The Annals of Probability
1986, Vol. 14, No. 3, 860–872

COUPLING OF MULTIDIMENSIONAL DIFFUSIONS BY REFLECTION

By Torgny Lindvall and L. C. G. Rogers

University of Göteborg and University College of Swansea

If $x \neq x'$ are two points of $\mathbb{R}^d$, $d \geq 2$, and if X is a Brownian motion in $\mathbb{R}^d$ started at x, then by reflecting X in the hyperplane $L \equiv \{y: |y - x| = |y - x'|\}$ we obtain a Brownian motion X' started at x', which couples with X when X first hits L. This paper deduces a number of well-known results from this observation, and goes on to consider the analogous construction for a diffusion X in $\mathbb{R}^d$ which is the solution of an s.d.e. driven by a Brownian motion B; the essential idea is the reflection of the increments of B in a suitable (time-varying) hyperplane. A completely different coupling construction is given for diffusions with radial symmetry.

1. Introduction. Let the paths of two independent Markov processes, with the same transition probabilities but different initial distributions, run until the first instant they hit the same state; that instant is the *coupling time*, T say. From T on, the distributions of the two processes are the same, so if $T < \infty$ a.s. we should be able to conclude that the processes are asymptotically equally distributed. Indeed, if they are denoted by X and X' and $Q_\lambda(t), Q_\mu(t)$ are the distributions of X_t, X_t' when $X_0 =_{\mathscr{D}} \lambda$, $X_0' =_{\mathscr{D}} \mu$, respectively, then in general the basic coupling inequality

$$\|Q_\lambda(t) - Q_\mu(t)\| \leq 2P(T > t) \tag{1}$$

holds; if $T < \infty$ a.s. then $P(T > t) \to 0$ as $t \to \infty$, and (1) implies the ergodicity result $\|Q_\lambda(t) - Q_\mu(t)\| \to 0$ as $t \to \infty$ ($\|\cdot\|$ is the total variation norm). The coupling is *successful* if $T < \infty$ a.s.

The coupling method has enjoyed much interest during the last decade and has now also found its way into the textbooks: cf. Billingsley [2], Grimmett and Stirzaker [8], and Karlin and Taylor [11]. For several of the possibilities of the device, see Griffeath [7]. As is apparent from the latter account, a construction of suitably dependent processes X, X' is often the crucial step in order to obtain a successful coupling; notice that (1) holds if $X_t = X_t'$ for $t \geq T$.

The method is well fitted to the study of one-dimensional diffusion processes since they have continuous paths; two such cannot pass each other without hitting each other, an observation that simplifies estimates about T considerably, cf. Lindvall [13]. But the paths of two independent diffusions in $\mathbb{R}^d$, $d \geq 2$, do not in general ever meet. Hence, in order to apply the coupling method to multidimensional diffusions, we are forced to special constructions. The purpose

Received July 1984; revised February 1985.

AMS 1980 *subject classifications*. Primary 60J60, 60J65, 60H10; secondary 60J45, 60J70.

Key words and phrases. Coupling, Brownian motion, multidimensional diffusion, stochastic differential equation, reflection, skew product, stationary distribution, radial process, tail σ-field of a one-dimensional diffusion.

of this paper is to present a number of cases where such constructions may be carried out.

When considering Brownian motions, the task is easy: we let one of the paths be the mirror image of the other, with mirror surface = the plane (obvious which) between the two starting points, until they meet in that plane; after coupling we let the paths coincide. In Section 2 we present the details of that coupling and examples of its applications, to the asymptotics of the n free particles model of gas kinetics, simple potential theory, and other topics.

The reflexion method of coupling just described for Brownian motion is a special case of a general one for coupling solutions of the stochastic differential equation

$$dX_t = \sigma(X_t)\,dB_t + b(X_t)\,dt. \tag{2}$$

The idea is to construct X' by solving a companion equation

$$dX_t' = \sigma(X_t')H(X_t, X_t')\,dB(t) + b(X_t')\,dt,$$

where $H(x, x')$ is an orthogonal matrix of determinant -1, a reflexion, essentially. The geometric interpretation is less direct, but we shall consider the effect of this coupling method in Section 3, under the assumption that σ does not vary too much, and we derive sufficient conditions for successful coupling.

If $\sigma \equiv I$ and $b(x) = -x\beta(|x|)$ for some β: $\mathbb{R}_+ \to \mathbb{R}$, then the radial process $|X_t|$, $t \geq 0$, is a one-dimensional diffusion in its own right. This, together with a skew product representation of X, permits a somewhat different method of coupling which can be completely analysed; this is done in Section 4.

Notice that if there exists a stationary distribution π, then $Q_\pi(t) = \pi$ for all $t \geq 0$, so if a successful coupling is obtained then (1), with $\mu = \pi$, yields a result on convergence toward stationarity: $Q_\lambda(t) \to \pi$ (in total variation norm) as $t \to \infty$. Sufficient conditions on b and σ for the existence of a stationary distribution are established by Bhattacharaya [1]; these conditions are rather involved in general and depend on comparison with one-dimensional radial processes, but in the special case of radial drift this comparison is exact, and the scale and speed picture of one-dimensional diffusions permits considerable simplifications: see Section 4.

