STOCHASTIC DIFFERENTIAL EQUATIONS WITH CONSTANT DISPERSION[1]

S. Ramasubramanian
Stat.-Math. Unit
Indian Statistical Institute
8th Mile, Mysore Road
Bangalore - 560 059

These notes grew out of a compulsion. Because of the format and schedule of the Refresher course, the lectures on stochastic differential equations had to run more or less concurrently with those on Brownian motion and stochastic calculus; so the emphasis is on s.d.e.'s with constant dispersion. Once the reader is familiar with the results and methods involving such s.d.e's, after becoming conversant with stochastic calculus, he/she can anticipate (and even supply) the technicalities needed to complete the analysis in the general case involving non constant dispersion.

1 Diffusion

Diffusion phenomenon concerns movement of a particle (sometimes called "solute") in a medium (usually called "solvent"). This movement is attributed to two factors:
(i) external forces / temperature gradient acting on the system
(ii) collision of solute particle with the more numerous solvent particles.

Size of solute and solvent particles are such that solute particles do not float or sediment; also interaction among solute particles is assumed to be zero; (dilute gas solution, colloidal solutions are examples).

The component of motion due to external forces is called "drift" (sometimes also referred to as first order motion), and the component due to bombard-

[1]These notes form part of a series of lectures on stochastic differential equations given at a Refresher course in Probability, Stochastic Processes and Applications, sponsored by the Indian Academy of Sciences, Bangalore, held at the Department of Mathematics, Cochin University of Science and Technology, Kochi in September - October, 2005.

ment by solvent particles is called "dispersion" (also referred to as second order motion).

We know that Brownian motion is the idealized molecular movement purely due to dispersion; that is, Brownian motion is caused entirely by the fluctuations due to the bombardment of solute by the innumerable solvent particles. This motivates the following: If $X(t)$ denotes the position of the diffusing solute particle at time t, then its displacement may be written as

$$\begin{aligned} X(t+\Delta t) - X(t) &= b(t, X(t))\Delta(t) + \sigma(t, X(t))[B(t+\Delta t) - B(t)] \\ &= b(t, X(t))\Delta t + \sigma(t, X(t))\Delta B(t) \qquad (1.1) \end{aligned}$$

or in differential notation

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))\ dB(t). \qquad (1.2)$$

Here $b(\cdot,\cdot)$ is the *drift* representing external forces and $\sigma(\cdot,\cdot)$ is the *dipersion* representing inhomogeneity of the medium, and $B(\cdot)$ is a Brownian motion.

The problem in giving a meaning to (1.2) or its corresponding integral version is that almost all Brownian trajectories have infinite variation over any bounded interval.

Kiyosi Ito succeeded in giving a meaning to integrals of the form $\left(\int_0^t f(s,\cdot)dB(s)\right)(\omega)$ for a large class of integrands f which can even be random. The basic idea is to take limits of Riemann type sums in mean square sense. Such an integral is called Ito integral. Once Ito integrals are developed one can interpret (1.2) as the integral equation

$$X(t) = X(0) + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dB(s). \qquad (1.3)$$

If $\sigma(\cdot,\cdot) \equiv \sigma$ is a constant, as one naturally expects, $\int_{t_1}^{t_2} \sigma\ dB(s) = \sigma(B(t_2) - B(t_1))$; so when σ is a constant, the integral equation (1.3) becomes

$$X(t) = X(0) + \sigma[B(t) - B(0)] + \int_0^t b(s, X(s))ds. \qquad (1.4)$$

The formal expression (1.2), suitably interpreted, is called a *stochastic differential equation*; (1.2) actually stands for the integral equation (1.3) with the last term interpreted as an Ito integral. Once we understand what is meant by a (unique) solution to (1.2) or equivalently (1.3), the solution process $X(\cdot)$ may be called the *diffusion process* with drift $b(\cdot,\cdot)$ and dispersion $\sigma(\cdot,\cdot)$.

When $\sigma(\cdot,\cdot) \equiv \sigma$ is a constant, note that the corresponding integral equation (1.4) does not involve any stochastic / Ito integral. So analysing (1.4) is easier. Observe that (1.4) represents a diffusion in a homogeneous medium but with a drift that may depend on time and space.

2 Gronwall inequality

We prove an elementary but very useful inequality

Theorem 2.1 *Let $T > 0$ be fixed. Let g be a continuous function satisfying*

$$0 \leq g(t) \leq \alpha(t) + \beta \int_0^t g(s)\ ds, 0 \leq t \leq T \tag{2.1}$$

where $\beta \geq 0$ is a constant and $\alpha(\cdot)$ is an integrable real valued function on $[0, T]$. Then

$$g(t) \leq \alpha(t) + \beta \int_0^t \alpha(s)\ \exp(\beta(t-s))ds, 0 \leq t \leq T. \tag{2.2}$$

In case $\alpha(\cdot) \equiv C$, a constant, then

$$g(t) \leq C\ e^{\beta t}, 0 \leq t \leq T.$$

Proof: Put $f(t) = e^{-\beta t} \int_0^t g(s)\ ds$. Then by the fundamental theorem of calculus and (2.1),

$$f'(t) = e^{-\beta t} \left[g(t) - \beta \int_0^t g(s)\ ds \right] \leq e^{-\beta t} \alpha(t).$$

(Here as $g(\cdot)$ is continuous note that $t \mapsto \int_0^t g(s)\ ds$ is differentiable.) Integrating the two extremes of the above inequality over $[0,t]$ and noting that $f(0) = 0$ we get $f(t) \leq \int_0^t \alpha(s)\ e^{-\beta s} ds$. Hence by the definition of $f(\cdot)$

$$\int_0^t g(s)\ ds \leq \int_0^t \alpha(s)\ e^{\beta(t-s)} ds.$$

Multiplying both sides by the positive constant β and adding $\alpha(\cdot)$ to both sides we get

$$\alpha(t) + \beta \int_0^t g(s)\ ds \leq \alpha(t) + \beta \int_0^t \alpha(s)\ e^{\beta(t-s)} ds. \qquad (2.3)$$

Now (2.1) and (2.3) yield (2.2).

Remark 2.2 The preceding result can also be proved by using (2.1) and Fubini's theorem repeatedly. Indeed,

$$\begin{aligned}
g(t) - \alpha(t) &\leq \beta \int_0^t g(s)\ ds \leq \beta \int_0^t \left[\alpha(s) + \beta \int_0^s g(r)\ dr \right] ds \\
&= \beta \int_0^t \alpha(s)\ ds + \beta^2 \int_0^t \left(\int_r^t ds \right) g(r)\ dr \\
&= \beta \int_0^t \alpha(s)\ ds + \beta^2 \int_0^t (t-s)\ g(s)\ ds \\
&\leq \beta \int_0^t \alpha(s)[1 + \beta(t-s)] ds + \beta^3 \int_0^t (t-s) \left(\int_0^s g(r)\ dr \right) ds \\
&= \beta \int_0^t \alpha(s)[1 + \beta(t-s)] ds + \frac{\beta^3}{2} \int_0^t (t-s)^2 g(s)\ ds \\
&\leq \beta \int_0^t \alpha(s) \left[1 + \beta(t-s) + \frac{(\beta(t-s))^2}{2!} \right] ds
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\beta^4}{2!}\int_0^t (t-s)^2 \left(\int_0^s g(r)\ dr\right) ds \\
\leq\ & \beta \int_0^t \alpha(s) \left[\sum_{j=0}^{k} \frac{(\beta(t-s))^j}{j!}\right] ds \\
&+\frac{\beta^{k+1}}{(k-1)!}\int_0^t (t-s)^{k-1} \left(\int_0^s g(r)\ dr\right) ds. \qquad (2.4)
\end{aligned}
$$

As $g(\cdot)$ is bounded over $[0,T]$ it is now clear that (2.4) leads to (2.2). Besides being more transparent, the above proof carries over also when ds integral in (2.1) is replaced by Riemann-Stieltjes integral. $\square$

3 A useful lemma

We shall prove a result whose proof is similar to the argument in the remark of the preceding section.

Lemma 3.1 *Let $g_n(\cdot), n = 1, 2, \ldots$ be nonnegative bounded continuous functions on $[0,T]$; (the g_n's are not assumed a priori to be uniformly bounded). Suppose there is a constant $C > 0$ such that*

$$0 \leq g_{n+1}(t) \leq C \int_0^t g_n(s)\ ds, 0 \leq t \leq T, n = 1, 2, \ldots. \qquad (3.1)$$

Then
(i) there is a constant $K > 0$ such that

$$0 \leq g_{n+1}(t) \leq K\ \frac{C^n t^n}{n!}, 0 \leq t \leq T, n = 1, 2, \ldots \qquad (3.2)$$

(ii) $\{g_n(\cdot)\}$ converges to zero uniformly over $[0,T]$;
(iii) $f_n(\cdot) \triangleq \sum\limits_{i=1}^{n} g_i(\cdot), n = 1, 2, \ldots$ converges uniformly over $[0,T]$ to a continuous function f.

Proof: By repeatedly using (3.1) and Fubini's theorem we see that

$$
\begin{aligned}
0 &\leq g_{n+1}(t) \leq C^2 \int_0^t \left(\int_0^s g_{n-1}(r)dr \right) ds \\
&= C^2 \int_0^t \left(\int_r^t ds \right) g_{n-1}(r)dr = C^2 \int_0^t (t-s) g_{n-1}(s)\ ds \\
&\leq C^3 \int_0^t (t-s) \left(\int_0^s g_{n-2}(r)dr \right) ds = C^3 \int_0^t \left(\int_r^t (t-s)ds \right) g_{n-2}(r)\ dr \\
&= \frac{C^3}{2} \int_0^t (t-s)^2 g_{n-2}(s)\ ds \leq \frac{C^4}{2} \int_0^t (t-s)^2 \left(\int_0^s g_{n-3}(r)\ dr \right) ds \\
&= \frac{C^4}{3!} \int_0^t (t-s)^3 g_{n-3}(s)\ ds \\
&\leq \ldots \leq \frac{C^n}{(n-1)!} \int_0^t (t-s)^{n-1} g_1(s)\ ds \\
&= K\ \frac{C^n t^n}{n!}
\end{aligned}
$$

where K is a bound for $g_1(\cdot)$; this proves the first assertion. The second assertion is immediate from (3.2). Again by (3.2), for $n \geq m$

$$
\begin{aligned}
\sup_{0 \leq t \leq T} |f_n(t) - f_m(t)| &= \sup_{0 \leq t \leq T} \sum_{i=m+1}^{n} g_i(t) \\
&\leq K \sum_{j=m}^{n-1} \frac{C^j T^j}{n!},
\end{aligned}
$$

whence it follows that $\{f_k\}$ is a Cauchy sequence in $C[0,T]$ with the sup norm. As $C[0,T]$ is complete, it follows that there is $f \in C[0,T]$ such that $\sup_{0 \leq t \leq T} |f_n(t) - f(t)| \to 0$ as $n \to \infty$. This completes the proof. □

4 An integral equation

As a prelude to the study of stochastic differential equations, we first consider a (non stochastic) integral equation.

