Stochastic processes III
BMS Advanced course

Winter 2007/8, TUB

Anton Bovier

Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin, Germany

Version: 12.02.2008

Contents

1

Stochastic differential equations

1.1 Stochastic integral equations

We will define the notion of stochastic differential equations first.

We want to construct stochastic processes where the velocities are given as functions of time and position, and that have in addition a stochastic component. We will consider the case where the stochastic component comes from a Brownian motion, B_t. Such an equation should look like

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \tag{1.1}$$

with prescribed initial conditions $X_0 = x_0$. The interpretation of such an equation is not totally straightforward, due to the term $\sigma(t, X_t)dB_t$. We will interpret such an equation as the integral equation

$$X_t = x_0 + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dB_s, \tag{1.2}$$

where the integral with respect to B is understood as the Itô stochastic integral defined in the last chapter. The functions b, σ are in the most general setting assumed to be locally bounded and measurable.

The questions one is of course interested are those of existence and uniqueness of solutions to such equations, as well as that of properties of solutions. We begin by discussing the notions of strong and weak solutions.

1.2 Strong and weak solutions

We will denote by W the Polish space $C(\mathbb{R}_+, \mathbb{R}^n)$ of continuous paths and we denote by $\mathcal{H}$ the corresponding Borel-σ-algebra, and by $\mathcal{H}_t \equiv \sigma\{x_s, s \leq t\}$ the filtration generated by the paths up to time t.

The formal set-up for a stochastic differential equation involves an initial conditions and a Brownian motion, all of which require a probability space. We will denote this by

$$(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}, \xi, B), \tag{1.3}$$

where

(i) $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ is a filtered space satisfying the usual conditions;
(ii) B is a Brownian motion (on $\mathbb{R}^d$), adapted to $\mathcal{F}_t$,
(iii) ξ is a $\mathcal{F}_0$-measurable random variable.

The minimal or *canonical* set-up has $\Omega = \mathbb{R}^n \times W$, $\mathbb{P} = \mu \times \mathbb{Q}$, where μ is the law of ξ and $\mathbb{Q}$ is Wiener measure and $\mathcal{F}_t$ the usual augmentation of $\mathcal{F}_t^0 \equiv \sigma\{\xi, B_s, s \leq t\}$.

The precise definition of *path-wise uniqueness* of a SDE is as follows:

Definition 1.2.1 For a SDE, path-wise uniqueness holds, if the following holds: For any set-up $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}, \xi, B)$, and any two continuous semi-martingales X and X', such that

$$\int_0^t (|b(s, X_s)| + |\sigma(s, X_s)|^2) ds < \infty \tag{1.4}$$

solving the SDE with this initial condition ξ and this Brownian motion B,

$$\mathbb{P}[X_t = X'_t, \quad \forall t] = 1. \tag{1.5}$$

If a SDE admits for any setup $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}, \xi, B)$ exactly one continuous semi-martingale as solution, we say that the SDE is *exact*.

The notion of *strong solutions* is naturally associated with the setting of exact SDE's.

Definition 1.2.2 A strong solution of a SDE is a function,

$$F : \mathbb{R}^n \times W \to W, \tag{1.6}$$

such that

$$F^{-1}(\mathcal{H}_t) \subset \mathcal{B}(\mathbb{R}^n) \times \bar{\mathcal{H}}_t, \forall t \geq 0, \tag{1.7}$$

and on any set-up $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}, \xi, B)$, the process

$$X = F(\xi, B)$$

solves the SDE. $\bar{H}_t$ is the augmentation of $\mathcal{H}_t$ with respect to the Wiener measure.

Existence and uniqueness results in the strong sense can be proven in a very similar way as in the case of ordinary differential equations, using Gronwall's inequality and the Picard iteration scheme.

The general approach is to assume local Lipshitz conditions, to prove existence of solutions for finite times, and then glue solutions together until a possible explosion.

Let us give the basic uniqueness and existence results, essentially due to Itô.

Theorem 1.2.1 *Assume that σ and b are bounded measurable, and that in addition there exists an open set $U \subset \mathbb{R}$, and $T > 0$, such that there exists $K < \infty$, s.t.*

$$|\sigma(t,x) - \sigma(t,y)| + |b(t,x) - b(t,y)| \leq K|x - y|, \tag{1.8}$$

for all $x, y \in U, t < T$. Let X, Y be two solutions of (1.2) (with the same Brownian motion B), and set

$$\tau \equiv \inf\{t \geq 0 : X_t \notin U \, or Y_t \notin U\}. \tag{1.9}$$

Then, if $\mathbb{E}[X_0 - Y_0]^2 = 0$, it follows that

$$\mathbb{P}\left[X(t \wedge \tau) = Y(t \wedge \tau), \forall 0 \leq t \leq T\right] = 1. \tag{1.10}$$

Proof. The proof is based on Gronwall's lemma and very much like the deterministic analog. We compute

$$\begin{aligned}
&\mathbb{E}\left[\max_{0\leq s\leq t} \left(X(s\wedge\tau) - Y(s\wedge\tau)\right)^2\right] \qquad (1.11)\\
&\leq 2\mathbb{E}\left[\max_{0\leq s\leq t}\left(\int_0^{s\wedge\tau} (\sigma(u,X(u)) - \sigma(u,Y(u)))dB_u\right)^2\right]\\
&+2\mathbb{E}\left[\max_{0\leq s\leq t}\left(\int_0^{s\wedge\tau} (b(u,X(u)) - b(u,Y(u)))du\right)^2\right]\\
&\leq 8\mathbb{E}\left[\int_0^{t\wedge\tau} (\sigma(u,X(u)) - \sigma(u,Y(u)))^2 du\right]\\
&+2t\mathbb{E}\left[\int_0^{t\wedge\tau} (b(u,X(u)) - b(u,Y(u)))^2 du\right]\\
&\leq 2K^2(t+4)\mathbb{E}\left[\int_0^{t\wedge\tau} \left(X(u) - Y(u)\right)^2 du\right]\\
&\leq 2K^2(4+t)\int_0^t \mathbb{E}\left[\max_{0\leq u\leq s} \left(X(u\wedge\tau) - Y(u\wedge\tau)\right)^2 ds\right].
\end{aligned}$$

Note that in the first inequality we used that $(a+b)^2 \leq 2a^2 + 2b^2$, in the second we used the Schwartz inequality for the drift term and Doob's L^2-maximum inequality for the diffusion term; the next inequality uses the Lipshitz condition and in the last we used Fubini's theorem.

Gronwall's inequality then implies that

$$\mathbb{E}\left[\max_{0\leq t\leq T}\left(X(t\wedge\tau)-Y(t\wedge\tau)\right)^2\right]=0.$$

This is most easily proven as follows: Let f be a non-negative function that satisfies the integral equation $f(t)\leq K\int_0^t f(s)ds$. Set $F(t)=\int_0^t f(s)ds$. Then

$$0\leq\frac{d}{dx}\left(e^{-tK}F(t)\right)\leq e^{-Kt}\left(-KF(t)+f(t)\right)\leq 0,$$

and hence $e^{-tK}F(t)\leq 0$, meaning that $F(t)\leq 0$. But since F is the integral of the non-negative function f, this means that $f(t)=0$.

Thus we have in particular that $\mathbb{P}[\max_{0\leq t\leq T}|X_t-Y_t|=0]=1$ as claimed. □

Finally, existence of solutions (for finite times) can be proven by the usual Picard iteration scheme under Lipschitz and growth conditions.

Theorem 1.2.2 *Let b,σ satisfy the Lipshitz conditions (1.8) and assume that*

$$|b(t,x)|^2+|\sigma(t,x)|^2\leq K^2(1+|x|^2). \tag{1.12}$$

Let ξ be a random vector with finite second moment, independent of B_t, and let $\mathcal{F}_t$ be the usual augmentation, $\mathcal{F}_t$, of the filtration associated with B and ξ. Then there exists a continuous, $\mathcal{F}_t$-adapted process X which is a strong solution *of the SDE with initial condition ξ. Moreover, X is square integrable, i.e. for any $T>0$, there exists $C(T,K)$, such that, for all $t\leq T$,*

$$\mathbb{E}|X_t|^2\leq C(K,T)(1+\mathbb{E}|\xi|^2)e^{C(K,T)t}. \tag{1.13}$$

Proof. We define a map, F, from the space of continuous adapted processed X, uniformly square integrable on $[0,T]$, to itself, via

$$F(X)_t\equiv\xi+\int_0^t b(s,X_s)ds+\int_0^t\sigma(s,X_s)dB_s. \tag{1.14}$$

Note that the square integrability of $F(X)$ needs the growth conditions (1.12)

Exercise: Prove this!

As in (1.11)

$$\mathbb{E}\left(\sup_{0\le t\le T}(F(X)_t - F(Y)_t)\right)^2 \tag{1.15}$$
$$\le 2E\left(\sup_{0\le t\le T}\left(\int_0^t (\sigma(X_s)-\sigma(Y_s))dB_s\right)^2\right)$$
$$+2E\left(\sup_{0\le t\le T}\left(\int_0^t (b(X_s)-b(Y_s))d_s\right)^2\right)$$
$$\le 2K^2(1+T)\int_0^T \mathbb{E}\sup_{0\le s\le t}(X_s-Y_s)^2\,dt$$

and hence

$$\mathbb{E}\left(\sup_{0\le t\le T}(F^k(X)_t - F^k(Y)_t)\right)^2 \le \frac{C^kT^{2k}}{k!}\mathbb{E}\left(\sup_{0\le t\le T}(X_t-Y_t)\right)^2. \tag{1.16}$$

Thus, for n sufficiently large, F^n is a contraction, and hence has a unique fixed point which solves the SDE. □

Remark 1.2.1 The conditions for existence above are not necessary. In particular, growth conditions are important only when the solutions can actually reach the regions there the coefficients become too big. Formulations of weaker hypothesis for existence and uniqueness can be found for instance in [9], Chapter 14. Their verification in concrete cases can of course be rather tricky.

We will now consider a weaker form of solutions, in which the solution is not constructed from the BM, but the BM comes from the solution. This is like in the martingale problem formulation, and we will soon see the equivalence of the two concepts.

Definition 1.2.3 A stochastic integral equation

$$X_t = X_0 + \int_0^t \sigma(s,X_s)dB_s + \int_0^t b(s,X_s)ds \tag{1.17}$$

has a *weak solution* with initial distribution μ, if there exists a filtered space $(\Omega,\mathcal{F},\mathbb{P},\{\mathcal{F}_t\})$, satisfying the usual conditions, and continuous martingales X and B, such that

(i) B is an $\mathcal{F}_t$-Brownian motion;
(ii) X_0 has law μ;
(iii) $\int_0^t(|\sigma(s,X_s)|^2 + |b(s,X_s)|)ds < \infty$, a.s., for all t;
(iv) (1.17) holds.

Definition 1.2.4 A solution of (1.17) is unique in law (or *weakly unique*, if whenever X_t and X'_t are two solutions such that the laws of X_0 and X'_0 are the same, then the laws of X and X' coincide.

Example. The following simple example illustrates the difference between strong and weak solutions. Consider the equation

$$X_t = X_0 + \int_0^t \mathrm{sign}\,(\mathrm{X_s})\mathrm{dB_s}. \tag{1.18}$$

Here we define $\mathrm{sign}\,(\mathrm{x}) = -1$, if $x \leq 0$, and $\mathrm{sign}\,(\mathrm{x}) = +1$, if $x > 0$. Obviously, $[X]_t = \int_0^t dt = t,$, so for any solution, X_t, that is a continuous local martingale, Lévy's theorem implies that X_t is a Brownian motion, if it exists. In particular, we have weak uniqueness of the solution. Moreover, we can easily construct a solution: Let X_t be a Brownian motion and set

$$B_t \equiv \int_0^t \mathrm{sign}\,(\mathrm{X_s})\mathrm{dX_s}. \tag{1.19}$$

Then $dB_s = \mathrm{sign}\,(\mathrm{X_s})\mathrm{dX_s}$, and hence

$$\int_0^t \mathrm{sign}\,(\mathrm{X_s})\mathrm{dB_s} = \int_0^\mathrm{t} \mathrm{sign}\,(\mathrm{X_s})^2\mathrm{dX_s} = \int_0^\mathrm{t} \mathrm{dX_s} = \mathrm{X_t} - \mathrm{X_0},$$

so the pair (X, B) yields a weak solution! Note that the Brownian motion is constructed from X, not the other way around! On the other hand, there is no path-wise uniqueness: Let, say, $X_0 = 0$. Then, if X_t is a solution, so is $-X_t$. Of course being Brownian motions, they have the same law. Note that the corresponding B_t in the construction above would be the same. Moreover, the Brownian motion of (1.19) is measurable with respect to the filtration generated by $|X_t|$ which is smaller than that of X_t; thus, X_t is not adapted to the filtration generated by the Brownian motion. Hence we see that there is indeed not necessarily a solution of this SDE for any B, and so this SDE does not have a strong solution.

Remark 1.2.2 The example (and in particular the last remark) is hiding an interesting fact and concept, that of *local time*. This is the content of the following theorem due to Tanaka:

Theorem 1.2.3 *Let X be a continuous semi-martingale. Then there exists a continuous increasing adapted process, $\{\ell_t, t \geq 0\}$, called* the local time of X at 0*, such that*

$$|X_t| - |X_0| = \int_0^t \mathrm{sign}\,(\mathrm{X_s})\mathrm{dX_s} + \ell_\mathrm{t}. \tag{1.20}$$

ℓ_t grows only when X is zero, i.e.

$$\int_0^t 1_{X_s \neq 0} d\ell_s = 0. \tag{1.21}$$

The proof of this result is not very hard and proceeds by approximating the function $|x|$ by smooth functions, f_n and passing to the limit carefully.

Note that this theorem implies that in the example above, $B_t = |X_t| - \ell_t$, and since ℓ_t depends only on $|X|$, the measurability properties claimed above hold.

The connection between weak and strong solutions is clarified in the following theorem due to Yamada and Watanabe. It essentially says that weak existence and path-wise uniqueness imply the existence of a strong solution, and in turn weak uniqueness.

Theorem 1.2.4 *An SDE is exact if and only if*

(i) there exists a weak solution, and
(ii) solutions are path-wise unique.

Then uniqueness in law also holds.

The proof of this theorem may be found in [12]

1.3 Weak solutions and the martingale problem

We will now show a deep and important connection between weak solutions of SDEs and the martingale problem.

The remarkable thing is that these issues can be cooked down again to the study of martingale problems. We do the computations for the one-dimensional case, but clearly everything goes through in the d-dimensional case exactly in the same way.

Let us first observe that, using Itô's formula, given that the equation (1.2) has a solution, then it is a solution of a martingale problem.

Lemma 1.3.5 *Assume that X solves (1.2). Define the operator G on the space of C^∞-functions $f : \mathbb{R} \to \mathbb{R}$, as*

$$G_t \equiv \frac{1}{2}\sigma^2(t,x)\frac{d^2}{dx^2} + b(t,x)\frac{d}{dx}. \tag{1.22}$$

Then X is a solution of the martingale problem for G.

Proof. For later use we will derive a more general result. Let $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$. We use Itô's formula to express

$$f(t,X_t) - f(0,X_0) = \int_0^t \partial_s f(s,X_s)ds + \int_0^t \partial_x f(s,X_s)dX_s \quad (1.23)$$
$$+\frac{1}{2}\int_0^t \partial_x^2 f(s,X_s)d[X]_s.$$

Now

$$dX_s = b(s,X_s)ds + \sigma(s,X_s)dB_s.$$

We set

$$M(t) \equiv X_t - \int_0^t b(s,X_s)ds$$

and note that this is by (**??**) equal to $\int_0^t \sigma(s,X_s)dB_s$, and hence a martingale. Moreover,

$$[M]_t = \int_0^t \sigma(s,X_s)^2 d[B]_s = \int_0^t \sigma(s,X_s)^2 ds.$$

Hence

$$\begin{aligned} f(t,X_t) - f(0,X_0) = &\int_0^t \partial_x f(s,X_s)b(s,X_s)ds \\ &+ \int_0^t \partial_s f(s,X_s)ds + \frac{1}{2}\int_0^t \sigma(s,X_s)\partial_x^2 f(s,X_s)ds \\ &+ \int_0^t \partial_x f(s,X_s)dM_s, \end{aligned}$$

or

$$f(t,X_t) - f(0,X_0) - \int_0^t [\partial_s f(s,X_s + (Gf)(s,X_s)]ds = -\int_0^t \partial_x f(s,X_s)dM_s, \quad (1.24)$$

where the right-hand side is a martingale, which means that X solves the martingale problem, as desired. □

This observation becomes really useful through the converse result.

Theorem 1.3.6 *Assume that b and σ are locally bounded as above and assume that in addition σ^{-1} is locally bounded. Let G_t be given by (1.22). Assume that X is a continuous solution to the martingale problem for (G, δ_{x_0}), then there exists a Brownian motion, B, such that (X, B) is a solution to the stochastic integral equation (1.2).*

Proof. We know that for every $f \in C^\infty(\mathbb{R})$,

$$f(X_t) - f(X_0) - \int_0^t (G_s f)(s, X_s) ds \tag{1.25}$$

is a continuous martingale. Choosing $f(x) = x$, it follows that

$$X_t - X_0 - \int_0^t b(s, X_s) ds \equiv M_t \tag{1.26}$$

is a continuous martingale. Essentially we want to show that this martingale is precisely the stochastic integral term in (1.2). To do this, we need to compute the bracket of M. To do this, we consider naturally (1.1) with $f(x) = x^2$. To simplify the notation, let us assume without loss of generality that $X_0 = 0$. This gives

$$X_t^2 - 2\int_0^t X_s b(s, X_s) ds - \int_0^t \sigma^2(s, X_s) ds = \widehat{M}(t), \tag{1.27}$$

where $\widehat{M}$ is a martingale. Thus

$$\begin{aligned} M(t)^2 &= X_t^2 - 2X_t \int_0^t b(s, X_s) ds + \left(\int_0^t b(s, X_s) ds \right)^2 \\ &= 2\int_0^t X_s b(s, X_s) ds + \int_0^t \sigma^2(s, X_s) ds + \widehat{M}(t) \\ &\quad - 2X_t \int_0^t b(s, X_s) ds + \left(\int_0^t b(s, X_s) ds \right)^2 . \end{aligned} \tag{1.28}$$

I claim that

$$2\int_0^t X_s b(s, X_s) ds - 2X_t \int_0^t b(s, X_s) ds + \left(\int_0^t b(s, X_s) ds \right)^2 \tag{1.29}$$

is also a martingale. By partial integration,

$$\int_0^t X_s b(s, X_s) ds = X_t \int_0^t b(s, X_s) ds - \int_0^t \int_0^s b(u, X_u) du dX_s.$$

Thus (1.29) equals

$$\begin{aligned} &-2\int_0^t \int_0^s b(u, X_u) du dM(s) \\ &-2\int_0^t \int_0^s b(u, X_u) du b(s, X_s) ds + \left(\int_0^t b(s, X_s) ds \right)^2 \\ &= -2\int_0^t \int_0^s b(u, X_u) du dM(s), \end{aligned}$$

which is a martingale. Hence

$$M(t)^2 - \int_0^t \sigma^2(s, X_s) ds \tag{1.30}$$

is a martingale, so that by definition of the quadratic variation process,

$$\int_0^t \sigma^2(s, X_s) ds = [M]_t.$$

Now set

$$B(t) \equiv \int_0^t \frac{1}{\sigma(s, X_s)} dM_s.$$

Then

$$[B]_t = \int_0^t \frac{1}{\sigma(s, X_s)^2} d[M]_s = t,$$

so by Lévy's theorem **??**, $B(t)$ is Brownian motion, and it follows that X solves (1.2) with this particular realization of Brownian motion. □

We can summarize these findings in the following theorem.

