\mathbb{R}^n} G(x)dx = 1,
\]
so that \( f_\gamma(x) \) is smooth, and \( f_\gamma \to f \) in \( C(\mathbb{T}^n) \). In particular, there exists \( K_\gamma \) that depends on \( \gamma \in (0,1) \) so that
\[
\|f_\gamma\|_{L^\infty} \leq \|f\|_{L^\infty}, \quad \|f_\gamma\|_{C^1} \leq K_\gamma \|f\|_{L^\infty}. \tag{2.6.36}
\]
It is straightforward to adapt what we have done in Section 1.5.2 for the time-dependent problem with a positive diffusion coefficient to show that (2.6.34) admits a smooth solution \( u^{\gamma,\delta} \) for each \( \gamma > 0 \) and \( \delta > 0 \). The difficulty is to pass to the limit \( \delta \downarrow 0 \), followed by \( \gamma \downarrow 0 \) to construct in the limit a viscosity solution to (2.6.33). This will require a priori bounds on \( u^{\gamma,\delta} \) summarized in the following lemma.

**Lemma 2.6.9** There exists \( M_0 > 0 \) so that if \( M > M_0 \) then the solution \( u^{\gamma,\delta} \) to (2.6.34) obeys the following gradient bound, for all \( \delta \in (0,1) \):
\[
|\nabla u^{\gamma,\delta}(x)| \leq C_\gamma (1 + \|f\|_{L^\infty}) \text{ for all } x \in \mathbb{T}^n. \tag{2.6.37}
\]
Here, the constant \( C_\gamma \) may depend on \( \gamma \in (0,1) \) but not on \( \delta \in (0,1) \). There also exists a constant \( C > 0 \) that does not depend on \( \gamma \in (0,1) \) or \( \delta \in (0,1) \) so that
\[
|u^{\gamma,\delta}(x)| \leq \frac{C}{M} (1 + \|f\|_{L^\infty}) \text{ for all } x \in \mathbb{T}^n. \tag{2.6.38}
\]

**Proof.** Let us look at the point \( x_0 \) where \( |\nabla u^{\gamma,\delta}(x)|^2 \) attains its maximum. Note that (we drop the super-scripts \( \gamma \) and \( \delta \) for the moment)
\[
\frac{\partial}{\partial x_i} (|\nabla u|^2) = 2 \frac{\partial u}{\partial x_j} \frac{\partial^2 u}{\partial x_i \partial x_j},
\]
so that, using (2.6.34), we compute
\[
\Delta(|\nabla u|^2) = 2 \sum_{i,j=1}^n \left( \frac{\partial^2 u}{\partial x_i \partial x_j} \right)^2 + 2 \sum_{j=1}^n \frac{\partial u}{\partial x_j} \frac{\partial \Delta u}{\partial x_j} = 2 \sum_{i,j=1}^n \left( \frac{\partial^2 u}{\partial x_i \partial x_j} \right)^2 + \frac{2M}{\delta} |\nabla u|^2
\]
\[
+ \frac{2}{\delta} \sum_{j=1}^n \frac{\partial u}{\partial x_j} \frac{\partial H(x, \nabla u)}{\partial x_j} + \frac{2}{\delta} \sum_{k,j=1}^n \frac{\partial u}{\partial x_j} \frac{\partial H(x, \nabla u)}{\partial p_k} \frac{\partial^2 u}{\partial x_j \partial x_k} - \frac{2}{\delta} \sum_{j=1}^n \frac{\partial u}{\partial x_j} \frac{\partial f_\gamma}{\partial x_j}
\]
\[
= 2 \sum_{i,j=1}^n \left( \frac{\partial^2 u}{\partial x_i \partial x_j} \right)^2 + \frac{2M}{\delta} |\nabla u|^2 + \frac{2}{\delta} \sum_{j=1}^n \frac{\partial u}{\partial x_j} \frac{\partial H(x, \nabla u)}{\partial x_j} + \frac{1}{\delta} \sum_{k=1}^n \frac{\partial H(x, \nabla u)}{\partial p_k} \frac{\partial |\nabla u|^2}{\partial x_k}
\]
\[
- \frac{2}{\delta} \sum_{j=1}^n \frac{\partial u}{\partial x_j} \frac{\partial f_\gamma}{\partial x_j}.
\]
Thus, at the maximum $x_0$ of $|\nabla u|^2$ we have

$$0 \geq \Delta(|\nabla u|^2)(x_0) = 2 \sum_{i,j=1}^{n} \left( \frac{\partial^2 u}{\partial x_i \partial x_j} \right)^2 + \frac{2M}{\delta} |\nabla u|^2 + \frac{2}{\delta} \sum_{j=1}^{n} \frac{\partial u}{\partial x_j} \frac{\partial H(x, \nabla u)}{\partial x_j} - \frac{2}{\delta} \sum_{j=1}^{n} \frac{\partial u}{\partial x_j} \frac{\partial f_\gamma}{\partial x_j}. \tag{2.6.39}$$

Let us recall the gradient bound (2.6.6) on $H(x,p)$:

$$|\nabla x H(x,p)| \leq K_0(1 + |p|). \tag{2.6.40}$$

We see from (2.6.39) and (2.6.40) that

$$Q = |\nabla u(x_0)| = \sup_{x \in \mathbb{T}^n} |\nabla u(x)|$$

satisfies

$$MQ^2 \leq K_0 Q(1 + Q) + Q \|f_\gamma\|_{C^1} \leq 5K_0(1 + Q^2) + K_\gamma Q \|f\|_{L^\infty}. \tag{2.6.41}$$

We used (2.6.36) above. It follows from (2.6.41) that there exist $M_0 > 0$ and $C_1$ that depend on $K_0$ but not on $\gamma \in (0, 1)$ and $C_\gamma$ that depends on $\gamma \in (0, 1)$ so that for all $M > M_0$ we have

$$Q \leq C_1 + C_\gamma \|f\|_{L^\infty}. \tag{2.6.42}$$

This proves (2.6.37).

To prove (2.6.38) we look at the point $x_M$ where $u$ attains its maximum over $\mathbb{T}^n$. At this point we have

$$Mu(x_M) = f_\gamma(x_M) + \delta \Delta u(x_M) - H(x_M, 0) \leq \|f_\gamma\|_{L^\infty} + \|H(\cdot, 0)\|_{L^\infty}, \tag{2.6.43}$$

hence

$$u(x_M) \leq \frac{C}{M}(1 + \|f\|_{L^\infty}).$$

A similar estimate holds at the minimum of $u$, proving (2.6.38). $\square$

The Lipschitz bound (2.6.37) and (2.6.38) show that if $M > M_0$, after passing to a subsequence $\delta_k \downarrow 0$, the family $u^{\gamma, \delta_k}(x)$ converges uniformly in $x \in \mathbb{T}^n$, to a function $u^{\gamma}(x)$.

**Exercise 2.6.10** Show that $u^{\gamma}(x)$ is the viscosity solution to

$$H(x, \nabla u^{\gamma}) + Mu^{\gamma} = f_\gamma(x), \quad x \in \mathbb{T}^n. \tag{2.6.44}$$

Hint: Exercise 2.5.8 and its solution should be helpful here.

The next step is to send $\gamma \to 0$.

**Exercise 2.6.11** Mimic the proof of Lemma 2.6.6 to show that $u^{\gamma}(x)$ are uniformly Lipschitz: there exists a constant $C_\gamma > 0$ that may depend on $\|f\|_{L^\infty}$ but is independent of $\gamma \in (0, 1)$ and of $M > M_0$ such that

$$|\text{Lip } u^{\gamma}| \leq C_\gamma. \tag{2.6.45}$$

Also show that

$$\|u^{\gamma}\|_{L^\infty} \leq \frac{1}{M}(\|H(\cdot, 0)\|_{L^\infty} + \|f\|_{L^\infty}). \tag{2.6.46}$$
This exercise shows that as long as \( M \geq M_0 \), the family \( u^{\gamma_k} \) converges, along as subsequence \( \gamma_k \downarrow 0 \), uniformly in \( x \in \mathbb{T}^n \), to a limit \( u(x) \in C(\mathbb{T}^n) \) that obeys the same uniform Lipschitz and \( L^\infty \)-bounds in Exercise 2.6.11. Invoking again the stability result of Exercise 2.5.8 shows that \( u(x) \) is the unique viscosity solution to

\[
H(x, \nabla u) + Mu = f(x), \quad x \in \mathbb{T}^n. \tag{2.6.47}
\]

This finishes the proof of Lemma 2.6.8. \( \square \)

The end of the proof of Proposition 2.6.7

We now explain how this construction implies the conclusion of Proposition 2.6.7. Let us take \( \varepsilon < M \), and re-write equation (2.6.31)

\[
H(x, \nabla u) + \varepsilon u = 0, \quad x \in \mathbb{T}^n. \tag{2.6.48}
\]

for which we need to find a solution, as

\[
H(x, \nabla u) + Mu = (M - \varepsilon)u, \quad x \in \mathbb{T}^n. \tag{2.6.49}
\]

As we have mentioned, we define the map \( S : C(\mathbb{T}^n) \to C(\mathbb{T}^n) \) as follows: given \( v \in C(\mathbb{T}^n) \), let \( u = S[v] \) be the unique viscosity solution to

\[
H(x, \nabla u) + Mu = (M - \varepsilon)v, \quad x \in \mathbb{T}^n. \tag{2.6.50}
\]

We claim that \( S \) is a contraction in \( C(\mathbb{T}^n) \). We have shown that \( u = S[v] \) can be constructed via the above procedure of passing to the limit \( \delta \to 0 \), followed by \( \gamma \to 0 \) in the regularized problem

\[
-\delta \Delta u^{\gamma,\delta} + H(x, \nabla u^{\gamma,\delta}) + Mu^{\gamma,\delta} = (M - \varepsilon)v, \quad x \in \mathbb{T}^n. \tag{2.6.51}
\]

Given \( v_1, v_2 \in C(\mathbb{T}^n) \), consider the corresponding solutions to the regularized problems (2.6.51):

\[
-\delta \Delta u_1^{\gamma,\delta} + H(x, \nabla u_1^{\gamma,\delta}) + Mu_1^{\gamma,\delta} = (M - \varepsilon)v_{1,\gamma}, \quad x \in \mathbb{T}^n, \tag{2.6.52}
\]

and

\[
-\delta \Delta u_2^{\gamma,\delta} + H(x, \nabla u_2^{\gamma,\delta}) + Mu_2^{\gamma,\delta} = (M - \varepsilon)v_{2,\gamma}, \quad x \in \mathbb{T}^n. \tag{2.6.53}
\]

Assume that the difference

\[
w = u_1^{\gamma,\delta} - u_2^{\gamma,\delta}
\]

attains its maximum at a point \( x_0 \). The function \( w \) satisfies

\[
-\delta \Delta w + H(x, \nabla u_1^{\gamma,\delta}) - H(x, \nabla u_2^{\gamma,\delta}) + Mw = (M - \varepsilon)(v_{1,\gamma} - v_{2,\gamma}), \quad x \in \mathbb{T}^n. \tag{2.6.54}
\]

Evaluating this at \( x = x_0 \), as \( \nabla u_1^{\gamma,\delta}(x_0) = \nabla u_2^{\gamma,\delta}(x_0) \), we see that

\[
-\delta \Delta w(x_0) + Mw(x_0) = (M - \varepsilon)(v_{1,\gamma}(x_0) - v_{2,\gamma}(x_0)), \quad x \in \mathbb{T}^n. \tag{2.6.55}
\]

As \( x_0 \) is the maximum of \( w \), we deduce that

\[
w(x_0) \leq \frac{M - \varepsilon}{M} \|v_{1,\gamma} - v_{2,\gamma}\|_{C(\mathbb{T}^n)}.\]
Using a nearly identical computation for the minimum, we conclude that
\[ \|u_1^{\gamma,\delta} - u_2^{\gamma,\delta}\|_{C(T^n)} \leq \frac{M - \varepsilon}{M}\|v_{1,\gamma} - v_{2,\gamma}\|_{C(T^n)}. \] (2.6.56)
Passing to the limit \(\delta \downarrow 0\) and \(\gamma \downarrow 0\), we obtain
\[ \|u_1 - u_2\|_{C(T^n)} \leq \frac{M - \varepsilon}{M}\|v_1 - v_2\|_{C(T^n)}, \] (2.6.57)

hence \(S\) is a contraction on \(C(T^n)\), as claimed. Thus, this map has a fixed point, which is the viscosity solution to
\[ H(x, \nabla u) + \varepsilon u = 0, \quad x \in T^n. \] (2.6.58)
This completes the proof of Proposition 2.6.7. □

### 2.6.2 Existence of the solution to the Cauchy problem

We will now construct the viscosity solution to the Cauchy problem
\[
\begin{align*}
&u_t + H(x, \nabla u) = 0, \quad t > 0, \quad x \in T^n, \\
&u(0, x) = u_0(x), \quad x \in T^n,
\end{align*}
\] (2.6.59)
with a continuous initial condition \(u_0(x)\). Recall that Exercise 2.5.13 implies the uniqueness of the solution with a given initial condition, so we do not need to address that issue. We make the same assumptions as in Theorem 2.6.2: there exists \(C_L > 0\) so that
\[ |H(x, p_1) - H(x, p_2)| \leq C_L|p_1 - p_2|, \quad \text{for all } x, p_1, p_2 \in \mathbb{R}^n, \] (2.6.60)
and
\[ \lim_{|p| \to +\infty} H(x, p) = +\infty, \quad \text{uniformly in } x \in T^n. \] (2.6.61)
We will again assume the gradient bound (2.6.6):
\[ |\nabla_x H(x, p)| \leq K_0(1 + |p|), \quad \text{for all } x \in T^n, \text{ and } p \in \mathbb{R}^n. \] (2.6.62)

**Theorem 2.6.12** The Cauchy problem (2.6.59) has a unique viscosity solution \(u(t, x)\). Moreover, the weak contraction property holds: if \(u(t, x)\) and \(v(t, x)\) are two solutions to (2.6.59) with the corresponding initial conditions \(u_0 \in C(T^n)\) and \(v_0 \in C(T^n)\), then
\[ \|u(t, \cdot) - v(t, \cdot)\|_{L^\infty} \leq \|u_0 - v_0\|_{L^\infty}. \] (2.6.63)

The weak contraction property is recorded here simply for the sake of completeness: we have seen in Exercise 2.5.14 that it follows immediately from the comparison principle. Therefore, we will focus on the existence of the solutions.

An important consequence of the weak contraction principle is that we may restrict ourselves to initial conditions that are smooth. Indeed, suppose that we managed to prove the theorem for smooth initial conditions, and consider \(u_0 \in C(T^n)\). Let \(u_0^{(k)}\) be a sequence of
smooth functions converging to \( u_0 \) in \( C(\mathbb{T}^n) \) as \( k \to +\infty \), and \( u^{(k)}(t, x) \) be the corresponding sequence of solutions to (2.6.59), with the initial conditions \( u_0^{(k)} \). It follows from the weak contraction principle that

\[
\|u^{(k)} - u^{(m)}\|_{L^\infty(\mathbb{R}_+ \times \mathbb{T}^n)} \leq \|u_0^{(k)} - u_0^{(m)}\|_{L^\infty},
\]

ensuring that \( u^{(k)} \) is a uniformly Cauchy sequence on \( C([0, +\infty) \times \mathbb{T}^n) \). Hence, it converges uniformly to a continuous function \( u \in C(\mathbb{R}_+ \times \mathbb{T}^n) \). The stability result in Exercise 2.5.8 implies that \( u \) is a solution to the Cauchy problem (2.6.59) with the initial condition \( u_0(x) \).

