Chapter 2
An Introduction to Fully Nonlinear Parabolic Equations

Cyril Imbert and Luis Silvestre

Abstract These notes contain a short exposition of selected results about parabolic equations: Schauder estimates for linear parabolic equations with Hölder coefficients, some existence, uniqueness and regularity results for viscosity solutions of fully nonlinear parabolic equations (including degenerate ones), the Harnack inequality for fully nonlinear uniformly parabolic equations.

2.1 Introduction

The literature about parabolic equations is immense and it is very difficult to have a complete picture of available results. Very nice books such as [LSU67, Kryl87, Dong91,Lieb96] are attempt to gather and order the most significant advances in this wide field. If now one restricts himself to fully nonlinear parabolic equations, the task is still almost impossible. Indeed, many results proved for parabolic equations were first proved for elliptic equations and these results are numerous. We recall that many problems come from geometry; the reader is referred to the survey paper [Kryl97] where Krylov gives historical and bibliographical landmarks.

In these notes, we will focus on three specific topics concerning parabolic equations: Schauder estimates for linear parabolic equations (following Safonov [Saf84] and the textbook by Krylov [Kryl96]), viscosity solutions for fully nonlinear parabolic equations (see e.g. [CIL92]) and the Harnack inequality for fully nonlinear uniformly parabolic equations.

C. Imbert (✉)

CNRS, UMR8050, Université Paris-Est Créteil Val-de-Marne, Centre de mathématiques, UFR sciences et technologies, 61 avenue du Général de Gaulle, 94010 Créteil cedex, France
e-mail: cyril.imbert@u-pec.fr

L. Silvestre
Department of Mathematics, University of Chicago, Chicago, IL 60637, USA

2.1.1   Main Objects and Notation

Geometric Objects

We first consider a connected open bounded set $\Omega \subset \mathbb{R}^d$. We refer to such a set as a *domain*. A domain is $C^{2,\alpha}$ if, locally, the boundary of the domain can be represented as the graph of a function with two derivatives that are α-Hölder continuous.

Parabolic equations are considered in cylindrical domain of the form $(0, T) \times \Omega$. The *parabolic boundary* of $(0, T) \times \Omega$ is denoted by $\partial_p (0, T) \times \Omega$; we recall that it is defined as follows

$$\partial_p (0, T) \times \Omega = \{0\} \times \Omega \cup (0, T) \times \partial \Omega.$$

The open ball of $\mathbb{R}^d$ centered at x of radius ρ is denoted by $B_\rho(x)$. If $x = 0$, we simply write B_ρ. The following elementary cylindrical domains play a central role in the theory: for all $\rho > 0$ and $x \in \mathbb{R}^d$, we define

$$Q_\rho(t, x) = (t - \rho^2, t) \times B_\rho(x).$$

When we write Q_ρ, we mean $Q_\rho(0, 0)$. It is also convenient to write

$$Q_\rho(t, x) = (t, x) + Q_\rho$$

and

$$Q_\rho = \rho Q_1.$$

A Linear Operator

The general parabolic equation considered in Sect. 2.2 involves the following linear operator

$$Lu = \sum_{i,j} a_{ij}(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_i b_i(t, x) \frac{\partial u}{\partial x_i} + c(t, x)u.$$

The set of $d \times d$ real symmetric matrices is denoted by $\mathbb{S}_d$. The identity matrix is denoted by I. For $A, B \in \mathbb{S}_d$, $A \geq B$ means that all the eigenvalues of $A - B$ are non-negative.

Unknown functions $u : (0, T) \times \Omega \to \mathbb{R}$ depend on two (set of) variables: $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$. It is convenient to use a capital letter X to refer to $(t, x) \in \mathbb{R}^{d+1}$.

The time derivative of u is either denoted by $\frac{\partial u}{\partial t}$ or $\partial_t u$ or u_t. Du denotes the gradient of the function u with respect to the space variable x. $D^2 u$ denotes the Hessian matrix of the function u with respect to x.

The linear operator introduced above can be written as follows

$$Lu = \mathrm{trace}(AD^2 u) + b \cdot Du + cu$$

where $A = (a_{ij})_{ij}$.

Hölder Spaces and Semi-norms

We say that $u \in C^{0,\alpha}(Q)$ for $Q \subset (0, T) \times \Omega$ if u is $\frac{\alpha}{2}$-Hölder continuous with respect to time t and α-Hölder continuous with respect to space x. The corresponding semi-norm is denoted by $[u]_{\alpha,Q}$. See Sect. 2.1.4 for details.

2.1.2 Fully Nonlinear Parabolic Equations

We first emphasize the fact that we will not consider systems of parabolic equations; in other words, we will focus on scalar parabolic equations. This means that the unknown function u will always be real valued. We also restrict ourselves to second order parabolic equations.

We consider parabolic equations posed in a domain $\Omega \subset \mathbb{R}^d$; hence, unknown functions u are defined in $(0, T) \times \Omega$ with $T \in [0, \infty]$. In order to construct solutions and prove uniqueness for instance, initial and boundary conditions should be imposed. However, we will very often not specify them.

Fully nonlinear parabolic equations appear in optimal control theory and geometry. Here are several significant examples.

- The Bellman equation

$$\partial_t u + \sup_{\alpha \in A} \left\{ -\sum_{i,j} a_{ij}^\alpha(x) \frac{\partial^2 u}{\partial x_i \, \partial x_j} + \sum_i b_i^\alpha(x) \frac{\partial u}{\partial x_i} \right\} + \lambda u = 0.$$

- The mean curvature equation

$$\partial_t u = \Delta u = \frac{D^2 u Du \cdot Du}{|Du|^2}.$$

- The parabolic Monge–Ampère equations proposed by Krylov in [Kryl76]

$$-\frac{\partial u}{\partial t}\det(D^2 u) = H^{d+1}$$

$$-\det(D^2 u) + \left[\frac{\partial u}{\partial t} + H\right]^{d+1} = 0 \qquad (2.1)$$

$$-\det\left(D^2 u - \frac{\partial u}{\partial t} I\right) = H^d$$

where $H = H(t, x, Du)$ is a nonlinear first order term.

- For the study of the Kähler–Ricci flow, one would like to study:

$$\frac{\partial u}{\partial t} = \ln(\det(D^2 u)). \qquad (2.2)$$

2.1.3 Aim of These Notes

Our goal is to construct solutions and study their regularity. One would like to construct classical solutions, that is to say solutions such that the derivatives appearing in the equation exist in the classical sense and satisfy the equation. But this is not always possible and it is sometimes (very often?) necessary to construct weak solutions. They are different notions of weak solutions; we will focus in these notes on so-called viscosity solutions. The advantage is that it is easy to construct such solutions. One can next try to prove that these solutions are regular.

Before 1988 (date of publication of [Jens88]), it was popular (necessary) to construct solutions of fully nonlinear elliptic (or parabolic) equations by using the *continuity method*. To apply it, it is necessary to get appropriate a priori estimates (on third derivatives for instance, or on the modulus of continuity of the second ones).

The situation changed dramatically when Jensen [Jens88] managed to apply the viscosity solution techniques of Crandall–Lions [CL81] to second order elliptic and parabolic equations. In particular, he understood how to adapt the so-called doubling variable techniques to prove uniqueness. Ishii also contributed to this major breakthrough. The reader is referred to the survey paper [CIL92] for further details.

Before presenting the viscosity solution techniques and some selected regularity results for these weak solutions, we will present shortly the classical Schauder approach to linear parabolic equations.

2.1.4 Spaces of Hölder Functions

Because we study parabolic equations, Hölder continuity of solutions refers to uniform continuity with respect to

$$\rho(X, Y) = \sqrt{|t - s|} + |x - y|$$

where $X = (t, x)$ and $Y = (s, y)$. In other words, solutions are always twice more regular with respect to the space variable than with respect to the time variable.

Remark 2.1.1 (Important). The reader should keep in mind that, following Krylov [Kryl96], we choose to write $u \in C^{0,\alpha}$ for functions that are α-Hölder continuous in x and $\frac{\alpha}{2}$-Hölder continuous in t. This choice is made first to emphasize the link between regularities with respect to time and space variables, second to simplify notation.

Let $Q \subset (0, T) \times \Omega$ and $\alpha \in (0, 1]$.

- $u \in C^{0,\alpha}(Q)$ means that there exists $C > 0$ s.t. for all $(t, x), (s, y) \in Q$, we have

$$|u(t, x) - u(s, y)| \leq C(|t - s|^{\frac{\alpha}{2}} + |x - y|^{\alpha}).$$

  In other words, u is $\frac{\alpha}{2}$-Hölder continuous in t and α-Hölder continuous in x.
- $u \in C^{1,\alpha}(Q)$ means that u is $\frac{\alpha+1}{2}$-Hölder continuous in t and Du is α-Hölder continuous in x.
- $u \in C^{2,\alpha}(Q)$ means that $\frac{\partial u}{\partial t}$ is $\frac{\alpha}{2}$-Hölder continuous in t and $D^2 u$ is α-Hölder continuous in x.

We also consider the following norms and semi-norms.

$$[u]_{\alpha, Q} = \sup_{X, Y \in Q, X \neq Y} \frac{|u(X) - u(Y)|}{\rho(X, Y)}$$

$$|u|_{0, Q} = \sup_{X \in Q} |u(X)|$$

$$[u]_{2+\alpha, Q} = \left[\frac{\partial u}{\partial t}\right]_{\alpha, Q} + [D^2 u]_{\alpha, Q}$$

$$|u|_{2+\alpha, Q} = |u|_{0, Q} + \left|\frac{\partial u}{\partial t}\right|_{0, Q} + |Du|_{0, Q} + |D^2 u|_{0, Q} + [u]_{2+\alpha, Q}.$$

We will use repeatedly the following elementary proposition.

Proposition 2.1.2.

$$[uv]_{\alpha, Q} \leq |u|_{0, Q}[v]_{\alpha, Q} + |v|_{0, Q}[u]_{\alpha, Q}$$

and for $k = 0, 2$,

$$[u + v]_{k+\alpha,Q} \leq [u]_{k+\alpha,Q} + [v]_{k+\alpha,Q}.$$

The following proposition implies in particular that in order to control the norm $|u|_{2+\alpha,Q}$, it is enough to control $|u|_{0,Q}$ and $[u]_{2+\alpha,Q}$.

Proposition 2.1.3 (Interpolation inequalities). *For all $\varepsilon > 0$, there exists $C(\varepsilon) > 0$ s.t. for all $u \in C^{2,\alpha}$,*

$$\begin{cases} |\frac{\partial u}{\partial t}|_{0,Q} & \leq \varepsilon[u]_{2+\alpha,Q} + C(\varepsilon)|u|_{0,Q}, \\ [Du]_{\alpha,Q} & \leq \varepsilon[u]_{2+\alpha,Q} + C(\varepsilon)|u|_{0,Q}, \\ [u]_{\alpha,Q} & \leq \varepsilon[u]_{2+\alpha,Q} + C(\varepsilon)|u|_{0,Q}. \end{cases} \tag{2.3}$$

The following proposition is a precise parabolic statement of the following elliptic fact: in order to control the Hölder modulus of continuity of the gradient of u, it is enough to make sure that, around each point, the function u can be perturbed linearly so that the oscillation of u in a ball of radius $r > 0$ is of order $r^{1+\alpha}$.

Proposition 2.1.4 (An equivalent semi-norm). *There exists $C \geq 1$ such that for all $u \in C^{2,\alpha}(Q)$,*

$$C^{-1}[u]'_{2+\alpha,Q} \leq [u]_{2+\alpha,Q} \leq C[u]'_{2+\alpha,Q}$$

where

$$[u]'_{2+\alpha,Q} = \sup_{X \in Q} \sup_{\rho > 0} \rho^{-2-\alpha} \inf_{P \in \mathcal{P}_2} |u - P|_{0,Q_\rho(X) \cap Q}$$

where

$$\mathcal{P}_2 = \{\alpha t + p \cdot x + \frac{1}{2}Xx \cdot x + c : \alpha, c \in \mathbb{R}, p \in \mathbb{R}^d, X \in \mathbb{S}_d\}.$$

The reader is referred to [Kryl96] for proofs of the two previous propositions.

2.2 Schauder Estimates for Linear Parabolic Equations

In this first section, we state a fundamental existence and uniqueness result for linear parabolic equations with Hölder continuous coefficients.

The proof of this theorem is rather long and presenting it completely is out of the scope of the present lectures notes. Instead, we would like to focus on two particular aspects: uniqueness and interior estimates.

The uniqueness of the solution is proved by using a maximum principle (Sect. 2.2.3), the existence can be obtained through the *continuity method*. This method relies on the proof of the "good" a priori estimate (2.4) on any $C^{2,\alpha}$ solution. This estimate is global in the sense that it deals with what happens at the interior of $(0, T) \times \Omega$ and at its boundary. In Sect. 2.2.5, we focus on what happens in the interior of the domain. Precisely, we present a complete proof of the interior Schauder estimate in the general case. It relies on Schauder estimates for parabolic equations with constant coefficients. The derivation of these estimates are presented in Sect. 2.2.4 by studying first the heat equation. We present here an argument due to Safonov circa 1984.

2.2.1   Linear Parabolic Equations

The standing example of linear parabolic equations with constant coefficients is the *heat equation*

$$\frac{\partial u}{\partial t} - \Delta u = f$$

where f is a *source term*. The general form of a linear parabolic equation with variable coefficients is the following

$$\frac{\partial u}{\partial t} - \sum_{i,j} a_{ij}(X) \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_i b_i(X) \frac{\partial u}{\partial x_i} - c(X)u = 0$$

where

$$c \leq 0$$

and $A(X) = (a_{ij}(X))_{i,j}$ is a symmetric matrix satisfying one of the following assumptions

- (Degenerate ellipticity) for all X, $A(X) \geq 0$;
- (Strict ellipticity) there exists $\lambda > 0$ s.t. for all X, $1\ A(X) \geq \lambda I$;
- (Uniform ellipticity) there exists $\Lambda \geq \lambda > 0$ s.t. for all X, $\lambda I A(X) \leq \Lambda I$.

We recall that I denotes the identity matrix and if $A, B \in \mathbb{S}_d$, $A \geq B$ means that all the eigenvalues of $A - B$ are non-negative.

It is convenient to consider the linear differential operator L defined as follows

$$Lu = \sum_{i,j} a_{ij}(X) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_i b_i(X) \frac{\partial u}{\partial x_i} + c(X)u.$$

2.2.2 A Fundamental Existence and Uniqueness Result

In this subsection, we state a fundamental existence and uniqueness result for linear parabolic equation with Hölder continuous coefficients. Such a result together with its proof can be found in various forms in several classical monographs such as [LSU67, Kryl96]. We choose here to present the version given in [Kryl96].

In the following statement, $\mathbb{R}^{d+1}_+$ denotes $[0, +\infty) \times \mathbb{R}^d$.

Theorem 2.2.1. *If Ω is a $C^{2,\alpha}$ domain and the coefficients $A, b, c \in C^{\alpha}$ $((0,T) \times \Omega)$ and $f \in C^{\alpha}(\mathbb{R}^{d+1}_+)$, $g \in C^{2+\alpha}((0,T) \times \Omega)$, $h \in C^{2,\alpha}(\mathbb{R}^d)$, and g and h are* compatible *(see Remark 2.2.3 below), then there exists a unique solution $u \in C^{2,\alpha}(Q)$ of*

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{in } (0,T) \times \Omega \\ u = g & \text{on } (0, +\infty) \times \partial\Omega \\ u = h & \text{on } \{0\} \times \bar{\Omega}. \end{cases}$$

In addition,

$$|u|_{2+\alpha,(0,T)\times\Omega} \leq C(|f|_{\alpha,\mathbb{R}^{d+1}_+} + |g|_{2+\alpha,(0,T)\times\Omega} + |h|_{2+\alpha,\mathbb{R}^d}) \qquad (2.4)$$

where $C = C(d, \lambda, K, \alpha, \rho_0, \operatorname{diam}(\Omega))$ and $K = |A|_{\delta,(0,T)\times\Omega} + |b|_{\delta,(0,T)\times\Omega} + |c|_{\delta,(0,T)\times\Omega}$ and ρ_0 is related to the $C^{2,\alpha}$ regularity of the boundary of Ω.

Remark 2.2.2. The inequality (2.4) is called the (global) Schauder a priori estimate.

Remark 2.2.3. The fact that data g and h are *compatible* has to do with conditions ensuring that a solution which is regular up to the boundary can be constructed. Since we will not address these problems, we refer the interested reader to [LSU67, Kryl96] for a precise definition.

2.2.3 Maximum and Comparison Principles

Maximum principles are powerful tools to study elliptic and parabolic equations. There are numerous statements which are not equivalent. We choose the following one.

Theorem 2.2.4 (Maximum principle). *Consider a bounded continuous function $u : (0,T) \times \Omega \to \mathbb{R}$ such that $\frac{\partial u}{\partial t}$ exists at each point of $(0,T) \times \Omega$ and Du, D^2u exist and are continuous in $(0,T) \times \Omega$.*

If

$$\frac{\partial u}{\partial t} - Lu \leq 0 \text{ in } (0, T) \times \Omega$$

$$u \leq 0 \text{ on } \partial_p (0, T) \times \Omega$$

then $u \leq 0$ in $(0, T) \times \Omega$.

Remark 2.2.5. The set $\partial_p (0, T) \times \Omega$ is the parabolic boundary of the cylindrical domain $(0, T) \times \Omega$. Its definition is recalled in the section devoted to notation.

Proof. Fix $\gamma > 0$ and consider the function $v(t, x) = u(t, x) - \frac{\gamma}{T-t}$. Assume that v is not non-positive. Then its maximum M on $(0, T) \times \Omega$ is positive. It is reached, and it cannot be attained for $t = 0$ or $x \in \partial\Omega$ since $v \leq u \leq 0$ on $\partial_p (0, T) \times \Omega$. It can neither be attained for $t = T$ since $v \to -\infty$ as $t \to T-$. We conclude that the maximum is attained for some $t \in (0, T)$ and $x \in \Omega$. In particular,

$$0 = \frac{\partial v}{\partial t}(t, x) = \frac{\partial u}{\partial t}(t, x) - \frac{\gamma}{(T - t)^2}$$

$$0 = Dv(t, x) = Du(t, x)$$

$$0 \geq D^2 v(t, x) = D^2 u(t, x).$$

Remark that since A is (uniformly) elliptic, the linear operator satisfies

$$Lu(t, x) = \text{trace}(AD^2 u) + b \cdot Du + cu = \text{trace}(AD^2 u) + cu \leq \text{trace}(AD^2 u) \leq 0$$

since $u(t, x) \geq v(t, x) > 0$, $c \leq 0$, $A \geq 0$ and $D^2 u(t, x) \leq 0$. We now use the fact that u satisfies $\frac{\partial u}{\partial t} - Lu \leq 0$ in $(0, T) \times \Omega$ to get the desired contradiction:

$$\frac{\gamma}{(T - t)^2} = \frac{\partial u}{\partial t}(t, x) \leq Lu(t, x) \leq 0.$$

Since γ is arbitrary, the proof is complete.   □

We now state two corollaries. The first one will be the starting point of the second section (Sect. 2.3). In the framework of linear equation, it is a direct consequence of the previous result.

Corollary 2.2.6 (Comparison principle I). *Consider two bounded continuous functions u and v which are differentiable with respect to time and such that first and second derivatives with respect to space are continuous. If*

$$\frac{\partial u}{\partial t} - Lu \leq f \ in \ (0, T) \times \Omega \qquad\qquad (2.5)$$

$$\frac{\partial v}{\partial t} - Lv \geq f \ in \ (0, T) \times \Omega$$

and $u \leq v$ in $\partial_p Q$, then $u \leq v$ in $(0, T) \times \Omega$.

Remark 2.2.7. Remark that this corollary implies that as soon as u satisfies (2.5), it lies below any solution of $\frac{\partial u}{\partial t} - Lu = f$. This is the reason why it is referred to as a subsolution of the equation $\frac{\partial u}{\partial t} - Lu = f$. In the same way, v lies above any solution and is referred to as a supersolution.

Remark 2.2.8. In view of the previous remark, we can reformulate the result of the previous corollary as follows: if a subsolution lies below a supersolution at the parabolic boundary then it lies below in the whole cylindrical domain.

The next result contains a first estimate for solutions of linear parabolic equations.

Corollary 2.2.9 (A first estimate). *Consider a bounded continuous solution u of $\frac{\partial u}{\partial t} - Lu = f$ in $(0, T) \times \Omega$. Assume moreover that it is differentiable with respect to time and continuously twice differentiable with respect to space. Then*

$$|u|_{0,(0,T)\times\Omega} \leq T |f|_{0,(0,T)\times\Omega} + |g|_{0,\partial_p(0,T)\times\Omega}.$$

Sketch of proof. Consider $v^{\pm} = u \pm (|g|_{0,\partial_p(0,T)\times\Omega} + t |f|_{0,(0,T)\times\Omega})$ and check that v^+ is a supersolution and v^- is a subsolution. Then the previous corollary yields the desired result.   □

2.2.4   Schauder Estimate for the Heat Equation

2.2.4.1   Statement and Corollary

The "interior" Schauder estimate for the heat equation takes the following form.

Theorem 2.2.10. *Let $\alpha \in (0, 1)$ and consider a C^{∞} function $u : \mathbb{R}^{d+1} \to \mathbb{R}$ with compact support and define $f = \frac{\partial u}{\partial t} - \Delta u$. Then there exists a constant $C > 0$ only depending on dimension and α such that*

$$[u]_{2+\alpha,\mathbb{R}^{d+1}} \leq C [f]_{\alpha,\mathbb{R}^{d+1}}.$$

It is then easy to derive a similar "interior" Schauder estimate for linear uniformly parabolic equation with constant coefficients and no lower order term.

Corollary 2.2.11. *Let $\alpha \in (0, 1)$ and assume that $A \equiv A_0$ in $\mathbb{R}^{d+1}$ and $b \equiv 0$, $c \equiv 0$. Then there exists a constant $C > 0$ only depending on dimension and α such that for any C^∞ function u with compact support*

$$[u]_{2+\alpha, \mathbb{R}^{d+1}} \leq C [f]_{\alpha, \mathbb{R}^{d+1}}$$

where $f = \frac{\partial u}{\partial t} - Lu$.

Sketch of proof. The proof consists in performing an appropriate change of coordinates. Precisely, we choose $P \in \mathbb{S}_d$ such that $A_0 = P^2$ and consider $v(t, x) = u(t, Px)$. Then check that $\Delta v = \text{trace}(A_0 D^2 u) = Lu$ and use Theorem 2.2.10. ∎

2.2.4.2 Two Useful Facts

Before proving Theorem 2.2.10, we recall two facts about the heat equation. We recall first that a solution $u \in C^\infty$ of

$$\frac{\partial u}{\partial t} - \Delta u = f,$$

with compact support included in $(0, +\infty) \times \mathbb{R}^d$, can be represented as

$$u(t, x) = \int_0^t \int_{\mathbb{R}^d} G(s, y) f(t - s, x - y) ds \, dy$$

where

$$G(t, x) = \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|x|^2}{4t}}.$$

We write in short hand

$$u = G \star f,$$

keeping in mind that G should be extended by 0 for $t < 0$ in order to make this rigorous. This formula can be justified using Fourier analysis for instance.

Fact 1. *For any $0 \leq \rho \leq R$,*

$$|G \star \mathbf{1}_{Q_R(Z_0)}|_{0, Q_\rho(Z_0)} \leq CR^2$$

where $\mathbf{1}_{Q_R(Z_0)}(Z) = 1$ if $Z \in Q_R(Z_0)$ and 0 if not.

Fact 2. *There exists a constant $C > 0$ such that any solution of $\frac{\partial h}{\partial t} = \Delta h$ in $Q_R(0)$ satisfies*

$$\left| \frac{\partial^n}{\partial t^n} D^\alpha h(0) \right| \le C \frac{|h|_{0,Q_R(0)}}{R^{2n+|\alpha|}}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| = \sum_i \alpha_i$ and $D^\alpha h = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \ldots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}} h$.

This second fact can be proved by using Bernstein's techniques. See [Kryl96, Chap. 8, p. 116].

2.2.4.3 Proof of the Schauder Estimate

The following proof is due to Safonov circa 1984. It is presented in [Kryl96]. Krylov says in [Kryl97] that "[he] believes this proof should be part of a general knowledge for mathematicians even remotely concerned with the theory of PDEs".

