Note 2

Stochastic differential equations

From the point of view of modelisation is interesting to study a (continuous) stochastic processes whose local behaviour is prescribed. For example: the process $(X_t)_{t\geqslant 0}$ could describe the current state of a system and given that state we know what its evolution in small time is. In order to model also random influence we could consider that increments over small times intervals are actually distributed like Gaussians (this is actually more general than it seems)

$$X_{t+\Delta t}-X_t\approx b(X_t)\Delta t+\sigma(X_t)G$$

where b, σ are given functions and G is a Gaussian random variable. The fact that we assume X_t to be the state of a system justify somehow the assumption that the coefficients depends on the randomness only via X_t. Let us assume that G has variance of order $(\Delta t)^\alpha$ for some $\alpha>0$ and that disturbances are independent over disjoint intervals. Then is not difficult to realise that, unless $\alpha=1/2$ the evolution of X over a finite time interval will not have a limit as we refine our approximation and take $\Delta t\to 0$ (For example one can look at the variance of X_t-X_0). This of course hints to the fact that, as a stochastic process X_t must be described by a semimartingale whose finite variation part described the deterministic part of the evolution and whose martingale part account for the random fluctuations, moreover we model the martingale part via a stochastic integral wrt to Brownian motion and the finite variation part via a differentiable function and we obtain that the local dynamics of X can be rigorously described by

$$\mathrm{d}X_t=f_t\mathrm{d}t+g_t\mathrm{d}B_t \tag{1}$$

where $(B_t)_{t\geqslant 0}$ is a family of Brownian motions and f,g two adapted processes. In general X will not be differentiable, but this decomposition is all what we need. In particular is unique.

Prove that if $\mathrm{d}X_t=f_t\mathrm{d}t+g_t\mathrm{d}B_t$ and also $\mathrm{d}X_t=f_t'\mathrm{d}t+g_t'\mathrm{d}B_t$ for adapted, continuous f,g,f',g' then we must have $f=f'$ and $g=g'$ almost surely.

A process like the one described by eq. (1) is called an Itô process. The functions f,g play the role of "stochastic derivatives". Prescribing the local dynamics of an Itô process is then done by prescribing a relation between these derivatives and the function itself, for example letting

$$f_t=b(X_t),\quad g_t=\sigma(X_t).$$

This leads to the notion of stochastic differential equation.

Definition 1. *Given vector fields $b,\sigma_\alpha\colon\mathbb{R}^n\to\mathbb{R}^n$ for $\alpha=1,\ldots,m$, an m-dimensional Brownian motion $(B_t)_{t\geqslant 0}$ and an initial condition $x\in\mathbb{R}^n$, the solution to the stochastic differential equation*

$$\mathrm{d}X_t=b(X_t)\mathrm{d}t+\sigma(X_t)\mathrm{d}B_t,\qquad X_0=x, \tag{2}$$

is an adapted stochastic process $(X_t)_{t\geqslant 0}$ such that

$$\int_0^t|b(X_s)|\mathrm{d}s+\sum_\alpha\int_0^t|\sigma_\alpha(X_s)|^2\mathrm{d}s<\infty,\quad a.s.$$

and

$$X_t = x_0 + \int_0^t b(X_s)\mathrm{d}s + \sum_\alpha \int_0^t \sigma_\alpha(X_s)\mathrm{d}B_s^\alpha, \quad t \geqslant 0. \tag{3}$$

We can generalise this definition by taking vector fields which depends on the time and on the randomness ω (in an adapted way). We can also relax it by requiring eq. (3) to be statisfied only on a finite interval $[0,T]$ or on a random interval $[0,\tau]$ where τ is a stopping time. In this case we say that X_t solves the SDE in the interval $[0,\tau]$ iff

$$X_{t\wedge\tau} = x_0 + \int_0^{t\wedge\tau} b(X_s)\mathrm{d}s + \sum_\alpha \int_0^{t\wedge\tau} \sigma_\alpha(X_s)\mathrm{d}B_s^\alpha, \qquad t \geqslant 0.$$

Localization of an SDE in a random interval is a useful technique, as usual in stochastic calculus.

Note that a solution to an SDE is a semi-martingale, and in particular an Itô process.

1 Existence and uniqueness

The basic existence and uniqueness results for SDEs holds, like in the case of ODE, for coefficients which are globally Lipshitz continuous. For the sake of a bit of generality we will take them dependent also on time (one easily observe that they could also be taken random, as long as they are adapted, but we will not do it).

Theorem 2. *Assume that* $b, \sigma_\alpha: \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ *satisfy for some* $K > 0$*,*

$$\begin{aligned} &|b(s,x)| + \sum_\alpha |\sigma_\alpha(s,x)| \leqslant K(1+|x|), \\ &|b(s,x) - b(s,y)| + \sum_\alpha |\sigma_\alpha(s,x) - \sigma_\alpha(s,y)| \leqslant K|x-y|, \qquad s \geqslant 0, \quad x,y \in \mathbb{R}^n. \end{aligned} \tag{4}$$

Then for any $x \in \mathbb{R}$ *there exists a solution to the SDE (2). This solution is adapted to the filtration generated by the Brownian motion and there exists* $\lambda > 0$ *(depending only on* $|x_0|$ *and* K*) for which*

$$\mathbb{E}\left[\sup_{t\geqslant 0} e^{-\lambda t}|X_t|^2\right] < \infty. \tag{5}$$

Moreover the solution is unique in the class of continuous adapted processes satisfying (5) for some λ*.*

Proof. As in the case of ODEs we will use a fixpoint argument to prove existence.

Take $\lambda > 0$. Introduce the space $\mathscr{X}_\lambda$ of continuous, adapted stochastic processes $(X_t)_{t\geqslant 0}$ with values in $\mathbb{R}^n$ such that $X_0 = x$ and such that the following norm

$$\|X\|_\lambda = \sup_{T\geqslant 0} \quad e^{-\lambda T}\left(\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^2\right]\right)^{1/2}$$

is finite. The normed space $(\mathscr{X}_\lambda, \|*\|_\lambda)$ is a Banach space (prove it), $\|*\|_\lambda \leqslant \|*\|_{\lambda'}$ if $\lambda' < \lambda$ so the family $(\mathscr{X}_\lambda)_\lambda$ is decreasing and we have

$$\|X\|_{\lambda/2} \leqslant \left(\mathbb{E}\left[\sup_{t\geqslant 0} e^{-\lambda t}|X_t|^2\right]\right)^{1/2} \leqslant \|X\|_{\lambda/4}, \tag{6}$$

(prove it!).

Let $\Psi: \mathscr{X}_\lambda \to \mathscr{X}_\lambda$ be the map

$$\Psi(X)_t = x_0 + \int_0^t b(s, X_s)\mathrm{d}s + \sum_\alpha \int_0^t \sigma_\alpha(s, X_s)\mathrm{d}B_s^\alpha, \qquad t \geqslant 0.$$

We need to prove that it is well defined. Observe that, by Doob's L^2 inequality and by the Itô isometry we have

$$\left\| \int_0^t \sigma_\alpha(s, X_s)\mathrm{d}B_s^\alpha \right\|_\lambda^2 = \sup_{T \geqslant 0} e^{-2\lambda T}\, \mathbb{E}\left[\sup_{t \in [0,T]} \left| \int_0^t \sigma_\alpha(s, X_s)\mathrm{d}B_s^\alpha \right|^2 \right]$$

$$\leqslant \sup_{T \geqslant 0} e^{-2\lambda T}\, \mathbb{E}\left[\left| \int_0^T \sigma_\alpha(s, X_s)\mathrm{d}B_s^\alpha \right|^2 \right] \leqslant \sup_{T \geqslant 0} e^{-2\lambda T}\, \mathbb{E}\left[\int_0^T |\sigma_\alpha(s, X_s)|^2 \mathrm{d}s \right]$$

$$\leqslant \int_0^\infty e^{-2\lambda s} \mathbb{E}[|\sigma_\alpha(s, X_s)|^2]\mathrm{d}s \leqslant \left(\int_0^\infty e^{-2\lambda s}\mathrm{d}s \right) \|s \mapsto \sigma_\alpha(s, X_s)\|_\lambda^2 = \frac{1}{2\lambda}\|s \mapsto \sigma_\alpha(s, X_s)\|_\lambda^2.$$

