11

Stochastic perturbation of continuous time systems

In this, the last chapter, continuous time systems in the presence of noise are considered. This leads us to examine systems of stochastic differential equations and to a derivation of the forward Fokker–Planck equation, describing the evaluation of densities for these systems. We close with some results concerning the asymptotic stability of solutions to the Fokker–Planck equation.

11.1 One-dimensional Wiener processes (Brownian motion)

In this and succeeding sections of this chapter, we turn to a consideration of continuous time systems with stochastic perturbations. We are specifically interested in the behavior of the system

$$\frac{dx}{dt} = b(x) + \sigma(x)\xi, \tag{11.1.1}$$

where $\sigma(x)$ is the amplitude of the perturbation and $\xi = dw/dt$ is known as a "white noise" term that may be considered to be the time derivative of a Wiener process. The system (11.1.1) is the continuous time analog of the discrete time problem with a constantly applied stochastic perturbation considered in Section 10.5.

The consideration of continuous time problems such as (11.1.1) will offer new insight into the possible behavior of systems, but at the expense of introducing new concepts and techniques. Even though the remainder of this chapter is written to be self-contained, it does not constitute an exhaustive treatment of stochastic differential equations such as (11.1.1). A definitive treatment of this subject may be found in Gikhman and Skorokhod [1969].

In this section and the material following, we will denote stochastic processes by $\{\xi(t)\}, \{\eta(t)\}, \ldots$ as well as $\{\xi_t\}, \{\eta_t\}, \ldots$, depending on the situation. Remember that in this notation $\xi(t)$ or ξ_t denote, for each fixed t, a random variable, namely, a measurable function $\xi_t\colon \Omega \to R$. Thus $\xi(t)$ and ξ_t, are really abbreviations for $\xi(t, \omega)$ and $\xi_t(\omega)$, respectively. The symbol ξ will be reserved for white

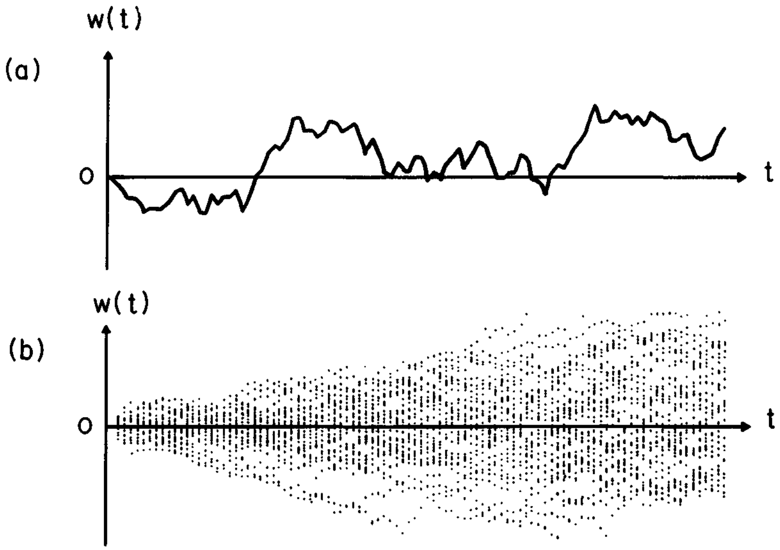

Figure 11.1.1. A process approximating a Wiener process. In (a) we show a single sample path for this process. In (b) we superimpose the points of many sample paths to show the progressive increase in the variance.

noise stochastic processes (to be described later), whereas η will be used for other stochastic processes.

Let a probability space $(\Omega, \mathcal{F}, \text{prob})$ be given. We start with a definition.

Definition 11.1.1. A stochastic process $\{\eta(t)\}$ is called **continuous** if, for almost all ω (except for a set of probability zero), the sample path $t \to \eta(t, \omega)$ is a continuous function.

A Wiener process can now be defined as follows.

Definition 11.1.2. A **one-dimensional normalized Wiener process** (or **Brownian motion**) $\{w(t)\}_{t\geq 0}$ is a continuous stochastic process with independent increments such that

(a) $w(0) = 0$; and
(b) for every s, t, $0 \leq s < t$, the random variable $w(t) - w(s)$ has the Gaussian density

$$g(t - s, x) = \frac{1}{\sqrt{2\pi(t - s)}} \exp[-x^2/2(t - s)]. \tag{11.1.2}$$

Figure 11.1.1a shows a sample path for a process approximating a Wiener process.

It is clear that a Wiener process has stationary increments since $w(t) - w(s)$ and $w(t + t') - w(s + t')$ have the same density function (11.1.2). Further, since $w(t) = w(t) - w(0)$, the random variable $w(t)$ has the density

$$g(t,x) = \frac{1}{\sqrt{2\pi t}} \exp(-x^2/2t). \tag{11.1.3}$$

An easy calculation shows

$$\begin{aligned} E((w(t) - w(s))^n) &= \frac{1}{\sqrt{2\pi(t - s)}} \int_{-\infty}^{\infty} x^n \exp[-x^2/2(t - s)]\,dx \\ &= \begin{cases} 1.3\cdots(n - 1)(t - s)^{n/2} & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd} \end{cases} \end{aligned} \tag{11.1.4}$$

and thus, in particular,

$$E(w(t) - w(s)) = 0 \tag{11.1.5}$$

and

$$D^2(w(t) - w(s)) = (t - s). \tag{11.1.6}$$

This last equation demonstrates that the variance of a Wiener process increases linearly with t.

Remark 11.1.1. The adjective **normalized** in our definition of the Wiener process is used because $D^2(w(t)) = t$. It is clear that multiplication of a normalized Wiener process by a constant $\sigma > 0$ again yields a process with properties similar to those of Definition 11.1.2, but now with the density

$$\frac{1}{\sqrt{2\pi\sigma^2 t}} \exp(-x^2/2\sigma^2 t).$$

and with the variance $\sigma^2 t$. These processes are also called Wiener processes. From this point on we will always refer to a normalized Wiener process as a Wiener process. □

In Figure 11.1.1b we have drawn a number of sample paths for a process approximating a Wiener process. Note that as time increases they all seem to be bounded by a convex envelope. This is due to the fact that the standard deviation of a Wiener process, from (11.1.6), increases as $\sqrt{t}$, that is,

$$[D^2(w(t))]^{1/2} = \sqrt{t}.$$

The highly irregular behavior of these individual trajectories is such that magnification of any part of the trajectory by a factor α^2 in the time direction and

α in the x direction yields a picture indistinguishable from the original trajectory. This procedure can be repeated as often as one wishes, and, indeed, the sample paths of a Wiener process are fractile curves [Mandelbrot, 1977]. To obtain some insight into the origin of this behavior consider the absolute value of the differential quotient

$$\left|\frac{\Delta w}{\Delta t}\right| = \frac{1}{|\Delta t|}|w(t_0 + \Delta t) - w(t_0)|.$$

We have

$$E\left(\left|\frac{\Delta w}{\Delta t}\right|\right) = \frac{1}{|\Delta t|}E(|w(t_0 + \Delta t) - w(t_0)|)$$

and, since the density of $w(t_0 + \Delta t) - w(t_0)$ is given by (11.1.3),

$$\begin{aligned} E(|w(t_0 + \Delta t) - w(t_0)|) &= \frac{1}{\sqrt{2\pi\,\Delta t}}\int_{-\infty}^{\infty} |x| \exp(-x^2/2\,\Delta t)\,dx \\ &= \sqrt{2\,\Delta t/\pi} \end{aligned}$$

or

$$E\left(\left|\frac{\Delta w}{\Delta t}\right|\right) = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{\Delta t}}.$$

Thus the mathematical expectation of $|\Delta w/\Delta t|$ goes to infinity, with a speed proportional to $(\Delta t)^{-1/2}$, when $|\Delta t| \to 0$. This is the origin of the irregular behavior shown in Figure 11.1.1.

Extending the foregoing argument, it can be proved that the sample paths of a Wiener process are not differentiable at any point almost surely. Thus, the white noise term $\xi = dw/dt$ in (11.1.1) does not exist as a stochastic process. However, since we do wish ultimately to consider (11.1.1) with such a perturbation, we must inquire how this can be accomplished. As shown in following sections, this is simply done by formally integrating (11.1.1) and treating the resulting system,

$$x(t) = \int_0^t b(x(s))\,ds + \int_0^t \sigma(x(s))\,dw(s) + x^0.$$

However, this approach leads to the new problem of defining what the integrals on the right-hand side mean, which will be dealt with in Section 11.3.

To obtain further insight into the nature of the process $w(t)$, examine the alternative sequence $\{z_n\}$ of processes, defined by

$$z_n(t) = w(t_{i-1}^n) + \frac{t - t_{i-1}^n}{t_i^n - t_{i-1}^n}[w(t_i^n) - w(t_{i-1}^n)] \qquad \text{for } t \in [t_{i-1}^n, t_i^n],$$

where $t_i^n = i/n$, $n = 1, 2, \ldots$, $i = 0, 1, 2, \ldots$. In other words, z_n is obtained by sampling the Wiener process $w(t)$ at times t_i^n and then applying a linear interpolation between t_i^n and t_{i+1}^n. Any sample path of the process $\{z_n(t)\}$ is differentiable, except at the points t_i^n, and the derivative $\eta_n = z_n'$ is given by

$$\eta_n(t) = n[w(t_i^n) - w(t_{i-1}^n)], \qquad \text{for } t \in (t_{i-1}^n, t_i^n).$$

The process $\eta_n(t)$ is piecewise constant. The heights of the individual segments are independent, have a mean value zero, and variance $D^2\eta_n(t) = n$. Thus, the variance grows linearly with n. If we look at this process approximating white noise, we see that it consists of a sequence of independent impulses of width $(1/n)$ and variance n. For very large n we will see peaks of almost all possible sizes uniformly spread along the t-axis.

Note that the random variable $z_n(t)$ for fixed t and large n is the sum of many independent increments. Thus the density of $z_n(t)$ must be close to a Gaussian by the central limit theorem. The limiting process $w(t)$ will, therefore, also have a Gaussian density, which is why we assumed that $w(t)$ had a Gaussian density in Definition 11.1.2.

Historically, Wiener processes (or Brownian motion) first became of interest because of the findings of the English biologist Brown, who observed the microscopic movement of pollen particles in water due to the random collisions of water molecules with the particles. The impulses coming from these collisions are almost ideal realizations of the process of white noise, somewhat similar to our process $\eta_n(t)$ for large n.

In other applications, however, much slower processes are admitted as "white noise" perturbations, for example, waves striking the side of a large ship or the influence of atmospheric turbulence on an airplane. In the example of the ship, the reason that this assumption is a valid approximation stems from the fact that waves of quite varied energies strike both sides of the ship almost independently with a frequency much larger than the free oscillation frequency of the ship.

Example 11.1.1. Having defined a one-dimensional Wiener process $\{w(t)\}_{t\geqslant 0}$, it is rather easy to construct an exact, continuous time, semidynamical system that corresponds to the partial differential equation

$$\frac{\partial u}{\partial t} + s\frac{\partial u}{\partial s} = \frac{1}{2}u. \tag{11.1.7}$$

Our arguments follow those of Rudnicki [In press], which generalize results of Lasota [1981], Brunovsky [1983], and Brunovsky and Komornik [1984].

The first step in this process is to construct the Wiener measure. Let X be the space of all continuous functions x: $[0, 1] \to R$ such that $x(0) = 0$. We are going to define some special subsets of X that are called cylinders. Thus, given a

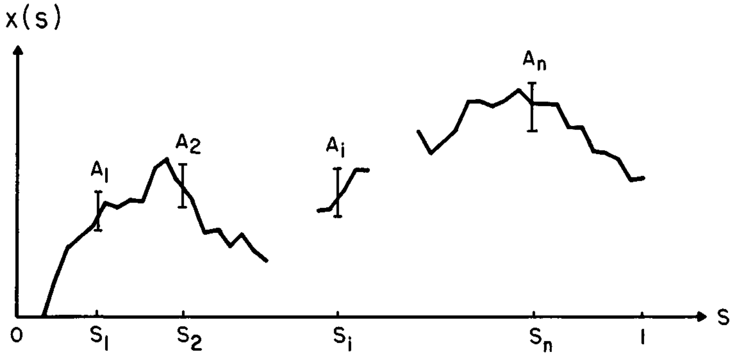

Figure 11.1.2. Schematic representation of implications of the cylinder definition [equation (11.1.8).]

sequence of real numbers,

$$0 < s_1 < \cdots < s_n \leqslant 1\,,$$

and a sequence of Borel subsets of R,

$$A_1, \ldots, A_n\,,$$

we define the corresponding **cylinder** by

$$C(s_1, \ldots, s_n; A_1, \ldots, A_n) = \{x \in X\colon x(s_i) \in A_i,\ i = 1, \ldots, n\}. \tag{11.1.8}$$

Thus the cylinder defined by (11.1.8) is the set of all functions $x \in X$ passing through the set A_i at s_i (see Figure 11.1.2). The **Wiener measure** μ_w of the cylinders (11.1.8) is defined by

$$\begin{aligned}\mu_w(C(s_1, \ldots, s_n; A_1, \ldots, A_n))\\ = \text{prob}\{w(s_1) \in A_1, \ldots, w(s_n) \in A_n\}.\end{aligned} \tag{11.1.9}$$

To derive an explicit formula for μ_w, consider a transformation $y = F(x)$ of R^n into itself given by

$$y_1 = x_1, y_2 = x_2 - x_1, \ldots, y_n = x_n - x_{n-1} \tag{11.1.10}$$

and set $A = A_1 \times \cdots \times A_n$. Then the condition

$$(w(s_1), \ldots, w(s_n)) \in A$$

is equivalent to the requirement that the random vector

$$(w(s_1), w(s_2) - w(s_1), \ldots, w(s_n) - w(s_{n-1})) \tag{11.1.11}$$

belong to $F(A)$. Since $\{w(t)\}_{t\geqslant 0}$ is a random process with independent increments, the density function of the random vector (11.1.11) is given by

$$g(s_1, y_1)g(s_2 - s_1, y_2), \ldots, g(s_n - s_{n-1}, y_n),$$

where, by the definition of the Wiener process [see equation (11.1.3)],

$$g(s, y) = \frac{1}{\sqrt{2\pi s}} \exp(-y^2/2s). \tag{11.1.12}$$

Thus we have

$$\begin{aligned}&\text{prob}\{w(s_1) \in A_1, \ldots, w(s_n) \in A_n\}\\ &\quad = \int \cdots \int_{F(A)} g(s_1, y_1)g(s_2 - s_1, y_2) \cdots g(s_n - s_{n-1}, y_n)\, dy_1 \cdots dy_n.\end{aligned}$$

Using the variables defined in (11.1.10), this becomes

$$\begin{aligned}&\text{prob}\{w(s_1) \in A_1, \ldots, w(s_n) \in A_n\}\\ &\quad = \int_{A_1} \cdots \int_{A_n} g(s_1, x_1)g(s_2 - s_1, x_2 - x_1)\\ &\qquad\qquad \cdots g(s_n - s_{n-1}, x_n - x_{n-1})\, dx_1 \cdots dx_n.\end{aligned}$$

By combining this expression with equations (11.1.9) and (11.1.12), we obtain the famous formula for the Wiener measure:

$$\begin{aligned}&\mu_w(C(s_1, \ldots, s_n; A_1, \ldots, A_n))\\ &\quad = \frac{1}{\sqrt{(2\pi)^n (s_1 - s_0) \cdots (s_n - s_{n-1})}}\\ &\cdot \int_{A_1} \cdots \int_{A_n} \exp\left[-\frac{1}{2}\sum_{k=1}^{n} \frac{(x_k - x_{k-1})^2}{s_k - s_{k-1}}\right] dx_1 \cdots dx_n.\end{aligned} \tag{11.1.13}$$

(We assume, for simplicity, that $s_0 = x_0 = 0$.)

