Chapter 5 Notes

5.1 Introduction

Consider a d-dimensional Markov family $X = \{X_t, \mathcal{F}_t; 0 \le t < \infty\}$, $(\Omega, \mathcal{F})$, $\{P^x\}_{x \in \mathbb{R}^d}$, and assume that X has continuous paths. We suppose that the the relation

$$\lim_{t \downarrow 0} \frac{1}{t}[E^x f(X_t) - f(x)] = (\mathcal{A}f)(x); \quad \forall x \in \mathbb{R}^d \tag{1.1}$$

holds for every f in a suitable subclass of the space $C^2(\mathbb{R}^d)$ of real-valued twice-continuously differentiable functions on $\mathbb{R}^d$; the operator $\mathcal{A}f$ is given by

$$(\mathcal{A}f)(x) \triangleq \frac{1}{2} \sum_{i=1}^{d} \sum_{k=1}^{d} a_{ik}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_k} + \sum_{i=1}^{d} b_i(x) \frac{\partial f(x)}{\partial x_i} \tag{1.2}$$

for suitable Borel-measurable functions $b_i, a_{ik} : \mathbb{R}^d \to \mathbb{R}, 1 \le i, k \le d$. The left-hand side of (1.1) is the infintesimal generator of the Markov family, applied to the test function f. On the other hand, the operator in (1.2) is called the second-order diffusion operator associated with the drift vector $b = (b_1, ..., b_d)$ and the diffusion matrix $a = \{a_{ik}\}_{1 \le i,k \le d}$ which is assumed to be symmetric and nonnegative-definite for every $x \in \mathbb{R}^d$.
The drift and diffusion coefficients can be interpreted in the following manner: fix $x \in \mathbb{R}^d$ and let $f_i(y) \triangleq y_i,\ f_{ik}(y) \triangleq (y_i - x_i)(y_k - x_k);\ y \in \mathbb{R}^d$. Assuming that (1.1) holds for these test functions, we obtain

$$E^x[X_t^{(i)} - x_i] = t b_i(x) + o(t) \tag{1.3}$$

$$E^x[(X_t^{(i)} - x_i)(X_t^{(k)} - x_k)] = t a_{ik}(x) + o(t) \tag{1.4}$$

as $t \downarrow 0$, for $1 \le i, k \le d$.
1.1 Definition. Let $X = \{X_t, \mathcal{F}_t; 0 \le t < \infty\}$, $(\Omega, \mathcal{F})$, $\{P^x\}_{x \in \mathbb{R}^d}$ be a d-dimensional Markov family, such that
(i) X has a continuous sample paths;
(ii) relation (1.1) holds for every $f \in C^2(\mathbb{R}^d)$ which is bounded and has bounded first- and second- order derivatives;
(iii) relations (1.3), (1.4) holds for every $x \in \mathbb{R}^d$; and
(iv) the tenets (a)-(d) of Definition 2.6.3 are satisfied, but only for stopping times S.
Then X is called a (Kolmogorov-Feller) diffusion process.

5.2 Strong Solutions

In this section, we introduce the concept of a stochastic differential equation w.r.t. Brownian motion and its solution in the strong sense. We discuss the questions of existence and uniqueness of such solutions, as well as some of their elementary properties.

Let us start with Borel-measurable functions $b_i(t,x),\ \sigma_{ij}(t,x);\ 1 \le i \le d,\ 1 \le j \le r$, from $[0,\infty) \times \mathbb{R}^d$ into $\mathbb{R}$, and define the $(d \times 1)$ drift vector $b(t,x) = \{b_i(t,x)\}_{1 \le i \le d}$ and the $(d \times r)$ dispersion matrix $\sigma(t,x) = \{\sigma_{ij}(t,x)\}_{1 \le i \le d; 1 \le j \le r}$. The intent is to assign a meaning to the stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{2.1}$$

written componentwise as

$$dX_t^{(i)} = b_i(t, X_t)dt + \sum_{j=1}^{r} \sigma_{ij}(t, X_t)dW_t^{(j)}; \quad 1 \le i \le d, \tag{2.1$'$}$$

