3.21′20

Chapter 5 Notes

5.1 Introduction

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

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

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 $\mathscr{A} f$ is given by

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

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

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

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

as $t \downarrow 0$, for $1 \le i, k \le d$.
1.1 Definition. Let $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$, $(\Omega, \mathscr{F})$, $\{P^x\}_{x\in\mathbb{R}^d}$ 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 $\{\Omega, \mathscr{F}, P\}$ and an r-dimensional Brownian motion $W = \{W_t, \mathscr{F}_t^W; 0 \le t < \infty\}$ 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 $\mathscr{F}_\infty^W$ and with given distribution

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

We consider the left-continuous filtration

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

as well as the collection of null sets

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

and create the augmented filtration

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

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

2.1 Definition. A **strong solution** of the stochastic differential equation (2.1), on the given probability space $\{\Omega, \mathscr{F}, P\}$ and w.r.t. fixed Brownian motion W and initial condition ξ, is a process $X = \{X_t; 0 \le t < \infty\}$ with continuous sample paths and with the following properties:
(i) X is adapted to the filtration $\{\mathscr{F}_t\}$ 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 $1 \le i \le d,\ 1 \le j \le r$ and $\le t < \infty$, and
(iv) the integral version of (2.1)

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

or equivalently,

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

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 $\{W_s, 0 \le s \le t\}$ 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 $(\Omega, \mathscr{F}, P)$, ξ is an independent, d-dimensional r.v., $\{\mathscr{F}_t\}$ is given by (2.3), and X, $\tilde{X}$ are two strong solutions of (2.1) relative to W with initial condition ξ, then $P[X_t = \tilde{X}_t; 0 \le t < \infty] = 1$. 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., $b(t, x) \leq b(t, y)$ for all $0 \leq t < \infty$ for all $0 \leq t < \infty$, $-\infty < y \leq x < \infty$. For this equation, strong uniqueness holds.
$\because$ for any two processes $X^{(1)}, X^{(2)}$ satisfying P-a.s.

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

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

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

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

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

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

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 $\alpha \geq 1$, namely, $X_t \equiv 0$; however, for $0 < \alpha < 1$ all functions of the form

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

with $\beta = \frac{1}{1-\alpha}$ and arbitrary $0 \leq s \leq \infty$, 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 $(\Omega, \mathscr{F}, P)$. 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 $(\Omega,\mathscr{F},P)$, let ξ be an $\mathbb{R}^d$-valued r.v., independent of the r-dimensional Brownian motion $W=\{W_t,\mathscr{F}_t^W; 0\leq t<\infty\}$, and with finite second moment:

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

Let $\{\mathscr{F}_t\}$ be as in (2.3). Then there exists a continuous, adapted process $X = \{X_t,\mathscr{F}_t; 0\leq t<\infty\}$ 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 $\{\mathscr{F}_t\}$. 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 $\{M_t = M_t^{(1)}, ..., M_t^{(d)}, \mathscr{F}_t; 0 \le t < \infty\}$ 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 $\Omega^* \in \mathscr{F}$ 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. $\square$

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 $(\Omega, \mathscr{F}, P)$ equipped with a filtration $\{\mathscr{F}_t\}$ which satisfies the usual conditions, we have a standard, one-dimensional Brownian motion $\{W_t, \mathscr{F}_t, 0 \leq t < \infty\}$ 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. $\square$

5.3. Weak Solutions

3.1. Definition. A **weak solution** of equation $dX_t = b(t, X_t)dt + \sigma(X_t, t)dW_t$ is a triple $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$, where

- (i) $(\Omega, \mathscr{F}, P)$ is a probability space, and $\mathscr{F}_t$ is a filtration of sub-σ-fields of $\mathscr{F}$ satisfying the usual conditions
- (ii) $X = \{X_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is a continuous, adapted $\mathbb{R}^d$-valued process, $W = \{W_t, \mathscr{F}_t, 0 \leq t < \infty\}$ is an r-dimensional Brownian motion, and $(iii), (iv)$ of Definition 2.1 are satisfied.

The probability measure $\mu(\Gamma) \triangleq P[X_0 \in \Gamma]$, $\Gamma \in \mathscr{B}(\mathbb{R}^d)$, is called the initial distribution of the solution.

The filtration $\{\mathscr{F}_t\}$ in Definition 3.1 is not necessarily the augmentation of the filtration $\mathscr{G}_t = \sigma(\xi) \vee \mathscr{F}_t^W$, $0 \leq t < \infty$, generated by the "driving Brownian motion" and by the "initial condition" $\xi = X_0$. Thus, the value of the solution $X_t(\omega)$ at time t is not necessarily given by a measurable functional of the Brownian path $\{W_s(\omega); 0 \leq s \leq t\}$ and the initial condition $\xi(\omega)$. On the other hand, because W is a Brownian motion relative to $\{\mathscr{F}_t\}$, the solution $X_t(\omega)$ at time t cannot anticipate the future of the Brownian motion; besides $\{W_s(\omega); 0 \leq s \leq t\}$ and $\xi(\omega)$, whatever extra information is required to compute $X_t(\omega)$ must be independent of $\{W_\theta(\omega) - W_t(\omega); t \leq \theta < \infty\}$.

One consequence of this arrangement is that the existence of a weak solution $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ does not guarantee, for a given Brownian motion $\{\tilde{W}_t, \tilde{\mathscr{F}}; 0 \leq t < \infty\}$ on a (possibly different) probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$, the existence of a process $\tilde{X}$ s.t. the triple $(\tilde{X}, \tilde{W}), (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}), \{\tilde{\mathscr{F}}_t\}$ is again a weak solution. However, strong solvability implies weak solvability.

A. Two Notion of Uniqueness

There are two reasonable concepts of uniqueness which can be associated with weak solutions.

3.2 Definition. Suppose that whenever $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ and $(\tilde{X}, W)$, $(\Omega, \mathscr{F}, P), \{\tilde{\mathscr{F}}_t\}$, are weak solutions to (2.1) with common Brownian motion W (relative to possibly different filtrations) on a common probability space $(\Omega, \mathscr{F}, P)$ and with common initial value, i.e., $P[X_0 = \tilde{X}_0] = 1$, the two processes X and $\tilde{X}$ are indistinguishable: $P[X_t = \tilde{X}_t;\ \forall 0 \le t < \infty] = 1$. We say then that pathwise uniqueness holds for equation (2.1).

3.3 Remark. All the strong uniqueness results of Section 2 are also valid for pathwise uniqueness; indeed, none of the proofs given there takes advantage of the special form of the filtration for a strong solution.

3.4 Definition. We say that **uniqueness in the sense of probability law** holds for equation (2.1) if, for any two weak solutions (X, W), $(\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$, and $(\tilde{X}, \tilde{W})$, $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}), \{\tilde{\mathscr{F}}_t\}$, with the same initial distribution, i.e.,

$$P[X_0 \in \Gamma] = \tilde{P}[\tilde{X}_0 \in \Gamma]; \quad \forall \Gamma \in \mathscr{B}(\mathbb{R}^d),$$

the two processes X, $\tilde{X}$ have the same law.

Existence of a weak solution does not imply that of a strong solution, and uniqueness in the sense of probability law does not imply pathwise uniqueness. However, pathwise uniqueness does not imply uniqueness in the sense of probability law (Proposition 3.20).

3.5 Example. (H. Tanaka (e.g. Zvonkin (1974))). Consider the one-dimensional equation

$$X_t = \int_0^t sgn(X_s) dW_s; \quad 0 \le t < \infty, \tag{3.1}$$

where

$$sgn(x) = \begin{cases} 1; & x > 0 \\ -1; & x \le 0 \end{cases}$$

If (X, W), $(\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ is a weak solution, then the process $X = \{X_t, \mathscr{F}_t, 0 \le t < \infty\}$ is a continuous, square-integrable martingale with quadratic variation process $\langle X \rangle_t = \int_0^t sgn^2(X_s) sd = t$. Therefore, X is a Brownian motion (Theorem 3.3.16), and uniqueness in the sense of probability law holds. On the other hand, $(-X, W)$, $(\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ is also a weak solution, so once we establish existence of a weak solution, we shall also have shown that pathwise uniqueness cannot hold for equation (3.1).

