SDEs with Jumps
Notes for Barrett Lectures, April 2009

Richard F. Bass

April 21, 2009

1 Introduction

This will be a course about stochastic differential equations with jumps.

Let me give just one example where this is important. In financial mathematics it is common to model the price of a stock by a geometric Brownian motion. However, this model suffers from some significant defects. One of the most important is that although the graph of a stock price looks a lot like the graph of a Brownian motion, the stock price will at times decrease or increase faster than a Brownian motion can. For example, if you look at the graph of a stock price on September 11, there is a sudden decrease, more than can be followed by a Brownian motion. Since one of the major goals of financial math is to determine option prices for stocks, it is crucial to have a good model for the stock price. Experimentally the model that best fits the data is a geometric Brownian motion with jumps at random times. The jumps can be caused by disasters, wars, discoveries, etc.

Processes with jumps are also important for some physical models, but I don't know too much about these.

Despite the importance of allowing jumps in the model, a lot of the basic theory of SDEs with jumps is still not very well understood. The purpose of this course is to give some ideas for the parts that are understood.

There are lots of possibilities for research here. Take your favorite result for diffusions or Brownian motion, and try to extend them to jump processes. Some of the results will be interesting and some will not, but there will be more interesting results than uninteresting ones.

As we go along, I will present lots of open problems.

My goal is to give the main ideas of the proofs and not burden you too much with ε's and δ's or integrability considerations. To get the details, you'll have to look at the references. A lot of the references can be found on my web page,

www.math.uconn.edu/∼bass

and then click on "Lecture notes" or "Bibliography," depending on what you are looking for.

1.1 Outline

I will select from the following.

1. Stochastic calculus. I will try to keep this relatively brief, as this is old stuff.

2. Pathwise uniqueness for SDEs and the lack thereof.

3. Potential theory for symmetric stable processes.
4. Weak uniqueness for SDEs with jumps.
5. Harmonic functions. The goal is to prove Harnack inequalities.
6. Symmetric jump processes. Here we obtain transition density estimates.

The last two are currently hot areas.

2 Stochastic calculus

The references here include the course by Meyer (in French) [25] in 1976, a book by Elliott [20] and my web page: *Stochastic calculus for discontinuous processes* in the Lecture notes section.

We'll talk about predictability, decomposition of martingales, the square brackets process, stochastic integrals, Ito's formula, Poisson point processes, and Lévy processes. As I said, this is old stuff, and I just want to give the main differences from the continuous case.

2.1 Notation

We let $\{\mathcal{F}_t\}_{\{t\geq 0\}}$ be a right continuous filtration that is complete with respect to some probability measure. Most of our processes will have paths that are right continuous with left limits (rcll), and in that case we let

$$X_{t-} = \lim_{s\uparrow t} X_s, \qquad \Delta X_t = X_t - X_{t-}.$$

2.2 Predictability

A stochastic process $X_t = X_t(\omega) = X(t,\omega)$ is a map from $[0,\infty) \times \Omega$ into (usually) $\mathbb{R}$. We define a σ-field, called the predictable σ-field or previsible σ-field on $[0,\infty) \times \Omega$ as the σ-field $\mathcal{P}$ generated by all processes of the form

$$H(s,\omega) = K(\omega) 1_{(a,b]}(s),$$

where K is bounded and $\mathcal{F}_{a-}$ measurable and $\mathcal{F}_{a-} = \bigvee_{s<a} \mathcal{F}_s$. A process adapted to $\mathcal{P}$ is called predictable.

It is easy to see that if H is left continuous, then it is predictable. In particular, if X is rcll, then X_{t-} is predictable.

Why do we care? When defining the stochastic integral with respect to a martingale with jumps, we need the integrand to be predictable in order to get a martingale out. It is also natural from the stock price point of view. If S_t is the stock price at time t and H_t is the number of shares held at time t, then

$$\int_0^t H_s \, dS_s$$

is the net profit. If S has jumps and we don't require H to be predictable, one can fix H_t to depend on ΔS_t, and get a guaranteed profit, without risk. That's not supposed to happen, and the number of shares held at time t should only depend on what we know up to time t, but not necessarily what happens at jump times.

2.3 Decomposition of martingales

Suppose M_t is a rcll square integrable martingale. Since it is rcll, there can only be finitely many jumps of size larger than

$1/k$ in absolute value for each positive integer k. Thus there are only countably many jumps for M_t. Let's list them as occurring at times $T_1, T_2, \ldots$. We certainly cannot assume that $T_1 < T_2 < \cdots$; we are just listing them. Let us also arrange it so that for each i, ΔM_{T_i} takes values in some $[a_i, b_i]$ where the interval does not include 0.

If we let

$$A_i(t) = \Delta M_{T_i} 1_{(t \geq T_i)},$$

then A_i is an increasing (or decreasing) process, hence a submartingale (or supermartingale). By the Doob-Meyer decomposition there exists $\widetilde{A}_i$ predictable and increasing (or decreasing) such that

$$M_i(t) = A_i(t) - \widetilde{A}_i(t)$$

is a martingale. Then

Theorem 2.1 *We can write*

$$M_t = \sum_{i=1}^{\infty} M_i(t) + M^c(t),$$

where M^c is a continuous martingale. The convergence is in L^2.

To give the idea, let us just show the orthogonality of M_i and M_j: We have

$$\begin{aligned}
\mathbb{E}\, M_i(\infty) M_j(\infty) &= \mathbb{E} \int_0^\infty M_i(\infty)\, dM_j(s) \qquad \text{(Lebesgue-Stieltjes integral)} \\
&= \mathbb{E} \int_0^\infty M_i(s)\, dM_j(s) \qquad \text{(condition on } \mathcal{F}_s\text{)} \\
&= \mathbb{E} \int_0^\infty \Delta M_i(s)\, dM_j(s) + \mathbb{E} \int_0^\infty M_i(s-)\, dM_j(s).
\end{aligned}$$

The first term on the last line is 0 because M_i and M_j do not jump at the same times. The second term is 0 because $M_i(s-)$ is predictable: even though we are doing Lebesgue-Stieltjes integrals here, the expectation of $\int_0^t H_s\,dM_j(s)$ with $H_s = K(\omega)1_{(a,b]}(s)$ is 0, and then we build up from this to get it is 0 for every predictable process (subject to some integrability conditions).

2.4 Square brackets

We define

$$[M] = [M, M] = \langle M^c, M^c\rangle + \sum_{s\le t} (\Delta M_s)^2.$$

We define $[M, N]$ by polarization.

Even for discontinuous martingales, one can define $\langle M, M\rangle$ as the predictable increasing part of the submartingale M_t^2. So why do we introduce $[M]$? It turns out it shows up in many places, e.g., the integration by parts formula, the Burkholder-Davis-Gundy inequalities, to mention just two.

Lemma 2.2 *$M_t^2 - [M]_t$ is a martingale.*

Proof. By orthogonality, it is enough to show $M_i(t)^2 - (\Delta M_i(T_i \wedge t))^2$ is a martingale. We can rewrite this expression as

$$2\int_0^t M_i(s-)\,dM_i(s),$$

where we view this as a Lebesgue-Stieltjes integral. But

$$\int_0^t H_s\,dM_i(s)$$

is a martingale for every predictable process H, modulo some integrability: first prove it for the simplest predictable processes $K1_{(a,b]}$, and then use a monotone class argument. $\square$

2.5 Stochastic integrals

This is almost identical to the construction for continuous martingales. If $H_s = K1_{(a,b]}(s)$ with K bounded and $\mathcal{F}_{a-}$ measurable, define

$$\int_0^t H_s \, dM_s = K(M_{t\wedge b} - M_{t\wedge a}).$$

This is a martingale whose square bracket is given by

$$\int_0^t H_s^2 \, d[M]_s.$$

If we have a linear combination of such integrands, we define the stochastic integral by linearity. And if H is predictable with $\mathbb{E} \int_0^t H_s^2 \, d[M]_s$, we approximate H by such linear combinations in L^2 (with norm $\|H\|^2 = \mathbb{E} \int_0^t H_s^2 \, d[M]_s$) and take limits.

We extend this definition to semimartingales $X_t = M_t + A_t$ where M is square integrable and A_t has bounded variation in the usual way.

By a stopping times argument, one can extend the definition to semimartingales where M is a local martingale and A has locally finite bounded variation. Because we may have jumps of X of arbitrarily large size, the arguments are much more complicated than in the continuous case. Some of the phraseology used is that certain stopping times "reduce" the semimartingale. But the arguments are fairly technical and not that interesting (to me).

2.6 Ito's formula

For semimartingales with jumps, Ito's formula goes as follows. If $f \in C^2$, then

$$f(X_t) = f(X_0) + \int_0^t f(X_{s-})\, dX_s + \tfrac{1}{2} \int_0^t f''(X_s)\, d\langle X^c \rangle_s \\ + \sum_{s \le t} [f(X_s) - f(X_{s-}) - f'(X_{s-}) \Delta X_s].$$

Let's make a few comments. First of all, the integrand in front of the dX_s is $f(X_{s-})$. Since X_{s-} is left continuous, it is predictable, so the integrand is predictable. This might not be the case if we had $f(X_s)$ as the integrand. Secondly, $\langle X^c \rangle$ is defined to be $\langle M^c \rangle$. Finally, if f and its first two derivatives are bounded, the sum on the last line is bounded by

$$\|f''\|_\infty \sum_{s \le t} (\Delta X_s)^2.$$

So, at least for reasonable semimartingales, the sum converges.

By applying this to $(X+Y)^2$, X^2, and Y^2, we get the product formula, also known as the integration by parts by formula.

$$X_t Y_t = X_0 Y_0 + \int_0^t X_{s-}\, dY_s + \int_0^t Y_{s-}\, dX_s + [X, Y]_t,$$

or

$$d(XY) = X\, dY + Y\, dX + d[X, Y].$$

For a semimartingale X, we let $[X, X] = [M, M]$. (There's always a nuisance here – one might be able to write $X = M + A$ in two different ways – but there is a canonical way if one requires A to be predictable.

The idea of the proof of Ito's formula is this. By taking limits, it suffices to assume there are only finitely many jumps. By looking at them sequentially, it suffices to assume there is only one jump, say, at time T. Apply the Ito formula for continuous processes to $X_t - \Delta X_T 1_{(t \geq T)}$, and then notice that the jump at time T contributes the same amount to both sides.

2.7 Poisson point processes

Let $\mathcal{S}$ be some state space, $\mathbb{S}$ a σ-field on $\mathbb{S}$, and λ an infinite measure on $(\mathcal{S}, \mathbb{S})$. Let $\mathcal{B}$ be the Borel σ-field on $[0, \infty)$ and

$$\mu : \Omega \times \mathcal{B} \times \mathbb{S} \to \mathbb{R}.$$

μ is a Poisson point process if for each $A \in \mathbb{S}$, $\mu([0,t] \times A)$ is a Poisson process with parameter $\lambda(A)$ and if $A_1, \ldots, A_n$ are disjoint with $\lambda(A_i) < \infty$ for each i, then $\mu([0,t] \times A_i)$ are independent processes. (We drop the ω from the notation as usual.)

Let $\nu([0,t] \times A) = t\lambda(A)$, which is nonrandom. Then for each A with $\lambda(A) < \infty$, we have that $(\mu - \nu)([0,t] \times A)$ is a martingale.

An example of a Poisson point process is this: if X_t is a Lévy process, that is, a rcll process with stationary, independent increments, then $\mu([0,t] \times A) = \sum_{s \leq t} 1_A(\Delta X_s)$ is a Poisson point process.

We can define stochastic integrals with respect to Poisson point processes more or less in the way stochastic integrals are defined with respect to martingales. If $H(\omega, s, z) = K(\omega) 1_{(a,b]}(s) 1_A(z)$, where K is bounded and $\mathcal{F}_{a-}$ measurable and $\lambda(A) < \infty$, we define

$$\int_0^t \int_{\mathcal{S}} H(s,z)\,(\mu-\nu)(ds\,dz) = K(\omega)(\mu-\nu)(([a,b]\cap[0,t])\times A).$$

We then use linearity and L^2 limits to define stochastic integrals for more general integrands. It is not hard to see that if

$$N_t = \int_0^t \int H(s,z)\,(\mu-\nu)(ds\,dz),$$

then

$$\langle N \rangle_t = \int_0^t \int H(s,z)^2\,\nu(ds\,dz) \qquad [N]_t = \int_0^t \int H(s,z)^2\,\mu(ds\,dz).$$

A very general SDE would be one like

$$dX_t = \sigma(X_t)\,dW_t + b(X_t)\,dt + \int_{\mathcal{S}} F(X_{t-},z)\,(\mu-\nu)(ds\,dz), \qquad X_0 = x_0. \tag{2.1}$$

We need such a general SDE to capture the behavior of some financial models, for instance. The idea is if μ assigns a point mass to (X_{s-},z), then X_t jumps an amount $F(X_{s-},z)$. We can get even more general by having a term which is an integral against $d\mu$ to take care of large jumps, provided some integrability conditions are satisfied.

2.8 Lévy processes

We already defined a Lévy process and how it is related to Poisson point processes. A Lévy process has Lévy measure n if $\lambda = n$ and $\int (1 \wedge |x|^2)\,n(dx) < \infty$.

Let us give a few examples of Lévy processes that have no continuous parts. (We can always add a constant multiple of Brownian motion or a constant times t.)

1. Let P_t^i be Poisson processes with intensities λ_i, $i = 1, 2, \ldots$; in particular P_t^i is Poisson with parameter $\lambda_i t$. If $\sum_{i=1}^\infty a_i \lambda_i < \infty$, then $\sum_{i=1}^\infty a_i P_t^i$ is a Lévy process. These are compound point processes. The Lévy measure is $n(dx) = \sum \lambda_i \delta_{a_i}(dx)$, where δ_a is point mass at a. One can show that if $\sum_i \lambda_i = \infty$, then the set of jump times, although countable, is dense in $[0, \infty)$.

