

# DIFFUSION PROCESSES

DANIEL W. STROOCK and S. R. S. VARADHAN  
COURANT INSTITUTE, NEW YORK UNIVERSITY

## 1. Introduction

One of the major problems in the theory of diffusion processes is to construct the process for a given set of diffusion coefficients. A diffusion process in  $R^d$  is hopefully determined by the two sets of coefficients

$$(1.1) \quad \begin{aligned} a &= a(t, x) = \{a_{ij}(t, x)\}, & 1 \leq i, j \leq d, t \in [0, \infty), x \in R^d, \\ b &= b(t, x) = \{b_j(t, x)\}, & 1 \leq j \leq d, t \in [0, \infty), x \in R^d. \end{aligned}$$

Here  $a$  is a positive semidefinite symmetric matrix for each  $t$  and  $x$ , and  $b$  is a  $d$  vector for each  $t$  and  $x$ . There are various ways of describing exactly what we mean by a diffusion process corresponding to the specified set of coefficients. We shall adopt the following approach.

Let  $\Omega$  be the space of  $R^d$  valued continuous functions on  $[0, \infty)$ . The value of a function  $\omega = x(\cdot)$  in  $\Omega$  at time  $t$  will be denoted by  $x(t)$ . The  $\sigma$ -field generated by  $x(s)$  for  $t_1 \leq s \leq t_2$  will be denoted by  $M_{t_1}^{t_2}$ . If  $t_1 = 0$ , we will denote this by  $M_{t_2}$  and by  $M^{t_1}$  in case  $t_2 = \infty$ , where  $M$  is the  $\sigma$ -field generated by  $x(s)$  for  $0 \leq s < \infty$ . The space  $\Omega$  can be viewed as a complete separable metric space, with uniform convergence on bounded intervals defining the topology. Then  $M$  is the Borel  $\sigma$ -field in  $\Omega$ . A stochastic process with values in  $R^d$ , defined for  $t \geq t_0$ , is a probability measure on  $(\Omega, M^{t_0})$ .

Given the coefficients  $\{a_{ij}(t, x)\}$  and  $\{b_j(t, x)\}$ , we define an operator  $L_t$  acting on functions  $f(x) \in C_0^\infty(R^d)$  by

$$(1.2) \quad (L_t f)(x) = \frac{1}{2} \sum a_{ij}(t, x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum b_j(t, x) \frac{\partial f}{\partial x_j}.$$

We say that a measure  $P$  is a solution to the Martingale problem corresponding to the given coefficients, starting at time  $t_0$  from the point  $x_0$  if

- (a)  $P$  is a probability measure on  $(\Omega, M^{t_0})$  such that  $P[x(t_0) = x_0] = 1$ , and
- (b) for each  $f \in C_0^\infty(R^d)$ ,  $f(x(t)) - \int_{t_0}^t (L_s f)(x(s)) ds$  is a martingale relative to  $(\Omega, M_t^{t_0}, P)$ .

Under suitable conditions on the coefficients  $a$  and  $b$ , one should attempt to answer the following questions:

- (1) For each  $t_0$  and  $x_0$ , does a solution  $P_{t_0, x_0}$  exist?

Results obtained at the Courant Institute of Mathematical Sciences, New York University; this research was sponsored by the US Air Force Office of Scientific Research, Contract AF-49(638)-1719.

- (2) Is it unique?  
 (3) Is the solution a strong Markov process?  
 (4) Does the solution depend continuously, in some suitable sense, on  $t_0$ ,  $x_0$ ,  $a$  and  $b$ ?  
 (5) If a diffusion process can be constructed in some other natural manner, does it coincide with the above construction?

Let us take up question (5) first. There are at least two other possible ways of constructing diffusion processes for given coefficients. If we use the stochastic differential equations of Itô, we take a matrix  $\sigma(t, x)$  such that

$$(1.3) \quad \sigma(t, x)\sigma^*(t, x) = a(t, x).$$

If  $\beta(t)$  is Brownian motion in  $d$  dimensions, one can set up the stochastic differential equation

$$(1.4) \quad dx(t) = \sigma(t, x(t)) d\beta(t) + b(t, x(t)) dt.$$

If  $\sigma$  and  $b$  are assumed to be bounded and uniformly (with respect to  $t$ ) Lipschitz continuous in  $x$ , one can show that the above equation has a unique solution  $x(t)$  for  $t \geq t_0$ , for each initial condition  $x(t_0) = x_0$ . The process so obtained is the diffusion process corresponding to  $a$  and  $b$  starting from the point  $x_0$  at time  $t_0$ .

