Non-local Diffusions, Drifts and Games

Luis Caffarelli

Abstract This is a brief discussion of the properties of solutions to several non-linear elliptic equations involving diffusive processes of non-local nature. These equation arise in several contexts: from continuum mechanics and phase transition, from population dynamics, from optimal control and game theory. The equations coming from continuum mechanics exhibit a variational structure and a theory parallel to the De Giorgi–Nash–Moser was necessary to show existence of regular solutions. Population dynamics suggests "porous media like equations" with a non-local pressure, and from optimal control we obtain fully non-linear equations that require methods of the type of the Krylov–Safonov–Evans theory. Finally, we discuss some non-local p and infinite Laplacian models coming from game theory.

1 Introduction

We are interested in integral diffusion equations:

$$u_t(x,t) = [L(u)](x,t)$$

where the operator L takes the form

$$L(u(x,t)) = \int [u(y) - u(x)] K(x,y)\, dy$$

for some positive kernel (or measure) $K(x,y)$ (or $K_x(y)$).

We call the equation a diffusion equation because solutions try to revert to some sort of "integral average" of u.

Indeed, if $u(x_0)$ is "smaller than" its surrounding values, as weighted by $K(x,y)$, $u(x_0,t)$ will tend to increase, if "bigger", to decrease (i.e., $u_t > 0$ or $u_t < 0$).

We may think of the heat equation as an infinitesimal version of this process.

L. Caffarelli (✉)
Department of Mathematics, The University of Texas at Austin, 1 University Station—C1200, Austin, TX 78712, USA
e-mail: caffarel@math.utexas.edu

H. Holden, K.H. Karlsen (eds.), *Nonlinear Partial Differential Equations*, Abel Symposia 7, DOI 10.1007/978-3-642-25361-4_3,

Indeed the Laplacian, Δu, is the limit of

$$\Delta u(x_0) = \lim_{r \to 0} \frac{1}{r^2} \fint_{B_r(x_0)} u(x) - u(x_0)\, dx.$$

In fact, if

$$K_\varepsilon(x, y) = \frac{1}{\varepsilon^2}\left(\frac{1}{\varepsilon^n}\varphi\left(\frac{x-y}{\varepsilon}\right)\right) = \frac{1}{\varepsilon^2}\varphi_\varepsilon(x-y)$$

for φ a probability density (a mollifier), the corresponding solutions u_ε should converge to a solution $u_0(x,t)$ of

$$(u_0)_t = a_{ij} D_{ij} u_0$$

where a_{ij} are the second moments of φ.

These types of equations (and the associated non-linear ones that we will discuss shortly) have roots in different phenomena and, as their second order counterpart, they naturally divide between those with variational structure and those coming from probabilistic considerations.

A familiar example for the first case is prescribing Neumann boundary data (for instance zero). Insulating a wall implies some temperature diffusivity along the surface, expressed by inverting the Dirichlet to Neumann map. A non-local related equation is the quasi-geostrophic equation that describes the evolution of temperature on the ocean surface, due to the (one-side) atmospheric conditions.

On the probabilistic side let us recall the Levy–Khintchine formula. In an informal "black-box" approach suppose we can observe the transition probability of a distribution of particles, for any sequence of times t_k, and we realize that the transition from t_1 to t_2 only depends on t_1 and t_2, in fact on $t_2 - t_1$.

Then, for any k, we can write the transition probability from t_1 to t_2 as the composition (convolution) k times of the transition from $t = 0$ to $t = \frac{1}{k}(t_2 - t_1)$. This suggests the possibility, as δt goes to zero, of describing the process through a "heat equation"—as a properly scaled infinitesimal limit of the δt transition.

This is what the Levy–Khintchine formula asserts: That the probability density evolves according to a heat equation

$$u_t = \cdots$$

consisting of a continuous part

$$\cdots = a_{ij} D_{ij} u + b_j \nabla u + \cdots$$

a symmetric jump process

$$\int [u(x+y) + u(x-y) - 2u(x)]\, d\mu(y)$$

$+\cdots$ an asymmetric part that we will discuss later

$$d\mu\text{“}=\text{”}K(y)\, dy.$$

Here, it is required to make sense for a C^2, bounded u, i.e.,

$$\int_{B_1(0)} \|y\|^2 \, d\mu(y) < \infty$$

and

$$\int_{\mathscr{C}(B_1(0))} d\mu(y) < \infty.$$

Between divergence and non-divergence lie the equations invariant under translations, i.e., where the kernel $K(x, y) = \widetilde{K}(x - y)$. In this case, the equation can be thought of as having both divergence and non-divergence structure and also, being of convolution type, they enjoy the advantage of allowing for methods of harmonic analysis.