where $W = \{W_t; 0 \le t < \infty\}$ is an r-dimensional Brownian motion and $X = \{X_t; 0 \le t < \infty\}$ is a suitable stochastic process with continuous sample paths and values in $\mathbb{R}^d$, the "solution" of the equation. The drift vector $b(t,x)$ and the dispersion matrix $\sigma(t,x)$ are the coefficients of this equation; the $(d \times d)$ matrix $a(t,x) \triangleq \sigma(t,x)\sigma^T(t,x)$ with elements

$$a_{ik}(t,x) \triangleq \sum_{j=1}^{r} \sigma_{ij}(t,x)\sigma_{kj}(t,x); \quad 1 \le i,k \le d \tag{2.2}$$

will be called the diffusion matrix.

A. Definitions

Choose a probability space $\{\Omega, \mathcal{F}, P\}$ and an r-dimensional Brownian motion $W = \{W_t, \mathcal{F}_t^W; 0 \le t < \infty\}$ on it. We assume also that this space is rich enough to accommodate a random vector ξ taking values in $\mathbb{R}^d$, independent of $\mathcal{F}_\infty^W$ and with given distribution

$$\mu(\Gamma) = P[\xi \in \Gamma]; \quad \Gamma \in \mathcal{B}(\mathbb{R}^d).$$

We consider the left-continuous filtration

$$\mathcal{G}_t \triangleq \sigma(\xi) \vee \mathcal{F}_t^W = \sigma(\xi, W_s; 0 \le s \le t); \quad 0 \le t < \infty,$$

as well as the collection of null sets

$$\mathcal{N} \triangleq \{N \subseteq \Omega;\ \exists G \in \mathcal{G}_\infty \text{ with } N \subseteq G \text{ and } P(G) = 0\},$$

and create the augmented filtration

$$\mathcal{F}_t \triangleq \sigma(\mathcal{G}_t \cup \mathcal{N}), \quad 0 \le t < \infty; \quad \mathcal{F}_\infty \triangleq \sigma\left(\bigcup_{t \ge 0} \mathcal{F}_t\right). \tag{2.3}$$

Then, $\{W_t, \mathcal{G}_t; 0 \leq t < \infty\}$ is an r-dimensional Brownian motion, and then so is $\{W_t, \mathcal{F}_t; 0 \leq t < \infty\}$ by Theorem 2.7.9. As in the proof of Proposition 2.7.7, the filtration $\{\mathcal{F}_t\}$ satisfies the usual conditions.

2.1 Definition. A **strong solution** of the stochastic differential equation (2.1), on the given probability space $\{\Omega, \mathcal{F}, P\}$ and w.r.t. fixed Brownian motion W and initial condition ξ, is a process $X = \{X_t; 0 \leq t < \infty\}$ with continuous sample paths and with the following properties:
(i) X is adapted to the filtration $\{\mathcal{F}_t\}$ of (2.3),
(ii) $P[X_0 = \xi] = 1$,
(iii) $P[\int_0^t \{|b_i(s, X_s)| + \sigma_{ij}^2(s, X_s)d\}s < \infty] = 1$ holds for every $1 \leq i \leq d$, $1 \leq j \leq r$ and $\leq t < \infty$, and
(iv) the integral version of (2.1)

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s; \quad 0 \leq t < \infty, \tag{2.4}$$

or equivalently,

$$X_t^{(i)} = X_0^{(i)} + \int_0^t b_i(s, X_s)ds + \sum_{j=1}^{r} \int_0^t \sigma_{ij}(s, X_s)dW_j^{(j)}; \quad 0 \leq t < \infty, \ 1 \leq i \leq d, \tag{2.4$'$}$$

holds almost surely.

2.2 Remark. The crucial requirement of this definition is captured in the condition (i); it corresponds to our intuitive understanding of:

W, ξ – "input" of a dynamical system,
X – "output" of a dynamical system,
(b, σ) – system parameters.

The **principle of causality** for dynamical systems requires that the output X, at time t depend only on ξ and the values of the input $\{W_s, 0 \leq s \leq t\}$ up to that time. This principle finds its mathematical expression in (i).