Theorem 4.1 *Let $x_0 \in I\!R, \sigma \in I\!R, T > 0$ be fixed; let $b : I\!R \to I\!R$ be a Lipschitz continuous function, that is, there is a constant $K > 0$ such that*

$$|b(y) - b(y')| \leq K|y - y'|, \ \forall \ y, y' \in I\!R. \tag{4.1}$$

Let $w : [0, T] \to I\!R$ be a continuous function such that $w(0) = 0$. Consider the integral equation

$$y(t) = x_0 + \sigma w(t) + \int_0^t b(y(s))ds, 0 \leq t \leq T. \tag{4.2}$$

Then (4.2) has a unique continuous solution.

Proof: Proof is by Picard's method. Put $y_0(t) \equiv x_0, 0 \leq t \leq T$; for $n = 1, 2, \ldots$ define $y_n(\cdot)$ inductively by

$$y_n(t) = x_0 + \sigma w(t) + \int_0^t b(y_{n-1}(s))ds, 0 \leq t \leq T. \tag{4.3}$$

Note that each $y_n(\cdot)$ is a well defined continuous function on $[0, T]$. Put

$$g_n(t) = |y_n(t) - y_{n-1}(t)|, 0 \leq t \leq T, n = 1, 2, \ldots$$

Using the Lipschitz condition (4.1) note that

$$0 \leq g_{n+1}(t) \leq K \int_0^t g_n(s) \ ds, 0 \leq t \leq T \tag{4.4}$$

for each $n = 1, 2, \ldots$. By (4.4) and Lemma 3.1 we get for $n \geq m$,

$$\begin{aligned} \sup_{0\leq t\leq T} |y_n(t) - y_m(t)| &\leq \sup_{0\leq t\leq T} \sum_{j=m+1}^{n} g_j(t) \\ &\leq K_1 \sum_{i=m}^{n-1} \frac{K^i T^i}{i!}, \end{aligned} \tag{4.5}$$

where $K_1 = |b(x_0)|T + |\sigma| \sup_{0 \le t \le T} |w(t)|$. It now follows that $\{y_n(\cdot)\}$ is a Cauchy sequence in $C[0,T]$, and hence has a uniform limit over $0,T]$, say, $y(\cdot)$. Clearly $y(\cdot)$ is a continuous. By (4.1) and the dominated convergence theorem

$$\begin{aligned}
\sup_{0\le t\le T}\left|\int_0^t b(y_n(s))ds - \int_0^t b(y(s))ds\right| &\le K\int_0^T |y_n(s)-y(s)|\ ds \\
&\to 0, \text{ as } n\to\infty.
\end{aligned}$$

So taking uniform limits on both sides of (4.3) as $n \to \infty$, we see that $y(\cdot)$ satisfies (4.2).

To prove uniqueness let $y^{(1)}(\cdot), y^{(2)}(\cdot)$ be two solutions to (4.2). Then by (4.1),

$$\begin{aligned}
|y^{(1)}(t) - y^{(2)}(t)| &= \left|\int_0^t [b(y^{(1)}(s)) - b(y^{(2)}(s))]ds\right| \\
&\le K\int_0^t |y^{(1)}(s) - y^{(2)}(s)|ds, 0 \le t \le T.
\end{aligned}$$

Gronwall inequality now implies $y^{(1)}(\cdot) \equiv y^{(2)}(\cdot)$, completing the proof. □

Note: In the iterative procedure given above one may start with any $y_0(\cdot) \in C[0,T]$.

Exercise 4.2 Let $0 \le T_1 \le T$. Let $w(\cdot), \hat{w}(\cdot) \in C[0,T]$ be such that $w(0) = \hat{w}(0) = 0$ and $w(s) = \hat{w}(s), 0 \le s \le T_1$. Let $y(\cdot), \hat{y}(\cdot)$ respectively denote the solution to (4.2) corresponding to $w(\cdot), \hat{w}(\cdot)$; and let $\{y_n(\cdot) : n = 0, 1, 2, \ldots\}, \{\hat{y}_n(\cdot) : n = 0, 1, 2, \ldots\}$ denote the respective members of Picard iteration scheme. (Of course, $x_0, \sigma, b(\cdot)$ are the same in both cases)
(i) Show that $\hat{y}_n(t) = y_n(t), 0 \le t \le T_1, n = 0, 1, 2, \ldots$
(ii) Show that, using (i) or uniqueness, $y(t) = \hat{y}(t), 0 \le t \le T_1$.
(iii) Let $b(\cdot), \sigma$ be fixed. Conclude from the above that for ay $0 \le t \le T$, the value of $y(t)$ depends only on $x_0, \{w(s) : 0 \le s \le t\}$.

Example 4.3 *Linear equation.* Suppose $b(y) = \beta y$ where β is a constant. Then the integral equation (4.2) becomes

$$y(t) = x_0 + \sigma w(t) + \beta \int_0^t y(s)\ ds, 0 \leq t \leq T. \tag{4.6}$$

This equation can be solved explicitly. Indeed put $\alpha(t) = x_0 + \sigma w(t), t \leq T$. Then (4.6) is

$$y(t) = \alpha(t) + \beta \int_0^t y(s)\ ds, 0 \leq t \leq T \tag{4.7}$$

which is like the second part of (2.1) with inequality replaced by equality. Proceeding as in Remark 2.2 replacing inequality by equality we get from (4.7)

$$\begin{aligned} y(t) - \alpha(t) &= \beta \int_0^t \alpha(s) \left[\sum_{j=0}^{k} \frac{(\beta(t-s))^j}{j!} \right] ds \\ &\quad + \frac{\beta^{k+1}}{(k-1)!} \int_0^t (t-s)^{k-1} \left(\int_0^s y(r)\ dr \right) ds \end{aligned} \tag{4.8}$$

which is the analogue of (2.4). Note that $y(\cdot)$ is bounded by Theorem 4.1. So it is easily seen that the last term on r.h.s. of (4.8) goes to zero uniformly over $[0, T]$ as $k \to \infty$. Therefore

$$\begin{aligned} y(t) &= \alpha(t) + \beta \int_0^t \alpha(s)\ e^{\beta(t-s)} ds \\ &= x_0 + \sigma w(t) + \int_0^t (x_0 + \sigma w(s))\ \beta\ e^{\beta(t-s)} ds \\ &= x_0 e^{\beta t} + \sigma w(t) + \sigma \int_0^t w(s)\ \beta\ e^{\beta(t-s)} ds, 0 \leq t \leq T. \end{aligned} \tag{4.9}$$

As uniqueness is guaranteed by Theorem 4.1 it follows that $y(\cdot)$ given by (4.9) is the unique solution to the linear equation (4.6).

Remark 4.4 If $f(\cdot)$ and $g(\cdot)$ are both continuous functions of bounded variation over an interval $[a, b]$ then it is known that

$$\int_a^b f(s)\ dg(s) = f(b)\ g(b) - f(a)\ g(a) - \int_a^b g(s)\ df(s). \qquad (4.10)$$

This suggests that if $f(\cdot)$ is a continuous function of bounded variation over $[a, b]$ and $g(\cdot)$ is an *arbitrary* continuous function on $[a, b]$, then r.h.s. of (4.10) can be used to *define* $\int_a^b f(s)\ dg(s)$. (In fact this was among the first efforts to define stochastic integrals; this approach has limited success, and was later subsumed by the more versatile Ito integrals.) Now that the function $s \mapsto e^{\beta(t-s)}$ is a continuous function of bounded variation over $[0, t]$. So, though $w(\cdot)$ is an arbitrary continuous function (with $w(0) = 0$), (4.9) can now be written as

$$y(t) = x_0 e^{\beta t} + \sigma \int_0^t e^{\beta(t-s)} dw(s), 0 \le t \le T. \qquad (4.11)$$

Of course, it is to be kept in mind that r.h.s of (4.11) is "defined" by r.h.s. of (4.9). □

Exercise 4.5 (*Time inhomogeneous case*): Let $b : [0, T] \times I\!R \to I\!R$ be a measurable function such that there is a constant $K > 0$ satisfying

$$\begin{aligned} |b(s, y) - b(s, y')| &\le K|y - y'| && (4.12) \\ |b(s, 0)| &\le K && (4.13) \end{aligned}$$

for all $y, y' \in I\!R, 0 \le s \le T$. With $x_0, \sigma, w(\cdot)$ as in Theorem 4.1 show that the integral equation

$$y(t) = x_0 + \sigma w(t) + \int_0^t b(s, y(s))ds,\ 0 \le t \le T \qquad (4.14)$$

has a unique continuous solution.

Exercise 4.6 (*Higher dimensional equation*). Let $d \ge 2$ and $w : [0, T] \to I\!R^d$ be a continuous function with $w(0) = 0$; let σ be a $(d \times d)$ matrix with real entries. Let $b : [0,T] \times I\!R^d \to I\!R^d$ be a measurable function satisfying the analogues of (4.12), (4.13) with $|\cdot|$ denoting the euclidean distance in $I\!R^d$. For any $x_0 \in I\!R^d$ show that there is a unique continuous function $y : [0,T] \to I\!R^d$ solving the analogue of (4.14). $\square$

Exercise 4.7 Let $x_n \to x$ in $I\!R, w_n(\cdot) \to w(\cdot)$ in $C[0,T]$. Let $y_n(\cdot)$ and $y(\cdot)$ respectively denote the solution to (4.14) corresponding to x_n, w_n and x, w. Show that $y_n(\cdot) \to y(\cdot)$ in $C[0,T]$.

5 Strong solution to SDE

We shall consider one dimensional stochastic differential equation of the form

$$dY(t) = \sigma\ dB(t) + b(t, Y(t))\ dt$$

in this section. We begin by describing the set up.

Let $(S, \mathcal{F}, P)$ be a probability space. Let $\{B(t) : t \geq 0\}$ be a standard one dimensional Brownian motion defined on $(S, \mathcal{F}, P)$. For $t \geq 0$, let

$$\begin{aligned} \mathcal{F}_t^B &= \sigma\{B(r) : 0 \leq r \leq t\} \\ &= \text{ the smallest } \sigma - \text{algebra making the family} \\ &\quad\ \{B(r) : 0 \leq r \leq t\} \text{ measurable.} \end{aligned}$$

Let ξ be a real valued random variable independent of the Brownian motion $\{B(t) : t \geq 0\}$. Denote

$$\begin{aligned} \mathcal{G}_t &= \sigma\{\xi, B(r) : 0 \leq r \leq t\}, 0 \leq t < \infty \\ \mathcal{N} &= \{\text{all } P - \text{null sets }\}, \\ \mathcal{F}_t &= \sigma(\mathcal{N} \vee \mathcal{G}_t), 0 \leq t < \infty, \mathcal{F}_\infty = \sigma\left(\bigcup_{t \geq 0} \mathcal{F}_t\right). \end{aligned}$$

For $0 \leq s < t$ note that the random variable $B(t) - B(s)$ and the σ-algebra $\mathcal{F}_s$ are independent; so it can be seen that $\{B(t) : t \geq 0\}$ is an $\{\mathcal{F}_t\}$-adapted Brownian motion.

