

## 26. METHODS FOR SOLVING SDES

We will now move to discussing methods for solving SDEs. In one of these we also establish a general theorem for change of variables in stochastic integrals.

**26.1 Reduction to ODE.**

The first method is based on reduction to an ordinary differential equation (ODE) assuming the variance of the diffusive term is constant (we set it to 1 for simplicity).

**Lemma 26.1** (Reduction to ODE) *Consider the SDE*

$$dX_t = a(t, X_t)dt + dB_t \quad (26.1)$$

*Then  $X$  is a strong solution if and only if*

$$V_t := X_t - B_t \quad (26.2)$$

*is a stochastic process with absolutely continuous paths whose weak (a.k.a. Lebesgue) derivative satisfies the ODE*

$$\frac{dV_t}{dt} = a(t, V_t + B_t) \quad (26.3)$$

*for Lebesgue-a.e.  $t \geq 0$ . In particular, if (26.3) has a solution for every  $B$ , then so does (26.1).*

*Proof.* Being a solution to the SDE means that

$$\forall t \geq 0: \quad X_t = X_0 + \int_0^t a(s, X_s)ds + B_t \quad (26.4)$$

This is equivalent to  $V$  satisfying

$$V_t = V_0 + \int_0^t a(s, V_s + B_s)ds \quad (26.5)$$

which, in particular, means that  $t \mapsto V_t$  is absolutely continuous and thus Lebesgue differentiable with the derivative satisfying the above ODE.  $\square$

We remark that, by Carathéodory's generalization of Peano existence theorem, (26.3) will have a solution as long as  $x \mapsto a(t, x)$  is continuous for each  $t \geq 0$ ,  $t \mapsto a(t, x)$  is measurable for each  $x \in \mathbb{R}$  and  $t \mapsto |a(t, x)|$  is dominated by a locally Lebesgue integrable function (of  $t$ ) on each compact set of  $x$ . (This uses that  $t \mapsto B_t$  is continuous.) A key point in any approach is that the ODE solves the SDE for *every* continuous path  $B$ , not just a.e. part as the general method presented earlier.

The reduction to an ODE is advantageous for numerical schemes because it effectively circumvents Itô integrals which, in light of Theorem 12.12, need to be simulated very carefully. Note that, besides the Bessel SDE (21.5), the class of equations (26.1) includes the *Langevin equation*,

$$dX_t = -\nabla H(X_t)dt + \sqrt{2} dB_t \quad (26.6)$$

where  $H: \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $C^1$ -function and  $B$  is a  $d$ -dimensional standard Brownian motion. An important feature of this SDE is the following:

**Lemma 26.2** Let  $H \in C^1(\mathbb{R}^d)$  be such that  $e^{-H(x)} \in L^1(\mathbb{R}^d)$  and let  $\mu$  be the Borel probability measure on  $\mathbb{R}^d$  given by

$$\mu(dx) := Ce^{-H(x)} dx \quad (26.7)$$

for a suitable normalizing constant  $C$ . If  $\{X_t : t \geq 0\}$  solves the SDE (26.6), then

$$\text{law of } X_0 = \mu \quad \Rightarrow \quad \forall t \geq 0: \quad X_t \stackrel{\text{law}}{=} X_0 \quad (26.8)$$

In short,  $\mu$  is a stationary distribution for the SDE (26.6).

Leaving the proof of this to homework we note that the Langevin equation arose originally in physics as an equation for the velocity of a particle that changes through random influences (represented by the  $dB_t$ -term) while being subject to friction due to contact with ambient fluid. In this case  $H$  is quadratic and the equation boils down to

$$dX_t = -\gamma X_t dt + dB_t \quad (26.9)$$

Langevin's original formulation was the equation  $\frac{dX_t}{dt} = \gamma X_t + W_t$  where  $W$  stands for the "noise" — which we interpret precisely using a white noise process.

A common use of Langevin evolution governed by the SDE (26.6) is often the opposite to what one might expect. Indeed, instead of solving the equation for the sake of studying the motion of a particle in a gradient force field, it is actually often used to study the "equilibrium" measure  $\mu$  itself.