Recall that the $C^{2,\alpha}$ regularity can be established "pointwise". Indeed, in view of Proposition 2.1.4, it is enough to be able to find a polynomial P which is linear in time and quadratic in space such that the oscillation of the difference between u and P decreases as $\rho^{2+\alpha}$ in a box of size ρ. The natural candidate for P is the "second order" Taylor polynomial of the function itself. The idea of Safonov is to perturb this natural candidate in order to reduce to the case where $f \equiv 0$.

Proof of Theorem 2.2.10. Without loss of generality, we can assume that the compact support of u is included in $(0, +\infty) \times \mathbb{R}^d$.

Take $X_0 \in \mathbb{R}^{d+1}$, $\rho > 0$ and $K \ge 1$ to be specified later. Let Q denote $Q_{(K+1)\rho}(X_0)$ and take $\zeta \in C^\infty(\mathbb{R}^{d+1})$ with compact support and such that $\zeta \equiv 1$ in Q.

We consider the "second order" Taylor polynomial associated with a function w at a point $X = (t, x)$

$$T_X w(s, y) = w(X) + w_t(X)(s-t) + Dw(X) \cdot (y-s) + \frac{1}{2} D^2 w(X)(y-x) \cdot (y-x).$$

We now consider

$$g = (\zeta T_{X_0} u)_t - \Delta(\zeta T_{X_0} u).$$

In view of properties of ζ,

$$g \equiv f(X_0) \text{ in } Q.$$

Keeping this piece of information in mind, we can write for $X \in Q$,

$$u - T_{X_0}u = u - \zeta T_{X_0}u = G \star (f - g)$$
$$= h + r$$

with

$$h = G \star ((f - g)\mathbf{1}_{Q^c}) \quad \text{and} \quad r = G \star ((f - f(X_0))\mathbf{1}_Q)$$

where $Q^c = \mathbb{R}^{d+1} \setminus Q$. Remark in particular that

$$h_t - \Delta h = 0 \text{ in } Q.$$

Now we estimate

$$|u - T_{X_0}u - T_{X_0}h|_{0,Q_\rho(X_0)} \leq |h - T_{X_0}h|_{0,Q_\rho(X_0)} + |r|_{0,Q_\rho(X_0)} \quad (2.6)$$

and we study the two terms of the right hand side.

We use Fact 1 to get first

$$|r|_{0,Q_\rho(X_0)} \leq [f]_{\alpha,Q}(K+1)^\alpha \rho^\alpha |G \star \mathbf{1}_Q|_{0,Q_\rho(X_0)}$$
$$\leq C(K+1)^{2+\alpha}\rho^{2+\alpha}[f]_{\alpha,Q}. \quad (2.7)$$

We now write for $X \in Q_\rho(X_0)$,

$$h(X) = h(X_0) + h_t(\theta, x)(t - t_0) + Dh(X_0) \cdot (x - x_0) + \frac{1}{2}D^2h(\Theta)(x - x_0) \cdot (x - x_0)$$

for some $\theta \in (t_0, t)$ and $\Theta = (t_0, y_0) \in Q_\rho(X_0)$. Hence, we have

$$h(X) - T_{X_0}h(X) = (h_t(\theta, x) - h_t(X_0))(t - t_0)$$
$$+ \frac{1}{2}(D^2h(\Theta) - D^2h(X_0))(x - x_0) \cdot (x - x_0)$$

from which we deduce

$$|h(X) - T_{X_0}h(X)| \leq \rho^2|h_t(\theta, x) - h_t(X_0)| + \rho^2|D^2h(\Theta) - D^2h(X_0)|. \quad (2.8)$$

We now use Fact 2 in order to get

$$|h - T_{X_0}h|_{0,Q_\rho(X_0)} \leq \rho^2 \left(\rho^2 \left| \frac{\partial^2}{\partial t^2} h \right|_{0,Q_\rho(X_0)} + \rho \left| \frac{\partial}{\partial t} Dh \right|_{0,Q_\rho(X_0)} \right)$$

$$+ C\rho^3 |D^3 h|_{0,Q_\rho(X_0)}$$

$$\leq C(\rho^4 (K\rho)^{-4} + \rho^3 (K\rho)^{-3} + \rho^3 (K\rho)^{-3}) |h|_{0,Q}$$

$$\leq C(K^{-4} + 2K^{-3}) |h|_{0,Q}$$

$$\leq CK^{-3} |h|_{0,Q}$$

by choosing $K \geq 1$. We next estimate $|h|_{0,Q}$ as follows

$$|h|_{0,Q} \leq |u - T_{X_0}u - r|_{0,Q} \leq |u - T_{X_0}u|_{0,Q} + |r|_{0,Q}$$

$$\leq C(K+1)^{2+\alpha} \rho^{2+\alpha} ([u]_{2+\alpha,Q} + |[f]_{\alpha,Q})$$

where we used (2.8) for u instead of h and we used (2.7). Then, we have

$$|h - T_{X_0}h|_{0,Q_\rho(X_0)} \leq C \frac{(K+1)^{2+\alpha}}{K^3} \rho^{2+\alpha} ([u]_{2+\alpha,Q} + [f]_{\alpha,Q}). \qquad (2.9)$$

Combining (2.6), (2.7) and (2.9), we finally get

$$\rho^{-(2+\alpha)} |u - T_{X_0}u - T_{X_0}h|_{0,Q_\rho(X_0)} \leq C(K+1)^{2+\alpha} [f]_{\alpha,Q}$$

$$+ C \frac{(K+1)^{2+\alpha}}{K^3} ([u]_{2+\alpha,Q} + [f]_{\alpha,Q}).$$

In view of Proposition 2.1.4, it is enough to choose $K \geq 1$ large enough so that

$$C \frac{(K+1)^{2+\alpha}}{K^3} \leq \frac{1}{2}$$

to conclude the proof of the theorem. $\qquad\qquad\square$

2.2.5   Schauder Estimate in the Case of Variable Coefficients

Theorem 2.2.12. *Consider a function* $u \in C^{2,\alpha}((0,T) \times \mathbb{R}^d)$ *for some* $\alpha \in (0,1)$. *Then there exists* $C = C(d, \alpha)$ *such that*

$$[u]_{2+\alpha,(0,T)\times\mathbb{R}^d} \leq C \left([f]_{\alpha,(0,T)\times\mathbb{R}^d} + |u|_{0,(0,T)\times\mathbb{R}^d} \right)$$

where $f = \frac{\partial u}{\partial t} - Lu$.

Remark 2.2.13 (Notation). In the remaining of this subsection, it is convenient to write semi-norms as $[\cdot]_{k+\alpha}$ instead of $[\cdot]_{k+\alpha,(0,T)\times\mathbb{R}^d}$, $k = 0, 2$. In the same way, $|\cdot|_0$ stands for $|\cdot|_{0,(0,T)\times\mathbb{R}^d}$.

Remark 2.2.14. Recall that by Corollary 2.2.9, one has

$$|u|_0 \leq T|u_t - Lu|_0 + |u(0,\cdot)|_{0,\mathbb{R}^d}.$$

Before giving a rigorous proof, we would like first to explain the main idea.

Main idea of the proof of Theorem 2.2.12. Assume first that there are no lower order terms ($c \equiv 0$ and $b \equiv 0$).

In a neighbourhood of $X_0 \in \mathbb{R}^{d+1}$, the coefficients of the linear operator L are frozen: the linear operator with constant coefficients is denoted by L_0. If X is close to X_0, then L is not very far from L_0 and this can be measured precisely thanks to the Hölder continuity of coefficients.

Use first Corollary 2.2.11:

$$[u]_{2+\alpha} \leq C[u_t - L_0u]_\alpha \leq C[u_t - Lu]_\alpha + C[Lu - L_0u]_\alpha.$$

Now control $[Lu - L_0u]_\alpha$ thanks to $[u]_{2+\alpha}$ and conclude.

Next, lower order terms are treated by using interpolation inequalities.   □

Let us now make this precise and rigorous.

Proof of Theorem 2.2.12. We first assume that $b \equiv 0$ and $c \equiv 0$. Let f denote $\frac{\partial u}{\partial t} - Lu$.

Let $\varepsilon \in (0, T/2)$ and $\gamma \leq \varepsilon/2$ be a positive real number to be fixed later and consider X_1 and X_2 such that

$$[u_t]_{\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} \leq 2\rho(X_1, X_2)^{-\alpha}|u_t(X_1) - u_t(X_2)|$$

where we recall that $\rho(X_1, X_2) = \sqrt{|t_1 - t_2|} + |x_1 - x_2|$ if $X_i = (t_i, x_i)$, $i = 1, 2$.

If $\rho(X_1, X_2) \geq \gamma$, then we use interpolation inequalities (2.3) in order to get

$$[u_t]_{\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} \leq 2\gamma^{-\alpha}|u_t|_0$$

$$\leq \frac{1}{4}[u]_{2+\alpha} + C(\gamma)|u|_0.$$

If $\rho(X_1, X_2) < \gamma$, we consider $\zeta \in C^\infty(\mathbb{R}^{d+1})$ with compact support such that $\zeta(X) = 1$ if $\rho(X, 0) \leq 1$ and $\zeta(X) = 0$ if $\rho(X, 0) \geq 2$. We next define $\xi(t, x) = \zeta(\gamma^{-2}(t - t_1), \gamma^{-1}(x - x_1))$. In particular, $\xi(X) = 1$ if $\rho(X, X_1) \leq \gamma$ and $\xi(X) = 0$ if $\rho(X, X_1) \geq 2\gamma$.

Now we use Corollary 2.2.11 in order to get

$$
\begin{aligned}
[u_t]_{\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} &\le 2\rho(X_1,X_2)^{-\alpha}|u_t(X_1)-u_t(X_2)| \\
&\le 2[(u\xi)]_{2+\alpha} \\
&\le 2C[(u\xi)_t - L(X_1)(u\xi)]_\alpha \\
&\le 2C[(u\xi)_t - L(u\xi)]_\alpha + 2C[(L(X_1)-L)(u\xi)]_\alpha. \qquad (2.10)
\end{aligned}
$$

We estimate successively the two terms of the right hand side of the last line. First, we write

$$
(u\xi)_t - L(u\xi) = \xi f + u(\xi_t - L\xi) - 2ADu \cdot D\xi
$$

since $L(u\xi) = uL\xi + \xi Lu + 2ADu \cdot D\xi$. Using interpolation inequalities (2.3), this implies

$$
\begin{aligned}
[(u\xi)_t - L(u\xi)]_\alpha &\le C(\gamma)([f]_\alpha + [u]_\alpha + [Du]_\alpha) \\
&\le \gamma^\alpha[u]_{2+\alpha} + C(\gamma)([f]_\alpha + |u|_0). \qquad (2.11)
\end{aligned}
$$

We next write

$$
(L(X_1)-L)(u\xi) = \text{trace}[(A(X_1)-A(X))D^2(u\xi)]
$$

and for X such that $\rho(X_1,X) \le 2\gamma$, we thus get thanks to interpolation inequalities (2.3)

$$
\begin{aligned}
[(L(X_1)-L)(u\xi)]_\alpha &\le C\gamma^\alpha[D^2(u\xi)]_\alpha + C|D^2(u\xi)|_0 \\
&\le C\gamma^\alpha[u]_{2+\alpha} + C(\gamma)|u|_0. \qquad (2.12)
\end{aligned}
$$

Combining (2.10)–(2.12), we finally get in the case where $\rho(X_1,X_2) \le \gamma$,

$$
[u_t]_{\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} \le C\gamma^\alpha[u]_{2+\alpha} + C(\gamma)([f]_\alpha + |u|_0).
$$

We conclude that we have in both cases

$$
[u_t]_{\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} \le (C\gamma^\alpha + 1/4)[u]_{2+\alpha} + C(\gamma)([f]_\alpha + |u|_0).
$$

We can argue in a similar way to get

$$
[D^2u]_{\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} \le (C\gamma^\alpha + 1/4)[u]_{2+\alpha} + C(\gamma)([f]_\alpha + |u|_0).
$$

Adding these two inequalities yield

$$[u]_{2+\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} \leq (C\gamma^\alpha + 1/2)[u]_{2+\alpha} + C(\gamma)([f]_\alpha + |u|_0).$$

Now choose γ such that $C\gamma^\alpha \leq 1/4$ and get

$$[u]_{2+\alpha,(\varepsilon,T-\varepsilon)\times\mathbb{R}^d} \leq \frac{3}{4}[u]_{2+\alpha} + C([f]_\alpha + |u|_0).$$

Taking the supremum over $\varepsilon \in (0, T/2)$ allows us to conclude in the case where $b \equiv 0$ and $c \equiv 0$.

If now $b \neq 0$ and $c \neq 0$, we apply the previous result and get

$$[u]_{2+\alpha} \leq C([f + b \cdot Du + cu]_\alpha + |u|_0).$$

Use now interpolation inequalities once again to conclude.   □

2.3 Viscosity Solutions: A Short Overview

Viscosity solutions were first introduced by Crandall and Lions [CL81]. This notion of weak solution enabled to characterize the value function of an optimal control problem as the unique solution of the corresponding first order Hamilton–Jacobi equation. An example of such an equation is the following one

$$\frac{\partial u}{\partial t} + \frac{1}{2}|Du|^2 + V(x) = 0 \tag{2.13}$$

for some continuous function V. The viscosity solution theory is also by now a fundamental tool for the study of nonlinear elliptic and parabolic equations.

2.3.1 Definition and Stability of Viscosity Solutions

2.3.1.1 Degenerate Ellipticity

We recall that linear parabolic equations in non-divergence form have the following general form

$$\frac{\partial u}{\partial t} - Lu = f$$

with

$$Lu = \text{trace}(AD^2u) + b \cdot Du + cu$$

with $A \geq 0$ (in the sense of symmetric matrices).

We now consider very general *nonlinear* parabolic equation of the form

$$\frac{\partial u}{\partial t} + F(t, x, Du, D^2u) = 0 \tag{2.14}$$

where we assume that the *nonlinearity* $F : (0, T) \times \Omega \times \mathbb{R}^d \times \mathbb{S}_d \to \mathbb{R}$ is continuous and satisfies the following condition

$$A \leq B \Rightarrow F(t, x, p, A) \geq F(t, x, p, B). \tag{2.15}$$

In other words, the nonlinearity F is non-increasing with respect to the matrix variable. We say that F is *degenerate elliptic*.

Remark 2.3.1. In the case of parabolic Monge–Ampère equations such as (2.1) or (2.2), the nonlinearity is well-defined and degenerate elliptic only on a subset of $\mathbb{S}_d$; precisely, it is only defined either on the subset $\mathbb{S}_d^+$ of semi-definite symmetric matrices or on the subset $\mathbb{S}_d^{++}$ of definite symmetric matrices. Hence, solutions should be convex or strictly convex.

2.3.1.2   Semi-continuity

Consider an open set $Q \subset \mathbb{R}^{d+1}$. We recall that u is lower semi-continuous at (t, x) if, for all sequences $(s_n, y_n) \to (t, x)$,

$$u(t, x) \leq \liminf_{n \to \infty} u(s_n, y_n).$$

In the same way, one can define upper semi-continuous functions. Very often, the previous inequality is written

$$u(t, x) \leq \liminf_{(s,y) \to (t,x)} u(s, y).$$

If u is bounded from below in a neighbourhood of Q, one can define the *lower semi-continuous envelope* of u in Q as the largest lower semi-continuous function lying below u. It is denoted by u_*. Similarly, the upper semi-continuous envelope u^* of a locally bounded from above function u can be defined.

2.3.1.3 Definition(s)

In this paragraph, we give the definition of a viscosity solution of the fully nonlinear parabolic equation (2.14). We give a first definition in terms of test functions. We then introduce the notion of subdifferentials and superdifferentials with which an equivalent definition can be given (see Remark 2.3.8 below).

In order to motivate the definition of a viscosity solution, we first derive necessary conditions for smooth solutions of (2.14).

Consider an open set $Q \subset \mathbb{R}^{d+1}$ and a function $u : Q \to \mathbb{R}$ which is C^1 with respect to t and C^2 with respect to x. Consider also a function ϕ with the same regularity and assume that $u \leq \phi$ in a neighbourhood of $(t, x) \in Q$ and $u = \phi$ at (t, x). Then

$$\frac{\partial \phi}{\partial t}(t, x) = \frac{\partial u}{\partial t}(t, x)$$

$$D\phi(t, x) = Du(t, x)$$

$$D^2\phi(t, x) \geq D^2 u(t, x).$$

Using the degenerate ellipticity of the nonlinearity F, we conclude that

$$\frac{\partial \phi}{\partial t}(t, x) + F(t, x, D\phi(t, x), D^2\phi(t, x))$$

$$\leq \frac{\partial u}{\partial t}(t, x) + F(t, x, Du(t, x), D^2 u(t, x)) = 0.$$

A similar argument can be used to prove that if $u \geq \phi$ in a neighbourhood of (t, x) with $u(t, x) = \phi(t, x)$ then the reserve inequality holds true. These facts motivate the following definitions.

Definition 2.3.2 (Test functions). A *test function* on the set Q is a function $\phi : Q \to \mathbb{R}$ which is C^1 with respect to t and C^2 with respect to x.

Given a function $u : Q \to \mathbb{R}$, we say that the test function ϕ *touches u from above (resp. below) at* (t, x) if $u \leq \phi$ (resp. $u \geq \phi$) in a neighbourhood of (t, x) and $u(t, x) = \phi(t, x)$.

Remark 2.3.3. If $u - \phi$ reaches a local maximum (resp. minimum) at (t_0, x_0), then $\phi + [u(t_0, x_0) - \phi(t_0, x_0)]$ touches u from above (resp. below).

Definition 2.3.4 (Viscosity solutions). Consider a function $u : Q \to \mathbb{R}$ for some open set Q.

- u is a *subsolution* of (2.14) if u is upper semi-continuous and if, for all $(t, x) \in Q$ and all test functions ϕ touching u from above at (t, x),

$$\frac{\partial \phi}{\partial t}(t, x) + F(t, x, D\phi(t, x), D^2\phi(t, x)) \leq 0.$$

- u is a *supersolution* of (2.14) if u is lower semi-continuous and if, for all $(t, x) \in Q$ and all test functions ϕ touching u from below at (t, x),

$$\frac{\partial \phi}{\partial t}(t, x) + F(t, x, D\phi(t, x), D^2\phi(t, x)) \geq 0.$$

- u is a *solution* of (2.14) if it is both a sub- and a supersolution.

Remark 2.3.5. Remark that a viscosity solution of (2.14) is a continuous function.

When proving uniqueness of viscosity solutions, it is convenient to work with the following objects.

Definition 2.3.6 (Second order sub-/super-differentials). The following set

$$\mathcal{P}^\pm(u)(t, x) = \{(\alpha, p, X) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d :$$
$$(\alpha, p, X) = (\partial_t \phi(t, x), D\phi(t, x), D^2\phi(t, x))$$
$$\text{s.t. } \phi \text{ touches } u \text{ from above (resp. below) at } (t, x)\}$$

is *the super-(resp. sub-)differential* of the function u at the point (t, x).

Remark 2.3.7. Here is an equivalent definition: $(\alpha, p, X) \in \mathcal{P}^+ u(t, x)$ if and only if

$$u(s, y) \geq u(t, x) + \alpha(s-t) + p \cdot (y-x) + \frac{1}{2}X(x-y) \cdot (x-y) + o\left(|s-t| + |y-x|^2\right)$$

for (s, y) in a neighbourhood of (t, x). A similar characterization holds for $\mathcal{P}^-$.

Remark 2.3.8. The definition of a viscosity solution can be given using sub- and super-differentials of u. Indeed, as far as subsolutions are concerned, in view of Definitions 2.3.4 and 2.3.6, u is a viscosity subsolution of (2.14) in the open set Q if and only if for all $(t, x) \in Q$ and all $(\alpha, p, X) \in \mathcal{P}^+ u(t, x)$,

$$\alpha + F(t, x, p, X) \leq 0.$$

When proving uniqueness, the following limiting versions of the previous objects are used.

Definition 2.3.9 (Limiting super-/sub-differentials).

$$\overline{\mathcal{P}}^\pm(u)(t, x) = \{(\alpha, p, X) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d : \exists(t_n, x_n) \to (t, x) \text{ s.t.}$$
$$(\alpha_n, p_n, X_n) \to (\alpha, p, X), u(t_n, x_n) \to u(t, x),$$
$$(\alpha_n, p_n, X_n) \in \mathcal{P}^\pm u(t_n, x_n)\}$$

Remark 2.3.10. Since F is assumed to be continuous, the reader can remark that u is a viscosity subsolution of (2.14) in Q if and only if for all $(t, x) \in Q$ and all $(\alpha, p, X) \in \overline{\mathcal{P}}^+ u(t, x)$,

$$\alpha + F(t, x, p, X) \leq 0.$$

An analogous remark can be made for supersolutions.

2.3.1.4 First Properties

In this section, we state without proofs some important properties of sub- and supersolutions. Proofs in the elliptic case can be found in [CIL92] for instance. These proofs can be readily adapted to the parabolic framework.

Proposition 2.3.11 (Stability properties).

- *Let $(u_\alpha)_\alpha$ be a family of subsolutions of (2.14) in Q such that the upper semi-continuous envelope u of $\sup_\alpha u_\alpha$ is finite in Q. Then u is also a subsolution of (2.14) in Q.*
- *If $(u_n)_n$ is a sequence of subsolutions of (2.14), then the upper relaxed-limit u of the sequence defined as follows*

$$\bar{u}(t, x) = \limsup_{(s,y) \to (t,x), n \to \infty} u_n(s, y) \tag{2.16}$$

is everywhere finite in Q, then it is a subsolution of (2.14) in Q.

Remark 2.3.12. An analogous proposition can be stated for supersolutions.

2.3.2 The Perron Process

In this subsection, we would like to give an idea of the general process that allows one to construct solutions for fully nonlinear parabolic equations.

2.3.2.1 General Idea

The Perron process is well known in harmonic analysis and potential analysis. It has been adapted to the case of fully nonlinear elliptic equations in non-divergence form by Ishii [Ish87].

    The general idea is the following one: assume that one can construct a subsolution u^- and a supersolution u^+ to a nonlinear parabolic equation of the form (2.14) such that $u^- \leq u^+$. Using Proposition 2.3.11, we can construct a maximal

subsolution u lying between u^- and u^+. Then a general argument allows one to prove that the lower semi-continuous envelope of the maximal subsolution u is in fact a supersolution.

Remark 2.3.13. Before making the previous argument a little bit more precise, we would like to point out that the function u constructed by this general method is not a solution in the sense of Definition 2.3.4. It is a so-called *discontinuous (viscosity) solution* of (2.14). We decided to stick to continuous viscosity solution in these lecture notes and to state the result of the Perron process as in Lemma 2.3.15 below. See also Sect. 2.3.2.3.

Example 2.3.14. In many important cases, $u^\pm$ are chosen in the following form: $u_0(x) \pm Ct$ where u_0 is the smooth initial datum and C is a large constant, precisely:

$$C \geq \sup_{x \in \mathbb{R}^d} |F(0, x, Du_0(x), D^2 u_0(x))|.$$

If non-smooth/unbounded initial data are to be considered, discontinuous stability arguments can be used next.

2.3.2.2 Maximal Subsolution and Bump Construction

We now give more details about the general process to construct a "solution". We consider a cylindrical domain $Q = (0, T) \times \Omega$ for some domain $\Omega \subset \mathbb{R}^d$.

Lemma 2.3.15. *Assume that $u^\pm$ is a super-(resp. sub-) solution of (2.14) in Q. Then there exists a function $u : Q \to \mathbb{R}$ such that $u^- \leq u \leq u^+$ and u^* is a subsolution of (2.14) and u_* is a supersolution of (2.14).*

Proof. Consider

$$\mathcal{S} = \{v : Q \to \mathbb{R} \text{ s.t. } u^- \leq v \leq u^+ \text{ and } v^* \text{ subsolution of (2.14)}\}.$$

By Proposition 2.3.11, we know that the upper semi-continuous envelope u^* of the function

$$u = \sup_{v \in \mathcal{S}} v$$

is a subsolution of (2.14).