Notation: We shall use σ to denote the dispersion coefficient. The symbol σ occurs also in $\sigma\{\cdots\}$, the sigma algebra generated by the family $\{\cdots\}$. We trust there will be no scope for confusion.

Note: We have augmented the filtration $\{\mathcal{G}_t\}$ with null sets to form $\{\mathcal{F}_t\}$ so that statements about almost sure equivalences do not become a nuisance to handle. $\Box$

Let $\sigma \in I\!R$ be a constant and $b : [0, \infty) \times I\!R \to I\!R$ be a measurable function. We consider the stochastic differential equation (s.d.e.)

$$\left.\begin{array}{rcl} dY(t) & = & \sigma\ dB(t) + b(t, Y(t))dt, t > 0 \\ & & \text{with the initial value } Y(0) = \xi. \end{array}\right\} \qquad (5.1)$$

A *strong solution* to the s.d.e. (5.1) on the given probability space $(S, \mathcal{F}, P)$ and w.r.t. the fixed Brownian motion $\{B(t)\}$ and the initial value ξ is a stochastic process $\{Y(t) : t \geq 0\}$ such that the following hold:
(i) $\{Y(t)\}$ has continuous sample paths; that is, for $\omega \in S$ the function $t \mapsto Y(t, \omega)$ is continuous;
(ii) $\{Y(t)\}$ is $\{\mathcal{F}_t\}$-adapted; that is, for $t \geq 0, Y(t)$ is $\mathcal{F}_t$-measurable;
(iii) $P(Y(0) = \xi) = 1$;
(iv) for every $t \geq 0, P\left(\int_0^t |b(s, Y(s))|ds < \infty\right) = 1$ holds;
(v) the integral version of (5.1) holds, viz. for $t \geq 0$,

$$Y(t) = \xi + \sigma(B(t) - B(0)) + \int_0^t b(s, Y(s))\ ds \qquad (5.2)$$

holds a.s.

Next we define a concept of uniqueness. Suppose that whenever $\{Y(t)\}, \{\tilde{Y}(t)\}$ are two strong solutions to (5.1) on the same probability space $(S, \mathcal{F}, P)$ and w.r.t. the same Brownian motion $\{B(t)\}$, same initial value ξ which is independent of $B(\cdot)$, and with $\{\mathcal{F}_t\}$ defined as above, then we have $P(Y(t) = \tilde{Y}(t)$ for all $t) = 1$. In such a case we say that *strong uniqueness holds for the s.d.e. (5.1)*; (to be very precise we should say that strong uniqueness holds for the coefficients $\sigma, b(\cdot, \cdot)$).

Remark 5.1 a) In an obviously analogous way, the notions of strong solution and strong uniqueness can be defined when we consider the processes

only on the bounded time interval $[0, T]$.
b) Suppose we want to consider the above on an interval of the form $[s, T]$ or $[s, \infty]$ with $s > 0$ as the initial time point. Then the initial value ξ may be taken to be independent of $\{B(r) - B(s) : r \in [s, \infty)\}$; the analogues $\mathcal{G}_t, \mathcal{F}_t$ would be

$$\begin{aligned} \mathcal{G}_t^s &= \sigma\{\xi, B(r) - B(s) : r \in [s, t]\}, t \geq s \\ \mathcal{F}_t^s &= \sigma(\mathcal{N} \vee \mathcal{G}_t^s), t \geq s; \end{aligned}$$

moreover the analogue of (5.2) would be

$$Y(t) = \xi + \sigma(B(t) - B(s)) + \int_s^t b(r, Y(r))\, dr, t \geq s \qquad (5.3)$$

we leave the details to the reader. □

Theorem 5.2 *(Existence & uniqueness): Let $T > 0$ be fixed. Let $\{B(t)\}, \xi$, $\{\mathcal{F}_t\}$ be as above; let $\sigma \in \mathbb{R}$ be a constant and $b : [0, T] \times \mathbb{R} \to \mathbb{R}$ be a measurable function such that there are constants $K_1, K_2 > 0$ such that (Linear growth condition):*

$$|b(t, x)| \leq K_1(1 + |x|), \qquad (5.4)$$

(Lipschitz continuity condition):

$$|b(t, x) - b(t, y)| \leq K_2|x - y| \qquad (5.5)$$

hold for all $x, y \in \mathbb{R}, 0 \leq t \leq T$. Suppose $E(\xi^2) < \infty$. Then the s.d.e (5.1) with the initial value ξ has a strong solution $\{Y(t)\}$; also strong uniqueness holds. Moreover $E(|Y(t)|^2) \leq C_1 e^{C_2 t}, 0 \leq t \leq T$ for some constants C_1, C_2.

□

To prove the theorem we need a lemma.

Lemma 5.3 *Let $t \geq 0$. A measurable function f on $(S, \mathcal{F}, P)$ is $\mathcal{G}_t$-measurable $\Leftrightarrow f(\omega) = f(\omega')$ for any $\omega, \omega' \in S$ with $\xi(\omega) = \xi(\omega'), B(s, \omega) = B(s, \omega')$ for $0 \leq s \leq t$.*

Proof: By continuity of sample paths we know that $\mathcal{F}^B$ is generated by finite dimensional cylinder sets of the form $H = \{\omega : (B(s_1,\omega), \dots, B(s_k,\omega)) \in A\}$ with $A \in \mathcal{B}(I\!R^k), 0 \le s_1 < s_2 < \dots < s_k < \infty,\ k = 1,2,\dots$ So it is enough to consider f of the form $f = I_{\xi^{-1}(C)}.\ I_F$ where $C \in \mathcal{B}(I\!R)$ and F a finite dimensional cylinder set in $\sigma\left(\bigcup_{s\ge 0} \mathcal{F}_s^B\right)$. The details of the proof are left to the reader. □

Proof of Theorem 5.2: We shall give two proofs (which have a lot of common features!)

1st proof: Note that (5.4), (5.5) together are equivalent to (4.12), (4.13). Let $\omega \in S$ be such that $B(0,\omega) = 0$; for a.a.ω this holds; for fixed ω, put $w(s) = B(s,\omega)$. In this case, equation (5.2) becomes

$$Y(t,\omega) = \xi(\omega) + \sigma[w(t) - w(0)] + \int_0^t b(s, Y(s,\omega))\ ds.$$

So the solution to (5.2) corresponding to ω is just the solution to (4.14) with $x_0 = \xi(\omega)$ and $w(\cdot) = B(\cdot,\omega)$. A unique continuous solution exists by Theorem 4.1 and Exercise 4.5. Next, if $\omega, \omega' \in S$ are such that $B(s,\omega) = B(s,\omega')$ for all $0 \le s \le t, \xi(\omega) = \xi(\omega')$, then by Exercise 4.2 $Y(s,\omega) = Y(s,\omega')$ for all $0 \le s \le t$; so by the preceding lemma we infer that $Y(t)$ is $\mathcal{F}_t$-measurable; thus $Y(\cdot)$ is $\{\mathcal{F}_t\}$-adapted. Strong uniqueness follows from uniqueness in Theorem 4.1 and Exercise 4.5. Also, by Schwartz inequality

$$\begin{aligned} E(|Y(t)|^2) &\le 3E|\xi|^2 + 3\sigma^2 E[(B(t))^2] + 3E\left|\int_0^t b(s,Y(s))ds\right|^2 \\ &\le 3E|\xi|^2 + 3\sigma^2 t + 3t\ E\int_0^t |b(s,Y(s))|^2 ds \\ &\le 3E|\xi|^2 + 3\sigma^2 t + 6K_1 t^2 + 6K_1 T\int_0^t E[|Y(s)|^2]ds \quad (5.6) \end{aligned}$$

where we have used the linear growth condition (5.4). An application of Gronwall inequality now yields the last assertion of the theorem. The requirement (iv) of strong solution is also easy to obtain.

2nd proof: Put $Y_0(t,\omega) \equiv \xi(\omega)$; inductively for $n = 1, 2, \ldots$ define

$$Y_n(t,\omega) = \xi(\omega) + \sigma(B(t,\omega) - B(0,\omega)) + \int_0^t b(s, Y_{n-1}(s,\omega))\, ds. \qquad (5.7)$$

For fixed ω, t, note that $\int_0^t b(s, Y_j(s,\omega))ds$ is a limit of Riemann sums involving $s \in [0,t]$. So for $t \geq 0$, the Riemann sums and hence $\int_0^t b(s, Y_j(s))ds$ are $\mathcal{F}_t$-measurable. Thus $Y_n(\cdot)$ is $\{\mathcal{F}_t\}$-adapted for $n = 0, 1, 2, \ldots$ For fixed ω, take $g_n(t) = |Y_n(t,\omega) - Y_{n-1}(t,\omega)|, n = 1, 2, \ldots$ and proceed as in the proof of Theorem 4.1; one can get a unique continuous solution $Y(\cdot,\omega)$ to (5.2) as the uniform limit of $Y_n(\cdot,\omega)$. Consequently $\{\mathcal{F}_t\}$-adaptedness of $Y(\cdot)$ follows. The other assertions can now be proved as in the first proof. This completes the proof. $\square$

Remark 5.4 a) In an entirely analogous manner, given $0 < s < T$, under analogues of (5.4), (5.5), one can prove that the s.d.e. (5.3) with $t \in [s,T]$ has an $\{\mathcal{F}_t^s\}$-adapted continuous strong solution and that strong uniqueness holds.

b) Let $s \geq 0$. Using existence of strong solution and strong uniqueness on $[s,N]$ for each $N = 1, 2, \ldots$, and letting $N \to \infty$, analogues on $[s,\infty)$ are easily obtained. Details are left to the reader.

Proposition 5.5 Let $\sigma, b(\cdot,\cdot)$ be as in Theorem 5.2. Let $\xi_n, n = 1, 2, \ldots$ and ξ be real valued square integrable random variables such that $\xi_n \to \xi$ a.s. as $n \to \infty$, and the family $\{\xi, \xi_n : n = 1, 2, \ldots\}$ is independent of the Brownian motion $\{B(t) - B(0) : t \geq 0\}$. Let $Y_n(\cdot), n = 1, 2, \ldots$ and $Y(\cdot)$ denote respectively the unique strong solution to s.d.e. (5.2) corresponding to the initial value $\xi_n, n = 1, 2, \ldots$ and ξ. Then $Y_n(\cdot,\omega) \to Y(\cdot,\omega)$ uniformly over $[0,T]$ as $n \to \infty$, for a.a.ω.

