Stochastic Differential Equations

Jun Wei Tan[*]

Julius-Maximilians-Universität Würzburg

(Dated: January 24, 2025)

> **Definition 1.** A stochastic differential equation is a (formal) equation of the form
>
> $$\frac{\mathrm{d}X_t}{\mathrm{d}t} = b(t, X_t) + \sigma(t, X_t)W_t, \tag{1}$$
>
> where W_t is white noise.

The above definition needs some explanation. Intuitively, the motivation is quite clear - we often need to perturb a (deterministic) differential equation with some random noise term. For example, a particle moving in a random environment will be subject to a force that, at each time, is random. As a candidate for white noise, we choose the (also formal) derivative of Brownian motion $W_t = \frac{\mathrm{d}B_t}{\mathrm{d}t}$. Physically, this can be justified by thinking of Brownian motion as the "fundamental" stochastic process - Brownian motion is the motion generated when the velocity is equal to simple white noise.

This definition allows is to write down the mathematically rigorous definition of a *solution* to an SDE:

> **Definition 2.** We say that the stochastic process X_t is a *solution* of the SDE (1) if
>
> $$X_t = X_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s$$

At this point, it must be said that there is some ambiguity in this - in particular, we must specify whether the second integral is to be interpreted in the Stratonovich or Itô sense. There are some other interpretations of SDEs proposed in current research [1], but we will also not consider them here. Here, we will consider the Itô integral only.

There are, as usual, two questions when it comes to such differential equations: Firstly, how do we solve them? Secondly, under what conditions do we have existence and unique-

---
[*] jun-wei.tan@stud-mail.uni-wuerzburg.de

ness?

The answer to the first question follows very much like the case of an ordinary differential equation. An ordinary differential equation is integrated most often by the fundamental theorem of calculus. Thus, we turn to the stochastic variation, the Itô formula, for answers.

Theorem 3 (The 1-Dimensional Itô Formula). Suppose X_t is an Itô process defined by the formula

$$\mathrm{d}X_t = u\,\mathrm{d}t + v\,\mathrm{d}B_t.$$

Let $g(t,x) \in C^2([0,\infty) \times \mathbb{R})$. Then

$$Y_t = g(t, X_t)$$

is also an Itô process, and

$$\mathrm{d}Y_t = \frac{\partial g}{\partial t}\,\mathrm{d}t + \frac{\partial g}{\partial x}\,\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\cdot(\mathrm{d}X_t)^2$$

where $(\mathrm{d}X_t)^2$ is computed according to the rules $\mathrm{d}t\cdot\mathrm{d}t = \mathrm{d}t\cdot\mathrm{d}B_t = 0$ and $\mathrm{d}B_t\cdot\mathrm{d}B_t = \mathrm{d}t$.

This leads us to our first, and most fundamental solution method, both of stochastic and of deterministic equations - guessing the solution. We begin with a simple population model as an example.

Example 4. A model of a population is given by the stochastic differential equation

$$\frac{\mathrm{d}N_t}{\mathrm{d}t} = rN_t + \alpha W_t N_t.$$

Here, r, α are constants and W_t is white noise.

This is the well known model for a population, except that we have allowed r to vary by a white noise term. The solution in the nonstochastic limit is given by a simple exponential. To solve this, we first rewrite the SDE in standard form:

$$\mathrm{d}N_t = rN_t\,\mathrm{d}t + \alpha N_t\,\mathrm{d}B_t,$$

or

$$\frac{\mathrm{d}N_t}{N_t} = r\,\mathrm{d}t + \alpha\,\mathrm{d}B_t.$$

Inspired by the solution in the deterministic case, we guess $g(t, x) = \ln x$ in Ito's formula, and let $Y_t = g(t, N_t)$. Then, we have

$$\mathrm{d}Y_t = \frac{\mathrm{d}N_t}{N_t} + \frac{1}{2}\alpha^2 \, \mathrm{d}t \, .$$

Substituting, we have

$$\mathrm{d}Y_t = \left(r - \frac{1}{2}\alpha^2\right)\mathrm{d}t + \alpha B_t$$

which integrates easily to yield

$$N_t = N_0 \exp\left(\left(r - \frac{1}{2}\alpha^2\right)t + \alpha B_t\right).$$

Clearly, for $\alpha = 0$, this reduces to the well known exponential solution. This process is known as <u>geometric Brownian motion</u>, which, for example, is a solution to the Black-Scholes equation [2].