2. Coupling of Brownian motions. Let Ω = the space of continuous functions ω on $[0, \infty)$ with values in $\mathbb{R}^d$, $d \geq 2$, endow Ω with the standard σ-field $\mathscr{F}$, and let P_x be the probability measure on $(\Omega, \mathscr{F})$ governing a Brownian motion starting at x. Our mirror surfaces are the hyperplanes

$$L_{xy} = \{u; (u - (x+y)/2, x - y) = 0\}$$

for $x, y \in \mathbb{R}^d$ if $x \neq y$ [$(\cdot,\cdot)$ denotes scalar product]; the definition of L_{xy} is arbitrary if $x = y$. With $T_{xy}(z)$ = the mirror image of z with respect to L_{xy}, we define $\beta_{xy}\omega \in \Omega$ through

$$(\beta_{xy}\omega)(t) = \begin{cases} T_{xy}(\omega(t)), & t \leq \kappa_{xy}, \\ \omega(t), & t \geq \kappa_{xy}, \end{cases}$$

where

$$\kappa_{xy} = \inf\{s > 0; \omega(s) \in L_{xy}\};$$

that infimum is set to ∞ if ω never hits L_{xy}.

With $X_t = \omega(t)$ and $X'_t = [\beta_{xy}X](t)$, it is easily seen that $X =_{\mathscr{D}} P_x$ implies $X' =_{\mathscr{D}} P_y$. With $Y_t = (X_t, X'_t)$, let P_{xy} be the distribution of Y when $X =_{\mathscr{D}} P_x$; the coupling of two Brownian motions starting at x and y, respectively is then the process Y with governing measure P_{xy}. We have random starting points with distributions λ, μ if we let Y be governed by $P_{\lambda,\mu}$ where

$$P_{\lambda,\mu}(\cdot) = \int P_{xy}(\cdot)[\lambda \times \mu](dx, dy).$$

In fact, any probability measure on $(\mathbb{R}^d)^2$ will do if it has marginals λ and μ; for certain problems, there are better choices than $\lambda \times \mu$.

This type of somewhat tedious definition will not appear again.

Under P_{xy}, we have $X_t =_{\mathscr{D}} N(x, tI)$ and $X'_t =_{\mathscr{D}} N(y, tI)$. It is easily proved that

$$\begin{aligned} |N(x, tI) - N(y, tI)| &= 2\Phi(-|x-y|/2t^{1/2}, |x-y|/2t^{1/2}) \\ &\le 2 \cdot |x-y|/(2\pi t)^{1/2}, \end{aligned} \tag{3}$$

where Φ is the standard normal distribution; " $\le$ " may be replaced by " $\sim$ ". It may amuse the reader to prove (3) by using the basic coupling inequality (1): we notice that $T = \kappa_{xy}$ and, actually,

$$\|N(x, tI) - N(y, tI)\| = 2P_{xy}(\kappa_{xy} > t). \tag{4}$$

EXAMPLE 1. For $t \ge 0$, let ξ_t be the position in $(\mathbb{R}^d)^n$ of an n-tuple of independent Brownian motions at time t, representing a configuration of n gas molecules at that time. In order to understand how quickly the initial configuration $\{x_1, \dots, x_n\}$ is forgotten, we shall estimate $\|\mathbb{P}\xi_t^{-1} - \mathbb{P}\xi_t'^{-1}\|$; here ξ'_t is the analogue of ξ_t with initial configuration $\{y_1, \dots, y_n\}$. For the estimate, let X_i, X'_i be Brownian motions with $X_i(0) = x_i$, $X'_i(0) = y_i$ and coupled as above. We find that

$$\begin{aligned} \|\mathbb{P}\xi_t^{-1} - \mathbb{P}\xi_t'^{-1}\| &\le 2\mathbb{P}(\xi_t \ne \xi'_t) \\ &\le 2\mathbb{P}\big(\kappa_{x_i y_i} > t \text{ for some } i, 1 \le i \le n\big) \le 2\sum_1^n \mathbb{P}\big(\kappa_{x_i y_i} > t\big), \end{aligned} \tag{5}$$

which is bounded by $2(\sum_1^n |x_i - y_i|)/(2\pi t)^{1/2}$ due to (3) and (4).

For a nonnegative random variable τ defined on $(\Omega, \mathscr{F})$, let

$$h_\tau(x, A) = P_x(\tau < \infty, X_\tau \in A)$$

and $\tau' = \tau \circ \beta_{xy}$, for $x, y \in \mathbb{R}^d$, $A \in \mathscr{R}^d$. For the next examples, we need an estimate of $\|h_\tau(x, \cdot) - h_\tau(y, \cdot)\|$ under the assumption

$$\tau = \tau' \quad \text{on } \{\kappa_{xy} \le \tau \wedge \tau'\}. \tag{6}$$

To that end, let

$$h_\tau(x, A) = P_{xy}(\tau < \infty, \tau' < \infty, X_\tau \in A) + P_{xy}(\tau < \infty, \tau' = \infty, X_\tau \in A)$$

and split $h_\tau(y, A)$ analogously. Using (6), it is easily deduced that

$$(7)\quad \|h_\tau(x, \cdot) - h_\tau(y, \cdot)\| \leq 2P_{xy}(\kappa_{xy} > \tau \wedge \tau') + P_{xy}(\{\tau = \infty\} \triangle \{\tau' = \infty\})$$

for all $x, y \in \mathbb{R}^d$.