We are now left with the actual construction of a solution to (2.6.59), with the assumption that \( u_0 \) is smooth. We are going to use the most pedestrian way to do it: a time discretization. Take a family of time steps \( \Delta t \to 0 \). For a fixed \( \Delta t > 0 \), consider the sequence \( u^n_{\Delta t}(x) \) defined by setting \( u^0(x) := u_0(x) \) and the recursion relation:

\[
\frac{u^{n+1}_{\Delta t} - u^n_{\Delta t}}{\Delta t} + H(x, \nabla u^{n+1}_{\Delta t}) = 0, \quad x \in \mathbb{T}^n,
\]

that is an implicit time discretization of (2.6.59). Given \( u^n_{\Delta t}(x) \), we look at (2.6.64) as a time-independent Hamilton-Jacobi equation

\[
H(x, \nabla u^{n+1}_{\Delta t}) + \frac{1}{\Delta t} u^{n+1}_{\Delta t} = \frac{1}{\Delta t} u^n_{\Delta t}, \quad x \in \mathbb{T}^n.
\]  

(2.6.65)

It is of the type, for which Proposition 2.6.7 guarantees existence of a unique continuous solution \( u^{n+1}_{\Delta t} \), as long as \( u^n_{\Delta t} \) is continuous. This produces the sequence \( u^n_{\Delta t}(x) \), for \( n \geq 0 \). An approximate solution \( u_{\Delta t} \) to the Cauchy problem (2.6.59) is then constructed by interpolating linearly between the times \( n \Delta t \) and \( (n + 1) \Delta t \):

\[
u_{\Delta t}(t, x) = u^n_{\Delta t}(x) + \frac{t - n \Delta t}{\Delta t} (u^{n+1}_{\Delta t}(x) - u^n_{\Delta t}(x)), \quad t \in [n \Delta t, (n + 1) \Delta t].
\]

(2.6.66)

The help provided by the smoothness assumption on \( u_0 \) manifests itself in the next proposition.

**Proposition 2.6.13** There is \( C > 0 \), depending on \( \|u_0\|_{L^\infty} \) and \( \text{Lip}(u_0) \) but not on \( \Delta t \in (0, 1) \), such that the function \( u_{\Delta t}(t, x) \) is uniformly Lipschitz continuous in \( t \) and \( x \) on \([0, +\infty) \times \mathbb{T}^n\), and the Lipschitz constant \( \text{Lip}(u_{\Delta t}) \) of \( u_{\Delta t} \) both in \( t \) and \( x \), over the set \([0, +\infty) \times \mathbb{T}^n\), satisfies

\[
\text{Lip}(u_{\Delta t}) \leq C.
\]

(2.6.67)

This ensures that there exists a sequence \( \Delta t_n \to 0 \), such that the corresponding sequence \( u_{\Delta t_n} \) converges as \( n \to \infty \) to a Lipschitz function \( u(t, x) \) with the Lipschitz constant \( \text{Lip}(u) \leq C \). The next step will be to prove

**Proposition 2.6.14** The function \( u(t, x) \) is a viscosity solution to the Cauchy problem (2.6.59):

\[
u_t + H(x, \nabla u) = 0, \quad t > 0, \quad x \in \mathbb{T}^n,
\]

\[
u(0, x) = u_0(x), \quad x \in \mathbb{T}^n.
\]

(2.6.68)
Proof. Let us prove this claim first, assuming the conclusion of Proposition 2.6.13. Note that the initial condition \( u(0, x) = u_0(x) \) is satisfied by construction, so we only need to check that \( u \) is a viscosity solution to (2.6.68). We will only prove that \( u \) is a super-solution, the sub-solution property of \( u \) can be proved identically. Let \( \varphi(t, x) \) be a \( C^1 \)-test function and \((t_0, x_0)\) be a minimum point for the difference \( u - \varphi \). As we have seen in the hint to Exercise 2.5.8, we may assume, possibly after subtracting a quadratic polynomial in \( t \) and \( x \) from the function \( \varphi \), that the minimum is strict. Consider the linearly interpolated time discretization \( \varphi_{\Delta t} \) of \( \varphi \): set \( \varphi^n(x) = \varphi(n\Delta t, x) \), for \( n \geq 0 \), and

\[ \varphi_{\Delta t}(t, x) = \varphi^n(x) + \frac{t - n\Delta t}{\Delta t} (\varphi^{n+1}(x) - \varphi^n(x)), \quad t \in [n\Delta t, (n+1)\Delta t). \]

Note a slight abuse of notation: the function \( \varphi_{\Delta t} \) is a linear interpolation of the function \( \varphi \), while \( u_{\Delta t} \) is not the linear interpolation of the function \( u \) but rather the linear interpolation of the solution to the time-discretized problem (2.6.64), with the time step \( \Delta t \). Nevertheless, as the minimum \((t_0, x_0)\) of \( u - \varphi \) is strict, and \( u_{\Delta t} \) converges to \( u \) uniformly, for \( \Delta t \) sufficiently small, there exists a minimum point \((t_{\Delta t}, x_{\Delta t})\) for \( u_{\Delta t} - \varphi_{\Delta t} \), such that

\[ \lim_{\Delta t \to 0} (t_{\Delta t}, x_{\Delta t}) = (t_0, x_0). \]

In addition, because both \((t_{\Delta t}, x_{\Delta t})\) are piecewise linear in \( t \), we have \( t_{\Delta t} = (n + 1)\Delta t \) for some \( n \geq 0 \). Then we have, again, because \((t_{\Delta t}, x_{\Delta t})\) is a minimum for \( u_{\Delta t} - \varphi_{\Delta t} \):

\[ \frac{u^{n+1}_{\Delta t}(x_{\Delta t}) - u^n_{\Delta t}(x_{\Delta t})}{\Delta t} = \partial_t^- u_{\Delta t}((n+1)\Delta t, x_{\Delta t}) \leq \partial_t^- \varphi_{\Delta t}((n+1)\Delta t, x_{\Delta t}) = \partial_t \varphi(t_0, x_0) + o(1), \]

as \( \Delta t \to 0 \). We also have, in the vicinity of \((t_0, x_0)\):

\[ \varphi(t, x) - \varphi_{\Delta t}(t, x) = O(\Delta t^2), \quad \partial_t \varphi(t, x) - \partial_t \varphi_{\Delta t}(t, x) = O(\Delta t), \quad \text{as} \quad \Delta t \to 0, \]

with the slight catch here that we have to speak of the left and right derivatives of \( \varphi_{\Delta t} \) at the discrete times \( n\Delta t \). On the other hand, the point \( x_{\Delta t} \) is a minimum of

\[ u^{n+1}_{\Delta t}(x) - \varphi_{\Delta t}((n + 1)\Delta t, x) \]

in the \( x \)-variable. Since \( u^{n+1}_{\Delta t} \) is a viscosity solution to (2.6.64), we have

\[ \frac{u^{n+1}_{\Delta t}(x_{\Delta t}) - u^n_{\Delta t}(x_{\Delta t})}{\Delta t} \geq -H(x_{\Delta t}, \nabla \varphi_{\Delta t}((n + 1)\Delta t, x_{\Delta t})) = -H(x_0, \nabla \varphi(t_0, x_0)) + o(1), \]

as \( \Delta t \to 0 \). Putting together (2.6.69)-(2.6.71) and sending \( \Delta t \) to 0, we obtain

\[ \partial_t \varphi(t_0, x_0) + H(x_0, \nabla \varphi(t_0, x_0)) \geq 0, \]

hence \( u \) is a super-solution to (2.6.68). This proves Proposition 2.6.14. □
Proof of Proposition 2.6.13

The reason behind this proposition is quite simple: if $u$ is a smooth solution to

$$u_t + H(x, \nabla u) = 0,$$  \hspace{1cm} (2.6.72)

then the function $v(t, x) = u_t(t, x)$ solves

$$v_t + \nabla_p H(x, \nabla u) \cdot \nabla v = 0,$$  \hspace{1cm} (2.6.73)

with the initial condition $v(0, x) = -H(x, \nabla u_0(x))$. It follows from the maximum principle, or the method of characteristics for smooth solutions, that

$$\|v(t, \cdot)\|_{L^\infty} \leq \|H(\cdot, \nabla u_0(\cdot))\|_{L^\infty}. $$  \hspace{1cm} (2.6.74)

Moreover, (2.6.72) and (2.6.74) together with the coercivity of $H(x, p)$ yield the uniform boundedness of $\nabla u$. The proof of the proposition consists in making this idea rigorous.

Let us recall that $u^n_{\Delta t}$ is the solution to the recursive equation (2.6.64)

$$\frac{u^{n+1}_{\Delta t} - u^n_{\Delta t}}{\Delta t} + H(x, \nabla u^{n+1}_{\Delta t}) = 0, \hspace{1cm} x \in \mathbb{T}^n, $$  \hspace{1cm} (2.6.75)

interpolated between the times of the form $n\Delta t$ as in (2.6.66):

$$u_{\Delta t}(t, x) = u^n_{\Delta t}(x) + \frac{t - n\Delta t}{\Delta t}(u^{n+1}_{\Delta t}(x) - u^n_{\Delta t}(x)), \hspace{1cm} t \in [n\Delta t, (n + 1)\Delta t]. $$  \hspace{1cm} (2.6.76)

The viscosity solution $u^{n+1}_{\Delta t}$ to (2.6.75) can be constructed using the by now familiar idea of a diffusive regularization:

$$-\delta \Delta u^{n+1}_{\Delta t} + H(x, \nabla u^{n+1}_{\Delta t}) + \frac{u^{n+1}_{\Delta t} - u^n_{\Delta t}}{\Delta t} = 0, \hspace{1cm} x \in \mathbb{T}^n, $$  \hspace{1cm} (2.6.77)

with $\delta > 0$, and then sending $\delta \downarrow 0$. As we have assumed that $u_0(x)$ is smooth, all $u^n_{\Delta t}(x)$ are also smooth, for all $\delta > 0$.

**Exercise 2.6.15** Show that

$$\|u^{n+1}_{\Delta t}\|_{L^\infty} \leq \|u^n_{\Delta t}\|_{L^\infty} + (\Delta t)\|H(\cdot, 0)\|_{L^\infty}. $$  \hspace{1cm} (2.6.78)

Hint: look at the maximum $x_0$ of the smooth function $u^{n+1}_{\Delta t}$ over $\mathbb{T}^n$.

**Exercise 2.6.16** Use the argument in the proof of Lemma 2.6.9 and Exercise 2.6.15 to show that there exists a constant $C_{n, \Delta t}$ that may depend on $n$ and $\Delta t$ but not on $\delta > 0$, so that

$$\|\nabla u^{n, \delta}_{\Delta t}\|_{L^\infty} \leq C_{n, \Delta t}. $$  \hspace{1cm} (2.6.79)
The bound (2.6.79) is quite poor as we did not track the dependence of \( C_{n,\Delta t} \) on \( n \) or \( \Delta t \), but we have extra help. The differential quotient

\[
v^{n,\delta}_{\Delta t} = \frac{u^{n+1,\delta}_{\Delta t} - u^{n,\delta}_{\Delta t}}{\Delta t}
\]

satisfies

\[
- \delta \Delta v^{n+1,\delta}_{\Delta t} + \frac{v^{n+1,\delta}_{\Delta t}}{\Delta t} + \frac{1}{\Delta t} \left( H(x, \nabla u^{n+1,\delta}_{\Delta t}) - H(x, \nabla u^{n,\delta}_{\Delta t}) \right) = \frac{v^{n,\delta}_{\Delta t}}{\Delta t}
\]

(2.6.80)

for all \( n \geq 0 \). At the maximum \( x_M \) and minimum \( x_m \) of the smooth function \( v^{n,\delta}_{\Delta t} \) we have

\[
\nabla u^{n+1,\delta}_{\Delta t}(x_M) = \nabla u^{n,\delta}_{\Delta t}(x_M), \quad \nabla u^{n+1,\delta}_{\Delta t}(x_m) = \nabla u^{n,\delta}_{\Delta t}(x_m).
\]

Using this in (2.6.80) we obtain

\[
\|v^{n+1,\delta}_{\Delta t}\|_{L^\infty} \leq \|v^{n,\delta}_{\Delta t}\|_{L^\infty} \leq \cdots \leq \|v^0,\delta\|_{L^\infty}.
\]

(2.6.81)

For the last term in the right side we observe that

\[
v^{0,\delta}_{\Delta t} = \frac{u^{1,\delta}_{\Delta t} - u_0}{\Delta t}
\]

satisfies, instead of (2.6.80), the equation

\[
- \delta \Delta v^{0,\delta}_{\Delta t} + \frac{v^{0,\delta}_{\Delta t}}{\Delta t} + \frac{1}{\Delta t} H(x, \nabla u^{1,\delta}_{\Delta t}) = \frac{\delta}{\Delta t} \Delta u_0.
\]

(2.6.82)

Again, the maximum principle implies

\[
\|v^{0,\delta}_{\Delta t}\|_{L^\infty} \leq \|H(\cdot, \nabla u_0)\|_{L^\infty} + \delta \|\Delta u_0\|_{L^\infty}.
\]

(2.6.83)

Using this in (2.6.81), we conclude that

\[
\|v^{n,\delta}_{\Delta t}\|_{L^\infty} \leq \|H(\cdot, \nabla u_0)\|_{L^\infty} + \delta \|\Delta u_0\|_{L^\infty},
\]

(2.6.84)

for all \( n \geq 0 \). This bound is the reason why we have assumed that \( u_0 \) is smooth.

We may now pass to the limit \( \delta \to 0 \) in (2.6.84) and recall the convergence of \( u^{n,\delta}_{\Delta t} \) to \( u^n_{\Delta t} \), to conclude that

\[
v^{n,\delta}_{\Delta t} \to v^n_{\Delta t} := \frac{u^{n+1}_{\Delta t} - u^n_{\Delta t}}{\Delta t} \quad \text{as} \quad \delta \downarrow 0.
\]

(2.6.85)

Combining this with the uniform bound (2.6.84), we conclude that

\[
\left\| \frac{u^{n+1}_{\Delta t} - u^n_{\Delta t}}{\Delta t} \right\|_{L^\infty} \leq \|H(\cdot, \nabla u_0)\|_{L^\infty},
\]

(2.6.86)

which is a uniform Lipschitz bound on \( u_{\Delta t} \) in the \( t \)-variable that we need. The reader should compare it to the bound (2.6.74) that we have obtained easily for smooth solutions.
The Lipschitz bound for $u^n_{\Delta t}$ in the $x$-variable follows easily. Recall that the functions $u^n_{\Delta t}$ satisfy (2.6.64):

$$H(x, \nabla u^{n+1}_{\Delta t}) + \frac{1}{\Delta t} u^{n+1}_{\Delta t} = \frac{1}{\Delta t} u^n_{\Delta t}, \quad x \in \mathbb{T}^n. \quad (2.6.87)$$

We know from Exercise 2.6.16 that $u^n_{\Delta t}$ are Lipschitz – even though we do not know if they have a Lipschitz constant that does not depend on $n$ or $\Delta t$. However, this already tells us that $u^n_{\Delta t}$ satisfy (2.6.87) almost everywhere. We write this equation in the form

$$H(x, \nabla u^{n+1}_{\Delta t}) = -v^n_{\Delta t}(x), \quad x \in \mathbb{T}^n. \quad (2.6.88)$$

The uniform bound on $v^n_{\Delta t}$ in (2.6.86) together with the coercivity of $H(x, p)$ imply that there exists a constant $K > 0$ that does not depend on $n$ or $\Delta t$ so that

$$\|\nabla u^{n+1}_{\Delta t}\|_{L^\infty} \leq K. \quad (2.6.89)$$

This finishes the proof of Proposition 2.6.13. □

Exercise 2.6.17 Prove the following elementary fact that we used in the very last step in the above proof: if $u(x)$ is a Lipschitz function then $\text{Lip}(u) = \|\nabla u\|_{L^\infty}$.