Theorem 1.3.7 *Let $\mathbb{P}^y$ be a solution of the martingale problem associated to the operator G defined in (1.22) starting in y. Then there exists a weak solution of the SDE (1.2) with law $\mathbb{P}^y$. Conversely, if there is a weak solution of (1.2), then there exists a solution of the martingale problem for (1.22). Uniqueness in law holds if and only if the associated martingale problem has a unique solution.*

In other words, solutions of our stochastic integral equation are Markov processes with generator given by the closure of the second order (elliptic) differential operator G given by (1.22). To study their existence and uniqueness, we can use the tools we developed in the theory of Markov processes. Note that we state the theorem without the boundedness assumption on σ^{-1} from Theorem 1.3.6, which in fact can be avoided with some extra work.

As a consequence, we sketch two existence and uniqueness results for weak solutions.

Theorem 1.3.8 *Consider the SDE with time-independent coefficients,*

$$dX_t = b(X_t) + \sigma(X_t) dB_t, \tag{1.31}$$

in $\mathbb{R}^d$ where the coefficients b_i and σ_{ij} are bounded and continuous. Then for any measure μ such that

$$\int \|x\|^{2m} \mu(dx) < \infty, \tag{1.32}$$

for some $m > 1$, there exists a weak solution to (1.31) with initial measure μ.

Proof. We only have to prove that the martingale problem with generator

$$Gf(y) = \sum_i b_i(y)\partial_i f(y) + \frac{1}{2}\sum_{i,j,k} \sigma_{ik}(y)\sigma_{kj}(y)\partial_i\partial_j f(y),$$

for $f \in C_0^2(\mathbb{R}^d)$ has a solution. To do this, we construct an explicit solution for a sequence of operators $G^{(n)}$ that converge to G and deduce from this the existence of the solution of the martingale problem for G.

To do this, let $t_j^{(n)} = j2^{-n}$ and set $\phi_n(t) = t_j^{(n)} \mathbb{1}_{t\in[t_j^{(n)}, t_{j+1}^{(n)})}$. Then set

$$b^{(n)}(t, y) \equiv b(y(\phi_n(t))), \quad \sigma^{(n)}(t, y) \equiv \sigma(y(\phi_n(t))).$$

Then define the processes $X_t^{(n)}$ by

$$X_0^{(n)} = \xi \tag{1.33}$$
$$X_t^{(n)} = X_{t_j^{(n)}}^{(n)} + b(X_{t_j^{(n)}}^{(n)})(t - t_j^{(n)}) + \sigma(X_{t_j^{(n)}}^{(n)})(B_t - B_{t_j^{(n)}}), t \in (t_j^{(n)}, t_{j+1}^{(n)}].$$

We will denote the laws of the processes $X^{(n)}$ by $P^{(n)}$. One easily verifies that the processes $X^{(n)}$ solves the integral equation

$$X_t^{(n)} = \xi + \int_0^t b^{(n)}(s, X^{(n)})ds + \int_0^t \sigma^{(n)}(s, X^{(n)})dB_s. \tag{1.34}$$

But then $X^{(n)}$ solves the martingale problem for the (time dependent) operator

$$(G_t^{(n)} f)(y) \equiv \sum_i b_i^{(n)}(t, y)\partial_i f(y(t)) + \frac{1}{2}\sum_{i,j,k} \sigma_{ik}^{(n)}(t, y)\sigma_{kj}^{(n)}(t, y)\partial_i\partial_j f(y(t)). \tag{1.35}$$

The first thing to show is that this family of probability measures is tight. For this one uses the criterion given by Theorem **??**. The basic ingredient is the following bound that is proven in a manner very similar to the bound (1.13):

$$\mathbb{E}\left\|X_t^{(n)} - X_s^{(n)}\right\|^{2m} \leq C_m(t - s)^m \tag{1.36}$$

for $0 \leq t, s \leq T$, with C_m uniform in n, and depends only on the bound on the coefficients. Moreover,

$$\mathbb{E}\|X_0^{(n)}\|^{2m} \leq C'_m < \infty \tag{1.37}$$

by assumption. The proof of (1.37) needs an otherwise very useful inequality that we take the opportunity to state here, the so-called Burkholder-Davis-Gundy inequality

Lemma 1.3.9 *Let M be a continuous local martingale. Then, for every $m > 0$, there exist universal constants k_m, K_m depending only on m, such that, for any stopping time T,*

$$k_m \mathbb{E}[M]_T^m \leq \mathbb{E}\left(\sup_{0\leq s\leq T} |M_s|\right)^{2m} \leq K_m \mathbb{E}[M]_T^m. \tag{1.38}$$

Proof. The key to the construction are moment inequalities for martingales combined with Doob's maximum inequality. One starts with the following observation: Define, for $\delta > 0, \epsilon \geq 0$

$$Y_t \equiv \delta + \epsilon[M]_t + M_t^2 = \delta + (1+\epsilon)[M]_t + 2\int_0^t M_s dM_s. \tag{1.39}$$

Using Itô's formula then gives that

$$Y_t^m = \delta^m + m(1+\epsilon)\int_0^t Y_s^{m-1} d[M]_s + 2(m(m-1)\int_0^t Y_s^{m-2} M_s^2 d[M]_s$$
$$+2m\int_0^t Y_s^{m-1} M_s dM_s. \tag{1.40}$$

The last integral is a uniformly integrable martingale so its mean vanishes (even if t is replaced by a stopping time. Hence

$$\mathbb{E}Y_T^m = \delta^2 + m(1+\epsilon)\mathbb{E}\int_0^T Y_s^{m-1} d[M]_s + 2m(m-1)\mathbb{E}\int_0^T Y_s^{m-2} M_s^2 d[M]_s. \tag{1.41}$$

In particular, seting $\epsilon = 0$ and letting $\delta \downarrow 0$, this gives

$$\mathbb{E}|M_T|^{2m} = 2m(m-1/2)\mathbb{E}\int_0^T |M_s|^{2(m-1)} d[M]_s. \tag{1.42}$$

Let us now only consider the case when $m > 1$. Letting $\delta \downarrow 0$, (1.41) gives

$$\mathbb{E}\left(\epsilon[M]_T + M_T^2\right)^m \geq m(1+\epsilon)\mathbb{E}\int_0^T \left(\epsilon[M]_s + M_s^2\right)^{m-1} d[M]_s$$
$$\geq m(1+\epsilon)\epsilon^{m-1}\mathbb{E}\int_0^T [M]_s^{m-1} d[M]_s. \tag{1.43}$$

The elementary inequality,

$$(a+b)^m \leq 2^{m-1}(a^m + b^m), \tag{1.44}$$

for any $a, b \geq 0$, applied to the left-hand side of (1.43) then yields the lower bound

$$\mathbb{E}M_T^{2m} \geq \left[\epsilon^{m-1}2^{1-m}(1+\epsilon) - \epsilon^m\right) \mathbb{E}[M]_T^m, \tag{1.45}$$

Chosing e.g. $\epsilon = 2^{-m+1}$ gives

$$\mathbb{E}M_T^{2m} \geq 2^{-m^2+1}\mathbb{E}[M]_T^m. \tag{1.46}$$

To get an upper bound, we deduce from (1.41) the alternative bound

$$2^{m-1}\left(\epsilon^m \mathbb{E}[M]_T^m + M_T^{2m}\right) \geq m(1+\epsilon)\epsilon^{m-1}\mathbb{E}\int_0^T |M_s|^{2m-2} d[M]_s. \tag{1.47}$$

Solving this for $\mathbb{E}|M_T|^{2m}$ gives

$$\mathbb{E}M_T^{2m} \leq \epsilon^m \left(\frac{(1+\epsilon)2^{1-m}}{2m-1} - 1\right)^{-1} \mathbb{E}[M]_T^m, \tag{1.48}$$

and with the choice $\epsilon = 2^{m-1}(2m-1)$,

$$\mathbb{E}M_T^{2m} \leq 2^{(m-1)^2}(2m-1)^m \mathbb{E}[M]_T^m, \tag{1.49}$$

Replacing M_T by $\sup_{t\leq T}$ can be done using Doob's max-inequalities and modifying the constants in the upper bound. This proves the lemma for $m > 1$. The arguments for the other cases are of similar nature and will be skipped. □

It is now clear how to obtain (1.36):

$$\mathbb{E}\left\|X_t^{(n)} - X_s^{(n)}\right\|^{2m} \leq \mathbb{E}\left\|\int_s^t b_n(u, X_u^{(n)})du\right\|^{2m} \tag{1.50}$$

$$+\mathbb{E}\left\|\int_s^t \sigma_n(u, X_u^{(n)})dB_u\right\|^{2m} \tag{1.51}$$

$$\leq (t-s)^{2m}\mathbb{E}\sup_{u\in[s,t]}\left\|b_n(u, X_u^{(n)})\right\|^{2m} \tag{1.52}$$

$$+\mathbb{E}\left(\int_s^t \left\|\sigma_n(u, X_u^{(n)})\right\|^2 du\right)^m \tag{1.53}$$

$$\leq C(m)(t-s)^m \tag{1.54}$$

Then Prohorov's theorem implies that the sequence is conditionally compact, so that we can at least extract a convergent subsequence. Hence we may assume that $P^{(n)}$ converges weakly to some probability measure P^*. We want to show that the process whose law is P^* solves the martingale problem for the operator G.

For $f \in C_0^2(\mathbb{R}^d)$, one checks that $G^{(n)}f(y) \to Gf(y)$ uniformly for y in compact subsets of Wiener space (recall the Arzelà-Ascoli theorem that states that conditionally compact subsets of this space are charcterized by boundedness and continuity modules; clearly if y is continuous, $f \in C^2$, since σ and b are continous functions, $(G_t^{(n)} f)(y)$ converges to $Gf(y_t))$, and from this one can use arguments like in the proof of Lemma (??). □

Remark 1.3.1 Note that we cheat a little here. Namely, the operators G^n and the form of the approximating integral equations are more general than what we have previously assumed in that the coefficients $b^{(n)}(t, y)$ and $\sigma^{(n)}(t, y)$ depend on the past of the function y and not only on the value of y at time t. There is, however, no serious difficulty in generalising the entire theory to that case. The only crucial property that needs to be maintained is that the coefficients remain progressive processes with respect to the filtration $\mathcal{F}_t$.

Remark 1.3.2 The preceeding theorem can be extended rather easily to the case when b and σ are time-dependent, and even to the case when they are bounded, continuous progressive functionals.

A uniqueness result is interestingly tied to a Cauchy problem.

Lemma 1.3.10 *If for every $f \in \mathbb{C}_0^\infty(\mathbb{R}^d)$ the Cauchy problem*

$$\begin{aligned} \frac{\partial u(t,x)}{\partial t} &= (Gu)(t,x), \quad (t,x) \in (0,\infty) \times \mathbb{R}^d \\ u(0,x) &= f(x), \quad x \in \mathbb{R}^d \end{aligned} \tag{1.55}$$

has a solution in $C([0,\infty) \times \mathbb{R}^d) \cap C^{(1,2)}((0,\infty) \times \mathbb{R}^d)$ that is bounded in any strip $[0,T] \times \mathbb{R}^d$, then any two solutions of the martingale problem for G with the same initial distribution have the same finite dimensional distributions.

Proof. Given the solution u let $g(t,x) \equiv u(T-t,x)$. Then g solves, for $0 \le t \le T$,

$$\begin{aligned} \frac{\partial g(t,x)}{\partial t} + (G_s g)(t,x) = 0, \quad (t,x) \in (0,\infty) \times \mathbb{R}^d \\ g(T,x) = f(x), \quad x \in \mathbb{R}^d \end{aligned} \tag{1.56}$$

Then it follows from (1.24) that $g(t, X_t)$ is a local martingale for any solution of the martingale problem. Hence

$$\mathbb{E}_x f(X_T) = \mathbb{E}_x g(T, X_T) = \mathbb{E}_x g(0, X_0) = g(0,x), \tag{1.57}$$

is the same for any solution. This implies uniqueness of the one-dimensional distributions. □

Now Theorem **??** implies immediately the following corollary:

Corollary 1.3.11 *Under the assumptions of the preceeding lemma, weak uniqueness holds for the SDE corresponding to the generator G.*

1.4 Weak solutions from Girsanov's theorem

Girsanov's theorem **??** provides a very efficient and explicite way of constructing weak solutions of certain SDE's.

Theorem 1.4.12 *Consider the stochastic differential equation*

$$dX_t = b(t, X_t) + dB_t, \quad 0 \leq t \leq T, \tag{1.58}$$

for fixed T. Assume that $b : [0, T] \times \mathbb{R}^d$ is measurable and satisfies, for some $K < \infty$,

$$\|b(t, x)\| \leq K(1 + \|x\|). \tag{1.59}$$

Then for any probability measure μ on $\mathbb{R}^d$ there exists a weak solution of (1.58) with initial law μ.

Proof. Let X be a family of Brownian motions starting in $x \in \mathbb{R}$ under laws P_x. Then

$$Z_t \equiv \exp\left(\int_0^t b(s, X_s) \cdot dX_s - \frac{1}{2}\int_0^t \|b(s, X_s)\|^2 ds\right) \tag{1.60}$$

is a martingale under P_x. Thus Girsanov's theorem says that under the measure $\mathbb{Q}_x$ such that $\frac{d\mathbb{Q}_x}{d\mathbb{P}_x} = Z_T$, the process

$$W_t \equiv X_t - X_0 - \int_0^t b(s, X_s) ds \tag{1.61}$$

for $0 \leq t \leq T$ is a Brownian motion starting in 0. Thus we have a pair (X_t, W_t) such that

$$X_t = X_0 + \int_0^t b(s, X_s) ds + W_t, \tag{1.62}$$

holds for $0 \leq t \leq T$, and W_t is a Brownian motion, under $\mathbb{Q}_x$. This shows that we have a weak solution of (1.58). ☐

A complementary result also provided criteria for uniqueness in law.

Theorem 1.4.13 *Assume that we have weak solutions $(X^{(i)}, W^{(i)})$, $i = 1, 2$, on filtered spaces $(\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbb{P}^{(i)}, \mathcal{F}_t^{(i)})$, of the SDE (1.4.12) with the same initial distribution. If*

$$\mathbb{P}^{(i)}\left[\int_0^T \|b(t,X_t^{(i)}\|^2 dt < \infty\right] = 1, \tag{1.63}$$

for $i = 1, 2$, then $(X^{(1)}, W^{(1)})$ and $(X^{(2)}, W^{(2)})$ have the same distribution under their respective probability measures $\mathbb{P}^{(i)}$.

Proof. Define stopping times

$$\tau_k^{(i)} \equiv T \wedge \inf\left\{0 \le t \le T : \int_0^t \|b(t,X_t^{(i)}\|^2 dt = k\right\}. \tag{1.64}$$

We define the martingales

$$\xi_t^{(k)}(X^{(i)}) \equiv \exp\left(-\int_0^{t\wedge\tau_k^{(i)}} b(s,X_s^{(i)})dW_s^{(i)} - \frac{1}{2}\int_0^{t\wedge\tau_k^{(i)}} \|b(s,X_s^{(i)})\|^2 ds\right), \tag{1.65}$$

and the corresponding transformed measures $\widetilde{\mathbb{P}}_k^{(i)}$. Then by Girsanov's theorem,

$$X_{t\wedge\tau_k^{(i)}}^{(i)} \equiv X_0^{(i)} + \int_0^{\tau_k^{(i)}} b(s,X_s^{(i)})ds + W_{t\wedge\tau_k^{(i)}}^{(i)} \tag{1.66}$$

is a Brownian motion with unital distribution μ, stopped at $\tau_k^{(i)}$. In particular, these processes have the same law for $i = 1, 2$. Now the $W^{(i)}$ and the stopping times $\tau_k^{(i)}$ can be expressed in terms of these processes, and probabilities of events of the form

$$\{((X_{t_1}^{(i)}, W_{t_1}^{(i)}), \dots (X_{t_n}^{(i)}, W_{t_n}^{(i)})) \in \Gamma, \tau_k^{(i)} = t_n\},$$

for any collections $t_1 < t_2 < \cdots < t_n$ thus have the same probabilities. Passing to the limit $k \uparrow \infty$ using that due to our assumption, $\mathbb{P}^{(i)}[\tau_k^{(i)} = T] \to 1$ we get uniqueness in law for the entire time interval $[0, T]$. □

1.5 Large deviations

In this section we will give a short glimpse in what is know as the *theory of large deviations large deviations* in the context of simple diffusions. I will emphasize the use of Girsanov's theorem and skid over numerous other interesting issues. There are many nice books on large deviation theory, in particular [3, 4, 6].

We begin with a discussion of *Schilder's theorem* for Brownian motion.

A we know very well, a Brownian motion B_t starting at the origin will, at time t, typically be found at a distance not greater than $\sqrt{t}$ from the origin, in particular, B_t/t converges to zero a.s. We will be interested in computing the probabilities that the BM follows an exceptional path that lives on the sale t. To formalize this idea, we fix a time scale T (which we might also call $1/\varepsilon$), and a smooth path $\gamma : [0,1] \to \mathbb{R}^d$. We want to estimate

$$\mathbb{P}\left[\sup_{0\le s\le 1} \|T^{-1}B_{sT} - \gamma(s)\| \le \epsilon\right]. \tag{1.67}$$

It will be convenient to adopt the notation $\|f\|_\infty \equiv \sup_{0\le s\le 1} \|f(s)\|$. We will first prove a lower bound on the probabilities of the form (1.67).