2. Let P_t^i and λ_i be as in the first example, but now suppose $\sum_{i=1}^\infty a_i^2 \lambda_i < \infty$. Then $\sum_{i=1}^\infty a_i (P_t^i - \lambda_i t)$ is a Lévy process. This allows the a_i, which are small, to be larger than in the first example. The Lévy measure is the same as in the first example.

3. X_t is a symmetric stable process of order α if

$$n(dx) = \frac{c}{|x|^{1+\alpha}}\, dx.$$

The symmetric stable process of order α in d dimensions is the same, except with $1 + \alpha$ replaced by $d + \alpha$.

It will be useful for later on to apply Ito's formula to a Lévy process. If X_t has no continuous part and is itself a martingale, we can write

$$X_t = X_0 + \int_0^t \int z\, (\mu - \nu)(ds\, dz) = X_0 + \int_0^t \int z\, (\mu(ds\, dz) - n(dz)\, ds).$$

If $f \in C^2$, we have

$$
\begin{aligned}
f(X_t) - f(X_0) &= \int_0^t f'(X_{s-})\,dX_s + \sum_{s\le t}[f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s] \\
&= \text{ martingale } + \int_0^t \int [f(X_{s-}+z) - f(X_{s-}) - f'(X_{s-})z]\,\mu(ds\,dz) \\
&= \text{ martingale } + \int_0^t \int [f(X_{s-}+z) - f(X_{s-}) - f'(X_{s-})z]\,n(dz)\,ds \\
&= \text{ martingale } + \int_0^t \mathcal{L}f(X_{s-})\,ds \\
&= \text{ martingale } + \int_0^t \mathcal{L}f(X_s)\,ds,
\end{aligned}
$$

where

$$
\mathcal{L}f(x) = \int [f(x+z) - f(x) - f'(x)z]\,n(dz).
$$

The way to read the last expression is that $f(x+z) - f(x)$ represents a jump of size z with intensity $n(dz)$, while the $f'(x)z$ represents a drift that compensates for the jumps.

3 Pathwise uniqueness

3.1 SDEs with Poisson point processes

Look at (2.1) with the Brownian motion and drift terms removed. (There is no harm doing something similar with them included, but for simplicity let's suppose they are not there.) A result due to Skorokhod is the following.

Theorem 3.1 *Suppose F is bounded and*

$$\int |F(x,z)-F(y,z)|^2\,\lambda(dz) \le c|x-y|^2.$$

Then the solution to (2.1) exists and pathwise uniqueness holds.

Proof. To do uniqueness, if X and Y are two solutions with the same starting points,

$$\begin{aligned}
\mathbb{E}\,(X_t-Y_t)^2 &= \mathbb{E}\Big[\int_0^t\int [F(X_{s-},z)-F(Y_{s-},z)]\,(\mu-\nu)(ds\,dz)\Big]^2\\
&= \mathbb{E}\int_0^t\int |F(X_{s-},z)-F(Y_{s-},z)|^2\,\mu(ds\,dz)\\
&= \mathbb{E}\int_0^t\int |F(X_{s-},z)-F(Y_{s-},z)|^2\,\lambda(dz)\,ds\\
&\le c\mathbb{E}\int_0^t |X_{s-}-Y_{s-}|^2\,ds\\
&= c\mathbb{E}\int_0^t |X_s-Y_s|^2\,ds.
\end{aligned}$$

The second equality is the expression for $\mathbb{E}\,[X-Y]_t$, the last since X and Y have only countably many jumps. Using Gronwall's inequality proves uniqueness.

Existence is the standard Picard iteration argument: let $X_0 \equiv x_0$, let

$$X_t^{n+1} = x_0 + \int_0^t F(X_{s-}^n, z)\,(\mu-\nu)(ds\,dz),$$

and prove that the sequence X^n converges.

3.2 SDEs with stable processes

For one-dimensional Brownian motion, one can do much better than Lipschitz continuity. The Yamada-Watanabe condition says pathwise uniqueness holds to $dX_t = \sigma(X_t)\, dW_t$ if $|\sigma(x) - \sigma(y)| \le \rho(|x - y|)$ for some function ρ satisfying

$$\int_{0+} \frac{1}{\rho(x)^2}\, dx = \infty.$$

There exists an analogue for stable processes of index between 1 and 2. (References: [6, Section 2] and [23].)

Theorem 3.2 *Suppose Z_t is a stable symmetric process of order α with $1 < \alpha < 2$. Suppose ρ is a nondecreasing continuous function on $[0, \infty)$ with*

$$\int_{0+} \frac{1}{\rho(x)^\alpha}\, dx = \infty.$$

Suppose F is bounded and satisfies $|F(x) - F(y)| \le \rho(|x - y|)$ for all x, y. Then the solution to

$$dX_t = F(X_{t-})\, dZ_t$$

is pathwise unique.

Before proving this, a few remarks. First, pathwise uniqueness automatically implies strong existence and weak uniqueness; the argument for Brownian motion goes through. Second, the proofs I know for Brownian motion do not go through for the stable case with $\alpha < 2$. Third, if $|F(x) - F(y)| \le c|x - y|^\beta$, i.e., F is Hölder of order β, then we need

$$\int_{0+} \frac{1}{(|x|^\beta)^\alpha}\, dx = \infty,$$

or $\alpha\beta \geq 1$, or $\beta \geq 1/\alpha$. This agrees with the Brownian motion case: $\beta \geq 1/2$. For stable α processes, we see we need more regularity (smoother F) as $\alpha \downarrow 1$. When $\alpha = 1$, our condition says we almost need Lipschitz continuous F.

Proof. Let X^1 and X^2 be two solutions, $Y = X^1 - X^2$, $H_t = F(X^1_{t-}) - F(X^2_{t-})$. So

$$Y_t = \int_0^t H_s \, dZ_s.$$

Let $a_n \downarrow 0$ such that

$$\int_{a_{n+1}}^{a_n} \frac{1}{\rho(x)^\alpha} \, dx = n.$$

Let h_n be nonnegative, C^2 with support in $[a_{n+1}, a_n]$, with $\int h_n = 1$ and $h_n(x) \leq 2/n\rho(x)^\alpha$. So far, the proof is exactly similar to the Brownian motion case.

Let $p_t(x, 0)$ be the transition densities for a symmetric stable process in one dimension,

$$g_\lambda(x) = \int_0^\infty e^{-\lambda t} p_t(x, 0) \, dt, \qquad G_\lambda f(x) = \int f(y) g_\lambda(x - y) \, dy.$$

The Fourier transform of $p_t(x, 0)$ is $e^{-t|u|^\alpha}$, hence the Fourier transform of g_λ is $1/(\lambda + |u|^\alpha)$, and it follows that g_λ is bounded and continuous, and $g_\lambda(x) < g_\lambda(0)$ if $x \neq 0$.

Let $f_n = G_\lambda h_n$, which is in C^2, and let $A_t = \int_0^t |H_s|^\alpha \, ds$. If one goes through a calculation very similar to the one where we calculated the generator for symmetric stable processes, one sees that

$$f(Y_t) = f(Y_0) + \text{ martingale } + \int_0^t |H_s|^\alpha \mathcal{L} f(Y_{s-})\, ds.$$

By the product formula, using the fact that A is continuous and of bounded variation,

$$\begin{aligned}\mathbb{E}\, e^{-\lambda A_t} f_n(Z_t) - f_n(0) &= \mathbb{E} \int_0^t e^{-\lambda A_s}\, d[f_n(Y_s)] - \mathbb{E} \int_0^t \lambda |H_s|^\alpha f_n(Y_{s-})\, ds \\ &= \mathbb{E} \int_0^t e^{-\lambda A_s} |H_s|^\alpha \mathcal{L} f_n(Y_{s-})\, ds - \mathbb{E} \int_0^t \lambda |H_s|^\alpha f_n(Y_{s-})\, ds.\end{aligned}$$

$f_n = G_\lambda h_n$, so $\mathcal{L} f_n = \lambda G_\lambda h_n - h_n = \lambda f_n - h_n$. Substituting, we obtain

$$\mathbb{E}\, e^{-\lambda A_t} f_n(Y_t) - f_n(0) = - \int_0^t e^{-\lambda A_t} |H_s|^\alpha h_n(Y_{s-})\, ds.$$

The last term is bounded in absolute value by

$$\mathbb{E} \int_0^t \frac{1}{\rho(Y_{s-})^\alpha} h_n(Y_{s-})\, ds \le t \frac{2}{n},$$

which tends to 0 as $n \to \infty$. $h_n(x)\, dx$ tends weakly to point mass at 0, so $f_n = G_\lambda h_n \to g_\lambda$. Therefore

$$g_\lambda(0) - \mathbb{E}\, e^{-\lambda A_t} g_\lambda(Y_t) = 0.$$

Since $g_\lambda(x) < g_\lambda(0)$ if $x \neq 0$, this implies $Y_t = 0$ a.s. (and $A_t = 0$, a.s.) This is true for each rational t, and since Y is rcll, that does it. □

When $\alpha \le 1$, then g_λ is no longer bounded, and the proof breaks down.

The analogue of the Yamada-Watanabe condition is sharp: if the integral is finite, one lets $F(x) = \rho(|x|)$, and one gets two solutions, the identically zero one, and a time change of a symmetric stable process as the other. The proofs are very similar to the Brownian motion case.

It is now known (despite what I wrote in one of my papers; ignore Section 3 of [6]), that when $\alpha \leq 1$ and $\beta < 1$, there are functions F that are bounded above and below by positive constants that are Hölder continuous of order β and for which pathwise uniqueness for the SDE

$$dX_t = F(X_{t-})\, dZ_t$$

fails. More about that soon.

The above theorem does raise some questions. If σ is bounded above and below by positive constants, then the SDE

$$dX_t = \sigma(X_t)\, dW_t$$

has a pathwise unique solution if σ is (locally) of bounded variation (this is due to Nakao), and even if σ has locally finite quadratic variation (due to LeGall). Is there any analogue of this for stable processes with $\alpha > 1$? Komatsu [23] states a theorem analogous to that of Nakao, but there is an error in his proof.

Another question: What about pathwise uniqueness for solutions to (2.1) when F is not Lipschitz?

One can establish pathwise uniqueness for the SDE

$$dX_t = dW_t + f(X_t)\, dt$$

when f is in L^1, or even more singular (e.g., a constant times the

δ function provided the constant is strictly less than 1). What about if W_t is replaced by a stable process?

If Z_t (not necessarily stable here) has only finitely jumps in each finite interval, pathwise uniqueness is almost trivial. What if Z has countably many jumps in each finite interval, but Z_t is increasing?

Although the remark above shows that the analogue of the Yamada-Watanabe condition is sharp, can one get a better theorem on pathwise uniqueness if we assume F is bounded above and below by positive constants? The answer is, essentially no; [8].

Theorem 3.3 *If $\beta < \frac{1}{\alpha} \wedge 1$, then there exists F that is bounded above and below by positive constants such that F is Hölder continuous of order β, but pathwise uniqueness does not hold for the SDE*

$$dX_t = F(X_{t-})\, dZ_t.$$

Notice that if $\alpha < 1$, we can take β up to but not including the value 1.

Let us give the idea of the proof. First, let Z and Z' be independent stable processes of order α, and let

$$dX_t = F(X_{t-})\, dZ_t,$$
$$dX'_t(\varepsilon) = F(X'_{t-}(\varepsilon))\, dZ_t + \varepsilon\, dZ'_t,$$

and

$$Y(\varepsilon) = X'(\varepsilon) - X.$$

As $\varepsilon \to 0$, one can show that $(X'(\varepsilon), Y(\varepsilon), Z)$ converges (weakly), and the limit Y solves

$$dY_t = [F(X_{t-} + Y_{t-}) - F(X_{t-})]\, dZ_t,$$

and then $X+Y$ is another solution to our SDE. If $T_b^\varepsilon = \inf\{t : |Y_t(\varepsilon)| \geq b\}$ and we show that $\mathbb{E}\, T_b^\varepsilon$ is bounded uniformly in ε, say, by c, then

$$\mathbb{P}(T_b^\varepsilon \geq 2c) \leq \frac{\mathbb{E}\,(T_b^\varepsilon)}{2c} \leq \frac{1}{2},$$

or with probability at least $\frac{1}{2}$, $T_b^\varepsilon \leq 2c$, or $|Y(\varepsilon)_t| \geq b$ for some $t \leq 2c$ with probability at least $\frac{1}{2}$. This is preserved in the limit, so Y_t is not identically 0. Therefore it suffices to get a bound on $\mathbb{E}\, T_b^\varepsilon$. Let us now drop the ε's from the notation.

Let $I_k = [2^{-k}, 2^{-k+1}]$ and $I_k^* = [2^{-k-1}, 2^{-k+2}]$. We get an estimate of the amount of time spent in I_k by Y_t up to time T_b, and then sum over k. We do the same for $-I_k$, show the amount of time spent at 0 is 0, and we get our bound. By the strong Markov property, the expected amount of time in I_k up to time T_b is bounded by

$$\begin{aligned}&[\text{expected number of times } Y_t \text{ goes from } I_k \text{ to } I_k^* \text{ before } T_b]\\ &\qquad \times [\text{maximum of the expected time for } Y_t \text{ to exit } I_k^* \text{ starting in } I_k].\end{aligned} \tag{3.1}$$

Y_t is a time change of a stable process W_t, so it is enough to estimate the number of crossings of a stable process. If R_i, S_i are the successive times the process enters I_k and leaves I_k^* and N_k is the number of crossings,

$$\begin{aligned}\mathbb{E} \int_0^{T_b} 1_{I_k^*}(W_s)\, ds &\geq \sum_i \mathbb{E}\,(S_i \wedge T_b - R_i \wedge T_b)\\ &\geq \sum_i \mathbb{E}\left[\mathbb{E}^{\,W_{R_i}} S_1; R_i < T_b\right]\\ &\geq [\mathbb{E}\, N_k - 1]\, [\inf_{x \in I_k} \mathbb{E}^{\,x} S_1].\end{aligned}$$

The first term on the left can be estimated by the Green function for W (which is bounded if $\alpha > 1$, is of order $-\log|x-y|$ if $\alpha = 1$, and is of order $|x-y|^{\alpha-1}$ if $\alpha < 1$), and similarly the last factor on the right in the last line. We end up with

$$\mathbb{E}\, N_k \le \begin{cases} c2^k(\alpha-1), & \alpha > 1, \\ ck, & \alpha = 1, \\ c, & \alpha < 1. \end{cases} \tag{3.2}$$

The different behavior for different α is what gives the constraints on β.

