

Comments to Chapter 7 in B.Ø.

Harald E. Krogstad, IMF

1 Introduction

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

Diffusion is originally the chaotic way molecules of fluids are mixed due to the motion of the molecules. In fact, physical Brownian motion may be viewed as 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 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

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

where, e.g. $u(\mathbf{x}, t)$ is the concentration of the material that is spread out into a fluid.

Brownian motion is, as we visualized by the numerical simulations, also a model 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 $t = 0$ has spread into a cloud with density proportional to the probability density of B_t , that is,

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

The distribution broadens as t increases, and tends, for $t \rightarrow 0$, 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

$$\frac{\partial p}{\partial t} = \frac{1}{2} \nabla^2 p \quad (3)$$

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

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

$$dX_t = dB_t \quad (4)$$

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 $0 \leq s \leq t$,*

$$E(X_t | \mathcal{M}_s) = E(X_t | X_s) \quad (5)$$

(where $E(X_t | X_s)$ is a short way of writing $E(X_t | \mathcal{F}_{X_s})$).

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

$$E(f(X_t) | \mathcal{M}_s) = E(f(X_t) | X_s), \quad (6)$$

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

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

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

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

(Note that $E(B_t - B_s) = E(B_t - B_s | B_s) = 0$).

More generally, Markov processes are constructed by means of so-called *transition functions* (generalizing the transition matrix of the Markov chain). Without going into too much detail, a (time homogeneous) transition function

$$p(t, u, A),$$

where $t \in [0, \infty)$, $u \in \mathbb{R}^n$, and $A \in \mathcal{B}(\mathbb{R}^n)$, fulfils

1. $p(t, \cdot, A)$ is a Borel function from $\mathbb{R}^n \rightarrow \mathbb{R}$,

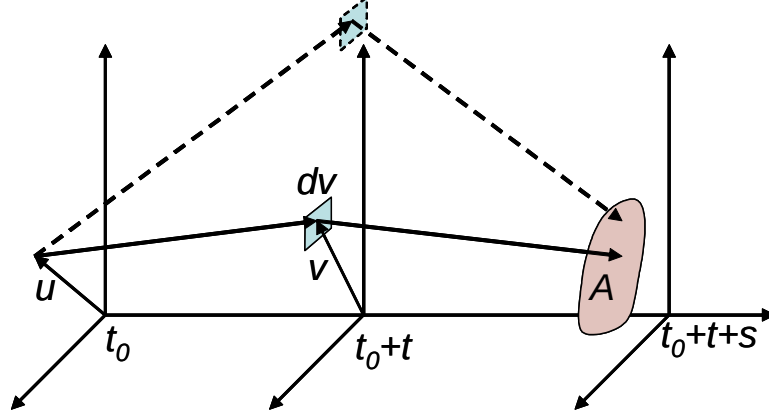


Figure 1: Illustration of the Chapman-Kolmogorov equation.

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

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

Note that the integral is a regular Lebesgue integral $\int_{\mathbb{R}^n} h dv$, where $h(v) = p(s, v, A)$ is a Borel function on $\mathbb{R}^n$, and $d\mu$ is the probability measure defined by $\mu(B) = p(t, u, B)$. In essence, the Chapman-Kolmogorov equation says that we go from the 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 . 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.

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

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

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

$$P(X_{t_1} \in F_1, X_{t_2} \in F_2, \dots, X_{t_k} \in F_k). \quad (10)$$

Taking the corresponding process X_t for granted, one may now say that

$$p(t, u, A) = P(X_t \in A | X_0 = u) \quad (11)$$

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

$$p(t, u, A) = \int_A \phi(t, u, v) dv, \quad (12)$$

In this case the Chapman-Kolmogorov equation takes the form

$$\begin{aligned}
p(t+s, u, A) &= \int_{\mathbb{R}^n} \phi(t, u, v) dv \int_A \phi(s, v, w) dw \\
&= \int_A \left(\int_{\mathbb{R}^n} \phi(t, u, v) \phi(s, v, w) dv \right) dw \\
&= \int_A \phi(t+s, u, w) dw.
\end{aligned} \tag{13}$$

Hence,

$$\phi(t+s, u, w) = \int_{\mathbb{R}^n} \phi(t, u, v) \phi(s, v, w) dv. \tag{14}$$

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

$$p(t, u, A) = \int_A \frac{1}{(2\pi t)^{n/2}} e^{-|u-v|^2/2t} dv, \tag{15}$$

and then Eqn. 14 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

We shall follow the notation in B.Ø. and let

$$X_t^{s,x} \tag{16}$$

denote the solution of the time homogeneous *Itô diffusion*

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t \tag{17}$$

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., such solutions are *time invariant* in the sense that $Z_t = X_t^{s,x}$ and $Y_t = X_{t+\gamma}^{s+\gamma,x}$ are *equivalent as stochastic processes*: Both start at x when $t = s$, and all their (finite dimensional) probability measures are the same.