That is the case, for instance, with the family of fractional Laplacians: For $0 < \alpha < 1$

$$\text{“}\Delta^{\alpha}\text{”}(x) = C(\alpha) \int [u(y) - u(x)] \frac{1}{|x - y|^{n+2\alpha}}, dy = (\hat{u}(\xi)|\xi|^{2\alpha})^{v}.$$

The constant $C(\alpha) \sim (1 - \alpha)$ to recuperate the standard Δu, as α goes to one.

Notice that the range of α's is such that it makes these kernels satisfy the Levy–Khintchine condition to be an infinite divisible distribution.

In fact the fractional Laplacians are also called “stable processes”.

On the other hand, the fractional Laplacians is what we obtain as an Euler–Lagrange equation for the energy integral corresponding to the $W^{\alpha,2}$ (the L^2 norm of the “alpha” fractional derivatives of u):

$$\text{“}D^{\alpha} u(x) = \int [u(y) - u(x)] \frac{1}{|x - y|^{n+\alpha}} \, dy\text{”}.$$

And finally, convolution with the Δ^{α} kernel corresponds after Fourier transform to the multiplier

$$(\widehat{\Delta^2}) = -|\xi|^{2\alpha}.$$

In that sense, the fractional Laplacian serves as a basic model for the three classical methods of second order PDE's.

- Superposition (potential theory, harmonic analysis)
- Energy method (calculus of variations, DeGiorgi–Nash–Moser)
- Probabilistic (optimal control-Krylov–Safonov)

Since we are interested in regularity properties of solutions to such an “elliptic” or “parabolic” equation, the kernel $K(x)$ should be singular at the origin to force u to be somewhat “special” in order to satisfy the equation: To know that after convolution with a smooth function u is smooth does not reflect so much on the regularity of u, at least at first glance.

In that sense, the fractional Laplacians provide a natural comparison scale of “order of differentiation” of the operator to help us develop a general setting.

2 Divergence Structure

Equations with "divergence" structure arise from continuum mechanics and calculus of variations.

A rough characterization would be that the kernel $K(x, y)$ is symmetric.

That makes the equation

$$\int [u(y) - u(x)]K(x, y)\, dy = 0$$

the Euler–Lagrange equation of

$$E(u)T(u) = \iint [u(x) - u(y)]^2 K(x, y)\, dx\, dy$$

and thus puts the problem in the framework of weak variational solutions test functions methods, etc.:

For a test function $\varphi(x)$, the bilinear form

$$B(u, \varphi) = \iint [u(y) - u(x)]K(x, y)[\varphi(y) - \varphi(x)],$$

depending on the problem at hand, must be zero, or prescribed or equal to

$$\int \varphi(y)u_t(x, t)$$

in the parabolic setting.

The general "non-linear calculus of variations" framework becomes then the study of the minimizers of the form:

$$\int \phi(u(x) - u(y))K(x - y)\, dx\, dy$$

with ϕ convex (quadratic for "uniform" fractional ellipticity).

The first, natural problem to study is that of regularity of local minimizers (the equivalent of the DeGiorgi solution of the Hilbert problem and the development of the DeGiorgi–Nash–Moser theory of regularity of solutions). Let us recall that in the second order case, the theory proceeds as follows:

A local minimizer, u, of the functional

$$E(w) = \int F(\nabla w)\, dx$$

satisfies the Euler–Lagrange equation

$$D_{x_i} F_i(\nabla u) = 0$$

or, in non-divergence form:

$$F_{ij}(\nabla u)D_{x_i x_j} u = 0.$$

If we would known that ∇u is continuous Shauder estimates would allow us to bootstrap the solution to higher regularity. In turn, first derivatives $D_e u = w$ satisfy

$$D_{x_i} F_{ij}(\nabla u) D_{x_j} w = 0.$$

But at this point we only know that ∇u is in L^2 and, from the uniform convexity of F, that the matrix $F_{ij}(\cdot) = A_{ij}(x)$ is strictly positive:

$$\lambda I \leq F_{ij}(\cdot) \leq \Lambda I.$$

But then, the celebrated DeGiorgi theorem establishes that solutions of an elliptic equation

$$D_i a_{ij}(x) D_j w = 0$$

with *no regularity* assumption on a_{ij} are Hölder continuous.

In particular, ∇u is Hölder continuous and higher regularity follows.

In this context, with Chan and Vasseur [9], we develop the DeGiorgi regularity theory for the parabolic case:

Let $u(x,t)$ be the solution of

$$u_t(x,t) = \int \phi'(u(x) - u(y)) K(x-y)$$

with "ϕ symmetric and quadratic" (i.e., $\lambda \leq \phi'' \leq \Lambda$) and

$$(1-\alpha)m|z|^{-(n+2\alpha)} \leq K(z) \leq (1-\alpha)M|z|^{-(n+2\alpha)}.$$

Then u becomes instantaneously smooth.