Next we construct F. Let F_0 be a sawtooth function with the width of the teeth equal to 1 and the height 1. Let F_k be a sawtooth function with the width of the teeth equal to 2^{-k} and the height 2^{-k}, and then let

$$F = \sum_k 2^{-\beta' k} F_k,$$

where β' is some number slightly larger than β. It is not hard to show F is Hölder continuous of order β.

Let's now look at the second factor on the last line of (3.1). If $|F(x+y) - F(x)|$ was bounded below for all x and y, life would be easier, but this isn't possible to do. We can show, though, that if

$$A_k(\theta) = \{x : |F(x+y) - F(x)| \ge \theta 2^{-k\beta} \text{ for all } y \in I_k^*\},$$

then the Lebesgue measure of $A_k(\theta)$ is at least a fixed proportion of the length of I_k. So since F is bounded below, a time change argument shows that the solution to $dY_t = [F(X_{t-} - Y_{t-}) - F(X_{t-})]\, dZ_t$ spends at least a certain amount of time in $A_k(\theta)$.

But if the process spends enough time in $A_k(\theta)$, then there is a reasonable chance of having a jump larger than 2^{-k+2} in size, which means the process exits I_k^* before too long. This leads to an upper bound on the exit time.

Beyond the above results, there are some results on weak existence, but not too much else. It would be nice to also have some results where the driving process is not necessarily of fixed order.

4 Potential theory

We look next at some potential theory for functions that are α-harmonic. Let u be defined on all of $\mathbb{R}^d$ with some integrability conditions. We say u is α-harmonic in a domain D if $u(x) = \mathbb{E}^x u(X_{\tau_B})$ for every $x \in B$ and every ball $B \subset \overline{B} \subset D$, where X_t is a d-dimensional symmetric stable process and τ_B is the time of the first exit from B.

We say u is regular α-harmonic in D if $u(x) = \mathbb{E}^x u(X_{\tau_D})$ for all $x \in D$. We will give an example of an α-harmonic function that is not regular in a moment.

One can explicitly calculate the "Poisson kernel" for a ball B: that is, find $k(x, y)$ such that

$$\mathbb{P}^x(X_{\tau_B} \in A) = \int_A k(x, y)\, dy$$

for every $A \subset B^c$. This is done in the same way as for Brownian motion, by using a Kelvin transformation, i.e., inversion in a sphere. Then if we fix $y \in B^c$ and let $u(x) = k(x, y)$, we have a function that is α-harmonic in B but not regular α-harmonic.

Using the explicit formula for $k(x, y)$, we can prove the Harnack inequality just as in the Brownian motion case.

Symmetric stable processes of order α have a scaling property. To see this, look at the generator. A time change multiplies the generator by a scalar, while a dilation does a change of variables. There is the issue of

$$\int [\ldots + 1_{(|h| \le r)} \nabla f(x) \cdot h] \frac{1}{|h|^{d+\alpha}} \, dh$$

compared with the same thing with r replaced by 1, but by symmetry,

$$\int h \, 1_{(r \le |h| \le 1)} \frac{1}{|h|^{d+\alpha}} \, dh = 0.$$

4.1 Boundary Harnack principle

Let's first talk about what the boundary Harnack principle (BHP) says for Brownian motion. If we have a non-negative harmonic function u in a domain D, then the Harnack inequality says that the values of u are comparable in a ball away from the boundary. In general, we cannot say much when we are in balls very close to the boundary, that is, the constant of comparability gets worse. If, however, the boundary values of u are zero on a part of the boundary and the domain is reasonably regular, then not only does $u(x) \to 0$ as x tends to that part of the boundary, but all such non-negative harmonic functions decay at the same rate. More specifically,

Theorem 4.1 *Let D be a reasonable domain, $z_0 \in \partial D$, u, v continuous on $\overline{D}$, u, v non-negative and harmonic in D, and for* *some r, u and v are 0 on $\partial D \cap B(z_0, 2r)$. Let $x_0 \in D \cap B(z_0, r)$. Then there exists a constant c such that*

$$\frac{u(x)}{v(x)} \le c\frac{u(x_0)}{v(x_0)}, \qquad x \in D \cap B(z_0, r).$$

This is true for some domains but not all. An example of a reasonable domain for which this is true is a Lipschitz domain, one that agrees locally with the region above the graph of a Lipschitz function.

Let us give an example of how the BHP can be used. If we have a thin strip of length R and width about r with $R \gg r$, then starting at x in the middle, the probability of exiting the sides before hitting the lower boundary starting at x can be shown to be of order $e^{-cR/r}$. If x is near the middle in the long direction, but δ above the lower boundary, the probability of hitting the upper boundary before hitting the lower boundary is of order δ. Using the BHP, we can get a more precise estimate.

Let $u(x)$ be the probability of hitting the sides before the top or bottom, and $v(x)$ the probability of hitting the top before the bottom. Let x_0 be in the middle. Then $v(x_0) \approx \frac{1}{2}$, while $u(x_0) \approx e^{-cR/r}$. So if x is near the middle in the long direction and δ above the bottom, then

$$\frac{u(x)}{\delta} \approx \frac{u(x)}{v(x)} \approx \frac{e^{-cR/r}}{1/2},$$

or

$$u(x) \approx \delta e^{-cR/r}.$$

The key to the proof of the BHP is to show that when exiting $B(z_0, r) \cap D$ the probability of the exit position being in D but not too close to ∂D is not too much smaller than the probability of the exit position being in D and close to ∂D. Of course, the largest probability is when the exit position is in ∂D. This is also the key to the proof for α-harmonic functions and we will show how it goes.

First we need the following fact. Let D be a domain and $A \subset D^c$. Let $F(x,z)$ be 1 if $x \in D, z \in A$ and 0 otherwise. We know

$$\int_0^{\tau_D \wedge t} \int F(X_{s-}, z)(\mu - \nu)(ds\, dz)$$

is a martingale. Let $g_D(x,y)$ be the Green function for D:

$$\mathbb{E} \int_0^{\tau_D} 1_C(X_s)\, ds = \int_C g_D(x,y)\, dy.$$

Then

$$\begin{aligned}
\mathbb{P}^x(X_{\tau_D} \in A) &= \mathbb{E}^x \int_0^{\tau_D} \int 1_D(X_{s-}) 1_A(X_{s-} + z)\, \mu(ds\, dz) \\
&= \mathbb{E}^x \int_0^{\tau_D} \int 1_D(X_{s-}) 1_A(X_{s-} + z)\, n(dz)\, ds \\
&= \int \int g_D(x,y) 1_A(y+z)\, \frac{1}{|z|^{d+\alpha}}\, dz\, dy \\
&= \int \int_A \frac{g_D(x,y)}{|y-z|^{d+\alpha}}\, dz\, dy.
\end{aligned}$$

Using Fubini, we then see that

$$\mathbb{P}^x(X_{\tau_D} \in dz) = \Big(\int \frac{g_D(x,y)}{|y-z|^{d+\alpha}} dy \Big)\, dz. \tag{4.1}$$

Now let us look at the key fact for the BHP for α-harmonic functions that are 0 in $B(z_0, r) \cap D^c$, due to [27]. Suppose D

is the intersection of the region above the graph of a Lipschitz function with a large ball. Let $z_0 \in \partial D$ be not too close to the circumference of the large ball.

Lemma 4.2 *Let $r > 0$ and x_0 such that $x_0 \in B(z_0, r) \cap D$ and $\operatorname{dist}(x_0, \partial D) = r/2$. There exists c such that if $x \in D \cap B(z_0, r/4)$, then*

$\mathbb{P}^x(X_{\tau_D} \in B(z_0, r)^c) \le c\mathbb{P}^x(X_t \text{ hits } B(x_0, r/4) \text{ before exiting } D).$

To prove this, write B for $B(z_0, r)$. Then from the formula (4.1)

$$\begin{aligned}
\mathbb{P}^x(X_{\tau_D} \in B^c) &= \int_{B^c} \int_D \frac{g_D(x,y)}{|y-z|^{d+\alpha}}\, dy\, dz \\
&= \int_D \int_{B^c} \frac{g_D(x,y)}{|y-z|^{d+\alpha}}\, dz\, dy \\
&\le c \int_D g_D(x,y)\, dy \qquad (4.2)
\end{aligned}$$

For x fixed, $g_D(x, y)$ is α-harmonic as a function of y, except when $y = x$, and so by Harnack

$$g_D(x, x_0) \ge c \int_{B(x_0, r/4)} g_D(x,y)\, dy. \qquad (4.3)$$

Let $C = B(x_0, r/4)$. Since $g_D(x, y) = 0$ if $y \notin D$,

$$\begin{aligned}
g_D(x, x_0) &= \mathbb{E}^{x_0} g_D(x, X_{\tau_C}) \\
&= \int_{C^c \cap D} g_D(x,z) \Big[\int_C \frac{g_C(x_0,y)}{|y-z|^{d+\alpha}}\, dy \Big]\, dz \\
&\ge c \int_{C^c \cap D} g_D(x,z)\, dz \\
&= c \int_{D \cap C^c} g_D(x,y)\, dy. \qquad (4.4)
\end{aligned}$$

Combining (4.2), (4.3), and (4.4)

$$g_D(x, x_0) \geq c \int g_D(x, y)\, dy \geq \mathbb{P}^x(X_{\tau_D} \in B^c).$$

Finally, by Harnack,

$$\begin{aligned} g_D(x, x_0) &= c \int_{B(x_0, r/4)} g_D(x, z)\, dz \\ &\leq \mathbb{E}^{x_0}\Big[\int_0^{\tau_D} 1_C(X_s)\, ds; X_t \text{ hits } C \text{ before exiting } D\Big] \\ &\leq c\mathbb{P}^x(X_t \text{ hits } C \text{ before exiting } D). \end{aligned}$$

4.2 Fatou theorem

References here are [16] and [17].

Let D be the upper half space (or more generally, a Lipschitz domain, although we'll restrict attention to upper half spaces). If f is bounded (or non-negative or in L^p for some p) and $u(x) = \mathbb{E}^x f(X_{\tau_D})$, where X is Brownian motion, then it is known that the nontangential limits of u exist and equal f a.e. More precisely, if D is the upper half space in $\mathbb{R}^d$ and for each $z \in \partial D$ we let

$$C_z = \{(w, y) : w \in \mathbb{R}^{d-1}, y \geq 0, |w - z| < y\},$$

then for almost every $z \in \partial D$ (with respect to $(d-1)$-dimensional Lebesgue measure,

$$\lim_{x \in C_z, x \to z} u(x) = f(z).$$

What about when X is a symmetric stable process? Of course here we need f defined on D^c.

But even when f is bounded, the limit theorem is not necessarily true. For example, let $d = 2$ and define f on D^c to be 0 or 1. Set it equal to 1 on the strip from $y = -\varepsilon$ to $y = -\varepsilon^{-1}$. If ε is small enough, and $w = (x, 1)$, then with high probability X_{τ_D} will lie in this strip and so $u(w)$ will be close to 1. Define f to be 0 when $y < -1$. On the strip $-\varepsilon^3 \le y < -\varepsilon$ define f to be 0.

So starting at (x, ε^2), with high probability the exit position will be in this second strip, and so $u(x, \varepsilon^2)$ is close to 0. Continuing to define f on the strips $-\varepsilon^{2k+1} < y < -\varepsilon^{2k-1}$ with alternating 1's and 0's, we see we do not have convergence for any $z \in \partial D$.

Therefore we need some regularity for f. Let us define L^p_β to be the set of functions f such that

$$\|f\|_{L^p_\beta} = \|f\|_p + \sup_{t>0} \frac{\|f(\,\cdot\, + t) - f(\,\cdot\,)\|_p}{|t|^\beta} < \infty.$$

This is a well known space, known as the L^p-Hölder continuous functions. Here $p \ge 1$ and $\beta \in (0, 1)$. Such f need not be continuous unless $\beta p > d$. Note that L^∞_β are the usual Hölder continuous functions of order β.

Our theorem is

Theorem 4.3 *If f is bounded and in L^p_β and $\beta p > 1$, then nontangential limits of u exist for a.e. $z \in \partial D$.*

When $\beta p \in (1, d)$, then f need not be continuous, but one can still define a notion of the trace of f on ∂D. Define $Tf(z) = f(z)$ if $z \in \partial D$ and f is continuous. It is possible to show that the operator T can be extended in one and only one way to a continuous operator on L^p_β. In this case, it turns out that the nontangential limit of u is Tf a.e.

The example above shows that $\beta p > 1$ is sharp. First take f as above, except 0 outside some large ball of radius R and multiply it by a mollifier that is supported in the ball of radius R and is 1 on the ball of radius $R/2$. This large ball modification doesn't really affect things, so we might as well assume we are in dimension 1.

Next modify f slightly. Let f be as before, but make it linear between $-(1-\delta)\varepsilon^{2k+1}$ and $-(1+\delta)\varepsilon^{2k+1}$. If we take δ much smaller than ε, this doesn't mess up the oscillation of u.

Let us look at $|f(x+t)-f(x)|$. This is bounded by 2, and we use that bound when $\varepsilon^{2k} < t/4$. When $\varepsilon^{2k} \geq t/4$, then the difference is either 0 or is bounded by t times the derivative: $ct\varepsilon^{-2k}$. The measure of such x's is at most ct. If k_0 is such that $\varepsilon^{2k_0} \approx t$ and $p > 1$, then we have

$$\begin{aligned}\|f(\,\cdot\,+t) - f(\,\cdot\,)\|_p &\leq c\Big(\sum_{k\leq k_0} (ct\varepsilon^{-2k})^p t + 2t\Big)^{1/p}\\ &\leq ct^{1/p}\Big(\varepsilon^{-2k_0 p} t^p\Big)^{1/p}\\ &= ct^{1/p}.\end{aligned}$$

If $\beta = 1/p$, then we have that our f is in L^p_β.

