The regularity theory for solutions to non-local, fully nonlinear equation

Luis Caffarelli
The University of Texas at Austin

May 2009
Part I

- Non-divergence second order equations and fully nonlinear equations from optimal control
- The non local version
- The regularity theory for second order equations (a review)
  - A.B.P
  - Krylov Safanov
  - Initial regularity theory
  - Evans-Krylov
  - Classical solutions
\[ x_0 \xrightarrow{\Delta t} \varphi_{\varepsilon,x_0}(x_0 + y) \, dy = \frac{1}{\varepsilon^n} \varphi_{1,x_0}(x_0 + \varepsilon y) \, dy \]

with \( \varphi \) a symmetric probability density continuous in \( x \).

If we let \( \Delta t, \varepsilon \) go to zero at the appropriate rate

\[ \Delta t \sim \varepsilon^2 \]

the \( \varphi_\varepsilon \) concentrates as a density in the unit sphere and
the probability density $u(x, t)$ of a particle being at position $x$ at time $t$ evolves according to an equation

$$D_t u(x, t) = \int_{S^1} \theta_x(\sigma) D_{\sigma\sigma} u(x, t) d\sigma$$

$$= \lim_{\varepsilon \to 0} \int d\sigma \int \ell^2 \varphi_{\varepsilon, x}(\ell \sigma) \left[ \frac{u(x + \ell \sigma) + u(x - \ell \sigma) - 2u(x)}{\ell^2} \right] d\ell$$

(i.e., $\theta(\sigma)$ is the second radial moment of $\varphi$

$$\theta(\sigma) = \int_0^\infty \ell^2 \varphi(\ell \sigma) d\ell$$
In the non local case, when the process is not continuous and “non infinitesimal” jumps take place

\[ \int \varphi_{\varepsilon,x}(y) \left[ \frac{u(x + y) + u(x - y) - 2u(x)}{|y|^2} \right] dy \]

does not converge anymore to a distribution on the sphere.

For instance, if \( \varphi_{\varepsilon,x}(y) \) is a “Cauchy-like” distribution

\[
\varphi(y) = \frac{1}{(1 + |y|^2)^{n+\alpha/2}}
\]
and $\alpha < 2$, 

$$
\varphi_\varepsilon(y) \sim \frac{\varepsilon^\alpha}{(\varepsilon^2 + |y|^2)^{\frac{n+\alpha}{2}}}
$$

the natural time-space scaling becomes $\Delta t = \varepsilon^\alpha$ and the corresponding diffusion equation becomes 

$$
D_t u(x, t) = (2 - \alpha) \int \frac{u(x + y) + u(x - y) - 2u(x)}{|y|^{n+\alpha}} dy
$$

\[\text{“$\Delta^\alpha(u)$”}\]

More generally, we will be interested in the equation 

$$
D_t u = \int [u(x + y) + u(x - y) - 2u(x)]K(x, y)
$$
Back to the second order case

The family of equations

\[ L(u) = \int \theta_x(\sigma) D_{\sigma\sigma} u \]

(\(\lambda < \theta < \Lambda\)) appears to be a very large one, but in fact, for any given weight \(\theta\)

\[ \int \theta(\sigma) D_{\sigma\sigma} u \]

defines a standard second order elliptic equation, since

\[ D_{\sigma\sigma} u = \sum \sigma_i \sigma_j u_{ij} \]

so

\[ Lu = \sum a_{ij} D_{ij} u, \quad \text{with} \quad a_{ij} = \int \theta \cdot \sigma_i \sigma_j \]
Optimal control