As in the second order case, the central step is to prove that first derivatives, $w = D_x u$, satisfy a "rough equation" and are Hölder continuous:

$$w_t(x,t) = \int [w(y,t) - w(x,t)] \underbrace{\phi''(u(y,t) - u(x,t)) K(x-y)}_{\substack{\text{"symmetric, measurable} \\ \text{fractional Laplacian like} \\ \text{kernel" } K(x,y,t)}} \, dy$$

(see also related articles by Barlow, Bass, Chen, Kassman, and of Komatsu [1, 3, 14, 15]).

The study of non-local, non-linear equations with "variational structure" has several motivations:

- What we could call surface diffusion: the quasigeostrophic equation that models ocean atmosphere interaction, the theory of semi-permeable membranes, planar fracture dynamics (see [5, 11]).
- Problems in statistical mechanics, like phase transition problems with long range interactions (as opposed to neighbor to neighbor). See for instance the work of Giacomin–Lebowitz and of Presutti [12].
- Material sciences, for instances polymers where many scales interact.
- Image processing, see for instance the work of Gilboa and Osher [13].

3 Non-divergence Equations

"Non-divergence" equations arise instead from probability (Levy processes), optimal control and game theory.

Suppose for instance that particles generate at some point x_0 of a domain Ω and bounce randomly until they exit Ω.

At that moment they release an amount of energy $u(y)$ depending on the point y where they land.

In principle to find out the expectation for future released energy $u(x_0)$ when starting at x_0, we should just solve $Lu = 0$ in Ω with external data $u(y)$ and the diffusion associated to the process.

In the case of optimal control we are able to "design" the jump process (the media) to maximize the expected value $u(x_0)$.

That is: We have a family of possible diffusion processes given by the kernels

$$L_\alpha u(x) = \int [u(x+y) - u(x)] K_\alpha(y)\, dy$$

and at each x we want to chose the optimal jump distribution

$$L_{\alpha(x)} = \int [u(x+y) - u(x)] K_{\alpha(x)}(y)\, dy.$$

In order to achieve that we have to find a solution u_0 of the equation

$$F(u_0) = \sup_\alpha L_\alpha u_0 = 0$$

with exterior data $u(y)$.

Indeed, this equation means that "u_0 is a supersolution of all the admissible operators, and at each point is the solution of at least one of the L_α." Therefore on one hand it is better than any choice and at the same time is an admissible distribution.

In the case of second order equations, the central result of the theory is the Evans–Krylov theorem:

In that case, the family of operators are second order

$$L_\alpha(u) = \sum a_{ij}^\alpha D_{ij} u,$$

the non-linear equation is

$$F(D^2 u) = \sup_\alpha \sum a_{ij}^\alpha D_{ij} u$$

and the Evans–Krylov theorem asserts that solutions to $F(D^2u) = 0$ are $C^{2,\beta}$ and thus **classical** (i.e., the derivatives involved are continuous).

In collaboration with Silvestre, we reproduce their theory for the corresponding non-local equations [6–8].

If the kernels K_α are all comparable to the s-Laplacian:

$$\lambda(1-s)|y|^{-(n+2s)} \le K_\alpha(y) \le \Lambda(1-s)|y|^{-(n+2s)}$$

and they are **symmetric** in y (no "drift"), then solutions to $F(u)=0$ are $C^{2s+\beta}$ that makes the corresponding integrals convergent and the solutions "classical".

One of the main features of the work is the proof of a theorem equivalent to the Krylov–Safonov Harnack inequality for "bounded measurable" kernels:

If w is for every x a solution of a different equation

$$L_x(w) = \int [w(x+y) - w(x)] K_x(y)\,dy = 0$$

with K_x changing discontinuously with x "bounded measurable coefficients", w is still Hölder continuous.

4 Drifts

What I want to discuss now is the relation, or interaction between diffusion and drift in the optimal control context:

For second order equations, when addressing gradient dependence of an equation, we have two different issues. On one hand semilinear equations, say, for instance

$$\Delta u = g(u, \nabla u)$$

with an associated idea of drift or transport and on the other quasilinear equations:

$$a_{ij}(\nabla u) D_{ij} u = 0$$

for instance those coming from the calculus of variations.

Semilinear equations with fractional diffusions arise for instance in the case of the quasigeostrophic equation:

$$\text{“}u_t - \Delta^s u = g(u, \nabla u)\text{”}$$

and assuming nice dependence on u, there is here a clear competition between diffusion and transport that becomes critical where $s = 1/2$.

But there is a second, implicit form of drift in the asymmetry of the kernel for a Levy process:

The most general "heat equation" for an infinite divisible distribution, leaving aside the continuous part and the standard drift is

$$\begin{aligned} u_t &= \frac{1}{2}\int [u(x+y) + u(x-y) - 2u(x)]\,d\mu(y) \\ &\quad + \frac{1}{2}\int \big([u(x+y) - u(x-y)] - 2(\nabla u(x), y)\chi_{B_1}\,d\mu\big)\,d\mu \\ &= \text{symmetric} + \text{antisymmetric}. \end{aligned}$$

Note that the antisymmetric part has in it an extra cancellation to ensure that the process does not drift to infinity.

