Regularity Theory for Fully Nonlinear Integro-Differential Equations

LUIS CAFFARELLI
University of Texas at Austin

LUIS SILVESTRE
University of Chicago

Abstract

We consider nonlinear integro-differential equations like the ones that arise from stochastic control problems with purely jump Lévy processes. We obtain a nonlocal version of the ABP estimate, Harnack inequality, and interior $C^{1,\alpha}$ regularity for general fully nonlinear integro-differential equations. Our estimates remain uniform as the degree of the equation approaches 2, so they can be seen as a natural extension of the regularity theory for elliptic partial differential equations. © 2008 Wiley Periodicals, Inc.

1 Introduction

Integro-differential equations appear naturally when studying discontinuous stochastic processes. The generator of an n-dimensional Lévy process is given by an operator with the general form

$$\begin{aligned} Lu(x) = &\sum_{ij} a_{ij}\partial_{ij}u + \sum_i b_i \partial_i u \\ &+ \int_{\mathbb{R}^n} \big(u(x+y) - u(x) - \nabla u(x)\cdot y\chi_{B_1}(y)\big)\mathrm{d}\mu(y). \end{aligned} \tag{1.1}$$

The first term corresponds to the diffusion, the second to the drift, and the third to the jump part. In this paper we focus on the equations obtained when we consider purely jump processes—processes without diffusion or drift part. The operators have the general form

$$Lu(x) = \int_{\mathbb{R}^n} \big(u(x+y) - u(x) - \nabla u(x)\cdot y\chi_{B_1}(y)\big)\mathrm{d}\mu(y), \tag{1.2}$$

where μ is a measure such that $\int_{\mathbb{R}^n} |y|^2/(1+|y|^2)\,\mathrm{d}\mu(y) < +\infty$.

The value of $Lu(x)$ is well-defined as long as u is bounded in $\mathbb{R}^n$ and $C^{1,1}$ at x. These concepts will be made more precise later.

Communications on Pure and Applied Mathematics, Vol. LXII, 0597–0638 (2009)

The operator L described above is a linear integro-differential operator. In this paper we want to obtain results for nonlinear equations. We obtain this kind of equations in stochastic control problems [11]. If in a stochastic game a player is allowed to choose from different strategies at every step in order to maximize the expected value of some function at the first exit point of a domain, a convex nonlinear equation emerges,

$$Iu(x) = \sup_{\alpha} L_{\alpha} u(x). \tag{1.3}$$

In a competitive game with two or more players, more complicated equations appear. We can obtain equations of the type

$$Iu(x) = \inf_{\beta} \sup_{\alpha} L_{\alpha\beta} u(x). \tag{1.4}$$

The difference between (1.3) and (1.4) is convexity. Alternatively, an operator like $Iu(x) = \sup_\alpha \inf_\beta L_{\alpha\beta} u(x)$ can be considered. A characteristic property of these operators is that

$$\inf_{\alpha\beta} L_{\alpha\beta} v(x) \le I(u+v)(x) - Iu(x) \le \sup_{\alpha\beta} L_{\alpha\beta} v(x). \tag{1.5}$$

A more general and better description of the nonlinear operators we want to deal with is the operators I for which (1.5) holds for some family of linear integro-differential operators $L_{\alpha\beta}$. The idea is that an estimate on $I(u+v) - Iu$ by a suitable extremal operator can be a replacement for the concept of ellipticity. Indeed, if we consider the extremal Pucci operators (see [7]) $M^+_{\lambda,\Lambda}$ and $M^-_{\lambda,\Lambda}$, and we have $M^-_{\lambda,\Lambda} v(x) \le I(u+v) - Iu \le M^+_{\lambda,\Lambda} v(x)$, then it is easy to see that I must be an elliptic second-order differential operator. If instead we compare with suitable nonlocal extremal operators, we will have a concept of ellipticity for nonlocal equations. We will give a precise definition in Section 3 (Definition 3.1).

We now explain the natural Dirichlet problem for a nonlocal operator. Let Ω be an open domain in $\mathbb{R}^n$. We are given a function g defined in $\mathbb{R}^n \setminus \Omega$, which is the boundary condition. We look for a function u such that

$$\begin{aligned} Iu(x) &= 0 && \text{for every } x \in \Omega, \\ u(x) &= g(x) && \text{for } x \in \mathbb{R}^n \setminus \Omega. \end{aligned}$$

Notice that the boundary condition is given in the whole complement of Ω and not only $\partial\Omega$. This is because of the nonlocal character of the operator I. From the stochastic point of view, it corresponds to the fact that a discontinuous Lévy process can exit the domain Ω for the first time jumping to any point in $\mathbb{R}^n \setminus \Omega$.

In this paper we will focus mainly on the regularity properties of solutions to an equation $Iu = 0$. In order to obtain regularity results, we must assume some *nice* behavior of the measures μ. Basically, our assumption is that they are symmetric, absolutely continuous, and not too degenerate. To fix ideas, we can think of integro-differential operators with a kernel comparable to the respective kernel of the fractional Laplacian $-(-\triangle)^{\sigma/2}u(x) = \int (u(x+y) - u(x))|y|^{-n-\sigma}\, \mathrm{d}y$.

This is the first of a series of papers where we plan to extend the existing theory for (fully nonlinear) second-order elliptic equations to the case of discontinuous processes in a seamless fashion, i.e., with methods and estimates that reach uniformly the second-order case. Our results in this paper are as follows:

- a comparison principle for a general nonlinear integro-differential equation,
- a nonlocal version of the Aleksandrov-Bakel′man-Pucci estimate,
- the Harnack inequality for integro-differential equations with kernels that are comparable with the ones of the fractional Laplacian but can be very discontinuous,
- a Hölder regularity result for the same class of equations as the Harnack inequality, and
- a $C^{1,\alpha}$ regularity result for a large class of not necessarily convex, nonlinear integro-differential equations.

Although there are some known results about Harnack inequalities and Hölder estimates for integro-differential equations with either analytical proofs [10] or probabilistic proofs [3, 4, 5, 12], the estimates in all these previous results blow up as the order of the equation approaches 2. In this way, they do not generalize to elliptic differential equations. We provide estimates that remain uniform in the degree and therefore make the theory of integro-differential equations and elliptic differential equations appear somewhat unified. Consequently, our proofs are more involved than the ones in the bibliography.

In this paper we consider only nonlinear operators that are translation invariant (independent of x or *constant coefficients*). The *variable coefficient* case will be considered in future work. In future papers, we are also planning to address the problem of the interior regularity of the integro-differential Hamilton-Jacobi-Bellman equation. This refers to the equation involving a convex nonlocal operator like (1.3). In that case we obtain an analogue of the Evans-Krylov theorem proving that the solutions to the equation have enough regularity to be classical solutions.

In Sections 2 to 6 we develop the technical results that we need for the regularity results presented in the second half of the paper. In this first part, definitions, approximation, stability, and comparison results are discussed. The most important results are in the second part of the paper, starting in Section 7, where the regularity theory of solutions is developed.

In Section 2 we discuss the appropriate definition of viscosity solution given the global nature of the problem. In Section 3 we introduce the notion of extremal operators associated to a family of kernels. These operators play the same role as the Pucci extremal operators in the second-order theory and allow us to define, in a parallel fashion to [7], *solutions to equations with bounded measurable coefficients* as those functions that are super- and subsolutions of the lower and upper extremal operators. In Section 4 we show how the classes of equations introduced above persist under appropriate limiting processes of (super- or sub-) solutions. This provides a powerful compactness tool for a priori estimates. In Section 5 we discuss approximation of solutions by sup (or inf) convolution and the resulting comparison theorems. In particular, the fundamental conclusion is that the difference of two solutions to the fully nonlinear, translation-invariant equation satisfies an *equation with bounded measurable coefficients* in the sense discussed above. In Section 6 we show how to obtain an elliptic partial differential equation as a limit of integro-differential equations.

In Section 7, for the reader's convenience, we provide a quick overview of the regularity results we will prove in the following sections. The most interesting part of the paper starts in Section 8 with the nonlocal ABP estimate. In Sections 9 and 10 we construct a special function and prove some pointwise estimates that will help in proving the Harnack inequality and Hölder estimates in Sections 11 and 12. In Section 13 we show the $C^{1,\alpha}$ estimates. Finally, in Section 14 we show how to generalize our previous results when our operators have truncated kernels. This last section is important for applications since very often the kernels of an integro-differential equation are comparable to the ones of the fractional Laplacian only in a neighborhood of the origin.

2 Definitions of Viscosity Solutions

As we mentioned in the introduction, equation (1.2) was given in too much generality for our purposes. We will restrict our attention to the operators where μ is given by a symmetric kernel K. It takes the form

$$Lu(x) = \mathrm{PV} \int_{\mathbb{R}^n} (u(x+y) - u(x)) K(y) \mathrm{d}y. \tag{2.1}$$

The kernel K must be a positive function, satisfy $K(y) = K(-y)$, and also

$$\int_{\mathbb{R}^n} \frac{|y|^2}{|y|^2+1} K(y) \mathrm{d}y < +\infty. \tag{2.2}$$

The assumption (2.2) is the standard Lévy-Khintchine condition. It is not necessary to subtract the term $-\nabla u(x) \cdot y \chi_{B_1}$ if we think of the integral in the principal value sense. Alternatively, due to the symmetry of the kernel K, the operator can also be written as

$$Lu(x) = \frac{1}{2} \int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x)) K(y) \mathrm{d}y.$$

In order to simplify the notation, we will write

$$\delta(u,x,y) := u(x+y) + u(x-y) - 2u(x).$$

The expression for L can be written concisely as

$$Lu(x) = \int_{\mathbb{R}^n} \delta(u, x, y) K(y) \mathrm{d}y \tag{2.3}$$

for some kernel K (which would be half of the one of (2.1)). We will alternate between writing the operators in the form (2.1) and (2.3) whenever it is convenient.

The nonlinear integro-differential operators that arise in stochastic control have the form (1.4) where we think that for each $L_{\alpha\beta}$ we have a kernel $K_{\alpha\beta}$ so that $L_{\alpha\beta}$ has the form (2.3). We will define a more general form for nonlinear integro-differential operators in Section 3.

The minimum assumption in order to have Iu well-defined is that every kernel $K_{\alpha\beta}$ must satisfy (2.2) in a uniform way. More precisely,

$$\text{if } K(y) := \sup_{\alpha\beta} K_{\alpha\beta}(y), \quad \text{then } \int_{\mathbb{R}^n} \frac{|y|^2}{|y|^2 + 1} K(y) \mathrm{d}y < +\infty. \tag{2.4}$$

The value of Iu can be evaluated in the classical sense if $u \in C^{1,1}$. If we want to evaluate the value of $Iu(x)$ at only one point x, we need u to be punctually $C^{1,1}$ in the sense of the following definition.

DEFINITION 2.1 A function φ is said to be $C^{1,1}$ *at the point* x, and we write $\varphi \in C^{1,1}(x)$, if there is a vector $v \in \mathbb{R}^n$ and a number $M > 0$ such that

$$|\varphi(x + y) - \varphi(x) - v \cdot y| \le M|y|^2 \quad \text{for } |y| \text{ small enough.}$$

We say that a function is $C^{1,1}$ in a set Ω if the previous definition holds at every point with a uniform constant M.

As in the second-order case, we give a definition of viscosity sub- and supersolutions for integro-differential equations by testing the operators in $C^{1,1}$ functions that *touch* the function u from either above or below. Often for nonlocal equations the definition is given by test functions that remain on one side of u in the whole space $\mathbb{R}^n$. We take a slightly more general approach. We consider a test function φ that touches u at a point x and remains on one side of u but is only defined locally in a neighborhood N of x. Then we complete φ with the tail of u to evaluate the integrals (2.3). We do this in order to allow arbitrary discontinuities in the function u outside of the domain Ω where it may be a solution of the equation.

DEFINITION 2.2 A function $u : \mathbb{R}^n \to \mathbb{R}$, upper (lower) semicontinuous in $\bar{\Omega}$, is said to be a *subsolution* (*supersolution*) to $Iu = f$, and we write $Iu \ge f$ ($Iu \le f$), if every time all the following happen:

- x is any point in Ω.
- N is a neighborhood of x in Ω.
- φ is some C^2 function in $\bar{N}$.
- $\varphi(x) = u(x)$.
- $\varphi(y) > u(y)$ ($\varphi(y) < u(y)$) for every $y \in N \setminus \{x\}$.

Then if we let

$$v := \begin{cases} \varphi & \text{in } N \\ u & \text{in } \mathbb{R}^n \setminus N, \end{cases}$$

we have $Iv(x) \geq f(x)$ ($Iv(x) \leq f(x)$).

A *solution* is a function u that is both a subsolution and a supersolution.

Note that Definition 2.2 is essentially the same as definition 2 in [2]. In this paper we will only consider continuous right-hand sides f.

For the set of test functions, we could also use a function φ that is $C^{1,1}$ only at the contact point x. This is a larger set of test functions, so a priori it may provide a stronger concept of solution. In Section 4 we will show that the two approaches are actually equivalent.

Usually the nonlocal operators I allow some growth at infinity for the Dirichlet data. For instance, if the value of $Iu(x)$ is well-defined every time $u \in C^{1,1}(x)$ and $u \in L^1(\mathbb{R}^n, w)$ for any weight w that bounds at infinity the tails of the kernels K_α, then the above definition would apply for semicontinuous functions in $\overline{\Omega}$ that are in $L^1(\mathbb{R}^n, w)$ but not necessarily bounded. In most cases, our regularity results in this paper can be extended to the unbounded case by truncating the function and adding an error term in the right-hand side.

3 Extremal Operators of Pucci Type

In (1.3) and (1.4) we consider the supremum or an inf sup of a collection of linear operators. Let us consider a collection of linear operators $\mathcal{L}$ that includes all of them. The maximal and a minimal operator with respect to $\mathcal{L}$ are defined as:

$$\mathrm{M}^+_{\mathcal{L}} v(x) = \sup_{L \in \mathcal{L}} Lu(x), \tag{3.1}$$

$$\mathrm{M}^-_{\mathcal{L}} v(x) = \inf_{L \in \mathcal{L}} Lu(x). \tag{3.2}$$

For example, an important class that we will use for regularity results is given by the class $\mathcal{L}_0$ of operators L of the form (2.3) with

$$(2-\sigma)\frac{\lambda}{|y|^{n+\sigma}} \leq K(y) \leq (2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}} \quad \text{where } 0 < \sigma < 2; \tag{3.3}$$

then $\mathrm{M}^+_{\mathcal{L}_0}$ and $\mathrm{M}^-_{\mathcal{L}_0}$ take a very simple form:

$$\mathrm{M}^+_{\mathcal{L}_0} v(x) = (2-\sigma)\int_{\mathbb{R}^n} \frac{\Lambda\delta(v,x,y)^+ - \lambda\delta(v,x,y)^-}{|y|^{n+\sigma}}\,\mathrm{d}y, \tag{3.4}$$

$$\mathrm{M}^-_{\mathcal{L}_0} v(x) = (2-\sigma)\int_{\mathbb{R}^n} \frac{\lambda\delta(v,x,y)^+ - \Lambda\delta(v,x,y)^-}{|y|^{n+\sigma}}\,\mathrm{d}y. \tag{3.5}$$

We will use these maximal operators to obtain regularity estimates. The factor $(2-\sigma)$ is important when $\sigma \to 2$. We need such a factor if we want to obtain second-order differential equations as limits of integro-differential equations. In terms of the regularity, we need the factor $(2-\sigma)$ for the estimates not to blow up as $\sigma \to 2$. This can be easily checked for fractional Laplacians.