Now using the assumption on the growth of σ_α we have $\|s \mapsto \sigma_\alpha(s, X_s)\|_\lambda \leqslant K(1 + \|X_s\|_\lambda)$ and therefore

$$\left\| t \mapsto \int_0^t \sigma_\alpha(s, X_s)\mathrm{d}B_s^\alpha \right\|_\lambda \leqslant \frac{K}{(2\lambda)^{1/2}}(1 + \|X\|_\lambda).$$

A similar but easier computation for $\int_0^t b(s, X_s)\mathrm{d}s$ gives

$$\left\| t \mapsto \int_0^t b(s, X_s)\mathrm{d}s \right\|_\lambda \leqslant \frac{K}{\lambda}(1 + \|X\|_\lambda).$$

Putting these results together we conclude that there exists a constant C such that for all $\lambda \geqslant 1$ we have

$$\|\Psi(X)\|_\lambda \leqslant \|x_0\|_\lambda + \left\| \int_0^t b(s, X_s)\mathrm{d}s \right\|_\lambda + \sum_\alpha \left\| \int_0^t \sigma_\alpha(s, X_s)\mathrm{d}B_s^\alpha \right\|_\lambda \leqslant \|x_0\|_\lambda + C\frac{K}{\lambda^{1/2}}(1 + \|X\|_\lambda).$$

So the map Ψ is well defined on $\mathscr{X}_\lambda$. A similar computation also gives that (for the same constant C)

$$\|\Psi(X) - \Psi(Y)\|_\lambda \leqslant C\frac{K}{\lambda^{1/2}}\|X - Y\|_\lambda, \qquad \forall X, Y \in \mathscr{X}_\lambda.$$

So if we take λ large enough so that $C\frac{K}{\lambda^{1/2}} \leqslant 1/2$ we have that the map Ψ is a contraction and therefore must have a unique fixed point $X = \Psi(X)$. By the way Ψ is defined, this fixpoint is a solution to the SDE (2), moreover we have

$$\|X\|_\lambda = \|\Psi(X)\|_\lambda \leqslant \|x_0\|_\lambda + C\frac{K}{\lambda^{1/2}}(1 + \|X\|_\lambda) \leqslant \|x_0\|_\lambda + \frac{1}{2}(1 + \|X\|_\lambda)$$

which means that $\|X\|_\lambda \leqslant 1 + 2\|x_0\|_\lambda$ giving (5) after using (6). Note that $\|x_0\|_\lambda \leqslant |x_0|$.

To prove that this solution X is adapted to the Brownian motion we reason as follows: the fixpoint is unique in $\mathscr{X}_\lambda$, and to obtain it we can just take the sequence of Picard's iterations $X^{(0)} = x_0$ and $X^{(n)} = \Psi(X^{(n-1)})$. Then $X^{(n)} \to X$ in $\mathscr{X}_\lambda$ and is easy to see that each $X^{(n)}$ is a process adapted to B. By taking a subsequence we can guarantee that $(X^{(n)})_n$ converges uniformly almost surely (in any finite interval) and therefore we can conclude that X is also adapted to B.

Uniqueness in $\mathscr{X}_\lambda$ is clear by the contractivity of Ψ and we can take any large λ keeping contractivity, so if a process satisfy (5) it will belong to some $\mathscr{X}_\lambda$ and therefore must coincide with the unique fixpoint. □

In the previous result, requiring a growth condition is not natural for uniqueness. We have the following stronger statement

Theorem 3. *If the coefficients satisfy the globally Lipshitz condition (4) then any two solutions X, Y to the SDE (2) defined on the same probability space and driven by the same Brownian motion satisfy $\mathbb{P}(X_t = Y_t \text{ for all } t \geqslant 0) = 1$, i.e they are indistinguishable.*

Proof. Assume that X, Y are two solutions, let $Z_t = X_t - Y_t$ and consider the stopping time $\tau = \inf\{t \geqslant 0: |Z_t| > 1\}$ then we have

$$Z_t = \int_0^t (b(s, X_s) - b(s, Y_s))\mathrm{d}s + \sum_\alpha \int_0^t (\sigma_\alpha(s, X_s) - \sigma_\alpha(s, Y_s))\mathrm{d}B^\alpha$$

and therefore

$$\mathbb{E}[Z_{t\wedge\tau}^2] \lesssim t\int_0^{t\wedge\tau} K^2 \mathbb{E}[Z_{s\wedge\tau}^2]\mathrm{d}s + \int_0^{t\wedge\tau} K^2 \mathbb{E}[Z_{s\wedge\tau}^2]\mathrm{d}s.$$

As a consequence $f(t) = \mathbb{E}[Z_{t\wedge\tau}^2]$ satisfies

$$f(0) = 0, \qquad f(t) \leqslant C\int_0^t f(s)\mathrm{d}s$$

and by Gronwall's lemma we have $f(t) = 0$ for all $t \geqslant 0$. This implies $\mathbb{P}(|Z_{t\wedge\tau}| = 0) = 1$ so $\tau = +\infty$ and we conclude since the processes are continuous. □

The kind of uniqueness in the previous theorem is called path-wise uniqueness.

Definition 4. *An SDE has path-wise uniqueness when any two solutions defined on the same probability space and driven by the same Brownian motion are indistinguishable.*

There are other concepts of uniqueness which are useful in the theory of SDEs. There are also various possible concepts of solution.

Definition 5. *A solution of an SDE is strong when it is adapted to the filtration generated by the Brownian motion.*

A solution which is not necessarily adapted to the driving Brownian motion is called a weak solution (not to be confused with weak solution in the sense of analysis).

In particular we have proven that an SDE with globally Lipshitz coefficients have path-wise uniqueness and strong solutions.

The globally Lipshitz assumption can be relaxed. In particular if the vectorfields satisfy a local Lipshitz assumption

$$\begin{aligned} &|b(s,x)| + \sum_\alpha |\sigma_\alpha(s,x)| \leqslant K_{U,T}(1+|x|), \qquad s \in [0,T], \quad x \in U. \\ &|b(s,x) - b(s,y)| + \sum_\alpha |\sigma_\alpha(s,x) - \sigma_\alpha(s,y)| \leqslant K_{U,T}|x-y|, \qquad s \in [0,T], \quad x, y \in U. \end{aligned} \tag{7}$$

in a given open set $U \subseteq \mathbb{R}^n$ and in a given time interval $[0, T]$ with a constant $K_{U,T}$ which may depend on both, then we have existence and uniqueness of strong solutions for any initial condition $x_0 \in U$ and up to the exit time of U or to T.

Theorem 6. *Assume (7). There exists a unique process X which satisfies the SDE in the random time interval* $[0, \sigma]$ *with* $\sigma = \tau_U \wedge T$ *and* $\tau_U = \inf\{t \geqslant 0: X_t \in U^c\}$*:*

$$X_{t\wedge\sigma} = x_0 + \int_0^{t\wedge\sigma} b(X_s)\mathrm{d}s + \sum_\alpha \int_0^{t\wedge\sigma} \sigma_\alpha(X_s)\mathrm{d}B_s^\alpha, \qquad t \geqslant 0.$$

Moreover it is adapted to the Brownian motion (i.e. a strong solution).

Proof. The proof of this statement is similar to the global statement, just use the stopped fixpoint map

$$\Psi_\sigma(X) = x_0 + \int_0^{t\wedge\sigma} b(X_s)\mathrm{d}s + \sum_\alpha \int_0^{t\wedge\sigma} \sigma_\alpha(X_s)\mathrm{d}B_s^\alpha.$$

The process constructed in this way is constant after σ. Moreover uniqueness holds up to σ, namely if X and Y are (weak) solutions to the SDE up to time σ_X and σ_Y respectively, then they must coincide up to $\rho = \sigma_X \wedge \sigma_Y$ amost surely: $\mathbb{P}(X_{t\wedge\rho} = Y_{t\wedge\rho}) = 1$ and from this is easy to conclude that $\rho = \sigma_X = \sigma_Y$ almost surely, so they coincide in the whole random interval on which each of them satisfies the SDE. □

Remark 7. If X solves SDE in $\mathbb{R}^n$ with time dependent deterministic coefficients, then the process $Y_t = (t, X_t)$ in $\mathbb{R}^{n+1}$ solves an SDE with time-independent coefficients, i.e. coefficients which depends only on Y. So considering autonomous equations (i.e. equations with time-independent deterministic coefficients) is not really a restriction as long as one is willing to enlarge the state space.

1.1 Explosion

This local theory is enough for most of the applications. It does not exclude the possibility of *explosion* for the solutions.