Lemma 1.5.14 *Let B be Brownian motion, set $B^T_s \equiv T^{-1}B_{Ts}$, and let γ be a smooth path in $\mathbb{R}^d$ starting in the origin. Then*

$$\lim_{\epsilon\downarrow 0}\lim_{T\uparrow\infty} T^{-1}\ln\mathbb{P}\left[\|B^T - \gamma\|_\infty \le \epsilon\right] \ge -I(\gamma) \equiv -\frac{1}{2}\int_0^1 \|\dot\gamma(s)\|^2 ds. \tag{1.68}$$

Proof. For notational simplicity we consider the case $d = 1$ only. Note that $B^T_s = T^{-1}B_{sT}$ has the same distribution as $T^{-1/2}B_s$. Thus we must estimate the probabilities

$$\mathbb{P}\left[\sup_{t\le 1}\|B_t - \sqrt{T}\gamma(t)\| \le \sqrt{T}\epsilon\right]. \tag{1.69}$$

To do this, we observe that by Girsanov's theorem, the process

$$\widehat{B}_t \equiv B_t - \sqrt{T}\gamma(t) \tag{1.70}$$

is a Brownian motion under the measure $\mathbb{Q}$ defined through

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\sqrt{T}\int_0^t \dot\gamma(s)dB_s - \frac{T}{2}\int_0^t \|\dot\gamma(s)\|^2 ds\right). \tag{1.71}$$

Hence

$$\begin{aligned}
&\mathbb{P}\left[\|B - \sqrt{T}\gamma\|_\infty \le \sqrt{T}\epsilon\right] \qquad (1.72)\\
&= \mathbb{P}\left[\|\widetilde{B}\|_\infty \le \sqrt{T}\epsilon\right]\\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\sqrt{T}\int_0^1 \dot\gamma(s)dB_s + \frac{T}{2}\int_0^1\|\dot\gamma(s)\|^2 ds}\mathbb{1}_{\|\widetilde{B}\|_\infty\le\sqrt{T}\epsilon}\right]\\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\sqrt{T}\int_0^1 \dot\gamma(s)d\widetilde{B}_s - \frac{T}{2}\int_0^1\|\dot\gamma(s)\|^2 ds}\mathbb{1}_{\|\widetilde{B}\|_\infty\le\sqrt{T}\epsilon}\right]\\
&= e^{-\frac{T}{2}\int_0^1\|\dot\gamma\|^2(s)ds}\mathbb{Q}\left[\|\widetilde{B}\|_\infty \le \sqrt{T}\epsilon\right]\mathbb{E}_{\mathbb{Q}}\left[e^{-\sqrt{T}\int_0^1\dot\gamma(s)d\widetilde{B}_s}\middle|\|\widetilde{B}\|_\infty \le \sqrt{T}\epsilon\right]\\
&= e^{-\frac{T}{2}\int_0^1\|\dot\gamma(s)\|^2 ds}\mathbb{P}\left[\|B\|_\infty \le \sqrt{T}\epsilon\right]\mathbb{E}_{\mathbb{P}}\left[e^{-\sqrt{T}\int_0^1\dot\gamma(s)dB_s}\middle|\|B\|_\infty \le \sqrt{T}\epsilon\right].
\end{aligned}$$

Now we may use Jensen's inequality to get that

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}}\left[e^{-\sqrt{T}\int_0^1\dot\gamma(s)dB_s}\middle|\|B\|_\infty\leq\sqrt{T}\epsilon\right] \\ &\geq\exp\left(-\sqrt{T}\mathbb{E}_{\mathbb{P}}\left[\int_0^1\dot\gamma(s)dB_s\middle|\|B\|_\infty\leq\sqrt{T}\epsilon\right]\right)=1.\end{aligned}\tag{1.73}$$

On the other hand, it is easy to see, using e.g. the maximum inequality, that, for any $\epsilon>0$,

$$\lim_{T\uparrow\infty}\mathbb{P}\left[\|B\|_\infty\leq\sqrt{T}\epsilon\right]=1.\tag{1.74}$$

Hence,

$$\liminf_{T\uparrow\infty}T^{-1}\ln\mathbb{P}\left[\|B-\sqrt{T}\gamma\|_\infty\leq\sqrt{T}\epsilon\right]\geq-\frac{1}{2}\int_0^t\|\dot\gamma(s)\|^2ds,\tag{1.75}$$

which is the desired lower bound. □

To prove a corresponding upper bound, we proceed as follows. Fix $n\in\mathbb{N}$ and set $t_k=k/n$, $k=0,\dots n$. Set $\alpha\equiv T/n$. Let L be the linear interpolation of B^T_s such that for all t_k, $B^T_{t_k}=L_{t_k}$. Then

$$\begin{aligned}\mathbb{P}\left[\|B^T-L\|_\infty>\delta\right]&\leq\sum_{k=1}^n\mathbb{P}\left[\max_{t_{k-1}\leq t\leq t_k}\|B^T_t-L_t\|>\delta\right]\\ &\leq n\mathbb{P}\left[\max_{0\leq t\leq\alpha}\|B^T_t-L_t\|>\delta\right]\\ &=n\mathbb{P}\left[\max_{0\leq t\leq\alpha}\|B_t-\frac{t}{\alpha}B_\alpha\|>\delta\sqrt{T}\right]\\ &\leq n\mathbb{P}\left[\max_{0\leq t\leq\alpha}\|B_t-\frac{t}{\alpha}B_\alpha\|>\delta\sqrt{T}\right]\\ &\leq n\mathbb{P}\left[\max_{0\leq t\leq\alpha}\|B_t\|>\delta\sqrt{T}/2\right],\end{aligned}$$

where we used that $\max_{0\leq t\leq\alpha}\|B_t-\frac{t}{\alpha}B_\alpha\|>x$ implies that $\max_{0\leq t\leq\alpha}\|B_t\|>x/2$. The last probability can be estimated using the following exponential inequality (for one-dimensional Brownian motion)

$$\mathbb{P}[\sup_{0\leq s\leq t}|B_s|>xt]\leq2\exp\left(-\frac{x^2t}{2}\right)\tag{1.76}$$

which is obtained easily using that $Z_t\equiv\exp\left(\alpha B_t-\frac{1}{2}\alpha^2t\right)$ is a martingale and applying Doob's submartingale inequality (see the proof of the Law of the iterated logarithm in [1]).

This gives us

$$\mathbb{P}\left[\max_{0\leq t\leq\alpha}\|B_t\| > \delta\sqrt{T}/2\right] \leq d\mathbb{P}\left[\max_{0\leq t\leq\alpha}|B_t| > \delta\sqrt{T}/2\sqrt{d}\right] \leq = 2de^{-\frac{\delta^2 n^2 T}{8d}} \quad (1.77)$$

and so

$$\mathbb{P}\left[\|B^T - L\|_\infty > \delta\right] \leq n2e^{-\frac{\delta^2 n^2 T}{8d}} \quad (1.78)$$

which can be made as small as desired by choosing n large enough.

The simplest way to proceed now is to estimate the probability that the value of the *action functional*, I, on L, has an exponential tail with rate T, i.e. that, for n large enough,

$$\limsup_{T\uparrow\infty} T^{-1}\ln\mathbb{P}\left[I(L)\geq\lambda\right]\leq\lambda. \quad (1.79)$$

This is proven easily using the exponential Chebyshev inequality, since

$$I(L) = \frac{n}{2}\sum_{k=1}^{n}\left\|B^T_{t_{k+1}} - B^T_{t_k}\right\|^2 = \frac{1}{2T}\sum_{i=1}^{dn}\eta_i^2$$

where η_i are iid standard normal random variables. But

$$\mathbb{E}e^{\rho\eta_i^2}\leq C_\lambda\leq\infty,$$

for all $\rho<1$, and so

$$\mathbb{P}\left[\frac{1}{2T}\sum_{i=1}^{dn}\eta_i^2>\lambda\right]\leq e^{-\rho\lambda T}\mathbb{E}e^{\rho\sum_{i=1}^{nd}\eta_i^2/2} \leq e^{-\rho\lambda T}C_\rho^{nd} \quad (1.80)$$

for all $\rho<1$, and so (1.79) follows, for any n.

We can deduce from the two estimates the following version of the upper bound:

Proposition 1.5.15 *Let $K_\lambda\equiv\{\phi : I(\phi)\leq\lambda\}$. Then*

$$\limsup_{T\uparrow\infty} T^{-1}\ln\mathbb{P}\left[\mathrm{dist}(\mathrm{B^T},\mathrm{K_\lambda})\geq\delta\right]\leq-\lambda. \quad (1.81)$$

Clearly the meaning of this proposition is that the probability to find a Brownian that is not near a path whose action is less than λ has probability less than $\exp(-\lambda T)$.

The two bounds, together with the fact that the levels sets K_λ (of I are compact (a fact we will not prove), imply the usual formulation of a *large deviation principle*:

Theorem 1.5.16 *For any Borel set $A \subset W$,*

$$-\inf_{\gamma \in \operatorname{int} A} I(\phi) \leq \liminf_{T \uparrow \infty} T^{-1} \ln \mathbb{P}\left[B^T \in A\right] \leq \limsup_{T \uparrow \infty} T^{-1} \ln \mathbb{P}\left[B^T \in A\right] \leq -\inf_{\gamma \in \bar{A}} I(\phi), \tag{1.82}$$

where $\operatorname{int} A$ *and* $\bar{A}$ *denote the interior respectively closure of* A.

The next step will be to pass to an analogous result for the solution of the SDE (1.58) with a scaled down Brownian term. i.e. we want to consider the equation

$$X_t = T^{-1/2} B_t + \int_0^t b(X_s) ds. \tag{1.83}$$

(for notational simplicity we take zero initial conditions). The easiest (although somewhat particular) way to do this is to construct the map $F : W \to W$, as

$$F(\gamma) = f, \tag{1.84}$$

where f is the solution of the integral equation

$$f(t) = \int_0^t b(f(s)) ds + \gamma(t). \tag{1.85}$$

We may use Gronwall's lemma to show that this mapping is continuous. Then $X = F(B^T)$, and

$$\mathbb{P}[X \in A] = \mathbb{P}[B^T \in F^{-1}(A)]. \tag{1.86}$$

Hence, since the continuous map maps open/reps. closed sets in open/resp. closed sets, we can use LDP for Brownian motion to see that

$$\mathbb{P}[X \in A] \leq \sup_{\gamma \in F^{-1}(\bar{A})} I(\gamma) = \sup_{F(\gamma) \in \bar{A}} I(\gamma) = \sup_{\gamma \in \bar{A}} I(F^{-1}(\gamma)), \tag{1.87}$$

and similarly for the lower bound. Hence the process X^T satisfies a large deviation principle with rate function $\widetilde{I}(\gamma) = I(F^{-1}(\gamma))$, and since

$$F^{-1}(\gamma)(t) = \gamma(t) - \int_0^t b(\gamma_s) ds,$$

$$\widetilde{I}(\gamma) = \frac{1}{2} \int_0^1 \|\dot{\gamma}_s - b(\gamma_s)\|^2 ds \tag{1.88}$$

This transportation of a rate function from one family of processes to their image is called sometimes a *contraction principle*.

Properties of action functionals . The rate function $I(\gamma)$ has the form of a classical action functional in Newtonian mechanics, i.e. it is of the form

$$I(\gamma) = \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s), s) ds, \tag{1.89}$$

where the Lagrangian, $\mathcal{L}$, takes the special form

$$\mathcal{L}(\gamma(s), \dot{\gamma}(s), s) = \|\dot{\gamma}(s) - b(\gamma(s), s)\|_2^2. \tag{1.90}$$

The principle of least action in classical mechanics then states that the systems follows a the trajectory of minimal action subject to boundary conditions. This leads to the Euler-Lagrange equations,

$$\frac{d}{dt}\frac{\partial}{\partial\dot{\gamma}}\mathcal{L}(\gamma, \dot{\gamma}, s) = \frac{\partial}{\partial\gamma}\mathcal{L}(\gamma, \dot{\gamma}, s). \tag{1.91}$$

In our case, these take the form

$$\frac{d^2}{dt^2}\gamma(t) = \frac{\partial}{\partial t}b(\gamma(t), t) + b(\gamma(t), t)\frac{\partial}{\partial\gamma(t)}b(\gamma(t), t). \tag{1.92}$$

One can readily identify a special class of solution of this second order equation, namely solutions of the first order equations

$$\dot{\gamma}(t) = b(\gamma(t), t), \tag{1.93}$$

which have the property that they yield absolute minima of the action, $I(\gamma) = 0$. Of course, being first order equations, they admit only one boundary or initial condition.

Typical questions one will ask in the probabilistic context are: what is the probability of a solution connecting a and b in time t. The large deviation principle yields the ansert

$$\mathbb{P}\left[|X_0 - a| \leq d, |X_t - b| \leq \delta\right] \sim \exp\left(-\epsilon^{-1} \inf_{\gamma:\gamma(0)=a,\gamma(t)=b} I(\gamma)\right), \tag{1.94}$$

which leads us to solve (1.92) subject to boundary conditions $\gamma(0) = a, \gamma(t) = b$. In general this will not solve (1.93), and thus the optimal solution will have positive action, and the event under consideration will have an exponentially small probability. On the other hand, under certain conditions one may find a zero-action solution if one does not fix the time of arrival at the endpoint:

$$\begin{aligned} &\mathbb{P}\left[|X_0 - a| \leq d, |X_t - b| \leq \delta, \text{for some}\, t < \infty\right] \\ &\sim \exp\left(-\epsilon^{-1} \inf_{\gamma:\gamma(0)=a,\gamma(t)=b,\text{for some}\, t\infty} I(\gamma)\right). \end{aligned} \tag{1.95}$$

Clearly the infimum will be zero, if the solution of the initial value problem (1.93) with $\gamma(0) = a$ has the property that for some $t < \infty$, $\gamma(t) = b$, or if $\gamma(t) \to b$, as $t \uparrow \infty$.

Exercise. Consider the case of one dimension with $b(x) = -x$. Compute the minimal action for the problem (1.94) and characterize the situations for which a minimal action solution exists.

A particularly interesting question is related to the so called *exit problem*. Assume that we we consider an event as in (1.95) that admits an zero-action path γ, such that $\gamma(0) = a, \gamma(T) = (b)$. Define the time reversed path $\hat{\gamma}(t) \equiv \gamma(T - t)$. Clearly $\frac{d}{dt}\hat{\gamma}(t) = -\dot{\gamma}(T - t)$. Hence a simple calculation shows that

$$I(\hat{\gamma}) - I(\gamma) = 2\int_0^T b(\gamma(s)) \cdot \dot{\gamma}(s) ds = \int_\gamma b(\gamma) d\gamma. \tag{1.96}$$

Let us now specialize to the case when the vector field b is the gradient of a potentia, $b(x) = \nabla F(x)$. Then

$$\int_\gamma b(\gamma) d\gamma = F(\gamma(T)) - F(\gamma(0)) = F(b) - F(a). \tag{1.97}$$

Hence

$$I(\hat{\gamma}) = I(\gamma) + F(b) - F(a), \tag{1.98}$$

If $I(\gamma) = 0$, then $I(\hat{\gamma}) = F(b) - F(a)$, and this is the miminal possible value for any curve going from b to a. This shows the remarkable fact that the most likely path going uphill against a potential is the time-reversal of the solution of the gradient flow. Estimates of this type are the basis of the so-called Wentzell-Freidlin theory [6].

1.6 SDE's from conditioning: Doob's h-transform

With Girsanov's theorem we have seen that drift can be produced through a change of measure. Another important way in which drift can arise is conditioning. We have seen this already in the case of discrete time Markov chains. Again we will see that the martingale formulation plays a useful rôle.

As in the discrete case, the key result is the following.

Theorem 1.6.17 *Let X be a Markov process, i.e. a solution of the martingale problem for an operator G and let h be a strictly positive harmonic function. Define the measure $\mathbb{P}^h$ s.t. for any $\mathcal{F}_t$ measurable random variable,*

$$\mathbb{E}_x^h[Y] = \frac{1}{h(x)}\mathbb{E}_x[h(X_t)Y]. \tag{1.99}$$

Then $\mathbb{P}^h$ is the law of a solution of the martingale problem for the operator G^h defined by

$$(G^h f)(x) \equiv \frac{1}{h(x)}(Lhf)(x). \tag{1.100}$$

As an important example, let us consider the case of Brownian motion in a domain $D \subset \mathbb{R}^d$, killed in the boundary of D. We will assume that D is a harmonic function in D and let τ_D the first exit time of D. Then

$$G^h = \frac{1}{2}\Delta + \frac{\nabla h}{h} \cdot \nabla,$$

and hence under the law $\mathbb{P}^h$, the Brownian motion becomes the solution of the SDE

$$dX_t = \frac{\nabla h(X_t)}{h(X_t)}dt + dB_t. \tag{1.101}$$

On the other hand, we have seen that, if h is the probability of some event, e.g.

$$H(x) = \mathbb{P}_x[X_{\tau_D} \in A],$$

for some $A \in \partial D$, then

$$\mathbb{P}^h[\cdot] = \mathbb{P}[\cdot|X_{\tau_D} \in A] \tag{1.102}$$

This means that the Brownian motion conditioned to exit D in a given place can be represented as a solution of an SDE with a particular drift. For instance, let $d = 1$, and let $D = (0, R)$. Consider the Brownian motion conditioned to leave D at R. It is elementary to see that

$$\mathbb{P}_x[X_{\tau_D} = R] = x/R.$$

Thus the conditioned Brownian motion solves

$$dX_t = \frac{1}{X_t}dt + dB_t. \tag{1.103}$$

Note that we can take $R \uparrow \infty$ without changing the SDE. Thus, the solution of (1.103) is Brownian motion conditioned to never return to the origin. This is understandable, as the strength of the drift away form zero goes to infinity (quickly) near 0. Still, it is quite a remarkable fact that conditioning can be exactly reproduced by the application of the right drift.