Another interesting class is given when the kernels have the form

$$K(y) = (2-\sigma)\frac{y^t A y}{|y|^{n+2+\sigma}}$$

for symmetric matrices A such that $\lambda I \leq A \leq \Lambda I$. This is a smaller class than the $\mathcal{L}_0$ above if we choose the respective constants λ and Λ accordingly, but it is a large enough class to recover the classical Pucci extremal operators [7] as $\sigma \to 2$.

Let $K(x)$ be the supremum of $K_\alpha(x)$ where K_α are all the kernels of all operators $L \in \mathcal{L}$. As a replacement for (2.4), for any class $\mathcal{L}$ we will assume

$$\int_{\mathbb{R}^n} \frac{|y|^2}{|y|^2+1} K(y)\mathrm{d}y < +\infty. \tag{3.6}$$

Using the extremal operators, we give a general definition of ellipticity for nonlocal equations. The following is the kind of operators for which the results in this paper apply:

DEFINITION 3.1 Let $\mathcal{L}$ be a class of linear integro-differential operators. We always assume (3.6). An *elliptic operator I with respect to $\mathcal{L}$* is an operator with the following properties:

- If u is any bounded function, $Iu(x)$ is well-defined for all $u \in C^{1,1}(x)$.
- If u is C^2 in some open set Ω, then $Iu(x)$ is a continuous function in Ω.
- If u and v are bounded functions $C^{1,1}$ at x, then

$$\mathrm{M}^-_{\mathcal{L}}(u-v)(x) \leq Iu(x) - Iv(x) \leq \mathrm{M}^+_{\mathcal{L}}(u-v)(x). \tag{3.7}$$

This last hypothesis allows us to linearize the equation in the translation-invariant case.

Definition 2.2 applies for the general nonlocal elliptic operators of Definition 3.1 mutatis mutandis.

Definition 3.1 may apply to operators I whether or not they are translation invariant. However, in this paper we will only focus on the translation-invariant case. In other words, for all nonlinear operators I in this paper we assume that $\tau_z I u = I(\tau_z u)$, where τ_z is the translation operator $\tau_z u(x) := u(x-z)$.

We will show that any operator of the form (1.4) is elliptic with respect to any class that contains all the operators $L_{\alpha\beta}$ as long as condition (2.4) is satisfied (Lemma 3.2 and Lemma 4.2). However, Definition 3.1 allows a richer class of equations. For example, we can consider an operator I given by

$$Iu(x) = \int_{\mathbb{R}^n} \frac{G(u(x+y)-u(x))}{|y|^{n+\sigma}}\,\mathrm{d}\sigma$$

for any function G such that $G(0) = 0$ and $\lambda \leq G'(x) \leq \Lambda$. This operator I would be elliptic with respect to the class $\mathcal{L}_0$.

LEMMA 3.2 *Let I be an operator as in* (1.4) *and $\mathcal{L}$ be any collection of integro-differential operators. Assume every $L_{\alpha\beta}$ belongs to the class $\mathcal{L}$. Then for every $u, v \in C^{1,1}(x)$ we have*

$$\mathrm{M}^-_{\mathcal{L}}(u-v)(x) \leq Iu(x) - Iv(x) \leq \mathrm{M}^+_{\mathcal{L}}(u-v)(x).$$

PROOF: Since $u, v \in C^{1,1}(x)$, $L_{\alpha\beta}u(x)$ and $L_{\alpha\beta}v(x)$ are defined classically for any $L_{\alpha\beta}$. Clearly, for each α, β we have

$$L_{\alpha\beta}v + \mathrm{M}^-_{\mathcal{L}}(u-v) \leq L_{\alpha\beta}u \leq L_{\alpha\beta}v + \mathrm{M}^+_{\mathcal{L}}(u-v).$$

Applying $\sup_\alpha$ and then $\inf_\beta$, we conclude the proof

$$Iv + \mathrm{M}^-_{\mathcal{L}}(u-v) \leq Iu \leq Iv + \mathrm{M}^+_{\mathcal{L}}(u-v).$$

(We thank the referee for helping us to simplify the proof of this lemma.) □

The family of operators that satisfy condition (3.3) have another important property. Definition 2.2 is made so that we never have to evaluate the operator I in the original function u. Every time we touch u with a smooth function φ from above, we construct a test function $v \in C^{1,1}(x)$ to evaluate I. But it is implicit in the viscosity method that a quadratic contact from one side forces some infinitesimal flatness from the other. In our case, since $C^{1,1}$ exceeds the necessary regularity for the convergence of the integral, if I is any nonlinear operator that is an inf sup (or a sup inf) of linear operators that satisfy (3.3), then I can be evaluated classically in u at those points x where u can be touched from above with a paraboloid. This is explained in the next lemma.

LEMMA 3.3 *Let I be an operator as in* (1.4) *so that for every $K_{\alpha\beta}$ the equation* (3.3) *holds. If we have a subsolution, $Iu \geq f$ in Ω and φ is a C^2 function that touches u from above at a point $x \in \Omega$, then $Iu(x)$ is defined in the classical sense and $Iu(x) \geq f(x)$.*

PROOF: For any $r > 0$, we define

$$v_r = \begin{cases} \varphi & \text{in } B_r, \\ u & \text{in } \mathbb{R}^n \setminus B_r, \end{cases}$$

and we have $\mathrm{M}^+ v_r(x) \geq I v_r(x) \geq f(x)$. Thus

$$(2-\sigma)\int \delta(v_r, x, y)^+ \frac{\Lambda}{|y|^{n+\sigma}} - \delta(v_r, x, y)^- \frac{\lambda}{|y|^{n+\sigma}}\, \mathrm{d}y \geq f(x).$$

Since φ touches u from above at x, for any $y \in \mathbb{R}^n$, $\delta(v_r, x, y) \geq \delta(u, x, y)$. Since $v_r \in C^{1,1}(x)$, $|\delta(v_r, x, y)|/|y|^{n+\sigma}$ is integrable, so is $\delta(u, x, y)^+/|y|^{n+\sigma}$.

We have

$$(2-\sigma)\int \delta(v_r,x,y)^- \frac{\lambda}{|y|^{n+\sigma}}\,\mathrm{d}y \le$$
$$(2-\sigma)\int \delta(v_r,x,y)^+ \frac{\Lambda}{|y|^{n+\sigma}}\,\mathrm{d}y - f(x).$$

Since φ touches u from above at x, $\delta(v_r,x,y)$ will decrease as r decreases. Therefore, for every $r < r_0$

$$(2-\sigma)\int_{\mathbb{R}^n} \delta(v_r,x,y)^- \frac{\lambda}{|y|^{n+\sigma}}\,\mathrm{d}y \le \tag{3.8}$$
$$(2-\sigma)\int_{\mathbb{R}^n} \delta(v_{r_0},x,y)^+ \frac{\Lambda}{|y|^{n+\sigma}}\,\mathrm{d}y - f(x).$$

But $\delta(v_r,x,y)^-$ is monotone increasing as r decreases, and it converges to $\delta(u,x,y)^-$ as $r \to 0$. From the monotone convergence theorem

$$\lim_{r\to 0}(2-\sigma)\int_{\mathbb{R}^n} \delta(v_r,x,y)^- \frac{\lambda}{|y|^{n+\sigma}}\,\mathrm{d}y = (2-\sigma)\int_{\mathbb{R}^n} \delta(u,x,y)^- \frac{\lambda}{|y|^{n+\sigma}}\,\mathrm{d}y.$$

And from (3.8), the integrals are uniformly bounded and thus

$$(2-\sigma)\int_{\mathbb{R}^n} \delta(u,x,y)^- \frac{\lambda}{|y|^{n+\sigma}}\,\mathrm{d}y \le$$
$$(2-\sigma)\int_{\mathbb{R}^n} \delta(v_{r_0},x,y)^+ \frac{\Lambda}{|y|^{n+\sigma}}\,\mathrm{d}y - f(x) < +\infty.$$

Therefore, $\delta(u,x,y)/|y|^{n+\sigma}$ is integrable, and $L_{\alpha\beta}u$ is well-defined in the classical sense for any α and β. Thus, $Iu(x)$ is computable in the classical sense. The difference $\delta(v_r-u,x,y)/|y|^{n+\sigma}$ is monotone decreasing as $r \searrow 0$, converges to 0, and it is bounded by the integrable function $\delta(v_{r_0}-u,x,y)/|y|^{n+\sigma}$. We can pass to the limit in the following expression:

$$\begin{aligned}\lim_{r\to 0} \mathrm{M}^+(v_r-u)(x) &= \lim_{r\to 0}(2-\sigma)\int \delta(v_r-u,x,y)^+ \frac{\Lambda}{|y|^{n+\sigma}}\,\mathrm{d}y\\ &= 0\end{aligned}$$

Now we use Lemma 3.2 to conclude

$$Iu(x) \ge Iv_r(x) + \mathrm{M}^-(u-v_r) = f(x) - \mathrm{M}^+(v_r-u) \to f(x).$$

So $Iu(x) \ge f(x)$. □

Lemma 3.3 is convenient for making proofs involving M^+ and M^- because it allows us to deal with viscosity solutions almost as if they were classical solutions.

4 Stability Properties of Solutions and Equations under Appropriate Limits

In this section we show a few technical properties of the operators I like (1.4). First, if $u \in C^{1,1}(\Omega)$, then Iu is continuous in Ω. As mentioned in the previous sections, it is necessary to justify that operators of the form (1.4) satisfy the conditions of Definition 3.1. Next, we will show that our notion of viscosity solutions allows us to use test functions that are only punctually $C^{1,1}$ instead of C^2 in a neighborhood of the point where the functions touch. Then we will show the important stability property of Definition 2.2. Namely, we show that if a sequence of subsolutions (or supersolutions) in Ω converges in a suitable way on any compact set in $\mathbb{R}^n$, then the limit is also a subsolution (or supersolution).

We start with a technical real analysis lemma.

LEMMA 4.1 *Let $f \in L^\infty(\mathbb{R}^n)$ and g_α be a family of functions so that $|g_\alpha(x)| \le g(x)$ for some L^1 function g. Then the family $f * g_\alpha$ is equicontinuous in every compact set.*

PROOF: Let K be a compact set in $\mathbb{R}^n$. Let $\varepsilon > 0$. Since $g \in L^1$, we can pick a large R so that $K \subset B_R$ and

$$\|f\|_{L^\infty} \left(\int_{\mathbb{R}^n \setminus B_R(x)} g(y)\mathrm{d}y \right) \le \frac{\varepsilon}{8}$$

for any $x \in K$. We write $f = f_1 + f_2$, where $f_1 = f\chi_{B_{2R}}$ and $f_2 = f\chi_{\mathbb{R}^n \setminus B_{2R}}$. From the above inequality, we have $|f_2 * g_\alpha| \le \frac{\varepsilon}{8}$ in K.

Since $g \in L^1$, there is a $\delta_0 > 0$ so that

$$\int_A g(x)\mathrm{d}x < \frac{\varepsilon}{16\,\|f\|_{L^\infty}} \quad \text{for any set } |A| < \delta_0. \tag{4.1}$$

Let η_t be a standard mollifier with compact support. We have $f_1 * \eta_t \to f_1$ a.e. (in every Lebesgue point of f_1). Recall that the support of f_1 is in B_{2R}. For t large, $f_1 * \eta_t = 0$ outside B_{4R}. By Egorov's theorem, there is a set $A \subset B_{4R}$ such that

$$|A| < \delta_0, \tag{4.2}$$

$$f_1 * \eta_t \to f_1 \quad \text{uniformly in } \mathbb{R}^n \setminus A. \tag{4.3}$$

In particular, there is a $\tilde{f_1} = f_1 * \eta_{t_0}$ such that $|f_1 - \tilde{f_1}| < \varepsilon/(8\,\|g\|_{L^1})$ in $\mathbb{R}^n \setminus A$. We have

$$\|(f_1 - \tilde{f_1})(1-\chi_A) * g_\alpha\|_{L^\infty} \le \|(f_1 - \tilde{f_1})(1-\chi_A)\|_{L^\infty}\|g_\alpha\|_{L^1} < \frac{\varepsilon}{8}. \tag{4.4}$$

On the other hand, from (4.1) and (4.2), we also get

$$\|(f_1 - \tilde{f_1})\chi_A * g_\alpha\|_{L^\infty} < \frac{\varepsilon}{8}. \tag{4.5}$$

Since $\tilde{f}_1$ is continuous and $\|g_\alpha\|_{L^1}$ is bounded, the family $\tilde{f}_1 * g_\alpha$ is equicontinuous. There is a $\delta > 0$ so that $|\tilde{f}_1 * g_\alpha(x) - \tilde{f}_1 * g_\alpha(y)| < \frac{\varepsilon}{4}$ every time $|x-y| < \delta$. Moreover,

$$\begin{aligned}
&|f * g_\alpha(x) - f * g_\alpha(y)| \\
&\quad \le |\tilde{f}_1 * g_\alpha(x) - \tilde{f}_1 * g_\alpha(y)| + |(f_1 - \tilde{f}_1) * g_\alpha(x) - (f_1 - \tilde{f}_1) * g_\alpha(y)| \\
&\qquad + |f_2 * g_\alpha(x) - f_2 * g_\alpha(y)| \\
&\quad \le \frac{\varepsilon}{4} + |(f_1 - \tilde{f}_1)\chi_A * g_\alpha(x)| + |(f_1 - \tilde{f}_1)\chi_A * g_\alpha(y)| \\
&\qquad + |(f_1 - \tilde{f}_1)(1 - \chi_A) * g_\alpha(x)| + |(f_1 - \tilde{f}_1)(1 - \chi_A) * g_\alpha(y)| \\
&\qquad + |f_2 * g_\alpha(x)| + |f_2 * g_\alpha(y)| \\
&\quad \le \varepsilon
\end{aligned}$$

for any α and every time $|x - y| < \delta$. □

LEMMA 4.2 *Let I be an operator as in* (1.4)*, and assume only* (2.4)*. Let v be a bounded function in $\mathbb{R}^n$ and $C^{1,1}$ in some set Ω. Then Iv is continuous in Ω.*

PROOF: We must prove the $L_{\alpha\beta}v$ in (1.4) are equicontinuous. As in (2.4), we write $K = \sup_{\alpha\beta} K_{\alpha\beta}$. Let $\varepsilon > 0$ and $x_0 \in \Omega$. Since v is $C^{1,1}$ in Ω, there is a constant C so that

$$|\delta(v, x, y)| < C|y|^2 \quad \text{if } x \in \Omega \text{ and } |y| < \operatorname{dist}(x, \partial\Omega).$$

Let $r > 0$ be such that

$$\int_{B_r} C|y|^2 K(y)\mathrm{d}y < \frac{\varepsilon}{3}$$

(note that we write $B_r = B_r(0)$).

We have

$$\begin{aligned}
L_{\alpha\beta}v(x) &= \int_{\mathbb{R}^n} \delta(v, x, y) K_{\alpha\beta}(y)\mathrm{d}y \\
&= \int_{B_r} \delta(v, x, y) K_{\alpha\beta}(y)\mathrm{d}y + \int_{\mathbb{R}^n \setminus B_r} \delta(v, x, y) K_{\alpha\beta}(y)\mathrm{d}y \\
&=: w_1(x) + w_2(x)
\end{aligned}$$

where

$$|w_1| = \left| \int_{B_r} \delta(v, x, y) K_{\alpha\beta}(y)\mathrm{d}y \right| \le \int_{B_r} C|y|^2 K(y)\mathrm{d}y < \frac{\varepsilon}{3}$$

and

$$w_2 = \int_{\mathbb{R}^n \setminus B_r} (v(x+y) + v(x-y) - 2v(x)) K_{\alpha\beta}(y) \mathrm{d}y$$
$$= v * g_{\alpha\beta} + v * \hat{g}_{\alpha\beta} - 2 \left(\int g_{\alpha\beta} \, \mathrm{d}y \right) v$$

where $g_{\alpha\beta}(y) = \chi_{\mathbb{R}^n \setminus B_r}(y) K_{\alpha\beta}(y)$ and $\hat{g}_{\alpha\beta}(y) = g_{\alpha\beta}(-y)$. For any α and β, $g_{\alpha\beta} \leq \chi_{\mathbb{R}^n \setminus B_r} K$, which is in L^1. From Lemma 4.1, w_2 is equicontinuous. So there is a $\delta > 0$ such that

$$|w_2(x) - w_2(x_0)| < \frac{\varepsilon}{3} \quad \text{if } |x - x_0| < \delta.$$

Therefore

$$L_{\alpha\beta} v(x) - L_{\alpha\beta} v(x_0)| \leq |w_1(x)| + |w_1(x_0)| + |w_2(x) - w_2(x_0)| < \varepsilon$$

uniformly in α and β. Thus $|I v(x) - I v(x_0)| < \varepsilon$ every time $|x - x_0| < \delta$. □

When we gave the definition of viscosity solutions in Section 2, we used C^2 test functions. Now we show that it is equivalent to use punctually $C^{1,1}$ functions.