Consider for example the SDE in one dimension

$$\mathrm{d}X_t = X_t^3 + \mathrm{d}B_t, \qquad X_0 = x_0.$$

The coefficients are locally Lipshitz so we know that it has a solution $(X_t)_{t\geqslant 0}$ which exists up to the time $\tau_n = \inf\{t \geqslant 0: |X_t| \geqslant n\}$ for any $n \geqslant 1$. The family $(\tau_n)_n$ is increasing and the question then is whether

$$\tau_\infty = \sup_n \tau_n$$

is finite or infinite. In case $\tau_\infty = +\infty$ we can solve the SDE for all times, while in the other case it must happen that as $t \nearrow \tau_\infty < \infty$ we have $|X_t| \to \infty$, i.e. the process runs away towards infinity in a finite time. In this case we call τ_∞ the explosion time.

That this could happen is clear form the particular case of the ODE

$$\mathrm{d}X_t = X_t^3, \qquad X_0 = x_0,$$

which has explicit solutions since by separation of variables:

$$\mathrm{d}t = \frac{\mathrm{d}X_t}{X_t^3} = \left(-\frac{1}{2}\right)\mathrm{d}\frac{1}{X_t^2}$$

so

$$\frac{1}{x_0^2} - \frac{1}{X_t^2} = t$$

and as $t \nearrow x_0^{-2}$ we have $X_t^{-2} \to 0$, i.e. $X_t \to +\infty$. On the other hand the equation

$$\mathrm{d}X_t = -X_t^3, \qquad X_0 = x_0,$$

has solutions such that

$$\frac{1}{x_0^2} + \frac{1}{X_t^2} = t$$

which therefore behave quite well for all times, and in particular satisfy the bound $X_t^{-2} \leqslant t$ independent of the initial condition. In this situation the *sign* of the vectorfields plays a fundamental role to decide occourrence of explosion or not. A more sophisticate approach is then needed to analyse this problem.

1.2 Examples

Consider the SDE in two dimensions

$$\begin{cases} \mathrm{d}X_t &= -Y_t \mathrm{d}B_t \\ \mathrm{d}Y_t &= X_t \mathrm{d}B_t \end{cases}$$

driven by one Brownian motion. The general theory guarantees existence of global solutions which grow at most exponentially in time. Consider $E_t = X_t^2 + Y_t^2$. A computation with Itô formula for this solutions gives

$$\mathrm{d}E_t = 2X_t \mathrm{d}X_t + 2Y_t \mathrm{d}Y_t + (Y_t^2 + X_t^2)\mathrm{d}t = E_t \mathrm{d}t$$

so we conclude that $(E_t)_{t \geqslant 0}$ satisfies an ODE which can be solved explicitly to get $E_t = e^t E_0$. Note that if we imagine to replace the Brownian motion by a smooth approximation $(B_t^\varepsilon)_{t \geqslant 0}$ and interpret the resulting equation as an ODE of the form

$$\begin{cases} \mathrm{d}X_t &= -Y_t \mathrm{d}B_t^\varepsilon \\ \mathrm{d}Y_t &= X_t \mathrm{d}B_t^\varepsilon \end{cases}$$

where the integrals are taken as Riemman–Lebesgue integrals we can identify them with usual Lebsegue integrals since B_t^ε is smooth and therefore $\mathrm{d}B_t^\varepsilon = \dot{B}_t^\varepsilon \mathrm{d}t$ where $\dot{B}_t^\varepsilon$ denotes the derivative. But now $\dot{E}_t = 2X_t \dot{X}_t + 2Y_t \dot{Y}_t = 0$ so $E_t = E_0$!! Somewhat SDEs are not compatible with approximating the Brownian motion and one has to be wary of these formal manipulations.

Stratonovich integration (not for us now) will solve this problem.

Still the intuiting gathered by looking at a similar ODE can be used to guess if an SDE has an explicit solution and how it looks like.

2 Tanaka's example

We have introduced strong solutions introducing a distinction between them and weak solutions (i.e. solutions which are not necessarily adapted to the driving Brownian motion).

Here's an example of a weak solution which is not strong. Let X be a one-dimensional Brownian motion and consider

$$B_t = \int_0^t \operatorname{sign}(X_s)\mathrm{d}X_s$$

where $\operatorname{sign}(x) = 1$ if $x \geqslant 0$ and $\operatorname{sign}(x) = 1$ otherwise. By Lèvy's theorem the process $(B_t)_t$ is again a Brownian motion. Moreover we have

$$\int_0^t \operatorname{sign}(X_s)\mathrm{d}B_s = \int_0^t \operatorname{sign}(X_s)^2\mathrm{d}X_s = \int_0^t \mathrm{d}X_s = X_t - X_0$$

so X satisfies an SDE

$$\mathrm{d}X_t = \operatorname{sign}(X_t)\mathrm{d}B_t, \tag{8}$$

with B as driving noise and $\sigma(x) = \operatorname{sign}(x)$. This coefficient is of not Lipshitz (not even continuous…). Moreover if we consider the initial condition $X_0 = 0$ then we have that both $(X_t)_{t\geqslant 0}$ and $(-X_t)_{t\geqslant 0}$ are solutions, i.e. path-wise uniqueness do not hold for (8).

However solutions of the SDE (8) cannot be arbitrary: again be Lévy's theorem any solution has to be a Brownian motion, so its law is unique. We call this kind of uniqueness, uniqueness in law.

Definition 8. *Uniqueness in law holds for an SDE if any solution (on any probability space) has the same law.*

Recall that if φ is a smooth function, by Itô formula

$$\varphi(X_t) = \varphi(X_0) + \int_0^t \varphi'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t \varphi''(X_s)\mathrm{d}s.$$

Take now $\varphi = \varphi_\varepsilon$ even with $\varphi_\varepsilon(x) = (\varepsilon + x^2)^{1/2}$. Then $\varphi_\varepsilon'(x) = x(\varepsilon + x^2)^{-1/2}$ and

$$\varphi_\varepsilon''(x) = (\varepsilon + x^2)^{-1/2} - x^2(\varepsilon + x^2)^{-3/2} = \frac{\varepsilon}{(\varepsilon + x^2)^{3/2}}.$$

If we let $Z_t^\varepsilon = \int_0^t \varphi_\varepsilon'(X_s)\mathrm{d}X_s$ we have

$$\mathbb{E}\sup_{t\leqslant T}|Z_t^\varepsilon - B_t|^2 = \mathbb{E}|Z_T^\varepsilon - B_T|^2 = \mathbb{E}\int_0^T |\varphi_\varepsilon'(X_s) - \operatorname{sign}(X_s)|\mathrm{d}s = \int_0^T \mathbb{E}|\varphi_\varepsilon'(X_s) - \operatorname{sign}(X_s)|\mathrm{d}s \to 0$$

as $\varepsilon \to 0$ by dominated convergence since $\varphi_\varepsilon'(x) = x(\varepsilon + x^2)^{-1/2} \to \operatorname{sign}(x)$ if $x \neq 0$ and is uniformly bounded so the pointwise (in s) convergence $\mathbb{E}|\varphi_\varepsilon'(X_s) - \operatorname{sign}(X_s)| \to \mathbb{E}[\mathbb{1}_{X_s=0}] = \mathbb{1}_{s=0}$ allows to conclude.

As a consequence and by subsequences $Z_t^\varepsilon \to B_t$ uniformly almost surely in any bounded interval. Since φ_ε is even we also have

$$Z_t^\varepsilon = \varphi_\varepsilon(X_t) - \frac{1}{2}\int_0^t \varphi_\varepsilon''(X_s)\mathrm{d}s = \varphi_\varepsilon(|X_t|) - \frac{1}{2}\int_0^t \varphi_\varepsilon''(|X_s|)\mathrm{d}s.$$

Therefore $(Z_t^\varepsilon)_{t\geqslant 0}$ is actually a function of $|X_t|$. We conclude that $(B_t)_{t\geqslant 0}$ is measurable wrt. $(|X_t|)_{t\geqslant 0}$. In particular this proves that $(X_t)_{t\geqslant 0}$ cannot be a strong solution to the SDE (8) since otherwise we will have the following inclusion of completed filtrations

$$(\mathscr{F}_t^X)_{t\geqslant 0} \subseteq (\mathscr{F}_t^B)_{t\geqslant 0} \subseteq (\mathscr{F}_t^{|X|})_{t\geqslant 0}$$

which is absurd since knowing the modulus of a Brownian motion does not allow to recover its sign! (think about it). We must conclude that X is strictly a weak solution. And that this holds for all (!) weak solutions. So no strong solutions exists.

