MA8109

Informal Comments on Itô Diffusions

Harald E. Krogstad, IMF

2007/rev. 2009/2011

1 Introduction

The introductory Chapter 7 on diffusion in B.Ø. is hard, partly because of a somewhat detailed (but correct, of course) notation. The results are also given in a general form, and sometimes perhaps too general for most applications.

Diffusion is originally the chaotic way molecules of fluids are mixed due to their motion. In fact, physical Brownian motion is a visible example of diffusion (spreading) of tiny particles in a solution. Over time, diffusion has been proposed as a useful way of modelling a lot of physical phenomena, like heat transfer and spreading of bacteria, deceases, rumors, and even intelligent life in the universe. Diffusion is material transport induced by differences in concentration, leading to a special type of partial differential equations called *diffusion equations*. The simplest (and by far the most studied) diffusion equation has the form

$$\frac{\partial u}{\partial t} = \nabla^2 u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}, \tag{1}$$

where, *e.g.* $u(\mathbf{x}, t)$ is the concentration of some material spreading out in a fluid.

Brownian motion is, as visualized by numerical simulations, also a model of how particles spread out from a location, say $\mathbf{x} = \mathbf{0}$, as time increases. After a certain time t, a number of particles, starting at $\mathbf{x} = \mathbf{0}$ for $t = 0$ have spread into a cloud with density proportional to the probability density of B_t, that is,

$$p(\mathbf{x}, t) = \frac{1}{(2\pi t)^{n/2}} \exp\left(-\frac{|\mathbf{x}|^2}{2t}\right). \tag{2}$$

The distribution broadens as t increases, and tends, for $t \to 0$, to what is called a δ-pulse at $\mathbf{x} = \mathbf{0}$. Moreover, the function in Eqn. 2 is, for $t > 0$, a solution to the equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\nabla^2 p \tag{3}$$

(verification left as an exercise to the reader!).

Thus, we observe an interesting connection between the solutions of the simple stochastic differential equation

$$dX_t = dB_t \tag{4}$$

and the partial differential Eqn. 3: If we are only interested in the probability law of Eqn. 4, we could as well solve Eqn. 3. In fact, as will be more clear below, the solution of Eqn. 3 defines a stochastic process which turns out to be Brownian motion! This connection represents the core of Chapters 7 and 8 in Øksendal, but we first need a few comments about what is called *Markov processes.*

2 Markov Processes

A *Markov process* is a stochastic process which generalizes the perhaps more familiar concept of a *Markov chain.* The Markov chain is a discrete Markov process, and both the definition and the general properties of Markov processes resemble the discrete case. Below we summarize some of the more basic properties of Markov processes.

There are various (equivalent) ways of defining a Markov process, and for us the following definition appears to be most suitable. Let X_t be a stochastic process on $[0, \infty)$ and $\mathcal{M}_t$ be the corresponding filtration, that is, $\mathcal{M}_t$ is the σ-algebra generated by $\{X_s\}$, $s \leq t$.

A stochastic process is a Markov process if for all $0 \leq s \leq t$,

$$\mathsf{E}\left(X_t | \mathcal{M}_s\right) = \mathsf{E}\left(X_t | X_s\right). \tag{5}$$

Here, $\mathsf{E}\left(X_t | X_s\right)$ is just a short way of writing expectation conditional the σ-algebra generated X_s, namely $\mathsf{E}\left(X_t | \mathcal{F}_{X_s}\right)$.

It is possible to prove that the definition implies that for all bounded Borel functions f and $0 \leq s \leq t$,

$$\mathsf{E}\left(f\left(X_t\right) | \mathcal{M}_s\right) = \mathsf{E}\left(f\left(X_t\right) | X_s\right), \tag{6}$$

and sometimes this is seen as the definition of a Markov process. In fact, the following 3 different definitions are equivalent ($0 \leq s \leq t$):

1. $P\left(X_t \in B | \mathcal{M}_s\right) = P\left(X_t \in B | \mathcal{F}_{X_s}\right)$ for all $B \in \mathcal{B}$,
2. $\mathsf{E}\left(X_t | \mathcal{M}_s\right) = \mathsf{E}\left(X_t | X_s\right)$ $\left(= \mathsf{E}\left(X_t | \mathcal{F}_{X_s}\right)\right)$,
3. If f is any bounded *Borel function*, $\mathsf{E}\left(f\left(X_t\right) | \mathcal{M}_s\right) = \mathsf{E}\left(f\left(X_t\right) | \mathcal{F}_{X_s}\right)$.