LEMMA 4.3 *Let I be elliptic with respect to some class $\mathcal{L}$ in the sense of Definition* 3.1. *Let $u : \mathbb{R}^n \to \mathbb{R}$ be an upper-semicontinuous function such that $I u \geq f(x)$ in Ω in the viscosity sense. Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a bounded function, punctually $C^{1,1}$ at a point $x \in \Omega$. Assume φ touches u from above at x. Then $I\varphi(x)$ is defined in the classical sense and $I\varphi(x) \geq f(x)$.*

PROOF: Since φ is $C^{1,1}$, the expression (2.3) is clearly integrable for every α and β and $I\varphi(x)$ is defined classically. Also, because φ is $C^{1,1}$, there is a quadratic polynomial q such that $|q(y) - \varphi(y)| \leq c|y - x|^2$ for $y \in B_r(x)$, where c can be chosen as small as we want as $r \to 0$. Let

$$v_r(x) = \begin{cases} q & \text{in } B_r, \\ u & \text{in } \mathbb{R}^n \setminus B_r. \end{cases}$$

Since $I u \geq f$ in Ω in the viscosity sense, then $I v_r(x) \geq f(x)$ with $I v_r(x)$ well-defined. Moreover, let

$$\varphi_r(x) = \begin{cases} q & \text{in } B_r, \\ \varphi & \text{in } \mathbb{R}^n \setminus B_r. \end{cases}$$

We have

$$\begin{aligned} I\varphi(x) &\geq I\varphi_r(x) + \mathrm{M}^-_{\mathcal{L}}(\varphi - \varphi_r)(x) \\ &\geq I v_r(x) + \mathrm{M}^-_{\mathcal{L}}(\varphi_r - v_r)(x) + \mathrm{M}^-_{\mathcal{L}}(\varphi - \varphi_r)(x) \\ &\geq I v_r(x) + \mathrm{M}^-_{\mathcal{L}}(\varphi - \varphi_r)(x) \quad \text{(since } \varphi_r - v_r \text{ has a minimum at } x) \geq \end{aligned}$$

$$\geq f(x) + \mathrm{M}^-_{\mathcal{L}}(\varphi - \varphi_r)(x)$$

$$\geq f(x) + \int_{B_r} \delta(q-\varphi, x, y)^- K(y)\,\mathrm{d}y \quad \text{(where K is the one from (3.6))}$$

$$\geq f(x) - \int_{B_r} 2c|y|^2 K(y)\,\mathrm{d}y.$$

Since $|y|^2 K(y)$ is integrable in a neighborhood of the origin, the expression

$$\int_{B_r} 2c|y|^2 K(y)\,\mathrm{d}y$$

goes to 0 as $r \to 0$. Thus, for any $\varepsilon > 0$, we can find a small r so that

$$I\varphi(x) \geq f(x) - \varepsilon.$$

Therefore $I\varphi(x) \geq f(x)$. □

One of the most useful properties of viscosity solutions is their stability under uniform limits on compact sets. We will prove a slightly stronger result. We show that the notion of viscosity supersolution is stable with respect to the natural limits for lower-semicontinuous functions. This type of limit is well-known in variational analysis and usually called a Γ-limit. In the viscosity solution community it is sometimes referred to as the "half-relaxed limit."

DEFINITION 4.4 (Γ-convergence) A sequence of lower-semicontinuous functions u_k Γ-*converges* to u in a set Ω if the two following conditions hold:

- For every sequence $x_k \to x$ in Ω, $\liminf_{k\to\infty} u_k(x_k) \geq u(x)$.
- For every $x \in \Omega$, there is a sequence $x_k \to x$ in Ω such that

$$\limsup_{k\to\infty} u_k(x_k) = u(x).$$

Naturally, a uniformly convergent sequence u_k would also converge in the Γ sense. An important property of Γ-limits is that if u_k Γ-converges to u, and u has a strict local minimum at x, then u_k will have a local minimum at x_k for a sequence $x_k \to x$.

LEMMA 4.5 *Let I be elliptic in the sense of Definition* 3.1 *and u_k be a sequence of functions that are uniformly bounded in $\mathbb{R}^n$ and lower-semicontinuous in Ω such that*

(i) *$Iu_k \leq f_k$ in Ω,*
(ii) *$u_k \to u$ in the Γ sense in Ω,*
(iii) *$u_k \to u$ a.e. in $\mathbb{R}^n$, and*
(iv) *$f_k \to f$ locally uniformly in Ω for some continuous function f.*

Then $Iu \leq f$ in Ω.

PROOF: Let φ be a test function from below for u touching at a point x in a neighborhood N. Since u_k Γ-converges to u in Ω, for large k we can find x_k and d_k such that $\varphi + d_k$ touches u_k at x_k. Moreover, $x_k \to x$ and $d_k \to 0$ as $k \to +\infty$. Since $I u_k \leq f_k$, if we let

$$v_k = \begin{cases} \varphi + d_k & \text{in } N, \\ u_k & \text{in } \mathbb{R}^n \setminus N, \end{cases}$$

we have $I v_k(x_k) \leq f_k(x_k)$.

Let $z \in N$ be such that $\operatorname{dist}(z, \partial N) > \rho > 0$. We have

$$\begin{aligned} |I v_k(z) - I v(z)| &\leq \max\left(|\mathrm{M}^+_{\mathcal{L}}(v_k - v)(z)|, |\mathrm{M}^+_{\mathcal{L}}(v - v_k)(z)|\right) \\ &\leq \sup_{L \in \mathcal{L}} |L(v_k - v)(z)|, \end{aligned}$$

$$y \leq \int_{\mathbb{R}^n} |\delta(v_k - v, z, y)| K(y) \mathrm{d}y \leq \int_{\mathbb{R}^n \setminus B_\rho} |\delta(v_k - v, z, y)| K(y) \mathrm{d}y.$$

The sequence v_k is bounded and $\delta(v_k - v, z, y)$ converges to 0 almost everywhere. Since $K \in L^1(\mathbb{R}^n \setminus B_\rho)$, we can use the dominated convergence theorem to show that the above expression goes to 0 as $k \to +\infty$. Moreover, the convergence is uniform in z. We obtain $I v_k \to I v$ locally uniformly in N.

From Definition 3.1, we have that $I v$ is continuous in N. We now compute

$$|I v_k(x_k) - I v(x)| \leq |I v_k(x_k) - I v(x_k)| + |I v(x_k) - I v(x)| \to 0.$$

So $I v_k(x_k)$ converges to $I v(x)$ as $k \to +\infty$. Since $x_k \to x$ and $f_k \to f$ locally uniformly, we also have $f_k(x_k) \to f(x)$, which finally implies $I v(x) \leq f(x)$. $\square$

Remark 4.6. In the previous lemma it was used that since the sequence of f_k converges locally uniformly to a continuous f, then for any sequence $x_k \to x$ we have $f_k(x_k) \to f(x)$. The hypothesis of the lemma could be relaxed to the case where f is only upper-semicontinuous and $-f_k$ Γ-converges to $-f$ since in that case we would have $\limsup f_k(x_k) \leq f(x)$. For simplicity, in this paper we are only considering continuous right-hand sides.

In the previous lemma we showed the stability of supersolutions under Γ-limits. Naturally, we also have the corresponding result for subsolutions. In that case we would consider the natural limit in the space of upper-semicontinuous functions, which is the same as the Γ-convergence of $-u_k$ to $-u$. As a corollary, we obtain the stability under uniform limits.

COROLLARY 4.7 *Let I be elliptic in the sense of Definition* 3.1 *and u_k be a sequence of functions that are bounded in $\mathbb{R}^n$ and continuous in Ω such that*

(i) *$I u_k = f_k$ in Ω,*
(ii) *$u_k \to u$ locally uniformly in Ω,*

(iii) $u_k \to u$ *a.e. in* $\mathbb{R}^n$*, and*

(iv) $f_k \to f$ *locally uniformly in* Ω *for some continuous function* f.

Then $Iu = f$ *in* Ω.

Remark 4.8. Γ-*convergence* was introduced by De Giorgi in the framework of variational analysis to study convergence of sequences of functionals in Banach spaces. Here we are using the same notion of convergence for functions in $\mathbb{R}^n$. This type of limit usually appears in viscosity solution theory in one form or another, even though the term Γ-*convergence* is rarely used.

5 Comparison Principle

The purpose of this section is to establish a basic comparison principle for (super- and sub-) solutions to our equations.

It is, in principle, not possible to compare two solutions at any given point since they may not have the appropriate *classical* behavior simultaneously. For the second-order case that means C^2 behavior; for the case of integro-differential equations, a quadratic behavior from one side would be enough. In our case, then, the classical sup- and inf-convolution method of Jensen works in a straightforward fashion. It is also important that in Theorem 5.9 we establish that the difference of two (super- and sub-) solutions satisfies itself an *elliptic* equation.

The method of Jensen [9] has been successfully applied to prove uniqueness results for integro-differential equations already [1]. In [2] a very general proof was given. Our definitions do not quite fit into the previous framework mainly because we consider the abstract class of operators given by Definition 3.1, and we allow discontinuities outside of the domain of the equation Ω. However, with small modifications, the same techniques can be adapted to our equations. We sketch the important ideas to prove the comparison principle in this section.

In order to have a comparison principle for a nonlinear operator I, we need to impose a minimal ellipticity condition to our collection of linear operators $\mathcal{L}$. The following assumption will suffice:

Assumption 5.1. There is a constant $R_0 \geq 1$ so that for every $R > R_0$, there exists a $\delta > 0$ (which could depend on R) such that for any operator L in $\mathcal{L}$, we have that $L\varphi > \delta$ in B_R, where φ is given by

$$\varphi(x) = \min(1, |x|^2/R^3).$$

In later sections we will need stronger assumptions to prove further regularity properties of the solutions. But for the comparison principle Assumption 5.1 is enough. Note that Assumption 5.1 is very mild. It just says that given the particular function $\min(1, |x|^2/R^3)$, the value of the operator will be strictly positive in B_R, but it does not require any uniform estimate on how that happens. If the operators $L \in \mathcal{L}$ are scale invariant, it just means that when we apply them to $\min(1, |x|^2)$, they are strictly positive in some neighborhood of the origin.

THEOREM 5.2 *Let $\mathcal{L}$ be some class satisfying Assumption* 5.1. *Let I be elliptic with respect to $\mathcal{L}$ in the sense of Definition* 3.1. *Let Ω be a bounded open set, and u and v be two functions such that*

(i) *u, v are bounded in $\mathbb{R}^n$,*
(ii) *u is upper-semicontinuous at every point in $\bar{\Omega}$,*
(iii) *v is lower-semicontinuous at every point in $\bar{\Omega}$,*
(iv) *$Iu \geq f$ and $Iv \leq f$ in Ω for some continuous function f, and*
(v) *$u \leq v$ in $\mathbb{R}^n \setminus \Omega$.*

Then $u \leq v$ in Ω.

By u being upper-semicontinuous at every point in $\bar{\Omega}$, we mean that u is semicontinuous in $\bar{\Omega}$ with respect to $\mathbb{R}^n$. The same applies for the function v.

We will use the usual idea of sup- and inf-convolutions in order to prove comparison. We start by defining these concepts.

DEFINITION 5.3 Given an upper-semicontinuous function u, the *sup-convolution approximation* u^ε is given by

$$u^\varepsilon(x) = \sup_y u(x+y) - \frac{|y|^2}{\varepsilon}. \tag{5.1}$$

On the other hand, if u is lower-semicontinuous, the *inf-convolution* u_ε is given by

$$u_\varepsilon(x) = \inf_y u(x+y) + \frac{|y|^2}{\varepsilon}. \tag{5.2}$$

Notice that $u^\varepsilon \geq u$ and $u_\varepsilon \leq u$. Note also that u^ε is a supremum of translations of u, and u_ε is an infimum of translations of u. Finally, notice that if for x the supremum is attained at y, i.e., $u^\varepsilon(x) = u(x+y) - |y|^2/\varepsilon$, then

$$\begin{aligned} u^\varepsilon(x+z) &\geq u(x+z+y-z) - \frac{|y-z|^2}{\varepsilon} \\ &= u^\varepsilon(x) - \frac{|z|^2}{}\varepsilon + 2\,\frac{z \cdot y}{\varepsilon} \end{aligned}$$

so u^ε is semiconvex.

The following two propositions are very standard, so we skip their proofs:

PROPOSITION 5.4 *If u is bounded and lower-semicontinuous in $\mathbb{R}^n$, then u_ε Γ-converges to u. If u is bounded and upper-semicontinuous in $\mathbb{R}^n$, then $-u^\varepsilon$ Γ-converges to $-u$.*

PROPOSITION 5.5 *If f is a continuous function and $Iu \geq f$, then $Iu^\varepsilon \geq f - d_\varepsilon$. And if $Iv \leq f$, then $Iv_\varepsilon \leq f + d_\varepsilon$, where $d_\varepsilon \to 0$ as $\varepsilon \to 0$ and depends on the modulus of continuity of f.*

Remark 5.6. Proposition 5.4 is a straightforward generalization of the fact that $u^\varepsilon \to u$ locally uniformly if u is continuous.

LEMMA 5.7 *Let $u : \mathbb{R}^n \to \mathbb{R}$ be a lower-semicontinuous function in $\mathbb{R}^n$ such that $Iu \leq 0$ in Ω in the viscosity sense. Let x be a point in Ω so that $u \in C^{1,1}(x)$. Then $Iu(x)$ is defined in the classical sense and $Iu(x) \leq 0$.*

PROOF: Use u as a test function for itself with Lemma 4.3. □

LEMMA 5.8 *Let I be elliptic in the sense of Definition* 3.1. *Let u and v be two bounded functions such that*

(i) *u is upper-semicontinuous and v is lower-semicontinuous in $\mathbb{R}^n$, and*
(ii) *$Iu \geq f$ and $Iv \leq g$ in the viscosity sense in Ω for two continuous functions f and g.*

Then $\mathrm{M}^+_{\mathcal{L}}(u - v) \geq f - g$ in Ω in the viscosity sense.

PROOF: By Proposition 5.5, we also have that $Iu^\varepsilon \geq f - d_\varepsilon$ and $Iv_\varepsilon \leq g + d_\varepsilon$. Moreover, $-u^\varepsilon \to -u$ and $v_\varepsilon \to v$ in the Γ sense. By the stability of viscosity solutions under Γ-limits and since $d_\varepsilon \to 0$, it is enough to show that $\mathrm{M}^+_{\mathcal{L}}(u^\varepsilon - v_\varepsilon) \geq f - g - 2d_\varepsilon$ in Ω for every $\varepsilon > 0$.