To extend the definition of μ_w, we can define the σ-algebra $\mathscr{A}$ to be the smallest σ-algebra of the subsets of X that contains all the cylinders defined by (11.1.8) for arbitrary n. By definition, the Wiener measure μ_w is the (unique) extension of μ_w, given by (11.1.13) on cylinders, to the entire σ-algebra $\mathscr{A}$. The proof that μ_w given by (11.1.13) on cylinders can be extended to the entire σ-algebra is technically difficult, and we omit it. However, note that if a Wiener process $\{w(t)\}_{t\geq 0}$ is given, then it is a direct consequence of our construction of the Wiener measure for cylinders that

$$\mu_w(E) = \text{prob}(\tilde{w} \in E) \qquad \text{for } E \in \mathscr{A}, \tag{11.1.14}$$

where $\tilde{w}$ is the restriction of w to the interval $[0, 1]$. (Incidentally, from this equation, it also follows that the assumption that a Wiener process $\{w(t)\}_{t\geq 0}$ exists is not trivial, but, in fact, is equivalent to the existence of the Wiener measure.)

With the element of the measure space $(X, \mathscr{A}, \mu_w)$ defined, we now turn to a definition of the semidynamical system $\{S_t\}_{t\geq 0}$ corresponding to (11.1.7). With the initial condition

$$u(0,s) = x(s)\,, \tag{11.1.15}$$

equation (11.1.7) has the solution

$$u(t,s) = e^{t/2}x(se^{-t})\,.$$

Thus, if we set

$$S_t x(s) = e^{t/2}x(se^{-t})\,, \tag{11.1.16}$$

this equation defines $\{S_t\}_{t\geq 0}$.

We first show that $\{S_t\}_{t\geq 0}$ preserves the Wiener measure μ_w. Since the measures μ_w on cylinders generate the Wiener measure on the entire σ-algebra $\mathcal{A}$, we will only verify the measure-preservation condition

$$\mu_w(S_t^{-1}(C)) = \mu_w(C) \tag{11.1.17}$$

for cylinders. First observe that for every $\alpha \in (0,1)$,

$$\begin{aligned}\mu_w(C(\alpha^2 s_1, \ldots, \alpha^2 s_n;\, \alpha A_1, \ldots, \alpha A_n))& \\ = \mu_w(C(s_1, \ldots, s_n;\, A_1, \ldots, A_n))&\end{aligned} \tag{11.1.18}$$

This follows directly from equation (11.1.13) if we set $y_i = \alpha x_i$ in the integral on the right-hand side. Further, from (11.1.16), it is clear that $(S_t x)(s_i) \in A_i$ if and only if $x(s_i e^{-t}) \in e^{-t/2}A_i$. Thus,

$$\begin{aligned}&S_t^{-1}(C(s_1, \ldots, s_n;\, A_1, \ldots, A_n))\\ &\quad = \{x \in X\colon (S_t x)(s_i) \in A_i,\ i = 1, \ldots, n\}\\ &\quad = C(e^{-t}s_1, \ldots, e^{-t}s_n;\, e^{-t/2}A_1, \ldots, e^{-t/2}A_n)\,.\end{aligned}$$

From this relation and (11.1.18) with $\alpha = e^{-t/2}$, we immediately obtain (11.1.17), thereby verifying that $\{S_t\}_{t\geq 0}$ preserves the Wiener measure μ_w.

To demonstrate the exactness of $\{S_t\}_{t\geq 0}$, we will be content to show that

$$\lim_{t\to\infty} \mu_w(S_t(C)) = 1 \qquad \text{if } \mu_w(C) > 0 \tag{11.1.19}$$

for cylinders. In this case we have

$$\begin{aligned}S_t(C) &= S_t(C(s_1, \ldots, s_n;\, A_1, \ldots, A_n))\\ &= \{S_t x\colon x \in C\} = \{e^{t/2}x(se^{-t})\colon x \in C\}\,.\end{aligned}$$

Set $y(s) = e^{t/2}x(se^{-t})$ so this becomes

$$\begin{aligned}S_t(C) &= S_t(C(s_1, \ldots, s_n;\, A_1, \ldots, A_n))\\ &= \{y \in X\colon y(s) = e^{t/2}x(se^{-t}), x(s_i) \in A_i,\ i = 1, \ldots, n\}\,.\end{aligned} \tag{11.1.20}$$

Since $s \in [0,1]$ and, thus, $se^{-t} \in [0, e^{-t}]$, the conditions $x(s_i) \in A_i$ are irrelevant for $s_i > e^{-t}$. Thus

$$S_t(C(s_1, \ldots, s_n; A_1, \ldots, A_n)) = C(s_1 e^t, \ldots, s_k e^t; e^{t/2}A_1, \ldots, e^{t/2}A_k)$$

where $k = k(t)$ is the largest integer $k \leqslant n$ such that $s_k \leqslant e^{-t}$. Once t becomes sufficiently large, that is, $t > -\log s_1$, then from (11.1.20) we see that the last condition $x_1 \in A_1$ disappears and we are left with

$$S_t(C(s_1, \ldots, s_n; A_1, \ldots, A_n)) = \{y \in X: y(s) = e^{t/2}x(se^{-t})\}.$$

However, since X is the space of all possible continuous functions x: $[0,1] \to R$, the set on the right-hand side is just X and, as a consequence,

$$\mu_w(S_t(C(s_1, \ldots, s_n; A_1, \ldots, A_n))) = 1 \qquad \text{for } t > -\log s_1,$$

which proves equation (11.1.19) for cylinders.

In the general case, for an arbitrary $C \in \mathscr{A}$ the demonstration that (11.1.19) holds is more difficult, but the outline of the argument is as follows. Starting with the equality

$$\mu_w(S_t(C)) = \mu_w(S_t^{-1}S_t(C)),$$

and using the fact that the family $\{S_t^{-1}S_t(C)\}_{t \geqslant 0}$ is increasing with t, we obtain

$$\lim_{t\to\infty} \mu_w(S_t(C)) = \mu_w(B), \tag{11.1.21}$$

where

$$B = \bigcup_{t \geqslant t_0} S_t^{-1}S_t(C) \tag{11.1.22}$$

and t_0 is an arbitrary nonnegative number. From (11.1.22), it follows that

$$B \in \mathscr{A}_\infty = \bigcap_{t \geqslant 0} S_t^{-1}(\mathscr{A}).$$

From the Blumenthal zero–one law [see Remark 11.2.1] it may be shown that the σ-algebra $\mathscr{A}_\infty$ contains only trivial sets. Thus, since $\mu_w(B) \geqslant \mu_w(C)$, we must have $\mu_w(B) = 1$ whenever $\mu_w(C) > 0$. Thus (11.1.19) follows immediately from (11.1.21).

A proof of exactness may also be carried out for equations more general than the linear version (11.1.7). The nonlinear equation

$$\frac{\partial u}{\partial t} + c(s)\frac{\partial u}{\partial s} = f(s, u), \tag{11.1.23}$$

has been used to model the dynamics of a population of cells undergoing simultaneous proliferation and maturation [Lasota, Mackey, and Ważewska-Czyżewska, 1981; Mackey and Dörmer, 1982], where s is the maturation variable. When the coefficients c and f satisfy some additional conditions, it can be shown that all the solutions of (11.1.23) with the initial condition (11.1.15) converge to the same limit if $x(0) > 0$. However, if $x(0) = 0$, then the solutions of (11.1.23) will exhibit extremely irregular behavior that can be identified with the exactness of the semidynamical system $\{S_t\}_{t\geq 0}$ corresponding to $u(t, s)$. This latter situation [$x(0) = 0$] corresponds to the destruction of the most primitive cell type (maturity $= 0$), and in such situations the erratic behavior corresponding to exactness of $\{S_t\}_{t\geq 0}$ is noted clinically. □

11.2 d-Dimensional Wiener processes (Brownian motion)

In considering d-dimensional Wiener processes we will require an extension of our definition of independent sets. Suppose we have a finite sequence

$$\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n, \qquad \mathcal{F}_i \subset \mathcal{F} \tag{11.2.1}$$

of σ-algebras. We define the independence of (11.2.1) as follows.

Definition 11.2.1. A sequence (11.2.1) consists of **independent σ-algebras** if all possible sequences of sets $A_1, \ldots, A_n$ such that

$$A_1 \in \mathcal{F}_1, \ldots, A_n \in \mathcal{F}_n$$

are independent.

Further, for every random variable ξ we denote by $\mathcal{F}(\xi)$ the σ-algebra of all events of the form $\{\omega: \xi(\omega) \in B\}$, where the B are Borel sets, or, more explicitly,

$$\mathcal{F}(\xi) = \{\xi^{-1}(B): B \text{ is a Borel set}\}.$$

Having a stochastic process $\{\eta(t)\}_{t\in\Delta}$ on an interval Δ, we denote the smallest σ-algebra that contains all sets of the form

$$\{\omega: \eta(t, \omega) \in B\},\ t \in \Delta,\ B \text{ is a Borel set},$$

by $\mathcal{F}(\eta(t): t \in \Delta)$.

With this notation we can restate our definition of independent random variables as follows. The random variables $\xi_1, \ldots, \xi_n$ are independent if $\mathcal{F}(\xi_1), \ldots, \mathcal{F}(\xi_n)$ are independent. In an analogous fashion, stochastic processes $\{\eta_1(t)\}_{t\in\Delta_1}, \ldots, \{\eta_n(t)\}_{t\in\Delta_n}$ are independent, if

$$\mathcal{F}(\eta_1(t): t \in \Delta_1), \ldots, \mathcal{F}(\eta_n(t): t \in \Delta_n)$$

are independent.

Finally, having m random variables $\xi_1, \ldots, \xi_m$ and n stochastic processes $\{\eta_1(t)\}_{t\in\Delta_1}, \ldots, \{\eta_n(t)\}_{t\in\Delta_n}$, we say that they are independent if the σ-algebras

$$\mathcal{F}(\xi_1), \ldots, \mathcal{F}(\xi_m), \mathcal{F}(\eta_1(t): t \in \Delta_1), \ldots, \mathcal{F}(\eta_n(t): t \in \Delta_n)$$

are independent. We will also say that a stochastic process $\{\eta(t)\}_{t\in\Delta}$ and a σ-algebra $\mathscr{F}_0$ are independent if $\mathscr{F}(\eta(t): t \in \Delta)$ and $\mathscr{F}_0$ are independent.

Now it is straightforward to define a d-dimensional Wiener process.

Definition 11.2.2. A d-dimensional vector valued process

$$w(t) = \{w_1(t), \dots, w_d(t)\}, \qquad t \geqslant 0$$

is a **d-dimensional Wiener process (Brownian motion)** if its components $\{w_1(t)\}_{t\geqslant 0}, \dots, \{w_d(t)\}_{t\geqslant 0}$ are one-dimensional independent Wiener processes (Brownian motion).

From this definition it follows that for every fixed t the random variables $w_1(t), \dots, w_d(t)$ are independent. Thus, it is an immediate consequence of Theorem 10.1.1 that the joint density of the random vector $(w_1(t), \dots, w_d(t))$ is given by

$$\begin{aligned} g(t, x_1, \dots, x_d) &= g(t, x_1) \cdots g(t, x_d) \\ &= \frac{1}{(2\pi t)^{d/2}} \exp\left[-\frac{1}{2t}\sum_{i=1}^{d} x_i^2\right]. \end{aligned} \tag{11.2.2}$$

The joint denisty g has the following properties:

$$\int_{R^d}\cdots\int g(t, x_1, \dots, x_d)\, dx_1 \cdots dx_d = 1\,, \tag{11.2.3}$$

$$\int_{R^d}\cdots\int x_i g(t, x_1, \dots, x_d)\, dx_1 \cdots dx_d = 0, \qquad i = 1, \dots, d, \tag{11.2.4}$$

and

$$\int_{R^d}\cdots\int x_i x_j g(t, x_1, \dots, x_d)\, dx_1 \cdots dx_d = \delta_{ij} t, \qquad i, j = 1, \dots, d, \tag{11.2.5}$$

where δ_{ij} is the Kronecker delta ($\delta_{ij} = 0$, $i \neq j$, $\delta_{ii} = 1$).

Remark 11.2.1. The family $\mathscr{F}(w(u): 0 \leqslant u \leqslant t)$ of σ-algebras generated by the Wiener process (or d-dimensional Wiener process) has the interesting property that it is right-hand continuous: We have (modulo zero)

$$\mathscr{F}(w(u): 0 \leqslant u \leqslant t) = \bigcap_{h>0} \mathscr{F}(w(u): 0 \leqslant u \leqslant t + h). \tag{11.2.6}$$

In particular at $t = 0$, since $w(0) = 0$ and the σ-algebra generated by $w(0)$ is trivial, we can see from equality (11.2.6) that the product

$$\bigcap_{h>0} \mathcal{F}(w(u)\colon 0 \leqslant u \leqslant h)$$

contains only sets of measure zero or one. The last statement is referred as the **Blumenthal zero–one law** (Friedman [1975].) □

11.3 The stochastic Itô integral: development

To understand what is meant by a solution to the stochastic differential equation (11.1.1), it is necessary to introduce the concept of the stochastic Itô integral. In this section we offer a simple but precise definition of this integral and calculate some specific cases so that a comparison with the usual Lebesgue integral may be made.

Let a probability space $(\Omega, \mathcal{F}, \text{prob})$ be given, and let $\{w(t)\}_{t\geqslant 0}$ be a one-dimensional Wiener process. If $\{\eta(t)\}_{t\in[\alpha,\beta]}$ is another stochastic process defined for $t \in [\alpha, \beta]$, $\alpha \geqslant 0$, we wish to know how to interpret the integral

$$\int_\alpha^\beta \eta(t)\,dw(t)\,. \tag{11.3.1}$$

Proceeding naively from the classical rules of calculus would suggest that (11.3.1) should be replaced by

$$\int_\alpha^\beta \eta(t)w'(t)\,dt\,.$$

However, this integral is only defined if $w(t)$ is a differentiable function, which we have already observed is not the case for a Wiener process.

Another possibility suggested by classical analysis is to consider (11.3.1) as the limit of approximating sums $\bar{s}$ of the form

$$\bar{s} = \sum_{i=1}^{k} \eta(\bar{t}_i)[w(t_i) - w(t_{i-1})]\,, \tag{11.3.2}$$

where

$$\alpha = t_0 < t_1 < \cdots < t_k = \beta$$

is a partition of the interval $[\alpha, \beta]$ and the intermediate points $\bar{t}_i \in [t_i, t_{i+1}]$. This turns out to be a more fruitful idea but has the surprising consequence that the limit of the approximating sums $\bar{s}$ of the form (11.3.2) depends on the choice of the intermediate points $\bar{t}_i$, in sharp contrast to the situation for the Riemann and Stieltjes integrals. This occurs because $w(t)$, at fixed ω, is not a function of bounded variation.

With these preliminary remarks in mind, we now proceed to develop some concepts of use in the definition of the Itô integral.

Definition 11.3.1. A family $\{\mathcal{F}_t\}$, $\alpha \leqslant t \leqslant \beta$, of σ-algebras contained in $\mathcal{F}$ is called **nonanticipative** if the following three conditions are satisfied:

(1) $\mathcal{F}_u \subset \mathcal{F}_t$ for $u \leqslant t$, so $\mathcal{F}_t$ increases as t increases;
(2) $\mathcal{F}_t \supset \mathcal{F}(w(u)\colon \alpha \leqslant u \leqslant t)$, so $w(u)$, $\alpha \leqslant u \leqslant t$, is measurable with respect to $\mathcal{F}_t$;
(3) $w(t+h) - w(t)$ is independent of $\mathcal{F}_t$ for $h \geqslant 0$, so all pairs of sets A_1, A_2 such that $A_1 \in \mathcal{F}_t$ and $A_2 \in \mathcal{F}(w(t+h) - w(t))$ are independent.

From this point on we will assume that a Wiener process $w(t)$ and a family of nonanticipative σ-algebras $\{\mathcal{F}_t\}$, $\alpha \leqslant t \leqslant \beta$, are given.