For quasilinear equations, one equivalent framework to the second order case is, of course, through the calculus of variations.

For instance, one defines the p- (s-Laplacian), i.e., s-derivatives in L^p, as the Euler–Lagrange equation of the L^p norm of the s-derivatives of a function

$$\|u\|^p_{W^{s,p}} = \iint \frac{[u(x) - u(y)]^p}{|x - y|^{n+sp}}\,dx\,dy.$$

This p-fractional Laplacian is naturally studied through "energy" and "test functions methods" (see [10]). But the p-Laplacian also can also be written in non-divergence form as

$$(p - \Delta)u = |\nabla u|^{p-2}(\Delta u + (p-2)u_{nn})$$

where u_{nn} denotes the second derivative in the direction of the gradient of u.

And this has a game-theoretical interpretation (Peres–Sheffield [16]): Let us go back to the example of expected energy release $u(x)$ of the random particle.

Suppose that as before the random process has the ("almost continuous") diffusion equation ($\delta_t \sim \varepsilon^2$)

$$\delta_t u(x,t) = \int [u(y+x)(y,t) - u(x,t)]\frac{1}{\varepsilon^2}\varphi_\varepsilon(y)\,dy$$

i.e., the particle at position x at time t, jumps, by time $t + \varepsilon^2$, to a position epsilon-away, according to the radially symmetric probability density $\varphi_\varepsilon(y) = \frac{1}{\varepsilon^n}\varphi(y/\varepsilon)$.

Then, as discussed before, when ε goes to zero, we would get the standard "heat" equation.

But, assume now that competing players P_1, P_2 are able to impose on the jump an epsilon-drift in their preferred direction, randomly in time, trying to maximize, respectively minimize, the expected value u.

That is, depending on which player has the input, the particle at x will jump to the position $(x + y)$, with probability density ($\tau_i = \tau_1$ or τ_2, a unit vector)

$$\varphi_\varepsilon(y + \lambda\tau_i) = \frac{1}{\varepsilon^n}\left(\varphi\left(\frac{y + \lambda\tau_i}{\varepsilon}\right)\right).$$

As a consequence, the jump probability density φ_ε has drifted in the direction τ_1 or τ_2 depending on which player imposed the drift. Here λ is the intensity of the drift and the expected value u will then satisfy the Isaac's equation

$$\inf_{\tau_1 \in S^1} \sup_{\tau_2 \in S^1} \left[\frac{1}{2\varepsilon^2}\int [u(x+y) - u(x)]\varphi_\varepsilon(y + \lambda\tau_1) + \varphi_\varepsilon(y + \lambda\tau_2)\,dy\right] = 0.$$

The natural choice for τ_2 is to push the drift in the direction of ∇u, and for τ_1 in that of $-\nabla u$. Therefore, if both players use the optimal strategy, the combination of

$$\varphi_\varepsilon(y + \lambda\tau_1) + \varphi_\varepsilon(y + \lambda\tau_2)$$

will shift the mass of φ_ε symmetrically in the directions of $\pm\nabla u$, increasing the second moment in that direction so that the limiting equation, as epsilon goes to zero becomes

$$\Delta u + C(\lambda)u_{nn}$$

i.e., the non-divergence form of the fractional Laplacian.

A similar argument can be made for jump processes:

In work with Bjorland and Figalli, we have studied existence and regularity properties of this "tug of war" game for jump processes. Let me start by pointing out that there are different ways to "influence the drift" that give rise to structurally different mathematical problems. A possible one is for shifted kernels:

That is, for kernels of the form

$$K_{e_1}(y) = K_0(y)[1 + A(y_1)]$$

with $K_0(y)$ a symmetric kernel of the size of a fractional Laplacian, and $A(y_1)$ a smooth odd function, $|A(y_1)| \le 1 - \delta$.

That is, we look at the Isaac's equation:

$$\inf_{\nu_1}\sup_{\nu_2}\frac{1}{2}\int [u(x+y) - u(x)]K_0(y)[2 + A(y\cdot\nu_1) + A(y\cdot\nu_2)]\,dy$$

(i.e., each player adds the implicit drift $A(y\cdot\nu)$ in his optimal direction ν_1 or ν_2).

Another possible way is that the player chooses a direction and it is this direction that suffers a random deviation (an "unsteady hand"). In that case the corresponding basic kernel $K_{e_1}(y)$ should be of the form

$$K_{e_1}(y) = K_0(y)\eta(\sigma\cdot e_1)$$

where η may vanish outside a neighborhood of e_1.