2.3 Definition. Let the drift vector $b(t, x)$ and dispersion matrix $\sigma(t, x)$ be given. Suppose that, whenever W is an r-dimensional Brownian motion on some $(\Omega, \mathcal{F}, P)$, ξ is an independent, d-dimensional r.v., $\{\mathcal{F}_t\}$ is given by (2.3), and X, $\tilde{X}$ are two strong solutions of (2.1) relative to W with initial condition ξ, then $P[X_t = \tilde{X}_t; 0 \leq t < \infty] = 1$. Under these conditions, we say that **strong uniqueness holds for the pair** (b, σ).

2.4 Example. Consider the one-dimensional equation

$$dX_t = b(t, X_t)dt + dW_t,$$

where $b : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ is bounded, Borel-measurable, and nonincreasing in the space variable; i.e., $b(t,x) \le b(t,y)$ for all $0 \le t < \infty$ for all $0 \le t < \infty$, $-\infty < y \le x < \infty$. For this equation, strong uniqueness holds.
$\because$ for any two processes $X^{(1)}, X^{(2)}$ satisfying P-a.s.

$$X_t^{(i)} = X_0 + \int_0^t b(s, X_s^{(i)})ds + W_t; \quad 0 \le t < \infty \text{ and } i = 1, 2,$$

we may define the continuous process $\Delta_t = X_t^{(1)} - X_t^{(2)}$ and observe that

$$\Delta_t^2 = \left(\int_0^t b(s, X_s^{(1)}) - b(s, X_s^{(2)})ds \right)^2 =$$

$$= 2\int_0^t (X_s^{(1)} - X_s^{(2)})[b(s, X_s^{(1)}) - b(s, X_s^{(2)})]ds \le 0; \quad 0 \le t < \infty, \text{ a.s. } P.$$

B. The Ito Theory

If the dispersion matrix $\sigma(t,x)$ is identically zero, (2.4) reduces to the ordinary (nonstochastic, except possibly in the initial condition) integral equation

$$X_t = X_0 + \int_0^t b(s, X_s)ds. \tag{2.5}$$

In the theory of such equations, it is customary to impose the assumption that the vector field $b(t,x)$ satisfies a local Lipschitz condition in the space variable x and is bounded on compact subsets of $[0,\infty) \times \mathbb{R}^d$. These conditions ensure that for sufficiently small $t > 0$, the Picard-Lindelöf iterations

$$X_t^{(0)} \equiv X_0; \quad X_t^{(n+1)} = X_0 + \int_0^t b(s, X_s^{(n)})ds, \quad n \ge 0, \tag{2.6}$$

converges to a solution of (2.5) and that this solution is unique. In the absence of such conditions the equation miight fail to be solvable or might have a continuum of solutions. For instance, the one-dimensional equation

$$X_t = \int_0^t |X_s|^\alpha ds \tag{2.7}$$

has only one solution for $\alpha \ge 1$, namely, $X_t \equiv 0$; however, for $0 < \alpha < 1$ all functions of the form

$$X_t = \begin{cases} 0; & 0 \le t \le s, \\ \left(\frac{t-s}{\beta}\right)^\beta; & s \le t < \infty \end{cases}$$

with $\beta = \frac{1}{1-\alpha}$ and arbitrary $0 \le s \le \infty$, solve (2.7).

2.5 Theorem. Suppose that the coefficients $b(t,x)$ $\sigma(t,x)$ are locally Lipschitz-continuous in the space variable, i.e., for every integer $n \geq 1$ there exists a constant $K_n > 0$ s.t. for every $t \geq 0$, $\|x\| \leq n$ and $\|y\| \leq n$:

$$\|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\| \leq K_n \|x - y\|. \tag{2.8}$$

Then strong uniqueness holds for equation (2.1).

2.6 Remark. For every $(d \times r)$ matrix σ, we write

$$\|\sigma\|^2 \triangleq \sum_{i=1}^{d} \sum_{j=1}^{r} \sigma_{ij}^2. \tag{2.9}$$

Before proceeding with the proof, let us recall the useful Gronwall inequality.

