PATH PROPERTIES OF ITÔ PROCESSES

GUOHUAN ZHAO

1. Some Properties of Brownian Motion

Theorem 1.1 (Exponential martingale inequality)**.** *Let M_t be a continuous martingale, τ a bounded stopping time, then*

$$\mathbf{P}\left(\sup_{t\leqslant\tau}|M_t|>\lambda\ \&\ \langle M\rangle_\tau\leqslant\mu\right)\leqslant 2\mathrm{e}^{-\frac{\lambda^2}{2\mu}}.$$

We need

Lemma 1.2. *Let W be a 1-dimensional Brownian motion. Then for any $\lambda, t>0$*

$$\mathbf{P}\left(\sup_{s\in[0,t]}|W_s|>\lambda\right)\leqslant 2\mathrm{e}^{-\frac{\lambda^2}{2t}}$$

Proof. Let $X_t=\mathrm{e}^{a|W_t|}$ with $a>0$. Since $x\mapsto \mathrm{e}^{a|x|}$ is a convex function, X_t is a submartingale. By Doob's inequality, we have

$$\mathbf{P}\left(W_t^*>\lambda\right)=\mathbf{P}\left(X_t^*>\mathrm{e}^{a\lambda}\right)\leqslant \mathrm{e}^{-a\lambda}\mathbf{E}X_t=\frac{\mathrm{e}^{-a\lambda}}{\pi t}\int_0^\infty \mathrm{e}^{ax-\frac{x^2}{2t}}\mathrm{d}x=2\mathrm{e}^{\frac{a^2t}{2}-a\lambda}.$$

Taking $a=\lambda/t$, we obtain

$$\mathbf{P}\left(W_t^*>\lambda\right)\leqslant 2\mathrm{e}^{-\frac{\lambda^2}{2t}}.$$

□

Proof of Theorem 1.1. By Dambis-Dubins-Schwarz Theorem, M_t is a time change of a Brownian motion W_t. So the desired probability is bounded by

$$\mathbf{P}\left(\sup_{t\leqslant T}|W_t|>\lambda\ \&\ \langle W\rangle_T<\mu\right),$$

where T is a stopping time. Since $\langle W\rangle_T=T$, the probability above is in turn bounded by

$$\mathbf{P}\left(\sup_{t\leqslant\mu}|W_t|>\lambda\right)\leqslant 2\mathrm{e}^{-\frac{\lambda^2}{2\mu}},$$

due to Lemma 1.2. □

The next result, which is known as the law of iterated logarithm shows in particular that Brownian paths are not $\frac{1}{2}$-Hölder continuous.

Theorem 1.3 (law of iterated logarithm)**.** *Let $(W_t)_{t\geq 0}$ be a Brownian motion. For $s\geq 0$,*

$$\mathbf{P}\left(\liminf_{t\to 0}\frac{W_{t+s}-W_s}{\sqrt{2t\log\log\frac{1}{t}}}=-1,\limsup_{t\to 0}\frac{W_{t+s}-W_s}{\sqrt{2t\log\log\frac{1}{t}}}=1\right)=1.$$

3 weeks.

Proof. Thanks to the symmetry and invariance by translation of the Brownian motion, it suffices to show that:

$$\mathbf{P}\left(\limsup_{t\to 0}\frac{W_t}{\sqrt{2t\log\log\frac{1}{t}}}=1\right)=1.$$

Let us first prove that

$$\mathbf{P}\left(\limsup_{t\to 0}\frac{W_t}{\sqrt{2t\log\log\frac{1}{t}}}\leqslant 1\right)=1.$$

Let us denote $h(t)=\sqrt{2t\log\log\frac{1}{t}}$. Let $\alpha,\beta>0$, from Doob's maximal inequality applied to the martingale $\left(e^{\alpha W_t-\frac{\alpha^2}{2}t}\right)_{t\geq 0}$, we have for $t\geq 0$:

$$\mathbf{P}\left(\sup_{0\leqslant s\leqslant t}\left(W_s-\frac{\alpha}{2}s\right)>\beta\right)=\mathbf{P}\left(\sup_{0\leqslant s\leqslant t}e^{\alpha W_s-\frac{\alpha^2}{2}s}>e^{\alpha\beta}\right)\leqslant e^{-\alpha\beta}.$$

Let now $\theta,\delta\in(0,1)$. Using the previous inequality for every $n\in\mathbb{N}$ with $t=\theta^n,\alpha=\frac{(1+\delta)h(\theta^n)}{\theta^n},\beta=\frac{1}{2}h(\theta^n)$, yields when $n\to+\infty$,

$$\mathbf{P}\left(\sup_{0\leqslant s\leqslant\theta^n}\left(W_s-\frac{(1+\delta)h(\theta^n)}{2\theta^n}s\right)>\frac{1}{2}h(\theta^n)\right)=O\left(\frac{1}{n^{1+\delta}}\right).$$

Therefore from Borel-Cantelli lemma, for almost every $\omega\in\Omega$, we may find $N(\omega)\in\mathbb{N}$ such that for $n\geq N(\omega)$,

$$\sup_{0\leqslant s\leqslant\theta^n}\left(W_s(\omega)-\frac{(1+\delta)h(\theta^n)}{2\theta^n}s\right)\leqslant\frac{1}{2}h(\theta^n).$$

But,

$$\sup_{0\leqslant s\leqslant\theta^n}\left(W_s(\omega)-\frac{(1+\delta)h(\theta^n)}{2\theta^n}s\right)\leqslant\frac{1}{2}h(\theta^n)$$

implies that for $\theta^{n+1}\leqslant t\leqslant\theta^n$,

$$W_t(\omega)\leqslant\sup_{0\leqslant s\leqslant\theta^n}W_s(\omega)\leqslant\frac{1}{2}(2+\delta)h(\theta^n)\leqslant\frac{(2+\delta)h(t)}{2\sqrt{\theta}}.$$

We conclude:

$$\mathbf{P}\left(\limsup_{t\to 0}\frac{W_t}{\sqrt{2t\log\log\frac{1}{t}}}\leqslant\frac{2+\delta}{2\sqrt{\theta}}\right)=1.$$

Letting now $\theta\to 1$ and $\delta\to 0$ yields

$$\mathbf{P}\left(\limsup_{t\to 0}\frac{W_t}{\sqrt{2t\log\log\frac{1}{t}}}\leqslant 1\right)=1.$$

