PDE from a probability point of view

Richard F. Bass
University of Connecticut

These notes are for a course I gave while on sabbatical at UBC. It presupposes the reader is familiar with stochastic calculus; see the notes on my web page for Stochastic Calculus, for example. These notes for the most part are based on my book *Diffusions and Elliptic Operators*, Springer-Verlag, 1997.

The topics covered are: stochastic differential equations, solving PDEs using probability, Harnack inequalities for nondivergence form elliptic operators, martingale problems, and divergence form elliptic operators.

We use ∂_i and ∂_{ij} for $\partial/\partial x_i$ and $\partial^2/\partial x_i \partial x_j$, resp.

1. Pathwise solutions.

Let W_t be a one-dimensional Brownian motion. We will be concerned with the *stochastic differential equation* (SDE)

$$dX_t = \sigma(X_t)\, dW_t + b(X_t)\, dt, \qquad X_0 = x. \tag{1.1}$$

This is a shorthand way of writing

$$X_t = x + \int_0^t \sigma(X_s)\, dW_s + \int_0^t b(X_s)\, ds. \tag{1.2}$$

Here σ and b are measurable real-valued functions. We will say (1.1) or (1.2) has a solution if there exists a continuous adapted process X_t satisfying (1.2). X_t is necessarily a semimartingale. Later on we will talk about various types of solutions, so to be more precise, we say that X_t is a *pathwise solution*. We say that we have *pathwise uniqueness* for (1.1) or (1.2) if whenever X_t and X'_t are two solutions, then there exists a set N such that $\mathbb{P}(N) = 0$ and for all $\omega \notin N$, we have $X_t = X'_t$ for all t.

The definitions for the higher-dimensional analogues of (1.1) and (1.2) are the same. Let σ_{ij} be measurable functions for $i, j = 1, \ldots, d$ and b_i measurable functions for $i = 1, \ldots, d$. Let W_t be a d-dimensional Brownian motion. We consider the equation

$$dX_t = \sigma(X_t)\, dW_t + b(X_t)\, dt, \qquad X_0 = x, \tag{1.3}$$

or equivalently, for $i = 1, \ldots, d$,

$$X_t^i = x_i + \int_0^t \sum_{j=1}^d \sigma_{ij}(X_s)\, dW_s^j + \int_0^t b_i(X_s)\, ds. \tag{1.4}$$

Here $X_t = (X_t^1, \ldots, X_t^d)$ is a semimartingale on $\mathbb{R}^d$.

The connection between stochastic differential equations and partial differential equations comes about through the following theorem, which is simply an application of Itô's formula. Let σ^T denote the transpose of the matrix σ and let a be the matrix $\sigma\sigma^T$. Let $C^2(\mathbb{R}^d)$ be the functions on $\mathbb{R}^d$ whose first and second partial derivatives are continuous, let $C_b^2(\mathbb{R}^d)$ be those functions in $C^2(\mathbb{R}^d)$ that are bounded and have bounded first and second derivatives, and let $\mathcal{L}$ be the operator on $C^2(\mathbb{R}^d)$ defined by

$$\mathcal{L}f(x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \partial_{ij} f(x) + \sum_{i=1}^d b_i(x) \partial_i f(x). \tag{1.5}$$

Proposition 1.1. *Suppose X_t is a solution to (1.3) with σ and b bounded and measurable and let $f \in C_b^2(\mathbb{R}^d)$. Then*

$$f(X_t) = f(X_0) + M_t + \int_0^t \mathcal{L}f(X_s)\, ds, \tag{1.6}$$

where

$$M_t = \int_0^t \sum_{i,j=1}^d \partial_i f(X_s) \sigma_{ij}(X_s)\, dW_s^j \tag{1.7}$$

is a martingale.

Proof. Since the components of the Brownian motion W_t are independent, we have $d\langle W^k, W^\ell \rangle_t = 0$ if $k \neq \ell$. Therefore

$$\begin{aligned} d\langle X^i, X^j \rangle_t &= \sum_k \sum_\ell \sigma_{ik}(X_t) \sigma_{jl}(X_t)\, d\langle W^k, W^\ell \rangle_t \\ &= \sum_k \sigma_{ik}(X_t) \sigma_{kj}^T(X_t)\, dt = a_{ij}(X_t)\, dt. \end{aligned}$$

We now apply Itô's formula:

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_i \int_0^t \partial_i f(X_s)\, dX_s^i + \frac{1}{2} \int_0^t \sum_{i,j} \partial_{ij} f(X_s)\, d\langle X^i, X^j \rangle_s \\ &= f(X_0) + M_t + \sum_i \int_0^t \partial_i f(X_s) b_i(X_s)\, ds \\ &\qquad + \frac{1}{2} \int_0^t \sum_{i,j} \partial_{ij} f(X_s) a_{ij}(X_s)\, ds \\ &= f(X_0) + M_t + \int_0^t \mathcal{L}f(X_s)\, ds. \end{aligned}$$

□

We will say that a process X_t and an operator $\mathcal{L}$ are *associated* if X_t satisfies (1.3) for $\mathcal{L}$ given by (1.5) and $a = \sigma\sigma^T$. We call the functions b the *drift coefficients* of X_t and of $\mathcal{L}$, and we call σ and a the *diffusion coefficients* of X_t and $\mathcal{L}$, respectively.

2. Lipschitz coefficients.

We now proceed to show existence and uniqueness for the SDE (1.1) when the coefficients σ and b are Lipschitz continuous. For notational simplicity, we first consider the case where the dimension is one. Recall that a function f is *Lipschitz* if there exists a constant c_1 such that $|f(x) - f(y)| \le c_1|x - y|$ for all x, y.

Theorem 2.1. *Suppose σ and b are Lipschitz and bounded. Then there exists a pathwise solution to the SDE (1.1).*

Proof. We use Picard iteration. Define $X^0(t) \equiv x$ and define inductively

$$X^{i+1}(t) = x + \int_0^t \sigma(X^i(s))\, dW_s + \int_0^t b(X^i(s))\, ds \tag{2.1}$$

for $i = 0, 1, \ldots$. Note

$$\begin{aligned} X^{i+1}(t) - X^i(t) = &\int_0^t [\sigma(X^i(s)) - \sigma(X^{i-1}(s))]\, dW_s \\ &+ \int_0^t [b(X^i(s)) - b(X^{i-1}(s))]\, ds. \end{aligned} \tag{2.2}$$

Let $g_i(t) = \mathbb{E}\,[\sup_{s\le t} |X^{i+1}(s) - X^i(s)|^2]$.

If F_t denotes the first term on the right-hand side of (2.2), then by Doob's inequality,

$$\begin{aligned} \mathbb{E} \sup_{s\le t} F_s^2 &\le c_1 \mathbb{E} \int_0^t |\sigma(X^i(s)) - \sigma(X^{i-1}(s))|^2\, ds \\ &\le c_2 \int_0^t \mathbb{E}\,|X^i(s) - X^{i-1}(s)|^2\, ds \\ &\le c_2 \int_0^t g_{i-1}(s)\, ds. \end{aligned} \tag{2.3}$$

If G_t denotes the second term on the right-hand side of (2.2), then by the Cauchy-Schwarz inequality,

$$\mathbb{E} \sup_{s\le t} G_s^2 \le \mathbb{E} \Big(\int_0^t |b(X^i(s)) - b(X^{i-1}(s))|\, ds \Big)^2 \tag{2.4}$$

$$\leq \mathbb{E}\, t \int_0^t |b(X^i(s)) - b(X^{i-1}(s))|^2 ds$$
$$\leq c_3 t \int_0^t \mathbb{E}\, |X^i(s) - X^{i-1}(s)|^2\, ds$$
$$\leq c_3 t \int_0^t g_{i-1}(s)\, ds.$$

So (2.2), (2.3), (2.4), and the inequality $(x+y)^2 \leq 2x^2 + 2y^2$ tell us that there exists A such that

$$g_i(t) \leq 2\mathbb{E} \sup_{s \leq t} F_s^2 + 2\mathbb{E} \sup_{s \leq t} G_s^2 \leq A(1+t) \int_0^t g_{i-1}(s)\, ds. \tag{2.5}$$

Since σ and b are bounded, arguments to those in the derivation of (2.3) and (2.4) show that $g_0(t)$ is bounded by $B(1+t)$ for some constant B. Iterating (2.5),

$$g_1(t) \leq A \int_0^t B(1+s)\, ds \leq AB(1+t)^2/2$$

for all t, so

$$g_2(t) \leq A \int_0^t (AB(1+s)^2)/2\, ds \leq A^2 B(1+t)^3/3!$$

for all t. By induction,

$$g_i(t) \leq A^i B(1+t)^{i+1}/(i+1)!$$

Hence $\sum_{i=0}^\infty g_i(t)^{1/2} < \infty$. Fix t and define the norm

$$\|Y\| = (\mathbb{E} \sup_{s \leq t} |Y_s|^2)^{1/2}. \tag{2.6}$$

One can show that this norm is complete. We then have that

$$\|X^n - X^m\| \leq \sum_{i=n}^{m-1} g_i(t)^{1/2} \to 0$$

if $m > n$ and $m, n \to \infty$. Therefore X^n is a Cauchy sequence with respect to this norm. Then there is a process X such that $\|X^n - X\| \to 0$ as $n \to \infty$. For each t, we can look at a subsequence so that $\sup_{s \leq t} |X(s) - X^{n_j}(s)| \to 0$ a.s., which implies that $X(s)$ has continuous paths. Letting $i \to \infty$ in (2.1), we see that $X(s)$ satisfies (1.2). $\square$

Uniqueness will be shown next. We first examine a portion of the proof that is known as Gronwall's lemma.

Lemma 2.3. (Gronwall's lemma) *Suppose $g : [0,\infty) \to \mathbb{R}$ is bounded on each finite interval, is measurable, and there exist A and B such that for all t*

$$g(t) \le A + B \int_0^t g(s)\,ds. \tag{2.7}$$

Then $g(t) \le Ae^{Bt}$ for all t.

Proof. Iterating the inequality for g,

$$\begin{aligned} g(t) &\le A + B\int_0^t \Big[A + B\int_0^s g(r)\,dr\Big]\,ds \\ &\le A + ABt + B^2 \int_0^t \int_0^s \Big[A + B\int_0^r g(q)\,dq\Big]\,ds\,dt \\ &= A + ABt + AB^2t^2/2 + B^3 \int_0^t \int_0^s \int_0^r g(q)\,dq\,dr\,ds \\ &\le \cdots . \end{aligned}$$

Since g is bounded on $[0,t]$, say by C, then

$$\int_0^t g(s)\,ds \le Ct, \qquad \int_0^t \int_0^s g(r)\,dr\,ds \le \int_0^t Cs\,ds \le Ct^2/2!,$$

and so on. Hence

$$g(t) \le Ae^{Bt} + B^n C t^n / n!$$

for each n. Letting $n \to \infty$ completes the proof. □

Theorem 2.4. *Suppose σ and b are Lipschitz and bounded. Then the solution to the SDE (1.1) is pathwise unique.*

Proof. Suppose X_t and X'_t are two pathwise solutions to (1.1). Let

$$g(t) = \mathbb{E} \sup_{s \le t} |X_s - X'_s|^2.$$

Since X_t and X'_t both satisfy (1.1), their difference satisfies

$$X_t - X'_t = H_t + I_t,$$

where

$$H_t = \int_0^t [\sigma(X_s) - \sigma(X'_s)]\,dW_s, \qquad I_t = \int_0^t [b(X_s) - b(X'_s)]\,ds.$$

As in the proof of Theorem 2.1, there exist c_1 and c_2 such that

$$\mathbb{E} \sup_{s\le t} H_s^2 \le c_1 \int_0^t g(s)\,ds, \qquad \mathbb{E} \sup_{s\le t} I_s^2 \le c_2 t \int_0^t g(s)\,ds.$$

Hence, if t_0 is a positive real and $t \le t_0$, there exists a constant c_3 depending on t_0 such that

$$g(t) \le 2\mathbb{E} \sup_{s\le t} H_s^2 + 2\mathbb{E} \sup_{s\le t} I_s^2 \le c_3 \int_0^t g(s)\,ds.$$

By Lemma 2.3, $g(t) = 0$ for $t \le t_0$. Since t_0 is arbitrary, uniqueness is proved. □

It is often useful to be able to remove the boundedness assumption on σ and b. We still want σ and b to be Lipschitz, so this can be phrased as follows.

Theorem 2.5. *Suppose σ and b are Lipschitz and there exists a constant c_1 such that*

$$|\sigma(x)| + |b(x)| \le c_1(1 + |x|).$$

Then there exists a pathwise solution to (1.1) and the solution is pathwise unique.

We omit the proof.

We have considered the case of $\mathbb{R}$-valued processes for simplicity, but with only trivial changes the proofs work when the state space is $\mathbb{R}^d$ (and even infinite dimensions if properly formulated), so we can state

Theorem 2.6. *Suppose $d \ge 1$ and suppose σ and b are Lipschitz. Then the SDE (1.3) has a pathwise solution and this solution is pathwise unique.*

In the above, we required σ and b to be functions of X_t only. Only cosmetic changes are required if we allow σ and b to be functions of t and X_t and consider

$$dX_t = \sigma(t, X_t)\,dW_t + b(t, X_t)\,dt. \tag{2.8}$$

3. Types of uniqueness.

When the coefficients σ and b fail to be Lipschitz, it is sometimes the case that (1.3) may not have a pathwise solution at all, or it may not be unique. We define some other notions of existence and uniqueness that are useful. We now assume that the dimension of the state space may be larger than one.

We say a *strong solution exists* to the SDE (1.3) if given the Brownian motion W_t there exists a process X_t satisfying (1.3) such that X_t is adapted to the filtration generated by W_t. A *weak solution exists* if there exists a pair of processes (X_t, W_t) such that W_t is a Brownian motion and the equation (1.3) holds. There is *weak uniqueness* holding if whenever (X_t, W_t) and (X'_t, W'_t) are two weak solutions, then the joint laws of the processes (X, W) and (X', W') are equal. When this happens, we also say that the solution to (1.3) is *unique in law.*

Let us explore some of the relationships between the various definitions just given. Pathwise existence and the existence of a strong solution are very close, differing only in unimportant measurability concerns. If the solution to (1.3) is pathwise unique, then weak uniqueness holds. In the case that σ and b are Lipschitz, the proof is much simpler.

Proposition 3.1. *Suppose σ and b are Lipschitz and bounded. Then the solution to (1.3) is a strong solution. Weak uniqueness holds for (1.3).*

Proof. For notational simplicity we consider the case of dimension one. The Picard iteration in Theorem 2.1 preserves measurability, so the solution constructed in these two theorems is adapted to the filtration generated by W_t. Thus the solution is a strong solution.

Suppose (X_t, W_t) and (X'_t, W'_t) are two solutions to (1.3). Let X''_t be the process that is constructed from W'_t analogously to how X_t was constructed from W_t, namely, by Picard iteration and stopping times. It follows that (X, W) and (X'', W') have the same law. By the pathwise uniqueness, $X'' = X'$, so the result follows. □

We now give an example to show that weak uniqueness might hold even if pathwise uniqueness does not. Let $\sigma(x)$ be equal to 1 if $x \geq 0$ and -1 otherwise. We take b to be identically 0. We consider solutions to

$$X_t = \int_0^t \sigma(X_s)\, dW_s. \tag{3.1}$$

Weak uniqueness holds since X_t must be a martingale, and the quadratic variation of X is $d\langle X\rangle_t = \sigma(X_t)^2\, dt = dt$; by a theorem of Lévy, X_t is a Brownian motion. Given a Brownian motion X_t and letting $W_t = \int_0^t \sigma^{-1}(X_s)\, dX_s$ where $\sigma^{-1} = 1/\sigma$, then again by Lévy's theorem, W_t is a Brownian motion; thus weak solutions exist.

On the other hand, pathwise uniqueness does not hold. To see this, let $Y_t = -X_t$. We have

$$Y_t = \int_0^t \sigma(Y_s)\, dW_s - 2\int_0^t 1_{\{0\}}(X_s)\, dW_s. \tag{3.2}$$

The second term on the right has quadratic variation $4\int_0^t 1_{\{0\}}(X_s)\, ds$, which is equal to 0 almost surely because X is a Brownian motion. Therefore the second term on the right of (3.2) equals 0 almost surely, or Y is another pathwise solution to (3.1).

This example is not satisfying because one would like σ to be positive and even continuous if possible. Such examples exist, however.

4. One-dimensional case.

Although we have often looked at the case where the state space is $\mathbb{R}$ instead of $\mathbb{R}^d$ for the sake of simplicity of notation, everything we have done so far has been valid in $\mathbb{R}^d$ for any d. We now look at some stronger results that hold only in the one-dimensional case.

Theorem 4.1. *Suppose b is bounded and Lipschitz. Suppose there exists a continuous function $\rho:[0,\infty)\to[0,\infty)$ such that $\rho(0)=0$, $\int_{0+}\rho^{-2}(u)\,du=\infty$, and σ is bounded and satisfies*

$$|\sigma(x)-\sigma(y)|\le\rho(|x-y|)$$

for all x and y. Then the solution to (1.3) is pathwise unique.

Proof. Let $a_n\downarrow 0$ be selected so that $\int_{a_n}^{a_{n-1}}du/\rho^2(u)=n$. Let h_n be continuous, supported in (a_n,a_{n-1}), $0\le h_n(u)\le 2/n\rho^2(u)$, and $\int_{a_n}^{a_{n-1}}h_n(u)\,du=1$ for each n. Let g_n be such that $g_n(0)=g_n'(0)=0$ and $g_n''=h_n$. Note $|g_n'(u)|\le 1$ and $g_n'(u)=1$ if $u\ge a_{n-1}$, hence $g_n(u)\uparrow u$ for $u\ge 0$.

Let X_t and X_t' be two solutions to (1.3). The function g_n is in C^2 and is 0 in a neighborhood of 0. We apply Itô's formula to $g_n((\varepsilon^2+|X_t-X_t'|^2)^{1/2})$ and let $\varepsilon\to 0$ to obtain

$$\begin{aligned}g_n(|X_t-X_t'|)=\text{ martingale }&+\int_0^t g_n'(|X_s-X_s'|)[b(X_s)-b(X_s')]\,ds\\&+\frac12\int_0^t g_n''(|X_s-X_s'|)[\sigma(X_s)-\sigma(X_s')]^2\,ds.\end{aligned}$$

We take the expectation of the right-hand side. The martingale term has 0 expectation. The next term has expectation bounded by

$$c_1\int_0^t\mathbb{E}\,|X_s-X_s'|\,ds.$$

The final term on the right-hand side is bounded in expectation by

$$\frac12\mathbb{E}\int_0^t\frac{2}{n(\rho|X_s-X_s'|)^2}(\rho|X_s-X_s'|)^2\,ds\le\frac{t}{n}.$$

Letting $n\to\infty$,

$$\mathbb{E}\,|X_t-X_t'|\le c_1\int_0^t\mathbb{E}\,|X_s-X_s'|\,ds.$$

By Gronwall's lemma, $\mathbb{E}\,|X_t - X_t'| = 0$ for each t. By the continuity of X_t and X_t', we deduce the uniqueness. $\square$

5. Examples.

Ornstein-Uhlenbeck process. The Ornstein-Uhlenbeck process is the solution to the SDE

$$dX_t = dW_t - \frac{X_t}{2}\,dt, \qquad X_0 = x. \tag{5.1}$$

The existence and uniqueness follow from Theorem 2.5, so $(\mathbb{P}^x, X_t)$ is a strong Markov process.

The equation (5.1) can be solved explicitly. Rewriting it and using the product rule,

$$e^{t/2}\,dW_t = e^{t/2}\,dX_t + e^{t/2}\frac{X_t}{2}dt = d[e^{t/2}X_t],$$

or

$$X_t = e^{-t/2}x + e^{-t/2}\int_0^t e^{s/2}\,dW_s. \tag{5.2}$$

Since the integrand of the stochastic integral is deterministic, it follows that X_t is a Gaussian process and the distribution of X_t is that of a normal random variable with mean $e^{-t/2}x$ and variance equal to $e^{-t}\int_0^t e^s\,ds = 1 - e^{-t}$.

If we let $Y_t = \int_0^t e^{s/2}\,dW_s$ and $V_t = Y(\log(t+1))$, then Y_t is a mean 0 continuous Gaussian process with independent increments, and hence so is V_t. Since the variance of $V_u - V_t$ is $\int_{\log(t+1)}^{\log(u+1)} e^s\,ds = u - t$, then V_t is a Brownian motion. Hence $X_t = e^{-t/2}x + e^{-t/2}V(e^t - 1)$. This representation of an Ornstein-Uhlenbeck process in terms of a Brownian motion is useful for, among other things, calculating the exit probabilities of a square root boundary.

Bessel processes. A Bessel process of order $\nu \geq 0$ will be defined to be a nonnegative solution of the SDE

$$dX_t = dW_t + \frac{\nu - 1}{2X_t}dt, \qquad X_0 = x. \tag{5.3}$$

Bessel processes have the same scaling properties as Brownian motion. That is, if X_t is a Bessel process of order ν started at x, then $aX_{a^{-2}t}$ is a Bessel process of order ν started at ax. In fact, from (5.3),

$$d(aX_{a^{-2}t}) = a\,dW_{a^{-2}t} + a^2\frac{\nu - 1}{2aX_{a^{-2}t}}\,d(a^{-2}t),$$

and the assertion follows from the uniqueness and the fact that $aW(a^{-2}t)$ is again a Brownian motion.

Bessel processes are useful for comparison purposes, and so the following is worth recording.

Proposition 5.1. *Suppose X_t is a Bessel process of order ν.*

(i) If $\nu > 2$, X_t never hits 0 and $|X_t| \to \infty$ a.s.

(ii) If $\nu = 2$, X_t hits every neighborhood of 0 infinitely often, but never hits 0.

(iii) If $0 < \nu < 2$, X_t hits 0 infinitely often.

(iv) If $\nu = 0$, then X_t hits 0 and then remains at 0 thereafter.

When we say that X_t hits 0, we consider only times $t > 0$.

Proof. When $\nu = 2$, X_t has the same law as a 2-dimensional Brownian motion, and (ii) follows from the corresponding facts about 2-dimensional Brownian motion. Suppose $\nu \neq 2$; by Itô's formula, $(X_t)^{2-\nu}$ is a martingale. Assertions (i) and (iii) now follow from a standard proof. Similarly, a Bessel process of order 0 hits 0. If X_t is such a process and $Y_t = X_t^2$, then $dY_t = Y_t^{1/2}\, dW_t$. Starting from 0, $Y_t \equiv 0$ is evidently a solution, so by the uniqueness any solution starting at 0 must remain at 0 forever; (iv) now follows by the strong Markov property. □

Brownian bridge. Brownian motion conditioned to be at 0 at time 1 is called Brownian bridge. Brownian bridge has the same law as $W_t - tW_1$. To see this, the covariance of $W_t - tW_1$ and W_1 is 0; hence they are independent. Therefore the law of W_t conditional on W_1 being 0 is the same as the law of $W_t - tW_1 + tW_1$ conditional on W_1 being 0, which is $W_t - tW_1$ by independence.

We will see shortly that Brownian bridge can be represented as the solution of a SDE

$$dX_t = dW_t - \frac{X_t}{1-t}dt, \qquad X_0 = 0. \tag{5.5}$$

Although Theorem 2.5 does not apply because the drift term depends on s as well as the position X_s, similar proofs to those given above guarantee uniqueness and existence for the solution of (5.5) for $s \le t$ for any $t < 1$. As with the Ornstein-Uhlenbeck process, (5.5) may be solved explicitly. We have

$$dW_t = dX_t + \frac{X_t}{1-t}dt = (1-t)\, d\Big[\frac{X_t}{1-t}\Big],$$

or

$$X_t = (1-t)\int_0^t \frac{dW_s}{1-s}.$$

Thus X_t is a continuous Gaussian process with mean 0. The variance of X_t is

$$(1-t)^2 \int_0^t (1-s)^{-2}\, ds = t - t^2,$$

the same as the variance of Brownian bridge. A similar calculation shows that the covariance of X_t and X_s is the same as the covariance of $W_t - tW_1$ and $W_s - sW_1$. Hence the law of X_t and Brownian bridge are the same.

Linear equations. The equation $dX_t = AX_t\,dW_t + BX_t\,dt$ may be written $dX_t = X_t\,dY_t$, where $Y_t = AW_t + Bt$, and then we can check that the solution is

$$X_t = X_0 \exp(Y_t - \langle Y \rangle_t/2).$$

6. Markov properties.

One of the more important applications of SDEs is to Markov processes. A Markov process is one where the probability of future events depends on the past history only through the present position. In order to be more precise, we need to introduce some notation. Rather than having one probability measure and a collection of processes, it is more convenient to have one process and a collection of measures.

Define Ω' to be the set of all continuous functions from $[0,\infty)$ to $\mathbb{R}^d$. We define $Z_t(\omega) = \omega(t)$ for $\omega \in \Omega'$. We call Z_t the *canonical process.* Suppose that for each starting point x the SDE (1.3) has a solution that is unique in law. Let us denote the solution by $X(x,t,\omega)$. For each x define a probability measure $\mathbb{P}^x$ on Ω' so that

$$\begin{aligned}\mathbb{P}^x(Z_{t_1} \in & A_1, \ldots, Z_{t_n} \in A_n)\\ &= \mathbb{P}(X(x,t_1,\omega) \in A_1, \ldots, X(x,t_n,\omega) \in A_n)\end{aligned}$$

whenever $t_1, \ldots, t_n \in [0,\infty)$ and $A_1, \ldots, A_n$ are Borel sets in $\mathbb{R}^d$. The measure $\mathbb{P}^x$ is determined on the smallest σ-field containing these cylindrical sets. Let $\mathcal{G}_t^{00}$ be the σ-algebra generated by Z_s, $s \le t$. We complete these σ-fields by considering all sets that are in the $\mathbb{P}^x$ completion of $\mathcal{G}_t^{00}$ for all x. (This is not quite the same as the completion with respect to $\mathbb{P}^x$, but it will be good enough for our purposes.) Finally, we obtain a right continuous filtration by letting $\mathcal{F}'_t = \cap_{\varepsilon>0} \mathcal{G}_{t+\varepsilon}^{00}$. We then extend $\mathbb{P}^x$ to $\mathcal{F}'_\infty$.