Exercise 2.6.18 (Hamiltonians that are coercive in $u$). So far, we have been remarkably silent about Hamilton-Jacobi equations of the form

$$u_t + H(x, u, \nabla u) = 0, \quad t > 0, x \in \mathbb{T}^n, \quad (2.6.90)$$

with the Hamiltonian that depends also on the function $u$ itself. There is one case when the above theory can be developed without any real input of new ideas: assume that $H(x, u, p)$ is non-decreasing in $u$, and that there exists $C_0 > 0$ so that for all $R > 0$, there exists $\delta_{1,2}(R)$ such that

$$0 < \delta_1(R) \leq \delta_2(R) < C_0,$$

and, for all $u \in [-R, R]$, we have

$$\delta_1(R)(|p| - 1) \leq H(x, u, p) \leq \delta_2(R)(|p| + 1) \text{ for all } |u| \leq R, x \in \mathbb{T}^n \text{ and } p \in \mathbb{R}^n.$$

Prove a well-posedness theorem analogous to Theorem 2.6.12. How far can one stretch the assumptions on $H(x, u, p)$? Hint: coercivity is really something one has to assume, one way or another.

2.7 When the Hamiltonian is strictly convex: the Lagrangian theory

Let us recall that in Section 2.4 we considered the Cauchy problem

$$u_t + \frac{1}{2} |\nabla u|^2 - R(x) = 0, \quad (2.7.1)$$

143
with an initial condition \( u(0, x) = u_0(x) \). We have shown that when both \( R(x) \) and \( u_0(x) \) are convex, this problem has a smooth solution given by the (at first sight) strange looking expression (2.4.19)

\[
    u(t, x) = \inf_{\gamma(t) = x} \left( u_0(\gamma(0)) + \int_0^t \left( \frac{|\gamma'(s)|^2}{2} + R(\gamma(s)) \right) ds \right).
\]

(2.7.2)

Moreover, this expression is well-defined even if the boundary value problem for the characteristic curves may be not well-posed. Hence, a natural idea is to generalize this formula to other Hamiltonians and take this generalization as the definition of a solution. On the other hand, we already have the notion of a viscosity solution, so an issue is if these objects agree. In this section, we investigate when the variational approach is possible and whether the solution you construct in this way is, indeed, a viscosity solution. We also discuss how the strict convexity of the Hamiltonian gives an improved regularity of the solution.

### 2.7.1 The Lax-Oleinik formula and viscosity solutions

In the construction of the viscosity solutions, we assumed very little about the Hamiltonian \( H \): all we really needed was coercivity and continuity. The other regularity assumptions we have made are mostly of the technical nature and can be avoided. From now on, we will adopt an even stronger technical assumption that \( H(x, p) \) is \( C^\infty(T^n \times \mathbb{R}^n) \) smooth but more crucially we will assume that \( H(x, p) \) is uniformly strictly convex in its second variable: there exists \( \alpha > 0 \) so that

\[
    D^2 p H(x, p) \geq \alpha I, \quad [D^2 p H(x, p)]_{ij} = \frac{\partial^2 H(x, p)}{\partial p_i \partial p_j},
\]

(2.7.3)

in the sense of quadratic forms, for all \( x \in T^n \) and \( p \in \mathbb{R}^n \). Unlike the regularity assumptions, the convexity of \( H(x, p) \) in \( p \) is essential not only for this section, but also for many results on the Hamilton-Jacobi equations.

**Exercise 2.7.1** The reader may be naturally concerned that in the construction of the viscosity solutions we have assumed that \( H(x, p) \) is uniformly Lipschitz:

\[
    |H(x, p_1) - H(x, p_2)| \leq C_L |p_1 - p_2| \quad \text{for all } x \in T^n \text{ and } p_1, p_2 \in \mathbb{R}^n,
\]

(2.7.4)

and differentiable in \( x \):

\[
    |\nabla_x H(x, p)| \leq C_0 (1 + |p|) \quad \text{for all } x \in T^n \text{ and } p \in \mathbb{R}^n,
\]

(2.7.5)

These assumptions are, of course, incompatible with the strict convexity assumption on \( H(x, p) \) in (2.7.3). Go through the proofs of existence and uniqueness of the viscosity solutions and show that the coercivity assumption

\[
    \lim_{|p| \to +\infty} H(x, p) = +\infty
\]

(2.7.6)

together with the assumption that (2.7.4) and (2.7.5) hold locally in \( p \), in the sense that for every compact set \( K \subset \mathbb{R}^n \) there exist two constants \( C_L(K) \) and \( C_0(K) \) such that

\[
    |H(x, p_1) - H(x, p_2)| \leq C_L |p_1 - p_2| \quad \text{for all } x \in T^n \text{ and } p_1, p_2 \in K,
\]

(2.7.7)

\[
    |\nabla_x H(x, p)| \leq C_0 (1 + |p|) \quad \text{for all } x \in T^n \text{ and } p \in K,
\]
are sufficient to prove existence and uniqueness of the viscosity solutions both in the Lions-
Papanicolaou-Varadhan Theorem 2.6.2 and in Theorem 2.6.12 for the solutions to the Cauchy
problem.

The Legendre transform and extremal paths

Recall that in Section 2.2 we have informally argued as follows: given a path \( \gamma(s), \ t \leq s \leq T \),
with the starting point \( \gamma(t) = x \), we can define its cost as

\[
C(\gamma)(t) = \int_t^T \tilde{L}(\dot{\gamma}(s)) ds + f(x(T)).
\]

Here, the function \( \tilde{L}(v) \) represents the running cost, and the function \( f(x) \) is the terminal
cost. The corresponding value function is

\[
\tilde{u}(t,x) = \inf_{\gamma: \gamma(t)=x} C(\gamma)(t), \tag{2.7.9}
\]

with the infimum taken over all curves \( \gamma \in C^1 \) such that \( \gamma(t) = x \). We have shown, albeit
very informally, that \( \tilde{u}(t,x) \) satisfies the Hamilton-Jacobi equation

\[
\tilde{u}_t + \tilde{H}(\nabla \tilde{u}) = 0, \tag{2.7.10}
\]

with the terminal condition \( u(T,x) = f(x) \). The Hamiltonian \( \tilde{H}(p) \) is given in terms of the
running cost \( \tilde{L}(v) \) by (2.2.9):

\[
\tilde{H}(p) = \inf_{v \in A} [\tilde{L}(v) + v \cdot p]. \tag{2.7.11}
\]

It is convenient to reverse the direction of time and set

\[
u(t,x) = \tilde{u}(T-t,x). \tag{2.7.12}
\]

This function satisfies the forward Cauchy problem

\[
u_t + H(\nabla \nu) = 0, \tag{2.7.13}
\]

with the initial condition \( u(0,x) = f(x) \) and the Hamiltonian given by

\[
H(p) = -\tilde{H}(p) = -\inf_{v \in \mathbb{R}^n} [\tilde{L}(v) + v \cdot p] = \sup_{v \in \mathbb{R}^n} [-p \cdot v - \tilde{L}(v)] = \sup_{v \in \mathbb{R}^n} [p \cdot v - L(v)], \tag{2.7.14}
\]

with the time-reversed cost function

\[
L(v) = \tilde{L}(-v). \tag{2.7.15}
\]

The natural questions are, first, if the above construction, using the minimizer in (2.7.9),
indeed, produces a solution to the initial value problem for (2.7.13) – so far, our arguments
were rather informal, and, second, how it is related to the notion of the viscosity solution.

This bring us to the terminology of the Legendre transforms. One of the standard refer-
ences for the basic properties of the Legendre transform is [123], where an interested reader
may find much more information on this beautiful subject. Given a function \( L(v) \), known as
the Lagrangian, we define its Legendre transform as in (2.7.14)

\[
H(p) = \sup_{v \in \mathbb{R}^n} (p \cdot v - L(v)). \tag{2.7.16}
\]
Exercise 2.7.2 Show the function $H(p)$ defined by (2.7.16) is convex. Hint: use the fact that $H(p)$ is the supremum of a family of linear functions in $p$.

This shows that if we hope to connect the Hamilton-Jacobi equations to the above optimal control problem, this can only be done for convex Hamiltonians. Hence, our assumption (2.7.3) that the Hamiltonian $H(x, p)$ is convex in $p$.

If the function $L(v)$ is smooth and strictly convex, then, for a given $p \in \mathbb{R}^n$, the maximizer $\bar{v}(p)$ in (2.7.16) is explicit: it is the unique solution to

$$p = \nabla L(\bar{v}). \quad (2.7.17)$$

Exercise 2.7.3 Show that if $L(v)$ is strictly convex, and $H(p)$ is its Legendre transform given by (2.7.16), then we have the duality

$$L(v) = \sup_{p \in \mathbb{R}^n} (p \cdot v - H(p)),$$

so that the Lagrangian $L$ is the Legendre transform of the Hamiltonian $H$. Hint: this is easier to verify if $L(v)$ is smooth, in addition to being convex.

As a consequence, if a function $H(p)$ is strictly convex, then we can define the Lagrangian $L$ as the Legendre transform of $H$. If the Hamiltonian $H(x, p)$ depends, in addition, on a variable $x \in \mathbb{T}^n$ as a parameter, then the Lagrangian $L(x, v)$ is defined as the Legendre transform of $H(x, p)$ in the variable $p$:

$$L(x, v) = \sup_{p \in \mathbb{R}^n} (p \cdot v - H(x, p)), \quad (2.7.18)$$

with the dual relation

$$H(x, p) = \sup_{v \in \mathbb{R}^n} (p \cdot v - L(x, v)). \quad (2.7.19)$$

We usually refer to $x$ as the spatial variable, and to $p$ as the momentum variable.

Exercise 2.7.4 Compute the Lagrangian $L(x, v)$ for the classical mechanics Hamiltonian

$$H(x, p) = \frac{|p|^2}{2m} + U(x),$$

with a given $m > 0$. Why is it called the classical mechanics Hamiltonian? What is the meaning of the two terms in its definition? Hint: consider the characteristic curves for this Hamiltonian.

Exercise 2.7.5 Consider a sequence of smooth strictly convex Hamiltonians $H_\varepsilon(p)$ that converges locally uniformly, as $\varepsilon \to 0$, to $H(p) = |p|$. What happens to the corresponding Lagrangians $L_\varepsilon(v)$ as $\varepsilon \to 0$?
In the context of the forward in time Hamilton-Jacobi equations, with the Hamiltonian that depends on the spatial variable as well, the variational problem \((2.7.8)-(2.7.9)\) is defined as follows. For \(t > 0\), and two points \(x, y \in \mathbb{T}^n\), we define the function
\[
h_t(y, x) = \inf_{\gamma(0)=y, \gamma(t)=x} \int_0^t L(\gamma(s), \dot{\gamma}(s)) \, ds.
\] (2.7.20)
Here, the infimum is taken over all paths \(\gamma\) on \(\mathbb{T}^n\), that are piecewise \(C^1[0, t]\), and \(L(x, v)\) is the Lagrangian given by \((2.7.18)\). The quantity
\[
A(\gamma) = \int_0^t L(\gamma(s), \dot{\gamma}(s)) \, ds
\]
is usually referred to as the Lagrangian action, or simply the action. This is a classical minimization problem, which admits the following result (Tonelli’s theorem).

**Proposition 2.7.6** Given any \((t, x, y) \in \mathbb{R}^*_+ \times \mathbb{T}^n \times \mathbb{T}^n\), there exists at least one minimizing path \(\gamma(s) \in C^2([0, t]; \mathbb{T}^n)\), such that
\[
h_t(y, x) = \int_0^t L(\gamma(s), \dot{\gamma}(s)) \, ds.
\]
Moreover there is \(C(t, |x - y|) > 0\) such that
\[
\|\dot{\gamma}\|_{L^\infty([0, t])} + \|\ddot{\gamma}\|_{L^\infty([0, t])} \leq C(t, |x - y|).
\] (2.7.21)
The function \(C\) tends to \(+\infty\) as \(t \to 0 -\) keeping \(|x - y|\) fixed. The function \(\gamma(s)\) solves the Euler-Lagrange equation
\[
\frac{d}{ds} \nabla_v L(\gamma(s), \dot{\gamma}(s)) - \nabla_x L(\gamma(s), \dot{\gamma}(s)) = 0.
\] (2.7.22)

We leave the proof as an exercise but give a hint for the proof. Think of how we proceeded in Section 2.4.2 as blueprint. Consider a minimizing sequence \(\gamma_n\). First, use the strict convexity of \(L\) to obtain the \(H^1\)-estimates for \(\gamma_n\), thus ensuring compactness in the space of continuous paths and weak convergence to \(\gamma \in H^1([0, t])\) with fixed ends. Next, show that the convexity of \(L\) implies that \(\gamma\) is, indeed, a minimizer. Finally, derive the Euler-Lagrange equation and show that \(\gamma\) is actually \(C^\infty\). Such a curve \(\gamma\) is called an extremal.

**The Lax-Oleinik semigroup and viscosity solutions**

We now relate the solutions to the Cauchy problem for the Hamilton-Jacobi equations
\[
\begin{align*}
  u_t + H(x, \nabla u) &= 0, \ t > 0, \ x \in \mathbb{R}^n, \\
  u(0, x) &= u_0(x),
\end{align*}
\] (2.7.23)
with a strictly convex Hamiltonian \(H(x, p)\), to the minimization problem. We let \(L(x, v)\) be the Legendre transform of \(H(x, p)\), and define the corresponding function \(h_t(y, x)\). Given the initial condition \(u_0 \in C(\mathbb{T}^n)\), we define the function
\[
u(t, x) = \mathcal{T}(t)u_0(x) = \inf_{y \in \mathbb{T}^n} (u_0(y) + h_t(y, x)).
\] (2.7.24)
The following exercise gives the dynamic programming principle, the continuous in time analog of relation \((2.2.5)\) in the time-discrete case we have considered in Section 2.2.
Exercise 2.7.7 Show that the infimum in (2.7.24) is attained. Also show that $(\mathcal{T}(t))_{t>0}$ is a semi-group: one has

$$\mathcal{T}(t+s)u_0 = \mathcal{T}(t)\mathcal{T}(s), \quad \text{for all } t \geq 0 \text{ and } s \geq 0,$$

that is,

$$u(t, x) = \inf_{y \in \mathbb{T}^n} (u(s, y) + h_{t-s}(y, x)),$$

(2.7.25)

for all $0 \leq s \leq t$, and $\mathcal{T}(0) = I$.

This semigroup is sometimes referred to as the Lax-Oleinik semigroup. Here is its link to the Hamilton-Jacobi equations and the viscosity solutions.