Let us now prove that

$$\mathbf{P}\left(\limsup_{t\to 0}\frac{W_t}{\sqrt{2t\log\log\frac{1}{t}}}\geq 1\right)=1.$$

Let $\theta\in(0,1)$. For $n\in\mathbb{N}$, we denote

$$A_n=\left\{\omega,W_{\theta^n}(\omega)-W_{\theta^{n+1}}(\omega)\geq(1-\sqrt{\theta})h(\theta^n)\right\}.$$

Let us prove that $\sum \mathbf{P}(A_n) = +\infty$. The basic inequality

$$\int_a^{+\infty} e^{-\frac{u^2}{2}} du \geq \frac{a}{1+a^2} e^{-\frac{a^2}{2}},$$

implies

$$\mathbf{P}(A_n) = \frac{1}{\sqrt{2\pi}} \int_{a_n}^{+\infty} e^{-\frac{u^2}{2}} du \geq \frac{a_n}{1+a_n^2} e^{-\frac{a_n^2}{2}},$$

with

$$a_n = \frac{(1-\sqrt{\theta})h(\theta^n)}{\theta^{n/2}\sqrt{1-\theta}}.$$

When $n \to +\infty$,

$$\frac{a_n}{1+a_n^2} e^{-\frac{a_n^2}{2}} = O\left(\frac{1}{n^{\frac{1+\theta-2\sqrt{\theta}}{1-\theta}}}\right),$$

therefore,

$$\sum \mathbf{P}(A_n) = +\infty.$$

As a consequence of the independence of the Brownian increments and of Borel-Cantelli lemma, the event

$$W_{\theta^n} - W_{\theta^{n+1}} \geq (1-\sqrt{\theta})h(\theta^n)$$

will occur almost surely for infinitely many n's. But, thanks to the first part of the proof, for almost every ω, we may find $N(\omega)$ such that for $n \geq N(\omega)$,

$$W_{\theta^{n+1}} > -2h(\theta^{n+1}) \geq -2\sqrt{\theta}h(\theta^n).$$

Thus, almost surely, the event $W_{\theta^n} > h(\theta^n)(1-3\sqrt{\theta})$ will occur for infinitely many n's. This implies

$$\mathbf{P}\left(\limsup_{t\to 0} \frac{W_t}{\sqrt{2t\log\log\frac{1}{t}}} \geq 1-3\sqrt{\theta}\right) = 1.$$

We finally get

$$\mathbf{P}\left(\limsup_{t\to 0} \frac{W_t}{\sqrt{2t\log\log\frac{1}{t}}} \geq 1\right) = 1.$$

by letting $\theta \to 0$. □

As a straightforward consequence, we may observe that the time inversion invariance property of Brownian motion implies:

Corollary 1.4. *Let $(W_t)_{t\geq 0}$ be a standard Brownian motion.*

$$\mathbf{P}\left(\liminf_{t\to+\infty} \frac{W_t}{\sqrt{2t\log\log t}} = -1, \limsup_{t\to+\infty} \frac{W_t}{\sqrt{2t\log\log t}} = 1\right) = 1.$$

2. Support theorem

Let

$$\sigma : \mathbb{R}_+ \times \Omega \to \mathbb{R}^{d\times d},\ b : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d \ \text{ and } \ a_t = \frac{1}{2}\sigma_t\sigma_t^T.$$

Set

$$x_t = \int_0^t \sigma_s \cdot \mathrm{d}W_s + \int_0^t b_s \mathrm{d}s. \tag{2.1}$$

For simplicity, we always assume that $a \in \mathbb{S}^d_\delta$.

The following result is a simplify version of Stroock-Varadhan's support theorem, which is taken from [Bas98].

Theorem 2.1 (Support theorem)**.** *Suppose σ, σ^{-1} and b are bounded, x_t is given by (2.1). Suppose $\varphi : [0,1] \to \mathbb{R}^d$ is continuous with $\varphi(0) = 0$. Then for each $\varepsilon > 0$, there exists a constant $c > 0$ depending only on ε, the modulus of continuity of φ, and the bounds on b, σ and σ^{-1} such that*

$$\mathbf{P}\left(\sup_{t\in[0,1]} |x_t - \varphi(t)| \leqslant \varepsilon\right) \geqslant c. \tag{2.2}$$

This can be interpreted as saying that the graph of x_t remains within an ε-tube around φ with positive probability.

To prove Theorem 2, we need some auxiliary lemmas.

By Lemma 1.2, there is a constant $\delta_0 > 0$ such that

$$\inf_{|x|\leqslant 1/3} \mathbf{P}\left(\sup_{t\in[0,\delta_0]} |W^x_t| \leqslant 1\right) \geqslant 5/6. \tag{2.3}$$

Lemma 2.2. *Let W be a 1-dimensional Brownian motion. For any $\varepsilon > 0$ and $T > 0$, there is a constant $c(\varepsilon, T) > 0$ such that*

$$\mathbf{P}\left(\sup_{s\in[0,T]} |W_t| \leqslant \varepsilon\right) \geqslant c(\varepsilon, T).$$

Proof. We assume $T = 1$. By the scaling property of Brownian motion ($\varepsilon^{-1}W_t \overset{d}{=} W_{\varepsilon^{-2}t}$), we only need to show

$$\mathbf{P}\left(\sup_{t\in[0,\varepsilon^{-2}]} |W_t| \leqslant 1\right) \geqslant c(\varepsilon) > 0.$$

It is easy to see that $\inf_{|x|\leqslant 1/3} \mathbf{P}(|W^x_{\delta_1}| \leqslant 1/3) \geqslant 1/3$ (for some $0 < \delta_1 \ll 1$), which together with (2.3) implies that

$$\inf_{|x|\leqslant 1/3} \mathbf{P}\left(\sup_{t\in[0,\delta_2]} |W^x_t| \leqslant 1,\ |W^x_{\delta_2}| \leqslant 1/3\right) \geqslant \frac{1}{6}, \quad \delta_2 = \delta_0 \wedge \delta_1 > 0.$$

By the Markov property of W,

$$\inf_{|x|\leqslant 1/3} \mathbf{P}\left(\sup_{t\in[0,k\delta_2]} |W^x_t| \leqslant 1,\ |W^x_{\delta_2}| \leqslant 1/3\right) \geqslant 6^{-k}.$$