We next define a fourth condition.

Definition 11.3.2. A stochastic **process** $\{\eta(t)\}$, $\alpha \leqslant t \leqslant \beta$, is called **non-anticipative** with respect to $\{\mathcal{F}_t\}$ if

(4) $\mathcal{F}_t \supset \mathcal{F}\{\eta(u)\colon \alpha \leqslant u \leqslant t\}$, so $\eta(u)$ is measurable with respect to $\mathcal{F}_t$.

For every random process $\{\eta(t)\}$, $\alpha \leqslant t \leqslant \beta$, we define the **Itô sum** s by

$$s = \sum_{i=1}^{k} \eta(t_{i-1})[w(t_i) - w(t_{i-1})]. \tag{11.3.3}$$

Note that in the definition of the Itô sum (11.3.3), we have specified the intermediate points $\bar{t}_i$ of (11.3.2) to be the left end of each interval, $\bar{t}_i = t_{i-1}$. For a given Itô sum s, we define

$$\delta(s) = \max_i (t_i - t_{i-1})$$

and call a sequence of Itô sums $\{s_n\}$ **regular** if $\delta(s_n) \to 0$ as $n \to \infty$.

We now define the Itô integral as follows.

Definition 11.3.3. Let $\{\eta(t)\}$, $\alpha \leqslant t \leqslant \beta$, be a nonanticipative stochastic process. If there exists a random variable ζ such that

$$\zeta = \text{st-lim}\, s_n \tag{11.3.4}$$

for every regular sequence of the Itô sums $\{s_n\}$, then we say that ζ is the **Itô integral** of $\{\eta(t)\}$ on the interval $[\alpha, \beta]$ and denote it by

$$\zeta = \int_\alpha^\beta \eta(t)\, dw(t). \tag{11.3.5}$$

Remark 11.3.1. It can be proved that for every continuous nonanticipative process the limit (11.3.4) always exists. □

Remark 11.3.2. Definition 11.3.1 of a nonanticipative σ-algebra is complicated, and the reason for introducing each element of the definition, as well as the implication of each, may appear somewhat obscure. Condition (1) is easy, for it merely means that the σ-algebra $\mathcal{F}_t$ of events grows as time proceeds. The second condition ensures that $\mathcal{F}_t$ contains all of the events that can be described by the Wiener process $w(s)$ for times $s \in [\alpha, t]$. Finally, condition (3) says that no information concerning the behavior of the process $w(u) - w(t)$ for $u > t$ can influence calculations involving the probability of the events in $\mathcal{F}_t$. Definition 11.3.2 gives to a stochastic process $\eta(u)$ the same property that condition (2) of Definition 11.3.1 gives to $w(u)$. Thus, all of the information that can be obtained from $\eta(u)$ for $u \in [\alpha, t]$ is contained in $\mathcal{F}_t$.

Taken together, these four conditions ensure that the integrand $\eta(t)$ of the Itô integral (11.3.5) does not depend on the behavior of $w(t)$ for times greater than β and aid in the proof of the convergence of the Itô approximating sums. Further, the nonanticipatory assumption plays an important role in the proof of the existence and uniqueness of solutions to stochastic differential equations since it guarantees that the behavior of a solution in a time interval $[0, t]$ is not influenced by the Wiener process for times larger than t. □

Example 11.3.1. For our first example of the calculation of a specific Itô integral, we take

$$\int_0^T dw(t).$$

In this case the integrand of (11.3.5) is $\eta(t) \equiv 1$. Thus $\mathcal{F}(\eta(t){:}\ 0 \leqslant t \leqslant T)$ is a trivial σ-algebra that contains the whole space Ω and the empty set $\emptyset$. To see this, note that, if $1 \in B$, then $\{\omega{:}\ \eta(t) \in B\} = \Omega$ and, if $1 \notin B$ then $\{\omega{:}\ \eta(t) \in B\} = \emptyset$. This trivial σ-algebra $\{\emptyset, \Omega\}$ is contained in any other σ-algebra, and thus condition (4) of Definition 11.3.2 is satisfied.

By definition

$$s = \sum_{i=1}^{k} [w(t_i) - w(t_{i-1})] = w(t_k) - w(t_0) = w(T)$$

and, thus,

$$\int_0^T dw(t) = w(T). \quad \square$$

Example 11.3.2. In this example we will evaluate

$$\int_0^T w(t)\, dw(t),$$

which is not as trivial as our previous example.

In this case, $\eta(t) = w(t)$, so that condition (4) of Definition 11.3.2 follows from condition (2) of Definition 11.3.1. The Itô sum,

$$s = \sum_{i=1}^{k} w(t_{i-1})[w(t_i) - w(t_{i-1})],$$

may be rewritten as

$$\begin{aligned} s &= \frac{1}{2}\sum_{i=1}^{k}[w^2(t_i) - w^2(t_{i-1})] - \frac{1}{2}\sum_{i=1}^{k}[w(t_i) - w(t_{i-1})]^2 \\ &= \frac{1}{2}w^2(T) - \frac{1}{2}\sum_{i=1}^{k}\gamma_i, \end{aligned} \tag{11.3.6}$$

where

$$\gamma_i = [w(t_i) - w(t_{i-1})]^2.$$

To evaluate the last summation in (11.3.6), observe that, from the Chebyshev inequality (10.2.10),

$$\begin{aligned} \mathrm{prob}\left\{\left|\frac{1}{2}\sum_{i=1}^{k}\gamma_i - \frac{1}{2}\sum_{i=1}^{k}m_i\right| \geqslant \varepsilon\right\} &\leqslant \frac{1}{\varepsilon^2}D^2\left(\frac{1}{2}\sum_{i=1}^{k}\gamma_i\right) \\ &= \frac{1}{4\varepsilon^2}\sum_{i=1}^{k}D^2(\gamma_i), \end{aligned} \tag{11.3.7}$$

where $m_i = E(\gamma_i)$. Further, by (11.1.4),

$$E(\gamma_i) = E([w(t_i) - w(t_{i-1})]^2) = t_i - t_{i-1}$$

and, by equations (10.2.6) and (11.1.4),

$$D^2(\gamma_i) \leqslant E(\gamma_i^2) = E([w(t_i) - w(t_{i-1})]^4) = 3(t_i - t_{i-1})^2.$$

Thus,

$$\frac{1}{2}\sum_{i=1}^{k}m_i = \frac{1}{2}\sum_{i=1}^{k}(t_i - t_{i-1}) = \frac{T}{2}$$

and

$$\sum_{i=1}^{k}D^2(\gamma_i) \leqslant 3\sum_{i=1}^{k}(t_i - t_{i-1})^2 \leqslant 3T\max_i(t_i - t_{i-1}).$$

Setting $\delta(s) = \max_i(t_i - t_{i-1})$ as before and using (11.3.7), we finally obtain

$$\mathrm{prob}\left\{\left|\frac{1}{2}\sum_{i=1}^{k}\gamma_i - \frac{T}{2}\right| \geqslant \varepsilon\right\} \leqslant \frac{3T}{4\varepsilon^2}\delta(s)$$

or, from (11.3.6),

$$\text{prob}\left\{\left|s - \left(\frac{w^2(T)}{2} - \frac{T}{2}\right)\right| \geqslant \varepsilon\right\} \leqslant \frac{3T}{4\varepsilon^2}\delta(s).$$

If $\{s_n\}$ is a regular sequence, then $\delta(s_n)$ converges to zero as $n \to \infty$ and

$$\text{st-lim } s_n = \tfrac{1}{2}w^2(T) - \tfrac{1}{2}T.$$

Thus we have shown that

$$\int_0^T w(t)\,dw(t) = \tfrac{1}{2}w^2(T) - \tfrac{1}{2}T,$$

clearly demonstrating that the stochastic Itô integral does not obey the usual rules of integration. □

This last example illustrates the fact that the calculation of stochastic integrals is, in general, not an easy matter and requires many analytical tools that may vary from situation to situation. What is even more interesting is that the sufficient conditions for the existence of stochastic integrals related to the construction of nonanticipative σ-algebras are quite complicated in comparison with the Lebesgue integration of deterministic functions.

Remark 11.3.3. From Example 11.3.2, it is rather easy to demonstrate how the choice of the intermediate point $\bar{t}_i$ influences the value of the integral. For example, picking $\bar{t}_i = \frac{1}{2}(t_{i-1} + t_i)$, we obtain, in place of the Itô sum, the **Stratonovich sum,**

$$\begin{aligned} s &= \sum_{i=1}^{k} w(\tfrac{1}{2}(t_{i-1} + t_i))[w(t_i) - w(t_{i-1})] \\ &= \frac{1}{2}w^2(T) - \frac{1}{2}\sum_{i=1}^{k}\gamma_i + \frac{1}{2}\sum_{i=1}^{k}\rho_i, \end{aligned}$$

where

$$\gamma_i = [w(t_i) - w(\tfrac{1}{2}(t_{i-1} + t_i))]^2$$

and

$$\rho_i = [w(\tfrac{1}{2}(t_{i-1} + t_i)) - w(t_{i-1})]^2.$$

Since the variables $\gamma_1, \ldots, \gamma_k$ are independent as are $\rho_1, \ldots, \rho_k$, we may use the Chebyshev inequality as in the previous example to show that

$$\text{st-lim} \sum_{i=1}^{k}\gamma_i = \frac{1}{2}T = \text{st-lim} \sum_{i=1}^{k}\rho_i.$$

Thus the Stratonovich sums $\{s_n\}$ converge to $\frac{1}{2}w^2(T)$, and the Stratonovich integral gives a result more in accord with our experience from calculus. However, the use of the Stratonovich integral in solving stochastic differential equations leads to other more serious problems. □

To close this section, we extend our definition of the Itô integral to the multidimensional case. If $G(t) = (\eta_{ij}(t))$, $i,j = 1,\ldots,d$ is a $d \times d$ matrix of continuous stochastic processes, defined for $\alpha \leqslant t \leqslant \beta$, and $w(t) = (w_i(t))$, $i = 1,\ldots,d$, is a d-dimensional Wiener process, then

$$\int_\alpha^\beta G(t)\,dw(t) = \begin{pmatrix} \zeta_1 \\ \cdot \\ \cdot \\ \zeta_d \end{pmatrix}, \tag{11.3.8}$$

where

$$\zeta_i = \sum_{j=1}^{d} \int_\alpha^\beta \eta_{ij}(t)\,dw_j(t)$$

defines the Itô integral. Thus, equation (11.3.8) is integrated term by term. In this case the family $\{\mathcal{F}_t\}$ of nonanticipative σ-algebras must satisfy conditions (2) and (3) of Definition 11.3.1 with respect to all $\{w_i(t)\}$, $i = 1,\ldots,d$, and condition (4) of Definition 11.3.2 must be satisfied by all $\{\eta_{ij}(t)\}$, $i,j = 1,\ldots,d$.

11.4 The stochastic Itô integral: special cases

In the special case when the integrand of the Itô integral does not depend on ω, that is to say, it is not a stochastic process, the convergence of the approximating sums is quite strong. This section is devoted to an examination of this situation and one in which we are simply integrating a stochastic process with respect to t.

Before stating our first proposition, we note that, if $f\colon [\alpha,\beta] \to R$ is a continuous function, then every regular sequence $\{s_n\}$ of approximating sums

$$s_n = \sum_{i=1}^{k_n} f(\bar{t}_i^n)[w(t_i^n) - w(t_{i-1}^n)], \qquad \bar{t}_i^n \in [t_{i-1}^n, t_i^n]$$

converges in the mean [i.e., strongly in $L^2(\Omega)$] to the integral

$$\zeta = \int_\alpha^\beta f(t)\,dw(t). \tag{11.4.1}$$

Although we will not prove this assertion, it suffices to say that the proof proceeds in a fashion similar to the proof of the following proposition.

Proposition 11.4.1. If $f\colon [\alpha,\beta] \to R$ is a continuous function, then

$$E\left(\int_\alpha^\beta f(t)\,dw(t)\right) = 0 \tag{11.4.2}$$

and

$$D^2\left(\int_\alpha^\beta f(t)\,dw(t)\right) = \int_\alpha^\beta [f(t)]^2\,dt. \tag{11.4.3}$$

Proof: Set

$$s = \sum_{i=1}^{k} f(t_{i-1})[w(t_i) - w(t_{i-1})] = \sum_{i=1}^{k} f(t_{i-1})\,\Delta w_i\,,$$

where $\Delta w_i = w(t_i) - w(t_{i-1})$. Then

$$s^2 = \sum_{i,j=1}^{k} f(t_{i-1})f(t_{j-1})\,\Delta w_i\,\Delta w_j\,.$$

We have immediately that

$$E(s) = \sum_{i=1}^{k} f(t_{i-1})E(\Delta w_i) = 0$$

and, since $w(t)$ is a Wiener process with independent increments,

$$E(\Delta w_i\,\Delta w_j) = \begin{cases} E(\Delta w_i)E(\Delta w_j) = 0 & \text{if } i \neq j \\ E(\Delta w_i^2) = t_i - t_{i-1} & \text{if } i = j. \end{cases}$$

We also have

$$D^2(s) = E(s^2) = \sum_{i,j=1}^{k} f(t_{i-1})f(t_{j-1})E(\Delta w_i\,\Delta w_j)$$
$$= \sum_{i=1}^{k} [f(t_{i-1})]^2(t_i - t_{i-1})\,.$$

Thus for any regular sequence $\{s_n\}$,

$$\lim_{n\to\infty} E(s_n) = 0 \tag{11.4.4}$$

and

$$\lim_{n\to\infty} D^2(s_n) = \int_\alpha^\beta [f(t)]^2\,dt. \tag{11.4.5}$$

Since, from the remarks preceding the proposition, $\{s_n\}$ converges in mean to the integral ζ given in equation (11.4.1), we have $\lim_{n\to\infty} E(s_n) = E(\zeta)$ and $\lim_{n\to\infty} D^2(s_n) = D^2(\zeta)$, which, by (11.4.4) and (11.4.5), completes the proof. ∎

A second special case of the stochastic integral occurs when the integrand is a stochastic process but it is desired to have the integral only with respect to time. Hence we wish to consider

$$\zeta = \int_\alpha^\beta \eta(t)\,dt \tag{11.4.6}$$

when $\{\eta(t)\}$, $\alpha \leqslant t \leqslant \beta$, is a given stochastic process. To define (11.4.6) we consider approximating sums of the form

$$\bar{s} = \sum_{i=1}^{k} \eta(\bar{t}_i)(t_i - t_{i-1}),$$

corresponding to the partition

$$\alpha = t_0 < t_1 < \cdots < t_k = \beta$$

with arbitrary intermediate points $\bar{t}_i \in [t_{i-1}, t_i]$. We now have the following definition.

Definition 11.4.1. If every regular $[\delta(\bar{s}_n) \to 0]$ sequence $\{\bar{s}_n\}$ of approximating sums is stochastically convergent and

$$\zeta = \text{st-lim}\, \bar{s}_n, \tag{11.4.7}$$

then this common limit is called the **integral** of $\eta(t)$ on $[\alpha, \beta]$ and is denoted by (11.4.6).

Observe that, when $\eta(t, \omega)$ possesses continuous sample paths, that is, it is a continuous function of t, the limit

$$\lim_{n\to\infty} \bar{s}_n(\omega)$$

exists as the classical Riemann integral. Thus when $\{\eta(t)\}$, $\alpha \leqslant t \leqslant \beta$, is a continuous stochastic process, this limit exists for almost all ω. Further, since, by Proposition 10.3.2, almost sure convergence implies stochastic convergence, the limit (11.4.7) must exist.

There is an interesting connection between the Itô integral (11.3.5) and the integral of (11.4.6) reminiscent of the classical "integration by parts" formula. It can be stated formally as follows.