Theorem 2.7.8 Given $u_0 \in C(\mathbb{T}^n)$, the function $u(t, x) := \mathcal{T}(t)u_0(x)$ is the unique viscosity solution to the Cauchy problem

$$u_t + H(x, \nabla u) = 0,$$

$$u(0, x) = u_0(x).$$

(2.7.26)

Proof. The initial condition for $u(t, x)$ holds essentially automatically so we only need to check that $u$ is the viscosity solution. We first show the super-solution property: take $t_0 > 0$ and $x_0 \in \mathbb{T}^n$ and let $\phi$ be a test function such that $(t_0, x_0)$ is a minimum for $u - \phi$. As usual, without loss of generality, we may assume that $u(t_0, x_0) = \phi(t_0, x_0)$. Consider the minimizing point $y_0$ such that

$$u(t_0, x_0) = u_0(y_0) + h_{t_0}(y_0, x_0).$$

Let also $\gamma$ be an extremal of the action between the times $t = 0$ and $t = t_0$, going from $y_0$ to $x_0$: $\gamma(0) = y_0, \gamma(t_0) = x_0$. We have, for all $0 \leq t \leq t_0$:

$$\phi(t, \gamma(t)) \leq u(t, \gamma(t)) \leq u_0(y_0) + \int_0^t L(\gamma(s), \dot{\gamma}(s)) \, ds.$$

(2.7.27)

The first inequality above holds because $(t_0, x_0)$ is a minimum of $u - \phi$ and $u(t_0, x_0) = \phi(t_0, x_0)$, and the second follows from the definition of $u(t, \gamma(t))$ in terms of the Lax-Oleinik semigroup. Note that at $t = t_0$ both inequalities in (2.7.27) become equalities: the first one because $u(t_0, x_0) = \phi(t_0, x_0)$, and the second because the curve $\gamma$ is a minimizer for $u(t_0, x_0)$. This implies

$$\frac{d}{dt} \left( u_0(y_0) + \int_0^t L(\gamma(s), \dot{\gamma}(s)) \, ds - \phi(t, \gamma(t)) \right) \bigg|_{t=t_0} \leq 0,$$

(2.7.28)

or, in other words

$$\phi_t(t_0, x_0) + \dot{\gamma}(t_0) \cdot \nabla \phi(t_0, x_0) - L(\gamma(t_0), \dot{\gamma}(t_0)) \geq 0.$$

(2.7.29)

Using the test point $v = \dot{\gamma}(t_0)$ in the definition (2.7.19) of $H(x, p)$, we then obtain

$$\phi_t(t_0, x_0) + H(x_0, \nabla \phi(t_0, x_0)) \geq 0.$$

(2.7.30)

Hence, $u(t, x)$ is a viscosity super-solution to (2.7.26).

To show the sub-solution property, consider a test function $\phi(t, x)$, as well as $t_0 > 0$ and $x_0 \in \mathbb{T}^n$, such that the difference $u - \phi$ attains its maximum at $(t_0, x_0)$, and assume,
once again, that \( u(t_0, x_0) = \phi(t_0, x_0) \). Using the semigroup property \((2.7.25)\), we obtain, for all \( t \leq t_0 \) and any curve \( \gamma(t) \) such that \( \gamma(t_0) = x_0 \):

\[
u(t_0, x_0) \leq u(t, \gamma(t)) + h_{t_0-t}(\gamma(t), x_0) \leq \phi(t, \gamma(t)) + h_{t_0-t}(\gamma(t), x_0).
\]

\[(2.7.31)\]

Given \( v \in \mathbb{R}^n \), we take the test curve

\[
\gamma(s) = x_0 - (t_0 - s)v
\]

in \((2.7.31)\), so that

\[
\gamma(t) = x_0 - (t_0 - t)v.
\]

Note that the curve

\[
\gamma_1(s) = x_0 - (t_0 - t)v + sv,
\]

can be used as a test curve in the definition of \( h_{t_0-t}(\gamma(t), x_0) \) because we have \( \gamma_1(0) = \gamma(t) \), and \( \gamma_1(t_0 - t) = x_0 \). Using this in \((2.7.31)\) gives

\[
u(t_0, x_0) \leq \phi(t, x_0 - (t_0 - t)v) + \int_0^{t_0-t} L(x_0 - (t_0 - t)v + sv, v)ds
\]

\[
= \phi(t, x_0 - (t_0 - t)v) + \int_0^{t_0-t} L(x_0 - sv, v)ds,
\]

and, once again, this inequality becomes an equality at \( t = t_0 \), since \( u(t_0, x_0) = \phi(t_0, x_0) \). Just as before, differentiating in \( t \) at \( t = t_0 \) gives

\[
\phi_t(t_0, x_0) + v \cdot \nabla \phi(t_0, x_0) - L(x_0, v) \leq 0.
\]

\[(2.7.33)\]

As \((2.7.33)\) holds for all \( v \in \mathbb{R}^n \), it follows that

\[
\phi_t(t_0, x_0) + H(x_0, \nabla \phi(t_0, x_0)) \leq 0.
\]

\[(2.7.34)\]

Therefore, \( u \) is also a viscosity sub-solution to \((2.7.26)\), and the proof is complete. \( \square \)

**Exercise 2.7.9** Show the weak contraction and the finite speed of propagation properties, directly from the Lax-Oleinik formula.

**Instant regularization to Lipschitz**

We conclude this section with a remarkable result on instant smoothing. We will show that if the initial condition \( u_0 \) is continuous on \( \mathbb{T}^n \), then the solution to the Cauchy problem

\[
u_t + H(x, \nabla u) = 0, \quad t > 0, \quad x \in \mathbb{T}^n,
\]

\[
u(0, x) = u_0(x)
\]

\[(2.7.35)\]

becomes instantaneously Lipschitz. The improved regularity comes from the strict convexity of the Hamiltonian: indeed, nothing of that sort is true without this assumption, as can be seen from the following exercise.
Exercise 2.7.10 Consider the initial value problem
\[ u_t + |u_x| = 0, \quad t > 0, \quad x \in \mathbb{T}^1, \tag{2.7.36} \]
\[ u(0, x) = u_0(x). \]

(i) Show that the solution to (2.7.36) is given by
\[ u(t, x) = \inf_{|x-y| \leq t} u_0(y). \tag{2.7.37} \]

Hint: one may do this directly but also by considering a family of strictly convex Hamiltonians $H_\varepsilon(p)$ that converges to $H(p) = |p|$ as $\varepsilon \to 0$, and using the Lax-Oleinik semi-group for
\[ u_\varepsilon(t, x) = 0, \quad t > 0, \quad x \in \mathbb{T}^1, \tag{2.7.38} \]
\[ u_\varepsilon(0, x) = u_0(x). \]

Exercise 2.7.5 may be useful here.

(ii) Given an example of a continuous initial condition $u_0(x)$ such that the viscosity solution to (2.7.36) is not Lipschitz.

On the other hand, if the Hamiltonian is strictly convex we have the following result.

Theorem 2.7.11 Let $H(x, p)$ be strictly convex, and $u(t, x)$ be the unique solution to the Cauchy problem
\[ u_t + H(x, \nabla u) = 0, \]
\[ u(0, x) = u_0(x), \tag{2.7.39} \]
with $u_0 \in C(\mathbb{T}^n)$. For all $t > 0$, there is $C_t > 0$ such that $|\text{Lip } u(t, x)| \leq C_t$ on $[t, +\infty) \times \mathbb{T}^n$. The constant $C_t$ does not depend on the initial condition and blows up as $t \downarrow 0$.

Let us point the key difference with Proposition 2.6.13: as can be seen from the proof of that proposition, we used the Lipschitz property of the initial condition $u_0$, and showed that the solution remains Lipschitz at $t > 0$. Here, the initial condition is not assumed to be Lipschitz but only continuous, and the improved regularity comes from the convexity of the Hamiltonian.

Proof. It is sufficient to consider time intervals of length one, and repeat the argument on the subsequent intervals. Given $0 < t \leq 1$, and $x \in \mathbb{T}^n$, consider the extremal curve $\gamma(s)$ such that $\gamma(t) = x$, and
\[ u(t, x) = u_0(\gamma(0)) + \int_0^t L(\gamma(s), \dot{\gamma}(s)) \, ds. \tag{2.7.40} \]

As $0 \leq s \leq 1$, both $\gamma(s)$ and $\dot{\gamma}(s)$ are uniformly bounded. Of course, on the torus $\gamma(s)$ is always bounded but it would also be bounded for $0 \leq s \leq 1$ if we were considering the problem on $\mathbb{R}^n$. Take $h \in \mathbb{R}^n$, and define the curve
\[ \gamma_1(s) = \gamma(s) + \frac{s}{t} h, \quad 0 \leq s \leq t, \]
so that
\[ \gamma_1(0) = \gamma(0), \quad \gamma_1(t) = x + h. \tag{2.7.41} \]
We may use the Lax-Oleinik formula for \( u(t, x + h) \) and (2.7.40) for \( u(t, x) \), as well as (2.7.41), to write
\[
  u(t, x + h) = u(t, \gamma_1(t)) \leq u(\gamma_1(0)) + \int_0^t L(\gamma_1(s), \dot{\gamma}_1(s))ds \\
  = u(t, x) + \int_0^t (L(\gamma_1(s), \dot{\gamma}_1(s)) - L(\gamma(s), \dot{\gamma}(s))) ds.
\]  
(2.7.42)

The integral in the right side can be estimated as
\[
  \int_0^t (L(\gamma_1(s), \dot{\gamma}_1(s)) - L(\gamma(s), \dot{\gamma}(s))) ds = \int_0^t [L\left(\gamma(s) + \frac{s}{t}h, \dot{\gamma}(s) + \frac{1}{t}h\right) - L(\gamma(s), \dot{\gamma}(s))] ds \\
  \leq \int_0^t \frac{1}{t} \left( sh \cdot \nabla_x L(\gamma(s), \dot{\gamma}(s)) + h \cdot \nabla_v L(\gamma(s), \dot{\gamma}(s)) + C_t|h|^2 \right) ds + C_t|h|^2,
\]
with a constant \( C_t > 0 \) that may blow up as \( t \downarrow 0 \). We may now use the Euler-Lagrange
\[
  \frac{d}{ds} \nabla_v L(\gamma(s), \dot{\gamma}(s)) - \nabla_x L(\gamma(s), \dot{\gamma}(s)) = 0
\]
to rewrite (2.7.43) as
\[
  \int_0^t (L(\gamma_1(s), \dot{\gamma}_1(s)) - L(\gamma(s), \dot{\gamma}(s))) ds \\
  \leq \frac{1}{t} \int_0^t \left( sh \cdot \nabla_x L(\gamma(s), \dot{\gamma}(s)) + h \cdot \nabla_v L(\gamma(s), \dot{\gamma}(s)) \right) ds + C_t|h|^2 \\
  = h \cdot \nabla_v L(\gamma(t), \dot{\gamma}(t)) + C_t|h|^2.
\]
(2.7.44)

Using (2.7.44) in (2.7.42), we obtain
\[
  u(t, x + h) - u(t, x) \leq h \cdot \nabla_v L(\gamma(t), \dot{\gamma}(t)) + C_t|h|^2,
\]
(2.7.45)

which proves the Lipschitz regularity in the spatial variable for all \( 0 < t \leq 1 \), because both \( \gamma(t) \) and \( \dot{\gamma}(t) \) are bounded. Again, the boundedness of \( \gamma(t) \) would only play a role if we considered the problem on \( \mathbb{R}^n \), of course. Here, we use the fact that (2.7.45) holds for arbitrary \( x \) and \( y = x + h \) so that the role of \( x \) and \( y \) can be switched.

In order to prove the Lipschitz regularity in time, let us examine a small variation of \( t \), denoted by \( t + \tau \) with \( t + \tau > 0 \). Perturbing the extremal curve \( \gamma \) into
\[
  \gamma_2(s) = \gamma\left(\frac{t}{t + \tau}s\right),
\]
we still have
\[
  \gamma_2(0) = \gamma(0), \quad \gamma_2(t + \tau) = \gamma(t) = x.
\]
The same computation as above gives
\[
  u(t + \tau, x) = u(t + \tau, \gamma_2(t + \tau)) \leq u(\gamma_2(0)) + \int_0^{t+\tau} L(\gamma_2(s), \dot{\gamma}_2(s))ds \\
  = u(t, x) + \int_t^{t+\tau} (L(\gamma_2(s), \dot{\gamma}_2(s)) - L(\gamma(s), \dot{\gamma}(s))) ds + \int_t^{t+\tau} L(\gamma_2(s), \dot{\gamma}_2(s))ds.
\]
(2.7.46)
It is now straightforward to see that there exists $C'_t > 0$ that depends on $t$ so that

$$u(t + \tau, x) - u(t, x) \leq C'_t |\tau|.$$  

Once again, the role of $t$ and $t' = t + \tau$ can be switched, hence $u(t, x)$ is Lipschitz in $t$ as well, for any $t > 0$, finishing the proof. □

**Exercise 2.7.12** (i) Where did we use the strict convexity of the Hamiltonian in the above proof?

(ii) Consider again the initial value problem (2.7.36) with the convex but non strictly convex Hamiltonian $H(p) = |p|$ and a continuous initial condition $u_0(x)$ that is not Lipschitz continuous. Consider a sequence of smooth strictly convex Hamiltonians $H_\varepsilon(p)$ such that $H_\varepsilon(p) \to H(p)$ as $\varepsilon \to 0$, locally uniformly on $\mathbb{R}$. Review the above proof and see what will happen to the Lipschitz constant of the corresponding solution $u_\varepsilon(t, x)$ to the Cauchy problem

$$u_\varepsilon^t + H(u_\varepsilon^x) = 0, \quad t > 0, \quad x \in \mathbb{T},$$

constructed by the Lax-Oleinik formula. Hint: again, Exercise 2.7.5 may be useful here.

**Exercise 2.7.13** Take $t > 0$ and $\gamma(s)$ an extremal such that $u$ is differentiable at $x = \gamma(t)$. Show that

$$\nabla u(t, x) = \nabla_x L(x, \dot{\gamma}(t)).$$

and

$$u_t(t, x) = -H(x, \nabla u(t, x)).$$

(2.7.48)

### 2.7.2 Semi-concavity and $C^{1,1}$ regularity

As we have mentioned, the Cauchy problem for a Hamilton-Jacobi equation

$$u_t + H(x, \nabla u) = 0,$$

(2.7.49)

with a prescribed initial condition $u(0, x) = u_0(x)$, may have more than one Lipschitz solution, so it is worth asking whether the unique viscosity solution has some additional regularity when the Hamiltonian is strictly convex, so that the solution can be constructed by the Lax-Oleinik semigroup. A relevant notion is that of semi-concavity. Most of the material of this section comes from [62].

**Semi-concavity**

We begin with the following definition.

**Definition 2.7.14** If $B$ is an open ball in $\mathbb{R}^n$, $F$ a closed subset of $B$ and $K$ a positive constant, we say that $u \in C(B)$ is $K$-semi-concave on $F$ if for all $x \in F$, there is $l_x \in \mathbb{R}^n$ such that for all $h \in \mathbb{R}^n$ satisfying $x + h \in B$, we have:

$$u(x + h) \leq u(x) + l_x \cdot h + K|h|^2.$$  

(2.7.50)

The function $u$ is said to be $K$-semi convex on $F$ if $-u$ is $K$-semi-concave on $F$. 