Proof: Let $\omega \in S$ be such that $\xi_n(\omega) \to \xi(\omega)$. Then it is not difficult to verify that

$$\sup_{0\leq s\leq t} |Y_n(s,\omega) - Y(s,\omega)| \leq |\xi_n(\omega) - \xi(\omega)| + K_2 \int_0^t \left[\sup_{0\leq v\leq r} |Y_n(v,\omega) - Y(v,\omega)| \right] dr.$$

Consequently Gronwall inequality gives

$$\sup_{0\le s\le t} |Y_n(s,\omega) - Y(s,\omega)| \le |\xi_n(\omega) - \xi(\omega)|\ e^{K_2 t},\ 0 \le t \le T$$

whence the proposition follows. $\square$

An easy consequence of the preceding proposition and dominated convergence theorem is

Corollary 5.6 *Let the notation and hypotheses be as in the preceding proposition. Let $0 \le t_1 < t_2 < \ldots < t_k \le T$, and f a bounded uniformly continuous function on $\mathbb{R}^k$. Then $E[f(Y_n(t_1),\ldots,Y_n(t_k))] \to E[f(Y(t_1),\ldots,Y(t_k))]$ as $n \to \infty$; that is, $(Y_n(t_1),\ldots,Y_n(t_k)) \Rightarrow (Y(t_1),\ldots,Y(t_k))$ as $n \to \infty$ for every $0 \le t_1 < t_2 < \ldots < t_k \le T, k = 1,2,\ldots,$ where $\Rightarrow$ denotes convergence in distribution (or law) or weak convergence. This property is known as Feller continuity.*

Remark 5.7 It is known that two possibly different one dimensional standard Brownian motions induce the same probability measure, viz. the standard one dimensional Wiener measure on the function space $C[0,\infty)$. We amplify this as follows. For $i = 1,2$, let $(S^{(i)}, \mathcal{F}^{(i)}, P^{(i)})$ be a probability space, $\{B^{(i)}(t) : t \ge 0\}$ a standard one dimensional Brownian motion on $(S^{(i)}, \mathcal{F}^{(i)}, P^{(i)})$. Let $\Omega = C([0,\infty) : \mathbb{R}) = \{w : [0,\infty) \to \mathbb{R} : w$ continuous$\}$. Let Ω be endowed with the topology of uniform convergence on compact subsets of $[0,\infty)$; with this topology Ω is a complete separable metric space. For $t \ge 0$, let $X_t \triangleq X(t,\cdot)$ denote the t-th coordinate projection, that is, $X_t(w) = X(t,w) = w(t), w \in \Omega$. Let $\mathcal{B}_t^\circ = \sigma\{X(r) : 0 \le r \le t\}, \mathcal{B}^\circ \equiv \mathcal{B}_\infty^\circ = \sigma\left(\bigcup\limits_{t\ge 0} \mathcal{B}_t^\circ\right)$. Then it is known that $\mathcal{B}^\circ$ is the Borel σ-algebra on Ω and that $\mathcal{B}^\circ$ is generated by the finite dimensional cylinder sets; also for $t > 0, \mathcal{B}_t^\circ$ can be identified with the Borel σ-algebra of $C[0,t]$. For each $\omega^{(i)} \in S^{(i)}$, note that $s \mapsto B^{(i)}(s,\omega^{(i)})$ is a continuous function on $[0,\infty)$. It is not difficult to prove that $B^{(i)}$ given by $\omega^{(i)} \mapsto B^{(i)}(\cdot,\omega^{(i)})$ is a measurable map from $(S^{(i)}, \mathcal{F}^{(i)})$ into $(\Omega, \mathcal{B}^\circ)$. So $P^{(i)}(B^{(i)})^{-1}$ is a probability measure on $(\Omega,\mathcal{B}^\circ)$; that is, $P^{(i)}(B^{(i)})^{-1}$ is the distribution of the continuous process $\{B^{(i)}(t) : t \ge 0\}$ with $B^{(i)}$ being viewed upon as a $C[0,\infty)$-valued random variable on $(S^{(i)}, \mathcal{F}^{(i)}, P^{(i)})$. Our initial statement of this paragraph means just that $P^{(1)}(B^{(1)})^{-1} = P^{(2)}(B^{(2)})^{-1} =$ standard one dimensional

Wiener measure on Ω. (Why is this true? Look at the finite dimensional distributions!)

The preceding paragraph raises the following natural question: Do two possibly different strong solutions to the s.d.e. (5.3) on different probability spaces, with different standard one dimensional Brownian motions, but with the same $\sigma, b(\cdot,\cdot)$ and initial values (in law) induce the same distribution on the function space $C[s,\infty)$?

Let $s \geq 0$; let $\sigma, b(\cdot,\cdot)$ be as in Theorem 5.2 / Remark 5.4. For $i = 1, 2$, let $(S^{(i)}, \mathcal{F}^{(i)}, P^{(i)}), \{B^{(i)}(t)\}$ be as in the preceding paragraph; let $\xi^{(i)}$ be a real valued random variable such that $\xi^{(i)}$ is independent of $\{B^{(i)}(r) - B^{(i)}(s) : r \geq s\}$. Let $\{Y^{(i)}(t) : t \geq s\}$ denote the strong solution to the s.d.e. (5.3) starting at time s on $(S^{(i)}, \mathcal{F}^{(i)}, P^{(i)})$ with $\{B^{(i)}(t)\}$ as the driving Brownian motion, dispersion σ, drift $b(\cdot,\cdot)$ and initial value $\xi^{(i)}$ at time s. If $\xi^{(1)}$ and $\xi^{(2)}$ have the same distribution on $I\!R$, do the continuous processes $\{Y^{(1)}(t) : t \geq s\}$ and $\{Y^{(2)}(t) : t \geq s\}$ induce the same distribution on $\Omega_s = C[s,\infty)$?

The answer is yes; the proof is not difficult but notationally a bit tedious; the idea is to chase the solution via Picard method of approximation as in the second proof of Theorem 5.2; for each n one can show that the approximating sequences $\{Y_n^{(1)}(t)\}$ and $\{Y_n^{(2)}(t)\}$ induce the distribution on $\Omega_s = C[s,\infty)$; finally let $n \to \infty$ to get the desired conclusion. Thus the distribution of the strong solution to (5.3) is uniquely determined by the initial time point, initial distribution, dispersion and drift. This property is called *uniqueness in law*.

With $\sigma, b(\cdot,\cdot)$ kept fixed, for initial time point s, initial distribution (that is, distribution of the initial value ξ) μ, we shall denote by $Q_{s,\mu}$ the probability measure induced on $C[s,\infty)$ by the strong solution to the s.d.e. (5.3). When $\xi(\cdot) \equiv x$, that is, $\mu = \delta_x$, we shall write $Q_{s,x}$ for $Q_{s,\mu}$. Expectation w.r.t. $Q_{s,\mu}$ will be denoted $E_{s,\mu}(\cdots)$; that is, $E_{s,\mu}(h) = \int_{\Omega_s} h(w)\, dQ_{s,\mu}(w)$; similarly, $E_{s,x}(h) = \int_{\Omega_s} h(w)\, dQ_{s,x}(w)$. □

Remark 5.8 Let $\xi_n, Y_n, n = 1, 2, \ldots, \xi, Y, f, t_1, \ldots, t_k$ be as in Corollary 5.6; note that $\sigma, b(\cdot,\cdot)$ are fixed. Let $\xi_n \sim \mu_n, n = 1, 2, \ldots, \xi \sim \mu$ denote the respective distributions on $\mathbb{R}$. We use the notation of the preceding remark. Observe that by change of variables we have

$$E[f(Y_n(t_1), \ldots, Y_n(t_k))] = E_{s,\mu_n}[f(X(t), \ldots, X(t_k))] \qquad (5.8)$$
$$E[f(Y(t_1), \ldots, Y(t_k))] = E_{s,\mu}[f(X(t_1), \ldots, X(t_k))]. \qquad (5.9)$$

So Corollary 5.6 can be restated as

$$Q_{s,\mu_n}(X(t_1), \ldots, X(t_k))^{-1} \Rightarrow Q_{s,\mu}(X(t_1), \ldots, X(t_k))^{-1}. \qquad (5.10)$$

Another equivalent way of stating *Feller continuity* is that

$$x \mapsto E_{s,x}[f(X(t_1), \ldots, X(t_k))]$$

is continuous, for any bounded continuous function f on $\mathbb{R}^k$, $s \leq t_1 < t_2 < \ldots < t_k \leq T, k = 1, 2, \ldots$ This is the same as saying that $x \mapsto Q_{s,x}$ is continuous as a map from $\mathbb{R}$ into the space of all probability measures on Ω endowed with the topology of weak convergence.

Remark 5.9 In (5.8), (5.9) notice the difference between the l.h.s. and r.h.s. On the l.h.s. the probability measure is kept fixed while the process varies, whereas on the r.h.s. the process, viz. $\{X(t)\}$ is kept fixed while the probability measure varies. This seemingly pedantic comment is, in fact, a manifestation of a deeper aspect: All the probabilistically relevant information about the process is contained in the measure $Q_{s,\mu}$, or rather in the family $\{Q_{s,x} : s \geq 0, s \in \mathbb{R}\}$ of probability measures on Ω. This has been the thrust of Remark 5.7, and an illustration was given in Remark 5.8 in rephrasing Feller continuity. This aspect is very similar to something we have already encountered in basic probability theory, viz.: We focus our attention on the distributions of random variables rather than on the random variables themselves. This point of view has far reaching implications. □

Remark 5.10 The preceding set up, s.d.e.'s, results and remarks have extensions to higher dimensions in an obvious fashion: take $B(\cdot)$ to be a d-dimensional Brownian motion, σ to be a $(d \times d)$ matrix, $b(\cdot, \cdot)$ to be an $\mathbb{R}^d$-valued function, etc. Details are left to the reader. □

6 Markov property

Markov property is a very important characteristic of diffusions. We begin with a lemma.