Proposition 11.4.2. If f: $[\alpha, \beta] \to R$ is differentiable with a continuous derivative f', then

$$\int_\alpha^\beta f(t)\,dw(t) = -\int_\alpha^\beta f'(t)w(t)\,dt + f(\beta)w(\beta) - f(\alpha)w(\alpha). \tag{11.4.8}$$

Proof: Since the integrals in (11.4.8) both exist we may pick special approximating sums of the form

$$\bar{s}_n = \sum_{i=1}^{k_n} f'(\bar{t}_i^n) w(\bar{t}_i^n)(t_i^n - t_{i-1}^n), \tag{11.4.9}$$

where the intermediate points $\bar{t}_i$ are chosen in such a way that

$$f(t_i^n) - f(t_{i-1}^n) = f'(\bar{t}_i^n)(t_i^n - t_{i-1}^n).$$

Substituting this expression into (11.4.9), we may rewrite $\bar{s}_n$ as

$$\begin{aligned} \bar{s}_n &= \sum_{i=1}^{k_n} [f(t_i^n) - f(t_{i-1}^n)] w(\bar{t}_i^n) \\ &= -\sum_{i=1}^{k_n - 1} [w(\bar{t}_{i+1}^n) - w(\bar{t}_i^n)] f(t_i^n) + f(t_{k_n}^n) w(\bar{t}_{k_n}^n) - f(t_0^n) w(\bar{t}_1^n). \end{aligned} \tag{11.4.10}$$

The sum on the right-hand side of (11.4.10) corresponds to the partition

$$\bar{t}_1^n < \cdots < \bar{t}_{k_n}^n$$

that does not contain intervals $(\alpha, \bar{t}_1^n)$ and $(\bar{t}_{k_n}^n, \beta)$. Setting $\bar{t}_0^n = \alpha$ and $\bar{t}_{k_n+1}^n = \beta$, we may rewrite (11.4.10) in the form

$$\bar{s}_n = -s_n + w(\beta) f(\beta) - w(\alpha) f(\alpha), \tag{11.4.11}$$

where

$$s_n = \sum_{i=0}^{k_n} [w(\bar{t}_{i+1}^n) - w(\bar{t}_i^n)] f(t_i^n).$$

The sequence $\{\bar{s}_n\}$ converges to

$$\int_\alpha^\beta f'(t) w(t)\, dt,$$

whereas $\{s_n\}$ converges, by our remarks preceding Proposition 11.4.1, to the integral

$$\int_\alpha^\beta f(t)\, dw(t).$$

Thus, passing to the limit in (11.4.11) finishes the proof. ■

Remark 11.4.1. In our short development of the Itô integral and presentation of its main properties, we have restricted ourselves to the special situation where the integrand is a continuous stochastic process. This allowed us to define the Itô integral in a relatively simple and direct way as the limit of the Itô sums (11.3.3). Generally, such an approach is inconvenient because of the restrictive nature of the assumption concerning the continuity of the integrand. Usually, the definition of the Itô integral is given in a more sophisticated manner: It is first defined for stochastic processes that are piecewise constant in time, and then, by using a limiting procedure in $L^2(\Omega)$ the definition is extended to a quite general class of integrands. An exhaustive treatment of this procedure may be found in Gikhman and Skorokhod [1969, 1975]. □

11.5 Stochastic differential equations

All the material developed in the previous sections was a necessary prelude to be able to study the stochastic differential equation

$$\frac{dx}{dt} = b(x) + \sigma(x)\xi \tag{11.5.1}$$

with initial condition

$$x(0) = x^0, \tag{11.5.2}$$

where

$$b(x) = \begin{pmatrix} b_1(x) \\ \vdots \\ b_d(x) \end{pmatrix} \quad \text{and} \quad \sigma(x) = \begin{pmatrix} \sigma_{11}(x) & \cdots & \sigma_{1d}(x) \\ \vdots & & \vdots \\ \sigma_{d1}(x) & \cdots & \sigma_{dd}(x) \end{pmatrix}$$

are given functions of x and

$$x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_d(t) \end{pmatrix}$$

is the unknown. In (11.5.1), the "white noise" vector

$$\xi = \begin{pmatrix} \dfrac{dw_1}{dt} \\ \vdots \\ \dfrac{dw_d}{dt} \end{pmatrix}$$

should be considered, from a mathematical point of view, as a pure symbol much like the letters "dt" in the notation for the derivative. However, from an application standpoint, ξ denotes a very specific process consisting of "infinitely" many independent, or random, impulses as discussed in Section 11.1. We assume that the initial vector x^0 and the Wiener process $\{w(t)\}$ are independent. To examine the solution of equations (11.5.1) and (11.5.2), we formally integrate (11.5.1) over the interval $[0, t]$ to give

$$x(t) = \int_0^t b(x(s))\, ds + \int_0^t \sigma(x(s))\, dw(s) + x^0. \tag{11.5.3}$$

Since the integrals that appear on the right-hand side of (11.5.3) are defined from our considerations of the previous sections, we are close to a formal definition of the solution.

First, however, it is necessary to choose a specific family of nonanticipative σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$. We may, for example, assume that $\mathcal{F}_t$ is the smallest σ-algebra containing all events of the form $\{\omega\colon w(u, \omega) \in B\}$ and $(x^0)^{-1}(B)$ for $0 \leq u \leq t$ and Borel sets B, that is, $\mathcal{F}_t$ is the smallest σ-algebra with respect to which $w(u)$, $0 \leq u \leq t$, and x^0 are measurable. This family is nonanticipative since conditions (1) and (2) of Definition 11.3.1 are evidently satisfied, and condition (3) follows from the fact that $\{w(t)\}$ is a process with independent increments and that x^0 and $\{w(t)\}$ are independent.

With this family of nonanticipative σ-algebras, we define the solution to equations (11.5.1) and (11.5.2).

Definition 11.5.1. A continuous stochastic process $\{x(t)\}_{t\geq 0}$ is called a **solution** of equations (11.5.1) and (11.5.2) if:

(a) $\{x(t)\}$ is nonanticipative, that is, it satisfies condition (4) of Definition 11.3.2; and
(b) For every $t \geq 0$, equation (11.5.3) is satisfied with probability 1.

It is well known from the theory of ordinary differential equations that it is necessary to assume some special conditions on the right-hand side in order to guarantee the existence and uniqueness of a solution. It is interesting that analogous conditions are also sufficient for stochastic differential equations. Thus we have the following theorem.

Theorem 11.5.1. If $b(x)$ and $\sigma(x)$ satisfy the Lipschitz conditions

$$|b(x) - b(y)| \leq L|x - y|, \qquad x, y \in R^d \tag{11.5.4}$$

and

$$|\sigma(x) - \sigma(y)| \leq L|x - y|, \qquad x, y \in R^d \tag{11.5.5}$$

with some constant L, then the initial value problem, equations (11.5.1) and (11.5.2), has a unique solution $\{x(t)\}_{t\geq 0}$.

Theorem 11.5.1 can be proved by the method of successive approximations as can the corresponding result for ordinary differential equations. Thus a sequence $\{x^i(t)\}_{i\geq 0}$ of stochastic processes would be defined with $x^0(t) = x^0$ and

$$x^i(t) = \int_0^t b(x^{i-1}(s))\,ds + \int_0^t \sigma(x^{i-1}(s))\,dw(s) + x^0.$$

Then, using the Lipschitz conditions (11.5.4) and (11.5.5), it is possible to evaluate the series

$$x(t) = \sum_{i=1}^{\infty} [x^i(t) - x^{i-1}(t)] + x^0$$

in $L^2(\Omega)$ norm by a convergent series of the form,

$$\sum_{n=0}^{\infty} k^n t^n / n!\,,$$

and to prove that $x(t)$ is, indeed, the desired solution. We omit the details as this proof is quite complicated, but a full proof may be found in Gikhman and Skorokhod [1969].

An alternative way to generate an approximating solution is to use the Euler linear extrapolation formula. Suppose that the solution $x(t)$ is given on the interval $[0, t_0]$. Then for values $t_0 + \Delta t$ larger than, but close to, t_0, we write

$$x(t_0 + \Delta t) = x(t_0) + b(x(t_0))\,\Delta t + \sigma(x(t_0))\,\Delta w, \tag{11.5.6}$$

where $\Delta w = w(t_0 + \Delta t) - w(t_0)$. (Observe that for an ordinary differential equation, this equation defines a ray tangent to the solution on $[0, t_0]$ at t_0.) In particular, when an interval $[0, T]$ is given, we may take a partition

$$0 = t_0 < \cdots < t_n = T$$

and define

$$\Delta x(t_i) = b(x(t_{i-1}))\,\Delta t_i + \sigma(x(t_{i-1}))\,\Delta w_i\,, \tag{11.5.7}$$

where $\Delta x(t_i) = x(t_i) - x(t_{i-1})$, $\Delta t_i = t_i - t_{i-1}$, $\Delta w_i = w(t_i) - w(t_{i-1})$, and $x(t_0) = x^0$.

It is evident that in some respects this approach is much simpler than the method of successive approximations, since no knowledge concerning the Itô integral is even necessary. Indeed, S. Bernstein employed this technique in his original investigations into stochastic differential equations, so we will call equations (11.5.6) and (11.5.7) the **Euler–Bernstein equations.**

Example 11.5.1. The oldest and best-known example of a stochastic differential equation is probably the **Langevin equation,**

$$\frac{dx}{dt} = -bx + \sigma\xi, \qquad x(0) = x^0, \tag{11.5.8}$$

where x is a scalar and the coefficients b and σ are constant.

By definition, the solution of (11.5.8) satisfies

$$x(t) = -b\int_0^t x(s)\,ds + \sigma\int_0^t dw(s) + x^0$$

or, using our calculations of Example 11.3.1,

$$x(t) = -b\int_0^t x(s)\,ds + \sigma w(t) + x^0. \tag{11.5.9}$$

Equation (11.5.9) is rather easy to deal with since it does not contain an Itô integral, and, since the one integral that does appear exists for almost all ω taken separately, we may use the usual rules of calculus.

Setting

$$z(t) = \int_0^t x(s)\,ds, \tag{11.5.10}$$

equation (11.5.9) becomes, for almost all ω,

$$\frac{dz}{dt} = -bz(t) + \sigma w(t) + x^0.$$

For fixed ω, this is an ordinary differential equation and, thus,

$$z(t) = \int_0^t e^{-b(t-s)}(\sigma w(s) + x^0)\,ds. \tag{11.5.11}$$

Combining equations (11.5.9) through (11.5.11) after some manipulation, yields

$$x(t) = x^0 e^{-bt} - b\sigma\int_0^t e^{-b(t-s)} w(s)\,ds + \sigma w(t).$$

Using the integration by parts formula (11.4.8), this becomes

$$x(t) = x^0 e^{-bt} + \sigma\int_0^t e^{-b(t-s)}\,dw(s). \tag{11.5.12}$$

From (11.5.12) and (11.4.2), it follows that

$$E(x(t)) = e^{-bt}E(x^0)$$

and, taking note of the independence of x^0 and $w(t)$,

$$D^2(x(t)) = e^{-2bt}D^2(x^0) + \sigma^2 D^2\left(\int_0^t e^{-b(t-s)}\,dw(s)\right).$$

With (11.4.3), this finally reduces to

$$D^2(x(t)) = e^{-2bt}D^2(x^0) + \sigma^2\int_0^t e^{-2b(t-s)}\,ds$$
$$= e^{-2bt}D^2(x^0) + (\sigma^2/2b)[1 - e^{-2bt}].$$

Thus, for the Langevin equation,

$$\lim_{t\to\infty} D^2(x(t)) = \sigma^2/2b.$$

This asymptotic property of the variance is a special case of a more general result that we will establish in the next section where we examine uses of the Fokker–Planck equation. □

11.6 The Fokker–Planck (Kolmogorov forward) equation

The preceding sections were aimed at obtaining an understanding of the dynamical system

$$\frac{dx}{dt} = b(x) + \sigma(x)\xi \tag{11.6.1}$$

with

$$x(0) = x^0 \tag{11.6.2}$$

under a stochastic perturbation ξ. This required us to first introduce the abstract concept of nonanticipative σ-algebras. Then we had to define the Itô integral, which is generally quite difficult to calculate. Finally we gave the solution to equations (11.6.1)–(11.6.2) in terms of a general formula, generated by the method of successive approximations, which contains infinitely many Itô integrals.

In this section we extend this to a discussion of the density function of the process $x(t)$, which is a solution of (11.6.1) and (11.6.2). This density is defined as the function $u(t, x)$ that satisfies

$$\text{prob}\{x(t) \in B\} = \int_B u(t, z)\,dz. \tag{11.6.3}$$

The uniqueness of $u(t, x)$ follows immediately from Proposition 2.2.1, but the existence requires some regularity conditions on the coefficients $b(x)$ and $\sigma(x)$, which are given in the following. We will also show how $u(t, x)$ can be found without any knowledge concerning the solution $x(t)$ of the stochastic differential equations (11.6.1) with (11.6.2). It will turn out that $u(t, x)$ is given by the solution of a partial differential equation, known as the Fokker–Planck (or Kolmogorov forward) equation and that it is completely specified by the coefficients $b(x)$ and $\sigma(x)$ of equation (11.6.1).

Now set

$$a_{ij}(x) = \sum_{k=1}^{d} \sigma_{ik}(x)\sigma_{jk}(x). \tag{11.6.4}$$

From (11.6.4) it is clear that $a_{ij} = a_{ji}$ and, thus, the quadratic form,

$$\sum_{i,j=1}^{d} a_{ij}(x)\lambda_i\lambda_j , \tag{11.6.5}$$

is symmetric. Further, since

$$\sum_{i,j=1}^{d} a_{ij}(x)\lambda_i\lambda_j = \sum_{k=1}^{d}\left(\sum_{i=1}^{d} \sigma_{ik}(x)\lambda_i\right)^2,$$

(11.6.5) is nonnegative.

We are now ready to state the main theorem of this section, which gives the Fokker–Planck equation.

Theorem 11.6.1. If the functions σ_{ij}, $\partial\sigma_{ij}/\partial x_k$, $\partial^2\sigma_{ij}/\partial x_k\partial x_l$, b_i, $\partial b_i/\partial x_j$, $\partial u/\partial t$, $\partial u/\partial x_i$, and $\partial^2 u/\partial x_i\,\partial x_j$ are continuous for $t > 0$ and $x \in R^d$, and if b_i, σ_{ij} and their first derivatives are bounded, then $u(t,x)$ satisfies the equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i\,\partial x_j}(a_{ij}u) - \sum_{i=1}^{d}\frac{\partial}{\partial x_i}(b_i u), \qquad t > 0,\ x \in R^d. \tag{11.6.6}$$

Equation (11.6.6) is called the **Fokker–Planck equation** or **Kolmogorov forward equation.**

Remark 11.6.1. In Theorem 11.6.1 we assumed $\partial b_i/\partial x_j$ and $\partial\sigma_{ij}/\partial x_k$ were bounded since this implies the Lipschitz conditions (11.5.4) and (11.5.5) which, in turn, guarantee the existence and uniqueness of the solution to the stochastic equations (11.6.1) with (11.6.2). In order to assure the existence and differentiability of u, it is sufficient, for example, that a_{ij} and b_i, together with their derivatives up to the third order, are continuous, bounded, and satisfy the uniform parabolicity condition (11.7.5). □

Proof of Theorem 11.6.1: We will use the Euler–Bernstein approximation formula (11.5.6) in the proof of this theorem as it allows us to derive (11.6.6) in an extremely simple and transparent fashion.

Thus let $t_0 > 0$ be arbitrary, and let $x(t)$ be the solution to equations (11.6.1) and (11.6.2) on the interval $[0, t_0]$. Define $x(t)$ on $[t_0, t_0 + \varepsilon]$ by

$$x(t_0 + \Delta t) = x(t_0) + b(x(t_0))\,\Delta t + \sigma(x(t_0))\,[w(t_0 + \Delta t) - w(t_0)]\,, \tag{11.6.7}$$

where $0 \leqslant \Delta t \leqslant \varepsilon$ and ε is a positive number. We assume (and this is the only additional assumption needed for simplifying the proof) that $x(t)$, extended according to (11.6.7), has a density $u(t, x)$ for $0 \leqslant t \leqslant t_0 + \varepsilon$ and that for $t = t_0$, $u_t(t, x)$ exists. Observe that at the point $t = t_0$, $u(t, x)$ (and $u_t(t, x)$) is simultaneously the density (and its derivative) for the exact and for the extended solution.