One advantage of Ω' is that it is equipped with *shift operators* $\theta_t : \Omega' \to \Omega'$ defined by $\theta_t(\omega)(s) = \omega(t+s)$. Another way of writing this is $Z_t \circ \theta_s = Z_{t+s}$. For stopping times T we set $\theta_T(\omega) = \theta_{T(\omega)}(\omega)$.

The *strong Markov property* is the assertion that

$$\mathbb{E}^{\,x}[Y \circ \theta_T \mid \mathcal{F}'_T] = \mathbb{E}^{\,Z_T}[Y], \qquad \text{a.s. } (\mathbb{P}^x) \tag{6.1}$$

whenever $x \in \mathbb{R}^d$, Y is bounded and $\mathcal{F}'_\infty$ measurable, and T is a finite stopping time. The *Markov property* holds if the above equality holds whenever T is a fixed (i.e., nonrandom) time. If the strong Markov property holds, we say $(\mathbb{P}^x, Z_t)$ is a *strong Markov process.*

To prove the strong Markov property it suffices to show

$$\mathbb{E}^{\,x}[f(Z_{T+t}) \mid \mathcal{F}'_T] = \mathbb{E}^{\,Z_T} f(Z_t), \qquad \text{a.s. } (\mathbb{P}^x) \tag{6.2}$$

for all $x \in \mathbb{R}^d$, f a bounded and continuous function on $\mathbb{R}^d$, and T a bounded stopping time. This is (6.1) with $Y = f(X_t)$.

It turns out that if pathwise uniqueness or weak uniqueness holds for (1.3) for every x, then $(\mathbb{P}^x, Z_t)$ form a strong Markov process.

Let T be a bounded stopping time. A *regular conditional probability* for $\mathbb{E}[\cdot \mid \mathcal{F}_T]$ is a kernel $Q_T(\omega, d\omega')$ such that

(i) $Q_T(\omega, \cdot)$ is a probability measure on Ω' for each ω;

(ii) for each $\mathcal{F}'_\infty$ measurable set A, $Q_T(\cdot, A)$ is a $\mathcal{F}'_\infty$ measurable random variable;

(iii) for each $\mathcal{F}'_\infty$ measurable set A and each $\mathcal{F}'_T$ measurable set B,

$$\int_B \mathbb{Q}_T(\omega, A)\, \mathbb{P}(d\omega) = \mathbb{P}(A \cap B).$$

Regular conditional probabilities need not always exist, but if the probability space is regular enough, as Ω' is, then they do.

We have the equation

$$Z_t = Z_0 + \int_0^t \sigma(Z_r)\, dW_r + \int_0^t b(Z_r)\, dr, \tag{6.3}$$

where W_r is a Brownian motion, not necessarily the same as the one in (1.3). If we let $\widetilde{Z}_t = Z_{T+t}$ and $\widetilde{W}_t = W_{T+t} - W_T$, it is plausible that $\widetilde{W}$ is a Brownian motion with respect to the measure $Q_T(\omega, \cdot)$ for almost every ω. We write (6.3) with t replaced by $T + t$ and then write (6.3) with t replaced by T. Taking the difference and using a change of variables, we obtain

$$\widetilde{Z}_t = \widetilde{Z}_0 + \int_0^t \sigma(\widetilde{Z}_r)\, \widetilde{W}_r + \int_0^t b(\widetilde{Z}_r)\, dr. \tag{6.4}$$

Theorem 6.1. *Suppose the solution to (1.3) is weakly unique for each x. Then $(\mathbb{P}^x, Z_t)$ is a strong Markov process.*

Proof. Fix x and let Q_T denote the regular conditional probability for $\mathbb{E}^x[\cdot \mid \mathcal{F}'_T]$. Except for ω in a null set, under $Q_T(\omega, \cdot)$ we have from (6.4) and Proposition 6.3 that $\widetilde{Z}$ is a solution to (1.3) with starting point $\widetilde{Z}_0 = Z_T$. So if $\mathbb{E}_{Q_T}$ denotes the expectation with respect to Q_T, the uniqueness in law tells us that

$$\mathbb{E}_{Q_T} f(\widetilde{Z}_t) = \mathbb{E}^{Z_T} f(Z_t), \qquad \text{a.s. } (\mathbb{P}^x).$$

On the other hand,

$$\mathbb{E}_{Q_T} f(\widetilde{Z}_t) = \mathbb{E}_{Q_T} f(Z_{T+t}) = \mathbb{E}^x[f(Z_{T+t}) \mid \mathcal{F}'_T], \qquad \text{a.s. } (\mathbb{P}^x),$$

which proves (6.2). □

By a slight abuse of notation, we will say $(\mathbb{P}^x, X_t)$ is a strong Markov family when $(\mathbb{P}^x, Z_t)$ is a strong Markov family.

7. Poisson's equation.

Let

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial_{ij}f(x) + \sum_{i=1}^d b_i(x)\partial_i f(x). \tag{7.1}$$

We assume the a_{ij} and b_i are bounded and at least C^1. We also assume that the operator $\mathcal{L}$ is uniformly strictly elliptic. An operator $\mathcal{L}$ is *strictly elliptic* if for each x there exists $\Lambda(x)$ such that

$$\sum_{i,j=1}^d a_{ij}(x)y_iy_j \geq \Lambda(x)\sum_{i=1}^d y_i^2 \qquad\qquad y = (y_1, \ldots, y_d) \in \mathbb{R}^d. \tag{7.2}$$

The operator $\mathcal{L}$ is *uniformly strictly elliptic* or *uniformly elliptic* if Λ can be chosen to be independent of x. We also call the matrix a strictly elliptic if (7.2) holds and uniformly elliptic if (7.2) holds with $\Lambda(x)$ not depending on x. We also assume throughout that the dimension d is greater than or equal to 3.

We emphasize that the uniform ellipticity of $\mathcal{L}$ is used only to show that the exit times of the domains we consider are finite a.s. For many nonuniformly elliptic operators, it is often the case that the finiteness of the exit times is known for other reasons, and the results then apply to equations involving these operators.

Suppose σ is a matrix such that $a = \sigma\sigma^T$ and each component of σ is bounded and in C^1. Let X_t be the solution to

$$X_t = x + \int_0^t \sigma(X_s)\,dW_s + \int_0^t b(X_s)\,ds. \tag{7.3}$$

We will write $(\mathbb{P}^x, X_t)$ for the strong Markov process corresponding to σ and b.

We consider first Poisson's equation in $\mathbb{R}^d$. Suppose $\lambda > 0$ and f is a C^1 function with compact support. Poisson's equation is

$$\mathcal{L}u(x) - \lambda u(x) = -f(x), \qquad x \in \mathbb{R}^d. \tag{7.4}$$

Theorem 7.1. *Suppose u is a C_b^2 solution to (7.4). Then*

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

Proof. Let u be the solution to (7.4). By Itô's formula,

$$u(X_t) - u(X_0) = M_t + \int_0^t \mathcal{L}u(X_s)\,ds,$$

where M_t is a martingale. By the product formula,

$$\begin{aligned} e^{-\lambda t}u(X_t) - u(X_0) = \int_0^t e^{-\lambda s} dM_s + \int_0^t e^{-\lambda s}\mathcal{L}u(X_s)\,ds \\ - \lambda \int_0^t e^{-\lambda s} u(X_s)\,ds. \end{aligned}$$

Taking $\mathbb{E}^{\,x}$ expectation and letting $t \to \infty$,

$$-u(x) = \mathbb{E}^{\,x} \int_0^\infty e^{-\lambda s}(\mathcal{L}u - \lambda u)(X_s)\,ds.$$

Since $\mathcal{L}u - \lambda u = -f$, the result follows. □

Let us now let D be a nice bounded domain, e.g., a ball. Poisson's equation in D requires one to find a function u such that $\mathcal{L}u - \lambda u = -f$ in D and $u = 0$ on ∂D, where $f \in C^2(\overline{D})$ and $\lambda \geq 0$. Here we can allow λ to be equal to 0. We will see later on that the time to exit D, namely, $\tau_D = \inf\{t : X_t \notin D\}$, is finite almost surely.

Theorem 7.2. *Suppose u is a solution to Poisson's equation in a bounded domain D that is C^2 in D and continuous on $\overline{D}$. Then*

$$u(x) = \mathbb{E}^{\,x} \int_0^{\tau_D} e^{-\lambda s} f(X_s)\,ds.$$

Proof. The proof is nearly identical to that of Theorem 7.1. $\tau_D < \infty$ a.s. Let $S_n = \inf\{t : \text{dist}\,(X_t, \partial D) < 1/n\}$. By Itô's formula,

$$u(X_{t\wedge S_n}) - u(X_0) = \text{ martingale } + \int_0^{t\wedge S_n} \mathcal{L}u(X_s)\,ds.$$

By the product formula,

$$\begin{aligned} &\mathbb{E}^{\,x} e^{-\lambda(t\wedge S_n)} u(X_{t\wedge S_n}) - u(x) \\ &\qquad = \mathbb{E}^{\,x} \int_0^{t\wedge S_n} e^{-\lambda s}\mathcal{L}u(X_s)\,ds - \mathbb{E}^{\,x} \int_0^{t\wedge S_n} e^{-\lambda s} u(X_s)\,ds \\ &\qquad = -\mathbb{E}^{\,x} \int_0^{t\wedge S_n} e^{-\lambda s} f(X_s)\,ds. \end{aligned}$$

Now let $n \to \infty$ and then $t \to \infty$ and use the fact that u is 0 on ∂D. $\square$

8. Dirichlet problem.

Let D be a ball (or other nice bounded domain) and let us consider the solution to the Dirichlet problem: given f a continuous function on ∂D, find $u \in C(\overline{D})$ such that u is C^2 in D and

$$\mathcal{L}u = 0 \text{ in } D, \qquad u = f \text{ on } \partial D. \tag{8.1}$$

Theorem 8.1. *The solution to (8.1) satisfies*

$$u(x) = \mathbb{E}^x f(X_{\tau_D}).$$

Proof. We use the fact that $\tau_D < \infty$ a.s. Let $S_n = \inf\{t : \mathrm{dist}\,(X_t, \partial D) < 1/n\}$. By Itô's formula,

$$u(X_{t\wedge S_n}) = u(X_0) + \text{ martingale } + \int_0^{t\wedge S_n} \mathcal{L}u(X_s)\,ds.$$

Since $\mathcal{L}u = 0$ inside D, taking expectations shows

$$u(x) = \mathbb{E}^x u(X_{t\wedge S_n}).$$

We let $t \to \infty$ and then $n \to \infty$. By dominated convergence, we obtain $u(x) = \mathbb{E}^x u(X_{\tau_D})$. This is what we want since $u = f$ on ∂D. $\square$

There are some further facts that can be deduced from Theorem 8.1. One is the *maximum principle*: if $x \in D$,

$$\sup_{\overline{D}} u \le \sup_{\partial D} u. \tag{8.2}$$

This follows from

$$u(x) = \mathbb{E}^x f(X_{\tau_D}) \le \sup_{\partial D} f.$$

If $\mathcal{L}v = 0$ in D, we say v is *$\mathcal{L}$-harmonic* in D.

9. Cauchy problem.

The related parabolic partial differential equation $\partial_t u = \mathcal{L}u$ is often of interest. Here $\partial_t u$ denotes $\partial u/\partial t$.

Suppose for simplicity that the function f is a continuous function with compact support. The Cauchy problem is to find u such that u is bounded, u is in C_b^2 in x, u is C^1 in t for $t>0$, and

$$\begin{aligned} \partial_t u(x,t) &= \mathcal{L}u(x,t), \qquad t>0, x\in\mathbb{R}^d, \\ u(x,0) &= f(x), \qquad x\in\mathbb{R}^d. \end{aligned} \tag{9.1}$$

Theorem 9.1. *The solution to (9.1) satisfies*

$$u(x,t)=\mathbb{E}^x f(X_t).$$

Proof. Fix t_0 and let $M_t=u(X_t,t_0-t)$. The solution u to (9.1) is known to be C^2 in x and C^1 in t for $t>0$. Note $\partial_t[u(x,t_0-t)]=-(\partial_t u)(x,t_0-t)$. By Itô's formula on $\mathbb{R}^d\times[0,t_0)$,

$$\begin{aligned} u(X_t,t_0-t) = \text{ martingale } &+ \int_0^t \mathcal{L}u(X_s,t_0-s)\,ds \\ &+ \int_0^t (-\partial_t u)(X_s,t_0-s)\,ds. \end{aligned}$$

Since $\partial_t u=\mathcal{L}u$, M_t is a martingale, and $\mathbb{E}^x M_0=\mathbb{E}^x M_{t_0}$. On the one hand,

$$\mathbb{E}^x M_{t_0}=\mathbb{E}^x u(X_{t_0},0)=\mathbb{E}^x f(X_{t_0}),$$

while on the other,

$$\mathbb{E}^x M_0=\mathbb{E}^x u(X_0,t_0)=u(x,t_0).$$

Since t_0 is arbitrary, the result follows. □

For bounded domains D, the Cauchy problem is to find u such that $\partial_t=\mathcal{L}u$ on D, $u(x,0)=f(x)$ for $x\in D$, and $u(x,t)=0$ for $x\in\partial D$. The solution is given by

$$u(x,t)=\mathbb{E}^x[f(X_t);t<\tau_D],$$

where τ_D is the exit time of D. The proof is very similar to the case of $\mathbb{R}^d$.

10. Schrödinger operators.

We next look at what happens when one adds a potential term, that is, when one considers the operator

$$\mathcal{L}u(x)+q(x)u(x). \tag{10.1}$$

This is known as the *Schrödinger operator*, and $q(x)$ is known as the *potential.* Equations involving the operator in (10.1) are considerably simpler than the quantum mechanics Schrödinger equation because here all terms are real-valued.

If X_t is the diffusion corresponding to $\mathcal{L}$, then solutions to PDEs involving the operator in (10.1) can be expressed in terms of X_t by means of the *Feynman-Kac formula.* To illustrate, let D be a nice bounded domain, e.g., a ball, q a C^2 function on $\overline{D}$, and f a continuous function on ∂D; q^+ denotes the positive part of q.

Theorem 10.1. *Let D, q, f be as above. Let u be a C^2 function on $\overline{D}$ that agrees with f on ∂D and satisfies $\mathcal{L}u + qu = 0$ in D. If*

$$\mathbb{E}^{\,x} \exp\Big(\int_0^{\tau_D} q^+(X_s)\, ds\Big) < \infty,$$

then

$$u(x) = \mathbb{E}^{\,x}\Big[f(X_{\tau_D}) e^{\int_0^{\tau_D} q(X_s)\, ds}\Big]. \tag{10.2}$$

Proof. Let $B_t = \int_0^{t\wedge\tau_D} q(X_s)\, ds$. By Itô's formula and the product formula,

$$\begin{aligned} e^{B(t\wedge\tau_D)} u(X_{t\wedge\tau_D}) = u(X_0) + \text{ martingale } &+ \int_0^{t\wedge\tau_D} u(X_r) e^{B_r}\, dB_r \\ &+ \int_0^{t\wedge\tau_D} e^{B_r}\, d[u(X)]_r. \end{aligned}$$

Taking $\mathbb{E}^{\,x}$ expectation,

$$\begin{aligned} \mathbb{E}^{\,x} e^{B(t\wedge\tau_D)} u(X_{t\wedge\tau_D}) = u(x) &+ \mathbb{E}^{\,x} \int_0^{t\wedge\tau_D} e^{B_r} u(X_r) q(X_r)\, dr \\ &+ \mathbb{E}^{\,x} \int_0^{t\wedge\tau_D} e^{B_r} \mathcal{L}u(X_r)\, dr. \end{aligned}$$

Since $\mathcal{L}u + qu = 0$,

$$\mathbb{E}^{\,x} e^{B(t\wedge\tau_D)} u(X_{t\wedge\tau_D}) = u(x).$$

If we let $t \to \infty$ and use the exponential integrability of q^+, the result follows. □

The existence of a solution to $\mathcal{L}u + qu = 0$ in D depends on the finiteness of $\mathbb{E}^{\,x} e^{\int_0^{\tau_D} q^+(X_s)\, ds}$, an expression that is sometimes known as the *gauge*.

Even in one dimension with $D = (0, 1)$ and q a constant function, the gauge need not be finite. For Brownian motion it is known that $\mathbb{P}^x(\tau_D > t) \geq ce^{-\pi^2 t/2}$ for t sufficiently large. Hence

$$\begin{aligned}\mathbb{E}^x \exp\Big(\int_0^{\tau_D} q\,ds\Big) &= \mathbb{E}^x e^{q\tau_D}\\ &= \int_0^\infty qe^{qt}\mathbb{P}^x(\tau_D > t)\,dt;\end{aligned}$$

this is infinite if $q \geq \pi^2/2$.

A very similar proof to that of Theorem 10.1 shows that under suitable assumptions on q, g, and D, the solution to $\mathcal{L}u + qu = -g$ in D with boundary condition $u = 0$ on ∂D is given by

$$u(x) = \mathbb{E}^x\Big[\int_0^{\tau_D} g(X_s)e^{\int_0^s q(X_r)\,dr}\,ds\Big]. \tag{10.3}$$

There is also a parabolic version of Theorem 10.1. The equation $\partial_t u = \mathcal{L}u + qu$ with initial condition $u(x, 0) = f(x)$ is solved by

$$u(x,t) = \mathbb{E}^x\Big[f(X_t)e^{\int_0^t q(X_s)\,ds}\Big]. \tag{10.4}$$

11. Fundamental solutions and Green functions.

The function $p(t, x, y)$ is the *fundamental solution* for $\mathcal{L}$ if the solution to

$$\partial_t u = \mathcal{L}u, \qquad u(x, 0) = f(x) \tag{11.1}$$

is given by

$$u(x,t) = \int p(t,x,y)f(y)\,dy$$

for all continuous f with compact support. We have seen that the solution is also given by $\mathbb{E}^x f(X_t)$. So

$$\int p(t,x,y)f(y)\,dy = \mathbb{E}^x f(X_t) = \int f(y)\mathbb{P}^x(X_t \in dy).$$

Thus the fundamental solution is the same as the transition density for the associated process.

An operator $\mathcal{L}$ in a nice domain D has a *Green function* $G_D(x, y)$ if $G_D(x, y) = 0$ if either x or y is in ∂D and the solution to

$$\mathcal{L}u = f \text{ in } D, \qquad u = 0 \text{ on } \partial D$$

is given by

$$u(x) = -\int G_D(x,y) f(y)\, dy$$

when f is continuous. We have also seen that the solution is given by

$$u(x) = -\mathbb{E}^x \int_0^{\tau_D} f(X_s)\, ds.$$

Thus $G_D(x,y)$ is the same as the occupation time density for X_t. That is, $G_D(x,y)$ is the Radon-Nikodym derivative of the measure $\mu(A) = \mathbb{E}^x \int_0^{\tau_D} 1_A(X_s)\, ds$ with respect to Lebesgue measure.

12. Adjoints.

The *adjoint* operator to $\mathcal{L}$ is the operator

$$\mathcal{L}^* f(x) = \sum_{i,j=1}^d \partial_{ij}\big(a_{ij}(x) f(x)\big) - \sum_{i=1}^d \partial_i\big(b_i(x) f(x)\big). \tag{12.1}$$

The reason for the name is that

$$\int_{\mathbb{R}^d} f(x) \mathcal{L} g(x)\, dx = \int_{\mathbb{R}^d} g(x) \mathcal{L}^* f(x)\, dx,$$

as integrations by parts show, provided f and g satisfy suitable regularity conditions. The adjoint operator corresponds to the process that is the dual of X_t. Roughly speaking, the dual of X_t is the process run backwards: X_{t_0-t}.

13. Black-Scholes formula.

We aren't going to derive the Black-Scholes PDE here, but we will show as an example how to solve it using probability.

The equation is

$$f_t(x,t) = \tfrac{1}{2}\sigma^2 x^2 f_{xx}(x,t) + rx f_x(x,t) - r f(x,t) \tag{13.1}$$

with initial condition

$$f(x,0) = (x-K)^+. \tag{13.2}$$

Here $f(x,t)$ tells you the price of a European call at time $T-t$ if the stock price is x, where K is the strike price, σ is the volatility, r is the interest rate, and T is the exercise time.

Let us first look at

$$f_t = \tfrac{1}{2}\sigma^2 x^2 f_{xx} + rx f_x. \tag{13.3}$$

Solve

$$dX_t = \sigma X_t\, dW_t + rX_t\, dt.$$

We have an explicit solution of this because it is linear:

$$X_t = x_0 e^{\sigma W_t + rt - \frac{1}{2}\sigma^2 t}.$$

To calculate $\mathbb{E}^x g(X_t)$, we can write

$$\begin{aligned}\mathbb{E}^x g(X_t) &= \int g(y)\mathbb{P}^x(X_t \in dy)\\ &= \int g(xe^{\sigma z + rt - \frac{1}{2}\sigma^2 t})\mathbb{P}(W_t \in dz)\\ &= \int g(xe^{\sigma z + rt - \frac{1}{2}\sigma^2 t})\tfrac{1}{\sqrt{2\pi t}} e^{-z^2/2t}\, dz.\end{aligned}$$

The solution to (13.3) is

$$f(x,t) = \mathbb{E}^x (X_t - K)^+.$$

Finally, to solve (13.1) we use the Feynman-Kac formula.

14. Nondivergence operators.

We consider operators in *nondivergence form*, that is, operators of the form

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial_{ij} f(x) + \sum_{i=1}^d b_i(x)\partial_i f(x). \tag{14.1}$$

These operators are sometimes said to be of *nonvariational form.*

We assume throughout this chapter that the coefficients a_{ij} and b_i are bounded and measurable. Unless stated otherwise, we also assume that the operator $\mathcal{L}$ is uniformly elliptic. The coefficients a_{ij} are called the *diffusion* coefficients and the b_i are called the *drift* coefficients. We let $\mathcal{N}(\Lambda_1, \Lambda_2)$ denote the set of operators of the form (14.1) with $\sup_i \|b_i\|_\infty \le \Lambda_2$ and

$$\Lambda_1 |y|^2 \le \sum_{i,j=1}^d y_i a_{ij}(x) y_j \le \Lambda_1^{-1}|y|^2, \qquad y \in \mathbb{R}^d, x \in \mathbb{R}^d. \tag{14.2}$$

We saw that if X_t is the solution to

$$dX_t = \sigma(X_t)\, dW_t + b(X_t)\, dt, \qquad X_0 = x_0, \tag{14.3}$$

where σ is a $d \times d$ matrix, b is a vector, and W_t is a Brownian motion, then X_t is associated to the operator $\mathcal{L}$ with $a = \sigma\sigma^T$. If $f \in C^2$, then

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,ds \tag{14.4}$$

is a local martingale under $\mathbb{P}$.

A very fruitful idea of Stroock and Varadhan is to phrase the association of X_t to $\mathcal{L}$ in terms which use (14.4) as a key element. Let Ω consist of all continuous functions ω mapping $[0, \infty)$ to $\mathbb{R}^d$. Let $X_t(\omega) = \omega(t)$ and let $\mathcal{F}_t$ be the right continuous modification of the σ-field generated by the X_s, $s \le t$. A probability measure $\mathbb{P}$ is a solution to the *martingale problem for $\mathcal{L}$ started at x_0* if

$$\mathbb{P}(X_0 = x_0) = 1 \tag{14.5}$$

and

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,ds$$

is a local martingale under $\mathbb{P}$ whenever $f \in C^2(\mathbb{R}^d)$. The martingale problem is *well posed* if there exists a solution and this solution is unique.

Uniqueness of the martingale problem for $\mathcal{L}$ is closely connected to weak uniqueness or uniqueness in law of (14.3). Recall that the cylindrical sets are ones of the form $\{\omega : \omega(t_1) \in A_1, \ldots, \omega(t_n) \in A_n\}$ for $n \ge 1$ and $A_1, \ldots, A_n$ Borel subsets of $\mathbb{R}^d$.

Theorem 14.1. *Suppose $a = \sigma\sigma^T$. Weak uniqueness for (14.3) holds if and only if the solution for the martingale problem for $\mathcal{L}$ started at x_0 is unique. Weak existence for (14.3) holds if and only if there exists a solution to the martingale problem for $\mathcal{L}$ started at x_0.*

Proof. We prove the uniqueness assertion. Let Ω be the continuous functions on $[0, \infty)$ and Z_t the coordinate process: $Z_t(\omega) = \omega(t)$. First suppose the solution to the martingale problem is unique. If (X_t^1, W_t^1) and (X_t^2, W_t^2) are two weak solutions to (14.3), define $\mathbb{P}_1^{x_0}$ and $\mathbb{P}_2^{x_0}$ on Ω by $\mathbb{P}_i^{x_0}(Z_\cdot \in A) = \mathbb{P}(X_\cdot^i \in A)$, $i = 1, 2$, for any cylindrical set A. Clearly $\mathbb{P}_i^{x_0}(Z_0 = x_0) = \mathbb{P}(X_0^i = x_0) = 1$. (14.4) is a local martingale under $\mathbb{P}_i^{x_0}$ for each i and each $f \in C^2$. By the hypothesis of uniqueness for the solution of the martingale problem, $\mathbb{P}_1^{x_0} = \mathbb{P}_2^{x_0}$. This implies that the laws of X_t^1 and X_t^2 are the same, or weak uniqueness holds.