EXAMPLE 2. For a $\mathscr{F}_\sigma$ set $B \subset \mathbb{R}^d$, let

$$\tau_B = \inf\{s > 0; X(s) \in B\}.$$

Following Port and Stone [14], we shorten $h_{\tau_B}(x, \cdot)$ to $h_B(x, \cdot)$; for concepts and results on Brownian motion and potential theory, see Chapters 2 and 4 of [14]. If D is an open set with a smooth boundary and D^c recurrent, then the Dirichlet problem has the unique solution

$$f(x) = E_x[\varphi(X(\tau_{D^c}))] = \int \varphi(y) h_{D^c}(x, dy),$$

where φ is the given bounded and continuous function defined on ∂D; E_x is expectation with respect to P_x.

With the help of (7), we find that

$$(8)\quad |f(x) - f(y)| \leq 2\|\varphi\| P_{xy}(\kappa_{xy} > \tau_{D^c} \wedge \tau'_{D^c})$$

for $x, y \in D$.

EXAMPLE 3. For $d = 2$, let μ_B be the equilibrium measure of a bounded nonpolar $\mathscr{F}_\sigma$ set B, cf. [14], page 77. We present a short proof of the result that

$$(9)\quad \|h_B(x, \cdot) - \mu_B\| \to 0 \quad \text{as } |x| \to \infty.$$

Now $\mu_B = \int h_B(z, \cdot)\sigma_r(dz)$ if $B \subset B_r$, where σ_r is the uniform distribution on the surface of the ball B_r with radius r, centered at 0. Hence, if $|x|$ is large enough,

$$\|h_B(x, \cdot) - \mu_B\| \leq 2P_{x, \sigma_{|x|}}(\kappa_{x, \sigma_{|x|}} > \tau_B \wedge \tau'_B)$$

due to an obvious extension of (7) covering the case when y is randomly chosen (with distribution $\sigma_{|x|}$). Now fix a ball B_r so large that $B \subset B_r$. Then, if $|x| \geq r$, we have

$$P_{x, \sigma_{|x|}}(\kappa_{x, \sigma_{|x|}} > \tau_B \wedge \tau'_B) \leq P_{x, \sigma_{|x|}}(\kappa_{x, \sigma_{|x|}} > \tau_{B_r}).$$

But the latter probability tends to 0 as $|x| \to \infty$ which is seen after a scale change such that the two starting points lie on the unit circle and use of $P_z(\tau_{B_\alpha} \to \infty$ as $\alpha \to 0) = 1$ for all z with $|z| = 1$.

EXAMPLE 4. For a set $B \in \mathscr{F}_\sigma$, the function $e_B(x) = P_x(\tau_B < \infty)$ is the equilibrium potential of B, cf. [14], page 58. Fix $x, y \in R^d$; if B is a subset of the half-space bounded by L_{xy} and containing x, then

$$e_B(y) \le e_B(x) \tag{10}$$

as follows directly from $\tau_B \le \tau'_B$.

It is left to the reader to couple Brownian motions on a circle, a torus, a cylinder, or a sphere. For spherical Brownian motion, see Itô and McKean [10], page 269f. Notice that the device is useful for proving ergodicity of any nontrivial diffusion on a circle.

3. Coupling of solutions of certain stochastic differential equations. Consider the diffusion $X = (X_t)_{t \ge 0}$ in $\mathbb{R}^d$ defined by the stochastic differential equation (s.d.e.)

$$dX_t = \sigma(X_t)\,dB_t + b(X_t)\,dt, \qquad X_0 = x. \tag{11}$$

Here, B is d-dimensional Brownian motion, $b\colon \mathbb{R}^d \to \mathbb{R}^d$, and $\sigma\colon \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is nonsingular. We shall always assume that

(12)(i) σ and b are Lipschitz,
(12)(ii) σ and b are bounded, and
(12)(iii) σ^{-1} is bounded.

The first condition ensures the existence of a pathwise unique strong solution to (11); the others can be relaxed by localisation techniques, but we avoid this for the sake of clarity.

In order to construct some other diffusion $X' = (X'_t)_{t \ge 0}$ with the same generator as X but started at $x' \ne x$ (y is used to denote other quantities in this section), with a coupling of X and X' in mind, the idea is to consider the stochastic differential equation

$$dX'_t = \sigma(X'_t)\,dB'_t + b(X'_t)\,dt, \qquad X'_0 = x', \tag{13}$$

where

$$dB'_t = H_t\,dB_t$$

and H is some previsible process with values in $\mathcal{O}(d) =$ the set of $d \times d$ orthogonal matrices. By cunning choice of H, it is possible to make X and X' couple if the coefficients σ and b are good.