152
Exercise 2.7.15  Examine the proof of Theorem 2.7.11 and check that it actually proves that for any $t > 0$ there exists $C_t > 0$ so that $u(t, x)$ is $C_t$-semi-concave in $x$.

The next theorem is crucial for the sequel. If $u$ is continuous in an open ball $B$ in $\mathbb{R}^n$, and $F$ is a closed subset of $B$, we say that $u \in C^{1,1}(F)$ if $u$ is differentiable in $F$ and $\nabla u$ is Lipschitz over $F$.

**Theorem 2.7.16**  Let $B$ be an open ball of $\mathbb{R}^n$ and $F$ closed in $B$. If $u \in C(B)$ is $K$-semi-concave and $K$-semi-convex in $F$, then $u \in C^{1,1}(F)$.

**Proof.** As $u$ is both $K$ semi-concave and $K$-semi-convex, for all $x \in F$, there are two vectors $l_x$ and $m_x$ such that for all $h$ such that $x + h \in B$ we have

$$
\begin{align*}
    u(x + h) &\leq u(x) + l_x \cdot h + K|h|^2, \\
    u(x + h) &\geq u(x) + m_x \cdot h - K|h|^2
\end{align*}
$$

(2.7.51)

which yields

$$(m_x - l_x) \cdot h \leq 2K|h|^2.$$ 

As this is true for all $h$, we conclude that $l_x = m_x$ and, therefore, $u$ is differentiable at $x$, and

$$l_x = m_x = \nabla u(x).$$

Next, we show that $\nabla u$ is Lipschitz over $F$. Given $(x, y, h) \in F \times F \times \mathbb{R}^n$, the semi-convexity and semi-concavity inequalities, written, respectively, between $x + h$ and $x$, $x$ and $y$, and $x + h$ and $y$, give:

$$
\begin{align*}
|u(x + h) - u(x) - \nabla u(x) \cdot h| &\leq K|h|^2 \\
|u(x) - u(y) - \nabla u(y) \cdot (x - y)| &\leq K|x - y|^2 \\
|u(y) - u(x + h) + \nabla u(y) \cdot (x + h - y)| &\leq K|x + h - y|^2.
\end{align*}
$$

Adding the three inequalities above, we obtain:

$$
|\nabla u(x) - \nabla u(y)| \cdot h \leq 3K(|h|^2 + |x - y|^2).
$$

(2.7.52)

Taking

$$h = |x - y| \frac{\nabla u(x) - \nabla u(y)}{|\nabla u(x) - \nabla u(y)|},$$

in the inequality (2.7.52) gives

$$|\nabla u(x) - \nabla u(y)| \leq 6K|x - y|,$$

which is the Lipschitz property of $\nabla u$ that we sought. □
Improved regularity of the viscosity solutions

Let us come back to the solution $u(t, x)$ to the Cauchy problem

$$
\begin{align*}
    u_t + H(x, \nabla u) &= 0, \\
    u(0, x) &= u_0(x).
\end{align*}
$$

(2.7.53)

We first prove that if $\gamma$ is a minimizing curve for $u(t, x)$, then it is also a minimizer for $u(s, \gamma(s))$ for all $0 \leq s \leq t$.

**Proposition 2.7.17** Fix $t > 0$ and $x \in \mathbb{T}^n$, and a minimizing curve $\gamma$ such that $\gamma(t) = x$, and

$$
    u(t, x) = u_0(\gamma(0)) + \int_0^t L(\gamma(s), \dot{\gamma}(s)) \, ds.
$$

(2.7.54)

Then for all $0 \leq s \leq s' \leq t$ we have

$$
    u(s', \gamma(s')) = u_0(\gamma(0)) + \int_0^{s'} L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma = u(s, \gamma(s)) + \int_s^{s'} L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma.
$$

(2.7.55)

**Exercise 2.7.18** Relate the result of this proposition to the dynamic programming principle.

**Proof.** The Lax-Oleinik formula implies that for all $0 < s < t$ we have

$$
    u(s, \gamma(s)) \leq u_0(\gamma(0)) + \int_0^s L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma.
$$

Assume that for some $0 < s < t$, we have a strict inequality

$$
    u(s, \gamma(s)) < u_0(\gamma(0)) + \int_0^s L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma.
$$

(2.7.56)

Then, there exists a curve $\gamma_1(s')$, $0 \leq s' \leq s$, such that $\gamma_1(s) = \gamma(s)$, and

$$
    u_0(\gamma_1(0)) + \int_0^s L(\gamma_1(\sigma), \dot{\gamma}_1(\sigma)) \, d\sigma < u_0(\gamma(0)) + \int_0^s L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma.
$$

Then, we can consider the concatenated curve $\gamma_2(s)$ so that $\gamma_2(s') = \gamma_1(s')$ for $0 \leq s' \leq s$, and $\gamma_2(s') = \gamma(s')$ for $s \leq s' \leq t$. The resulting curve is piece-wise $C^1[0, t]$, hence is an allowed trajectory. This would give

$$
\begin{align*}
    u(t, \gamma(t)) &= u_0(\gamma(0)) + \int_0^s L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma + \int_s^t L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma \\
    &> u_0(\gamma_2(0)) + \int_0^t L(\gamma_2(s), \dot{\gamma}_2(s)) \, ds,
\end{align*}
$$

(2.7.57)

which would contradict the extremal property of the curve $\gamma$ between the times $0$ and $t$. Therefore, (2.7.56) can not hold, and for all $0 \leq s \leq s' \leq t$ we have:

$$
    u(s', \gamma(s')) = u_0(\gamma(0)) + \int_0^{s'} L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma = u(s, \gamma(s)) + \int_s^{s'} L(\gamma(\sigma), \dot{\gamma}(\sigma)) \, d\sigma.
$$

(2.7.58)

This finishes the proof of Proposition 2.7.17. □
Definition 2.7.19  We say that \( \gamma : [0, t] \to \mathbb{T}^n \) is calibrated by \( u \) if (2.7.54) holds.

Let us define the conjugate semigroup of the Lax-Oleinik semigroup by:

\[
\tilde{T}(t)u_0(x) = \sup_{y \in \mathbb{T}^n} (u_0(y) - h_t(x, y)), \quad \forall u_0 \in C(\mathbb{T}^n), \; t > 0. (2.7.59)
\]

We will denote \( \tilde{u}(t, x) = \tilde{T}(t)u_0(x) \). The following lemma is proved exactly as Theorem 2.7.11.

Lemma 2.7.20  Let \( u_0 \in C(\mathbb{T}^n) \) and \( \sigma > 0 \). There is \( K(\sigma) > 0 \) such that \( \tilde{T}(\sigma)u_0 \) is \( K(\sigma) \)-semi-convex. The constant \( K(\sigma) \) blows up as \( \sigma \to 0 \).

Given \( 0 < s < s' \), we define the set \( \Gamma_{s,s'}[u_0] \) as the union of all points \( x \in \mathbb{T}^n \) so that the extremal calibrated by \( u \), which passes through the point \( x \) at a time \( s_1 \in [s, s'] \) can be continued forward in time until the time \( s' \), and backward in time until the time \( s \).

Corollary 2.7.21  Let \( u_0 \in C(\mathbb{T}^n) \) and \( u(t, x) = T(t)u_0(x), \) and \( 0 < s_1 < s_2 \), then for any \( \varepsilon > 0 \), the function \( u \in C^{1,1}([s_1, s_2] \times \Gamma_{s_1, s_2+\varepsilon}) \).

Proof. Let us first deal with the spatial regularity. Let us take \( s \in [s_1, s_2] \) and \( x_0 \in \Gamma_{s_1, s_2+\varepsilon} \), so that the extremal \( \gamma \) such that \( x_0 = \gamma(s) \) can be continued past the time \( s \), until the time \( s_2 + \varepsilon \).

As we have mentioned in Exercise 2.7.15, there is \( K > 0 \) depending on \( s_1 \) such that the function \( u(t, x) \) is \( K \)-semi-concave at all \( x \in \mathbb{T}^n \), in particular, at \( x_0 \). Hence, we only need to argue that \( u \) is semi-convex at \( x_0 \), and here we are going to use the fact that \( x_0 \in \Gamma_{s_1, s_2+\varepsilon} \). Note that for all \( y \in \mathbb{R}^n \) we have, by the Lax-Oleinik formula,

\[
u(s_2 + \varepsilon, y) \leq u(s_2 + \varepsilon, y). (2.7.60)\]

In addition, the calibration relation (2.7.58) implies that if \( x_0 = \gamma(s) \) and \( x_0 \in \Gamma_{s_1, s_2+\varepsilon} \), then equality is attained when \( y = \gamma(s_2 + \varepsilon) \). We conclude that in this case we have

\[
u(s, x_0) = \sup_{y \in \mathbb{T}^n} (u(s_2 + \varepsilon, y) - h_{s_2+\varepsilon - s}(x_0, y)) = \tilde{T}(s_2 + \varepsilon - s)[u(s_2 + \varepsilon, \cdot)](x_0). \]

It follows from Lemma 2.7.20 that there is a constant \( \tilde{K} \) depending on \( \varepsilon \), such that \( u(s, \cdot) \) is \( \tilde{K} \)-semi-convex in \( x \) on \( [s_1, s_2] \times \Gamma_{s_1, s_2+\varepsilon} \).

Theorem 2.7.16 now implies that the function \( u(s, \cdot) \) is \( C^{1,1} \) in \( x \) on the set \( \Gamma_{s_1, s_2+\varepsilon} \) for all \( s_1 \leq s \leq s_2 \). To end the proof, one just has to invoke relation (2.7.48) in Exercise 2.7.13 to obtain the corresponding regularity in the time variable.

This corollary may not, at first sight, look so striking. To enjoy its scope, let us specialize it to the solutions to the stationary equation

\[
H(x, \nabla u) = 0, \tag{2.7.61}
\]

assuming that they exist. Corollary 2.7.21 allows us to discover the following
Corollary 2.7.22 Consider a solution $u$ of (2.7.61), and let $F$ be the set of all points $x \in \mathbb{T}^n$ such that there exists $\varepsilon_x > 0$ and a $C^1$ curve $\gamma: (-\varepsilon_x, \varepsilon_x) \to \mathbb{T}^n$ such that $\gamma(0) = x$ and

$$u(\gamma(\varepsilon_x)) - u(\gamma(-\varepsilon_x)) = \int_{-\varepsilon_x}^{\varepsilon_x} L(\gamma(s), \dot{\gamma}(s))ds. \quad (2.7.62)$$

Then $u \in C^{1,1}(F)$.

In other words, $u$ is $C^{1,1}$ at every point through which an extremal of the Lagrangian passes, as opposed to ending at this point.

Let us examine some further consequences of this fact, in the form of a few exercises, just to give a glimpse of how far reaching these considerations can be. Their solution does not need more tools or ideas than the ones already presented, but the reader may find them fairly elaborate. We begin with an application of the finite speed of propagation property.

Exercise 2.7.23 Let $u(x)$ be a Lipschitz viscosity solution of

$$H(x, \nabla u) = 0$$

in a bounded open subset $\Omega$ of $\mathbb{R}^n$. Show that, for every open subset $\Omega_1$ of $\Omega$ such that $\overline{\Omega}_1$ is compactly embedded in $\Omega$, there is $\varepsilon > 0$ such that, for all $t \in [0, \varepsilon]$ and $x \in \Omega_1$ we have

$$u(x) = T(t)u(x).$$

We continue with a statement that looks surprisingly elementary. However its solution is not.

Exercise 2.7.24 Let $\Omega$ be an open subset of $\mathbb{R}^n$ and $u_p$ a sequence in $C^1(\Omega)$, such that

$$|\nabla u_p| = 1$$

for all $p$. Show that all uniform limits of $u_p$ are $C^1$ functions. Hint: if $x_0 \in \Omega$, then, for small $\varepsilon > 0$, the function $u_p(x)$ coincides, in a small neighborhood of $x$, with both $T(\varepsilon)u_p$ and $\tilde{T}(\varepsilon)u_p$.

We end the section with two regularity properties of the distance function. Recall that, if $S$ is a set of $\mathbb{R}^n$, the distance function to $S$ is given by

$$d_S(x) = \inf_{v \in S} |x - v|.$$ 

It is, obviously, a Lipschitz function with Lipschitz constant 1. We can say much more, just recalling the age-old fact that the shortest path between two points is the line joining these two points: this makes $d_S$ a viscosity solution of $|\nabla d| = 1$, or, even better:

$$|\nabla d|^2 = 1. \quad (2.7.63)$$

We may use the previous theory for the following results.

Exercise 2.7.25 If $S$ is a compact set, $x_0 \notin S$ and $v$ is such that

$$|x - v| = d_S(x),$$

then $d_S$ is $C^{1,1}$ on the line segment $[v, x]$.

Exercise 2.7.26 If $S$ is a convex set, then $d_S$ is $C^{1,1}$ outside $S$.

If you are stuck with any of the above three exercises, see [62].
2.8 Large time behavior in a particular case

For the rest of this chapter, we go back to the long term behavior of the solutions to the Hamilton-Jacobi equations but unlike in Section 2.3, we now we consider the inviscid case. In this section, we are going to prove that the solution to the Cauchy problem

$$u_t + R(x) \sqrt{1 + |\nabla u|^2} = 0, \quad t > 0, \quad x \in \mathbb{T}^n$$

$$u(0, x) = u_0(x),$$

converges, similarly to what happened in Theorem 2.3.3 for the viscous problem, to a wave solution, even though viscosity is not present anymore.

Let us first explain how equation (2.8.1) comes up from simple geometric principles. Consider a family of hypersurfaces $\Sigma(t)$ of $\mathbb{R}^n$, moving according to an imposed normal velocity $R(x)$:

$$V_n = R(x), \quad x \in \mathbb{R}^n,$$

the function $R$ being given and positive. Assume that, at each time $t$, the surface $\Sigma(t)$ is the level set of a function $v(t, x)$:

$$\Sigma(t) = \{x \in \mathbb{R}^{n+1} : v(t, x) = 0\}.$$  

It is easy to see that the normal velocity $V_n$ at the point $x$, at time $t$, is given by

$$V_n(t, x) = \frac{v_t(t, x)}{|\nabla v(t, x)|},$$

so that the evolution equation for the function $v(t, x)$ is

$$v_t = R(x) |\nabla v| \quad \text{on } \Sigma(t).$$

Assume that $\Sigma(t)$ is given in the form of a graph of a function $u(t, x')$, $x' \in \mathbb{R}^{n-1}$, that is:

$$v(t, x) = x_n - u(t, x'), \quad x' \in \mathbb{R}^{n-1}.$$  

Then we obtain (2.8.1) for the function $u(t, x)$:

$$u_t + R(x') \sqrt{1 + |\nabla_{x'} u|^2} = 0.$$  

We will assume that the function $R(x)$ is smooth and positive: there exists $R_0 > 0$ so that

$$R(x) \geq R_0 > 0 \quad \text{for all } x \in \mathbb{T}^n.$$  

We will also use the notation

$$\bar{R} = \|R\|_{L^\infty}.$$  

Notice that the assumptions for the Hamiltonian $H(x, p) = R(x) \sqrt{1 + |p|^2}$ obviously fall in line with those made in Section 2.3 on the convergence to the viscous waves, and in Section 2.6 on the existence of the inviscid waves and of the solutions to the inviscid Cauchy problem.
Theorem 2.8.1 Let \( u_0 \in C(\mathbb{T}^n) \) and assume that \( R(x) \) is smooth and satisfies (2.8.5). There is a solution \( u_\infty(x) \) to
\[
R(x)\sqrt{1 + |\nabla u_\infty|^2} = \bar{R}, \quad x \in \mathbb{T}^n,
\]
such that we have, uniformly with respect to \( x \in \mathbb{T}^n \):
\[
\lim_{t \to +\infty} \left( u(t,x) + t\bar{R} - u_\infty(x) \right) = 0.
\]
Let us comment that \( H(x,p) = R(x)\sqrt{1 + |p|^2} \) is locally strictly convex in its second variable (but not uniformly strictly convex), and we could choose to attack the problem via the Lax-Oleinik formula, with a little extra help due to the lack of the global strict convexity. We decide, however, not to rely on the strict convexity in any form: the arguments that we are going to display will work, at almost no additional cost, for the important class of Hamiltonians of the form
\[
H(x,p) = |\nabla u| - f(x),
\]
which are not strictly convex even locally. The proof is inspired by the arguments in [107].