Now suppose weak uniqueness holds for (14.3). Let

$$Y_t = Z_t - \int_0^t b(Z_s)\,ds.$$

Let $\mathbb{P}_1^{x_0}$ and $\mathbb{P}_2^{x_0}$ be solutions to the martingale problem. If $f(x) = x_k$, the kth coordinate of x, then $\partial_i f(x) = \delta_{ik}$ and $\partial_{ij} f = 0$, or $\mathcal{L}f(Z_s) = b_k(Z_s)$. Therefore the kth coordinate of Y_t is a local martingale under $\mathbb{P}_i^{x_0}$. Now let $f(x) = x_k x_m$. Computing $\mathcal{L}f$, we see that $Y_t^k Y_t^m - \int_0^t a_{km}(Z_s)\,ds$ is a local martingale. We set

$$W_t = \int_0^t \sigma^{-1}(Z_s)\,dY_s.$$

The stochastic integral is finite since

$$\begin{aligned} &\mathbb{E}\int_0^t \sum_j (\sigma^{-1})_{ij}(Z_s) \sum_k (\sigma^{-1})_{ik}(Z_s)\, d\langle Y^j, Y^k\rangle_s \qquad (14.6)\\ &\qquad = \mathbb{E}\int_0^t \sum_{i,k} (a^{-1})_{ik}(Z_s) a_{ik}(Z_s)\,ds = t < \infty. \end{aligned}$$

It follows that W_t is a martingale, and a calculation similar to (14.6) shows that $W_t^k W_t^m - \delta_{km} t$ is also a martingale under $\mathbb{P}_i^{x_0}$. So by Lévy's theorem, W_t is a Brownian motion under both $\mathbb{P}_1^{x_0}$ and $\mathbb{P}_2^{x_0}$, and (Z_t, W_t) is a weak solution to (14.3). By the weak uniqueness hypothesis, the laws of Z_t under $\mathbb{P}_1^{x_0}$ and $\mathbb{P}_2^{x_0}$ agree, which is what we wanted to prove.

A similar proof shows that the existence of a weak solution to (14.3) is equivalent to the existence of a solution to the martingale problem. □

Since pathwise existence and uniqueness imply weak existence and uniqueness, if the σ_{ij} and b_i are Lipschitz, then the martingale problem for $\mathcal{L}$ is well posed for every starting point.

15. Some estimates.

Diffusions corresponding to elliptic operators in nondivergence form do not have an exact scaling property as does Brownian motion, i.e., rX_{t/r^2} does not necessarily have the same law as X_t. However, they do have a weak scaling property that is nearly as useful: rX_{t/r^2} is again a diffusion corresponding to another elliptic operator of the same type.

Proposition 15.1. *Suppose $\mathcal{L}$ is an elliptic operator with zero drift coefficients. Suppose $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ started at x_0. Then the law of rZ_{t/r^2} is a solution to the martingale problem for $\mathcal{L}_r$ started at rx_0, where*

$$\mathcal{L}_r f(x) = \sum_{i,j=1}^d a_{ij}(x/r)\partial_{ij} f(x), \qquad f \in C^2.$$

Proof. It is obvious that rZ_{t/r^2} starts at rx_0 with $\mathbb{P}$ probability one. If $f \in C^2$, let $g(x) = f(rx)$. Setting $V_t = rZ_{t/r^2}$,

$$\begin{aligned} f(V_t) &= g(Z_{t/r^2}) \\ &= g(x_0) + \text{ martingale } + \int_0^t \sum_{i,j} \partial_{ij} g(Z_{s/r^2})\, d\langle Z^i, Z^j \rangle_{s/r^2}. \end{aligned} \tag{15.1}$$

By the definition of g, $\partial_{ij} g(x) = r^2 \partial_{ij} f(rx)$, so $\partial_{ij} g(Z_{s/r^2}) = r^2 \partial_{ij} f(V_s)$. From the definition of martingale problem applied to the function $x_i x_j$, we see that as in the proof of Theorem 14.1, $Z_t^i Z_t^j - \int_0^t a_{ij}(Z_s)\, ds$ is a local martingale under $\mathbb{P}$, and hence $d\langle Z^i, Z^j \rangle_s = a_{ij}(Z_s)\, ds$ and

$$d\langle Z^i, Z^j \rangle_{s/r^2} = r^{-2} a_{ij}(Z_{s/r^2})\, ds = r^{-2} a_{ij}(V_s/r)\, ds.$$

Substituting in (15.1),

$$f(V_t) = f(V_0) + \text{ martingale } + \int_0^t \sum_{i,j} a_{ij}(V_s/r) \partial_{ij} f(V_s)\, ds.$$

Thus the law of V_t under $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}_r$. □

The following elementary bounds on the time to exit a ball will be used repetitively. Recall that τ_A denotes the hitting time of A.

Proposition 15.2. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$, so that the drift coefficients of $\mathcal{L}$ are 0. Suppose $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ started at 0.*

(a) There exists c_1 depending only on Λ such that

$$\mathbb{P}(\tau_{B(0,1)} \le t) \le c_1 t.$$

(b) There exist c_2 and c_3 depending only on Λ such that

$$\mathbb{P}(\tau_{B(0,1)} \ge t) \le c_2 e^{-c_3 t}.$$

Proof. Write B for $B(0,1)$. Let f be a C^2 function that is zero at 0, one on ∂B, with $\partial_{ij} f$ bounded by a constant c_4. Since $\mathbb{P}$ is a solution to the martingale problem,

$$\mathbb{E}\, f(X_{t\wedge\tau_B}) = \mathbb{E} \int_0^{t\wedge\tau_B} \mathcal{L} f(X_s)\, ds \le c_5 t,$$

where c_5 depends on c_4 and Λ. Since $f(X_{t\wedge\tau_B}) \ge 1_{(\tau_B \le t)}$, this proves (a).

To prove (b), look at X_t^1. Since $\mathbb{P}$ is a solution to the martingale problem, taking $f(x) = x_1$ in (14.4) shows that X_t^1 is a local martingale. Taking $f(x) = x_1^2$ in (14.4) shows that $(X_t^1)^2 - \int_0^t a_{11}(X_s)\,ds$ is also a local martingale. So $d\langle X^1\rangle_t = a_{11}(X_t)\,dt$, and X_t^1 is a nondegenerate time change of a one-dimensional Brownian motion. X_s^1 stays in the interval $[-1,1]$ up until time t only if a Brownian motion stays in the interval $[-1,1]$ up until time $c_6 t$, and this is known to be bounded by $c_7 e^{-c_8 t}$. If X_s has not exited B by time t, then X_s^1 has not exited $[-1,1]$, and (b) follows. □

An important property of X_t is that it satisfies a support theorem. Suppose X_t satisfies (15.1). We suppose that σ, σ^{-1}, and b are bounded, but we impose no other smoothness conditions. Let $a = \sigma\sigma^T$.

Lemma 15.3. *Suppose $Y_t = M_t + A_t$ is a continuous semimartingale with dA_t/dt and $d\langle M\rangle_t/dt$ bounded above by c_1 and $d\langle M\rangle_t/dt$ bounded below by $c_2 > 0$. If $\varepsilon > 0$ and $t_0 > 0$, then*

$$\mathbb{P}(\sup_{s\le t_0} |Y_s| < \varepsilon) \ge c_3,$$

where $c_3 > 0$ depends only on c_1, c_2, ε, and t_0.

Proof. Let $B_t = \inf\{u : \langle M\rangle_u > t\}$. Then $W_t = M_{B_t}$ is a continuous martingale with quadratic variation equal to t; hence by Lévy's theorem, W_t is a Brownian motion. If $Z_t = Y_{B_t} = W_t + E_t$, then $E_t = \int_0^t e_s\,ds$ for some e_s bounded by c_4, where c_4 depends only on c_1 and c_2. Our assertion will follow if we can show

$$\mathbb{P}(\sup_{s\le c_1 t_0} |Z_s| < \varepsilon) \ge c_3.$$

We now use Girsanov's theorem. Define a probability measure $\mathbb{Q}$ by

$$d\mathbb{Q}/d\mathbb{P} = \exp\Big(-\int_0^{t_0} e_s\,dW_s - \frac{1}{2}\int_0^{t_0} e_s^2\,ds\Big)$$

on $\mathcal{F}_{t_0}$. Under $\mathbb{P}$, W_t is a martingale, so under $\mathbb{Q}$ we have that

$$W_t - \Big\langle \int_0^\cdot (-e_s)\,dW, W\Big\rangle_t = W_t + \int_0^t e_s\,ds$$

is a martingale with the same quadratic variation as W has under $\mathbb{P}$, namely, t. Then under $\mathbb{Q}$, Z_t is a Brownian motion. By a well known property of Brownian motion

$$\mathbb{Q}(\sup_{s\le c_1 t_0} |Z_s| < \varepsilon) \ge c_5,$$

for c_5 depending only on ε and $c_1 t_0$. So if C is the event $\{\sup_{s\le c_1 t_0} |Z_s| < \varepsilon\}$,

$$c_5 \le \mathbb{Q}(C) = \int_C (d\mathbb{Q}/d\mathbb{P})\, d\mathbb{P} \le \Big(\mathbb{E}\,(d\mathbb{Q}/d\mathbb{P})^2\Big)^{1/2}\Big(\mathbb{P}(C)\Big)^{1/2}$$

by the Cauchy-Schwarz inequality. The proof is concluded by noting that $d\mathbb{Q}/d\mathbb{P}$ has a second moment depending only on c_4 and t_0. □

We use this lemma to obtain an analogous result for X_t.

Theorem 15.4. *Let $\varepsilon \in (0,1)$, $t_0 > 0$. There exists c_1 depending only on the upper bounds of σ, b, and σ^{-1} such that*

$$\mathbb{P}(\sup_{s\le t_0} |X_s - X_0| < \varepsilon) \ge c_1.$$

Proof. For notational simplicity assume $X_0 = 0$. Let $y = (\varepsilon/4, 0, \dots, 0)$. Applying Itô's formula with $f(z) = |z-y|^2$ and setting $V_t = |X_t - y|^2$, then $V_0 = (\varepsilon/4)^2$ and

$$dV_t = 2\sum_i (X^i_t - y_i)\, dX^i_t + \sum_i d\langle X^i\rangle_t.$$

If we set Y_t equal to V_t for $t \le \inf\{u : |V_u| > (\varepsilon/2)^2\}$ and equal to some Brownian motion for t larger than this stopping time, then Lemma 15.3 applies and

$$\mathbb{P}(\sup_{s\le t_0} |V_s - V_0| \le (\varepsilon/8)^2) = \mathbb{P}(\sup_{s\le t_0} |Y_s - Y_0| \le (\varepsilon/8)^2) \ge c_2.$$

By the definition of y and V_t, this implies with probability at least c_2 that X_t stays inside $B(0,\varepsilon)$. □

We can now prove the *support theorem* for X_t.

Theorem 15.5. *Suppose σ and b are bounded, σ^{-1} is bounded, $x \in \mathbb{R}^d$,and X_t satisfies (15.1) with $X_0 = x$. Suppose $\psi : [0,t] \to \mathbb{R}^d$ is continuous with $\psi(0) = x$ and $\varepsilon > 0$. There exists c_1, depending only on ε, t, the modulus of continuity of ψ, and the bounds on b and σ such that*

$$\mathbb{P}(\sup_{s\le t} |X_s - \psi(s)| < \varepsilon) \ge c_1.$$

This can be phrased as saying the graph of X_s stays inside an ε-tube about ψ. By this we mean, if $G^\varepsilon_\psi = \{(s,y) : |y - \psi(s)| < \varepsilon, s \le t\}$, then $\{(s, X_s) : s \le t\}$ is contained in G^ε_ψ with positive probability.

Proof. We can find a differentiable function $\widehat{\psi}$ such that $\widehat{\psi}(0) = x$ and the $\varepsilon/2$ tube about $\widehat{\psi}$ (which is $G^{\varepsilon/2}_{\widehat{\psi}}$ in the above notation) is contained in G^{ε}_{ψ}, the ε-tube about ψ. So without loss of generality, we may assume ψ is differentiable with a derivative bounded by a constant, say c_2.

Define a new probability measure $\mathbb{Q}$ by

$$d\mathbb{Q}/d\mathbb{P} = \exp\Big(\int_0^t \psi'(s)\sigma^{-1}(X_s)\,dW_s - \frac{1}{2}\int_0^t |\psi'(s)\sigma^{-1}(X_s)|^2\,ds\Big)$$

on $\mathcal{F}_t$. We see that

$$\begin{aligned}\Big\langle \int_0^{\cdot} \psi'(s)\sigma^{-1}(X_s)\,dW_s, X\Big\rangle \\ &= \Big\langle \int_0^{\cdot} \psi'(s)\sigma^{-1}(X_s)\,dW_s, \int_0^{\cdot} \sigma(X_s)\,dW_s\Big\rangle \\ &= \int_0^t \psi'(s)\,ds = \psi(t) - \psi(0).\end{aligned}$$

So by the Girsanov theorem, under $\mathbb{Q}$ each component of X_t is a semimartingale and $X^i_t - \int_0^t b_i(X_s)\,ds - \psi_i(t)$ is a martingale for each i. Furthermore, if

$$\widehat{W}_t = \int_0^t \sigma^{-1}(X_t)\,[dX_t - b(X_t)\,dt - \psi'(t)\,dt],$$

each component of $\widehat{W}$ is a continuous martingale, and a calculation shows that $d\langle \widehat{W}^i, \widehat{W}^j\rangle_t = \delta_{ij}\,dt$ under $\mathbb{Q}$. Therefore $\widehat{W}$ is a d-dimensional Brownian motion under $\mathbb{Q}$. Since

$$d(X_t - \psi(t)) = \sigma(X_t)\,d\widehat{W}_t + b(X_t)\,dt,$$

then by Theorem 15.4,

$$\mathbb{Q}(\sup_{s\le t} |X_s - \psi(s)| < \varepsilon) \ge c_3.$$

Very similarly to the last paragraph of the proof of Lemma 15.3, we conclude

$$\mathbb{P}(\sup_{s\le t} |X_s - \psi(s)| < \varepsilon) \ge c_4. \qquad \square$$

16. Convexity.

In this section we will let the a_{ij} be smooth (C^2, say) and strictly elliptic, and assume that the drift coefficients are identically 0. Let D be either $B(0,1)$ or a unit cube centered at 0.

Suppose u is continuous. The *upper contact set* of u is the set

$$U_u = \{y \in D : \text{ there exists } p \in \mathbb{R}^d \text{ such that } \\ u(x) \le u(y) + p \cdot (x-y) \text{ for all } x \in D\}.$$

Here $p \cdot (x-y)$ denotes the inner product. In this definition p will depend on y. A point y is in U_u if there is a hyperplane, namely, $u(x) = u(y) + p \cdot (x-y)$, that lies above the graph of u but touches the graph at $(y, u(y))$. With this interpretation we see that when u is concave (i.e., $-u$ is convex), then $U_u = D$, and conversely, if $U_u = D$, then u is concave.

When $u \in C^1$, for $y \in U_u$ there is only one p such that $u(x) \le u(y) + p \cdot (x-y)$, namely, $p = \nabla u(y)$. For $u \in C^2$ let H_u denote the *Hessian* matrix:

$$(H_u)_{ij}(x) = \partial_{ij} u(x).$$

Proposition 16.1. *If $u \in C^2$ and $y \in U_u$, then $H_u(y)$ is nonpositive definite.*

Proof. Let h be a unit vector. $y \in U_u$ implies there exists p such that $u(y + \varepsilon h) \le u(y) + \varepsilon p \cdot h$ and $u(y - \varepsilon h) \le u(y) - \varepsilon p \cdot h$. Combining,

$$u(y + \varepsilon h) + u(y - \varepsilon h) - 2u(y) \le 0.$$

Dividing by ε^2 and letting $\varepsilon \to 0$ gives $h^T H_u(y) h \le 0$. □

Let $S_u(y)$ be the set of slopes of supporting hyperplanes to u at y. That is,

$$S_u(y) = \{p \in \mathbb{R}^d : u(x) \le u(y) + p \cdot (x-y) \text{ for all } x \in D\}.$$

As we noted above, $S_u(y) \ne \emptyset$ if and only if $y \in U_u$, and if $u \in C^1$ and $y \in U_u$, then $S_u(y) = \{\nabla u(y)\}$. Let

$$S_u(A) = \bigcup_{y \in A} S_u(y).$$

Let $|A|$ denote the Lebesgue measure of A and $\det H$ the determinant of H. Recall that if V is a neighborhood in D, $v : D \to \mathbb{R}^d$ is in C^1, and $v(V)$ is the image of V under v, then

$$|v(V)| \le \int_V |\det J_v|, \tag{16.1}$$

where J_v is the Jacobian of v. (We have inequality instead of equality because we are not assuming v is one-to-one.)

Proposition 16.2. *Suppose u is continuous on $\overline{D}$ and C^2 in D. There exists c_1 not depending on u such that*

$$\sup_D u \le \sup_{\partial D} u + c_1 \Big(\int_{U_u} |\det H_u| \Big)^{1/d}.$$

Proof. Replacing u by $u - \sup_{\partial D} u$, we may assume $\sup_{\partial D} u = 0$. We first show

$$|S_u(D)| = |S_u(U_u)| \le \int_{U_u} |\det H_u|. \tag{16.2}$$

Since $S_u(y) = \{\nabla u(y)\}$, the Jacobian matrix of the mapping S_u is H_u.

Next suppose u takes a positive maximum at $y \in D$. Let v be the function such that the region below the graph of v is the cone with base D and vertex $(y, u(y))$. More precisely, let G_1 be the smallest convex set in $D \times [0, \infty)$ containing $\partial D \times \{0\}$ and the point $(y, u(y))$; let $v(x) = \sup\{z \ge 0 : (x, z) \in G_1\}$ for $x \in D$.

Suppose $p \in S_v(D)$. We look at the family of hyperplanes $\alpha + p \cdot (x - y)$. If we start with α large and let α decrease to $-\infty$, there is a first hyperplane that touches the graph of u (not necessarily at $(y, u(y))$). Consequently $p \in S_u(D)$. We have thus shown that $S_v(D) \subseteq S_u(D)$.

We see that

$$|S_v(D)| \le |S_u(D)| \le \int_{U_u} |\det H_u|. \tag{16.3}$$

We now compute $|S_v(\{y\})|$. If each coordinate of p is between $-u(y)/d$ and $+u(y)/d$, then $p \in S_v(y)$. So

$$|S_v(D)| \ge |S_v(\{y\})| \ge c_2 (u(y)/d)^d.$$

Combining with (16.2),

$$u(y)^d \le c_2^{-1} d^d |S_v(D)| \le c_3 \int_{U_u} |\det H_u|. \qquad \square$$

We will use the inequality

$$\frac{1}{d} \sum_{j=1}^d \lambda_j \ge \prod_{j=1}^d \lambda_j^{1/d}, \qquad \lambda_j \ge 0, \quad j = 1, \ldots, d. \tag{16.4}$$

One way to prove (16.4) is to let $\Omega = \{1, 2, \ldots, d\}$, let $\mathbb{P}$ assign mass $1/d$ to each point of Ω, let X be the random variable defined by $X(j) = \lambda_j$, and apply Jensen's inequality to the convex function $-\log x$. We then have

$$-\log \Big(\sum_{j=1}^d \lambda_j \frac{1}{d} \Big) \le \frac{1}{d} \sum_{j=1}^d (-\log \lambda_j),$$

which implies (16.4).

We now prove a key estimate due to Alexandroff-Bakelman-Pucci.

Theorem 16.3. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$, the coefficients of $\mathcal{L}$ are in C^2, $u \in C^2$, and $\mathcal{L}u = f$ in D. There exists c_1 independent of u such that*

$$\sup_D u \le \sup_{\partial D} u + c_1 \Big(\int_D |f(x)|^d \, dx \Big)^{1/d}.$$

Proof. Fix $y \in U_u$, let $B = -H_u(y)$, and let A be the matrix $a(y)$. Let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of B. Since H_u is nonpositive definite, $\lambda_j \ge 0$. Let P be an orthogonal matrix and C a diagonal matrix such that $B = P^T CP$. Note $|\det H_u| = \det B = \lambda_1 \cdots \lambda_d$ and

$$(AB)_{ii} = \sum_{j=1}^d A_{ij} B_{ji} = -\sum_j a_{ij}(y) \partial_{ij} u(y).$$

Then

$$\begin{aligned} -f(y) &= -\sum_{i,j} a_{ij}(y) \partial_{ij} u(y) = \text{ trace}\,(AB) \\ &= \text{ trace}\,(AP^T CP) = \text{ trace}\,(CPAP^T) = \sum_{j=1}^d \lambda_j (PAP^T)_{jj}. \end{aligned} \tag{16.5}$$

Since A is uniformly positive definite, there exists c_2 such that $(PAP^T)_{jj} \ge c_2$, so by (16.4),

$$\begin{aligned} -f(y) &\ge \sum_j c_2 \lambda_j = c_2 d \sum_j (\lambda_j / d) \\ &\ge c_2 d (\prod_j \lambda_j)^{1/d} = c_2 d |\det H_u|^{1/d}. \end{aligned}$$

Taking dth powers, integrating over U_u, and using Proposition 16.2 completes the proof.
□

17. Green functions.

Let $\mathbb{P}$ be a solution to the martingale problem for $\mathcal{L}$ started at x (assuming one exists) and let $\mathbb{E}$ be the corresponding expectation. If D is a domain, a function $G_D(x,y)$ is called a *Green function* for the operator $\mathcal{L}$ in the domain D if

$$\mathbb{E} \int_0^{\tau_D} f(X_s) \, ds = \int_D G_D(x,y) f(y) \, dy \tag{17.1}$$

for all nonnegative Borel measurable functions f on D. The function $G^\lambda(x,y)$ is called the *λ-resolvent density* if

$$\mathbb{E} \int_0^\infty e^{-\lambda s} f(X_s) \, ds = \int_{\mathbb{R}^d} G^\lambda(x,y) f(y) \, dy \tag{17.2}$$

for all nonnegative Borel measurable f on $\mathbb{R}^d$.

An immediate consequence of the Alexandroff-Bakelman-Pucci estimate is the following.

Theorem 17.1. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$ and the diffusion coefficients are in C^2. Then there exists c_1 depending only on Λ such that*

$$\Big|\mathbb{E} \int_0^{\tau_{B(0,1)}} f(X_s)\,ds\Big| \le c_1 \Big(\int_{B(0,1)} |f(y)|^d\,dy\Big)^{1/d}.$$

Proof. We prove this inequality for f that are C^2 in $\overline{B(0,1)}$; a limit argument then yields the inequality for arbitrary f. Let $u(y) = \mathbb{E}^{\,y} \int_0^{\tau_{B(0,1)}} f(X_s)\,ds$. u is C^2 in $B(0,1)$, continuous on the closure of $B(0,1)$, and $\mathcal{L}u = -f$. In fact, u is 0 on the boundary of $B(0,1)$. Now apply Theorem 4.3. $\square$

Corollary 17.2.

$$G_B(x,\cdot) \in L^{d/(d-1)}(B).$$

Proof. By Theorem 17.1 and (17.1),

$$\Big|\int_B G_B(x,y) f(y)\,dy\Big| \le c_1 \|f\|_{L^d(B)}$$

for all $f \in L^d(B)$. The result follows by the duality of L^d and $L^{d/(d-1)}$. $\square$

We also have

Theorem 17.3. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$ and the diffusion coefficients are in C^2. There exists c_1 not depending on f such that if $f \in L^d$, then*

$$\Big|E \int_0^\infty e^{-\lambda t} f(X_t)\,dt\Big| \le c_1 \Big(\int_{\mathbb{R}^d} |f(y)|\,dy\Big)^{1/d}.$$

Proof. By the smoothness of the diffusion coefficients, there is a unique solution to the martingale problem for $\mathcal{L}$ starting at each $x \in \mathbb{R}^d$; we denote it $\mathbb{P}^x$. Moreover, $(\mathbb{P}^x, X_t)$ forms a strong Markov family.

Let $S_0 = 0$ and $S_{i+1} = \inf\{t > S_i : |X_t - X_{S_i}| > 1\}$, $i = 0, 1, \ldots$ Then $S_{i+1} = S_i + S_1 \circ \theta_{S_i}$. By Proposition 2.3, there exists t_0 such that $\sup_x \mathbb{P}^x(S_1 \le t_0) \le 1/2$. Then

$$\begin{aligned}\mathbb{E}^{\,x} e^{-\lambda S_1} &\le \mathbb{P}^x(S_1 \le t_0) + e^{-\lambda t_0}\mathbb{P}^x(S_1 > t_0)\\ &= (1 - e^{-\lambda t_0})\mathbb{P}^x(S_1 \le t_0) + e^{-\lambda t_0}.\end{aligned}$$

So if $\rho = \sup_x \mathbb{E}^x e^{\lambda S_1}$, then $\rho < 1$. By the strong Markov property,

$$\mathbb{E}^x e^{-\lambda S_{i+1}} = \mathbb{E}^x \Big(e^{-\lambda S_i} \mathbb{E}^x (e^{-\lambda S_1 \circ \theta_{S_i}} \mid \mathcal{F}_{S_i}) \Big) \le \rho \mathbb{E}^x e^{-\lambda S_i},$$

and by induction $\mathbb{E}^x e^{-\lambda S_i} \le \rho^i$.

We now write

$$\mathbb{E}^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt = \sum_{i=0}^\infty \int_{S_i}^{S_{i+1}} e^{-\lambda t} f(X_t)\, dt. \tag{17.3}$$

By the strong Markov property at time S_i and Theorem 17.1,

$$\begin{aligned} \Big| \mathbb{E}^x \int_{S_i}^{S_{i+1}} e^{-\lambda t} f(X_t)\, dt \Big| &= \Big| \mathbb{E}^x \Big(e^{-\lambda S_i} \mathbb{E}^{X_{S_i}} \int_0^{S_1} e^{-\lambda t} f(X_t)\, dt \Big) \Big| \\ &\le c_2 E^x e^{-\lambda S_i} \|f\|_d \le c_2 \rho^i \|f\|_d. \end{aligned}$$

Substituting in (17.3) proves the theorem. □

This implies $G^\lambda(x, \cdot) \in L^{d/(d-1)}$.

One disadvantage of Theorems 17.1 and 17.3 is that we required the diffusion coefficients to be smooth. We will remove this restriction in the next section by an approximation procedure due to Krylov. Earlier Krylov had also proved, however, that Theorems 17.1 and 17.3 hold whenever $X_t = x + \int_0^t \sigma_s\, dW_s$, where $\sigma_s(\omega)$ is an adapted, matrix-valued process that is bounded and is uniformly positive definite (that is, there exists c_1 such that $y^T \sigma_s(\omega) y \ge c_1 |y|^2$ for all $y \in \mathbb{R}^d$, where c_1 is independent of s and y).