We next prove that the lower semi-continuous envelope u_* of u is a supersolution of (2.14) in Q. Arguing by contradiction, one can assume that there exists $(\alpha, p, X) \in \mathcal{P}^- u_*(t, x)$ such that

$$\alpha + F(t, x, p, X) =: -\theta < 0. \tag{2.17}$$

Remark that at (t, x), we have necessarily

$$u_*(t, x) < u^+(t, x).$$

Indeed, if this is not the case, then $(\alpha, p, X) \in \mathcal{P}^- u^+(t, x)$ and (2.17) cannot be true since u^+ is a supersolution of (2.14). Up to modifying the constant θ, we can also assume that

$$u_*(t, x) - u^+(t, x) \leq -\theta < 0. \tag{2.18}$$

Without loss of generality, we can also assume that $(t, x) = (0, 0)$ and $u_*(t, x) = 0$. Let us consider the following "paraboloid"

$$P(s, y) = \tau s + p \cdot y + \frac{1}{2} X y \cdot y + \delta - \gamma \left(\frac{1}{2} |y|^2 + |s| \right)$$

with δ and γ to be chosen later. Compute next

$$\frac{\partial P}{\partial s}(s, y) + F(s, y, DP(s, y), D^2 P(s, y))$$

$$= \tau - \gamma \frac{s}{|s|} + F(s, y, p + X y - \gamma y, X - \gamma I)$$

(if $s = 0$, $\frac{s}{|s|}$ should be replaced with any real number $\sigma \in [-1, 1]$). Hence, for r and γ small enough, we have

$$\frac{\partial P}{\partial s} + F(s, y, DQ, D^2 Q) \leq -\frac{\theta}{2} < 0$$

for all $(s, y) \in V_r$. Moreover, since $(\tau, p, X) \in \mathcal{P}^- u_*(t, x)$, we have

$$u_*(s, y) \geq \tau s + p \cdot y + \frac{1}{2} X y \cdot y + o(|y|^2 + |s|)$$

$$\geq P(s, y) - \delta + \gamma \left(\frac{1}{2} |y|^2 + |s| \right) + o(|y|^2 + |s|).$$

Choose now $\delta = \frac{\gamma r}{4}$ and consider $(s, y) \in V_r \setminus V_{r/2}$:

$$u_*(s, y) \geq P(s, y) - \frac{\gamma r}{4} + \frac{\gamma r}{2} + o(r) = P(s, y) + \frac{\gamma r}{4} + o(r).$$

Consequently, for r small enough,

$$u(s, y) - P(s, y) \geq \frac{\gamma r}{8} > 0 \text{ in } V_r \setminus V_{r/2},$$

$$P(s, y) < u^+(s, y) \text{ in } V_r$$

where we used (2.18) to get the second inequality.

We next consider

$$U(s, y) = \begin{cases} \max\{u(s, y), P(s, y)\} & \text{if } (s, y) \in V_r, \\ u(s, y) & \text{if not.} \end{cases}$$

On one hand, we remark that the function U^* is still a subsolution of (2.14) and $U \geq u \geq u_-$ and $U \leq u^+$. Consequently, $U \in \mathcal{S}$ and in particular, $U \leq u$. On the other hand, $\sup_{\mathbb{R}^+ \times \mathbb{R}^d}\{U - u\} \geq \delta$; indeed, consider $(t_n, x_n) \to (0, 0)$ such that $u(t_n, x_n) \to u_*(0, 0) = 0$ and write

$$\lim_{n \to \infty} U(t_n, x_n) - u(t_n, x_n) \geq \lim_{n \to \infty} P(t_n, x_n) - u(t_n, x_n) = \delta > 0.$$

This contradicts the fact that $U \leq u$. The proof of the lemma is now complete. □

2.3.2.3 Continuous Solutions from Comparison Principle

As mentioned above, the maximal subsolution u^* is not necessarily continuous; hence, its lower semi-continuous envelope u_* does not coincide necessarily with it. In particular, we cannot say that u is a solution in the sense of Definition 2.3.4 (cf. Remark 2.3.13 above).

We would get a (continuous viscosity) solution if $u^* = u_*$. On one hand, u^* is upper semi-continuous by construction and on the other hand $u_* \leq u^*$ by definition of the semi-continuous envelopes. Hence, u is a solution of (2.14) if and only if $u^* \leq u_*$ in Q. Since u^* is a subsolution of (2.14) in Q and u_* is a supersolution of (2.14) in Q, it is thus enough that (2.14) satisfies a *comparison principle* and that the barriers $u^\pm$ satisfy some appropriate inequality on the parabolic boundary. More precisely, we would like on one hand that

Comparison principle. *If u is a subsolution of* (2.14) *in Q and v is a supersolution of* (2.14) *in Q and $u \leq v$ on the parabolic boundary $\partial_p Q$, then $u \leq v$ in Q.*

and on the other hand, we would like that $u^* \leq u_*$ on $\partial_p Q$. This boundary condition would be true if

$$(u^+)^* \leq (u^-)_* \text{ on } \partial_p Q.$$

We emphasize that the lower and upper semi-continuous envelopes appearing in the previous inequality are performed with respect to time and space.

Example 2.3.16. If for instance $Q = (0, T) \times \mathbb{R}^d$, then barriers should satisfy

$$(u^+)^*(0, x) \leq (u^-)_*(0, x) \text{ for } x \in \mathbb{R}^d.$$

This condition is fullfilled for such a Q if $u^\pm = u_0 \pm Ct$ (see Example 2.3.14).

In the next subsection, we will present general techniques for proving comparison principles. The reader should be aware of the fact that, in many practical cases, general theorems from the viscosity solution theory do not apply to the equation under study. In those cases, one has to adapt the arguments presented below in order to take into account the specific difficulties implied by the specific equation. The reader is referred to [CIL92] for a large review of available tools.

2.3.3 Introduction to Comparison Principles

In this subsection, we present classical techniques to prove comparison principles in some typical cases.

2.3.3.1 First Order Equations

In this paragraph, we first study first order Hamilton–Jacobi equations of the following form

$$\frac{\partial u}{\partial t} + H(x, Du) = 0. \tag{2.19}$$

As we will see, a comparison principle holds true if H satisfies the following structure condition: for all $x, y, p \in \mathbb{R}^d$,

$$|H(x, p) - H(y, p)| \leq C|x - y|. \tag{2.20}$$

In order to avoid technicalities and illustrate main difficulties, we assume that $x \mapsto H(x, p)$ is $\mathbb{Z}^d$-periodic; hence, solutions should also be $\mathbb{Z}^d$-periodic for $\mathbb{Z}^d$-periodic initial data.

Theorem 2.3.17 (Comparison principle II). *Consider a continuous $\mathbb{Z}^d$-periodic function u_0. If u is a $\mathbb{Z}^d$-periodic subsolution of (2.19) in $(0, T) \times \mathbb{R}^d$ and v is a $\mathbb{Z}^d$-periodic supersolution of (2.19) in $(0, T) \times \mathbb{R}^d$ such that $u(0, x) \leq u_0(x) \leq v(0, x)$ for all $x \in \mathbb{R}^d$, then $u \leq v$ in $(0, T) \times \mathbb{R}^d$.*

Proof. The beginning of the proof is the same as in the proof of Theorem 2.2.4: we assume that

$$M = \sup_{t \in (0,T), x \in \mathbb{R}^d} \left\{ u(t, x) - v(t, x) - \frac{\gamma}{T - t} \right\} > 0.$$

Here, we cannot use the equation directly, since it is not clear wether $u - v$ satisfies a nonlinear parabolic equation or not (recall that the equation is nonlinear). Hence, we should try to duplicate the (time and space) variables.

Doubling Variable Technique

Consider

$$M_\varepsilon = \sup_{t,s\in(0,T),x,y\in\mathbb{R}^d} \left\{ u(t,x) - v(s,y) - \frac{(t-s)^2}{2\varepsilon} - \frac{|x-y|^2}{2\varepsilon} - \frac{\eta}{T-t} \right\}.$$

Remark that $M_\varepsilon \geq M > 0$. This supremum is reached since u is upper semi-continuous and v is lower semi-continuous and both functions are $\mathbb{Z}^d$-periodic. Let $(t_\varepsilon, s_\varepsilon, x_\varepsilon, y_\varepsilon)$ denote a maximizer. Then we have

$$\frac{(t_\varepsilon - s_\varepsilon)^2}{2\varepsilon} + \frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon} \leq u(t_\varepsilon, x_\varepsilon) - v(s_\varepsilon, y_\varepsilon) \leq |u^+|_0 + |v_-|_0$$

where we recall that $|w|_0 = \sup_{(t,x)\in(0,T)\times\mathbb{R}^d} |w(t,x)|$. In particular, up to extracting subsequences, $t_\varepsilon \to t$, $s_\varepsilon \to t$ and $x_\varepsilon \to x$, $y_\varepsilon \to y$ and $t_\varepsilon - s_\varepsilon = O(\sqrt{\varepsilon})$ and $x_\varepsilon - y_\varepsilon - O(\sqrt{\varepsilon})$.

Assume first that $t = 0$. Then

$$0 < M \leq \limsup_{\varepsilon\to 0} M_\varepsilon \leq \limsup_\varepsilon u(t_\varepsilon, x_\varepsilon) - \liminf_\varepsilon v(s_\varepsilon, y_\varepsilon)$$

$$\leq u(0,x) - v(0,x) \leq 0.$$

This is not possible. Hence $t > 0$.

Since $t > 0$, for ε small enough, $t_\varepsilon > 0$ and $s_\varepsilon > 0$. Now remark that the function ϕ_u

$$(t,x) \mapsto v(s_\varepsilon, y_\varepsilon) + \frac{(t-s_\varepsilon)^2}{2\varepsilon} + \frac{|x-y_\varepsilon|^2}{2\varepsilon} + \frac{\eta}{T-t}$$

is a test function such that $u - \phi_u$ reaches a maximum at $(t_\varepsilon, x_\varepsilon)$. Hence (recall Remark 2.3.3),

$$\frac{\eta}{(T-t_\varepsilon)^2} + \frac{t_\varepsilon - s_\varepsilon}{\varepsilon} + H(x_\varepsilon, p_\varepsilon) \leq 0$$

with $p_\varepsilon = \frac{x_\varepsilon - y_\varepsilon}{\varepsilon}$. Similarly, the function ϕ_v

$$(s,y) \mapsto u(t_\varepsilon, x_\varepsilon) - \frac{(s-t_\varepsilon)^2}{2\varepsilon} - \frac{|y-x_\varepsilon|^2}{2\varepsilon} - \frac{\eta}{T-t_\varepsilon}$$

is a test function such that $v - \phi_v$ reaches a minimum at $(s_\varepsilon, y_\varepsilon)$; hence

$$\frac{t_\varepsilon - s_\varepsilon}{\varepsilon} + H(y_\varepsilon, p_\varepsilon) \leq 0$$

with the same p_ε! Substracting the two viscosity inequalities yields

$$\frac{\eta}{(T - t_\varepsilon)^2} \leq H(y_\varepsilon, p_\varepsilon) - H(x_\varepsilon, p_\varepsilon).$$

In view of (2.20), we conclude that

$$\frac{\eta}{T^2} \leq C|x_\varepsilon - y_\varepsilon| = O(\sqrt{\varepsilon}).$$

Letting $\varepsilon \to 0$ yields the desired contradiction. $\qquad\square$

Remark 2.3.18. Condition (2.20) is satisfied by (2.13) if the potential V is Lipschitz continuous. On the contrary, if $H(x, p) = c(x)|p|$, then the Hamilton–Jacobi equation is the so-called eikonal equation and it does not satisfy (2.20) even if c is globally Lipschitz. Such an Hamiltonian satisfies

$$|H(x, p) - H(y,)| \leq C(1 + |p|)|x - y|. \tag{2.21}$$

For such equations, the penalization should be studied in greater details in order to prove that

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon} \to 0 \text{ as } \varepsilon \to 0.$$

With this piece of information in hand, the reader can check that the same contradiction can be obtained for Lipschitz c's. See for instance [Barl94] for details.

Since we will use once again this additional fact about penalization, we state it now in a lemma.

Lemma 2.3.19. *Consider $\tilde{u}(t, x) = u(t, x) - \eta(T - t)^{-1}$. Assume that*

$$M_\varepsilon = \sup_{\substack{x, y \in \mathbb{R}^d \\ t, s \in (0, T)}} \tilde{u}(t, x) - v(s, y) - \frac{|x - y|^2}{2\varepsilon} - \frac{|t - s|^2}{2\varepsilon}$$

is reached at $(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon)$. Assume moreover that $(x_\varepsilon, y_\varepsilon, t_\varepsilon, s_\varepsilon) \to (x, y, t, s)$ as $\varepsilon \to 0$. Then

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon} \to 0 \text{ as } \varepsilon \to 0.$$

Remark 2.3.20. The reader can check that the previous lemma still holds true if $v(s, y)$ is replaced with $v(t, y)$ and if the term $\varepsilon^{-1}|t - s|^2$ is removed.

Proof. Remark first that $\varepsilon \mapsto M_\varepsilon$ is non-decreasing and $M_\varepsilon \geq M := \sup_{\mathbb{R}^d}(\tilde{u} - v)$. Hence, as $\varepsilon \to 0$, M_ε converges to some limit $l \geq M$. Moreover,

$$M_{2\varepsilon} \geq \tilde{u}(t_\varepsilon, x_\varepsilon) - v(s_\varepsilon, y_\varepsilon) - \frac{|x_\varepsilon - y_\varepsilon|^2}{4\varepsilon} - \frac{|t_\varepsilon - s_\varepsilon|^2}{4\varepsilon}$$

$$\geq M_\varepsilon + \frac{|x_\varepsilon - y_\varepsilon|^2}{4\varepsilon} + \frac{|t_\varepsilon - s_\varepsilon|^2}{4\varepsilon}.$$

Hence,

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{4\varepsilon} + \frac{|t_\varepsilon - s_\varepsilon|^2}{4\varepsilon} \leq M_{2\varepsilon} - M_\varepsilon \to l - l = 0. \qquad \square$$

2.3.3.2 Second Order Equations with No x Dependance

In this subsection we consider the following equation

$$\frac{\partial u}{\partial t} + H(x, Du) - \Delta u = 0 \tag{2.22}$$

still assuming that $x \mapsto H(x, p)$ is $\mathbb{Z}^d$-periodic and satisfies (2.20). The classical parabolic theory implies that there exists smooth solutions for such an equation. However, we illustrate viscosity solution techniques on this (too) simple example.

Theorem 2.3.21 (Comparison principle III). *Consider a continuous $\mathbb{Z}^d$-periodic function u_0. If u is a $\mathbb{Z}^d$-periodic subsolution of (2.22) in $(0, T) \times \mathbb{R}^d$ and v is a $\mathbb{Z}^d$-periodic supersolution of (2.19) in $(0, T) \times \mathbb{R}^d$ such that $u(0, x) \leq u_0(x) \leq v(0, x)$ for all $x \in \mathbb{R}^d$, then $u \leq v$ in $(0, T) \times \mathbb{R}^d$.*

Remark 2.3.22. A less trivial example would be

$$\frac{\partial u}{\partial t} + H(x, Du) - \text{trace}(A_0 D^2 u) = 0$$

for some degenerate matrix $A_0 \in \mathbb{S}_d$, $A_0 \geq 0$. We prefer to keep it simple and study (2.22).

First attempt of proof. We follow the proof of Theorem 2.3.17. If one uses the two test functions ϕ_u and ϕ_v to get viscosity inequalities, this yields

$$\frac{1}{(T - t_\varepsilon)^2} + \frac{t_\varepsilon - s_\varepsilon}{\varepsilon} + H(x_\varepsilon, p_\varepsilon) \leq \text{trace}(\varepsilon^{-1} I),$$

$$\frac{t_\varepsilon - s_\varepsilon}{\varepsilon} + H(y_\varepsilon, p_\varepsilon) \geq -\text{trace}(\varepsilon^{-1} I).$$

Substracting these two inequalities, we get

$$\frac{1}{T^2} \leq O(\sqrt{\varepsilon}) + \frac{2d}{\varepsilon}$$

and it is not possible to get a contradiction by letting $\varepsilon \to 0$.   □

In the previous proof, we lost a very important piece of information about second order derivatives; indeed, assume that u and v are smooth. As far as first order equations are concerned, using the first order optimality condition

$$Du(t_\varepsilon, x_\varepsilon) - p_\varepsilon = 0 \quad \text{and} - Dv(s_\varepsilon, y_\varepsilon) + p_\varepsilon = 0$$

is enough. But for second order equations, one has to use second order optimality condition

$$\begin{pmatrix} Du(t_\varepsilon, x_\varepsilon) & 0 \\ 0 & -Dv(s_\varepsilon, y_\varepsilon) \end{pmatrix} \leq \begin{pmatrix} \varepsilon^{-1}I & -\varepsilon^{-1}I \\ -\varepsilon^{-1}I & \varepsilon^{-1}I \end{pmatrix}.$$

It turns out that for *semi-continuous functions*, the previous inequality still holds true up to an arbitrarily small error in the right hand side.

Uniqueness of viscosity solutions for second order equations where first obtained by Lions [Lions83] by using probabilistic methods. The analytical breakthrough was achieved by Jensen [Jens88]. Ishii's contribution was also essential [Ish89]. In particular, he introduced the matrix inequalities contained in the following lemma. See [CIL92] for a detailed historical survey.

We give a first version of Jensen–Ishii's lemma for the specific test function $(2\varepsilon)^{-1}|x - y|^2$.

Lemma 2.3.23 (Jensen–Ishii's lemma I). *Let U and V be two open sets of $\mathbb{R}^d$ and I an open interval of $\mathbb{R}$. Consider also a bounded subsolution u of (2.14) in $I \times U$ and a bounded supersolution v of (2.14) in $I \times V$. Assume that $u(t, x) - v(t, y) - \frac{|x-y|^2}{2\varepsilon}$ reaches a local maximum at $(t_0, x_0, y_0) \in I \times U \times V$. Letting p denote $\varepsilon^{-1}(x_0 - y_0)$, there exists $\tau \in \mathbb{R}$ and $X, Y \in \mathbb{S}_d$ such that*

$$(\tau, p, X) \in \overline{\mathcal{P}}^+ u(t_0, x_0), (\tau, p, Y) \in \overline{\mathcal{P}}^- v(t_0, y_0)$$

$$-\frac{2}{\varepsilon}\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{3}{\varepsilon}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix}. \tag{2.23}$$

Remark 2.3.24. As a matter of fact, it is not necessary to assume that u and v are sub- and supersolution of an equation of the form (2.14). We chose to present first the result in this way to avoid technicalities. Later on, we will need the standard version of this lemma, so we will state it. See Lemma 2.3.30 below.

Remark 2.3.25. Such a result holds true for more general test functions $\phi(t, x, y)$ than $(2\varepsilon)^{-1}|x - y|^2$. However, this special test function is a very important one and many interesting results can be proven with it. We will give a more general version of this important result, see Lemma 2.3.30.

Remark 2.3.26. The attentive reader can check that the matrix inequality (2.23) implies in particular $X \leq Y$.

Remark 2.3.27. This lemma can be used as a black box and one does so very often. But we mentioned above that some times, one has to work more to get a uniqueness result for some specific equation. In this case, it could be necessary to consider more general test functions, or even to open the black box and go through the proof to adapt it in a proper way.

With such a lemma in hand, we can now prove Theorem 2.3.21.

Proof of Theorem 2.3.21. We argue as in the proof of Theorem 2.3.17 but we do not duplicate the time variable since it is embedded in Lemma 2.3.23. Instead, we consider

$$M_\varepsilon = \sup_{\substack{x,y \in \mathbb{R}^d \\ t \in (0,T)}} \left\{ u(t,x) - v(t,y) - \frac{|x-y|^2}{2\varepsilon} - \frac{\eta}{T-t} \right\},$$

let $(t_\varepsilon, x_\varepsilon, y_\varepsilon)$ denote a maximiser and apply Lemma 2.3.23 with $\tilde{u}(t,x) = u(t,x) - \frac{\eta}{T-t}$ and v and we get τ, X, Y such that

$$\left(\tau + \frac{\eta}{(T-t)^2}, p_\varepsilon, X \right) \in \overline{\mathcal{P}}^+ u(t_\varepsilon, x_\varepsilon), (\tau, p_\varepsilon, Y) \in \overline{\mathcal{P}}^- v(t_\varepsilon, y_\varepsilon), \quad X \leq Y$$

(see Remark 2.3.26 above). Hence, we write the two viscosity inequalities

$$\frac{\gamma}{(T-t)^2} + \tau + H(x_\varepsilon, p_\varepsilon) \leq \text{trace } X$$

$$\tau + H(y_\varepsilon, p_\varepsilon) \geq \text{trace } Y \geq \text{trace } X$$

and we substract them in order to get the desired contradiction

$$\frac{\gamma}{T^2} \leq O(\sqrt{\varepsilon}).$$

The proof is now complete.   □

2.3.3.3   Second Order Equations with x Dependance

In this paragraph, we prove a comparison principle for the following *degenerate* elliptic equation

$$\frac{\partial u}{\partial t} + H(x, Du) - \text{trace}(\sigma(x)\sigma^T(x)D^2u) = 0 \qquad (2.24)$$

under the following assumptions

- $x \mapsto H(x, p)$ is $\mathbb{Z}^d$-periodic and satisfies (2.21);
- $\sigma : \mathbb{R}^d \to \mathbb{M}_{d,m}(\mathbb{R})$ is Lipschitz continuous and $\mathbb{Z}^d$-periodic, $m \le d$.

Here, $\mathbb{M}_{d,m}(\mathbb{R})$ denotes the set of real $d \times m$-matrices. We make precise that σ^T denotes the transpose matrix of the $d \times m$-matrix σ.

   The following theorem is, to some respects, the nonlinear counterpart of the first comparison principle we proved in Sect. 2.2 (see Corollary 2.2.6). Apart from the nonlinearity of the equation, another significant difference with Corollary 2.2.6 is that (2.24) is degenerate elliptic and not uniformly elliptic.

Theorem 2.3.28 (Comparison principle IV). *Consider a continuous $\mathbb{Z}^d$-periodic function u_0. If u is a $\mathbb{Z}^d$-periodic subsolution of (2.22) in $(0, T) \times \mathbb{R}^d$ and v is a $\mathbb{Z}^d$-periodic supersolution of (2.19) in $(0, T) \times \mathbb{R}^d$ such that $u(0, x) \le u_0(x) \le v(0, x)$ for all $x \in \mathbb{R}^d$, then $u \le v$ in $(0, T) \times \mathbb{R}^d$.*

Proof. We argue as in the proof of Theorem 2.3.21. The main difference lies after writing viscosity inequalities thanks to Jensen–Ishii's lemma. Indeed, one gets

$$\frac{\eta}{T^2} \le -H(x_\varepsilon, p_\varepsilon) + H(y_\varepsilon, p_\varepsilon) + \text{trace}(\sigma(x_\varepsilon)\sigma^T(x_\varepsilon)X) - \text{trace}(\sigma(y_\varepsilon)\sigma^T(y_\varepsilon)Y)$$

$$\le C\left(1 + \frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon}\right)|x_\varepsilon - y_\varepsilon|$$

$$+ \text{trace}(\sigma(x_\varepsilon)\sigma^T(x_\varepsilon)X) - \text{trace}(\sigma(y_\varepsilon)\sigma^T(y_\varepsilon)Y).$$

The first term can be handled thanks to Lemma 2.3.19. But one cannot just use $X \le Y$ obtained from the matrix inequality (2.23) to handle the second one. Instead, consider an orthonormal basis $(e_i)_i$ of $\mathbb{R}^m$ and write

$$\text{trace}(\sigma(x_\varepsilon)\sigma^T(x_\varepsilon)X) - \text{trace}(\sigma(y_\varepsilon)\sigma^T(y_\varepsilon)Y)$$

$$= \text{trace}(\sigma^T(x_\varepsilon)X\sigma(x_\varepsilon)) - \text{trace}(\sigma^T(y_\varepsilon)Y\sigma(y_\varepsilon))$$

$$= \sum_{i=1}^m (X\sigma(x_\varepsilon)e_i \cdot \sigma(x_\varepsilon)e_i - Y\sigma(y_\varepsilon)e_i \cdot \sigma(y_\varepsilon)e_i)$$

$$\le \frac{3}{\varepsilon} \sum_{i=1}^m |\sigma(x_\varepsilon)e_i - \sigma(y_\varepsilon)e_i|^2;$$

we applied (2.23) to vectors of the form $(\sigma(x_\varepsilon)e_i, \sigma(y_\varepsilon)e_i) \in \mathbb{R}^d \times \mathbb{R}^d$ to get the last line. We can now use the fact that σ is Lipschitz continuous and get

$$\text{trace}(\sigma(x_\varepsilon)\sigma^T(x_\varepsilon)X) - (\sigma(y_\varepsilon)\sigma^T(y_\varepsilon)Y) \leq C \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon}.$$

We thus finally get

$$\frac{\eta}{T^2} \leq C|x_\varepsilon - y_\varepsilon| + C \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon}.$$

We can now get the contradiction $\eta < 0$ by using Lemma 2.3.19 and letting $\varepsilon \to 0$. The proof is now complete. $\square$

2.3.4 Hölder Continuity Through the Ishii–Lions Method

In this subsection, we want to present a technique introduced by Ishii and Lions in [IL90] in order to prove Hölder continuity of solutions of very general fully nonlinear elliptic and parabolic equations. On one hand, it is much simpler than the proof we will present in the next section; on the other hand, it cannot be used to prove further regularity such as Hölder continuity of the gradient.