Let φ be a C^2 function touching $(u^\varepsilon - v_\varepsilon)$ from above at the point x. For any $\varepsilon > 0$ both functions u^ε and $-v_\varepsilon$ are semiconvex, which means that for each of them there is a paraboloid touching it from below at every point x. If a C^2 function touches $(u^\varepsilon - v_\varepsilon)$ from above at the point x, then both u^ε and $-v_\varepsilon$ must be $C^{1,1}(x)$. But by Lemma 5.7 and Definition 3.1, this means that we can evaluate $Iu^\varepsilon(x)$ and $Iv_\varepsilon(x)$ in the classical sense and

$$\mathrm{M}^+_{\mathcal{L}}(u^\varepsilon - v_\varepsilon)(x) \geq Iu^\varepsilon(x) - Iv_\varepsilon(x) \geq f - g - 2d_\varepsilon,$$

which clearly implies that also $\mathrm{M}^+_{\mathcal{L}}\varphi(x) \geq f - g - 2d_\varepsilon$ since φ touches $u^\varepsilon - v_\varepsilon$ from above. Thus $\mathrm{M}^+_{\mathcal{L}}(u^\varepsilon - v_\varepsilon) \geq f - g - 2d_\varepsilon$ in Ω in the viscosity sense.

Taking $\varepsilon \to 0$ and using Lemma 4.5 we finish the proof. □

The result of Lemma 5.8 is almost the result we need to prove the comparison principle, except that we want to allow functions u and v that are discontinuous outside of the domain Ω. We fix this last detail in the following theorem.

THEOREM 5.9 *Let I be elliptic in the sense of Definition* 3.1. *Let u and v be two bounded functions in $\mathbb{R}^n$ such that*

(i) *u is upper-semicontinuous and v is lower-semicontinuous in $\overline{\Omega}$, and*
(ii) *$Iu \geq f$ and $Iv \leq g$ in the viscosity sense in Ω for two continuous functions f and g.*

Then $\mathrm{M}^+_{\mathcal{L}}(u - v) \geq f - g$ in Ω in the viscosity sense.

PROOF: First we will show that there exist two sequences u_k and v_k, upper- and lower-semicontinuous, respectively, such that

- $u_k = u$ in $\overline{\Omega}$ for every n,
- $v_k = v$ in $\overline{\Omega}$ for every n,

- $u_k \to u$ and $v_k \to v$ a.e. in $\mathbb{R}^n \setminus \overline{\Omega}$, and
- $I u_k \geq f_k$ and $I v_k \leq g_k$ with $f_k \to f$ and $g_k \to g$ locally uniformly in Ω.

It is clear that we can find two sequences u_k and v_k satisfying the first three items above by doing a standard mollification of u and v away from Ω and then *filling the gap* in a semicontinuous way. What we will show is that then there are functions f_k and g_k for which the fourth item also holds.

The function $u_k - u$ vanishes in Ω and thus $\mathrm{M}^-_{\mathcal{L}}(u_k - u)$ is defined in the classical sense in Ω. Moreover,

$$\mathrm{M}^-_{\mathcal{L}}(u_k - u)(x) \geq \\ -2 \int_{\mathbb{R}^n \setminus B_{\mathrm{dist}(x,\partial\Omega)}(x)} |u_k(x+y) - u(x+y)| K(y) \mathrm{d}y =: h_k(x).$$

Note that h_k is continuous in Ω, and by dominated convergence $h_k \to 0$ locally uniformly in Ω as $k \to \infty$.

Let φ be a function touching globally u_k from above at a point x; assume only that $\varphi \in C^{1,1}(x)$. Then also $\varphi + u - u_k \in C^{1,1}(x)$. But $\varphi + u - u_k$ touches u from above at x, so by Lemma 4.3 $I(\varphi + u - u_k)(x) \geq f(x)$. But now

$$I\varphi(x) \geq I(\varphi + u - u_k)(x) + \mathrm{M}^-_{\mathcal{L}}(u - u_k)(x) \geq f(x) + h_k(x),$$

so we prove the fourth item above for u_k by choosing $f_k = f + h_k$. Similarly, we prove it for v_k.

Now that we have such sequences u_k and v_k, we apply Lemma 4.5 and finish the proof. □

LEMMA 5.10 *Let u be a bounded function, upper-semicontinuous at every point in $\overline{\Omega}$, such that $\mathrm{M}^+_{\mathcal{L}} u \geq 0$ in the viscosity sense in Ω. Then $\sup_\Omega u \leq \sup_{\mathbb{R}^n \setminus \Omega} u$.*

PROOF: Let us choose $R > R_0$ large enough so that $\Omega \subset B_R$. For any $\varepsilon > 0$, let φ_M be the function

$$\varphi_M(x) = M + \varepsilon(1 - \min(1, |x|^2/R^3)).$$

Note that $M \leq \varphi_M(x) \leq M + \varepsilon$ for every $x \in \mathbb{R}^n$. Also, by Assumption 5.1, there is a $\delta > 0$ such that $\mathrm{M}^+_{\mathcal{L}} \varphi_M(x) \leq -\varepsilon\delta$ for any $x \in B_R$.

Let M_0 be the smallest value of M for which $\varphi_M \geq u$ in $\mathbb{R}^n$. We will show that $M_0 \leq \sup_{\mathbb{R}^n \setminus \Omega} u$. Otherwise, if $M_0 > \sup_{\mathbb{R}^n \setminus \Omega} u$, there must be a point $x_0 \in \Omega$ for which $u(x_0) = \varphi_{M_0}(x_0)$. But in that case φ_{M_0} would touch u from above at $x_0 \in \Omega$, and by the definition of $\mathrm{M}^+_{\mathcal{L}} u \geq 0$ in the viscosity sense we would have that $\mathrm{M}^+_{\mathcal{L}} \varphi_{M_0} \geq 0$, a contradiction. Therefore, for every $x \in \mathbb{R}^n$, we have

$$u(x) \leq \varphi_{M_0}(x) \leq M_0 + \varepsilon \leq \sup_{\mathbb{R}^n \setminus \Omega} u + \varepsilon.$$

We finish the proof by making $\varepsilon \to 0$. □

PROOF OF THEOREM 5.2: By Theorem 5.9, $M^+_{\mathcal{L}}(u-v) \geq 0$ in Ω. Then Lemma 5.10 says that $\sup_\Omega(u-v) \leq \sup_{\mathbb{R}^n\setminus\Omega}(u-v)$, which finishes the proof. $\square$

Once we have the comparison principle for semicontinuous sub- and supersolutions, existence of the solution of the Dirichlet problem follows using Perron's method [8] as long as we can construct suitable barriers.

6 Second-Order Elliptic Equations as Limits of Fractional Diffusions

In this section we briefly discuss how second-order equations appear as limits of fractional diffusions as σ goes to 2. This shows in particular how much larger and complex the family of fractional, fully nonlinear equations is compared to the second-order case.

It is well-known that

$$\lim_{\sigma\to 2} \int_{\mathbb{R}^n} \frac{c_n(2-\sigma)}{|y|^{n+\sigma}}\,\delta(u,x,y)\mathrm{d}y = \lim_{\sigma\to 2} -(-\triangle)^{\sigma/2}u(x) = \triangle u(x).$$

With a simple change of variables $z = Ay$, we arrive at the following identity:

$$\lim_{\sigma\to 2} \int_{\mathbb{R}^n} \frac{c_n(2-\sigma)}{\det A|A^{-1}z|^{n+\sigma}}\,\delta(u,x,z)\mathrm{d}z = \sum a_{ij}u_{ij}(x), \tag{6.1}$$

where $\{a_{ij}\}$ are the entries of AA^{T}. This means that we can recover any linear second-order elliptic operator as a limit of integro-differential ones like (6.1). Moreover, let us say we have a fully nonlinear operator of the form $F(D^2u)$. Let us assume the function F is Lipschitz and monotone in the space of symmetric matrices. Then F can be written as

$$F(M) = \inf_\alpha \sup_\beta \left(\sum a^{\alpha\beta}_{ij} M_{ij} + b^{\alpha\beta} \right)$$

for some collection of positive matrices $\{a^{\alpha\beta}_{ij}\} = A_{\alpha\beta}A^{\mathsf{T}}_{\alpha\beta}$ and constants $b^{\alpha\beta}$. Thus any elliptic fully nonlinear operator can be recovered as a limit of integro-differential operators as

$$F(D^2u) = \lim_{\sigma\to 2}\left(\inf_\alpha \sup_\beta \left(\int \frac{c_n(2-\sigma)}{\det A_{\alpha\beta}|A^{-1}_{\alpha\beta}z|^{n+\sigma}}\,\delta(u,x,z)\mathrm{d}z + b^{\alpha\beta} \right)\right)$$

as long as the limit commutes with the operations of infimum and supremum. That is going to be the case every time the convergence is uniform in α and β, which is the case, for example, if the matrices $A_{\alpha\beta}$ are uniformly elliptic.

Another possibility is to take a family $A_{\alpha\beta}$ so that

$$F(D^2u) = \lim_{\sigma\to 2}\left(\inf_\alpha \sup_\beta \int \frac{\delta(u,x,A_{\alpha\beta}y)}{|y|^{n+\sigma}}\,\mathrm{d}y \right).$$

Note that we can also consider operators of the form

$$Iu(x) := (2-\sigma)\int \frac{1}{|y|^{n+\sigma-2}}\, G\left(\frac{\delta(u,x,y)}{|y|^2}, y\right)\mathrm{d}y$$

with $G(d, y)$ being an arbitrary function, Lipschitz and monotone in d, such that $G(0, y) = 0$. This suggests an unusual family of second-order nonlinear equations: for P a quadratic polynomial,

$$F(D^2 P) = \int\limits_{\partial B_1} G(P(\sigma), \sigma)\mathrm{d}\sigma.$$

7 Overview of the Regularity Results

In the following sections we establish for integral diffusions the main techniques and theorems to reproduce the regularity theory for fully nonlinear second-order equations. A brief review of these steps follows (see [7]).

An essential step in the regularity theory for the second-order case is a Harnack inequality and Hölder continuity for solutions of *elliptic equations with bounded measurable coefficients*; this is the Krylov-Safonov theory.

By "solutions of elliptic equations with bounded measurable coefficients" we understand the class S of continuous functions u satisfying in the viscosity sense the extremal inequalities

$$\mathrm{M}^-(u) \le 0 \le \mathrm{M}^+(u)$$

where M^- and M^+ denote the extremal Pucci operators.

$$\mathrm{M}^-(u) \text{ (resp., } \mathrm{M}^+(u)) = \inf_{\lambda I \le a_{ij} \le \Lambda I} \text{(resp. sup) } a_{ij}\partial_{ij}u(x).$$

Heuristically these inequalities imply that for some pointwise choice of $a_{ij}(x)$, we have $a_{ij}(x)\partial_{ij}u(x) = 0$.

The Krylov-Safonov theorem asserts that if $u \in S$ and it is nonnegative in B_1, then $\sup_{B_{1/2}} u \le C \inf_{B_{1/2}} u$. This implies in particular that u is Hölder-continuous.

In Sections 11 and 12 we develop the corresponding nonlocal Harnack inequality. A fundamental tool in this theorem is a restricted version of the Aleksandrov-Bakel′man-Pucci (ABP) theorem.

The classical ABP estimate allows us to relate a pointwise estimate with an estimate in measure. It says that if w satisfies

(1) $a_{ij}(x)\partial_{ij}w(x) \ge -C$ in B_1,
(2) $w \le 0$ on ∂B_1, and
(3) $w(0) \ge 1$,

then $|\{w > 0\}| \ge \gamma_0 > 0$.

Using this estimate, we can prove an estimate in L^ε through a Calderon-Zygmund-type decomposition (Section 10). For that we need a *localization* argument for the set where $w > 0$ in the ABP above (Section 9). The estimate in L^ε says that if

(1) $u \in S$,
(2) $u \geq 0$ in B_1, and
(3) $u(0) = 1$,

then $|\{u > \lambda\}| < \lambda^{-\varepsilon}$.

The use of the L^ε estimate from both sides (for u and for $C - u$) gives a decrease-in-oscillation lemma that implies Hölder continuity of u (Section 12) and a Harnack inequality (Section 11) under somewhat more stringent conditions.

With these tools at hand we can pass to the regularity of the solution to translation-invariant, fully nonlinear equations. Indeed, if v is a solution of such an equation, for any h

$$\frac{v(x+h) - v(x)}{|h|}$$

is in the class S (because of Theorem 5.9). Therefore we should expect the derivatives of v to be Hölder continuous, i.e., $u \in C^{1,\alpha}$ for some $\alpha > 0$ (Section 13).

Finally, if the fully nonlinear, translation-invariant operator is convex (i.e., a supremum of linear operators L_α), then the average of two solutions $(v_1 + v_2)/2$ is a supersolution. In particular, the second-order differential quotient $(v(x+h) + v(x-h) - 2v(x))/2$ is the difference of a supersolution and a solution, and thus it is itself a supersolution of an extremal operator. The Evans-Krylov theorem should imply that the function is more regular than the order of the equation, that is, $v \in C^{\sigma+\alpha}$. This is the topic of a forthcoming paper. We now start to develop the plan presented above. We start with a nonlocal ABP estimate.

8 A Nonlocal ABP Estimate

The Aleksandrov-Bakel′man-Pucci (ABP) estimate is a key ingredient in the proof of the Harnack inequality by Krylov and Safonov. It is the relation that allows us to pass from an estimate in measure to a pointwise estimate. In this section we obtain an estimate for integro-differential equations that converges to the ABP estimate as σ approaches 2. In a later section, we will use this nonlocal version of the ABP theorem to prove the Harnack inequality.

In this and the next few sections we will consider the class $\mathcal{L}_0$ defined by condition (3.3). We write M^+ and M^- to denote $\mathrm{M}^+_{\mathcal{L}_0}$ and $\mathrm{M}^-_{\mathcal{L}_0}$.

Let u be a function that is not positive outside the ball B_1. Consider its concave envelope Γ in B_3 defined as

$$\Gamma(x) := \begin{cases} \min\left\{p(x) : \text{for all planes } p \geq u^+ \text{ in } B_3\right\} & \text{in } B_3, \\ 0 & \text{in } \mathbb{R}^n \setminus B_3. \end{cases}$$

We will concentrate in the contact set $\Sigma = \{u = \Gamma\} \cap B_1$. The first lemma establishes that if $x_0 \in \Sigma$, u stays quadratically close to the tangent plane to Γ at x_0 in a large portion of a neighborhood around x_0.

LEMMA 8.1 *Let $u \leq 0$ in $\mathbb{R}^n \setminus B_1$. Let Γ be its concave envelope in B_3. Assume $\mathrm{M}^+ u(x) \geq -f(x)$ in B_1. Let $\rho_0 = 1/(8\sqrt{n})$, $r_k = \rho_0 2^{-1/(2-\sigma)-k}$, and $R_k(x) = B_{r_k}(x) \setminus B_{r_{k+1}}(x)$.*

There is a constant C_0 depending only on n and λ (but not on σ) such that for any $x \in \{u = \Gamma\}$ and any $M > 0$, there is a k such that

$$|R_k(x) \cap \{u(y) < u(x) + (y-x)\cdot\nabla\Gamma(x) - M r_k^2\}| \leq C_0 \frac{f(x)}{M} |R_k(x)| \tag{8.1}$$

where $\nabla\Gamma$ stands for any element of the superdifferential of Γ at x that will coincide with its gradient, and also the gradient of u, when these functions are differentiable.