Note that the process defined by (1.103) has also another interpretation. Let $W = (W_1, \dots, W_d)$ be d-dimensional Brownian motion. Set $R_t = \|W(t)\|_2$. Then R_t is called the *Bessel process* with dimension d. It turns out that this process is also the (weak) solution of a stochastic differential equation, namely:

Proposition 1.6.18 *The Bessel process in dimension d is a weak solution of*

$$dR_t = \frac{d-1}{2R_t} + dB_t. \tag{1.104}$$

Proof. Let us first construct the Brownian motion B_t from the d-dimensional Brownian motions W as follows. Set

$$B_t^{(i)} \equiv \int_0^t \frac{W_i(s)}{R_s} dW_i(s)$$

and

$$B_t \equiv \sum_{i=1}^d B_t^{(i)}.$$

The processes in $B_t^{(i)}$ are continuous square integrable martingales since

$$\mathbb{E}\left(\int_0^t \frac{W_i(s)}{R_s} dW_i(s)\right) = \mathbb{E}\int_0^t \left(\frac{W_i(s)}{R_s}\right)^2 ds \leq t;$$

Moreover the

$$[B]_t = \sum_{i,j} [B^{(i)}, B^{(j)}]_t = \sum_i \int_0^t \left(\frac{W_i(s)}{R_s}\right)^2 ds = t,$$

so by Lévy's theorem, B is Brownian motion. Thus we can write (1.104) as

$$dR_t = \sum_i \frac{1}{R_t} dW_i(t) + \frac{1}{2}\frac{d-1}{R_t} dt.$$

But this is precisely the result of applying Itô's formula to the function $f(W) = \|W\|_2$. Note that this derivation is slightly sloppy, since the function f is not differentiable at zero, but the result is correct anyway (for a fully rigorous proof see e.g. [10], Chapter 3.3). □

In particular, we see that the one-dimensional Brownian motion conditioned to stay strictly positive for all positive times is the 3-dimensional Bessel process. This shows in particular that in dimension 3 (and trivially higher), Brownian motion never returns to the origin. Looking at the SDE describing the Bessel process, one might guess that the value of d, as soon as $d > 1$, should not be so important for this property, since there is always a divergent drift away from 0. We will now show that this is indeed the case.

Proposition 1.6.19 *Let R_t be the solution of the SDE (1.104) with $d \geq 2$ and initial condition $R_0 = r \geq 0$. Then*

$$\mathbb{P}\left[\forall t > 0 : R_t > 0\right] = 1. \tag{1.105}$$

Proof. Let first $r > 0$. Let

$$\tau_k \equiv \inf\left\{t \geq 0 : R_t = k^{-k}\right\},$$

$$\sigma_k \equiv \inf\left\{t \geq 0 : R_t = k\right\}$$

and $T_k \equiv \tau_k \wedge \sigma_k \wedge n$. Now use Itô's formula for the function $h(R_{T_k})$, where $h(x) = \frac{1}{1-\alpha}x^{-\alpha+1}$, if $(d-1)/2 = \alpha \neq 1$, and $h(x) = \ln x$, if $d = 2$. The point is that h is a harmonic function w.r.t. the operator $G = \frac{d^2}{dx^2} + \alpha \frac{1}{x}\frac{d}{dx}$, and hence $h(R_t)$ is a martingale. Moreover, since T_k is a bounded stopping time, it follows that

$$\mathbb{E}_r\left[h\left(R_{T_k}\right)\right] = h(r). \tag{1.106}$$

Finally,

$$\mathbb{E}_r\left[h\left(R_{T_k}\right)\right] = h(k)\mathbb{P}_r[T_k = \sigma_k] + h(k^{-k})\mathbb{P}_r[T_k = \tau_k] + h(B_n)\mathbb{P}_r[T_k = n]. \tag{1.107}$$

Hence

$$\mathbb{P}_r[T_k = \tau_k] \leq \frac{h(r)}{h(k^{-k})} \leq \begin{cases} k^{-(\alpha-1)k} r^{-\alpha+1}, & \text{if } d \neq 2, \\ \frac{\ln r}{k \ln k}, & \text{if } d = 2. \end{cases} \tag{1.108}$$

Now all what is left to show is that $\mathbb{P}[n < \tau_k \wedge \sigma_k] \downarrow 0$, as $n \uparrow \infty$. But this is obvious from the fact that $R_t \geq r + B_t$, and $\mathbb{P}_0[B_t \leq n]$ tends to zero as $n \uparrow \infty$. Hence,

$$\lim_{n\uparrow\infty} \mathbb{P}_r[T_k = \tau_k] = \mathbb{P}_r[\tau_k < \sigma_k]$$

which in turn tends to zero with k. Now set $\tau \equiv \inf\{t > 0 : B_t = 0\}$. For every k, $\tau_k < \tau$, so that, again since $\sigma_k \uparrow \infty$, a.s.,

$$\mathbb{P}\left[\tau < \infty\right] \leq \lim_{k\uparrow\infty} \mathbb{P}_r\left[\tau < \sigma_k\right] \leq \lim_{k\uparrow\infty} \mathbb{P}_r\left[\tau_k < \sigma_k\right] = 0. \tag{1.109}$$

This proves the case $r > 0$. For $r = 0$, just use that, by the strong Markov property, for any $\epsilon > 0$,

$$\mathbb{P}_0\left[R_t > 0, \forall \epsilon < t < \infty\right] = \mathbb{E}_0 \mathbb{P}_{B_\epsilon}\left[[R_t > 0, \forall 0 < t < \infty\right] = 1, \qquad (1.110)$$

since $\mathbb{P}_0[R_\epsilon > 0] = 1$. Finally let $\epsilon \downarrow 0$ to complete the proof. □

Remark 1.6.1 The method used above is important beyond this example. It has a useful generalization in that one need not chose for h a harmonic function. In fact all goes through if h is chosen to be super-harmonics. In many situations it may be difficult to find a harmonic function, whereas one may well be able to to find a useful super-harmonic function.

2

SDE's and partial differential equations

Already in the context of discrete time Markov processes [1] we have seen that the martingale problem formulation of Markov processes leads to an interesting connection between probability theory and linear boundary value problems. In the case of stochastic differential equations, this connections become even more profound leads to the connection between diffusion processes and potential theory which can be seen as one of the mathematical highlights of stochastic analysis. The classical case relates only to Brownian motion, but the extension to more general second order stochastic differential equations is quite straight-forward. Note that we will henceforth switch notation and denote generators by $\mathcal{L}$ rather than G, as the letter G will be needed to denote Green's functions. For analytic background on elliptic partial differential equations the standard reference is the textbook [8] by Gilbarg and Trudiger.

2.1 The Dirichlet problem

We consider the stochastic differential equation of the previous chapter with time-independent drift and dispersion matrix

$$dX_t = b(X_t) + \sigma(X_t)dB_t, \tag{2.1}$$

in $\mathbb{R}^d$. We have seen that the (weak) solutions of this equation are a strong Markov process with generator whose restriction to $C^2(\mathbb{R}^d)$ is given by

$$\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i}, \tag{2.2}$$

where the *diffusion matrix* a is given by

$$a_{ij}(x) = \sum_{k=1}^{d} \sigma_{ik}(x)\sigma_{kj}(x). \tag{2.3}$$

In the sequel we will always assume that the dispersion matric σ is non-degenerate and hence the diffusion matrix is strictly positive, i.e. for all $x \in \mathbb{R}^d$, $a(x)$ defines a strictly positive quadratic form. If a is strictly positive, then the operator $\mathcal{L}$ is called *elliptic*. If in some domain $D \subset \mathbb{R}^d$,

$$\sum_{i,j} a_{ij}(x)\xi_i\xi_j \geq \delta\|\xi\|^2,$$

for all $x \in D$, then we call $\mathcal{L}$ *uniformly elliptic* in D.

The classical *Dirichlet problem* associated with an elliptic operator $\mathcal{L}$ and a domain, D, is described as follows (we assume here that D is bounded). Let $D \subset \mathbb{R}$ and continuous functions $g\bar{D} \to \mathbb{R}$, $k : \bar{D} \to [0, \infty)$, and $u : \partial D \to \mathbb{R}$ be given. can we find a continuous function $f : \bar{D} \to \mathbb{R}$, such that

$$-(\mathcal{L}f)(x) + k(x)f(x) = g(x), \forall x \in D \tag{2.4}$$

$$f(x) = u(x), \forall x \in \partial D. \tag{2.5}$$

Remark 2.1.1 The Dirichlet problem can also be posed if u is not a continuous function on the boundary of D. In that case the condition that f be continuous on $\bar{D}$ must be replaced by that condition that, for all $x \in \partial D$, whenever a sequence $x_n \in D$ converges to x, then $f(x_n) \to u(x)$.

It is rather straightforward to see that the existence of a solution of such a problem implies a stochastic representation. Namely:

Theorem 2.1.1 *Assume that f solves the Dirichlet problem above, and let X be a weak solution of the SDE (2.1). Let $\tau_D \equiv \inf\{t \geq 0 : X_t \notin D\}$. If*

$$\mathbb{E}_x \tau_D < \infty, \quad \forall x \in D, \tag{2.6}$$

then

$$f(x) = \mathbb{E}_x \left[f(X_{\tau_D}) \exp\left(-\int_0^{\tau_D} k(X_s)ds \right) + \int_0^{\tau_D} g(X_s) \exp\left(-\int_0^t k(X_s)ds \right) dt \right] \tag{2.7}$$

Proof. The key to this result is the following lemma:

Lemma 2.1.2 *Let $\mathcal{F}_t$ be a filtration and X an adapted process. Let $f, g, k \in B(S)$. Then $f(X_t) - \int_0^t (\mathcal{L}f)(X_s)ds$ is a martingale if and only if*

$$M_t \equiv e^{-\int_0^t k(X_s)ds} f(X_t) + \int_0^t e^{-\int_0^s k(X_r)dr} \left(k(X_s)f(X_s) - (\mathcal{L}f)(X_s) \right) ds \tag{2.8}$$

is a martingale.

Proof. The proof of this lemma follows from Proposition 4.1.1 in [2] . Just choose for $M(t)$ the martingale $f(X_t) - \int_0^t (\mathcal{L}f)(X_s)ds$ and for $V(t)$ the process $\exp\left(-\int_0^t k(X_s)ds\right)$. Then it is a slightly tedious but straightforward computation (that uses Fubini's theorem at the right moment) to show that the expression in (2.8) is of the form $V(t)M(t) - \int_0^t M(t)dV(t)$ and hence a martingale. □

We use Lemma 2.1.2 with $g = \mathcal{L}f$ where f solves the Dirichlet problem. Now since $\mathbb{E}_x\tau_D < \infty$ by assumption, we get from the optional sampling theorem that

$$\mathbb{E}_x M_{\tau_D} = \mathbb{E}_x M_0. \tag{2.9}$$

But $\mathbb{E}_x M_0 = f(x)$, while, due to the fact that

$$\begin{aligned} \mathbb{E}_x M_{\tau_D} &= \mathbb{E}_x \Big[e^{-\int_0^{\tau_D} k(X_s)ds} f(X_D) \\ &\quad + \int_0^{\tau_D} e^{-\int_0^s k(X_r)dr} \left(k(X_s)f(X_s) - \mathcal{L}f(X_s) \right) ds \Big] \\ &= \mathbb{E}_x \left[e^{-\int_0^{\tau_D} k(X_s)ds} f(X_D) + \int_0^{\tau_D} e^{-\int_0^s k(X_r)dr} g(X_s) ds \right] \end{aligned} \tag{2.10}$$

which is what we claimed. □

It is interesting to note that the finiteness of the expectation of the exit time τ_D is quite easily ensured (for bounded domains) from a rather weak ellipticity condition.

Lemma 2.1.3 *Let D be open and bounded in $\mathbb{R}^d$, and assume that for some $1 \leq \ell \leq d$,*

$$\min_{x \in \bar{D}} a_{\ell\ell}(x) > 0. \tag{2.11}$$

Then $\mathbb{E}_x \tau_D < \infty$, for all $x \in D$.

Proof. Set $a = \min_{x\in\bar{D}} a_{\ell\ell}(x)$, $b \equiv \max_{x\in\bar{D}} \|b(x)\|$, and $q \equiv min_{x\in\bar{D}} x_\ell$. Let $\nu > 2b/a$. Consider the smooth function $h(x) = -\mu e^{\nu x_\ell}$, with $\mu > 0$ to be chosen later. Clearly

$$-\mathcal{L}h(x) = \mu e^{\nu x_\ell} \left(\frac{1}{2}\nu^2 a_{\ell\ell}(x) + \nu b_\ell(x) \right) \geq \frac{1}{2}\mu\nu a e^{\nu b} \left(\nu - 2b/a\right).$$

Now we can chose μ such that the right-hand side is larger than 1, and so $\mathcal{L}h(x) \leq -1$, for all $x \in D$. But

$$h(X_{t\wedge\tau_D}) - \int_0^{t\wedge\tau_D} \mathcal{L}h(X_s)ds$$

is a martingale, and so

$$-\mathbb{E}_x \int_0^{t\wedge\tau_D} \mathcal{L}h(X_s)ds = h(x) - \mathbb{E}_x h(X_{t\wedge\tau_D}),$$

or

$$h(x) - \mathbb{E}_x h(X_{t\wedge\tau_D}) \geq \mathbb{E}_x(t \wedge \tau_D)$$

and hence

$$\mathbb{E}_x(t \wedge \tau_D) \leq \max_{y\in\bar{D}} |h(y)| < \infty.$$

Passing to the limit $t \uparrow \infty$ implies $\mathbb{E}_x \tau_D < \infty$. □

The previous results give a stochastic representation formula for solutions of the Dirichlet problem, assuming that a solution to the Dirichlet problem and a weak solution of the SDE exist. One may ask whether one can use this representation to prove the existence of solutions of the Dirichlet problem? We will address this question in the simpler context of Brownian motion.

Brownian motion and potential theory. Let us now consider the setting where $\mathcal{L} = \frac{1}{2}\Delta$ and X_t is Brownian motion in $\mathbb{R}^d$. Let us begin with the simplest boundary value problem

$$\begin{aligned} \Delta f(x) &= 0, \quad x \in D, \\ f(x) &= u(x), \quad x \in \partial D. \end{aligned} \tag{2.12}$$

We assume that u is bounded and continuous. From the theorem above, an obvious candidate solution is

$$f(x) = \mathbb{E}_x u(B_{\tau_D}). \tag{2.13}$$

Now f clearly satisfies the boundary conditions, and it is also not hard to show hat $\Delta f(x) = 0$ for $x \in D$. There are various ways ti show this.

Note first that we can write, with P_t the semi-group corresponding to the Brownian motion starting at x that

$$\begin{aligned}(P_t f)(x) &= \mathbb{E}_x\left[\left(\mathbb{1}_{\tau_D>t} + \mathbb{1}_{\tau_D\leq t}\right)\mathbb{E}_{X_t}[u(B_{\tau_D})]\right] \\ &= \mathbb{E}_x\left[\mathbb{1}_{\tau_D>t}\mathbb{E}_{X_t}[u(B_{\tau_D})]\right] + \mathbb{E}_x\left[\mathbb{1}_{\tau_D\leq t}\mathbb{E}_{X_t}[u(B_{\tau_D})]\right].\end{aligned} \tag{2.14}$$

Now in the first term we can use the Markov property to see that

$$\mathbb{E}_x\left[\mathbb{1}_{\tau_D>t}\mathbb{E}_{X_t}[u(B_{\tau_D})]\right] = \mathbb{E}_x[u(B_{\tau_D})] = f(x) \tag{2.15}$$

while the second satisfies the bound

$$\left|\mathbb{E}_x\left[\mathbb{1}_{\tau_D\leq t}\mathbb{E}_{X_t}[u(B_{\tau_D})]\right]\right| \leq \max_{x\in\partial D} u(x)\mathbb{P}_x\left[\tau_D \leq t\right]. \tag{2.16}$$

Using (e.g.) the estimates from (**??**), one can easily show that

$$\lim_{t\downarrow 0} t^{-1}\mathbb{P}_x\left[\tau_D \leq t\right] = 0,$$

for any $x \in D$. This then implies that

$$\frac{1}{2}\Delta f(x) = \lim_{t\downarrow 0} t^{-1}([P_t - \mathbb{1}]f)(x) = 0. \tag{2.17}$$

We see that all that remains to show to establish that f solves the Dirichlet problem is the continuity of f at the boundary of D. [1]

As we will see, the continuity property is linked to regularity properties of the boundary of D.

Definition 2.1.1 Define the stopping time $\sigma_D \equiv \inf\{t > 0 : B_t \in D^c\}$ (note the difference to τ_D when we start the process in the boundary of D!). A point, $z \in \partial D$, is called *regular* , if $\mathbb{P}_z[\sigma_D = 0] = 1$.

Thus a regular point has the property that the Brownian motion starting at it will essentially immediately return to the boundary. An irregular point is one from which Brownian motion can immediately escape into D.

Remark 2.1.2 It follows from the so-called *Blumenthal-Getoor* $0-1$-*law* (Lemma 2.1.4 below) that If a point z is not regular, then $\mathbb{P}_z[\sigma_D = 0] = 0$.

Lemma 2.1.4 *[Blumenthal-Getoor* 0-1-*law] Let B_t be a d-dimensional Brownian motion, starting in x, on a filtered space $(\Omega, \widetilde{\mathcal{F}}, \mathbb{P}_x, \widetilde{\mathcal{F}}_t)$ where $\widetilde{\mathcal{F}}$ is the usual augmentation of the natural filtration, $\mathcal{F}_t$, generated be the Brownian motion. Then, if $F \in \widetilde{\mathcal{F}}_0$, $\mathbb{P}_x[F] \in \{0, 1\}$.*

[1] Clearly continuity is essential: without asking it there is no point in the problem, since it would admit lots of solutions, e.g. zero in D and u on ∂D.

Proof. If $F \in \widetilde{\mathcal{F}}_0$, then F differs from some set $G \in \mathcal{F}_0$ only by a $\mathbb{P}_x$-null set. But since G must be of the form $G = \{B_0 \in A\}$ for some Borel set A, it follows that

$$\mathbb{P}_x[F] = \mathbb{P}_x[G] = \mathbb{1}_A(x) \in \{0, 1\}.$$

□

The following theorem establishes that the Dirichlet problem is solvable (uniquely) for bounded regular domains.

Theorem 2.1.5 *Let $d \geq 2$ and let $z \in \partial D$ be fixed. Then the following statements are equivalent:*

(i) For any bounded measurable function $u : \partial D \to \mathbb{R}$ which is continuous at z,

$$\lim_{D \ni x \to z} \mathbb{E}_x u(B_{\tau_D}) = u(z). \tag{2.18}$$

(ii) z is a regular point for D.