The fundamental assumptions is that the equation is uniformly elliptic (see below for a definition). For pedagogical purposes, we do not want to prove a theorem for the most general case. Instead, we will look at (2.24) for $\mathbb{S}_d$-valued σ's and special H's

$$\frac{\partial u}{\partial t} + c(x)|Du| - \text{trace}(\sigma(x)\sigma(x)D^2u) = 0 \qquad (2.25)$$

Assumptions (A)

- c is bounded and Lipschitz continuous in Q;
- $\sigma : Q \to \mathbb{S}_d$ is bounded and Lipschitz continuous in x and constant in t;
- There exists $\lambda > 0$ such that for all $X = (t, x) \in Q$,

$$A(x) := \sigma(x)\sigma(x) \geq \lambda I.$$

Under these assumptions, the equation is *uniformly elliptic*, i.e. there exist two positive numbers $0 < \lambda \leq \Lambda$, called *ellipticity constants*, such that

$$\forall X = (t, x) \in Q, \quad \lambda I \leq A(x) \leq \Lambda I. \qquad (2.26)$$

Theorem 2.3.29. *Under Assumptions (A) on H and σ, any viscosity solution u of (2.25) in an open set $Q \subset \mathbb{R}^{d+1}$ is Hölder continuous in time and space.*

When proving Theorem 2.3.29, we will need to use Jensen–Ishii's lemma for a test function which is more general than $(2\varepsilon)^{-1}|x - y|^2$. Such a result can be found in [CIL92].

Lemma 2.3.30 (Jensen–Ishii's Lemma II). *Let U and V be two open sets of $\mathbb{R}^d$ and I an open interval of $\mathbb{R}$. Consider also a bounded subsolution u of (2.14) in $I \times U$ and a bounded supersolution v of (2.14) in $I \times V$. Assume that $u(t, x) - v(t, y) - \phi(x - y)$ reaches a local maximum at $(t_0, x_0, y_0) \in I \times U \times V$. Letting p denote $D\phi(x_0 - y_0)$, for all $\beta > 0$ such that $\beta Z < I$, there exists $\tau \in \mathbb{R}$ and $X, Y \in \mathbb{S}_d$ such that*

$$(\tau, p, X) \in \overline{\mathcal{P}}^+ u(t_0, x_0), (\tau, p, Y) \in \overline{\mathcal{P}}^- v(t_0, y_0)$$

$$-\frac{2}{\beta}\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \le \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le \begin{pmatrix} Z^\beta & -Z^\beta \\ -Z^\beta & Z^\beta \end{pmatrix} \tag{2.27}$$

where $Z = D^2\phi(x_0 - y_0)$ and $Z^\beta = (I - \beta Z)^{-1} Z$.

We can now turn to the proof of Theorem 2.3.29.

Proof of Theorem 2.3.29. We first prove that u is Hölder continuous with respect to x. Without loss of generality, we can assume that Q is bounded. We would like to prove that for all $X_0 = (t_0, x_0) \in Q$ and $(t, x), (t, y) \in Q$,

$$u(t, x) - u(t, y) \le L_1 |x - y|^\alpha + L_2 |x - x_0|^2 + L_2 (t - t_0)^2$$

for $L_1 = L_1(X_0)$ and $L_2 = L_2(X_0)$ large enough. We thus consider

$$M = \sup_{(t,x),(t,y)\in Q} \{u(t, x) - u(t, y) - \phi(x - y) - \Gamma(t, x)\}$$

with $\phi(z) = L_1 |z|^\alpha$ and $\Gamma(t, x) = L_2 |x - x_0|^2 + L_2 (t - t_0)^2$ and we argue by contradiction: we assume that for all $\alpha \in (0, 1)$, $L_1 > 0$, $L_2 > 0$, we have $M > 0$.

Since Q is bounded, M is reached at a point denoted by $(\bar{t}, \bar{x}, \bar{y})$. The fact that $M > 0$ implies first that $\bar{x} \ne \bar{y}$. It also implies

$$\begin{cases} |\bar{x} - \bar{y}| \le \left(\frac{2|u|_{0,Q}}{L_1}\right)^{\frac{1}{\alpha}} =: A < d(X_0, \partial Q), \\ |\bar{X} - X_0| < \sqrt{\frac{2|u|_{0,Q}}{L_2}} =: R_2 \le \frac{d(X_0, \partial Q)}{2} \end{cases} \tag{2.28}$$

if L_1 and L_2 are chosen so that

$$L_1 > \frac{2|u|_{0,Q}}{(d(X_0, \partial Q))^\alpha}, \quad L_2 \geq \frac{8|u|_{0,Q}}{(d(X_0, \partial Q))^2}.$$

In particular we have $\bar{x}, \bar{y} \in \Omega$. We next apply Jensen–Ishii's Lemma 2.3.30 to $\tilde{u}(t, x) = u(t, x) - \Gamma(t, x)$ and $v(s, y)$. Then there exists $\tau \in \mathbb{R}$ and $X, Y \in \mathbb{S}_d$ such that

$$(\tau + 2L_2(\bar{t} - t_0), \bar{p} + 2L_2(\bar{x} - x_0), X + 2L_2 I) \in \overline{\mathcal{P}}^+ u(\bar{t}, \bar{x}), \quad (\tau, \bar{p}, Y) \in \overline{\mathcal{P}}^- u(\bar{t}, \bar{y})$$

where $\bar{p} = D\phi(\bar{x} - \bar{y})$ and $Z = D^2\phi(\bar{x} - \bar{y})$ and (2.27) holds true. In particular, $X \leq Y$. We can now write the two viscosity inequalities

$$2L_2(\bar{t} - t_0) + \tau + H(\bar{x}, \bar{p} + 2L_2(\bar{x} - x_0)) \leq \text{trace}(A(\bar{x})(X + 2L_2 I))$$
$$\tau + H(\bar{y}, \bar{p}) \geq \text{trace}(A(\bar{y})Y)$$

and combine them with (2.28) and (2.26) to get

$$-CL_2 \leq 2L_2(\bar{t} - t_0) \leq c(\bar{y})|\bar{p}| - c(\bar{x})|\bar{p} + 2L_2(\bar{x} - x_0)|$$
$$+ CL_2 + \text{trace}(A(\bar{x})X) - \text{trace}(A(\bar{y})Y). \tag{2.29}$$

We next estimate successively the difference of first order terms and the difference of second order terms. As far as first order terms are concerned, we use that c is bounded and Lipschitz continuous and (2.28) to get

$$c(\bar{y})|\bar{p}| - c(\bar{x})|\bar{p} + 2L_2(\bar{x} - x_0)| \leq C|\bar{x} - \bar{y}||\bar{p}| + CL_2|\bar{x} - x_0|$$
$$\leq C|\bar{x} - \bar{y}||\bar{p}| + CL_2. \tag{2.30}$$

As far as second order terms are concerned, we use (2.26) to get

$$\text{trace}(A(\bar{x})X) - \text{trace}(A(\bar{y})Y) \leq \text{trace}(A(\bar{x})(X - Y)) + \text{trace}((A(\bar{x}) - A(\bar{y}))Y)$$
$$\leq \lambda \text{trace}(X - Y)$$
$$+ \sum_i (\sigma(\bar{x})Y\sigma(\bar{x})e_i \cdot e_i - \sigma(\bar{y})Y\sigma(\bar{y})e_i \cdot e_i)$$
$$\leq \lambda \text{trace}(X - Y) + C\|Y\||\bar{x} - \bar{y}|.$$

We should next estimate $|\bar{p}|$, $\text{trace}(X - Y)$ and $\|Y\|$. In order to do so, we compute $D\phi$ and $D^2\phi$. It is convenient to introduce the following notation

$$a = \bar{x} - \bar{y}, \qquad \hat{a} = \frac{a}{|a|}, \qquad \varepsilon = |a|.$$

$$\bar{p} = D\phi(a) = L_1\alpha|a|^{\alpha-2}a \tag{2.31}$$

$$Z = D^2\phi(a) = L_1\alpha(|a|^{\alpha-2}I + (\alpha-2)|a|^{\alpha-4}a \otimes a)$$

$$= \gamma^{-1}(I - (2-\alpha)\hat{a} \otimes \hat{a}). \tag{2.32}$$

with $\gamma = (L_1\alpha)^{-1}\varepsilon^{2-\alpha}$. The reader can remark that if one chooses $\beta = \gamma/2$, then

$$Z^\beta = (I - \beta Z)^{-1}Z = \frac{2}{\gamma}\left(I - 2\frac{2-\alpha}{3-\alpha}\hat{a} \otimes \hat{a}\right). \tag{2.33}$$

Since Y is such that $-\frac{1}{\beta}I \leq -Y \leq Z^\beta$, we conclude that

$$\|Y\| \leq \frac{2}{\gamma}.$$

We next remark that (2.27) and (2.33) imply that all the eigenvalues of $X - Y$ are non-positive and that one of them is less than

$$4Z^\beta\hat{a} \cdot \hat{a} = -\frac{8}{\gamma}\frac{1-\alpha}{3-\alpha}.$$

Hence

$$\text{trace}(X - Y) \leq -\frac{8}{\gamma}\frac{1-\alpha}{3-\alpha}.$$

Finally, second order terms are estimated as follows

$$\text{trace}(A(\bar{x})X) - \text{trace}(A(\bar{y})Y) \leq -\frac{C}{\gamma} + C\frac{\varepsilon}{\gamma} \leq -\frac{C}{2\gamma} \tag{2.34}$$

(choosing L_1 large enough so that $\varepsilon \leq 1/2$). Combining now (2.29), (2.30) and (2.34) and recalling the definition of γ and ε, we finally get

$$-CL_2 \leq C\varepsilon^\alpha - \frac{CL_1}{\varepsilon^{2-\alpha}} \leq \frac{C}{L_1} - CL_1^{\frac{2}{\alpha}}.$$

Since L_2 is fixed, it is now enough to choose L_1 large enough to get the desired contradiction. The proof is now complete.   □

2.4 Harnack Inequality

In this section, we consider the following special case of (2.14)

$$\frac{\partial u}{\partial t} + F(x, D^2 u) = f \qquad (2.35)$$

for some uniformly elliptic nonlinearity F (see below for a definition) and some continuous function f. The goal of this section is to present and prove the Harnack inequality (Theorem 2.4.35). This result states that the supremum of a non-negative solution of (2.35) can be controlled from above by its infimum times a universal constant plus the L^{d+1}-norm of the right hand side f. The estimates that will be obtained do not depend on the regularity of F with respect to x.

We will see that it is easy to derive the Hölder continuity of solutions from the Harnack inequality, together with an estimate of the Hölder semi-norm.

The Harnack inequality is a consequence of both the L^ε-estimate (Theorem 2.4.15) and of the local maximum principle (Proposition 2.4.34). Since this local maximum principle is a consequence of the L^ε-estimate, the heart of the proof of the Harnack inequality thus lies in proving that a (small power of) non-negative supersolution is integrable, see Theorem 2.4.15 below.

The proof of the L^ε estimate relies on various measure estimates of the solution. These estimates are obtained through the use of a *maximum principle* due to Krylov in the parabolic case.

The proof of the L^ε estimate also involves many different sets, cylinders and cubes. The authors are aware of the fact that it is difficult to follow the corresponding notation. Some pictures are provided and the authors hope they are helpful with this respect.

Pucci's Operators

Given ellipticity constants $0 < \lambda \le \Lambda$, we consider

$$P^+(M) = \sup_{\lambda I \le A \le \Lambda I} \{-\operatorname{trace}(AM)\},$$

$$P^-(M) = \inf_{\lambda I \le A \le \Lambda I} \{-\operatorname{trace}(AM)\}.$$

Some model fully nonlinear parabolic equations are

$$\frac{\partial u}{\partial t} + P^+(D^2 u) = f, \qquad (2.36)$$

$$\frac{\partial u}{\partial t} + P^-(D^2 u) = f. \qquad (2.37)$$

Remark that those nonlinear operators only depend on ellipticity constants λ, Λ and dimension d. They are said *universal*. Similarly, constants are said *universal* if they only depend on λ, Λ and d.

Uniform Ellipticity

Throughout the remaining of this section, we make the following assumptions on F: for all $X, Y \in \mathbb{S}_d$ and $x \in \Omega$,

$$P^-(X - Y) \leq F(x, X) - F(x, Y) \leq P^+(X - Y).$$

This condition is known as the *uniform ellipticity* of F. Remark that this condition implies in particular that F is degenerate elliptic in the sense of Sect. 2.3.1.1 (see Condition 2.15).

2.4.1 A Maximum Principle

In order to state and prove the maximum principle, it is necessary to define first the parabolic equivalent of the convex envelope of a function, which we will refer to as the monotone envelope.

2.4.1.1   Monotone Envelope of a Function

Definition 2.4.1 (Monotone envelope). If Ω is a convex set of $\mathbb{R}^d$ and (a, b) is an open interval, then the *monotone envelope* of a lower semi-continuous function $u : (a, b) \times \Omega \to \mathbb{R}$ is the largest function $v : (a, b) \times \Omega \to \mathbb{R}$ lying below u which is non-increasing with respect to t and convex with respect to x. It is denoted by $\Gamma(u)$.

Combining the usual definition of the convex envelope of a function with the non-increasing envelope of a function of one real variable, we obtain a first representation formula for $\Gamma(u)$.

Lemma 2.4.2 (Representation formula I).

$$\Gamma(u)(t, x) = \sup\{\xi \cdot x + h : \xi \cdot x + h \leq u(s, x) \text{ for all } s \in (a, t], x \in \Omega\}.$$

The set where $\Gamma(u)$ coincides with u is called the *contact set*; it is denoted by C_u. The following lemma comes from convex analysis, see e.g. [HUL].

Lemma 2.4.3. *Consider a point (t, x) in the contact set C_u of u. Then $\xi \cdot x + h = \Gamma(u)(t, x)$ if and only if ξ lies in the convex subdifferential $\partial u(t, x)$ of $u(t, \cdot)$ at x and $-h$ equals the convex conjugate $u^*(t, x)$ of $u(t, \cdot)$ at x.*

Recall that a convex function is locally Lipschitz continuous and in particular a.e. differentiable, for a.e. contact points, $(\xi, h) = (Du(t, x), u(t, x) - x \cdot Du(t, x))$. This is the reason why we next consider for $(t, x) \in (a, b) \times \Omega$ the following function

$$G(u)(t, x) = (Du(t, x), u(t, x) - x \cdot Du(t, x)).$$

The proof of the following elementary lemma is left to the reader.

Lemma 2.4.4. *If u is $C^{1,1}$ with respect to x and Lipschitz continuous with respect to t, then the function $G : (a, b) \times \Omega \to \mathbb{R}^{d+1}$ is Lipschitz continuous in (t, x) and for a.e. $(t, x) \in (a, b) \times \Omega$,*

$$\det D_{t,x} G(u) = u_t \det D^2 u.$$

We now give a second representation formula for $\Gamma(u)$ which will help us next to describe viscosity subdifferentials of the monotone envelope (see Lemma 2.4.6 below).

Lemma 2.4.5 (Representation formula II).

$$\Gamma(u)(t, x) = \inf \left\{ \sum_{i=1}^{d+1} \lambda_i u(s_i, x_i) : \sum_{i=1}^{d+1} \lambda_i x_i = x, s_i \in [a, t], \right.$$
$$\left. \sum_{i=1}^{d+1} \lambda_i = 1, \lambda_i \in [0, 1] \right\}. \qquad (2.38)$$

In particular, if

$$\Gamma(u)(t_0, x_0) = \sum_{i=1}^{d+1} \lambda_i u(t_i^0, x_i^0),$$

then

- *for all $i = 1, \ldots, d + 1$, $\Gamma(u)(t_i, x_i) = u(t_i, x_i)$;*
- *$\Gamma(u)$ is constant with respect to t and linear with respect to x in the convex set $\mathrm{co}\{(t, x_i^0), (t_i^0, x_i^0), i = 1, \ldots d + 1\}$.*

Proof. Let $\tilde{\Gamma}(u)$ denote the function defined by the right hand side of (2.38). First, we observe that $\tilde{\Gamma}(u)$ lies below u and is non-increasing with respect to t and convex with respect to x. Consider now another function v lying below u which is non-increasing with respect to t and convex with respect to x. We then have

$$u(t, x) \geq \tilde{\Gamma}(u)(t, x) \geq \tilde{\Gamma}(v)(t, x) \geq v(t, x).$$

The proof is now complete. □

We next introduce the notion of *harmonic sum of matrices*. For $A_1, A_2 \in \mathbb{S}_d$ such that $A_1 + A_2 \geq 0$, we consider

$$(A_1 \Box A_2)\zeta \cdot \zeta = \inf_{\zeta_1 + \zeta_2 = \zeta} \{A_1\zeta_1 \cdot \zeta_1 + A_2\zeta_2 \cdot \zeta_2\}.$$

The reader can check that if A_1 and A_2 are not singular, $A_1 \Box A_2 = (A_1^{-1} + A_2^{-1})^{-1}$. We can now state and prove

Lemma 2.4.6. *Let* $(\alpha, p, X) \in \mathcal{P}^- \Gamma(u)(t_0, x_0)$ *and*

$$\Gamma(u)(t_0, x_0) = \sum_{i=1}^{d+1} \lambda_i u(t_i^0, x_i^0). \tag{2.39}$$

Then for all $\varepsilon > 0$ *such that* $I + \varepsilon X > 0$, *there exist* $(\alpha_i, X_i) \in (-\infty, 0] \times \mathbb{S}_d$, $i = 1, \ldots, d+1$, *such that*

$$\begin{cases} (\alpha_i, p, X_i) \in \overline{\mathcal{P}^-} u(t_i^0, x_i^0) \\ \sum_{i=1}^{d+1} \lambda_i \alpha_i = \alpha \\ X_\varepsilon \leq \lambda_1^{-1} X_1 \Box \cdots \Box \lambda_{d+1}^{-1} X_{d+1} \end{cases} \tag{2.40}$$

where $X_\varepsilon = X \Box \varepsilon^{-1} I = (I + \varepsilon X)^{-1} X$.

Proof. We first define for two arbitrary functions $v, w : \mathbb{R}^d \to \mathbb{R}$,

$$v \overset{x}{\Box} w(x) = \inf_{y \in \mathbb{R}^d} v(x - y) + w(y).$$

For a given function $v : [0, +\infty) \times \mathbb{R}^d \to \mathbb{R}$, we also consider the non-increasing envelope $M[v]$ of v:

$$M[v](t, x) = \inf_{s \in [0,t]} v(s, x).$$

We now can write

$$\Gamma(u)(t, x) = \overset{x}{\underset{1 \leq i \leq d+1}{\Box}} M[u_i](t, x)$$

where

$$u_i(t, x) = \lambda_i u\left(t, \frac{x}{\lambda_i}\right).$$

Consider also $t_i^0 \in [0, t_0]$ such that

$$M[u_i](t_0, x_i^0) = u_i(t_i^0, x_i^0) = \lambda_i u\left(t_i^0, \frac{x_i^0}{\lambda_i}\right).$$

Lemma 2.4.6 is a consequence of the two following ones.

Lemma 2.4.7. *Consider* $(\alpha, p, X) \in \mathcal{P}^- V(t_0, x_0)$ *where*

$$V(t, x) = \underset{1 \le i \le d+1}{\overset{x}{\square}} v_i(t, x)$$

$$V(t_0, x_0) = \sum_{i=1}^{d+1} v_i(t_0, x_i^0).$$

Then for all $\varepsilon > 0$ *such that* $I + \varepsilon X > 0$, *there exist* $(\beta_i, Y_i) \in \mathbb{R} \times \mathbb{S}_d$ *such that we have*

$$(\beta_i, p, Y_i) \in \overline{\mathcal{P}^-} v_i(t_0, x_i^0)$$

$$\sum_{i=1}^{d+1} \beta_i = \alpha$$

$$X_\varepsilon \le \square_{i=1}^{d+1} Y_i.$$

Proof. We consider a test function ϕ touching V from below at (t_0, x_0) such that

$$(\alpha, p, X) = (\partial_t \phi, D\phi, D^2\phi)(t_0, x_0).$$

We write for (t, x_i) in a neighborhood of (t_0, x_i^0),

$$\phi\left(t, \sum_{i=1}^{d+1} x_i\right) - \phi\left(t_0, \sum_{i=1}^{d+1} x_i^0\right) \le \sum_{i=1}^{d+1} v_i(t, x_i) - \sum_{i=1}^{d+1} v_i(t_0, x_i^0).$$

Following [ALL97, Imb06], we conclude through Jensen–Ishii's lemma for $d + 1$ functions and general test functions (see Lemma 2.5.6 in appendix) that for all $\varepsilon > 0$ such that $I + d\varepsilon X > 0$, there exist $(\beta_i, Y_i) \in \mathbb{R} \times \mathbb{S}_d, i = 1, \ldots, d + 1$ such that

$$(\beta_i, p, Y_i) \in \overline{\mathcal{P}^-} v_i(t_0, x_i^0)$$

$$\sum_{i=1}^{d+1} \beta_i = \alpha$$

and

$$\begin{pmatrix} X & \dots & X \\ \vdots & \ddots & \vdots \\ X & \dots & X \end{pmatrix}_{\varepsilon} \leq \begin{pmatrix} Y_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & Y_{d+1} \end{pmatrix}$$

where, for any matrix A, $A_{\varepsilon} = (I + \varepsilon A)^{-1} A$. A small computation (presented e.g. in [Imb06, p. 796]) yields that the previous matrix inequality is equivalent to the following one

$$X_{d\varepsilon} \zeta \cdot \zeta \leq \sum_{i=1}^{d+1} Y_i \zeta_i \cdot \zeta_i$$

where $\zeta = \sum_{i=1}^{d+1} \zeta_i$. Taking the infimum over decompositions of ζ, we get the desired matrix inequality. $\qquad \square$

Lemma 2.4.8. *Consider* $s_1 \in [0, s_0]$ *such that*

$$M[v](s_0, y_0) = v(s_1, y_0).$$

Then for all $(\beta, q, Y) \in \mathcal{P}^- M[v](s_0, y_0)$,

$$(\beta, q, Y) \in \mathcal{P}^- v(s_1, y_0) \quad \text{and} \quad \beta \leq 0.$$

Proof. We consider the test function ϕ associated with (β, q, Y) and we write for h and δ small enough

$$\phi(s_0 + h, y_0 + \delta) - \phi(s_0, y_0) \leq M[v](s_0 + h, y_0 + \delta) - M[v](s_0, y_0)$$
$$\leq v(s_1 + h, y_0 + \delta) - v(s_1, y_0).$$

This implies $(\beta, q, Y) \in \mathcal{P}^- v(s_1, y_0)$. Moreover, choosing $\delta = 0$, we get

$$\phi(s_0 + h, y_0) \leq \phi(s_0, y_0)$$

and $\beta \leq 0$ follows. $\qquad \square$

The proof is now complete. $\qquad \square$

2.4.1.2   Statement

The following result is the first key result in the theory of regularity of fully non-linear parabolic equations. It is the parabolic counterpart of the famous Alexandroff estimate, also called Alexandroff–Bakelman–Pucci (ABP) estimate, see [CafCab] for more details about this elliptic estimate. The following one was first proved for linear equations by Krylov [Kryl76] and then extended by Tso [Tso85]. The following result appears in [Wang92a].

Theorem 2.4.9 (Maximum principle). *Consider a supersolution of* (2.36) *in* $Q_\rho = Q_\rho(0, 0)$ *such that* $u \geq 0$ *on* $\partial_p(Q_\rho)$. *Then*

$$\sup_{Q_\rho} u^- \leq C\rho^{\frac{d}{d+1}} \left(\int_{u=\Gamma(u)} (f^+)^{d+1} \right)^{\frac{1}{d+1}} \tag{2.41}$$

where C is universal and $\Gamma(u)$ is the monotone envelope of $\min(0, u)$ *extended by* 0 *to* $Q_{2\rho}$.

Remark 2.4.10. This is a maximum principle since, if $f \leq 0$, then u cannot take negative values.

Proof. We prove the result for $\rho = 1$ and the general one is obtained by considering $v(t, x) = u(\rho^2 t, \rho x)$. Moreover, replacing u with $\min(0, u)$ and extending it by 0 in $Q_2 \setminus Q_1$, we can assume that $u = 0$ on $\partial_p Q_1$ and $u \equiv 0$ in $Q_2 \setminus Q_1$.

We are going to prove the three following lemmas. Recall that $G(u)$ is defined page 44.