Now, we turn to the questions of existence and uniqueness. Recall the basic theorem for existence and uniqueness of a deterministic differential equation

Theorem 5 (Picard-Lindelöf). Let $\dot{x} = f(t, x)$ be a differential equation with f defined on a rectangle $[a, b] \times \mathbb{R}^n$. If f is locally Lipschitz continuous in x, with Lipschitz constant independent of time, and continuous in time, then the differential equation has a unique global solution on $[a, b]$.

We prove the uniqueness similar to in [3]. First, we require an inequality for deterministic differential equations known as Grönwall's inequality:

Lemma 6 (Grönwall's Inequality). Suppose $\psi(t)$ satisfies

$$\psi'(t) \leq \beta(t)u(t),$$

on an interval $[a, b]$. Then $\psi(t)$ is bounded by the solution of the ODE

$$v'(t) = \beta(t)v(t)$$

for all time:

$$\psi'(t) \leq \psi(a)\exp\left(\int_a^t \beta(s)\,\mathrm{d}s\right)$$

Proof. We define $v(t) = \exp\left(\int_a^t \beta(s)\,\mathrm{d}s\right)$. Note that v satisfies the differential equation $v'(t) = \beta(t)v(t)$ and is always strictly greater than 0.

Then we consider the derivative of $u(t)/v(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{u(t)}{v(t)} = \frac{u'(t)v(t) - v'(t)u(t)}{v(t)^2} = \frac{u'(t) - \beta(t)u(t)}{v(t)} \le 0$$

Thus the function $u(t)/v(t)$ starts at 1 and is strictly decreasing, which proves the above inequality. $\square$

Lemma 7 (Generalized Grönwall's Inequality). Statement from `https://math.stackex` `change.com/questions/4773818/generalized-gronwall-inequality-covering-m` `any-different-applications`:

Suppose ψ is a function on $[a, b]$ that satisfies the differential inequality

$$\psi'(t) \le F(\psi(t))$$

with $F \in C^1(\mathbb{R})$, while v satisfies the differential equation

$$v'(t) = F(v(t)), \qquad v(a) = \psi(a).$$

Then $\psi(t) \le v(t)$.

Proof. Define $G : \mathbb{R}^2 \to \mathbb{R}$ by

$$G = \begin{cases} \frac{F(x) - F(y)}{x - y} & x \neq y, \\ F'(x) & x = y. \end{cases}$$

Since F is C^1, G is continuous. Then we define $\beta(t) = G(\psi(t), v(t))$ and $w(t) = u(t) - v(t)$. Then

$$w'(t) = u'(t) - v'(t) \le F(u(t)) - F(v(t)) = [u(t) - v(t)]G(u(t), v(t)) = \beta(t)w(t).$$

By the classical version of Grönwall's Inequality 6, we have

$$w(t) \le w(0)\exp\left(\int_a^t \beta(s)\,\mathrm{d}s\right) = 0,$$

as desired. $\square$

Corollary 8. Suppose $\psi(t)$ satisfies

$$\psi'(t) \leq \alpha'(t) + \beta(t)\psi(t).$$

Then

$$\psi(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s) \exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s$$

Proof. We just need to show that the second function is a solution to the differential equation obtained by replacing the inequality with an inequality:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\alpha(t) + \int_a^t \alpha(s)\beta(s)\exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s\right]$$

$$=\alpha'(t) + \alpha(t)\beta(t) + \int_a^t \alpha(s)\beta(s)\beta(t)\exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s$$

$$=\alpha'(t) + \beta(t)\left[\alpha(t) + \int_a^t \alpha(s)\beta(s)\exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s\right]. \qquad \square$$

Then, we define a slightly weaker notion of Lipschitz continuity:

Definition 9 (Linear Growth). A function $g(t,x)$ on $[a,b] \times \mathbb{R}$ is said to satisfy the <u>linear growth condition in x</u> if there exists $K > 0$ such that

$$|g(t,x)| \leq K(1 + |x|)$$

for all $t \in [a,b]$ and $x \in \mathbb{R}$.