To prove Theorem 4.3, there are two main steps, bounding a maximal function, and then using the maximal function. Unlike the Brownian motion case, we construct the maximal function in D^c. Let L be a large integer, and for $z \in \partial D$, let $A_{ij}(z)$ be a box of side length 2^{-i} and whose center has horizontal component $z + j2^{-i}$ and vertical component -2^{2-2i}. Let $F_{ij}(x)$

be the average of $|f|$ over $A_{ij}(x)$. Let $G_i(x) = \sup_{|j|\leq L} F_{ij}(x)$ and $M(x) = \sup_{i\geq 0} G_i(x)$.

Let $\beta' < \beta$ but with $\beta' p > 1$. Suppose we can show that the L^1 norm of M (with respect to $(d-1)$-dimensional Lebesgue measure) is bounded be a constant times the $L^p_{\beta'}$ norm of f. Given ε and $f \in L^p_\beta$, we can find g continuous with compact support and $h \in L^p_{\beta'}$ with h having $L^p_{\beta'}$ norm less than ε and L^∞ norm less than a constant multiple of the L^∞ norm of f such that $f = g + h$. Since g is continuous, its harmonic extension converges nontangentially, so we need to look at the nontangential convergence of the harmonic extension of h. The L^1 norm of the maximal function for h is bounded by a constant times ε. Choose δ small and R large such that

$$\mathbb{P}^{(x,y)}(|X_{\tau_D} - x| > Ry) < \varepsilon$$

and

$$\mathbb{P}^{(x,y)}(X_{\tau_D} > -\delta y) < \varepsilon.$$

By taking L large enough, we have

$$\{(v,w) : |(v,w) - (x,0)| < Ry, w < -\delta y\} \subset \bigcup_{|j|\leq L, i\geq 0} A_{ij}(x).$$

Since h is bounded, we see that up to some multiple of ε, the oscillation of h is bounded by the maximal function, which is small except possibly on a small set.

So the proof comes down to getting an estimate on the maximal function. The main estimate is to get a bound on the L^1 norm of

$$|F_{i+1,j_1}(x) - F_{i,j_2}(x)|$$

with $|j_1|, |j_2| \leq L$. Since the sup over $2L+1$ elements is comparable to the sum over $2L+1$ elements, this will give us essentially the same bound on the L^1 norm of

$$|G_{i+1}(x) - G_i(x)|.$$

It will turn out that this bound is of the form $c2^{-\gamma i}$ for some $\gamma > 0$. Since $\|G_1\|_1$ is easily seen to be finite, then from

$$\|M\|_1 \le \|G_1\|_1 + \sum_{i\ge 0} \|G_{i+1} - G_i\|_1,$$

we get our bound on $\|M\|_1$.

So let's look at $|F_{i+1,j_1}(x) - F_{i,j_2}(x)|$ Let us suppose $j_1 = j_2 = 0$, though this is just for simplification. By dividing A_{i0} into 2^d equal subcube, we see that we need to look at the difference between two cubes of side length 2^{-i} that are about 2^{-i} apart. Let t be the vector that goes from the center of one cube to the center of the other. the length of t is of order 2^{-i}. If we integrate over the x variable, $f(w+t) - f(w)$ for a given w occurs in the integrand for many x's, in fact for a set of measure $2^{-i(d-1)}$. We also get a factor 2^{id} because we are looking at averages, not integrals. So if $B(x)$ and $B'(x)$ are our two cubes, then

$$2^{id} \int \Big| \int_{B(x)} |f| - \int_{B'(x)} |f| \Big|\, dx \le c2^i \int 1_E(z) |f(z+t) - f(z)|\, dz,$$

where E is the product of some big ball with a strip of width $c2^{-i}$. So using Hölder's inequality, the above is bounded by

$$c2^i \|f(\,\cdot\, + t) - f(\,\cdot\,)\|_p (2^{-i})^{1/q} \le c2^{i(1-\frac{1}{q})} |t|^\beta = c2^{i/p} 2^{-i\beta}.$$

This is summable in i if $\beta > 1/p$, or if $\beta p > 1$.

Besides the BHP (which has been shown to hold much more generally than in Lipschitz domains), results are known on the

Martin boundary, intrinsic ultracontractivity, and on the conditional gauge.

Take almost any potential result for Brownian motion, and the analogue for symmetric stable processes would be interesting.

Here are a few open problems. First of all, it is known that the Fatou theorem holds in Lipschitz domains. Can it be shown to hold in more general domains? A class of domains to consider is the following: a Lipschitz domain is locally the region above the graph of a Lipschitz function Φ, and Φ Lipschitz means $\nabla\Phi$ is in L^∞. What if L^∞ is replaced by L^p? These are known as L_1^p domains. Results are known for harmonic functions in these domains with respect to the Laplacian, but not with respect to α-harmonic functions.

An irritating feature of the Fatou theorem above is that we required f to be bounded. It would be nice to eliminate this restriction.

The hitting distribution of Brownian motion in a Lipschitz domain is known to be equivalent to surface measure, and the Radon-Nikodym derivative is in $L^{2+\varepsilon}$ with respect to surface measure for some $\varepsilon > 0$. What is the analogue for symmetric stable processes? Here one would replace surface measure by Lebesgue measure on the complement of the domain. What about extensions to L_1^p domains?

5 Weak uniqueness

5.1 The perturbation method

The most common way of proving weak uniqueness is the perturbation approach. Let R_λ be a resolvent, that is,

$$R_\lambda f(x) = \mathbb{E}^x \int_0^\infty e^{-\lambda t} f(X_t)\,dt,$$

for some known process. Formally, $R_\lambda = (\lambda - \mathcal{A})^{-1}$, where $\mathcal{A}$ is the generator of the process. Now consider the generator $\mathcal{A}+\mathcal{B}$, and let S_λ be the resolvent of the corresponding process. We expect

$$S_\lambda f = R_\lambda f + R_\lambda \mathcal{B} R_\lambda f + R_\lambda (\mathcal{B} R_\lambda)^2 f + \cdots .$$

This is plausible, since

$$\begin{aligned}(\lambda - \mathcal{A} - \mathcal{B})\Big[R_\lambda\Big(\sum_{i=0}^\infty (\mathcal{B}R_\lambda)^i\Big)\Big]& \\ &= (\lambda - \mathcal{A}) R_\lambda \sum_{i=0}^\infty (\mathcal{B}R_\lambda)^i - \mathcal{B}R_\lambda \sum_{i=0}^\infty (\mathcal{B}R_\lambda)^i \\ &= \text{ the identity.}\end{aligned}$$

This is more or less fine if $\|\mathcal{B}R_\lambda\| < 1$ for some norm. An additional complication is if the norm is not L^∞ or C^α or something like that, the the formula for $S_\lambda f$ only holds a.e., and an additional argument is needed to go from almost everywhere to everywhere.

Let's look at a simple example. Suppose we are in one dimension, $\alpha > 1$, $\mathcal{A}$ is the generator of a symmetric stable process of index α and $\mathcal{B}f(x) = b(x)f'(x)$. Then $R_\lambda f(x) = \int f(y)r_\lambda(x,y)\,dy$, where $r_\lambda(x,y)$ dies off rapidly as $|x-y|$ tends to infinity and is asymptotically

$$\frac{c}{|x-y|^{1-\alpha}}$$

when $|x-y|$ is small. Not worrying about the behavior at infinity,

$$\mathcal{B}R_\lambda f(x) \approx -b(x)\int_{|x-y|\le 1} f(y)\frac{c}{|x-y|^{2-\alpha}}\,dy.$$

This is bounded in absolute value by

$$c\|b\|_\infty\,\|f\|_\infty \int_{|x-y|\le 1}\frac{1}{|x-y|^{2-\alpha}} \le \tfrac{1}{2}\|f\|_\infty$$

if $\alpha > 1$ and $\|b\|_\infty$ is sufficiently small.

How can we make this all more precise? Let $\mathcal{L}$ be a generator of the form

$$\mathcal{L}f(x) = \int [f(x+h) - f(x) - 1_{(|h|\le 1)}\nabla f(x)\cdot h]\frac{A(x,h)}{|h|^{d+\alpha}}\,dh.$$

The processes corresponding to this class of generators are sometimes called stable-like processes; [19].

We say a probability measure $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ started at a point x_0 if $\mathbb{P}(X_0 = x_0)$ and $f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,ds$ is a $\mathbb{P}$-martingale whenever $f \in C_b^2$. Here $\Omega = D[0,\infty)$, $X_t(\omega) = \omega(t)$, and $\mathcal{F}_t$ is the filtration generated by X.

It is known ([24]) that existence and uniqueness of the a solution to the martingale problem is equivalent to weak existence and weak uniqueness for the solution of an affiliated SDE.

Let us give a special case of a more general theorem ([15]).

Theorem 5.1 *Suppose for some $\varepsilon > 0$ we have $|A(x,h) - 1| \le ch^\varepsilon$ for all $|h| \le 1$. Suppose also that $A(x,-h) = A(x,h)$ for all x and h. Then for each x there is a unique solution to the martingale problem for $\mathcal{L}$ started at x.*

One consequence of this is that there is a strong Markov process associated with $\mathcal{L}$.

Proof. First of all, jumps of size larger than δ come at discrete times, and if we have uniqueness up to the first such jump time, this completely determines the distribution of the jump, and we can then run until the second jump larger than h, and so on. So without loss of generality we may assume $A(x,h) = 1$ for all $|h| \ge \delta$ for some δ.

Let $\mathcal{P}$ be a solution to the martingale problem started at x_0. Let

$$S_\lambda f = \mathbb{E} \int_0^\infty e^{-\lambda t} f(X_t)\, dt.$$

We don't know anything about the solution being a Markov process, so we cannot consider S_λ a resolvent, and so don't write $S_\lambda f(x_0)$. If $f \in C_b^2$, then

$$\mathbb{E}\, f(X_t) - f(x_0) = \mathbb{E} \int_0^t \mathcal{L} f(X_s)\, ds.$$

Multiply by $e^{-\lambda t}$ and integrate over t:

$$\begin{aligned} S_\lambda f - \frac{1}{\lambda} f(x_0) &= \int_0^\infty e^{-\lambda t} \mathbb{E} \int_0^t \mathcal{L} f(X_s)\, ds\, dt \\ &= \int_0^\infty \int_s^\infty e^{-\lambda t} \mathbb{E}\, \mathcal{L} f(X_s)\, dt\, ds \\ &= \int_0^\infty \frac{e^{-\lambda s}}{\lambda} \mathbb{E}\, \mathcal{L} f(X_s)\, ds \\ &= \frac{1}{\lambda} S_\lambda(\mathcal{L} f). \end{aligned}$$

Let $\mathcal{A}$ be the generator of symmetric stable process of index α and let R_λ be the corresponding resolvent. Then if $f = R_\lambda g$, we have

$$\mathcal{A} f = \mathcal{A}(R_\lambda g) = \lambda R_\lambda g - g.$$

Letting $\mathcal{B} = \mathcal{L} - \mathcal{A}$, then

$$\lambda S_\lambda R_\lambda g = R_\lambda g(x_0) + S_\lambda \mathcal{B} R_\lambda g + S_\lambda(\lambda R_\lambda g - g),$$

or

$$S_\lambda g = R_\lambda g(x_0) + S_\lambda \mathcal{B} R_\lambda g.$$

This is the analogue of the series expansion for S_λ that we wrote earlier.

If S_λ^i, $i = 1, 2$, correspond to two solutions, then

$$(S_\lambda^1 - S_\lambda^2) g = (S_\lambda^1 - S_\lambda^2) \mathcal{B} R_\lambda g.$$

If $\|\mathcal{B} R_\lambda\| < 1$ for some norm and

$$\Theta = \sup_{\|g\| \le 1} |(S_\lambda^1 - S_\lambda^2) g|,$$

we get

$$\Theta \le \Theta \|\mathcal{B} R_\lambda\|,$$

which implies that Θ equals either 0 or infinity.

If $\Theta = 0$, then we have $S^1_\lambda g = S^2_\lambda g$, and by the uniqueness of the Laplace transform, we get $\mathbb{E}^1 g(X_t) = \mathbb{E}^2 g(X_t)$ for almost every t. If g is continuous, using the right continuity of X_t shows this holds for all t, and we thus get the uniqueness of the one-dimensional distributions. An argument of Stroock-Varadhan shows that this is enough to imply uniqueness of the finite dimensional distributions, which is what we want.

So one of the things we need to do is rule out $\Theta = \infty$. Stroock and Varadhan have an argument to do this, but to make it go through, one needs

$$|R_\lambda g(x)| \le c\|g\|$$

for every x. We are going to use the L^2 norm, and this won't be true in general. The way to get around it is to take the initial distribution to be random with a nice density. Since one has uniqueness for all such starting distributions, one can conclude uniqueness for almost every starting point. Finally, we can use the Harnack inequality and regularity of resolvents (see the next section!) to go to uniqueness for every starting point.