2.7 Problem. Suppose that the continuous function $g(t)$ satisfies

$$0 \leq g(t) \leq \alpha(t) + \beta \int_0^t g(s)ds; \quad 0 \leq t \leq T, \tag{2.10}$$

with $\beta \geq 0$ and $\alpha : [0,T] \to \mathbb{R}$ integrable. Then,

$$g(t) \leq \alpha(t) + \beta \int_0^t \alpha(s) e^{\beta(t-s)} ds; \quad 0 \leq t \leq T. \tag{2.11}$$

Proof of Theorem 2.5. Let us suppose that X and $\tilde{X}$ are both strong solutions, defined for all $t \geq 0$, of (2.1) relative to the same Brownian motion W and the same initial condition ξ, on some $(\Omega, \mathcal{F}, P)$. We define the stopping times $\tau_n = \inf\{t \geq 0; \|X_t\| \geq n\}$ for $n \geq 1$, as well as their tilded counterparts, and we set $S_n \triangleq \tau_n \wedge \tilde{\tau}_n$. Clearly $\lim_{n\to\infty} S_n = \infty$ a.s. P, and

$$X_{t\wedge S_n} - \tilde{X}_{t\wedge S_n} = \int_0^{t\wedge S_n} \{b(u, X_u) - b(u, \tilde{X}_u)\} du =$$

$$+ \int_0^{t\wedge S_n} \{\sigma(u, X_u) - \sigma(u, \tilde{X}_u)\} dW_u.$$

Using the vector inequality $\|v_1 + ... + v_k\|^2 \leq k^2(\|v_1\|^2 + ... + \|v_k\|^2)$, the Holder inequality for Lebesgue integrals, the basic property (3.2.27) of stochastic integrals, and (2.8), we may write for $0 \leq t \leq T$:

$$E\|X_{t\wedge S_n} - \tilde{X}_{t\wedge S_n}\|^2 \leq 4E\| \int_0^{t\wedge S_n} \{b(u, X_u) - b(u, \tilde{X}_u)\} du\|^2 +$$

$$4E\| \int_0^{t\wedge S_n} \{\sigma(u, X_u) - \sigma(u, \tilde{X}_u)\} dW_u\|^2 \leq 4E\left[\int_0^{t\wedge S_n} \|b(u, X_u) - b(u, \tilde{X}_u)\| du \right]^2$$

$$+4E\sum_{i=1}^{d}\left[\sum_{j=1}^{r}\int_0^{t\wedge S_n}(\sigma_{ij}(u,X_u)-\sigma_{ij}(u,\tilde{X}_u))dW_u^{(j)}\right]^2\leq$$

$$\leq 4tE\int_0^{t\wedge S_n}\|b(u,X_u)-b(u,\tilde{X}_u)\|^2du+4E\int_0^{t\wedge S_n}\|\sigma(u,X_u)-\sigma(u,\tilde{X}_u)\|^2\leq$$

$$\leq 4(T+1)K_n^2\int_0^t E\|X_{u\wedge S_n}-\tilde{X}_{u\wedge S_n}\|^2du.$$

We now apply Problem 2.7 with $g(t)\triangleq E\|X_{t\wedge S_n}-\tilde{X}_{t\wedge S_n}\|^2$, $\alpha(t)=0$, $\beta=4(T+1)K_n^2$ to conclude that $g(t)\leq\alpha(t)+\beta\int_0^t\alpha(s)e^{-\beta(t-s)}ds\Rightarrow$ $g(t)=E\|X_{s\wedge S_n}-\tilde{X}_{s\wedge S_n}\|^2=0$, so $\{X_{t\wedge S_n};0\leq t<\infty\}$ and $\{\tilde{X}_{t\wedge S_n};0\leq t<\infty\}$ are modifications of one another, and thus are indistinguishable. Letting $n\to\infty$, we see that the same is true for $\{X_t;0\leq t<\infty\}$ and $\{\tilde{X}_t;0\leq t<\infty\}$. $\square$

2.8 Remark It is worth noting that even for ordinary differential equations, a local Lipschitz condition is not sufficient to guarantee global existence of a solution. For example, the unique (by Theorem 2.5) solution to the equation

$$X_t=1+\int_0^t X_s^2ds$$

is $X_t=\frac{1}{1-t}$, which "explodes" at $t\uparrow 1$. We thus impose stronger conditions in order to obtain an existence result.