Because of the inequality

$$1 + x^2 \leq (1+x)^2 \leq 2(1+x^2),$$

the earlier definition is equivalent to the following definition:

Proposition 10. A function $g(t,x)$ satisfies the linear growth condition if there exists $K > 0$ such that

$$|g(t,x)|^2 \leq K(1 + x^2)$$

The second definition will be more useful in the following proofs. Note that it is indeed stronger

Proposition 11. Globally Lipschitz continuous functions satisfy the linear growth condition.

Proof. If $f(x)$ is Lipschitz, we have

$$|f(x) - f(0)| \leq K|x|,$$

or

$$|f(x)| \leq |f(0)| + K|x|$$

by the reverse triangle inequality. $\qquad\square$

Conversely, linear growth does not imply Lipschitz continuity, or any kind of continuity

Example 12. Let

$$f(x) = \begin{cases} x & x \in \mathbb{Q}, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, f grows linearly, but is not Lipschitz continuous, or even continuous at any point.

Now, we continue to the proof of uniqueness:

Theorem 13. Let $\sigma(t, x)$ and $f(t, x)$ be measurable functions on $[a, b] \times \mathbb{R}$ satisfying the Lipschitz condition in x. Suppose ξ is an $\mathcal{F}_a$-measurable random variable satisfying $\mathbb{E}[\xi^2] < \infty$. Then the stochastic differential equation

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = f(t, X_t) + \sigma(t, X_t)W_t$$

has at most one solution on $[a, b]$.

Proof. As one might expect, we assume that we have two solutions X_t and Y_t to the stochastic differential equation. Then $Z_t = X_t - Y_t$ is a stochastic process satisfying the stochastic differential equation, or equivalently integral equation

$$Z_t = \int_a^t (\sigma(s, X_s) - \sigma(s, Y_s)) \, \mathrm{d}B(s) + \int_a^t (f(s, X_s) - f(s, Y_s)) \, \mathrm{d}s \,.$$

Then we consider the expectation value of the square Z_t^2. Because $(a+b)^2 < 2(a^2+b^2)$, we have

$$Z_t^2 \leq 2\left[\left(\int_a^t (\sigma(s, X_s) - \sigma(s, Y_s))\, \mathrm{d}B_s\right)^2 + \left(\int_a^t (f(s, X_s) - f(s, Y_s))\, \mathrm{d}s\right)^2\right].$$

Now, we aim to estimate the expectation value of the two quantities in parentheses. As with deterministic differential equations, this is done using the Lipschitz condition:

$$E\left(\int_a^t (\sigma(s, X_s) - \sigma(s, Y_s))\, \mathrm{d}B(s)\right)^2 = \int_a^t E[(\sigma(s, X_s) - \sigma(s, Y_s))^2]\, \mathrm{d}s$$
$$\leq K^2 \int_a^t E(Z_s^2)\, \mathrm{d}s$$

where the first line comes from the Ito isometry and the second comes from the Lipschitz condition. Similarly, for the second term, we have

$$\left(\int_a^t (f(s, X_s) - f(s, Y_s))\, \mathrm{d}s\right)^2 \leq (t-a) \int_a^t (f(s, X_s) - f(s, Y_s))^2\, \mathrm{d}s$$
$$\leq (b-a)K^2 \int_a^t Z_s^2\, \mathrm{d}s.$$

Putting these together, we get

$$E(Z_t^2) \leq 2K^2(1 + b - a) \int_a^t E(Z_s^2)\, \mathrm{d}s.$$

This is a classical differential equation, and it is well known that the result is $E(Z_t^2) = 0$ for all $t \in [a, b]$.