2.8.1 Counting the waves

In this section, we discuss the wave solutions to (2.8.1). They satisfy
\[
R(x)\sqrt{1 + |\nabla u|^2} = c, \quad x \in \mathbb{T}^n,
\]
an equation that can be alternatively stated as
\[
|\nabla u(x)|^2 = f(x), \quad x \in \mathbb{T}^n,
\]
with
\[
f(x) = \frac{c^2}{R^2(x)} - 1.
\]
Thus, we will also look at the closely related question of existence of the solutions to (2.8.11) with a general function \( f(x) \).

Identification of the speed

We begin with the following.

Proposition 2.8.2 A solution to an equation of the form
\[
|\nabla u(x)|^2 = f(x) + \gamma, \quad x \in \mathbb{T}^n,
\]
with a smooth function \( f \) exists if and only if
\[
\gamma = -\min_{x \in \mathbb{T}^n} f(x).
\]
In other words, a solution to
\[ |\nabla u(x)|^2 = f(x), \quad x \in \mathbb{T}^n, \tag{2.8.15} \]
exists if and only if
\[ \min_{x \in \mathbb{T}^n} f(x) = 0. \tag{2.8.16} \]

A consequence of this proposition is that the only \( c \) such that equation
\[ R(x)\sqrt{1 + |\nabla u|^2} = c, \quad x \in \mathbb{T}^n, \tag{2.8.17} \]
has a solution \( u_{\infty}(x) \) is \( c = \bar{R} \), as seen from (2.8.11)-(2.8.12). Thus, a solution to (2.8.7) exists, as claimed in the first part of Theorem 2.8.1. Let us elaborate this point a little further. Note that the unique \( c \) for which (2.8.17) has a solution, can be alternatively defined as the only value of \( \gamma \) such that each solution to the Cauchy problem
\[ u_t + R(x)\sqrt{1 + |\nabla u|^2} = \gamma, \quad t > 0, \quad x \in \mathbb{T}^n, \]
\[ u(0, x) = u_0(x), \tag{2.8.18} \]
is uniformly bounded in time. This is an immediate consequence of the comparison principle for the viscosity solutions.

**Exercise 2.8.3** (i) Explain this point: show that if \( \gamma \neq c \) then the solution to the Cauchy problem (2.8.18) can not remain bounded as \( t \to +\infty \), and, conversely, if \( \gamma = c \) then it remains bounded as \( t \to +\infty \).

(ii) Show also that \( c \) is the unique value \( \gamma \) such that there exists both a sub-solution \( \underline{u} \) and a super-solution \( \bar{u} \) to
\[ R(x)\sqrt{1 + |\nabla u|^2} = \gamma, \quad x \in \mathbb{T}^n. \tag{2.8.19} \]

**Proof of Proposition 2.8.2.** We know that for each \( f \in C(\mathbb{T}^n) \) there exists some \( \gamma \in \mathbb{R} \) such that a solution to
\[ |\nabla u(x)|^2 = f(x) + \gamma, \quad x \in \mathbb{T}^n, \tag{2.8.20} \]
exists. We need to show that
\[ \gamma = -\min_{x \in \mathbb{T}^n} f(x). \tag{2.8.21} \]

Thus, as in Exercise 2.8.3(ii), we only need to construct a sub-solution and a super-solution to (2.8.20) for \( \gamma \) as in (2.8.21). First, observe that if
\[ \gamma + \min_{x \in \mathbb{T}^n} f(x) \geq 0, \tag{2.8.22} \]
then all constants are sub-solutions to (2.8.20).

On the other hand, a quadratic function of the form
\[ \bar{u}(x) = \frac{\alpha}{2} (x - x_0)^2, \tag{2.8.23} \]
with some \( x_0 \in \mathbb{T}^n \), is a super-solution to (2.8.20) if
\[ \alpha^2 |x - x_0|^2 \geq f(x) + \gamma, \quad \text{for all } x \in \mathbb{T}^n. \tag{2.8.24} \]
It follows that, in particular, 
\[ f(x_0) + \gamma \leq 0, \]
hence such super-solution can exist only if 
\[ \gamma + \min_{x \in \mathbb{T}^n} f(x) \leq 0. \] (2.8.25)

On the other hand, if (2.8.25) does hold, \( x_0 \) is a minimum of \( f(x) \), and \( f \) is smooth, as we assume here, then (2.8.24) does hold if we choose \( \alpha > 0 \) to be sufficiently large.

Thus, if \( \gamma = -\min_{x \in \mathbb{T}^n} f(x) \) then we can find both a sub-solution and a super-solution to (2.8.20), finishing the proof. \( \square \)

**Exercise 2.8.4** Note that the super-solution we have constructed in (2.8.23) is not periodic. Explain why this is not an issue.

**Exercise 2.8.5** Note that we did use the assumption that \( f(x) \) is smooth in the construction of the super-solution in the above proof. Show that nevertheless the conclusion of Proposition 2.8.2 holds for \( f \in C(\mathbb{T}^n) \). Hint: approximate \( f \in C(\mathbb{T}^n) \) by a sequence of smooth functions \( f_k \) that converges uniformly to \( f \) and obtain a uniform Lipschitz bound for the solutions to

\[ |\nabla u_k|^2 = f_k(x) + \gamma_k, \quad \gamma_k := -\min_{x \in \mathbb{T}^n} f_k(x), \]
such that \( u_k(0) = 0 \). Finally, use the stability property of the viscosity solutions to show that \( u_k \) converges, along a subsequence, to a viscosity solution to

\[ |\nabla u|^2 = f(x) + \gamma, \quad \gamma := -\min_{x \in \mathbb{T}^n} f(x). \] (2.8.26)

**A simple example of the non-uniqueness of the waves**

We now discuss examples when the solutions to

\[ H(x, \nabla u) = c \] (2.8.27)

may be not unique. Recall that there can only be one \( c \) for which solutions to (2.8.27) may exist – this is a simple consequence of the comparison principle for the viscosity solutions. However, we will see that for a given \( c \in \mathbb{R} \), solutions to (2.8.27) may be not unique. This is a big difference with the viscous case

\[ -\Delta u + H(x, \nabla u) = c, \] (2.8.28)
described in Theorem 2.3.1, where both \( c \) and \( u \) are unique.

To see that uniqueness may fail, consider a very simple example in one dimension, of the same nature as in Proposition 2.8.2.

\[ |u'| = f(x), \quad x \in \mathbb{T}^1. \] (2.8.29)

Assume that \( f \in C^1(\mathbb{T}^1) \) is 1/2-periodic, satisfies

\[ f(x) > 0 \text{ on } (0, 1/2) \cup (1/2, 1), \text{ and } f(0) = f(1/2) = f(1) = 0. \]
and is symmetric with respect to $x = 1/4$ (and thus $x = 3/4$). Let $u_1$ and $u_2$ be 1-periodic and be defined, over a period, as follows:

$$u_1(x) = \begin{cases} 
\int_0^x f(y) \, dy, & 0 \leq x \leq \frac{1}{2}, \\
\int_x^1 f(y) \, dy, & \frac{1}{2} \leq x \leq 1,
\end{cases}$$

$$u_2(x) = \begin{cases} 
\int_x^{1/2} f(y) \, dy, & 0 \leq x \leq \frac{1}{4}, \\
\int_{1/2}^x f(y) \, dy, & \frac{1}{4} \leq x \leq \frac{1}{2}, \\
u_2 \text{ is } \frac{1}{2}\text{-periodic.}
\end{cases}$$

Note that $u_1(x)$ is continuously differentiable but $u_2(x)$ is only Lipschitz: its graph has corners at $x = 1/4$ and $x = 3/4$.

**Exercise 2.8.6** Verify that both $u_1$ and $u_2$ are viscosity solutions of (2.8.29), and $u_2$ cannot be obtained from $u_1$ by the addition a constant. Pay attention to what happens at $x = 1/4$ and $x = 3/4$ with $u_2(x)$. Why can’t you construct a solution that would have a corner at a minimum rather than the maximum?

### Trajectories at very negative times

We are now going to carry out a more systematic study of the steady solutions to

$$|\nabla u|^2 = f(x), \quad x \in \mathbb{T}^n. \quad (2.8.30)$$

We assume that the function $f$ is smooth and non-negative and vanishes at a finite number of points $x_1, \ldots, x_N$. The smoothness assumption is adopted merely for convenience, continuity of $f$ would certainly suffice. What follows is a (much simplified) adaptation of the last chapter of the book of Fathi [63].

The viscosity solution to (2.8.30) exists by Proposition 2.8.2 and our assumptions on $f$. It is given by the Lax-Oleinik formula: for any $t < 0$ we have

$$u(x) = \inf_{\gamma(0)=x} \left( u(\gamma(t)) + \int_t^0 \left( \frac{|\gamma(s)|^2}{4} + f(\gamma(s)) \right) ds \right). \quad (2.8.31)$$

We know from the preceding section that the infimum is, in fact, a minimum, attained at an extremal of the Lagrangian, that we denote $\gamma_t(s), t \leq s \leq 0$. The Lagrangian associated to the Hamiltonian $H(x,p) = |p|^2 - f(x)$ is

$$L(x,v) = \frac{|v|^2}{4} + f(x). \quad (2.8.32)$$

This means, in particular, that $L(x,v)$ is nonnegative and vanishes only at the points of the form $(x,v) = (x_i, 0), i \in \{1, \ldots, N\}$. Hence, we expect that the minimizers in (2.8.31) should prefer to stay near the points where $f$ vanishes, and move very slowly around those points. To formalize this idea, we would like to send the starting time $t \to -\infty$ and say that each minimizing curve $\gamma_t(s)$ is near one of $x_i$, for $s$ sufficiently large and negative.
Lemma 2.8.7 The function $u(x)$ can be written as

$$u(x) = \inf_{i \in \{1, \ldots, N\}} \inf_{\gamma(0) = x_i} \left( u(x_i) + \int_{-\infty}^{0} \left( \frac{|\dot{\gamma}(s)|^2}{4} + f(\gamma(s)) \right) ds \right),$$

(2.8.33)

with the infimum taken over all curves $\gamma(s)$ such that $\gamma(0) = x$ and $\gamma(s) \to x_i$ as $s \to -\infty$.

Proof. First, note that $u(x)$ is bounded by the right side of (2.8.33), as follows immediately from the Lax-Oleinik formula (2.8.31). We need to show that equality is actually attained. Let us take $t < 0$ large and negative and consider the corresponding minimizer $\gamma_t(s)$, so that

$$u(x) = u(\gamma(-t)) + \int_{t}^{0} \left( \frac{|\dot{\gamma}(s)|^2}{4} + f(\gamma(s)) \right) ds.$$  

(2.8.34)

Take $\varepsilon > 0$ and consider the set

$$D_\varepsilon = \{ y \in \mathbb{T}^n : |y - x_i| \leq \varepsilon \text{ for some } 1 \leq i \leq N \}.$$

If $\varepsilon > 0$ is sufficiently small than $D_\varepsilon$ is a union of $N$ pairwise disjoint balls

$$B_\varepsilon^{(k)} = \{ y \in \mathbb{T}^n : |y - x_k| \leq \varepsilon \}.$$

The function $f(y)$ is strictly positive outside of $D_\varepsilon$: there exists $\lambda_\varepsilon > 0$ so that $f(y) > \lambda_\varepsilon$ for all $y \notin D_\varepsilon$. It follows that the total time that $\gamma_t(s)$ spends outside of $D_\varepsilon$ is also bounded:

$$|\{ s \in [t, 0] : \gamma(s) \notin D_\varepsilon \}| \leq \frac{2\|u\|_{L^\infty}}{\lambda_\varepsilon}. $$

(2.8.35)

Exercise 2.8.8 Show that there exists $\mu_\varepsilon$ so that if $t < s_1 < s_2 < 0$, and $\gamma(s_1) \in B_\varepsilon^{(k)}$ while $\gamma(s_2) \in B_\varepsilon^{(k')}$ with $k \neq k'$, then

$$\int_{s_1}^{s_2} \left( \frac{|\dot{\gamma}(s)|^2}{4} + f(\gamma(s)) \right) ds \geq \mu_\varepsilon. $$

(2.8.36)

Hint: show that if the switch from $B_\varepsilon^{(k)}$ to $B_\varepsilon^{(k')}$ happens ”quickly” then the contribution of the first term inside the integral is bounded from below, and if this switch happens slowly, then the contribution of the second term inside the integral is bounded from below.

We see from (2.8.35) and (2.8.36) that there exists $T_\varepsilon < 0$ such that if $t < s < T_\varepsilon$, then there exists $1 \leq k \leq N$ so that $\gamma_t(s) \in B_\varepsilon^{(k)}$. In other words, $\gamma_t(s)$ remains in one of these balls for all $s < T_\varepsilon$. It follows that there is $1 \leq k \leq N$ and a sequence $t_m \to -\infty$ such that $\gamma_{t_m}(s) \in B_k^{(e)}$ for all $s < T_\varepsilon$. The uniform $H^1$-bound on the minimizers allows us, after extraction of a subsequence, to pass to the limit $t_m \to -\infty$ and obtain a curve $\gamma(s)$, defined for all $s \leq 0$, such that $\gamma_{t_m} \to \gamma(s)$, $\gamma(-\infty) = x_k$, $\gamma(0) = x$, and

$$u(x) = u(x_k) + \int_{-\infty}^{0} \left( \frac{|\dot{\gamma}(s)|^2}{4} + f(\gamma(s)) \right) ds.$$  

(2.8.37)

It follows that $u(x)$ is bounded from below by the right side of (2.8.33), and the proof of Lemma 2.8.7 is complete. □
Classification of steady solutions

We can now classify all solutions to
\[ |\nabla u|^2 = f(x), \quad x \in \mathbb{T}^n. \] (2.8.38)

Let us set
\[ S(x_i, x) = \inf_{\gamma(-\infty)=x_i} \int_{-\infty}^{0} \left( \frac{|\dot{\gamma}(s)|^2}{4} + f(\gamma(s)) \right) ds. \] (2.8.39)

It may be seen as the energy of a connection between \( x_i \) and \( x \), or, in a more mathematically precise way, as a sort of distance between \( x_i \) and \( x \). This fruitful point of view, developed in [63], will not be pushed further here.