Lemma 2.4.11. *The function $\Gamma(u)$ is $C^{1,1}$ with respect to x and Lipschitz continuous with respect to t in Q_1. In particular, $G\Gamma(u) := G(\Gamma(u))$ is Lipschitz continuous with respect to (t, x).*

The second part of the statement of the previous lemma is a consequence of Lemma 2.4.4 above. We will prove the previous lemma together with the following one.

Lemma 2.4.12. *The partial derivatives $(\partial_t \Gamma(u), D^2\Gamma(u))$ satisfy for a.e. $(t, x) \in Q_1 \cap C_u$,*

$$-\partial_t \Gamma(u) + \lambda \Delta(\Gamma(u)) \leq f^+(x)$$

where $C_u = \{u = \Gamma(u)\}$.

The key lemma is the following one.

Lemma 2.4.13. *If M denotes $\sup_{Q_1} u^-$, then*

$$\{(\xi, h) \in \mathbb{R}^{d+1} : |\xi| \leq M/2 \leq -h \leq M\} \subset G\Gamma(u)(Q_1 \cap C_u)$$

where $C_u = \{u = \Gamma(u)\}$.

Before proving these lemmas, let us derive the conclusion of the theorem. Using successively Lemma 2.4.13, the area formula for Lipschitz maps (thanks to Lemma 2.4.11) and Lemma 2.4.4, we get

$$CM^{d+1} = |\{(\xi, h) \in \mathbb{R}^{d+1} : |\xi| \le M/2 \le -h \le M\}|$$
$$\le |G\Gamma(u)(Q_1 \cap C_u)|$$
$$\le \int_{Q_1 \cap C_u} |\det G\Gamma(u)|$$
$$\le \int_{Q_1 \cap C_u} -\partial_t \Gamma(u) \det(D^2 \Gamma(u)).$$

Now using the geometric–arithmetic mean inequality and Lemma 2.4.12, we get

$$CM^{d+1} \le \lambda^{-d} \int_{Q_1 \cap C_u} -\partial_t \Gamma(u) \det(\lambda D^2 \Gamma(u))$$
$$\le \frac{1}{\lambda^d (d+1)^{d+1}} \int_{Q_1 \cap C_u} (-\partial_t \Gamma(u) + \lambda \Delta(\Gamma(u))^{d+1}$$
$$\le C \int_{Q_1 \cap C_u} (f^+)^{d+1}$$

where C's are universal. □

We now turn to the proofs of Lemmas 2.4.11–2.4.13.

Proof of Lemmas 2.4.11 and 2.4.12. In order to prove that $\Gamma(u)$ is Lipschitz continuous with respect to t and $C^{1,1}$ with respect to x, it is enough to prove that there exists $C > 0$ such that

$$\forall (t, x) \in Q_2, \ \forall (\alpha, p, X) \in \mathcal{P}^- \Gamma(u)(t, x), \qquad \begin{cases} -\alpha \le C \\ X \le CI. \end{cases} \qquad (2.42)$$

Indeed, since $\Gamma(u)$ is non-increasing with respect to t and convex with respect to x, (2.42) yields that $\Gamma(u)$ is Lipschitz continuous with respect to t and $C^{1,1}$ with respect to x. See Lemma 2.5.8 in appendix for more details.

In order to prove (2.42), we first consider $(\alpha, p, X) \in \mathcal{P}^- \Gamma(u)(t, x)$ such that $X \ge 0$. Recall (cf. Lemma 2.4.6 above) that $\alpha \le 0$. We then distinguish two cases.

Assume first that $\Gamma(u)(t, x) = u(t, x)$. In this case, $(\alpha, p, X) \in \mathcal{P}^- u(t, x)$ and since u is a supersolution of (2.36), we have

$$\alpha - \lambda \operatorname{trace}(X) = \alpha + P^+(X) \ge f(x) \ge -C$$

where $C = |f|_{0;Q_1}$. Hence, we get (2.42) since $X \geq 0$ implies that $X \leq \text{trace}(X)I$. We also remark that the same conclusion holds true if $(\alpha, p, X) \in \overline{\mathcal{P}}^- \Gamma(u)(t, x)$ such that $X \geq 0$.

Assume now that $\Gamma(u)(t, x) < u(t, x)$. In this case, there exist $\lambda_i \in [0, 1]$, $i = 1, \ldots, d+1$, and $x_i \in Q_2$, $i = 1, \ldots, d+1$, such that (2.39) holds true with (t_0, x_0) and (t_i^0, x_i^0) replaced with (t, x) and (t_i, x_i). If $(t_i, x_i) \in Q_2 \setminus Q_1$ for two different i's, then Lemma 2.4.5 implies that $M = 0$ which is false. Similarly, $t_i > -1$ for all i. Hence, there is at most one index i such that $(t_i, x_i) \in Q_2 \setminus Q_1$ and in this case $(t_i, x_i) \in \partial_p Q_2$ and $t_i > -1$. In particular, $|x_i| = 2$. We thus distinguish two subcases.

Assume first that $(t_{d+1}, x_{d+1}) \in \partial_p Q_2$ with $t_{d+1} > -1$ and $(t_i, x_i) \in Q_1$ for $i = 1, \ldots, d$. In particular $|x_{d+1}| = 2$ and since $x \in Q_1$, we have $\lambda_{d+1} \leq \frac{2}{3}$. This implies that there exists λ_i such that $\lambda_i \geq (3d)^{-1}$. We thus can assume without loss of generality that $\lambda_1 \geq (3d)^{-1}$. Then from Lemma 2.4.6, we know that for all $\varepsilon > 0$ such that $I + \varepsilon X > 0$, there exist $(\alpha_i, X_i) \in \mathbb{R} \times \mathbb{S}_d$, $i = 1, \ldots, d+1$ such that (2.40) holds true. In particular,

$$X_\varepsilon \leq \frac{1}{\lambda_1} X_1 \leq 3d X_1.$$

Since $(\alpha_1, p, X_1) \in \overline{\mathcal{P}}^- u(t_1, x_1)$ and $\Gamma(u)(t_1, x_1) = u(t_1, x_1)$, we know from the discussion above that $X_1 \leq CI$. Hence for all ε small enough,

$$X_\varepsilon \leq 3d\, CI.$$

Letting $\varepsilon \to 0$ allows us to conclude that $X \leq 3d CI$ in the first subcase. As far as α is concerned, we remark that $\alpha_{d+1} = 0$ and $-\alpha_i \leq C$ for all $i = 1, \ldots, d+1$ so that

$$-\alpha = \sum_{i=1}^{d+1} \lambda_i(-\alpha_i) \leq C.$$

Assume now that all the points (t_i, x_i), $i = 1, \ldots, d+1$, are in Q_1. In this case, we have for all i that $-\alpha_i \leq C$ and $X_i \leq CI$ which implies

$$-\alpha = \sum_{i=1}^{d+1} \lambda_i(-\alpha_i) \leq C,$$

$$X_\varepsilon \leq \square_{i=1}^{d+1} \lambda_i^{-1} CI = CI.$$

We thus proved (2.42) in all cases where $X \geq 0$. Consider now a general subdifferential $(\alpha, p, X) \in \mathcal{P}^- \Gamma(u)(t, x)$. We know from Lemma 2.5.9 in appendix that there exists a sequence (α_n, p_n, X_n) such that

$$(\alpha_n, p_n, X_n) \in \mathcal{P}^-\Gamma(u)(t_n, x_n)$$

$$(t_n, x_n, \alpha_n, p_n) \to (t, x, \alpha, p)$$

$$X \le X_n + o_n(1), X_n \ge 0.$$

From the previous discussion, we know that

$$\alpha = \alpha_n + o_n(1) \le (C + 1)$$

$$X \le X_n + o_n(1) \le (C + 1)I$$

for all n. The proof is now complete.   □

Proof of Lemma 2.4.13. The supersolution $u \le 0$ is lower semi-continuous and the minimum $-M < 0$ in Q_2 is thus reached at some $(t_0, x_0) \in Q_1$ (since $u \equiv 0$ outside Q_1). Now pick (ξ, h) such that

$$|\xi| \le M/2 \le -h \le M.$$

We consider $P(y) = \xi \cdot y + h$. We remark that $P(y) < 0$ for $y \in Q_1$, hence $P(y) < u(0, y)$ in Q_1. Moreover, since $|x_0| < 1$,

$$P(x_0) - u(t_0, x_0) = \xi \cdot x_0 + h + M > h - |\xi| + M \ge 0$$

hence $\sup_{y \in Q_2}(P(y) - u(t_0, y)) \ge 0$. We thus choose

$$t_1 = \sup\{t \ge 0 : \forall s \in [0, t], \sup_{Q_2}(P(y) - u(s, y)) < 0\}.$$

We have $0 \le t_1 \le t_0$ and

$$0 = \sup_{Q_2}(P(y) - u(t_1, y)) = P(y_1) - u(t_1, y_1).$$

In particular, $\xi = Du(t_1, y_1)$ and $h = u(t_1, x_1) - \xi \cdot x_1$, that is to say, $(\xi, h) = G(u)(t_1, y_1)$ with $(t_1, y_1) \in C_u$.   □

2.4.2   The L^ε-Estimate

This subsection is devoted to the important "L^ε estimate" given in Theorem 2.4.15. This estimate is sometimes referred to as the weak Harnack inequality.

Theorem 2.4.15 claims that the L^ε-"norm" in a neighbourhood $\tilde{K}_1$ of $(0, 0)$ of a non-negative (super-)solution u of the model equation (2.36) can be controlled by its infimum over a neighbourhood $\tilde{K}_2$ of $(1, 0)$ plus the L^{d+1}-norm of f.
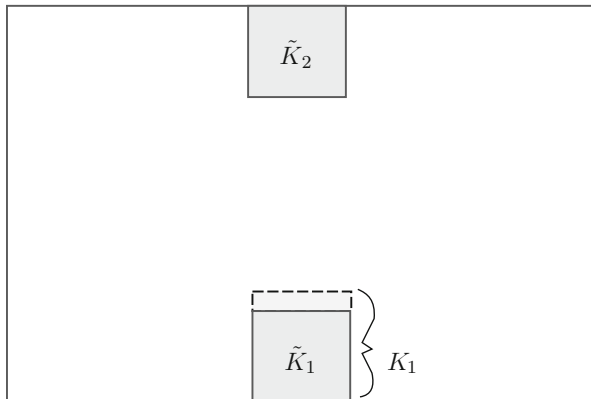
Fig. 2.1 The two neighbourhoods $\tilde{K}_1$ and $\tilde{K}_2$

Remark 2.4.14. Since ε can be smaller than 1, the integral of u^ε is in fact not the (ε-power of) a norm.

We introduce the two neighbourhoods mentioned above (see Fig. 2.1).

$$\tilde{K}_1 = (0, R^2/2) \times (-R, R)^d,$$
$$\tilde{K}_2 = (1 - R^2, 1) \times (-R, R)^d.$$

Theorem 2.4.15 (L^ε estimate). *There exist universal positive constants R, C and ε, such that for all non-negative supersolution u of*

$$\frac{\partial u}{\partial t} + P^+(D^2 u) \geq f \quad \text{in } (0, 1) \times B_{\frac{1}{R}}(0),$$

the following holds true

$$\left(\int_{\tilde{K}_1} u^\varepsilon \right)^{\frac{1}{\varepsilon}} \leq C (\inf_{\tilde{K}_2} u + \|f\|_{L^{d+1}((0,1) \times B_{\frac{1}{R}}(0))}). \tag{2.43}$$

The proof of this theorem is difficult and lengthy; this is the reason why we explain the main steps of the proof now.

First, one should observe that it is possible to assume without loss of generality that $\inf_{\tilde{K}_2} u \leq 1$ and $\|f\|_{L^{d+1}((0,1) \times B_{\frac{1}{R}}(0))} \leq \varepsilon_0$ (for some universal constant ε_0 to be determined) and to prove

$$\int_{\widetilde{K}_1} u^\varepsilon(t, x) dx \leq C$$

where $\varepsilon > 0$ and $C > 0$ are universal. We recall that a constant is said to be universal if it only depends on ellipticity constants λ and Λ and dimension d. Getting such an estimate is equivalent to prove that

$$|\{u > t\} \cap \tilde{K}_1| \le C t^{-\varepsilon}$$

(see page 69 for more details). To get such a decay estimate, it is enough to prove that

$$|\{u > N^k\} \cap \tilde{K}_1| \le C N^{-k\varepsilon}$$

for some universal constant $N > 1$. This inequality is proved by induction thanks to a covering lemma (see Lemma 2.4.27 below). This amounts to cut the set $\{u > N^k\} \cap \tilde{K}_1$ in small pieces (the dyadic cubes) and make sure that the pieces where u is very large ($u \ge t$, $t \gg 1$) have small measures.

This will be a consequence of a series of measure estimates obtained from a basic one. The proof of the basic measure estimate is a consequence of the maximum principle proved above and the construction of an appropriate barrier we will present soon. But we should first introduce the parabolic cubes we will use in the decomposition. We also present the choice of parameters we will make.

2.4.2.1   Parabolic Cubes and Choice of Parameters

We consider the following subsets of $Q_1(1, 0)$.

$$K_1 = (0, R^2) \times (-R, R)^d,$$
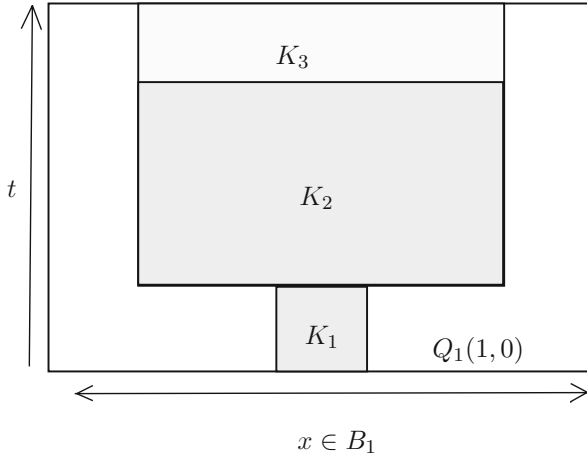$$K_2 = (R^2, 10R^2) \times (-3R, 3R)^d,$$
$$K_3 = (R^2, 1) \times (-3R, 3R)^d.$$

The constant R will be chosen as follows

$$R = \min\left(\frac{1}{3\sqrt{d}}, 3 - 2\sqrt{2}, \frac{1}{\sqrt{10(m+1)}}\right) \tag{2.44}$$

where m will be chosen in a universal way in the proof of the L^ε estimate.

2.4.2.2   A Useful Barrier

The following lemma will be used to derive the basic measure estimate. This estimate is the cornerstone of the proof of the L^ε estimate.

Fig. 2.2 The sets K_1, K_2 and K_3

Lemma 2.4.16. *For all* $R \in (0, \min((3\sqrt{d})^{-1}, (10)^{-1/2}))$*, there exists a nonnegative Lipschitz function* $\phi : Q_1(1,0) \to \mathbb{R}$*,* C^2 *with respect to* x *where it is positive, such that*

$$\frac{\partial \phi}{\partial t} + P^+(D^2\phi) \leq g$$

for some continuous bounded function $g : Q_1(1,0) \to \mathbb{R}$ *and such that*

$$\operatorname{supp} g \subset K_1$$
$$\phi \geq 2 \text{ in } K_3$$
$$\phi = 0 \text{ in } \partial_p Q_1(1,0).$$

Remark 2.4.17. Recall the definitions of K_1, K_2 and K_3 (see Fig. 2.2).

$$K_1 = (0, R^2) \times (-R, R)^d,$$
$$K_2 = (R^2, 10R^2) \times (-3R, 3R)^d,$$
$$K_3 = (R^2, 1) \times (-3R, 3R)^d.$$

The proof of the lemma above consists in constructing the function ϕ more or less explicitly. It is an elementary computation. However, it is an important feature of non divergence type equations that this type of computations can be made. Consider in contrast the situation of parabolic equations with measurable coefficients in divergence form. For that type of equations, a result like the one of Lemma 2.4.16 would be significantly harder to obtain.

Proof. We will construct a function φ which solves the equation

$$\varphi_t + P^+(D^2\varphi) \leq 0 \tag{2.45}$$

in the whole cylinder $Q_1(1,0)$, such that φ is positive and unbounded near $(0,0)$ but $\varphi = 0$ in $\partial_p Q_1(1,0)$ away from $(0,0)$, and moreover $\varphi > 0$ in K_2. Note that if the equation were linear, φ could be its heat kernel in the cylinder. Once we have this function φ, we obtain ϕ simply by taking

$$\phi(t,x) = 2\frac{\varphi(t,x)}{\min_{K_2}\varphi} \text{ for } (t,x) \in \setminus K_1,$$

and making ϕ equal any other smooth function in K_1 which is zero on $\{t = 0\}$.

We now construct this function φ. We will provide two different formulas for $\varphi(t,x)$. The first one will hold for $t \in (0,T)$ for some $T \in (0,1)$. Then the second formula provides a continuation of the definition of φ on $[T,1]$.

For some constant $p > 0$ and a function $\Phi : \mathbb{R}^d \to \mathbb{R}$, we will construct the function φ in $(0,T)$ with the special form

$$\varphi(t,x) = t^{-p}\Phi\left(\frac{x}{\sqrt{t}}\right).$$

Let us start from understanding what conditions Φ must satisfy in order for φ to be a subsolution to (2.45).

$$0 \geq \varphi_t + P^+(D^2\varphi) = t^{-1-p}\left(-p\Phi\left(\frac{x}{\sqrt{t}}\right)\right.$$
$$\left. -\frac{1}{2}\frac{x}{\sqrt{t}} \cdot \nabla\Phi\left(\frac{x}{\sqrt{t}}\right) + P^+(D^2\Phi)\left(\frac{x}{\sqrt{t}}\right)\right).$$

Therefore, we need to find a function $\Phi : \mathbb{R}^d \to \mathbb{R}$ and some exponent p such that

$$-p\Phi(x) - \frac{1}{2}x \cdot \nabla\Phi(x) + P^+(D^2\Phi)(x) \leq 0. \tag{2.46}$$

For some large exponent q, we choose Φ like this

$$\Phi(x) = \begin{cases} \text{something smooth and bounded between 1 and 2 if } |x| \leq 3\sqrt{d}, \\ (6\sqrt{d})^q(2^q-1)^{-1}\left(|x|^{-q} - (6\sqrt{d})^{-q}\right) & \text{if } 3\sqrt{d} \leq |x| \leq 6\sqrt{d}, \\ 0 & \text{if } |x| \geq 6\sqrt{d}. \end{cases}$$

For $3\sqrt{d} < |x| < 6\sqrt{d}$, we compute explicitly the second and third terms in (2.46),

$$-\frac{1}{2}x \cdot \nabla\Phi(x) = (6\sqrt{d})^q (2^q - 1)^{-1}\frac{q}{2}|x|^{-q}$$

$$P^+(D^2\Phi)(x) = (6\sqrt{d})^q (2^q - 1)^{-1} q(\Lambda(d-1) - \lambda(q+1))|x|^{-q-2}.$$

By choosing q large enough so that $\lambda(q + 1) > \Lambda(d - 1) + 18d$, we get that

$$-\frac{1}{2}x \cdot \nabla\Phi(x) + P^+\Phi(x) \le 0.$$

In order for (2.46) to hold in $B_{3\sqrt{d}}$, we just have to choose the exponent p large enough, since at those points $\Phi \ge 1$. Furthermore, since $\Phi \ge 0$ everywhere and $\Phi = 0$ outside $B_{6\sqrt{d}}$, then the inequality (2.46) holds in the full space $\mathbb{R}^d$ in the viscosity sense.

Since Φ is supported in $B_{6\sqrt{d}}$, then $\varphi = 0$ on $(0, T) \times \partial B_1$, for $T = (36d)^{-1}$. Thus, $\varphi = 0$ on the lateral boundary $(0, T) \times \partial B_1$. Moreover,

$$\lim_{t \to -1} \varphi(t, x) = 0,$$

uniformly in $B_1 \setminus B_\varepsilon$ for any $\varepsilon > 0$.

We have provided a value of φ up to time $T \in (0, 1)$. In order to continue φ in $[T, 1]$ we can do the following. Observe that by the construction of Φ, we have $P^+(D^2\varphi(T, x)) \le 0$ for $x \in B_1 \setminus B_{1/2}$ and $\varphi(x, T) \ge T^{-p}$ for $x \in B_{1/2}$. Therefore, let

$$C = \sup_{x \in B_1} \frac{P^+(D^2\varphi(T, x))}{\varphi(T, x)} < +\infty,$$

then we define $\varphi(t, x) = e^{-C(t-T)}\varphi(T, x)$ for all $t > T$, which is clearly a positive subsolution of (2.45) in $(T, 1] \times B_1$ with $\varphi = 0$ on $[T, 1] \times \partial B_1$.

The constructed function φ vanishes only on the set $\{(t, x) : t < T$ and $|x| \ge 6\sqrt{dt}\}$. Since the set $K_3 = (R^2, 1) \times (-3R, 3R)^d$ has no intersection with this set, then

$$\inf_{K_3} \varphi > 0.$$

This is all that was needed to conclude the proof. $\square$

2.4.2.3   The Basic Measure Estimate

As in the elliptic case, the basic measure estimate is obtained by combining the maximum principle of Theorem 2.4.9 and the barrier function constructed in Lemma 2.4.16. For the following proposition, we use the notation from Remark 2.4.17.
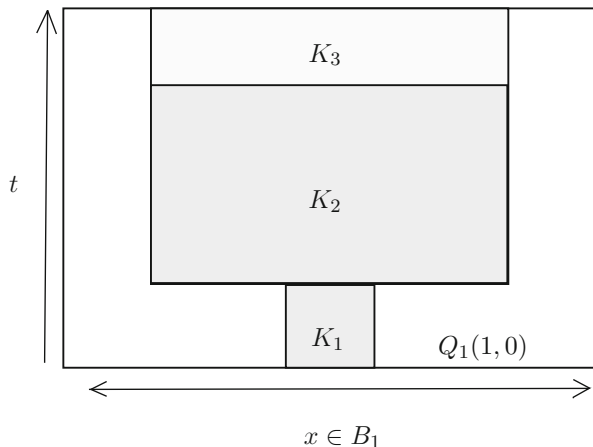
Fig. 2.3 Basic measure estimate in $Q_1(1,0)$

Proposition 2.4.18 (Basic measure estimate). *There exist universal constants $\varepsilon_0 \in (0,1)$, $M > 1$ and $\mu \in (0,1)$ such that for any non-negative supersolution of*

$$\frac{\partial u}{\partial t} + P^+(D^2u) \geq f \ \text{in } Q_1(1,0),$$

the following holds true: if

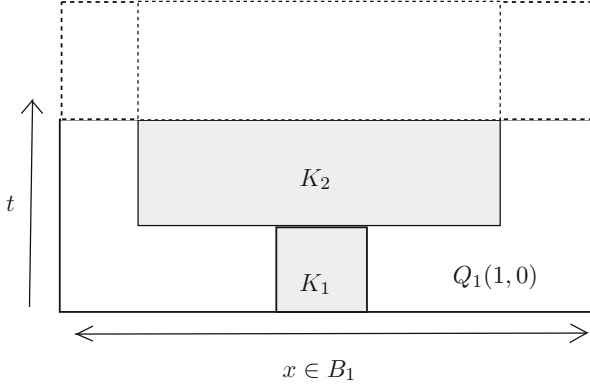$$\begin{cases} \inf_{K_3} u \leq 1 \\ \|f\|_{L^{d+1}(Q_1(1,0))} \leq \varepsilon_0 \end{cases}$$

then

$$|\{u \leq M\} \cap K_1| \geq \mu|K_1|.$$

Remark 2.4.19. Since $K_2 \subset K_3$ (see Fig. 2.3), the result also holds true if $\inf_{K_3} u$ is replaced with $\inf_{K_2} u$. This will be used in order to state and prove the stacked measure estimate.

Remark 2.4.20. If u is a non-negative supersolution of

$$\frac{\partial u}{\partial t} + P^+(D^2u) \geq f \ \text{in } (0,T) \times B_1,$$

Fig. 2.4 A supersolution in a smaller cylinder

for some $T \in (R^2, 1)$ (see Fig. 2.4), we still get

$$\left. \begin{array}{l} \inf_{(R^2,T) \times (-3R,3R)^d} u \leq 1 \\ \|f\|_{L^{d+1}((0,T) \times B_1)} \leq \varepsilon_0 \end{array} \right\} \Rightarrow |\{u \leq M\} \cap K_1| \geq \mu|K_1|.$$

The reason is that such a solution could be extended to $Q_1(1,0)$ (for example giving any boundary condition on $(T,1) \times \partial B_1$ and making f quickly become zero for $t > T$), and then Proposition 2.4.18 can be applied to this extended function. This remark will be useful when getting the "stacked" measure estimate in the case where the stack of cubes reaches the final time.