2.9 Theorem. Suppose that the coefficients $b(t,x)$, $\sigma(t,x)$ satisfy the global Lipschitz and linear growth conditions

$$\|b(t,x)-b(t,y)\|+\|\sigma(t,x)-\sigma(t,y)\|\leq K\|x-y\|, \tag{2.12}$$

$$\|b(t,x)\|^2+\|\sigma(t,x)\|^2\leq K^2(1+\|x\|^2), \tag{2.13}$$

for every $0\leq t<\infty$, $x\in\mathbb{R}^d$, $y\in\mathbb{R}^d$, where K is a positive constant. On some probability space $(\Omega,\mathcal{F},P)$, let ξ be an $\mathbb{R}^d$-valued r.v., independent of the r-dimensional Brownian motion $W=\{W_t,\mathcal{F}_t^W;0\leq t<\infty\}$, and with finite second moment:

$$E\|\xi\|^2<\infty. \tag{2.14}$$

Let $\{\mathcal{F}_t\}$ be as in (2.3). Then there exists a continuous, adapted process $X=\{X_t,\mathcal{F}_t;0\leq t<\infty\}$ which is a strong solution of equation (2.1) relative to W, with initial condition ξ. Moreover,this process is square-integrable: for every $T>0$, there exists a constant C, depending only on K and T, s.t.

$$E\|X_t\|^2\leq C(1+E\|\xi\|^2)e^{Ct};\quad 0\leq t\leq T. \tag{2.15}$$

The idea of the proof is to mimic the deterministic situation and to construct recursively, by analogy with (2.6), a sequence of successive approximations by setting $X_t^{(0)} = \xi$ and

$$X_t^{(k+1)} \triangleq \xi + \int_0^t b(s, X_s^{(k)})ds + \int_0^t \sigma(s, X_s^{(k)})dW_s; \quad 0 \le t < \infty, \tag{2.16}$$

for $k \ge 0$. These processes are continuous and adapted to the filtration $\{\mathcal{F}_t\}$. The hope is that the sequence $\{X^{(k)}\}_{k=1}^{\infty}$ will converge to a solution of equation (2.1).

2.10 Problem. For every $T > 0$, there exists a positive constant C depending only on K and T, s.t. for the iterations in (2.16) we have

$$E\|X_t^{(k)}\|^2 \le C(1 + E\|\xi\|^2)e^{Ct}; \quad 0 \le t \le T, \ k \ge 0. \tag{2.17}$$

Proof of Theorem 2.9) We have $X_t^{(k+1)} - X_t^{(k)} = B_t + M_t$ from (2.16), where

$$B_t \triangleq \int_0^t \{b(s, X_s^{(k)} - b(s, X_s^{(k-1)})\}ds, \quad M_t \triangleq \int_0^t \{\sigma(s, X_s^{(k)}) - \sigma(s, X_s^{(k-1)})\}dW_s.$$

Thanks to the inequalities (2.13) and (2.17), the process $\{M_t = M_t^{(1)}, ..., M_t^{(d)}, \mathcal{F}_t; 0 \le t < \infty\}$ is seen to be a vector of square-integrable martingales, for which Problem 3.3.29 and Remark 3.3.30 give

$$E\left[\max_{0\le s\le t} \|M_s\|^2\right] \le \Lambda_1 E \int_0^t \|\sigma(s, X_s^{(k)}) - \sigma(s, X_s^{(k-1)})\|^2 ds \le$$

$$\le \Lambda_1 K^2 E \int_0^t \|X_s^{(k)} - X_s^{(k-1)}\|^2 ds.$$

On the other hand, we have by Holder's inequality for Lebesgue integrals, $E\|B_t\|^2 \le \int_0^t 1^2 ds E \int_0^t \|b(s, X_s^{(k)}) - b(s, X_s^{(k-1)})\|^2 \le K^2 t \int_0^t E\|X_s^{(k)} - X_s^{(k-1)}\|^2 ds$, and therefore, with $L = 4K^2(\Lambda_1 + T)$,

$$E\left[\max_{0\le s\le t} \|X_s^{(k+1)} - X_s^{(k)}\|^2\right] = E\left[\max_{0\le s\le t} \|B_t + M_t\|^2\right] \le \tag{2.18}$$

$$\le E\left[4 \max_{0\le s\le t} \|B_t\|^2 + 4 \max_{0\le s\le t} \|M_t\|^2\right] \le L \int_0^t E\|X_s^{(k)} - X_s^{(k-1)}\|^2 ds; \quad 0 \le t \le T.$$