18. Resolvents.

In this section we present a theorem of Krylov on approximating resolvents and then apply it to extend Theorem 17.3 to arbitrary solutions of the martingale problem for an elliptic operator $\mathcal{L}$. We suppose that $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$ for some $\Lambda > 0$, but make no smoothness assumptions on the coefficients. Let $\mathbb{P}$ be any solution to the martingale problem for $\mathcal{L}$ started at x_0.

Recall that $f * g(x) = \int f(y) g(x-y)\, dy$. Let φ be a nonnegative radially symmetric C^∞ function with compact support such that $\int_{\mathbb{R}^d} \varphi = 1$ and $\varphi > 0$ on $B(0, r)$ for some r. Let $\varphi_\varepsilon(x) = \varepsilon^{-d} \varphi(x/\varepsilon)$.

Theorem 18.1. *Let $\lambda > 0$. There exist a_{ij}^ε in C^∞ with the following properties:*

(i) if $\mathcal{L}^\varepsilon$ is defined by

$$\mathcal{L}^\varepsilon f(x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}^\varepsilon(x) \partial_{ij} f(x), \tag{18.1}$$

then $\mathcal{L}^\varepsilon \in \mathcal{N}(\Lambda, 0)$, *and*

(ii) if $\mathbb{P}^x_\varepsilon$ *is the solution to the martingale problem for* $\mathcal{L}^\varepsilon$ *started at* x *and*

$$G^\lambda_\varepsilon h(x) = \mathbb{E}^{\,x}_{\,\varepsilon} \int_0^\infty e^{-\lambda t} h(X_t)\,dt \tag{18.2}$$

for h *bounded, then*

$$(G^\lambda_\varepsilon f * \varphi_\varepsilon)(x_0) \to \mathbb{E} \int_0^\infty e^{-\lambda t} f(X_t)\,dt \tag{18.3}$$

whenever f *is continuous.*

It turns out that $G^\lambda_\varepsilon f$ is equicontinuous in ε, so that in fact $G^\lambda_\varepsilon f(x_0)$ converges to the right-hand side of (18.3).

The a^ε_{ij} depend on $\mathbb{P}$, and different solutions to the martingale problem could conceivably give us different sequences a^ε_{ij}.

Proof. Define a measure μ by

$$\mu(C) = \mathbb{E} \int_0^\infty e^{-\lambda t} 1_C(X_t)\,dt. \tag{18.4}$$

By the support theorem, for each $y \in \mathbb{R}^d$ and $s > 0$, there is positive probability under $\mathbb{P}$ that X_t enters the ball $B(y,s)$ and stays there a positive length of time. So $\mu(B(y,s)) > 0$ for all y and s. Define

$$a^\varepsilon_{ij}(x) = \frac{\int \varphi_\varepsilon(x-y) a_{ij}(y)\,\mu(dy)}{\int \varphi_\varepsilon(x-y)\,\mu(dy)}. \tag{18.5}$$

By our assumptions on φ, the denominator is not zero. It is clear that (i) holds.

Suppose u is in C^2 and bounded. By the product formula and Itô's formula,

$$\begin{aligned} e^{-\lambda t} u(X_t) &= u(X_0) - \int_0^t u(X_s)\lambda e^{-\lambda s}\,ds + \int_0^t e^{-\lambda s}\,d[u(X)]_s \\ &= u(X_0) - \int_0^t u(X_s)\lambda e^{-\lambda s}\,ds + \text{ martingale} \\ &\qquad + \int_0^t e^{-\lambda s} \mathcal{L}u(X_s)\,ds. \end{aligned}$$

Taking expectations and letting $t \to \infty$,

$$u(x_0) = \mathbb{E} \int_0^\infty e^{-\lambda s}(\lambda u - \mathcal{L}u)(X_s)\,ds = \int (\lambda u - \mathcal{L}u)(x)\,\mu(dx). \tag{18.6}$$

We next apply (18.6) to $u = v * \varphi_\varepsilon$, where v is a bounded and C^2 function. On the left-hand side we have $\int v(x_0 - y)\varphi_\varepsilon(y)\,dy$. Note that

$$\begin{aligned}
\mathcal{L}(v * \varphi_\varepsilon)(z) &= \frac{1}{2}\sum_{i,j} a_{ij}(z)\partial_{ij}(v * \varphi_\varepsilon)(z) \\
&= \frac{1}{2}\sum_{i,j} a_{ij}(z)((\partial_{ij}v) * \varphi_\varepsilon)(z) \\
&= \frac{1}{2}\sum_{i,j} \int a_{ij}(z)\partial_{ij}v(x)\varphi_\varepsilon(x-z)\,dx.
\end{aligned} \tag{18.7}$$

However, by (18.5),

$$\int a_{ij}(z)\varphi_\varepsilon(x-z)\,\mu(dz) = a^\varepsilon_{ij}(x)\int \varphi_\varepsilon(x-y)\,\mu(dy). \tag{18.8}$$

Combining (18.6), (18.7), and (18.8),

$$\begin{aligned}
\int v(x_0 - y)\varphi_\varepsilon(y)\,dy &= \int [\lambda(v * \varphi_\varepsilon) - \mathcal{L}(v * \varphi_\varepsilon)](x)\,\mu(dx) \\
&= \int\int (\lambda - \mathcal{L}^\varepsilon)v(x)\varphi_\varepsilon(x-y)\,\mu(dy)\,dx.
\end{aligned} \tag{18.9}$$

Suppose f is smooth, and let $v(x) = G^\lambda_\varepsilon f(x)$. v is in C^2 and bounded and $(\lambda - \mathcal{L}^\varepsilon)v = f$. Substituting in (18.9),

$$\begin{aligned}
\int G^\lambda_\varepsilon f(x_0 - y)\varphi_\varepsilon(y)\,dy &= \int\int f(x)\varphi_\varepsilon(x-y)\,\mu(dy)\,dx \\
&= \int f * \varphi_\varepsilon(y)\,\mu(dy).
\end{aligned} \tag{18.10}$$

By a limit argument, we have (18.10) when f is continuous. Since f is continuous, $f * \varphi_\varepsilon$ is bounded and converges to f uniformly. Hence

$$\int f * \varphi_\varepsilon(y)\,\mu(dy) \to \int f(y)\mu(dy) = \mathbb{E}\int_0^\infty e^{-\lambda t} f(X_t)\,dt. \qquad \square$$

It is easy to see that if the a_{ij} are continuous, then a^ε_{ij} converges to a_{ij} pointwise.

Defining b^ε_i by the analogue of (18.5), there is no difficulty extending this theorem to the case $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$, $\Lambda_2 > 0$.

Theorem 18.2. *Let $\mathbb{P}$ be as above. There exists c_1 not depending on f such that*

$$\Big|\mathbb{E}\int_0^\infty e^{-\lambda t} f(X_t)\,dt\Big| \le c_1 \|f\|_d.$$

Proof. By Theorem 18.1, the left-hand side is the limit of $|G^\varepsilon_\lambda f * \varphi_\varepsilon(x_0)|$ if f is continuous and bounded. The coefficients in $\mathcal{L}^\varepsilon$ are smooth, so by Theorem 17.3 $\|G^\lambda_\varepsilon f\|_\infty \le c_1\|f\|_d$, c_1 independent of ε. This proves the proposition for f smooth, and the case of general f follows by a limit argument. □

Corollary 18.3. *Under the assumptions of Theorem 18.1,*

$$(G^\lambda_\varepsilon f * \varphi_\varepsilon)(x_0) \to \mathbb{E}\int_0^\infty e^{-\lambda t} f(X_t)\,dt,$$

if f is bounded.

Proof. We start with (18.10). By a limit argument, we have (18.10) holding for f bounded. So we need to show that the right-hand side of (18.10) converges to $\int f(y)\,\mu(dy)$. Since f is bounded, $f * \varphi_\varepsilon$ converges to f almost everywhere and boundedly. By Theorem 18.2 and (18.4), μ is absolutely continuous with respect to Lebesgue measure. Then

$$\begin{aligned}\int f * \varphi_\varepsilon(y)\,\mu(dy) &= \int f * \varphi_\varepsilon(y)(d\mu/dy)\,dy\\ &\to \int f(y)(d\mu/dy)\,dy = \int f(y)\,\mu(dy)\end{aligned}$$

by dominated convergence. □

19. Harnack inequality.

In this section we prove some theorems of Krylov and Safonov concerning positive $\mathcal{L}$-harmonic functions. Recall that a function h is $\mathcal{L}$-harmonic in a domain D if $h \in C^2$ and $\mathcal{L}h = 0$ in D. These results were first proved probabilistically by Krylov and Safonov and are a good example of the power of the probabilistic point of view.

In this section we assume that $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$ so that the drift coefficients are 0. We assume that for each $x \in \mathbb{R}^d$ we have a solution to the martingale problem for $\mathcal{L}$ started at x and that $(\mathbb{P}^x, X_t)$ forms a strong Markov family.

Let $Q(x,r)$ denote the cube of side length r centered at x. Our main goal is to show that X_t started at x must hit a set A before exiting a cube with positive probability if A has positive Lebesgue measure and x is not too near the boundary. The first proposition starts things off by handling the case when A nearly fills the cube. Recall that we are using $|A|$ to denote the Lebesgue measure of A.

Proposition 19.1. *There exist ε and $c_1 = c_1(\varepsilon)$ such that if $x \in Q(0,1/2)$, $A \subseteq Q(0,1)$, and $|Q(0,1) - A| < \varepsilon$, then $\mathbb{P}^x(T_A < \tau_{Q(0,1)}) \geq c_1$.*

Proof. Let us write τ for $\tau_{Q(0,1)}$. Recall there exist c_2 and c_3 not depending on x such that $\mathbb{E}^x \tau \geq c_2$ and $\mathbb{E}^x \tau^2 \leq c_3$.

Note that $\mathbb{E}^x \int_0^\tau 1_{A^c}(X_s)\, ds = \mathbb{E}^x \int_0^\tau 1_{(Q(0,1)-A)}(X_s)\, ds$. Since

$$\mathbb{E}^x(\tau - (\tau \wedge t_0)) \leq \mathbb{E}^x(\tau; \tau \geq t_0) \leq \mathbb{E}^x \tau^2/t_0,$$

we can choose t_0 large enough so that $\mathbb{E}^x(\tau - (\tau \wedge t_0)) \leq c_2/4$. Then

$$\begin{aligned}
\mathbb{E}^x &\int_0^\tau 1_{(Q(0,1)-A)}(X_s)\, ds \qquad (19.1)\\
&\leq c_2/4 + e^{t_0} \mathbb{E}^x \int_0^{t_0} e^{-s} 1_{(Q(0,1)-A)}(X_s)\, ds\\
&\leq c_2/4 + e^{t_0} \mathbb{E}^x \int_0^{\infty} e^{-s} 1_{(Q(0,1)-A)}(X_s)\, ds\\
&\leq c_2/4 + c_5 e^{t_0} \|1_{Q(0,1)-A}\|_d\\
&\leq c_2/4 + c_5 e^{t_0} \varepsilon^{1/d}.
\end{aligned}$$

If ε is chosen small enough, then $\mathbb{E}^x \int_0^\tau 1_{A^c}(X_s)\, ds < c_2/2$.

On the other hand,

$$\begin{aligned}
c_2 \leq \mathbb{E}^x \tau &= \mathbb{E}^x(\tau; T_A < \tau) + \mathbb{E}^x \int_0^\tau 1_{A^c}(X_s)\, ds\\
&\leq (\mathbb{E}^x \tau^2)^{1/2} (\mathbb{P}^x(T_A < \tau))^{1/2} + c_2/2\\
&\leq c_3^{1/2} (\mathbb{P}^x(T_A < \tau))^{1/2} + c_2/2,
\end{aligned}$$

and the result follows with $c_1 = c_2^2/4c_3$. □

We used Theorem 18.2 because it applies to arbitrary solutions to the martingale problem, whereas Theorem 17.1 requires the a_{ij} to be smooth. As noted at the end of Section 5, Theorem 17.1 actually holds for arbitrary solutions to the martingale problem; if we used that fact, we then could have obtained the estimate in (19.1) more directly.

Next we decompose $Q(0,1)$ into smaller subcubes such that a set A fills up a certain percentage of each of the smaller subcubes. If Q is a cube, let $\widehat{Q}$ denote the cube with the same center as Q but side length three times as long.

Proposition 19.2. *Let $q \in (0,1)$. If $A \subseteq Q(0,1)$ and $|A| \le q$, then there exists D such that (i) D is the union of cubes $\widehat{R}_i$ such that the interiors of the R_i are pairwise disjoint, (ii) $|A| \le q|D \cap Q(0,1)|$, and (iii) for each i, $|A \cap R_i| > q|R_i|$.*

Proof. We will do the case $d = 2$; the higher-dimensional case differs only in the notation. We form a collection of subsquares $\mathcal{R} = \{R_i\}$ as follows. Divide $Q(0,1)$ into four equal squares Q_1, Q_2, Q_3, and Q_4 with disjoint interiors. For $j = 1,2,3,4$, if $|A \cap Q_j| > q|Q_j|$, we let Q_j be one of the squares in $\mathcal{R}$. If not, we split Q_j into four equal subsquares $Q_{j1}, Q_{j2}, Q_{j3}, Q_{j4}$ and repeat; Q_{jk} will be one of the R_i if $|A \cap Q_{jk}| > q|Q_{jk}|$, and otherwise we divide Q_{jk}. To be more precise, let $\mathcal{Q}_n$ be the collection of squares of side lengths 2^{-n} with vertices of the form $[j/2^n, k/2^n]$ for integers j and k. An element Q' of $\mathcal{Q}_n$ will be in $\mathcal{R}$ if $|A \cap Q'| > q|Q'|$ and Q' is not contained in any $Q'' \in \mathcal{Q}_0 \cup \mathcal{Q}_1 \cup \cdots \cup \mathcal{Q}_{n-1}$ with $|A \cap Q''| > q|Q''|$.

We let $D = \cup_i \widehat{R}_i$ where the union is over $R_i \in \mathcal{R}$. Assertions (i) and (iii) are clear and it remains to prove (ii). Recall that almost every point $z \in A$ is a point of density of A, that is, $|B(z,r) \cap A|/|B(z,r)| \to 1$ a.e. for $z \in A$; this follows by the Lebesgue density theorem. If z is a point of density of A and T_n denotes the element of $\mathcal{Q}_n$ containing z, then $|T_n \cap A|/|T_n| \to 1$. If z is a point of density of A and z is not on the boundary of some square in $\mathcal{Q}_n$ for some n, it follows that z must be in some $R_i \in \mathcal{R}$. We conclude that $|A - D| = 0$.

We form a new collection of squares $\mathcal{S}$. We divide $Q(0,1)$ into four equal subsquares Q_1, Q_2, Q_3, Q_4. If $Q_j \subseteq D$, it will be in $\mathcal{S}$; otherwise split Q_j into four subsquares and continue. More exactly, $Q' \in \mathcal{Q}_n$ will be in $\mathcal{S}$ if $Q' \subseteq D$ but Q' is not contained in any $Q'' \in \mathcal{Q}_0 \cup \cdots \mathcal{Q}_{n-1}$ for which $Q'' \subseteq D$.

Since D is the union of cubes $\widehat{R}_i$, then $|D \cap Q(0,1)| = \sum_i |S_i|$ where the sum is over $S_i \in \mathcal{S}$. Since almost every point of A is contained in D and almost every point of D is in one of the S_i's that are in $\mathcal{S}$, we conclude $|A| = \sum_i |S_i \cap A|$. It thus suffices to show that

$$|A \cap S_i| \le q|S_i| \tag{19.2}$$

for each $S_i \in S$. We then sum over i and (ii) will be proved.

Consider $S_i \in \mathcal{S}$. If $S_i = Q(0,1)$, we are done by the hypotheses on A. Otherwise S_i is in $\mathcal{Q}_n$ for some $n \ge 1$ and is contained in a square $Q' \in \mathcal{Q}_{n-1}$. Let C_1, C_2, C_3 denote the other three squares of $\mathcal{Q}_n$ that are contained in Q'. Since $S_i \in \mathcal{S}$, then $Q' = S_i \cup C_1 \cup C_2 \cup C_3$ is not in $\mathcal{S}$. Since $S_i \subseteq D$, at least one of the squares C_1, C_2, C_3 cannot be contained in D. We have $S_i \cup C_1 \cup C_2 \cup C_3 \subseteq \widehat{S}_i$. $\widehat{S}_i$ is not contained in D, which implies that $S_i \notin \mathcal{R}$. We thus have $S_i \cup C_1 \cup C_2 \cup C_3$ is not contained in D but $S_i \notin \mathcal{R}$; this could only happen if $|S_i \cap A| \le q|S_i|$, which is (19.2). $\square$

Lemma 19.3. *Let $r \in (0,1)$. Let $y \in Q(0,1)$ with* $\operatorname{dist}(y, \partial Q(0,1)) > r$, $\mathcal{L}' \in \mathcal{N}(\Lambda, 0)$, *and P be a solution to the martingale problem for $\mathcal{L}'$ started at y. If $Q(z,r) \subseteq Q(0,1)$, then $\mathbb{P}(T_{Q(z,r)} < \tau_{Q(0,1)}) \geq \zeta(r)$ where $\zeta(r) > 0$ depends only on r and Λ.*

Proof. This follows easily from the support theorem. □

We now prove the key result, that sets of positive Lebesgue measure are hit with positive probability.

Theorem 19.4. *There exists a nondecreasing function $\varphi : (0,1) \to (0,1)$ such that if $B \subseteq Q(0,1)$, $|B| > 0$, and $x \in Q(0,1/2)$, then*

$$\mathbb{P}^x(T_B < \tau_{Q(0,1)}) \geq \varphi(|B|).$$

Proof. Again we suppose the dimension d is 2 for simplicity of notation. Set

$$\begin{aligned}\varphi(\varepsilon) = \inf\{\mathbb{P}^y(T_B <&\tau_{Q(z_0,R)}) : z_0 \in \mathbb{R}^d, R > 0, y \in Q(z_0, R/2),\\ &|B| \geq \varepsilon|Q(z_0,R)|, B \subseteq Q(z_0,R)\}.\end{aligned}$$

By Proposition 19.1 and scaling, $\varphi(\varepsilon) > 0$ for ε sufficiently close to 1. Let q_0 be the infimum of those ε for which $\varphi(\varepsilon) > 0$. We suppose $q_0 > 0$, and we will obtain our contradiction.

Choose $q > q_0$ such that $(q + q^2)/2 < q_0$. This is possible, since $q_0 < 1$ and so $(q_0 + q_0^2)/2 < q_0$. Let $\eta = (q - q^2)/2$. Let $\beta = (q \wedge (1-q))/16$ and let ρ be equal to $\zeta((1-\beta)/6)$ as defined in Lemma 19.3. There exist $z_0 \in \mathbb{R}^d$, $R > 0$, $B \subseteq Q(z_0,R)$, and $x \in Q(z_0, R/2)$ such that $q > |B|/|Q(z_0,R)| > q - \eta$ and $\mathbb{P}^x(T_B < \tau_{Q(z_0,R)}) < \rho\varphi(q)^2$. Without loss of generality, let us assume $z_0 = 0$ and $R = 1$, and so we have $\mathbb{P}^x(T_B < \tau_{Q(0,1)}) < \rho\varphi(q)^2$.

We next use Proposition 19.2 to construct the set D (with A replaced by B). Since $|B| > q - \eta$ and

$$|B| \leq q|D \cap Q(0,1)|,$$

then

$$|D \cap Q(0,1)| \geq \frac{|B|}{q} > \frac{q-\eta}{q} = \frac{q+1}{2}.$$

Let $\widetilde{D} = D \cap Q(0, 1-\beta)$. Then $|\widetilde{D}| > q$. By the definition of φ, this implies that

$$\mathbb{P}^x(T_{\widetilde{D}} < \tau_{Q(0,1)}) \geq \varphi(q).$$

We want to show that if $y \in \widetilde{D}$, then

$$\mathbb{P}^y(T_B < \tau_{Q(0,1)}) \geq \rho\varphi(q). \tag{19.3}$$

Once we have that, we write

$$\begin{aligned}\mathbb{P}^x(T_B < \tau_{Q(0,1)}) &\geq \mathbb{P}^x(T_{\widetilde{D}} < T_B < \tau_{Q(0,1)})\\ &\geq \mathbb{E}^x(\mathbb{P}^{X(T(\widetilde{D}))}(T_B < \tau_{Q(0,1)}); T_{\widetilde{D}} < \tau_{Q(0,1)})\\ &\geq \rho\varphi(q)\mathbb{P}^x(T_{\widetilde{D}} < \tau_{Q(0,1)}) \geq \rho\varphi(q)^2,\end{aligned}$$

our contradiction.

We now prove (19.3). If $y \in \partial\widetilde{D}$, then $y \in \widehat{R}_i$ for some $R_i \in \mathcal{R}$ and dist $(y, \partial Q(0,1)) \geq 1 - \beta$. Let R_i^* be the cube with the same center as R_i but side length half as long. By Lemma 19.3,

$$\mathbb{P}^y(T_{R_i^*} < \tau_{Q(0,1)}) \geq \rho.$$

By the definition of q and the fact that $R_i \in \mathcal{R}$, then $|B \cap R_i| \geq q|R_i|$. By the definition of $\varphi(q)$, we have $\mathbb{P}^z(T_{B\cap R_i} < \tau_{R_i}) \geq \varphi(q)$ if $z \in R_i^*$. So by the strong Markov property,

$$\begin{aligned}\mathbb{P}^y(T_B < \tau_{Q(0,1)}) &\geq \mathbb{E}^y(\mathbb{P}^{X(T_{R_i^*})}(T_B < \tau_{R_i}); T_{R_i^*} < \tau_{Q(0,1)})\\ &\geq \rho\varphi(q).\end{aligned}$$

□

Theorem 19.4 is the key estimate. We now proceed to show that $\mathcal{L}$-harmonic functions are Hölder continuous and that they satisfy a Harnack inequality. A function h is $\mathcal{L}$-harmonic in D if $h \in C^2$ and $\mathcal{L}h = 0$ in D. If h is $\mathcal{L}$-harmonic, then by Itô's formula, $h(X_{t\wedge\tau_D})$ is a martingale. There may be very few $\mathcal{L}$-harmonic functions unless the coefficients of $\mathcal{L}$ are smooth, so we will use the condition that $h(X_{t\wedge\tau_D})$ is a martingale as our hypothesis.

Theorem 19.5. *Suppose h is bounded in $Q(0,1)$ and $h(X_{t\wedge\tau_{Q(0,1)}})$ is a martingale. Then there exist α and c_1 not depending on h such that*

$$|h(x) - h(y)| \leq c_1\|h\|_\infty |x-y|^\alpha, \qquad x, y \in Q(0,1/2).$$

Proof. Define $\operatorname{Osc}_B h = \sup_{x\in B} h(x) - \inf_{x\in B} h(x)$. To prove the theorem, it suffices to show there exists $\rho < 1$ such that for all $z \in Q(0,1/2)$ and $r \leq 1/4$,

$$\operatorname*{Osc}_{Q(z,r/2)} h \leq \rho \operatorname*{Osc}_{Q(z,r)} h. \tag{19.4}$$

If we look at $Ch + D$ for suitable constants C and D, we see that it is enough to consider the case where $\inf_{Q(z,r)} h = 0$ and $\sup_{Q(z,r)} h = 1$. Let $B = \{x \in Q(z,r/2) : h(x) \geq 1/2\}$. We may assume that $|B| \geq (1/2)|Q(z,r/2)|$, for if not, we replace h by $1-h$.

If $x \in Q(z, r/2)$, then $h(x) \leq 1$. On the other hand, since we know $h(X_{t\wedge\tau_{Q(0,1)}})$ is a martingale,

$$\begin{aligned} h(x) &= \mathbb{E}^x[h(X(\tau_{Q(z,r)} \wedge T_B))] \\ &\geq (1/2)\mathbb{P}^x(T_B < \tau_{Q(z,r)}) \geq (1/2)\varphi(2^{-(d+1)}), \end{aligned}$$

from Theorem 19.4 and scaling. Hence $\operatorname{Osc}_{Q(z,r/2)} h \geq 1 - \varphi(2^{-(d+1)})/2$. Setting $\rho = 1 - \varphi(2^{-(d+1)/2})/2$ proves (19.4). □

Theorem 19.6. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$. There exists c_1 depending only on Λ such that if h is nonnegative, bounded in $Q(0,16)$, and $h(X(t \wedge \tau_{Q(0,16)}))$ is a martingale, then $h(x) \leq c_1 h(y)$ if $x, y \in Q(0,1)$.*

Proof. If we look at $h + \varepsilon$ and let $\varepsilon \to 0$, we may assume $h > 0$. By looking at Ch, we may assume $\inf_{Q(0,1/2)} h = 1$. By Theorem 19.5, we know that h is Hölder continuous in $Q(0,8)$, so there exists $y \in Q(0,1/2)$ such that $h(y) = 1$. We want to show that h is bounded above by a constant in $Q(0,1)$, where the constant depends only on Λ.