Proof. Consider the function $w = u - \phi$ where ϕ is the barrier function from Lemma 2.4.16. Then w satisfies (in the viscosity sense)

$$\frac{\partial w}{\partial t} + P^+(D^2 w) \geq \frac{\partial u}{\partial t} + P^+(D^2 u) - \frac{\partial \phi}{\partial t} - P^+(D^2 \phi) \geq f - g.$$

Remark also that

- $w \geq u \geq 0$ on $\partial_p Q_1(1,0)$;
- $\inf_{K_3} w \leq \inf_{K_3} u - 2 \leq -1$ so that $\sup_{K_3} w^- \geq 1$;
- $\{\Gamma(w) = w\} \subset \{w \leq 0\} \subset \{u \leq \phi\}$.

We recall that $\Gamma(w)$ denotes the monotone envelope of $\min(w, 0)$ extended by 0 to $Q_2(1,0)$. We now apply the maximum principle (Theorem 2.4.9) and we get

$$1 \leq \sup_{K_3} w^- \leq \sup_{Q_1} w^- \leq C_{\max} \|f\|_{L^{d+1}(Q_1(1,0))} + C_{\max} \left(\int_{\{u \leq \phi\}} |g|^{d+1} \right)^{\frac{1}{d+1}}.$$

Remember now that supp $g \subset K_1$ and get

$$1 \le C_{\max}\varepsilon_0 + C_{\max}|\{u \le M\} \cap K_1|$$

with $M > \max(\sup_{K_1} \phi, 1)$. Choose now ε_0 so that $C_{\max}\varepsilon_0 \le 1/2$ and get the result with $\mu = \frac{1}{1+2C_{\max}|K_1|}$. The proof is now complete. $\qquad \square$

Corollary 2.4.21 (Basic measure estimate scaled). *For the same constants ε_0, M and μ of Proposition 2.4.18 and any $x_0 \in \mathbb{R}^d$, $t_0 \in \mathbb{R}$ and $h > 0$, consider any nonnegative supersolution of*

$$\frac{\partial u}{\partial t} + P^+(D^2 u) \ge f \ \text{in} \ (t_0, x_0) + \rho Q_1(1, 0).$$

If

$$\|f\|_{L^{d+1}((t_0, x_0) + \rho Q_1(0,1))} \le \varepsilon_0 \frac{h}{M \rho^{d/(d+1)}}$$

then

$$|\{u > h\} \cap \{(t_0, x_0) + \rho K_1\}| > (1 - \mu)|(t_0, x_0) + \rho K_1| \Rightarrow u > \frac{h}{M} \ \text{in} \ (t_0, x_0) + \rho K_3.$$

Here, we recall that by ρK we mean $\{(\rho^2 t, \rho x) : (t, x) \in K\}$.

Remark 2.4.22. As in Remark 2.4.20, $(t_0, x_0) + \rho(0, 1) \times B_{\frac{1}{R}}(0)$ can be replaced with $(t_0, x_0) + \rho(0, T) \times B_{\frac{1}{R}}(0)$ for any $T \in (0, 1)$.

Proof. We consider the scaled function

$$v(t, x) = M h^{-1} u(t_0 + \rho^2 t, x_0 + \rho x).$$

This function solves the equation

$$\frac{\partial v}{\partial t} + P^+(D^2 v) \ge \tilde{f} \ \text{in} \ Q_1(1, 0)$$

where $\tilde{f}(t, x) = M h^{-1} \rho^2 f(t_0 + \rho^2 t, x_0 + \rho x)$. Note that

$$\|\tilde{f}\|_{L^{d+1}(Q_1(1,0))} = M h^{-1} \rho^{d/(d+1)} \|f\|_{L^{d+1}((t_0, x_0) + \rho Q_1(1,0))} \le \varepsilon_0.$$

We conclude the proof applying Proposition 2.4.18 to v. $\qquad \square$

2.4.2.4   Stacks of Cubes

Given $\rho \in (0, 1)$, we consider for all $k \in \mathbb{N}$, $k \geq 1$,

$$K_2^{(k)} = (\alpha_k R^2, \alpha_{k+1} R^2) \times (-3^k R, 3^k R)^d$$

where $\alpha_k = \sum_{i=0}^{k-1} 9^i = \frac{9^k - 1}{8}$.

The first stack of cubes that we can consider is the following one

$$\cup_{k \geq 1} K_2^{(k)}.$$

This stack is obviously not contained in $Q_1(1, 0)$ since time goes to infinity. It can spill out of $Q_1(1, 0)$ either on the lateral boundary or at the final time $t = 1$. We are going to see that at the final time, the "x-section" is contained in $(-3, 3)^d$.

We consider a scaled version of K_1 included in K_1 and we stack the corresponding $K_2^{(k)}$'s. The scaled versions of K_1, K_2 and $K_2^{(k)}$ are

$$\rho K_1 = (0, \rho^2 R^2) \times B_{\rho R}(0),$$

$$\rho K_2 = (\rho^2 R^2, 10\rho^2 R^2) \times B_{\rho R}(0),$$

$$\rho K_2^{(k)} = (\alpha_k \rho^2 R^2, \alpha_{k+1} \rho^2 R^2) \times (-3^k \rho R, 3^k \rho R)^d.$$

We now consider

$$L_1 = (t_0, x_0) + \rho K_1 \subset K_1$$

and

$$L_2^{(k)} = (t_0, x_0) + \rho K_2^{(k)}.$$

Lemma 2.4.23 (Stacks of cubes). *Choose $R \leq \min(3 - 2\sqrt{2}, \sqrt{\frac{2}{5}}) = 3 - 2\sqrt{2}$. For all $L_1 = (t_0, x_0) + \rho K_1 \subset K_1$, we have*
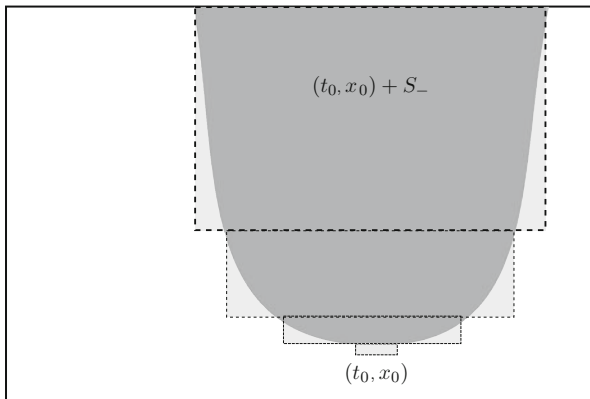
$$\tilde{K}_2 \subset \left(\cup_{k \geq 1} L_2^{(k)} \right) \cap (0, 1) \times (-3, 3)^d = \left(\cup_{k \geq 1} L_2^{(k)} \right) \cap \{0 < t < 1\}.$$

In particular, if moreover $R \leq (3\sqrt{d})^{-1}$,

$$\left(\cup_{k \geq 1} L_2^{(k)} \right) \subset (0, 1) \times B_{\frac{1}{R}}(0).$$

Moreover, the first $k^* = k$ such that $L_2^{(k+1)} \cap \{t = 1\} = \emptyset$ satisfies

$$\rho^2 R^2 \leq \frac{1}{\alpha_{k*}}.$$

Fig. 2.5 Stacks of cubes

Proof. We first remark that the stack of cubes lies between two "square" paraboloids (see Fig. 2.5)

$$(t_0, x_0) + S_- \subset \cup_{k \geq 1} L_2^{(k)} \subset (t_0, x_0) + S_+$$

where

$$S_\pm = \cup_{s \geq s_\pm} \{p_\pm(s)\} \times (-s, s)^d$$

and $p_\pm(s_\pm) = \rho^2 R^2$ and $p_\pm(z) = a_\pm z^2 + b_\pm \rho^2 R^2$ are such that

$$p_+(3^k \rho R) = \alpha_k \rho^2 R^2$$
$$p_-(3^k \rho R) = \alpha_{k+1} \rho^2 R^2.$$

This is equivalent to

$$a_+ = \frac{1}{8} \quad \text{and} \quad a_- = \frac{9}{8} \quad \text{and} \quad b_+ = b_- = -\frac{1}{8} \quad \text{and} \quad s_\pm = \sqrt{\frac{9}{8}} \rho R.$$

Remark now that

$$[(t_0, x_0) + S_+] \cap Q_1(1, 0) \subset [0, 1] \times (-R - a_+^{-\frac{1}{2}}, R + a_+^{-\frac{1}{2}})^d.$$

We thus choose R such that $(R + a_+^{-\frac{1}{2}}) \leq 3$. This condition is satisfied if

$$R \leq 3 - 2\sqrt{2}.$$

Remark next that

$$(t_0, x_0) + S_- \supset \cap_{x \in (-R,R)^d} [(R^2, x) + S_-].$$

Hence

$$[(t_0, x_0) + S_-] \cap Q_1(1, 0) \supset \tilde{K}_2$$

as soon as

$$a_+ (2R)^2 \leq 1 - 2R^2.$$

It is enough to have

$$\frac{5}{2} R^2 = (4a_+ + 2)R^2 \leq 1.$$

Finally, the integer k^* satisfies

$$t_0 + \alpha_{k^*} R^2 \rho^2 \leq 1 < t_0 + \alpha_{k^*+1} R^2 \rho^2. \qquad \square$$

2.4.2.5   The Stacked Measure Estimate

In this paragraph, we apply repeatedly the basic measure estimate obtained above and get an estimate in the finite stacks of cubes we constructed in the previous paragraph.

Proposition 2.4.24 (Stacked measure estimate). *For the same universal constants $\varepsilon_0 \in (0, 1)$, $M > 1$ and $\mu \in (0, 1)$ from Proposition 2.4.18, the following holds true: consider a non-negative supersolution u of*

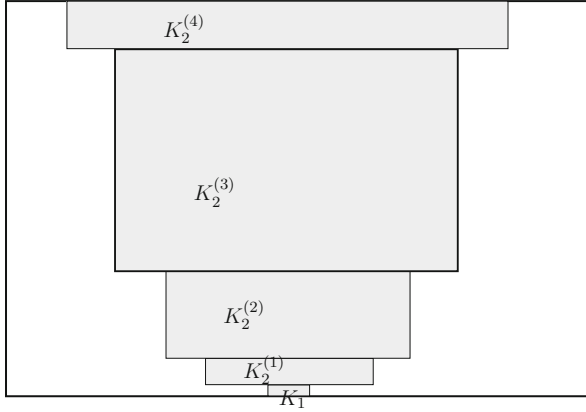$$\frac{\partial u}{\partial t} + P^+ (D^2 u) \geq f \ in \ (0, 1) \times B_{\frac{1}{R}}(0)$$

and a cube $L_1 = (t_0, x_0) + \rho K_1 \subset K_1$. Assume that for some $k \geq 1$ and $h > 0$

$$\|f\|_{L^{d+1}((0,1) \times B_{\frac{1}{R}}(0))} \leq \varepsilon_0 \frac{h}{M^k \rho^{d/(d+1)}}.$$

Then

$$|\{u > h\} \cap L_1| > (1 - \mu)|L_1| \Rightarrow \inf_{L_2^{(k)} \cap \{0 < t < 1\}} u > \frac{h}{M^k}.$$

Fig. 2.6 Stacks of neighbourhoods $K_2^{(k)}$

Remark 2.4.25. Remember that $L_2^{(k)} = (t_0, x_0) + \rho K_2^{(k)}$ and see Fig. 2.6. Thanks to Lemma 2.4.23, we know that $L_2^{(k)} \cap \{0 < t < 1\} \subset (0, 1) \times B_{\frac{1}{R}}(0)$.

Proof. We prove the result by induction on k. Corollary 2.4.21 corresponds to the case $k = 1$ if we can verify that

$$\|f\|_{L^{d+1}((t_0, x_0) + \rho Q_1(1,0))} \leq \varepsilon_0 \frac{h}{M \rho^{d/(d+1)}}.$$

It is a consequence of the fact that $Q_1(1, 0) \subset (0, 1) \times B_{\frac{1}{R}}(0)$.

For $k > 1$, the inductive hypothesis reads

$$\inf_{L_2^{(k-1)} \cap \{0 < t < 1\}} u > \frac{h}{M^{k-1}}.$$

If $L_2^{(k-1)}$ is not contained in $(0, 1) \times B_{\frac{1}{R}}(0)$, there is nothing to prove at rank k since $L_2^{(k)} \cap \{0 < t < 1\} = \emptyset$. We thus assume that $L_2^{(k-1)} \subset (0, 1) \times B_{\frac{1}{R}}(0)$.

In particular

$$|\{u > \frac{h}{M^{k-1}}\} \cap L_2^{(k-1)}| = |L_2^{(k-1)}|. \tag{2.47}$$

Note that $L_2^{(k-1)} = (t_1, 0) + \rho_1 K_1$ and $L_2^{(k)} = (t_1, 0) + \rho_1 K_2$ with $t_1 = t_0 + \alpha_{k-1} R^2 \rho^2$ and $\rho_1 = 3^{k-1}\rho$. In particular (2.47) implies

$$|\{u > \frac{h}{M^{k-1}}\} \cap \{(t_1, 0) + \rho_1 K_1\}| > (1 - \mu)|(t_1, 0) + \rho_1 K_1|.$$

So we apply Corollary 2.4.21 again to obtain

$$\inf_{L_2^{(k)} \cap \{0 < t < 1\}} u = \inf_{\{(t_1, 0) + \rho_1 K_2\} \cap \{0 < t < 1\}} u > \frac{h}{M^k}.$$

We can do so since $\rho_1 \geq \rho$ and Lemma 2.4.23 implies that $L_2^{(k)} \subset (0, 1) \times (-3, 3)^d$. In particular, the corresponding domain in which the supersolution is considered is contained in $(0, 1) \times B_{\frac{1}{R}}(0)$. We used here Remark 2.4.20 when $(t_1, 0) + \rho_1 K_2$ is not contained in $\{0 < t < 1\}$. Thus, we finish the proof by induction. □

Before turning to the proof of Theorem 2.4.15, we observe that the previous stacked measure estimate implies in particular the following result.

Corollary 2.4.26 (Straight stacked measure estimate). *Assume that* $R \leq \frac{1}{\sqrt{10(m+1)}}$. *Under the assumptions of Proposition 2.4.24 with* $k = m$, *for any cube* $L_1 \subset K_1$

$$|\{u > h\} \cap L_1| > (1 - \mu)|L_1| \Rightarrow u > \frac{h}{M^m} \ in \ \overline{L_1}^{(m)} \subset Q_1(1, 0).$$

Proof. Apply Proposition 2.4.24 with $k = m$ and remark that $\overline{L_1}^{(m)} \subset L_2^{(m)}$. The fact that $\overline{L_1}^{(m)} \subset Q_1(1, 0)$ (see Fig. 2.7) comes from the fact that $10(m + 1)$ $R^2 \leq 1$. □

2.4.2.6   A Stacked Covering Lemma

When proving the fundamental L^ε-estimate (sometimes called the weak Harnack inequality) for fully nonlinear elliptic equations, the Calderón–Zygmund decomposition lemma plays an important role (see [CafCab] for instance). It has to be adapted to the parabolic framework.

We need first some definitions. A *cube* Q is a set of the form $(t_0, x_0) + (0, s^2) \times (-s, s)^d$. A *dyadic cube* K *of* Q is obtained by repeating a finite number of times the following iterative process: Q is divided into 2^{d+2} cubes by considering all the translations of $(0, s^2/4) \times (0, s)^d$ by vectors of the form $(l(s^2/4), sk)$ with $k \in \mathbb{Z}^d$ and $l \in \mathbb{Z}$ included in Q. When a cube K_1 is split in different cubes including K_2, K_1 is called a *predecessor* of K_2.

Given $m \in \mathbb{N}$, and a dyadic cube K of Q, the set $\bar{K}^{(m)}$ is obtained by "stacking" m copies of its predecessor $\bar{K}$. More rigorously, if the predecessor $\bar{K}$ has the form $(a, b) \times L$, then we define $\bar{K}^{(m)} = (b, b + m(b - a)) \times L$. Figure 2.8 corresponds to the case $m = 3$.

Lemma 2.4.27 (Stacked covering lemma). *Let* $m \in \mathbb{N}$. *Consider two subsets* A *and* B *of a cube* Q. *Assume that* $|A| \leq \delta|Q|$ *for some* $\delta \in (0, 1)$. *Assume also the following: for any dyadic cube* $K \subset Q$,

Fig. 2.7 L_1 and the predecessors $\overline{L_1}$ and $\overline{L_1}^{(3)}$

$$|K \cap A| > \delta|A| \Rightarrow \bar{K}^m \subset \mathcal{B}.$$

Then $|A| \le \delta \frac{m+1}{m}|\mathcal{B}|$.

Remark 2.4.28. This lemma is implicitly used in [Wang92a] (see e.g. Lemma 3.23 of this paper) but details of the proof are not given.

The proof uses a Lebesgue's differentiation theorem with assumptions that are not completely classical, even if we believe that such a generalization is well-known. For the sake of completeness, we state and prove it in appendix (see Theorem 2.5.1 and Corollary 2.5.2).

Proof of Lemma 2.4.27. By iterating the process described to define dyadic cubes, we know that there exists a countable collection of dyadic cubes K_i such that

$$|K_i \cap A| \ge \delta|K_i| \quad \text{and} \quad |\bar{K}_i \cap A| \le \delta|\bar{K}_i|$$

where $\bar{K}_i$ is a predecessor of K_i. We claim that thanks to Lebesgue's differentiation theorem (Corollary 2.5.2), there exists a set N of null measure such that

$$A \subset (\cup_{i=1}^{\infty} K_i) \cup N.$$
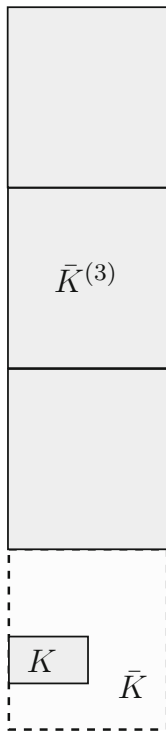
Fig. 2.8 A dyadic cube K and stacked predecessors $\bar{K}^{(m)}$

Indeed, consider $(t, x) \in A \setminus \cup_{i=1}^{\infty} K_i$. On one hand, since $(t, x) \in Q$, it belongs to a sequence of closed dyadic cubes of the form $L_j = (t_j, x_j) + [0, r_j^2] \times [-r_j, r_j]^d$ with $r_j \to 0$ as $j \to +\infty$ such that

$$|A \cap L_j| \leq \delta |L_j|$$

that is to say

$$\fint_{L_j} \mathbf{1}_A \leq \delta < 1.$$

On the other hand, for $(t, x) \in A \setminus \cup_{i=1}^{\infty} K_i$,

$$0 < 1 - \delta \leq 1 - \fint_{L_j} \mathbf{1}_A = \fint_{L_j} |\mathbf{1}_A - \mathbf{1}_A(t, x)|.$$

We claim that the right hand side of the previous equality goes to 0 as $j \to \infty$ as soon as $(t, x) \notin N$ where N is a set of null measure. Indeed, Corollary 2.5.2 implies that for (t, x) outside of such a set N,

$$\fint_{L_j} |\mathbf{1}_A - \mathbf{1}_A(t, x)| \le \fint_{\tilde{L}_j} |\mathbf{1}_A - \mathbf{1}_A(t, x)| \to 0$$

where $\tilde{L}_j = (t, x) + [0, 4r_j^2] \times [-2r_j, 2r_j]^d$. We conclude that $A \setminus \cup_i K_i \subset N$.

We can relabel predecessors $\bar{K}_i$ so that they are pairewise disjoint. We thus have $A \subset \cup_{i=1}^\infty K_i \cup N$ with $\bar{K}_i^m \subset \mathcal{B}$ thanks to the assumption; in particular,

$$A \subset \cup_{i=1}^\infty K_i \cup N \subset \cup_{i=1}^\infty \bar{K}_i \cup \bar{K}_i^m \cup N$$

with $\cup_{i=1}^\infty \bar{K}_i^m \subset \mathcal{B}$. Classically, we write

$$|A| \le \sum_{i \ge 1} |A \cap \bar{K}_i| \le \delta \sum_{i \ge 1} |\bar{K}_i| \le \delta| \cup_{i=1}^\infty \bar{K}_i|. \tag{2.48}$$

In order to conclude the proof of the lemma, it is thus enough to prove that for a countable collection $(\bar{K}_i)_i$ of disjoint cubes, we have

$$| \cup_{i=1}^\infty \bar{K}_i \cup \bar{K}_i^m | \le \frac{m}{m + 1} | \cup_{i=1}^\infty \bar{K}_i^m |. \tag{2.49}$$

Indeed, combining (2.48) and (2.49) yields the desired estimate (keeping in mind that $\cup_i \bar{K}_i^m \subset \mathcal{B}$).

Estimate (2.49) is not obvious since, even if the $\bar{K}_i$'s are pairwise disjoint, the stacked cubes $\bar{K}_i^m$ can overlap. In order to justify (2.49), we first write

$$\cup_{i=1}^\infty \bar{K}_i \cup \bar{K}_i^m = \cup_{j=1}^\infty J_j \times L_j$$

where L_j are disjoint cubes of $\mathbb{R}^d$ and J_j are open sets of $\mathbb{R}$ of the form

$$J = \cup_{k=1}^\infty (a_k, a_k + (m + 1)h_k).$$

Remark that

$$\cup_{i=1}^\infty \bar{K}_i^m = \cup_{j=1}^\infty \tilde{J}_j \times L_j$$

where $\tilde{J}_j$ has the general form

$$\tilde{J} = \cup_{k=1}^\infty (a_k + h_k, a_k + (m + 1)h_k).$$

Hence, the proof is complete once Lemma 2.4.29 below is proved. □

Lemma 2.4.29. *Let* $(a_k)_{k=1}^N$ *and* $(h_k)_{k=1}^N$ *be two (possibly infinite) sequences of real numbers for* $N \in \mathbb{N} \cup \{\infty\}$ *with* $h_k > 0$ *for* $k = 1, \ldots, N$. *Then*

$$\left| \cup_{k=1}^N (a_k, a_k + (m+1)h_k) \right| \leq \frac{m}{m+1} \left| \cup_{k=1}^N (a_k + h_k, a_k + (m+1)h_k) \right|.$$

Proof. We first assume that N is finite. We write $\cup_{k=1}^N (a_k + h_k, a_k + (m+1)h_k)$ as $\cup_{l=1}^L I_l$ where I_l are disjoint open intervals. We can write them as

$$I_l = \cup_{k=1}^{N_l} (b_k + l_k, b_k + (m+1)l_k) = (\inf_{k=1,\ldots,N_l} (b_k + l_k), \sup_{k=1,\ldots,N_l} (b_k + (m+1)l_k)).$$

Pick k_l such that $\inf_{k=1,\ldots,N_l} (b_k + l_k) = b_{k_l} + l_{k_l}$. In particular,

$$|I_l| = \sup_{k=1,\ldots,N_l} (b_k + (m+1)l_k)) - \inf_{k=1,\ldots,N_l} (b_k + l_k)$$

$$\geq m l_{k_l}.$$

Then

$$\left| \cup_{k=1}^N (a_k + h_k, a_k + (m+1)h_k) \right| \geq m \sum_l l_{k_l} = \frac{m}{m+1} \sum_l (m+1) l_{k_l}.$$

It is now enough to remark that $(m+1)l_{k_l}$ coincide with the length of one of the intervals $\{(a_k, a_k + (m+1)h_k)\}_k$ and they are distinct since so are the I_l's. The proof is now complete in the case where N is finite.