Lemma 6.1 *Let $(S, \mathcal{F}, P)$ be a probability space. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be sub σ-algebras of $\mathcal{F}$ such that $\mathcal{B} \subseteq \mathcal{A}$, and $\mathcal{A}$ and $\mathcal{C}$ are independent (that is, if $A \in \mathcal{A}, C \in \mathcal{C}$ then $P(A \cap C) = P(A) \cdot P(C)$). Let $\mathcal{B} \vee \mathcal{C} = \sigma\{\mathcal{B} \cup \mathcal{C}\}$ be the smallest σ-algebra containing $\mathcal{B} \cup \mathcal{C}$. If U is a $(\mathcal{B} \vee \mathcal{C})$-measurable integrable random variable then $E(U|\mathcal{A}) = E(U|\mathcal{B})$ a.s.*

Proof: Let $\mathcal{D} = \left\{ \bigcup\limits_{i=1}^{n} (B_i \cap C_i) : B_i \in \mathcal{B}, C_i \in \mathcal{C}, 1 \leq i \leq n, n = 1, 2, \ldots \right\}$. It can be shown that $\mathcal{D}$ is an algebra, and hence $\mathcal{B} \vee \mathcal{C} = \sigma(\mathcal{D})$. So by standard measure theoretic arguments it is enough to consider $U = U_1 U_2$ where U_1 is bounded $\mathcal{B}$-measurable function and U_2 is bounded $\mathcal{C}$-measurable function. In this case note that

$$\begin{aligned} E(U|\mathcal{A}) &= E(U_1 U_2|\mathcal{A}) = U_1 E(U_2|\mathcal{A}) \\ &= U_1 E(U_2) \ a.s. \\ \text{and } E(U|\mathcal{B}) &= E(U_1 U_2|\mathcal{B}) = U_1 E(U_2|\mathcal{B}) \\ &= U_1 E(U_2) \ a.s. \end{aligned}$$

This completes the proof. □

Let $\sigma, b(\cdot, \cdot)$ be as in Theorem 5.2 / Remark 5.4. Let $s \geq 0$ be fixed. For $x \in \mathbb{R}$, let $\{Y^{(s,x)}(t) : t \geq s\}$ denote the unique strong solution such that

$$Y^{(s,x)}(t) = x + \sigma(B(t) - B(s)) + \int_s^t b(r, Y^{(s,x)}(r)) dr, t \geq s. \quad (6.1)$$

See Remark 5.1(b), Remark 5.4. Set $Y^{(s,x)}(r) = x$ for $r \leq s$. Let $\mathcal{F}_t = \sigma(\mathcal{N} \vee \mathcal{F}_t^B), t \geq s$ as the initial value is constant. Set

$$\mathcal{E}^t = \sigma\{Y^{(s,x)}(r) : r \geq t\}, t \geq s. \quad (6.2)$$

Notation: For any random variables V, ζ we shall denote $E(V|\zeta) = E(V|\sigma\{\zeta\})$.

Theorem 6.2 *(Markov property): Let the notation and hypotheses of Theorem 5.2 / Remark 5.4 hold. Let $t \geq s$. Let V be a bounded $\mathcal{E}^t$-measurable function. Then*

$$E(V|\mathcal{F}_t) = E(V|Y^{(s,x)}(t)), \ a.s. \tag{6.3}$$

Proof: For any $r \geq t$ observe that

$$\begin{aligned} Y^{(s,x)}(r) &= x + \sigma(B(r) - B(s)) + \int_t^r b(\theta, Y^{(s,x)}(\theta)) \, d\theta \\ &= Y^{(s,x)}(t) + \sigma(B(r) - B(t)) + \int_t^r b(\theta, Y^{(s,x)}(\theta)) \, d\theta. \end{aligned} \tag{6.4}$$

Strong uniqueness and (6.4) now imply (in the notation of (6.1)) that

$$Y^{(s,x)}(r) = Y^{(t,Y^{(s,x)}(t))}(r), r \geq t. \tag{6.5}$$

That is, $\{Y^{(s,x)}(r) : r \geq t\}$ is the unique strong solution to the s.d.e.

$$\left.\begin{aligned} dY(r) &= \sigma \ dB(r) + b(r, Y(r))dr, r > t \\ & \text{ with the initial value } Y(t) = Y^{(s,x)}(t) \end{aligned}\right\}. \tag{6.6}$$

In the above, it may be noted that the iteration (Picard method) used in the second proof of Theorem 5.2 implies that

$$Y^{(s,x)}(r) = Y^{(t,Y^{(s,x)}(t))}(r), r \geq t$$

is measurable w.r.t. $\sigma\{Y^{(s,x)}(t)\} \vee \sigma\{B(\theta) - B(t) : \theta \geq t\}$. See also Remark 5.1 (b). Now in the preceding lemma take $\mathcal{A} = \mathcal{F}_t, \mathcal{B} = \sigma\{Y^{(s,x)}(t)\}, \mathcal{C} = \sigma\{B(\theta) - B(t) : \theta \geq t\}$. Clearly $\mathcal{B} \subseteq \mathcal{A}$, and $\mathcal{A}$ and $\mathcal{C}$ are independent; also $\mathcal{E}^t \subseteq (\mathcal{B} \vee \mathcal{C})$. The required result (6.3) is now immediate from the preceding lemma. □

Exercise 6.3 Notation as above. For $t \geq s$ let $\mathcal{E}_t = \sigma\{Y^{(s,x)}(r) : r \leq t\}$. Show that for $t \geq s$ and any bounded $\mathcal{E}^t$-measurable function V

$$E(V|\mathcal{E}_t) = E(V|Y^{(s,x)}(t)), \ a.s. \tag{6.7}$$

Remark 6.4 Notation as above. Let $s \leq t < t_1 < t_2 < \ldots < t_k, F \in \mathcal{B}(I\!R^k), k = 1, 2, \ldots$. Then (6.3) implies (and in fact equivalent to)

$$\begin{aligned} & P((Y^{(s,x)}(t_1), \ldots, Y^{(s,x)}(t_k)) \in F | \mathcal{F}_t) \\ = \;& P((Y^{(s,x)}(t_1), \ldots, Y^{(s,x)}(t_k)) \in F | Y^{(s,x)}(t)), \; a.s. \end{aligned}$$

which may be a more familiar version of the Markov property.

Remark 6.5 Let $A \in \mathcal{F}_t, C \in \mathcal{E}^t$. Then from (6.3) we get

$$\begin{aligned} P(A \cap C | Y^{(s,x)}(t)) &= E[I_A I_C | Y^{(s,x)}(t)] \\ &= E[E(I_A I_C | \mathcal{F}_t) | Y^{(s,x)}(t)] \\ &= E[I_A E(I_C | \mathcal{F}_t) | Y^{(s,x)}(t)] \\ &= E[I_A E(I_C | Y^{(s,x)}(t)) | Y^{(s,x)}(t)] \\ &= E(I_C | Y^{(s,x)}(t)) \cdot E(I_A | Y^{(s,x)}(t)) \\ &= P(A | Y^{(s,x)}(t)) \cdot P(C | Y^{(s,x)}(t)), \; a.s. \quad (6.8) \end{aligned}$$

Here $\mathcal{F}_t$ can be regarded as the repository of all information regarding the "past", $\mathcal{E}^t$ as that of the "future" and $\sigma\{Y^{(s,x)}(t)\}$ as the "present". So (6.8) means that conditional on the "present" (that is, given the "present"), the "past" and the "future" are independent. In fact this notion is equivalent to the Markov property; that is, if equality of the extremes of (6.8) is known for all $A \in \mathcal{F}_t, C \in \mathcal{E}^t$, then by standard measure theoretic arguments (6.3) can be obtained. □

We next want to describe the Markov property of the diffusion in terms of the distribution or law of the process. We first need a lemma.

Lemma 6.6 *Let $(S, \mathcal{F}, P)$ be a probability space, and $(\hat{S}, \hat{\mathcal{F}})$ be a measurable space. Let $U : (S, \mathcal{F}) \to (\hat{S}, \hat{\mathcal{F}})$ be a measurable function; let $\hat{P} = PU^{-1}$. Let $\mathcal{A} \subset \mathcal{F}, \hat{\mathcal{A}} \subset \hat{\mathcal{F}}$ be sub σ-algebras such that $U^{-1}(\hat{\mathcal{A}}) = \mathcal{A}$. Let f be an integrable function on $(\hat{S}, \hat{\mathcal{F}}, \hat{P})$. Then*

$$\hat{E}(f | \hat{\mathcal{A}})(U(\omega)) = E(f \circ U | \mathcal{A})(\omega), \; \textit{for } P - a.a.\omega.$$

(Here $E(\cdot|\cdot), \hat{E}(\cdot|\cdot)$ denote respectively conditional expectation w.r.t. $P, \hat{P}$).

Proof: Note that any $A \in \mathcal{A}$ can be written $A = U^{-1}(\hat{A})$ for some $\hat{A} \in \hat{\mathcal{A}}$. Now for $A \in \mathcal{A}$, by change of variables and by definition of conditional expectation we get,

$$\begin{aligned}
\int_A \hat{E}(f|\hat{\mathcal{A}})(U(\omega))\, dP(\omega) &= \int_{U^{-1}(\hat{A})} \hat{E}(f|\hat{\mathcal{A}})(U(\omega))\, dP(\omega) \\
&= \int_{\hat{A}} \hat{E}(f|\hat{\mathcal{A}})(\hat{\omega})\, dPU^{-1}(\hat{\omega}) = \int_{\hat{A}} \hat{E}(f|\hat{\mathcal{A}})(\hat{\omega})\, d\hat{P}(\hat{\omega}) \\
&= \int_{\hat{A}} f(\hat{\omega})\, d\hat{P}(\hat{\omega}) = \int_{\hat{A}} f(\hat{\omega})\, dPU^{-1}(\hat{\omega}) \\
&= \int_{U^{-1}(\hat{A})} f(U(\omega))\, dP(\omega) = \int_A f(U(\omega))\, dP(\omega)
\end{aligned}$$

whence the required result follows. $\square$

Let $\Omega, X_t(\cdot) \stackrel{\Delta}{=} X(t,\cdot), \mathcal{B}_t \equiv \mathcal{B}_t^\circ, t \geq 0$ be as in Remark 5.7. For $t \geq 0$ put $\mathcal{B}^t = \sigma\{X(r) : r \geq t\}$. Let $s \geq 0, x \in I\!R$ be fixed. For $t \geq s$ let $\mathcal{E}^t$ be defined by (6.2), $\mathcal{E}_t$ as in Exercise 6.3. Define $Y^{(s,x)} : S \to \Omega$ by $Y^{(s,x)}(\omega) = Y^{(s,x)}(\cdot,\omega)$; that is, $Y^{(s,x)}(\omega)(r) = Y^{(s,x)}(r,\omega), r \geq 0, \omega \in S$; (remember $Y^{(s,x)}(r) = x$ for $r \leq s$). Note that, as indicated in Remark 5.7, $Q_{s,x} = P(Y^{(s,x)})^{-1}$ is the *distribution* (or *law*) of the diffusion process $\{Y^{(s,x)}(r)\}$.