## 26.2 Coordinate change.

Attempts to solve equations such as (26.6) naturally point to another method for solving SDEs which is based on a *coordinate change*. A deficiency of this particular approach is that we need to assume that the coefficients do not depend explicitly on time and that, for practical reasons, it is more or less restricted to one-dimensional processes.

Consider an  $\mathbb{R}$ -valued process  $X$  solving the SDE

$$dX_t = a(X_t)dt + \sigma(X_t)dB_t \quad (26.10)$$

The point of the method is to eliminate the drift term by instead looking at the process  $f \circ X$ , for some  $f \in C^2(\mathbb{R})$ . Indeed, assuming that  $X$  obeys (26.10), the Itô formula gives

$$\begin{aligned} df(X_t) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X \rangle_t \\ &= \left[ a(X_t)f'(X_t) + \frac{1}{2}\sigma(X_t)^2 f''(X_t) \right] dt + f'(X_t)\sigma(X_t)dB_t \end{aligned} \quad (26.11)$$

Setting the term in the large bracket to zero leads to an ODE for  $f'$  which is solved by

$$f'(x) = \exp\left\{-2 \int_{x_0}^x \frac{a(u)}{\sigma(u)^2} du\right\} \quad (26.12)$$

where  $x_0$  is arbitrary — albeit with a natural choice being the initial value of  $X$ . Assuming the integral exists at least locally, since  $f' > 0$ , one more integration gives us a strictly increasing function that can be regarded as a coordinate transformation on  $\mathbb{R}$ . We now formulate the conclusion with all the needed conditions spelled out explicitly:

**Lemma 26.3** (Coordinate change) *Consider the SDE (26.10) for  $\mathbb{R}$ -valued process  $X$ . Assume that, for an interval  $(\alpha, \beta) \subseteq \mathbb{R}$ ,*

$$x \mapsto \frac{a(x)}{\sigma(x)^2} \in L^{1,\text{loc}}(\alpha, \beta) \quad (26.13)$$

and that, given any  $x_0 \in (\alpha, \beta)$ ,

$$f(x) := \int_{x_0}^x \exp\left\{-2 \int_{x_0}^v \frac{a(u)}{\sigma(u)^2} du\right\} dv \quad (26.14)$$

obeys  $f(x) \rightarrow -\infty$  as  $x \downarrow \alpha$  and  $f(x) \rightarrow +\infty$  as  $x \uparrow \beta$ . Then  $X$  solves (26.10) for all times with initial value  $X_0 = x_0$  if and only if the process  $\{Z_t : t \geq 0\}$ , with  $Z_t := f(X_t)$  solves the SDE

$$dZ_t = \tilde{\sigma}(Z_t) dB_t \quad \text{for } \tilde{\sigma}(z) := f' \circ f^{-1}(z) \sigma \circ f^{-1}(z) \quad (26.15)$$

(same Brownian motion and filtration) with initial value  $Z_0 = f(x_0)$ . In particular, we have

$$\forall t \geq 0: \quad X_t \in (\alpha, \beta) \quad \text{a.s.} \quad (26.16)$$

*Proof.* Assume  $X$  is a solution to (26.10). Under (26.13),  $f \in C^1(\mathbb{R})$  with  $f'$  absolutely continuous. The Itô formula still applies in this case (see Lemma 24.4) up to the stopping time  $\tau := \inf\{t \geq 0: X_t \notin (\alpha, \beta)\}$ . The computation (26.11) then shows that  $Z$  indeed obeys (26.15). To prove (26.16), we note a fact whose proof we leave to homework:

**Lemma 26.4** *Let  $x \mapsto \sigma(x)$  be a Borel measurable function and suppose that  $Z$  is a strong solution to the SDE*

$$dZ_t = \sigma(Z_t) dB_t \quad (26.17)$$

up to the stopping time  $\tau_M := \inf\{t \geq 0: |Z_t| \geq M\}$ . Then  $\tau_M \rightarrow \infty$  as  $M \rightarrow \infty$  a.s. In particular, strong solutions to the SDE (26.17) cannot blow up in finite time.

In light of  $f^{-1}(\mathbb{R}) = (\alpha, \beta)$ , this means that  $\tau = \infty$  and so  $X$  solves the SDE (26.10) for all positive times.