Now let $h: R^d \to R$ be a C^3 function with compact support. We wish to calculate the mathematical expectation of $h(x(t_0 + \Delta t))$. First note that since $u(t_0 + \Delta t, x)$ is the density of $x(t_0 + \Delta t)$, we have, by (10.2.2),

$$E(h(x(t_0 + \Delta t))) = \int_{R^d} h(x)u(t_0 + \Delta t, x)\,dx\,. \tag{11.6.8}$$

However, using equation (11.6.7), we may write the random variable $h(x(t_0 + \Delta t))$ in the form

$$h(x(t_0 + \Delta t)) = h(Q(x(t_0), w(t_0 + \Delta t) - w(t_0)))\,, \tag{11.6.9}$$

where

$$Q(x, y) = x + b(x)\,\Delta t + \sigma(x)y\,.$$

The variables $x(t_0)$ and $\Delta w(t_0) = w(t_0 + \Delta t) - w(t_0)$ are independent for each $0 \leqslant \Delta t \leqslant \varepsilon$ since $x(t_0)$ is $\mathcal{F}_{t_0}$ measurable and $\Delta w(t_0)$ is independent with respect to $\mathcal{F}_{t_0}$. Thus the random vector $(x(t_0), \Delta w(t_0))$ has the joint density

$$u(t_0, x)g(\Delta t, y)\,,$$

where g is given by (11.1.3). As a consequence, the mathematical expectation of (11.6.9) is given by

$$\int_{R^d}\int_{R^d} h(Q(x, y))u(t_0, x)g(\Delta t, y)\,dx\,dy$$
$$= \int_{R^d}\int_{R^d} h(x + b(x)\,\Delta t + \sigma(x)y)u(t_0, x)g(\Delta t, y)\,dx\,dy\,.$$

From this and (11.6.8), we obtain

$$\int_{R^d} h(x)u(t_0 + \Delta t, x)\,dx$$
$$= \int_{R^d}\int_{R^d} h(x + b(x)\,\Delta t + \sigma(x)y)u(t_0, x)g(\Delta t, y)\,dx\,dy\,.$$

By developing h in a Taylor expansion, we have

$$\int_{R^d} h(x)u(t_0 + \Delta t, x)\,dx = \int_{R^d}\int_{R^d}\Big\{h(x) + \sum_{i=1}^{d}\frac{\partial h}{\partial x_i}[b_i(x)\,\Delta t + (\sigma(x)y)_i]$$
$$+ \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2 h}{\partial x_i\,\partial x_j}[b_i(x)\,\Delta t + (\sigma(x)y)_i]$$
$$\cdot [b_j(x)\,\Delta t + (\sigma(x)y)_j] + r(\Delta t)\Big\}$$
$$\cdot\, u(t_0, x)g(\Delta t, y)\,dx\,dy, \tag{11.6.10}$$

where $r(\Delta t)$ denotes the remainder and $(\sigma(x)y)_i$ is the ith coordinate of the vector $\sigma(x)y$.

On the right-hand side of (11.6.10) we have a finite collection of integrals that we will first integrate with respect to y. Observe that

$$(\sigma(x)y)_i(\sigma(x)y)_j = \sum_{k,l=1}^{d}\sigma_{ik}(x)\sigma_{jl}(x)y_k y_l\,.$$

By equation (11.2.3)

$$\int_{R^d} g(\Delta t, y)\,dy = 1\,,$$

whereas from (11.2.4)

$$\int_{R^d} (\sigma(x)y)_i g(\Delta t, y)\,dy = 0\,.$$

Finally, from (11.2.5), we have

$$\int_{R^d} (\sigma(x)y)_i(\sigma(x)y)_j g(\Delta t, y)\,dy = a_{ij}(x)\,\Delta t,$$

where a_{ij} is as defined in (11.6.4). By combining all of these results, we can write equation (11.6.10) as

$$\int_{R^d} h(x)\,[u(t_0 + \Delta t, x) - u(t_0, x)]\,dx$$
$$= \Delta t\int_{R^d}\Big\{\sum_{i=1}^{d}\frac{\partial h}{\partial x_i}b_i(x) + \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2 h}{\partial x_i\,\partial x_j}a_{ij}(x)\Big\}u(t_0, x)\,dx + R(\Delta t), \tag{11.6.11}$$

where the new remainder $R(\Delta t)$ is

$$R(\Delta t) = \frac{1}{2}\int_{R^d}\sum_{i,j=1}^{d}\frac{\partial^2 h}{\partial x_i\,\partial x_j}b_i(x)b_j(x)\,(\Delta t)^2 u(t_0, x)\,dx$$
$$+ \int_{R^d}\int_{R^d} r(\Delta t)u(t_0, x)g(\Delta t, y)\,dx\,dy. \tag{11.6.12}$$

It is straightforward to show that $R(\Delta t)/\Delta t$ goes to zero as $\Delta t \to 0$. The first integral on the right-hand side of (11.6.12) contains $(\Delta t)^2$, so this is easy. The second integral may be evaluated by using the classical formula for the remainder $r(\Delta t)$:

$$r(\Delta t) = \sum_{i,j,k=1}^{d} \frac{\partial^3 h}{\partial x_i\, \partial x_j\, \partial x_k}\bigg|_z [b_i\, \Delta t + (\sigma y)_i] \cdot [b_j\, \Delta t + (\sigma y)_j][b_k\, \Delta t + (\sigma y)_k].$$

The third derivatives of h are evaluated at some intermediate point z, which is irrelevant because we only use the fact that these derivatives are bounded since h is of compact support.

All of the components appearing in $r(\Delta t)$ can be evaluated by terms of the form

$$M(\Delta t)^3,\ M(\Delta t)^2 |y_i|,\ M(\Delta t)\,|y_i y_j|,\ M|y_i y_j y_k|,$$

where M is a constant. To evaluate $R(\Delta t)$ we must integrate these terms with respect to x and y. Using

$$\int_{-\infty}^{\infty} |z|^n g(\Delta t, z)\, dz = \alpha_n (\Delta t)^{n/2},$$

where the constants α_n depend only on n, integration of $M(\Delta t)^3$ again gives $M(\Delta t)^3$ since $u(t_0, x)$ and $g(\Delta t, y)$ are both densities. Integration of $M(\Delta t)^2|y_i|$ gives $M(\Delta t)^2 C_i (\Delta t)^{1/2}$, where $C_i = \alpha_1$. Analogously, integration of the third term gives $M(\Delta t) C_{ij}(\Delta t)$, whereas the fourth yields $MC_{ijk}(\Delta t)^{3/2}$, where C_{ij} depends on α_1 and α_2, and C_{ijk} depends on α_1, α_2, and α_3. All these terms divided by Δt approach zero as $\Delta t \to 0$.

Returning to (11.6.11), dividing by Δt and passing to the limit as $\Delta t \to 0$, we obtain

$$\int_{R^d} h(x) \frac{\partial u}{\partial t}\, dx = \int_{R^d} \left\{ \sum_{i=1}^{d} \frac{\partial h}{\partial x_i} b_i(x) + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 h}{\partial x_i\, \partial x_j} a_{ij}(x) \right\} u(t_0, x)\, dx. \tag{11.6.13}$$

Since h has compact support we may easily integrate the right-hand side of (11.6.13) by parts. Doing this and shifting all terms to the left-hand side, we finally have

$$\int_{R^d} h(x) \left\{ \frac{\partial u(t_0, x)}{\partial t} + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} [b_i(x) u(t_0, x)] - \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i\, \partial x_j} [a_{ij}(x) u(t_0, x)] \right\} dx = 0. \tag{11.6.14}$$

Since $h(x)$ is a C^3 function with compact support, but otherwise arbitrary, the integral condition (11.6.14), which is satisfied for every such h implies that the term in braces vanishes. This completes the proof that $u(t_0, x)$ satisfies equation (11.6.6). ■

Remark 11.6.2. To deal with the stochastic differential equations (11.6.1) with (11.6.2), we were forced to introduce many abstract and difficult concepts. It is ironic that, once we pass to a consideration of the density function $u(t, x)$ of the random process $x(t)$, all this material becomes unnecessary, as we must only insert the appropriate coefficients a_{ij} and b_i into the Fokker–Planck equation (11.6.6)! □

11.7 Properties of the solutions of the Fokker–Planck equation

As we have shown in the previous section, the density function $u(t, x)$ of the solution $x(t)$ of the stochastic differential equation (11.6.1) with (11.6.2) satisfies the partial differential equation (11.6.6). Moreover, if the initial condition $x(0) = x^0$, which is a random variable, has a density f then $u(0, x) = f(x)$. Thus, to understand the behavior of the densities $u(t, x)$, we must study the initial-value (Cauchy) problem:

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \, \partial x_j} [a_{ij}(x)u] - \sum_{i=1}^{d} \frac{\partial}{\partial x_i} [b_i(x)u], \qquad t > 0, \, x \in R^d, \tag{11.7.1}$$

$$u(0, x) = f(x), \qquad x \in R^d. \tag{11.7.2}$$

Observe that equation (11.7.1) is of second order and may be rewritten in the form

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \, \partial x_j} + \sum_{i=1}^{d} \tilde{b}_i(x) \frac{\partial u}{\partial x_i} + \tilde{c}(x)u, \; t > 0, \quad x \in R^d, \tag{11.7.3}$$

where

$$\tilde{b}_i(x) = -b_i(x) + \sum_{j=1}^{d} \frac{\partial a_{ij}(x)}{\partial x_j}$$

and (11.7.4)

$$\tilde{c}(x) = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 a_{ij}(x)}{\partial x_i \, \partial x_j} - \sum_{i=1}^{d} \frac{\partial b_i}{\partial x_i}.$$

As was shown in Section 11.6, the quadratic form

$$\sum_{i,j=1}^{d} a_{ij}(x) \lambda_i \lambda_j,$$

corresponding to the term in (11.7.3) with second-order derivatives, is always nonnegative. We will assume the somewhat stronger inequality,

$$\sum_{i,j=1}^{d} a_{ij}(x)\lambda_i\lambda_j \geqslant \rho \sum_{i=1}^{d} \lambda_i^2\,, \tag{11.7.5}$$

where ρ is a positive constant, holds. This is called the **uniform parabolicity condition.**

It is known that, if the coefficients a_{ij}, $\tilde{b}_i$, and $\tilde{c}_i$ are smooth and satisfy the growth conditions

$$|a_{ij}(x)| \leqslant M, \quad |\tilde{b}_i(x)| \leqslant M(1 + |x|), \quad |\tilde{c}(x)| \leqslant M(1 + |x|^2)\,, \tag{11.7.6}$$

then the classical solution of the Cauchy problem, equations (11.7.2) and (11.7.3), is unique and given by the integral formula

$$u(t,x) = \int_{R^d} \Gamma(t,x,y) f(y)\,dy\,, \tag{11.7.7}$$

where the kernel Γ, called the **fundamental solution,** is independent of the initial density function f.

To state these results precisely, we need to give a formal definition of the classical solution.

Definition 11.7.1. Let $f\colon R^d \to R$ be a bounded continuous function. A function $u(t,x)$, $t > 0$, $x \in R^d$, is called a **classical solution** of equations (11.7.2) and (11.7.3) if it satisfies the following conditions:

(a) For every $T > 0$, $u(t,x)$ is bounded for $t \in (0,T]$, $x \in R^d$;
(b) $u(t,x)$ has continuous derivatives u_t, u_x, $u_{x_i x_j}$ and satisfies equation (11.7.3) for every $t > 0$, $x \in R^d$; and
(c) $$\lim_{t\to 0} u(t,x) = f(x)\,. \tag{11.7.8}$$

Condition (a) is necessary because in the space of unbounded functions, the Cauchy problem, even for the heat equation $u_t = \frac{1}{2}\sigma^2 u_{xx}$, is not uniquely determined. Condition (b) is obvious, and (c) is necessary since (11.7.3) is satisfied only for $t > 0$ and, thus, the values of $u(t,x)$ for $t > 0$ must be related to the initial condition $u(0,x) = f(x)$.

Eidelman [1969] proves the following simple, but quite general, result concerning the existence and uniqueness of the classical solution (see also Friedman [1964]).

Theorem 11.7.1. Assume that the coefficients of equation (11.7.3) are C^3 functions and satisfy the following conditions:

(a) The a_{ij} satisfy the uniform parabolicity condition (11.7.5);
(b) The a_{ij}, $\tilde{b}_i$, and $\tilde{c}$ satisfy the growth conditions (11.7.6).

Then, for every continuous bounded f, there is a unique classical solution of the problem, equations (11.7.2) and (11.7.3), which is given by formula (11.7.7). The kernel $\Gamma(t,x,y)$, defined for $t > 0$, $x,y \in R^d$, is continuous and differentiable with respect to t, is twice differentiable with respect to x_i, and satisfies (11.7.3) as a function of (t,x) for every fixed y. Further, in every strip $0 < t \leqslant T$, $x,y \in R^d$, Γ satisfies the inequalities

$$\begin{aligned} &0 < \Gamma(t,x,y) \leqslant \Phi(t,x-y), &&\left|\frac{\partial\Gamma}{\partial t}\right| \leqslant \Phi(t,x-y), \\ &\left|\frac{\partial\Gamma}{\partial x_i}\right| \leqslant \Phi(t,x-y), &&\left|\frac{\partial^2\Gamma}{\partial x_i\,\partial x_j}\right| \leqslant \Phi(t,x-y), \end{aligned} \tag{11.7.9}$$

where

$$\Phi(t,x-y) = kt^{-(n+2)/2}\exp[-\delta(x-y)^2/t] \tag{11.7.10}$$

and the constants δ and k depend on T.

The explicit construction of the fundamental solution Γ for general coefficients a_{ij}, $\tilde{b}_i$, and $\tilde{c}$ is usually impossible. It is easy only for some special cases, such as the heat equation,

$$u_t = (\sigma^2/2)u_{xx}.$$

In this case, Γ is the familiar kernel

$$\Gamma(t,x,y) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp[-(x-y)^2/2\sigma^2 t].$$

Nevertheless, the properties (11.7.9) of Γ given in Theorem 11.7.1 allow us to deduce some very interesting properties of the solutions $u(t,x)$.

Let f be a continuous function with bounded support, say on the ball $B_r = \{x: |x| \leqslant r\}$, and let u be the corresponding solution of equations (11.7.2) and (11.7.3). Then, from the first inequality (11.7.9), we have

$$|u(t,x)| \leqslant \int_{B_r}\Gamma(t,x,y)\,|f(y)|\,dy \leqslant M\int_{B_r}\Phi(t,x-y)\,dy$$

where $M = \max_y |f|$. Further, since $|x-y|^2 \geqslant \frac{1}{2}x^2 - r^2$ for $|y| \leqslant r$, we have

$$\int_{B_r}\Phi(t,x-y)\,dy \leqslant kt^{-(n+2)/2}\exp(-\delta[\tfrac{1}{2}|x|^2 - r^2]/t)\,|B_r|$$

and, consequently,

$$|u(t,x)| \leqslant Kt^{-(n+2)/2} \exp(-\tfrac{1}{2}\delta|x|^2/t),$$

where $K = kMe^{\delta r^2}|B_r|$. By using the remaining inequalities (11.7.9), we may derive analogous inequalities for the derivatives of u as summarized in

$$|u|, |u_t|, |u_{x_i}|, |u_{x_ix_j}| \leqslant Kt^{-(n+2)/2} \exp(-\tfrac{1}{2}\delta|x|^2/t) \tag{11.7.11}$$

These inequalities are quite important for they allow us to multiply equation (11.7.3) by any function that increases more slowly than $\exp(-\frac{1}{2}|x|^2)$ decreases (e.g., $x, x^2, \ldots, e^{rx}$), and then to integrate term by term to, for example, calculate the moments of $u(t,x)$.