One can prove that if the solution to the stochastic differential equation exists, then the solution to the martingale problem also exists. Moreover, if the former is unique, then so is the latter. Furthermore, one shows that the two solutions are the same in the sense that the solution to the martingale problem is the distribution of the solution to the stochastic differential equation.

Another possibility is to use the theory of partial differential equations. We consider the fundamental solution  $p(s, x, t, y)$  of the equation

$$(1.5) \quad \frac{\partial p}{\partial s} + \frac{1}{2} \sum a_{ij}(s, x) \frac{\partial^2 p}{\partial x_i \partial x_j} + \sum b_j(s, x) \frac{\partial p}{\partial x_j} = 0,$$

where  $p(s, x, t, y)$  serves as the transition probability density of a Markov process. This is taken as the diffusion process corresponding to  $a$  and  $b$ . One can verify that, in this case also, the solution to the martingale problem exists, is unique, and coincides with the process constructed through the fundamental solution. The existence of a fundamental solution is proved under the assumption that  $a$  and  $b$  are bounded, satisfy a Hölder condition and  $a$  is uniformly positive definite.

## 2. Existence and uniqueness: the general case

Let us now assume that  $a(t, x)$  is bounded continuous and positive definite for each  $t$  and  $x$ . Assume  $b(t, x)$  bounded and measurable. Under these assumptions we can answer the questions (1) through (4) raised in the introduction, affirmatively.

**THEOREM 2.1.** *The solution  $P_{s,x}$  to the martingale problem, starting from the point  $x$  at some  $s$ , exists and is unique for every  $s$  and  $x$ . Moreover,  $P_{s,x}$  is a strong Markov process with transition probabilities  $P(s, x, t, A) = P_{s,x}[x(t) \in A]$ . Further, the solution depends continuously on  $s, x, a(\cdot, \cdot)$ , and  $b(\cdot, \cdot)$  in the following sense: if  $a_n(\cdot, \cdot) \rightarrow a(\cdot, \cdot)$  uniformly on compact sets,  $b_n(\cdot, \cdot) \rightarrow b(\cdot, \cdot)$  in measure,  $s_n \rightarrow s, x_n \rightarrow x$  and if  $a_n(\cdot, \cdot), b_n(\cdot, \cdot)$  are uniformly bounded, then the solution  $P_n$  corresponding to  $a_n, b_n$  starting from  $x_n$  at time  $s_n$  converges weakly to the solution  $P$  for the limiting coefficients starting from  $x$  at time  $s$ .*

This theorem can be found in [2] and [3].

**3. Boundary conditions**

Let us suppose that  $G \subset R^d$  is a smooth region. More precisely, there is a function  $\phi(x) \in C_b^2(R^d)$  such that

$$(3.1) \quad \begin{aligned} G &= [x: \phi(x) > 0], \\ \delta G &= [x: \phi(x) = 0], \\ \|\nabla G\| &\geq 1 \text{ on } \delta G. \end{aligned}$$

As before we shall assume that  $a$  is bounded, positive definite for each  $t$  and  $x$  and is continuous on  $[0, \infty) \times \bar{G}$ . Assume  $b$  is bounded and measurable. One can check that some sort of boundary condition on  $\delta G$  is needed to describe what happens to the process when it reaches the boundary  $\delta G$ . A class of these boundary conditions are of the form

$$(3.2) \quad \rho(s, x) \frac{\partial}{\partial s} + \sum_{j=1}^d \gamma_j(s, x) \frac{\partial}{\partial x_j} = 0,$$

where  $\rho$  and  $\gamma$  are suitable function on the boundary  $[0, \infty) \times \delta G$ . We shall assume the following regarding  $\rho$  and  $\gamma$ :

- (i)  $\gamma$  is Lipschitz continuous in  $t$  and  $x$ , is bounded, and there is a constant  $\beta > 0$ , such that  $\langle \gamma, \nabla \phi \rangle \geq \beta > 0$ ; for all  $t, x \in [0, \infty) \times \delta G$ ;
- (ii) either  $\rho$  is identically zero or it is everywhere positive, is bounded, and satisfies a Lipschitz condition in  $t$  and  $x$ .

