Probab. Th. Rel. Fields 82, 187–211 (1989)

Probability Theory and Related Fields

Asymptotic Properties of Solutions of Multi-dimensional Stochastic Differential Equations

G. Kersting

Fachbereich Mathematik, J.W. Goethe-University, Postfach 111932, D-6000 Frankfurt, Federal Republic of Germany

Summary. Let $X_t \in \mathbb{R}^d$ be the solution of the stochastic equation $dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t$, where W_t denotes a standard Wiener process. The aim of the paper is to clarify under which conditions the drift term or the diffusion term is of negligible significance for the long term behaviour of X_t.

I. Introduction

In this paper we present several results which may be useful for judging the long term behaviour of diffusion processes. The object of this study is some $\mathbb{R}^d$-valued diffusion X_t, $t \geqq 0$, given by the stochastic equation

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t, \tag{1.1}$$

where W_t, $t \geqq 0$, is a standard Wiener process in $\mathbb{R}^n$. We assume that the vector $b(x)$ and the $d \times n$-matrix $\sigma(x)$ are uniformly Lipschitz continuous functions of $x \in \mathbb{R}^d$. Then, given X_0, (1.1) has a unique solution in the Ito sense, which does not explode in finite time. (See Durrett (1984) for the relevant facts about stochastic integration.)

In order to obtain information on the long term properties of X_t one might think of trying to compare X_t with other processes, which possibly are easier to analyze. An obvious candidate here is the deterministic process x_t, $t \geqq 0$, given by the ordinary differential equation

$$dx_t = b(x_t)\,dt. \tag{1.2}$$

It would be of great use to know not only, when X_t and x_t show the same kind of behaviour, but also, when the properties of both processes differ substantially from each other. Apparently the latter will occur, if the behaviour of X_t is close to that of the diffusion Z_t, given by the stochastic equation

$$dZ_t = \sigma(Z_t)\,dW_t. \tag{1.3}$$

Let us introduce the diffusion matrix

$$a(x)=\sigma(x)\,\sigma(x)^T.$$

It is natural to conjecture that X_t and x_t have similar properties, if there is little noise in the system in the sense that $a(x)$ is small (with respect to $b(x)$; recall that $a(x)$ measures the amount of random oscillation at point x). If on the other hand $a(x)$ is large, one would expect that X_t and Z_t are close to each other. Now the results of this paper suggest that these two domains adjoin each other. To be more precise let us introduce the extremal eigenvalues of $a(x)$

$$\lambda_{\max}(x)=\max_{|\xi|=1}\xi^T a(x)\,\xi,$$

$$\lambda_{\min}(x)=\min_{|\xi|=1}\xi^T a(x)\,\xi.$$

(ξ denotes a d-dimensional column vector.) We conjecture that X_t and x_t have similar properties, if the *low-noise condition*

$$\lambda_{\max}(x)=o(|x|\cdot|b(x)|), \tag{1.4}$$

as $|x|\to\infty$, holds, whereas under the *high-noise condition*

$$|x|\cdot|b(x)|=o(\lambda_{\min}(x)) \tag{1.5}$$

we expect that X_t and Z_t behave similarly. The aim of this paper is to provide some support for these conjectures.

Of course not every diffusion of interest belongs to the domain, covered by (1.4) and (1.5). Prominent examples are the solutions of the one-dimensional linear stochastic equation

$$dX_t=b\cdot X_t\,dt+\sigma\cdot X_t\,dW_t,$$

with $b,\sigma\in\mathbb{R}$. The higher dimensional versions of this equation are fairly difficult to analyze, both the drift and the diffusion component contribute significantly to the properties of the corresponding solutions (compare for instance Arnold et al. (1986)). In the light of our conjecture this is not difficult to understand: These linear models are located on the border of both domains defined by (1.4) and (1.5).

The paper is organized as follows: In the next section we prove some results for certain diffusion models, which belong to the low-noise class (1.4). It turns out that not only X_t and x_t possess similar properties but also Y_t and x_t, where Y_t denotes the Stratonovitch solution of (1.1). This is not surprising. If the noisy part is on the whole negligible, then it does not matter, how the noise is added to the system. In Sect. 3 we analyze models which satisfy (1.5). Here our strategy is as follows: We construct a bijection u: $\mathbb{R}^d\to\mathbb{R}^d$ such that the drift term of $u(X_t)$ vanishes. More precisely we like to achieve that

$$du(X_t)=\tau(X_t)\,dW_t \tag{1.6}$$

for $|X_t| \geqq C$ and a suitable $C > 0$, furthermore

$$u(x) \sim x \quad \text{and} \quad Du(x) \sim E,$$

as $|x| \to \infty$. Here Du denotes the Jacobian of u and E the identity matrix. From Ito's formula $\tau = Du \cdot \sigma$. Since by means of the chain rule

$$du(x_t) = Du(x_t) \cdot b(x_t)\, dt,$$

it is fairly obvious in view of (1.6) that in general $u(X_t)$ and $u(x_t)$ will behave completely different. Since $u(x) \sim x$, this translates to X_t and x_t. Our main conjecture is that just a slight strengthening of (1.5) is sufficient for the existence of such a mapping u. In Sect. 3, which contains the main results of this paper, this will be discussed in detail. In the appendix two auxiliary analytical lemmas are proven.