Example 11.7.1. Again consider the Langevin equation

$$\frac{dx}{dt} = -bx + \sigma\xi$$

first introduced in Example 11.5.1. The corresponding Fokker–Planck equation is

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sigma^2\frac{\partial^2 u}{\partial x^2} + b\frac{\partial}{\partial x}(xu). \tag{11.7.12}$$

Multiply (11.7.12) by x^n and integrate to obtain

$$\frac{d}{dt}\int_{-\infty}^{\infty} x^n u\,dx = \frac{1}{2}\sigma^2\int_{-\infty}^{\infty} x^n\frac{\partial^2 u}{\partial x^2}\,dx + b\int_{-\infty}^{\infty} x^n\frac{\partial}{\partial x}(xu)\,dx.$$

Since by our foregoing discussion, u and its derivatives decay exponentially as $|x| \to \infty$, we can integrate by parts to give

$$\frac{d}{dt}\int_{-\infty}^{\infty} x^n u\,dx = \frac{1}{2}\sigma^2 n(n-1)\int_{-\infty}^{\infty} x^{n-2}u\,dx - nb\int_{-\infty}^{\infty} x^n u\,dx. \tag{11.7.13}$$

Let

$$m_n(t) = \int_{-\infty}^{\infty} x^n u(t,x)\,dx$$

be the nth moment of the function $u(t,x)$. From (11.7.13) we thus have an infinite system of ordinary differential equations in the moments,

$$\frac{dm_0}{dt} = 0, \qquad \frac{dm_1}{dt} = -bm_1,$$

$$\frac{dm_n}{dt} = \tfrac{1}{2}\sigma^2 n(n-1)m_{n-2} - nbm_n, \qquad n \geqslant 2,$$

which can be solved sequentially. Assuming that the initial function f is a density, we have

$$m_0(t) = m_0(0) = \int_{-\infty}^{\infty} f\,dx = 1\,,$$

$$m_1(t) = C_1 e^{-bt}, \qquad C_1 = m_1(0)\,,$$

$$m_2(t) = \frac{\sigma^2}{2b} + C_2 e^{-2bt}, \qquad C_2 = m_2(0) - \frac{\sigma^2}{2b}\,,$$

$$m_3(t) = C_3 e^{-3bt} + \frac{3C_1\sigma^2}{2b}(e^{-bt} - e^{-3bt}), \qquad C_3 = m_3(0)\,.$$

Successive formulas for higher moments become progressively more complicated. However, it is straightforward to demonstrate inductively that

$$\lim_{t\to\infty} m_n(t) = \begin{cases} 1\cdot 3\cdot 5\cdots(n-1)\left(\dfrac{\sigma^2}{2b}\right)^{n/2}, & \text{for } n \text{ even} \\ 0, & \text{for } n \text{ odd}\,. \end{cases}$$

Thus the limiting moments are the same as the moments of the Gaussian density

$$g_{\sigma b}(x) = \sqrt{b/\pi\sigma^2}\,\exp(-bx^2/\sigma^2)\,.$$

At the end of the next section it will become clear that not only do the moments of the solution of equation (11.7.12) converge to the moments of the Gaussian density $g_{\sigma b}$, but also that $u(t,x) \to g_{\sigma b}(x)$ as $t \to \infty$. □

Remark 11.7.1. A comparison of the discrete and continuous time systems with stochastic perturbations considered here reveals a close analogy between the dynamical laws, equations (10.4.1) and (10.5.1), and the stochastic differential equation (11.6.1) as well as between equations (10.4.3) and (10.5.4), for the evolution of densities, and the Cauchy problem, equations (11.7.1) and (11.7.2). □

11.8 Semigroups of Markov operators generated by parabolic equations

In this section we examine the solutions of the Fokker–Planck equation as a flow of densities governed by a semigroup of Markov operators. We start with the following definition.

Definition 11.8.1. Assume that the coefficients a_{ij}, $\tilde{b}_i$, and $\tilde{c}$ of (11.7.3) satisfy the conditions of Theorem 11.7.1. Then, for every $f \in L^1$, not necessarily continuous, the function

$$u(t,x) = \int_{R^d} \Gamma(t,x,y) f(y)\, dt \tag{11.8.1}$$

will be called a **generalized solution** of equations (11.7.2) and (11.7.3).

Since $\Gamma(t,x,y)$, as a function of (t,x), satisfies (11.7.3) for $t > 0$, $u(t,x)$ has the same property. However, if f is discontinuous, then condition (11.7.8) might not hold at a point of discontinuity.

Having a generalized solution Γ of equations (11.7.2) and (11.7.3), we define a family of operators $\{P_t\}_{t\geq 0}$ by

$$P_0 f(x) = f(x), \qquad P_t f(x) = \int_{R^d} \Gamma(t,x,y) f(y)\, dy. \tag{11.8.2}$$

We will now show that, from the properties of Γ stated in Theorem 11.7.1, we obtain the following corollary.

Corollary 11.8.1. The family of operators $\{P_t\}_{t\geq 0}$ is a stochastic semigroup, that is,

(1) $P_t(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 P_t f_1 + \lambda_2 P_t f_2, f_1 f_2 \in L^1$;
(2) $P_t f \geq 0$ for $f \geq 0$;
(3) $\|P_t\| = \|f\|$ for $f \geq 0$;
(4) $P_{t_1+t_2} f = P_{t_1}(P_{t_2} f), f \in L^1$.

Proof: Properties (1) and (2) follow immediately from equation (11.8.1) since the right-hand side is an integral operator with a positive kernel.

To verify (3), first assume that f is continuous with compact support. By multiplying the Fokker–Planck equation by a C^2 bounded function $h(x)$ and integrating, we obtain

$$\int_{R^d} h(x) u_t\, dx = \int_{R^d} h(x) \left\{ \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i\, \partial x_j}(a_{ij} u) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}(b_i u) \right\} dx$$

and integration by parts gives

$$\int_{R^d} h(x) u_t\, dx = \int_{R^d} \left\{ \frac{1}{2} \sum_{i,j=1}^{d} a_{ij} \frac{\partial^2 h}{\partial x_i\, \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial h}{\partial x_i} \right\} u\, dx.$$

Setting $h = 1$, we have

$$\frac{d}{dt} \int_{R^d} u\, dx = \int_{R^d} u_t\, dx = 0.$$

Since $u \geqslant 0$ for $f \geqslant 0$, we have

$$\frac{d}{dt}\|u\| = 0 \qquad \text{for } t > 0.$$

Further, the initial condition (11.7.8), inequality (11.7.11) and the boundedness of u imply, by the Lebesgue dominated convergence theorem, that $\|P_t f\|$ is continuous at $t = 0$. This proves that $\|P_t f\|$ is constant for all $t \geqslant 0$. If $f \in L^1$ is an arbitrary function, we can choose a sequence $\{f_k\}$ of continuous functions with compact support that converges strongly to f. Now,

$$|(\|P_t f\| - \|f\|)| \leqslant |(\|P_t f\| - \|P_t f_k\|)| + |(\|P_t f_k\| - \|f_k\|)| + \|f_k - f\|. \tag{11.8.3}$$

Since, as we just showed, P_t preserves the norm, the term $\|P_t f\| - \|f_k\|$ is zero. To evaluate the first term, note that

$$\begin{aligned} |(\|P_t f\| - \|P_t f_k\|)| &\leqslant \|P_t f - P_t f_k\| \\ &\leqslant \int_{R^d} \Gamma(t,x,y)\|f - f_k\|\,dy \leqslant M_t\|f - f_k\|, \end{aligned}$$

where $M_t = \sup_{x,y} \Gamma$. Thus the right-hand side of (11.8.3) converges strongly to zero as $k \to \infty$. Since the left-hand side is independent of k, we have $\|P_t f\| = \|f\|$, which completes the proof of (3). As we know, conditions (1)–(3) imply that $\|P_t f\| \leqslant \|f\|$ for all f and, thus, the operators P_t are continuous.

Finally to prove (4), again assume f is a continuous function with compact support and set $\bar{u}(t,x) = u(t + t_1, x)$. An elementary calculation shows that $\bar{u}(t,x)$ satisfies the Fokker–Planck equation with the initial condition $\bar{u}(0,x) = u(t_1,x)$. Thus, by the uniqueness of solutions to the Fokker–Planck equation,

$$u(t + t_1, x) = P_t u(t_1, x)$$

and, at the same time,

$$u(t + t_1, x) = P_{t+t_1} f(x) \quad \text{and} \quad u(t_1, x) = P_{t_1} f(x).$$

From these it is immediate that

$$P_{t+t_1} f = P_t(P_{t_1} f),$$

which proves (4) for all continuous f with compact support. If $f \in L^1$ is arbitrary, we again pick a sequence $\{f_k\}$ of continuous functions with compact supports that converges strongly to f and for which

$$P_{t_2+t_1} f_k = P_{t_2}(P_{t_1} f_k)$$

holds. Since the P_t have been shown to be continuous, we may pass to the limit of $k \to \infty$ and obtain (4) for arbitrary f. ■

Remark 11.8.1. In developing the material of Theorems 11.6.1, 11.7.1, and Corollary 11.8.1, we have passed from the description of $u(t, x)$ as the density of the random variable $x(t)$, through a derivation of the Fokker–Planck equation for $u(t, x)$ and then shown that the solutions of the Fokker–Planck equation define a stochastic semigroup $\{P_t\}_{t\geq 0}$. This semigroup describes the behavior of the dynamical system, equations (11.6.1) and (11.6.2). In actuality, our proof of Theorem 11.6.1 shows that the right-hand side of the Fokker–Planck equation is the infinitesimal operator for $P_t f$, although our results were not stated in this fashion. Further, Theorem 11.7.1 and Corollary 11.8.1 give the construction of the semigroup generated by this infinitesimal operator. □

Remark 11.8.2. Observe that, when the stochastic perturbation disappears ($\sigma_{ij} = 0$), then the Fokker–Planck equation reduces to the Liouville equation and $\{P_t\}$ is simply the semigroup of Frobenius–Perron operators corresponding to the dynamical system

$$\frac{dx_i}{dt} = b_i(x), \qquad i = 1, \ldots, d. \quad \square$$

11.9 Asymptotic stability of solutions of the Fokker–Planck equation

As we have seen, the fundamental solution Γ may be extremely useful. However, since a formula for Γ is not available in the general case, it is not of much use in the determination of asymptotic stability properties of $u(t, x)$. Thus, we would like to have other techniques available, and in this section we develop the use of Liapunov functions for this purpose, following Dłotko and Lasota [in press].

Here, by a **Liapunov function** we mean any function $V: R^d \to R$ that satisfies the following four properties:

(1) $V(x) \geq 0$ for all x;
(2) $\lim_{|x|\to\infty} V(x) = \infty$;
(3) V has continuous derivatives $(\partial V/\partial x_i)$, $(\partial^2 V/\partial x_i\, \partial x_j)$, $i, j = 1, \ldots, d$; and
(4) $$V(x) \leq \rho e^{\delta|x|} \quad \text{and} \quad \left|\frac{\partial V(x)}{\partial x_i}\right| \leq \rho e^{\delta|x|} \tag{11.9.1}$$

for some constants ρ, δ.

Conditions (1)–(4) are not very restrictive, for example, any positive definite quadratic form (of even order m)

$$V(x) = \sum_{i_1,\ldots,i_m=1}^{d} a_{i_1\ldots i_m} x_{i_1} \cdots x_{i_m}$$

is a Liapunov function. Our main purpose will be to use a Liapunov function V that satisfies the differential inequality

$$\sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 V}{\partial x_i\, \partial x_j} + \sum_{i=1}^{d} b_i(x) \frac{\partial V}{\partial x_i} \leqslant -\alpha V(x) + \beta \tag{11.9.2}$$

with positive constants α and β. Specifically, we can state the following theorem.

Theorem 11.9.1. Assume that the coefficients a_{ij}, $\tilde{b}_i$, and $\tilde{c}$ of equation (11.7.3) satisfy conditions (a) and (b) of Theorem 11.7.1 and that there is a Liapunov function V satisfying (11.9.2). Then the stochastic semigroup $\{P_t\}_{t\geqslant 0}$ defined by the generalized solution of the Fokker–Planck equation and given in (11.8.2) is asymptotically stable.

Proof: The proof is similar to that of Theorem 5.7.1. First pick a continuous density f with compact support and then consider the mathematical expectation of V calculated with respect to the solution u of equations (11.7.1) and (11.7.2). We have

$$E(V \mid u) = \int_{R^d} V(x) u(t,x)\, dx. \tag{11.9.3}$$

By inequalities (11.7.11) and (11.9.1), $u(t,x)V(x)$ and $u_t(t,x)V(x)$ are integrable. Thus, differentiation of (11.9.3) with respect to t gives

$$\begin{aligned} \frac{dE(V \mid u)}{dt} &= \int_{R^d} V(x) u_t(t,x)\, dx \\ &= \int_{R^d} V(x) \left\{ \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i\, \partial x_j} [a_{ij}(x) u] - \sum_{i=1}^{d} \frac{\partial}{\partial x_i} [b_i(x) u] \right\} dx. \end{aligned}$$

Integrating by parts and using the fact that the products uV, $u_{x_i}V$, and uV_{x_i} vanish exponentially as $|x| \to \infty$, we obtain

$$\frac{dE(V \mid u)}{dt} = \int_{R^d} \left\{ \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 V}{\partial x_i\, \partial x_j} + \sum_{i=1}^{d} b_i(x) \frac{\partial V}{\partial x_i} \right\} u(t,x)\, dx.$$

From this and inequality (11.9.2), we have

$$\frac{dE(V \mid u)}{dt} \leqslant -\alpha E(V \mid u) + \beta.$$

To solve this differential inequality, multiply through by $e^{\alpha t}$, which gives

$$\frac{d}{dt}[E(V\,|\,u)e^{\alpha t}] \leqslant \beta e^{\alpha t}.$$

Since $E(V\,|\,u)$ at $t = 0$ equals $E(V\,|\,f)$, integration on the interval $[0,t]$ yields

$$E(V\,|\,u)e^{\alpha t} - E(V\,|\,f) \leqslant (\beta/\alpha)(e^{\alpha t} - 1)$$

or

$$E(V\,|\,u) \leqslant e^{-\alpha t}E(V\,|\,f) + (\beta/\alpha)(1 - e^{-\alpha t}).$$

Since $E(V\,|\,f)$ is finite, we can find a $t_0 = t_0(f)$ such that

$$E(V\,|\,u) \leqslant (\beta/\alpha) + 1 \qquad \text{for } t \geqslant t_0.$$

Now let $G_q = \{x\colon V(x) < q\}$. From the Chebyshev inequality (5.7.9), we have

$$\int_{G_q} u(t,x)\,dx \geqslant 1 - \frac{E(V\,|\,u)}{q},$$

and taking $q > 1 + (\beta/\alpha)$ gives

$$\int_{G_q} u(t,x)\,dx \geqslant 1 - \frac{1}{q}\left[1 + \frac{\beta}{\alpha}\right] \doteqdot \varepsilon > 0$$

for $t \geqslant t_0$. Since $V(x) \to \infty$ as $|x| \to \infty$, there is an $r > 0$ such that $V(x) \geqslant q$ for $|x| \geqslant r$. Thus the set G_q is contained in the ball B_r and, as a consequence,

$$\begin{aligned} u(t,x) &= \int_{R^d} \Gamma(1,x,y)u(t-1,y)\,dy \geqslant \int_{B_r} \Gamma(1,x,y)u(t-1,y)\,dy \\ &\geqslant \inf_{|y|\leqslant r} \Gamma(1,x,y) \int_{B_r} u(t-1,y)\,dy \geqslant \varepsilon \inf_{|y|\leqslant r} \Gamma(1,x,y), \\ &\qquad\qquad t \geqslant t_0 + 1,\ x \in R^d. \end{aligned} \tag{11.9.4}$$

Since $\Gamma(1,x,y)$ is strictly positive and continuous, the function

$$h(x) = \varepsilon \inf_{|y|\leqslant r} \Gamma(1,x,y)$$

is also positive. From (11.9.4), we have

$$P_t f(x) = u(t,x) \geqslant h \qquad \text{for } t \geqslant t_0 + 1,$$

which shows that $\{P_t\}$ has a nontrivial lower-bound function. Hence, by Theorem 7.4.1, the proof is complete. ■

When $\{P_t\}$ is asymptotically stable, the next problem is to determine the limiting function

$$\lim_{t\to\infty} P_t f(x) = u_*(x), \qquad f \in D. \tag{11.9.5}$$

This may be accomplished by using the following proposition.