We formulate the problem in the following manner. A solution corresponding to  $a, b, \rho$ , and  $\gamma$ , starting a time  $t_0$  from the point  $x_0$ , is a measure  $P$  on  $(\Omega, M^{t_0})$ , where now  $\Omega = C[[0, \infty), \bar{G}]$  and  $M^{t_0}$  is the natural  $\sigma$ -field as before. The measure  $P$  is such that

- (a)  $P[x(t_0) = x_0] = 1$ , and
- (b) for every  $u \in C_0^\infty[[0, \infty) \times \bar{G}]$  with  $\rho(\partial u / \partial s) + \gamma \cdot \nabla u \geq 0$  on  $[0, \infty) \times \delta G$ ,

$$(3.3) \quad u(t, x(t)) - \int_{t_0}^t (U_s + L_s u)(s, x(s)) \chi_G(x(s)) ds,$$

is a submartingale relative to  $(\Omega, M_t^{t_0}, P)$ . We will call  $P$ , if it exists, a solution to the submartingale problem.

**THEOREM 3.1.** *Under the assumptions (a) and (b) on  $\rho$  and  $\gamma$ , along with the assumptions on  $a, b$  mentioned in the earlier section, for each  $x \in \bar{G}$  and  $s \geq 0$ , there is a unique solution  $P_{s,x}$  to the submartingale problem. The solution is a strong Markov process. Moreover, the solution depends continuously on  $s, x, a, b, \rho$ , and  $\gamma$ .*

**REMARK 3.1.** In the homogeneous case, that is, when  $a, b, \rho$ , and  $\gamma$  are independent of  $t$ , assumption (b) on  $\rho$  can be replaced by the assumption that  $\rho$  is bounded, continuous and nonnegative.

These results can be found in [4].

#### 4. Invariance principle

Using the results of Sections 2 and 3, one can prove general theorems concerning the convergence of Markov chains to diffusion processes. The conditions for convergence are very natural and involve essentially the first two moments. For simplicity, we shall treat the homogeneous case and assume further that  $d = 1$ . The general case is quite similar. Let us suppose that there is no boundary.

For each  $\delta > 0$ ,  $\pi_\delta(x, dy)$  is the transition probability of a Markov chain with  $R$  as its state space. The transitions of the chain occur in multiples of time  $\delta$ . Let us define

$$(4.1) \quad \begin{aligned} b_\delta(x) &= \frac{1}{\delta} \int (y - x) \pi_\delta(x, dy), \\ a_\delta(x) &= \frac{1}{\delta} \int (y - x)^2 \pi_\delta(x, dy), \\ \theta_\delta(x) &= \frac{1}{\delta} \int |y - x|^\mu \pi_\delta(x, dy), \end{aligned} \quad \mu > 2.$$

We assume:

(a)  $|b_\delta(x)| \leq C$  and  $b_\delta(x)$  converges as  $\delta$  tends to zero to a continuous limit  $b(x)$  uniformly on bounded intervals;

(b)  $|a_\delta(x)| \leq C$  and  $a_\delta(x)$  converges as  $\delta$  tends to zero to a continuous limit  $a(x)$  uniformly on bounded intervals;

(c)  $\theta_\delta(x) \rightarrow 0$ , as  $\delta$  tends to zero, uniformly on bounded intervals;

(d) the solution to the martingale problem for any starting point, corresponding to the coefficients  $a$  and  $b$  is unique.

**THEOREM 4.1.** *Under the above assumptions the Markov chain converges as  $\delta$  tends to zero, to the diffusion corresponding to  $a$  and  $b$ .*

The general case for the  $d$  dimensional case, where  $a$  and  $b$  could depend on  $t$  is treated in [3]. There are similar results when a boundary is involved and these can be found in [4].