**Theorem 2.8.9** Let \( x_1, \ldots, x_N \) be the zeros of a smooth non-negative function \( f(x) \). Given a collection of numbers \( \{a_1, \ldots, a_N\} \) there is a unique solution \( u(x) \) to (2.8.38), such that
\[ u(x_i) = a_i \text{ for all } 1 \leq i \leq N, \] (2.8.40)
if and only if for all \((i, j) \in \{1, \ldots, N\}^2:\)
\[ a_j - a_i \leq S(x_i, x_j). \] (2.8.41)

**Proof.** Lemma 2.8.7 already shows that the values of \( u(x_i) \) determine the value of \( u(x) \) for all \( x \in \mathbb{T}^n \), and that if a solution exists and (2.8.40) holds, then
\[ a_j = \inf_{i \in \{1, \ldots, N\}} (a_i + S(x_i, x_j)). \] (2.8.42)

This implies (2.8.41).

To prove existence if \( a_i \) are given and satisfy (2.8.41), set
\[ u(x) = \inf_{i \in \{1, \ldots, N\}} (a_i + S(x_i, x)). \] (2.8.43)

Using the by now familiar arguments, it is easy to see that \( u \) is a solution to (2.8.38). Moreover, we have
\[ u(x_j) = \inf_{i \in \{1, \ldots, N\}} (a_i + S(x_i, x_j)). \]
This, together with (2.8.41) implies \( u(x_j) = a_j. \)

**Exercise 2.8.10** Apply the above theory to the equation \(|u'| = f(x)\) on \( \mathbb{T}^1 \), with the non-negative function \( f(x) \) vanishing at 2 or 3 distinct points. Find out how many different solutions one may have.

**Exercise 2.8.11** Let \( \Omega \) be a smooth bounded subset of \( \mathbb{R}^n \). Assume that \( f \) is nonnegative and vanishes only at a finite number of points and \( u_0 \in C(\partial \Omega) \). Find a necessary and sufficient condition on the values of \( u_0 \) so that the boundary value problem
\[ |\nabla u|^2 = f(x), \quad x \in \Omega, \]
\[ u(x) = u_0(x), \quad x \in \partial \Omega, \] (2.8.44)
is well-posed. Count its solutions. If you have difficulty, we recommend that you read the very remarkable study of the non-uniqueness in Lions [93].
Exercise 2.8.12 Let $f$ be a smooth non-negative function of $\mathbb{R}^n$ such that $f(0) = 0$, and

$$\alpha I \leq D^2 f \leq \frac{1}{\alpha} I,$$

in the sense of matrices. Describe the characteristics for

$$|\nabla u|^2 = f(x), \quad x \in \mathbb{R}^n,$$

and classify its solutions.

2.8.2 The large time behavior

In the remainder of the section, we will prove that every solution of the Cauchy Problem (2.8.1) converges to a wave solution. Subtracting the linear term $t\|R\|_{\infty}$ from $u$, and still calling $u(t, x)$ the new unknown, recall that we are solving the equation

$$u_t + R(x)\sqrt{1 + |\nabla u|^2} - \|R\|_{\infty} = 0 \quad x \in \mathbb{T}^n \quad u(0, x) = u_0(x) \quad (2.8.45)$$

To end the proof of Theorem 2.8.1, it suffices to prove the

Proposition 2.8.13 Assume $u_0 \in C^1(\mathbb{T}^n)$. There is a solution $u_\infty(x)$ of (2.8.7) such that $u(t, x)$ of (2.8.45) satisfies

$$\lim_{t \to +\infty} u(t, x) = u_\infty(x),$$

uniformly in $x \in \mathbb{T}^n$.

Indeed, the weak contraction property implies that Proposition 2.8.13 carries over to continuous initial data $u_0$, thus ending the proof of Theorem 2.8.1. So, restricting ourselves to $C^1$ initial data, we obtain a unique viscosity solution to (2.8.45) that satisfy the bounds

$$|u(t, x)|, \ |u_t(t, x)|, \ |\nabla u(t, x)| \leq C,$$

for some universal $C > 0$.

Proof of Proposition 2.8.13. Let $\mathcal{Z}$ be the set where $R$ reaches its maximum:

$$\mathcal{Z} = \{x \in \mathbb{T}^n : R(x) = \|R\|_{\infty}\}.$$

We are going to prove that $u$ is nonincreasing on $\mathcal{Z}$, it is intuitively obvious: assume $x \in \mathcal{Z}$ and suppose that the equation for $u$ holds at the point $(t, x)$, then we have

$$u_t(t, x) = \|R\|_{\infty} \left(1 - \sqrt{1 + |\nabla u(t, x)|^2}\right) \leq 0.$$

The trouble is that of course, the equation for $u$ holds almost everywhere and we have no guarantee that it holds at $(t, x)$. To make this simple argument rigorous, consider $t_0 > 0$ and $x_0 \in \mathcal{Z}$ and

$$\overline{u}(t, x) = u(t_0, x) - (t - t_0)(R(x) - \|R\|_{\infty}).$$
We claim that \( u \) is a super-solution to (2.8.45) on \([t_0, +\infty) \times \mathbb{T}^n\), that is equal to \( u \) at \( t = t_0 \). This implies \( u(t, x) \leq \overline{u}(t, x) \) for \((t, x) \in [t_0, +\infty) \times \mathbb{T}^n\); writing this at \( x = x_0 \in \mathcal{Z} \) yields \( u(t, x_0) \leq u(t_0, x_0) \). To prove the super-solution property, let \( \varphi(t, x) \) be a test function and \((t_1, x_1) \in [t_0, +\infty)\) be a minimum point for \( u - \varphi \). The function \( u(t, x) \) being obviously a viscosity solution of
\[
  u_t + R(x) - \|R\|_\infty = 0, \quad x \in \mathbb{T}^n,
\]
we have
\[
  \varphi_t(t_1, x_1) + R(x_1)\sqrt{1 + |\nabla \varphi(t_1, x_1)|^2} - \|R\|_\infty \\
  \geq \varphi_t(t_1, x_1)R(x) - \|R\|_\infty \\
  = 0.
\]
This proves the super-solution property.

Thus, \( u(t, \cdot) \) converges on \( \mathcal{Z} \). The large time convergence of \( u(t, \cdot) \) outside \( \mathcal{Z} \) will follow from the fact that (a transform of) \( u \) solves a Hamilton-Jacobi that is more complex than (2.8.45), but that has the merit of carrying an absorption term, that will force all the terms of \( u(t, \cdot) \) to accumulate around a single one. We will use the Kruzhkov transform, that has proved to be surprisingly effective in the study of first order Hamilton-Jacobi equations:
\[
  v(t, x) = -e^{-u(t, x)}.
\]
We have
\[
v_t = |v|u_t, \quad \nabla v = |v|\nabla u.
\]
Because of the \( L^\infty \) and gradient bounds for \( u \), the function \( v \) satisfies \( L^\infty \) and gradient bounds of the same type. Moreover, because the function \( u \mapsto -e^{-u} \) is increasing in \( u \), the function \( v \) is a viscosity solution of
\[
v_t + R(x)\frac{|\nabla v|^2}{|v| + \sqrt{|v|^2 + |\nabla v|^2}} + (\|R\|_\infty - R(x))v = 0, \quad t > 0, \quad x \in \mathbb{T}^n. \tag{2.8.46}
\]
For \( \delta > 0 \), let \( \mathcal{Z}_\delta \) be the set of all points that are at distance at most \( \delta \) from \( \mathcal{Z} \), there is \( \rho_\delta > 0 \) such that
\[
  \|R_\infty - R(x)\| \geq \rho_\delta \quad \text{for } x \text{ outside } \mathcal{Z}_\delta.
\]
Let \( t_\delta \) be such that, for all \( x \in \mathcal{Z}, \ h > 0 \) and \( t \geq t_\delta \) we have
\[
  0 \leq v(t, x) - v(t + h, x) \leq \delta.
\]
Such a \( t_\delta \) exists because, even though the convergence of \( v \) on \( \mathcal{Z} \) seems to be only pointwise, the uniform gradient bounds for \( v \) imply that it is in fact uniform. And this entails, for a universal constant \( C > 0 \):
\[
  v(t, x) - v(t + h, x) \leq C\delta \quad \text{for } t \geq t_\delta \text{ and } x \in \mathcal{Z}_\delta.
\]
Our task is now to extend this inequality outside \( \mathcal{Z}_\delta \). For that, let us set
\[
  v_\delta(t, x) = v(t + h, x) - C\delta - \|v(t_\delta, \cdot)\|_\infty e^{-\rho_\delta(t - t_\delta)}, \quad t \geq t_\delta, \quad x \notin \mathcal{Z}_\delta.
\]
We are going to prove that it is a sub-solution to (2.8.46) for \( t \geq t_\delta \), and \( x \) outside \( Z_\delta \), that is additionally less than \( v \) for \((t, x) \in [t_\delta, +\infty) \times Z_\delta\), and \( t = t_\delta, \ x \in T^N\). This will imply \( v(t, x) \geq v_\delta(t, x) \) for \( t \geq t_\delta \) and \( x \) outside \( Z_\delta \), which entails in turn
\[
v(t, x) \geq v(t + h) - C(\delta + e^{-\rho_\delta(t-t_\delta)})
\]
for \( t \geq t_\delta, \ x \in T^n \) and \( h > 0 \). This implies the pointwise convergence of \( v(t, x) \) to \( \liminf_{t \to +\infty} v(t, x) \)
and, consequently, its uniform convergence from the gradient bound. So, to finish the proof, it remains to prove the above inequalities.

It is obvious that \( v(t, x) \geq v_\delta(t, x) \) for \( t \geq t_\delta \) and \( x \in Z_\delta \). At time \( t = t_\delta \) we have
\[
v(t_\delta, x) - v_\delta(t_\delta, x) = v(t_\delta, x) + \|v(t_\delta, .)\|_\infty + C\delta - v(t + h, x) \geq +C\delta - v(t + h, x) \geq 0.
\]

So, the last effort consists in proving the sub-solution property for \( v_\delta \), let \( \varphi \) be our favorite test function and \((t_1, x_1)\) a maximum point of \( v_\delta - \varphi \). Then, \((t_1, x_1)\) is a maximum point of
\[
(t, x) \mapsto v(t, x + h) - \left( \varphi + C\delta + \|v(t_\delta, .)\|_\infty e^{-\rho_\delta(t-t_\delta)} \right).
\]
Set \( \tau_h v(t, x) = v(t + h, x) \), we have the following inequality, taken at \((t_1, x_1)\):
\[
\varphi_t - \|v(t_\delta, .)\|\rho_\delta + R(x_1) \frac{\|\nabla \varphi\|^2}{\|\tau_h v\| + \sqrt{(\tau_h v)^2 + \|\nabla \varphi\|^2}} + (\|R\|_\infty - R(x_1))\tau_h v \leq 0.
\]

However we have \( \tau_h v(t, x) \geq v_\delta(t, x) \) on \([t_\delta, +\infty) \times T^n\), hence \( |v_\delta(t, x)| \geq |v(t, x)| \). Also, recall that \( \|R\|_\infty - R \geq \rho_\delta \) outside \( Z_\delta \). This brings us to
\[
\varphi_t + R(x_1) \frac{\|\nabla \varphi\|^2}{\|v_\delta\| + \sqrt{(v_\delta)^2 + \|\nabla \varphi\|^2}} + (\|R\|_\infty - R(x_1))v_\delta \leq 0,
\]
at \((t_1, x_1)\). This is the requested inequality for \( v_\delta \), which finishes the proof of the proposition. \( \square \)

**Exercise 2.8.14** Carry out the same analysis to the equation
\[
u_t + \|\nabla u\| = f(x), \quad t > 0, \ x \in T^N,
\]
where \( f \in C(T^n) \) satisfies the usual assumptions: continuous, nonnegative, with a nontrivial zero set.

**Exercise 2.8.15** This exercise shows, in the strictly convex case, that the Lax-Oleinik formula yields an (almost) explicit form of the asymptotic limit of the solution of the Cauchy Problem. Let us deal here with the equation
\[
u_t + \|\nabla u\|^2 = f(x), \quad t > 0, \ x \in T^N, u(0, x) = u_0(x) \quad (2.8.47)
\]
with \( f \) continuous, vanishing on a finite set \( \{x_1, ..., x_N\} \).
1. Recall the expression of the Lax-Oleinik formula for the solution \( u(t, x) \) of the above Cauchy Problem.

2. For a given \( y \in \mathbb{T}^n \), let us define
\[
S(t, y, x) = \inf_{\gamma(t)=x, \gamma(0)=y} \int_0^t \left( \frac{|\dot{\gamma}|^2}{4} + f(\gamma(s)) \right) ds.
\]
- Show that \((t, x) \mapsto S(t, y, x)\) solves the time-dependent Hamilton-Jacobi equation on \( \mathbb{T}^n \setminus y \).
- Show that \((t, x) \mapsto S(t, y, x)\) has a large time limit.
- If \( y \) is one of the \( x_i \)'s, then this large time limit is the function \( S(x_i, x) \) defined by (2.8.39) (thus, in particular, a steady solution of (2.8.47) on the whole \( T^n \)).

3. Show that
\[
\lim_{t \to +\infty} u(t, x) = \inf_{i \in \{1, \ldots, N\}} \left( u_0(x_i) + S(x_i, x) \right).
\]

**Exercise 2.8.16** Let \( H(x, p) = R(x) \sqrt{1 + |p|^2} - \|R\|_{\infty} \).

1. Compute its Lagrangian \( L(x, v) \). Is it defined everywhere? What is the reason for that?
2. Propose a Lax-Oleinik type formula for the solution of (2.8.45) and reprove Proposition 2.8.13.

**Exercise 2.8.17** Consider the solution \( u(t, x) \) of (??).

1. Show that \( u(t, x) \) converges to a steady solution.
2. Show that \( u(t, x) \) has a unique maximum \( \bar{x}(t) \).
3. Prove that \( \bar{x} \) solves the differential system
\[
\dot{x} = -D^2 u(t, \bar{x}(t)) \cdot \nabla R(\bar{x}(t)).
\]
4. Study the convergence rate between \( \bar{x} \) and its large time limit.

If you do not know how to start, you may consult [102].

### 2.9 Convergence of the Lax-Oleinik semigroup

In this last section, we prove that the solutions of
\[
u_t + H(x, \nabla u) = \begin{cases} 0 & x \in \mathbb{T}^n \bigl(0, x\bigr) = u_0(x) \\
\end{cases}
\]
will converge to a wave solution, under the (by now) usual assumptions of uniform strict convexity. So far, we have seen a very particular mechanism for convergence: the dynamics on a special set (namely, the zero set of \( f \), or, equivalently, the set where \( R \) attains its maximum) dictates in turn the convergence in the area where the equation is coercive. The next exercise shows a seemingly different mechanism.
Exercise 2.9.1 Let $c$ be a positive real number, and $f$ a smooth function on $\mathbb{T}^1$. Let us consider the Cauchy Problem

$$u_t + cu_x + u_x^2 = f(x), \quad t > 0, \quad x \in \mathbb{T}^1$$

$$u(0, x) = u_0(x)$$

(2.9.2)

We choose $c > 2\sqrt{\|f\|_{\infty}}$.

1. Show directly the existence of $\lambda \in [-\|f\|_{\infty}, \|f\|_{\infty}]$ such that the problem has a wave solution $\phi(x)$ of speed $\lambda$.