What we have discussed is Tanaka's example of a weak solution of an SDE with bounded coefficients which is however not strong. This shows that some regularity of the coefficients is needed to ensure existence of strong solutions.

3 Tanaka's formula

Let now $(X_t)_{t\geqslant 0}$ be a continuous semimartingale. Using the same function φ_ε we have

$$\frac{1}{2}\int_0^t \varphi_\varepsilon''(X_s)\mathrm{d}[X]_s = \varphi_\varepsilon(X_t) - \varphi_\varepsilon(X_0) + \int_0^t \varphi_\varepsilon'(X_s)\mathrm{d}X_s$$

and again we can show that

$$\int_0^t \varphi_\varepsilon'(X_s)\mathrm{d}X_s \to \int_0^t \mathrm{sign}_0(X_s)\mathrm{d}X_s$$

a.s. uniformly in bounded intervals where $\mathrm{sign}_0(x) = \mathbb{1}_{x>0} - \mathbb{1}_{x<0}$, since (maybe stopping X if necessary)

$$\mathbb{E}\left|\int_0^t [\varphi_\varepsilon'(X_s) - \mathrm{sign}_0(X_s)]\mathrm{d}X_s\right|^2 = \mathbb{E}\int_0^t [\varphi_\varepsilon'(X_s) - \mathrm{sign}_0(X_s)]^2\mathrm{d}[X]_s$$

and $\varphi_\varepsilon'(X_s) - \mathrm{sign}_0(X_s) \to 0$ pointwise, while $|\varphi_\varepsilon'(X_s) - \mathrm{sign}_0(X_s)| \leqslant 2$, so dominated convergence allows to conclude. Of course we have also $\varphi_\varepsilon(X_t) \to |X_t|$ uniformly in t in bounded intervals a.s. and as a consequence we deduce that the limit $\lim_{\varepsilon\to 0}\frac{1}{2}\int_0^t \varphi_\varepsilon''(X_s)\mathrm{d}[X]_s$ exists a.s. (maybe up to subsequences) and defined a continuous function. More generally, for any $x\in\mathbb{R}$ we can define a continuous process $(\ell_t^X(x))_{t\geqslant 0}$ by

$$\ell_t^X(x) = \lim_{\varepsilon\to 0}\frac{1}{2}\int_0^t \varphi_\varepsilon''(x-X_s)\mathrm{d}[X]_s = \lim_{\varepsilon\to 0}\left[\varphi_\varepsilon(x-X_t) - \varphi_\varepsilon(x-X_0) - \int_0^t \varphi_\varepsilon'(x-X_s)\mathrm{d}X_s\right]$$

$$= |x-X_t| - |x-X_0| - \int_0^t \mathrm{sign}_0(x-X_s)\mathrm{d}X_s.$$

This process is increasing (as pointwise limit of increasing functions) and therefore of bounded variation and adapted to the filtration generated by X.

Note that $\frac{1}{2}\varphi_\varepsilon''(x) \to \delta(x)$ in the sense of distributions, therefore, for any continuous funciton f: $\mathbb{R}\to\mathbb{R}$ we have

$$f_\varepsilon(x) = \left(f * \frac{1}{2}\varphi_\varepsilon''\right)(x) = \int_{\mathbb{R}} f(y)\frac{1}{2}\varphi_\varepsilon''(y-x)\mathrm{d}y \to f(x), \qquad \varepsilon\to 0,$$

and f_ε is smooth for any $\varepsilon > 0$. By Fubini

$$\int_0^t f_\varepsilon(X_s)\mathrm{d}[X]_s = \int_{\mathbb{R}} f(y)\left(\int_0^t \frac{1}{2}\varphi_\varepsilon''(y - X_s)\mathrm{d}[X]_s\right)\mathrm{d}y$$

using the almost sure convergence of $\int_0^t \frac{1}{2}\varphi_\varepsilon''(y - X_s)\mathrm{d}[X]_s$ and a function f which compactly supported we deduce that

$$\int_0^t f(X_s)\mathrm{d}[X]_s = \lim_{\varepsilon\to 0}\int_0^t f_\varepsilon(X_s)\mathrm{d}[X]_s = \int_{\mathbb{R}} f(y)\ell_t^X(y)\mathrm{d}y \tag{9}$$

which can be extended to all bounded continuous functions by continuity since

$$\left|\int_{\mathbb{R}} f(y)\ell_t^X(y)\mathrm{d}y\right| = \|f\|_{L^\infty}\int_0^t \mathrm{d}[X]_s = \|f\|_{L^\infty}[X]_t$$

and $[X]_t < \infty$ almost surely. The formula (9) justify the terminology *local time* for $\ell_t^X(x)$. We can introduce the occupation measure μ_t^X of X as the random measure such that

$$\mu_t^X(A) = \int_0^t \mathbb{1}_{X_s\in A}\mathrm{d}[X]_s \qquad \text{for all } A\in\mathscr{B}(\mathbb{R}).$$

Then eq. (9) shows that every semimartingale satisfies (by dominated convergence)

$$\mu_t^X(A) = \int_0^t \mathbb{1}_{X_s\in A}\mathrm{d}[X]_s = \int_A \ell_t^X(y)\mathrm{d}y.$$

so the occupation measure is absolutely continuous wrt. to the Lebsegue measure of $\mathbb{R}$ and the density is precisely ℓ_t^X.

Another interesting property of the local time is that, for all x we have

$$\int_0^t \mathbb{1}_{X_s\neq x}\mathrm{d}\ell_s(x) = 0$$

that is, the measure $\mathrm{d}\ell_s(x)$ is supported on the closed set $\{s: X_s = x\}$. To prove this let

$$Z_t^\varepsilon = \frac{1}{2}\int_0^t \varphi_\varepsilon''(X_s)\mathrm{d}s$$

we have (why?) for all $\delta > 0$

$$\int_0^t \frac{N|X_s - x|}{1 + N|X_s - x|}\mathrm{d}\ell_s^X(0) = \lim_{\varepsilon\to 0}\int_0^t \frac{N|X_s - x|}{1 + N|X_s - x|}\mathrm{d}Z_s^\varepsilon$$

$$= \lim_{\varepsilon\to 0}\int_0^t \frac{N|X_s - x|}{1 + N|X_s - x|}\frac{1}{2}\varphi_\varepsilon''(X_s)\mathrm{d}s = 0$$

and therefore by monotone convergence

$$\int_0^t \mathbb{1}_{X_s\neq x}\mathrm{d}\ell_s(x) = \lim_{N\to\infty}\int_0^t \frac{N|X_s - x|}{1 + N|X_s - x|}\mathrm{d}\ell_s^X(0) = 0$$

which means that $(\ell_t^X(0))_{t\geqslant 0}$ increases only when $X_t = x$.

3.1 A digression on SDEs with reflection

In the case of Brownian motion Tanaka's formula reads

$$|B_t| = |B_0| + \int_0^t \mathrm{sign}_0(B_s)\mathrm{d}B_s + \ell_t^B(0),$$

in particular if we let $|B_t| = X_t$ and $W_t = \int_0^t \mathrm{sign}_0(B_s)\mathrm{d}B_s$ we have that $(W_t)_{t\geqslant 0}$ is a Brownian motion, that X_t is a semimartingale and that

$$X_t = X_0 + W_t + \ell_t^X(0) \tag{10}$$

indeed note that

$$\ell_t^B(0) = \lim_{\varepsilon\to 0}\frac{1}{2}\int_0^t \varphi_\varepsilon''(B_s)\mathrm{d}[B]_s = \lim_{\varepsilon\to 0}\frac{1}{2}\int_0^t \varphi_\varepsilon''(B_s)\mathrm{d}s = = \lim_{\varepsilon\to 0}\frac{1}{2}\int_0^t \varphi_\varepsilon''(|B_s|)\mathrm{d}s = \ell_t^X(0)$$

since $[X]_t = [B]_t = t$.