The final operator is as before, the inf sup over all rotations of $K_{e_1}(y)$.

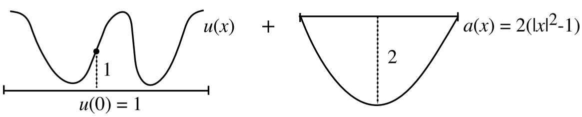
In both cases, it follows from the non-local Harnack inequality and ABP theorem [6] that solutions are C^α for some α.

In fact, let me take this opportunity to discuss informally the non-local ABP theorem, that is central to many of the developments for non-local optimal control.

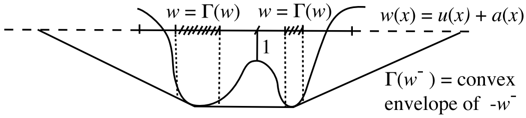
The local version of the ABP theorem needed for the Harnack inequality (as presented in [4]) is the following:

Theorem 1 *$u \ge 0$ in B_1, $Lu = a_{ij}(x)D_{ij}u \le 0$, $u(0) \le 1$. Then, $\exists\varepsilon_0$, such that $|\{u < 2\}| \ge \varepsilon_0(\lambda, \Lambda) > 0$.*

Proof We add to u a negative paraboloid in B_1:



and construct its convex envelope in B_1:

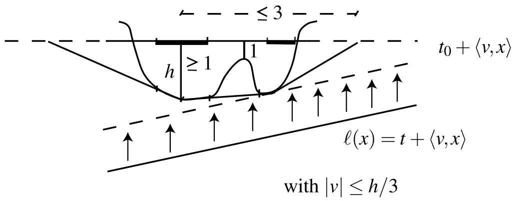


We will estimate $|\{w = \Gamma(w)\}|$ by below.

Indeed, in this set w is negative and

$$u = w - a \leq 0 + 2.$$

For this purpose, we use the classical A-B-P argument, i.e., we estimate the volume of the image of the gradient map: $\nabla\Gamma : B_1 \to \mathbb{R}^n$. To do that, we lift from minus infinity a plane with generic slope v:



If $\ell(x) = t + \langle v, x\rangle$ with $|v| \leq h/3$, for some value t_0, ℓ is a supporting plane of $\Gamma(w^-)$ at some interior point $x_0 \in \{\Gamma = w\}$. Thus "the image of $\{w^- = \Gamma(w)\}$ by the map: $x \to \nabla\Gamma(x)$ contains the ball of "v's" of radius $\frac{h}{3} = \frac{\sup w^-}{3} \geq \frac{1}{3}$, i.e.,

$$\left(\frac{1}{3}\right)^n \leq C\,\mathrm{Vol}[\nabla\Gamma(\{w^- = \Gamma(w)\})].$$

We now "change variables", from v to x

$$1 \leq C\,\mathrm{Vol}[\nabla\Gamma(\{w^- = \Gamma w\})] = \int 1\,dv = \int_{\{w=\Gamma(w)\}} |\det D\nabla\Gamma|\,dx.$$

But $D\nabla\Gamma = D^2\Gamma$, a non-negative matrix, since Γ is convex, so

$$\left(\frac{1}{3}\right)^n \leq \int_{\{w=\Gamma(w)\}} \det D^2\Gamma \leq \int_{\{w=\Gamma(w)\}} \det D^2 w \leq \int_{\{w=\Gamma w\}} [\mu_{\max}(D^2(w))]^n$$

with μ the largest eigenvalue of $D^2 w$ (at a contact point $w = \Gamma(w)$, $D^2 w \geq D^2\Gamma \geq 0$). Since all max eigenvalues of $D^2 w \geq 0$,

$$Lw \cong \lambda\mu_{\max}, \quad \text{but also} \quad Lw \leq Lh \leq 2n\Lambda.$$

We then get

$$1 \leq C\left(\frac{\Lambda}{\lambda}\right)^n \int_{\{w=\Gamma(w)\}} 1 = C\left(\frac{\Lambda}{\lambda}\right)^n |\{w = \Gamma(w)\}|.$$

This is "almost" the proof of the ABP version we need for the Harnack inequality.