Digression: For a *Borel function* f, the inverse sets, $f^{-1}(B)$, of Borel sets, $B \in \mathcal{B}$, are Borel sets. Typically for us, a Borel function is a mapping from $\mathbb{R}^m$ to $\mathbb{R}^n$ for some m and n. Therefore, a mapping h, consisting of a measurable function g composed with a Borel function,

$$h(x) = (f \circ g)(x) = f(g(x)),$$

will also be measurable (Recall the definition of a measurable function, where the inverse image of all Borel sets are measurable sets).

The first thing to check out is whether there really exist Markov processes, and Brownian motion itself (with $\mathcal{M}_t = \mathcal{F}_t$) is an obvious example:

$$\begin{aligned} \mathsf{E}\left(B_t | \mathcal{F}_s\right) &= B_s \\ &= \mathsf{E}\left(B_t - B_s | B_s\right) + \mathsf{E}\left(B_s | B_s\right) \\ &= \mathsf{E}\left(B_t | B_s\right). \end{aligned} \tag{7}$$

(We already know that B_t is a Martingale, $\mathsf{E}\left(B_s | B_s\right) = B_s$, and that $\mathsf{E}\left(B_t - B_s | B_s\right) = 0$).

More generally, Markov processes are constructed by means of so-called *transition functions* (generalizing the transition matrix of the Markov chain). In simple terms, the transition

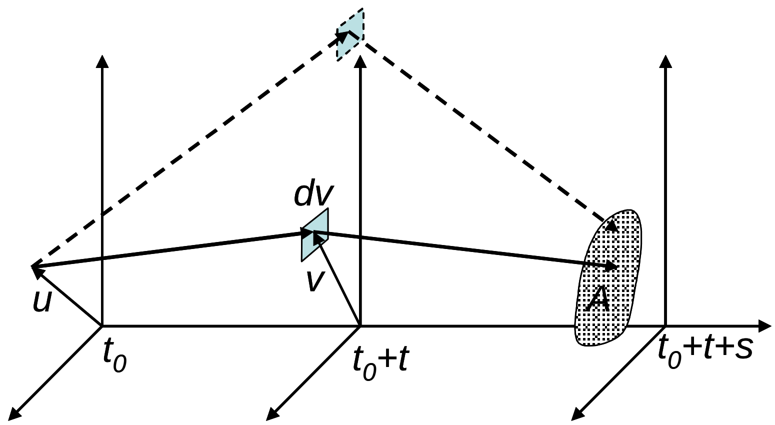

Figure 1: Illustration of the Chapman-Kolmogorov equation.

function, $p\left(t,u,A\right)$, gives the probability that $X_t \in A$, assuming that $X_0 = u$. Formally, a (time homogeneous) transition function for a Markov process, where $t \in [0,\infty)$, $u \in \mathbb{R}^n$, and $A \in \mathcal{B}\left(\mathbb{R}^n\right)$, fulfils

1. $p\left(t,\cdot,A\right)$ is a Borel function from $\mathbb{R}^n \rightarrow \mathbb{R}$,
2. $p\left(t,u,\cdot\right)$ is a probability measure on $\mathbb{R}^n$,
3. $p\left(t,u,A\right)$ satisfies the *Chapman-Kolmogorov equation* for all positive t and s:

$$p\left(t+s,u,A\right) = \int_{\mathbb{R}^n} p\left(s,v,A\right) p\left(t,u,dv\right) \tag{8}$$

Note that the integral is a regular Lebesgue integral of the form $\int_{\mathbb{R}^n} h d\mu$, where $h\left(v\right) = p\left(s,v,A\right)$ is a Borel function on $\mathbb{R}^n$, and $d\mu$ is the probability measure defined by $\mu\left(B\right) = p\left(t,u,B\right)$. In essence, the Chapman-Kolmogorov equation says that we go from location u at $t = t_0$ into the set A at a later time t_0+t+s by passing through dv at time t_0+t, and then continuing on to A. For each separate step, we apply the transition function. The integral represents all possible ways this may be done by integrating over all $v \in \mathbb{R}^n$, and since passing through different points at this time are exclusive events, the integral in Eqn. 8 must be valid. The argument is visualized in Fig. 1.