Eq. (10) is a kind of equation for a process X driven by a Brownian motion W and involving the local time of the process itself. By construction we know that $X_t \geqslant 0$. Somehow the local time acts as to keep the value of X_t positive, against the action of the Brownian motion W which with positive probability would push the process into the negative semiaxis. Note that

$$\int_0^t \mathbb{1}_{X_s>0}\mathrm{d}\ell_s^X(0) = 0.$$

Observe the following nice property. There exists only one pair (X,ρ) such that X is a continuous adapted process such that $X_t \geqslant 0$ and ρ a continuous increasing process and

$$\begin{cases} X_t = x_0 + B_t + \rho_t, & t\geqslant 0, \\ \int_0^\infty \mathbb{1}_{X_s>0}\mathrm{d}\rho_s = 0, & \end{cases} \tag{11}$$

where $(B_t)_{t\geqslant 0}$ is a Brownian motion and $x_0 \geqslant 0$.

Indeed assume there is another pair (X',ρ') sharing the same properties with respect to the same Brownian motion and same initial condition, then letting $Z_t = X_t - Y_t$ we have

$$X_t - Y_t = \rho_t - \rho_t'.$$

Compute

$$\mathrm{d}|X_t - Y_t|^2 = (X_t - Y_t)\mathrm{d}(X_t - Y_t) = (X_t - Y_t)(\mathrm{d}\rho_t - \mathrm{d}\rho_t') = X_t\mathrm{d}\rho_t - Y_t\mathrm{d}\rho_t - X_t\mathrm{d}\rho_t' + Y_t\mathrm{d}\rho_t'$$

$$= -Y_t\mathrm{d}\rho_t - X_t\mathrm{d}\rho_t'$$

since (by localization via stopping times and a passage to the limit)

$$\int_0^\infty X_t\mathrm{d}\rho_t + Y_t\mathrm{d}\rho_t' = 0.$$

We conclude that $\mathrm{d}|X_t - Y_t|^2 \leqslant 0$ meaning that $|X_t - Y_t|^2 \leqslant |X_0 - Y_0|^2 = 0$, that is $X_t = Y_t$ a.s. and indeed there exists a unique such process up to indistinguishability. This implies also that $\rho_t = \rho_t'$ a.s.

This is a striking uniqueness result since somehow the equations (11) prescribe, not only the behaviour of the process X but also that of the measure $\mathrm{d}\rho$ at the same time.

The unique process satisfying (10) is called *reflected Brownian motion*: it behaves as a Brownian motion most of the time but as soon as it reaches zero it is reflected always on the upper half plane. Note that we have proven that the law of the reflected Brownian motion is that of the modulus of a standard Brownian motion since $X_t = |B_t|$ by construction and our uniqueness result ensures that there are no other possible solutions.

4 Weak Solution to SDEs via Girsanov's transformation

Let $(X_t)_{t\geqslant 0}$ be a n-dimensional Brownian motion starting at $X_0 = x_0$ and b: $\mathbb{R}^n \to \mathbb{R}^n$ a globally Lipshitz vectorfield. Then the process

$$Z_t = \exp\left(\int_0^t b(X_s)\mathrm{d}X_s - \frac{1}{2}\int_0^t |b(X_s)|^2 \mathrm{d}s\right), \quad t \geqslant 0, \tag{12}$$

is a positive local martingale (and therefore a supermartingale). By Novikov's criterion we know that it is actually a true martingale since for all $0 \leqslant s \leqslant t$

$$\mathbb{E}[Z_t|\mathscr{F}_s] = \mathbb{E}\left[\exp\left(\int_s^t b(X_r)\mathrm{d}X_r - \frac{1}{2}\int_s^t |b(X_r)|^2 \mathrm{d}r\right)|\mathscr{F}_s\right] Z_s = \varphi_{t-s}(X_s) Z_s$$

with

$$\varphi_t(x) = \mathbb{E}\left[\exp\left(\int_0^t b(x+X_r)\mathrm{d}X_r - \frac{1}{2}\int_0^t |b(x+X_r)|^2 \mathrm{d}r\right)\right].$$

By the Lipschiz condition on b and Jensen's inequality

$$\mathbb{E}\exp\left(\frac{1}{2}\int_0^t |b(x+X_r)|^2 \mathrm{d}r\right) \lesssim \mathbb{E}\exp\left(\frac{C}{2}\int_0^t |X_r|^2 \mathrm{d}r\right) \lesssim \int_0^t \mathbb{E}\exp\left(\frac{C}{2}|X_r|^2 t\right)\frac{\mathrm{d}r}{t}$$

$$\lesssim \int_0^t \mathbb{E}\exp\left(\frac{C}{2}|X_1|^2 r T\right)\frac{\mathrm{d}r}{t} \lesssim \mathbb{E}\exp\left(\frac{C}{2}|X_1|^2 t^2\right)$$

which is finite for $0 < t \leqslant \delta$ with $C\delta^2 < 1$. By Novikov's condition we have $\varphi_t(x) = 1$ for all $t \leqslant \delta$ and $x \in \mathbb{R}^n$. But then a simple telescoping argument gives that $(Z_t)_{t\geqslant 0}$ is a true martingale in any finite interval $[0, T]$. (Warning: in general is not true that is a closed martingale!!!)

We can then consider the measure $\mathbb{Q}$ defined as

$$\mathbb{Q}(A) = \mathbb{E}[Z_T \mathbb{1}_A], \qquad A \in \mathscr{F}_T.$$

This is a good definition since if $A \in \mathscr{F}_S \subseteq \mathscr{F}_T$ for $S \leqslant T$ we have $\mathbb{E}[Z_T \mathbb{1}_A] = \mathbb{E}[Z_S \mathbb{1}_A]$. $\mathbb{Q}$ is a new measure on Ω whose restrictions to $\mathscr{F}_T$ is absolutely continuous wrt. the restriction of $\mathbb{P}$:

$$\frac{\mathrm{d}\mathbb{Q}|_{\mathscr{F}_T}}{\mathrm{d}\mathbb{P}|_{\mathscr{F}_T}} = Z_T.$$

By Girsanov's theorem

$$B_t = X_t - \int_0^t b(X_s)\mathrm{d}s, \qquad t \geqslant 0$$

is a local martingale and since $[B]_t = [X]_t = t$ it is also a Brownian motion. As a consequence we have shown that the pair (X, B) under the measure $\mathbb{Q}$ is a weak solution of the SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t, \qquad t \geqslant 0$$

with initial condition $X_0 = x_0$.

Example 9. (Brownian motion with constant drift) Let X be a d–dimensional $\mathbb{P}$–Brownian motion and $\gamma \in \mathbb{R}^d$ a fixed vector. Consider the martingale

$$Z_t^\gamma := \exp\left(\gamma . X_t - \frac{1}{2} |\gamma|^2 t \right) = \mathscr{E}(\gamma . X)_t, \qquad t \geqslant 0$$

and for any $t \geqslant 0$ the measure $\mathbb{P}^\gamma$ defined on $(\Omega, \mathscr{F}_t, (\mathscr{F}_s)_{\leqslant s \leqslant t})$ by $\mathrm{d}\mathbb{P}^\gamma|_{\mathscr{F}_t} = Z_t^\gamma \mathrm{d}\mathbb{P}|_{\mathscr{F}_t}$. By Girsanov's theorem we have that, the process

$$\tilde{X}_s = X_s - \gamma\, s, \qquad s \geqslant 0$$

is a d–dimensional $\mathbb{P}^\gamma$–Brownian motion, so under $\mathbb{P}_t^\gamma$ the process X is a Brownian motion with drift γ. The $\mathscr{F}_\infty$ measurable event

$$\lim_{s \to +\infty} \frac{(X_s - \gamma s)}{s} = -\gamma,$$

has $\mathbb{P}$ probability 1 (e.g. for the law of iterated logarithm applied to the $\mathbb{P}$–BM X) while it has $\mathbb{P}^\gamma$ probability 0 since $s \mapsto X_s - \gamma s$ is a $\mathbb{P}^\gamma$–Brownian motion.

Another consequence of Girsanov's theorem is a criterion for uniqueness in law.

Theorem 10. *All the weak solutions (X, B) of the SDE*

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t, \qquad t \geqslant 0 \tag{13}$$

satisfying for all $T \geqslant 0$

$$\frac{1}{2} \int_0^T |b(X_s)|^2 \mathrm{d}s < \infty, \quad a.s. \tag{14}$$

have the same law.