This shows that $Z_t = 0$, or $X_t = Y_t$ almost surely for all $t \in [a, b]$.

The last step is to extend this to show that $X_t = Y_t$ almost surely.

To do so, first we enumerate the rationals in $[a, b]$ as $r_1, r_2, \ldots$. Because $Z_t = 0$ almost surely for each t, we can find Ω_n with $\mathbb{P}(\Omega_n) = 1$ and $Z_{r_n} = 0$ on Ω_n. Let $\Omega' = \bigcap_{k=1}^{\infty} \Omega_k$. Then $\mathbb{P}(\Omega') = 1$.

Then, by (almost surely path) continuity, we choose Ω'' such that $\mathbb{P}(\Omega'') = 1$ and $t \mapsto Z_t(\omega)$ is continuous for all $\omega \in \Omega''$.

Then, we let $\Omega_0 = \Omega' \cap \Omega''$. Clearly, Z_T is a continuous function that vanishes at all rationals in $[a, b]$. Thus, Z_t is 0 for all $t \in [a, b]$ and $\omega \in \Omega_0$. $\qquad \square$

Note that the last part of the proof is necessary. The first part, which emulated the classical case, showed that the two processes are the same for all time almost certainly. However, unlike the classical case, knowing the distributions for all time t does not suffice as a solution. As an example,

Example 14. Consider the two stochastic processes

$$X_t(\omega) = \begin{cases} 1 & t = \omega \\ 0 & \text{otherwise,} \end{cases}$$

and

$$Y_t = 0.$$

Then it is clear that for all $t \in [0, \infty)$ we have $X_t = Y_t$ almost certainly, but these two processes are clearly different. For example, X_t is sample path discontinuous, while Y_t is.

Now that we have proven the weaker result (only of uniqueness), we can move on to prove the result of existence and uniqueness. It is instructive to first consider the proof for a constant σ in Eq. (1)

Theorem 15. The differential equation (1) has a solution with σ constant, if $b(t, x)$ is Lipschitz continous in x.

Proof. Uniqueness follows from the previous theorem. For existence, we iterate

$$X_t^{(n+1)}(\omega) = X_t + \int_0^t b(s, X_s^{(n)}(\omega)) \, ds + \sigma B(t).$$

By hypothesis, b is continuous. Thus, we can show by induction that X_n has continuous paths — if we suppose so, then we can find Ω such that $\mathbb{P}(\Omega) = 1$ and $t \mapsto X_n(t, \omega)$ is continuous for all $\omega \in \Omega$.

For all such ω, $\int_0^t b(X_n(s, \omega)) \, ds$ is continuous as a function of t by the fundamental theorem of calculus, and $B(t)$ is known to be continuous. Thus, all the X_ns are almost certainly sample path continuous.

Then, we must show that the limit process satisfies the integral equation. To do so, first we show uniform convergence on compact intervals. This follows as it does in a deterministic differential equation: Since b is Lipschitz, we can find B such that

$$|X_t^{(n+1)}(\omega) - X_t^{(n)}(\omega)| \le B \int_0^t |X_s^{(n)}(\omega) - X_s^{(n-1)}(\omega)| \, ds.$$

Then, we proceed by induction: Since $|X_0| \leq \omega$, we can show that

$$|X_t^{(n)}| \leq \frac{C\omega B^n t^n}{n!}$$

for some $C > 0$. Thus, $X_t(\omega)$ converges uniformly for all $t \in [0, T]$ for all $T > 0$.

Then, by the dominated convergence theorem, we have integrability with respect to t, as well as measurability with respect to ω. The dominated convergence theorem also shows that the limit function satisfies the integral equation. $\qquad\square$

Now, we are ready to handle the general case

Theorem 16. The stochastic differential equation (1) has a unique solution where σ and b are uniformly Lipschitz in x.