Note that even if it is possible to define transition functions for any process, it is the Chapman-Kolmogorov equation that is the key for obtaining Markov processes. Referring again to Fig. 1, for a general process the transition function $p\left(s,v,A\right)$ will also depend on what happened prior to t_0+t, *e.g.* the position u.

From the transition functions, the existence of the corresponding stochastic process follows from Kolmogorov's Existence Theorem (B.Ø., Thm. 2.1.5), where the finite dimensional distributions are defined by

$$\begin{aligned} &\nu_{t_1,t_2,\cdots,t_k}\left(F_1 \times F_2 \times \cdots \times F_k\right) \\ &= \int_{F_1\times\cdots\times F_k} p\left(t_1,x,dx_1\right) p\left(t_2-t_1,x_1,dx_2\right) \cdots p\left(t_k-t_{k-1},x_{k-1},dx_k\right). \end{aligned} \tag{9}$$

These measures, which depend on x, define the finite dimensional probability measures for a (Markov) process X_t starting at x, i.e. the probability of the simultaneous events

$$P\left(X_{t_1} \in F_1, X_{t_2} \in F_2, \cdots, X_{t_k} \in F_k\right). \tag{10}$$

In many practical situations, the transition function may be written in term of a probability density $\phi\left(t,u,v\right)$ so that

$$p\left(t,u,A\right) = \int_A \phi\left(t,u,v\right) dv, \tag{11}$$

In this case, the Chapman-Kolmogorov equation takes the form

$$\begin{aligned} p\left(t+s,u,A\right) &= \int_{\mathbb{R}^n} \phi\left(t,u,v\right) dv \int_A \phi\left(s,v,w\right) dw \\ &= \int_A \left(\int_{\mathbb{R}^n} \phi\left(t,u,v\right)\phi\left(s,v,w\right) dv\right) dw \\ &= \int_A \phi\left(t+s,u,w\right) dw. \end{aligned} \tag{12}$$

Hence,

$$\phi\left(t+s,u,w\right) = \int_{\mathbb{R}^n} \phi\left(t,u,v\right)\phi\left(s,v,w\right) dv. \tag{13}$$

The transition function for the standard Brownian motion is of this form:

$$p\left(t,u,A\right) = \int_A \frac{1}{\left(2\pi t\right)^{n/2}} e^{-\left|u-v\right|^2/2t} dv, \tag{14}$$

and then Eqn. 13 is a *convolution* integral (proof left to the reader!). See B.Ø., Sec. 2.2 how this is utilized for constructing the Brownian motion by first stating the finite dimensional simultaneous distributions, and then using Kolmogorov's Existence Theorem.

3 The Markov Property of Itô Diffusion

A stochastic differential equation

$$dX_t = b\left(t,X_t\right) dt + \sigma\left(t,X_t\right) dB_t \tag{15}$$

is also called an *Itô diffusion*. Note that the functions b and σ are dependent only on X_t and t, and not on anything happening earlier than time t. In this chapter, we always assume that b and σ satisfy the conditions stated in Thm. 5.2.1 (for existence and uniqueness of the solution). If an explicit dependence on t is missing in b and σ, the diffusion is called *autonome* or *time homogeneous*. This is similar to what we say about an ordinary differential equation where the right hand side does not depend explicitly on t. For an autonome equation there is no preferred origin of time. All diffusions below are autonome,

$$dX_t = b\left(X_t\right) dt + \sigma\left(X_t\right) dB_t. \tag{16}$$

Not all Itô processes are diffusions: Since it is only required that b and σ should be adapted to the filtration $\mathcal{F}_t$, b and σ may well depend on something happening before time t, e.g. $dX_t = B_{t/2} dB_t$.

Recall that we say that two processes X_t and Y_t are *equivalent* (as stochastic processes) if $X_0 = Y_0 = Z$, and all finite dimensional distributions are equal. For Brownian motion, $(B_t,\ B_0 = 0)$ and $\tilde{B}_t = B_{t+s} - B_s$ are equivalent.

We follow the notation in B.Ø. and let

$$X_t^{s,x} \tag{17}$$

denote the solution of Eqn. 16 at time t, given the starting time s and the starting location x. Occasionally, if $s = 0$, we may write

$$X_t^x = X_t^{s,x}, \tag{18}$$

and $X_t = X_t^0$.

As discussed in B.Ø., the solutions of an autonome Itô diffusion are time invariant in the sense that $Z_t = X_t^{0,x}$ and $Y_t = X_{s+t}^{s,x}$ are equivalent: Both start at x at $t = 0$, and all their (finite dimensional) probability measures will be the same.