Proposition 11.9.1. If the assumptions of Theorem 11.9.1 are satisfied, then the limiting function u of (11.9.5) is the unique density satisfying the elliptic equation

$$\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\,\partial x_j}[a_{ij}(x)u] - \sum_{i=1}^{d}[b_i(x)u] = 0\,. \tag{11.9.6}$$

Proof: Assume that $\bar{u}(x)$ is a solution of (11.9.6). To prove the uniqueness of $\bar{u}(x)$, note that, because $\bar{u}$ is a solution of (11.9.6), it follows that $u(t,x) = \bar{u}(x)$ is a time-independent solution of the Fokker–Planck equation (11.7.1). Thus, by Theorem 11.9.1,

$$\bar{u}(x) = \lim_{t\to\infty} u(t,x) = u_*(x)$$

and $u_*(x) = \bar{u}(x)$ is unique.

Next we show that u_* satisfies (11.9.6). Let $f \in D(R^d)$ be a continuous function with compact support. We have $u(t+s,x) = P_t u(s,x)$, or

$$u(t+s,x) = \int_{R^d}\Gamma(t,x,y)u(s,y)\,dy\,.$$

Passing to the limit as $s \to \infty$, we obtain

$$u_*(x) = \int_{R^d}\Gamma(t,x,y)u_*(y)\,dy\,.$$

Since Γ is a fundamental solution of the Fokker–Planck equation, $u_*(x)$ is also a solution, and, since $u_*(x)$ is independent of t, it must satisfy equation (11.9.6). Thus the proof is complete. ∎

Example 11.9.1. Again consider the Langevin equation

$$\frac{dx}{dt} = -bx + \sigma\xi$$

and the corresponding Fokker–Planck equation

$$\frac{\partial u}{\partial t} = \tfrac{1}{2}\sigma^2\frac{\partial^2 u}{\partial x^2} + b\frac{\partial}{\partial x}(xu)\,.$$

Inequality (11.9.2) becomes

$$\tfrac{1}{2}\sigma^2\frac{\partial^2 V}{\partial x^2} - bx\frac{\partial V}{\partial x} \leqslant -\alpha V + \beta,$$

which is satisfied with $V(x) = x^2$, $\alpha = 2b$, and $\beta = \sigma^2$. Thus all solutions $u(t,x)$, such that $u(0,x) = f(x)$ is a density, converge to the unique (nonnegative and normalized) solution u_* of

$$\tfrac{1}{2}\sigma^2 \frac{d^2u}{dx^2} + b\frac{d}{dx}(xu) = 0. \tag{11.9.7}$$

The function

$$u_*(x) = g_{\sigma b}(x) = (1/\sigma)\sqrt{b/\pi}\,\exp(-bx^2/\sigma^2),$$

which is the Gaussian density with mean zero and variance $\sigma^2/2b$, satisfies (11.9.7), and, by Proposition 11.9.1, it is the unique solution. □

Example 11.9.2. Next consider the system of stochastic differential equations

$$\frac{dx}{dt} = Bx + \sigma\xi, \tag{11.9.8}$$

where $B = (b_{ij})$ and $\sigma = (\sigma_{ij})$ are constant matrices. Assume that the matrix (a_{ij}) with

$$a_{ij} = \sum_{k=1}^{d} \sigma_{ik}\sigma_{jk}$$

is nonsingular and that the unperturbed system

$$\frac{dx}{dt} = Bx \tag{11.9.9}$$

is asymptotically stable, that is, all solutions converge to zero as $t \to \infty$.

The Fokker–Planck equation corresponding to (11.9.8) has the form

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 u}{\partial x_i\,\partial x_j} - \sum_{i=1}^{d}\frac{\partial}{\partial x_i}[b_i(x)u]$$

where

$$b_i(x) = \sum_{j=1}^{d} b_{ij}x_j\,.$$

Since the coefficients a_{ij} are constant and the matrix (a_{ij}) is nonsingular, this guarantees that the uniform parabolicity condition (11.7.5) is satisfied. All of the remaining conditions appearing in Theorem 11.7.1 in this case are obvious. Since (11.9.9) is asymptotically stable, the real parts of all eigenvalues of B are negative and from the classical results of Liapunov stability theory there is a Liapunov function V such that

 https://doi.org/10.1017/CBO9780511897474.012

$$\sum_{i=1}^{d} b_i(x)\frac{\partial V}{\partial x_i} \leqslant -\alpha V(x)\,, \tag{11.9.10}$$

where V is a positive definite quadratic form

$$V(x) = \sum_{i',j'=1}^{d} k_{i'j'}x_{i'}x_{j'}\,. \tag{11.9.11}$$

Differentiating (11.9.11) with respect to x_i and then x_j, multiplying by $\frac{1}{2}a_{ij}$, summing over i and j, and adding the result to (11.9.10) gives

$$\frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 V}{\partial x_i\,\partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial V}{\partial x_i} \leqslant -\alpha V(x) + \sum_{i,j=1}^{d} a_{ij}k_{ij}\,.$$

Thus inequality (11.9.2) is satisfied. Hence the semigroup $\{P_t\}$ generated by the perturbed system (11.9.8) is asymptotically stable.

To summarize, if the unperturbed system (11.9.9) is asymptotically stable, then any stochastic perturbation with a nonsingular matrix (a_{ij}) leads to a stochastic semigroup that is also asymptotically stable. In this case the limiting density is also Gaussian and can be found by the method of undetermined coefficients by substituting

$$u(x) = c\,\exp\left(\sum_{i,j=1}^{d} \rho_{ij}x_ix_j\right)$$

into the equation

$$\frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 u}{\partial x_i\,\partial x_j} - \sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left(u\sum_{j=1}^{d} b_{ij}x_j\right) = 0\,. \quad \square$$

Example 11.9.3. Consider the second-order system

$$m\frac{d^2x}{dt^2} + \beta\frac{dx}{dt} + F(x) = \sigma\xi \tag{11.9.12}$$

with constant coefficients m, β, and σ. Equation (11.9.12) describes the dynamics of many mechanical and electrical systems in the presence of "white noise." In the mechanical interpretation, m would be the mass of a body whose position is x, β is a friction coefficient, and $F = \partial\phi/\partial x$ is a conservative force (with a corresponding potential function ϕ) acting on the body. Introducing the velocity $v = dx/dt$ as a new variable, equation (11.9.12) is equivalent to the system

$$\frac{dx}{dt} = v \quad \text{and} \quad m\frac{dv}{dt} = -\beta v - F(x) + \sigma\xi\,. \tag{11.9.13}$$

The Fokker–Planck equation corresponding to (11.9.13) is

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2m^2}\frac{\partial^2 u}{\partial v^2} - \frac{\partial}{\partial x}(vu) + \frac{1}{m}\frac{\partial}{\partial v}\{[\beta v + F(x)]u\}. \tag{11.9.14}$$

Unfortunately, the asymptotic stability of the solutions of (11.9.14) cannot be studied by Theorem 11.9.1 as the quadratic form associated with the second-order term is

$$0 \cdot \lambda_1^2 + (\sigma^2/m^2) \cdot \lambda_2^2,$$

which is clearly not positive definite. Using more sophisticated techniques, it is possible to prove that the solutions to some parabolic equations with associated semidefinite quadratic forms are asymptotically stable. However, in this example we wish only to derive the steady-state solution to (11.9.14) and to bypass the question of asymptotic stability.

In a steady state, $(\partial u/\partial t) = 0$, so (11.9.14) becomes

$$\frac{\sigma^2}{2m^2}\frac{\partial^2 u}{\partial v^2} - \frac{\partial}{\partial x}(vu) + \frac{1}{m}\frac{\partial}{\partial v}\{[\beta v + F(x)]u\} = 0,$$

which may be written in the alternate form

$$\left(\frac{\beta}{m}\frac{\partial}{\partial v} - \frac{\partial}{\partial x}\right)\left[vu + \frac{\sigma^2}{2m\beta}\frac{\partial u}{\partial v}\right] + \frac{\partial}{\partial v}\left[\frac{1}{m}F(x)u + \frac{\sigma^2}{2m\beta}\frac{\partial u}{\partial x}\right] = 0.$$

Set $u(x, v) = X(x)V(v)$, so that the last equation becomes

$$\left(\frac{\beta}{m}\frac{\partial}{\partial v} - \frac{\partial}{\partial x}\right)\left[X\left(vV + \frac{\sigma^2}{2m\beta}\frac{dV}{dv}\right)\right] + \left[\frac{1}{m}F(x)X + \frac{\sigma^2}{2m\beta}\frac{dX}{dx}\right]\frac{dV}{dv} = 0,$$

which will certainly be satisfied if X and V satisfy

$$\frac{dX}{dx} + \frac{2\beta}{\sigma^2}F(x)X = 0 \tag{11.9.15}$$

and

$$\frac{dV}{dv} + \frac{2m\beta}{\sigma^2}vV = 0, \tag{11.9.16}$$

respectively.

Integrating equations (11.9.15) and (11.9.16) and combining the results gives

$$u(x, v) = c\exp\{-(2\beta/\sigma^2)[\tfrac{1}{2}mv^2 + \phi(x)]\}. \tag{11.9.17}$$

The constant c in (11.9.17) is determined from the normalization condition

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x, v)\,dx\,dv = 1.$$

The velocity integration is easily carried out and we have

$$c = c_1\sqrt{\beta m/\pi\sigma^2},$$

where

$$\frac{1}{c_1} = \int_{-\infty}^{\infty} \exp[(-2\beta/\sigma^2)\phi(x)]\,dx. \tag{11.9.18}$$

Thus (11.9.17) becomes

$$u(x,v) = c_1\sqrt{\beta m/\pi\sigma^2}\exp\{-(2\beta/\sigma^2)[\tfrac{1}{2}mv^2 + \phi(x)]\}. \tag{11.9.19}$$

The interesting feature of (11.9.19) is that the right-hand side may be written as the product of two functions, one dependent on v and the other on x. This can be interpreted to mean that in the steady state the positions and velocities are independent. Furthermore, observe that for every ϕ for which the integral (11.9.18) is convergent, $u(x, v)$, as given by (11.9.19), is a well-defined solution of the steady-state equation and that the distribution of velocities is Maxwellian independent of the nature of the potential function ϕ. The Maxwellian nature of the velocity distribution is a natural consequence of the characteristics of the noise perturbation term in the force balance equation (11.9.13). □

11.10 An extension of the Liapunov function method

A casual inspection of the proofs of Theorems 5.7.1 and 11.9.1 shows that they are based on the same idea: We first prove that the mathematical expectation $E(V\,|\,P_tf)$ is bounded for large t and then show, by the Chebyshev inequality, that the density P_tf is concentrated on some bounded region. With these facts we are then able to construct a lower-bound function. This technique may be formalized as follows.

Let a stochastic semigroup $\{P_t\}_{t\geq 0}$, $P_t\colon L^1(G) \to L^1(G)$, be given, where G is an unbounded measurable subset of R^d. Further, let $V\colon G \to R$ be a continuous nonnegative function such that

$$\lim_{|x|\to\infty} V(x) = \infty. \tag{11.10.1}$$

Also set, as before,

$$E(V\,|\,P_tf) = \int_G V(x)P_tf(x)\,dx. \tag{11.10.2}$$

With these definitions it is easy to prove the following proposition.

Proposition 11.10.1. Assume there exists a linearly dense subset $D_0 \subset D(G)$ and a constant $M < \infty$ such that

$$E(V \mid P_t f) \leqslant M \tag{11.10.3}$$

for every $f \in D_0$ and sufficiently large t, say $t \geqslant t_1(f)$. Let r be such that $V(x) \geqslant M + 1$ for $|x| \geqslant r$ and $x \in G$. If, for some $t_0 > 0$, there is a nontrivial function h_r with $h_r \geqslant 0$ and $\|h_r\| > 0$ such that

$$P_{t_0} f \geqslant h_r \qquad \text{for } f \in D \tag{11.10.4}$$

whose support is contained in the ball $B_r = \{x \in R^d\colon |x| \leqslant r\}$, then the stochastic semigroup $\{P_t\}_{t\geqslant 0}$ is asymptotically stable.

Proof: Pick $f \in D_0$. From the Chebyshev inequality and (11.10.3), it follows that

$$\int_{G_a} P_t f(x)\,dx \geqslant 1 - \frac{M}{a}, \qquad \text{for } t \geqslant t_1, \tag{11.10.5}$$

where $G_a = \{x \in G\colon V(x) < a\}$. Pick $a = M + 1$ so $V(x) \geqslant a$ for $|x| \geqslant r$. Then $G_a \subset B_r$ and

$$P_t f = P_{t_0} P_{t-t_0} f \geqslant P_{t_0} f_t = \|f_t\| P_{t_0} \tilde{f}, \tag{11.10.6}$$

where $f_t = (P_{t-t_0} f) 1_{G_a}$ and $\tilde{f} = f_t/\|f_t\|$. From (11.10.5), we have

$$\|f_t\| = \int_{G_a} P_{t-t_0} f(x)\,dx \geqslant 1 - \frac{M}{a} \qquad \text{for } t \geqslant t_0 + t_1,$$

and, by (11.10.4), $P_{t_0}\tilde{f} \geqslant h_r$. Thus, using (11.10.6), we have shown that

$$[1 - (M/a)]h_r$$

is a lower-bound function for the semigroup $\{P_t\}_{t\geqslant 0}$. Since, by assumption, h_r is a nontrivial function and we took $a > M$, then it follows that the lower-bound function for the semigroup $\{P_t\}_{t\geqslant 0}$ is also nontrivial. Application of Theorem 7.4.1 completes the proof. ∎

Example 11.10.1. As an example of the application of Proposition 11.10.1, we will first prove the asymptotic stability of the semigroup generated by the integro-differential equation

$$\frac{\partial u(t,x)}{\partial t} + u(t,x) = \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} + \int_{-\infty}^{\infty} K(x,y)u(t,y)\,dy, \qquad t > 0,\ x \in R \tag{11.10.7}$$

with the initial condition

$$u(0,x) = \phi(x) \qquad x \in R, \tag{11.10.8}$$

which we first considered in Example 7.9.1. As in that example, we assume that K is a stochastic kernel, but we also assume that K satisfies

$$\int_{-\infty}^{\infty} |x| K(x, y)\, dx \leq \alpha |y| + \beta \qquad \text{for } y \in R \tag{11.10.9}$$

where α and β are nonnegative constants and $\alpha < 1$.

To slightly simplify an intricate series of calculations we assume, without any loss of generality, that $\sigma = 1$. (This is equivalent to defining a new $\bar{x} = x/\sigma$.) Our proof of the asymptotic stability of the stochastic semigroup, corresponding to equations (11.10.7) and (11.10.8), follows arguments given by Jama [In press] in verifying (11.10.3) and (11.10.4) of Proposition 11.10.1.