Proof. Unlike in the previous theorem, we cannot do this by estimating the solutions directly. Instead, we apply our estimates to 2nd moments. For brevity, we assume that both functions are Lipschitz with the same Lipschitz constant; taking the maximum of the constants yields this result.

$$\mathbb{E}|X_{n+1}(t) - X_n(t)|$$
$$\leq \mathbb{E}\left(\int_0^t (b(s, X_s^{(n)}) - b(s, X_s^{(n)})) + \int_0^t \sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)}) \, ds\right)^2$$
$$\leq 2\mathbb{E}\left(\int_0^t (b(s, X_s^{(n)}) - b(s, X_s^{(n-1)})) \, ds\right)^2 + 2\mathbb{E}\left(\int_0^t (\sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)}))\right)^2$$
$$\leq 2B^2 \mathbb{E}\left(\int_0^t |X_s^{(n)} - X_s^{(n-1)}| \, ds\right)^2 + 2B^2 \int_0^t \mathbb{E}|X_s^{(n)} - X_s^{(n-1)}|^2 \, ds$$
$$\leq 2B^2 \mathbb{E}\left(t \int_0^t |X_s^{(n)} - X_s^{(n-1)}|^2 \, ds\right) + 2B^2 \int_0^t \mathbb{E}|X_s^{(n)} - X_s^{(n-1)}|^2 \, ds$$
$$\leq 2B^2(T+1) \int_0^t \mathbb{E}|X_s^{(n)} - X_s^{(n-1)}|^2 \, ds$$

for all $t \in [0, T]$, where we have used the inequality $(x + y)^2 \leq 2x^2 + 2y^2$ in the first line, the Ito isometry in the second, and the Cauchy Schwarz inequality in the third.

Now, we are basically done. Since $X_t^{(n)} \to X_t$ uniformly, Lipschitz continuity implies that $\mu(X_t^{(n)}) \to \mu(X_t)$ and $b(X_t^{(n)}) \to b(X_t)$ almost surely (since their squared expectations vanish).

By the Ito isometry, we then have

$$\lim_{n\to\infty} \int_0^t \sigma(s, X_s^{(n)}) \, dB(s) = \int_0^t \sigma(s, X(s)) \, dB(s)$$

and

$$\lim_{n\to\infty} \int_0^t b(s, X_s^{(n)}) \, ds = \int_0^t b(s, X_s) \, ds \,.$$

These imply that the limit function satisfies the stochastic integral equation. □

I. THE MARKOV PROPERTY

It is a known property of initial value problems that the future solution is not dependent on the past. In particular, we can imagine that we have some initial state $x(0)$ and let it evolve a time t to $x(t)$. Then we can let it evolve further. Alternatively, we can consider an initial value problem that has the value $x(t)$ at time t. We expect that these two solutions are identical. In a stochastic differential equation, this "'memory" property is known as the Markov property. Before we define it, let us recall a few definitions:
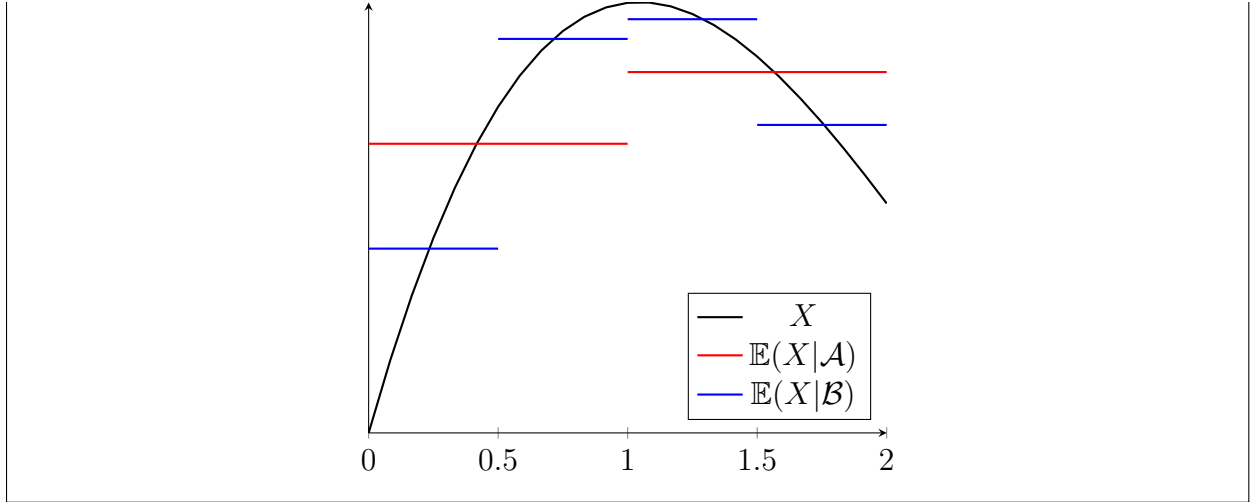