For $t \geq s$, the following are easily verified.

$$\sigma\{Y^{(s,x)}(t)\} = (Y^{(s,x)})^{-1}(\sigma\{X(t)\}) \qquad (6.9)$$
$$\mathcal{E}_t = (Y^{(s,x)})^{-1}(\mathcal{B}_t) \qquad (6.10)$$
$$\mathcal{E}^t = (Y^{(s,x)})^{-1}(\mathcal{B}^t). \qquad (6.11)$$

Theorem 6.7 *(Markov property): Let $s \geq 0, x \in I\!R$ be fixed; let $E_{s,x}(\cdot|\cdot)$ denote conditional expectation w.r.t. $Q_{s,x}$. For $t \geq s$ and any bounded $\mathcal{B}^t$-measurable function Z on Ω, we have*

$$E_{s,x}(Z|\mathcal{B}_t) = E_{s,x}(Z|\sigma\{X(t)\}), \ a.s.\ Q_{s,x}. \qquad (6.12)$$

In other words, on the probability space $(\Omega, \mathcal{B}_\infty, Q_{s,x})$, with the filtration $\{\mathcal{B}_t\}$, the stochastic process $\{X(t) : t \geq s\}$ is a continuous Markov process.

Proof: Note that $V = Z \circ Y^{(s,x)}$ is $\mathcal{E}^t$-measurable by (6.11). Then for P-a.a.ω, we have

$$\begin{aligned}
E_{s,x}(Z|\mathcal{B}_t)(Y^{(s,x)}(\omega)) &= E(Z \circ Y^{(s,x)}|\mathcal{E}_t)(\omega) \text{ (by (6.10) and Lemma 6.6)} \\
&= E(V|\mathcal{E}_t)(\omega) = E(V|\sigma\{Y^{(s,x)}(t)\})(\omega) \text{ (by Exercise 6.3)} \\
&= E_{s,x}(Z|\sigma\{X(t)\})(Y^{(s,x)}(\omega)) \text{ (by (6.9) and Lemma 6.6)}
\end{aligned}$$

which implies (6.12) by the definition of $Q_{s,x}$. □

Note that the r.h.s. of (6.3) or (6.7) is a function of $Y^{(s,x)}(t)$, say of the form $g(Y^{(s,x)}(t))$. Our next objective is to identify the function g. As $\mathcal{E}^t$ is identified as the storehouse of information about the future, intuitively we may guess that $g(z)$ would correspond to taking expectation w.r.t. $Q_{t,z}$. We begin with a lemma.

Lemma 6.8 *Let $(S, \mathcal{F}, P)$ be a probability space and $\{B(t) : t \geq 0\}$ a standard one-dimensional Brownian motion on $(S, \mathcal{F}, P)$. Let $s \geq 0$. Let ξ be a real valued random variable independent of $\{B(t) - B(s) : t \geq s\}$. Let ξ take only the values $y_1, y_2, \ldots$ such that $P(\xi = y_i) = p_i > 0 \ \forall \ i$, and $\sum_i p_i = 1$. Denote $A_i = \xi^{-1}(\{y_i\}), i \geq 1$. For fixed i, define the probability measure P_i and the stochastic process $\{B^{(i)}(t) : t \geq s\}$ by*

$$\begin{aligned}
P_i(C) &= P(C|A_i) = \frac{P(C \cap A_i)}{P(A_i)}, C \in \mathcal{F}, \\
B^{(i)}(t, \omega) &= I_{A_i}(\omega)[B(t, \omega) - B(s, \omega)], t \geq s, \omega \in S.
\end{aligned}$$

Then on the probability space $(S, \mathcal{F}, P_i)$, the process $\{B^{(i)}(t) - B^{(i)}(s) : t \geq s\}$ is a standard one dimensional Brownian motion.

Proof: It is enough to prove that

$$\begin{aligned}
& P_i(B^{(i)}(t_1) - B^{(i)}(s) \in F_1, \ldots, B^{(i)}(t_k) - B^{(i)}(s) \in F_k) \\
= \ & P(B(t_1) - B(s) \in F_1, \ldots, B(t_k) - B(s) \in F_k)
\end{aligned}$$

for any $t_1, \ldots, t_k \in [s, \infty), F_1, \ldots, F_k \in \mathcal{B}(I\!R), k = 1, 2, \ldots$. The above equality can be proved easily using the independence of ξ and $\{B(t) - B(s) : t \geq s\}$, and the definition of P_i. □

Theorem 6.9 *Let $\sigma, b(\cdot,\cdot)$ be as in Theorem 5.2 and Remark 5.4. For $s \geq 0, x \in I\!R$ let $Q_{s,x}$ denote the distribution in Ω of the diffusion process with drift $b(\cdot,\cdot)$ and dispersion σ as indicated in Remark 5.7; (see also the discussion preceding Theorem 6.7). Let $\{B(t)\}$ be a standard one dimensional Brownian motion on a probability space $(S, \mathcal{F}, P)$. Let $s \geq 0$ be fixed, and ξ be a real valued random variable independent of $\{B(t) - B(s) : t \geq s\}$. Let $\{Y(t) : t \geq s\}$ be the unique strong solution to the s.d.e.*

$$Y(t) = \xi + \sigma(B(t) - B(s)) + \int_s^t b(r, Y(r))dr, t \geq s. \tag{6.13}$$

Then for $k = 1, 2, \ldots, t_1, t_2, \ldots, t_k \in [s, \infty), F_j \in \mathcal{B}(I\!R), 1 \leq j \leq k$ we have

$$P(Y(t_1) \in F_1, Y(t_2) \in F_2, \ldots, Y(t_k) \in F_k | \sigma\{\xi\}) = \psi(\xi) \tag{6.14}$$

where $\psi : I\!R \to I\!R$ is the measurable function given by

$$\psi(y) = Q_{s,y}(X(t_1) \in F_1, X(t_2) \in F_2, \ldots, X(t_k) \in F_k). \tag{6.15}$$

Moreover if μ is the distribution of the initial value ξ and $Q_{s,\mu}$ is the distribution in Ω of the process $\{Y(t) : t \geq s\}$, then

$$Q_{s,\mu}(H) = \int_{I\!R} Q_{s,y}(H)\ d\mu(y), H \in \mathcal{B}_\infty. \tag{6.16}$$

Proof: Step 1: We first consider the case when ξ takes only countably many values; we use the notation of the preceding lemma in this step. For fixed i, put $Y^{(i)}(t, \omega) = I_{A_i}(\omega) Y(t, \omega), t \geq s, \omega \in S$. Then using the preceding lemma it is seen that $Y^{(i)}(\cdot)$ is a strong solution to the s.d.e. (6.13) with initial value $Y^{(i)}(s) = y_i$. Since the distribution of a diffusion process is uniquely determined by Remark 5.7, it follows that $P^{(i)}(Y^{(i)})^{-1} = Q_{s,y_i}$. Consequently for $k = 1, 2, \ldots, t_k, \ldots, t_k \in [s, \infty), F_j \in \mathcal{B}(I\!R), 1 \leq j \leq k$ we get

$$\begin{aligned} & P(Y(t_1) \in F_1, \ldots, Y(t_k) \in F_k | \{\xi = y_i\}) \\ = \ & \frac{1}{P(A_i)}\ P(\{\omega \in A_i : Y(t_j, \omega) \in F_j, 1 \leq j \leq k\}) \\ = \ & P^{(i)}(Y^{(i)}(t_1) \in F_1, \ldots, Y^{(i)}(t_k) \in F_k) \\ = \ & Q_{s,y_i}(X(t_1) \in F_1, \ldots, X(t_k) \in F_k) \end{aligned} \tag{6.17}$$

proving (6.14) in this case. Measurability of (6.15) will be dealt with in the general case later; (anyway it is not a major issue here as $\{y_1, y_2, \ldots\}$ is countable). As $Q_{s,\mu} = PY^{-1} = \sum_i p_i \; P^{(i)}(Y^{(i)})^{-1}$ the last assertion is clear from (6.17) above.

Step 2: Let ξ and notation be as in Step 1. Let f, g be bounded continuous functions on $\mathbb{R}^k, \mathbb{R}$ respectively. Then by (6.17) and standard measure theoretic arguments

$$
\begin{aligned}
& \int_S f(Y(t_1,\omega),\ldots,Y(t_k,\omega)) \; g(\xi(\omega)) \; dP(\omega) \\
= & \sum_i \left[\int_{A_i} f(Y^{(i)}(t_1,\omega),\ldots,Y^{(i)}(t_k,\omega)) \; dP^{(i)}(\omega) \right] \; g(y_i) \; p_i \\
= & \sum_i \; E_{s,y_i}[f(X(t_1),\ldots,X(t_k))] \; g(y_i) \; p_i \\
= & \int_{\mathbb{R}} E_{s,y}[f(X(t_1),\ldots,X(t_k))] \; g(y) \; d\mu(y). \qquad (6.18)
\end{aligned}
$$

Though measurability in y of the integrand on r.h.s. of (6.18) is not an issue, note that it follows by Feller continuity.

Step 3: Now we consider a general ξ. For $n = 1, 2, \ldots$ let ξ_n be real valued random variable such that ξ_n takes only countably many values, $\sigma(\xi_n\} \subseteq \sigma\{\xi\}$ for each n, and $\xi_n \to \xi$ a.s.; (how can one get such $\{\xi_n\}$?). Note that each ξ_n is independent of $\sigma\{B(t) - B(s) : t \geq s\}$. Let μ_n denote the distribution of $\xi_n, n \geq 1$.

Let $\{Y^{(n)}(t) : t \geq s\}$ denote the unique strong solution to the s.d.e.

$$
Y^{(n)}(t) = \xi_n + \sigma(B(t) - B(s)) + \int_s^t b(r, Y^{(n)}(r)) \; dr, t \geq s.
$$

Let Q_{s,μ_n} denote the distribution of the process $\{Y^{(n)}(t) : t \geq s\}$.

Let $k \geq 1$; let f, g respectively be bounded continuous functions on $\mathbb{R}^k, \mathbb{R}$.

Claim: The extremes of (6.18) are equal for general ξ, Y as in our hypothesis.