Now start with a probability space $(\Omega, \mathscr{F}, P)$ and a one-dimensional Brownian motion $X = \{X_t, \tilde{\mathscr{F}}_t^X; 0 \le t < \infty\}$ on it; we assume $P[X_0 = 0] = 1$ and denote by $\{\tilde{\mathscr{F}}_t^X\}$ the augmentation of the filtration $\{\mathscr{F}_t^X\}$ under P. The same argument as before shows that

$$W_t \triangleq \int_0^t sgn(X_s)dX_s; \quad 0 \le t < \infty$$

is a Brownian motion adapted to $\tilde{\mathscr{F}}_t^X\}$. Corollary 3.2.20 shows that $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ is a weak solution to (3.1). With $\{\tilde{\mathscr{F}}_t^W\}$ denoting the augmentation of $\{\mathscr{F}_t^W\}$, this construction gives $\tilde{\mathscr{F}}_t^W \subseteq \tilde{\mathscr{F}}_t^X$, which is the opposite inclusion from that required of a strong solution.

Let us now show that equation (3.1) does not admit a strong solution. Assume the contrary, i.e., let X satisfy (3.1) on a given $(\Omega, \mathscr{F}, P)$ w.r.t. a given Brownian motion W, and assume $\tilde{\mathscr{F}}_t^X \subseteq \tilde{\mathscr{F}}_t^W$ for every $t \ge 0$. Then X is necessarily a Brownian motion, and from Tanaka's formula (3.6.13) with $a = 0$, we have

$$W_t = \int_0^t sgn(X_s)dX_s = |X_s| - 2L_t^X(0) =$$

$$= |X_t| - \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} meas\{0 \le s \le t; |X_s| \le \varepsilon\}, \quad 0 \le t < \infty, \text{ a.s. } P$$

where $L_t^X(0)$ is the local time at the origin for X. Consequently, $\tilde{\mathscr{F}}_t^W \subseteq \tilde{\mathscr{F}}_t^{|X|}$, and thus also $\tilde{\mathscr{F}}_t^X \subseteq \tilde{\mathscr{F}}_t^{|X|}$ holds for every $t \ge 0$. But the last inclusion is absurd.

B. Weak Solution by Means of the Girsanov Theorem

The principal method for creating weak solutions to stochastic differential equations is transformation of drift via Girsanov theorem.

3.6 Proposition. Consider the stochastic differential equation

$$dX_t = b(t, X_t)dt + dW_t; \quad 0 \le t \le T, \tag{3.2}$$

where T is a fixed positive number, W is a d-dimensional Brownian motion, and $b(t, x)$ is a Borel-measurable, $\mathbb{R}^d$-valued function on $[0, T] \times \mathbb{R}^d$ which satisfies

$$\|b(t, x)\| \le K(1 + \|x\|); \quad 0 \le t \le T, \ x \in \mathbb{R}^d \tag{3.3}$$

for some positive constant K. For any probability measure μ on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, equation (3.2) has weak solution with initial distribution μ.

Proof) We begin with a d-dimensional Brownian family $X = \{X_t, \mathscr{F}_t; 0 \le t \le T\}, (\Omega, \mathscr{F}), \{P^x\}_{x \in \mathbb{R}^d}$. According to Corollary 3.5.16,

$$Z_t \triangleq \exp\left\{ \sum_{j=1}^d \int_0^t b_j(s, X_s)dX_s^{(j)} - \frac{1}{2}\int_0^t \|b(s, X_s)\|^2 ds \right\}$$

is a martingale under each measure P^x, so the Girsanov theorem 3.5.1 implies that, under Q^x given by $\frac{dQ^x}{dP^x} = Z_t$, the process

$$W_t \triangleq X_t - X_0 - \int_0^t b(s, X_s)ds; \quad 0 \le t \le T \tag{3.4}$$

is a Brownian motion with $Q^x[W_0 = 0] = 1, \forall x \in \mathbb{R}^d$. Rewriting (3.4) as

$$X_t = X_0 + \int_0^t b(s, X_s)ds + W_t; \quad 0 \le t \le T,$$

we see that, with $Q^\mu(A) \triangleq \int_{\mathbb{R}^d} Q^x(A)\mu(dx)$, the triple $(X, W), (\Omega, \mathscr{F}, Q^\mu), \{\mathscr{F}_t\}$ is a weak solution to (3.2). $\square$

Corollary 3.5.16 (Benes(1971)). Let the vector $\mu = (\mu^{(1)}, ..., \mu^{(d)})$ of progressively measurable functionals on $C[0,\infty)^d$ satisfy, for each $0 \le T < \infty$ and some $K_T > 0$ depending on T, the condition

$$\|\mu(t,x)\| \le K_T(1 + x^*(t)); \quad 0 \le t \le T,$$

where $x^*(t) \triangleq \max_{0 \le s \le t} \|x(s)\|$. Then, with $X_t = (X_t^{(1)}, ..., X_t^{(d)})$ defined by $X_t^{(i)}(\omega) = \mu^{(i)}(t, W.(\omega))$, we have $Z(X) = \exp\left[\sum_{i=1}^d \int_0^t X_s^{(i)} dW_s^{(i)} - \frac{1}{2}\int_0^t \|X_s\|^2 ds\right]$ is a martingale.

3.7 Remark. If we seek a solution to (3.2) defined for all $0 \le t < \infty$, we can repeat the preceding argument using the filtration $\{\mathscr{F}_t^X\}$ instead of $\{\mathscr{F}_t\}$ and citing Corollary 3.5.2 rather than Girsanov (3.5.1) theorem. Whereas $\{\mathscr{F}_t\}$ in Proposition 3.6 can be chosen to satisfy the usual conditions, $\{\mathscr{F}_t^X\}$ does not have property. Thus, as a last step in this construction, we take $\mathscr{N}$ to be the collection of null sets of $(\Omega, \mathscr{F}_\infty^X, Q^\mu)$, set $\tilde{\mathscr{G}}_t = \sigma(\mathscr{F}_t^X \cup \mathscr{N})$ and $\mathscr{G}_t = \tilde{\mathscr{G}}_{t+}$. The filtration $\{\mathscr{G}_t\}$ satisfies the usual conditions.

3.8 Remark. It is apparent from Corollary 3.5.16 that Proposition 3.6 can be extended to include the case

$$X_t = X_0 + \int_0^t b(s, X_s)ds + W_t; \quad 0 \le t \le T, \tag{3.5}$$

where $b(t,x)$ is a vector of progressively measurable functionals on $C[0,\infty)^d$; see Definition 3.14.

3.9 Remark. Even when the initial distribution μ degenerates to unit point mass at some $x \in \mathbb{R}^d$, the filtration $\{\mathscr{F}_t^W\}$ of the driving Brownian motion in (3.5) may be strictly smaller than the filtration $\{\mathscr{F}_t^X\}$ of the solution process (discussion following Definition 3.1). This is shown by the celebrated example of Cirel'son (1975).

Girsanov theorem is also helpful in the study of uniqueness in law of weak solutions. We use it to establish a companion to Proposition 3.6.

3.10 Proposition. Assume that $(X^{(i)}, W^{(i)}), (\Omega^{(i)}, \mathscr{F}^{(i)}, P^{(i)}), \{\mathscr{F}_t^{(i)}\}; i = 1, 2$, are weak solutions to (3.2) with the same initial distribution. If

$$P^{(i)}\left[\int_0^T \|b(t, X_t^{(i)})\|^2 dt < \infty\right] = 1; \quad i = 1, 2, \tag{3.6}$$

then $(X^{(1)}, W^{(1)})$ and $(X^{(2)}, W^{(2)})$ have the same law under their respective probability measures.