PROOF: Since u can be touched by a plane from above at x, from Lemma 3.3, $\mathrm{M}^+ u(x)$ is defined classically and we have

$$\mathrm{M}^+ u(x) = (2-\sigma)\int_{\mathbb{R}^n} \frac{\Lambda\delta^+ - \lambda\delta^-}{|y|^{n+\sigma}}\,\mathrm{d}y.$$

Recall $\delta = \delta(u,x,y) := u(x+y) + u(x-y) - 2u(x)$. Note that if both $x+y \in B_3$ and $x-y \in B_3$, then $\delta(u,x,y) \leq 0$, since $u(x) = \Gamma(x) = p(x)$ for some plane p that remains above u in the whole ball B_3. Moreover, if either $x+y \notin B_3$ or $x-y \notin B_3$, then both $x+y$ and $x-y$ are not in B_1, so $u(x+y) \leq 0$ and $u(x-y) \leq 0$. Therefore, in any case $\delta(u,x,y) \leq 0$. Thus we have

$$\begin{aligned}
-f(x) \leq \mathrm{M}^+ u(x) &= (2-\sigma)\int_{\mathbb{R}^n} \frac{-\lambda\delta^-}{|y|^{n+\sigma}}\,\mathrm{d}y \\
&\leq (2-\sigma)\int_{B_{r_0}(x)} \frac{-\lambda\delta^-}{|y|^{n+\sigma}}\,\mathrm{d}y
\end{aligned}$$

where $r_0 = \rho_0 2^{-1/(2-\sigma)}$.

Splitting the integral in the rings R_k and reorganizing terms, we obtain

$$f(x) \geq (2-\sigma)\lambda \sum_{k=0}^{\infty} \int_{R_k(x)} \frac{\delta^-}{|y|^{n+\sigma}}\,\mathrm{d}y.$$

Let us assume that equation (8.1) does not hold. We will arrive at a contradiction. We can use the opposite of (8.1) to estimate each integral in the terms of the previous equation:

$$\begin{aligned}
f(x) &\geq (2-\sigma)\lambda \sum_{k=0}^{\infty} \int_{R_k(x)} \frac{\delta^-}{|y|^{n+\sigma}}\,\mathrm{d}y \\
&\geq c(2-\sigma)\sum_{k=0}^{\infty} M \frac{r_k^2}{r_k^\sigma} C_0 \frac{f(x)}{M} \geq
\end{aligned}$$

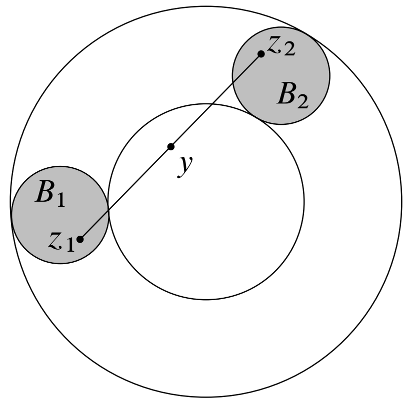

FIGURE 8.1. The balls B_1 and B_2.

$$\geq c(2-\sigma)\frac{\rho_0^2}{1-2^{-(2-\sigma)}}\,C_0 f(x)$$

$$\geq c C_0 f(x)$$

where the last inequality holds because $(2-\sigma)/(1-2^{-(2-\sigma)})$ remains bounded below for $\sigma \in (0,2)$. By choosing C_0 large enough, we obtain a contradiction. $\square$

Remark 8.2. Note that Lemma 8.1 implies that if $\mathrm{M}^+u(x) \geq g(x)$, then $u(x) \neq \Gamma(x)$ at every point where $g(x) > 0$.

Remark 8.3. Lemma 8.1 would hold for any particular choice of ρ_0 (modifying C_0 accordingly). The particular choice $\rho_0 = 1/(8\sqrt{n})$ is convenient for the proofs in Section 10 in this paper.

The next lemma says that the *large-portion* estimate for u in Lemma 8.1 implies a uniform quadratic detachment of Γ from its tangent plane at x_0 in a smaller ball.

LEMMA 8.4 *Let Γ be a concave function in B_r. Assume that for a small ε*

$$|\{y : \Gamma(y) < \Gamma(x) + (y-x)\cdot\nabla\Gamma(x) - h\} \cap (B_r(x) \setminus B_{r/2}(x))| \leq \varepsilon |B_r(x) \setminus B_{r/2}(x)|, \tag{8.2}$$

then $\Gamma(y) \geq \Gamma(x) + (y-x)\cdot\nabla\Gamma(x) - h$ in the whole ball $B_{r/2}(x)$.

PROOF: Let $y \in B_{r/2}(x)$. There are two points y_1 and y_2 in $B_r(x) \setminus B_{r/2}(x)$ such that

(1) $y = (y_1 + y_2)/2$ and
(2) $|y_1 - x| = |y_2 - x| = \frac{3}{4}r$.

Let us consider the balls $B_1 = B_{r/4}(y_1)$ and $B_2 = B_{r/4}(y_2)$ (see Figure 8.1). They are symmetrical with respect to y, and they are completely contained in $B_r(x) \setminus B_{r/2}(x)$. If ε is small enough, there will be two points $z_1 \in B_1$ and $z_2 \in B_2$ so that

(1) $y = (z_1 + z_2)/2$,
(2) $\Gamma(z_1) \geq \Gamma(x) + (z_1 - x) \cdot \nabla\Gamma(x) - h$, and
(3) $\Gamma(z_2) \geq \Gamma(x) + (z_2 - x) \cdot \nabla\Gamma(x) - h$,

and by the concavity of Γ we finish the proof since $\Gamma(y) \geq (\Gamma(z_1) + \Gamma(z_2))/2$. $\square$

COROLLARY 8.5 *For any $\varepsilon_0 > 0$ there is a constant C such that for any function u with the same hypothesis as in Lemma* 8.1*, there is an $r \in (0, \rho_0 2^{-1/(2-\sigma)})$ such that*

$$\frac{\left|\{y \in B_r \setminus B_{r/2}(x) : u(y) < u(x) + (y-x) \cdot \nabla\Gamma(x) - Cf(x)r^2\}\right|}{|B_r(x) \setminus B_{r/2}(x)|} \leq \varepsilon_0, \tag{8.3}$$

and

$$|\nabla\Gamma(B_{r/4}(x))| \leq Cf(x)^n |B_{r/4}(x)|. \tag{8.4}$$

Recall $\rho_0 = 1/(8\sqrt{n})$.

PROOF: By choosing $M = Cf(x)/\varepsilon_0$, we have (8.3) right away from Lemma 8.1. Equation (8.4) then follows as a consequence of Lemma 8.4 and concavity. $\square$

The previous lemmas allow us to get a lower bound on the volume of the union of balls where Γ (and u) detach quadratically from the corresponding tangent planes to Γ by the volume of the image of the gradient map, as in the standard ABP theorem. A rough estimate would be the following:

LEMMA 8.6 *Let $B_r(x)$ with $x \in \Sigma$ be the family of balls $B_{r_k}(x)$ constructed in Lemma* 8.1; *then*

$$\left|\bigcup_{x \in \Sigma} B_r(x)\right| \geq C(\sup u)^n.$$

PROOF: We may extract a countable subcovering of Σ, B_j, with finite overlapping. Then the volume of the image of B_3 by the gradient map of Γ coincides with the volume of the gradient map restricted to Σ and thus with the gradient map restricted to the intermediate set $\bigcup B_j$.

In each B_j, Γ has quadratic growth and therefore

$$|\nabla\Gamma(B_j)| \leq C|B_j|.$$

Thus
$$(\sup u)^n = (\sup \Gamma)^n \le C |\nabla\Gamma(B_3)| \le C \sum_j |B_j|.$$
□

A more precise theorem that reproduces the classical ABP in the limit is the following:

THEOREM 8.7 *Let u and Γ be functions as in Lemma* 8.1. *There is a finite family of* (*open*) *cubes Q_j ($j = 1, \dots, m$) with diameters d_j such that the following hold* (*see Figure* 8.2):

(i) *Any two cubes Q_i and Q_j in the family do not intersect.*
(ii) $\{u = \Gamma\} \subset \bigcup_{j=1}^m \bar{Q}_j$.
(iii) $\{u = \Gamma\} \cap \bar{Q}_j \ne \varnothing$ *for any Q_j.*
(iv) $d_j \le \rho_0 2^{-1/(2-\sigma)}$, *where* $\rho_0 = 1/(8\sqrt{n})$.
(v) $|\nabla\Gamma(\bar{Q}_j)| \le C(\max_{\bar{Q}_j} f)^n |Q_j|$.
(vi) $|\{y \in 8\sqrt{n} Q_j : u(y) > \Gamma(y) - C(\max_{\bar{Q}_j} f) d_j^2\}| \ge \mu |Q_j|$.

The constants $C > 0$ and $\mu > 0$ depend on n, Λ, and λ (but not on σ).

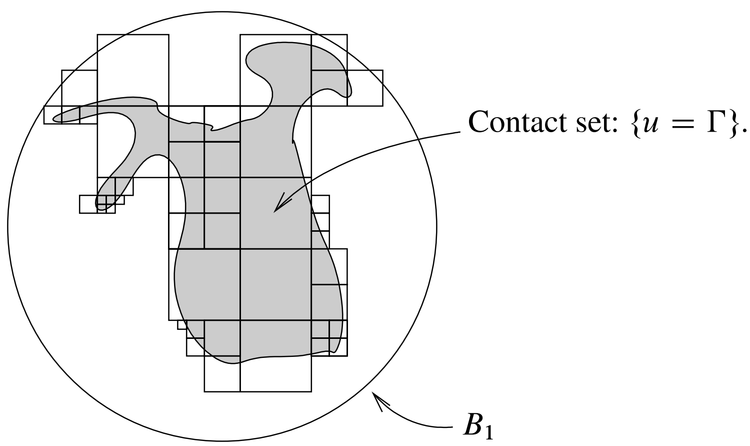

FIGURE 8.2. The family of cubes covering $\{u = \Gamma\}$.

PROOF: In order to obtain such a family we start by covering B_1 with a tiling of cubes of diameter $\rho_0 2^{-1/(2-\sigma)}$. We discard all those that do not intersect $\{u = \Gamma\}$. Whenever a cube does not satisfy (v) and (vi), we split it into 2^n cubes of half-diameter and discard those whose closure does not intersect $\{u = \Gamma\}$. The problem is to prove that eventually all cubes satisfy (v) and (vi), and this process finishes after a finite number of steps.

Let us assume the process does not finish in a finite number of steps. We assume it produces an infinite sequence of nested cubes. The intersection of their closures will be a point x_0. Since all of them intersect the contact set $\{u = \Gamma\}$, which is a closed set, then $u(x_0) = \Gamma(x_0)$. We will now find a contradiction by showing that eventually one of these cubes containing x_0 will not split.

Given $\varepsilon_0 > 0$, by Corollary 8.5, there is a radius r with $0 < r < \rho_0 2^{-1/(2-\sigma)}$ such that

$$\frac{|\{y \in B_r(x_0) \setminus B_{r/2}(x_0) : u(y) < u(x_0) + (y - x_0) \cdot \nabla\Gamma(x_0) - Cf(x_0)r^2\}|}{|B_r(x_0) \setminus B_{r/2}(x_0)|} \le \varepsilon_0. \tag{8.5}$$

and

$$|\nabla\Gamma(B_{r/4}(x_0))| \le Cf(x_0)^n |B_{r/4}(x_0)|. \tag{8.6}$$

There is a cube Q_j, with $x_0 \in \bar{Q}_j$, with diameter d_j, such that $\frac{r}{4} \le d_j < \frac{r}{2}$. Therefore (see Figure 8.3)

$$B_{r/2}(x_0) \supset \bar{Q}_j, \qquad B_r(x_0) \subset 8\sqrt{n}Q_j.$$

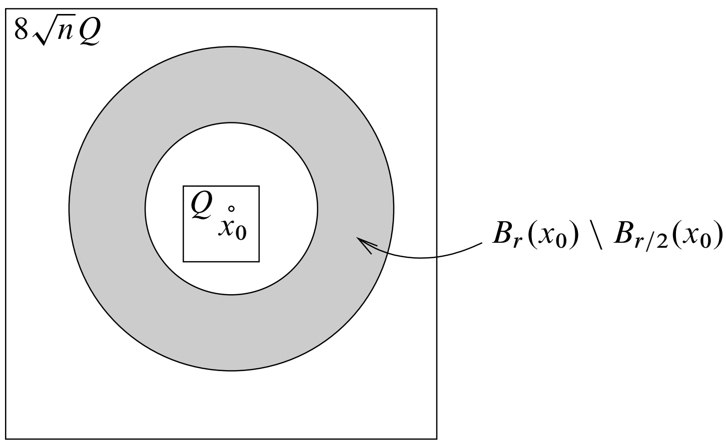

FIGURE 8.3. The largest cube in the family containing x_0 and contained in $B_{r/2}$.

Recall that in B_2, $\Gamma(y) \le u(x_0) + (y - x_0) \cdot \nabla\Gamma(x_0)$ simply because Γ is concave and $\Gamma(x_0) = u(x_0)$. Using (8.5) and that d_j and r are comparable, we get

$$\begin{aligned}
&\Big|\Big\{y \in 8\sqrt{n}Q_j : u(y) \ge \Gamma(y) - C(\max_{\bar{Q}_j} f)d_j^2\Big\}\Big| \\
&\ge |\{y \in 8\sqrt{n}Q_j : u(y) \ge u(x_0) + (y - x_0) \cdot \nabla\Gamma(x_0) - Cf(x_0)r^2\}| \\
&\ge (1 - \varepsilon_0)|B_r(x_0) \setminus B_{r/2}(x_0)| \ge \mu|Q_j|.
\end{aligned}$$

Thus (vi) follows. Moreover, since $\bar{Q}_j \subset B_r$, (v) also holds for Q_j. Therefore Q_j would not be split and the process must stop. $\square$

Remark 8.8. Note that the upper bound for the diameters $\rho_0 2^{-1/(2-\sigma)}$ becomes very small when σ is close to 2. If we add $\sum |\nabla\Gamma(Q_j)|$ and let $\sigma \to 2$, we obtain the classical Aleksandrov estimate as the limit of the Riemann sums. For each $\sigma > 0$ we have

$$|\nabla\Gamma(\{u = \Gamma\})| \le \sum_j C(\max_{\bar{Q}_j} f^+)^n |Q_j|.$$

As $\sigma \to 2$, the cube covering of $\{u = \Gamma\}$ becomes thinner and the above becomes the integral

$$|\nabla\Gamma(\{u = \Gamma\})| \le C \int_{\{u=\Gamma\}} f^+(x)^n \, dx.$$

9 A Special Function

Following the proof of the Harnack inequality in [6] or [7] for the second-order case, we need to show that under the hypotheses of Lemma 8.1, u is nonnegative not just in a positive portion of B_1 but more precisely in a positive portion of any *middle-size* cube, say of diameter $\frac{1}{100}$, centered in $B_{1/2}$. For that purpose, in this section we construct a special function that is a subsolution of a minimal equation outside a small ball and it is strictly positive in a larger ball. The importance of this function is that by adding it to u we will force the contact set with Γ to stay inside one of the *intermediate* cubes.

LEMMA 9.1 *There is a $p > 0$ and $\sigma_0 \in (0, 2)$ such that the function*

$$f(x) = \min(2^p, |x|^{-p})$$

is a subsolution to

$$\mathrm{M}^- f(x) \ge 0 \tag{9.1}$$

for every $\sigma_0 < \sigma < 2$ and $|x| > 1$.

PROOF: It is enough to show (9.1) for $x = e_1 = (1, 0, \dots, 0)$. For every other x such that $|x| = 1$, the relation follows by rotation. If $|x| > 1$, we can consider the function $\tilde{f}(y) = |x|^p f(|x|y) \ge f(y)$; thus $\mathrm{M}^- f(x) = C\mathrm{M}^- \tilde{f}(x/|x|) \ge C\mathrm{M}^- f(x/|x|) > 0$.