Since we may write

$$\begin{aligned}
X_t^x &= \int_0^t \left[b\left(X_u^x\right) du + \sigma\left(X_u^x\right) dB_u \right] \\
&= \int_0^s \left[b\left(X_u^x\right) du + \sigma\left(X_u^x\right) dB_u \right] + \int_s^t \left[b\left(X_u^x\right) du + \sigma\left(X_u^x\right) dB_u \right] \\
&= X_s^x + \int_s^t \left[b\left(X_u^x\right) du + \sigma\left(X_u^x\right) dB_u \right],
\end{aligned} \tag{19}$$

it is intuitively clear that the last integral depends on X_s^x and the Brownian motion from s to t, but is independent of anything happening before time s. Therefore, we would expect that an Itô diffusion is a Markov process. However, a complete proof of, say,

$$\mathsf{E}\left(X_t^x | \mathcal{M}_s\right) = \mathsf{E}\left(X_t^x | X_s^x\right), \tag{20}$$

seems to be cumbersome (There are alternate routes to this result, requiring theory we have not covered).

As far as I have been able to sort out, a simple proof only exists for diffusions without drift,

$$X_t^x = \int_0^t \sigma\left(X_u^x\right) dB_u. \tag{21}$$

We may then utilize two important properties of the Itô integral. First of all, X_t^x is $\mathcal{F}_t$ adapted (True here because of the Itô integral, and, in fact, true for any diffusion by Theorem 5.2.1). Moreover, the Itô integral is a martingale with respect to $\mathcal{F}_t$. It then follows, since all sets in $\mathcal{M}_t$ is also in $\mathcal{F}_t$, that $\mathcal{M}_t \subset \mathcal{F}_t$. Together with the *telescope rule* for conditional expectation

(Theorem B.3 in B.Ø.), the argument goes as follows ($s < t$):

$$\begin{aligned} \mathsf{E}\left(X_t^x|\mathcal{M}_s\right) &= \mathsf{E}\left(\mathsf{E}\left(X_t^x|\mathcal{F}_s\right)|\mathcal{M}_s\right) \\ &= \mathsf{E}\left(X_s^x|\mathcal{M}_s\right) \\ &= X_s^x \\ &= \mathsf{E}\left(X_s^x|\mathcal{F}_{X_s^x}\right) \\ &= \mathsf{E}\left(\mathsf{E}\left(X_t^x|\mathcal{F}_s\right)|\mathcal{F}_{X_s^x}\right) \\ &= \mathsf{E}\left(X_t^x|\mathcal{F}_{X_s^x}\right). \end{aligned} \tag{22}$$

Theorem 7.1.2 in B.Ø. is the basic Markov property of the Itô diffusion. The proof of the theorem looks to be more cumbersome than necessary, but, as one could expect, it has not been possible to simplify it significantly.

We may state the theorem by first introducing the function

$$G_h(y) = \mathsf{E}\left(f\left(X_h^y\right)\right), \tag{23}$$

where f is a bounded *Borel function*. The stochastic variable $\omega \to X_h^y(\omega)$ has a probability distribution $\mu_{y,h}$, and

$$G_h(y) = \mathsf{E}\left(f\left(X_h^y\right)\right) = \int_{\mathbb{R}^n} f(x)\, d\mu_{y,h}(x). \tag{24}$$

Note that the composite function $f \circ X_h^y$ is $\mathcal{M}_{X_h^y}$-measurable, as well as $\mathcal{F}_h$-measurable, since f is a Borel function (comments above).

It is proved later in B.Ø. (Lemma 8.1.4) that G_h will be continuous if f is continuous.

Theorem 7.1.2 can now be stated as follows:

$$\mathsf{E}\left(f\left(X_{t+h}\right)|\mathcal{F}_t\right) = G_h\left(X_t\right). \tag{25}$$

Both sides are stochastic variables, and the expression should be read

$$\mathsf{E}\left(f\left(X_{t+h}\right)|\mathcal{F}_t\right)(\omega) = G_h\left(X_t(\omega)\right), \omega \in \Omega. \tag{26}$$

We thus have an explicit expression for the conditional expectation on the left side if we know the function G_h and X_t. As shown in Eqn. 24, G_h is may be computed from the distribution of X_h^y.