The equations of optimal control arise when we can choose in the continuous case the weight $\theta$ at each point $x$, among a family of weights $\theta_\alpha \in \mathcal{A}$, in order to enhance the diffusion process:

$$D_t u = \sup_{\theta_\alpha, x \in \mathcal{A}} \int \theta_{\alpha, x} u_{\sigma \sigma} d\sigma = \sup [a_{ij, \alpha}(x)D_{ij} u] = F(D^2 u)$$

and from a family of kernels $K \subset \mathbb{K}$:

$$D_t u = \sup_{K \in \mathbb{K}} \int [u(x + y) + u(x - y) - 2u(x)] K(x, y) dy = \Phi(u)$$

for the non local case.
In this lecture we will consider only the stationary case, i.e.,

\[ D_t(u) = 0 = F(D^2u) \quad \text{(resp. } \phi(u)) \]

Heuristically, we may think that \( u \) solves the maximal equation above if, for each \( x \), there exists a particular \( \theta_\alpha \) (resp. \( K \)) such that

\[ L_{\theta_\alpha}(u)(x) = 0 \quad \text{(resp. } L_K(u)(x) = 0) \]

and for any other \( \theta \), (\( K \)), \( L_\theta(u)(x) \) (resp. \( L_K(u)(x) \)) \( \leq 0 \)
Three basic properties of the resulting function $F$ under the sup. process are

**a) Ellipticity:** If $M, N$ are symmetric matrices, and $N$ is positive

$$F(M + N) - F(M) \sim \|N\|$$

In particular if

$$F(M_1) = F(M_2) = 0$$

we have

$$F(M_1) = F(M_1 - M_2 + M_2) = F((M_1 - M_2)^+ - (M_1 - M_2)^- + M_2)$$

and this implies that

$$\|(M_1 - M_2)^+\| \sim \|M_1 - M_2\|^\_$$
b) From a) it follows that the differential of $F$ with respect to $M$,

$$F_{ij}(M) = \frac{\partial F}{\partial M_{ij}}(M)$$

is an “elliptic matrix”

$$\lambda I \leq F_{ij} \leq \Lambda I.$$  

In particular, if $u$ is a solution of $F(D^2 u) = 0$ and we differentiate the equation with respect to a direction $e$, we get

$$F_{ij}(D^2 u)D_{ij}(D_e u) = 0$$

“first derivatives, $w$, of the solution $u$, satisfy an elliptic equation with measurable coefficients.”
c) $F$ is a convex function of $D^2$, since it is the supremum of linear functions

$$\int \theta_\alpha D_{\sigma\sigma}u$$

d) From c) if we take a second derivative of the equation in the direction $e$

$$0 = D_{ee}F = F_{ij}(D^2u)D_{xixj} D_{eeu} + D_{ije}F_{ij,\ell} D_{kle}$$

$D_{ee}u$ (and all its convex combinations) are all supersolutions of the same linearized equation

$$F_{ij}D_{ij} \ldots \leq 0$$
Regularity theory of solutions to linear equations with measurable coefficients and to optimal control

- **Linear equations:** Solutions of a linear equation $u$ satisfy the Harnack inequality (Krylov-Safanov):

\[
\begin{align*}
    u & \geq 0 \text{ in } B_1, \\
    \sup_{B_{1/2}} u & \leq C \inf_{B_{1/2}} u
\end{align*}
\]

and therefore are Hölder continuous:

If

\[
\text{osc}_{B_r} u \leq C (\text{osc}_{B_1} u) r^\alpha
\]
Two important properties implicit in the proof

**a)** weak-$L^\varepsilon$-estimate

1) $u$ a supersolution in $B_1$

2) $u \geq 0$, $u(0) = 1$.

Then

$$|\{u > t\}| \leq Ct^{-\varepsilon}$$

**b)** (oscillation theorem)

1) $u$ a subsolution in $B_1$, $u \leq 1$, then

$$u(0) \leq C|\{u > 0\} \cap B_{1/2}|$$

(compare with mean value theorem for harmonic functions)
Regularity for optimal control

1) Bounded solutions are $C^{1,\alpha}$:

Why?: First derivatives, $\omega = D_e u$, of the solution $u$, satisfy the linear equation

$$F_{ij}(D^2 u)D_{ij}\omega = 0$$

At this point we know nothing about $D^2 u$, but $a_{ij}(x) = F_{ij}(D^2 u(x))$ is an elliptic matrix.