Proof. Let (X, B) any weak solution of (13) satisfying (14). Define the increasing sequence of stopping times $(\tau_n)_{n \geqslant 1}$ as

$$\tau_n = \inf \left\{ t \geqslant 0 \colon \frac{1}{2} \int_0^t |b(X_s)|^2 \mathrm{d}s \leqslant n \right\}$$

and note that (14) implies $\tau_n \to \infty$ almost surely. Now consider $(Z_t)_{t \geqslant 0}$ as in eq. (12) above and observe that the process $(Q_t)_{t \geqslant 0}$ defined as $Q_t = Z_t^{-1}$ satisfies

$$Q_{t \wedge \tau_n} := Z_{t \wedge \tau_n}^{-1} = \exp\left(- \int_0^{t \wedge \tau_n} b(X_s)\mathrm{d}X_s + \frac{1}{2} \int_0^{t \wedge \tau_n} |b(X_s)|^2 \mathrm{d}s \right)$$

$$= \exp\left(- \int_0^{t \wedge \tau_n} b(X_s)\mathrm{d}B_s - \frac{1}{2} \int_0^{t \wedge \tau_n} |b(X_s)|^2 \mathrm{d}s \right).$$

Due to the presence of the stopping time, the Novikov's criterion is trivially satisfied, and we can define the measure $\mathbb{Q}^{(n)}$ such that $\mathrm{d}\mathbb{Q}^{(n)}|_{\mathscr{F}_t} = Q_{t\wedge\tau_n}\mathrm{d}\mathbb{P}|_{\mathscr{F}_t}$ for all $t \geqslant 0$ and under which

$$\tilde{B}_t^{(n)} = B_t + \int_0^{t\wedge\tau_n} b(X_s)\mathrm{d}s$$

is a Brownian motion. However by the SDE (13) we have $\tilde{B}_{t\wedge\tau_n}^{(n)} = X_{t\wedge\tau_n}$ so indeed $(X_t)_{t\geqslant 0}$ is a $\mathbb{Q}$-Brownian motion in the random interval $[0, \tau_n]$. As a consequence, for any $T \geqslant 0$ and any $\mathbb{1}_A(X, B) \in \mathscr{F}_T$ we have

$$\mathbb{E}[\mathbb{1}_A(X,B)\mathbb{1}_{T\leqslant\tau_n}] = \mathbb{E}_{\mathbb{Q}}[\mathbb{1}_{T\leqslant\tau_n}\mathbb{1}_A(X,B)Q_T^{-1}]$$

$$= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_{T\leqslant\tau_n}\mathbb{1}_A(X,B)\exp\left(\int_0^T b(X_s)\mathrm{d}X_s - \frac{1}{2}\int_0^T |b(X_s)|^2\mathrm{d}s\right)\right]$$

$$= \mathbb{E}\left[\mathbb{1}_{T\leqslant\sigma_n(B)}\mathbb{1}_A\left(B, B - \int_0^{\cdot} b(B)\mathrm{d}s\right)\exp\left(\int_0^T b(B_s)\mathrm{d}B_s - \frac{1}{2}\int_0^T |b(B_s)|^2\mathrm{d}s\right)\right] = \mathbb{E}[h_A(B)]$$

where $\sigma_n(B) = \inf\left\{t \geqslant 0\colon \frac{1}{2}\int_0^t |b(B_s)|^2\mathrm{d}s \leqslant n\right\}$ and where $h_A\colon C(\mathbb{R}_+;\, \mathbb{R}^n) \to \mathbb{R}_+$ is a suitable measurable function depending on A (and T). For any two weak solutions $(X^{(i)}, B^{(i)})$ $i = 1, 2$ we have

$$\mathbb{E}[\mathbb{1}_A(X^{(1)}, B^{(1)})\mathbb{1}_{T\leqslant\tau_n^{(1)}}] = \mathbb{E}[h_A(B^{(1)})] = \mathbb{E}[h_A(B^{(2)})] = \mathbb{E}[\mathbb{1}_A(X^{(2)}, B^{(2)})\mathbb{1}_{T\leqslant\tau_n^{(2)}}]$$

where $\tau_n^{(i)} = \sigma_n(X^{(i)})$ for $i = 1, 2$. Letting $n \to \infty$ and using (14) to prove that $\tau_n^{(i)} \to \infty$ a.s. for $i = 1$, 2, we deduce by dominated convergence that

$$\mathbb{E}[\mathbb{1}_A(X^{(1)}, B^{(1)})] = \mathbb{E}[\mathbb{1}_A(X^{(2)}, B^{(2)})].$$

Since T is arbitrary this equality holds for all $A \in \mathscr{B}(C(\mathbb{R}_+; \mathbb{R}^n))$ and therefore we conclude that $\mathrm{Law}(X^{(1)}, B^{(1)}) = \mathrm{Law}(X^{(2)}, B^{(2)})$ (as measures on $C(\mathbb{R}_+; \mathbb{R}^n)$). □

5 Conditioning: Brownian Bridge

We want to study solutions of SDEs conditioned to satisfy certain properties. The simplest case is Brownian motion started at 0 and conditioned to reach a given point x at a given time T. We will assume that we are in the Wiener space $\Omega = C(\mathbb{R}_+; \mathbb{R})$ and that the Brownian motion B is the canonical process (i.e. $B_t(\omega) = \omega$). We will denote as usual by $(\Omega, \mathscr{F}, \mathbb{P})$ the probability triple and consider the canonical filtration $(\mathscr{F}_t)_{t\geqslant 0}$.

The difficulty is given by the fact that we try to condition on a probability 0 event, and therefore we cannot use the elementary formula to define the conditioned measure. To avoid this problem we use the idea of *disintegration* of the probability measure $\mathbb{P}$ according to the random variable B_T, namely we want to find a probability kernel $\mathbb{P}^x = \mathbb{P}^{x,T}$ such that $\mathbb{P}^x(B_T = x) = 1$ for all $x \in \mathbb{R}$ and

$$\mathbb{P}(A) = \int_{\mathbb{R}} \mathbb{P}^x(A)\mathbb{P}(B_T \in \mathrm{d}x), \qquad A \in \mathscr{F}, \tag{15}$$

where $\mathbb{P}(B_T \in \mathrm{d}x)$ is a convenient notation for the law of B_T under $\mathbb{P}$. Recall that, in general, a probability kernel $(\mathbb{P}^x)_{x\in\mathbb{R}}$ is a way to associate to any $x \in \mathbb{R}$ a probability measure on Ω in such a way that $x \mapsto \mathbb{P}^x(A)$ is measurable for any $A \in \mathscr{F}$. If such a kernel exists we will say that $\mathbb{P}^x$ is the conditional probability we were looking for.

An heuristic way to justify this definition is to note that we have

$$\mathbb{P}(A|B_T\in[x+\delta,x-\delta])=\frac{\mathbb{P}(A\cap B_T\in[x+\delta,x-\delta])}{\mathbb{P}(B_T\in[x+\delta,x-\delta])}=\frac{\int_{[x+\delta,x-\delta]}\mathbb{P}^x(A)\mathbb{P}(B_T\in\mathrm{d}x)}{\int_{[x+\delta,x-\delta]}\mathbb{P}(B_T\in\mathrm{d}x)}$$

and if we formally take the limit $\delta\to 0$ we obtain

$$\mathbb{P}(A|B_T\in[x+\delta,x-\delta])\to\mathbb{P}^x(A).$$

Rigorously we can observe that if such kernel exists then we must have

$$\mathbb{P}^{B_T}(A)=\mathbb{E}[\mathbb{1}_A|B_T]$$

almost surely. So we could also write symbolically $\mathbb{P}^x(A)=\mathbb{E}[\mathbb{1}_A|B_T=x]$. In order to explicitly find $\mathbb{P}^x$ we reason as follows. The kernel has to satisfy

$$\mathbb{E}[\mathbb{1}_{B_T\in B}\mathbb{P}^{B_T}(A)]=\mathbb{E}[\mathbb{1}_A\mathbb{1}_{B_T\in B}]$$

for all $B\in\mathscr{B}(\mathbb{R})$ and $A\in\mathscr{F}$. Now if $A\in\mathscr{F}_s$ with $s<t$ we have

$$\mathbb{E}[\mathbb{1}_A\mathbb{1}_{B_T\in B}]=\mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_{B_T\in B}|\mathscr{F}_s]]=\mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_{B_T\in B}|B_s]]=\mathbb{E}\left[\mathbb{1}_A\int_B p(s,B_s;T,y)\mathrm{d}y\right]$$