Letting $k = [\varepsilon^{-2}\delta_2^{-1}] + 1$, we get

$$\mathbf{P}\left(\sup_{t\in[0,\varepsilon^{-2}]} |W_t| \leqslant 1\right) \geqslant \inf_{|x|\leqslant 1/3} \mathbf{P}\left(\sup_{t\in[0,\varepsilon^{-2}]} |W_t^x| \leqslant 1,\ |W_{\delta_2}^x| \leqslant 1/3\right)$$
$$\geqslant 6^{-\varepsilon^{-2}\delta_2^{-1}-1} =: c(\varepsilon) > 0.$$

□

Lemma 2.3. *Suppose $X_0 = 0$, $X_t = M_t + A_t$ is a continuous semimartingale with $\mathrm{d}A_t/\mathrm{d}t$ and $\mathrm{d}\langle M\rangle_t/\mathrm{d}t$ bounded above by N_1 and $\mathrm{d}\langle M\rangle_t/\mathrm{d}t$ bounded below by $N_2 > 0$. If $\varepsilon > 0$ and $T > 0$, then*

$$\mathbf{P}\left(\sup_{t\in[0,T]} |X_t| < \varepsilon\right) \geqslant c(\varepsilon, T, N_1, N_2) > 0.$$

Proof. Let $\tau(t) = \inf\{u > 0 : \langle M\rangle_u > t\}$. Then $\tau(t) \asymp t$, and $B_t = M_{\tau(t)}$ is a Brownian motion due to Lemma 2.2. Then $Y_t := X_{\tau(t)} = B_t + \int_0^t b_s \mathrm{d}s$ with $|b_s| \leqslant C(N_1, N_2)$. Our assertion will follow if we can show

$$\mathbf{P}\left(\sup_{t\in[0,T]} |Y_t| \leqslant \varepsilon\right) \geqslant c > 0.$$

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

$$\mathrm{d}\mathbf{Q}/\mathrm{d}\mathbf{P} = \mathcal{E}_T(-b) := \exp\left(-\int_0^T b_s \mathrm{d}B_s - \frac{1}{2}\int_0^T |b_s|^2 \mathrm{d}s\right) \quad \text{on} \quad \mathcal{F}_T.$$

By Girsanov's theorem, under $\mathbf{Q}$, Y_t is a Brownian motion. Therefore,

$$\mathbf{Q}(A) \geqslant c > 0, \quad A = \left\{\sup_{t\in[0,T]} |Y_t| \leqslant \varepsilon\right\}.$$

By Hölder's inequqlity,

$$c \leqslant \mathbf{Q}(A) \leqslant E^{\mathbf{P}}(\mathcal{E}_T(-b)\mathbf{1}_A) \leqslant [E^{\mathbf{P}}\mathcal{E}_T^2(-b)]^{\frac{1}{2}}[\mathbf{P}(A)]^{\frac{1}{2}}.$$

Since b is bounded, it is easy to verify that $E^{\mathbf{P}}\mathcal{E}_T^2(-b) < \infty$. This yields $\mathbf{P}(A) \geqslant c > 0$. □

Now we are on the point to give

Proof of Theorem 2. Step 1: We first consider the case and $\varphi = 0$. Let $z \in \partial B_{\varepsilon/4}$. Applying Itô's formula with $f(x) = |x - z|^2$ and setting $y_t = |x_t - z|^2$, then

$$y_t = z^2 + \int_0^t (x_s - z)\cdot \mathrm{d}x_s + \int_0^t \mathrm{tr}a_s \mathrm{d}s, \quad \langle y\rangle_t = \int_0^t (x_s - z)^T a_s (x_s - z)\mathrm{d}s \asymp y_t$$

$\langle y\rangle_t \geqslant c\varepsilon^2$ before $\tau := \inf\{s > 0 : |y_s - y_0| \geqslant (\varepsilon/8)^2\}$. If we set z_t equal to y_t for $t \leqslant \tau$ and equal to some Brownian motion for t larger than this stopping time, then Lemma 2.3 applies (for z_t) and

$$\mathbf{P}\left(\sup_{t\in[0,T]} |x_t| \leqslant \varepsilon\right) \geqslant \mathbf{P}\left(\sup_{t\in[0,T]} |y_t - y_0| \leqslant (\varepsilon/8)^2\right) = \mathbf{P}\left(\sup_{t\in[0,T]} |z_t - z_0| \leqslant (\varepsilon/8)^2\right) > 0.$$

Step 2: Without loss of generality, we may assume φ is differentiable with a derivative bounded by a constant. Define a new probability measure $\mathbf{Q}$ by

$$\mathrm{d}\mathbf{Q}/\mathrm{d}\mathbf{P} = \exp\left(-\int_0^T \varphi'(s)\sigma_s^{-1}\mathrm{d}W_s - \frac{1}{2}\int_0^T |\varphi'(s)\sigma_s^{-1}|^2\mathrm{d}s\right) \quad \text{on} \quad \mathcal{F}_T.$$

Noting that

$$\left\langle -\int_0^\cdot \varphi'(s)\sigma_s^{-1}\mathrm{d}W_s, x\right\rangle_t = \int_0^t \varphi'(s)\mathrm{d}s = -\varphi(t).$$

So by the Girsanov theorem, under $\mathbf{Q}$ each component of x_t is a semimartingale and $n_t^i := x_t^i - \int_0^t b_s^i\mathrm{d}s - \varphi^i(t)$ is a martingale for each $i = 1, \cdots, d$. Therefore,

$$B_t := \int_0^t \sigma_s^{-1}\mathrm{d}n_s$$

is a continuous local martingale with $\langle B^i, B^j\rangle_t = \delta_{ij}t$ under $\mathbf{Q}$. Therefore B_t is a d-dimensional Brownian motion udner $\mathbf{Q}$. Since

$$x_t - \varphi(t) = \int_0^t \sigma_s\mathrm{d}B_s + \int_0^t b_s\mathrm{d}s,$$

by *Step 1*, $\mathbf{Q}(\sup_{t\in[0,T]}|x_t - \varphi(t)| < \varepsilon) \geqslant c > 0$. similarly to the last paragraph of the proof for Lemma 2.3, we conclude

$$\mathbf{P}\left(\sup_{t\in[0,T]}|x_t - \varphi(t)| < \varepsilon\right) \geqslant c > 0.$$

□

3. ABP estimate and Generalized Itô's formula

Below we will use the an analytic result due to Alexsandroff to study the Itô process given by (2.1). For simplicity, *in this section, we assume that $b = 0$.*

Proposition 3.1 (Alexsandroff)**.** *Let f be a nonnegative function on B_2 such that f^d has finite integral over B_2 and $f = 0$ outside B_2. Then there exists a* **nonpositive convex** *function u on B_2 such that*

(i) for any $x \in B_2$,

$$|u(x)| \leqslant C\left(\int_{B_2} f^d dx\right)^{\frac{1}{d}}; \tag{3.1}$$

(ii) for any constant $a \in \mathbb{S}_+^d$, $\varepsilon > 0$ and $x \in B_2$,

$$a_{ij}\partial_{ij}u_\varepsilon(x) \geqslant d\sqrt[d]{\det a}\, f_\varepsilon(x), \tag{3.2}$$

where $u_\varepsilon = u * \zeta_\varepsilon$, and ζ_ε is a standard mollifier.