**5. Special case:  $d = 2$**

The method of the main theorem mentioned in Section 2, involves several steps. The hypothesis of continuity of  $a(t, x)$  is superfluous. What one needs is that the discontinuities of  $a(t, x)$  should be small compared to the eigenvalues of  $a(t, x)$ . The actual constants are hard to follow through in general. But when  $d = 2$ , in the homogeneous case, this leads to the following theorem.

**THEOREM 5.1.** *If  $a(x)$  is uniformly positive definite on compact sets and if trace  $a(x) \equiv 1$ , then for any bounded measurable  $b(x)$ , existence and uniqueness hold for the martingale problem.*

By a random time change, the condition that trace  $a(x) \equiv 1$  can be removed and replaced by the boundedness of  $a$ . Once uniqueness is established, one shows by a standard reasoning that the process is strongly Markovian. With a little more work, one can establish the continuous dependence of the solution on the coefficients and the starting point.

These results can be found essentially in [1].

**6. Special case:  $d = 1$**

When  $d = 1$ , the results are even stronger. If  $a(t, x)$  is bounded measurable and uniformly positive on compact sets and if  $b(t, x)$  is bounded and measurable, then the solution to the martingale problem exists and is unique for every starting point. Since a proof has not appeared anywhere, we will give a quick sketch. From the way uniqueness is proved in [2] for the general case the basic step is a perturbation in  $L_p$ . When  $d = 1$ ,  $p$  can be taken to be 2. We write the operator

$$(6.1) \quad \frac{1}{2} a(t, x) \frac{\partial}{\partial x^2} + \frac{\partial}{\partial t},$$

as

$$(6.2) \quad \frac{1}{2} \ell \frac{\partial}{\partial x^2} + \frac{1}{2} (a(t, x) - \ell) \frac{\partial}{\partial x^2} + \frac{\partial}{\partial t},$$

where  $\ell$  is chosen to be a suitable large number. For this to work in  $L_2$ , where the norms of the various operators are explicitly computable, it suffices that

$$(6.3) \quad |a(t, x) - \ell| \leq \ell' < \ell.$$

If  $a$  is such that  $0 < \alpha_1 \leq a \leq \alpha_2 < \infty$ , then  $\ell$  can be taken as  $\alpha_2$  and  $\ell'$  can be  $\alpha_2 - \alpha_1$ . Since uniqueness is purely a local property, our assertion follows.

We fix  $0 < \alpha_1 < \alpha_2 < \infty$  and consider the totality of all bounded measurable  $a(t, x)$  satisfying  $\alpha_1 \leq a(t, x) \leq \alpha_2$ . For simplicity, we shall assume that  $b$  is identically zero. By existence and uniqueness mentioned above, we know that there are transition probabilities  $\{p_a(s, x, t, dy)\}$  depending on  $a$ . As  $a$  varies over the above class, one can check that  $\{p_a\}$  varies over a compact family. The

convergence notion is weak convergence in  $dy$  which is uniform over compact sets of  $s, x$  and  $t$ . The limits are all again transition probabilities corresponding to some  $a$  in the same class.

This means that there is a notion of “weak” convergence such that  $a_n \rightarrow a$  “weakly” if and only if  $p_{a_n} \rightarrow p_a$  as described above. It will of course be very interesting to know precisely what this convergence is. There are two special cases worth noticing. If  $a_n(t, x)$  are purely functions of  $x$ , then  $p_{a_n} \rightarrow p_a$  if and only if

$$(6.4) \quad \int_{\ell_1}^{\ell_2} \frac{dx}{a_n(x)} \rightarrow \int_{\ell_1}^{\ell_2} \frac{dx}{a(x)}$$

for  $-\infty < \ell_1 < \ell_2 < \infty$ . If  $a_n(t, x)$  are purely functions of  $t$ , then  $p_{a_n} \rightarrow p_a$  if and only if

$$(6.5) \quad \int_{\ell_1}^{\ell_2} a_n(s) ds \rightarrow \int_{\ell_1}^{\ell_2} a(s) ds$$

for  $0 \leq \ell_1 \leq \ell_2 < \infty$ .

For the general case when  $a_n$  depends on both  $t$  and  $x$ , these special cases provide conflicting clues. The problem is more involved than one expects to begin with.