Inequality (2.18) can be iterated to yield the successive upper bounds

$$E\left[\max_{0\le s\le t} \|X_s^{(k+1)} - X_s^{(k)}\|^2\right] \le C^* L^k \int_0^t \int_0^t ... \int_0^t ds_1 ... ds_k \le C^* \frac{(Lt)^k}{k!}; \quad 0 \le t \le T, \tag{2.19}$$

where $C^* = \max_{0\le t\le T} E\|X_t^{(1)} - \xi\|^2$, a finite quantity because of (2.17). Relation (2.19) and the Chebyshev inequality now gives

$$P\left[\max_{0\le t\le T} \|X_t^{(k+1)} - X_t^{(k)}\| > \frac{1}{2^{k+1}}\right] \le 4C^* \frac{(4LT)^k}{k!}; \quad k = 1, 2, ..., \tag{2.20}$$

and this upper bound is the general term in a convergent series. From the Borel-Cantelli lemma, we conclude that there exists an event $\Omega^* \in \mathcal{F}$ with $P(\Omega^*)$ and an integer-valued random variable $N(\omega)$ s.t. for every $\omega \in \Omega^*$: $\max_{0\le t\le T} \|X_t^{(k+1)}(\omega) - X_t^{(k)}(\omega)\| \le 2^{-(k+1)},\ \forall k \ge N(\omega)$. Consequently,

$$\max_{0\le t\le T} \|X_t^{(k+m)}(\omega) - X_t^{(k)}(\omega)\| \le 2^{-k}, \quad \forall m \ge 1,\ k \ge N(\omega). \tag{2.21}$$

We see then that the sequence of sample paths $\{X_t^{(k)}(\omega);\ 0 \le t \le T\}_{k=1}^\infty$ is convergent in the supremum norm on continuous functions, from which follows the existence of a continuous limit $\{X_t^{(k)}(\omega);\ 0 \le t \le T\}$ for all $\omega \in \Omega^*$. Since T is arbitrary we have the existence of a continuous process $X_t = \{X_t(\omega); 0 \le t \le T\}$ with the property that for P-a.e. ω, the sample paths $\{X_\cdot^{(k)}(\omega)\}_{k=1}^\infty$ converge to $X_\cdot(\omega)$ uniformly on compact subsets of $[0, \infty)$. Inequality (2.15) is a consequence of (2.17) and Fatou's lemma. From (2.15) and (2.13) we have condition (iii) of Definition 2.1. Conditions (i) and (ii) are clearly satisfied by X. The following problem concludes the proof. $\square$

2.11 Problem. Show that the just constructed process

$$X_t \triangleq \lim_{k\to\infty} X_t^{(k)}; \quad 0 \le t < \infty \tag{2.22}$$

satisfies requirement (iv) of Definition 2.1.

C. Comparison Results and Other Refinements

In the one-dimensional case, the Lipschitz condition on the dispersion coefficient can be relaxed considerably.

2.13 Proposition.(Yamada & Watanabe (1971)) Let us suppose that the coefficients of the one-dimensional equation ($d = r = 1$)

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{2.1}$$

satisfy the conditions

$$|b(t, x) - b(t, y)| \le K|x - y| \tag{2.23}$$

$$|\sigma(t, x) - \sigma(t, y)| \le h(|x - y|), \tag{2.24}$$

for every $0 \le t < \infty$ and $x \in \mathbb{R},\ y \in \mathbb{R}$, where K is a positive constant and $h : [0, \infty) \to [0, \infty)$ is a strictly increasing function with $h(0) = 0$ and

$$\int_{(0,\varepsilon)} \frac{1}{h^2(u)} du = \infty; \quad \forall \varepsilon > 0. \tag{2.25}$$

Then strong uniqueness holds for the equation (2.1).