What we are missing is the localization property:

$$\text{“}\{w = \Gamma(w)\} \cap Q_{1/4}(y) \quad \text{for } |y| \le 1/4\text{”} \quad \text{instead of} \quad \{w = \Gamma(w)\},$$

i.e., we need the extra fact that we can get the contact set to be inside of any cube of size $1/4$ close to the origin in order to make a C-Z decomposition. For that, all we need is to change h by an h' with: $Lh' \le 0$ outside $Q_{1/4}$ (so that $Lw \le 0$ outside $Q_{1/4}$ and contact cannot occur), $h'(0) \le -2$ so $\inf w \le -1$, and Lh' still bounded above, so Lw is bounded above. □

5 The Corresponding ABP for Integral Diffusions [6]

As before, we assume $u \ge 0$ in B_3, $Lu \le 0$, $u(0) \le 1$. Now

$$Lu(x) = \int [u(x+y) + u(x-y) - 2u(x)] K_x(y)\, dy = \int \delta_2 u(x,y) K_x.$$

For simplicity we will truncate K_x:

$$\lambda(2-s)|y|^{-n+s}\chi_{B_1}(y) \le K_x(y) \le \Lambda(2-s)|y|^{-(n+s)}\chi_{B_1}(y)$$

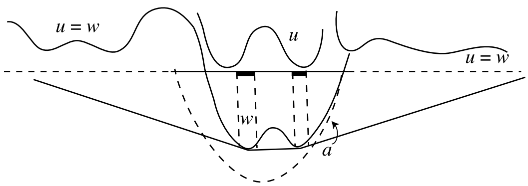
and restrict ourselves to $x \in B_1(0)$, so L is well defined.

We want to show:

$$\text{“}\exists M, \varepsilon > 0,\ M, \varepsilon(\lambda, \Lambda, s), \quad \text{such that} \quad |\{u < M\} \cap B_1| \ge \varepsilon\text{”}$$

$$M, \varepsilon \text{ deteriorate with } s \text{ only for } \boxed{s \to 0}.$$

We proceed as before. Consider $w = u + a$, with $a = 2(|x|^2 - 1) \wedge 0$ and construct the convex envelope $\Gamma(w^-)$ in B_3

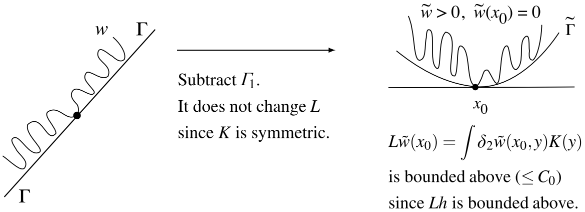


As before

$$\text{Vol}\ \nabla\Gamma(\{w = \Gamma(2)\}) \ge \left(\frac{|\inf w|}{4}\right)^n.$$

The problem is how to relate $|\{w = \Gamma(w)\}|$ with its image (no good change of variables formula).

Consider a point x_0 in $w = \Gamma(w)$. We have the following local geometry



We are going to prove the following family of *steps*. Consider x_0 as above.

(a) For some small diadic ring $\mathbb{R}_k = B_{2^{-k}} \setminus B_{2^{-(k+1)}}$ $\tilde{w}$ "grows quadratically on average" in the sense that

$$\fint_{R_k} \tilde{w} \le C_1 (r_k)^2 \quad (r_k = 2^{-k}).$$

(b) Of course, this does not imply that $\tilde{w} \le c_1 r_k^2$, but since $0 \le \tilde{\Gamma}(\tilde{w}) \le \tilde{w}$, and $\tilde{\Gamma}$ is convex. (a) does imply that

$$\tilde{\Gamma}|_{B_{2^{-(k+1)}}} \le C_2 2^{-2k} \quad \text{and} \quad \nabla \tilde{\Gamma}|_{B_{2^{-(k+2)}}} \le C_2 2^{-k}$$

($k = k(x_0)$ of course).

(c) In particular:

$$\operatorname{Vol} \nabla \Gamma(B \ldots) = \operatorname{Vol} \nabla \tilde{\Gamma}(B_{2^{-(k+2)}}(x_0)) \le C |B_{2^{-(k+2)}}(x_0)|.$$

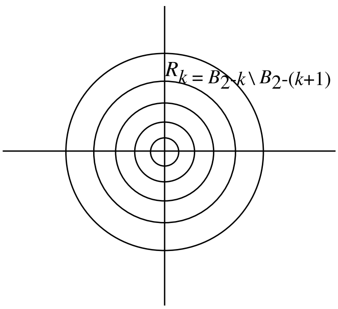
We now extract a covering of $\{w = \Gamma(w)\}$ with the family of these balls $B_{r(x)}(x)$ and we have

$$1 \le \operatorname{Vol} \nabla \Gamma(\{w = \Gamma\}) \le C \sum |B_{r(x_j)}(x_j)|.$$

But in each B_{r_j}, $(\cdot)\tilde{w}$ differs from Γ by at most $(r_j)^2$ in a large portion of $B_{r(x_j)}$ since

$$u \le w + 2; \quad |\{u \le 3\}| \ge C \sum |B_{r(x_j)}| \ge 1.$$

We divide the integration in diadic rings around x_0



$$L\tilde{w} = \sum_k \int_{\mathbb{R}_k} \delta \tilde{w}\, K \le C.$$

Since $\tilde{w}(x_0) = 0$, and $\tilde{w} \geq 0$, the integrand in all of the rings is positive and

$$Lw \sim (2-s) \sum (r_k)^{-(n+s)} \int_{\mathbb{R}_k} \tilde{w} \leq C.$$

(a) We first show that if $C_1 = MC$ is a large multiple of C there is at least one ring where

$$\fint_{\mathbb{R}_k} \tilde{w} \leq C_1 r_k^2.$$

If not

$$C \geq L\tilde{w} = (2-s) \sum r_k^{-(n+s)} \int \tilde{w} \geq (2-s) \sum r_k^{-s} \fint \tilde{w} \geq C_1 \frac{2-s}{1-2^{(s-2)}}$$

$\sim C_1$, a contradiction.