If now $N = \infty$, we get from the previous case that for any $N \in \mathbb{N}$,

$$\left| \cup_{k=1}^N (a_k, a_k + (m+1)h_k) \right| \leq \frac{m}{m+1} \left| \cup_{k=1}^N (a_k + h_k, a_k + (m+1)h_k) \right|$$

$$\leq \frac{m}{m+1} \left| \cup_{k=1}^\infty (a_k + h_k, a_k + (m+1)h_k) \right|.$$

It is now enough to let $N \to \infty$ to conclude. $\qquad\qquad\qquad\qquad\qquad\qquad$ □

2.4.2.7 Proof of the L^ε-Estimate

The proof of the L^ε estimate consists in obtaining a decay in the measure of the sets $\{u > M^k\} \cap \tilde{K}_1$ (see Fig. 2.9). As in the elliptic case, the strategy is to apply the covering Lemma 2.4.27 iteratively making use of Corollary 2.4.26. The main difficulty of the proof (which is not present in the elliptic case) comes from the fact that if K is a cube contained in $\tilde{K}_1$, then nothing prevents $\bar{K}^{(m)}$ spilling out of K_1.
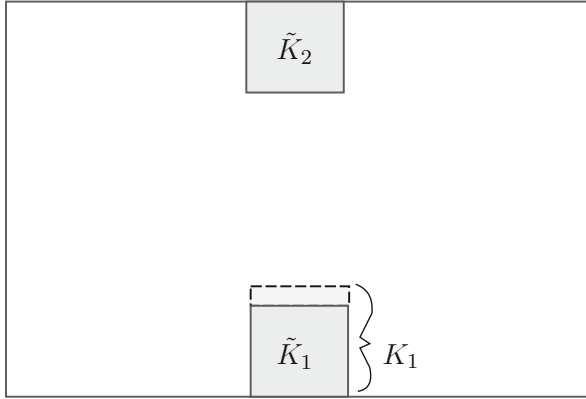
Fig. 2.9 The two neighbourhoods $\tilde{K}_1$ and $\tilde{K}_2$

Proof of Theorem 2.4.15. First, we can assume that

$$\inf_{\tilde{K}_2} u \leq 1 \quad \text{and} \quad \|f\|_{L^{d+1}((0,1)\times B_{\frac{1}{R}}(0))} \leq \varepsilon_0$$

(where ε_0 comes from Proposition 2.4.24) by considering

$$v^\delta(t,x) = \frac{u}{\inf_{\tilde{K}_2} u + \varepsilon_0^{-1}\|f\|_{L^{d+1}((0,1)\times B_{\frac{1}{R}}(0))} + \delta}.$$

We thus want to prove that there exits a universal constant $C > 0$ such that

$$\int_{\tilde{K}_1} u^\varepsilon(t,x)\, dt\, dx \leq C. \tag{2.50}$$

In order to get (2.43), it is enough to find universal constants $m, k_0 \in \mathbb{N}$ and $B > 1$ such that for all $k \geq k_0$,

$$|\{u > M^{km}\} \cap (0, R^2/2 + C_1 B^{-k}) \times (-R, R)^d| \leq C(1-\mu/2)^k \tag{2.51}$$

where C is universal and M and μ comes from Proposition 2.4.24. Indeed, first for $t \in [M^{km}, M^{(k+1)m})$, we have

$$|\{u > t\} \cap (0, R^2/2 + C_1 B^{-k}) \times (-R, R)^d| \leq C(1-\mu/2)^k \leq Ct^{-\varepsilon}$$

with $\varepsilon = -\frac{\ln(1-\mu/2)}{m \ln M} > 0$. We deduce that for all $t > 0$, we have

$$|\{u > t\} \cap \tilde{K}_1| \leq Ct^{-\varepsilon}.$$

Now we use the formula

$$\int_{\tilde{K}_1} u^{\varepsilon}(t, x) \, dt \, dx = \varepsilon \int_0^{\infty} \tau^{\varepsilon-1} |\{u > \tau\} \cap \tilde{K}_1| d\tau$$

$$\leq \varepsilon |\tilde{K}_1| \int_0^1 \tau^{\varepsilon-1} d\tau + \varepsilon \int_1^{\infty} \tau^{\varepsilon-1} |\{u > \tau\} \cap \tilde{K}_1| d\tau$$

and we get (2.50) from (2.51).

We prove (2.51) by induction on k. For $k = k_0$, we simply choose

$$C \geq (1 - \mu/2)^{-k_0} |(0, R^2/2 + C_1 B^{-1}) \times (-R, R)^d|.$$

Now we assume that $k \geq k_0$, that the result holds true for k and we prove it for $k + 1$. In order to do so, we want to apply the covering Lemma 2.4.27 with

$$A = \{u > M^{(k+1)m}\} \cap (0, R^2/2 + C_1 B^{-k-1}) \times (-R, R)^d$$

$$B = \{u > M^{km}\} \cap (0, R^2/2 + C_1 B^{-k}) \times (-R, R)^d$$

$$Q = K_1 = (0, R^2) \times (-R, R)^d$$

for some universal constants B and C_1 to be chosen later. We can choose k_0 (universal) so that $B \subset K_1$. For instance

$$2C_1 B^{-k_0} \leq R^2.$$

The induction assumption reads

$$|B| \leq C(1 - \mu/2)^k.$$

Lemma 2.4.30. *We have $|A| \leq (1 - \mu)|Q|$.*

Proof. Since, $\inf_{\tilde{K}_2} u \leq 1$, we have in particular $\inf_{K_3} u \leq 1$. The basic measure estimate (Proposition 2.4.18) then implies that

$$|A| \leq |\{u > M\} \cap K_1\}| \leq (1 - \mu)|K_1| = (1 - \mu)|Q|. \qquad \square$$

Lemma 2.4.31. *Consider any dyadic cube $K = (t, x) + \rho K_1$ of Q. If*

$$|K \cap \{u > M^{(k+1)m}\} \cap (0, R^2/2 + C_1 B^{-k-1}) \times (-R, R)^d\}| > (1-\mu)|K|, \quad (2.52)$$

then

$$\bar{K}^m \subset \{u > M^{km}\} \cap (0, R^2/2 + C_1 B^{-k}) \times (-R, R)^d$$

where $\bar{K}^m$ is defined at the beginning of Sect. 2.4.2.6.

Proof. We remark that the straight stacked measure estimate, Corollary 2.4.26, applied with $h = M^{(k+1)m} \geq M^m$, implies

$$\bar{K}^m \subset \{u > M^{km}\}.$$

We thus have to prove that

$$\bar{K}^m \subset [0, R^2/2 + C_1 B^{-k}] \times (-R, R)^d. \qquad (2.53)$$

Because of (2.52), we have

$$K \cap (0, R^2/2 + C_1 B^{-k-1}) \times (-R, R)^d \neq \emptyset.$$

Hence

$$\bar{K}^m \subset [0, R^2/2 + C_1 B^{-k-1} + \text{height}(\bar{K}) + \text{height}(\bar{K}^m)] \times (-R, R)^d$$

where $\text{height}(L) = \sup\{t : \exists x, (t, x) \in L\} - \inf\{t : \exists x, (t, x) \in L\}$. Moreover,

$$\text{height}(K) = R^2 \rho^2$$
$$\text{height}(\bar{K}) = 4 \, \text{height}(K)$$
$$\text{height}(\bar{K}^m) = m \, \text{height}(\bar{K}).$$

Hence, (2.53) holds true if

$$R^2/2 + C_1 B^{-k-1} + 4(m+1)R^2 \rho^2 \leq R^2/2 + C_1 B^{-k}$$

i.e.

$$R^2 \rho^2 \leq \frac{C_1(B-1)}{4(m+1)} B^{-k-1}. \qquad (2.54)$$

In order to estimate $R^2 \rho^2$ we are going to use the stacked measure estimate given by Proposition 2.4.24 together with the fact that K is a cube for which (2.52) holds.

On one hand, Proposition 2.4.24 and (2.52) imply that as long as $l \leq (k+1)m$, we have

$$u > M^{(k+1)m-l} \text{ in } L_2^{(l)} \cap \{0 < t < 1\};$$

in particular,

$$\inf_{\cup_{l=1}^{(k+1)m} L_2^{(l)} \cap \{0<t<1\}} u > 1.$$

On the other hand, using notation from Lemma 2.4.23,

$$\inf_{\cup_{l=1}^{k^*+1} L_2^{(l)} \cap \{0<t<1\}} u \leq \inf_{\tilde{K}_2} u \leq 1$$

Hence $(k+1)m < k^* + 1$. Moreover, Lemma 2.4.23 implies

$$R^2 \rho^2 \leq (1 - t_0)(\alpha_{k^*})^{-1} \leq \frac{9}{9^{(k+1)m}}.$$

Hence, we choose $B = 9^m$ and $C_1 = \frac{36(m+1)}{9^m - 1}$.   □

We can now apply the covering lemma and conclude that

$$|A| \leq \delta \frac{m+1}{m} |\mathcal{B}|.$$

We choose m large enough (universal) such that

$$(1 - \mu)\frac{m+1}{m} \leq 1 - \mu/2.$$

Recalling that we chose μ such that $\frac{1}{\mu} = 1 + 2C_{\max} R^{d+2}$ (where $C_{\max}$ is the universal constant appearing in the maximum principle), the previous condition is equivalent to

$$m \geq 4C_{\max} R^{d+2}.$$

Since $R \leq 1$, it is enough to choose $m \geq 4C_{\max}$.

Thanks to the induction assumption, we thus finally get

$$|\{u > M^{(k+1)m}\} \cap (0, R^2/2 + C_1 B^{-k-1}) \times (-R, R)^d| \leq C(1 - \mu/2)^{k+1}.$$

The proof is now complete.   □

2.4.3 Harnack Inequality

The main result of this subsection is the following theorem.

Theorem 2.4.32 (Harnack inequality). *For any non-negative function u such that*

$$\left.\begin{array}{l} \frac{\partial u}{\partial t} + P^+(D^2 u) \geq -f \\ \frac{\partial u}{\partial t} + P^-(D^2 u) \leq f \end{array}\right\} \tag{2.55}$$

in Q_1, we have

$$\sup_{\tilde{K}_4} u \le C(\inf_{Q_{R^2}} u + \|f\|_{L^{d+1}(Q_1)})$$

where $\tilde{K}_4 = (-R^2 + \frac{3}{8}R^4, -R^2 + \frac{1}{2}R^4) \times B_{\frac{R^2}{2\sqrt{2}}}(0)$.

Remark 2.4.33. The case where u solves (2.55) in Q_ρ instead of Q_1 follows by scaling. Indeed, consider $v(t,x) = u(\rho^2 t, \rho x)$ and change constants accordingly.

We will derive Theorem 2.4.32 combining Theorem 2.4.15 with the following proposition (which in turn also follows from Theorem 2.4.15).

Proposition 2.4.34 (Local maximum principle). *Consider a function u such that*

$$\frac{\partial u}{\partial t} + P^-(D^2 u) \le f \ in \ Q_1. \tag{2.56}$$

Then for all $p > 0$, we have

$$\sup_{Q_{1/2}} u \le C\left(\left(\int_{Q_1} (u^+)^p\right)^{\frac{1}{p}} + \|f\|_{L^{d+1}(Q_1)}\right).$$

Proof. First we can assume that $u \ge 0$ by remarking that u^+ satisfies (2.56) with f replaced with $|f|$.

Let Ψ be defined by

$$\Psi(t,x) = h \max((1 - |x|)^{-2\gamma}, (1+t)^{-\gamma})$$

where γ will be chosen later. We choose h minimal such that

$$\Psi \ge u \ in \ Q_1.$$

In other words

$$h = \min_{(t,x) \in Q_1} \frac{u(t,x)}{\max((1 - |x|)^{-2\gamma}, (1+t)^{-\gamma})}.$$

We want to estimate h from above. Indeed, we have

$$\sup_{Q_{\frac{1}{2}}} u \le Ch$$

for some constant C depending on γ and $Q_{\frac{1}{2}}$.

In order to do estimate h, we consider a point (t_0, x_0) realizing the minimum in the definition of h. We consider

$$\delta^2 = \min((1 - |x_0|)^2, (1 + t_0)).$$

In particular

$$u(t_0, x_0) = h\delta^{-2\gamma}$$

and $Q_\delta(t_0, x_0) \subset Q_1$.

We consider next the function $v(t, x) = C - u(t, x)$ where

$$C = \sup_{Q_{\beta\delta}(t_0, x_0)} \Psi$$

for some parameter $\beta \in (0, 1)$ to be chosen later. Remark first that

$$h\delta^{-2\gamma} \leq C \leq h((1 - \beta)\delta)^{-2\gamma}.$$

Remark next that v is a supersolution of

$$\frac{\partial v}{\partial t} + P^+(D^2 v) + |f| \geq 0 \quad \text{in } Q_1$$

and $v \geq 0$ in $(t_0 - (R\beta\delta)^2, t_0) \times B_{\beta\delta}(x_0) \subset Q_{\beta\delta}(t_0, x_0)$. From the L^ε estimate (Theorem 2.4.15 properly scaled and translated), we conclude that

$$\int_L v^\varepsilon \leq C(\beta\delta)^{d+2} \left(\inf_{(t_0 - \beta\delta, x_0) + \beta\delta\tilde{K}_2} v + (\beta\delta)^{\frac{d}{d+1}} \|f\|_{L^{d+1}(Q_1)} \right)^\varepsilon$$

where $L = (t_0 - \beta\delta, x_0) + \beta\delta\tilde{K}_1$. Moreover,

$$\inf_{(t_0 - \beta\delta, x_0) + \beta\delta\tilde{K}_2} v \leq v(t_0, x_0)$$

$$= C - u(t_0, x_0)$$

$$\leq h\left((1 - \beta)^{-2\gamma} - 1\right)\delta^{-2\gamma}.$$

Hence, we have

$$\int_L v^\varepsilon \leq C(\beta\delta)^{d+2} \left[h\left((1 - \beta)^{-2\gamma} - 1\right)\delta^{-2\gamma} + (\beta\delta)^{\frac{d}{d+1}} \|f\|_{d+1} \right]^\varepsilon. \qquad (2.57)$$

We now consider the set

$$A = \left\{ (t, x) \in L : u(t, x) < \frac{1}{2} u(t_0, x_0) = \frac{1}{2} h \delta^{-2\gamma} \right\}.$$

We have

$$\int_A v^\varepsilon \geq |A| \left(h \delta^{-2\gamma} - \frac{1}{2} h \delta^{-2\gamma} \right)^\varepsilon = |A| \left(\frac{h \delta^{-2\gamma}}{2} \right)^\varepsilon.$$

We thus get from (2.57) the following estimate

$$|A| \leq C |L| \left[\left((1 - \beta)^{-2\gamma} - 1 \right)^\varepsilon + (\delta^{2\gamma} h^{-1})^\varepsilon (\beta \delta)^{\frac{d\varepsilon}{d+1}} \|f\|_{d+1}^\varepsilon \right].$$

Finally, we estimate $\int_{Q_1} u^\varepsilon$ from below as follows

$$\int_{Q_1} u^\varepsilon \geq \int_{L \setminus A} u^\varepsilon \geq (|L| - |A|) 2^{-\varepsilon} (h \delta^{-2\gamma})^\varepsilon.$$

Hence, choosing $\gamma = \frac{d+2}{2\varepsilon}$ and combining the two previous inequalities, we get

$$\beta^{2+d} C_1 h^\varepsilon = |L| 2^{-\varepsilon} (h \delta^{-2\gamma})^\varepsilon \leq \int_{Q_1} u^\varepsilon$$

$$+ \beta^{2+d} C_2 h^\varepsilon \left((1 - \beta)^{-2\gamma} - 1 \right)^\varepsilon$$

$$+ \beta^{2+d+\frac{d\varepsilon}{d+1}} C_2 \|f\|_{d+1}^\varepsilon.$$

We used $\delta \leq 1$. Choose now β small enough so that

$$C_2 \left((1 - \beta)^{-2\gamma} - 1 \right)^\varepsilon \leq C_1/2$$
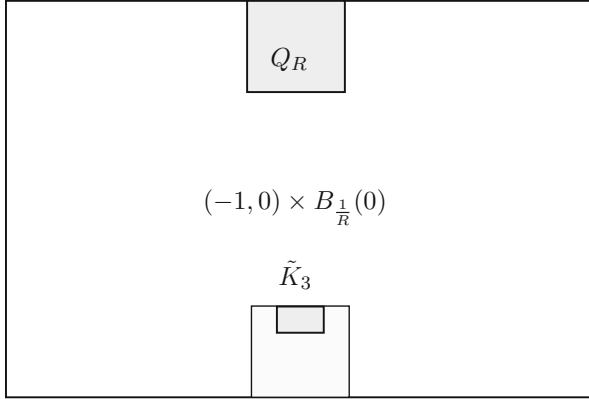
and conclude in the case $p = \varepsilon$. The general case follows by interpolation.   □

Theorem 2.4.32 is a direct consequence of the following one.

Theorem 2.4.35. *For any non-negative function u satisfying (2.55) in $(-1, 0) \times B_{\frac{1}{R}}(0)$, we have*

$$\sup_{\tilde{K}_3} u \leq C (\inf_{Q_R} u + \|f\|_{L^{d+1}((-1,0) \times B_{\frac{1}{R}}(0))})$$

where $\tilde{K}_3 = (-1 + \frac{3}{8} R^2, -1 + R^2/2) \times B_{\frac{R}{2\sqrt{2}}}(0)$ (see Fig. 2.10).

Fig. 2.10 The set $\widetilde{K}_3$

Proof of Theorem 2.4.35. On the one hand, from Theorem 2.4.15 (the L^ε estimate) applied to $u(t+1, x)$ we know that

$$\left(\fint_{(-1, -1+R^2/2) \times B_{R/\sqrt{2}}} u(x)^\varepsilon dx \right)^{1/\varepsilon} \leq C(\inf_{Q_R} u + \|f\|_{L^{d+1}(Q_1)}). \tag{2.58}$$

On the other hand, we apply Proposition 2.4.34 to the scaled function $v(t, x) = u((t + 1 - R^2/2)/(R^2/2), \sqrt{2}x/R) \geq 0$ and $p = \varepsilon$ to obtain

$$\sup_{Q_{\frac{1}{2}}} v \leq C \left(\left(\fint_{Q_1} v^\varepsilon \right)^{\frac{1}{\varepsilon}} + \|f\|_{L^{d+1}(Q_1)} \right).$$

Scaling back to the original variables, we get

$$\sup_{\widetilde{K}_3} u \leq C \left(\left(\fint_{(-1, -1+R^2/2) \times B_{R/\sqrt{2}}} u^\varepsilon \right)^{\frac{1}{\varepsilon}} + \|f\|_{L^{d+1}(Q_1)} \right). \tag{2.59}$$

Combining (2.58) with (2.59) we get

$$\sup_{\widetilde{K}_3} u \leq C \left(\inf_{Q_R} u + \|f\|_{L^{d+1}(Q_1)} \right),$$

which finishes the proof. $\square$

2.4.4 Hölder Continuity

An important consequence of Harnack inequality (Theorem 2.4.32) is the Hölder continuity of functions satisfying (2.55).

Theorem 2.4.36. *If u satisfies* (2.55) *in Q_ρ then u is α-Hölder continuous in Q_ρ and*

$$[u]_{\alpha,Q_{\rho/2}} \le C\rho^{-\alpha} \left(|u|_{0,Q_\rho} + \rho^{\frac{d}{d+1}} \|f\|_{L^{d+1}(Q_\rho)} \right).$$

Proof. We only deal with $\rho = 1$. We prove that if u satisfies (2.55) in Q_1 then u is α-Hölder continuous at the origin, i.e.

$$|u(t,x) - u(0,0)| \le C \left(|u|_{0,Q_1} + \|f\|_{L^{d+1}(Q_1)} \right) (|x| + \sqrt{t})^\alpha. \qquad (2.60)$$

To get such an estimate, it is enough to prove that the oscillation of the function u in Q_ρ decays as ρ^α; more precisely, we consider

$$M_\rho = \sup_{Q_\rho} u,$$

$$m_\rho = \inf_{Q_\rho} u,$$

$$\mathrm{osc}_{Q_\rho} u = M_\rho - m_\rho.$$

Then (2.60) holds true as soon as

$$\mathrm{osc}_{Q_\rho} u \le C \left(|u|_{0,Q_1} + \|f\|_{L^{d+1}(Q_1)} \right) \rho^\alpha. \qquad (2.61)$$

Indeed, consider $(t,x) \in Q_\rho \setminus Q_{\rho/2}$ and estimate $|u(t,x) - u(0,0)|$ from above by $\mathrm{osc}_{Q_\rho} u$ and $\rho/2$ from above by $|x|_\infty + \sqrt{t}$.

In order to prove (2.61), we consider the two functions $u - m_\rho \ge 0$ and $M_\rho - u \ge 0$ in Q_ρ. They both satisfy (2.55) in Q_ρ. From the Harnack inequality, we thus get

$$\sup_{\rho \tilde{K}_4}(u - m_\rho) \le C(\inf_{Q_{R^2\rho}} (u - m_\rho) + \rho^{\frac{d}{d+1}} \|f\|_{d+1})$$

$$\sup_{\rho \tilde{K}_4}(M_\rho - u) \le C(\inf_{Q_{R^2\rho}} (M_\rho - u) + \rho^{\frac{d}{d+1}} \|f\|_{d+1})$$

where $\rho \tilde{K}_4 \subset Q_\rho$ follows from $\tilde{K}_4 \subset (-1,0) \times B_1$. We next add these two inequalities which yields

$$\mathrm{osc}_{Q_\rho} u \le C(\mathrm{osc}_{Q_\rho} u - \mathrm{osc}_{Q_{\gamma\rho}} u + \rho^{\frac{d}{d+1}} \|f\|_{d+1})$$

with $C > 1$ and where γ denotes R^2. Rearranging terms, we get

$$\mathrm{osc}_{Q_{\gamma\rho}} u \le \frac{C-1}{C} \mathrm{osc}_{Q_\rho} u + \rho^{\frac{d}{d+1}} \|f\|_{d+1}$$

where C is universal. Then an elementary iteration lemma allows us to achieve the proof of the theorem; see Lemma 2.5.13 in appendix with $h(\rho) = \mathrm{osc}_{Q_\rho} u$ and $\delta = (C-1)/C$ and $\beta = d/(d+1)$.   □

Appendix: Technical Lemmas

A.1 Lebesgue's Differentiation Theorem

The purpose of this appendix is to prove a version of Lebesgue's differentiation theorem with parabolic cylinders. Recall that the usual version of the result says that if $f \in L^1(\Omega, dt \otimes dx)$ where Ω is a Borel set of $\mathbb{R}^{d+1}$, then for a.e. $(t, x) \in \Omega$,

$$\lim_{j \to \infty} \fint_{G_j} |f - f(t, x)| = 0$$

as long as the sequence of sets G_j satisfies the *regularity* condition:

$$G_j \subset B_j$$
$$|G_j| \ge c|B_j|$$

where B_j is a sequence of balls $B_{r_j}(t, x)$ with $r_j \to 0$.

A sequence of parabolic cylinders $Q_{r_j}(t, x)$ cannot satisfy the regularity condition because of the different scaling between space and time. Indeed $|Q_{r_j}(t, x)| = r_j^{d+2}$ which is an order of magnitude smaller than r_j^{d+1}.

Fortunately, the classical proof of Lebesgue's differentiation theorem can be repeated and works for parabolic cylinders as well, as it is shown below.

Theorem 2.5.1 (Lebesgue's differentiation theorem). *Consider an integrable function $f \in L^1(\Omega, dt \otimes dx)$ where Ω is an open set of $\mathbb{R}^{d+1}$. Then for a.e. $(t, x) \in \Omega$,*

$$\lim_{r \to 0+} \fint_{(t-r^2, t) \times B_r(x)} |f - f(t, x)| = 0$$

where $\fint_O g = \frac{1}{|O|} \int_O g$ for any Borel set $O \subset \mathbb{R}^{d+1}$ and integrable function g.

In the proof, we will in fact use the following corollary.

Corollary 2.5.2 (Generalized Lebesgue's differentiation theorem). *Let G_j be a family of sets which is regular in the following sense: there exists a constant $c > 0$ and $r_j \to 0$ such that*

$$G_j \subset (t - r_j^2, t) \times B_{r_j}(x),$$

$$|G_j| \geq c r_j^{d+2}.$$

Then, except for a set of measure zero which is independent of the choice of $\{G_j\}$, we have

$$\lim_{j \to +\infty} \fint_{G_j} |f - f(t, x)| = 0.$$

Remark 2.5.3. It is interesting to point out that if the parabolic cylinders were replaced by other families of sets not satisfying the *regularity* condition, the result of Lemma 2.5.5 may fail. For example if we take

$$\tilde{M} f(t, x) = \sup_{(a,b) \times B_r(y) \ni (t,x)} \fint_{(a,b) \times B_r(y) \cap \Omega} |f|$$

then Lemma 2.5.5 would fail for $\tilde{M} f$.

Proof of Corollary 2.5.2. We obtain Corollary 2.5.2 as an immediate consequence of Theorem 2.5.1 by noting that since $G_j \subset (t - r_j^2, t) \times B_{r_j}(x)$.

$$\fint_{G_j} |f - f(t, x)| \leq \frac{r^2 |B_r|}{|G_j|} \fint_{(t-r^2, t) \times B_r(x)} |f - f(t, x)|.$$

Thus, the result holds at all points where this right hand side goes to zero, which is a set of full measure by Theorem 2.5.1 and that $\frac{r^2 |B_r|}{|G_j|} \geq c > 0$. □

In order to prove Theorem 2.5.1, we first need a version of Vitali's covering lemma.