We define

$$\begin{aligned} Y_t &\equiv X_t - X'_t, & V_t &\equiv |Y_t|^{-1} Y_t, \\ \alpha_t &\equiv \sigma(X_t) - \sigma(X'_t) H_t, & \beta_t &\equiv b(X_t) - b(X'_t). \end{aligned} \tag{14}$$

Then by Itô's formula,

$$d(|Y_t|) = (V_t, \alpha_t\,dB_t) + \tfrac{1}{2}|Y_t|^{-1}\left[2(Y_t, \beta_t) + \operatorname{tr} \alpha_t \alpha_t^T - |\alpha_t^T V_t|^2\right] dt. \tag{15}$$

REMARK. The right-hand side makes no sense if $Y_t = 0$, nor is V_t well defined, but these are essentially irrelevant points, since we are only interested in behaviour of Y up to the first hit of zero.

We are attempting to make Y hit zero in finite time, and there are potentially three reasons why it might not do so:

(16)(a) the drift $|Y_t|^{-1}(Y_t, \beta_t)$ on the right side of (15) might be too big for large values of $|Y|$,
(16)(b) the drift $|Y_t|^{-1}(\operatorname{tr} \alpha_t \alpha_t^T - |\alpha_t V_t|^2)$ might be unbounded for small values of $|Y|$ (it is always nonnegative), or
(16)(c) the covariance $|\alpha_t^T V_t|^2$ of the martingale part of $|Y|$ might tend to zero so that, although Y might *tend* to zero it will never hit it in finite time.

We shall present a construction analogous to the reflexion couplings of Section 2 which takes care of (b) and (c), at least in examples where σ is suitably well behaved. In general, localisation is needed to extend the construction; details of this will appear later.

This then leaves only (a), which is essentially the old problem of deciding recurrence of a diffusion in $\mathbb{R}^d$, which is handled by the old techniques. The following lemma is very similar to results of Khasminskii [12], Friedman [5], and Ikeda and Watanabe [9]; we leave the proof as an exercise.

LEMMA 1. *Let $R = (R_t)_{t \geq 0}$ be a continuous semimartingale such that $R_0 > 0$ a.s. and*

$$dR_t = K_t\, dW_t + \theta_t\, dt,$$

where $(W_t)_{t \geq 0}$ is one-dimensional Brownian motion, $(K_t)_{t \geq 0}, (\theta_t)_{t \geq 0}$ are previsible processes, and for some $K > 0$, $K_t > K$ for all t.

Suppose there exists $\gamma\colon (0, \infty) \to \mathbb{R}$ which is locally integrable and such that

$$\theta_t \leq K_t^2 \gamma(R_t) \quad \text{for all } t, \quad \text{a.s.} \tag{17}$$

Let $f\colon (0, \infty) \to R$ be defined by

$$f'(r) = \exp\left(-2\int_1^r \gamma(u)\, du\right), \qquad f(1) = 0. \tag{18}$$

For $\tau_\varepsilon \equiv \inf\{t;\, R_t \leq \varepsilon\}$, $\tau_0 \equiv \lim_{\varepsilon \searrow 0} \tau_\varepsilon$ we then have

(i) *if* $\lim_{x \to \infty} f(x) = \infty$, *then* $P(\tau_\varepsilon < \infty) = 1$ *for all* $\varepsilon > 0$, *and*
(ii) *if* $\lim_{x \to \infty} f(x) = \infty$, $\lim_{x \to 0} f(x) > -\infty$ *and* $\liminf_{x \to 0} f'(x) > 0$, *then*

$$P(\tau_0 < \infty) = 1.$$

We now give the construction of $H_t \equiv H(X_t, X_t')$. It is easiest to understand if σ is *constant*, but will work equally well for sufficiently small perturbation of such a σ; let

$$c \equiv \sup\left\{ \|\sigma(x) - \sigma(x')\|;\, x, x' \in R^d \right\}.$$

The $\mathcal{O}(d)$-valued matrix H is given by

$$H(x, x') \equiv I - 2u(x, x')u(x, x')^T, \tag{19}$$

where $u(x, x')$ is the unit vector $\sigma(x')^{-1}y/|\sigma(x')^{-1}y|$, and H is locally Lipschitz on $(\mathbb{R}^d)^2 \setminus \{(x, x); x \in \mathbb{R}^d\}$. The matrix H is reflexion in the plane orthogonal to $\sigma(x')^{-1}y$, and as such generalises the reflexion construction of Section 2; if $\sigma = I$, $b = 0$, it yields the same pair (X, X') as was considered there.