In fact, if M is large, we can start the sum from $k = k_0$ and we get

$$C \geq C_1 \frac{(2-s)}{1-2^{(s-2)}} \cdot 2^{(2-s)k_0}$$

still a contradiction.

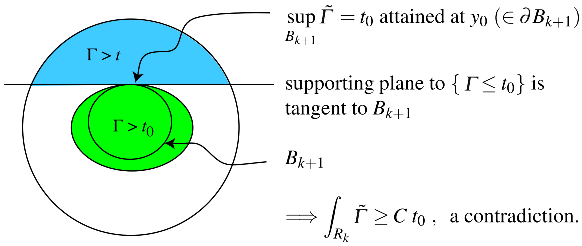
Of course, $\tilde{w}$ may still be highly oscillatory but

(a) In 99% of the rings, $\tilde{w} \leq 100 C_1 r_k^2$, that is, in the original configuration w stays close to its convex envelope Γ.
(b) Further, since $0 \leq \tilde{\Gamma} \leq \tilde{w}$

$$\fint_{\mathbb{R}_k} \tilde{\Gamma} \leq C_1 r_k^2.$$

But $\tilde{\Gamma}$ is convex, so this implies a bound $\tilde{\Gamma} \leq C_2 r_k^2$ in B_{k+1} and $\nabla \Gamma \leq r_k$ in B_{k+2}.

Let $t_0 = r_k^2$

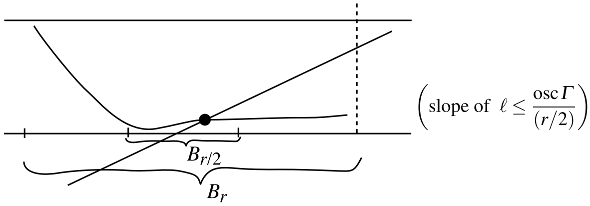


In turn, this implies that $\sup_{B_{k+2}} \nabla \tilde{\Gamma} < C\, r_k^2$.

Finally, it is a general fact that

If Γ is convex in B_r,

$$\nabla\Gamma\Big|_{B_{r/2}} \leq \left(\frac{\operatorname{osc}\Gamma}{B_r}\right) 2r^{-1}$$



A covering lemma completes the proof.

This proof, of course, requires in principle that the kernels be symmetric (some asymmetry is "tolerated" by the fact that the gradient of Γ is bounded, as in the second order case).

But the nature of the "game" symmetrizes the kernel:

From the "inf sup" property, for any x_0, there exists a direction ν^+ so that

$$0 \leq \int [u(x+y)+u(x-y)-2(u(x))]K_{\nu^+}+K_\mu$$

for any μ (in particular $-\nu^+$) and vice versa, a ν^- so that

$$0 \geq \int [u(x+y)+u(x-y)-2u(x)]K_{\nu^-}+K_\mu$$

for any μ, and this property is all that's needed.

Going back to the two possible "integral drifts", in the first case it is also possible to prove that solutions are in fact $C^{2s+\sigma}$, i.e., the integrals converge and the solution is classical (see [7]).

This is because the nature of the drift is such that, as the problem is rescaled the perturbation term $A(x_1)$ drifts to infinity.

6 Non-local Infinite Laplacian

Finally, I would like to discuss briefly the "tug of war" non-local "infinite Laplacian".

The Infinite Laplacian appears in the case when there is no diffusion left, i.e., when it is just the players taking random turns in choosing the direction of the drift (tug of war).

For the infinitesimal case, when the length of the jump is predetermined you formally get "$u_{nn}=0$", n the direction of the gradient (Peres–Schramm–Sheffield–Wilson [17]).

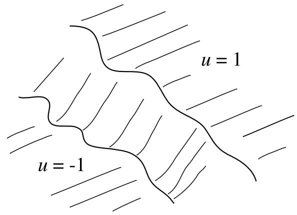
Also, in collaboration with Bjorland and Figalli [2], we consider the case in which the jump of the particle follows the distribution of the s-Laplacian

$$\inf_{\nu_1 \in s^1} \sup_{\nu_2 \in s^2} \int \frac{u(x+\nu_1 t) + u(x+\nu_2 t) - 2u(x)}{t^{1+2s}}\, dt = 0.$$

(That is, each player pulls in the directions ν_1 and ν_2.)