Conversely, if  $Z$  is a solution to (26.15), then setting  $X_t := h(Z_t)$  for  $h \in C^1(\mathbb{R})$  with  $h'$  absolutely continuous gives

$$dX_t = h'(Z_t) \tilde{\sigma}(Z_t) dB_t + \frac{1}{2} h''(Z_t) \tilde{\sigma}(Z_t)^2 dt \quad (26.18)$$

For  $h := f^{-1}$  we get  $h'(z) \tilde{\sigma}(z) = \sigma \circ h(z)$  and

$$h''(z) \tilde{\sigma}(z)^2 = \left(\frac{1}{f' \circ h}\right)'(z) \tilde{\sigma}(z)^2 = \frac{-1}{f' \circ h(z)^3} f'' \circ h(z) \tilde{\sigma}(z) = -\frac{f'' \circ h(z)}{f' \circ h(z)} \sigma \circ h(z)^2 \quad (26.19)$$

Using also that (26.14) solves the ODE

$$a(x) f'(x) + \frac{1}{2} \sigma(x)^2 f''(x) = 0 \quad (26.20)$$

on  $\text{Ran}(h) = (\alpha, \beta)$ , we get  $h''(z) \tilde{\sigma}(z)^2 = 2a \circ h(z)$ . Plugging these observations in (26.18), we get (26.10).  $\square$

The coordinate change works even in higher dimensions. Indeed, here the ODE (26.20) becomes

$$\sum_{i,j=1}^m \sigma_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^m a_i(x) \frac{\partial f}{\partial x_i}(x) = 0 \quad (26.21)$$

The problem here is that such equations have many general solutions and work is thus needed to choose one that defines a proper coordinate transform of  $\mathbb{R}^m \rightarrow \mathbb{R}^m$ . A slight exception is the case of  $m = 2$  where one can sometimes cast the problem using complex variables which makes the equation effectively one-dimensional.

### 26.3 Time change.

The coordinate change (as well as other general considerations) have naturally led us to the class of “pure” SDEs:

$$dX_t = \sigma(X_t)dB_t \quad (26.22)$$

We will now present a method for solving these which is based on re-parametrizing the time and that so even depending on the path of the underlying Brownian motion. For this we need a result allowing us to perform a time change inside stochastic integrals.

**Theorem 26.5** (Change of variables in Itô integral) *Given a probability space supporting a Brownian motion  $\{B_t: t \geq 0\}$  that is adapted to a filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  with  $\mathcal{F}_0$  containing all null sets, let  $\{Z_t: t \geq 0\}$  be a jointly-measurable, adapted process such that*

$$\forall t \geq 0: \quad U_t := \int_0^t Z_s^2 ds < \infty \quad \text{a.s.} \quad (26.23)$$

and

$$\lim_{t \rightarrow \infty} U_t = \infty \quad \text{a.s.} \quad (26.24)$$

Denote  $T(u) := \inf\{t \geq 0: U_t \geq u\} < \infty$  and recall from Theorem 18.2 that that  $T(u)$  is a stopping time for  $\{\mathcal{F}_t\}_{t \geq 0}$  satisfying  $T(u) < \infty$  for all  $u < \infty$  with  $T(u) \rightarrow \infty$  as  $u \rightarrow \infty$  a.s. and that the process

$$\forall u \geq 0: \quad \tilde{B}_u := \int_0^{T(u)} Z_s dB_s \quad (26.25)$$

is a standard Brownian motion adapted to the filtration  $\{\mathcal{F}_{T(u)}\}_{u \geq 0}$ . Then, given a left-continuous process  $\{Y_t: t \geq 0\}$  adapted to  $\{\mathcal{F}_t\}_{t \geq 0}$ , the process  $\{Y_{T(u)}: u \geq 0\}$  is also left-continuous, adapted to  $\{\mathcal{F}_{T(u)}\}_{u \geq 0}$ , the identity

$$\forall u \geq 0: \quad \int_0^u Y_{T(v)}^2 dv = \int_0^{T(u)} Y_s^2 Z_s^2 ds \quad (26.26)$$

holds and, assuming these integrals are finite a.s. for all  $u \geq 0$ , we have

$$\forall u \geq 0: \quad \int_0^u Y_{T(v)} d\tilde{B}_v = \int_0^{T(u)} Y_s Z_s dB_s \quad \text{a.s.} \quad (26.27)$$

with the stochastic integrals well defined and finite a.s.