### 7. An example

Let us define the coefficients

$$(7.1) \quad a_+^h(x) = \begin{cases} \alpha & \text{if } \left[ \frac{x}{h} \right] \text{ is even,} \\ \beta & \text{if } \left[ \frac{x}{h} \right] \text{ is odd.} \end{cases}$$

and

$$(7.2) \quad a_-^h(x) = \begin{cases} \alpha & \text{if } \left[ \frac{x}{h} \right] \text{ is odd,} \\ \beta & \text{if } \left[ \frac{x}{h} \right] \text{ is even,} \end{cases}$$

$\beta > \alpha > 0$ . (For  $x < 0$ ,  $[x] = -[-x] - 1$ .)

Let  $\pi_{\pm}^h(t, x, dy)$  be the transition probabilities corresponding to the homogeneous coefficients  $a_{\pm}^h(x)$ . Let us define

$$(7.3) \quad a^{k,h}(t, x) = \begin{cases} a_+^h(x) & \text{if } \left[ \frac{t}{k} \right] \text{ is even,} \\ a_-^h(x) & \text{if } \left[ \frac{t}{k} \right] \text{ is odd.} \end{cases}$$

Let  $\pi^{k,h}(s, x, t, dy)$  be the transition probability corresponding to  $a^{k,h}$ . We denote by  $P_{s,x}^{k,h}$ , the process corresponding to  $a^{k,h}$  starting off at time  $s$  from  $x$ .

**THEOREM 7.1.** *Let  $h_n$  and  $k_n$  tend to zero such that  $\rho_n = k_n h_n^{-2}$  converges to a limit  $\rho, 0 \leq \rho \leq \infty$ . Then  $\pi^{k_n, h_n}$  converges to a limit  $\pi$  which is Brownian motion*

$$(7.4) \quad \pi(s, x, t, dy) = [2\pi(t - s)\sigma^2]^{-1/2} \exp \left\{ -\frac{|y - x|^2}{2\sigma^2(t - s)} \right\},$$

where  $\sigma^2 = \sigma^2(\rho)$  is a continuous function of  $\rho$  with  $\sigma^2(0) = \frac{1}{2}(\alpha + \beta)$  and  $\sigma^2(\infty) = 2\alpha\beta/(\alpha + \beta)$ .

**PROOF.** There are always convergent subsequences because of compactness. The limits are uniform in  $s, x$ , and  $t$ , so that the periodicity of  $\pi^{k,h}$  leads to the invariance of the limit  $\pi$ , with respect to space and time translations. Therefore, any possible limit  $\pi$  is Brownian motion. We only have to compute  $\sigma^2$ .

Let us denote by  $\pi_n$  the transition probability  $\pi^{k_n, h_n}$ , by  $a_n$  the corresponding coefficient and by  $P_n$  the process starting from 0 at time 0. Let us define

$$(7.5) \quad \sigma_n^2 = E^{P_n}|x(1)|^2 = E^{P_n} \int_0^1 a_n(s, x(s)) ds.$$

The computation of  $\sigma_n^2$  really involves an idea of how much time, on the average, the process spends in the regions where  $a = \alpha$  and  $a = \beta$ . We shall suppose that  $N$  is such that  $2Nk_n = 1$ . This involves at most an error of magnitude  $k_n$  in the computation of  $\sigma_n^2$ . Instead of considering the process  $x(s)$ , let us consider the process

$$(7.6) \quad y(s) = h_n^{-1}x(k_n s).$$

We denote by  $\theta_n$  the measure corresponding to it. The generator of the  $y(s)$  process can be written as

$$(7.7) \quad \frac{1}{2} \rho_n a(t, x) \frac{\partial^2}{\partial x^2},$$

where  $p_n = k_n h_n^{-1}$  and  $a(t, x) = a^{1,1}(t, x)$ . Now for the  $y$  process it is a question of how much time is spent in regions  $\{a = \alpha\}$  and  $\{a = \beta\}$  relatively up to time  $2N$ . Since the problem is periodic, we consider the reduced problem on the circle.  $A^+$  and  $A^-$  are the upper and lower semicircles.  $\pi^+$  and  $\pi^-$  are the transition probabilities for the homogeneous processes corresponding to  $a^+(x)$  and  $a^-(x)$ , where

$$(7.8) \quad a^+(x) = \begin{cases} \alpha & \text{on } A^+, \\ \beta & \text{on } A^-. \end{cases}$$

and

$$(7.9) \quad a^-(x) = \begin{cases} \beta & \text{on } A^+, \\ \alpha & \text{on } A^-. \end{cases}$$