2.14 Example In the above proposition, we can take the function h to be $h(u) = u^\alpha; \quad \alpha \geq \frac{1}{2}$.

Proof of Proposition 2.13) By (2.25) and the properties of the function h, there exists a strictly decreasing sequence $\{a_n\}_{n=0}^\infty \subseteq (0,1]$ with $a_0 = 1$, $\lim_{n\to\infty} a_n = 0$ and $\int_{a_n}^{a_{n-1}} \frac{1}{h^2(u)} du = n$, for every $n \geq 1$. (For a_n chosen, choose a_{n+1} large enough so that this integral is equal to $n+1$. If we cannot choose such a_{n+1}, then (2.25) becomes finite.) For each $n \geq 1$, there exists a continuous function ρ_n on $\mathbb{R}$ with support in (a_n, a_{n-1}) so that $0 \leq \rho_n(x) \leq \frac{2}{nh^2(x)}$ holds for every $x > 0$, and $\int_{a_n}^{a_{n-1}} \rho_n(x)dx = 1$. Then the function

$$\psi_n(x) \triangleq \int_0^{|x|} \int_0^y \rho_n(u) du dy; \quad x \in \mathbb{R} \tag{2.26}$$

is even and twice continuously differentiable, with $|\psi_n^{'}(x)| \leq 1$ and $\lim_{n\to\infty} \psi_n(x) = \int_0^{|x|} 1 dy = |x|$ for $x \in \mathbb{R}$. Furthermore, the seq. $\{\psi_n\}_{n=1}^\infty$ is nondecreasing.
Now suppose that there are two strong solutions $X^{(1)}$ and $X^{(2)}$ of (2.1) with $X_0^{(1)} = X_0^{(2)}$ a.s. It suffices to prove the indistinguishability of $X^{(1)}$ and $X^{(2)}$ under the assumption

$$E\int_0^t |\sigma(s, X_s^{(i)})|^2 ds < \infty; \quad 0 \leq t < \infty, i = 1,2; \tag{2.27}$$

otherwise, we may use condition (iii) of Definition 2.1 and a localization argument to reduce the situation to one in which (2.27) holds. We have

$$\Delta_t \triangleq X_t^{(1)} - X_t^{(2)} = \int_0^t \{b(s, X_s^{(1)}) - b(s, X_s^{(2)})\} ds + \int_0^t \{\sigma(s, X_s^{(1)}) - \sigma(s, X_s^{(2)})\} dW_s;$$

and by the Ito rule,

$$\begin{aligned} \psi_n(\Delta_t) = & \int_0^t \psi_n^{'}(\Delta_s)[b(s, X_s^{(1)}) - b(s, X_s^{(2)})] ds + \\ & + \frac{1}{2}\int_0^t \psi_n^{''}(\Delta_s)[\sigma(s, X_s^{(1)}) - \sigma(s, X_s^{(2)})]^2 ds + \\ & + \int_0^t \psi_n^{'}(\Delta_s)[\sigma(s, X_s^{(1)}) - \sigma(s, X_s^{(2)})] dW_s. \end{aligned} \tag{2.28}$$

The expectation of the stochastic integral in (2.28) is zero because of assumption (2.27), whereas the expectation of the second integral in (2.28) is bounded above by $E\int_0^t \psi_n^{''}(\Delta_s) h^2(|\Delta_s|) ds \leq \frac{2t}{n}$. We conclude that

$$E\psi_n(\Delta_t) \leq E\int_0^t \psi_n^{'}(\Delta_s)[b(s, X_s^{(1)}) - b(s, X_s^{(2)})] + \frac{t}{n} \leq \tag{2.29}$$

$$\leq K \int_0^t E|\Delta_s|ds + \frac{t}{n}; \quad t \geq 0, n \geq 1.$$

A passage to the limit as $n \to \infty$ yields $E|\Delta_t| \leq K \int_0^t E|\Delta_s|ds;\ t \geq 0$ and by the Gronwall inequality, we have $E|\Delta_t| = E|X_t^{(1)} - X_t^{(2)}| = 0$. Furthermore, the sample path continuity gives us the indistinguishability. □