So we need an L^2 bound on $\mathcal{B}R_\lambda$. For simplicity, let us take $\alpha < 1$ and d, the dimension, equal to 1. Using the fact that $A(x,h) = 1$ if $|h| \ge 1$ and symmetry, we have that for all x

$$\begin{aligned}
|\mathcal{B}R_\lambda f(x)| &= \Big| \int_{-\delta}^{\delta} [R_\lambda f(x+h) - R_\lambda f(x)] \frac{A(x,h)-1}{|h|^{1+\alpha}}\,dh \Big| \\
&\le \int_0^\delta |R_\lambda f(x+h) + R_\lambda f(x-h) - 2R_\lambda f(x)| \frac{|A(x,h)-1|}{|h|^{1+\alpha}}\,dh \\
&\le \int_0^\delta |R_\lambda f(x+h) + R_\lambda f(x-h) - 2R_\lambda f(x)| \frac{h^\varepsilon}{|h|^{1+\alpha}}\,dh.
\end{aligned}$$

Therefore

$$\|\mathcal{B}R_\lambda f\|_2 \le c\int_0^\delta \|R_\lambda f(\,\cdot+h) + R_\lambda f(\,\cdot-h) - 2R_\lambda f(\,\cdot\,)\|_2 \frac{h^\varepsilon}{h^{1+\alpha}}\,dh.$$

By Parseval,

$$\begin{aligned}\|R_\lambda f(\,\cdot+h) &+ R_\lambda f(\,\cdot-h) - 2R_\lambda f(\,\cdot\,)\|_2^2\\ &= c\int |e^{iuh} + e^{-iuh} - 2|^2 |\widehat{R_\lambda f}(u)|^2\,du\\ &\le c\int \frac{|\widehat{f}(u)|^2}{(\lambda+|u|^\alpha)^2}(|u|^\alpha|h|^\alpha)^2\,du\\ &\le c|h|^{2\alpha}\|f\|_2^2.\end{aligned}$$

Therefore

$$\|\mathcal{B}R_\lambda f\|_2 \le c\int_0^\delta \|f\|_2 \frac{h^\alpha h^\varepsilon}{h^{1+\alpha}}\,dh \le c\|f\|_2\delta^\varepsilon.$$

This will be less than $\frac{1}{2}\|f\|_2$ if we take δ small enough.

5.2 Variable order

The perturbation approach is a powerful one, but it does not handle jump operators like

$$\mathcal{L}f(x) = \int_{\mathbb{R}^d} [f(x+h) - f(x) - 1_{(|h|\le 1)}\nabla f(x)\cdot h]\frac{1}{|h|^{d+\alpha(x)}}\,dh.$$

The corresponding processes are also sometimes called stable-like; [4]. There are variable order operators, and are a special case of

$$\mathcal{L}f(x) = \int_{\mathbb{R}^d} [f(x+h) - f(x) - 1_{(|h|\le 1)}\nabla f(x)\cdot h]n(x,h)\,dh,$$

or even

$$\mathcal{L}f(x) = \int_{\mathbb{R}^d} [f(x+h) - f(x) - 1_{(|h|\le 1)} \nabla f(x) \cdot h] \, n(x, dh).$$

Let's consider a special case: $d = 1$, $0 < \eta < \alpha(x) < 1 - \eta < 1$ for all x (so we don't have to to consider the ∇f term), and then

$$\mathcal{L}f(x) = \int_{\mathbb{R}} [f(x+h) - f(x)] \frac{1}{|h|^{1+\alpha(x)}} \, dh.$$

The following is from [4].

Theorem 5.2 *Suppose $\alpha(x)$ is Hölder continuous of order γ. Then the martingale problem for $\mathcal{L}$ started at x_0 has a unique solution for each x_0.*

Proof. The idea is a sort of perturbation argument, but we don't perturb off a Markov process. Let r_λ^β be the resolvent density for a stable symmetric process of oder β. We don't need to worry about behavior at infinity, and near 0, we have

$$r_\lambda^\beta(x, y) \sim \frac{c}{|x-y|^{1-\beta}}.$$

Let

$$H(x, y) = r_\lambda^{\alpha(y)}(x, y) \approx \frac{c}{|x-y|^{1-\alpha(y)}}.$$

If g is smooth with compact support, let

$$f(x) = \int H(x, y) g(y) \, dy.$$

Now f might not be in C_b^2, but let's pretend it is. What one does is do an additional smoothing by an approximation to the identity.

With this proviso, we have as before

$$S_\lambda(\lambda - \mathcal{L})f = f(x_0),$$

where $S_\lambda f = \mathbb{E}\int_0^\infty e^{-\lambda t} f(X_t)\, dt$. If we have two solutions, then

$$(S_\lambda^1 - S_\lambda^2)(\lambda - \mathcal{L})f = 0.$$

Suppose we show

$$(\lambda - \mathcal{L})f(x) = g(x) + Kg(x) \tag{5.1}$$

with $Kg(x) = \int K(x,y)g(y)\, dy$ and

$$\int |K(x,y)|\, dy \le \tfrac{1}{2}$$

for all x. Then defining

$$\Theta = \sup_{\|F\|_\infty \le 1} |(S_\lambda^1 - S_\lambda^2)F|,$$

we have

$$|(S_\lambda^1 - S_\lambda^2)g| \le |(S_\lambda^1 - S_\lambda^2)Kg| \le \Theta \|Kg\|_\infty \le \tfrac{1}{2}\Theta\|g\|_\infty,$$

and hence $\Theta \le \frac{1}{2}\Theta$. Unlike the previous theorem, here we know $|S_\lambda g| \le \|g\|_\infty/\lambda$, and hence $\Theta = 0$. We then proceed as before.

So we need to show (5.1). Let

$$\mathcal{M}^\beta f(x) = \int [f(x+h) - f(x)] \frac{1}{|h|^{1+\beta}}\, dh,$$

and we then have

$$\begin{aligned}
(\lambda - \mathcal{L})f(x) &= (\lambda - \mathcal{M}^{\alpha(x)})f(x)\\
&= \int (\lambda - \mathcal{M}^{\alpha(y)}) \frac{1}{|x-y|^{1-\alpha(y)}} g(y)\, dy\\
&\quad + \int [\mathcal{M}^{\alpha(y)} - \mathcal{M}^{\alpha(x)}] \frac{1}{|x-y|^{1-\alpha(y)}} g(y)\, dy\\
&= I_1 + I_2.
\end{aligned}$$

Look at I_1. The operator $\mathcal{M}^\beta$ operates on the x variable, so the $\alpha(y)$ doesn't play an important role, and

$$I_1 \approx \int \delta_0(x-y)g(y)\,dy = g(x).$$

Using the approximate identity helps make this rigorous.

We let

$$K(x,y) = [\mathcal{M}^{\alpha(y)} - \mathcal{M}^{)\alpha(x)}]\frac{1}{|x-y|^{1-\alpha(y)}}.$$

So we need to consider

$$\int_{\mathbb{R}} \Big[\frac{1}{|x-y+h|^{1-\alpha(y)}} - \frac{1}{|x-y|^{1-\alpha(y)}}\Big]\Big[\frac{1}{|h|^{1+\alpha(y)}} - \frac{1}{|h|^{1+\alpha(x)}}\Big]\,dh. \tag{5.2}$$

Let $R = |x-y|$. We break the integral into a number of pieces: $h \in [-3R/2, -R/2]$, $h \in (-\infty, -3R/2)$, etc. The first piece is the hardest and most important. So $h \approx -R$, and the second factor in the integrand in (5.2) is equal to

$$\Big[\frac{1}{|h|^{1+\alpha(y)}}\Big]\Big[1 - |h|^{\alpha(y)-\alpha(x)}\Big].$$

This in turn is approximately

$$\begin{aligned}
R^{-1-\alpha(y)}&\Big[1 - e^{(\alpha(y)-\alpha(x))\log|h|}\Big]\\
&\le cR^{-1-\alpha(y)}|\alpha(y)-\alpha(x)|\,|\log|h|\,|\\
&\le cR^{-1-\alpha(y)}R^\gamma \log R\\
&\le cR^{-1-\alpha(y)+\gamma/2}.
\end{aligned}$$

So we are now looking at

$$\int_{-3R/2}^{-R/2} \frac{1}{|R+h|^{1-\alpha(y)}} R^{-1-\alpha(y)+\gamma/2} \approx R^{-1+\gamma/2}.$$

Hence

$$|K(x,y)| \approx |x-y|^{\gamma/2-1},$$

which integrates (locally) to something bounded in x.

To get the integral less than $\frac{1}{2}$, this can be done by taking λ large enough and including its influence in the calculations.

From the proof we see that it is enough that for some $\varepsilon > 0$

$$|\alpha(x) - \alpha(x)| \le \frac{1}{|\log|x-y|\,|^{2+\varepsilon}}.$$

By using Fourier analysis, one can do even a bit better and have the 2 in the above replaced by 1.

Theorem 5.2 has been greatly generalized in [28]. On the other hand, it is known that uniqueness for the martingale problem does not always hold; see [1, Section 6]. When exactly does uniqueness hold?

Consider the system of SDEs

$$dX_t^i = \sum_{j=1}^d A_{ij}(X_{t-})\, dZ_t^j,$$

where the Z^j are independent 1-dimensional symmetric stable processes of order α, $A_{ij}(x)$ is continuous in x, and for each x the matrix $A(x)$ is bounded and nondegenerate. This is the analogue of the SDEs used to define diffusions on $\mathbb{R}^d$. Under the above conditions it was proved in [9] that weak existence and uniqueness of the solution to the above system of SDEs holds. The case where the Z^j are independent but stable symmetric processes of ordeer α_j, where α_j depends on j, would be of interest to people in financial mathematics.

What about martingale problems where the operator is derived from an SDE defined in terms of Poisson point processes?

If the $A(x,h)$ are not continuous, but only measurable, can one show existence of a solution of the martingale problem? One needs an analogue of Krylov's inequality (essentially equivalent to the Alexandrov-Bakel'man-Pucci estimate from PDE). Some progress on Krylov's inequality in this context has been made by Foondun.

6 Harnack inequalities

We now turn to Harnack inequalities. These have a long history in PDE, notable being the one of Moser for divergence form operators and the one of Krylov-Safonov for nondivergence form operators. Harnack inequalities have a number of important uses, one being the proof of regularity of solutions to a PDE.

All those Harnack inequalities were for operators that were local, i.e., were differential in form. Here we are going to look at Harnack for non-local operators.

The class of operators we are going to consider are ones with generator of the form

$$\mathcal{L}f(x) = \int [f(x+h) - f(x) - 1_{(|h|\le 1)} \nabla f(x) \cdot h] \frac{A(x,h)}{|h|^{d+\alpha}}\, dh. \quad (6.1)$$

We assume

$$0 < c_1 \le A(x,h) \le c_2, \qquad x, h \in \mathbb{R}^d,$$

and

$$A(x,-h) = A(x,h), \qquad x, h \in \mathbb{R}^d.$$

The second condition can be relaxed, but the first is crucial – Harnack need not hold without it.

We will also assume that there is a solution to the martingale problem for $\mathcal{L}$ started at each x, and that these solutions form a strong Markov process.

The reference for this section is [13].

6.1 Preliminaries

We need a few preliminary facts.

1. Scaling: if $Y_t = aX_{ta^{-\alpha}}$, then Y is also stable-like with
$$A_Y(x,h) = a^{-(d+\alpha)}A_X(a^{-1}x, a^{-1}h).$$

2. Lévy system formula: Suppose $\overline{A} \cap \overline{B} = \emptyset$. Then
$$\sum_{s\le t} 1_{(X_{s-}\in A, X_s\in B)} - \int_0^t 1_A(X_s)\int_B \frac{A(X_s, u-X_s)}{|u-X_s|^{d+\alpha}}\,du\,ds$$

is a $\mathbb{P}^x$-martingale for each x.

Proof. For simplicity let us take $\alpha < 1$, so we can drop the ∇f term. Let $f \in C^2$ with $f = 0$ on A and $f = 1$ on B. The process $M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,ds$ is a martingale, so $\int_0^t 1_A(X_{s-})\,dM_s$ is also a martingale.

If $x \in A$, then $f(x) = 0$, so
$$\mathcal{L}f(x) = \int f(x+h)\frac{A(x,h)}{|h|^{d+\alpha}}\,dh = \int f(u)\frac{A(x,u-x)}{|u-x|^{d+\alpha}}\,du.$$

On the other hand
$$f(X_t) - f(X_0) = \sum_{s\le t}[f(X_s) - f(X_{s-})],$$

so

$$\sum_{s\le t}[1_A(X_{s-})(f(X_s)-f(X_{s-}))]-\int_0^t 1_A(X_{s-})\mathcal{L}f(X_s)\,ds$$

is a martingale. In the last integral we can change X_{s-} to X_s since the two integrands differ on a countable set.

Taking limits completes the proof.

3. Estimates on exit times:

Let τ_A be the times of the first exit from A.

a. There exists c such that

$$\mathbb{P}^x(\sup_{s\le t}|X_s-X_0|>1)\le ct.$$

Proof. We may assume $x=0$. Take $f\in C^2$ taking values in $[0,1]$ such that $f(x)=0$ and $f(y)=1$ if $|y|\ge 1$. Then

$$\begin{aligned}|\mathcal{L}f(z)|\le c\int_{|h|\le 1}|h|^2\frac{A(z,h)|}{|h|^{d+\alpha}}\,dh\\ +c\int_{|h|>1}\frac{A(z,h)|}{|h|^{d+\alpha}}\,dh\\ \le c.\end{aligned}$$

Hence

$$\begin{aligned}\mathbb{P}^0(\tau_{B(0,1)}\le t)&\le \mathbb{E}^{\,0}f(X_{\tau_{B(0,1)}\wedge t})-f(0)\\ &=\mathbb{E}^{\,0}\int_0^{\tau_{B(0,1)}\wedge t}\mathcal{L}f(X_s)\,ds\le ct.\end{aligned}$$

b. Let $\varepsilon>0$. There exists c such that

$$\mathbb{E}^{\,x}\tau_{B(x,r)}\ge cr^\alpha$$

if $|z-x| \le (1-\varepsilon)r$.

Proof. We may take $x=0$ and by scaling, $r=1$. Then

$$\mathbb{P}^z(\tau_{B(0,1)} \le \varepsilon^\alpha t) \le \mathbb{P}^z(\sup_{s\le \varepsilon^\alpha t} |X_s - X_0| \ge \varepsilon) \le ct \le \tfrac{1}{2}$$

if t is small enough. Hence

$$\mathbb{E}^{\,z}\tau_{B(x,1)} \ge \varepsilon^\alpha t \mathbb{P}^z(\tau_{B(x,1)} \ge \varepsilon^\alpha t) \ge \varepsilon^\alpha t/2.$$

c. There exists c such that

$$\mathbb{E}^{\,z}\tau_{B(x,r)} \le cr^\alpha$$

for all z.