Abbreviate $\sigma(x')$ to σ, $u(x, x')$ to u and so forth. If we define $\delta = \delta(x, x') = \sigma(x) - \sigma(x')$ then

$$\alpha \equiv \sigma(x) - \sigma(x')H(x, x') = \delta + 2\sigma u u^T, \tag{20}$$

and a few calculations yield the identity

$$\operatorname{tr} \alpha\alpha^T - |\alpha^T v|^2 = \operatorname{tr} \delta\delta^T - |\delta^T v|^2 = O(|y|^2) \tag{21}$$

since σ is assumed Lipschitz. The whole point of this construction was that now the drift term on the right side of (15) is

$$(v, \beta) + \tfrac{1}{2}|y|^{-1}\{\operatorname{tr} \alpha\alpha^T - |\alpha^T v|^2\},$$

which *remains bounded* as $y \to 0$. This takes care of (16)(b), and, provided c is small enough, it also takes care of (16)(c). Indeed, if $\Lambda \equiv \sup_x \|\sigma^{-1}(x)\|$, then

$$\begin{aligned} |\alpha^T v| &= |\delta^T v + 2uu^T\sigma^T v| \\ &\geq 2|u^T\sigma^T v| - |\delta^T v| \\ &= 2|y|/|\sigma^{-1}y| - |\delta^T v| \\ &\geq 2\Lambda^{-1} - c. \end{aligned}$$

Thus provided $c < 2\Lambda^{-1}$, $|\alpha^T v|$ is bounded away from zero.

Now define for $r > 0$

$$\begin{aligned} \gamma(r) &= \sup\{|\alpha^T v|^{-2}[(v, \beta) + (\operatorname{tr} \alpha\alpha^T - |\alpha^T v|^2)/2|y|]; \\ &\qquad\qquad x, x' \text{ such that } |x - x'| = r\} \\ &= \sup\{|\alpha^T v|^{-2}[(v, \beta) + (\operatorname{tr} \delta\delta^T - |\delta^T v|^2)/2|y|^2]; \\ &\qquad\qquad x, x' \text{ such that } |x - x'| = r\}. \end{aligned} \tag{22}$$

[Here, of course, $\beta = b(x) - b(x')$.]

We are now in position to use Lemma 1 on the continuous semimartingale $R_t = |Y_t|$.

THEOREM 1. *Suppose σ, b are Lipschitz and bounded, and that*

$$\Lambda \equiv \sup_x \|\sigma^{-1}(x)\| < \infty.$$

Suppose also that

$$\sup_{x, x'} \|\sigma(x) - \sigma(x')\| < 2\Lambda^{-1}.$$

If

$$\int_1^\infty dr \left\{\exp\left[-2\int_1^r \gamma(u)\, du\right]\right\} = \infty, \tag{23}$$

where γ is defined at (22), *then we have a successful coupling of X and X': if $\tau_0 = \inf\{t; |Y_t| = 0\}$, then $P(\tau_0 < \infty) = 1$.*

PROOF. Taking R_t to be $|Y_t|$ in Lemma 1, we identify $K_t = |\alpha_t^T V_t| \geq 2\Lambda^{-1} - c > 0$ from (15). Moreover, by the definition of γ,

$$\theta_t \leq K_t^2 \gamma(R_t) \quad \text{for all } t.$$

By the key fact (21), γ remains bounded in a neighbourhood of 0, and so $\liminf_{x \to 0} f'(x) > 0$, and $\lim_{x \to 0} f(x) > -\infty$. The condition (23) is now simply the condition $\lim_{x \to \infty} f(x) = \infty$ of conclusion (ii) of Lemma 1, which completes the proof. □

EXAMPLE 5. We begin by considering the special case where $\sigma(x) = \sigma(0)$ for all x, and so $c = 0$, $\delta = 0$ and, more simply,

$$\gamma(r) = \sup\left\{|\alpha^T v|^{-2}(v, \beta); x, x' \text{ such that } |x - x'| = r\right\}.$$

There is one particularly obvious case where (23) is satisfied, namely, that where $(v, \beta) \leq 0$ always, or, put another way, where

$$(24) \qquad (x - x', b(x) - b(x')) \leq 0$$

for all x, x'. Notice the interpretation of this condition on the vector field b; if x_t, x_t' are two solutions of the ordinary differential equation

$$dx_t = b(x_t)\, dt,$$

then $|x_t - x_t'|$ is *decreasing*, so the integral curves of b do not spread out as time passes. This is intuitively very appealing: if b is such as to keep solutions together, then by including a stochastic term we can ensure a successful coupling.

If the vector field b is assumed differentiable, then the condition (24) holds iff for all $z \in \mathbb{R}^d$

$$\sum_{i,j=1}^{d} z_i z_j\, \partial_j b_i(x) \leq 0, \quad \forall x.$$

This condition will be satisfied if, for example, b is the gradient of a concave function, or, again, if $b(x) = Ax$ where A is antisymmetric.

EXAMPLE 6. Let us suppose $b \equiv 0$. If σ is constant, then coupling takes place, as in the discussion of the example above. How far can one perturb σ from a constant matrix and still be certain of coupling? In this case,

$$\gamma(r) = \sup\left\{|\alpha^T v|^{-2}(\operatorname{tr} \delta\delta^T - |\delta^T v|^2)/2r; x, x' \text{ such that } |x - x'| = r\right\},$$

so if we had

$$(25) \qquad |\alpha^T v|(\operatorname{tr} \delta\delta^T - |\delta^T v|^2) \leq 1$$

everywhere, then $\gamma(r) \leq 1/2r$, and condition (23) holds; the coupling is successful. Now

$$0 \leq \operatorname{tr} \delta\delta^T - |\delta^T v|^2 \leq (d-1)c^2$$

and $|\alpha^T v| \geq 2\Lambda^{-1} - c$, so condition (25) is satisfied if

$$c \leq 2\Lambda^{-1}/\left(1 + (d-1)^{1/2}\right),$$

that is, if the perturbation from the case $c = 0$ of a constant covariance is not too large, we still have a successful coupling.