From Krylov-Safanov $\omega$ is $C^\alpha$

(no convexity assumption)
2) (Evans-Krylov)

Bounded solutions are \( C^{2,\alpha} \) (classical) solutions

\[
\left( \| u \|_{C^{2,\alpha}(B_{1/2})} \leq C \| u \|_{L^\infty(B_1)} \right)
\]

**Why?** All second derivatives \( \omega = D_{ee}u \) are supersolutions of the same linearized operator,

\[
F_{ij}(D^2u)D_{ij}w \leq 0
\]

That means that the value of \( w \) at a point, controls the “average” of \( w \) by above \((L^\varepsilon \) lemma).

But \( D^2u = M \) “lives” in the Lipschitz, elliptic surface

\[
F(M) = 0
\]

providing the remaining control, from the opposite side that forces oscillation of \( D^2u \) to decay.
Part II

Non local linear equations and optimal control

(in collaboration with Luis Silvestre)

Typical linear equation

\[ L_K(u)(x) = \int [u(x + y) - u(x - y) - 2u(x)]K(y) \, dy \]

(or more generally \( K(x, y) \))

The kernel \( K \) is non negative, integrable at infinity and in order to make sense, near zero we “can afford” a divergent singularity as long as it is integrable against \(|y|^2\) (at least if \( u \) is \( C^2 \)), i.e.,

\[ \int K(y) \min(|y|^2, 1) < \infty \]
A family of examples is given by the “fractional Laplacians”

\[ \Delta^\tau u = (1 - \alpha) \int \frac{[u(x + y) + u(x - y) - 2u(x)]}{|y|^{n+2\tau}} \, dy \]

so called because its symbol is \(-|\xi|^{2\tau}\) (i.e., has the homogeneity of a 2\(\tau\)-derivative and \(-\Delta^\tau\) is a positive operator, and also because it is the first variation of minimizing the \(H_\tau\) norm (“\(\tau\) derivatives in \(L^2\)”).
As in the second order equations, solutions of

$$\Delta^{\tau} u = f$$

are expected to be “$2\tau$ derivatives” better than $f$ and this indicates that if we expect the equation $L_K(u)$ to be “regularizing”, $K$ should be divergent at the origin.

(If $K$ is in $L^1$, $L_K(u)$ is “at least as good” as $u$, so it will not force regularity of $u$!)
In fact, to expect “$h$-derivatives regularization” by $(L_K)^{-1}f$, the kernel $K(y)$ should diverge at zero by a factor of $|y|^{-h}$. So, we fix the “order of the integrate equation $L_K(u)$” to be $h$ by asking the kernel $K_x(y)$ to satisfy

$$
\lambda(2 - h)|y|^{-(n+h)} \leq K_x(y) \leq \Lambda(2 - h)|y|^{-(n+h)}
$$
Regularity of linear equations

We consider a “linear equation of order $h$, with measurable coefficients” to be

$$Lu = \int_{\mathbb{R}^n} [u(x + y) + u(x - y) - 2u(x)]K(x, y)$$

with $K$ having no regularity in $\chi$ and

$$\lambda(2 - h)|y|^{-(n+h)} \leq K(x, y) \leq \Lambda(2 - h)|y|^{-(n+h)}$$

Note that to compute $Lu$ we need the values of $u$ in all of $\mathbb{R}^n$, so in order to study

$$L(u) = 0 \text{ in } B_1$$

we still need to prescribe (as Dirichlet data) the values of $u$ in all of $\mathbb{R}^n$. (For a jump process the particle does not “hit the boundary”. It may jump to anywhere in $\mathbb{R}^n$.)
Harnack inequality

If $u$ is a non negative solution in $B_1$ of

$$0 = L_{K_x}(u) = \int [u(x + y) + u(x - y) - 2u(x)]K(x, y)$$

with $K$ of order $h \geq h_0 > 0$ and $u$ outside $B_1$ is “reasonable”.