Lemma 2.5.4 (Vitali's covering lemma). *Consider a bounded collection of cubes $(Q_\alpha)_\alpha$ of the form $Q_\alpha = (t_\alpha - r_\alpha^2, t_\alpha) \times B_{r_\alpha}(x_\alpha)$ and a set A such that $A \subset \cup_\alpha Q_\alpha$. Then there is a finite number of cubes $Q_1, \ldots, Q_N$ such that $A \subset \cup_{j=1}^N 5Q_j$ where $5Q_j = (t_\alpha - 25 r_\alpha^2, t_\alpha) \times B_{5 r_\alpha}(x_\alpha)$.*

Consider next the maximal function Mf associated with a function $f \in L^1(\Omega, dt \otimes dx)$

$$Mf(t, x) = \sup_{Q \ni (t,x)} \fint_{Q \cap \Omega} |f|$$

where the supremum is taken over cubes Q of the form $(s, y) + (-r^2, 0) \times B_r$.

Lemma 2.5.5 (The maximal inequality). *Consider* $f \in L^1(\Omega, dt \otimes dx)$, f *positive, and* $\lambda > 0$, *we have*

$$|\{Mf > \lambda\}| \leq \frac{C}{\lambda} \|f\|_{L^1}$$

for some constant C depending only on dimension d.

Proof. For all $x \in \{Mf > \lambda\}$, there exists $Q \ni x$ such that

$$\inf_Q f \geq \frac{\lambda}{2} |Q|.$$

Hence, the set $\{Mf > \lambda\}$ can be covered by cubes Q. From Vitali's covering lemma, there exists a finite cover of $\{Mf > \lambda\}$ with some $5Q$'s:

$$\{Mf > \lambda\} \subset \cup_{j=1}^N 5Q_j$$

with Q_j that are disjoint and such that

$$\int_{Q_j \cap \Omega} f \geq \frac{\lambda}{2} |Q_j \cap \Omega|.$$

Hence

$$\int_\Omega f \geq \int_{\cup_j Q_j \cap \Omega} f = \sum_j \int_{Q_j \cap \Omega} f$$

$$\geq \frac{\lambda}{2} |\cup_j Q_j \cap \Omega| = \frac{\lambda}{2} \times \frac{1}{5^{d+2}} |\cup_j 5Q_j \cap \Omega| \geq \frac{\lambda}{C} |\{Mf > \lambda\}|$$

with $C = 2 \times 5^{d+2}$.   □

We can now prove Lebesgue's differentiation theorem (Theorem 2.5.1).

Proof of Theorem 2.5.1. We can assume without loss of generality that the set Ω is bounded. We first remark that the result is true if f is continuous. If f is not continuous, we consider a sequence $(f_n)_n$ of continuous functions such that

$$\|f - f_n\|_{L^1} \leq \frac{C}{2^n}.$$

Moreover, up to a subsequence, we can also assume that for a.e. $(t, x) \in \Omega$,

$$f_n(t, x) \to f(t, x) \quad \text{as } n \to \infty.$$

Thanks to the maximal inequality (Lemma 2.5.5), we have in particular

$$|\{M(f - f_n) > \lambda\}| \leq \frac{C}{\lambda 2^n}.$$

By Borel–Cantelli's Lemma, we conclude that for all $\lambda > 0$, there exists $n_\lambda \in \mathbb{N}$ such that for all $n \geq n_\lambda$,

$$M(f - f_n) \leq \lambda \quad \text{a.e. in } \Omega.$$

We conclude that for a.e. $(t, x) \in \Omega$ and all $k \in \mathbb{N}$, there exists a strictly increasing sequence n_k such that for all $r > 0$ such that $Q_r(t, x) \subset \Omega$,

$$\fint_{Q_r(t,x)} |f - f_{n_k}| \leq M(f - f_{n_k}) \leq \frac{1}{k}.$$

Moreover, since f_n is continuous and Ω is bounded, there exists $r_k > 0$ such that for $r \in (0, r_k)$, we have

$$\fint_{Q_r(t,x)} |f_{n_k} - f_{n_k}(t, x)| \leq \frac{1}{k}.$$

Moreover, for a.e. $(t, x) \in \Omega$,

$$|f_{n_k}(t, x) - f(t, x)| \to 0 \quad \text{as } k \to \infty.$$

These three facts imply that for a.e. $(t, x) \in \Omega$, for all $\varepsilon > 0$, there exists $r_\varepsilon > 0$ such that $r \in (0, r_\varepsilon)$,

$$\fint_{Q_r(t,x)} |f - f(t, x)| \leq \varepsilon.$$

This achieves the proof of the lemma.   □

A.2   Jensen–Ishii's Lemma for N Functions

When proving Theorem 2.4.9 (more precisely, Lemma 2.4.6), we used the following generalization of Lemmas 2.3.23 and 2.3.30 whose proof can be found in [CIL92].

Lemma 2.5.6 (Jensen–Ishii's Lemma III). *Let U_i, $i = 1, \ldots, N$ be open sets of $\mathbb{R}^d$ and I an open interval of $\mathbb{R}$. Consider also lower semi-continuous functions $u_i : I \times U_i \to \mathbb{R}$ such that for all $v = u_i$, $i = 1, \ldots, N$, $(t, x) \in I \times U_i$, there exists $r > 0$ such that for all $M > 0$ there exists $C > 0$,*

$$\left.\begin{array}{r}(s,y) \in Q_r(t,x) \\ (\beta, q, Y) \in \mathcal{P}^- v(s,y) \\ |v(s,y)| + |q| + |Y| \leq M\end{array}\right\} \Rightarrow -\beta \leq C.$$

Let $x = (x_1, \ldots, x_N)$ and $x_0 = (x_1^0, \ldots, x_N^0)$. Assume that $\sum_{i=1}^N u_i(t, x_i) - \phi(t, x)$ reaches a local minimum at $(t_0, x_0) \in I \times \Pi_i U_i$. If α denotes $\partial_t \phi(t_0, x_0)$ and p_i denotes $D_{x_i} \phi(x_0)$ and A denotes $D^2 \phi(t_0, x_0)$, then for any $\beta > 0$ such that $I + \beta A > 0$, there exist $(\alpha_i, X_i) \in \mathbb{R} \times \mathbb{S}_d$, $i = 1, \ldots, N$, such that for all $i = 1, \ldots, N$,

$$(\alpha_i, p_i, X_i) \in \overline{\mathcal{P}}^- u(t_0, x_i^0)$$

$$\sum_{i=1}^N \alpha_i = \alpha$$

and

$$\frac{1}{\beta}\begin{pmatrix} I & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & I \end{pmatrix} \geq \begin{pmatrix} X_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & X_N \end{pmatrix} \geq A_\beta$$

where $A_\beta = (I + \beta A)^{-1} A$.

Remark 2.5.7. The condition on the functions u_i is satisfied as soon as the u_i's are supersolutions of a parabolic equation. This condition ensures that some compactness holds true when using the doubling variable technique in the time variable. See [CIL92, Theorem 8.2, p. 50] for more details.

A.3 Technical Lemmas for Monotone Envelopes

When proving the maximum principle (Theorem 2.4.9), we used the two following technical lemmas.

Lemma 2.5.8. *Consider a convex set Ω of $\mathbb{R}^d$ and a lower semi-continuous function $v : [a, b] \times \bar{\Omega} \to \mathbb{R}$ which is non-increasing with respect to $t \in (a, b)$ and convex with respect to $x \in \Omega$. Assume that v is bounded from above and that for all $(\alpha, p, X) \in \mathcal{P}^- v(t, x)$, we have*

$$-\alpha \leq C \quad and \quad X \leq CI.$$

Then v is Lipschitz continuous with respect $t \in (a, b)$ and $C^{1,1}$ with respect to $x \in \Omega$.

Proof of Lemma 2.5.8. We assume without loss of generality that Ω is bounded. In this case, v is bounded from above and from below, hence is bounded. Next, we also get that v is Lipschitz continuous with respect to x in $[a, b] \times F$ for all closed convex set $F \subset \Omega$ such that $d(F, \partial\Omega) > 0$.

Step 1.

We first prove that v is Lipschitz continuous with respect to t: for all $(t_0, x_0) \in (a, b) \times \Omega$,

$$M = \sup_{s, t \in (a,b), x, y \in \Omega} \left\{ v(t, x) - v(s, y) - L|t - s| - \frac{L}{4\varepsilon}|x - y|^2 - L\varepsilon \right.$$

$$\left. - L_0|x - x_0|^2 - L_0(t - t_0)^2 \right\} \leq 0$$

for L large enough only depending on C and the Lipschitz constant of v with respect to x around (t_0, x_0) and for L_0 large enough. We argue by contradiction by assuming that $M > 0$. Consider $(\bar{s}, \bar{t}, \bar{x}, \bar{y})$ where the maximum M is reached. Remark first that

$$L_0|\bar{y} - x_0|^2 + L_0(\bar{s} - t_0)^2 + L|\bar{t} - \bar{s}| + \frac{L}{4\varepsilon}|\bar{x} - \bar{y}|^2 + L\varepsilon \leq v(\bar{t}, \bar{x}) - v(\bar{s}, \bar{y})$$

$$\leq 2|v|_{0,[a,b] \times \bar{\Omega}}.$$

In particular, we can choose L_0 and L large enough so that $(\bar{s}, \bar{y}), (\bar{t}, \bar{x}) \in (a, b) \times \Omega$. Remark next that $\bar{t} \neq \bar{s}$. Indeed, if $\bar{t} = \bar{s}$, then

$$0 < M \leq v(\bar{t}, \bar{x}) - v(\bar{t}, \bar{y}) - \frac{L}{4\varepsilon}|\bar{x} - \bar{y}|^2 - L\varepsilon$$

and choosing L larger than the Lipschitz constant of v with respect to x yields a contradiction. Hence the function v is touched from below at $(\bar{s}, \bar{y})$ by the test function

$$(s, y) \mapsto C_0 - \frac{L}{4\varepsilon}|\bar{x} - y|^2 - L|\bar{t} - s|$$

where C_0 is a constant depending on $(\bar{t}, \bar{x})$. In particular,

$$(L \operatorname{sign}(\bar{t} - \bar{s}), L(4\varepsilon)^{-1}(\bar{x} - \bar{y}), L(4\varepsilon)^{-1}I) \in \mathcal{P}^- v(\bar{s}, \bar{y}).$$

We thus should have $L \leq C$. Choosing $L > C$ yields also the desired contradiction.

Step 2.

In order to prove that for all $t \in (a, b)$, $u(t, \cdot)$ is $C^{1,1}$ with respect to x, it is enough to prove that for all $(p, X) \in D^{2,-}u(t, x)$ (see below), $X \leq CI$. Indeed, this implies that $u(t, \cdot) + \frac{C}{2}|\cdot|^2$ is concave [ALL97]. Since $u(t, \cdot)$ is convex, this implies that it is $C^{1,1}$ [CanSin04].

$(p, X) \in D^{2,-}u(t, x)$ means that there exists $\psi \in C^2(\mathbb{R}^d)$ such that $p = D\psi(x)$ and $X = D^2\psi(x)$ and

$$\psi(y) - \psi(x) \leq u(t, y) - u(t, x)$$

for $y \in B_r(x)$. We can further assume that the minimum of $u(t, \cdot) - \psi$ is strict. We then consider the minimum of $u(s, x) - \psi(x) + \varepsilon^{-1}(s-t)^2$ in $(t-r, t+r) \times B_r(x)$. For ε small enough, this minimum is reached in an interior point $(t_\varepsilon, x_\varepsilon)$ and $(t_\varepsilon, x_\varepsilon) \to (t, x)$ as $\varepsilon \to 0$. Then

$$(\varepsilon^{-1}(s_\varepsilon - t), D\psi(x_\varepsilon), D^2\psi(x_\varepsilon)) \in \mathcal{P}^- u(t_\varepsilon, x_\varepsilon).$$

Hence, $D^2\psi(x_\varepsilon) \leq CI$. Letting $\varepsilon \to 0$ yields $X \leq CI$. This achieves Step 2.

The proof of the lemma is now complete. $\qquad\qquad\qquad\qquad\qquad\square$

Lemma 2.5.9. *Consider a convex set Ω of $\mathbb{R}^d$ and $v : (a, b) \times \Omega \to \mathbb{R}$ which is non-increasing with respect to $t \in (a, b)$ and convex with respect to $x \in \Omega$. Then for all $(\alpha, p, X) \in \mathcal{P}^- v(t, x)$, that there exists (α_n, p_n, X_n) such that*

$$(\alpha_n, p_n, X_n) \in \mathcal{P}^- v(t_n, x_n)$$
$$(t_n, x_n, \alpha_n, p_n) \to (t, x, \alpha, p)$$
$$X \leq X_n + o_n(1), X_n \geq 0.$$

The proof of this lemma relies on Alexandroff theorem in its classical form. A statement and a proof of this classical theorem can be found for instance in [EG92]. We will only use the following consequence of this theorem.

Theorem 2.5.10. *Consider a convex set Ω of $\mathbb{R}^d$ and a function $v : (a, b) \times \Omega \to \mathbb{R}$ which is convex with respect to $(t, x) \in (a, b) \times \Omega$. Then for almost $(t, x) \in (a, b) \times \Omega$, there exists $(\alpha, p, X) \in \mathcal{P}^- \cap \mathcal{P}^+ v(t, x)$, that is to say such that,*

$$v(s, y) = v(t, x) + \alpha(s-t) + p \cdot (y-x) + \frac{1}{2}X(y-x) \cdot (y-x) + o(|s-t| + |y-x|^2). \tag{2.62}$$

Jensen's lemma is also needed (stated here in a "parabolic" version for the sake of clarity).

Lemma 2.5.11 (Jensen). *Consider a convex set Ω of $\mathbb{R}^d$ and a function $v : (a, b) \times \Omega \to \mathbb{R}$ such that there exists $(\tau, C) \in \mathbb{R}^2$ such that $u(t, x) + \tau t^2 + C|x|^2$ is convex*

with respect to $(t, x) \in (a, b) \times \Omega$. *If u reaches a strict local maximum at* (t_0, x_0), *then for* $r > 0$ *and* $\delta > 0$ *small enough, the set*

$$K = \{(t, x) \in (t_0 - r, t_0 + r) \times B_r(x_0) : \exists (\tau, p) \in (-\delta, \delta) \times B_\delta,$$

$$(s, y) \mapsto u(s, y) - \tau s - p \cdot y \text{ reaches a local maximum at } (t, x)\}$$

has a positive measure.

See [CIL92] for a proof. We can now turn to the proof of Lemma 2.5.8. The proof of Lemma 2.5.9 below mimics the proof of [ALL97, Lemma 3] in which there is no time dependence.

Proof of Lemma 2.5.9. Consider a test function ϕ such that $u - \phi$ reaches a local maximum at (t, x) and

$$(\alpha, p, X) = (\partial_t \phi, D\phi, D^2\phi)(t, x).$$

Without loss of generality, we can assume that this maximum is strict; indeed, replace ϕ with $\phi(s, y) - |y - x|^2 - (s - t)^2$ for instance. Then consider the function

$$v_\varepsilon(t, x) = \inf_{y \in \mathbb{R}^d, s \geq 0} \left\{ v(s, y) + \frac{1}{\varepsilon}|y - x|^2 + \frac{1}{\varepsilon}(s - t)^2 \right\}.$$

One can check that v_ε is still convex with respect to x and non-increasing with respect to t and that

$$(t, x) \mapsto v_\varepsilon(t, x) + \frac{1}{\varepsilon}|x|^2 + \frac{1}{\varepsilon}t^2$$

is concave with respect to (t, x). Moreover, $v_\varepsilon \leq v$ and

$$\lim_{\varepsilon \to 0} v_\varepsilon(t, x) = v(t, x).$$

This implies that there exists $(t_\varepsilon, x_\varepsilon) \to 0$ as $\varepsilon \to 0$ such that $v_\varepsilon - \phi$ reaches a local maximum at $(t_\varepsilon, x_\varepsilon)$. Remarking that $v_\varepsilon - \phi$ satisfies the assumptions of Jensen's lemma, Lemma 2.5.11 above, we combine it with Theorem 2.5.10 and we conclude that we can find slopes $(\tau_n, p_n) \to (0, 0)$ and points $(t_n, x_n) \to (t_\varepsilon, x_\varepsilon)$ as $n \to \infty$ where $v_\varepsilon - \phi$ satisfies (2.62) and $v_\varepsilon - \phi - \tau_n s - p_n y$ reaches a local maximum at (t_n, x_n). In other words,

$$(\tau_n + \partial_t \phi(t_n, x_n), p_n + D\phi(t_n, x_n), D^2 v_\varepsilon(t_n, x_n)) \in \mathcal{P}^- v_\varepsilon(t_n, x_n)$$

with

$$D^2 v_\varepsilon(t_n, x_n) \geq 0$$

and

$$D^2\phi(t_n, x_n) \le D^2 v_\varepsilon(t_n, x_n).$$

In order to conclude, we use the classical following result from viscosity solution theory (see [CIL92] for a proof):

Lemma 2.5.12. *Consider* (s_n, y_n) *such that*

$$v_\varepsilon(t_n, x_n) = v(s_n, y_n) + \varepsilon^{-1}|y_n - x_n|^2 + \varepsilon^{-1}(t_n - s_n)^2.$$

Then

$$|y_n - x_n|^2 + (t_n - s_n)^2 \le \varepsilon|v^+|_{0,(a,b)\times\Omega}$$

and

$$\mathcal{P}^- u_\varepsilon(t_n, x_n) \subset \mathcal{P}^- u(s_n, y_n).$$

We used in the previous lemma that v is bounded from above since Ω is bounded. Putting all the previous pieces of information together yields the desired result. □

A.4   An Elementary Iteration Lemma

The following lemma is classical, see for instance [GT01, Lemma 8.23].

Lemma 2.5.13. *Consider a non-decreasing function* $h : (0, 1) \to \mathbb{R}^+$ *such that for all* $\rho \in (0, 1)$,

$$h(\gamma\rho) \le \delta h(\rho) + C_0\rho^\beta$$

for some $\delta, \gamma, \beta \in (0, 1)$. *Then for all* $\rho \in (0, 1)$,

$$h(\rho) \le C_\alpha \rho^\alpha$$

for all $\alpha = \frac{1}{2}\min(\frac{\ln\delta}{\ln\gamma}, \beta) \in (0, 1)$.

Proof. Consider $k \in \mathbb{N}, k \ge 1$, and get by induction that for all $\rho_0, \rho_1 \in (0, 1)$ with $\rho_1 \le \rho_0$,

$$h(\gamma^k \rho_1) \le \delta^k h(\rho_1) + C_0\rho_1^\beta \sum_{j=0}^{k-1} \gamma^{\beta j}.$$

Then write

$$h(\gamma^k \rho_1) \leq \delta^k h(\rho_0) + C_0 \frac{\rho_1^\beta}{1 - \gamma^\beta}$$

$$\leq (\gamma^k)^{\tilde{\beta}} h(\rho_0) + C_0 \frac{\rho_1^\beta}{1 - \gamma^\beta}$$

$$\leq (\gamma^k)^{2\alpha} h(\rho_0) + C_0 \frac{\rho_1^{2\alpha}}{1 - \gamma^\beta}$$

where $\tilde{\beta} = \frac{\ln \delta}{\ln \gamma}$. Now pick $\rho \in [\gamma^{k+1}\rho_1, \gamma^k \rho_1)$ and choose $\rho_1 = \sqrt{\rho_0 \rho}$ and get from the previous inequality the desired result for $\rho \in (0, \rho_0)$. Choose next $\rho_0 = \frac{1}{2}$ and conclude for $\rho \in (0, 1)$. □

References

[ALL97] O. Alvarez, J.-M. Lasry, P.-L. Lions, Convex viscosity solutions and state constraints. J. Math. Pures Appl. (9) **76**(3), 265–288 (1997)

[Barl94] G. Barles, in *Solutions de viscosité des équations de Hamilton-Jacobi*. Mathématiques & Applications (Berlin), vol. 17 (Springer, Paris, 1994), x+194 pp.

[CafCab] L.A. Caffarelli, X. Cabre, in *Fully Nonlinear Elliptic Equations*. American Mathematical Society Colloquium Publications, vol. 43 (American Mathematical Society, Providence, 1995), vi+104 pp.

[CanSin04] P. Cannarsa, C. Sinestrari, Semiconcave functions, in *Hamilton-Jacobi Equations, and Optimal Control*. Progress in Nonlinear Differential Equations and Their Applications, vol. 58 (Birkhäuser, Boston, 2004), xiv+304 pp.

[CL81] M.G. Crandall, P.-L. Lions, Condition d'unicité pour les solutions généralisées des équations de Hamilton-Jacobi du premier ordre. C. R. Acad. Sci. Paris Sér. I Math. **292**(3), 183–186 (1981)

[CIL92] M.G. Crandall, H. Ishii, P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. (N.S.) **27**(1), 1–67 (1992)

[Dong91] G.C. Dong, in *Nonlinear Partial Differential Equations of Second Order*. Translated from the Chinese by Kai Seng Chou [Kaising Tso]. Translations of Mathematical Monographs, vol. 95 (American Mathematical Society, Providence, 1991), viii+251 pp.

[EG92] L.C. Evans, R.F. Gariepy, in *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics (CRC Press, Boca Raton, 1992), viii+268 pp.

[GT01] D. Gilbarg, N.S. Trudinger, in *Elliptic Partial Differential Equations of Second Order*, Reprint of the 1998 edn. Classics in Mathematics (Springer, Berlin, 2001), xiv+517 pp.

[HUL] J.-B. Hirriart-Urruty, C. Lemaréchal, Fundamentals of convex analysis. Abridged version of Convex analysis and minimization algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and II [Springer, Berlin, 1993; MR1295240 (95m:90002)]. Grundlehren Text Editions (Springer, Berlin, 2001), x+259 pp.

[Imb06] C. Imbert, Convexity of solutions and $C^{1,1}$ estimates for fully nonlinear elliptic equations. J. Math. Pure Appl. (9) **85**(6), 791–807 (2006)

[Ish87] H. Ishii, Perron's method for Hamilton-Jacobi equations. Duke Math. J. **55**(2), 369–384 (1987)

[Ish89] H.Ishii, On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs. Comm. Pure Appl. Math. **42**(1), 15–45 (1989)

[IL90] H. Ishii, P.-L. Lions, Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. J. Differ. Equat. **83**(1), 26–78 (1990)

[Jens88] R. Jensen, The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. Arch. Ration. Mech. Anal. **101**(1), 1–27 (1988)

[Kryl76] N.V. Krylov, Sequences of convex functions, and estimates of the maximum of the solution of a parabolic equation. Sibirsk. Mat. Z. **17**(2), 290–303, 478 (1976)

[Kryl87] N.V. Krylov, in *Nonlinear Elliptic and Parabolic Equations of the Second Order*. Translated from the Russian by P.L. Buzytsky. Mathematics and Its Applications (Soviet Series), vol. 7 (D. Reidel Publishing Co., Dordrecht, 1987), xiv+462 pp.

[Kryl96] N.V. Krylov, in *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. Graduate Studies in Mathematics, vol. 12 (American Mathematical Society, Providence, 1996), xii+164 pp.

[Kryl97] N.V. Krylov, Fully nonlinear second order elliptic equations: recent development. Dedicated to Ennio De Giorgi. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **25**(3–4), 569–595 (1997)

[LSU67] O.A. Ladyzenskaja, V.A. Solonnikov, N.N. Uralceva, in *Linear and Quasilinear Equations of Parabolic Type* (Russian). Translated from the Russian by S. Smith. Translations of Mathematical Monographs, vol. 23 (American Mathematical Society, Providence, 1967), xi+648 pp.

[Lieb96] G.M. Lieberman, *Second Order Parabolic Differential Equations* (World Scientific, River Edge, 1996)

[Lions83] P.-L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, II. Viscosity solutions and uniqueness. Comm. Partial Differ. Equat. **8**(11), 1229–1276 (1983)

[Saf84] M.V. Safonov, The classical solution of the elliptic Bellman equation (Russian). Dokl. Akad. Nauk SSSR **278**(4), 810–813 (1984)

[Tso85] K. Tso, On an Aleksandrov-Bakelman type maximum principle for second-order parabolic equations. Comm. Partial Differ. Equat. **10**(5), 543–553 (1985)

[Wang92a] L. Wang, On the regularity theory of fully nonlinear parabolic equations, I. Comm. Pure Appl. Math. **45**(1), 27–76 (1992)