(iii) For all $\epsilon > 0$,

$$\lim_{D \ni x \to z} \mathbb{P}_x[\tau_D > \epsilon] = 0. \tag{2.19}$$

Proof. We first proof that (i) implies (ii). From the remark 2.1.2, we know that if the origin is irregular, then $\mathbb{P}_z[\sigma_D = 0] = 0$. We will use the fact that in $d \geq 2$, the probability that Brownian motion visits any given point is zero, and in particular the probability that it returns to its starting point is zero. Thus, if K_r denotes the ball of radius r around z,

$$\lim_{r \downarrow 0} \mathbb{P}_z[B_{\sigma_D} \in K_r] = \mathbb{P}_z[B_{\sigma_D} = z] = 0.$$

Now fix r such that $\mathbb{P}_z[B_{\sigma_D} \in K_r] < 1/4$ and chose a sequence δ_n, $0 < \delta_n < r$, tending to zero. Let $\tau_n \equiv \inf\{t \geq 0 : \|B_t\| \geq \delta_n\}$. Then $\mathbb{P}_z[\tau_n \downarrow 0] = 1$, and so $\lim_n \mathbb{P}_z[\tau_n < \sigma_D] = 1$. Moreover, on $\{\tau_n < \sigma_D\}$ we have that $B_{\tau_n} \in D$. Thus for n so large that $\mathbb{P}_z[\tau_n < \sigma_D] \geq 1/2$, we have then that

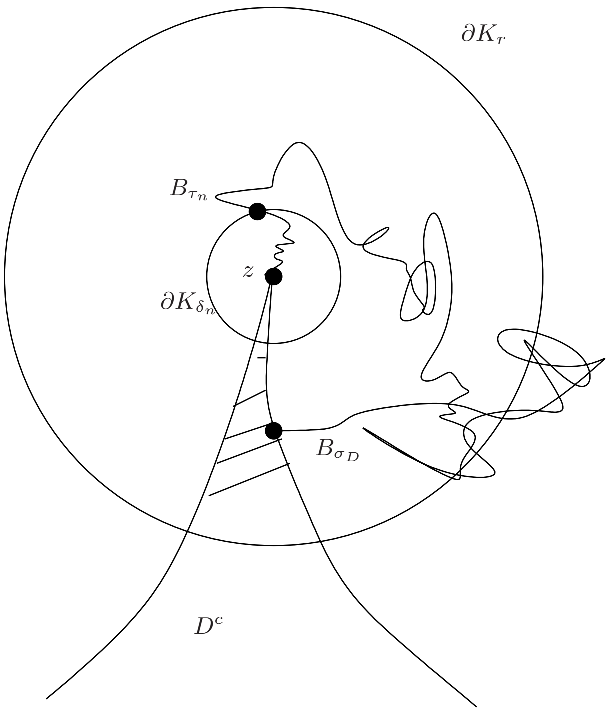

Fig. 2.1. Setting in the proof of (i) implies (ii)

$$
\begin{aligned}
\frac{1}{4} &\geq \mathbb{P}_z[B_{\sigma_D} \in K_r] \geq \mathbb{P}_z[B_{\sigma_D} \in K_r, \tau_n < \sigma_n] \qquad (2.20)\\
&= \mathbb{E}_z[1\!\!1_{\tau_n < \sigma_D} \mathbb{P}_z[B_{\sigma_D} \in K_t | \mathcal{F}_{\tau_n}]\\
&= \int_{D \cap \partial K_{\delta_n}} \mathbb{P}_z[\tau_n < \sigma_D, B_{\tau_n} \in dx] \mathbb{P}_x[B_{\sigma_D} \in K_r]\\
&\geq \inf_{x \in D \cap K_{\delta_n}} \mathbb{P}_x[B_{\tau_D} \in K_r] \int_{D \cap K_{\delta_n}} \mathbb{P}_z[\tau_n < \sigma_D, B_{\tau_n} \in dx]\\
&= \inf_{x \in D \cap \partial K_{\delta_n}} \mathbb{P}_x[B_{\tau_D} \in K_r] \mathbb{P}_z[\tau_n < \sigma_D]\\
&\geq \frac{1}{2} \inf_{x \in D \cap \partial K_{\delta_n}} \mathbb{P}_x[B_{\tau_D} \in K_r].
\end{aligned}
$$

Hence $\mathbb{P}_{x_n}[B_{\tau_D} \in K_r] \leq \frac{1}{2}$ for some $x_n \in D \cap K_{\delta_n}$. Now choose a continuous bounded function, f, with $f(z) = 1$, $f(x) \leq 1$, $x \in K_r$, and

$f(x) = 0$, $x \notin K_r$. For such functions we get

$$\limsup_n \mathbb{E}_{x_n} f(B_{\tau_D}) \leq \limsup_n \mathbb{P}_{x_n}[B_{\tau_D} \in K_r] \leq \frac{1}{2} < 1 = f(z),$$

so that (i) cannot hold. Therefore (i) implies (ii).

Let us now show that (ii) implies (iii). Notice that the function

$$\begin{aligned} g_\delta(x) &\equiv \mathbb{P}_x[B_s \in D; \delta \leq s \leq \epsilon] = \mathbb{E}_x[\mathbb{P}_{B_\delta}[\tau_D > \epsilon - \delta]] \\ &= \int \mathbb{P}_y[\tau_D > \epsilon - \delta]\mathbb{P}_x[B_\delta \in dy] \end{aligned} \tag{2.21}$$

is continuous in x. But

$$g_\delta(x) \downarrow g(x) \equiv \mathbb{P}_x[B_s \in D; 0 < s \leq \epsilon] = \mathbb{P}_x[\sigma_D > \epsilon],$$

as $\delta \downarrow 0$, so that g is upper semi-continuous. This implies that

$$\limsup_{x \to z} P_x[\tau_D > \epsilon] \leq \limsup_{x \to z} P_x[\sigma_D > \epsilon] = \limsup_{x \to z} g(x) \leq g(z) = 0,$$

where the last inequality comes from the regularity of z, i.e. (ii). Thus we have (iii) from (ii).

Finally we show that (iii) implies (i). We start from the observation that $\mathbb{P}_x[\max_{0 \leq t \leq \epsilon} \|B_t - B_0\| < r]$ is independent of x and converges to one as $\epsilon \downarrow 0$. Now

$$\begin{aligned} \mathbb{P}_x[\|B_{\tau_D} - B_0\| < r] &\geq \mathbb{P}_x \left[\left\{ \max_{0 \leq t \leq \epsilon} \|B_t - B_0\| < r \right\} \cap \{\tau_D \leq \epsilon\} \right] \\ &\geq \mathbb{P}_0 \left[\left\{ \max_{0 \leq t \leq \epsilon} \|B_t\| < r \right\} \right] - \mathbb{P}_x \left[\tau_D \leq \epsilon\right]. \end{aligned}$$

When $x \to z$, by (iii) the second term vanishes for all ϵ, and letting $\epsilon \downarrow 0$, the first term tends to one. Thus we get that

$$\lim_{D \ni x \to z} \mathbb{P}_x \left[\|B_{\tau_D} - x\| < r\right] = 1.$$

Thus

$$\begin{aligned} |\mathbb{E}_x f(B_{\tau_D}) - f(z)| &\leq |\mathbb{E}_x f(B_{\tau_D}) - f(z)| \\ &\leq |\mathbb{E}_x \left[\mathbb{1}_{\|B_{\tau_D} - x\| < r} (f(B_{\tau_D}) - f(z)) \right] | \\ &\quad + 2 \max_{y \in \partial D} |f(y)| \mathbb{P}_x[\|B_{\tau_D} - x\| \geq r] \end{aligned} \tag{2.22}$$

Clearly all three terms vanish as $x \to z$ and $r \downarrow 0$ by the continuity at zero and boundedness of f. $\square$

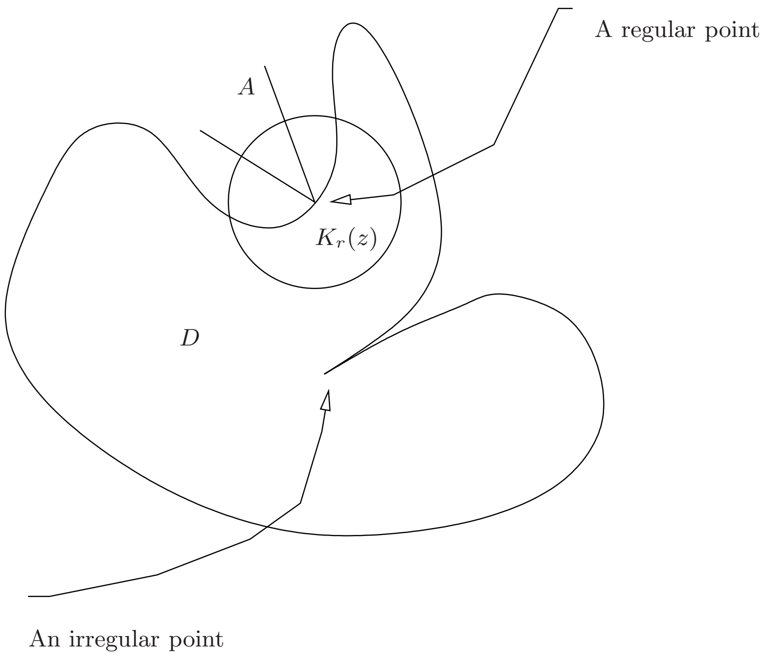

Fig. 2.2. A domain with one irregular point violating the cone-condition

The preceeding theorems imply that the Dirichlet problem has a unique solution if and only if any point in the boundary of D is regular. Otherwise, no solution exists. Moreover, the solution has the stochastic representation (2.14).

Th following proposition gives a sufficient verifiable criteria for regularity.

Proposition 2.1.6 *A point, $z \in \partial D$ is regular if there exists a cone, A, with vertex z, such that, for some $r > 0$, $A \cap K_r(z) \subset D^c$.*

Proof. Let $C > 0$ denote the fraction of the surface of $K_r(z)$ that lies within A. Let $K^{(n)} \equiv K_{r/n}(z)$, and $A_n \equiv A \cap \partial K^{(n)}$. Now $\tau_D = 0$ if $B_{\tau_{K^{(n)}}} \in A_n$ for arbitrary large n, i.e. $\{\tau_D = 0\} \supset \limsup_n \{B_{\tau_{K^{(n)}}} \in A_n\}$. Thus

$$\begin{aligned} \mathbb{P}_z[\tau_D = 0] &\geq \mathbb{P}_z[\limsup_n \{B_{\tau_{K^{(n)}}} \in A_n\}] \\ &\geq \limsup_n \mathbb{P}_z[B_{\tau_{K^{(n)}}} \in A_n] = C > 0. \end{aligned} \tag{2.23}$$

The fact that then $\mathbb{P}_z[\tau_D = 0] = 1$ follows from the fact that the event in question is in $\mathcal{F}_{0+}$ and the Blumenthal-Getoor zero-one law. □

A slightly more abstract criterion is interesting because it involves the notion of a *barrier*.

Definition 2.1.2 Let $D \subset \mathbb{R}^d$ be open and $a \in \partial D$. A continuous function $v : \bar{D} \to \mathbb{R}$ that is harmonic in D, positive in $\bar{D}\backslash\{a\}$, and zero at a is called a *barrier*.

Proposition 2.1.7 *Let D be bounded and $a \in \partial D$. If there exists a barrier at a, then a is regular.*

Proof. Let v be a barrier. Let $f : \partial D \to \mathbb{R}$ and define $M \equiv \sup_{x \in \partial D} |f(x)|$. For any $\epsilon > 0$, we can find $\delta > 0$, such that for $x \in \partial D$ and $|x - a| \leq \delta$, $|f(x) - f(a)| \leq \epsilon$. Choose k such that $kv(x) \geq 2M(x)$, for $x \in \bar{D}$ and $|x - a| \geq \delta$. Then $|f(x) - f(a)| \leq \epsilon + kv(x)$, for all $x \in \partial D$. Thus

$$|\mathbb{E}_x f(B_{\tau_D}) - f(a) \leq \epsilon + k\mathbb{E}_x v(B_{\tau_D}) \leq \epsilon + kv(x),$$

for all $x \in D$. Now since v is continuous and $v(a) = 0$, it follows that

$$\limsup_{d \ni x \to a} |\mathbb{E}_x f(B_{\tau_d}) - f(a)| \leq \epsilon,$$

for all $\epsilon > 0$, hence condition (i) of Theorem 2.1.5 holds and a is regular. □

To show that the discussion of regular points is not empty, let us look at a classical example of a point that is not regular. This is called *Lebesgue's thorn*. Let $d = 3$, and define, for ϵ_n, $n \in \mathbb{N}$ such that $\epsilon_n \downarrow 0$, the sets

$$E \equiv \left\{(x_1, x_2, x_3) : -1 < x_1 < 1; x_2^2 + x_3^2 < 1\right\}; \tag{2.24}$$

$$F_n \equiv \left\{(x_1, x_2, x_3) : 2^{-n} \leq x_1 \leq 2^{-n+1}; x_2^2 + x_3^2 \leq \epsilon_n\right\}; \tag{2.25}$$

$$D \equiv E \backslash \bigcup_{n \in \mathbb{N}} F_n. \tag{2.26}$$

Let $B_t \equiv (B_t^{(1)}, B_t^{(s)}, B_t^{(3)})$ be three dimensional Brownian motion. We know form our discussion of the Bessel-processes that $(B_t^{(2)}, B_t^{(3)})$ will never hit the point $(0, 0)$, i.e.

$$\mathbb{P}\left[\exists t > 0 : (B_t^{(2)}, B_t^{(3)}) = (0, 0)\right] = 0.$$

Thus B_t will never hit the compact set

$$K_n \equiv \left\{(x_1, x_2, x_3) : 2^{-2} \le x_1 \le 2^{-n+1}; x_2^2 + x_3^2 = 0\right\};$$

Since moreover $\|B_t\| \to \infty$, a.s., almost all paths remain some positive distance away form the set K_n, and hence, the probability that a path enters an ϵ-neighborhood of it can be made as small as desired by choosing ϵ small enough. In particular, one can choose ϵ_n so small that

$$\mathbb{P}\left[\exists t > 0 : B_t \in F_n\right] \le 3^{-n}.$$

But unless B_t (starting at 0) immediately returns to D, i.e. if $\sigma_D = 0$, B_t must enter the set $\bigcup_{n\in\mathbb{N}} F_n$, so that

$$\begin{aligned}\mathbb{P}_0[\sigma_D = 0] &\le \mathbb{P}\left[\exists t > 0, \exists_n : B_t \in F_n\right] \\ &\le \sum_{n=1}^{\infty} \mathbb{P}\left[\exists t > 0 : B_t \in F_n\right] \le \sum_{n=1}^{\infty} 3^{-n} < 1.\end{aligned} \tag{2.27}$$

Hence 0 is not regular.

2.2 Maximum principle and Harnack-inequalities

The relation between harmonic functions and martingales has a number of further implications.

The first of these is the *mean value property.*

Lemma 2.2.8 *Let D be a bounded domain, τ_D the first hitting time of a Markov process with generator $\mathcal{L}$. Let $z \in D$ be fixed. Assume that $\mathbb{E}_z \tau_D < \infty$. Define μ_D as the probability measure on ∂D given by*

$$\mu_D(dx) = \mathbb{P}_z\left[X_{\tau_D} \in dx\right]. \tag{2.28}$$

Then, a if a function $h : D \to \mathbb{R}$ is harmonic in D, it holds that,

$$\int_{\partial D} \mu_D(dx) h(x) = h(z). \tag{2.29}$$

Proof. Use the fact that $h(x_{\tau_D})$ is a martingale. □

The measure $\mu_D(dx)$ is called the *exit distribution.* It is absolutely continuous with respect to the Euclidean surface measure, $n_D(dx)$ on ∂D.

An immediate consequence of the mean value property is the *maximum principle:*

Theorem 2.2.9 *Let h be harmonic in an open, connected domain D. If h achieves its supremum in D, then it is constant.*

Proof. Let $h(x) = \sup_{y \in D} h(y) = M$. Let $D_M \equiv \{y \in D : h(y) = M\}$. Since h is continuous, this set is closed. Moreover, by the mean value property, for any $y \in D_M$, for any ball $B_r(y) \subset D$

$$M = h(y) = \int_{\partial B_r(y)} \mu_{B_r(y)}(dz) h(z),$$

which implies that for $\mu_{B_r(y)}$-almost all $z \in \partial B_r(y)$, $h(z) = M$. Since both h is continuous (and $\mu_{B_r(y)}$ is is absolutely continuous with respect to the surface measure, $h(y) = M$ for all $y \in \partial B_r(y)$. But then D_M is open, and being open and closed, it must coincide with D. □

A more subtle consequence of the martingale property for harmonic functions are the *Harnack-inequalities*.

We consider first the case of Brownian motion. Let $R > 0$ and let $B_R/x)$ the ball of radius R centered at x. By symmetry, we have the $\mathbb{P}_x \left[\tau_{B_R(x)} \in dz\right]$ is the uniform distribution on $\partial B_R(x)$. Hence, by the mean value property,

$$\begin{aligned} h(x) &= \frac{1}{\int_{B_R(x)} d^d y} \int_0^R dr \int_{\partial B_r(x)} h(z) \sigma_{B_r(x)}(dz) \\ &= \frac{1}{V(B_R)} \int_{B_R(x)} h(y) d^d y. \end{aligned} \tag{2.30}$$

I.e., the value of $h(x)$ equals to its spatial average over the ball of radius R. Now let $y \in B_R(x)$ and r sucht that $B_r(y) \subset B_R(x)$. Clearly we have again that

$$h(y) = \frac{1}{V(B_r)} \int_{B_r(y)} h(z) d^d z. \tag{2.31}$$

Now let h be a *non-negative* harmonic function. Then it follows that

$$h(x) \geq \frac{1}{V(B_R)} \int_{B_r(y)} h(z) d^d z = \frac{V(B_r)}{V(B_R)} h(y) = \left(\frac{r}{R}\right)^d h(y). \tag{2.32}$$

From these basic estimates one can now derive the the Harnack inequality.