(3.1) is called Alexandroff–Bakelman–Pucci estimate in PDE literature.

In Appendix A, we provide the proof for Proposition 3.1 based on the very initial knowledge of the solvability of the following Monge–Ampère equations and estimates of its solutions:

$$\det \nabla^2 u(x) = f \quad \text{in } D, \tag{3.3}$$

which, actually, after a long development became also one of the cornerstones of the theory of fully nonlinear elliptic partial differential equations.

Set

$$\tau_R(x) = \inf\{t > 0 : x + x_t \notin B_R\}.$$

Proposition 3.1 implies

Theorem 3.2 (Krylov [Kry09]). *There is a constant $C(d)$ such that for any $R > 0$, and nonnegative Borel f given on $\mathbb{R}^d$, we have*

$$\mathbf{E}\int_0^{\tau_R(x)} f(x + x_t)\sqrt[d]{\det a_t}\,\mathrm{d}t \leqslant C(d)R\|f\|_{L^d(B_R)}. \tag{3.4}$$

Proof. By scaling, we only need to consider the case $R = 1$. We can assume $f \in C_c^\infty(B_1)$.

By Itô's formula,

$$u_\varepsilon(x + x_{t\wedge\tau_1(x)}) - u_\varepsilon(x) = \int_0^{t\wedge\tau_1(x)} a_s^{ij}\partial_{ij}u_\varepsilon(x + x_s)\mathrm{d}s + m_{t\wedge\tau_1(x)}$$

Taking expectation, letting $t \to \infty$ and using Proposition 3.1, we get

$$\int_0^{\tau_1(x)} \sqrt[d]{\det a_t}\, f_\varepsilon(x + x_t)\mathrm{d}t \leqslant d^{-1}\int_0^{\tau_1(x)} a_t^{ij}\partial_{ij}u_\varepsilon(x + x_t)\mathrm{d}t$$
$$\leqslant \frac{2}{d}\sup_{x\in B_1}|u(x)| \leqslant C(d)\|f\|_{L^d(B_2)} = C(d)\|f\|_{L^d(B_1)}.$$

Letting $\varepsilon \to 0$, we obtain our assertion. □

We should point out that here we do not need to assume $a \in \mathbb{S}_\delta^d$.

Remark 3.1. *(i) (3.4) implies that if x_t is a Itô's process given by (2.1) with σ non-degenerate, then the process $t \mapsto \int_0^t f(x_s)\mathrm{d}s$ is well-defined.*

(ii) Suppose x_t is a Itô process given by (2.1), $a \in \mathbb{S}_\delta^d$ and b satisfying $|b_t| \leqslant \mathfrak{b}(x_t)$ with some $\mathfrak{b} \in L^d$. In this case, Krylov [Kry21] also proved (3.4) with $\|f\|_{L^d(D)}$ replaced by $\|f\|_{L^{d-\varepsilon}(D)}$ for some $\varepsilon = \varepsilon(d, \delta, \|\mathfrak{b}\|) > 0$.

Theorem 3.2 as many results below admits a natural generalization with conditional expectations. This generalization is obtained by tedious and not informative repeating the proof with obvious changes. We mean the following which we call the conditional version of Theorem 3.2 . Let γ be a finite stopping time, then

$$\mathbf{E}\left[\int_\gamma^{\tau_R(x)} f(x + x_t)\sqrt[d]{\det a_t}\,\mathbf{1}_{\{\gamma\leqslant\tau_R(x)\}}\,\mathrm{d}t\Big|\mathcal{F}_\gamma\right] \leqslant C(d)R\|f\|_{L_d(B_R)}. \tag{3.5}$$

Theorem 3.3. *There are constants C, μ depending only on d, such that*

$$\mathbf{E}\exp\left(\frac{\mu\tau_R(x)}{\delta R^2}\right) \leqslant C, \quad \forall R \in (0,\infty) \quad \text{and} \quad x \in B_R. \tag{3.6}$$

In particular, for each $\lambda > 0$,

$$\mathbf{P}(\tau_R(x) \geqslant \lambda) \leqslant C\exp\left(-\frac{\mu\lambda}{\delta R^2}\right). \tag{3.7}$$

Lemma 3.4.

$$\mathbf{E}\tau_R(x)^n \leqslant n!(CR^2/\delta)^n.$$

Proof. We can assume $x = 0$.

We claim that
$$I_n(t) := \mathbf{E}\left([\tau_R - t]_+^n | \mathcal{F}_t\right) \leqslant n!(CR^2/\delta)^n. \tag{3.8}$$

Of course, (3.8) implies our desired result.

When $n = 1$, (3.5) implies (3.8). If our assertion is true for a given n, then

$$\begin{aligned}
I_{n+1}(t) =&(n+1)!\mathbf{E}\left(\int \mathbf{1}_{t<t_1<\cdots<t_{n+1}<\tau_R} \mathrm{d}t_1 \cdots \mathrm{d}t_{n+1} \Big| \mathcal{F}_t\right) \\
=&(n+1)! \int \mathrm{d}t_1 \cdots t_{n+1} \mathbf{E}\left(\mathbf{1}_{t<t_1<\cdots<t_n<\tau_R} \mathbf{1}_{t_n<t_{n+1}<\tau_R} \Big| \mathcal{F}_t\right) \\
=&(n+1)! \int \mathrm{d}t_1 \cdots t_{n+1} \mathbf{E}\left[\mathbf{1}_{t<t_1<\cdots<t_n<\tau_R} \mathbf{E}\left(\mathbf{1}_{t_n<t_{n+1}<\tau_R} \Big| \mathcal{F}_{t_n}\right) \Big| \mathcal{F}_t\right] \\
=&(n+1)\mathbf{E}\left[n! \int \mathbf{1}_{t<t_1<\cdots<t_n<\tau_R} \mathrm{d}t_1 \cdots t_n \int \mathbf{E}\left(\mathbf{1}_{t_n<t_{n+1}<\tau_R} \Big| \mathcal{F}_{t_n}\right) \mathrm{d}t_{n+1} \Big| \mathcal{F}_t\right] \\
=&(n+1)\mathbf{E}\left\{[\tau_R - t]_+^n \mathbf{E}\left[\int_{t_n}^{\tau_R} \mathbf{1}_{B_R}(x_{t_{n+1}}) \mathrm{d}t_{n+1} | \mathcal{F}_{t_n}\right] \Big| \mathcal{F}_t\right\} \\
\overset{(3.5)}{\leqslant}& (n+1)C\delta^{-1}R^2 I_n(t) \overset{(3.8)}{\leqslant} (n+1)!(CR^2/\delta)^{n+1}.
\end{aligned}$$