*Proof.* As  $T(0) = 0$ , also  $\mathcal{F}_{T(0)}$  contains all null sets and so both filtrations allow us to select adapted continuous versions. Since  $s \mapsto Y_s$  is left continuous, and so is  $u \mapsto T(u)$ , the process  $u \mapsto Y_{T(u)}$  is left-continuous and thus jointly measurable. In order to prove (26.26), note that for  $Y_s := Y_a 1_{(a,b]}(s)$ , this boils down to

$$\begin{aligned} \int_0^u Y_{T(v)}^2 dv &= Y_a^2 \lambda(\{v \geq 0: v < u \wedge a < T(v) \leq b\}) = Y_a^2 (U_b \wedge u - U_a \wedge u) \\ &= Y_a^2 \left( \int_0^{b \wedge T(u)} Z_s^2 ds - \int_0^{a \wedge T(u)} Z_s^2 ds \right) = \int_0^{T(u)} Y_s^2 Z_s^2 ds \end{aligned} \quad (26.28)$$

based on the equivalence  $T(v) > a \Leftrightarrow U_a < v$  and the facts that  $U_{T(u)} = u$  by continuity of  $U$ . Additivity then extends (26.26) to all simple processes. Since left-continuous  $Y$  can be approximated in  $L^2$  by simple processes, we get (26.26) in the stated generality.

The approximation arguments for Itô integrals reduce (26.27) to  $Y$  simple and, by additivity,  $Y$  of the form  $Y_s := 1_{(0,a]}(s)$ . In this case the left-integral equals

$$\int_0^u Y_{T(v)} d\tilde{B}_v = \int_0^u 1_{(0,U_a]}(v) d\tilde{B}_v = \tilde{B}_{u \wedge U_a} = \int_0^{T(u) \wedge a} Z_s dB_s \quad (26.29)$$

thus equating it with the integral on the right.  $\square$

As stated above, the change of variables might seem somewhat confusing. It is often better to implement it using the formal calculus rules:

$$\begin{aligned} du &= Z_t^2 dt \\ d\tilde{B}_u &= Z_t dB_t \end{aligned} \quad (26.30)$$

which then change the integral of interest as

$$\int_0^{t(u)} Y_s Z_s dB_s = \int_0^u Y_{t(u)} dB_u \quad (26.31)$$

We also mention that the assumption of left-continuity can be weakened to the assumption of progressive measurability but the above proof becomes more complicated.

We can now move to the solution of the SDE (26.22). The main idea is as follows: Since  $X$  is a local martingale, it is a time change of Brownian motion. So we may try to start by defining  $X$  as a Brownian motion which then needs to be time-changed so that the process  $B$  on the right hand side of (26.22) becomes a Brownian motion. This naturally leads to:

**Theorem 26.6** *Let  $\sigma: \mathbb{R} \rightarrow \mathbb{R}$  be Borel measurable and such that*

$$\forall M \geq 0: \int_{[-M,M]} \frac{1}{\sigma(x)^2} dx < \infty \quad (26.32)$$

*Let  $\{W_t: t \geq 0\}$  be a Brownian motion started from  $x \in \mathbb{R}$ . Then*

$$\forall t \geq 0: U_t := \int_0^t \frac{1}{\sigma(W_s)^2} ds < \infty \quad \text{a.s.} \quad (26.33)$$

with

$$\lim_{t \rightarrow \infty} U_t = +\infty \quad \text{a.s.} \quad (26.34)$$

and so, for  $T(u) := \inf\{t \geq 0: U_t \geq u\}$ ,

$$B_u := \int_0^{T(u)} \frac{1}{\sigma(W_s)} dW_s \quad (26.35)$$

is a standard Brownian motion for which

$$X_u := W_{T(u)} \quad (26.36)$$

solves the SDE  $dX_u = \sigma(X_u)dB_u$  with initial value  $X_0 = x$ .