Proof. By scaling we take $r=1$. Let S be the time of the first jump larger than 2 in size. Then

$$\begin{aligned}\mathbb{P}^z(S \le 1) &= \mathbb{E}^{\,z} \sum_{s\le S\wedge 1} 1_{(|X_s - X_{s-}|>2)}\\ &= \mathbb{E}^{\,z}\int_0^{S\wedge 1}\int_{|h|>2} \frac{A(X_s,h)}{|h|^{d+\alpha}}\,dh\,ds\\ &\ge c\mathbb{E}^{\,z}(S\wedge 1) \ge c\mathbb{P}^z(S>1).\end{aligned}$$

This implies $\mathbb{P}^z(S\le 1) \ge c'$, and a standard argument finishes the proof.

4. Hitting sets: Suppose $F \subset B(x,1)$. There exists c such that

$$\mathbb{P}^y(X \text{ hits } F \text{ before exiting } B(x,3)) \ge c|F|, \qquad y \in B(x,2).$$

Proof. Write τ for the first exit time from $B(x,3)$ and T for the first hitting time of F. If $X_s \in B(x,3)$ and $u \in B(x,1)$, then $|X_s - u|$ is bounded above. Hence

$$\begin{aligned}\mathbb{P}^y(T<\tau) &\geq \mathbb{E}^y \sum_{s\leq T\wedge\tau\wedge t} 1_{(X_s\neq X_{s-}, X_s\in F)} \\ &= \mathbb{E}^y \int_0^{T\wedge\tau\wedge t}\int_F \frac{A(X_s, u-X_s)}{|u-X_s|^{d+\alpha}}\,du\,ds \\ &\geq c|F|\mathbb{E}^y(T\wedge\tau\wedge t).\end{aligned}$$

Now

$$\begin{aligned}\mathbb{E}^y(T\wedge\tau\wedge t_0) &\geq t_0\mathbb{P}^y(T\geq\tau\geq t_0) \\ &\geq t_0[1-\mathbb{P}^y(T<\tau)-\mathbb{P}^y(\tau<t_0)].\end{aligned}$$

Choose t_0 such that $\mathbb{P}^y(\tau<t_0)\leq 1/4$. If $\mathbb{P}^y(T<\tau)\geq 1/4$, we are done, so assume $\mathbb{P}^y(T<\tau)\leq 1/4$. But then $\mathbb{E}^y(T\wedge\tau\wedge t_0)\geq t_0$.

5. Hitting distributions:

If H is bounded, nonnegative, and supported in $B(z,2r)^c$, and $z,z'\in B(x,(1-\varepsilon)r)$, then

$$\mathbb{E}^z H(X_{\tau_{B(x,r)}}) \leq c\mathbb{E}^{z'} H(X_{\tau_{B(x,r)}}).$$

Proof. Let τ be the exit time from $B(x,r)$. It suffices to prove the result for $H=1_F$ with $F\subset B(x,2r)^c$. Since $1_F(X_\tau)$ is non-zero only if there is a jump from $B(x,r)$ to $B(x,2r)^c$, we have

$$\begin{aligned}\mathbb{E}^z 1_F(X_{t\wedge\tau}) &= \mathbb{E}^z \int_0^{t\wedge\tau}\int_F \frac{A(X_s,u-X_s)}{|u-X_s|^{d+\alpha}}\,du\,ds \\ &\leq c\mathbb{E}^z\int_0^{t\wedge\tau}\int_F \frac{1}{|u-x|^{d+\alpha}}\,du\,ds,\end{aligned}$$

using the fact that $|u - X_s|$ is comparable to $|u - x|$ if $X_s \in B(x,r)$ and $|u - x| \geq 2r$. If we let $t \to \infty$, we get

$$\mathbb{E}^{z} 1_F(X_\tau) \leq c(\mathbb{E}^{z}\tau) \int_F \frac{du}{|u-x|^{d+\alpha}} \leq cr^\alpha \int_F \frac{du}{|u-x|^{d+\alpha}}.$$

We get the same lower bound (with a different c) for $\mathbb{E}^{z'} 1_F(X_\tau)$ in the same way.

6.2 Harnack inequality

We want to show

Theorem 6.1 *There exists c such that if h is nonnegative and bounded in $\mathbb{R}^d$ and $\mathcal{L}$-harmonic in $B(x_0, 16)$, then*

$$h(x) \leq ch(y), \qquad x, y \in B(x_0, 1).$$

We make a few remarks.

1. We require h bounded on $\mathbb{R}^d$ only to make sure the notion of $\mathcal{L}$-harmonic is well defined. c does not depend on the bound on h.

2. Saying h is $\mathcal{L}$-harmonic means

$$h(X_{t \wedge \tau_{B(x_0,16)}})$$

is a martingale with respect to each starting point.

3. If h is not nonnegative on all of $\mathbb{R}^d$, examples show that the Harnack inequality might not hold.

4. Scaling and usual arguments show that if D is a connected domain, E is a compact subset of D, and h is nonnegative and

bounded on $\mathbb{R}^d$ and $\mathcal{L}$-harmonic on D, then there exists c such that

$$h(x) \le ch(y), \qquad x, y \in E.$$

Proof. Multiplying by a constant, we may assume $\inf_{B(x_0,1)} h = 1/2$. Choose $z_0 \in B(x_0, 1)$ such that $h(z_0) < 1$. (We don't yet know that h is continuous.)

Suppose there exists a point $x_1 \in B(x_0, 1)$ such that $h(x_1) = K$, where K will be chosen later. We will then show that we can find $R = R(K) < 1$, $q < 1$, $\beta > 0$ and a point x_2 such that $|x_2 - x_1| \le Rq$ and $h(x_2) \ge (1 + \beta)K$, and then a point x_3 such that $|x_3 - x_2| \le Rq^2$ with $h(x_3) \ge (1 + \beta)^2 K$, and so on. If we take K large enough so that R is small enough, this will imply that h is unbounded in $B(x_0, 4)$, a contradiction.

We will also choose ζ, η later. Write B_r for $B(x_1, r)$ and τ_r for the corresponding exit time. Choose r such that $|B_{r/3}| = c/(\eta K)$. If A is compact and contained in $A' = \{w \in B_{r/3} : h(w) \ge \zeta K\}$ and T_A is the hitting time of A, then

$$\begin{aligned} 1 \ge h(z_0) &\ge \mathbb{E}^{z_0}[h(X_{T_A \wedge \tau_{16}}); T_A < \tau_{16}] \\ &\ge \zeta K \mathbb{P}^{z_0}(T_A < \tau_{16}) \\ &\ge c\zeta K |A|. \end{aligned}$$

This is true for any compact set contained in A', so is true for A' as well. So we want to choose ζ, K such that $|A'|/|B_{r/3}| \le 1/2$.

With $H = h1_{B^c_{2r}}$, if $x \in B_r$ and

$$\mathbb{E}^x[h(X_{\tau_r}); X_{\tau_r} \notin B_{2r}] \ge \eta K,$$

then for all $y \in B_r$

$$h(y) \ge c\mathbb{E}^x[h(X_{\tau_r}); X_{\tau_r} \notin B_{2r}] \ge c\eta K \ge \zeta K,$$

which can't happen since A' is not all of $B_{r/3}$.

Let $M = \sup_{B_{2r}} h$. Let C be compact and contained in $B(x_1, r/3) \setminus A'$ with $|C|$ at least a third of the volume of $B_{r/3}$. Then

$$\begin{aligned} K = h(x_1) &= \mathbb{E}^{x_1}[h(X_{T_C}); T_C < \tau_r] \\ &\quad + \mathbb{E}^{x_1}[h(X_{\tau_r}); \tau_r < T_C; X_{\tau_r} < T_C, X_{\tau_r} \in B_{2r}] \\ &\quad + \mathbb{E}^{x_1}[h(X_{\tau_r}); \tau_r < T_C; X_{\tau_r} < T_C, X_{\tau_r} \notin B_{2r}] \\ &\le \zeta K \mathbb{P}^{x_1}(T_C < \tau_r) + M\mathbb{P}^{x_1}(\tau_r < T_C) + \eta K, \end{aligned}$$

and hence

$$\frac{M}{K} \ge \frac{1 - \eta - \zeta \mathbb{P}^{x_1}(T_C < \tau_r)}{1 - \mathbb{P}^{x_1}(T_C < \tau_r)} \ge 1 + 2\beta.$$

Thus there must exist $x_2 \in B(x_1, 2r)$ such that $h(x_2) \ge K(1+\beta)$. We have η, ζ small, and $K \approx r^{-d}$. So if K is large, then r is small, and we stay within B_2.

6.3 Regularity

We will prove that harmonic functions are Hölder continuous. Once we have that, we can prove

Theorem 6.2 *If f is bounded and $S_\lambda f$ is bounded, then*

$$S_\lambda f(x) = \mathbb{E}^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt$$

is Hölder continuous for $\lambda \ge 0$.

Proof. Let's first do the case $\lambda = 0$. We write

$$S_0 f(x) - S_0 f(x') = \mathbb{E}^{\,x} \int_0^{\tau_\varepsilon} f(X_s)\,ds + \mathbb{E}^{\,x} S_0 f(X_{\tau_\varepsilon}) \\ - \mathbb{E}^{\,x'} \int_0^{\tau_\varepsilon} f(X_s)\,ds - \mathbb{E}^{\,x'} S_0 f(X_{\tau_\varepsilon}).$$

The first and third terms are bounded by $c\|f\|_\infty \varepsilon^\alpha$ and the second and fourth terms together are less than $c|x-x'|^\beta \|f\|_\infty$ since $S_0 f$ is harmonic in B_ε. We take $\varepsilon = 4|x-x'|^{1\wedge(\beta/\alpha)}$, which gives the Hölder continuity of S_0.

If g has compact support, let $f = g - \lambda S_\lambda g$. Then $S_0 f = S_\lambda g$, which has L^∞ norm bounded by $\|g\|_\infty/\lambda$, and $\|f\|_\infty \le 2\|g\|_\infty$. By taking limits, we can eliminate the assumption of g having compact support.

We now look at harmonic functions.

Theorem 6.3 *If h is bounded on $\mathbb{R}^d$ and $\mathcal{L}$ harmonic in $B(x_0, 2)$, then h is Hölder continuous in $B(x_0, 1)$:*

$$|h(x) - h(y)| \le c\|h\|_\infty |x-y|^\beta, \qquad x, y \in B(x_0, 1).$$

Proof. We may suppose $0 \le h \le M$ on $\mathbb{R}^d$. We will find $\rho < 1$ small and $\gamma < 1$ close to 1 such that

$$\sup_{B(x,\rho^k)} h - \inf_{B(x,\rho^k)} h \le M\gamma^k \tag{6.2}$$

for each k, which will prove the theorem. Let $B_i = B(x, \rho^i)$, $\tau_i = \tau_{B_i}$, $a_i = \inf_{B_i} h$, $b_i = \sup_{B_i} h$.

We use induction. Suppose we have (6.2) for all $i \le k$ and let's prove it for $i = k+1$.

$a_k \le h \le b_k$ on B_{k+1}. Let $A' = \{z \in B_{k+1} : h(z) \le \frac{a_k+b_k}{2}\}$ and suppose $|A'|/|B_{k+1}| \ge 1/2$; if not, look at $M-h$. Pick $y, z \in B_{k+1}$ such that $h(y) \ge b_{k+1} - \varepsilon$, $h(z) \le a_{k+1} + \varepsilon$. Take A compact contained in A' with $|A'|/|B_{k+1}| \ge 1/3$. Then

$$\begin{aligned} h(y) - h(z) = \mathbb{E}^y[h(X_{T_A}) - h(z); T_A < \tau_k] \\ + \mathbb{E}^y[h(X_{\tau_k}) - h(z); \tau_k < T_A, X_{\tau_k} \in B_{k-1}] \\ + \sum_{i=1}^{\infty} \mathbb{E}^y[h(X_{\tau_k}) - h(z); \tau_k < T_A, X_{\tau_k} \in B_{k-i-1} \setminus B_{k-i}]. \end{aligned}$$

The first term is bounded by

$$\Big(\frac{a_k+b_k}{2} - a_k\Big)\mathbb{P}^y(T_A < \tau_k).$$

The second term is bounded by

$$(b_k - a_k)\mathbb{P}^y(\tau_k < T_A).$$

The sum is bounded by

$$\begin{aligned} \sum_{i=1}^{\infty} (b_{k-i-1} - a_{k-i-1})\mathbb{P}^y(X_{\tau_k \notin B_{k-i}}) \\ \le \sum_{i=1}^{\infty} cM\gamma^{k-i-1}(\rho^k)^\alpha/(\rho^{k-i})^\alpha \\ = cM\gamma^k \sum_{i=1}^{\infty} (\rho^\alpha/\gamma)^i \\ \le cM\gamma^k \end{aligned}$$

with c small if we take γ close to 1 and ρ small. Then

$$\begin{aligned} h(y) - h(z) &\le \tfrac{1}{2}(b_k - a_k)\mathbb{P}^y(T_A < \tau_k) + (b_k - a_k)(1 - \mathbb{P}^y(T_A < \tau_k) + cM\gamma^k \\ &\le M\gamma^k(\tfrac{1}{2}P + (1-P) + c) \le M\gamma^{k+1} \end{aligned}$$

if γ and ρ are chosen appropriately; here $P = \mathbb{P}^y(T_A < \tau_k)$. Therefore $b_{k+1} - a_{k+1} \le M\gamma^{k+1} + 2\varepsilon$. Since ε is arbitrary, that does it.

Foondun has considered operators that consist of a diffusion coefficient and a jump part and proved the Harnack inequality for such operators.

In [10] and [11] the Harnack inequality and regularity of harmonic functions with respect to operators of variable order was considered. In the first, the class of operators considered was of the form

$$\mathcal{L}f(x) = \int [f(x+h) - f(x) - 1_{(|h|\le 1)} \nabla f(x) \cdot h] n(x,h)\, dh,$$

where $n(x,h)$ satisfies

$$c_1 \frac{1}{|h|^{d+\alpha}} \le n(x,h) \le c_2 \frac{1}{|h|^{d+\beta}}, \qquad |h| \le 1, \tag{6.3}$$

with an additional assumption when $|h| > 1$. It was shown that the Harnack inequality holds if $0 < \alpha < \beta < 2$ and $\beta - \alpha < 1$. Is the last assumption necessary? Some condition on the large jumps (where $|h| > 1$) was shown to be necessary, but the form seems unduly restrictive. It would be nice to have a better condition.