Theorem 2.2.10 *Let $D' \subset D$ be two connected open sets. Let h be a non-negative harmonic fucntion with respect to Brownian motion on $D \subset \mathbb{R}^d$. Then there exists a constant K, depending only on D, D', such that*

$$\sup_{x \in D'} h(x) \leq K \inf_{x \in D'} h(x). \tag{2.33}$$

Proof. For $y \in D$ choose R such that $B_{4R}(y) \subset D$. Then for any two points, $x_1, x_2 \in B_R(y)$, the previous inequalities imply that

$$h(x_1) = \frac{1}{V(B_R)} \int_{B_R(y)} h(z) d^d z \leq \frac{1}{V(B_R)} \int_{B_{2R}(y)} h(z) d^d z, \quad (2.34)$$
$$h(_2) = \frac{1}{V(B_{3R})} \int_{B_{3R}(y)} h(z) d^d z \geq \frac{1}{V(B_{3R})} \int_{B_{2R}(y)} h(z) d^d z,$$

Hence

$$\sup_{x \in B_R(y)} h(x) \leq 3^d \inf_{x \in B_R(y)} h(x). \quad (2.35)$$

Now let x_1 and x_2 in $\bar{D}'$ be such that $h(x_1) = \sup_{x \in D')} h(x)$, $h(x_2) = \inf_{x \in D'} h(x)$. Now let γ be a closed arc joining x_1 and x_2 in D. Choose R such that $4R < \mathrm{dist}(\gamma, \mathrm{D}^c)$. This arc can, by the Heine-Borel theorem, be covered by a finite number, N, of balls of radius R, where N depends only on D and D'. Then we an compare $h(x_1)$ and $h(X_2)$ by using he estimate (2.35) not more than N times, hence

$$h(x_1) \leq 3^{dN} h(x_2). \quad (2.36)$$

This proves the theorem. □

There are obvious extensions of the Harnack inequality beyond Brownian motion (for analytic proofs in the general case of elliptic SDE's, see [8]). In fact, inspecting the proof all we used on Brownian motion beyond the martingale property of harmonic functions was the uniformity of the exit distribution on balls. Moreover, it is clear that to get a Harnack inequalty, we do not really need uniformity, but upper and lower bounds on the density of the exit distribution are sufficient.

Theorem 2.2.11 *Let X be a continuous strong solution of an SDE. Let $D \subset \mathbb{R}^d$ be a bounded open domain. Assume that there exist constants, $0 < c < C < \infty$, depending only on D, such that, for any ball $B_R(x) \subset D$,*

$$c \leq \frac{\mathbb{P}_x \left(X_{\tau_d} \in dy \right)}{n_D(dy)} \leq C. \quad (2.37)$$

Then any harmonic function h satisfies a Harnack inequality in D, in the sense that for any $D' \subset D$, there exists a constant K, such that (2.33) holds.

The proof is on the exact same lines as that of the previous theorem and will be left as an exercise.

3

Reversible diffusions

In this Chapter we turn to more explicit computations in the context of diffusion processes with small diffusivity. We will exploit the some special structures in the context of reversible processes.

3.1 Reversibility

The theory of Markov processes that we have developed so far can be seen as a theory of operators acting either on bounded functions (the semi-group action of functions), or on measures. In special cases we can replace this by a L^2 theory with respect to certain measures.

Let P_t by a strongly continuous contraction semi-group acting on a space $B(\mathcal{S})$. Assume that a measure, μ, on S, is invariant with respect to P_t. Then the action of P_t can be extended to the L^2 space $L^2(S, \mu)$.

Lemma 3.1.1 *Let f be in $L^2(S, \mu)$ where μ is invariant with respect to P_t. Then $(P_t f) \in L^2(S, \mu)$.*

Proof. We will show that the L^2-norm of $P_t f$ is controlled by that of f. Namely,

$$\begin{aligned} \int \mu(dx) \left[(P_t f)(x)\right]^2 &= \int \mu(dx) \left[\int P_t(x, dy) f(y)\right]^2 \\ &\leq \int \mu(dx) \int P_t(x, dy) f(y)^2 \int P_t(x, dy) \\ &\leq \int \mu(dx) \int P_t(x, dy) f(y)^2 = \int \mu(dx) f(x)^2 \end{aligned} \tag{3.1}$$

Note that we used the Cauchy-Schwarz inequality and the invariance of μ. □

Having an L^2-action of P_t, we can naturally define its adjoint, P_t^*, via

$$\int \mu(dx) f(x)(P_t g)(x) = \int \mu(dx) \left(P_t^* f\right)(x) g(x), \tag{3.2}$$

for all $f, g \in L^2(S, \mu)$. One may check that P^* is itself a Markov semi-group that generates the time-reversed process to X, in the sense that $\left(P_t^* f\right)(X_t) = f(X_0)$.

Definition 3.1.1 A measure, μ, on S is called reversible with respect to P_t, if, for all functions $f, g \in L^2(S, \mu)$,

$$\int f(x)(P_t g)(x)\mu(dx) = \int g(x)(P_t f)(x)\mu(dx) \tag{3.3}$$

Lemma 3.1.2 *If μ is a reversible probability measure for P_t, then μ is an invariant probability measure for P_t.*

Proof. Clearly $f = 1$ is in $L^2(S, \mu)$. Hence we have

$$\int (\mu P_t)(dx) g(x) = \int (P_t g)(x)\mu(dx) = \int g(x)\mu(dx). \tag{3.4}$$

for all bounded measurable functions g, hence μ is invariant. □

Note that the converse is not true in general, i.e. an invariant measure is not necessarily reversible.

Thus, we may also say that that a measure is reversible with respect to P_t, if P_t is *self-adjoint* on the space $L^2(S, \mu)$.

The terminology "reversible measure" is customary, but actually irritating. The reversibility property is one of the Markov process, resp. the semi-group, and not one of the measure. So I prefer to call a Markov semi-group reversible, if there exits a measure, μ, such that P_t is symmetric in the space $L^2(S, \mu)$, i.e. that (3.3) holds.

One of the nice things is that a SCCSG that is reversible is a contraction in the L^2-space, by Lemma 3.1.1.

The notions above introduced through the semi-group extend to the generator of a Markov process. Thus, for an invariant measure μ, we can define the adjoint, $\mathcal{L}^*$ of a generator $\mathcal{L}$, such that

$$\int \mu(dx)(\mathcal{L}^* g)(x) f(x) = \int \mu(dx) g(x)(\mathcal{L} f)(x), \tag{3.5}$$

for all $f, g \in \mathcal{D}(\mathcal{L})$ such that $\mathcal{L}f, \mathcal{L}g \in L^2(S, \mu)$. Note that, if μ is a probability measure, the second condition is automatically verified. A reversible Markov process is then characterized by the fact that its generator is self-adjoint in $L^2(S, \mu)$ for some invariant measure μ.

Theorem 3.1.3 *Let μ be a reversible measure for a Markov process. Then the generator, $\mathcal{L}$, defines a non-negative definite quadratic form,*

$$\mathcal{E}(f,g) \equiv -\int \mu(dx)g(x)(\mathcal{L}f)(x), \tag{3.6}$$

called the Dirichlet form .

Proof. First, due to the fact that $\mathcal{L}$ is self-adjoint, $\mathcal{E}(f,f)$ is real for all f in $\mathcal{D}(\mathcal{L})$. Moreover, by definition, we have that for such f and if $\mathcal{E}(f,f) < \infty$,

$$\mathcal{E}(f,f) = \lim_{t\downarrow 0} t^{-1} \int \mu(dx)f(x)\left(f(x) - (P_t f)(x)\right). \tag{3.7}$$

But

$$\begin{aligned} \int \mu(dx)f(x)\left(f(x) - (P_t f)(x)\right) &= \|f\|_{2,\mu}^2 - \int \mu(dx)f(x)(P_t f)(x) \\ &\geq \|f\|_{2,\mu}^2 - \|f\|_{2,\mu}\|P_t f\|_{2,\mu}^2 \geq \|f\|_{2,\mu}^2 - \|f\|_{2,\mu}\|f\|_{2,\mu}^2 = 0, \end{aligned} \tag{3.8}$$

where we used Cauchy-Schwarz and Lemma 3.1.1. This implies that the limit, too, is non-negative. □

Remark 3.1.1 The form $\mathcal{E}$ can be extended to the set $\{f : \mathcal{E}(f,f) < \infty\}$ which mostly is larger than the domain if $\mathcal{L}$. There is an entire theory that allows to use this fact to construct a Markov process from a Dirichlet form. For a detailed treatment, see e.g. the book [7] by Fukushima et al..

Since $\mathcal{L}$ is positive and self-adjoint, it can be written in the form $\mathcal{L} = AA$, with A positive, and the Dirichlet form then has the form

$$\mathcal{E}(f,g) = \int \mu(dx)Af(x)Ag(x). \tag{3.9}$$

3.2 Reversible diffusions

We will now look at reversibility issues in the context of diffusions. The formal adjoint of the operator $\mathcal{L}$ given in (2.2) is

$$\mathcal{L}^* g(x) = \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}a_{ij}(x)g(x) - \sum_i \frac{\partial}{\partial x_i}b_i(x)g(x) \tag{3.10}$$
$$= \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}g(x)$$
$$+\sum_i \left(\sum_j \frac{\partial a_{ij}(x)}{\partial x_j} - b_i(x)\right)\frac{\partial}{\partial x_i}g(x)$$
$$+\left(\frac{1}{2}\sum_{ij}\frac{\partial^2}{\partial x_i \partial x_j}a_{ij}(x) - \sum_i \frac{\partial}{\partial x_i}b_i(x)\right)g(x).$$

We can see that this is equal to $\mathcal{L}$ if and only if

$$\sum_j \frac{\partial a_{ij}(x)}{\partial x_j} = 2b_i(x), \tag{3.11}$$

for all $i = 1, \dots d$. Thus (3.11) is a condition for the diffusion to be reversible with respect to Lebesgue measure.

Next we may want to look for a reversible measure $\mu(dx) = e^{F(x)}dx$, i.e. a reversible measure that is absolutely continuous w.r.t. Lebesgue measure. This will be the case if

$$(\mathcal{L}^*(ge^F)(x) = e^{F(x)}\mathcal{L}g.$$

But

$$(\mathcal{L}^*(ge^F)(x) = e^{F(x)}\frac{1}{2}\sum_{i,j}a_{ij}(x)\frac{\partial^2 g(x)}{\partial x_i \partial x_j} \tag{3.12}$$
$$+e^{F(x)}\sum_{i,j}a_{ij}(x)\frac{\partial g(x)}{\partial x_i}\frac{\partial F(x)}{\partial x_j}$$
$$+e^{F(x)}\frac{1}{2}\sum_{i,j}a_{ij}(x)g(x)\left[\frac{\partial^2 F(x)}{\partial x_i \partial x_j} + \frac{\partial F(x)}{\partial x_i}\frac{\partial F(x)}{\partial x_j}\right]g(x)$$
$$+e^{F(x)}\sum_i\left(\sum_j \frac{\partial a_{ij}(x)}{\partial x_j} - b_i(x)\right)\left(\frac{\partial F(x)}{\partial x_i}g(x) + \frac{\partial g(x)}{\partial x_i}\right)$$
$$+e^{F(x)}\left(\frac{1}{2}\sum_{ij}\frac{\partial^2}{\partial x_i \partial x_j}a_{ij}(x) - \sum_i \frac{\partial}{\partial x_i}b_i(x)\right)g(x)$$

The first condition for reversibility is then that

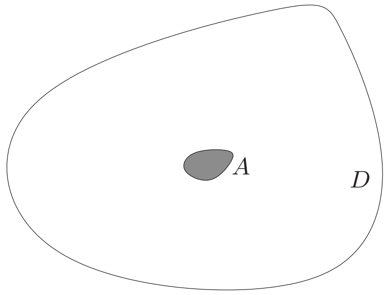

Fig. 3.1. Capacitor

$$\sum_j \left(a_{ij}(x) \frac{\partial F(x)}{\partial x_j} + \frac{\partial a_{ij}}{\partial x_j} \right) = 2b_i(x). \tag{3.13}$$

or

$$2b_i(x) = e^{-F(x)} \sum_j \frac{\partial}{\partial x_j} \left(a_{ij}(x) e^{F(x)} \right). \tag{3.14}$$

In particular, in the simples case when $a_{ij}(x) = \delta_{ij}$, we get a necessary and sufficient condition

$$2b_i(x) = \frac{\partial}{\partial x_i} F(x), \tag{3.15}$$

i.e. the drift must be the gradient of a potential F (up to the factor 2). In that case the generator takes the very suggestive form

$$\mathcal{L} = \frac{1}{2} e^{-F(x)} \nabla e^{F(x)} \nabla. \tag{3.16}$$

The corresponding Dirichlet form then can be written as

$$\mathcal{E}(f,g) = -\int \mu(dx) f(x) (\mathcal{L}g)(x) = \frac{1}{2} \int \mu(dx) \langle \nabla f(x), \nabla g(x) \rangle, \tag{3.17}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^d$.

3.3 Equilibrium measure, equilibrium potential, and capacity

In the following we will return to the general case of SDE corresponding to a generator that is a uniformly elliptic differential operator $\mathcal{L}$ with coefficients satisfying Lipschitz conditions (so that unique strong solutions to the SDE exist).

Let D be a open domain in $\mathbb{R}^d$ with $\partial D = A \cup B$, with $A \cap B = 0$. Then the solution of the Dirichlet problem

$$\begin{aligned} \mathcal{L}h(x) &= 0, x \in D \\ h(x) &= 1, x \in A \\ h(x) &= 0, x \in B \end{aligned} \tag{3.18}$$

is called the *equilibrium potential* of the *capacitor* (A, B). Recall that, for $x \in D$,

$$h(x) = \mathbb{P}_x[\tau_A < \tau_B]. \tag{3.19}$$

Remark 3.3.1 The boundary conditions here are not continuous, so recall Remark 2.1.1. We do not assume that D is connected.

Remark 3.3.2 The names here come from the classical case when $\mathcal{L} = \Delta/2$. Then the Dirichlet problem is a classical problem of electrostatics. The sets A and B correspond to two metal plates attached to a battery that imposes a constant voltage (potential) difference between the plates. The solution of this problem then describes the electrostatic potential (whose gradient is the electrostatic field).

Next we consider the inhomogeneous Dirichlet problem,

$$\begin{aligned} -(\mathcal{L}f)(x) &= g, \ x \in D \\ f(x) &= 0, \ x \in \partial D \end{aligned} \tag{3.20}$$

We have seen that, if this problem has a unique solution, then it has the probabilistic representation

$$f(x) = \mathbb{E}_x \int_0^{\tau_D} g(X_t) = \mathbb{E}_x \int_0^{\tau_D} \int_D P_t(x, dy) g(y) dt, \tag{3.21}$$

where $P_t(x, dy)$ is the semi-group associated to the generator $\mathcal{L}$. Thus we define the *Green kernel*,

$$G_D(x, dy) \equiv \mathbb{E}_x \int_0^{\tau_D} P_t(x, dy) dt \tag{3.22}$$

in terms of which the solution of (3.20) can be written as

$$f(x) = \int_D g(y) G_D(x, dy) \equiv (G_D g)(x). \tag{3.23}$$

Note the similarity with the *resolvent* of the semigroup. In fact, one may define

$$G_D^{(\lambda)}(x, dy) \equiv \mathbb{E}_x \int_0^{\tau_D} e^{-\lambda t} P_t(x, dy) dt \tag{3.24}$$

Then G_D^λ exists for all $\lambda > 0$ even if $\mathbb{E}_x \tau_D = \infty$, and

$$f_\lambda(x) = \int_D g(y) G_D^{(\lambda)}(x, dy) \equiv (G_D^{(\lambda)} g)(x) \tag{3.25}$$

solves the Dirichlet problem

$$\begin{aligned}(-\mathcal{L} - \lambda) f_\lambda(x) &= g, x \in D \\ h(x) &= 0, x \in \partial D\end{aligned} \tag{3.26}$$

Note that it is of course an interesting question (to which we will return), to ask for which values of λ we can still define $G_D^{(\lambda)}$ for given D.

The Green kernel will often have a density with respect to Lebesgue measure, i.e.

$$G_D(x, dy) = G_D(x, y) dy. \tag{3.27}$$

The function $G_D(x, y)$ is then called the *Green function*.

Example Let us return to the case of Brownian motion and $\mathcal{L} = \frac{\Delta}{2}$. Then we know that $P_t(x, dy) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{\|x-y\|^2}{2}} dy$, and hence

$$\begin{aligned}G_D(x, y) &= \mathbb{E}_x \int_0^{\tau_D} \frac{1}{(2\pi t)^{d/2}} e^{-\frac{\|x-y\|^2}{2}} dt \\ &= \pi^{-d/2} \|x - y\|^{2-d} \mathbb{E}_x \Gamma\left(\frac{d}{2} - 1, \frac{\|x-y\|^2}{2\tau_D}\right),\end{aligned} \tag{3.28}$$

where $\Gamma(a, b)$ denote the incomplete gamma-function,

$$\Gamma(a, b) \equiv \int_b^\infty e^{-s} s^{a-1}.$$

Let us now look at the relation between equilibrium potential and the Dirichlet form in the case of a reversible diffusion. Let us try to compute $\mathcal{E}(h, h)$. One might be tempted to think that $\mathcal{E}(h, h) = 0$, since $\mathcal{L}h(x) = 0$ except on the boundary of the sets A and B. But of course on these, $\mathcal{L}h$ may be singular, since no differentiability assumptions are made on the boundary. So we may interpret $\mathcal{L}h$ as a measure that is concentrated on the boundaries of A and B. Since h vanishes on ∂B, we get that

$$\mathcal{E}(h, h) = -\int_{\partial A} \mu(x)(\mathcal{L}h)(dx). \tag{3.29}$$

The measure $(-\mathcal{L}h)(dx)$ is called the *equilibrium measure* associated to the capacitor A, B.

To understand this better, let us return to the case $a_{ij}(x) \equiv \delta_{ij}$. We then have the following *integral formula*, known as the *first Green's formula* .

Theorem 3.3.4 *Let D be a regular domain and let ϕ, ψ be in $C^2(D)$. Let $\mathcal{L}$ be given by (2.2). Then*

$$\int_D dx\, e^{F(x)} \big(\langle \nabla\phi(x), \nabla\psi(x)\rangle - \psi(x)(2\mathcal{L}\phi)(x)\big) \tag{3.30}$$
$$= \int_{\partial D} e^{F(x)} \psi(x)\partial_{n(x)}\phi(x)\, d\sigma_D(x),$$

where $\partial_{n(x)}$ denotes the inner normal derivative at $x \in \partial D$.