By the support theorem and scaling, if $x \in Q(0,2)$, there exists δ such that

$$\mathbb{P}^y(T_{Q(x,1/2)} < \tau_{Q(0,8)}) \geq \delta.$$

By scaling, if $w \in Q(x,1/2)$, then $\mathbb{P}^w(T_{Q(x,1/4)} < \tau_{Q(0,8)}) \geq \delta$. So by the strong Markov property,

$$\mathbb{P}^y(T_{Q(x,1/4)} < \tau_{Q(0,8)}) \geq \delta^2.$$

Repeating and using induction,

$$\mathbb{P}^y(T_{Q(x,2^{-k})} < \tau_{Q(0,8)}) \geq \delta^k.$$

Then

$$\begin{aligned} 1 = h(y) &\geq \mathbb{E}^y[h(X_{T(Q(x,2^{-k}))}); T_{Q(x,2^{-k})} < \tau_{Q(0,8)}] \\ &\geq \delta^k \Big(\inf_{Q(x,2^{-k})} h \Big), \end{aligned}$$

or

$$\inf_{Q(x,2^{-k})} h \leq \delta^{-k}. \tag{19.5}$$

By (19.4) there exists $\rho < 1$ such that

$$\operatorname*{Osc}_{Q(x,2^{-(k+1)})} h \leq \rho \operatorname*{Osc}_{Q(x,2^{-k})} h. \tag{19.6}$$

Take m large so that $\rho^{-m} \geq \delta^{-2}/(\delta^{-1} - 1)$. Let $M = 2^m$. Then

$$\operatorname*{Osc}_{Q(x,M2^{-k})} h \geq \rho^{-m} \operatorname*{Osc}_{Q(x,2^{-k})} h \geq \frac{\delta^{-2}}{\delta^{-1} - 1} \operatorname*{Osc}_{Q(x,2^{-k})} h. \tag{19.7}$$

Take K large so that $\sqrt{d}M2^{-K} < 1/8$. Suppose there exists $x_0 \in Q(y,1)$ such that $h(x_0) \geq \delta^{-K-1}$. We will construct a sequence $x_1, x_2, \ldots$ by induction. Suppose we have $x_j \in Q(x_{j-1}, M2^{-(K+j-1)})$ with $h(x_j) \geq \delta^{-K-j-1}$, $j \leq n$. Since $|x_j - x_{j-1}| < \sqrt{d}M2^{-(K+j-1)}$, $1 \leq j \leq n$, and $|x_0 - y| \leq 1$, then $|x_n - y| < 2$. Since $h(x_n) \geq \delta^{-K-n-1}$ and by (19.5), $\inf_{Q(x_n, 2^{-K-n})} h \leq \delta^{-K-n}$,

$$\operatorname*{Osc}_{Q(x_n,2^{-K-n})} h \geq \delta^{-K-n}(\delta^{-1} - 1).$$

So $\operatorname{Osc}_{Q(x_n, M2^{-K-n})} h \geq \delta^{-K-n-2}$, which implies that there exists $x_{n+1} \in Q(x_n, M2^{-K-n})$ with $h(x_{n+1}) \geq \delta^{-K-n-2}$ because h is nonnegative. By induction we obtain a sequence x_n with $x_n \in Q(y,4)$ and $h(x_n) \to \infty$. This contradicts the boundedness of h on $Q(0,8)$. Therefore h is bounded on $Q(0,1)$ by δ^{-K-1}. □

Corollary 19.7. *Suppose D is a bounded connected open domain and $r > 0$. There exists c_1 depending only on D, Λ, and r such that if h is nonnegative, bounded in D, and $h(X_{t\wedge\tau_D})$ is a martingale, then $h(x) \leq c_1 h(y)$ if $x, y \in D$ and* $\operatorname{dist}(x, \partial D)$ *and* $\operatorname{dist}(y, \partial D)$ *are both greater than r.*

Proof. We form a sequence $x = y_0, y_1, y_2, \ldots, y_m = y$ such that $|y_{i+1} - y_i| < (a_{i+1} \wedge a_i)/32$, where $a_i = \operatorname{dist}(y_i, \partial D)$ and each $a_i < r$. By compactness we can choose M depending only on r so that no more than M points y_i are needed. By scaling and Theorem 19.6, $h(y_i) \leq c_2 h(y_{i+1})$ with $c_2 > 1$. So

$$h(x) = h(y_0) \leq c_2 h(y_1) \leq \cdots \leq c_2^m h(y_m) = c_2^m h(y) \leq c_2^M h(y).$$ □

20. Existence.

In this section we discuss the existence of solutions to the martingale problem for an elliptic operator in nondivergence form. Let $\mathcal{L}$ be the elliptic operator in nondivergence form defined by

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\partial_{ij}f(x) + \sum_{i=1}^{d} b_i(x)\partial_i f(x), \qquad f \in C^2. \tag{20.1}$$

We assume throughout that the a_{ij} and b_i are bounded and measurable. Since the coefficient of $\partial_{ij}f(x)$ is $(a_{ij}(x) + a_{ji}(x))/2$, there is no loss of generality in assuming that $a_{ij} = a_{ji}$. We let

$$\begin{aligned}\mathcal{N}(\Lambda_1, \Lambda_2) = \{&\mathcal{L} : \sup_{i \leq d} \|b_i\|_\infty \leq \Lambda_2 \text{ and} \\ &\Lambda_1 |y|^2 \leq \sum_{i,j=1}^{d} y_i y_j a_{ij}(x) \leq \Lambda_1^{-1}|y|^2 \text{ for all } x, y \in \mathbb{R}^d\}.\end{aligned} \tag{20.2}$$

If $\mathcal{L} \in \mathcal{N}(A, B)$ for some $A > 0$, then we say $\mathcal{L}$ is uniformly elliptic.

A probability measure $\mathbb{P}$ is a *solution to the martingale problem for $\mathcal{L}$ started at x* if

$$\mathbb{P}(X_0 = x) = 1 \tag{20.3}$$

and

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,ds \tag{20.4}$$

is a local martingale under $\mathbb{P}$ whenever f is in $C^2(\mathbb{R}^d)$.

We begin by showing that continuity of the coefficients of $\mathcal{L}$ is a sufficient condition for the existence of a solution to the martingale problem. For simplicity let us assume that the a_{ij} are uniformly continuous and that the b_i are 0.

Theorem 20.1. *Suppose the a_{ij} are bounded and uniformly continuous and the b_i are zero. Then there exists a solution to the martingale problem for $\mathcal{L}$ started at x.*

Proof. Let a_{ij}^n be uniformly bounded C^2 functions on $\mathbb{R}^d$ that converge to a_{ij} uniformly on $\mathbb{R}^d$. Let

$$\mathcal{L}_n f(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}^n(x)\partial_{ij}f(x), \tag{20.5}$$

let σ_n be a Lipschitz square root of a^n, and let X^n be the solution to

$$dX_t^n = \sigma^n(X_t^n)\,dW_t, \qquad X_0^n = x,$$

where W_t is a d-dimensional Brownian motion. Let $\mathbb{P}_n$ be the law of X^n. Our desired $\mathbb{P}$ will be a limit point of the sequence $\{\mathbb{P}_n\}$.

Each $\mathbb{P}^n$ is a probability measure on $\Omega = C([0, \infty))$. A collection of continuous functions on a compact set has compact closure if they are uniformly bounded at one point and they are equicontinuous. This implies easily that the $\mathbb{P}_n$ are tight.

Let $\mathbb{P}_{n_k}$ be a subsequence that converges weakly and call the limit $\mathbb{P}$. We must show that $\mathbb{P}$ is a solution to the martingale problem. If g is a continuous function on $\mathbb{R}^d$ with compact support, $g(X_0)$ is a continuous function on Ω, so

$$g(x) = \mathbb{E}_{n_k} g(X_0) \to \mathbb{E}\, g(X_0).$$

Since this is true for all such g, we must have $\mathbb{P}(X_0 = x) = 1$.

Next let $f \in C^2(\mathbb{R}^d)$ be bounded with bounded first and second partial derivatives. To show

$$\mathbb{E}\Big[f(X_t) - f(X_s) - \int_s^t \mathcal{L}f(X_r)\,dr; A\Big] = 0$$

whenever $A \in \mathcal{F}_s$, it suffices to show

$$\mathbb{E}\Big[\Big\{f(X_t) - f(X_s) - \int_s^t \mathcal{L}f(X_r)\,dr\Big\}\prod_{i=1}^m g_i(X_{r_i})\Big] = 0 \tag{20.6}$$

whenever $m \geq 1$, $0 \leq r_1 \leq \cdots \leq r_m \leq s$, and the g_i are continuous functions with compact support on $\mathbb{R}^d$. Setting

$$Y(\omega) = \Big\{f(X_t) - f(X_s) - \int_s^t \mathcal{L}f(X_r)\,dr\Big\}\prod_{i=1}^m g_i(X_{r_i}),$$

Y is a continuous bounded function on Ω, so $\mathbb{E}\,Y = \lim_{k\to\infty}\mathbb{E}_{n_k}Y$. Since $\mathbb{P}_{n_k}$ is a solution to the martingale problem for $\mathcal{L}_{n_k}$,

$$\mathbb{E}_{n_k}\Big[\Big\{f(X_t) - f(X_s) - \int_s^t \mathcal{L}_{n_k}f(X_r)\,dr\Big\}\prod_{i=1}^m g_i(X_{r_i})\Big] = 0.$$

Since the g_i are bounded, it suffices to show

$$\mathbb{E}_{n_k}\Big[\int_s^t |(\mathcal{L}f - \mathcal{L}_{n_k}f)(X_r)|\,dr\Big] \to 0 \tag{20.7}$$

as $k \to \infty$.

Let $\varepsilon > 0$. Choose k large so that $|a_{ij}(y) - a_{ij}^{n_k}(y)| < \varepsilon$ if $i,j = 1,\ldots,d$. Since $f \in C^2$, there exists c_2 such that

$$\mathbb{E}_{n_k}\int_s^t |(\mathcal{L}f - \mathcal{L}_{n_k}f)(X_r)|\,dr \leq c_2(t-s)\varepsilon \leq c_3\varepsilon,$$

which proves (20.7).

Finally, suppose $f \in C^2$ but is not necessarily bounded. Let f_M be a C^2 function that is bounded with bounded first and second partial derivatives and that equals f on $B(0,M)$. If $T_M = \inf\{t : |X_t| \geq M\}$, the above argument applied to f_M shows that $f_M(X_t) - f_M(X_0) - \int_0^t \mathcal{L}f_M(X_s)\,ds$ is a martingale, and hence so is $f(X_{t\wedge T_M}) - f(X_0) - \int_0^{t\wedge T_M}\mathcal{L}f(X_s)\,ds$. Since X_t is continuous, $T_M \to \infty$ a.s., and therefore $f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,ds$ is a local martingale.

□

If the operator $\mathcal{L}$ is uniformly elliptic, we can allow the b_i to be bounded without requiring any other smoothness. If $\mathcal{L}$ is given by (20.1), let $\mathcal{L}'$ be defined by

$$\mathcal{L}'f(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial_{ij}f(x). \tag{20.8}$$

Theorem 20.2. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$. If there exists a solution to the martingale problem for $\mathcal{L}'$ started at x, then there exists a solution to the martingale problem for $\mathcal{L}$ started at x.*

Proof. Let $\mathbb{P}'$ be a solution to the martingale problem for $\mathcal{L}'$ started at x. Let $\sigma(x)$ be a positive definite square root of $a(x)$. Then under $\mathbb{P}'$ X_t^i is a martingale and $d\langle X^i, X^j\rangle_t = a_{ij}(X_t)\,dt$. Letting $W_t = \int_0^t \sigma^{-1}(X_s)\,dX_s$, we see that W_t is a d-dimensional Brownian motion with quadratic variation $\langle W^i, W^j\rangle_t = \delta_{ij}t$. Hence under $\mathbb{P}'$ the process W_t is a Brownian motion and

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

Define a new probability measure $\mathbb{P}$ by setting the restriction of $d\mathbb{P}/d\mathbb{P}'$ to $\mathcal{F}_t$ equal to

$$M_t = \exp\Big(\int_0^t (b\sigma^{-1})(X_s)\,dW_s + \frac{1}{2}\int_0^t |(b\sigma^{-1})(X_s)|^2\,ds\Big). \tag{20.9}$$

Under $\mathbb{P}'$, M_t is a martingale. By the Girsanov theorem, under $\mathbb{P}$ each component of

$$X_t - \Big\langle \int_0^\cdot (b\sigma^{-1})(X_s)\,dW_s, X\Big\rangle_t = X_t - \int_0^t b(X_s)\,ds$$

is a martingale and the quadratic variation of X remains the same. If

$$\widetilde{W}_t = \int_0^t \sigma^{-1}(X_s)\,d\Big(X_s - \int_0^s b(X_r)\,dr\Big),$$

then under $\mathbb{P}$, $\widetilde{W}_t$ is a martingale with $\langle \widetilde{W}^i, \widetilde{W}^j\rangle_t = \delta_{ij}t$, and hence $\widetilde{W}$ is a Brownian motion. Thus

$$dX_t = \sigma(X_t)\,d\widetilde{W}_t + b(X_t)\,dt.$$

$\mathbb{P}$ is therefore a solution to the martingale problem for $\mathcal{L}$. □

As a consequence of Theorems 20.1 and 20.2, there exists a solution to the martingale problem if $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$ for some $\Lambda_1, \Lambda_2 > 0$ and the a_{ij} are continuous.

Even if the a_{ij} are not continuous, a solution to the martingale problem will exist if uniform ellipticity holds.

21. The strong Markov property.

We are not assuming that our solutions are part of a strong Markov family. As a substitute we have the following. Let $\mathbb{P}$ be a solution to the martingale problem for $\mathcal{L}$ started at x and let S be a finite stopping time. Define a probability measure $\mathbb{P}_S$ on $\Omega = C([0,\infty))$ by

$$\mathbb{P}_S(A) = \mathbb{P}(A \circ \theta_S). \tag{21.1}$$

Here θ_S is the shift operator that shifts the path by S. Let $\mathbb{Q}_S(\omega, d\omega')$ be a regular conditional probability for $\mathbb{P}_S[\,\cdot \mid \mathcal{F}_S]$.

Proposition 21.1. *With probability one, $Q_S(\omega, \cdot)$ is a solution to the martingale problem for $\mathcal{L}$ started at $X_S(\omega)$.*

Proof. If $A(\omega) = \{\omega' : X_0(\omega') = X_S(\omega)\}$, we first show that $\mathbb{Q}_S(\omega, A(\omega)) = 1$ for almost every ω. To do this, it suffices to show that

$$\mathbb{P}(B) = \mathbb{E}_{\mathbb{P}}[\mathbb{Q}_S(\omega, A(\omega)); B]$$

whenever $B \in \mathcal{F}_S$. The right-hand side, by the definition of $\mathbb{Q}_S$, is equal to

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[\mathbb{P}_S(A \mid \mathcal{F}_S); B] &= \mathbb{E}_{\mathbb{P}}[\mathbb{P}(X_S = X_0 \circ \theta_S \mid \mathcal{F}_S); B] \\ &= \mathbb{P}(X_S = X_S; B) = \mathbb{P}(B).\end{aligned}$$

Next, if $f \in C^2$ and is bounded with bounded first and second partial derivatives, we need to show that

$$M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_r)\, dr$$

is a martingale under $\mathbb{Q}_S$ for almost every ω. Let $u > t$. Since $M_t \circ \theta_S = M_{t+S} - M_S$ is a martingale with respect to $\mathcal{F}_{S+t}$, then

$$\mathbb{E}_{\mathbb{P}}[M_u \circ \theta_S; B \circ \theta_S \cap A] = \mathbb{E}_{\mathbb{P}}[M_t \circ \theta_S; B \circ \theta_S \cap A]$$

whenever $B \in \mathcal{F}_t$ and $A \in \mathcal{F}_S$. This is the same as saying

$$\mathbb{E}_{\mathbb{P}}[(M_u 1_B) \circ \theta_S; A] = \mathbb{E}_{\mathbb{P}}[(M_t 1_B) \circ \theta_S; A]. \tag{21.2}$$

Since (21.2) holds for all $A \in \mathcal{F}_S$, by the definition of $\mathbb{Q}_S$,

$$\mathbb{E}_{\mathbb{Q}_S}[M_u; B] = \mathbb{E}_{\mathbb{Q}_S}[M_t; B]$$

whenever $B \in \mathcal{F}_t$, which is what we needed to show.

Finally, if $f \in C^2$, then M_t is a local martingale under Q_S by the same argument as in the last paragraph of the proof of Theorem 20.1. □

Essentially the same proof shows that

Corollary 21.2. *Let $\mathbb{Q}'_S$ be a regular conditional probability for $\mathbb{P}_S[\cdot \mid X_S]$. Then with probability one, $\mathbb{Q}'_S$ is a solution to the martingale problem for $\mathcal{L}$ started at $X_S(\omega)$.*

If $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$, we can in fact show that there exists a family of solutions to the martingale problem that is a strong Markov family. We take $\Lambda_2 = 0$ for simplicity.

Theorem 21.3. *Let $\Lambda > 0$ and suppose $\mathcal{L}_n \in \mathcal{N}(\Lambda, 0)$ with the $a^n_{ij} \in C^2$ and converging almost everywhere to the a_{ij}. Suppose $(\mathbb{P}^x_n, X_t)$ is a strong Markov family of solutions to the martingale problem for $\mathcal{L}_n$. Then there exists a subsequence n_k and a strong Markov family of solutions $(\mathbb{P}^x, X_t)$ to the martingale problem for $\mathcal{L}$ such that $\mathbb{P}^x_{n_k}$ converges weakly to $\mathbb{P}^x$ for all x.*

Note that part of the assertion is that the subsequence n_k does not depend on x.

Proof. Let $\{g_i\}$ be a countable dense subset of $C(\mathbb{R}^d)$, the continuous bounded functions on $\mathbb{R}^d$, and let $\{\lambda_j\}$ be a countable dense subset of $(0, \infty)$. Let

$$G^\lambda_n g(x) = \mathbb{E}^{\,x}_{\,n} \int_0^\infty e^{-\lambda t} g(X_t)\, dt.$$

Note that $\|G^\lambda_n g\|_\infty \le \|g\|_\infty/\lambda$. By the equicontinuity of $G^\lambda_n g$ in n for each g and a diagonalization argument, we can find a subsequence n_k such that $G^{\lambda_j}_{n_k} g_i$ converges boundedly and uniformly on compacts. Since

$$\|G^{\lambda_j}_n g - G^{\lambda_j}_n h\|_\infty \le \frac{1}{\lambda_j} \|g - h\|_\infty$$

it follows that $G^{\lambda_j}_{n_k} g$ converges uniformly on compacts for all $g \in C(\mathbb{R})$. Since

$$\|G^\lambda_n g - G^\mu_n g\|_\infty \le \frac{c_1}{\lambda - \mu} \|g\|_\infty,$$

it follows that $G^\lambda_{n_k} g$ converges uniformly on compacts for all bounded $g \in C(\mathbb{R}^d)$ and all $\lambda \in (0, \infty)$. Call the limit $G^\lambda g$.

Suppose $x_n \to x$. By the tightness estimate, $\mathbb{P}^{x_{n_k}}_{n_k}$ is a tight sequence. Let $\mathbb{P}$ be any subsequential limit point. By Corollary 20.4, $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ started at x. If n' is a subsequence of n_k such that $\mathbb{P}^{x'_n}_{n'}$ converges weakly to $\mathbb{P}$, by the equicontinuity of $G^\lambda_n g$,

$$\begin{aligned} \mathbb{E} \int_0^\infty e^{-\lambda t} g(X_t)\, dt &= \lim_{n' \to \infty} \mathbb{E}^{\,x_{n'}}_{\,n'} \int_0^\infty e^{-\lambda t} g(X_t)\, dt \\ &= \lim_{n' \to \infty} G^\lambda_{n'} g(x_{n'}) = G^\lambda g(x). \end{aligned}$$

This holds for all bounded and continuous g; hence we see that if $\mathbb{P}_1$ and $\mathbb{P}_2$ are any two subsequential limit points of $\mathbb{P}^{x_{n_k}}_{n_k}$, their one-dimensional distributions agree by the uniqueness of the Laplace transform and the continuity of $g(X_t)$.

We next show that the two-dimensional distributions of any two subsequential limit points must agree. If g and h are bounded and continuous and $\mu > \lambda$,

$$\begin{aligned}
\mathbb{E} \int_0^\infty \int_0^\infty & e^{-\lambda t} e^{-\mu s} g(X_t) h(X_{t+s})\, ds\, dt \\
&= \lim_{n'\to\infty} \mathbb{E}\,^{x_{n'}}_{n'} \int_0^\infty \int_0^\infty e^{-\lambda t} e^{-\mu s} g(X_t) h(X_{t+s})\, ds\, dt \\
&= \lim_{n'\to\infty} \mathbb{E}\,^{x_{n'}}_{n'} \int_0^\infty e^{-\lambda t} g(X_t) \mathbb{E}\,^{X_t}_{n'} \int_0^\infty e^{-\mu s} h(X_s)\, ds\, dt \\
&= \lim_{n'\to\infty} \mathbb{E}\,^{x_{n'}}_{n_k} \int_0^\infty e^{-\lambda t} g(X_t) G^\mu_{n'} h(X_t)\, dt \\
&= \lim_{n'\to\infty} G^\lambda_{n'} \big(g(G^\mu_{n'} h)\big)(x_{n'}).
\end{aligned}$$

By the equicontinuity of the $G^\mu_n h$ and the fact that $G^\mu_{n_k} h$ converges boundedly and uniformly on compacts to $G^\mu h$, the right-hand side converges to $G^\lambda(g(G^\mu h))(x)$. By the uniqueness of the Laplace transform, we see that any two subsequential limit points have the same two-dimensional distributions.

Repeating the argument, we see that any two subsequential limit points have the same finite dimensional distributions. Since X_t is continuous, this implies that $\mathbb{P}_1 = \mathbb{P}_2$. We have thus shown that if $x_n \to x$, then $\mathbb{P}^{x_{n_k}}_{n_k}$ converges weakly to a probability measure; we call the limit $\mathbb{P}^x$. By the proof of Theorem 20.3, we know that $\mathbb{P}^x$ is a solution to the martingale problem for $\mathcal{L}$ started at x.

We now want to show that $(\mathbb{P}^x, X_t)$ forms a strong Markov family of solutions. We will do this by first showing that $\mathbb{E}\,^x_{n_k} f(X_t)$ converges uniformly on compacts to $\mathbb{E}\,^x f(X_t)$ if f is bounded and continuous. We have pointwise convergence of $\mathbb{E}\,^x_{n_k} f(X_t)$ for each x since we have weak convergence of $\mathbb{P}^x_{n_k}$ to $\mathbb{P}^x$.

We claim that the maps $x \mapsto \mathbb{E}\,^x_n f(X_t)$ are equicontinuous on compacts. If not, there exists $\varepsilon > 0$, $R > 0$, a subsequence n_m, and $x_m, y_m \in B(0,R)$ such that $|x_m - y_m| \to 0$ but

$$|\mathbb{E}\,^{x_m}_{n_m} f(X_t) - \mathbb{E}\,^{y_m}_{n_m} f(X_t)| > \varepsilon. \tag{21.3}$$

By compactness, there exists a further subsequence such that $\mathbb{P}^{x_{m_j}}_{n_{m_j}}$ converges weakly and also $x_{m_j} \to x \in \overline{B(0,R)}$; it follows that $y_{m_j} \to x$ also. By what we have already proved, $\mathbb{P}^{x_{m_j}}_{n_{m_j}}$ converges weakly to $\mathbb{P}^x$; hence $\mathbb{E}\,^{x_{m_j}}_{n_{m_j}} f(X_t)$ converges to $\mathbb{E}\,^x f(X_t)$ and the same with x_{m_j} replaced by y_{m_j}, a contradiction to (21.3). We thus have that the maps $x \mapsto \mathbb{E}\,^x_n f(X_t)$ are equicontinuous.

This implies that the convergence of $\mathbb{E}\,^x_{n_k} f(X_t)$ is uniform on compacts. In particular, the limit $\mathbb{E}\,^x f(X_t)$ is a continuous function of x. The map $x \mapsto \mathbb{E}\,^x f(X_t)$ being continuous when f is continuous implies that $(\mathbb{P}^x, X_t)$ is a strong Markov family of solutions. □

22. Some useful techniques.

In this section we want to provide a number of results that make proofs of uniqueness for the martingale problem easier. First, we show that if the diffusion coefficients are uniformly elliptic, then the drift coefficients do not matter. Second, we show that it is enough to look at λ-resolvents. Third, we prove that uniqueness of the martingale problem is a local property. Fourth, we see that it suffices to look at strong Markov solutions.

Let us show that for uniformly elliptic operators we may assume the drift coefficients are 0.

Theorem 22.1. *Suppose $\mathcal{L}'$ is defined by (20.8) and suppose there is uniqueness for the martingale problem for $\mathcal{L}'$ started at x. If $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$, then there is uniqueness for the martingale problem for $\mathcal{L}$ started at x.*

Proof. Let $\mathbb{P}_1, \mathbb{P}_2$ be two solutions to the martingale problem for $\mathcal{L}$ started at x. From the definition of martingale problem, $\langle X^i, X^j \rangle_t = \int_0^t a_{ij}(X_s)\,ds$. Define $\mathbb{Q}_i$ on $\mathcal{F}_t$, $i = 1, 2$, by

$$d\mathbb{Q}_i/d\mathbb{P}_i = \exp\Big(-\int_0^t (ba^{-1})(X_s)\,dX_s - \frac{1}{2}\int_0^t (ba^{-1}b^T)(X_s)\,ds\Big),$$

where b^T denotes the transpose of b. A simple calculation shows that the quadratic variation of $\int_0^t (ba^{-1})(X_s)\,dX_s$ is $\int_0^t (ba^{-1}b^T)(X_s)\,ds$, so $d\mathbb{Q}_i/d\mathbb{P}_i$ is of the right form for use in the Girsanov theorem. If $f \in C^2$ and

$$M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,ds, \tag{22.1}$$

then M_t is a local martingale under $\mathbb{P}_i$. By Itô's formula, the martingale part of M_t is the same as the martingale part of $\int_0^t \nabla f(X_s)\cdot dX_s$. We calculate

$$\begin{aligned}\Big\langle \int_0^\cdot ba^{-1}(X_s)\,dX_s, M\Big\rangle_t &= \int_0^t \sum_{i,j=1}^d (ba^{-1})_j(X_s)\partial_i f(X_s)\,d\langle X^i, X^j\rangle_s \\ &= \int_0^t \sum_{i=1}^d b_i(X_s)\partial_i f(X_s)\,ds.\end{aligned}$$

Hence by the Girsanov theorem, under $\mathbb{Q}_i$ the process

$$M_t - \Big(-\int_0^t b(X_s)\cdot\nabla f(X_s)\,ds\Big) = f(X_t) - f(X_0) - \int_0^t \mathcal{L}'f(X_s)\,ds$$

is a local martingale. Clearly $\mathbb{Q}_i(X_0 = x) = 1$, so $\mathbb{Q}_i$ is a solution to the martingale problem for $\mathcal{L}'$ started at x. By the uniqueness assumption, $\mathbb{Q}_1 = \mathbb{Q}_2$. So if $A \in \mathcal{F}_t$,

$$\mathbb{P}_i(A) = \int_A \exp\Big(\int_0^t (ba^{-1})(X_s)\,dX_s + \frac{1}{2}\int_0^t (ba^{-1}b^T)(X_s)\,ds\Big)\,d\mathbb{Q}_i,$$

which implies $\mathbb{P}_1(A) = \mathbb{P}_2(A)$. $\square$

To prove uniqueness it turns out that it is sufficient to look at quantities which are essentially λ-potentials (that is, λ-resolvents). It will be convenient to introduce the notation

$$\mathcal{M}(\mathcal{L}, x) = \{\mathbb{P} : \mathbb{P} \text{ is a solution to the martingale problem for } \mathcal{L} \text{ started at } x\}. \tag{22.2}$$

Theorem 22.2. *Suppose for all $x \in \mathbb{R}^d$, $\lambda > 0$, and $f \in C^2(\mathbb{R}^d)$,*

$$\mathbb{E}_1 \int_0^\infty e^{-\lambda t} f(X_t)\, dt = \mathbb{E}_2 \int_0^\infty e^{-\lambda t} f(X_t)\, dt$$

whenever $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}(\mathcal{L}, x)$. Then for each $x \in \mathbb{R}^d$ the martingale problem for $\mathcal{L}$ has a unique solution.