So we get what we desired. $\square$

Exercise 3.1. *Let B be a one-dimensional BM. Let $I = (-1, 1)$. Prove that*
$$\mathbf{E}\tau_I^n \leqslant C^n n!.$$
Using this to give another proof for (3.6)*.*

Corollary 3.3 basically says that τ_R is smaller than a constant times R^2. We want to show that in a sense the converse is also true: R^2 is basically smaller than a constant times τ_R.

Lemma 3.5. *There exists C depending only on d such that*
$$\mathbf{P}(\tau_R/R^2 \leqslant t) \leqslant C\delta^{-1}t, \quad \forall t, R > 0. \tag{3.9}$$

Proof. We only need to prove the case $R = 1$. Let ϕ be a C^2 function that is zero at 0, one on ∂B_1, with $\partial_{ij}\phi$ bounded by a constant. By Itô's formula
$$\mathrm{d}\phi(x_t) = \nabla\phi(x_t) \cdot \sigma_t \mathrm{d}W_t + a_t^{ij}\partial_{ij}\phi(x_t)\mathrm{d}t,$$
which yields that
$$\phi(x_{t\wedge\tau_1}) = \mathbf{E}\int_0^{t\wedge\tau_1} a_s^{ij}\partial_{ij}\phi(x_s)\mathrm{d}s \leqslant C\delta^{-1}t.$$
Since $\phi(x_{t\wedge\tau_1}) \geqslant \mathbf{1}_{\{\tau_1\leqslant t\}}$, we get $\mathbf{P}(\tau_1 \leqslant t) \leqslant C\delta^{-1}t$. $\square$

Lemma 3.6. *There is a constant $\mathfrak{R} = R(d, \delta)$ such that*
$$\mathbf{E}\exp(-\tau_{\mathfrak{R}}) \leqslant 1/2.$$

Proof. Let X be a non-negative random variable, and let $F : \mathbb{R}_+ \to \mathbb{R}$ be a decreasing function with $F(\infty) = 0$. Then

$$\mathbf{E}F(X) = -\int_0^\infty F'(t)\mathbf{P}(X \leqslant t)\mathrm{d}t,$$

due to Fubini's theorem. Set $X = \tau_R$ and $F(t) = \mathrm{e}^{-t}$. Then

$$\mathbf{E}\mathrm{e}^{-\tau_R} = \int_0^\infty \mathrm{e}^{-t}\mathbf{P}(\tau_R \leqslant t)\mathrm{d}t \leqslant \int_0^\infty \mathrm{e}^{-t}[1 \wedge (C\delta^{-1}R^{-2}t)]\mathrm{d}t \leqslant C\delta^{-1}R^{-2}.$$

We set $\mathfrak{R} = \sqrt{2C/\delta}$. □

Exercise 3.2. *For any $R \in (0, \infty)$*

$$\mathbf{E}\exp\left(-\left(\mathfrak{R}/R\right)^2\tau_R\right) \leqslant 1/2. \tag{3.10}$$

Theorem 3.7. *For any $\kappa \in (0,1), R \in (0,\infty), x \in B_{\kappa R}$, and $\lambda \geqslant 0$,*

$$\mathbf{E}\exp\left(-\lambda\tau_R(x)\right) \leqslant 2e^{-\sqrt{\lambda}(1-\kappa)R/K}, \tag{3.11}$$

where $K = \mathfrak{R}/\log 2$.

Proof. Recall that $\tau_R(x)$ is the first exit time of $x + x_t$ from B_R. Let $\tau'_R(x)$ be the first exit time of $x + x_t$ from $B_{(1-\kappa)R}(x)$. It follows that in the proof of , we may assume that $\kappa = 0$ and $x = 0$. Then, as usual we may assume that R. In that case take N, to be specified later, and introduce τ^k, $k = 1, \cdots, N$, as the first exit time of x_t from $B_{k/N}$. Also set γ^k be the first exit times of x_t from $B_{N^{-1}}(x_{\tau^{k-1}})$ after τ^{k-1} ($\tau^{k-1} \leqslant \gamma^k \leqslant \tau^k$). obviously,

$$\tau_1 \geqslant (\gamma^1 - \tau_0) + (\gamma^2 - \tau^1) + \cdots + (\gamma^N - \tau^{N-1}).$$

By the conditional version of (3.10),

$$\mathbf{E}\left\{\exp\left[-\mathfrak{R}^2N^2(\gamma^k - \tau^{k-1})\right]|\mathcal{F}_{\tau^{k-1}}\right\} \leqslant 1/2.$$

Therefore,

$$\begin{aligned}
&\mathbf{E}\left[\exp\left(-\mathfrak{R}^2N^2\tau_1\right)\right]\\
\leqslant&\mathbf{E}\left[\prod_{k=1}^{N}\exp\left(-\mathfrak{R}^2N^2(\gamma^k - \tau^{k-1})\right)\right]\\
\leqslant&\mathbf{E}\left\{\prod_{k=1}^{N-1}\exp\left(-\mathfrak{R}^2N^2(\gamma^k - \tau^{k-1})\right)\mathbf{E}\left[\exp\left(-\mathfrak{R}^2N^2(\gamma^N - \tau^{N-1})\right)\Big|\mathcal{F}_{\tau^{N-1}}\right]\right\}\\
\leqslant&\frac{1}{2}\mathbf{E}\left[\prod_{k=1}^{N-1}\exp\left(-\mathfrak{R}^2N^2(\gamma^k - \tau^{k-1})\right)\right] \leqslant \cdots \leqslant (1/2)^N.
\end{aligned} \tag{3.12}$$

Choosing $N = [\sqrt{\lambda}/\mathfrak{R}]$, we get (3.11). □

Exercise 3.3. *For any $R, t > 0$,*

$$\mathbf{P}\left(\tau_R(x) \leqslant tR^2\right) \leqslant 2\exp\left(-\frac{\beta(1-\kappa)^2}{t}\right), \tag{3.13}$$

where $\beta = \beta(\mathfrak{R}) \in (0,1)$.