Proof. In the case $F = 0$ this formula a classical. The extension to the general case is by a straightforward computation. □

Remark 3.3.3 An immediate consequence of this identity is the so-called *second Green's formula*,

$$\int_D e^{F(x)} dx\, \big(\phi(x)(2\mathcal{L}\psi)(x) - \psi(x)(2\mathcal{L}\phi)(x)\big) \tag{3.31}$$
$$= \int_{\partial D} e^{F(x)} (\psi(x)\partial_{n(x)}\phi(x) - \phi(x)\partial_{n(x)}\psi(x))\, d\sigma_D(x)$$

The second Green's formula gives rise to the integral representation of a solution of the Dirichlet boundary value problem,

$$-(\mathcal{L}f)(x) = 0, \quad x \in D, \tag{3.32}$$
$$f(x) = u(x), \quad x \in \partial D,$$

in terms of the *Poisson kernel* , namely

$$f(x) = \int_{\partial D} e^{F(y)-F(x)} u(y)\partial_{n(y)} G_D(y, x) d\sigma_D(y). \tag{3.33}$$

Using the first Green's formula, we can give a precise relation between equilibrium potential and capacity. Namely, setting $\phi = \psi = h$ in (3.32), we see that

$$\int_D dx\, e^{F(x)} \langle \nabla h(x), \nabla h(x)\rangle = \int_A e^{F(x)} \partial_{n(x)} h(x) d\sigma_A(x), \tag{3.34}$$

i.e. we have that on A the equilibrium measure, $(-\mathcal{L}h)(x)$ is given by

$$e_{A,B}(dx) \equiv \partial_{n(x)} h(x) d\sigma_A(x). \tag{3.35}$$

The quantity

$$\mathrm{cap}(A, B) \equiv \int_A e^{F(x)} \partial_{n(x)} h(x) d\sigma_A(x), \tag{3.36}$$

is called the *capacity* of the capacitor A, B. In electrical language, it is the total charge on the plate A. Using relation (3.34), we see that alternatively, the capacity is also the total *energy* of the potential h.

Last exit distribution and equilibrium measure. It will be nice to have a probabilistic interpretation of the equilibrium measure that will at the same time explain why Lh really becomes a surface measure.

We see that, for x in A, we should have something like

$$\begin{aligned} -(\mathcal{L}h)(x) &= \lim_{t\downarrow 0} t^{-1}(1-P_t)(h(x)) \\ &= \lim_{t\downarrow 0} t^{-1}\mathbb{E}_x\left(1-\mathbb{P}_{X_t}\left[\tau_A<\tau_B\right]\right) \\ &= \lim_{t\downarrow 0} t^{-1}\mathbb{E}_x\mathbb{P}_{X_t}\left[\tau_B<\tau_A\right]. \end{aligned} \tag{3.37}$$

Let us define the *last exit time*, L_A, from A as

$$L_A \equiv \sup\{0<0<\tau_B : X_t \in A\}, \tag{3.38}$$

with the convention $\sup\emptyset = 0$ Note that this is obviously no stopping time and that

$$\mathbb{P}_x[L_A>0] = \mathbb{P}_x[\tau_A<\tau_B] \equiv h(x). \tag{3.39}$$

Note that we can write the expression in the last line of (3.37) as

$$\mathbb{E}_x\mathbb{P}_{X_t}\left[\tau_B<\tau_A\right] = \mathbb{P}_x[0<L_A<t].$$

Hence we set

$$\psi_t(z) \equiv t^{-1}\mathbb{P}_z\left[0<L_A<t\right]. \tag{3.40}$$

Let us also define the *last exit distribution*, $L(x,dy)$, on A, by

$$L(x,dy) \equiv \mathbb{P}_x\left[X_{L_A-}\in dy; L_A>0\right]. \tag{3.41}$$

We want to prove the following lemma:

Lemma 3.3.5 *Let f be a continuous function on $\bar{D}$. Then*

$$\lim_{t\downarrow 0}\int G_D(x,y)\psi_t(y)f(y)dy = \int_A L(x,dy)f(y). \tag{3.42}$$

Proof. Without loss let $f\geq 0$. Using the representation of the Green function through the semigroup (3.22) we get

$$\begin{aligned}\int G_D(x,y)\psi_t(y)f(y)dy &= \mathbb{E}_x \int_0^{\tau_B} \psi_t(X_s)f(X_s)ds \\ &= t^{-1}\int_0^\infty \mathbb{E}_x\left[f(X_s)\mathbb{P}_{X_s}[0<L<t]\right]ds \\ &= t^{-1}\int_0^\infty \mathbb{E}_x\left[f(X_s)\mathbb{1}_{s<L<s+t}\right]ds \\ &= \mathbb{E}_x\left[0<L_A\leq t; t^{-1}\int_0^{L_a} f(X_s)ds\right] \\ &\quad +\mathbb{E}_x\left[t<L_A; t^{-1}\int_{L_A-t}^{L_a} f(X_s)ds\right].\end{aligned} \tag{3.43}$$

First, both terms in the last line are obviously uniformly bounded as $t \downarrow 0$. Moreover,

$$\mathbb{E}_x\left[0<L_A\leq t; t^{-1}\int_0^{L_a} f(X_s)ds\right] \leq C\mathbb{E}_x\left[0<L_A\leq t\right]\downarrow 0, \tag{3.44}$$

as $t \downarrow 0$. Finally, by continuity of f,

$$\lim_{t\downarrow 0}\mathbb{E}_x\left[t<L_A; t^{-1}\int_{L_A-t}^{L_a} f(X_s)ds\right] = \mathbb{E}_x\left[0<L_A; f(X_{L_A-})ds\right]. \tag{3.45}$$

Integrating over A gives the claim of the lemma. □

From Lemma (3.3.5) one can deduce that the family of measures $\psi_t(y)dy$ converges to a measure $e(dy)$ on A. Moreover, this measure satisfies

$$G_D(x,y)e(dy) = L(x,dy). \tag{3.46}$$

Integrating this formula over A, we arrive at the expression

$$\int_A G_D(x,y)e(dy) = \int_A L(x,dy) = h(x). \tag{3.47}$$

Hence $e(dy)$ satisfies the defining relation of the equilibrium measure

Thus we have proven a very interesting and useful relation between the equilibrium potential, the equilibrium measure, and the Green function.

Theorem 3.3.6 *Let as before $A \subset D$ be open sets with smooth boundary. Then, for all $x \in D$,*

$$h(x) = \int_{\partial A} G_D(x,y)e_{A,D}(dy). \tag{3.48}$$

Remark 3.3.4 Is is instructive to think about this result in the following way. We have already seen that we may want to think of $\mathcal{L}h$ as a measure. Then we have that

$$\begin{aligned} -(\mathcal{L}h)(x)dx = e_{A,D}(dx), \quad x \in D, \\ h(x) = 0, \quad x \in \partial D. \end{aligned} \tag{3.49}$$

Then the solution of this problem in terms of the Green function is precisely the expression (3.48). Note that (3.3.6) holds also in A, with $h(x) = 1$.

This formula for the Green function gives of course corresponding formulas for solutions of Dirichlet problems. E.g., if we consider for some function g the Dirichlet problem

$$\begin{aligned} -(\mathcal{L}f)(x) = g(x), \quad x \in D \\ f(x) = 0, \quad x \in \partial D, \end{aligned} \tag{3.50}$$

then of course $f(x) = \int_D dy G_D(x,y)g(y)$. By symmetry, $G_D(x,y) = e^{F(y)-F(x)}G_D(y,x)$, and so

$$\begin{aligned} \int_D dx e^{F(x)} h(x)g(x) &= \int_D dx h(x) \int_{\partial A} e^{F(x)} g(x) G_D(y,x) e^{F(y)-F(x)} e_{A,D}(dy) \\ &= \int_{\partial A} e^{F(y)} \int_D G_D(y,g)g(x)e_{A,D}(dy) \\ &= \int_{\partial A} e^{F(y)} f(y) e_{A,D}(dy). \end{aligned} \tag{3.51}$$

Introducing the probability measure

$$\nu_{A,D}(dy) \equiv \frac{e^{F(y)}e_{A,D}(dy)}{\mathrm{cap}(A,D)}, \tag{3.52}$$

on ∂A, this gives

$$\int_{\partial A} \nu_{A,D}(dy)f(y) = \frac{1}{\mathrm{cap}(A,D)} \int_D dx e^{F(x)} h(x)g(x). \tag{3.53}$$

As a particular example we get, with $g(x) = 1$,

$$\int_{\partial A} \nu_{A,D}(dy)\mathbb{E}_y \tau_D = \frac{1}{\mathrm{cap}(A,D)} \int_D dx e^{F(x)} h(x). \tag{3.54}$$

Dirichlet principle . We have seen that the equilibrium Dirichlet form computed on the equilibrium potential gives the capacity. We will how show that the equilibrium potential is the solution of a variational problem.

Theorem 3.3.7 *With the notations and assumptions above, the following holds. Let $\mathcal{H}_{A,D}$ be the space of continuous functions,, f, on $\bar{D}$ such that,*

(i) $\mathcal{E}(f,f,) < \infty$*, and*
(ii) $f(x) \geq 1$*,* $x \in A$ *and* $f(x) \leq 0$*,* $x \in< D^c$*.*

$$\mathrm{cap}(A,d) = \inf_{f \in \mathcal{H}_{A,D}} \mathcal{E}(f,f). \tag{3.55}$$

Moreover, if $\mathcal{H}_{A,B} \neq \emptyset$, the infimum in (3.55) is achieved uniquely on the equilibrium potential, i.e. $\mathrm{cap}(A,D) = \mathcal{E}(h,h)$*.*

Proof. Let us assume that the set $\mathcal{H}_{A,B}$ is not empty. Consider a function g such that $\mathcal{E}(g,g) < \infty$ and that g vanishes on both ∂D and ∂A. Notice that, for $h \in \mathcal{H}_{A,D}$,

$$\mathcal{E}(h+\epsilon g, h+\epsilon g) - \mathcal{E}(h,h) = \epsilon \int_{D \setminus \bar{A}} \mu(dx) g(x) (\mathcal{L}h)(x) + \epsilon^2 \mathcal{E}(g,g). \tag{3.56}$$

This implies two things: first, if $\mathcal{L}h(x) = 0$, then h is a global minimum of $\mathcal{E}$ in $\mathcal{H}_{A,D}$. We know already that such a function exists, namely the equilibrium potential. Next assume that there is another function, f, such that $\mathcal{E}(f,f)) = \mathcal{E}(h,h)$. Then the identity

$$\mathcal{E}\left(\tfrac{f+h}{2}, \tfrac{f+h}{2}\right) + \mathcal{E}\left(\tfrac{f-h}{2}, \tfrac{f-h}{2}\right) = \frac{1}{2}\mathcal{E}(f,f) + \frac{1}{2}\mathcal{E}(h,h), \tag{3.57}$$

implies that

$$\mathcal{E}\left(\tfrac{f+h}{2}, \tfrac{f+h}{2}\right) \leq \mathcal{E}(h,h) - \mathcal{E}\left(\tfrac{f-h}{2}, \tfrac{f-h}{2}\right). \tag{3.58}$$

Since h is an absolute minimum, this can only hold if

$$\mathcal{E}\left(f-h, f-h\right) = 0. \tag{3.59}$$

But this means that $\|\nabla(f-g)(x)\|_2 = 0$ μ-almost surely. □

The Dirichlet principle is a powerful tool for asymptotic computations of capacities, and, hence, as we shall see, much more. To a large extend this is due to the fact that it allows for natural upper and lower bounds. The most immediate one of these is of course given by the elementary observation that

Corollary 3.3.8 *For any function,* $f \in \mathcal{H}_{A,B}$*,*

$$\mathrm{cap}(A,B) \leq \mathcal{E}(f,f). \tag{3.60}$$

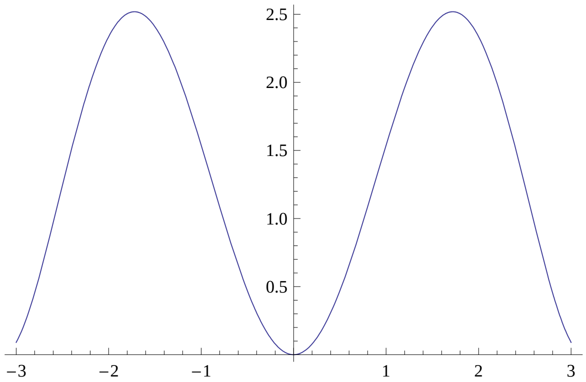

Fig. 3.2. A potential function on $[-3, 3]$

3.4 The case of dimension one.

The above considerations lead to very explicite answers in the case when $d = 1$. The first observation is that all homogeneous boundary value problems in this case can, by linearity, be reduced to computing the equilibrium potential for on interval (a, b),i.e.

$$\begin{aligned}(\mathcal{L}h)(x) &= 0, \quad x \in (a, b) \\ h(a) &= 0 \\ h(b) &= 1\end{aligned} \tag{3.61}$$

Note also that the general uniformly elliptic case, we can be reduced to the problem with generator

$$(\mathcal{L}h)(x) - \frac{1}{2}f''(x) + b(x)f'(x). \tag{3.62}$$

Note also that, in $d = 1$, any bounded function b can be written as a derivative of another function, $F/2$, where

$$F(x) = 2\int_0^x b(x)dx. \tag{3.63}$$

Thus we are always in the reversible case. Hence we are reduced to solving the ordinary differential equation

$$\frac{1}{2}h''(x) + b(x)h'(x) = 0, \tag{3.64}$$

which in turn reduces to the first order equation

$$\frac{1}{2}u'(x) + b(x)u(x) = 0 \tag{3.65}$$

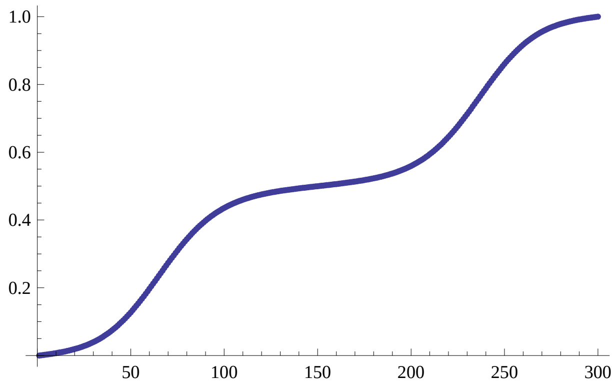

Fig. 3.3. The corresponding equilibrium potential $P_x[\tau_3 < \tau_{-3}]$

when we set $u = h'$. Clearly (3.65) has the general solution

$$u(x) = C_1 e^{-F(x)} \tag{3.66}$$

and so the general solution of (3.64) is

$$h(x) = C_1 \int_0^x e^{-F(y)} dy + C_2, \tag{3.67}$$

with C_1 and C_2 integration constants to be determined from the boundary conditions. In particular, for the equilibrium potential related to the interval (a, b) we have

$$h(x) = \frac{\int_a^x e^{-F(y)} dy}{\int_a^b e^{-F(y)} dy}. \tag{3.68}$$

Hence the capacity $\text{cap}(a, b)$ is readily computed as

$$\text{cap}(a, b) = \mathcal{E}(h, h) = \frac{1}{2 \int_a^b e^{-F(y)} dy}. \tag{3.69}$$

Some reflection shows that we can get from (3.3.6) the following formula for the Green function in (a, b): For $x < y$,

$$G_{(a,b)}(x,y) = \frac{h_{x,\{a,b\}}(y)}{e_{x,b}} \tag{3.70}$$
$$= {}^{F(x)} \frac{1 - h_{x,b}(y)}{\mathrm{cap}(x,b)}$$
$$= e^{F(x)} \frac{\frac{\int_y^b e^{-F(z)}dz}{\int_x^b e^{-F(z)}dz}}{\frac{1}{2\int_x^b e^{-F(z)}dz}}$$
$$= \frac{1}{2} e^{F(x)} \int_y^b e^{-F(z)} dz$$

and for $y < x$,

$$G_{(a,b)}(x,y) = \frac{1}{2} e^{F(y)} \int_a^y e^{-F(z)} dz. \tag{3.71}$$

3.5 Another view on one-dimensional diffusions

We have seen in the previous section that the computation of the equilibrium potential in one-dimensional case allows to compute the Green function and hence to essentially solve everything that can be expressed in terms of Dirichlet problems.

We will now take a different look at the same issue. The perspective will be more on the level of the process. We will see that the solution of a 1d SDE can be constructed from Brownian motion in a way that will exhibit again the crucial rôle.

Let us recall that the harmonic functions we encountered con be written in the form

$$\mathbb{P}_x(\tau_a < \tau_b) = \frac{s(x) - s(a)}{s(b) - s(a)}, \tag{3.72}$$

where $s(x)$ is an increasing function whose derivative is $e^{-F(x)}$. The function s is usually called the *scale function* Recall that in the case of Brownian motion, $s(x) = x$. Now let B be Brownian motion and consider the process $Y_t = s^{-1}(B_t)$. Clearly we have that

$$\mathbb{P}_x^Y(\tau_a < \tau_b) = \mathbb{P}_{s(x)}^B(\tau_{s(a)} < \tau_{s(b)}) = \frac{s(x) - s(a)}{s(b) - s(a)}, \tag{3.73}$$

(here the superscripts indicate that the probabilities are w.r.t. the corresponding processes) hence the process has the same harmonic function as the one solving $dX_t = \frac{1}{2}F'(X_t)dt + dB_t$. Is it the same process? No, but using Itô's formula, we see that $Z_t \equiv s(X_t)$ satisfies

$$
\begin{aligned}
dZ_t &= s'(X_t)dX_t + \frac{1}{2}s''(X_t)dt \\
&= s'(X_t)dB_t + \frac{1}{2}(s'(X_t)F'(X_t) + s''(X_t))dt \\
&= s'\left(s^{-1}(X_t)\right)dB_t,
\end{aligned} \tag{3.74}
$$

which is of the form

$$dZ_t = g(Z_t)dB_t. \tag{3.75}$$

We will show that any solution of an SDE of the form (3.75) is a time change of Brownian motion.

Theorem 3.5.9 *Let g be a measurable function such that $g(x) \geq \delta > 0$. Then (3.75) has a unique weak solution. Moreover, there exists a Brownian motion, B, such that*

$$Z_t = B(\gamma_t), \tag{3.76}$$

where $\gamma_t \equiv \inf\{u : A(u) > t\}$, and

$$A(t) \equiv \int_0^t g(B(u)^{-2}du. \tag{3.77}$$

Proof. Clearly we have that th if Z solves (3.75), then

$$d[Z]_t = g(Z_t)^2 dt. \tag{3.78}$$

On the other hand,

$$[B]_{[Z]_t} = [Z]_t. \tag{3.79}$$

Inverting this relation we may write

$$B(t) = Z_{\tau(t)}, \tag{3.80}$$

where $\tau(t)$ is the inverse time change, i.e.

$$[Z]_{\tau(t)} = \int_0^{\tau(t)} g(Z_u)^2 du = t. \tag{3.81}$$

Differentiating this latter relation we get

$$1 = g(Z_{\tau(t)})^2 \tau'(t) = g(B(t))^2 \tau'(t). \tag{3.82}$$

Hence

$$\tau'(t) = \frac{1}{g(B(t))^2}, \tag{3.83}$$

and hence

$$\tau(t) = \int_0^t \frac{1}{g(B(u))^2} du, \tag{3.84}$$

τ being the inverse of the time change, we see that the time change $[Z]_t$ is really the inverse of the function τ, which is given purely in terms of the Brownian motion B. □

It remains to generalize the first part of our construction. Thus consider the general form of the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \tag{3.85}$$

with $\sigma > 0$. We know already that the equilibrium potential will be of the form (3.72) with s being an integral of

$$s'(x) = e^{-\int_0^x 2b(z)/\sigma^2(z)dz}. \tag{3.86}$$

Again $Z_t = s(X_t)$ then has the same harmonic function as Brownian motion, and the same calculation as in (3.74) shows that Z_t is a solution of (3.75), this time with

$$g(x) = s'(s^{-1}(x))\sigma(s^{-1}(x)). \tag{3.87}$$

We summarize these results in the following theorem.