Proof. By the uniqueness of the Laplace transform and the continuity of f and X_t, our hypothesis implies that $\mathbb{E}_1 f(X_t) = \mathbb{E}_2 f(X_t)$ for all $t > 0$ and $f \in C^2$ if $x \in \mathbb{R}^d$ and $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}(\mathcal{L}, x)$. A limit argument shows that equality holds for all bounded f. In other words, the one-dimensional distributions of X_t under $\mathbb{P}_1$ and $\mathbb{P}_2$ are the same.

We next look at the two-dimensional distributions. Suppose f, g are bounded and $0 < s < t$. For $i = 1, 2$, let $\mathbb{P}_{i,s}(A) = \mathbb{P}_i(A \circ \theta_s)$, and let $\mathbb{Q}_i$ be a regular conditional probability for $\mathbb{E}_{i,s}(\cdot \mid X_s)$. By Corollary 21.2, $\mathbb{Q}_i$ is a solution to the martingale problem for $\mathcal{L}$ started at X_s. By the first paragraph of this proof,

$$\mathbb{E}_{\mathbb{Q}_1} g(X_{t-s}) = \mathbb{E}_{\mathbb{Q}_2} g(X_{t-s}), \qquad \text{a.s.}$$

Since $\mathbb{Q}_1(A)$ is measurable with respect to the σ-field generated by the single random variable X_s for each A, then $\mathbb{E}_{\mathbb{Q}_1} g(X_{t-s})$ is also measurable with respect to the σ-field generated by X_s. So $\mathbb{E}_{\mathbb{Q}_1} g(X_{t-s}) = \varphi(X_s)$ for some function φ. Then

$$\begin{aligned} \mathbb{E}_1 f(X_s) g(X_t) &= \mathbb{E}_1 [f(X_s) \mathbb{E}_1(g(X_t) \mid X_s)] \\ &= \mathbb{E}_1 f(X_s) \mathbb{E}_{\mathbb{Q}_1}(g(X_{t-s})) = \mathbb{E}_1 f(X_s) \varphi(X_s). \end{aligned}$$

By the uniqueness of the one-dimensional distributions, the right-hand side is equal to $\mathbb{E}_2 f(X_s)\varphi(X_s)$, which, similarly to the above, is equal to $\mathbb{E}_2 f(X_s) g(X_t)$. Hence the two-dimensional distributions of X_t under $\mathbb{P}_1$ and $\mathbb{P}_2$ are the same.

An induction argument shows that the finite dimensional distributions of X_t under $\mathbb{P}_1$ and $\mathbb{P}_2$ are the same. Since X_t has continuous paths, we deduce $\mathbb{P}_1 = \mathbb{P}_2$. $\square$

We now want to show that questions of uniqueness for martingale problems for elliptic operators are local questions. We start by giving a "piecing-together" lemma.

Lemma 22.3. *Suppose $\mathcal{L}_1, \mathcal{L}_2$ are two elliptic operators with bounded coefficients. Let $S = \inf\{t : |X_t - x| \geq r\}$ and let $\mathbb{P}_1, \mathbb{P}_2$ be solutions to the martingale problems for $\mathcal{L}_1, \mathcal{L}_2$, respectively, started at x. Let $\mathbb{Q}_2$ be a regular conditional probability for $\mathbb{E}_{\mathbb{P}_{2S}}[\cdot \mid \mathcal{F}_S]$, where $\mathbb{P}_{2S}(A) = \mathbb{P}_2(A \circ \theta_S)$. Define $\overline{\mathbb{P}}$ by*

$$\overline{\mathbb{P}}(B \circ \theta_S \cap A) = \mathbb{E}_{\mathbb{P}_1}[\mathbb{Q}_2(B); A], \qquad A \in \mathcal{F}_S, B \in \mathcal{F}_\infty.$$

If the coefficients of $\mathcal{L}_1$ and $\mathcal{L}_2$ agree on $B(x,r)$, then $\overline{\mathbb{P}}$ is a solution to the martingale problem for $\mathcal{L}$ started at x.

$\overline{\mathbb{P}}$ represents the process behaving according to $\mathbb{P}_1$ up to time S and according to $\mathbb{P}_2$ after time S.

Proof. It is clear that the restriction of $\overline{\mathbb{P}}$ to $\mathcal{F}_S$ is equal to the restriction of $\mathbb{P}_1$ to $\mathcal{F}_S$. Hence

$$\overline{\mathbb{P}}(X_0 = x) = \mathbb{P}_1(X_0 = x) = 1.$$

If $f \in C^2$,

$$\begin{aligned} M_t &= f(X_{t\wedge S}) - f(X_0) - \int_0^{t\wedge S} \mathcal{L}_1 f(X_s)\, ds \\ &= f(X_{t\wedge S}) - f(X_0) - \int_0^{t\wedge S} \mathcal{L}_2 f(X_s)\, ds \end{aligned}$$

is a martingale under $\mathbb{P}_1$. Since for each t these random variables are $\mathcal{F}_S$ measurable, M_t is also a martingale under $\overline{\mathbb{P}}$. It remains to show that $N_t = f(X_{S+t}) - f(X_S) - \int_S^{S+t} \mathcal{L}_2 f(X_s)\, ds$ is a martingale under $\overline{\mathbb{P}}$. This follows from Proposition 21.1 and the definition of $\overline{\mathbb{P}}$. □

Theorem 22.4. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$. Suppose for each $x \in \mathbb{R}^d$ there exist $r_x > 0$ and $\mathcal{K}(x) \in \mathcal{N}(\Lambda_1, \Lambda_2)$ such that the coefficients of $\mathcal{K}(x)$ agree with those of $\mathcal{L}$ in $B(x, r_x)$ and the solution to the martingale problem for $\mathcal{K}(x)$ is unique for every starting point. Then the martingale problem for $\mathcal{L}$ started at any point has a unique solution.*

Proof. Fix x_0 and suppose $\mathbb{P}_1$ and $\mathbb{P}_2$ are two solutions to the martingale problem for $\mathcal{L}$ started at x_0. Suppose x_1 is such that $x_0 \in B(x_1, r_{x_1}/4)$. Let $S = \inf\{t : |X_t - x_1| > r_{x_1}/2\}$. Write $\mathbb{P}^{\mathcal{K}}$ for the solution to the martingale problem for $\mathcal{K}(x_1)$ started at x_0. Let $\mathbb{Q}_S^{\mathcal{K}}$ be the regular conditional probability defined as in (21.1). For $i = 1, 2$, define

$$\overline{\mathbb{P}}_i(B \circ \theta_S \cap A) = \mathbb{E}_i[\mathbb{Q}_S^{\mathcal{K}}(B); A], \qquad i = 1, 2, \quad A \in \mathcal{F}_S, B \in \mathcal{F}_\infty. \tag{22.3}$$

Since the coefficients of $\mathcal{L}$ and $\mathcal{K}(x_1)$ agree on $B(x_1, r_{x_1})$, by Lemma 22.3 applied to $\mathbb{P}_i$ and $\mathbb{P}^{\mathcal{K}}$, $\overline{\mathbb{P}}_i$ is a solution to the martingale problem for $\mathcal{K}(x_1)$ started at x_0. By the uniqueness assumption, they must both be equal to $\mathbb{P}^{\mathcal{K}}$. Hence the restriction of $\mathbb{P}_1$ and $\mathbb{P}_2$ to $\mathcal{F}_S$ must be the same, namely, the same as the restriction of $\mathbb{P}^{\mathcal{K}}$ to $\mathcal{F}_S$. We have thus shown that any two solutions to the martingale problem for $\mathcal{L}$ started at a point x_0 agree on $\mathcal{F}_S$ if $x_0 \in B(x_1, r_{x_1}/4)$ and $S = \inf\{t : |X_t - x_1| > r_{x_i}/2\}$.

Let $N > 0$. $\overline{B(x_0, N)}$ is compact and hence there exist finitely many points $x_1, \ldots, x_m$ such that $\{B(x_i, r_{x_i}/4)\}$ is a cover for $\overline{B(x_0, N)}$. Let us define a measurable mapping $\psi : \overline{B(x_0, N)} \to \{1, \ldots, m\}$ by letting $\psi(x)$ be the smallest index for which $x \in B(x_{\psi(x)}, r_{\psi(x)}/4)$. Let $S_0 = 0$ and $S_{i+1} = \inf\{t > S_i : X_t \notin B(\psi(X_{S_i}), r_{\psi(X(S_i))}/2)\}$. The S_i are thus stopping times describing when X_t has moved far enough to exit its current ball.

We now show that any two solutions $\mathbb{P}_1$ and $\mathbb{P}_2$ for the martingale problem for $\mathcal{L}$ started at x_0 agree on $\mathcal{F}_{S_i \wedge \tau(B(x_0,N))}$ for each i. We already have done the case $i = 1$ in the first paragraph of this proof.

Let $\mathbb{Q}_{i,S_1}$ be a regular conditional probability defined as in (21.1). If $A \in \mathcal{F}_{S_1}$ and $B \in (\mathcal{F}_\infty \circ \theta_{S_1}) \cap \mathcal{F}_{S_2}$, then

$$\mathbb{P}_i(A \cap B) = \mathbb{E}_i[\mathbb{Q}_{i,S_1}(B); A], \qquad i = 1, 2.$$

By Proposition 21.1, $\mathbb{Q}_{i,S_1}$ is a solution to the martingale problem for $\mathcal{L}$ started at X_{S_1}, so by what we have shown in the first paragraph $\mathbb{Q}_{1,S_1} = \mathbb{Q}_{2,S_1}$ on $(\mathcal{F}_\infty \circ \theta_{S_1}) \cap \mathcal{F}_{S_2}$. Since $\mathbb{Q}_{i,S_1}(B)$ is $\mathcal{F}_{S_1}$ measurable and $\mathbb{P}_1 = \mathbb{P}_2$ on $\mathcal{F}_{S_1}$, this shows $\mathbb{P}_1(A \cap B) = \mathbb{P}_2(A \cap B)$. The random variable $\int_0^{S_2} e^{-\lambda r} f(X_r)\, dr$ can be written

$$\int_0^{S_1} e^{-\lambda r} f(X_r)\, dr + e^{-\lambda S_1} \Big(\int_0^{S_1} e^{-\lambda r} f(X_r)\, dr \circ \theta_{S_1} \Big).$$

Hence $\mathbb{E}_1 \int_0^{S_2} e^{-\lambda r} f(X_r)\, dr = \mathbb{E}_2 \int_0^{S_2} e^{-\lambda r} f(X_r)\, dr$ whenever f is bounded and continuous and $\lambda > 0$. As in Theorem 22.2, this implies $\mathbb{P}_1 = \mathbb{P}_2$ on $\mathcal{F}_{S_2}$.

Using an induction argument, $\mathbb{P}_1 = \mathbb{P}_2$ on $\mathcal{F}_{S_i \wedge \tau(B(x_0,N))}$ for each i. Note that

$$r = \min_{1 \le i \le m} r_{x_i} > 0.$$

Since $S_{i+1} - S_i$ is greater than the time for X_t to move more than $r/4$, $S_i \uparrow \tau_{B(0,N)}$ by the continuity of the paths of X_t. Therefore $\mathbb{P}_1 = \mathbb{P}_2$ on $\mathcal{F}_{\tau(B(x_0,N))}$. Since N is arbitrary, this shows that $\mathbb{P}_1 = \mathbb{P}_2$. □

It is often more convenient to work with strong Markov families. Recall the definition of $\mathcal{M}(\mathcal{L}, x)$ from (22.2).

Theorem 22.5. *Let $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$. Suppose there exists a strong Markov family $(\mathbb{P}_1^x, X_t)$ such that for each $x \in \mathbb{R}^d$, $\mathbb{P}_1^x$ is a solution to the martingale problem for $\mathcal{L}$ started at x. Suppose whenever $(\mathbb{P}_2^x, X_t)$ is another strong Markov family for which $\mathbb{P}_2^x \in \mathcal{M}(\mathcal{L}, x)$ for each x, we have $\mathbb{P}_1^x = \mathbb{P}_2^x$ for all x. Then for each x the solution to the martingale problem for $\mathcal{L}$ started at x is unique.*

In other words, if we have uniqueness within the class of strong Markov families, then we have uniqueness.

Proof. Let f be bounded and continuous, $\lambda > 0$, and $x \in \mathbb{R}^d$. Let $\mathbb{P}$ be any solution to the martingale problem for $\mathcal{L}$ started at x. There exists a sequence a_{ij}^n converging to a_{ij} almost everywhere as $n \to \infty$ such that the coefficients of the a_{ij}^n are C^2, $\mathcal{L}^n \in \mathcal{N}(\Lambda, 0)$, and if $\widetilde{\mathbb{P}}_n^x$ is a solution to the martingale problem for $\mathcal{L}^n$ started at x,

$$\widetilde{\mathbb{E}}_n^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt \to \mathbb{E} \int_0^\infty e^{-\lambda t} f(X_t)\, dt. \tag{22.4}$$

By Theorem 21.3, there exists a subsequence n_k such that $\widetilde{\mathbb{P}}_{n_k}^x$ converges weakly for all x, and if we call the limit $\widetilde{\mathbb{P}}^x$, then $(\widetilde{\mathbb{P}}^x, X_t)$ is a strong Markov family of solutions. By our hypothesis, $\widetilde{\mathbb{P}}^x = \mathbb{P}_1^x$. Using the weak convergence of $\widetilde{\mathbb{P}}_{n_k}^x$ to $\mathbb{P}_1^x$,

$$\widetilde{\mathbb{E}}_{n_k}^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt \to \mathbb{E}_1^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt.$$

Combining with (22.4), $\int_0^\infty e^{-\lambda t} f(X_t)\, dt$ has the same expectation under $\mathbb{P}$ and $\mathbb{P}_1^x$. Our result now follows by Theorem 22.2. □

23. Uniqueness.

We present a case for which uniqueness of the martingale problem is known. We assume $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$ for some $\Lambda_1 > 0$, and by virtue of Theorem 22.1, we may take $\Lambda_2 = 0$ without loss of generality.

Theorem 23.1. *Suppose $d \geq 3$. There exists ε_d (depending only on the dimension d) with the following property: if*

$$\sup_{i,j} \sup_x |a_{ij}(x) - \delta_{ij}| < \varepsilon_d,$$

then there exists a unique solution to the martingale problem for $\mathcal{L}$ started at any $x \in \mathbb{R}^d$.

Proof. Let $\mathbb{P}_1$, $\mathbb{P}_2$ be any two solutions to the martingale problem for $\mathcal{L}$ started at x. Define $G_i^\lambda f(x) = \mathbb{E}_i^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt$. If $f \in C^2$ is bounded with bounded first and second partial derivatives, then by Itô's formula,

$$f(X_t) = f(X_0) + \text{ martingale } + \int_0^t \mathcal{L} f(X_s)\, ds.$$

Multiplying by $e^{-\lambda t}$, taking the expectation with respect to $\mathbb{P}_i$, and integrating over t from 0 to ∞,

$$\begin{aligned}\mathbb{E}_i \int_0^\infty e^{-\lambda t} f(X_t)\,dt &= \frac{1}{\lambda} f(x) + \mathbb{E}_i \int_0^\infty e^{-\lambda t} \int_0^t \mathcal{L} f(X_s)\,ds\,dt \qquad (23.1)\\ &= \frac{1}{\lambda} f(x) + \mathbb{E}_i \int_0^\infty \mathcal{L} f(X_s) \int_s^\infty e^{-\lambda t}\,dt\,ds\\ &= \frac{1}{\lambda} f(x) + \frac{1}{\lambda}\mathbb{E}_i \int_0^\infty e^{-\lambda s} \mathcal{L} f(X_s)\,ds.\end{aligned}$$

Set

$$u^\lambda(z) = \int_0^\infty e^{-\lambda t}\big((2\pi t)^{-d/2} e^{-z^2/2t}\big)\,dt,$$

the λ-potential density of Brownian motion. Let $U^\lambda f(x) = \int f(y) u^\lambda(x-y)\,dy$, the λ-potential of f with respect to Brownian motion. Then set

$$\mathcal{B} = \frac{1}{2} \sum_{i,j=1}^d (a_{ij}(x) - \delta_{ij}) \partial_{ij} f(x). \qquad (23.2)$$

If $f = U^\lambda g$ for $g \in C^2$ with compact support, then

$$U^\lambda g = U^0(g - \lambda U^\lambda g).$$

$$\frac{\Delta U^\lambda g}{2} = \lambda U^\lambda g - g = \lambda f - g.$$

Since $\mathcal{L} f = (1/2)\Delta f + \mathcal{B} f$, we have from (23.1) that

$$\begin{aligned}G_i^\lambda f &= \lambda^{-1} f(x) + \lambda^{-1} G_i^\lambda \Big(\frac{\Delta U^\lambda g}{2} + \mathcal{B} f\Big)(x)\\ &= \lambda^{-1} f(x) + \lambda^{-1} G_i^\lambda (\lambda f - g) + \lambda^{-1} G_i^\lambda \mathcal{B} f,\end{aligned}$$

or

$$G_i^\lambda g = f(x) + G_i^\lambda \mathcal{B} f(x).$$

Hence

$$G_i^\lambda g = U^\lambda g(x) + G_i^\lambda \mathcal{B} U^\lambda g(x), \qquad i = 1, 2. \qquad (23.3)$$

(We remark that if we were to iterate (23.3), that is, substitute for G_i^λ on the right-hand side, we would be led to

$$G_i^\lambda g = U^\lambda g + U^\lambda \mathcal{B} U^\lambda g + \cdots,$$

which indicates that (23.3) is essentially variation of parameters in disguise.)

We return to the proof. Let

$$\rho = \sup_{\|g\|_d \le 1} |G_1^\lambda g - G_2^\lambda g|.$$

$\rho < \infty$. Taking the difference of (23.3) with $i = 1$ and $i = 2$, we have

$$G_i^\lambda g - G_2^\lambda g = (G_1^\lambda - G_2^\lambda)(\mathcal{B}U^\lambda g). \tag{23.4}$$

The right-hand side is bounded by $\rho \|\mathcal{B}U^\lambda g\|_d$. We need the result from analysis that

$$\|\partial_{ij} U^\lambda g\|_d \le c_1 \|g\|_d.$$

Then

$$\|\mathcal{B}U^\lambda g\|_d \le \varepsilon_d \sum_{i,j=1}^d \|\partial_{ij} U^\lambda g\|_d \le \varepsilon_d c_1 d^2 \|g\|_d \le (1/2)\|g\|_d$$

if we take $\varepsilon_d < 1/2c_1 d^2$. Hence

$$|G_1^\lambda g - G_2^\lambda g| \le (\rho/2)\|g\|_d.$$

If we now take the supremum of the left-hand side over $g \in C^2$ with $\|g\|_d \le 1$, we obtain $\rho \le \rho/2$. Since we observed that $\rho < \infty$, this means that $\rho = 0$, or $G_1^\lambda g = G_2^\lambda g$ if $g \in L^d$. In particular, this holds if g is continuous with compact support. By a limit argument, this holds for all continuous bounded g. This is true for every starting point $x \in \mathbb{R}^d$, so by Theorem 22.2, $\mathbb{P}_1 = \mathbb{P}_2$. □

Corollary 23.2. *Let C be a positive definite matrix. There exists ε_d such that if*

$$\sup_{i,j} \sup_x |a_{ij}(x) - C_{ij}| < \varepsilon_d,$$

then there exists a unique solution to the martingale problem for $\mathcal{L}$ started at any $x \in \mathbb{R}^d$.

Proof. Let $\sigma(x)$ be a positive definite square root of $a(x)$ and $C^{1/2}$ a positive definite square root of C. By Theorem 20.1, to establish uniqueness it suffices to establish weak uniqueness of the stochastic differential equation $dX_t = \sigma(X_t)\, dW_t$. If X_t is a solution to this stochastic differential equation, it is easy to see that $Y_t = C^{-1/2} X_t$ is a solution to $dY_t = (\sigma C^{-1/2})(Y_t)\, dW_t$ and conversely. By Theorem 20.1 again, weak uniqueness for the latter stochastic differential equation will follow if we have weak uniqueness for the martingale problem for $\mathcal{L}^C$, where the coefficients of $\mathcal{L}^C$ are $C^{-1} a_{ij}$. The assumption $|a_{ij}(x) - C_{ij}| < \varepsilon_d$ implies $|C^{-1} a_{ij}(x) - \delta_{ij}| < c_1 \varepsilon_d$, where c_1 depends on C. The result follows by Theorem 23.1 by taking ε_d sufficiently small. □

We now can prove the important result due to Stroock and Varadhan.

Theorem 23.3. *If $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$ and the a_{ij} are continuous, then the martingale problem for $\mathcal{L}$ started at x has a unique solution.*

Proof. By Theorem 22.1, we may suppose that $\Lambda_2 = 0$. If $x \in \mathbb{R}^d$, let $C = a(x)$ and then choose r_x such that if $y \in B(x, 2r_x)$, then $|a_{ij}(y) - a_{ij}(x)| < \varepsilon_d$ for $i, j = 1, \ldots, d$, where ε_d is given by Corollary 23.2. Let $a^x_{ij}(y)$ be continuous functions that agree with $a_{ij}(y)$ on $B(x, r_x)$ and such that if

$$\mathcal{K}^x f(z) = \sum_{i,j=1}^d a^x_{ij}(z)\partial_{ij} f(z),$$

then $\mathcal{K}^x \in \mathcal{N}(\Lambda_1, 0)$, and

$$\sup_{i,j} \sup_{y} |a^x_{ij}(y) - a_{ij}(x)| < \varepsilon_d.$$

By Corollary 23.2, we have uniqueness of the martingale problem for $\mathcal{K}^x$ starting at any point in $\mathbb{R}^d$. Moreover, the coefficients of $\mathcal{K}^x$ agree with those of $\mathcal{L}$ inside $B(x, r_x)$. The conclusion now follows by Theorem 22.4. □

24. Consequences of uniqueness.

We mention some conclusions that one can draw when uniqueness holds.

Theorem 24.1. *Suppose there exists a unique solution $\mathbb{P}^x$ to the martingale problem for $\mathcal{L}$ started at x for each $x \in \mathbb{R}^d$. Then $(\mathbb{P}^x, X_t)$ forms a strong Markov family.*

Uniqueness implies some convergence results.

Theorem 24.2. *Suppose $\mathcal{L}_n \in \mathcal{N}(\Lambda_1, \Lambda_2)$ and the diffusion coefficients a^n_{ij} converge to a_{ij} almost everywhere, and similarly for the drift coefficients b^n_i. Suppose $x_n \to x$, $\mathbb{P}$ is the unique solution to the martingale problem for $\mathcal{L}$ started at x, and for each n, $\mathbb{P}_n$ is a solution to the martingale problem for $\mathcal{L}_n$ started at x_n. Then $\mathbb{P}_n$ converges weakly to $\mathbb{P}$.*

Proof. The probability measures $\mathbb{P}_n$ are tight. Any subsequential limit point is a solution to the martingale problem for $\mathcal{L}$ started at x. By the uniqueness hypothesis, any subsequential limit point must be equal to $\mathbb{P}$; this implies that the whole sequence converges to $\mathbb{P}$. □

25. Divergence form operators.

Elliptic operators in *divergence form* are operators $\mathcal{L}$ defined on C^2 functions by

$$\mathcal{L}f(x) = \frac{1}{2} \sum_{i,j=1}^d \partial_i(a_{ij}\partial_j f)(x), \tag{25.1}$$

where the a_{ij} are measurable functions of x and $a_{ij}(x) = a_{ji}(x)$ for all pairs i, j and all x. Let $\mathcal{D}(\Lambda)$ be the set of operators in divergence form such that for all x and all $y = (y_1, \ldots, y_d)$,

$$\Lambda|y|^2 \le \sum_{i,j=1}^d a_{ij}(x) y_i y_j \le \Lambda^{-1}|y|^2. \tag{25.2}$$

Throughout this chapter we assume the operator $\mathcal{L}$ is *uniformly elliptic*, that is, $\mathcal{L} \in \mathcal{D}(\Lambda)$ for some $\Lambda > 0$.

If the a_{ij} are not differentiable, an interpretation has to be given to $\mathcal{L}f$; see (25.6). For most of this chapter we will assume the a_{ij} are smooth. With this assumption,

$$\mathcal{L}f(x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \partial_{ij} f(x) + \frac{1}{2} \sum_{j=1}^d \Big(\sum_{i=1}^d \partial_i a_{ij}(x) \Big) \partial_j f(x), \tag{25.3}$$

and so $\mathcal{L}$ is equivalent to an operator in nondivergence form with

$$b_j(x) = (1/2) \sum_{i=1}^d \partial_i a_{ij}(x).$$

However, all of our estimates for $\mathcal{L} \in \mathcal{D}(\Lambda)$ will depend only on Λ and not on any smoothness of the a_{ij}. So by a limit procedure, our results and estimates will be valid for operators $\mathcal{L}$ where the a_{ij} are only bounded and strictly elliptic.

We refer to the conclusion of the following proposition as *scaling*.