where

$$p(s,x;t,y)=\frac{1}{\sqrt{2\pi(t-s)}}e^{-\frac{1}{2}\frac{(x-y)^2}{(t-s)}},\qquad x,y\in\mathbb{R},0\leqslant s<t$$

the probability density that a Brownian motion started from x at time s reaches to y at time t. By Fubini we can rewrite this as

$$\mathbb{E}[\mathbb{1}_A\mathbb{1}_{B_T\in B}]=\int_B\mathbb{E}[\mathbb{1}_A p(s,B_s;T,y)]\mathrm{d}y=\int_B\mathbb{E}\left[\mathbb{1}_A\frac{p(s,B_s;T,y)}{p(0,0;T,y)}\right]p(0,0;T,y)\mathrm{d}y$$

where $p(0,0;T,y)$ is the density of B_T so if we let

$$\mathbb{P}^x(A)=\mathbb{E}\left[\mathbb{1}_A\frac{p(s,B_s;T,x)}{p(0,0;T,x)}\right]\tag{16}$$

we have indeed obtained that

$$\mathbb{E}[\mathbb{1}_A\mathbb{1}_{B_T\in B}]=\mathbb{E}[\mathbb{P}^{B_T}(A)\mathbb{1}_{B_T\in B}]\tag{17}$$

and so $\mathbb{P}^{B_T}(A)=\mathbb{E}[\mathbb{1}_A|B_T]$ for all $A\in\mathscr{F}_s$ with $s<T$ and therefore also for all $A\in\mathscr{F}_{T-}=\sigma(\mathscr{F}_s\colon s\in[0,T))$.

From the continuity of the Brownian motion we deduce that $B_T=\lim_{h\downarrow 0}B_{T-h}$ and therefore that $\sigma(B_T)\subseteq\mathscr{F}_{T-}$ with the consequence that $\mathscr{F}_{T-}=\mathscr{F}_T$. So equation (17) holds indeed for all $A\in\mathscr{F}_T$.

By inspection of (16) one observe that $\mathbb{P}^x$ is a probability kernel (a measurable function in x and a probability in $A \in \mathscr{F}_T$). Moreover one note also that the process

$$Z_s^x := \frac{p(s, B_s; T, x)}{p(0, 0; T, x)}, \qquad s < T,$$

is a martingale in $[0, T)$ with $Z_0^x = 1$. Using the continuity of the paths of the Brownian motion we can prove that $Z_s^x \to 0$ as $s \to T$ a.s. Indeed $B_T \neq x$ almost surely, that is with probability 1 there exists $\delta > 0$ such that $|B_T - x| > \delta$ and by continuity $|B_s - x| > \delta / 2$ for $|s - T|$ small, but then $p(s, B_s; T, x) \to 0$ as $s \to T$.

By continuity of the paths we have

$$\mathbb{P}^x(|B_T - x| > \delta) = \lim_{\sigma \to 0} \mathbb{P}^x(|B_{T-\sigma} - x| > \delta) = \lim_{\sigma \to 0} \frac{\mathbb{E}[\mathbb{1}_{|B_{T-\sigma} - x| > \delta} p(T - \sigma, B_{T-\sigma}; T, x)]}{p(0, 0; T, x)} = 0$$

and as a consequence that $\mathbb{P}^x(B_T = x) = 1$.

In order to extend our definition of $\mathbb{P}^x$ to all $\mathscr{F}$ we reason as follows: any $A \in \mathscr{F}$ can be approximated arbitrarily well as $A_1 \cap A_2$ with $A_1 \in \mathscr{F}_{T-\delta}^B$ and $A_2 \in \sigma(B_t : t \geqslant T)$ for some $\delta > 0$. But now using the Markov property and letting $\varphi(B_T) = \mathbb{E}[\mathbb{1}_{A_2} | B_T]$ we have

$$\mathbb{E}[\mathbb{1}_{A_1} \mathbb{1}_{A_2} \mathbb{1}_{B_T \in B}] = \mathbb{E}[\mathbb{1}_{A_1} \varphi(B_T) \mathbb{1}_{B_T \in B}] = \mathbb{E}[\mathbb{P}^{B_T}(A_1) \varphi(B_T) \mathbb{1}_{B_T \in B}] = \mathbb{E}[\mathbb{P}^{B_T}(A_1 \cap A_2) \mathbb{1}_{B_T \in B}]$$

indeed

$$\mathbb{P}^x(A_1) \varphi(x) = \mathbb{E}^x[\mathbb{1}_{A_1} \varphi(B_T)] = \mathbb{E}\left[\mathbb{1}_{A_1} \frac{p(T - \delta, B_{T-\delta}; T, x)}{p(0, 0; T, y)} \varphi(B_T)\right]$$

$$= \mathbb{E}\left[\mathbb{1}_{A_1} \frac{p(T - \delta, B_{T-\delta}; T, x)}{p(0, 0; T, y)} \mathbb{1}_{A_2}\right] = \mathbb{E}^x[\mathbb{1}_{A_1} \mathbb{1}_{A_2}] = \mathbb{P}^x(A_1 \cap A_2)$$

Therefore eq. (17) is valid for all A of the form $A_1 \cap A_2$ and since δ is arbitrary we deduce that it is valid for all $A \in \mathscr{F}$ by a monotone class argument.

We want now to caracterise the canonical process under the measure $\mathbb{P}^x$. In order to do this observe that the positive martingale $(Z_s^x)_{s \in [0, T)}$ is the stochastic exponential of the process

$$L_t^x = \int_0^t \frac{p'(s, B_s; T, x)}{p(s, B_s; T, x)} \mathrm{d}B_s, \qquad t \in [0, T)$$

where $p'(s, x; t, x) = \partial_x p(s, x; t, x)$ (this can be checked via Ito's formula). As a consequence of this and of Girsanov's theorem we have that under $\mathbb{P}^x$ the canonical process $(B_t)_{t \geqslant 0}$ satisfies

$$\mathrm{d}B_t = \frac{p'(t, B_t; T, x)}{p(t, B_t; T, x)} \mathrm{d}t + \mathrm{d}W_t, \qquad t \in [0, T) \tag{18}$$

where W_t is a one dimensional Brownian motion. Moreover $B_T = x$ and $(B_t - B_T)_{t \geqslant T}$ is a Brownian motion. In this way we have completely caracterised the measure $\mathbb{P}^x$ from a pathwise perspective.

Writing down explicitly the expression for (18) we obtain

$$\mathrm{d}B_t = -\frac{B_t - x}{T - t} \mathrm{d}t + \mathrm{d}W_t, \qquad t \in [0, T) \tag{19}$$

The process $(B_t)_{t\in[0,T]}$ is called Brownian bridge (from $(0,0)$ to (t,y)) and since the coefficients are regular it is the unique pathwise solution of the SDE (19) in $[0,T)$. Existence of weak solutions of course follows by our explicit construction of $\mathbb{P}^x$. However, given the regularity of the coefficients it is also true that a strong solution exists. On the probability space $(\Omega,\mathscr{F},\mathbb{P}^x)$ we can therefore construct also a strong solution X measurable w.r.t. W which by path-wise uniqueness is indistinguishable from $(B_t)_{t\geqslant 0}$. We deduce that the law of this strong solution is given by the law of B under $\mathbb{P}^x$ and that the SDE (19) has uniqueness in law. By continuity of the sample paths of the measure $\mathbb{P}^x$ we deduce that any weak solution must also satisfy $B_t\to B_T$ as $t\to T$.

6 Conditioning: Bessel process

Another natural kind of conditioning is when we require the process to remain in a given domain for some time T or even forever.

Let us consider the simpler situation of a one dimensional Brownian motion and let us try to describe how does it looks like a Brownian motion which is conditioned to stay positive forever. In order to do so let $R>0$, $B_0=x$ for some $x\in(0,R)$ and consider the stopping times $\tau_y=\inf\{t\geqslant 0: B_t=y\}$ for $y\in\mathbb{R}$ and $\tau_D=\inf\{t\geqslant 0: B_t\notin D\}$ where $D=(0,R)$. Let

$$\mathbb{Q}(A)=\mathbb{P}(A|\tau_D<+\infty,B_{\tau_D}=R)=\frac{\mathbb{P}(A\cap\{\tau_D<+\infty,B_{\tau_D}=R\})}{\mathbb{P}(\tau_D<+\infty,B_{\tau_D}=R)}$$

which represents the probability of A conditioned on the event that the Brownian motion escapes the domain D at R, namely that it reaches R before ever touching 0.