Proof) For each $k \geq 1$, let

$$\tau_k^{(i)} \triangleq T \wedge \inf\left\{0 \leq t \leq T; \int_0^t \|b(s, X_s^{(i)})\|^2 ds = k\right\}. \tag{3.7}$$

According to Novikov's condition (Corollary 3.5.13),

$$\xi_t^{(k)}(X^{(i)}) \triangleq \exp\left\{-\int_0^{t\wedge\tau_k^{(i)}} b(s, X_s^{(i)}) dW_s^{(i)} - \frac{1}{2}\int_0^{\tau_k^{(i)}} \|b(s, X_s^{(i)})\|^2 ds\right\} \tag{3.8}$$

is a martingale, so we may define probability measures $\tilde{P}_k^{(i)}$ on $\mathscr{F}_T^{(i)}, i = 1, 2$, according to the prescription $\frac{d\tilde{P}_k^{(i)}}{dP^{(i)}} = \xi_T^{(k)}(X^{(i)})$. The Girsanov theorem 3.5.1 states that, under $\tilde{P}_k^{(i)}$, the process

$$X_{t\wedge\tau_k^{(i)}}^{(i)} = X_0^{(i)} + \int_0^{t\wedge\tau_k^{(i)}} b(s, X_s^{(i)}) ds + W_{t\wedge\tau_k^{(i)}}^{(i)}; \quad 0 \leq t \leq T \tag{3.9}$$

is a d-dimensional Brownian motion with initial distribution μ, stopped at time $\tau_k^{(i)}$. But $\tau_k^{(i)}$, $\{W_t^{(i)}; 0 \leq t \leq \tau_k^{(i)}\}$, and $\xi_T^{(k)}(X^{(i)})$ can all be defined in terms of the process in (3.9) (Problem 3.5.6). Therefore, for $0 = t_0 < t_1 < ... < t_n = T$ and $\Gamma \in \mathscr{B}(\mathbb{R}^{2d(n+1)})$, we have

$$\begin{aligned}
&P^{(1)}[(X_{t_0}^{(1)}, W_{t_0}^{(1)}, ..., X_{t_n}^{(1)}, W_{t_n}^{(1)}) \in \Gamma; \tau_k^{(1)} = T] = \\
&= \int_{\Omega^{(1)}} \frac{1}{\xi_T^{(k)}(X^{(1)})} 1_{\{(X_{t_0}^{(1)}, W_{t_0}^{(1)}, ..., X_{t_n}^{(1)}, W_{t_n}^{(1)}) \in \Gamma; \tau_k^{(1)} = T\}} d\tilde{P}_k^{(1)} = \\
&= \int_{\Omega^{(2)}} \frac{1}{\xi_T^{(k)}(X^{(2)})} 1_{\{(X_{t_0}^{(2)}, W_{t_0}^{(2)}, ..., X_{t_n}^{(2)}, W_{t_n}^{(2)}) \in \Gamma; \tau_k^{(2)} = T\}} d\tilde{P}_k^{(2)} = \\
&= P^{(2)}[(X_{t_0}^{(2)}, W_{t_0}^{(2)}, ..., X_{t_n}^{(2)}, W_{t_n}^{(2)}) \in \Gamma; \tau_k^{(2)} = T].
\end{aligned} \tag{3.10}$$

By assumption (3.6), $\lim_{k\to\infty} P^{(i)}[\tau_k^{(i)} = T] = 1, i = 1, 2$, so the passage to the limit as $k \to \infty$ in (3.10) gives the desired conclusion. $\square$

5.13 Corollary (Novikov (1972). Let $W = \{(W_t^{(1)}, ..., W_t^{(n)}),\ \mathscr{F}_t;\ 0 \le t < \infty\}$ be a d-dimensional Brownian Motion, and let $X = (X_t^{(1)}, ..., X_t^{(d)}), \mathscr{F}_t;\ 0 \le t < \infty\}$ be a vector of measurable, adapted processes satisfying $P[\int_0^T (X_t^{(i)})^2 dt < \infty] = 1;\ 1 \le i \le d,\ 0 \le T < \infty$. If

$$E\left[\exp(\frac{1}{2}\int_0^T \|X_s\|^2 ds)\right] < \infty;\ 0 \le T < \infty,$$

then $Z(X)$ is a martingale.

3.11 Corollary. If the drift term $b(t,x)$ in (3.2) is uniformly bounded, then uniqueness in the sense of probability law holds for the equation (3.2). Furthermore, with $0 \le t_1 < t_2 < ... < t_n \le T$ and with the notation developed in the proof of Proposition 3.6, we have then

$$Q^\mu[(X_{t_1}, ..., X_{t_n}) \in \Gamma] = \int_{\mathbb{R}^d} E^x[1_{\{(X_{t_1},...,X_{t_n})\in\Gamma\}} Z_T]\mu(dx); \quad \Gamma \in \mathscr{B}(\mathbb{R}^{dn}). \tag{3.11}$$

3.14 Definition. Let $b_i(t,y)$ and $\sigma_{ij}(t,y); 1 \le i \le d, 1 \le j \le r$, be progressively measurable functionals from $[0,\infty)\times C[0,\infty)^d$ into $\mathbb{R}$ (Definition 3.5.15). A weak solution to the functional stochastic differential equation

$$dX_t = b(t,X)dt + \sigma(t,X)dW_t; \quad 0 \le t < \infty, \tag{3.15}$$

is a triple $(X,W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ satisfying (i), (ii) of Definition 3.1, as well as

$$(iii)\ \int_0^t \{|b_i(s,X)| + \sigma_{ij}^2(s,X)\}ds < \infty; \quad 1 \le i \le d, \quad 1 \le j \le r \quad t \ge 0,$$

$$(iv)\ X_t = X_0 + \int_0^t b(s,X)ds + \int_0^t \sigma(s,X)dW_s; \quad 0 \le t < \infty,$$

almost surely.

C. A Digression on Regular Conditional Probabilities

3.16 Definition Let $(\Omega, \mathscr{F}, P)$ be a probability space and $\mathscr{G}$ a sub-σ-field of $\mathscr{F}$. A function $Q(\omega; A) : \Omega \times \mathscr{F} \to [0,1]$ is called a **regular conditional probability** for $\mathscr{F}$ given $\mathscr{G}$ if

(i) for each $\omega \in \Omega,\ Q(\omega, \cdot)$ is a probability measure on $(\Omega, \mathscr{F})$,

(ii) for each $A \in \mathscr{F},$ the mapping $\omega \mapsto Q(\omega, A)$ is $\mathscr{G}$-measurable, and

(iii) for each $A \in \mathscr{F},\ Q(\omega, A) = P[A|\mathscr{G}];\ P$-a.e. $\omega \in \Omega$.

Suppose that, whenever $Q'(\omega, A)$ is another function with these properties, there exists a null set $N \in \mathscr{F}$ s.t. $Q(\omega, A) = Q'(\omega, A)$ for all $A \in \mathscr{F}$ and $\omega \in \Omega - N$. We then say that the regular conditional probability for $\mathscr{F}$ given $\mathscr{G}$ is unique.

3.17 Definition. Let $(\Omega, \mathscr{F})$ be a measurable space. We say that $\mathscr{F}$ is **countably determined** if there exists a countable collection of sets $\mathscr{M} \subseteq \mathscr{F}$ s.t., whenever two probability measures agree on $\mathscr{M}$, they also agree on $\mathscr{F}$. We say that $\mathscr{F}$ is **countably generated** if there exists a countable collection of sets $\mathscr{C} \subseteq \mathscr{F}$ s.t. $\mathscr{F} = \sigma(\mathscr{C})$.