Proposition 25.1. *Let $\mathcal{L} \in \mathcal{D}(\Lambda)$ and let $(\mathbb{P}^x, X_t)$ be the associated process (in the sense of Section I.2). If $r > 0$, $a^r_{ij}(x) = a_{ij}(x/r)$, and $\mathcal{L}^r f(x) = \sum_{i,j=1}^d \partial_i (a^r_{ij} \partial_j f)(x)$, then $\mathcal{L}^r \in \mathcal{D}(\Lambda)$ and $(\mathbb{P}^{x/r}, rX_{t/r^2})$ is the process associated to $\mathcal{L}^r$.*

Proof. Using (25.3), this is proved entirely analogously to the nondivergence case. □

An important example of operators in divergence form is given by the Laplace-Beltrami operators on Riemannian manifolds. Such an operator is the infinitesimal generator of a Brownian motion on the manifold. After a time change, the Laplace-Beltrami operator in local coordinates is an operator in divergence form, where the a_{ij} matrix is the inverse of the matrix g_{ij} that determines the Riemannian metric.

Recall the divergence theorem. Suppose D is a nice region, F is a smooth vector field, $\nu(x)$ is the outward pointing normal vector at $x \in \partial D$, and σ is surface measure on ∂D. The divergence theorem then says that

$$\int_{\partial D} F \cdot \nu(y)\, \sigma(dy) = \int_D \operatorname{div} F(x)\, dx. \tag{25.4}$$

Proposition 25.2. *Let g be a C^∞ function with compact support and f a bounded C^∞ function. Then*

$$\int_{\mathbb{R}^d} g(x)\mathcal{L}f(x)\,dx = -\frac{1}{2}\int_{\mathbb{R}^d}\Big(\sum_{i,j=1}^d \partial_i g(x) a_{ij}(x)\partial_j f(x)\Big)\,dx.$$

The integrand on the right could be written $\nabla g \cdot a\nabla f$.

Proof. We apply the divergence theorem. Let D be a ball large enough to contain the support of g and let $F(x)$ be the vector field whose ith component is

$$\frac{g(x)}{2}\sum_{j=1}^d a_{ij}(x)\partial_j f(x).$$

Since g is 0 on ∂D, then $F\cdot\nu = 0$ on ∂D, and also,

$$\begin{aligned}\operatorname{div} F(x) &= \frac{1}{2}\sum_{i=1}^d \partial_i\Big(g(x)\sum_{j=1}^d a_{ij}(x)\partial_j f(x)\Big)\\ &= \frac{1}{2}\sum_{i,j=1}^d \partial_i g(x) a_{ij}(x)\partial_j f(x) + g(x)\mathcal{L}f(x).\end{aligned}$$

We now substitute into (25.4). □

Applying Proposition 25.2 twice, if f and g are smooth with compact support,

$$\int g(x)\mathcal{L}f(x)\,dx = \int f(x)\mathcal{L}g(x)\,dx. \tag{25.5}$$

This equation says that $\mathcal{L}$ is *self-adjoint* with respect to Lebesgue measure.

Note that Proposition 25.2 allows us to give an interpretation to $\mathcal{L}f = 0$ even when the a_{ij} are not differentiable. We say f is a solution to $\mathcal{L}f = 0$ if f is differentiable in some sense, e.g., $f \in W^{1,p}$ for some p, and

$$\int \sum_{i,j=1}^d \partial_i f(x) a_{ij}(x)\partial_j g(x)\,dx = 0 \tag{25.6}$$

whenever g is in C^∞ with compact support. Here $W^{1,p}$ is the closure of $C^2\cap L^\infty$ with respect to the norm

$$\|f\|_{W^{1,p}} = \|f\|_p + \sum_{i=1}^d \|\partial_i f\|_p.$$

The expression

$$\int \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\partial_i f(x)\partial_j g(x)\,dx = \frac{1}{2}\int \nabla f(x)\cdot a(x)\nabla g(x)\,dx$$

is an example of what is known as a *Dirichlet form.* If we denote it by $\mathcal{E}(f,g)$, then Proposition 25.2 says that

$$\int g\mathcal{L}f\,dx = -\mathcal{E}(f,g)$$

for g with compact support. In the case of Brownian motion, the Dirichlet form is

$$\mathcal{E}_{BM}(f,g) = \frac{1}{2}\int \nabla f(x)\cdot\nabla g(x)\,dx.$$

Part of defining a Dirichlet form is specifying the domain. For example, the Dirichlet form for Brownian motion in $\mathbb{R}^d$ has domain $\{f\in L^2 : \mathcal{E}_{BM}(f,f)<\infty\}$. The Dirichlet form for reflecting Brownian motion in a domain $D\subseteq\mathbb{R}^d$ operates on $\{f\in L^2(D) : \int_D |\nabla f(x)|^2\,dx<\infty\}$, whereas the Dirichlet form for Brownian motion killed on exiting a set D has domain $\{f\in L^2(D) : \int_D |\nabla f(x)|^2\,dx<\infty, f=0 \text{ on } \partial D\}$.

Note that the uniform ellipticity of $\mathcal{L}$ implies that

$$\Lambda\mathcal{E}_{BM}(f,f)\le \mathcal{E}_{\mathcal{L}}(f,f)\le \Lambda^{-1}\mathcal{E}_{BM}(f,f). \tag{25.7}$$

26. Inequalities.

We will make use of several classical inequalities. The first is the Sobolev inequality.

Theorem 26.1. *Suppose $d>2$. There exists c_1 such that if $f\in C^2$ and $\nabla f\in L^2$, then*

$$\Big(\int_{\mathbb{R}^d} |f(x)|^{2d/(d-2)}\,dx\Big)^{(d-2)/2d} \le c_1\Big(\int_{\mathbb{R}^d} |\nabla f(x)|^2\,dx\Big)^{1/2}.$$

A variant of the Sobolev inequality is the following for bounded domains.

Corollary 26.2. *Suppose $d>2$. Let Q be the unit cube. Suppose f is C^2 on Q and $\nabla f\in L^2(Q)$. There exists c_1 such that*

$$\Big(\int_Q |f|^{2d/(d-2)}\Big)^{(d-2)/d} \le c_1\Big[\int_Q |\nabla f|^2 + \int_Q |f|^2\Big].$$

Proof. Let Q^* be the cube with the same center as Q but side length twice as long. By reflecting over the boundaries of Q, we can extend f to Q^* so that $\int_{Q^*} |f|^p \le c_2 \int_Q |f|^p$ for $p = 2d/(d-2)$ and also $\int_{Q^*} |f|^2 \le c_2 \int_Q |f|^2$ and $\int_{Q^*} |\nabla f|^2 \le c_2 \int_Q |\nabla f|^2$, where c_2 is a constant not depending on f. Let φ be a C^∞ function taking values in $[0,1]$ with support in Q^* and so that $\varphi = 1$ on Q. Applying Theorem 26.1 to φf,

$$\Big(\int_Q |f|^p\Big)^{2/p} \le \Big(\int |\varphi f|^p\Big)^{2/p} \le c_1 \int |\nabla(\varphi f)|^2,$$

where $p = 2d/(d-2)$. Since

$$|\nabla(\varphi f)|^2 \le 2|\nabla \varphi|^2 |f|^2 + 2|\varphi|^2 |\nabla f|^2,$$

and φ and $\nabla\varphi$ are bounded by constants independent of f and have support in Q^*, the result follows. $\square$

Another closely related inequality is the Nash inequality.

Theorem 26.3. *Suppose $d \ge 2$. There exists c_1 such that if $f \in C^2$, $f \in L^1 \cap L^2$, and $\nabla f \in L^2$, then*

$$\Big(\int |f|^2\Big)^{1+2/d} \le c_1 \Big(\int |\nabla f|^2\Big)\Big(\int |f|\Big)^{4/d}.$$

Proof. If $\widehat{f}(\xi) = \int e^{ix\cdot\xi} f(x)\,dx$ is the Fourier transform of f, then the Fourier transform of $\partial_j f$ is $i\xi_j \widehat{f}(\xi)$. Recall $|\widehat{f}(\xi)| \le \int |f|$. By the Plancherel theorem, $\int |f|^2 = c_2 \int |\widehat{f}(\xi)|^2\,d\xi$ and $\int |\nabla f|^2 = c_2 \int |\xi|^2 |\widehat{f}(\xi)|^2\,d\xi$. We have

$$\begin{aligned}\int |f|^2 = c_2 \int |\widehat{f}(\xi)|^2\,d\xi &\le c_2 \int_{|\xi|\le R} |\widehat{f}|^2 + c_2 \int_{|\xi|>R} \frac{|\xi|^2}{R^2} |\widehat{f}|^2\\ &\le c_3 R^d \Big(\int |f|\Big)^2 + c_4 R^{-2} \int |\nabla f|^2.\end{aligned}$$

We now choose R to minimize the right-hand side. $\square$

The Poincaré inequality states the following.

Theorem 26.4. *Suppose Q is a unit cube of side length h and f is C^2 on Q with $\nabla f \in L^2(Q)$. There exists c_1 not depending on f such that*

$$\int_Q |f(x) - f_Q|^2\,dx \le c_1 h^2 \int_Q |\nabla f(x)|^2\,dx,$$

where $f_Q = |Q|^{-1} \int_Q f(x)\,dx$.

Proof. By a translation of the coordinate axes, we may suppose Q is centered at the origin. Since $\nabla(f - f_Q) = \nabla f$, by subtracting a constant from f we may suppose without loss of generality that $f_Q = 0$. Let us also suppose for now that $h = 1$.

If $m = (m_1, \ldots, m_d)$, let C_m denote the Fourier coefficient of $e^{2\pi i m\cdot x}$, that is,

$$C_m = \int_Q e^{-2\pi i m\cdot x} f(x)\,dx.$$

Since $\int_Q f = 0$, then $C_0 = 0$. The mth Fourier coefficient of $\partial_j f$ is $2\pi i m_j C_m$. By the Parseval identity and the fact that $C_0 = 0$,

$$\begin{aligned}\int_Q |\nabla f|^2 &= \sum_m (2\pi)^2 |m|^2 |C_m|^2 \\ &\geq c_2 \sum_m |C_m|^2 = c_2 \int_Q |f|^2.\end{aligned} \tag{26.1}$$

We eliminate the supposition that $h = 1$ by a scaling argument, namely, we apply (26.1) to $f(x) = g(xh)$ for x in the unit cube, and then replace g by f. □

Finally, we will need the John-Nirenberg inequality. We continue to use the notation

$$f_Q = |Q|^{-1} \int_Q f. \tag{26.2}$$

Theorem 26.5. *Suppose* Q_0 *is a cube,* $f \in L^1(Q_0)$*, and for all cubes* $Q \subseteq Q_0$*,*

$$\frac{1}{|Q|} \int_Q |f(x) - f_Q| \leq 1. \tag{26.3}$$

Then there exist c_1 *and* c_2 *independent of* f *such that*

$$\int_{Q_0} e^{c_1[f(x) - f_{Q_0}]}\,dx \leq c_2.$$

An f satisfying (26.3) is said to be in BMO, the space of functions of bounded mean oscillation.

27. Moser's Harnack inequality.

Let $Q(h)$ denote the cube centered at the origin with side length h. Moser's Harnack inequality (Theorem 27.5) says that if $\mathcal{L} \in \mathcal{D}(\Lambda)$, there exists c_1 depending only on Λ such that if $\mathcal{L}u = 0$ and $u \geq 0$ in $Q(4)$, then

$$\sup_{Q(1)} u \leq c_1 \inf_{Q(1)} u.$$

We begin proving this important fact by establishing a sort of converse to Poincaré's inequality for powers of u. Recall that u is $\mathcal{L}$-harmonic in $Q(r)$ if u is C^2 on $Q(r)$ and $\mathcal{L}u = 0$ on $Q(r)$.

Proposition 27.1. *Suppose $r > 1$ and u is nonnegative and $\mathcal{L}$-harmonic in $Q(r)$. There exists c_1 depending only on the ellipticity bound Λ such that if $v = u^p$ for $p \in \mathbb{R}$, then*

$$\int_{Q(1)} |\nabla v|^2 \le c_1 \Big(\frac{2p}{2p-1}\Big)^2 \frac{1}{(r-1)^2} \int_{Q(r)} |v|^2.$$

Proof. The result is trivial if $p = 1/2$. The result is also trivial if $p = 0$, for then v is identically 1 and $\nabla v = 0$. So we suppose p is some value other than 0 or 1/2. Let φ be a smooth function taking values in $[0, 1]$ with support in $Q(r)$ such that $\varphi = 1$ on $Q(1)$ and $|\nabla\varphi| \le c_2/(r-1)$. Let $w = u^{2p-1}\varphi^2$. Since u is $\mathcal{L}$-harmonic and $w = 0$ outside of $Q(r)$, Proposition 25.2 tells us that

$$\begin{aligned} 0 = 2\int_{Q(r)} w\mathcal{L}u &= -\int_{Q(r)} \nabla w \cdot a\nabla u \\ &= -(2p-1)\int_{Q(r)} u^{2p-2}\varphi^2 \nabla u \cdot a\nabla u - 2\int_{Q(r)} u^{2p-1}\varphi\nabla\varphi \cdot a\nabla u. \end{aligned}$$

We then have, using (25.2) and the Cauchy-Schwarz inequality,

$$\begin{aligned} \int_{Q(r)} |\nabla v|^2\varphi^2 &= \int_{Q(r)} p^2 u^{2p-2}|\nabla u|^2\varphi^2 \\ &\le \Lambda p^2 \int_{Q(r)} u^{2p-2}\varphi^2\nabla u \cdot a\nabla u \\ &= c_2 \frac{2p^2}{|2p-1|} \int_{Q(r)} u^{2p-1}\varphi\nabla\varphi \cdot a\nabla u \\ &= \frac{2c_2p^2}{|2p-1|} \int_{Q(r)} u^p\varphi\nabla\varphi \cdot a u^{p-1}\nabla u \\ &= \frac{2c_2p}{|2p-1|} \int_{Q(r)} v\nabla\varphi \cdot a\varphi\nabla v \\ &\le \frac{2c_3|p|}{|2p-1|} \Big(\int_{Q(r)} |\nabla v|^2\varphi^2\Big)^{1/2} \Big(\int_{Q(r)} v^2|\nabla\varphi|^2\Big)^{1/2}. \end{aligned}$$

Dividing both sides by $(\int_{Q(r)} |\nabla v|^2\varphi^2)^{1/2}$, we obtain

$$\begin{aligned} \int_{Q(1)} |\nabla v|^2 &\le \int_{Q(r)} |\nabla v|^2\varphi^2 \\ &\le c_3^2\Big(\frac{2p}{2p-1}\Big)^2 \int_{Q(r)} v^2|\nabla\varphi|^2 \\ &\le c_3^2\Big(\frac{2p}{2p-1}\Big)^2 \frac{1}{(r-1)^2}\int_{Q(r)} v^2. \end{aligned}$$

□

Let us define

$$\Phi(p,h) = \Big(\int_{Q(h)} u^p \Big)^{1/p}.$$

Proposition 27.2. *Suppose $d \geq 3$. If $u \geq 0$ in $Q(2)$ and $\mathcal{L}u = 0$ in $Q(2)$, then for all $q_0 > 0$ there exists c_1 (depending on q_0 but not u) such that*

$$\sup_{Q(1)} u \leq c_1 \Phi(q_0, 2).$$

Proof. Let $R = d/(d-2)$, $p > 0$, and $2 > r > 1$. By Corollary 26.2 and Proposition 27.1,

$$\begin{aligned}
\Big(\int_{Q(1)} u^{2pR} \Big)^{1/R} &\leq c_2 \Big[\int_{Q(1)} |\nabla(u^p)|^2 + \int_{Q(1)} |u^p|^2 \Big] \\
&\leq c_3 \Big[\frac{1}{(r-1)^2} \Big(\frac{2p}{2p-1} \Big)^2 \int_{Q(r)} u^{2p} + \int_{Q(1)} |u^p|^2 \Big] \\
&\leq \frac{c_4}{(r-1)^2} \Big(\frac{2p}{2p-1} \Big)^2 \int_{Q(r)} u^{2p}.
\end{aligned}$$

Taking both sides to the $1/2p$ power and using scaling, if $r < s < 2r$,

$$\Phi(2Rp, r) \leq \Big(\frac{c_4}{(s/r-1)^2} \frac{(2p)^2}{(2p-1)^2} \Big)^{1/2p} \Phi(2p, s). \tag{27.1}$$

Suppose $p_0 = R^{-m-1/2}/2$, where m is the smallest positive integer such that $2p_0 < q_0$. Let $p_n = R^n p_0$, $r_n = 1 + 2^{-n}$. Then

$$\frac{r_n}{r_{n-1}} - 1 = \frac{2^{-n-1}}{1 + 2^{-(n-1)}} \geq 2^{-n}/2$$

and by our assumption on p_0,

$$\Big(\frac{2p_n}{2p_n - 1} \Big)^2 \leq c_5,$$

where c_5 depends only on R. Substituting in (27.1),

$$\Phi(2p_{n+1}, r_{n+1}) \leq (c_6 2^{2n})^{1/(2R^n p_0)} \Phi(2p_n, r_n).$$

By induction,

$$\Phi(2p_n, r_n) \leq c_6^\alpha 2^\beta \Phi(2p_0, 2),$$

where

$$\alpha = \sum_{j=0}^\infty \frac{1}{2R^j p_0} < \infty, \qquad \beta = \sum_{j=0}^\infty \frac{2j}{2R^j p_0} < \infty.$$

Therefore $\Phi(2p_n, r_n) \le c_7 \Phi(2p_0, 2)$. By Hölder's inequality,

$$\Phi(2p_0, 2) \le c_8 \Phi(q_0, 2).$$

The conclusion now follows from the fact that

$$\sup_{Q(1)} u \le \limsup_{n\to\infty} \Phi(2p_n, r_n). \qquad \square$$

Proposition 27.3. *Suppose u is bounded below by a positive constant on $Q(2)$ and $q_0 > 0$. Then there exists c_1 (depending only on q_0 but not u) such that*

$$\inf_{Q(1)} u \ge \Big(\int_{Q(2)} u^{-q_0} \Big)^{-1/q_0}.$$

Proof. The proof is almost identical to the above, working with

$$\Phi(-p, h) = \Big(\int_{Q(h)} u^{-p} \Big)^{-1/p} \tag{27.2}$$

instead of $\Phi(p, h)$. $\square$

To connect $\Phi(p, h)$ for $p > 0$ and $p < 0$, we look at $\log u$.

Proposition 27.4. *Suppose u is positive and $\mathcal{L}$-harmonic in $Q(4)$. There exists c_1 independent of u such that if $w = \log u$, then*

$$\int_Q |\nabla w|^2 \le c_1 h^{d-2}$$

for all cubes Q of side length h contained in $Q(2)$.

Proof. Let Q^* be the cube with the same center as Q but side length twice as long. Note $Q^* \subseteq Q(4)$. Let φ be C^∞ with values in $[0, 1]$, equal to 1 on Q, supported in Q^*, and such that $\|\nabla\varphi\|_\infty \le c_2/h$. Since $\nabla w = \nabla u/u$ and u is $\mathcal{L}$-harmonic in $Q(4)$,

$$\begin{aligned} 0 = 2\int \frac{\varphi^2}{u} \mathcal{L}u &= -\int \nabla(\varphi^2/u) \cdot a\nabla u \\ &= -\int \frac{2\varphi\nabla\varphi}{u} \cdot a\nabla u + \int \frac{\varphi^2}{u^2} \nabla u \cdot a\nabla u \\ &= -2\int \varphi\nabla\varphi \cdot a\nabla w + \int \varphi^2 \nabla w \cdot a\nabla w. \end{aligned}$$

So by the Cauchy-Schwarz inequality and (25.2),

$$\int_{Q^*} \varphi^2 |\nabla w|^2 \le c_3 \int_{Q^*} \varphi^2 \nabla w \cdot a \nabla w = c_4 \int_{Q^*} \nabla\varphi \cdot a\varphi\nabla w$$
$$\le c_5 \Big(\int_{Q^*} |\nabla\varphi|^2 \Big)^{1/2} \Big(\int_{Q^*} \varphi^2 |\nabla w|^2 \Big)^{1/2}.$$

Dividing by the second factor on the right, squaring, and using the bound on $|\nabla\varphi|$,

$$\int_Q |\nabla w|^2 \le \int_{Q^*} \varphi^2 |\nabla w|^2 \le c_5^2 |Q^*| (c_2/h)^2,$$

which implies our result. □

Putting all the pieces together, we have Moser's Harnack inequality.

Theorem 27.5. *There exists c_1 such that if u is $\mathcal{L}$-harmonic and nonnegative in $Q(4)$, then*

$$\sup_{Q(1)} u \le c_1 \inf_{Q(1)} u.$$

Proof. By looking at $u + \varepsilon$ and letting $\varepsilon \to 0$, we may suppose u is bounded below in $Q(4)$. Set $w = \log u$. Multiplying u by a constant, we may suppose that $\int_{Q(2)} w = 0$. By Proposition 27.4 and Theorem 26.4, there exists c_3 such that if Q is a cube contained in $Q(2)$, then

$$\Big(\frac{1}{|Q|} \int_Q |w - w_Q| \Big)^2 \le \frac{1}{|Q|} \int_Q |w - w_Q|^2 \le c_2 \frac{h^2}{|Q|} \int_Q |\nabla w|^2 \le c_3.$$

By the John-Nirenberg inequality applied to $w/c_3^{1/2}$ and $-w/c_3^{1/2}$, there exist c_4 and q_0 such that

$$\int_{Q(2)} e^{q_0 w} \le c_4, \qquad \int_{Q(2)} e^{-q_0 w} \le c_4.$$

This can be rewritten as

$$\int_{Q(2)} u^{q_0} \int_{Q(2)} u^{-q_0} \le c_4^2,$$

or

$$\Phi(q_0, 2) \le c_4^{2/q_0} \Phi(-q_0, 2). \tag{27.3}$$

This and Propositions 27.2 and 27.3 show

$$\sup_{Q(1)} u \le c_5 \Phi(q_0, 2) \le c_6 \Phi(-q_0, 2) \le c_7 \inf_{Q(1)} u.$$ □

An easy corollary proved by repeated use of Theorem 27.5 to a suitable overlapping sequence of cubes is the following.

Corollary 27.6. *Suppose $D_1 \subseteq \overline{D_1} \subseteq D_2$, where D_1 and D_2 are bounded connected domains in $\mathbb{R}^d$ and $d \geq 3$. There exists c_1 depending only on D_1 and D_2 such that if u is nonnegative and $\mathcal{L}$-harmonic in D_2, then*

$$\sup_{D_1} u \leq c_1 \inf_{D_2} u.$$

Another corollary of the Moser Harnack inequality is that $\mathcal{L}$-harmonic functions must be Hölder continuous with a modulus of continuity independent of the smoothness of the a_{ij}.

Theorem 27.7. *Suppose $d \geq 3$ and suppose u is $\mathcal{L}$-harmonic in $Q(2)$. There exist c_1 and α not depending on u such that if $x, y \in Q(1)$,*

$$|u(x) - u(y)| \leq c_1 |x - y|^\alpha \sup_{Q(2)} |u|.$$

Proof. Fix x and let $r < 1$. Our result will follow if we show there exists $\rho < 1$ independent of r such that

$$\operatorname*{Osc}_{B(x,r/2)} u \leq \rho \operatorname*{Osc}_{B(x,r)} u. \tag{27.4}$$

By looking at $Cu+D$ for suitable C and D, we may suppose that the infimum of $Cu+D$ on $B(x,r)$ is 0, the supremum is 1, and there exists $x_0 \in B(x, r/2)$ such that $(Cu+D)(x_0) \geq 1/2$. By Corollary 27.6 with $D_1 = B(x, r/2)$ and $D_2 = B(x, r)$, there exists c_2 such that

$$(Cu + D)(y) \geq c_2 (Cu + D)(x_0) \geq c_2/2, \qquad y \in B(x, r/2).$$

On the other hand, if $(\mathbb{P}^x, X_t)$ is the process associated with $\mathcal{L}$, then

$$(Cu + D)(y) = \mathbb{E}^{\,y} (Cu + D)(X_{\tau(B(x,r))}) \leq 1$$

by optional stopping. Hence $\operatorname{Osc}_{B(x,r/2)} (Cu + D) \leq 1 - c_2/2$, and (27.4) follows. □

28. Upper bounds on heat kernels.

We are now going to investigate bounds on the transition densities of X_t, where $(\mathbb{P}^x, X_t)$ is the process associated to an operator $\mathcal{L} \in \mathcal{D}(\Lambda)$. Let P_t be the operator defined by

$$P_t f(x) = \mathbb{E}^{\,x} f(X_t).$$

We shall see that there exists a symmetric function $p(t,x,y)$ such that

$$P_t f(x) = \int f(y) p(t,x,y)\, dy$$

and that $p(t,x,y)$ has upper and lower bounds similar to those of Brownian motion. Recall that $\partial_t u$ means $\partial u/\partial t$. Since $u(x,t) = \mathbb{E}^x f(X_t)$ is also a solution to the Cauchy problem $\partial_t u = \mathcal{L}u$ in $\mathbb{R}^d \times (0,\infty)$ with initial condition $u(x,0) = f(x)$ and, $u(x,t) = \int f(y)p(t,x,y)\, dy$, then $p(t,x,y)$ is also the fundamental solution to the Cauchy problem for $\mathcal{L}$. The equation $\partial_t u = \mathcal{L}u$ is a model for heat flow in a nonhomogeneous medium, which leads to the name *heat kernel* for $p(t,x,y)$.

First, we derive some properties of P_t. We continue to assume that the coefficients a_{ij} are smooth and that $\mathcal{L} \in \mathcal{D}(\Lambda)$ for some $\Lambda > 0$.

Proposition 28.1. *If $f \in C^\infty$ is bounded and in L^1, then $P_t f$ is differentiable in t and*

$$\partial_t P_t f = P_t \mathcal{L} f = \mathcal{L} P_t f.$$

Proof. By Itô's formula,

$$P_{t+h} f(x) - P_t f(x) = \mathbb{E}^x \int_t^{t+h} \mathcal{L} f(X_s)\, ds,$$

so

$$\partial_t P_t f(x) = \mathbb{E}^x \mathcal{L} f(X_t) = P_t \mathcal{L} f(x)$$

by the continuity of $\mathcal{L} f$.