The above estimates for first exit times have many important applications.

Proposition 3.8. *For any $\kappa \in (0,1)$ there is a function $q(\gamma), \gamma \in (0,1)$, depending only on d, δ, κ and naturally, also on γ, such that for any $R \in (0,\infty), x \in B_{\kappa R}$, and closed $\Gamma \subset B_R$ satisfying $|\Gamma| \geqslant \gamma |B_R|$ we have*

$$\mathbf{P}\left(\sigma_\Gamma(x) \leqslant \tau_R(x)\right) \geqslant q(\gamma),$$

where $\sigma_\Gamma(x)$ is the first time the process $x + x_t$ hits Γ. Furthermore, $q(\gamma) \to 1$ as $\gamma \uparrow 1$.

Proof. By using scaling we reduce the general case to the one in which $R = 1$. In that case for any $\varepsilon > 0$ we have

$$\begin{aligned}\mathbf{P}\left(\sigma_\Gamma(x) > \tau_1(x)\right) &\leqslant \mathbf{P}\left(\tau_1(x) = \int_0^{\tau_1(x)} \mathbf{1}_{B_1\setminus\Gamma}(x + x_t)\,\mathrm{d}t\right)\\ &\leqslant \mathbf{P}\left(\tau_1(x) \leqslant \varepsilon\right) + \varepsilon^{-1}\mathbf{E}\int_0^{\tau_1(x)} I_{B_1\setminus\Gamma}(x + x_t)\,\mathrm{d}t.\end{aligned}$$

In light of Theorem 3.2, we can estimate the right-hand side and then obtain

$$\begin{aligned}\mathbf{P}\left(\sigma_\Gamma(x) > \tau_1(x)\right) &\leqslant 2e^{-C/\varepsilon} + C\varepsilon^{-1}\left|B_1\setminus\Gamma\right|^{1/d}\\ &\leqslant 2e^{-C/\varepsilon} + C\varepsilon^{-1}(1-\gamma)^{1/d}\end{aligned}$$

where the constants C depend only on d, δ, κ. By denoting

$$q(\gamma) = 1 - \inf_{\varepsilon>0}\left(2e^{-C/\varepsilon} + C\varepsilon^{-1}(1-\gamma)^{1/d}\right).$$

□

Note that in the above result, we have no assumption on the shape of the set Γ.

Exercise 3.4. *For any $\kappa \in (0,1)$, $R \in (0,\infty)$. For any $x \in B_1$ and $B_{\kappa R}(y) \subseteq B_R$, we have*

$$\mathbf{P}\left(\sigma_{B_{\kappa R}(y)}(x) < \tau_R(x)\right) \geqslant \zeta(\kappa) > 0,$$

where $\zeta(\kappa) > 0$ depends only on d, δ, and naturally, also on κ.

Hint: Using support theorem.

Theorem 3.9. *Let $p \geqslant d$. Then there exists constants C depending only on d, δ, such that for any $\lambda > 0$ and Borel nonnegative f given on $\mathbb{R}^d$ we have*

$$\mathbf{E}\int_0^\infty e^{-\lambda t} f\left(X_t\right)\mathrm{d}t \leqslant C\lambda^{\frac{d}{2p}-1}\left\|f\right\|_p. \tag{3.14}$$

Proof. Let γ be a stopping time and γ' be the first exit time of x_t from $B_R(x_\gamma)$ after γ. By the conditional version of (3.11),

$$\mathbf{E}\left[\exp\left(-\lambda(\gamma' - \gamma)\right)\Big|\mathcal{F}_\gamma\right] \leqslant 2e^{-\sqrt{\lambda}R/K}.$$

Choosing $R = K/\sqrt{\lambda}$, then

$$\mathbf{E}\left[\exp\left(-\lambda(\gamma' - \gamma)\right)\Big|\mathcal{F}_\gamma\right] \leqslant 2/e < 1.$$

Let $\tau^0 = 0$ and τ^k be the first exit time of x_t from $B_R(x_{\tau^{k-1}})$ after τ^{k-1}. As the proof for (3.12), we have

$$\mathbf{E}\mathrm{e}^{-\lambda\tau^k} = \mathbf{E}\prod_{i=1}^k \mathrm{e}^{-\lambda(\tau^k - \tau^{k-1})} \leqslant (2/e)^k. \tag{3.15}$$

If (3.15) holds, then

$$
\begin{aligned}
\mathbf{E}\int_0^\infty e^{-\lambda t} f\left(x_t\right)\mathrm{d}t \leqslant & \sum_{k=1}^\infty \mathbf{E}\left[e^{-\lambda\tau^{k-1}}\mathbf{E}\left(\int_{\tau^{k-1}}^{\tau^k} f\left(x_t\right)\mathrm{d}t\Big|\mathcal{F}_{\tau^{k-1}}\right)\right] \\
\overset{(3.4)}{\leqslant} & \sum_{k=1}^\infty \mathbf{E}\left(C\delta^{-1}R\|f\|_{L^d(B_R(x_{\tau^{k-1}}))}\mathrm{e}^{-\lambda\tau^{k-1}}\right) \\
\leqslant & C\delta^{-1}R^{2-\frac{d}{p}}\|f\|_p\sum_{k=0}^\infty \mathbf{E}\mathrm{e}^{-\lambda\tau^k} \leqslant C\delta^{-1}\|f\|_p\sum_{k=0}^\infty (2/e)^k \\
\leqslant & C\|f\|_p/\lambda^{1-(d/2p)}.
\end{aligned}
$$

□

Theorem 3.10 (Generalized Itô's formula, see Krylov-[Kry09]). *Let x_t be a Itô process given by (2.1). Suppose that $a \in \mathbb{S}_\delta^d$ and b are bounded, then for any $u \in W^{2,p}_{loc}$ with $p \geqslant d$, we have*

$$
u(x_t) - u(x_0) = \int_0^t \nabla u(x_s)\sigma_s \mathrm{d}W_s + \int_0^t a_s^{ij}\partial_{ij}u(x_s)\mathrm{d}s \tag{3.16}
$$