*Proof.* We start by noting that the local integrability requirement (26.32) ensures the a.s. finiteness of  $U_t$  for every  $t \geq 0$ . For this let  $\tilde{\tau}_M := \inf\{t \geq 0: |W_t - x| \geq M\}$ . Then Fubini-Tonelli and the fact that the probability density of  $W_s$  is bounded by  $s^{-1/2}$  tell us

$$\begin{aligned} E\left(1_{\{\tilde{\tau}_M > t\}} \int_0^t \frac{1}{\sigma(W_s)^2} ds\right) &\leq \int_0^t E\left(\frac{1}{\sigma(W_s)^2} 1_{\{|W_s - x| \leq M\}}\right) ds \\ &\leq \int_0^t \frac{ds}{\sqrt{s}} \int_{-M}^M \frac{1}{\sigma(x+y)^2} dy \end{aligned} \quad (26.37)$$

This is finite which proves  $U_t < \infty$  a.s. on  $\bigcup_{M \geq 1} \{\tilde{\tau}_M > t\}$ . By continuity of  $W$  the latter event is of full measure, thus proving  $U_t < \infty$  a.s.

With (26.33) in hand, we now observe that, thanks to the recurrence of standard Brownian motion and the strong Markov property,  $U_\infty := \lim_{t \rightarrow \infty} U_t$  equals the sum of an infinite number of i.i.d. copies of the random variable

$$\int_0^{\tau_x} \frac{1}{\sigma(W_s)^2} ds \quad (26.38)$$

where  $\tau_x := \inf\{t \geq 1: W_t = x\}$ . This random variable is positive with positive probability (in fact, a.s.) and so  $U_\infty = \infty$  a.s. It follows that the process in (26.35) is a standard Brownian motion by Theorem 18.2. Theorem 26.5 in turn shows

$$\int_0^u \sigma(W_{T(v)}) dB_v = \int_0^{T(u)} \sigma(W_s) \frac{1}{\sigma(W_s)} dW_s = \int_0^{T(u)} dW_s = W_{T(u)} \quad (26.39)$$

Writing (26.36), the process  $X$  indeeds solves the desired SDE.  $\square$

The above theorem, albeit rather straightforward, is not even close to the end of the story. The first question that naturally springs in mind is uniqueness. The Tanaka equation — to which Theorem 26.6 applies, being modeled on our analysis thereof — shows that pathwise uniqueness cannot generally be expected, so the question is whether uniqueness in law might hold. Here we note that, if  $\sigma(x_0) = 0$  for some  $x_0 \in \mathbb{R}$ , then  $X_t = x_0$  is a solution. However, whenever (26.32) applies and  $X_t$  is the solution with  $X_0 = x_1 \neq x_0$  constructed in Theorem 26.6, then setting

$$\tilde{X}_t := \begin{cases} X_t & \text{if } t < \tau_{x_0} \\ x_0 & \text{if } t \geq \tau_{x_0} \end{cases} \quad (26.40)$$

where  $\tau_{x_0} := \inf\{t \geq 0: X_t = x_0\}$  gives us another solution that, since  $\tau_{x_0} < \infty$  a.s., will be different in law from  $X$ . Taking  $\tau'_{x_0} := \inf\{t \geq \tau_x + 1: X_t = x_0\}$ , we can even set

$$\tilde{X}'_t := \begin{cases} X_t & \text{if } t < \tau_{x_0} \vee t > \tau'_{x_0} \\ x_0 & \text{if } \tau_{x_0} \leq t \leq \tau'_{x_0} \end{cases} \quad (26.41)$$

and still get a solution. There are two ways to prevent such situations:

- (1)  $x_0 \in \mathbb{R}$  with  $\sigma(x_0) = 0$  will not be hit by a solution started away from  $x_0$ , or
- (2) such an  $x_0$  can be hit but then the solution must get stuck there forever.