In the space $C[0,\infty)^m$, for an arbitrary integer $m \geq 1$ let us introduce the σ-fields

$$\mathscr{B}(C[0,\infty)^m) \triangleq \sigma(z(s); 0 \leq s \leq t) = \varphi_t^{-1}(\mathscr{B}(C[0,\infty)^m)) \tag{3.19}$$

for $0 \leq t < \infty$, where

3.18 Theorem. Suppose that Ω is a complete, separable metric space, and denote the Borel σ-field $\mathscr{F} = \mathscr{B}(\Omega)$. Let P be a probability measure on $(\Omega, \mathscr{F})$, and let $\mathscr{G}$ be a sub σ-field of $\mathscr{F}$. Then a regular conditional probability Q for $\mathscr{F}$ given $\mathscr{G}$ exists and is unique. Furthermore, if $\mathscr{H}$ is a countable determined sub-σ-field of $\mathscr{G}$, then there exists a null set $N \in \mathscr{G}$ s.t.

$$Q(\omega; A) = 1_A(\omega); \quad A \in \mathscr{H},\ \omega \in \Omega - N \tag{iv}$$

In particular, if X is a measurable r.v. taking values in another complete, separable metric space, then with $\mathscr{H}$ denoting the σ-field generated by X, (iv) implies

$$Q(\omega; \{\omega' \in \Omega; X(\omega') = X(\omega)\}) = 1; \quad P\text{-a.e. } \omega \in \Omega. \tag{iv$'$}$$

When the σ-field $\mathscr{G}$ is generated by a r.v., we may recast the assertions of Theorem 3.18 as follows.

3.19 Theorem. Let $(\Omega, \mathscr{F}, P)$ be as in Theorem 3.18, and let X be a measurable mapping from this space into a measurable space $(S, \mathscr{S})$, on which it induces the distribution $PX^{-1}(B) \triangleq P[\omega \in \Omega; X(\omega) \in B], B \in \mathscr{S}$. There exists then a function $Q(x; A) : S \times \mathscr{F} \to [0,1]$, called a regular conditonal probability for $\mathscr{F}$ given X, s.t.

(i) for each $x \in S, Q(x; \cdot)$ is a probability measure on $(\Omega, \mathscr{F})$,

(ii) for each $A \in \mathscr{F}$, the mapping $x \mapsto Q(x; A)$ is $\mathscr{S}$-measurable, and

(iii) for each $A \in \mathscr{F}, Q(x; A) = P[A|X = x], PX^{-1}$-a.e., $x \in S$.

If $Q'(x; A)$ is another function with these properties, then there exists a set $N \in \mathscr{S}$ with $PX^{-1} = 0$ s.t. $Q(x; A) = Q'(x; A)$ for all $A \in \mathscr{F}$ and $x \in S - N$. Furthermore, if S is also a complete, separable metric space and $\mathscr{S} = \mathscr{B}(S)$, then N can be chosen so that we have an additional property:

$$Q(x; \{\omega \in \Omega, X(\omega) \in B\}) = 1_B(x); \quad B \in \mathscr{S},\ x \in S - N. \tag{iv}$$

In particular,

$$Q(x; \{\omega \in \Omega, X(\omega) = x\}) = 1; \quad PX^{-1}\text{-a.e. } x \in S. \tag{iv$'$}$$

D. Results of Yamada and Watanabe on Weak and Strong Solutions

Returning to our initial question about the relation between pathwise uniqueness and uniqueness in the sense of probability law, let us consider two weak solutions $(X^{(j)}, W^{(j)}), (\Omega_j, \mathscr{F}_j, v_j), \{\mathscr{F}_t^{(j)}\}; j = 1, 2$ of equation (2.1) with

$$\mu(B) \triangleq v_1[X_0^{(1)} \in B] = v_2[X_0^{(2)} \in B]; \quad B \in \mathscr{B}(\mathbb{R}^d). \tag{3.20}$$

We set $Y_t^{(j)} = X_t^{(j)} - Y_t^{(j)}$; $0 \leq t < \infty$, and we regard the j-th solution as consisting of three parts; $X_0^{(j)}, W^{(j)}$, and $Y^{(j)}$. This triple induces a measure P_j on

$$(\Theta, \mathscr{B}(\Theta)) \triangleq (\mathbb{R}^d \times C[0,\infty)^r \times C[0,\infty)^d,$$
$$\mathscr{B}(\mathbb{R}^d) \otimes \mathscr{B}(C[0,\infty)^r) \otimes \mathscr{B}(C[0,\infty)^d))$$

according to the prescription

$$P_j(A) \triangleq v_j[(X_0^{(j)}, W^{(j)}, Y^{(j)}) \in A]; \quad A \in \mathscr{B}(\Theta), j = 1, 2. \tag{3.21}$$

We denote by $\theta = (x, w, y)$ the generic element of Θ. The marginal of each P_j on the x-coordinate of θ is μ, the marginal on the w-coordinate is Wiener measure P_* and the distribution of the (x, w) pair is the product measure $\mu \times P_*$ because $X_0^{(j)}$ is $\mathscr{F}_0^{(j)}$-measurable and $W^{(j)}$ is independent of $\mathscr{F}_0^{(j)}$ (Problem 2.5.5). Furthermore, under P_j, the initial value of the y-coordinate is zero, a.s.

The two weak solutions $(X^{(1)}, W^{(1)})$ and $(X^{(2)}, W^{(2)})$ are defined on (possibly) different sample spaces. Our first task is to bring them together on the same, canonical space, while preserving their joint distributions. Towards this end, we note that on $(\Theta, \mathscr{B}(\Theta), P_j)$, there exists a regular conditional probability for $\mathscr{B}(\Theta)$ given (x, w). We shall be interested only in conditional probabilities of sets in $\mathscr{B}(\Theta)$ of the form $\mathbb{R}^d \times C[0,\infty)^r \times F$, where $F \in \mathscr{B}(C[0,\infty)^d)$. Thus, with a slight abuse of terminology, we speak of

$$Q_j(x, w; F) : \mathbb{R}^d \times C[0,\infty)r \times \mathscr{B}(C[0,\infty)^d) \to [0,1]$$

as the regular

conditional probability for $\mathscr{B}(C[0,\infty)^d)$ given (x, w). According to Theorem 3.19, this regular conditional probability enjoys the following properties:

(i) for each $x \in \mathbb{R}^d, w \in C[0,\infty)^r, Q_j(x, w, \cdot)$ is a probability measure on $(C[0,\infty)^d, \mathscr{B}(C[0,\infty)^d))$, (3.22)

(ii) for each $F \in \mathscr{B}(C[0,\infty)^d)$, the mapping $(x, w) \mapsto Q_j(x, w; F)$ is $\mathscr{B}(\mathbb{R}^d) \otimes \mathscr{B}(C[0,\infty)^r)$-measurable, and (3.22)

$(iii) P_j(G \times F) = \int_G Q_j(x, w; F)\mu(dx)P_*(dw); F \in \mathscr{B}(C[0,\infty)^d)$, $G \in \mathscr{B}(\mathbb{R}^d) \otimes \mathscr{B}(C[0,\infty)^r)$. (3.22)

Finally, we consider the measurable space $(\Omega, \mathscr{F})$, where $\Omega = \Theta \times C([0,\infty)^d)$ and $\mathscr{F}$ is the completion of the σ-field $\mathscr{B}(\Theta) \otimes \mathscr{B}(C[0,\infty)^d)$ by the collection $\mathcal{N}$ of null sets under the probability measure

$$P(d\omega) \triangleq Q_1(x, w; dy_1)\mu(dx)Q_2(x, w; dy_2)\mu(dx)P_*(dw). \tag{3.23}$$