Here is a problem I particularly like. For operators in nondivergence form:

$$\mathcal{L}f(x) = \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x),$$

with a first order term, it is known that $\mathcal{L}$-harmonic functions are at least Hölder continuous in the interior of the domain.

The same is true for solutions to $\mathcal{L}u = f$. But if the a_{ij} are Hölder continuous of order β, then much more is true: harmonic functions are $C^{2+\beta}$ in the interior, and similarly for the solution to $\mathcal{L}u = f$ is f is also Hölder continuous. The smoother the a_{ij} and f, the smoother the solution. What is the analogue if the nondivergence form operator is replaced by (6.1)? I think this problem is doable and interesting. Harder but even more interesting is if we let $\mathcal{L}$ be of variable order.

7 Symmetric jump processes

Both when studying the heat equation in nonhomogeneous regions and when studying nonlinear PDEs, operators in divergence form (variational form) come up:

$$\mathcal{L}f(x) = \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}(a_{ij}(\cdot)\frac{\partial}{\partial x_j} f(\cdot))(x).$$

Here $a_{ij}(x)$ are the entries in a $d \times d$ matrix for each x. We assume that for each x, the matrix is bounded and positive definite, and the constant that comes up in the definition of positive definite can be taken independently of x. We also assume $a_{ij} = a_{ji}$. However we do not want to assume any smoothness on the a_{ij} in x.

How are we to interpret the PDE

$$\mathcal{L}u = f$$

when the a_{ij} are not smooth? If we multiply both sides by g

and integrate by parts, we get

$$-\int \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial}{\partial x_i} g(x) \frac{\partial}{\partial x_j} u(x)\, dx = \int f(x) g(x)\, dx. \qquad (7.1)$$

So a weak solution is a function u such that (7.1) holds for every smooth g. The left hand side of (7.1) (without the minus sign) is called the Dirichlet form associated with $\mathcal{L}$, and is written $\mathcal{E}(g, u)$. The Bible on Dirichlet forms is [21].

One of the properties of a process associated with a Dirichlet form is symmetry of the transition densities (with respect to the appropriate measure): $p(t, x, y) = p(t, y, x)$.

There is another class of processes that are associated with Dirichlet forms that is easier to visualize. Let the state space be the lattice $\mathbb{Z}^d$, and between any two neighboring states x, y let a conductance C_{xy} be given. A symmetric Markov chain is one where the process waits an exponential time at the current state and then jumps to a neighboring state chosen with probability

$$\mathbb{P}^x(\text{next state is } y) = \frac{C_{xy}}{\sum_z C_{xz}}.$$

As you can imagine, one could allow conductances and jumps between non-neighboring states as well.

Once one has $\mathcal{E}(f, f)$ defined for all f in the domain, we can define $\mathcal{E}(f, g)$ by polarization.

A symmetric stable process of index α is a symmetric Markov process and its Dirichlet form is

$$\mathcal{E}(f, f) = \int\int \frac{(f(x) - f(x))^2}{|x - y|^{d+\alpha}}\, dx\, dy.$$

If f is C^1, then the numerator is bounded by $c|x-y|^2$, which insures convergence of the integral.

In this section we will consider processes whose Dirichlet form is given by

$$\mathcal{E}(f,f) = \int\int \frac{(f(x)-f(x))^2}{|x-y|^{d+\alpha}} N(x,y)\,dx\,dy,$$

where $N(x,y)$ is symmetric, bounded above, and bounded below by a positive constant. We sometimes write

$$J(x,y) = \frac{N(x,y)}{|x-y|^{d+\alpha}}.$$

We will show (sketch!) the main result concerning the processes associated with this class of Dirichlet forms. The proofs are adapted from those in [1].

Theorem 7.1 *There exist constants c_1, c_2 such that the transition densities of the process satisfy*

$$c_1\Big(t^{-d/\alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\Big) \le p(t,x,y) \le c_2\Big(t^{-d/\alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\Big).$$

Once one has a good handle on the transition densities, one is usually able to say almost everything about the process that one would like.

We will sometimes look at

$$J_\delta(x,y) = J(x,y)1_{(|x-y|\le\delta)},$$

which is the truncated jump kernel. This corresponds to the process with all jumps larger than δ removed.

We point out that the corresponding processes need not be semimartingales.

7.1 Upper bound

We first make an observation. If we have an upper bound for $p(t,x,x)$ for all x, then we automatically have a bound for $p(t,x,y)$, although it is unlikely to be a good one when x and y are far apart. This follows from the Chapman-Kolmogorov equations.

$$\begin{aligned} p(2t,x,y) &= \int p(t,x,z)p(t,z,y)\,dz \\ &\le \Big(\int p(t,x,z)^2\,dz\Big)^{1/2}\Big(\int p(t,y,z)^2\,dz\Big)^{1/2}. \end{aligned}$$

But

$$\int p(t,x,z)^2\,dz = \int p(t,x,z)p(t,z,x)\,dz = p(2t,x,x).$$

In proving the upper bound, there are two tools we need. The first is a result of [18], which is an extension of a method of Davies, which in turn is based on ideas of Nash.

Let

$$\begin{aligned} \Gamma(f,f)(x) &= \int (f(x)-f(y))^2 J(x,y)\,dx\,dy, \\ \Lambda(\psi) &= \|e^{-2\psi}\Gamma(e^{\psi},e^{\psi})\|_\infty \vee \|e^{2\psi}\Gamma(e^{-\psi},e^{-\psi}))\|_\infty, \\ E(t,x,y) &= \sup\{|\psi(x)-\psi(y)| - t\Lambda(\psi)^2 : \psi\in C_b, \Lambda(\psi)<\infty\}. \end{aligned}$$

Then we have [18, Theorem 3.25]:

Theorem 7.2 *Suppose*

$$\|f\|_2^{2+4/\nu} \le c(\mathcal{E}(f,f) + B\|f\|_2^2)\|f\|_1^{4/\nu}.$$

Then for all t, x, y,

$$p(t,x,y) \le ct^{-\nu/2} e^{Bt} e^{-E(2t,x,y)}.$$

The basic idea of the proof is the following, due to Nash. Fix x_0 and let $F(t) = \int p(t,x_0,x)^2\,dx$. Suppose $B = 0$. Then

$$\begin{aligned} F'(t) &= \int 2p(t,x_0,x)\frac{\partial p(t,x,x_0)}{\partial t}\,dx \\ &= 2\int p(t,x_0,x)\mathcal{L}p(t,x_0,x)\,dx \\ &= -2\mathcal{E}(p(t,x_0,\cdot),p(t,x_0,\cdot)) \\ &\le -c\|p(t,x_0,\cdot)\|_2^{2+4/\nu}, \end{aligned}$$

since the L^1 norm of $p(t,x_0,\cdot)$ is 1. Therefore

$$F'(t) \le -cF(t)^{2+4/\nu},$$

and solving this differential inequality (just the way you solve a differential equation of this sort), we get

$$F(t) \le ct^{-\nu/2}.$$

But $F(t) = p(2t,x_0,x_0)$ by the semigroup property.

Davies' idea was to extend this by applying the above argument to

$$p^\psi(t,x,y) = e^{\psi(x)}p(t,x,y)e^{-\psi(y)}$$

and using L^p instead of L^2. This gives the off-diagonal bounds, after some complicated calculations and writing the L^{2p} bound in terms of the L^p bound. (See [5, Section VII.5] for the proof in the diffusion case.)

We will apply this to $J_\delta(x,y)$, a truncated version of J. So we also need a way to go from the transition densities for the process corresponding to J_δ to the ones for J. There are two ways of doing this. One, more analytic, is to notice that if one takes the infinitesimal generator for the process corresponding to J, it is a perturbation of the one for J_δ by a bounded operator, and just as we can write the resolvent for a perturbed operator in terms of the resolvent for the unperturbed operator, there are formulas (more complicated) for writing the semigroup for the perturbed operator in terms of the unperturbed one.

A more probabilistic approach is one of Meyer [26]. Let X_t^δ be the process corresponding to J_δ. Let $W(x) = \int (J(x,y) - J_0(x,y))\,dy$. Let S_1 be an exponential random variable with parameter 1 independent of X^δ. Run X^δ until the first time T_1 that $\int_0^t W(X_s^\delta)\,ds$ exceeds S_1. At this time introduce a jump from $X_{T_1-}^\delta$ to y, where the distribution of y is given according to

$$\frac{(J - J_0)(X_{T_1-}^\delta, y)}{W(X_{T_1-}^\delta)}.$$

Then start the process at y, let it run according to the law of X^δ, and repeat the process. The new process, X_t, will have Dirichlet form defined in terms of J. (Meyer's construction is more general, and applies to any strong Markov process, not just symmetric ones.)

We then have a result due to [3]:

Lemma 7.3

$$p(t,x,y) \le p^\delta(t,x,y) + t\|(J - J_\delta)\|_\infty.$$

Proof. We first observe that

$$\mathbb{P}^x(X_t \in B) = \mathbb{P}^x(X_t^\delta \in B, T_1 > t) + \mathbb{E}^x \int_0^t \int_B r_{t-s}(X_s^\delta, z) W(X_s^\delta)\,dz\,ds,$$

where

$$r_t(y,z) = \int \frac{(J-J_\delta)(y,w)}{W(y)} p(t,w,z)\,dw.$$

The first term on the right is the contribution if the first introduced jump hasn't happened yet. The second term comes from jumping at time s for some $s \le t$ (the rate of jumping is where the $W(X^\delta_s)$ term comes from), then moving according to X for a time $t-s$, and then ending up in B. It follows that

$$p(t,x,y) \le p^\delta(t,x,y) + \mathbb{E}^x \int_0^t r_{t-s}(X^\delta_s, y) W(X^\delta_s)\,ds.$$

The second term is bounded by $t\|J-J_\delta\|_\infty$ because

$$\begin{aligned} r_{t-s}(v,y)W(v) &= \int (J-J_\delta)(v,w)p(t-s,w,z)\,dw \\ &\le \|J-J_\delta\|_\infty \int p(t-s,w,z)\,dw = \|J-J_\delta\|_\infty. \end{aligned}$$

We apply our two tools as follows. We have

$$\mathcal{E}(f,f) = \mathcal{E}_\delta(f,f) + \int\int_{|x-y|>\delta} \frac{(f(x)-f(y))^2}{|y-x|^{d+\alpha}} N(x,y)\,dy\,dx.$$

To bound the second term, note $(f(x)-f(y))^2 \le 2f(x)^2 + 2f(y)^2$. We have

$$\int\int_{|x-y|>\delta} \frac{f(x)^2}{|x-y|^{d+\alpha}} N(x,y)\,dy\,dx \le c\delta^{-\alpha} \int f(x)^2\,dx = c\delta^{-\alpha}\|f\|_2^2,$$

and similarly for the term with $f(y))^2$ in the numerator.

If we can prove a Nash inequality for $\mathcal{E}$, we can thus obtain a Nash inequality for $\mathcal{E}_\delta$. To get the Nash inequality for $\mathcal{E}$ we use a fact from the theory of Besov spaces, or more simply, the facts that for a symmetric stable process with Dirichlet form $\mathcal{E}_{SS}$: (a) $\mathcal{E}_{SS}$ and $\mathcal{E}$ are comparable; (b) the transition densities for a symmetric stable process are bounded by $ct^{-d/\alpha}$ (this comes from scaling and the fact that $p(1,\cdot,\cdot)$ is bounded: the Fourier transform is $e^{-|u|^\alpha}$); and (c) another (easier) theorem from [18] which says that if the transition densities are bounded by $t^{-\mu}$ for some μ, then the corresponding Nash inequality holds.

Now let us obtain bounds on $p(t,x_0,y_0)$. By scaling it is enough to take $t=1$, and by the Nash argument we gave above, we see that $p(1,x_0,x_0)$ is bounded. Now let $R=|y_0-x_0|$,

$$\psi(x)=\lambda(R-|x-x_0|)^+,$$

where λ will be chosen later, let M be larger than $(d+\alpha)/\alpha$, and let us truncate at $\delta=R/3M$. Since $|\psi(x)-\psi(y)|\le\lambda|x-y|$ and $|e^t-1|^2\le t^2e^{2t}$, we have (here all the Λ,Γ,E's are defined in terms of J_δ),

$$\begin{aligned} e^{-2\psi(x)}\Gamma(e^\psi,e^\psi)(x) &= \int_{|x-y|\le\delta}(e^{\psi(x)-\psi(y)}-1)^2\frac{N(x,y)}{|x-y|^{d+\alpha}}\,dy\\ &\le ce^{2\lambda\delta}\lambda^2\int_{|x-y|\le\delta}|x-y|^2\frac{dy}{|x-y|^{d+\alpha}}\\ &\le ce^{2\lambda\delta}\lambda^2\delta^{2-\alpha}\le ce^{3\lambda\delta}\delta^{-\alpha}. \end{aligned}$$

So

$$-E(2,x_0,y_0)\le-\lambda R+c\delta^{-\alpha}e^{3\lambda\delta}.$$

Take

$$\lambda=\frac{1}{3\delta}\log(\delta^\alpha).$$

Observing that $\lambda R=M\log(\delta^\alpha)$, we obtain

$$p^\delta(1,x_0,y_0)\le cR^{-(d+\alpha)}.$$

Since we could have taken M even larger, the term involving p^δ is not the important one. We then calculate $\|J - J_0\|_\infty \le c\delta^{-(d+\alpha)}$, and see that

$$p(1, x_0, y_0) \le c\delta^{-(d+\alpha)/\alpha} \le cR^{-(d+\alpha)},$$

which is what we want.