Proof of claim: By Feller continuity (see Remark 5.8), $y \mapsto E_{s,y}[f(X(t_1), \ldots, X(t_k))]$ is bounded continuous and hence Borel measurable. Again by Feller continuity (see (5.10)), $Q_{s,\mu_n}(X(s), X(t_1), \ldots, X(t_k))^{-1} \Rightarrow Q_{s,\mu}(X(s), X(t_1), \ldots, X(t_k))^{-1}$, as $n \to \infty$. As (6.18) holds for $\xi_n, Y^{(n)}$ we now have by change of variables

$$\begin{aligned}
& \int_S f(Y(t_1,\omega), \ldots, Y(t_k,\omega))\ g(\xi(\omega))\ dP(\omega) \\
= & \int_\Omega f(X(t_1,w), \ldots, X(t_k,w))\ g(X(s,w))\ dQ_{s,\mu}(w) \\
= & \lim_{n\to\infty} \int_\Omega f(X(t_1,w), \ldots, X(t_k,w))\ g(X(s,w))\ dQ_{s,\mu_n}(w) \\
= & \lim_{n\to\infty} \int_S f(Y^{(n)}(t_1,\omega), \ldots, Y^{(n)}(t_k,\omega))\ g(\xi_n(\omega))\ dP(\omega) \\
= & \lim_{n\to\infty} \int_{\mathbb{R}} E_{s,y}[f(X(t_1), \ldots, X(t_k))]\ g(y)\ d\mu_n(y) \\
= & \int_{\mathbb{R}} E_{s,y}[f(X(t_1), \ldots, X(t_k))]\ g(y)\ d\mu(y) \qquad (6.19)
\end{aligned}$$

where we have used $\mu_n \Rightarrow \mu$ in the last step. This proves the claim.

Now let C, D be closed sets in $\mathbb{R}^k, \mathbb{R}$ respectively. One can choose bounded continuous functions f_m, g_m respectively on $\mathbb{R}^k, \mathbb{R}$ such that $f_m(z) \downarrow I_C(z)\ \forall$ $z \in \mathbb{R}^k$ and $g_m(y) \downarrow I_D(y)\ \forall\ y \in \mathbb{R}$. So (6.18) or (6.19) applied to f_m, g_m and the dominated convergence theorem imply that the extremes of (6.19) are equal even if $f = I_C, g = I_D$. As closed sets generate the Borel σ-algebra, by standard arguments we conclude that for any $F \in \mathcal{B}(\mathbb{R}^k), G \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}
& \int_S I_F(Y(t_1,\omega), \ldots, Y(t_k,\omega))\ I_G(\xi(\omega))\ dP(\omega) \\
= & \int_{\mathbb{R}} E_{s,y}[I_F(X(t_1), \ldots, X(t_k))]\ I_G(y)\ d\mu(y). \qquad (6.20)
\end{aligned}$$

Now (6.20), $P\xi^{-1} = \mu$ and the definition of conditional expectation imply (6.14); measurability of ψ also follows by the above arguments. As finite dimensional cylinder sets generate $\mathcal{B}_\infty$, (6.16) is an immediate consequence of (6.14), (6.15). This completes the proof. □

For $s \geq 0, x \in I\!R$ let $\{Y^{(s,x)}(t)\}$ be given by (6.1); let $Q_{s,x}$ be the distribution in Ω of the process $Y^{(s,x)}$. Let $\{X(t) : t \geq 0\}$ denote the process defined on Ω by the coordinate projections. Define the *transition probability function* by

$$\begin{aligned} P(s,x;t,A) &\stackrel{\Delta}{=} Q_{s,x}(X(t) \in A) \\ &= P(Y^{(s,x)}(t) \in A), 0 \leq s \leq t, x \in I\!R, A \in \mathcal{B}(I\!R). \end{aligned} \quad (6.21)$$

Note that $(s,x,t) \mapsto P(s,x;t,A)$ is a bounded measurable function for fixed $A \in \mathcal{B}(I\!R)$, and $A \mapsto P(s,x;t,A)$ is a probability measure on $\mathcal{B}(I\!R)$ for fixed s,x,t. The transition probability $P(s,x;t,A)$ is clearly the probability of finding the diffusion process in the set A at time t given that it started from x at time s. By Theorems 6.2 and 6.9, for $s \leq t \leq r, x \in I\!R, A \in \mathcal{B}(I\!R)$

$$\begin{aligned} P(Y^{(s,x)}(r) \in A | \mathcal{F}_t) &= P(Y^{(s,x)}(r) \in A | \sigma\{Y^{(s,x)}(t)\}) \\ &= P(t, Y^{(s,x)}(t); r, A), \ a.s. \end{aligned} \quad (6.22)$$

or equivalently on the probability space $(\Omega, \mathcal{B}_\infty, Q_{s,x})$

$$\begin{aligned} Q_{s,x}(X(r) \in A | \mathcal{B}_t) &= Q_{s,x}(X(r) \in A | \sigma\{X(t)\}) \\ &= P(t, X(t); r, A), \ a.s. \ Q_{s,x}. \end{aligned} \quad (6.23)$$

In the above $P(t, X(t); r, A)$ or $Q_{t,X(t)}(X(r) \in A)$ is a notation for $P(t,y;r,A)$ or $Q_{t,y}(X(r) \in A)$ respectively evalued at $y = X(t)$.

Proposition 6.10 The transition probability function satisfies the *Chapman-Kolmogorov equation*, viz. for $s \leq t \leq r, x \in I\!R, A \in \mathcal{B}(I\!R)$,

$$P(s,x;r,A) = \int_{I\!R} P(t,y;r,A) \ P(s,x;t,dy). \quad (6.24)$$

Proof:

$$\begin{aligned} P(s,x;r,A) &= Q_{s,x}(X(r) \in A) = E_{s,x}[I_A(X(r))] \\ &= E_{s,x}[E_{s,x}[I_A(X(r)) | \sigma\{X(t)\}]] \\ &= E_{s,x}[P(t, X(t); r, A)] \ \ \text{(by 6.23)} \\ &= \int_\Omega P(t, X(t,w); r, A) \ dQ_{s,x}(w) \end{aligned}$$

$$= \int_{\mathbb{R}} P(t, y; r, A)\ dQ_{s,x}(X(t))^{-1}(y) \text{ (by change of variables)}$$

$$= \int_{\mathbb{R}} P(t, y; r, A)\ P(s, x; t, dy)$$

as desired. □

Exercise 6.11 Using Theorems 6.7, 6.9 and standard measure theoretic arguments, for any $0 \le s \le t, x \in \mathbb{R}$ and any bounded $\mathcal{B}^t$-measurable random variable Z show that

$$E_{s,x}[Z|\mathcal{B}_t] = E_{s,x}[Z|X(t)] = \zeta(X(t)),\ a.s.\ Q_{s,x} \qquad (6.25)$$

where $\zeta : \mathbb{R} \to \mathbb{R}$ is the measurable function given by $\zeta(y) = E_{t,y}(Z)$. It is customary to write (6.25) as

$$E_{s,x}[Z|\mathcal{B}_t] = E_{s,x}[Z|X(t)] = E_{t,X(t)}(Z),\ a.s.\ Q_{s,x}.$$

7 Strong Markov property

To elucidate the strong Markov property, it is better to work with $\Omega, \{X(t)\}, \{\mathcal{B}_t : t \ge 0\}, \{Q_{s,x} : s \ge 0, x \in \mathbb{R}\}$.

Recall that a measurable function $\tau : \Omega \to [0, \infty) \cup \{+\infty\}$ is a *stopping time* w.r.t. $\{\mathcal{B}_t\}$ if $\{\tau \le t\} \in \mathcal{B}_t$ for all $t \ge 0$. The associated σ-algebra $\mathcal{B}_\tau$ is defined by

$$\mathcal{B}_\tau = \{A \in \mathcal{B}_\infty : A \cap \{\tau \le t\} \in \mathcal{B}_t \text{ for all } t \ge 0\}. \qquad (7.1)$$

It can be shown that $\mathcal{B}_\tau$ is a σ-algebra; also $\tau, X_\tau(\cdot) = X(\tau, \cdot), X_{\tau\wedge t}(\cdot), t \ge 0$ are all $\mathcal{B}_\tau$-measurable. It can in fact be shown that

$$\mathcal{B}_\tau = \sigma\{X(\tau \wedge t) : t \ge 0\}. \qquad (7.2)$$

Because of (7.2) the σ-algebra $\mathcal{B}_\tau$ can be taken as the repository of all information about the process upto the stopping time τ.

Remark 7.1 (i) As sample paths of $X(\cdot)$ are continuous for any closed set F, τ_F defined by

$$\begin{aligned} \tau_F(w) &= \inf\{t \geq 0 : X(t,w) \in F\} \\ &= \text{ first hitting time of } F \end{aligned}$$

is a stopping time w.r.t. $\{\mathcal{B}_t\}$.

(ii) If η, τ are stopping times w.r.t. $\{\mathcal{B}_t\}$ and $\eta \leq \tau$ then $\mathcal{B}_\eta \subseteq \mathcal{B}_\tau$, and hence η is also $\mathcal{B}_\tau$-measurable. $\square$

For proofs of the various assertions above, see the first chapters of [KS],[SV].

Roughly speaking, strong Markov property is the Markov property where deterministic time t can be replaced by stopping time τ. It is clear from the above that $\mathcal{B}_\tau$ is the analogue of $\mathcal{B}_t$ representing the "past". What are the analogues of $\mathcal{B}^t, \sigma\{X(t)\}$ standing respectively for the "future", "present"? Again (7.2) gives us the clue. Define

$$\mathcal{B}^\tau = \sigma\{\tau + r, X(\tau + r) : r \geq 0\}. \tag{7.3}$$

Clearly $\mathcal{B}^\tau$ contains all information about the process after time τ. It is also now apparent that $\sigma\{\tau, X(\tau)\}$ is the "present". (It may be kept in mind that $X_\tau(w) \equiv X(\tau)(w) = X(\tau(w), w)$.)

Theorem 7.2 *(Strong Markov property): Let $\sigma, b(\cdot,\cdot)$ be as in Sections 5 and 6. Let $Q_{s,x}$ be the distribution of the diffusion process with drift $b(\cdot,\cdot)$ and dispersion σ, starting at (s,x). Let $s \geq 0$ be fixed; let τ be a stopping time w.r.t. $\{\mathcal{B}_t\}$ such that $\tau \geq s$. For any bounded $\mathcal{B}^\tau$-measurable function Z, any $x \in \mathbb{R}$*

$$\begin{aligned} E_{s,x}[I_{\{\tau<\infty\}} Z | \mathcal{B}_\tau] &= E_{s,x}[I_{\{\tau<\infty\}} \cdot Z | \sigma\{\tau, X(\tau)\}] \\ &= I_{\{\tau<\infty\}} \cdot \varphi(\tau, X(\tau)), \ \textit{a.s. } Q_{s,x} \end{aligned} \tag{7.4}$$

where φ is the measurable function given by $\varphi(t,y) = E_{t,y}(Z)$. Equation (7.4) is also sometimes written informally as

$$E_{s,x}[I_{\{\tau<\infty\}} Z | \mathcal{B}_\tau] = I_{\{\tau<\infty\}} \cdot E_{\tau, X(\tau)}(Z). \tag{7.5}$$

Proof: Step 1: We may take $s = 0$. We may also assume that $\tau < \infty$ a.s. For $n = 1, 2, \ldots$ define

$$\tau_n(w) = \frac{k+1}{2^n}, \text{ if } \frac{k}{2^n} \leq \tau(w) < \frac{k+1}{2^n}, k = 0, 1, 2, \ldots$$

Then τ_n's are stopping times w.r.t. $\{\mathcal{B}_t\}$; also $\tau_n(w) \downarrow \tau(w)$ for all w, and $\{\tau_n = \frac{k}{2^n}\} \in \mathcal{B}_{k/2^n}$, for $n, k = 1, 2, \ldots$.