We have denoted by $\omega = (x, w, y_1, y_2)$ a generic element of Ω. In order to endow $(\Omega, \mathscr{F}, P)$ with a filtration that satisfies the usual conditions, we take

$$\mathscr{G}_t \triangleq \sigma\{(x, w(s), y_1(s), y_2(s)); 0 \le s \le t\}, \quad \tilde{\mathscr{G}_t} \triangleq \sigma(\mathscr{G}_t \cup \mathcal{N}), \quad \mathscr{F}_t \triangleq \tilde{\mathscr{G}}_{t+},$$

for $0 \le t < \infty$. From (3.21), (3.22) (iii), (3.23), we have

$$P[\omega \in \Omega; (x, w, y_j) \in A] = v_j[(X_0^{(j)}, W^{(j)}, Y^{(j)}) \in A]; \quad A \in \mathscr{B}(\Theta), j = 1, 2, \tag{3.21'}$$

and so the distribution of $(x + y_j, w)$ under P is the same as distribution of $X^{(j)}, W^{(j)}$ under v_j. In particular, the w-coordinate process $\{w(t), \mathscr{G}_t; 0 \le t < \infty\}$ is an r-dimensional Brownian motion on $(\Omega, \mathscr{F}, P)$, and it is then not difficult to see that the same is true for $\{w(t, \mathscr{F}_t; 0 \le t < \infty)\}$.

3.20 Proposition (Yamada & Watanabe (1971)). Pathwise uniqueness implies uniqueness in the sense of probability law.

Proof) We start with two weak solutions $(X^{(j)}, W^{(j)}), (\Omega_j, \mathscr{F}_j, v_j), \{\mathscr{F}_t^{(j)}\}; j = 1; 2$, of equation (2.1), with (3.20) satisfied. We have created two weak solutions $(x + y_j, w)$, $j = 1, 2$, on a single probability space $(\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$, s.t. $(X^{(j)}, W^{(j)})$ under v_j has the same law as $(x+y_j, w)$ under P. Pathwise uniqueness implies $P[x + y_1(t) = x + y_2(t), \forall 0 \le t < \infty] = 1$, or equivalently,

$$P[\omega = (x, w, y_1, y_2) \in \Omega; y_1 = y_2] = 1 \tag{3.24}$$

It develops from (3.21′), (3.24) that

$$\begin{aligned} v_1[(X_0^{(1)}, W^{(1)}, Y^{(1)}) \in A] &= P[\omega \in \Omega; (x, w, y_1) \in A] = \\ &= P[\omega \in \Omega; (x, w, y_2) \in A] = \\ &= v_2[(X_0^{(2)}, W^{(2)}, Y^{(2)}) \in A]; \quad A \in \mathscr{B}(\Theta) \end{aligned}$$

and this is uniqueness in the sense of probability law. $\square$

Proposition 3.20 has the remarkable corollary that weak existence and pathwise uniqueness imply strong existence.

3.23 Corollary. Suppose that the stochastic differential equation (2.1) has a weak solution $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ with initial distribution μ, and suppose that pathwise uniqueness holds for (2.1). Then there exists a $\mathscr{B}(\mathbb{R}^d) \otimes \mathscr{B}(C[0,\infty)^r)/\mathscr{B}(C[0,\infty)^d)$-measurable function $h : \mathbb{R}^d \times C[0,\infty)^r \to C[0,\infty)^d$, which is also $\hat{\mathscr{B}}_t/\mathscr{B}_t(C[0,\infty)^d)$-measurable for every fixed $0 \le t < \infty$, s.t.

$$\tilde{X}_{\cdot} = h(\xi, \tilde{W}_{\cdot}), \quad \text{a.s. } P. \tag{3.29}$$

Moreover, given any probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ rich enough to support an $\mathbb{R}^d$-valued r.v. ξ with distribution μ and an independent Brownian motion $\{\tilde{W}_t, \mathscr{F}_t^{\tilde{W}}; 0 \le t < \infty\}$, the process

$$\tilde{X}. \triangleq h(\xi, \tilde{W}.) \tag{3.30}$$

is a strong solution of equation (2.1) with initial condition ξ.

Proof) Let $h(x,w) = x + k(x,w)$, where k is as in Problem 3.22. From (**??**), and (3.21$'$), we see that (3.29) holds. For ξ and $\tilde{W}$ as described, both $(X_0, W.)$ and $(\xi, \tilde{W}.)$ induce the same measure $\mu \times P_*$ on $\mathbb{R}^d \times \mathscr{B}(C[0,\infty)^r)$, and since $(X. = h(X_0, W.), W.)$ satisfies (2.1), so does $(\tilde{X}. = h(\xi, \tilde{W}.), \tilde{W}.)$. The process $\tilde{X}$ is adapted to $\{\mathscr{F}_t\}$ given by (2.3), because h is $\hat{\mathscr{B}}_t/\mathscr{B}_t(C[0,\infty)^d)$-measurable. $\square$

5.4 The Martingale Problem of Stroock and Varadhan

A. Some Fundamental Martingales

Suppose that $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ is a weak solution to the stochastic differential equation (2.1). For every $t \ge 0$, we introduce the second-order differential operator

$$(\mathscr{A}_t f)(x) \triangleq \frac{1}{2}\sum_{i=1}^{d}\sum_{k=1}^{d} a_{ik}(t,x)\frac{\partial^2 f(x)}{\partial x_i \partial x_k} + \sum_{i=1}^{d} b_i(t,x)\frac{\partial f(x)}{\partial x_i}; \quad f \in C^2(\mathbb{R}^d), \tag{4.1}$$

where $a_{ik}(t,x) = \sum_{j=1}^{r} \sigma_{ij}(t,x)\sigma_{kj}(t,x)$ are the components of the diffusion matrix (2.2). If, as in the next proposition, f is a function of $t \in [0,\infty)$ and $x \in \mathbb{R}^d$, then $(\mathscr{A}_t f)(t,x)$, is obtained by applying $\mathscr{A}_t$ to $f(t,\cdot)$.

4.2 Proposition. For every continuous function $f(t,x) : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ which belongs to $C^{1,2}([0,\infty) \times \mathbb{R}^d)$, the process $M^f = \{M_t^f, \mathscr{F}_t; 0 \le t < \infty\}$ given by

$$M_t^f \triangleq f(t, X_t) - f(0, X_0) - \int_0^t \left(\frac{\partial f}{\partial s} + \mathscr{A}_s f\right)(s, X_s) ds \tag{4.2}$$

is a continuous, local martingale; i.e., $M^f \in \mathscr{M}^{c,loc}$. If g is another member of $C^{1,2}([0,\infty) \times \mathbb{R}^d)$, then $M^g \in \mathscr{M}^{c,loc}$ and

$$\langle M^f, M^g \rangle_t = \sum_{i=1}^{d}\sum_{k=1}^{d} \int_0^t a_{ik}(s, X_s)\frac{\partial}{\partial x_i} f(s, X_s)\frac{\partial}{\partial x_k} g(s, X_s) ds. \tag{4.3}$$

Furthermore, if $f \in C_0([0,\infty) \times \mathbb{R}^d)$ and the coefficients $\sigma_{ij}; 1 \le i \le d, 1 \le j \le r$, are bounded on the support of f, then $M^f \in \mathscr{M}^c$.