From Example 7.9.1, we know that the stochastic semigroup $\{P_t\}_{t\geq 0}$ generated by equations (11.10.7) and (11.10.8) is defined by (with $\sigma^2 = 1$)

$$P_t \phi = e^{-t} \sum_{n=0}^{\infty} T_n(t) \phi, \tag{11.10.10}$$

where

$$\begin{aligned} T_n(t) f &= \int_0^t T_0(t - \tau) P T_{n-1}(\tau) f \, d\tau, \\ T_0(t) f(x) &= \int_{-\infty}^{\infty} g(t, x - y) f(y)\, dy \end{aligned} \tag{11.10.11}$$

and

$$Pf(x) = \int_{-\infty}^{\infty} K(x, y) f(y)\, dy, \qquad g(t, x) = \frac{1}{\sqrt{2\pi t}} \exp(-x^2/2t). \tag{11.10.12}$$

Let $f \in D(R)$ be a continuous function with compact support. Define

$$E(t) = E(|x| \,|\, P_t f) = \int_{-\infty}^{\infty} |x| P_t f(x)\, dx,$$

which may be rewritten using (11.10.10) as

$$E(t) = e^{-t} \sum_{n=0}^{\infty} e_n(t),$$

where

$$e_n(t) = \int_{-\infty}^{\infty} |x| T_n(t) f(x)\, dx.$$

We are going to show that $E(t)$, as given here, satisfies condition (11.10.3).

If we set

$$f_{nt} = P T_{n-1}(t) f \quad \text{and} \quad q_{n\tau}(t) = \int_{-\infty}^{\infty} |x| T_0(t - \tau) f_{n\tau}(x)\, dx$$

then, using (11.10.11), we may write $e_n(t)$ as

$$e_n(t) = \int_0^t q_{n\tau}(t)\,d\tau. \tag{11.10.13}$$

Using the second relation in equations (11.10.11), $q_{n\tau}(t)$ can be written as

$$q_{n\tau}(t) = \int_{-\infty}^{\infty} f_{n\tau}(y)\left[\int_{-\infty}^{\infty} |x| g(t-\tau, x-y)\,dx\right] dy.$$

Since $|x| \leqslant |x-y| + |y|$, it is evident that

$$\int_{-\infty}^{\infty} |x| g(t-\tau, x-y)\,dx \leqslant \sqrt{\frac{2(t-\tau)}{\pi}} + |y| \tag{11.10.14}$$

and, as a consequence,

$$q_{n\tau}(t) \leqslant \int_{-\infty}^{\infty} |y| f_{n\tau}(y)\,dy + \sqrt{\frac{2(t-\tau)}{\pi}} \int_{-\infty}^{\infty} f_{n\tau}(y)\,dy. \tag{11.10.15}$$

By using equation (7.9.18) from the proof of the Phillips perturbation theorem and noting that P is a Markov operator (since K is a stochastic kernel) and $\|f\| = 1$, we have

$$\int_{-\infty}^{\infty} f_{n\tau}(y)\,dy = \|PT_{n-1}(\tau)f\| = \|T_{n-1}(\tau)f\| \leqslant \frac{\tau^{n-1}}{(n-1)!}. \tag{11.10.16}$$

Furthermore, from equations (11.10.9) and (7.9.18),

$$\begin{aligned}\int_{-\infty}^{\infty} |y| f_{n\tau}(y)\,dy &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |y| K(y,z) T_{n-1}(\tau) f(z)\,dy\,dz \\ &\leqslant \alpha \int_{-\infty}^{\infty} |z| T_{n-1}(\tau) f(z)\,dz + \beta \int_{-\infty}^{\infty} T_{n-1}(\tau) f(z)\,dz \\ &\leqslant \alpha e_{n-1}(\tau) + \beta \frac{\tau^{n-1}}{(n-1)!}.\end{aligned}$$

Substituting this and (11.10.16) into (11.10.15) gives

$$q_{n\tau}(t) \leqslant \alpha e_{n-1}(\tau) + \left[\beta + \sqrt{\frac{2(t-\tau)}{\pi}}\right] \frac{\tau^{n-1}}{(n-1)!}$$

so that (11.10.13) becomes

$$e_n(t) \leqslant \alpha \int_0^t e_{n-1}(\tau)\,d\tau + \beta \frac{t^n}{n!} + \sqrt{\frac{2}{\pi}} \int_0^t \sqrt{t-\tau}\, \frac{\tau^{n-1}}{(n-1)!}\,d\tau, \qquad n = 1, 2, \ldots \tag{11.10.17}$$

To obtain $e_0(t)$ we again use (11.10.14) to give

$$e_0(t) = \int_{-\infty}^{\infty} |x| T_0(t) f(x)\, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x| g(t, x - y) f(y)\, dx\, dy$$

$$\leqslant \sqrt{\frac{2t}{\pi}} + m_1, \qquad m_1 = \int_{-\infty}^{\infty} |y| f(y)\, dy. \tag{11.10.18}$$

With equations (11.10.17) and (11.10.18) we may now proceed to examine $E(t)$. Sum (11.10.17) from $n = 1$ to m and add (11.10.18). This gives

$$\sum_{n=0}^{m} e_n(t) \leqslant m_1 + \sqrt{\frac{2t}{\pi}} + \beta e^t + \sqrt{\frac{2}{\pi}} \int_0^t \sqrt{t-\tau}\, e^\tau\, d\tau + \int_0^t \sum_{n=0}^{m} e_n(\tau)\, d\tau,$$

where we used the fact that

$$\sum_{n=1}^{m} \frac{t^n}{n!} \leqslant \sum_{n=0}^{\infty} \frac{t^n}{n!} = e^t.$$

Define $E_m(t) = e^{-t} \sum_{n=0}^{m} e_n(t)$; hence we can write

$$E_m(t) \leqslant m_1 e^{-t} + \rho + \alpha \int_0^t e^{-(t-\tau)} E_m(\tau)\, d\tau, \tag{11.10.19}$$

where

$$\rho = \beta + \max_t \left[\sqrt{\frac{2t}{\pi}}\, e^{-t} \right] + \int_0^\infty \sqrt{u}\, e^{-u}\, du.$$

To solve the integral inequality (11.10.19), it is enough to solve the corresponding equality and note that $E_m(t)$ is below this solution [Walter, 1970]. This process leads to

$$E_m(t) \leqslant [\rho/(1-\alpha)] + m_1 e^{-(1-\alpha)t},$$

or passing to the limit as $m \to \infty$,

$$E(t) \leqslant [\rho/(1-\alpha)] + m_1 e^{-(1-\alpha)t}. \tag{11.10.20}$$

Since the constant ρ does not depend on f, (11.10.20) proves that the semigroup $\{P_t\}_{t \geqslant 0}$, generated by (11.10.7) and (11.10.8), satisfies equation (11.10.3) with $V(x) = |x|$.

Next we verify equation (11.10.4). Assume that $f \in D(R)$ is supported on $[-r, r]$. Then we have

$$P_1 f \geqslant e^{-1} T_0(1) f = e^{-1} \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} f(y) \exp[-\tfrac{1}{2}(x-y)^2]\, dy$$

$$\geqslant \frac{1}{\sqrt{2\pi}} \exp[-(x^2 + r^2 + 1)] \int_{-r}^{r} f(y)\, dy$$

$$= \frac{1}{\sqrt{2\pi}} \exp[-(x^2 + r^2 + 1)],$$

and the function on the right-hand side is clearly nontrivial.

Thus we have shown that the semigroup $\{P_t\}_{t\geq 0}$ generated by equations (11.10.7) and (11.10.8) is asymptotically stable, and therefore the solution with every initial condition $\phi \in D$ converges to the same limit. □

Example 11.10.2. Using a quite analogous approach, we now prove the asymptotic stability of the semigroup generated by the equation

$$\frac{\partial u(t,x)}{\partial t} + c\frac{\partial u(t,x)}{\partial x} + u(t,x) = \int_x^\infty K(x,y)u(t,y)\,dy \tag{11.10.21}$$

with the conditions

$$u(t,0) = 0 \quad \text{and} \quad u(0,x) = \phi(x) \tag{11.10.22}$$

(see Example 7.9.2). However, in this case some additional constraints on kernel K will be introduced at the end of our calculations. The necessity of these constraints is related to the fact that the smoothing properties of the semigroup generated by the infinitesimal operator (d^2/dx^2) of the previous example are not now present (see Example 7.4.1). Rather, in the present example the operator (d/dx) generates a semigroup that merely translates functions (see Example 7.4.2). Thus, in general the properties of equations (11.10.7) and (11.10.21) are quite different in spite of the fact that we are able to write the explicit equations for the semigroups generated by both equations using the formulas of the Phillips perturbation theorem. Our treatment follows that of Dłotko and Lasota [in press].

To start, we assume K is a stochastic kernel and satisfies

$$\int_0^y xK(x,y)\,dx \leq \alpha y + \beta \qquad \text{for } y > 0, \tag{11.10.23}$$

where α and β are nonnegative constants and $\alpha < 1$. In the Chandrasekhar–Münch equation, $K(x,y) = \psi(x/y)/y$, and (11.10.23) is automatically satisfied since

$$\int_0^y xK(x,y)\,dx = \int_0^y (x/y)\psi(x/y)\,dx = y\int_0^1 z\psi(z)\,dz$$

and

$$\int_0^1 z\psi(z)\,dz < \int_0^1 \psi(z)\,dz = 1.$$

As in the preceding example, the semigroup $\{P_t\}_{t\geq 0}$ generated by equations (11.10.21) and (11.10.22) is given by equations (11.10.10) and (11.10.11), but now (assuming $c = 1$ for ease of calculations)

$$T_0(t)f(x) = 1_{[0,\infty)}(x - t)f(x - t) \tag{11.10.24}$$

and

$$Pf(x) = \int_x^\infty K(x,y)f(y)\,dy. \tag{11.10.25}$$

To verify condition (11.10.3), assume that $f \in D([0,\infty))$ is a continuous function with compact support contained in $(0,\infty)$ and consider

$$E(t) = \int_0^\infty xP_t f(x)\,dx.$$

By using notation similar to that introduced in Example 11.10.1, we have

$$E(t) = e^{-t}\sum_{n=0}^{\infty} e_n(t), \qquad e_n(t) = \int_0^\infty xT_n(t)f(x)\,dx$$

and

$$e_n(t) = \int_0^t q_{n\tau}(t)\,d\tau, \qquad q_{n\tau}(t) = \int_0^\infty xT_0(t - \tau)Pf_{n\tau}(x)\,dx,$$

where $f_{n\tau} = T_{n-1}(\tau)f$. From equations (11.10.24) and (11.10.25), we have

$$q_{n\tau}(t) = \int_{t-\tau}^\infty x\left[\int_{x-t+\tau}^\infty K(x - t + \tau, y)f_{n\tau}(y)\,dy\right]dx,$$

or, setting $x - t + \tau = z$ and using (11.10.23),

$$\begin{aligned} q_{n\tau}(t) &= \int_0^\infty\left[\int_z^\infty zK(z,y)f_{n\tau}(y)\,dy\right]dz \\ &\quad + (t - \tau)\int_0^\infty\left[\int_z^\infty K(z,y)f_{n\tau}(y)\,dy\right]dz \\ &\leqslant \alpha\int_0^\infty yf_{n\tau}(y)\,dy + \beta\int_0^\infty f_{n\tau}(y)\,dy \\ &\quad + (t - \tau)\int_0^\infty\left[\int_z^\infty K(z,y)f_{n\tau}(y)\,dy\right]dz. \end{aligned}$$

Since K is stochastic and

$$\|f_{n\tau}\| = \|T_{n-1}(\tau)f\| \leqslant \tau^{n-1}/(n - 1)!,$$

this inequality reduces to

$$q_{n\tau}(t) \leqslant \alpha e_{n-1}(\tau) + [\beta + t - \tau][\tau^{n-1}/(n - 1)!], \qquad n = 1,2,\ldots$$

Thus

$$e_n(t) \leqslant \alpha \int_0^t e_{n-1}(\tau)\,d\tau + \beta\frac{t^n}{n!} + \int_0^t (t-\tau)\frac{\tau^{n-1}}{(n-1)!}\,d\tau. \qquad (11.10.26)$$

Further,

$$\begin{aligned} e_0(t) &= \int_0^\infty xT_0(t)f(x)\,dx = \int_t^\infty xf(x-t)\,dx \\ &= \int_0^\infty zf(z)\,dz + t\int_0^\infty f(z)\,dz \end{aligned}$$

or

$$e_0(t) = m_1 + t, \qquad m_1 = \int_0^\infty zf(z)\,dz. \qquad (11.10.27)$$

Observe the similarity between equations (11.10.26)–(11.10.27) and equations (11.10.17)–(11.10.18). Thus, proceeding as in Example 11.10.1, we again obtain (11.10.20) with

$$\rho = \beta + \int_0^\infty ue^{-u}\,du + \max_t(te^{-t}).$$

Thus we have shown that the semigroup generated by equations (11.10.21)–(11.10.22) satisfies condition (11.10.3).

However, the proof that (11.10.4) holds is more difficult for the reasons set out at the beginning of this example. To start, pick $r > 0$ as in Proposition 11.10.1, that is,

$$r = M + 1 = [\rho/(1-\alpha)] + 1.$$

For an arbitrary $f \in D([0,r])$ and $t_0 > 0$, we have

$$\begin{aligned} P_{t_0}f(x) \geqslant e^{-t_0}T_1(t_0)f(x) &= e^{-t_0}\int_0^{t_0} T_0(t_0-\tau)PT_0(\tau)f(x)\,d\tau \\ &= e^{-t_0}\int_0^{t_0}\Bigg[1_{[0,\infty)}(x-t_0+\tau) \\ &\qquad \cdot \int_{x-t_0+\tau}^\infty K(x-t_0+\tau,y) \\ &\qquad \cdot 1_{[0,\infty)}(y-\tau)f(y-\tau)\,dy\Bigg]\,d\tau. \end{aligned}$$

In particular, for $0 \leqslant x \leqslant t_0$,

$$P_{t_0}f(x) \geqslant e^{-t_0}\int_{t_0-x}^{t_0}\left[\int_{\tau}^{\infty}K(x - t_0 + \tau, y)f(y - \tau)\,dy\right]d\tau.$$

Now set $z = y - \tau$ and $s = x - t_0 + \tau$ and remember that $f \in D([0,r])$ to obtain

$$P_{t_0}f(x) \geqslant e^{-t_0}\int_0^x\left[\int_0^r K(s, z + s + t_0 - x)f(z)\,dz\right]ds$$

$$\geqslant h_r(x)\int_0^r f(z)\,dz = h_r(x) \qquad \text{for } 0 \leqslant x \leqslant t_0,$$

where

$$h_r(x) = e^{-t_0}\inf_{0\leqslant z\leqslant r}\int_0^x K(s, z + s + t_0 - x)\,ds.$$

It is therefore clear that $h_r \geqslant 0$, and it is easy to find a sufficient condition for h_r to be nontrivial. For example, if $K(s,u) = \psi(s/u)/u$, as in the Chandrasekhar–Münch equation, then

$$h_r(x) = e^{-t_0}\inf_z\int_0^x \psi\left(\frac{s}{z + s + t_0 - x}\right)\frac{ds}{z + s + t_0 - x}.$$

If we set $q = s/(z + s + t_0 - x)$ in this expression, then

$$h_r(x) = e^{-t_0}\inf_z\int_0^{x/(z+t_0)}\frac{\psi(q)}{1-q}\,dq$$

$$\geqslant e^{-t_0}\int_0^{x/(r+t_0)}\psi(q)\,dq.$$

Since $\psi(q)$ is a density, we have

$$\lim_{t_0\to\infty}\int_0^{x/(r+t_0)}\psi(q)\,dq = 1$$

uniformly for $x \in [t_0 - 1, t_0]$. Thus, for some sufficiently large t_0, we obtain

$$h_r(x) \geqslant e^{-t_0}\int_0^{x/(r+t_0)}\psi(q)\,dq > 0 \qquad \text{for } x \in [t_0 - 1, t_0],$$

showing that h_r is a nontrivial function. Therefore all the assumptions of Proposition 11.10.1 are satisfied and the semigroup $\{P_t\}$ generated by the Chandrasekhar–Münch equation is asymptotically stable. □