Let $x = e_1 = (1, 0, \dots, 0)$. We use the following elementary relations that hold for any $a > b > 0$ and $q > 0$:

$$(a+b)^{-q} \ge a^{-q}\left(1 - q\,\frac{b}{a}\right), \tag{9.2}$$

$$(a+b)^{-q} + (a-b)^{-q} \ge 2a^{-q} + q(q+1)b^2 a^{-q-2}; \tag{9.3}$$

then for $|y| < \frac{1}{2}$,

$$\begin{aligned}\delta &= |x+y|^{-p} + |x-y|^{-p} - 2|x|^{-p} \\ &= (1+|y|^2+2y_1)^{-p/2} + (1+|y|^2-2y_1)^{-p/2} - 2 \\ &\geq 2(1+|y|^2)^{-p/2} + p(p+2)y_1^2(1+|y|^2)^{-p/2-2} - 2 \\ &\geq p\left(-|y|^2 + (p+2)y_1^2 - \frac{1}{2}(p+2)(p+4)y_1^2|y|^2\right).\end{aligned}$$

We choose p large such that

$$(p+2)\lambda \int_{\partial B_1} y_1^2 \, \mathrm{d}\sigma(y) - \Lambda|\partial B_1| = \delta_0 > 0. \tag{9.4}$$

We use the above relation to bound the part of the integral in the definition of M^- for which y stays in a small ball B_r (with $r < \frac{1}{2}$). We estimate $\mathrm{M}^- f(e_1)$:

$$\begin{aligned}\mathrm{M}^- f(e_1) &= (2-\sigma)\int_{B_r} \frac{\lambda\delta^+ - \Lambda\delta^-}{|y|^{n+\sigma}} \, \mathrm{d}y + (2-\sigma)\int_{\mathbb{R}^n \setminus B_r} \frac{\lambda\delta^+ - \Lambda\delta^-}{|y|^{n+\sigma}} \, \mathrm{d}y \\ &\geq (2-\sigma)C\int_0^r \frac{\lambda p\delta_0 s^2 - \frac{1}{2}p(p+2)(p+4)C\Lambda s^4}{s^{1+\sigma}} \, \mathrm{d}s \\ &\quad - (2-\sigma)\int_{\mathbb{R}^n \setminus B_r} \Lambda \frac{2^p}{|y|^{n+\sigma}} \, \mathrm{d}y \\ &\geq c r^{2-\sigma} p\delta_0 - p(p+2)(p+4)C \frac{2-\sigma}{4-\sigma} r^{4-\sigma} - \frac{2-\sigma}{\sigma} C 2^{p+1} r^{-\sigma}\end{aligned}$$

where we used (9.4) to bound the first integral and the fact that $0 \leq f(x) \leq 2^p$ to bound the second integral. Now we choose (and fix) $r \in (0, \frac{1}{2})$, and then take σ_0 close enough to 2 so that if $2 > \sigma > \sigma_0$, the factor $(2-\sigma)$ makes the second and third terms small enough so that we get

$$\mathrm{M}^- f(e_1) \geq \frac{c r^{2-\sigma} p\delta_0}{2} > 0,$$

which finishes the proof. □

COROLLARY 9.2 *Given any $\sigma_0 \in (0,2)$ and $r > 0$ there is a $p > 0$ and δ such that the function*

$$f(x) = \min(\delta^{-p}, |x|^{-p})$$

is a subsolution to

$$\mathrm{M}^- f(x) \geq 0 \tag{9.5}$$

for every $\sigma_0 < \sigma < 2$ and $|x| > r$.

PROOF: Without loss of generality, we prove it the corollary for $r = 1$. The general case follows by rescaling.

The only difference from Lemma 9.1 is that now we are given the value of σ_0 beforehand. Let σ_1 and p_0 be the σ_0 and p of Lemma 9.1. So we know that for $\sigma > \sigma_1$, the result of the Corollary holds if $\delta = \frac{1}{2}$ and $p = p_0$. If we take $\delta < \frac{1}{2}$, we are only making the function larger away from x, so the result will still hold for $\sigma > \sigma_1$. Now we will pick δ smaller so that the result also holds for $\sigma_0 < \sigma \leq \sigma_1$.

The key is that if $p \geq n$, $|x|^{-p}$ is not integrable around the origin. So we take $p = \max(p_0, n)$. Now, let $x = e_1$ as in the proof of Lemma 9.1. Assume $\sigma_0 < \sigma \leq \sigma_1$. We write

$$\begin{aligned} \mathrm{M}^- f(e_1) &= (2-\sigma) \int_{\mathbb{R}^n} \frac{\lambda \delta^+}{|y|^{n+\sigma}}\, \mathrm{d}y - (2-\sigma) \int_{\mathbb{R}^n} \frac{\Lambda \delta^-}{|y|^{n+\sigma}}\, \mathrm{d}y \\ &=: I_1 + I_2 \end{aligned}$$

where I_1 and I_2 represent the two terms in the right-hand side above. Since $\sigma > \sigma_0$, $f \in C^2(x)$, and f is bounded below, we have $I_2 \geq -C$ for some constant C depending on σ_0, λ, Λ, and dimension. On the other hand, since $\sigma \leq \sigma_1$ and $(|x+y|^{-p} + |x-y|^{-p} - 2|x|^{-p})^+$ is not integrable, if we choose δ small enough, we can make I_1 be as large as we wish. In particular, we can choose δ such that $I_1 > C > -I_2$; thus $\mathrm{M}^- f(e_1) > 0$. □

COROLLARY 9.3 *Given any $\sigma_0 \in (0,2)$, there is a function Φ such that*

(i) *Φ is continuous in $\mathbb{R}^n$,*
(ii) *$\Phi(x) = 0$ for x outside $B_{2\sqrt{n}}$,*
(iii) *$\Phi(x) > 2$ for $x \in Q_3$, and*
(iv) *$\mathrm{M}^-\Phi > -\psi(x)$ in $\mathbb{R}^n$ for some positive function $\psi(x)$ supported in $\bar{B}_{1/4}$*

for every $\sigma > \sigma_0$.

PROOF: Let p and δ be as in Corollary 9.2 with $r = \frac{1}{4}$. We consider

$$\Phi = c \begin{cases} 0 & \text{in } \mathbb{R}^n \setminus B_{2\sqrt{n}}, \\ |x|^{-p} - (2\sqrt{n})^{-p} & \text{in } B_{2\sqrt{n}} \setminus B_\delta, \\ q & \text{in } B_\delta, \end{cases}$$

where q is a quadratic paraboloid chosen so that Φ is $C^{1,1}$ across ∂B_δ. We choose the constant c so that $\Phi(x) > 2$ for $x \in Q_3$ (recall that $Q_3 \subset B_{3\sqrt{n}/2} \subset B_{2\sqrt{n}}$). Since $\Phi \in C^{1,1}(B_{2\sqrt{n}})$, $\mathrm{M}^-\Phi$ is continuous in $B_{2\sqrt{n}}$, and from Corollary 9.2, $\mathrm{M}^-\Phi \geq 0$ outside $B_{1/4}$. □

10 Point Estimates

The main ingredient in the proof of the Harnack inequality, as shown in [7], is a lemma that links a pointwise estimate with an estimate in measure. With the estimates developed in Sections 8 and 9, we are ready to obtain such estimates. The corresponding lemma in our context is the following:

LEMMA 10.1 *Let $\sigma > \sigma_0 > 0$. There exist constants $\varepsilon_0 > 0$, $0 < \mu < 1$, and $M > 1$ (depending only on σ_0, λ, Λ, and dimension) such that if*

(i) $u \geq 0$ *in* $\mathbb{R}^n$,
(ii) $\inf_{Q_3} u \leq 1$, *and*
(iii) $\mathrm{M}^- u \leq \varepsilon_0$ *in* $Q_{4\sqrt{n}}$,

then $|\{u \leq M\} \cap Q_1| > \mu$.

By $Q_r(x)$ we mean the open cube $\{y : |y_j - x_j| \leq \frac{r}{2}$ for every $j\}$ and $Q_r := Q_r(0)$. We will also use the following notation for dilations: if $Q = Q_r(x)$, then $\lambda Q := Q_{\lambda r}(x)$.

If we assume $\sigma \leq \sigma_1 < 2$, there is a simpler proof of Lemma 10.1 using the ideas from [10]. The result here is more involved because we want an estimate that remains uniform as $\sigma \to 2$.

PROOF: Consider $v := \Phi - u$, where Φ is the special function constructed in Corollary 9.3. We want to apply Theorem 8.7 (rescaled) to v. Note that $\mathrm{M}^+ v \geq \mathrm{M}^- \Phi - \mathrm{M}^- u \geq -\psi - \varepsilon_0$. Let Γ be the concave envelope of v in $B_{6\sqrt{n}}$.

Let Q_j be the family of cubes given by Theorem 8.7. We have

$$\begin{aligned}
\max v &\leq C|\nabla\Gamma(B_{2\sqrt{n}})|^{1/n} \leq C\Big(\sum_j |\nabla\Gamma(\bar{Q}_j)|\Big)^{1/n} \\
&\leq \Big(C\sum_j (\max_{Q_j}(\psi + \varepsilon_0)^+)^n |Q_j|\Big)^{1/n} \\
&\leq C\varepsilon_0 + C\Big(\sum_j (\max_{Q_j} \psi^+)^n |Q_j|\Big)^{1/n}.
\end{aligned}$$

However, since $\inf_{Q_3} u \leq 1$ and $\Phi \geq 2$ in Q_3, then $\max v \geq 1$ and we have

$$1 \leq C\varepsilon_0 + C\Big(\sum_j (\max_{Q_j} \psi^+)^n |Q_j|\Big)^{1/n}.$$

If we choose ε_0 small enough, this will imply

$$\frac{1}{2} \leq C\Big(\sum_j (\max_{Q_j} \psi^+)^n |Q_j|\Big)^{1/n}.$$

Recall that ψ is supported in $\bar{B}_{1/4}$ and is bounded; thus:

$$\frac{1}{2} \leq C\Big(\sum_{Q_j \cap B_{1/4} \neq \varnothing} |Q_j|\Big)^{1/n},$$

which provides a bound below for the sum of the volumes of the cubes Q_j that intersect $B_{1/4}$,

$$\sum_{Q_j \cap B_{1/4} \neq \varnothing} |Q_j| \geq c. \tag{10.1}$$

The diameters of all cubes Q_j are bounded by $\rho_0 2^{-1/(2-\sigma)}$, which is always smaller than $\rho_0 = 1/(8\sqrt{n})$. Therefore, every time Q_j intersects $B_{1/4}$, the cube $4\sqrt{n}\, Q_j$ will be contained in $B_{1/2}$.

Let $M_0 := \max_{B_{1/2}} \Phi$. By Theorem 8.7, we have

$$|\{x \in 4\sqrt{n} Q_j : v(x) \geq \Gamma(x) - C d_j^2\}| \geq c|Q_j| \tag{10.2}$$

and $C d_j^2 < C\rho_0^2$.

Let us consider the cubes $4\sqrt{n}\, Q_j$ for every cube Q_j that intersects $B_{1/4}$. It provides an open cover of the union of the corresponding cubes $\overline{Q}_j$, and it is contained in $B_{1/2}$. We take a subcover with finite overlapping that also covers the union of the original $\overline{Q}_j$. Combining (10.1) with (10.2), we obtain

$$|\{x \in B_{1/2} : v(x) \geq \Gamma(x) - C\rho_0^2\}| \geq c.$$

Then

$$|\{x \in B_{1/2} : u(x) \leq M_0 + C\rho_0^2\}| \geq c$$

Let $M = M_0 + C\rho_0^2$. Since $B_{1/2} \subset Q_1$, we have

$$|\{x \in Q_1 : u(x) \leq M\}| \geq c,$$

which finishes the proof. □

Lemma 10.1 is the key to the proof of the Harnack inequality. The following lemma is a consequence of Lemma 10.1 as it is shown in lemma 4.6 in [7]. We have intentionally written Lemma 10.1 and the following one identically to their corresponding versions in [7].

LEMMA 10.2 *Let u be as in Lemma* 10.1. *Then*

$$|\{u > M^k\} \cap Q_1| \leq (1-\mu)^k$$

for $k = 1, 2, \ldots$, where M and μ are as in Lemma 10.1. *As a consequence, we have that*

$$|\{u \geq t\} \cap Q_1| \leq d t^{-\varepsilon} \quad \forall t > 0$$

where d and ε are positive universal constants.

By a standard covering argument we obtain the following theorem:

THEOREM 10.3 *Let $u \geq 0$ in $\mathbb{R}^n$, $u(0) \leq 1$, and $\mathrm{M}^- u \leq \varepsilon_0$ in B_2 (supersolution). Assume $\sigma \geq \sigma_0$ for some $\sigma_0 > 0$. Then*

$$|\{u > t\} \cap B_1| \leq C t^{-\varepsilon} \quad \text{for every } t > 0$$

where the constant C and ε depend on λ, Λ, n, and σ_0.

Scaling the above theorem, we obtain the following version:

THEOREM 10.4 *Let $u \geq 0$ in $\mathbb{R}^n$ and $\mathrm{M}^- u \leq C_0$ in B_{2r} (supersolution). Assume $\sigma \geq \sigma_0$ for some $\sigma_0 > 0$. Then*

$$|\{u > t\} \cap B_r| \leq C r^n (u(0) + C_0 r^\sigma)^\varepsilon t^{-\varepsilon} \quad \textit{for every } t$$

where the constant C and ε depend on λ, Λ, n, and σ_0.

For second-order equations, Theorems 10.3 and 10.4 are referred in the literature as u being in L^ε (see [7]).

11 Harnack Inequality

The Harnack inequality is a very important tool in analysis. In this section we obtain a version for integro-differential equations. Our estimate depends only on a lower bound $\sigma \geq \sigma_0 > 0$, but it remains uniform as $\sigma \to 2$. In that respect, we can consider this estimate as a generalization of the Krylov-Safonov Harnack inequality.

This section is not needed for the rest of the paper because we will prove our regularity results using Theorem 10.4 only. A reader interested only in the regularity results can skip this section.

THEOREM 11.1 *Let $u \geq 0$ in $\mathbb{R}^n$, $\mathrm{M}^- u \leq C_0$, and $\mathrm{M}^+ u \geq -C_0$ in B_2. Assume $\sigma \geq \sigma_0$ for some $\sigma_0 > 0$. Then $u(x) \leq C(u(0) + C_0)$ for every $x \in B_{1/2}$.*

PROOF: Dividing by $u(0) + C_0$, it is enough to consider $u(0) \leq 1$ and $C_0 = 1$. Let $\varepsilon > 0$ be the one from Theorem 10.4. Let $\gamma = \frac{n}{\varepsilon}$. Let us consider the minimum value of t such that

$$u(x) \leq h_t(x) := t(1 - |x|)^{-\gamma} \quad \text{for every } x \in B_1.$$

There must be an $x_0 \in B_1$ such that $u(x_0) = h_t(x_0)$; otherwise we could make t smaller. Let $d = (1 - |x_0|)$ be the distance from x_0 to ∂B_1.

For $r = \frac{d}{2}$, we want to estimate the portion of the ball $B_r(x_0)$ covered by $\{u < u(x_0)/2\}$ and by $\{u > u(x_0)/2\}$. We will show that t cannot be too large. In this way we obtain the result of the theorem, since the upper bound $t < C$ implies that $u(x) < C(1 - |x|)^{-\gamma}$.