If  $\pi^\pm(t)$  denotes the semigroup corresponding to  $a^\pm$ , then the  $\rho_n$  in front of the generator changes  $\pi^\pm(t)$  to  $\pi^\pm(t\rho_n)$ . We can now write

$$\begin{aligned}
 (7.10) \quad \sigma_n^2 &= E^{P_n} \left[ \int_0^1 a_n(s, x(s)) ds \right] \\
 &= k_n E^{Q_n} \left[ \int_0^{2N} a(s, y(s)) ds \right] \\
 &= k_n \sum_{r=1}^N E^{Q_n} \int_{(2r-1)}^{2r} a(s, y(s)) ds + k_n \sum_{r=0}^{N-1} E^{Q_n} \int_{2r}^{(2r+1)} a(s, y(s)) ds \\
 &= k_n \sum_{r=1}^N E^{Q_n} \int_{(2r-1)}^{2r} a^-(y(s)) ds + k_n \sum_{r=0}^{N-1} E^{Q_n} \int_{2r}^{(2r+1)} a^+(y(s)) ds.
 \end{aligned}$$

Since  $k_n$  is nearly  $(2N)^{-1}$ , it suffices to look at  $E^{Q_n} \int_{2r}^{(2r+1)} a^+(y(s)) ds$  and  $E^{Q_n} \int_{(2r-1)}^{2r} a^-(y(s)) ds$  for large  $r$  and  $n$ .

We have

$$(7.11) \quad E^{Q_n} \int_{2r}^{(2r+1)} a^+(y(s)) ds = [\pi^+(\rho_n)\pi^-(\rho_n)]^r \left( \int_0^1 \pi^+(s\rho_n)a^+ ds \right)$$

and

$$(7.12) \quad E^{Q_n} \int_{(2r-1)}^{2r} a^-(y(s)) ds = [\pi^-(\rho_n)\pi^+(\rho_n)]^{r-1} \left( \int_0^1 \pi^-(s\rho_n)a^- ds \right).$$

As  $\rho_n$  tends to  $\rho$  and  $n \rightarrow \infty$ , it is clear that we have to look at the invariant measures  $\mu_\rho^+$  and  $\mu_\rho^-$  solving  $\mu_\rho^+ \pi^+(\rho)\pi^-(\rho) = \mu_\rho^+$  and  $\mu_\rho^- \pi^-(\rho)\pi^+(\rho) = \mu_\rho^-$ . We can then write

$$(7.13) \quad \lim_{n \rightarrow \infty} \sigma_n^2 = \sigma^2 = \frac{1}{2} \left[ \int_0^1 \mu_\rho^+ \pi^+(s\rho)a^+ ds + \int_0^1 \mu_\rho^- \pi^-(s\rho)a^- ds \right],$$

where  $\sigma^2$  is of course only a function of  $\rho$ . By standard techniques, one can justify all the steps and even compute  $\sigma^2(\rho)$  when  $\rho = 0$  or  $\infty$ .

REFERENCES

[1] N. V. KRYLOV, "On the first boundary value problem for elliptic equations of second order," *Differentsial'nye Uravneniya*, Vol. 3 (1967), pp. 315-326. (In Russian.)  
 [2] D. W. STROOCK and S. R. S. VARADHAN, "Diffusion processes with continuous coefficients I," *Comm. Pure Appl. Math.*, Vol. 22 (1969), pp. 345-400.  
 [3] ———, "Diffusion processes with continuous coefficients II," *Comm. Pure Appl. Math.*, Vol. 22 (1969), pp. 479-530.  
 [4] ———, "Diffusion processes with boundary conditions," *Comm. Pure Appl. Math.*, Vol. 24 (1971), pp. 147-225.