Theorem 3.5.10 *Assume that $\sigma(x) > 0$, as long as $\int_0^x b(z)\sigma^{-2}(z) < \infty$ exists of $x \in I$, then the SDE has a unique weak solution given by*

$$X_t = s^{-1}\left(B(\gamma_t)\right), \tag{3.88}$$

where γ_t is the continuous inverse of the function

$$A_t \equiv \int_0^t \frac{1}{g(B(u))^2} du, \tag{3.89}$$

where g is given by (3.87), and B is Brownian motion.

The strong point of this result is that very little regularity is required for the drift or difussivity. For example, this theorem allows to make sense of the *Brownian motion in a Brownian potential*: Let $W(t)$ be a realization of a Brownian motion, and consider the formal expression

$$dX_t = W'(X_t)dt + dB_t. \tag{3.90}$$

Since W is not differentiable, this expression is formal. However, the corresponding potential, W, is well defined, and so is the scale function $s(x) = \exp(W(x))$. Thus we can interpret the process obtained from Eqs. (3.88)-(3.89) as the solution of (3.90).

3.6 Brownian local time and speed measures

The aim of this section is to give an alternative representation of the time change formula that will give rise to the possibility to construct even larger classes of one dimensional diffusion processes. At the same time we will deepen the discussion of *local time* that was initiated in Theorem 1.2.3. The following discussion draws on lecture notes by Steve Lalley.

Let us first observe what would be the natural notion of a occupation measure. Let $A \in \mathcal{B}(\mathbb{R})$ be a Borel subset of the real line. Then we can introduce

$$\Gamma_t(A) \equiv \int_0^t \mathbb{1}_A(B_s)ds \tag{3.91}$$

as a random measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The main result we will use and need, is the following theorem.

Theorem 3.6.11 *With probability one, for each $t < \infty$, the occupation measure Γ_t is absolutely continuous with respect to Lebesgue measure, and it density, l_t^x, is jointly continuous in t and x.*

l_t^x is called the *local time* of Brownian motion at x. We have already seen that the local time at zero can be represented as a stochastic integral via an extension of Itô's formula (Tanaka's formula). This gives the representation

$$l_t^a \equiv |B_t - a| - |B_0 - a| - \int_0^t \mathrm{sign}\,(\mathrm{B_s - a})\mathrm{dB_s}. \tag{3.92}$$

We will first show that

Theorem 3.6.12 *There exists a version of the process $\{l_t^a, a \in \mathbb{R}, t \in \mathbb{R}_+\}$ that is jointly continuous in t and a.*

Proof. We will deal with a fixed time horizon $T < \infty$. Define

$$\xi_1(a,t) \equiv \int_0^t \mathrm{sign}\,(\mathrm{B_s - a})\mathrm{dB_s} \tag{3.93}$$

and

$$\xi_2(a,t) \equiv |B_t - a| - |B_0 - a|. \tag{3.94}$$

Obviously ξ_2 is jointly continuous, since B is continuous. Thus we need to prove that ξ_1 has a jointly continuous version. The tool to prove this is a lemma due to Kolmogorov.

Lemma 3.6.13 *Let ξ be a stochastic process indexed by $\mathbb{R}^n$ with values in a complete metric space with metric ρ. If there exist positive constants, α, β, ϵ, such that*

$$\mathbb{E}\left(\rho(\xi_x, \xi_y)\right)^{\alpha} \leq \beta|x-y|^{n+\epsilon}, \tag{3.95}$$

for all $x, y \in \mathbb{R}^n$, then there exists a continuous version of X.

We apply this theorem to the process $\xi_1(\cdot, t)$, $t \in [0, T]$. We may also consider the process on a bounded interval. Then it will be enough to show that

$$\begin{aligned} \mathbb{E}\left|\xi_1(x,t) - \xi_1(y,t)\right|^p &\leq C|x-y|^{2+\delta} \\ \mathbb{E}\left|\xi_1(x,t) - \xi_1(x,t')\right|^p &\leq C|t-t'|^{2+\delta} \end{aligned} \tag{3.96}$$

Now

$$\begin{aligned} |\xi_1(x,t) - \xi_1(y,t)| &= \left|\int_0^t (\,\mathrm{sign}\,(\mathrm{B_s} - \mathrm{a}) - \,\mathrm{sign}\,(\mathrm{B_s} - \mathrm{b}))\mathrm{dB_s}\right| \\ &\leq 2\int_0^t \mathbb{1}_{(x,x')}(B_s)dB_s. \end{aligned} \tag{3.97}$$

Hence, using the Burkholder inequalities,

$$\begin{aligned} \mathbb{E}\left|\xi_1(x,t) - \xi_1(y,t)\right|^{2m} &\leq 2^{2m}\mathbb{E}\left|\int_0^t \mathbb{1}_{(x,x')}(B_s)dB_s\right|^{2m} \\ &\leq C_m 2^{2m}\mathbb{E}\left|\int_0^t \mathbb{1}_{(x,x')}(B_s)ds\right|^{2m}. \end{aligned} \tag{3.98}$$

In the final expression we can now estimate

$$\begin{aligned} &\mathbb{E}\left|\int_0^t \mathbb{1}_{(x,x')}(B_s)ds\right|^{2m} \\ &= m!\int_{0\leq t_1\leq t_{m-1}\leq\cdots\leq t_m\leq t} \mathbb{P}\left[B_{t_1} \in (x,y), \dots, B_{t_m} \in (x,y)\right] \\ &\leq m!\int_{0\leq t_1\leq t_{m-1}\leq\cdots\leq t_m\leq t} \mathbb{E}\Big[P_{B_0}(B_{t_1} \in (x,y)]P_{B_{t_1}}(B_{t_2} \in (x,y))\dots \\ &\dots P_{B_{t_{m-1}}(B_{t_m}} \in (x,y))\Big] \\ &\leq m!(y-x)^m\sqrt{t}^m. \end{aligned} \tag{3.99}$$

The corresponding estimate for different times is similar. □

With this absolutely continuous local time process we can of course write the time change function (3.89) in the form

$$A_t = \int_0^t \frac{1}{g(B_u)^2} du = \int dx \frac{1}{g(x)^2} \int_0^t du = \int m(dz) l_t^z, \qquad (3.100)$$

where $m(dz) = g^{-2}(z)dz$ and l_t^z is the density of the Brownian local time process. The measure $m(z)$ is called he *speed measure*(essentially it tells us how the local time of Brownian motion is transformed to the time of the new process in the point z). . This formulation gives rise to an even wider class of one-dimensional diffusions that can be constructed as time changes of Brownian motion through more general speed measures.

3.7 The one-dimensional trap model and a singular diffusion

In the following we show that these processes are not totally hypothetical, but that they can arise from more or less reasonable discrete models.

In the following we will give the construction of a random motion in a random environment that was studied by Fontes, Isopi, and Newman some years ago [5].

We begin be prescribing a random environment on $\mathbb{Z}$ as a family of iid random variables, τ_i, $i \in \mathbb{Z}$, whose distribution will be be assumed to satisfy

$$\lim_{t \uparrow \infty} t^\alpha \mathbb{P}[\tau_1 > t] = 1 \qquad (3.101)$$

for $\alpha < 1$. Note that this implies in particular that $\mathbb{E}\tau_1 = +\infty$. Our next ingredient will be a continuous time, unbiased simple random walk, Z_k, $k \in \mathbb{N}$, on $\mathbb{Z}$.

(Note that a continuous time random walk can be described as follows: Let Y_k, $k \in \mathbb{N}$, be a discrete time simple random walk on $\mathbb{Z}$ (i.e. $Y_k = \sum_{i=1}^k u_i$, where u_i are iid with $\mathbb{P}[u_i = \pm 1] = \frac{1}{2}$). Let $C(k) = \sum_{i=0}^{k-1} e_i$, where e_i, $i \in \mathbb{N}$, are iid exponential random variables with rate 1. Then

$$Z_t 0 = Y_{C^{-1}(t)} \qquad (3.102)$$

where C^{-1} is the inverse of C,

$$f^{-1}(t) = \inf\{k : f(k) > t\}. \qquad (3.103)$$

We will now construct a continuous time process a time change of the simple radom walk Z as follows. Define the so-called *clock process*

$$S(u) \equiv \int_0^u \tau_{Z_r} dr \qquad (3.104)$$

The process X is then defined, for a given realization of the random variables τ_i, as

$$X_t = Z_{S^{-1}(u)} \tag{3.105}$$

We now want to re-write the clock process in terms of a speed measure. For this we define the local time process of Z as

$$L(j,t) = \int_0^t \mathbb{1}_{Z_u=j} du. \tag{3.106}$$

Then we can re-write the clock process as

$$S(u) = \sum_{j\in\mathbb{Z}} t_j L(j,u). \tag{3.107}$$

One sees easily that there is a complete analogy between the construction of a diffusion from Brownian motion.

We now consider a rescaling of space and time to obtain a continuous process limit. Clearly we have from (a complete analog of) Donsker's invariance principle that

$$\lim_{\epsilon\downarrow 0} \epsilon Z_{t/\epsilon^2} = B_t. \tag{3.108}$$

Now assume that for some β,

$$\epsilon^2 S^{-1}(\epsilon^{-\beta} t) \equiv S_\epsilon(t) \to \Sigma(t), \tag{3.109}$$

then

$$X_t^\epsilon \equiv \epsilon X_{\epsilon^{-\beta} t} = \epsilon Z_{\epsilon^{-2} S_\epsilon(t)} \equiv Z^\epsilon_{S_\epsilon(t)} \tag{3.110}$$

and we may expect that

$$Z^\epsilon_{S_\epsilon(t)} \to B_{\Sigma(t)}. \tag{3.111}$$

The question is thus to see whether and to what the process $\epsilon^2 S^{-1}(\epsilon^{-\beta} t)$, respectively its inverse,

$$S_\epsilon(u) \equiv \epsilon^\beta S(u/\epsilon^2), \tag{3.112}$$

converges. Now

$$\epsilon^\beta S(u/\epsilon^2) = \sum_{i\in\mathbb{Z}} \epsilon^\beta \tau_i L(i, u/\epsilon^2) \equiv \sum_{i\in\mathbb{Z}} \epsilon^\beta \tau_i L_\epsilon(\epsilon i, u), \tag{3.113}$$

where by definition $L_\epsilon(\epsilon i, u) = L(i, u/\epsilon^2)$ is the the local time process for the rescaled process $Z_\epsilon(t) \equiv \epsilon Z_{t/\epsilon^2}$ and we may expect this to converge to the local time process of Brownian motion. On the other hand, we can think of the sum as an integral over the random measure

$$m_\epsilon(dx) \equiv \sum_{i\in\mathbb{Z}} \delta_{i\epsilon}(dx)\tau_i\epsilon^\beta, \tag{3.114}$$

i.e.

$$S_\epsilon(t) = \int m_\epsilon(dx)L_\epsilon(x,t), \tag{3.115}$$

It is a curious fact that

$$\int m_\epsilon(dx)L_\epsilon(x,t) = \int m_\epsilon(dx)l_t^x. \tag{3.116}$$

This is due to the fact that the local time density of Brownian motion on an integer point i before visiting one of its neighbors is an exponential random variable with mean one. Thus we have actually immediately an expression of our (rescaled) process X^ϵ immediately as a time change of Brownian motion with speed measure m_ϵ.

Thus the key question is whether $m_\epsilon(dx)$ converges. this will be the case, due to (3.101), if $\beta = 1/\alpha$. This follows from a more general result about the convergence of so-called extremal processes.

(for a proof see, e.g., [11]):

Theorem 3.7.14 *Assume that X_i are iid random variables that satisfy*

$$\lim_{n\uparrow\infty} \epsilon^{-1}\mathbb{P}[X_i > u_\epsilon(c)] = \nu(c). \tag{3.117}$$

where ν is an increasing (respectively decreasing) function. Then, the point process

$$\sum_{i\in\mathbb{Z}} \delta_{(i\epsilon, u_\epsilon^{-1}(X_i))} \tag{3.118}$$

converges in distribution to the Poisson point process, $\mathcal{R}$ on $\mathbb{R}_+\times\mathbb{R}$ with intensity measure $dt \times d\nu(x)$ (respectively $-d\nu$ if ν is decreasing).

Using the property (3.101), we see that in our case, with $u_\epsilon(c) \equiv \epsilon^{-1/\alpha}c$, we have that

$$\epsilon^{-1}\mathbb{P}[\tau_1 > \epsilon^{-1/\alpha}c] = c^{-\alpha}[\epsilon^{-1/\alpha}c]^\alpha\mathbb{P}[\tau_1 > \epsilon^{-1/\alpha}c] \to c^{-\alpha}.$$

Thus the theorem yields

Corollary 3.7.15 *The point process*

$$R_\epsilon \equiv \sum_{i\in\mathbb{Z}} \delta_{(i\epsilon, \epsilon^{1/\alpha}\tau_i)} \to \mathcal{R} \tag{3.119}$$

converges to the Poisson point process on $\mathbb{R}\times\mathbb{R}_+$ with intensity measure $dt \times \alpha c^{-1-\alpha}dc$.

One can show that this implies that, if $\alpha < 1$, the measures

$$m_\epsilon(dx) = \int R_\epsilon(dx, dt)t \tag{3.120}$$

converge to the measure

$$m(dx) \equiv \int R(dx, dt)t. \tag{3.121}$$

Note of course that we are speaking of random measures here. So what converges is the distribution of these random measures. In proper terms, we would have to equip the space of measures with a topology (e.g. the vague topology) and speak of *weak convergence* of the family of random measures with respect to this underlying topology.

One can easily check that the measure $m(dx)$ is singular (in fact it is a pure point measure) with respect to Lebesgue's measure. Nonetheless one can use it to construct a singular diffusion as a time change of Brownian motion form it, that will be the natural candidate for the limit process in our model.

It is known that if a sequence of (point) measures, ν_n, converges to a point measure ν (in a suitable topology that I will not discuss here), then the corresponding time-changed processes converge to the process with time change obtained from the speed measure ν.

Can we apply this fact in our case, when the measures μ_ϵ converge only in weakly? The answer is yes, in general due to Skorohod's theorem, that states that weak convergence of a family of random variables, X_n, is equivalent to the existence of a another family, $\bar{X}_n$, such that for each n, X_n and $\bar{X}_n$ have the same distribution, while $\bar{X}_n$ converges almost surely.

A coupling . It is an amusing observation that in the case of our random measures m_ϵ, this construction can be made in a very explicit way. It will also exhibit a deep relation between these measures and Lévy processes.

Let us first briefly recall what an α-stable Lévy subordinator, U, is. There are in fact at least two ways to describe it: one is to say that U is a non-decreasing stationary process with independent increments whose Laplace transform is given by

$$\mathbb{E}e^{-\lambda U(x)} = \exp\left[x\alpha \int_0^\infty (e^{-\lambda w} - 1)w^{-1-\alpha} dw\right]. \tag{3.122}$$

Another way to characterize it is to say that it is the distribution function on $\mathbb{R}$ associated with the measure $m(dx)$, normalized s.t. $U(0) = 0$ (see my lecture notes on ageing [**?**]).

Now introduce the scaling function G such that

$$\mathbb{P}[U(1) \leq G(a)] = \mathbb{P}[\tau_0 \leq a]. \tag{3.123}$$

Then define

$$\tau_i^\epsilon \equiv G^{-1}\left(e^{-1/\alpha}U(\epsilon(i+1) - U(\epsilon i)\right). \tag{3.124}$$

Lemma 3.7.16 *The family of random variables τ_i^ϵ, $i \in \mathbb{Z}$ is iid and τ_i^ϵ has the same distribution as τ_1.*

Proof. The proof of this lemma follows form the fact that the subordinator U is α-stable, i.e. that $\epsilon^{-1/\alpha}U(\epsilon)$ has the same law as $U(1)$. □

Using these random variables we can construct measures

$$\bar{m}_\epsilon \equiv \sum_{i \in \mathbb{Z}} \epsilon^\alpha \tau_i^\epsilon \delta_{\epsilon i}, \tag{3.125}$$

which now converge almost surely to m, where of course the the distribution function of this m is used as U in the construction of the τ_i^ϵ.

The existence of a non-trivial scaling limit for this model have far-reaching consequences for its long time asymptotics. In particular, it implies so-called *aging behavior*. This notion refers to the long-time behavior of of certain *correlation functions* of the process, e.g.

$$R(t_w, t) \equiv \mathbb{P}[X_{t+t_w} = X_{t_w}]. \tag{3.126}$$

One says that a process shows aging, if

$$\lim_{t_w \uparrow \infty} R(t_w, \theta t) = f(\theta), \tag{3.127}$$

for some non-trivial function f. Now in our case we have that

$$R(t_w, \theta t) = \mathbb{P}[X_{t_w(1+\theta} = X_{t_w}] = \mathbb{P}[X^{t_w^{-1/\alpha}}_{1+\theta} = X^{t_w^{-1/\alpha}}_1]. \tag{3.128}$$

Again we may expect this to converge to

$$\mathbb{P}\left[B_{\Sigma(1+\theta)} = B_{\Sigma(1)}\right], \tag{3.129}$$

which would be our desired *ageing function* expressed in term of the limiting process.

Bibliography

[1] A. Bovier. Stochastic processes 1. Discrete time. Lecture notes, Berlin Mathematical School, 2006.

[2] A. Bovier. Markov processes. Lecture notes, Bonn University, 2009.

[3] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.

[4] Frank den Hollander. *Large deviations*, volume 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000.

[5] L. R. G. Fontes, M. Isopi, and C. M. Newman. Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension. *Ann. Probab.*, 30(2):579–604, 2002.

[6] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.

[7] Masatoshi Fukushima, Yōichi Ōshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1994.

[8] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[9] J. Jacod. *Calcul stochastique et problèmes de martingales*. Springer, New York, 1979.

[10] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*. Graduate Texts in Mathematics. Springer, New York, 1988.

[11] M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York, 1983.

[12] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge Univer-

sity Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

Index