4. The radial drift case. In this final section, we consider a diffusion $X = (X_t)_{t \geq 0}$ defined as the solution of an s.d.e. of the form

$$dX_t = dB_t - X_t \beta(|X_t|)\,dt, \tag{26}$$

where $\beta\colon \mathbb{R}^+ \to \mathbb{R}$ is Lipschitz and, say, bounded. Itô's formula gives

$$d|X_t| = (U_t, dB_t) + \left\{\tfrac{1}{2}(d-1)|X_t|^{-1} - |X_t|\beta(|X_t|)\right\} dt \tag{27}$$

and

$$dU_t = |X_t|^{-1}\left(I - U_t U_t^T\right) dB_t - \tfrac{1}{2}(d-1) U_t |X_t|^{-2}\,dt, \tag{28}$$

where $U_t = X_t/|X_t|$. In particular, if the Brownian motion W is defined by $dW_t = (U_t, dB_t)$, then $|X_t|$ solves the stochastic differential equation

$$d|X_t| = dW_t + \left\{\tfrac{1}{2}(d-1)|X_t|^{-1} - |X_t|\beta(|X_t|)\right\} dt.$$

This s.d.e. has a pathwise unique strong solution which is a time-homogeneous diffusion process, by general results on s.d.e.'s; see Yamada and Watanabe [16]. As in the skew-product representation of Brownian motion, we deduce from (28) that if Z is a Brownian motion on S^{d-1}, and if R is a diffusion on $\mathbb{R}^+$ solving the s.d.e. (27), then

$$X_t = R_t Z\left(\int_0^t R_s^{-2}\,ds\right).$$

So the study of X reduces essentially to the study of the one-dimensional diffusion $|X|$.

The coupling construction will be as follows. Let X and X' be two *independent* processes distributed as the solution of (26), but with different starting points. We run the two processes until $\tau \equiv \inf\{t;\ |X_t| = |X'_t|\}$, after which we make X' follow the path of X reflected in the hyperplane orthogonal to $X_\tau - X'_\tau$. This hyperplane passes through 0, and the coupling time is the first time at which X hits the hyperplane after τ. Thus there are two problems: firstly to show that $\tau < \infty$, and secondly to show that the hyperplane must be hit in finite time. Whether these stopping times are finite or not depends on properties of the diffusion $|X_t|$. If we make the assumption that $\beta \geq 0$, then the methods of the previous section guarantee successful coupling, since

$$(x - x', b(x) - b(x')) \leq 0$$

for all x, x'. This is an improvement over the results we shall now obtain, but we should emphasise that we do not here assume $\beta \geq 0$.

Let $s\colon (0,\infty) \to R$ be the scale function of $|X|$; explicitly,

$$s(r) = \int_1^r \exp\left[-2\int_1^x \left\{\tfrac{1}{2}(d-1)u^{-1} - u\beta(u)\right\} du\right] dx$$
$$= \int_1^r x^{-(d-1)} \exp\left[2\int_1^x u\beta(u)\,du\right] dx; \tag{29}$$

see, for example, Breiman [3]. Since β is assumed bounded, it is clear that $s(0+) = -\infty$, so the diffusion is recurrent iff $s(\infty) = \infty$. If $s(\infty) < \infty$, we shall suppose without loss of generality that $s(\infty) = 0$.

Now let $(Y_t^1)_{t\geq 0}$ and $(Y_t^2)_{t\geq 0}$ be two independent copies of the diffusion with law $(s(|X_t|))_{t\geq 0}$, with starting points y_1 and y_2, respectively. Define

$$\mathscr{G}_t^i \equiv \sigma\left(Y_s^i\colon s \geq t\right), \qquad \mathscr{G}_\infty^i \equiv \bigcap_{t\geq 0} \mathscr{G}_t^i, \qquad i = 1,2,$$

and

$$\mathscr{G}_t \equiv \mathscr{G}_t^1 \vee \mathscr{G}_t^2, \qquad \mathscr{G}_\infty \equiv \bigcap_{t\geq 0} \mathscr{G}_t.$$

We defer the proof of the following result, which holds in much greater generality than the present context.

LEMMA 2. $\mathscr{G}_\infty = \mathscr{G}_\infty^1 \vee \mathscr{G}_\infty^2$.

REMARK. The inclusion $\mathscr{G}_\infty \supseteq \mathscr{G}_\infty^1 \vee \mathscr{G}_\infty^2$ is immediate; without independence, the reverse inclusion is false.

The finiteness of τ is decided by the following result.