Let us first consider $A := \{u > u(x_0)/2\}$. By the L^ε estimate (Theorem 10.3) we have

$$|A \cap B_1| \leq C \left| \frac{2}{u(x_0)} \right|^\varepsilon \leq C t^{-\varepsilon} d^n,$$

whereas $|B_r| = C d^n$, so if t is large, A can cover only a small portion of $B_r(x_0)$ at most,

$$\left| \left\{ u > \frac{u(x_0)}{2} \right\} \cap B_r(x_0) \right| \leq C t^{-\varepsilon} |B_r|. \tag{11.1}$$

In order to get a contradiction, we will show that $|\{u < u(x_0)/2\} \cap B_r(x_0)| \leq (1-\delta)B_r$ for a positive constant δ independent of t. We estimate $|\{u < u(x_0)/2\} \cap B_{\theta r}(x_0)|$ for $\theta > 0$ small. For every $x \in B_{\theta r}(x_0)$ we have

$$u(x) \leq h_t(x) \leq t\left(\frac{d - \theta d}{2}\right)^{-\gamma} \leq u(x_0)\left(\frac{1-\theta}{2}\right)^{-\gamma},$$

with $(1 - \theta/2)^{-\gamma}$ close to 1.

Let us consider

$$v(x) = \left(1 - \frac{\theta}{2}\right)^{-\gamma} u(x_0) - u(x)$$

so that $v \geq 0$ in $B_{\theta r}$, and also $\mathrm{M}^- v \leq 1$ since $\mathrm{M}^+ u \geq -1$. We would want to apply Theorem 10.4 to v. The only problem is that v is not positive in the whole domain but only on $B_{\theta r}$. In order to apply the theorem we have to consider $w = v^+$ instead, and estimate the change in the right-hand side due to the truncation error.

We want to find an upper bound for $\mathrm{M}^- w = \mathrm{M}^- v^+$ instead of $\mathrm{M}^- v$. We know that

$$\mathrm{M}^- v(x) = (2-\sigma) \int_{\mathbb{R}^n} \frac{\lambda \delta(v, x, y)^+ - \Lambda \delta(v, x, y)^-}{|y|^{n+\sigma}} \, \mathrm{d}y \leq 1.$$

Therefore, if $x \in B_{\theta r/2}(x_0)$,

$$\begin{aligned} \mathrm{M}^- w &= (2-\sigma) \int_{\mathbb{R}^n} \frac{\lambda \delta(w, x, y)^+ - \Lambda \delta(w, x, y)^-}{|y|^{n+\sigma}} \, \mathrm{d}y \\ &\leq 1 + 2(2-\sigma) \int_{\mathbb{R}^n \cap \{v(x+y) < 0\}} -\Lambda \frac{v(x+y)}{|y|^{n+\sigma}} \, \mathrm{d}y \qquad (11.2) \\ &\leq 1 + 2(2-\sigma) \int_{\mathbb{R}^n \setminus B_{\theta r/2}(x_0 - x)} \Lambda \frac{(u(x+y) - (1-\frac{\theta}{2})^{-\gamma} u(x_0))^+}{|y|^{n+\sigma}} \, \mathrm{d}y. \end{aligned}$$

Notice that the restriction $u \geq 0$ does not provide an upper bound for this last expression. We must obtain it in a different way.

Let us consider the largest value $\tau > 0$ such that $u(x) \geq g_\tau := \tau(1 - |4x|^2)$. There must be a point $x_1 \in B_{1/4}$ such that $u(x_1) = \tau(1 - |4x_1|^2)$. The value of τ cannot be larger than 1 since $u(0) \leq 1$. Thus we have the upper bound

$$\begin{aligned} &(2-\sigma) \int_{\mathbb{R}^n} \frac{\delta(u, x_1, y)^-}{|y|^{n+\sigma}} \, \mathrm{d}y \qquad (11.3) \\ &\leq (2-\sigma) \int_{\mathbb{R}^n} \frac{\delta(g_\tau, x_1, y)^-}{|y|^{n+\sigma}} \, \mathrm{d}y \leq C \qquad (11.4) \end{aligned}$$

for a constant C that is independent of σ.

Since $\mathrm{M}^- u(x_1) \le 1$, then

$$(2-\sigma) \int_{\mathbb{R}^n} \frac{\delta(u, x_1, y)^+}{|y|^{n+\sigma}} \, \mathrm{d}y \le C.$$

In particular, since $u(x_1) \le 1$ and $u(x_1 - y) \ge 0$,

$$(2-\sigma) \int_{\mathbb{R}^n} \frac{(u(x_1 + y) - 2)^+}{|y|^{n+\sigma}} \, \mathrm{d}y \le C.$$

We can use the inequality above to estimate (11.2). We can assume $u(x_0) > 2$, since otherwise t would not be large:

$$\begin{aligned}
(2-\sigma) & \int_{\mathbb{R}^n \setminus B_{\theta r}(x_0 - x)} \Lambda \frac{(u(x+y) - (1-\theta/2)^{-\gamma} u(x_0))^+}{|y|^{n+\sigma}} \, \mathrm{d}y \\
&\le (2-\sigma) \int_{\mathbb{R}^n \setminus B_{\theta r/2}(x_0 - x)} \Lambda \frac{(u(x_1 + y + x - x_1) - (1-\theta/2)^{-\gamma} u(x_0))^+}{|y + x - x_1|^{n+\sigma}} \\
&\qquad\qquad \cdot \frac{|y + x - x_1|^{n+\sigma}}{|y|^{n+\sigma}} \, \mathrm{d}y \\
&\le C(\theta r)^{-n-\sigma}.
\end{aligned}$$

So, finally, we obtain

$$\mathrm{M}^- w \le C(\theta r)^{-n-\sigma} \quad \text{in } B_{\theta r/2}(x_0).$$

Now we can apply Theorem 10.4 to w in $B_{\theta r/2}(x_0)$. Recalling that $w(x_0) = ((1-\theta/2)^{-\gamma} - 1) u(x_0)$, we have

$$\begin{aligned}
&\left| \left\{ u < \frac{u(x_0)}{2} \right\} \cap B_{\frac{\theta r}{4}}(x_0) \right| \\
&\quad = \left| \{ w > u(x_0) \left(\left(1 - \frac{\theta}{2}\right)^{-\gamma} - \frac{1}{2} \right) \} \cap B_{\theta r/4}(x_0) \right| \\
&\quad \le C(\theta r)^n \left(\left(\left(1 - \frac{\theta}{2}\right)^{-\gamma} - 1 \right) u(x_0) + C(\theta r)^{-n-\sigma} (r\theta)^\sigma \right)^\varepsilon \\
&\qquad \cdot \left(u(x_0) \left(\left(1 - \frac{\theta}{2}\right)^{-\gamma} - \frac{1}{2} \right) \right)^{-\varepsilon} \\
&\quad \le C(\theta r)^n \left(\left(\left(1 - \frac{\theta}{2}\right)^{-\gamma} - 1 \right)^\varepsilon + \theta^{-n\varepsilon} t^{-\varepsilon} \right).
\end{aligned}$$

Now let us choose $\theta > 0$ so that the first term is small:

$$C(\theta r)^n \left(\left(1 - \frac{\theta}{2}\right)^{-\gamma} - 1 \right)^\varepsilon \le \frac{1}{4} |B_{\theta r/4}(x_0)|.$$

Notice that the choice of θ is independent of t. For this fixed value of θ we observe that if t is large enough, we will also have

$$C(\theta r)^n \theta^{-n\varepsilon} t^{-\varepsilon} \leq \frac{1}{4} \left| B_{\theta r/2} \right|$$

and therefore

$$\left| \left\{ u < \frac{u(x_0)}{2} \right\} \cap B_{\theta r/4}(x_0) \right| \leq \frac{1}{2} |B_{\theta r/4}(x_0)|,$$

which implies that for t large

$$\left| \left\{ u > \frac{u(x_0)}{2} \right\} \cap B_{\theta r/4}(x_0) \right| \geq c \, |B_r| \, .$$

But this contradicts (11.1). Therefore t cannot be large and we finish the proof. □

12 Hölder Estimates for Equations with "Bounded Measurable Coefficients"

The purpose of this section is to prove the following Hölder regularity result.

THEOREM 12.1 *Let $\sigma > \sigma_0$ for some $\sigma_0 > 0$. Let u be a bounded function in $\mathbb{R}^n$ such that*

$$\begin{aligned} \mathrm{M}^+ u &\geq -C_0 \quad \text{in } B_1, \\ \mathrm{M}^- u &\leq C_0 \quad \text{in } B_1; \end{aligned}$$

then there is an $\alpha > 0$ (depending only on λ, Λ, n, and σ_0) such that $u \in C^\alpha(B_{1/2})$ and

$$u_{C^\alpha(B_{1/2})} \leq C(\sup_{\mathbb{R}^n} |u| + C_0)$$

for some constant $C > 0$.

Even though this result could be obtained as a consequence of the Harnack inequality, we will prove it using only Theorem 10.4. We do it in this way because it looks potentially simpler to generalize since we proved the Harnack inequality (Theorem 11.1) using Theorem 10.4.

Theorem 12.1 follows from the following lemma by a simple scaling.

LEMMA 12.2 *Let $\sigma > \sigma_0$ for some $\sigma_0 > 0$. Let u be a function such that*

$$-\frac{1}{2} \leq u \leq \frac{1}{2} \quad \text{in } \mathbb{R}^n, \qquad \mathrm{M}^+ u \geq -\varepsilon_0 \quad \text{in } B_1, \qquad \mathrm{M}^- u \leq \varepsilon_0 \quad \text{in } B_1;$$

then there is an $\alpha > 0$ (depending only on λ, Λ, n, and σ_0) such that $u \in C^\alpha$ at the origin. More precisely,

$$|u(x) - u(0)| \leq C|x|^\alpha$$

for some constant C.

PROOF: We will show that there exists sequences m_k and M_k such that $m_k \le u \le M_k$ in $B_{4^{-k}}$ and

$$M_k - m_k = 4^{-\alpha k} \tag{12.1}$$

so that the theorem holds with $C = 4^\alpha$.

For $k = 0$ we choose $m_0 = -\frac{1}{2}$ and $M_0 = \frac{1}{2}$. By assumption we have $m_0 \le u \le M_0$ in the whole space $\mathbb{R}^n$. We want to construct the sequences M_k and m_k by induction.

Assume we have the sequences up to m_k and M_k. We want to show we can continue the sequences by finding m_{k+1} and M_{k+1}.

In the ball $B_{4^{-k-1}}$, either $u \ge (M_k + m_k)/2$ in at least half of the points (in measure), or $u \le (M_k + m_k)/2$ in at least half of the points. Let us say that

$$\left|\left\{u \ge \frac{M_k + m_k}{2}\right\} \cap B_{4^{-k-1}}\right| \ge \frac{|B_{4^{-k-1}}|}{2}.$$

Consider

$$v(x) := \frac{u(4^{-k}x) - m_k}{(M_k - m_k)/2}$$

so that $v(x) \ge 0$ in B_1 and $|\{v \ge 1\} \cap B_{1/4}| \ge |B_{1/4}|/2$. Moreover, since $\mathrm{M}^- u \le \varepsilon_0$ in B_1,

$$\mathrm{M}^- v \le \frac{4^{-k\sigma}\varepsilon_0}{(M_k - m_k)/2} = 2\varepsilon_0 4^{k(\sigma-\alpha)} \le 2\varepsilon_0 \quad \text{in } B_{4^k}$$

if α is chosen less than σ.

From the inductive hypothesis, for any $j \ge 1$, we have

$$\begin{aligned} v \ge \frac{(m_{k-j} - m_k)}{(M_k - m_k)/2} &\ge \frac{(m_{k-j} - M_{k-j} + M_k - m_k)}{(M_k - m_k)/2} \\ &\ge -2 \cdot 4^{\alpha j} + 2 \ge 2(1 - 4^{\alpha j}) \quad \text{in } B_{2^j}. \end{aligned}$$

Therefore $v(x) \ge -2(|4x|^\alpha - 1)$ outside B_1. If we let $w(x) = \max(v, 0)$, then $\mathrm{M}^- w \le \mathrm{M}^- v + 2\varepsilon_0$ in $B_{3/4}$ if α is small enough. We still have $|\{w \ge 1\} \cap B_{1/4}| \ge |B_{1/4}|/2$. Given any point $x \in B_{1/4}$, we can apply Theorem 10.4 in $B_1(x)$ to obtain

$$C(w(x) + 2\varepsilon_0)^\varepsilon \ge |\{w > 1\} \cap B_{1/2}(x)| \ge \frac{1}{2}|B_{1/4}|.$$

If we have chosen ε_0 small, this implies that $w \ge \theta$ in $B_{1/4}$ for some $\theta > 0$. Thus if we let $M_{k+1} = M_k$ and $m_{k+1} = m_k + \theta(M_k - m_k)/2$, we have $m_{k+1} \le u \le M_{k+1}$ in $B_{2^{k+1}}$. Moreover, $M_{k+1} - m_{k+1} = (1 - \theta/2)4^{-\alpha k}$. So we must choose α and θ small and so that $(1 - \frac{\theta}{2}) = 4^{-\alpha}$, and we obtain $M_{k+1} - m_{k+1} = 4^{-\alpha(k+1)}$.

On the other hand, if $|\{u \leq (M_k + m_k)/2\} \cap B_{4^{-k}}| \geq |B_{4^{-k}}|/2$, we define

$$v(x) := \frac{M_k - u(4^{-k}x)}{(M_k - m_k)/2}$$

and continue in the same way using that $\mathrm{M}^+ u \geq -\varepsilon_0$. □

13 $C^{1+\alpha}$ Estimates

In this section we prove an interior $C^{1,\alpha}$ regularity result for the solutions to a general class of fully nonlinear integro-differential equations. The idea of the proof is to use the Hölder estimates of Theorem 12.1 to incremental quotients of the solution. There is a difficulty in that we have no uniform bound in L^∞ for the incremental quotients outside of the domain. This becomes an issue since we are dealing with nonlocal equations. The way we solve it is by assuming some extra regularity of the family of integral operators $\mathcal{L}$. The extra assumption, compared to the assumptions for Hölder regularity (3.3), is a modulus of continuity of K in measure, so as to make sure that faraway oscillations tend to cancel out.

Given $\rho_0 > 0$, we define the class $\mathcal{L}_1$ by the operators L with kernels K such that

$$(2-\sigma)\frac{\lambda}{|y|^{n+\sigma}} \leq K(y) \leq (2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}}, \tag{13.1}$$

$$\int_{\mathbb{R}^n \setminus B_{\rho_0}} \frac{|K(y) - K(y-h)|}{|h|}\, \mathrm{d}y \leq C \quad \text{every time } |h| < \frac{\rho_0}{2} \tag{13.2}$$

A simple condition for (13.2) to hold would be that $|\nabla K(y)| \leq \Lambda/|y|^{1+n+\sigma}$.

In the following theorem we give interior $C^{1,\alpha}$ estimates for fully nonlinear elliptic equations.

THEOREM 13.1 *Assume $\sigma > \sigma_0$. There is a $\rho_0 > 0$ (depending on λ, Λ, σ_0, and n) so that if I is a nonlocal elliptic operator with respect to $\mathcal{L}_1$ in the sense of Definition* 3.1 *and u is a bounded function such that $Iu = 0$ in B_1, then there is a universal $\alpha > 0$ (depending only on λ, Λ, n, and σ_0) such that $u \in C^{1+\alpha}(B_{1/2})$ and*

$$u_{C^{1+\alpha}(B_{1/2})} \leq C(\sup_{\mathbb{R}^n} |u| + |I0|)$$

for some constant $C > 0$ (where by $I0$ we mean the value we obtain when we apply I to the function that is a constant equal to 0*). The constant C depends on λ, Λ, σ_0, n, and the constant in* (13.2)*.*

PROOF: Because of the assumption (13.1), the class $\mathcal{L}_1$ is included in $\mathcal{L}_0$ given by (3.3). Since $Iu = 0$ in B_1, in particular $\mathrm{M}^+ u \geq Iu - I0 = -I0$ and also $\mathrm{M}^- u \leq I0$ in B_1, and therefore by Theorem 12.1 we have $u \in C^\alpha(B_{1-\delta})$ for any $\delta > 0$ with $\|u\|_{C^\alpha} \leq C(\sup |u| + |I0|)$.