As a result of such considerations, one then gets a full characterization of existence of a weak solution and uniqueness in law:

**Theorem 26.7** (Engelbert and Schmidt) *Let  $\sigma: \mathbb{R} \rightarrow \mathbb{R}$  be Borel measurable and set*

$$\mathcal{I}(\sigma) := \bigcap_{\epsilon > 0} \left\{ x \in \mathbb{R}: \int_{(-\epsilon, \epsilon)} \frac{1}{\sigma(x+y)^2} dy = \infty \right\} \quad (26.42)$$

and

$$\mathcal{Z}(\sigma) := \{x \in \mathbb{R}: \sigma(x) = 0\} \quad (26.43)$$

Then a weak solution to

$$dX_t = \sigma(X_t)dB_t \quad (26.44)$$

exists and is unique for every initial value if and only if  $\mathcal{I}(\sigma) = \mathcal{Z}(\sigma)$ .

To see that this is quite intuitive note that, in light of the time-change via (26.33), the divergence of the integral of  $y \mapsto \sigma(x+y)^{-2}$  over  $(0, \epsilon)$  means that no solution can reach  $x$  in finite time from the right, while the divergence on  $(-\epsilon, 0)$  means the same for solutions from the left. Somewhat harder is to convince oneself that, even when one of these integral diverges, a solution started from  $x$  has to remain at  $x$  forever.

### 26.4 Time change and reduction to ODE.

As our final technique we combine the methods of time change and reduction to ODE as follows. Consider the SDE

$$dX_u = a(u, X_u)du + \sigma(u, X_u)dB_u \quad (26.45)$$

Introducing the time change  $(u, B) \rightarrow (t, \tilde{B})$  by a process  $\{Z_t: t \geq 0\}$  where

$$\begin{aligned} du(t) &= Z_t^2 dt \\ dB_{u(t)} &= Z_t d\tilde{B}_t \end{aligned} \quad (26.46)$$

brings the SDE to the form

$$dX_{u(t)} = a(u(t), X_{u(t)})Z_u^2 du + \sigma(u(t), X_{u(t)})Z_u d\tilde{B}_u \quad (26.47)$$

We now set  $Z_u := \sigma(u(t), X_{u(t)})^{-1}$  to rewrite this as

$$dX_{u(t)} = \frac{a(u(t), X_{u(t)})}{\sigma(u(t), X_{u(t)})^2} du + d\tilde{B}_u \quad (26.48)$$

This is turned into an ODE by passing to  $V_t := X_{u(t)} - \tilde{B}_u$  with another ODE arising from the differential formula for  $u(t)$ . Hereby we get:

**Theorem 26.8** *Suppose that  $(u, x) \mapsto \sigma(u, x)$  is continuous and let  $\tilde{B}$  be a standard Brownian motion. Assume that  $t \mapsto (U_t, V_t)$  is a solution to the system of ODEs*

$$\frac{dU_t}{dt} = \frac{1}{\sigma(U_t, V_t + \tilde{B}_t)^2} \quad \wedge \quad \frac{dV_t}{dt} = \frac{a(U_t, V_t + \tilde{B}_t)}{\sigma(U_t, V_t + \tilde{B}_t)^2} \quad (26.49)$$

with initial values  $(U_0, V_0) := (0, x_0)$ , for some  $x_0 \in \mathbb{R}$ . Suppose that  $U_t < \infty$  for all  $t \geq 0$  and with  $U_t \rightarrow \infty$  as  $t \rightarrow \infty$ . Then, setting  $T(u) := \inf\{t \geq 0 : U_t \geq u\}$ , the process

$$B_u := \int_0^{T(u)} \frac{1}{\sigma(U_t, V_t + \tilde{B}_t)^2} dt \quad (26.50)$$

is a standard Brownian motion and

$$X_u := V_{T(u)} + \tilde{B}_{T(u)} \quad (26.51)$$

is a global (weak) solution to the SDE (26.45) with  $X_0 = x_0$ .

The proof of this theorem uses time change as indicated in the calculation above. The conditions ensure that Theorem 26.5 can be used as stated. The upshot of the theorem is that, for a fairly generic class of SDEs, a full-fledged SDE theory is actually not required. Note also that a solution to (26.49) will typically exist for *all* Brownian paths, not just a full-measure set thereof. Finally, working with ODE is easier when one wishes to compute the solutions numerically.

Further reading: Karatzas-Shreve, Section 5.5AB