LEMMA 3. (i) *If the diffusion* $s(|X|)$ *is recurrent* ($s(\infty) = \infty$), *or if the diffusion* $s(|X|)$ *is transient* ($s(\infty) < \infty$) *and*

$$\int_1^\infty \left(s(t)^2/s'(t)\right) \int_1^t du/s'(u)\,dt = \infty, \tag{30}$$

then τ *is finite a.s.*

(ii) *If the diffusion* $s(|X|)$ *is transient and* (30) *fails, then* τ *is not a.s. finite.*

PROOF. (i) According to Rösler [15], if $s(|X|)$ is recurrent then the tail σ-fields $\mathscr{G}_\infty^i$ are trivial, and, according to Fristedt and Orey [6], if $s(|X|)$ is transient and (30) holds, then the σ-fields $\mathscr{G}_\infty^i$ are trivial. If $P_{y_1 y_2}$ is the law of $(Y_t^1, Y_t^2)_{t\geq 0}$, and

$$A_+ = \left\{Y_t^1 - Y_t^2 > 0 \text{ for all large enough } t\right\},$$
$$A_- = \left\{Y_t^1 - Y_t^2 < 0 \text{ for all large enough } t\right\},$$
$$A_0 = \Omega \setminus (A_+ \cup A_-),$$

then by Lemma 2, since $\mathscr{G}_\infty^i$ are trivial and $A_\pm \in \mathscr{G}_\infty$, it must be that for any y

$$P_{yy}(A_+) = 0 \text{ or } 1.$$

But by symmetry, $P_{yy}(A_+) = P_{yy}(A_-)$, so both must be zero, and $P_{yy}(A_0) = 1$. Hence $P_{y_1y_2}(A_0) = 1$ for all y_1, y_2, and $\tau < \infty$ $P_{y_1y_2}$ a.s. since $A_0 \subseteq \{\tau < \infty\}$.

(ii) Fristedt and Orey [6] show that if $s(|X|)$ is transient and (30) fails, then $\mathscr{G}^i_\infty$ are not trivial, and there exists a continuous increasing function c such that

$$t - c(Y_t^i) \to \eta_i, \qquad i = 1,2,$$

where η_i are some nondegenerate random variables. Thus

$$P_{y_1y_2}\big(Y_t^1 > Y_t^2 \text{ for all large enough } t\big) \geq P_{y_1y_2}(\eta_2 > \eta_1) > 0,$$

and coupling is not certain: $P_{y_1y_2}(\tau < \infty) < 1$. □

Proof of Lemma 2. Take $X \in L^\infty(\mathscr{G}_\infty)$. Then $X \in L^\infty(\mathscr{G}_t^1 \vee \mathscr{G}_t^2)$ for each t, so given $\varepsilon > 0$ and t, there exists a simple random variable

$$Y = \sum_{j=1}^n \alpha_j 1_{A_j^1} 1_{A_j^2}$$

such that $E[|X - Y|] < \varepsilon$, and $A_j^i \in \mathscr{G}_t^i$ for $i = 1,2$ and $j = 1,2,\dots,n$. Moreover, if $s > t$,

$$\begin{aligned} E\big[\big|X - E\big[Y|\mathscr{G}_s^1 \vee \mathscr{G}_s^2\big]\big|\big] &= E\big[\big|E\big[X - Y|\mathscr{G}_s^1 \vee \mathscr{G}_s^2\big]\big|\big] \\ &\leq E\big[|X - Y|\big] < \varepsilon \end{aligned} \tag{31}$$

since $X \in L^\infty(\mathscr{G}_s^1 \vee \mathscr{G}_s^2)$. Now because of the independence of $\mathscr{G}_t^1$ and $\mathscr{G}_t^2$,

$$E\big[1_{A^1}1_{A^2}|\mathscr{G}_s^1 \vee \mathscr{G}_s^2\big] = E\big[1_{A^1}|\mathscr{G}_s^1\big]E\big[1_{A^2}|\mathscr{G}_s^2\big]$$

if $A^i \in \mathscr{G}_t^i$, and so

$$\begin{aligned} E\big[Y|\mathscr{G}_s^1 \vee \mathscr{G}_s^2\big] &= \sum_{j=1}^n \alpha_j E\big[1_{A_j^1}|\mathscr{G}_s^1\big]E\big[1_{A_j^2}|\mathscr{G}_s^2\big] \\ &\to \sum_{j=1}^n \alpha_j E\big[1_{A_j^1}|\mathscr{G}_\infty^1\big]E\big[1_{A_j^2}|\mathscr{G}_\infty^2\big] \quad \text{a.s.} \end{aligned}$$

by the reversed martingale convergence theorem. The limit random variable is measurable on $\mathscr{G}_\infty^1 \vee \mathscr{G}_\infty^2$, and (31) implies that X can be approximated in L^1 to within ε by a $\mathscr{G}_\infty^1 \vee \mathscr{G}_\infty^2$-measurable random variable. Hence X is a.s. equal to a $\mathscr{G}_\infty^1 \vee \mathscr{G}_\infty^2$-measurable random variable. □