Then

$$\sup_{B_{1/2}} u \leq C[u(0) + \text{“nice correction from the data outside” } B_1]$$
• The constant $C$ is uniform all the way to second order equations

• $L^\varepsilon$ lemma and oscillation lemma hold

• Main difficulty is to find the corresponding Alexandrov-Bakelmann-Pucci theorem, and in particular the estimates for the gradient map.
Regularity for optimal control equations

\[ 0 = \phi(u) = \sup_{K \in \mathbb{K}} \int [u(x + y) + u(x - y) - 2u(x)]K(y) \, dy \text{ for } x \text{ in } B_1 \]

All \( K \in \mathbb{K} \) uniformly of order \( h \geq h_0 > 0 \). Then

**a)** Initial regularity: \( u \) is of class \( C^{1,\alpha}(B_{1/2}) \) and

\[ |u|_{C^{1,\alpha}(B_{1/2})} \leq C|u|_{L^\infty(B_1)} + \text{“term from outside”} \]

(all uniform in \( h_0 \leq h \leq 2 \))
(Again, “first derivatives are solutions of a linearized operator” and as such they are Hölder continuous.)

\( \phi(u) \) does not have to be a sup \( L_K \), can be an inf sup \( L_K \)

(Isaacs equation from game theory)
**Higher regularity** (non local version of Evans-Krylov)

If $\phi(u) = 0$, $\phi$ convex, of order $h$, then $u$ is a classical solution, with regularity

$$\|u\|_{C^{h+\sigma}(B_{1/2})} \leq \|u\|_{L^\infty(B_1)} + \text{ outside}$$

The $\sigma$ and the constants are uniform for $0 < h_0 \leq h \leq 2$, all the way to second order.
**Heuristic of the proof**

The proof is based on the following idea.

For every $x$ we consider the set $A(x)$ of those $y$ for which

$$h(x, y) = [u(x + y) + u(x - y) - 2u(x) - [u(y) + u(-y) - 2u(0)] \leq 0$$

This set $A(x)$ is in some sense extremal.

Indeed for any other set $B$

$$0 \geq \int_B h(x, y)|x - y|^{-(n+s)} dy \geq \int_{A(x)} h(x, y)|x - y|^{-(n+s)} dy \quad (\ast)$$

We want to show that as $x$ goes to zero, $(\ast)$ goes to zero Hölder continuously. If we could keep $A$ fix, $(\ast)$ would be a supersolution and the $L^\varepsilon$ estimate applies.

We then do that by showing the convergence “for all $A$’s” at the same time.
Second order case

Let $u$ be a $C^2$ solution of

$$F(D^2u) = \sup_{\theta \in A} \int \theta(\sigma)u_{\sigma\sigma}d\sigma = 0$$

($F$ convex!).

Then

$$\|u\|_{C^{2,\alpha}(B_{1/2})} \leq C\|u\|_{C^{1,1}(B_1)}$$

**Proof.** We will prove that

$$|\Delta u(x) - \Delta u(0)| \leq \left| \int_S [u_{\sigma\sigma}(x) - u_{\sigma\sigma}(0)] d\sigma \right| \leq |x|^\alpha$$
This follows after rescaling from the following inductive argument:

For any measurable set $A$ of the unit sphere $S^1$, define

$$h(x, A) = \int_A [u_{\sigma\sigma}(x) - u_{\sigma\sigma}(0)] d\sigma$$

Assume that for any $A$,

$$h(x, A)|_{B_1} \geq -1$$

There exists a $\tau_0 > 0$ such that for any $A$,

$$h(x, A)|_{B_{1/2}} \geq -1 + \tau_0$$
Indeed, this would imply, by choosing $A(x) = \chi_{u_{\sigma\sigma} > 0}$, that

$$\int |u_{\sigma\sigma}(x) - u_{\sigma\sigma}(0)| \leq |x|^\alpha$$

and in particular $|\Delta u(x) - \Delta u(0)| \leq |x|^\alpha$.