Step 2: We shall denote $E[h : C] = \int_C h \, dP$. Now let $t \geq 0$ and f be a bounded continuous function on $[0, \infty) \times I\!R$. Let $A \in \mathcal{B}_\tau$. Observe that $A \cap \{\tau_n = \frac{k}{2^n}\} \in \mathcal{B}_{\frac{k}{2^n}}$. Now by the Markov property (Theorem 6.7) and the dominated convergence theorem

$$\begin{aligned}
& E_{0,x}[f(\tau + t, X(\tau + t)) : A] \\
= \; & \lim_{n\to\infty} E_{0,x}[f(\tau_n + t, X(\tau_n + t)) : A] \\
= \; & \lim_{n\to\infty} \sum_{k=0}^{\infty} E_{0,x}[f(\tau_n + t, X(\tau_n + t)) : A \cap \{\tau_n = \frac{k}{2^n}\}] \\
= \; & \lim_{n\to\infty} \sum_{k=0}^{\infty} E_{0,x}[f(\frac{k}{2^n} + t, X(\frac{k}{2^n} + t)) : A \cap \{\tau_n = \frac{k}{2^n}\}] \\
= \; & \lim_{n\to\infty} \sum_{k=0}^{\infty} E_{0,x}[E_{\frac{k}{2^n}, X(\frac{k}{2^n})}(f(\frac{k}{2^n} + t, X(\frac{k}{2^n} + t))) : A \cap \{\tau_n = \frac{k}{2^n}\}] \\
= \; & \lim_{n\to\infty} E_{0,x}[E_{\tau_n, X(\tau_n)}(f(\tau_n + t, X(\tau_n + t))) : A]. \qquad (7.6)
\end{aligned}$$

It can be proved that the function $(r, z, t) \mapsto E_{r,z}[f(r + t, X(r + t))]$ is a bounded continuous function on $\{r \geq 0, t \geq 0, z \in I\!R\}$; this is an extension of Feller continuity. So

$$E_{\tau_n, X(\tau_n)}(f(\tau_n + t, X(\tau_n + t))) \to E_{\tau, X(\tau)}(f(\tau + t, X(\tau + t))) \; a.s.$$

So by the dominated convergence theorem and (7.6) we get

$$E_{0,x}[f(\tau + t, X(\tau + t)) : A] = E_{0,x}[E_{\tau, X(\tau)}(f(\tau + t, X(\tau + t))) : A].$$

Thus for any bounded continuous function f

$$E_{0,x}[f(\tau + t, X(\tau + t)) | \mathcal{B}_\tau] = E_{\tau, X(\tau)}[f(\tau + t, X(\tau + t))], \; a.s. \qquad (7.7)$$

Step 3: Let $G \subseteq [0, \infty) \times I\!R$ be an open set. Choose a sequence $\{f_n\}$ of bounded continuous functions such that $f_n \uparrow I_G$ pointwise. Now (7.7) applied to f_n, and the monotone convergence theorem imply that (7.7) holds for $f = I_G$ and that $(r, z, t) \mapsto E_{r,z}(I_G(r + t, X(r + t)))$ is measurable. Standard arguments now give that (7.7) holds for any bounded measurable function f, along with the required measurability.

Step 4: To complete the proof of the equality of the first and third expressions in (7.4), note that it is enough to consider Z of the form $Z = \prod_{j=1}^{k} f_j(\tau + t_j, X(\tau + t_j))$, with $0 \leq t_1 < \ldots < t_k, f_1, \ldots, f_k$ being bounded Borel measurable functions on $[0, \infty) \times I\!\!R$. For this we apply induction on k. The case $k = 1$ has been taken care of in Step 3. Put $U = \prod_{j=1}^{k-1} f_j(\tau + t_j, X(\tau + t_j))$; note that U is $\mathcal{B}_{\tau+t_{k-1}}$-measurable. Note that

$$\begin{aligned} & E_{0,x}(Z|\mathcal{B}_\tau) = E_{0,x}(E_{0,x}(Z|\mathcal{B}_{\tau+t_{k-1}})|\mathcal{B}_\tau) \\ = \; & E_{0,x}[U \cdot E_{0,x}(f_k(\tau + t_k, X(\tau + t_k))|\mathcal{B}_{\tau+t_{k-1}})|\mathcal{B}_\tau] \\ = \; & E_{0,x}[U\{E_{\tau+t_{k-1},X(\tau+t_{k-1})}[f_k(\tau + t_k, X(\tau + t_k))]\}|\mathcal{B}_\tau] \qquad (7.8) \end{aligned}$$

by induction hypothesis corresponding to $k = 1$ applied to the stopping time $\tau + t_{k-1}$. Now set $g_j = f_j$ for $j = 1, 2, \ldots, k-2$ and

$$g_{k-1}(t, y) = E_{t,y}[f_k(\tau + t_k, X(\tau + t_k)] \cdot f_{k-1}(t, y).$$

Put $\hat{Z} = \prod_{j=1}^{k-1} g_j(\tau + t_j, X(\tau + t_j))$. Note that by Step 3, g_{k-1} is bounded measurable on $[0, \infty) \times I\!\!R$, and hence $\hat{Z}$ is a bounded $\mathcal{B}^\tau$-measurable function. So induction hypothesis corresponding to $(k-1)$ applied to $\hat{Z}$ now gives from (7.8),

$$\begin{aligned} E_{0,x}(Z|\mathcal{B}_\tau) &= E_{0,x}[\hat{Z}|\mathcal{B}_\tau] = E_{\tau,X(\tau)}(\hat{Z}) \\ &= E_{\tau,X(\tau)}[U \cdot E_{\tau+t_{k-1},X(\tau+t_{k-1})}(f_k(\tau + t_k, X(\tau + t_k)))] \\ &= E_{\tau,X(\tau)}[U \cdot E_{\tau,X(\tau)}[f_k(\tau + t_k, X(\tau + t_k))|\mathcal{B}_{\tau+t_{k-1}}]] \\ &= E_{\tau,X(\tau)}(Z) \qquad (7.9) \end{aligned}$$

where in the penultimate step of (7.9) we have used induction hypothesis corresponding to $k = 1$. Thus the first and the third expressions in (7.4) are equal; the same arguments also give measurability of φ.

Step 5: As φ is measurable, clearly $\varphi(\tau, X(\tau))$ is $\sigma\{\tau, X(\tau)\}$-measurable; also $I_{\{\tau<\infty\}}$ is $\sigma\{\tau, X(\tau)\}$-measurable. So the second and the third expressions in (7.4) are equal. This completes the proof. □

Remark 7.3 *Time homogeneous case.* Suppose $b : I\!R \to I\!R$ is a function independent of the time variable, satisfying the hypotheses of Section 5; of course σ is a constant. The solution to the s.d.e. in such a case is called a *time homogeneous diffusion.*

We know that $\{B(t'+h) - B(s'+h) : s' \leq t'\}$ has the same distribution as $\{B(t') - B(s') : s' \leq t'\}$ for every $h \geq 0$. So using the iteration procedure as in the second proof of Theorem 5.2, it can be seen that $\{Y^{(s,x)}(t) : t \geq s\}$ has the same distribution as $\{Y^{(0,x)}(t) : t \geq 0\}$. Hence $Q_{s,x} = Q_{0,x}$, and we denote the measure by Q_x; expectation w.r.t. Q_x is denoted $E_x(\cdot)$.

It can be easily seen that $P(s+h, x; t+h, A) = P(s, x; t, A)$ for all $h \geq 0$. In other words $P(s, x; t, A)$ depends only on $(t-s), x, A$.

To illustrate how the strong Markov property can be stated in the time homogeneous case let $\tau \geq 0$ be a stopping time w.r.t. $\{\mathcal{B}_t\}$. Define the *shift* $\theta_\tau : \Omega \to \Omega$ by $(\theta_\tau w)(t) = w(\tau(w) + t)$. We claim that θ_τ is $\mathcal{B}^\tau$-measurable. Indeed, let $C = \{w : X(t_j, w) = w(t_j) \in A_j, 1 \leq j \leq k\}$ where $0 \leq t_1 < t_2 < \ldots < t_k < \infty, A_i \in \mathcal{B}(I\!R)$. Then $\theta_\tau^{-1}(C) = \{w : X(\tau(w) + t_j, w) \in A_j, 1 \leq j \leq k\}$ and hence $\theta_\tau^{-1}(C) \in \mathcal{B}^\tau$. Thus $\theta_\tau^{-1}(\mathcal{B}_\infty) \subseteq \mathcal{B}^\tau$, proving the claim.

Consequently if Z is any $\mathcal{B}_\infty$-measurable function then $Z \circ \theta_\tau$ is $\mathcal{B}^\tau$-measurable. So strong Markov property can be stated as

Proposition 7.4 Let $\sigma, b(\cdot), \{Q_x\}$ be as above. Let τ be a stopping time w.r.t. $\{\mathcal{B}_t\}$. Then for any bounded Borel measurable function Z on Ω, and any $x \in I\!R$,

$$E_x[I_{\{\tau<\infty\}} \cdot (Z \circ \theta_\tau) | \mathcal{B}_\tau] = I_{\{\tau<\infty\}} \cdot \varphi(X(\tau)), \ a.s.\ Q_x \qquad (\ 7.10\ )$$

where $\varphi(y) = E_y(Z)$. $\square$

Remark 7.5 The results of Sections 6 and 7 can be extended to higher space dimensions without much difficulty.

Diffusion processes constitute one of the most extensively studied classes of continuous time Markov processes. In our discussion on Markov and strong

Markov property we have not given formal definitions from the theory of Markov processes. But should any reader subsequently venture into a study of Markov processes, familiarity with a rich class of examples would give a better insight.

References

[KS] I. Karatzas and S.E. Shreve: *Brownian motion and stochastic calculus.* Springer, New York, 1991.

[O] B. Oksendal: *Stochastic differential equations*, sixth edition. Springer International Edition, Springer, New Delhi, 2003.

[SV] D.W. Stroock and S.R.S. Varadhan: *Multidimensional diffusion processes*, Springer, New York, 1979.

[V] S.R.S. Varadhan: *Lectures on Diffusion problems and Partial Differential Equations.* TIFR Lecture Notes, Bombay, 1977.