For the success of the coupling construction, we still have to ensure that after time τ, the process X hits the hyperplane orthogonal to $X_\tau - X'_\tau$. This will happen if and only if the clock $C \equiv (C_t)_{t \geq 0}$,

$$C_t \equiv \int_0^t |X_s|^{-2}\, ds,$$

of the skew product representation diverges a.s. If $|X|$ is recurrent, this obviously happens, and if $|X|$ is transient, we have to ensure that if we time change Y by the inverse to C then Y does not reach $0 = s(\infty)$ in finite time. The boundary behaviour of scale and speed to guarantee this is well known: see, for example, Breiman [3] or Freedman [4]. Rephrased in terms of the present notation, the condition is

$$(32) \qquad \int_1^\infty |s(y)|/y^2 s'(y)\, dy \equiv \int_1^\infty (1/y^2 s'(y)) \left\{ \int_y^\infty s'(u)\, du \right\} dy = \infty.$$

Finally, if $|X|$ is recurrent and has a stationary distribution π_0, then by coupling with an independent copy of $|X|$ with initial distribution π_0 we can deduce that the distribution of $|X_t| \to \pi_0$ for every initial distribution of $|X|$. From this, by the skew product representation, the convergence of X_t to stationarity as $t \to \infty$ follows. The condition for $|X|$ to have a stationary distribution is that the speed measure should have finite mass; a few calculations reduce this to the simple criterion

$$\int_0^\infty s'(u)^{-1}\, du = \int_0^\infty u^{d-1} \exp\left(-2 \int_1^u v\beta(v)\, dv \right) du < \infty$$

which hence is necessary and sufficient for a stationary distribution to exist.

We summarize the result of this section.

THEOREM 2. *Suppose the diffusion $X = (X_t)_{t \geq 0}$ is defined by*

$$dX_t = dB_t - X_t \beta(|X_t|)\, dt,$$

where β is bounded and Lipschitz. The scale function of the one-dimensional diffusion $|X| = (|X_t|)_{t \geq 0}$ is

$$s(r) = \int_1^r x^{-(d-1)} \exp\left(2 \int_1^x u\beta(u)\, du \right) dx.$$

(a) *If $s(\infty) = \infty$, then $|X|$ is recurrent, and coupling of X and X' is certain using the method described above. If*

$$\int_0^\infty 1/s'(x)\, dx < \infty$$

then the diffusion has an invariant distribution to which the law of X_t (under any initial distribution) converges in norm as $t \to \infty$.

(b) *If $s(\infty) < \infty$, then taking $s(\infty) = 0$, coupling of X and X' is certain provided*

$$\text{(i)} \qquad \int_1^\infty \left(s(t)^2/s'(t) \right) \left\{ \int_1^t 1/s'(u)\, du \right\} dt = \infty$$

and

$$\text{(ii)} \qquad \int_1^\infty (1/t^2 s'(t)) \left\{ \int_t^\infty s'(u)\, du \right\} dt = \infty.$$

REFERENCES

[1] BHATTACHARYA, R. N. (1978). Criteria for recurrence and existence of invariant measures for multidimensional diffusions. *Ann. Probab.* **6** 541–553.

[2] BILLINGSLEY, P. (1979). *Probability and Measure.* Wiley, New York.

[3] BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, Mass.

[4] FREEDMAN, D. (1971). *Brownian Motion and Diffusion*. Holden-Day, San Francisco.

[5] FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications* **1**. Academic, New York.

[6] FRISTEDT, B. and OREY, S. (1978). The tail σ-field of one dimensional diffusions. In *Stochastic Analysis* 127–138. Academic, New York.

[7] GRIFFEATH, D. (1978). Coupling methods for Markov processes. In *Advances in Mathematics Supplementary Studies: Studies in Probability and Ergodic Theory* **2** 1–43. Academic, New York.

[8] GRIMMETT, G. R. and STIRZAKER, D. R. (1982). *Probability and Random Processes*. Oxford Univ. Press.

[9] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. North Holland–Kodansha, Amsterdam–Tokyo.

[10] ITÔ, K. and MCKEAN, H. P. (1965). *Diffusion Processes and Their Sample Paths*. Springer, Berlin.

[11] KARLIN, S. and TAYLOR, H. M. (1981). *A Second Course in Stochastic Processes*. Academic, New York.

[12] KHASMINSKII, R. Z. (1960). Ergodic properties of recurrent diffusion processes and stabilisation of the solution of the Cauchy problem for parabolic equations. *Theory Probab. Appl.* **5** 179–196.

[13] LINDVALL, T. (1983). On coupling of diffusion processes. *J. Appl. Probab.* **20** 82–93.

[14] PORT, S. C. and STONE, C. J. (1978). *Brownian Motion and Classical Potential Theory*. Academic, New York.

[15] RÖSLER, U. (1979). The tail σ-field of time-homogeneous one-dimensional diffusion processes. *Ann. Probab.* **7** 847–857.

[16] YAMADA, T. and WATANABE, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11** 155–167.

DEPARTMENT OF MATHEMATICS
CHALMERS INSTITUTE OF TECHNOLOGY
SVEN HULTINS GATA 6
S-412 96 GÖTEBORG
SWEDEN

DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE
UNIVERSITY COLLEGE
SINGLETON PARK
SWANSEA SA2 8PP
GREAT BRITAIN