Assume now that $A\in\mathscr{F}_s$ and use the Markov property of the Brownian motion to compute

$$\mathbb{E}[\mathbb{1}_A\mathbb{1}_{\tau_D<+\infty,B_{\tau_D}=R}]=\mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_{\tau_D<+\infty,B_{\tau_D}=R}|\mathscr{F}_s]]$$

$$=\mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_{\tau_D<s,B_{\tau_D}=R}|\mathscr{F}_s]]+\mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_{s\leqslant\tau_D<\infty,B_{\tau_D}=R}|\mathscr{F}_s]]$$

$$=\mathbb{E}[\mathbb{1}_A\mathbb{1}_{\tau_D<s,B_{\tau_D}=R}]+\mathbb{E}[\mathbb{1}_A\mathbb{1}_{s\leqslant\tau_D}\mathbb{E}[\mathbb{1}_{\tau_D<\infty,B_{\tau_D}=R}|B_s]]$$

But now $\mathbb{E}[\mathbb{1}_{\tau_D<\infty,B_{\tau_D}=R}|B_s]=h(B_s)$ where

$$h(x)=\mathbb{E}_x[\mathbb{1}_{\tau_D<\infty,B^x_{\tau_D}=R}]=\frac{x}{R},\qquad x\in(0,R),$$

and where we denoted $\mathbb{E}_x$ the expectation wrt. the probability measure under which $(B_t)_{t\geqslant 0}$ is a Brownian motion starting at $x\in(0,R)$. Therefore

$$\mathbb{E}[\mathbb{1}_A\mathbb{1}_{\tau_D<+\infty,B_{\tau_D}=R}]=\mathbb{E}[\mathbb{1}_A\mathbb{1}_{\tau_D<s,B_{\tau_D}=R}]+\mathbb{E}[\mathbb{1}_A\mathbb{1}_{s\leqslant\tau_D}h(B_s)]=\mathbb{E}[\mathbb{1}_Ah(B_{s\wedge\tau_D})].$$

Note that $s\mapsto h(B_s)$ is a martingale and therefore also $s\mapsto h(B_{s\wedge\tau_D})$ as a consequence we have obtained that for all $s\geqslant 0$

$$\frac{\mathrm{d}\mathbb{Q}^R|_{\mathscr{F}_s}}{\mathrm{d}\mathbb{P}|_{\mathscr{F}_s}}=h(B_{s\wedge\tau_D})$$

and we are now in the setting where we can use the Girsanov transform to obtain that under $\mathbb{Q}^R$ the process $(B_t)_{t\geqslant 0}$ is a semimartingale satisfying

$$\mathrm{d}B_t = \mathrm{d}W_t + \frac{h'(B_{t\wedge\tau_D})}{h(B_{t\wedge\tau_D})}\mathrm{d}t = \mathrm{d}W_t + \frac{1}{B_{t\wedge\tau_D}}\mathrm{d}t, \tag{20}$$

where $(W_t)_{t\geqslant 0}$ is a Brownian motion.

Strickly speaking we have constructed a family of measure $(\mathbb{Q}^R)_{R>0}$ indexed by R. Under each of them (20) is satisfied with $\tau_D = \tau_{(0,R)}$. Now note that

$$\mathbb{Q}^R(\tau_0 < \tau_R) = \frac{\mathbb{P}(\tau_0 < \tau_R < +\infty, B_{\tau_{(0,R)}} = R)}{\mathbb{P}(\tau_{(0,R)} < +\infty, B_{\tau_{(0,R)}} = R)} = 0$$

so under any of the $\mathbb{Q}^R$ we have $\tau_0 > \tau_R$ and $\tau_{(0,R)} = \tau_R$ almost surely. Moreover if $0 < R < R'$ we also have $\tau_{R'} > \tau_R$ and that $\mathbb{Q}^R|_{\mathscr{F}_{\tau_R}} = \mathbb{Q}^{R'}|_{\mathscr{F}_{\tau_R}}$. So the family of measure $(\mathbb{Q}^R, \mathscr{F}_{\tau_R})_{R\geqslant 0}$ is a consistent family of probabilities defined on an increasing sequence of σ-algebras $(\mathscr{F}_{\tau_R})_{R\geqslant 0}$ and they identify a unique probability measure $\mathbb{Q}$ on the σ-algebra $\mathscr{G} = \sigma((\mathscr{F}_{\tau_R})_{R\geqslant 0})$. However we have also

$$\lim_{R\to\infty} \mathbb{Q}(\tau_R < \infty) = \lim_{R\to\infty} \mathbb{Q}^R(\tau_R < \infty) = \lim_{R\to\infty} \frac{\mathbb{P}(\tau_R < +\infty, B_{\tau_{(0,R)}} = R)}{\mathbb{P}(\tau_{(0,R)} < +\infty, B_{\tau_{(0,R)}} = R)} = 0.$$

so under $\mathbb{Q}$ it happens almost surely that $\tau_R \to \infty$ as $R \to \infty$. We conclude that under $\mathbb{Q}$ the process B satisfies

$$\mathrm{d}B_t = \mathrm{d}W_t + \frac{1}{B_t}\mathrm{d}t, \tag{21}$$

where $(W_t)_{t\geqslant 0}$ is a BM.

As a conclusion we have constructed the law $\mathbb{Q}$ of a BM conditioned to stay positive forever and verified that under $\mathbb{Q}$ the canonical process B satisfies the SDE (20). This SDE has regular coefficients away from 0 and therefore pathwise uniqueness and existence of strong solutions until the first time the processes touch 0. We have also learned that $\mathbb{Q}(\tau_0 = +\infty) = 1$ so the process B never touches 0 from which is now easy to deduce that *any* weak solution to (20) has the same law $\mathbb{Q}$ and that any weak solution never touches zero.

Definition 11. *A weak solution of (20) is called a three-dimensional Bessel process.*

The reason for this strange names comes as follows. Let $(X_t)_{t\geqslant 0}$ be a d-dimensional Brownian motion and consider the process $R_t = |X_t|$ where the bars denote the Euclidean norm in $\mathbb{R}^d$. Using Ito formula is easy to deduce that whenever $|X_t| > 0$ we have

$$\mathrm{d}R_t = \frac{X_t}{|X_t|}\cdot \mathrm{d}X_t + \frac{d-1}{2}\frac{1}{|X_t|}\mathrm{d}t = \mathrm{d}W_t + \frac{d-1}{2}\frac{1}{R_t}\mathrm{d}t$$

where we let $W_t = \int_0^t \frac{X_s}{|X_s|}\mathrm{d}X_s$. The notation is justified by the fact that

$$[W,W]_t = \sum_{\alpha,\beta=1}^{d} \int_0^t \frac{X_s^\alpha}{|X_s|}\frac{X_s^\beta}{|X_s|}\mathrm{d}[X^\alpha, X^\beta]_s = \sum_{\alpha=1}^{d} \int_0^t \frac{X_s^\alpha}{|X_s|}\frac{X_s^\alpha}{|X_s|}\mathrm{d}s = t$$

and therefore $(W_t)_{t\geqslant 0}$ is indeed a Brownian motion by Lévy's theorem. We conclude that

Theorem 12. *Let $(X_t)_{t\geqslant 0}$ be a d–dimensional Brownian motion, then the one dimensional process $R_t = |X_t|$ is a weak solution of the SDE*

$$\mathrm{d}R_t = \mathrm{d}W_t + \frac{d-1}{2}\frac{1}{R_t}\mathrm{d}t. \tag{22}$$

In particular if $d = 3$ this process coincides with a Brownian motion conditioned to stay positive.

Definition 13. *For any $d \geqslant 0$ (non necessarily integer) the weak solution of (22) is called a d-dimensional Bessel process.*

Using our result on conditioned Brownian motion we deduce that the d-dimensional Brownian motion do not hit any point with probability 1 (why?) if $d \geqslant 3$. Namely for any $y \in \mathbb{R}^d$ we have

$$\mathbb{P}(\exists t \geqslant 0 : X_t = y) = 0.$$

Actually one can prove that the d-dimensional Bessel process never hit the origin as soon as $d \geqslant 2$, this is due to the strong repulsive drift appearing in (22).

Theorem 14. *Let $(R_t)_{t\geqslant 0}$ be the d-dimensional Bessel process with $d \geqslant 2$ and $R_0 = r \geqslant 0$. Then $R_t > 0$ for all $t > 0$ almost surely.*

Proof. (See Bovier notes) □