Proof) By Ito's rule on $f(t, X_t)$, we have

$$f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial f}{\partial s}(s,X_s)ds + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(s,X_s)dX_s^{(i)} +$$

$$+ \frac{1}{2}\int_0^t \sum_{i=1}^{d}\sum_{j=1}^{r} \frac{\partial^2 f}{\partial x_i \partial x_j}(s,X_s)d\langle X^{(i)}, X^{(j)}\rangle_s =$$

$$= f(0,X_0) + \int_0^t \frac{\partial f}{\partial s}(s,X_s)ds + \sum_{i=1}^{d}\int_0^t b_i(s,X_s)\frac{\partial f}{\partial x_i}(s,X_s)ds +$$

$$+ \sum_{i=1}^{d}\sum_{j=1}^{r}\int_0^t \sigma_{ij}(s,X_s)\frac{\partial f}{\partial x_i}(s,X_s)dW_s^{(j)} + \frac{1}{2}\int_0^t \sum_{i=1}^{d}\sum_{j=1}^{r} a_{ij}(s,X_s)\frac{\partial^2 f}{\partial x_i \partial x_j}(s,X_s)ds.$$

$$\therefore M_t^f = \sum_{i=1}^{d}\sum_{j=1}^{r} M_t^{(i,j)} \triangleq \sum_{i=1}^{d}\sum_{j=1}^{r}\int_0^t \sigma_{ij}(s,X_s)\frac{\partial}{\partial x_i}f(s,X_s)dW_s^{(j)}. \qquad (4.4)$$

Introducing the stopping times

$$S_n \triangleq \inf\left\{t \geq 0; \|X_t\| \geq n \text{ and } \int_0^t \sigma_{ij}^2(s,X_s)ds \geq n \text{ for some } (i,j)\right\}$$

and recalling that a weak solution must satisfy condition (iii) of Definition 2.1: $P\left[\int_0^t \{|b_i(s,X_s)| + \sigma_{ij}^2(s,X_s)\}ds < \infty\right] = 1$ holds for every $1 \leq i \leq d,\ 1 \leq j \leq r,\ 0 \leq t < \infty$. Thus, we see that $\lim_{n\to\infty} S_n = \infty$ a.s. The processes

$$M_t^f(n) \triangleq M_{t\wedge S_n}^f = \sum_{i=1}^{d}\sum_{j=1}^{r}\int_0^{t\wedge S_n} \sigma_{ij}(s,X_s)\frac{\partial}{\partial x_i}f(s,X_s)dW_s^{(j)}; \quad n \geq 1, \quad (4.5)$$

are continuous martingales, and so $M^f \in \mathscr{M}^{c,loc}$.
Furthermore,

$$\langle M^f, M^g\rangle_t = \lim_{n\to\infty}\langle M^f, M^g\rangle_{t\wedge S_n} = \lim_{n\to\infty}\langle M^f(n), M^g(n)\rangle_t =$$

$$= \lim_{n\to\infty}\sum_{i=1}^{d}\sum_{k=1}^{d}\int_0^t a_{ik}(s,X_s)\frac{\partial}{\partial x_i}f(s,X_s)\frac{\partial}{\partial x_k}g(s,X_s)ds.$$

If f has a compact support on which each σ_{ij} is bounded, then the integrand in the expression for $M^{(i,j)}$ in (4.4) is bounded, so $M^f \in \mathscr{M}_2^c$. $\square$

The simplest case in Proposition 4.2 is that of a d-dimensional Brownian motion, which corresponds to $b_i(t,x) \equiv 0$ and $\sigma_{ij}(t,x) \equiv \delta_{ij};\ 1 \leq i,j \leq d$. Then the operator in (4.1) becomes

$$\mathscr{A}f = \frac{1}{2}\Delta f = \frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2 f}{\partial x_i^2}; \quad f \in C^2(\mathbb{R}^d).$$

4.3 Problem. Let $b_i(t,y)$ and $\sigma_{ij}(t,y); 1 \le i \le d, 1 \le j \le r$, be progressively measurable functionals from $[0,\infty) \times C[0,\infty)^d$ into $\mathbb{R}$. By analogy with (2.2), we define the diffusion matrix $a(t,y)$ with components

$$a_{ik}(t,y) \triangleq \sum_{j=1}^{r} \sigma_{ij}(t,y)\sigma_{kj}(t,y); \quad 0 \le t < \infty, y \in C[0,\infty)^d. \tag{4.6}$$

Suppose that $(X,W),(\Omega,\mathscr{F},P),\{\mathscr{F}_t\}$, is a weak solution to the functional stochastic differential equation (3.15), and set

$$(\mathscr{A}_t^{'} u)(y) = \frac{1}{2}\sum_{i=1}^{d}\sum_{k=1}^{d} a_{ik}(t,y)\frac{\partial^2 u(y(t))}{\partial x_i \partial x_k} + \sum_{i=1}^{d} b_i(t,y)\frac{\partial u(y(t))}{\partial x_i}; \tag{4.1$'$}$$

$$0 \le t < \infty,\ u \in C^2(\mathbb{R}^d),\ y \in C[0,\infty)^d.$$

Then, for any functions $f, g \in C^{1,2}([0,\infty)\times\mathbb{R}^d)$, the process

$$M_t^f \triangleq f(t,X_t) - f(0,X_0) - \int_0^t \left[\frac{\partial f}{\partial s} + \mathscr{A}_s^{'} f\right](s,X_s)ds, \mathscr{F}_t; \quad 0 \le t < \infty \tag{4.2$'$}$$

is in $\mathscr{M}^{c,loc}$, and

$$\langle M^f, M^g\rangle_t = \sum_{i=1}^{d}\sum_{k=1}^{d} a_{ik}(s,X)\frac{\partial}{\partial x_i} f(s,X_s)\frac{\partial}{\partial x_k} g(s,X_s)ds. \tag{1}$$

Furthermore, if the first derivatives of f are bounded, and for each $0 < T < \infty$ we have

$$\|\sigma(t,y)\| \le K_T; \quad 0 \le t \le T \quad y \in C[0,\infty)^d, \tag{4.7}$$

where K_T is a constant depending on T, then $M^f \in \mathscr{M}_2^c$.

4.4 Problem. A continuous, adapted process $W = \{W_t, \mathscr{F}_t; 0 \le t < \infty\}$ is a d-dimensional Brownian motion if and only if

$$f(W_t) - f(W_0) - \frac{1}{2}\int_0^t \Delta f(W_s)ds, \mathscr{F}_t; \quad 0 \le t < \infty,$$

is in $\mathscr{M}^{c,loc}$ for every $f \in C^2(\mathbb{R}^d)$.

B. Weak Solutions and Martingale Problems

Problem 4.4 provides a novel martingale characterization of Brownian motion. The basic idea in the theory of Strock & Varadhan is to employ M^f of (4.2) in a similar fashion to characterize diffusions with general drift and dispersion coefficients.

4.5 Definition. A probability measure P on $(C[0,\infty)^d), \mathscr{B}(C[0,\infty)^d))$, under which

$$M_t^f = f(y(t)) - f(y(0)) - \int_0^t (\mathscr{A}_s^{'} f)(y) ds, \mathscr{F}_t; \quad 0 \le t < \infty, \tag{4.8}$$

is a continuous, local martingale for every $f \in C^2(\mathbb{R}^d)$, is called a **solution to the local martingale problem** associated with $\{\mathscr{A}^{'}\}$. Here, $\mathscr{F}_t = \mathscr{G}_{t+}$ and $\{\mathscr{G}_t\}$ is the augmentation under P of the canonical filration $\mathscr{B}_t \triangleq \mathscr{B}(C[0,\infty)^d)$ as in (3.19).

According to Problem 4.3, a weak solution to the functional stochastic differential equation (3.15) induces on $(C[0,\infty)^d), \mathscr{B}(C[0,\infty)^d))$ a probability measure P which solves the local martingale problem associated with $\{\mathscr{A}^{'}\}$. The converse of this assertion is also true.