We know $P_t f$ is a smooth function of x. Applying Itô's formula to $P_t f$,

$$\begin{aligned} P_h(P_t f)(x) - P_t f(x) &= \mathbb{E}^x P_t f(X_h) - \mathbb{E}^x P_t f(X_0) \\ &= \mathbb{E}^x \int_0^h \mathcal{L}(P_t f)(X_s)\, ds. \end{aligned}$$

However, $P_h(P_t f) = P_{t+h} f$ by the Markov property. Dividing by h, letting $h \to 0$, and using the continuity of $\mathcal{L}(P_t f)$,

$$\partial_t P_t f(x) = \mathbb{E}^x \mathcal{L}(P_t f)(X_0) = \mathcal{L} P_t f(x).$$ □

Next we show that P_t is a symmetric operator.

Proposition 28.2. *If f and g are bounded and in L^1,*

$$\int f(x) P_t g(x)\, dx = \int g(x) P_t f(x)\, dx.$$

Proof. Let $\overline{f}, \overline{g} \in L^1 \cap C^2$ be bounded with bounded first and second partial derivatives. By (25.5),

$$\int \overline{f}(\mathcal{L}\overline{g}) = \int \overline{g}(\mathcal{L}\overline{f}).$$

Therefore

$$\int \overline{f}\big((\lambda - \mathcal{L})\overline{g}\big) = \int \overline{g}\big((\lambda - \mathcal{L})\overline{f}\big). \tag{28.1}$$

If f, g are bounded C^∞ functions and $\lambda > 0$, let $\overline{f} = G^\lambda f$, $\overline{g} = G^\lambda g$, where G^λ is defined by

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

$\overline{f}$ and $\overline{g}$ can be shown to be smooth and $(\lambda - \mathcal{L})\overline{f} = f$, $(\lambda - \mathcal{L})\overline{g} = g$. By Jensen's inequality, $\overline{f}$ and $\overline{g}$ are in L^1. Substituting in (28.1),

$$\int (G^\lambda f) g = \int (G^\lambda g) f, \qquad \lambda > 0.$$

We have seen in Proposition 28.1 that $P_t f$ is differentiable in t, and hence continuous in t. Noting $G^\lambda f = \int_0^\infty e^{-\lambda t} P_t f\, dt$, the uniqueness of the Laplace transform tells us that

$$\int (P_t f) g = \int (P_t g) f, \qquad t > 0.$$

We now use a limit argument to extend this to the case where f and g are arbitrary bounded functions in L^1. □

With these preliminaries out of the way, we can now present Nash's method, which yields an upper bound for the transition density.

Theorem 28.3. *There exists a function $p(t, x, y)$ mapping $(0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$ to $[0, \infty)$ that is symmetric in x and y for almost every pair (x, y) (with respect to Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}^d$) and such that $P_t f(x) = \int f(y) p(t, x, y)\, dy$ for all bounded functions f. There exists c_1 depending only on Λ such that*

$$p(t, x, y) \le c_1 t^{-d/2}, \qquad t > 0, \quad x, y \in \mathbb{R}^d.$$

Proof. Let f be C^∞ with compact support with $\int f = 1$. We observe that

$$\int P_t f(x)\, dx = \int 1(P_t f) = \int (P_t 1) f = \int f = 1$$

because $P_t 1 = 1$.

Set

$$E(t) = \int (P_t f(x))^2 \, dx,$$

and note $E(0) = \int f(x)^2 \, dx < \infty$. By Proposition 28.1,

$$E'(t) = 2 \int P_t f(x) \partial_t (P_t f(x)) \, dx = 2 \int P_t f(x) \mathcal{L} P_t f(x) \, dx.$$

By Proposition 25.2, this is equal to

$$- \int \nabla (P_t f) \cdot a \nabla (P_t f)(x) \, dx \le -\Lambda \int |\nabla (P_t f)(x)|^2 \, dx,$$

since $\mathcal{L} \in \mathcal{D}(\Lambda)$. By Theorem 26.3 (the Nash inequality), we have the right-hand side bounded above in turn by

$$-c_2 \Big(\int (P_t f(x))^2 \Big)^{1+2/d} \Big(\int P_t f(x) \Big)^{4/d} = -c_2 E(t)^{1+2/d}.$$

Therefore

$$E'(t) \le -c_2 E(t)^{1+2/d}, \tag{28.2}$$

or

$$(E(t)^{-2/d})' \ge c_3.$$

(We are treating the differential inequality (28.2) by the same methods we would use if it were an equality and we had a first order separable differential equation.) An integration yields

$$E(t)^{-2/d} - E(0)^{-2/d} \ge c_3 t,$$

or

$$E(t)^{-2/d} \ge c_3 t.$$

We conclude from this that

$$E(t) \le c_4 t^{-d/2}.$$

Using the linearity of P_t, we thus have that

$$\|P_t f\|_2 \le c_4^{1/2} t^{-d/4} \|f\|_1 \tag{28.3}$$

for f smooth. A limit argument extends this to all $f \in L^1$. We now use a duality argument. If $g \in L^1$ and $f \in L^2$,

$$\int g(P_t f) = \int f(P_t g) \le \|f\|_2 \|P_t g\|_2 \le c_4^{1/2} t^{-d/4} \|g\|_1 \|f\|_2.$$

Taking the supremum over $g \in L^1$ with $\|g\|_1 \le 1$,

$$\|P_t f\|_\infty \le c_4^{1/2} t^{-d/4} \|f\|_2. \tag{28.4}$$

By the semigroup property, (28.4) applied to $P_{t/2}f$, and (28.3) applied to f,

$$\begin{aligned}\|P_t f\|_\infty = \|P_{t/2}(P_{t/2}f)\|_\infty &\le c_4^{1/2}(t/2)^{-d/4}\|P_{t/2}f\|_2\\ &\le c_4(t/2)^{-d/2}\|f\|_1.\end{aligned}$$

This says

$$|P_t f(x)| \le c_5 t^{-d/2} \int |f(y)|\, dy. \tag{28.5}$$

Applying this to $f = 1_B$, B a Borel set, we see that $\mathbb{P}^x(X_t \in dy)$ is absolutely continuous with respect to Lebesgue measure and the density, which we shall call $p(t,x,y)$, is nonnegative and bounded by $c_5 t^{-d/2}$ for almost all pairs (x,y). The symmetry (except for a null set of pairs) follows easily by Proposition 28.2. □

Recall $P_{t+s}f = P_t P_s f$, or

$$\begin{aligned}\int f(y)p(t+s,x,y)dy &= \int P_s f(z) p(t,x,z)dz\\ &= \int\int f(y)p(s,z,y)p(t,x,z)dy\, dz.\end{aligned}$$

This is true for every bounded f and it follows that

$$p(t+s,x,y) = \int p(t,x,z)p(t,z,y)\, dz, \qquad a.e.$$

29. Off-diagonal upper bounds.

One of the most important facts concerning divergence form operators is Aronson's bounds: if $p(t,x,y)$ is the fundamental solution to the heat equation $\partial u/\partial t = \mathcal{L}u$ in $\mathbb{R}^d$, then there exist constants c_1, c_2, c_3, c_4 such that

$$p(t,x,y) \le c_1 t^{-d/2} \exp(-c_2|x-y|^2/t), \qquad x,y \in \mathbb{R}^d, \quad t > 0, \tag{29.1}$$

and

$$c_3 t^{-d/2} \exp(-c_4|x-y|^2/t) \le p(t,x,y), \qquad x,y \in \mathbb{R}^d, \quad t > 0. \tag{29.2}$$

In this section we will prove (29.1). We will assume for convenience that the a_{ij} are smooth, but none of our estimates will depend on the smoothness; the case of non-smooth a_{ij} will follow from by an easy limit argument.

Under the assumption that the a_{ij} are smooth, the diffusion X_t associated with $\mathcal{L}$ can be written

$$X_t = X_0 + \int_0^t \sigma(X_s)dW_s + \int_0^t b(X_s)ds, \tag{29.3}$$

where W_t is a d-dimensional Brownian motion, the first integral on the right is an Itô stochastic integral, σ is a bounded positive definite square root of a, and the ith component of b is $\sum_{j=1}^d \partial a_{ji}/\partial x_j$. It is easy to deduce from this that for each x_0 and each t

$$\mathbb{E}^{x_0} \sup_{s\le t} |X_s - x_0|^2 < \infty. \tag{29.4}$$

Let $x_0 \in \mathbb{R}^d$,

$$M(t) = \int |y-x_0| p(t,x_0,y)dy, \qquad Q(t) = -\int p(t,x_0,y)\log p(t,x_0,y)dy. \tag{29.5}$$

Since $M(t) = \mathbb{E}^{x_0}|X_t - x_0|$, then $M(t)$ is finite by (29.4). The finiteness of $Q(t)$ will follow from (29.6) and (29.7) below.

Theorem 29.1. *There exists c_1 not depending on x_0 or t such that $M(t) \le c_1 t^{1/2}$.*

Proof. First, using Theorem 28.3 and the fact that $\int p(t,x_0,y)dy = 1$, we have

$$Q(t) \ge -c_2 + \tfrac{1}{2} d \log t. \tag{29.6}$$

Second, note $\inf_s(s\log s + \lambda s) = -e^{-\lambda-1}$. Using this with $\lambda = a|y-x_0| + b$, we obtain

$$\begin{aligned} -Q(t) + aM(t) + b &= \int [p(t,x_0,y)\log p(t,x_0,y) + (a|y-x_0|+b)p(t,x_0,y)]dy \\ &\ge -e^{-b-1}\int e^{-a|y-x_0|}dy = -c_3 e^{-b} a^{-d}. \end{aligned}$$

Setting $a = d/M(t)$ and $e^{-b} = (e/c_3)a^d$, after some algebra we obtain

$$M(t) \ge c_4 e^{Q(t)/d}. \tag{29.7}$$

Third, we differentiate $Q(t)$. Since the a_{ij} are smooth and uniform ellipticity holds, it is known that $p(t,x_0,y)$ is strictly positive and is C^∞ in each variable on $(0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$ and that $p(t,x_0,y)$ and its first and second partial derivatives have exponential decay at infinity. Performing the differentiation,

$$\begin{aligned} Q'(t) &= -\int (1+\log p(t,x_0,y))\frac{\partial}{\partial t} p(t,x_0,y)dy \\ &= -\int (1+\log p(t,x_0,y))\mathcal{L}p(t,x_0,y)dy. \end{aligned}$$

Note that $\int 1\mathcal{L}p(t,x,y)dy = \int(\mathcal{L}1)p(t,x,y)dy = 0$. Applying Proposition 25.2, the above expression is equal to

$$\int \nabla(\log p(t,x_0,y))\cdot a\nabla p(t,x_0,y)dy$$
$$= \int [\nabla(\log p(t,x_0,y)) \cdot a\nabla(\log p(t,x_0,y))]p(t,x_0,y)dy.$$

Using Cauchy-Schwarz and the uniform ellipticity bounds,

$$Q'(t) \geq c_5\Big(\int |\nabla \log p(t,x_0,y)|p(t,x_0,y)dy\Big)^2 = c_5\Big(\int |\nabla p(t,x_0,y)|dy\Big)^2.$$

Set $r(y) = |y-x_0|$. As $|\nabla r| \leq 1$, we have similarly that $M'(t) = -\int \nabla r \cdot a\nabla p(t,x_0,y)dy$, and so

$$|M'(t)| \leq \int |a\cdot\nabla p(t,x_0,y)|dy.$$

(Because r is not differentiable at x_0, to establish this we approximate r by smooth functions and use a simple limit argument.) We thus conclude

$$Q'(t) \geq c_6(M'(t))^2. \tag{29.8}$$

By the continuity of X_t, (29.4), and dominated convergence,

$$\lim_{t\to 0} M(t) = \lim_{t\to 0} \mathbb{E}^{x_0}|X_t - x_0| = 0,$$

so

$$c_4 e^{Q(t)/d} \leq M(t) \leq c_7\int_0^t (Q'(s))^{1/2}ds. \tag{29.9}$$

Finally, define $R(t) = d^{-1}[Q(t)+c_2-\frac{d}{2}\log t]$, and observe from (29.6) that $R(t)\geq 0$. Then

$$Q'(t) = dR'(t) + d/(2t).$$

Using (29.9) and the inequality $(a+b)^{1/2} \leq a^{1/2} + b/(2a^{1/2})$, we have

$$c_8 t^{1/2} e^{R(t)} \leq M(t) \leq c_9 d^{1/2}\int_0^t \Big(\frac{1}{2s} + R'(s)\Big)^{1/2} ds$$
$$\leq c_{10}\int_0^t \Big(\frac{1}{2s}\Big)^{1/2} ds + c_{10}\int_0^t \Big(\frac{s}{2}\Big)^{1/2} R'(s)ds.$$

By integration by parts and the fact that $R\geq 0$, this is less than

$$c_{11}(2t)^{1/2} + c_{11}R(t)\Big(\frac{t}{2}\Big)^{1/2},$$

which leads to

$$c_8 e^{R(t)} \le M(t)/t^{1/2} \le c_{12}(1 + R(t)).$$

The inequality $c_8 e^{R(t)} \le c_{12}(1+R(t))$ implies that $R(t)$ is bounded, and the result follows.
□

Theorem 29.2. *There exist c_1 and c_2 such that (29.1) holds for all $x, y \in \mathbb{R}^d$, $t > 0$.*

Proof. First, if $S_\lambda = \inf\{t : |X_t - X_0| \ge \lambda\}$, then

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

By Chebyshev's inequality and Theorem 1, the first term on the right hand side is bounded by

$$\frac{2\mathbb{E}^x|X_t - X_0|}{\lambda} \le \frac{2M(t)}{\lambda} \le \frac{c_3 t^{1/2}}{\lambda}.$$

By the strong Markov property, the second term is bounded by

$$\int_0^t \mathbb{E}^x[\mathbb{P}^{X_s}(|X_{t-s} - X_0| \ge \lambda/2); S_\lambda \in ds] \le \frac{2}{\lambda}\int_0^t M(t-s)\mathbb{P}^x(S_\lambda \in ds) \le \frac{c_3 t^{1/2}}{\lambda}.$$

Adding,

$$\mathbb{P}^x(\sup_{s\le t} |X_s - X_0| \ge \lambda) \le \frac{2c_3 t^{1/2}}{\lambda}. \tag{29.10}$$

Second, let $D > 0$, let $n = [aD^2]$, and let $b > 0$, where a, b will be chosen in a moment. By (29.10) we have

$$\begin{aligned}\mathbb{E}^x e^{-nS_{D/n}} &\le 1 \cdot \mathbb{P}^x(S_{D/n} < b/n) + e^{-n(b/n)}\mathbb{P}^x(S_{D/n} > b/n)\\ &= (1 - e^{-b})\mathbb{P}^x(S_{D/n} < b/n) + e^{-b}\\ &\le (1 - e^{-b})\frac{2c_3(b/n)^{1/2}}{D/n} + e^{-b}\\ &\le e^{-2}\end{aligned}$$

if we first choose b large and then a small; a, b can be chosen independently of x, D and n because

$$\frac{(b/n)^{1/2}}{D/n} = \frac{(bn)^{1/2}}{D} \le (ab)^{1/2}.$$

Let $T_0 = 0$ and define inductively $T_{i+1} = T_i + S_{D/n} \circ \theta_{T_i}$, where the θ_t are the usual shift operators for Markov processes; this means that T_i is the ith time that X_t moves a distance D/n. By the strong Markov property

$$\mathbb{E}^x e^{-nT_m} = \mathbb{E}^x[e^{-nT_{m-1}}; \mathbb{E}^{X_{T_{m-1}}} e^{-nS_{D/n}}] \le e^{-2}\mathbb{E}^x e^{-nT_{m-1}},$$

so by induction

$$\mathbb{E}^x e^{-nT_n} \le e^{-2n}.$$

Then

$$\begin{aligned} \mathbb{P}^x(\sup_{s\le 1}|X_s - X_0| \ge D) &\le \mathbb{P}^x(T_n \le 1) = \mathbb{P}^x(e^{-nT_n} \ge e^{-n}) \\ &\le e^n \mathbb{E}^x e^{-nT_n} \le e^n e^{-2n} = e^{-n} \le e^{-c_4 D^2}. \end{aligned} \tag{29.11}$$

Third, let $A = \{z : |z - x| > |z - y|\}$. By (29.11) with $D = |x - y|/2$ and Theorem 28.3

$$\begin{aligned} \int_A p(1, x, z)p(1, z, y)dz &\le c_5 \int_A p(1, x, z)dz = c_5\mathbb{P}^x(X_1 \in A) \\ &\le c_5\mathbb{P}^x(|X_1 - x| \ge |x - y|/2) \le c_5 e^{-c_6|x-y|^2}, \end{aligned}$$

while

$$\int_{A^c} p(1, x, z)p(1, z, y)dz \le c_5\mathbb{P}^y(X_1 \in A^c) \le c_5\mathbb{P}^y(|X_1 - y| \ge |x - y|/2) \le c_5 e^{-c_6|x-y|^2}.$$

Adding and using the semigroup property,

$$p(2, x, y) = \int p(1, x, z)p(1, z, y)dz \le 2c_5 e^{-c_6|x-y|^2}.$$

The theorem now follows by scaling. □

30. Lower bounds.

In this section we obtain a lower bound on the transition densities. We start with a lower bound for $p_t(x) = p(t, x_0, x)$ when x is close to x_0. To do that we need the following weighted Poincaré inequality.

Theorem 30.1. *Let $\varphi(x) = c_1 e^{-|x|^2/2}$, where c_3 is chosen so that $\int_{\mathbb{R}^d} \varphi(x)dx = 1$. Then*

$$\int_{\mathbb{R}^d} |f(x) - \overline{f}|^2\varphi(x)\, dx \le c_1 \int_{\mathbb{R}^d} |\nabla f(x)|^2\varphi(x)\, dx,$$

where $\overline{f} = \int_{\mathbb{R}^d} f(x)\varphi(x)\,dx$.

Proof. The proof is very similar to the proof of the ordinary Poincaré inequality. For the one dimensional case, instead of Fourier series, we express f in terms of Hermite polynomials. $H_0 = 0$ and $H_k' = c_k H_{k-1}$ where $|c_k| > 0$ for $k \neq 0$. For the higher dimensional case, we write f as a series in $\prod_{i=1}^d H_{k_i}(x_i)$. □

Theorem 30.2. *Let* $x_1 \in \mathbb{R}^d$*. There exist constants* c_1 *and* c_2 *such that if* $|x_1 - x_0| < c_1$*, then* $p(2, x_0, x_1) \geq c_2$*.*

Proof. Let $\varphi(x) = c_1 e^{-|x-x_1|^2/2}$, where c_3 is chosen so that $\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$. We will choose r later on and we suppose $|x_1 - x_0| < r/2$. Define

$$G(t) = \int \log p_t(x)\varphi(x)\,dx.$$

Then

$$\begin{aligned}
G'(t) &= \int \frac{\mathcal{L}p_t}{p_t}\varphi\,dx \\
&= -\int (\nabla p_t)\cdot a\Big(\nabla \frac{\varphi}{p_t}\Big) \\
&= \int \nabla p_t \cdot a\frac{\nabla p_t}{p_t^2}\varphi - \int \frac{\nabla p_t}{p_t}\cdot a\nabla\varphi \\
&= \frac{1}{2}\int \frac{\nabla p_t}{p_t}\cdot a\frac{\nabla p_t}{p_t}\varphi + \frac{1}{2}\int \Big[\frac{\nabla p_t}{p_t}\cdot a\frac{\nabla p_t}{p_t} + 2\frac{\nabla p_t}{p_t}\cdot ax + x\cdot ax\Big]\varphi \\
&\qquad - \frac{1}{2}\int x\cdot ax\,\varphi \\
&\geq \tfrac{1}{2}\int |\nabla \log p_t|^2\varphi - c_4 \\
&\geq c_5 \int |\log p_t - G(t)|^2\varphi - c_4 \\
&\geq c_5 \int_{D_t} |\log p_t - G(t)|^2\varphi - c_4,
\end{aligned}$$

where $D_t = \{x \in B(x_0, r) : p_t(x) \geq e^{-K}\}$ and r and K will be chosen in a moment.

First choose r such that

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

So

$$\begin{aligned}
\tfrac{3}{4} &\leq \int_{B(x_0,r)} p_t(x)\,dx = \int_{D_t} + \int_{B(x_0,r)-D_t} \\
&\leq c_6|D_t|t^{-d/2} + c_7 r^d e^{-K}.
\end{aligned}$$

Choose K so that $c_7 r^d e^{-K} \le 1/4$. Then $c_6|D_t|t^{-d/2} \ge 1/2$, or $|D_t| \ge t^{d/2}/(2c_6) \ge c_8$ if $t \in [1/2, 1]$.

For any t, by Jensen's inequality,

$$G(t) = \int (\log p_t(x))\varphi \le \log \int p_t \varphi \le \log \int p_t = \log 1 = 0.$$

So $G(t) \le 0$. If $\log p_t > 0$, then $|\log p_t - G(t)| \ge |G(t)|$. If $-K \le \log p_t \le 0$, then either (a): $\quad |G(t)| \ge 2K$ or (b): $\quad |G(t)| < 2K$. In case (a),

$$(\log p_t - G)^2 = (-|\log p_t| + |G(t)|)^2 \ge \tfrac{1}{4}|G(t)|^2.$$

In case (b),

$$(\log p_t - G)^2 = (-|\log p_t| + |G(t)|)^2 \ge 0 \ge \tfrac{1}{4}|G(t)|^2 - K^2.$$

Therefore, using the fact that φ is bounded below by a positive constant on $B(x_0, r)$ and $D \subset B(x_0, r)$,

$$\int_{D_t} |\log p_t - G(t)|^2 \varphi \ge c_9 |D_t|(|G(t)|^2 - K^2) \ge c_{10} G(t)^2 - c_{11}.$$

We thus have

$$G'(t) \ge BG(t)^2 - A, \qquad t \in [1/2, 1]$$

for some constants A and B.

Now we do some calculus. Suppose $G(1) \le -Q$ where $Q = \max(4A, (16A/B)^{1/2})$. Since $G'(t) \ge -A$,

$$G(1) - G(t) \ge -A/2, \qquad t \in [1/2, 1],$$

or

$$G(t) \le G(1) + A/2.$$

This implies $G(t) \le -Q/2$. Since $BQ^2/4 \ge 4A$, then $A < \frac{B}{2}G(t)^2$, and hence

$$G'(t) \ge \frac{B}{2} G(t)^2.$$

Solving, $G'/G^2 \ge B/2$, or $(1/G)' = -G'/G^2 \le -B/2$, and then

$$\frac{1}{G(1)} - \frac{1}{G(t)} \le -\frac{B}{2}(1 - t) \le -\frac{B}{4}.$$

Since $G(t) \le 0$, then $1/G(1) \le -B/4$, or $1 \ge -BG(1)/4$, or

$$G(1) \ge -\frac{4}{B}.$$

So either $G(1) \geq -Q$ or $G(1) \geq -4/B$. In either case there exists $R > 0$ such that $G(1) \geq -R$.

Finally, applying the above first with x_0 and then with x_0 replaced by x_1

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

This gives our result. □

We now complete the proof of the lower bound by what is known as a *chaining argument.*

Theorem 30.3. *There exist c_1 and c_2 depending only on Λ such that*

$$p(t, x, y) \geq c_1 t^{-d/2} e^{-c_2 |x-y|^2/t}.$$

Proof. By Theorem 30.2 with $x_0 = x$ and scaling, there exists c_3 such that if $|x - y| < c_3 t^{1/2}$, then $p(t, x, y) \geq c_4 t^{-d/2}$. Thus to prove the theorem, it suffices to consider the case $|x - y| \geq c_3 t^{1/2}$.

By Theorem 30.2 (with $x_0 = w$) and scaling, there exist c_4 and c_5 such that if $|z - w| \leq c_4 (t/n)^{1/2}$,

$$p(t/n, w, z) \geq c_5 (t/n)^{-d/2}. \tag{30.1}$$

Let $R = |x - y|$ and let n be the smallest positive integer greater than $9R^2/c_4^2 t$. So $3R/n \leq c_4 (t/n)^{1/2}$. Let $v_0 = x$, $v_n = y$, and $v_1, \ldots, v_{n-1}$ be points equally spaced on the line segment connecting x and y. Let $B_i = B(v_i, R/n)$. If $w \in B_i$ and $z \in B_{i+1}$, then $|z - w| \leq 3R/n \leq c_4 (t/n)^{1/2}$, and so $p(t/n, w, z) \geq c_5 (t/n)^{-d/2}$ by (30.1).

By the semigroup property,

$$\begin{aligned}&p(t, x, y) \\ &\quad = \int \cdots \int p(t/n, x, z_1) p(t/n, z_1, z_2) \cdots p(t/n, z_{n-1}, y)\, dz_1 \cdots dz_{n-1} \\ &\quad \geq \int_{B_{n-1}} \cdots \int_{B_1} p(t/n, x, z_1) p(t/n, z_1, z_2) \cdots p(t/n, z_{n-1}, y)\, dz_1 \cdots dz_{n-1} \\ &\quad \geq (c_6 (t/n)^{-d/2})^n \prod_{i=1}^{n-1} |B_i|.\end{aligned}$$

Since $|B_i| \geq c_7(R/n)^d$ and $(R/n)(t/n)^{-1/2}$ is bounded below by a positive constant, then

$$p(t,x,y) \geq c_8 c_9^n (n/t)^{d/2} \geq c_8 t^{-d/2} \exp(-n \log c_9^{-1}).$$

If $n > 2$, then $n/2 \leq 9R^2/c_4^2 t$, so

$$p(t,x,y) \geq c_8 t^{-d/2} \exp(-18R^2 \log c_9^{-1}/c_4^2 t).$$

If $n \leq 2$, then $9R^2/c_4^2 t \leq 2$, and

$$p(t,x,y) \geq c_8 t^{-d/2} \exp(-2 \log c_9^{-1}).$$

The result follows with $c_1 = c_8(c_9^2 \wedge 1)$ and $c_2 = 18(\log(c_9^{-1}) \wedge 1)/c_4^2$. □