Proof of Theorem 7.1.2

Let $r > t$ and consider $X_r^{t,x}(\omega)$. For a fixed x,

$$X_r^{t,x} = x + \int_t^r b\left(X_u\right) du + \int_t^r \sigma\left(X_u\right) dB_u \tag{27}$$

is independent of $\mathcal{F}_t$ since nothing on the right hand side depends on anything happening before or at time t (NB! In the Itô integral we only use differences of the Brownian motion, and $B_{t+\delta} - B_t$ is independent of B_t for all $\delta > 0$). Note also that we, according Eqn. 19 (and the uniqueness of the solution of the Itô diffusion, Thm. 5.2.1), may write

$$X_{t+h}(\omega) = X_{t+h}^{t,X_t(\omega)}(\omega). \tag{28}$$

Fix t and h and let

$$g(y,\omega) = f\left(X_{t+h}^{t,y}(\omega)\right). \tag{29}$$

Referring to the somewhat technical Exercise 7.6, B.Ø. shows that it is possible to approximate $g(y,\omega)$ by a sum

$$g(y,\omega) \backsim \sum_k \phi_k(y)\,\psi_k(\omega). \tag{30}$$

Following the rules of conditional expectation,

$$\begin{aligned}\mathsf{E}\left(\phi_k(X_t)\,\psi_k(\omega)\,|\mathcal{F}_t\right) &= \phi_k(X_t)\,\mathsf{E}\left(\psi_k(\omega)\,|\mathcal{F}_t\right)\\ &= \phi_k(y)\,\mathsf{E}\left(\psi_k(\omega)\,|\mathcal{F}_t\right)|_{y=X_t}\\ &= \mathsf{E}\left(\phi_k(y)\,\psi_k(\omega)\,|\mathcal{F}_t\right)|_{y=X_t}.\end{aligned} \tag{31}$$

By going to the limit (which we do not discuss in detail), this establishes that

$$\mathsf{E}\left(g(X_t,\omega)\,|\mathcal{F}_t\right) = \mathsf{E}\left(g(y,\omega)\,|\mathcal{F}_t\right)|_{y=X_t}. \tag{32}$$

We now observe that by Eqns. 28 and 29

$$\begin{aligned}\mathsf{E}\left(g(X_t,\omega)\,|\mathcal{F}_t\right) &= \mathsf{E}\left(f\left(X_{t+h}^{t,X_t}(\omega)\right)|\mathcal{F}_t\right)\\ &= \mathsf{E}\left(f(X_{t+h})\,|\mathcal{F}_t\right),\end{aligned} \tag{33}$$

that is, the *left hand side* of Theorem 7.1.2.

Let us then consider the *right hand side* of Eqn. 32:

$$\begin{aligned}\mathsf{E}\left(g(y,\omega)\,|\mathcal{F}_t\right)|_{y=X_t(\omega)} &= \mathsf{E}\left(f\left(X_{t+h}^{t,y}(\omega)\right)|\mathcal{F}_t\right)|_{y=X_t(\omega)}\\ &\overset{(1)}{=} \mathsf{E}\left(f\left(X_{t+h}^{t,y}(\omega)\right)\right)|_{y=X_t(\omega)}\\ &\overset{(2)}{=} \mathsf{E}\left(f\left(X_{h}^{0,y}(\omega)\right)\right)|_{y=X_t(\omega)} \qquad (34)\\ &= G_h(y)\,|_{y=X_t(\omega)}\\ &= G_h(X_t(\omega)). \qquad (35)\end{aligned}$$

Equality (1) follows since $f\left(X_{t+h}^{t,y}(\omega)\right)$ is independent of $\mathcal{F}_t$ according to the remark following Eqn. 27, and equality (2) by the time homogeneity of the Itô diffusion. This proves Theorem 7.1.2. However, the key result which makes the way for us is indeed Eqn. 32, and we have not really proved that in detail.

Theorem 7.1.2 has an interesting interpretation in that it gives a recipe for computing expectations conditional on events happening to X_t. First of all,

$$\mathsf{E}\left(f(X_{t+h});X_t=y\right) = G(y).$$

More generally, if A is a Borel set in $\mathbb{R}^n$ such that $\mu_{X_t}(A)>0$, then

$$\mathsf{E}\left(f(X_{t+h});X_t\in A\right) = \frac{\int_A G_h(y)\,d\mu_{X_t}(y)}{\mu_{X_t}(A)}. \tag{36}$$

(Check that!).