4.6 Proposition. Let P be a probability measure on $(C[0,\infty)^d), \mathscr{B}(C[0,\infty)^d))$ under which the process M^f of (4.8) is a continuous, local martingale for the choices $f(x) = x_i$ and $f(x) = x_i x_k; 1 \le i, k \le d$. Then there is a r-dimensional Brownian motion $W = \{W_t, \tilde{\mathscr{F}}_t; 0 \le t < \infty\}$ defined on an extension $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ of $(C[0,\infty)^d), \mathscr{B}(C[0,\infty)^d), P)$, s.t. $(X_t \triangleq y(t), W_t), (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}), \{\tilde{\mathscr{F}}_t\}$ is a weak solution to equation (3.15).

Proof) By assumption,

$$M_t^{(i)} \triangleq X_t^{(i)} - X_0^{(i)} - \int_0^t b_i(s,X) ds, \ \mathscr{F}_t; \quad 0 \le t < \infty$$

is a continuous, local martingale under P. In particular,

$$P\left[\int_0^t |b_i(s,X)| ds < \infty; 0 \le t < \infty\right] = 1; \quad 1 \le i \le d. \tag{4.9}$$

With $f(x) = x_i x_k$, we see that

$$M_t^{(i,k)} \triangleq X_t^{(i)} X_t^{(k)} - X_0^{(i)} X_0^{(k)} - \int_0^t [X_s^{(i)} b_k(s,X) + X_s^{(k)} b_i(s,X) + a_{ik}(s,X)] ds$$

is also a continuous, local martingale. But one can express

$$M_t^{(i)} M_t^{(k)} - \int_0^t a_{ik}(s,X) ds = \tag{4.10}$$

$$= \left[X_t^{(i)} - X_0^{(i)} - \int_0^t b_i(s,X)ds\right]\left[X_t^{(k)} - X_0^{(k)} - \int_0^t b_k(s,X)ds\right] - \int_0^t a_{ik}(s,X)ds$$

$$= X_t^{(i)}X_t^{(k)} - X_t^{(i)}X_0^{(k)} - X_0^{(i)}X_t^{(k)} + X_0^{(i)}X_0^{(k)} - (X_t^{(i)} - X_0^{(i)})\int_0^t b_k(s,X)ds -$$

$$-(X_t^{(k)} - X_0^{(k)})\int_0^t b_i(s,X)ds - \int_0^t a_{ik}(s,X)ds = M_t^{(i,k)} - X_0^{(i)}M_t^{(k)} - X_0^{(k)}M_t^{(i)}$$

$$+ X_0^{(i)}[M_t^{(k)} - X_t^{(k)} + X_0^{(k)} + \int_0^t b_k(s,X)ds] + X_0^{(k)}[M_t^{(i)} - X_t^{(i)} + X_0^{(i)} +$$

$$+ \int_0^t b_i(s,X)ds] + \int_0^t (X_s^{(i)} - X_t^{(i)})b_k(s,X)ds + \int_0^t (X_s^{(k)} - X_t^{(k)})b_i(s,X)ds.$$

where $M_t^{(i,k)} - X_0^{(i)}M_t^{(k)} - X_0^{(k)}M_t^{(i)}$ is a continuous local martingale and the process below is also a local martingale:

$$\int_0^t (X_s^{(i)} - X_t^{(i)})b_k(s,X)ds + \int_0^t (X_s^{(k)} - X_t^{(k)})b_i(s,X)ds + \tag{4.11}$$

$$+ \int_0^t b_i(s,X)ds \int_0^t b_k(s,X)ds =$$

$$= \int_0^t (M_s^{(i)} - M_t^{(i)})b_k(s,X)ds + \int_0^t (M_s^{(k)} - M_t^{(k)})b_i(s,X)ds =$$

$$= -\int_0^t \left[\int_0^s b_k(u,X)du\right] dM_s^{(i)} - \int_0^t \left[\int_0^s b_i(u,X)du\right] dM_s^{(k)}.$$

The first equality holds because

$$\int_0^t (M_s^{(i)} - M_t^{(i)})b_k(s,X)ds = \int_0^t (X_s^{(i)} - X_0^{(i)} - \int_0^s b_i(u,X)du - X_t^{(i)} + X_0^{(i)} +$$

$$+ \int_0^t b_i(u,X)du)b_k(s,X)ds = \int_0^t (X_s^{(i)} - X_t^{(i)})b_k(s,X)ds -$$

$$- \int_0^t \int_0^s b_i(u,X)b_k(s,X)duds + \int_0^t b_i(u,X)du \int_0^t b_k(s,X)ds,$$

and the same holds for i replaced by k. Thus, we have by change of variables,

$$\int_0^t \int_0^t b_i(u,X)b_k(s,X)duds - \int_0^t \int_0^s b_i(u,X)b_k(s,X)duds =$$

$$= \int_0^t \int_s^t b_i(u,X)b_k(s,X)duds = \int_0^t \int_u^t b_i(s,X)b_k(u,X)dsdu =$$

$$= \int_0^t \int_0^s b_k(u,X)b_i(s,X)duds.$$

The last equality is true by setting $Y_t = \int_0^t b_k(s,X)ds$ and using the integration by parts formula (Problem 3.12):

$$\int_0^t \left[\int_0^s b_k(u,X)du\right] dM_s^{(i)} = \int_0^t Y_s dM_s^{(i)} = M_t^{(i)} Y_t - M_0^{(i)} Y_0 - \int_0^t M_s^{(i)} dY_s =$$
$$= \int_0^t M_t^{(i)} b_k(s,X)ds - \int_0^t M_s^{(i)} b_k(s,X)ds = -\int_0^t (M_s^{(i)} - M_t^{(i)}) b_k(s,X)ds.$$

We see then that the process of (4.11) is a continuous local martingale. Therefore, the process of (4.10) is in $\mathscr{M}^{c,loc}$, and

$$\langle M^{(i)}, M^{(k)}\rangle_t = \int_0^t a_{ik}(s,X)ds; \quad 0 \le t < \infty, \text{ a.s.} \tag{4.12}$$

We may now invoke Theorem 3.4.2 to conclude the existence of a d-dimensional Brownian motion $\{\tilde{W}_t, \tilde{\mathscr{F}}_t; 0 \le t < \infty\}$ on an extension $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ of $(C[0,\infty)^d), \mathscr{B}(C[0,\infty)^d), P)$ endowed with a filtration $\{\tilde{\mathscr{F}}_t\}$ which satisfies the usual conditions, as well as the existence of a matrix $\rho = \{\rho_{ij}(t), \tilde{\mathscr{F}}; 0 \le t < \infty\}_{1\le i,j\le d}$ of measurable, adapted processes with

$$\tilde{P}\left[\int_0^t \rho_{ij}^2(s)ds < \infty\right] = 1; \quad 1 \le i,j \le d, 0 \le t < \infty \tag{4.13}$$

such that

$$M_t^{(i)} = \sum_{j=1}^d \int_0^t \rho_{ij}(s) d\tilde{W}_s^{(j)}; \quad 1 \le i,j \le d, 0 \le t < \infty \tag{4.14}$$

holds a.s. $\tilde{P}$. This last equation can be rewritten as

$$X_t = X_0 + \int_0^t b(s,X)ds + \int_0^t \rho(s) d\tilde{W}_s; \quad 0 \le t < \infty. \tag{4.15}$$

In order to complete the proof, we need to establish the existence of an r-dimensional Brownian motion $W = \{W_t, \tilde{\mathscr{F}}_t; 0 \le t < \infty\}$ on $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ s.t.

$$\int_0^t \rho(s) d\tilde{W}_s = \int_0^t \sigma(s,X) dW_s; \quad 0 \le t < \infty \tag{4.16}$$

holds $\tilde{P}$-almost surely. From (4.12), (4.14) and with notation (4.6), we have

$$\tilde{P}\left[\sum_{j=1}^d \rho_{ij}(t)\rho_{kj}(t) = a_{ik}(t,X), \text{for a.e. } t \ge 0\right] = 1; \quad 1 \le i,k \le d$$

and (4.13) will imply

$$\tilde{P}\left[\int_0^t \sigma_{ij}^2(s,X)ds < \infty\right] = 1; \quad 1 \le i \le d, 1 \le j \le r, 0 \le t < \infty. \tag{4.17}$$