Definition 17 (Conditional Expectation). Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and X is a real-valued random variable on that space with finite expectation. Let $\mathcal{H} \subseteq \mathcal{F}$ be a sub σ-algebra of $\mathcal{F}$. Then the conditional expectation of X relative to $\mathcal{H}$ is a $\mathcal{H}$-$\mathcal{B}$ measurable function $\mathbb{E}(X|\mathcal{H})$ such that

$$\int_H \mathbb{E}(X|\mathcal{H}) \, d\mathbb{P} = \int_H X \, d\mathbb{P}$$

for all $H \in \mathcal{H}$.

Examples are as follows

Example 18.

Where $Y_1, \ldots, Y_n$ is a collection of random variables, we use $\mathbb{E}(X|Y_1, \ldots, Y_n)$ to denote $\mathbb{E}(X|\sigma(Y_1, \ldots, Y_n))$, the conditional expectation on the σ-algebra generated by the Y_is.

Using this, we can define the conditional probability of an event,

---

Definition 19. The conditional probability of an event A, conditional on a σ-algebra $\mathcal{A}$, $\mathbb{P}(A|\mathcal{A})$ is given by $\mathbb{E}(1_A|\mathcal{A})$.

Now, we may define the Markov property

Definition 20. A stochastic process X_t, with $a \leq t \leq b$, is said to have the *Markov property* if for all sequences $a < t_1 < \cdots < t_n < t < b$ and corresponding $x_1, \ldots, x_n$, we have
$$\mathbb{P}(X_t \leq x|X_{t_1} = x_1, \ldots, X_{t_n} = x_n) = \mathbb{P}(X_t \leq x|X_{t_n} = x_n).$$

As an example, all processes with independent increments have the Markov property. Before we move on, let us prove a useful lemma

---

Lemma 21. Suppose a stochastic process X_t defined on $[a, b]$ is adapted to a filtration $\mathcal{F}_t$. Suppose also that
$$\mathbb{P}(X_t \leq x|\mathcal{F}_s) = \mathbb{P}(X_t \leq x|X_s).$$
Then X_t is a Markov process.

Finally, we move to characterise a Markov process by its marginal distributions. Recall that a discrete-time Markov process is characterised by its transition matrix — that is, the

probability to transition to a certain state given its current state.

We define a Markov process on $[a, b]$ and let $a = t_1 < t_2$. We have, by Bayes's Theorem

$$\mathbb{P}(X_{t_1} \leq c_1 \cap X_{t_2} \leq c_2) = \mathbb{P}(X_{t_2} \leq c_2 | X_{t_1} \leq c_1)\mathbb{P}(X_{t_1} \leq c_1)$$

[1] J. Shi, T. Chen, R. Yuan, B. Yuan, and P. Ao, "Relation of a new interpretation of stochastic differential equations to ito process," *Journal of Statistical Physics*, vol. 148, p. 579–590, July 2012.

[2] T. Sauer, *Numerical Solution of Stochastic Differential Equations in Finance*, pp. 529–550. 07 2012.

[3] W. Rudin, *Principles of Mathematical Analysis*. International series in pure and applied mathematics, McGraw-Hill, 1976.