To prove the inductive statement, let me recall the consequences from Krylov-Safanov:
a) weak-$L^\varepsilon$-estimate

1) $u$ a supersolution in $B_1$

2) $u \geq 0$, $u(0) = 1$.

Then

$$|\{u > t\}| \leq Ct^{-\varepsilon}$$

b) (oscillation theorem)

1) $u$ a subsolution in $B_1$, $u \leq 1$, then

$$u(0) \leq C|\{u > 0\} \cap B_{1/2}|$$
The proof is then as follows: Assume that for some \( x_0 \in B_{1/4} \), for some \( A_0 \),

\[
h(x_0, A_0) \leq -1 + \tau_0
\]

By convexity of \( F \), \( h(x, A_0) + 1 \) is a (non negative) supersolution in \( B_1 \). Since

\[
h(x_0, A_0) + 1 \leq \tau_0
\]

from the \( L^\varepsilon \) estimate, for \( \tau_0 \) very small, this forces

\[
h(x, A_0) + 1 \leq (\tau_0)^{1/2}
\]

in “almost all” of, say, \( B_{3/4} \):

\[
|\{h(x, A_0) + 1\} \geq \tau_0^{1/2}| \leq C \tau_0^{-\varepsilon/2}
\]
Also, if \( C \subset A \) (resp. \( C \subset B = CA \)) and \( x \in \Omega = \{ h(x, A) \leq \tau^{1/2} \} \) we must have

\[
h(x, C) \leq (\tau_0)^{1/2} \quad \text{(resp. } h(x, C) \geq -\tau_0^{1/2})
\]

(If no, by deleting (or adding) \( C \) from \( A \) we would make

\[
h(x, A \setminus C) < -1 \quad \text{(resp. } h(x, A \cup C) < -1)
\]

a contradiction.)

Therefore, in \( \Omega \),

\[
|h(x, A) - \left( -\int (u_{\sigma\sigma})^- (x) \right)| \leq C \tau^{1/2}
\]

and also

\[
|h(x, B) - \int (u_{\sigma\sigma})^+ (x)| \leq C \tau^{1/2}
\]
But remember that ellipticity implies that

$$\int (u_{\sigma\sigma})^- \sim \int (u_{\sigma\sigma})^+$$

Then, since in $\Omega$, $\int u^-_{\sigma\sigma} \sim -h(x, A) \sim -1$, this implies that

$$h(x, B) \sim \int u^+_{\sigma\sigma} \geq n_0 > 0 \quad \text{(say } 10^{-4})$$

But $h(x, B)$ is also a supersolution. Hence $v(x) = n_0 - h(x, B)$ is a subsolution,

$$v \leq n_0 + 1 \quad \text{(since } h \geq -1)$$

$$v \leq 0 \text{ in almost all of } B_{3/4} \quad \text{(i.e., in } \Omega)$$

and $v(0) = n_0$. 
This contradicts the oscillation lemma, since $\nu(0)$ must satisfy

$$\nu(0) \leq |\{v \geq 0\} \cap B_{3/4}|,$$

“almost nowhere” on $B_{3/4}$.

In the non local case the spirit of the proof is the same:

In this case for $A$ a symmetric set in $\mathbb{R}^n$ (instead of the sphere) we define

$$h(x, A) = \int \left\{ \left[ u(x + y) + u(x - y) - 2u(x) \right] - \left[ u(0 + y) + u(0 - y) - 2u(0) \right] \right\} K(y) \chi_A(y) \, dy$$

and the same ideas apply.