The relations (4.9), (4.15)-(4.17) will then yield $(X, W), (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}), \{\tilde{\mathscr{F}}\}$ as a weak solution to (3.15).
Is suffices to construct W satisfying (4.16) under the assumption $r = d$. Indeed, if $r > d$ we may augment X, b and σ by setting $X_t^{(i)} = b_i(t, y) = \sigma_{ij}(t, y) = 0; d+1 \leq i \leq r, 1 \leq j \leq r$. This r-dimensional process X satisfies an appropriately modified version of (4.8), and we may proceed as before except now we shall obtain a matrix ρ which, like σ, will be of dimension $(r \times r)$. On the other hand, if $r < d$, we need only augment σ by setting $\sigma_{ij}(t, y) = 0; 1 \leq i \leq d, r+1 \leq j \leq d$, and nothing else is affected. Both ρ and σ are then $(d \times d)$ matrices.
According to Problem 4.7 following this proof, there exists a Borel-measurable $(d \times d)$-matrix-valued function $R(\rho, \sigma)$ defined on the set

$$D \triangleq \{(\rho, \sigma); \rho \text{ and } \sigma \text{ are } (d \times d) \text{ matrices with } \rho\rho^T = \sigma\sigma^T\} \tag{4.18}$$

such that $\sigma = \rho R(\rho, \sigma)$ and $R(\rho, \sigma) R^T(\rho, \sigma) = I$, the $(d \times d)$ identity matrix. We set

$$W_t \triangleq \int_0^t R^T(\rho_s, \sigma(s, X)) d\tilde{W}_s; \quad 0 \leq t < \infty.$$

Then $W^{(i)} \in \mathscr{M}, 1 \leq i \leq d$, and

$$\langle W^{(i)}, W^{(j)} \rangle_t = t\delta_{ij}; \quad 1 \leq i, j \leq d, 0 \leq t < \infty.$$

It follows from Levy's Theorem 3.3.16 that $\{W_t, \tilde{\mathscr{F}}_t; 0 \leq t < \infty\}$ is a d-dimensional Brownian motion. $\square$

4.7 Problem. Show that there exists a Borel-measurable, $(d \times d)$-matrix-valued function $R(\rho, \sigma)$ defined on the set D of (4.18) and s.t.

$$\sigma = \rho R(\rho, \sigma), \quad R(\rho, \sigma) R^T(\rho, \sigma) = I; \quad (\rho, \sigma) \in D.$$

4.8 Corollary. The existence of a solution P to the local martingale problem associated with $\{\mathscr{A}'\}$ is equivalent to the existence of a weak solution $(X, W), (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}), \{\tilde{\mathscr{F}}_t\}$ to the equation (3.15). The two solutions are related by $P = \tilde{P} X^{-1}$; i.e., X induces the measure P on $(C[0, \infty)^d), \mathscr{B}(C[0, \infty)^d))$.

4.9 Corollary. The uniqueness of the solution P to the local martingale problem with fixed but arbitrary initial distribution

$$P[y \in C[0, \infty)^d; y(0) \in \Gamma] = \mu(\Gamma); \quad \Gamma \in \mathscr{B}(\mathbb{R}^d)$$

is equivalent to uniqueness in the sense of probability law for the equation (3.15).

Because of the difficulty in computing expectations for local martingales, it is helpful to introduce the following modification of Definition 4.5.

4.10 Definition (Martingale Problem of Stroock & Varadhan (1969)). A probability measure P on $(C[0, \infty)^d), \mathscr{B}(C[0, \infty)^d))$ under which M^f in (4.8) is a continuous martingale for every $f \in C_0^2(\mathbb{R}^d)$ is called a **solution to the martingale problem** associated with $\{\mathscr{A}_t^{'}\}$.

Given progressively measurable functionals $b_i(t,y), \sigma_{ij}(t,y) : [0,\infty)\times C[0,\infty)^d; 1 \leq i \leq d, 1 \leq j \leq r$, the associated family of operators $\{\mathscr{A}_t^{'}\}$, and a probability measure μ on $\mathscr{B}(\mathbb{R}^d)$, we can consider the following three conditions:

(A) There exists a weak solution to the functional stochastic differential equation (3.15) with initial distribution μ.
(B) There exists a solution P to the local martingale problem associated with $\{\mathscr{A}_t^{'}\}$ with $P[y(0) \in \Gamma] = \mu(\Gamma); \Gamma \in \mathscr{B}(\mathbb{R}^d)$.
(C) There exists a solution P to the martingale problem associated with $\{\mathscr{A}_t^{'}\}$ with $P[y(0) \in \Gamma] = \mu(\Gamma); \Gamma \in \mathscr{B}(\mathbb{R}^d)$.

4.11 Proposition. Conditions (A) and (B) are equivalent and are implied by (C). Furthermore, (A) implies (C) under either of the additional assumptions:
(A.1) For each $0 < T < \infty$, condition (4.7) holds.
(A.2) Each $\sigma_{ij}(t,y)$ is of the form $\sigma_{ij}(t,y) = \tilde{\sigma}_{ij}(t,y(t))$, where the Borel-measurable functions $\tilde{\sigma}_{ij} : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}$ are bounded on compact sets.

Proof) We have already established the equivalence of (A) and (B). If P is a solution to the martingale problem and $f \in C^2(\mathbb{R}^d)$ does not necessarily have compact support, we can define for every integer $k \geq 1$ the stopping time

$$S_k \triangleq \inf\{t \geq 0; \|y(t)\| \geq k\}. \tag{4.19}$$

Let $g_k \in C_0^2(\mathbb{R}^d)$ agree with f on $\{x \in \mathbb{R}^d; \|x\| \leq k\}$. Under P, each $M_t^{g_k}$ is a martingale which agrees with M_t^f for $t \leq S_k$. It follows that $M^f \in \mathscr{M}^{c,loc}$; thus (C)$\Rightarrow$(B). Under (A.1), Problem 4.3 shows (A)$\Rightarrow$(C). Under (A.2), the argument for (A)$\Rightarrow$(C) is given in Proposition 4.2. $\Box$

4.12 Remark. It is not always necessary to verify the martingale property of M^f under P for every $f \in C_0^2(\mathbb{R}^d)$, in order to conclude that P solves the martingale problem. For example, consider $f_i(x) \triangleq x_i$ and $f_{ij}(x_i x_j)$ for $1 \leq i,j \leq d$, and choose sequences $\{g_i^{(k)}\}_{k=1}^\infty$, $\{g_{ij}^{(k)}\}_{k=1}^\infty$ of functions in $C_0^2(\mathbb{R}^d)$ s.t. $g_i^{(k)}(x) = f_i(x), g_{ij}^{(k)} = f_{ij}(x)$ for $\|x\| \leq k$. If $M^{g_i^{(k)}}$ and $M^{g_{ij}^{(k)}}$ are martingales for $1 \leq i,j \leq d$ and $k \geq 1$, then M^{f_i} and $M^{f_{ij}}$ are local martingales. According to Proposition 4.6, there is a weak solution to (3.15) and the Corollary 4.8 now implies that P solves the local martingale problem. Under either of the assumptions (A.1) and (A.2), P must also solve the martingale problem.

It is also instructive to note that in the Definition 3.1 and 3.14 of weak solution to a stochastic differential equation, the Brownian motion W appears only as a "nuisance parameter". The Strook & Varadhan formulation eliminates this "parametric" process completely. Indeed, the essence of the implication (B)$\Rightarrow$(A) proved in Proposition 4.6 is the construction of this process.

C. Well-Posedness and the Strong Markov Property