Now we want to improve the obtained regularity iteratively by applying Theorem 12.1 again until we obtain Lipschitz regularity in a finite number of steps.

Assume we have proved that $u \in C^{\beta}(B_r)$ for some $\beta > 0$ and $\frac{1}{2} < r < 1$. We want to apply Theorem 12.1 for the difference quotient

$$w^h = \frac{u(x+h) - u(x)}{|h|^{\beta}}$$

to obtain $u \in C^{\beta+\alpha}(B_{r-\delta})$. By Theorem 5.9, $\mathrm{M}^+_{\mathcal{L}_1} w^h \geq 0$ and $\mathrm{M}^-_{\mathcal{L}_1} w^h \leq 0$ in B_r. In particular, $\mathrm{M}^+ w^h \geq 0$ and $\mathrm{M}^- w^h \leq 0$ in B_r.

The function w^h is uniformly bounded in B_r because $u \in C^{\beta}(B_r)$. Outside B_r the function w^h is not uniformly bounded, so we cannot apply Theorem 12.1 immediately. However, w^h has oscillations that cause cancellations in the integrals because of our assumption (13.2).

Let η be a smooth cutoff function supported in B_r such that $\eta \equiv 1$ in $B_{r-\delta/4}$, where δ is some small positive number that will be determined later. Let us write $w^h = w^h_1 + w^h_2$, where

$$w^h_1 = \frac{\eta u(x+h) - \eta u(x)}{|x|^{\beta}}, \qquad w^h_2 = \frac{(1-\eta)u(x+h) - (1-\eta)u(x)}{|x|^{\beta}}.$$

Let $x \in B_{r/2}$ and $|h| < \frac{\delta}{16}$. In this case $(1-\eta)u(x) = (1-\eta)u(x+h) = 0$ and $w^h(x) = w^h_1(x)$. We have to show that $w^h_1 \in C^{\beta+\alpha}(B_{r-\delta})$.

We have

$$\begin{aligned}
\mathrm{M}^+ w^h_1 &\geq \mathrm{M}^+_{\mathcal{L}_1} w^h_1 = \mathrm{M}^+_{\mathcal{L}_1}(w^h - w^h_2) \geq 0 - \mathrm{M}^+_{\mathcal{L}_1} w^h_2, \\
\mathrm{M}^- w^h_1 &\leq \mathrm{M}^-_{\mathcal{L}_1} w^h_1 = \mathrm{M}^-_{\mathcal{L}_1}(w^h - w^h_2) \leq 0 - \mathrm{M}^-_{\mathcal{L}_1} w^h_2.
\end{aligned}$$

In order to apply Theorem 12.1, we will show that $|\mathrm{M}^+_{\mathcal{L}_1} w^h_2|$ and $|\mathrm{M}^-_{\mathcal{L}_1} w^h_2|$ are bounded in $B_{r-\delta/2}$ by $C \sup|u|$ for some universal constant C. We must show those inequalities for any operator $L \in \mathcal{L}_1$.

Since $(1-\eta)u(x) = (1-\eta)u(x+h) = 0$ and $w^h(x) = w^h_1(x)$, we have the expression

$$Lw^h_2 = \int_{\mathbb{R}^n} \frac{(1-\eta)u(x+y+h) - (1-\eta)u(x+y)}{|h|^{\beta}} K(y)\,\mathrm{d}y,$$

and we notice that both terms $(1-\eta)u(x+y+h) = (1-\eta)u(x+y) = 0$ for $|y| < \frac{\delta}{8}$. We take $\rho_0 = \frac{\delta}{4}$; therefore we can integrate by parts the incremental quotient to obtain

$$\begin{aligned}
|Lw^h_2| &= \left| \int_{\mathbb{R}^n} (1-\eta)u(x+y) \frac{K(y) - K(y-h)}{|h|^{\beta}}\,\mathrm{d}y \right| \\
&\leq \int_{\mathbb{R}^n} |(1-\eta)u(x+y)|\, |h|^{1-\beta} \frac{|K(y) - K(y-h)|}{|h|}\,\mathrm{d}y \quad \text{using (13.2)} \leq
\end{aligned}$$

$$\leq |h|^{1-\beta} \int_{\mathbb{R}^n \setminus B_{\delta/4}} \frac{|K(y) - K(y-h)|}{|h|} \, dy \sup_{\mathbb{R}^n} |u|$$
$$\leq C|h|^{1-\beta}|u| \leq C \sup_{\mathbb{R}^n} |u|.$$

So, we have obtained $M^+ w_1^h \geq -C \sup |u|$ and $M^- w_1^h \leq C \sup |u|$ in $B_{r-\delta/2}$ for $|h| < \frac{\delta}{16}$. We can apply Theorem 12.1 to get that w_1^h (and thus also w^h) is uniformly C^α in $B_{r-\delta}$. By the standard telescopic sum argument [7], this implies that $u \in C^{\alpha+\beta}(B_{r-\delta})$.

Iterating the above argument, we obtain that u is Lipschitz in $[1/\alpha]$ steps. Then, for any unit vector e, we use the same reasoning for the incremental quotients

$$w^h = \frac{u(x+he) - u(x)}{h}$$

to conclude that $u \in C^{1,\alpha}$ in a smaller ball. If we choose the constant δ appropriately, we get $u \in C^{1,\alpha}(B_{1/2})$. □

Remark 13.2. Note that the value of ρ_0 in Theorem 13.1 is not scale invariant. If we want to scale the estimate to apply it to a function u such that $Iu = 0$ in B_r, then we also have to multiply the value of ρ_0 times r.

Remark 13.3. Note that the family $\mathcal{L}$ given by the operators L with the form

$$Lu(x) = \int_{\mathbb{R}^n} \frac{c_n(2-\sigma)}{\det A|A^{-1}z|^{n+\sigma}} \, \delta(u,x,z) dz$$

satisfies conditions (13.1) and (13.2). Thus, from the arguments in Section 6 and Theorem 13.1, we recover the $C^{1,\alpha}$ estimates for fully nonlinear elliptic equations.

14 Truncated Kernels

For applications, it is important to be able to deal with integro-differential operators whose kernels do not satisfy (3.3) in the whole space $\mathbb{R}^n$ but only in a neighborhood of the origin. For example, we want to be able to deal with the operators related to truncated α-stable Lévy processes. In this section we extend our regularity results for this kind of operator.

We consider the following class $\mathcal{L}$: We say that an operator L belongs to $\mathcal{L}$ if its corresponding kernel K has the form

$$K(y) = K_1(y) + K_2(y) \geq 0 \tag{14.1}$$

where

$$(2-\sigma)\frac{\lambda}{|x|^{n+\sigma}} \leq K_1(y) \leq (2-\sigma)\frac{\Lambda}{|x|^{n+\sigma}}$$

and $K_2 \in L^1(\mathbb{R}^n)$ with $\|K_2\|_{L^1} \leq \kappa$.

In this class $\mathcal{L}$ we can consider kernels that are comparable to $|y|^{-n-\sigma}$ near the origin but decay exponentially at infinity or even become zero outside some ball. For example,

$$K(y) = \frac{1}{|y|^{n+\sigma}} e^{-|y|^2} \quad \text{or} \quad K(y) = \frac{a(y)}{|y|^{n+\sigma}} \chi_{B_1}(y) \text{ where } \lambda \leq a(y) \leq \Lambda.$$

This class $\mathcal{L}$ is larger than the class $\mathcal{L}_0$ in (3.3). However, in the following lemma we show that the extremal operators $\mathrm{M}^+_{\mathcal{L}}$ and $\mathrm{M}^-_{\mathcal{L}}$ are controlled by the corresponding extremal operators of $\mathcal{L}_0$, M^+ and M^-, plus the L^∞ norm of u.

LEMMA 14.1 *Let u be a bounded function in $\mathbb{R}^n$ and $C^{1,1}$ at the point x. Then*

$$\mathrm{M}^-_{\mathcal{L}} u(x) \geq \mathrm{M}^- u(x) - 4\kappa \|u\|_{L^\infty}, \qquad \mathrm{M}^+_{\mathcal{L}} u(x) \leq \mathrm{M}^+ u(x) + 4\kappa \|u\|_{L^\infty}.$$

PROOF: All we have to do is show that for each $L \in \mathcal{L}$, we have $Lu(x) \geq \mathrm{M}^- u(x) - \kappa \inf_{\mathbb{R}^n} u$ and $Lu(x) \leq \mathrm{M}^+ u(x) + \kappa \sup_{\mathbb{R}^n} u$.

We have

$$\begin{aligned} Lu &= \int \delta(u,x,y)(K_1(y) + K_2(y))\mathrm{d}y \\ &= \int \delta(u,x,y) K_1(y)\mathrm{d}y + \int \delta(u,x,y) K_2(y)\mathrm{d}y \\ &\geq \mathrm{M}^- u(x) + \int (u(x+y) + u(x-y) - 2u(x)) K_2(y)\mathrm{d}y \\ &\geq \mathrm{M}^- u(x) - 4\|u\|_{L^\infty} \|K_2\|_{L^1} = \mathrm{M}^- u(x) - 4\kappa \|u\|_{L^\infty}. \end{aligned}$$

In a similar way the inequality for $\mathrm{M}^+_{\mathcal{L}} u(x)$ follows. □

COROLLARY 14.2 *If u is bounded in $\mathbb{R}^n$ and in an open set Ω, $\mathrm{M}^+_{\mathcal{L}} u \geq -C$, and $\mathrm{M}^-_{\mathcal{L}} u \leq C$, then*

$$\mathrm{M}^+ u \geq -C - 4\kappa \|u\|_{L^\infty}, \qquad \mathrm{M}^- u \leq C + 4\kappa \|u\|_{L^\infty}.$$

THEOREM 14.3 *Let $\sigma > \sigma_0$ for some $\sigma_0 > 0$. Let u be a bounded function in $\mathbb{R}^n$ such that*

$$\mathrm{M}^+_{\mathcal{L}} u \geq -C_0 \quad \text{in } B_1, \qquad \mathrm{M}^-_{\mathcal{L}} u \leq C_0 \quad \text{in } B_1;$$

then there is an $\alpha > 0$ (depending only on λ, Λ, n, and σ_0) such that $u \in C^\alpha(B_{1/2})$ and

$$u_{C^\alpha(B_{1/2})} \leq C(\|u\|_{L^\infty} + C_0)$$

for some constant $C > 0$ that depends on λ, Λ, n, σ_0, and κ.

PROOF: From Corollary 14.2

$$\mathrm{M}^+ u \geq -C_0 - 4\kappa \|u\|_{L^\infty}, \qquad \mathrm{M}^- u \leq C_0 + 4\kappa \|u\|_{L^\infty}.$$

Then, from Theorem 12.1,

$$u_{C^\alpha(B_{1/2})} \leq C(\|u\|_{L^\infty} + C_0 + 4\kappa \|u\|_{L^\infty}) \leq \tilde{C}(\|u\|_{L^\infty} + C_0).$$

□

If we use Theorem 14.3 instead of Theorem 12.1 in the proof of Theorem 13.1, we obtain a $C^{1,\alpha}$ result for a class $\mathcal{L}$ that includes kernels with exponential decay or compact support.

THEOREM 14.4 *Let $\mathcal{L}$ be the class of operators with kernels K such that*

$$\int_{\mathbb{R}^n \setminus B_{\rho_0}} \frac{|K(y) - K(y-h)|}{|h|} \, \mathrm{d}y \le C \quad \textit{every time } |h| < \frac{\rho_0}{2}, \tag{14.2}$$

$$K = K_1 + K_2, \tag{14.3}$$

$$(2-\sigma)\frac{\lambda}{|y|^{n+\sigma}} \le K_1(y) \le (2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}}, \tag{14.4}$$

$$\|K_2\|_{L^1} \le \kappa. \tag{14.5}$$

There is a $\rho_0 > 0$ so that if I is a nonlocal elliptic operator in the sense of Definition 3.1 *and u is a bounded function such that $Iu = 0$ in B_1, then there is an $\alpha > 0$ (depending only on λ, Λ, n, and σ) such that $u \in C^{1+\alpha}(B_{1/2})$ and*

$$u_{C^{1+\alpha}(B_{1/2})} \le C(\sup_{\mathbb{R}^n} |u| + |I0|)$$

for some constant $C > 0$.

Remark 14.5. We can prove Theorem 14.3 because in our C^α estimates we allow a bounded right-hand side. Theorem 14.4 would be more general if inequality (14.2) was required with K_1 instead of K. In order to prove such a result, we would need to have $C^{1,\alpha}$ estimates like the ones of Theorem 13.1 with a nonzero right-hand side. This type of result is well-known for elliptic partial differential equations [6], and we are planning to extend it to nonlocal equations in future work.

It is not hard to check that if assumption (14.2) involved K_1 instead of K, then the class $\mathcal{L}$ above would be the same as the larger class $\mathcal{L}_0$ of (3.3), and Theorem 14.4 would apply to a very large family of operators.

Acknowledgments. We would like to thank Prof. Ping'an Zhang and also the referee for pointing out several typos in an earlier version of this paper. Both authors were partially supported by National Science Foundation grants.

Bibliography

[1] Awatif, S. Équations d'Hamilton-Jacobi du premier ordre avec termes intégro-différentiels. I. Unicité des solutions de viscosité. *Comm. Partial Differential Equations* **16** (1991), no. 6-7, 1057–1074.

[2] Barles, G.; Imbert, C. Second-order elliptic integro-differential equations: viscosity solutions' theory revisited. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **25** (2008), no. 3, 567–585.

[3] Bass, R. F.; Kassmann, M. Harnack inequalities for non-local operators of variable order. *Trans. Amer. Math. Soc.* **357** (2005), no. 2, 837–850 (electronic).

[4] Bass, R. F.; Kassmann, M. Hölder continuity of harmonic functions with respect to operators of variable order. *Comm. Partial Differential Equations* **30** (2005), no. 7-9, 1249–1259.

[5] Bass, R. F.; Levin, D. A. Harnack inequalities for jump processes. *Potential Anal.* **17** (2002), no. 4, 375–388.

[6] Caffarelli, L. A. Interior a priori estimates for solutions of fully nonlinear equations. *Ann. of Math. (2)* **130** (1989), no. 1, 189–213.

[7] Caffarelli, L. A.; Cabré, X. *Fully nonlinear elliptic equations*. American Mathematical Society Colloquium Publications, 43. American Mathematical Society, Providence, R.I., 1995.

[8] Ishii, H. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs. *Comm. Pure Appl. Math.* **42** (1989), no. 1, 15–45.

[9] Jensen, R. The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. *Arch. Rational Mech. Anal.* **101** (1988), no. 1, 1–27.

[10] Silvestre, L. Hölder estimates for solutions of integro-differential equations like the fractional Laplace. *Indiana Univ. Math. J*. **55** (2006), no. 3, 1155–1174.

[11] Soner, H. M. Optimal control with state-space constraint. II. *SIAM J. Control Optim.* **24** (1986), no. 6, 1110–1122.

[12] Song, R.; Vondraček, Z. Harnack inequality for some classes of Markov processes. *Math. Z.* **246** (2004), no. 1-2, 177–202.

LUIS CAFFARELLI
University of Texas at Austin
Department of Mathematics
1 University Station C1200
Austin, TX 78712
E-mail: caffarel@math.utexas.edu

LUIS SILVESTRE
University of Chicago
Department of Mathematics
5734 S. University Avenue
Chicago, IL, 60637
E-mail: luis@math.uchicago.edu

Received October 2007.
Revised July 2008.