Formally, for $s > 1/2$, the direction of the jump is given by ∇u: Since the integrals diverge, each players is "forced" to take that choice.

We prove existence, uniqueness and (some) regularity, under a monotone geometry, for $s > 1/2$.

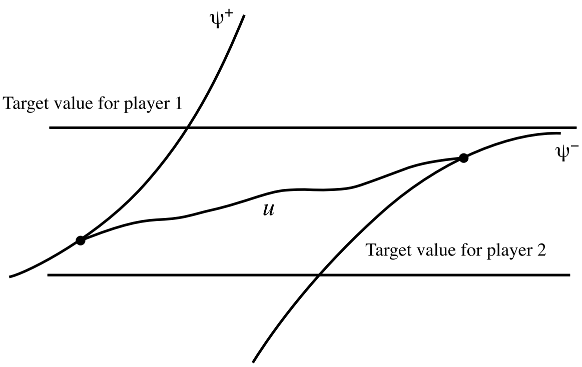


We assume that the domain Ω is "strip like", i.e., between bounded Lipschitz graphs with uniform separation and pay off is respectively 1 and -1. We show that there exists a unique viscosity solution (the least supersolution and larger subsolution coincide), and it is C^{2s-1}.

($|x|^{2s-1}$ are the "cones" for this problem, note that for $s = 1/2$ the theory breaks down.)

We end up with some comments:

- The case $s < 1/2$ seems very interesting since "∇u" does not fix the direction of the jump any more, and players will choose to jump in "non-opposite directions" most of the time.
- Instead of prescribing boundary values, it seems more natural to prescribe upper and lower obstacles where it would be optimal for one of the players to stop playing (execute an option).



We can prove in this case similar results as to the boundary value problem discussed before [2].

Acknowledgements The author was partially supported by National Science Foundation Grant DMS-0654267.

References

1. Barlow, M.T., Bass, R.F., Chen, Zh.Q., Kassmann, M.: Non-local Dirichlet forms and symmetric jump processes. Trans. Am. Math. Soc. **361**(4), 1963–1999 (2009)
2. Bjorland, C., Caffarelli, L., Figalli, A.: Non-local tug-of-war and the infinity fractional Laplacian. arXiv:1011.1966
3. Bass, R.F., Kassmann, M.: Harnack inequalities for non-local operators of variable order. Trans. Am. Math. Soc. **357**(2), 837–850 (2005)
4. Caffarelli, L., Cabre, X.: Fully Nonlinear Elliptic Equations. Am. Math. Soc. Colloq. Publ., vol. 43. Am. Math. Soc., Providence (1995)
5. Caffarelli, L., Vasseur, A.: Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. Ann. Math. (2011, in press)
6. Caffarelli, L.A., Silvestre, L.: Regularity theory for fully nonlinear integro-differential equations. Commun. Pure Appl. Math. **62**(5), 597–638 (2009)
7. Caffarelli, L.A., Silvestre, L.: Regularity results for nonlocal equations by approximation. Arch. Ration. Mech. Anal. (2011, in press). arXiv:0902.4030
8. Caffarelli, L.A., Silvestre, L.: The Evans-Krylov theorem for non local fully non linear equations. Ann. Math. (2011, in press). arXiv:0905.1339
9. Caffarelli, L.A., Chan, C.H., Vasseur, A.: Regularity theory for nonlinear integral operators. arXiv:1003.1699
10. Chambolle, A., Lindgren, E., Monneau, R.: The Hölder infinite Laplacian and Hölder extensions. Preprint (2010)
11. Duvant, G., Lions, J.-L.: Les inéquations en mécanique et en physique. Travaux et Recherches Mathématiques, vol. 21. Dunod, Paris (1972) (in French)
12. Giacomin, G., Lebowitz, J.L., Presutti, E.: Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems. In: Stochastic Partial Differential Equations: Six Perspectives. Math. Surveys Monogr., vol. 64, pp. 107–152. Am. Math. Soc., Providence (1999)
13. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. Multiscale Model. Simul. **7**(3), 1005–1028 (2008)
14. Kassmann, M.: A priori estimates for integro-differential operators with measurable kernels. Calc. Var. Partial Differ. Equ. **34**(1), 1–21 (2009)
15. Komatsu, T.: Uniform estimates for fundamental solutions associated with non-local Dirichlet forms. Osaka J. Math. **32**(4), 833–860 (1995)
16. Peres, Y., Sheffield, S.: Tug-of-war with noise: a game-theoretic view of the p-Laplacian. Duke Math. J. **145**(1), 91–120 (2008)
17. Peres, Y., Schramm, O., Sheffield, S., Wilson, D.B.: Tug-of-war and the infinity Laplacian. J. Am. Math. Soc. **22**(1), 167–210 (2009)