2.15 Example (Girsanov (1962)). From what we have just proved, it follows that strong uniqueness holds for the one-dimensional stochastic equation

$$X_t = \int_0^t |X_s|^\alpha dW_s; \quad 0 \leq t < \infty, \tag{2.30}$$

with $b(t, X_t) = 0,\ \sigma(t, X_t) = |X_t|^\alpha$, as long as $\alpha \geq \frac{1}{2}$, and the unique solution is the trivial one $X_t \equiv 0$. This is also a solution when $0 < \alpha, \frac{1}{2}$, but it is no longer the only solution (because the condition of Proposition 2.13 fails). We shall see in Remark 5.6 that not only does strong uniqueness fails when $0 < \alpha < \frac{1}{2}$, but we do not even have uniqueness in the weaker sense.

The methodology employed in the proof of Proposition 2.13 can be used to great advantage in establishing the **comparison results** for solutions of one-dimensional stochastic differential equations. Such results amount to a certain kind of "monotonicity" of the solution process X w.r.t. the drift coefficients $b(t, x)$, and they are useful in variety of situations, including the study of certain simple stochastic control problems. We develop some comparison results in the following proposition and problem.

2.18 Proposition. Suppose that on a certain probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\{\mathcal{F}_t\}$ which satisfies the usual conditions, we have a standard, one-dimensional Brownian motion $\{W_t, \mathcal{F}_t, 0 \leq t < \infty\}$ and two continuous adapted processes $X^{(j)}; j = 1, 2$ s.t.

$$X_t^{(j)} = X_0^{(j)} + \int_0^t b_j(s, X_s^{(j)})ds + \int_0^t \sigma(s, X_s^{(j)})dW_s; \quad 0 \leq t < \infty \tag{2.31}$$

holds a.s. for $j = 1, 2$. We assume that
(i) the coefficients $\sigma(t, x)$, $b_j(t, x)$ are continuous, real-valued functions on $[0, \infty) \times \mathbb{R}$.
(ii) the dispersion matrix $\sigma(t, x)$ satisfies condition (2.24), where h is as described in Proposition 2.13,
(iii) $X_0^{(1)} \leq X_0^{(2)}$ a.s.,
(iv) $b_1(t, x) \leq b_2(t, x), \forall 0 \leq t < \infty,\ x \in \mathbb{R}$, and
(v) either $b_1(t, x)$ or $b_2(t, x)$ satisfies the Lipschitz continuity condition (2.23).
Then

$$P[X_t^{(1)} \leq X_t^{(2)}, \forall 0 \leq t < \infty] = 1. \tag{2.32}$$

Proof) For concreteness, let us suppose that $|b(t, x) - b(t, y)| \leq K|x - y|$ is satisfied by $b_1(t, x)$. Proceeding as in the proof of Proposition 2.13, we assume

WLOG that (2.27) holds. We recall that the functions $\psi_n(x)$ of (2.26) and create a new seq. of auxiliary functions by setting $\phi_n(x) = \psi_n(x) \cdot 1_{(0,\infty)}(x)$; $x \in \mathbb{R}$, $n \geq 1$. With $\Delta_t = X_t^{(1)} - X_t^{(2)}$, the analogue of relation (2.29) is

$$E\phi_n(\Delta_t) - \frac{t}{n} \leq E \int_0^t \phi_n^{'}(\Delta_s)[b_1(s, X_s^{(1)}) - b_2(s, X_s^{(2)})]ds =$$

$$= E \int_0^t \phi_n^{'}(\Delta_s)[b_1(s, X_s^{(1)}) - b_1(s, X_s^{(2)})]ds +$$

$$= E \int_0^t \phi_n^{'}(\Delta_s)[b_1(s, X_s^{(2)}) - b_2(s, X_s^{(2)})]ds \leq K \int_0^t E(\Delta_s^+)ds,$$

by (iv) and (2.23). Now we can let $n \to \infty$ to obtain $E(\Delta_t^+) \leq K \int_0^t E(\Delta_s^+)ds$; $0 \leq t < \infty$, and by the Gronwall inequality, we have $E(\Delta_t^+) = 0$, i.e., $X_t^{(1)} \leq X_t^{(2)}$ a.s. P. $\square$

D. Approximations of Stochastic Differential Equations