For use in the lower bound, and also because it is useful, we obtain

Theorem 7.4 *We have*

$$\mathbb{P}^x(\sup_{s\le t} |X_s - X_0| > L) \le cL^{-\alpha}.$$

Proof. Let T be the first time X moves more than L. Then

$$\begin{aligned}\mathbb{P}^x(T \le t) &\le \mathbb{P}^x(|X_t - X_0| \ge L/2) + \mathbb{P}^x(T \le t, |X_t - X_0| < L/2)\\ &\le \mathbb{P}^x(|X_t - X_0| \ge L/2) + \int_0^t \mathbb{P}^x(|X_t - X_s| \ge L/2, T \in ds).\end{aligned}$$

The first term is bounded by integrating the transition probabilities. The second term is bounded by

$$\int_0^t \frac{c(t-s)}{L^\alpha}\mathbb{P}^x(T \in ds) \le \frac{c}{L^\alpha}.$$

7.2 Lower bounds

The original idea is again due to Nash, although Fabes and Stroock pushed it much further. We first want to show that $p(1, x, y)$ is not too small if $|x - y|$ is not too large. With positive probability the exponential random variable in Meyer's construction will be sufficiently large that, using independence, the lower bound on $p(1,x,y)$ will follow if we prove the same fact with p replaced by p^δ, where $\delta = 1$.

By scaling, this will prove $p(t,x,y) \geq ct^{-d/\alpha}$ if $|x-y|^{d+\alpha} \leq t$. For the case when $|x-y|^{d+\alpha} > t$, we then reason as follows: using scaling it is enough to take $t = 1$. Starting at x there is some probability that the process stays near x until it makes one large jump into $B(y, 1/2)$ before time $1/4$, and then stays in $B(y,1)$ for a time $1/2$. We thus get a lower bound on going from x to $B(y,1)$ in time $1/2$. Then

$$\begin{aligned} p(1,x,y) &= \int p(1/2,x,z)p(1/2,z,y)\,dz \\ &\geq \int_{B(y,1)} p(1/2,x,z)p(1/2,z,y)\,dz \\ &\geq \mathbb{P}^x(X_{1/2} \in B(y,1)) \inf_{z\in B(y,1)} p(1/2,z,y). \end{aligned}$$

Putting in the appropriate estimates gives the right bound.

It clearly suffices to get a lower bound on $\overline{p}(1,x,y)$ when $|x-y| \leq 1$ and $\overline{p}$ is the transition density for X^δ killed on exiting $B = B(x_0, 4)$.

Let

$$G(t) = \int \log \overline{p}(t,x_0,x)\varphi(x)\,dx,$$

where φ is an appropriate weight function that is bounded above and has integral 1. (We'll specify φ shortly.) If we show

$$G(1) \geq -R \tag{7.2}$$

for some positive real R and similarly when x_0 is replaced by a nearby x_1, then by Jensen's inequality

$$\begin{aligned}\log \overline{p}(2, x_0, x_1) &= \log \Big[\int \overline{p}(1, x_0, z) \overline{p}(1, x_1, z)\, dz \Big] \\ &\geq \log \Big[c \int \overline{p}(1, x_0, z) \overline{p}(1, x_1, z) \varphi(z)\, dz \Big] \\ &\geq \log c + \int \log(\overline{p}(1, x_0, z) \overline{p}(1, x_1, z)) \varphi(z)\, dz \\ &= \log c + \int \log \overline{p}(1, x_0, z) \varphi(z)\, dz + \int \log \overline{p}(1, x_1, z) \varphi(z)\, dz \\ &\geq \log c - 2R,\end{aligned}$$

and taking exponentials gives our result.

So it comes down to proving (7.2). Let $\varphi(x) = c(16 - |x - x_0|^2)^+$, where c is chosen so that the integral over $\mathbb{R}^d$ of φ is 1.

The usual Poincaré inequality says that

$$\int_B (f(x) - f_B)^2\, dx \leq c \int_B |\nabla f(x)|^2\, dx,$$

where f_B is the average of f over B, where B is a ball or a cube. A weighted Poincaré inequality is one where Lebesgue measure is replaced by some other measure. Our weighted inequality is the following:

Theorem 7.5 *We have*

$$\int (f(x) - \overline{f})^2 \varphi(x)\, dx \leq c \int \int (f(x) - f(y))^2 (\varphi(x) \wedge \varphi(y))\, dy\, dx,$$

where

$$\overline{f} = \int f(x) \varphi(x)\, dx.$$

There are standard ways to prove this, see, e.g., [22].

Now we work on (7.2). We have

$$\begin{aligned}
G'(t) &= \int \frac{\partial \log \overline{p}(t, x_0, x)}{\partial t} \varphi(x)\, dx \\
&= \int \frac{\mathcal{L}\overline{p}(t, x_0, x)}{\overline{p}(t, x_0, x)} \varphi(x)\, dx \\
&= -\mathcal{E}\Big(\overline{p}(t, x_0, \cdot), \frac{\varphi}{\overline{p}(t, x_0, \cdot)}\Big) \\
&= -\int\int \Big\{ [\overline{p}(t, x_0, y) - \overline{p}(t, x_0, x)] \\
&\qquad \times \Big[\frac{\varphi(y)}{\overline{p}(t, x_0, y)} - \frac{\varphi(x)}{\overline{p}(t, x_0, x)}\Big]\Big\} J_\delta(x, y)\, dx\, dy.
\end{aligned}$$

Now some algebra and calculus. Let $a = \overline{p}(t, x_0, y)/\overline{p}(t, x_0, x)$, $b = \varphi(y)/\varphi(x)$. The part of the integrand in the last line inside the braces is equal to

$$\varphi(x)\Big[\sqrt{b}\Big(\frac{a}{\sqrt{b}} + \frac{\sqrt{b}}{a} - 2\Big) - (1 - \sqrt{b})^2\Big].$$

Setting $A = a/\sqrt{b}$ and using the inequality

$$A + \frac{1}{A} - 2 \geq (\log A)^2,$$

the expression inside the braces is larger than

$$\varphi(x)\sqrt{b}(\log A)^2 - \varphi(x)(1 - \sqrt{b})^2.$$

Now $\varphi\sqrt{b} \geq \varphi(x) \wedge \varphi(y)$, and from the explicit expression for $\varphi(x)$ and $\varphi(y)$, we have

$$\int\int (\sqrt{\varphi(x)} - \sqrt{\varphi(y)})^2 J_\delta(x, y)\, dx\, dy \leq c.$$

Since $J_\delta(x,y)$ is 0 if $|x-y| \geq 1$ and is bounded below when $|x-y| < 1$, and $\varphi(x)$ and $\varphi(y)$ are comparable when $|x-y| \leq 1$, we obtain

$$G'(t) \geq \int_B \int_B \Big(\log \frac{\overline{p}(t,x_0,y)}{\sqrt{\varphi(y)}} - \log \frac{\overline{p}(t,x_0,x)}{\sqrt{\varphi(x)}} \Big)^2 \varphi(y)\, dy\, dx - c.$$

Let

$$K(t) = \int_B \log \Big(\frac{\overline{p}(t,x_0,y)}{\sqrt{\varphi(y)}} \Big) \varphi(y)\, dy.$$

Then using the Poincaré inequality,

$$G'(t) \geq \int_B \Big(\log \frac{\overline{p}(t,x_0,y)}{\sqrt{\varphi(y)}} - K(t) \Big)^2 \varphi(y)\, dy - c.$$

Using the inequality

$$(A-B)^2 \geq \frac{A^2}{2} - B^2$$

with $A = \log \overline{p}(t,x_0,y) - G(t)$ and $B = -\frac{1}{2} \log \varphi(y) - c$ and the fact that $\int B^2 \varphi(y)\, dy$ is a constant not depending on x or t, we now have

$$G'(t) \geq c_1 \int (\log \overline{p}(t,x_0,y) - G(t))^2 \varphi(y)\, dy - c_2.$$

This is greater than

$$c_1 \int_{D_t} (\log \overline{p}(t,x_0,y) - G(t))^2 \varphi(y)\, dy - c_2,$$

where we will choose D_t in a moment.

We use Theorem 7.4 to find r such that

$$\int_{B(x_0,r)^c} \overline{p}(t,x_0,x)\, dx \leq \tfrac{1}{4}, \qquad t \leq 1,$$

and then choose K so that $|B(x_0,r)|e^{-K} \le \frac{1}{4}$. If $D_t = \{y \in B(x_0,r) : \overline{p}(t,x_0,y) \ge e^{-K}\}$, we have

$$\begin{aligned} \tfrac{3}{4} &\le \int_{B(x_0,r)} \overline{p}(t,x_0,x)\,dx = \int_{D_t} + \int_{B(x_0,r)\setminus D_t} \\ &\le c|D_t|t^{-d/\alpha} + |B(x_0,r)|e^{-K}, \end{aligned}$$

or $|D_t| \ge c$ if $t \in [\frac{1}{2}, 1]$.

Some calculus (cf. the last part of Theorem 30.1 of my lecture notes *PDE from a probabilistic point of view* or the proof of Proposition 4.9 in [1]) eventually leads to

$$G'(t) \ge c|D_t|(G(t)^2 - S)^2 - c \ge IG(t)^2 - J, \qquad t \in [\tfrac{1}{2}, 1].$$

Some more calculus then gives us our desired lower bound.

Foondun has considered Dirichlet forms that have a diffusion component and a jump component. Chen and Kumagai [19] have considered more general state spaces. In [1] variable order operators were considered. Some of these reults were generalized by Kassmann and his co-authors.

In [12] and work in progress, the weak convergence of symmetric Markov chains to a process corresponding to elliptic operators in divergence form or to operators with jumps is considered. There is a lot to be done here.

In [1], a condition similar to (6.3) was imposed when the jumps were of size less than 1 and no jumps larger than 1 were allowed. (Here, though, the full range of α, β was premitted.) Which of the results generalize when jumps larger than 1 are included? I believe some interesting things occur.

At the end of the last section I mentioned the problem of showing harmonic functions and solutions to $\mathcal{L}u = f$ were smoother than Hölder continuous if the function $A(x,h)$ was smoother. There is the analoguous problem for the case of symmetric jump processes.

References

[1] M.T. Barlow, R.F. Bass, Z.-Q. Chen, and M. Kassmann, Non-local Dirichlet forms and symmetric jump processes, *Trans. Amer. Math. Soc.* **361** (2009) 1963–1999.

[2] M.T. Barlow, R.F. Bass, and T. Kumagai, Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps, *Math. Z.* **261** (2009) 297–320.

preprint.

[3] M.T. Barlow, A. Grigor'yan and T. Kumagai. Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.*, to appear.

[4] R.F. Bass, Uniqueness in law for pure jump Markov processes, *Probab. Th. & related Fields* **79** (1988), 271–287.

[5] R.F. Bass, *Diffusions and Elliptic Operators.* Springer, New York, 1997.

[6] R.F. Bass, Stochastic differential equations driven by symmetric stable processes, *Séminaire de Probabilités XXXVI*, 302–313, Springer, New York, 2003.

[7] R.F. Bass, Stochastic differential equations with jumps, *Prob. Surveys*, **1** (2004) 1–19.

[8] R.F. Bass, K. Burdzy, and Z.-Q. Chen, Stochastic differential equations driven by stable processes for which pathwise uniqueness fails, *Stoch. Proc. & their Applic.* **111** (2004) 1–15.

[9] R.F. Bass and Z.-Q. Chen, Systems of equations driven by stable processes, *Probab. Th. rel. Fields* **134** (2006) 175–214.

[10] R.F. Bass and M. Kassmann, Harnack inequalities for non-local operators of variable order, *Trans. Amer. Math. Soc.* **357** (2005) 837–850.

[11] R.F. Bass and M. Kassmann, Hölder continuity of harmonic functions with respect to operators of variable order, *Comm. PDE* **30** (2005) 1249–1259.

[12] R.F. Bass and T. Kumagai, Symmetric Markov chains on $\mathbb{Z}^d$ with unbounded range, *Trans. Amer. Math. Soc.* **360** (2008) 2041–2075.

[13] R.F. Bass and D.A. Levin, Harnack inequalities for jump processes, *Potential Anal.* **17** (2002) 375–388.

[14] R.F. Bass and D.A. Levin, Transition probabilities for symmetric jump processes, *Trans. Amer. Math. Soc.* **354** (2002) 2933–2953.

[15] R.F. Bass and H. Tang, The martingale problem for a class of stable-like processes, *Stoch. Proc. Appl.& their Applic.* **119** (2009) 1144–1167.

[16] R.F. Bass and D. You, A Fatou theorem for α-harmonic functions, *Bull. Sciences Math.* **127** (2003) 635–648.

[17] R.F. Bass and D. You, A Fatou theorem for α-harmonic functions in Lipschitz domains, *Prob. Th. rel. Fields* **133** (2005) 391–408.

[18] E. A. Carlen, S. Kusuoka, and D. W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré Probab. Statist.*, **23(2, suppl.)** (1987), 245–287.

[19] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for stable-like processes on d-sets. *Stoch. Proc. Applic.*, **108** (2003), 27–62.

[20] R.J. Elliott. *Stochastic calculus and applications.* Springer-Verlag, New York, 1982.

[21] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes*, de Gruyter, Berlin, 1994.

[22] D. Jerison, The Poincaré inequality for vector fields satisfying Hörmander's condition. *Duke Math. J.* **53** (1986) 503–523.

[23] T. Komatsu, On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations of jump type. *Proc. Japan Acad. Ser. A Math. Sci.* **58** (1982) 353–356.

[24] J.-P. Lepeltier and B. Marchal, Probléme des martingales et équations différentielles stochastiques associés à un opérateur intégro-différentiel. *Ann. Inst. H. Poincaré Sect. B* **12** (1976) 43–103.

[25] P.-A. Meyer, *Un cours sur les intégrales stochastiques*, In: Séminaire de Probabilités X, Berlin, Springer, 1976.

[26] P.-A. Meyer. Renaissance, recollements, mélanges, ralentissement de processus de Markov. *Ann. Inst. Fourier*, **25** (1975), 464–497.

[27] R. Song and J.-M. Wu. Boundary Harnack principle for symmetric stable processes. *J. Funct. Anal.* **168** (1999) 403–427.

[28] H. Tang, Ph.D. dissertation, University of Connecticut, 2006.

Richard F. Bass
Department of Mathematics
University of Connecticut
Storrs, CT 06269-3009, USA
bass@math.uconn.edu