Let us now consider

$$\begin{aligned}
X_{t+h}^x &= \int_0^{t+h} b(X_u^x) du + \int_0^{t+h} \sigma(X_u^x) dB_u \\
&= \int_0^t b(X_u^x) du + \int_0^t \sigma(X_u^x) dB_u + \int_t^{t+h} b(X_u^x) du + \int_t^{t+h} \sigma(X_u^x) dB_u \\
&= X_t^x + \int_t^{t+h} b(X_u^x) du + \int_t^{t+h} \sigma(X_u^x) dB_u.
\end{aligned} \tag{19}$$

The first term, X_t^x , is of course $\mathcal{F}_{X_t^x}$ -measurable (and hence $\mathcal{M}_t$ -measurable), whereas the last term is independent of X_s^x for $s < t$, but depends on X_t^x . Hence, it is possible to write

$$\begin{aligned}
E(X_{t+h}^x | \mathcal{M}_t) &= E\left(X_t^x + \int_t^{t+h} b(X_u^x) du + \int_t^{t+h} \sigma(X_u^x) dB_u | \mathcal{M}_t\right) \\
&= X_t^x + E\left(\int_t^{t+h} b(X_u^x) du + \int_t^{t+h} \sigma(X_u^x) dB_u | \mathcal{M}_t\right) \\
&= E\left(X_t^x + \int_t^{t+h} b(X_u^x) du + \int_t^{t+h} \sigma(X_u^x) dB_u | \mathcal{F}_{X_t^x}\right) \\
&= E(X_{t+h}^x | \mathcal{F}_{X_t^x}).
\end{aligned} \tag{20}$$

This shows that the solution of the Itô diffusion is a Markov process.

Theorem 7.1.2 in B.Ø. is an extension of 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 shall state the theorem by first introducing the function

$$G_h(y) = E(f(X_h^y)), \tag{21}$$

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

$$G_h(y) = E(f(X_h^y)) = \int_{\Omega} f(X_h^y(\omega)) d\mu_{y,h}(\omega). \tag{22}$$

Note that the composite function $f \circ X_h^y$ is $\mathcal{M}_{X_h^y}$ -measurable, and also $\mathcal{F}_h$ -measurable due to the properties of the Ito integral (Here $\{\mathcal{F}_h\}$ is as usual the filtration of the Brownian motion). In fact, 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:

$$E(f(X_{t+h}) | \mathcal{F}_t) = G_h(X_t). \tag{23}$$

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

$$E(f(X_{t+h}) | \mathcal{F}_t)(\omega) = G_h(X_t(\omega)). \tag{24}$$

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. 22, G_h is simply 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 ,

$$X_r^{t,x} = x + \int_t^r b(X_u) du + \int_t^r \sigma(X_u) dB_u, \tag{25}$$

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

$$X_{t+h}(\omega) = X_{t+h}^{t, X_t(\omega)}(\omega). \quad (26)$$

Fix t and h and let

$$g(y, \omega) = f\left(X_{t+h}^{t, y}(\omega)\right). \quad (27)$$

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

$$g(y, \omega) \rightsquigarrow \sum_k \phi_k(y) \psi_k(\omega). \quad (28)$$

Following the rules of conditional expectation,

$$\begin{aligned} E(\phi_k(X_t) \psi_k(\omega) | \mathcal{F}_t) &= \phi_k(X_t) E(\psi_k(\omega) | \mathcal{F}_t) \\ &= \phi_k(y) E(\psi_k(\omega) | \mathcal{F}_t) |_{y=X_t} \\ &= E(\phi_k(y) \psi_k(\omega) | \mathcal{F}_t) |_{y=X_t}. \end{aligned} \quad (29)$$

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

$$E(g(X_t, \omega) | \mathcal{F}_t) = E(g(y, \omega) | \mathcal{F}_t) |_{y=X_t}. \quad (30)$$

We now observe that by Eqns. 26 and 27

$$\begin{aligned} E(g(X_t, \omega) | \mathcal{F}_t) &= E\left(f\left(X_{t+h}^{t, X_t}(\omega)\right) | \mathcal{F}_t\right) \\ &= E(f(X_{t+h}) | \mathcal{F}_t), \end{aligned} \quad (31)$$

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

Let us then consider the *right hand side* of Eqn. 30:

$$\begin{aligned} E(g(y, \omega) | \mathcal{F}_t) |_{y=X_t(\omega)} &= E\left(f\left(X_{t+h}^{t, y}(\omega)\right) | \mathcal{F}_t\right) |_{y=X_t(\omega)} \\ &\stackrel{(1)}{=} E\left(f\left(X_{t+h}^{t, y}(\omega)\right)\right) |_{y=X_t(\omega)} \\ &\stackrel{(2)}{=} E\left(f\left(X_h^{0, y}(\omega)\right)\right) |_{y=X_t(\omega)} \end{aligned} \quad (32)$$

$$\begin{aligned} &= G_h(y) |_{y=X_t(\omega)} \\ &= G_h(X_t(\omega)). \end{aligned} \quad (33)$$

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. 25, and equality (2) by the time homogeneity of the Ito diffusion. This proves Theorem 7.1.2. However, the key result which makes the way for us is indeed Eqn. 30, 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,

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

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

$$E(f(X_{t+h}); X_t \in A) = \frac{\int_A G_h(y) d\mu_{X_t}(y)}{\mu_{X_t}(A)} \quad (34)$$

(check that!).