Proof. We only need to consider the case $u \in W^{2,d}$. Let $\eta \in C_c^\infty(B_1)$ with $\int \eta = 1$. Set $\eta_\varepsilon(x) = \varepsilon^{-d}\eta(x/\varepsilon)$ and $u_\varepsilon = u * \eta_\varepsilon$. By Itô's formula,

$$
u_\varepsilon(x_t) - u_\varepsilon(x_0) = \int_0^t \nabla u_\varepsilon(x_s)\sigma_s \mathrm{d}W_s + \int_0^t a_s^{ij}\partial_{ij}u_\varepsilon(x_s))\mathrm{d}s. \tag{3.17}
$$

Fact: by Sobolev embedding theorem, we have

$$
W^{2,d} \hookrightarrow C_b; \quad \|\nabla u\|_{2d} \leqslant C(\|\nabla^2 u\|_{L^d} + \|\nabla u\|_{L^d}). \tag{3.18}
$$

Since $u \in C_b$, by letting $\varepsilon \to 0$, one sees the left-hand side of (3.17) goes to $u(x_t) - u(x_0)$ as $\varepsilon \to 0$. For the right-hand side of (3.17). By Doob's maximal inequality

$$
\begin{aligned}
& \mathbf{E}\sup_{t\in[0,T]}\left|\int_0^t \nabla u_\varepsilon(x_s)\sigma_s \mathrm{d}W_s - \int_0^t \nabla u_{\varepsilon'}(x_s)\sigma_s\mathrm{d}W_s\right|^2 \\
\leqslant & C\mathbf{E}\int_0^T |\nabla u_\varepsilon - \nabla u_{\varepsilon'}|^2(x_s)\mathrm{d}s \leqslant C\|\nabla u_\varepsilon - \nabla u_{\varepsilon'}\|_{L^{2d}}^2 \\
\overset{(3.18)}{\leqslant} & C\|u_\varepsilon - u_{\varepsilon'}\|_{W^{2,d}} \to 0, \quad \varepsilon, \varepsilon' \to 0.
\end{aligned}
$$

Similarly, we can also show that the second integral on the right-hand side of (3.17) also converges to $\int_0^t a_s^{ij}\partial_{ij}u(x_s)\mathrm{d}s$ □

Remark 3.2. *The above generalized Itô's formula also holds for Itô process given by (2.1), where $a \in \mathbb{S}_\delta^d$, and b satisfying $|b_t| \leqslant \mathfrak{b}(x_t)$ with $\mathfrak{b} \in L^d$.*

References

[Bas98] Richard F Bass. *Diffusions and elliptic operators*. Springer Science & Business Media, 1998.

[Kry09] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.

[Kry21] Nicolai V Krylov. On diffusion processes with drift in L_d. *Probability Theory and Related Fields*, 179(1):165–199, 2021.

Appendix A. Monge-Ampère Equation

Lemma A.1 (Area formula)**.** *Consider a locally Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}^d$ and a Borel set $A \subseteq \mathbb{R}^d$. Then the function $y \mapsto N_A(y) := \operatorname{card}\{f^{-1}(y) \cap A\}\}$ is measurable and*

$$\int_A |\det(\nabla f(x))| \mathrm{d}x = \int_{\mathbb{R}^n} N_A(y) \mathrm{d}y \geqslant \mathscr{L}^d(f(A)).$$

Consequently, for any $g \geqslant 0$,

$$\int_{f(A)} g(y) \mathrm{d}y \leqslant \int_A g(f(x)) |\det \nabla f(x)| \mathrm{d}x. \tag{A.1}$$

To motivate the definition of weak solutions to (3.3), given an open set $D \subset \mathbb{R}^n$, consider $u : D \to \mathbb{R}$ a convex function of class C^2 satisfying (3.3) for some $f : D \to \mathbb{R}^+$. Then given any Borel set $E \subset D$, it follows by the area formula that

$$\int_E f \, \mathrm{d}x = \int_E \det D^2 u \, \mathrm{d}x = |\nabla u(E)|.$$

Notice that while the above argument needs u to be of class C^2, the identity

$$\int_E f = |\nabla u(E)|$$

makes sense if u is only of class C^1. To find a definition when u is merely convex one could try to replace the gradient $\nabla u(x)$ with the subdifferential $\partial u(x)$ and ask for the above equality to hold for any Borel set E. Here $\partial u(x)$ is given by

$$\partial u(x) := \left\{ p \in \mathbb{R}^d : u(y) \geqslant u(x) + \langle p, y - x \rangle \quad \forall y \in D \right\}.$$

This motivates the following definition:

Definition A.1. *Given an open set $D \subset \mathbb{R}^n$ and a convex function $u : D \to \mathbb{R}$, we define the Monge-Ampère measure associated to u by*

$$\mu_u(E) := \left| \bigcup_{x \in E} \partial u(x) \right|$$

The basic idea of Alexandrov was to say that u is a weak solution of (3.3) if $\mu_u|_D = \nu|_D$.

Lemma A.2. *Let $u, v : D \to \mathbb{R}$ be convex functions. Then*

$$\mu_{u+v} \geqslant \mu_u + \mu_v \quad \text{ and } \quad \mu_{\lambda u} = \lambda^n \mu_u \quad \forall \lambda > 0.$$

The following result is the celebrated Alexandrov maximum principle.

Theorem A.3. *Let D be an open bounded convex set, and let $u : D \to \mathbb{R}$ be a convex function such that $u|_{\partial D} = 0$. Then there exists a dimensional constant $C = C(d)$ such that*

$$|u(x)| \leqslant C(d) \operatorname{diam}(D)^{\frac{d-1}{d}} \operatorname{dist}(x, \partial D)^{\frac{1}{d}} |\partial u(D)|^{\frac{1}{d}}, \quad \forall x \in D. \tag{A.2}$$

Proof. Let $(x, u(x))$ be a point on the graph of u, and consider the convex "conical" function $y \mapsto \widehat{C}_x(y)$ with vertex at $(x, u(x))$ that vanishes on ∂D. Since $u \leqslant \widehat{C}_x$ in D (by the convexity of u), Lemma 2.7 implies that

$$\left| \partial \widehat{C}_x(x) \right| \leqslant \left| \partial \widehat{C}_x(D) \right| \leqslant |\partial u(D)|;$$

so, to conclude the proof, it suffices to bound $|\partial \widehat{C}_x(x)|$ from below. It is not hard to see