2. Show that $\phi$ is unique (up to the addition of a constant) and smooth.

3. For $x_0 \in \mathbb{T}^N$, let $\gamma(t)$ be the solution of

$$\dot{\gamma} = c + 2\phi'(\gamma), \quad \gamma(0) = x_0.$$  

(2.9.3)

Show that $\gamma$ visits each point of $\mathbb{T}^1$ an infinite number of times.

4. Set $v = u - \phi$. Show that we have, in the viscosity sense:

$$\partial_t v + (c + 2\phi')v_x \leq 0.$$ 

5. Conclude that there exists a constant $k(u_0)$ such that

$$\lim_{t \to +\infty} \|u(t,.) - \phi - k(u_0)\|_{\infty} = 0.$$ 

In fact, these two mechanisms are two different faces of the same coin. Indeed, (2.9.3) is nothing else than the equation of a characteristic curve associated to the steady solutions. Its $\omega$-limit set is the whole circle. In the case of steady solutions to (2.8.30), we rather talk about extremals but they still satisfy the characteristic equation, and we see that their $\omega$-limit sets is included in the zero set of $f$ which is, precisely, the set that organizes the whole convergence phenomenon.

Exercise 2.9.2 When (2.9.1) is posed on the one-dimensional circle $\mathbb{T}^1$, with a uniformly strictly convex Hamiltonian, show that only the above two mechanisms are present. In the second case, explain why the wave solutions are smooth and unique up to a constant.

It turns out that this is a general fact. For a general Hamilton-Jacobi equation of the type (2.9.1), there is indeed a set where all the extremals associated to the wave solutions accumulate, and which indeed orchestrate the convergence to a steady set. Several sets could be good candidates for it: Aubry set, Mather set, Aubry-Mather set... they were all identified by Fathi and the reader who wishes to learn more may consult [61] or [62], where their general theory is exposed. In this section, our scope is more limited: we want to identify a set where, following the ideas of the preceding section, the dynamics of $u$ will propagate to the whole torus. First, let us set the theorem that we want to prove. Its first proof is due to Fathi [61]. The proof provided here is inspired from [125].
Theorem 2.9.3 Let $H$ be uniformly strictly convex in $p$:

$$\alpha I \leq D^2_p H(x, p), \text{ in the sense of quadratic forms.}$$

For $u_0 \in C(T^n)$, the solution $u(t, x)$ of (2.9.1) converges to a wave solution: if $c$ is the wave speed, there exists a solution $u_\infty(x)$ of

$$H(x, \nabla u) = c, \quad x \in T^n, \quad (2.9.4)$$

such that

$$\lim_{t \to \infty} \| u(t, .) + ct - u_\infty \|_\infty = 0.$$ 

Let us, once and for all, assume that $c = 0$, even if it means replacing $H$ by $H - c$. Our goal is to recycle the main idea of Proposition 2.8.13, namely, to find a set, similar to $Z$, where convergence will hold because of some monotonicity property. It turns out that the following easy remark can be made: let $\phi$ be a steady solution, and $\gamma : [0, t] \to T^n$ an extremal path of $L$ calibrating $\phi$. For all $0 \leq s \leq s' \leq t$ we have

$$\phi(\gamma(s')) = \phi(\gamma(s)) + \int_s^{s'} L(\gamma, \dot{\gamma}) d\sigma, \quad (2.9.5)$$

whereas, by definition of the Lax-Oleinik semigroup we have

$$u(s', \gamma(s')) \leq u(s, \gamma(s)) + \int_s^{s'} L(\gamma, \dot{\gamma}) d\sigma. \quad (2.9.6)$$

Subtracting, we obtain

$$u(s', \gamma(s')) - \phi(\gamma(s')) \leq u(s, \gamma(s)) - \phi(\gamma(s)); \quad (2.9.7)$$

thus $u - \phi$ is nonincreasing along the extremal path. One may wonder whether this innocent observation can be of any use, and the remarkable fact is that it is useful indeed. For this we need to introduce a few more objects. Let $\omega(u_0)$ be the omega-limit set of $u_0$ with respect to $T$, namely the set of all functions $\psi \in C(T^n)$ for which there is a sequence $(t_n)_{n}$ going to $+\infty$ such that $(TT(t_n)u_0)_n$ converges to $\psi$ in the uniform norm. Such a function generates an entire orbit of $T$, in other words $T(t)\psi$ is defined for all real $t$.

Let $\phi$ be a steady solution, that will be chosen once and for all. Let us define $Z_\phi$ as the collection of all extremal paths calibrated by $\phi$ on $\mathbb{R}$, that is, the collection of all trajectories $\gamma : \mathbb{R} \to T^n$ such that (2.9.5) holds for all $-\infty < s \leq s' < +\infty$. Let $Z$ be the collection of all the limiting trajectories of the extremal paths of $M_0$, that is: if $\gamma \in M_\phi$ and $(s_n)_n$ is a sequence such that $(\gamma(s_n + .))_n$ converges to $\gamma_\infty$ uniformly on every bounded interval of $\mathbb{R}$ - such a sequence exists by virtue of Ascoli’s theorem - then the limit is in $Z$. If $\gamma_\infty$ is such a uniform limit, then, trivially, $\gamma_\infty$ calibrates $\phi$ on $\mathbb{R}$. But the additional information is that it calibrates every other steady solution on $\mathbb{R}$: this simply comes from (2.9.7), where $u$ is replaced by any steady solution.

The reader that has accepted to follow us up to this point has certainly guessed what will happen: the set $Z$ will have the same role as the zero set of $f$ in the equation $u_t + |\nabla u|^2 = f$, namely we have the
**Proposition 2.9.4** Let \( v(t, x) \) be an entire orbit of \( \mathcal{T} \), namely: \( v(t, .) = \mathcal{T}(t - s)v(s, .) \) for all \(-\infty < s \leq t < +\infty\), such that \( v(0, .) \in \omega(u_0) \). Then \( v \) does not depend on \( t \) on \( \mathcal{Z} \).

**Proof.** Choose \( x \in \mathcal{Z} \). We claim that it is sufficient to show that, if \( \gamma_\infty \) is an extremal path of \( \mathcal{Z} \), with \( \gamma_\infty(t_0) = x \), then the function \( t \mapsto v(t, \gamma_\infty(t)) - \phi(\gamma_\infty(t)) \) is constant. Indeed, \( \gamma_\infty \) is calibrated by \( v \) on \( \mathbb{R} \), so that \( v \) is \( C^{1,1} \) on \( \gamma_\infty \) and we have, by Remark 2.7:

\[
\nabla v(t, \gamma_\infty(t)) = \nabla vL(\gamma_\infty(t), \dot{\gamma}_\infty(t)),
\]

for all \( t \in \mathbb{R} \). But, since \( \phi \) also calibrates \( \gamma_\infty \), we also have

\[
\nabla \phi(\gamma(t)) = \nabla vL(\gamma_\infty(t), \dot{\gamma}_\infty(t)).
\]

And so, we have \( H(\gamma(t), \gamma_\infty(t)) = 0 \), which entails, specializing at \( t = t_0 \):

\[
\partial_tv(t_0, x) = 0.
\]

This implies the proposition.

It remains to prove the claim. By definition of \( \gamma_\infty \), there is a global extremal path \( \gamma \) calibrated by \( \phi \), and a sequence \((\tau_n)_n \) such that

\[
\gamma_\infty(\sigma) = \lim_{n \to +\infty} \gamma(\tau_n + \sigma),
\]

uniformly in every compact in \( \sigma \in \mathbb{R} \). Choose a sequence \((t_n)_n \) going to \(+\infty\) such that

\[
v(s, x) = \lim_{n \to +\infty} \mathcal{T}(t_n + \tau_n + s)u_0(x),
\]

uniformly in every compact in \( s \in \mathbb{R} \). Even if it means removing some terms of the sequence, we may assume the existence of \( \psi_1 \in \omega(u_0) \) such that

\[
\psi_1 = \lim_{n \to +\infty} \mathcal{T}(\tau_n)u_0,
\]

in the uniform topology. Let us define, for all \( \sigma \in \mathbb{R} \):

\[
l(\sigma) = \lim_{s \to +\infty} \left( \mathcal{T}(\sigma + s + t)u_0(\gamma(s + t)) - \phi(\gamma(s + t)) \right),
\]

by the monotonicity property along the extremals it is independent of \( t \). We may assume, without loss of generality, that the sequence \((l(t_n))_n \) has a limit \( l \). By the weak contraction property, we have

\[
\|\mathcal{T}(t + t_n + \tau_n)u_0 - \mathcal{T}(t + t_n)\psi_1\|_\infty \leq \|\mathcal{T}(\tau_n)u_0 - \psi_1\|_\infty,
\]

so that we have

\[
\mathcal{T}(t + t_n)\psi_1(\gamma(t + t_n)) - \phi(\gamma(t + t_n)) = l_n.
\]

Sending \( n \) to \(+\infty\) yields

\[
\mathcal{T}(t)\psi(\gamma_\infty(t)) - \phi(\gamma_\infty(t)) = l,
\]

which is the sought for result. □

To show the convergence, we will show that any orbit in \( \omega(u_0) \) do not depend on time, the weak contraction property will imply that \( u(t, .) \) converges uniformly, as \( t \to +\infty \), to a time independent function.
Exercise 2.9.5 Show that the limit is a viscosity solution of (2.9.4). Hint: this is not a joke, there is actually an argument (not so difficult, though) to write down.

So, let us pick \( x \in \mathbb{T}^n \) and let us prove that \( v(t, x) \) does not depend on \( t \). For all \( s > 0 \) we have

\[
u(t, x) = \inf_{\gamma(-s-t) = x} \left(u(\gamma(-s-t)) + \int_{-s-t}^{0} L(\gamma, \dot{\gamma}) d\sigma\right);
\]

if \( \gamma_s \) is an extremal realizing the above infimum, we claim the existence of two sequences \( (s_n)_n \) and \( \tau_n > \), going to \( +\infty \), such such that

\[
\lim_{s \to +\infty} (s_n - \tau_n) = +\infty, \quad \lim_{n \to +\infty} \text{dist}(\gamma_{s_n}(-\tau_n), \mathcal{Z}) = 0.
\]

This is elementary analysis: there is a sequence \( (s_n)_n \) going to \( +\infty \) such that \( (\gamma_{s_n}(-\tau_n))_n \) converges locally uniformly to an entire extremal path \( \gamma \), which is calibrated by \( v \) on \( \mathbb{R} \). For every \( \varepsilon > 0 \) there is, by definition of \( \mathcal{Z} \), a positive \( t_\varepsilon \) such that

\[
\text{dist}(\gamma_{t_\varepsilon}, \mathcal{Z}) \leq \varepsilon,
\]

so that we have, for \( n = n_\varepsilon \) large enough:

\[
\text{dist}(\gamma_{s_n}(-\frac{s_n}{2} + t_\varepsilon), \mathcal{Z}) \leq 2\varepsilon.
\]

Taking successively \( \varepsilon = 1, \frac{1}{2}, \frac{1}{3}, ..., \frac{1}{p}, ... \) yields the desired sequence. Coming back to \( v \) we have

\[
v(t, x) = v(-\tau_n, \gamma_{s_n}(-\tau_n)) + \int_{-\tau_n}^{0} L(\gamma_{s_n}, \dot{\gamma}_{s_n}) d\sigma.
\]

Even if it means removing some terms from the sequence \( (\tau_n)_n \), we may assume the existence of \( z \in \mathcal{Z} \) such that

\[
\lim_{n \to +\infty} \gamma_{s_n}(-\tau_n) = z.
\]

So we have

\[
v(t, x) = v(-\tau_n, z) + \int_{-\tau_n}^{0} L(\gamma_{s_n}, \dot{\gamma}_{s_n}) d\sigma + o_{n \to +\infty}(1). \quad (2.9.8)
\]

On the other hand, we have, by application of the semigroup property:

\[
v(t, x) \leq \inf_{z \in \mathcal{Z}} \left(v(-\tau_n, z) + \inf_{\gamma(-\tau_n) = z} \int_{-\tau_n}^{0} L(\gamma_{s_n}, \dot{\gamma}_{s_n}) d\sigma\right)
\]

\[
= \inf_{z \in \mathcal{Z}} \left(\psi(z) + \inf_{\gamma(-\tau_n) = z} \int_{-\tau_n}^{0} L(\gamma_{s_n}, \dot{\gamma}_{s_n}) d\sigma\right),
\]

and the last line is true because, for all \( z \in \mathcal{Z} \), we have \( v(-\tau_n, z) = \psi(z) \) by Proposition 2.9.4. So, in the end, we have

\[
v(t, x) \leq \inf_{z \in \mathcal{Z}} \left(\psi(z) + \liminf_{n \to +\infty} \inf_{\gamma(-\tau_n) = z} \int_{-\tau_n}^{0} L(\gamma_{s_n}, \dot{\gamma}_{s_n}) d\sigma\right),
\]

a quantity that does not depend on \( t \) anymore. However, (2.9.8) shows that the inequality is actually an equality: the theorem is proved.
Remark 2.9.6 What we have done here may seem rather elaborate and, perhaps, far-fetched. In fact it is quite natural. Indeed, consider a nice and simple transport equation

\[ u_t + b(x) \nabla u = 0, \quad t > 0, \; x \in \mathbb{R}^2, \]

and remember that the values of \( u(t, x) \) are determined from its characteristic curves \( \dot{\gamma} = b(\gamma) \). One reads the asymptotic values of the solutions where the characteristics accumulate. We see quite a similar phenomenon here.

We can actually write down an explicit expression of the limiting solution, that we still call \( \psi(x) \). Notice indeed that, in the course of our proof, we have encountered an old friend from the last section, namely the quantity

\[ S(y, x) = \liminf_{t \to +\infty} S_t(y, x), \quad S_t(y, x) = \left( \inf_{\gamma(-t) = y} \int_{-t}^0 L(\gamma, \dot{\gamma})ds \right), \tag{2.9.9} \]

the infimum being of course taken on all the paths \( \gamma \) such that \( \gamma(0) = x \).

Exercise 2.9.7 Consider the function \( (t, y) \mapsto S_t(y, x) \).

- Show that \( \liminf_{t \to +\infty} S_t(y, x) \) is in fact a limit.
- Exercise 2.8.15 gives, for the problem

\[ u_t + |\nabla u|^2 = f(x), \]

an explicit form of the limiting steady solution to the Cauchy Problem. Use its ideas to prove that, if \( \psi(x) \) is the limit of \( T u_0(x) \), we have

\[ \psi = \inf_{y \in \mathbb{Z}} \left( u_0(y) + S(y, x) \right). \]

With the results of this section, the tour ends here. We hope that the reader has enjoyed it. One could say much more on the organization of the steady solutions, and the reader should consult the book [62]. They would be, however, outside the scope of this book. Let us just notice that the results of the present section provide a complete parallel with the large time behavior of the solutions of viscous Hamilton-Jacobi equations, which was the goal we wanted to achieve: the viscosity solutions of the inviscid problem still converge to waves, although their organization, that we have largely uncovered, is much more complicated.
Bibliography


[38] P. Constantin and C. Foias, Navier-Stokes Equations, University of Chicago Press, 1988


[54] H. Dong, D. Du and D. Li, Finite time singularities and global well-posedness for fractal Burgers’ equation, Indiana Univ. Math. J., 58 (2009), 807-.


179


[120] G. Polya, On the zeros of an integral function represented by Fourier’s integral, Messenger of Math. 52, 1923, 185-188