- $\partial \widehat{C}_x(x)$ contains the ball B_ρ with $\rho = |u(x)|/\mathrm{diam}(D)$
- $\partial \widehat{C}_x(x)$ contains a vector of norm $|u(x)|/\mathrm{dist}(x, \partial D)$

Thus,

$$\partial \widehat{C}_x(x) \supset B_\varrho(0) \cup \{q\}, \quad |q| = |u(x)|/\mathrm{dist}(x, \partial D).$$

Since $\partial \widehat{C}_x(x)$ is convex, it follows that $\partial \widehat{C}_x(x)$ contains the cone $\mathcal{C}$ generated by q and $\Sigma_q := \{p \in B_\rho : \langle p, q\rangle = 0\}$. Therefore

$$c(d)\rho^{d-1}|q| = |\mathcal{C}| \leqslant |\partial u(D)|.$$

□

Theorem A.4. *Let D be an open bounded convex set, and let ν be a Borel measure on D with $\nu(D) < \infty$. Then there exists a unique convex function $u : D \to \mathbb{R}$ solving the Dirichlet problem*

$$\begin{cases} \mu_u = v & \text{in } D \\ u = 0 & \text{on } \partial D \end{cases}$$

Proof. By the stability result proved in Lemma below, since any finite measure can be approximated in the weak* topology by a finite sum of Dirac deltas, we only need to solve the Dirichlet problem when $\nu = \sum_{i=1}^N \alpha_i \delta_{x_i}$ with $x_i \in D$ and $\alpha_i > 0$. To prove existence of a solution, we use the so-called Perron method: we define

$$\mathcal{S}[\nu] := \{v : \Omega \to \mathbb{R} \text{ convex} : \ v|_{\partial\Omega} = 0, \mu_v \geqslant \nu \text{ in } \Omega\}$$

and we show that the largest element in $\mathcal{S}[\nu]$ is the desired solution. We split the argument into several steps.

Step 1: $\mathcal{S}[\nu] \neq \emptyset$. To construct an element of $\mathcal{S}[\nu]$, we consider the "conical" function C_{x_i}, that is 0 on $\partial\Omega$ and takes the value -1 at its vertex x_i. The Monge–Ampère measure of this function is concentrated at x_i and has mass equal to some positive number β_i corresponding to the measure of the set of supporting hyperplanes at x_i. Now, consider the convex function $\bar{v} = \sum_{i=1}^N \lambda C_{x_i}$, where λ has to be chosen. We notice that $\bar{v}|_{\partial\Omega} = 0$. In addition, provided λ is sufficiently large, Lemma below implies that

$$\mu_{\bar{v}} \geqslant \sum_{i=1}^N \mu_{\lambda \hat{C}_{x_i}} = \sum_{i=1}^N \lambda^d \mu_{\hat{C}_{x_i}} = \sum_{i=1}^N \lambda^d \beta_i \delta_{x_i} \geqslant \sum_{i=1}^N \alpha_i \delta_{x_i} = \nu.$$

This yields $\bar{v} \in \mathcal{S}[\nu]$.

Step 2: $v_1, v_2 \in \mathcal{S}[\nu] \Rightarrow w := \max\{v_1, v_2\} \in \mathcal{S}[\nu]$. Set

$$\Omega_0 := \{v_1 = v_2\}, \quad \Omega_1 := \{v_1 > v_2\}, \quad \text{and} \quad \Omega_2 := \{v_1 < v_2\}$$

Also, given a Borel set $E \subseteq \Omega$, consider $E_i = E \cap \Omega_i$.

Since Ω_1 and Ω_2 are open sets, $w|_{\Omega_1} = v_1$ and $w|_{\Omega_2} = v_2$,

$$\partial w(E_1) = \partial v_1(E_1), \quad \partial w(E_2) = \partial v_2(E_2).$$

In addition, since $w = v_1$ on Ω_0 and $w \geqslant v_1$ everywhere else, we have

$$\partial v_1(E) \subseteq \partial w(E_0).$$

Therefore,

$$\mu_w(E) \geqslant \mu_{v_1}(E_0 \cup E_1) + \mu_{v_2}(E_2) \geqslant \nu(E).$$

Step 3: $u := \sup_{v\in S[\nu]} v$ belongs to $S[\nu]$. Let $w_m \uparrow u$ locally uniformly. Then $\mu_{w_m} \rightharpoonup *\mu_u$. Also, we deduce immediately that $u|_{\partial\Omega} = 0$ by construction; hence, $u \in \mathcal{S}[\nu]$.

Step 4: The measure μ_u is supported at the points $\{x_1, \cdots x_N\}$. Otherwise, there exists a set $E \subseteq D$ such that

$$E \cap \{x_1, \ldots, x_N\} = \emptyset \quad \text{and} \quad |\partial u(E)| = \mu_u(E) > 0$$

Therefore,

$$\left|\partial u(E)\backslash[\cup_{i=1}^N \partial u(x_i) \cup \partial u(\partial D)]\right| = |\partial u(E)| > 0$$

Let $x_0 \in E$ and $p \in \partial u(x_0)\backslash[\cup_{i=1}^N \partial u(x_i) \cup \partial u(\partial D)]$. Then there exists $\delta > 0$ such that

$$u \geqslant \ell_{x_0,p} + 2\delta \quad \text{on } \{x_1, \ldots, x_N\} \cup \partial\Omega, \tag{A.3}$$

where $\ell_{x_0,p}(x) = u(x_0) + p \cdot (x - x_0)$. Set $\bar{u} := \max\{\ell_{x_0,p} + \delta, u\} \gneqq u$. Notice that $\bar{u}$ is convex, $\bar{u} \geqslant u$, and it follows by (A.3) that $\bar{u} = u$ in a neighborhood of $\{x_1, \ldots, x_N\} \cup \partial\Omega$. In particular, $\bar{u}|_{\partial\Omega} = 0$ and $\partial\bar{u}\,(x_i) = \partial u\,(x_i)$, which implies that $u \log eqq \bar{u} \in \mathcal{S}[\nu]$. This is a contradiction.

Step 5: $\mu_u = \nu$. By *Step 3* and *Step 4*, we know that $\mu_u = \sum_{i=1}^N \beta_i \delta_{x_i}$ with $\beta_i \geqslant \alpha_i$. Assume that $\beta_1 = \mu_u(x_1) > \nu(x_1) = \alpha_1$. $\square$