Notational Conventions. All vectors are column vectors. The transpose of a vector (or a matrix) ξ is written as ξ^T. x, y, z always denote elements of $\mathbb{R}^d$. $|\cdot|$ is the Euklidean norm, $\langle \cdot, \cdot \rangle$ the ordinary scalar product. Further we denote $\partial/\partial x_i$ and $\partial^2/\partial x_i \partial x_j$ by D_i and D_{ij}. – For convenience we sometimes suppress the index t in stochastic equations. Thus Eq. (1.1) is also written as

$$dX = b(X)\, dt + \sigma(X)\, dW,$$

or coordinatewise for every $1 \leqq i \leqq d$

$$dX_i = b_i(X)\, dt + \sum_{j=1}^{n} \sigma_{ij}(X)\, dW_j.$$

II. Some Results for the Low-noise Case

In this section we collect some results supporting our statement that in the low-loise case the properties of X_t are determined mainly by the drift component. For related results we refer the reader to Clark (1987), Cranston (1983), Keller et al. (1984) and Pinsky (1987).

A. Fix Points and Stationary Distributions

Let us suppose that

$$\langle b(x), x \rangle < 0 \quad \text{for all } x \neq 0.$$

Then 0 is a stable fix point of the vector field $(b(x))$. One might believe that the diffusion X_t should therefore have a limiting distribution. In general this is not the case, however, under the following 'low-noise type' condition this is true:

$$\operatorname{trace} a(x) \leqq -2\langle x, b(x) \rangle - \varepsilon$$

for some $\varepsilon>0$ and all $|x|$ large enough is sufficient for having a stationary limit. For the proof notice that

$$d|X|^2=(2\langle X, b(X)\rangle+\operatorname{trace} a(X))\,dt+2\sum_{i,j} X_i\,\sigma_{ij}(X)\,dW_j.$$

By assumption the drift term does not exceed $-\varepsilon$ for large $|X|$, which is enough to guarantee the existence of a stationary limit. This is all well-known, so we do not go further into it but refer the reader to the literature (compare Chapter III, Th. 7.1 and Chapter IV, Th. 4.1 in Has'minskii (1980)).

B. Convergence of the Angular Process

Here we discuss a situation, where $|x_t|\to\infty$ and $x_t/|x_t|$ converges. We shall show that in the low-noise case these statements remain valid for X_t. Let us introduce the decomposition

$$b(x)=b_p(x)+b_0(x)$$

of the drift vector, where $b_p(x)$ is the projection of $b(x)$ onto the subspace generated by the vector x, i.e.

$$b_p(x)=\langle x, b(x)\rangle\frac{x}{|x|^2}.$$

We assume

(A1) $\langle x, b(x)\rangle>0$ *for all* $|x|\geqq 1$,

(A2) *there is a non-negative, monotone decreasing function* $h(t)$ *such that* $\int_1^\infty h(t)\frac{dt}{t}<\infty$ *and, as* $|x|\to\infty$,

$$|b_0(x)|\leqq h(|x|)\,|b(x)| \qquad \textit{for } |x|\geqq 1.$$

In particular $|b_0(x)|=o(|b(x)|)$. It is easy to show that $|x_t|$ diverges and $x_t/|x_t|$ converges, if $|x_0|\geqq 1$. In fact from $dx_t=b(x_t)\,dt$ and the chain rule

$$\frac{d}{dt}|x_t|=\frac{\langle x_t, b(x_t)\rangle}{|x_t|}=|b_p(x_t)|\sim|b(x_t)|,$$

thus $|x_t|$ is strictly increasing and going to infinity. Further from (A2)

$$\left|\frac{d}{dt}\frac{x_t}{|x_t|}\right|=\frac{|b_0(x_t)|}{|x_t|}=O\left(\frac{h(|x_t|)}{|x_t|}\frac{d}{dt}|x_t|\right),$$

and from (A2) the convergence of $x_t/|x_t|$ follows. It turns out that these properties are preserved, if a small amount of noise in added. We assume the following:

(A 3) *As* $|x|\to\infty$,

$$\lambda_{\max}(x)=O(|b(x)|\cdot|x|\cdot h(|x|)),$$

where $h(t)$ *is as in* (A 2).

Since $h(t)=o(1)$, (A 3) is slightly stronger than the low-noise condition (1.4). For later use we mention that (A 3) implies

$$\operatorname{trace} a(x)=O(|b(x)|\cdot|x|\cdot h(|x|)).$$

Proposition 1. *Under* (A 1)–(A 3), *as* $t\to\infty$, *almost surely* $|X_t|\to\infty$ *and* $X_t/|X_t|$ *has a finite limit.*

Proof. (i) From (A 1) and (A 2) $\langle x, b(x)\rangle \sim |x|\cdot|b(x)|$ for $|x|\to\infty$. Now $|X_t|\to\infty$ follows from (A 3) and the well-known recurrence/transience criteria due to Has'minskii and Battacharya (compare Battacharya (1978)).

(ii) Define the function

$$H(t)=\int_t^\infty h(s)\frac{ds}{s}, \qquad t\geqq 0,$$

where $h(s)$ is as in (A 2). As shown in Lemma A 1 in the appendix we may assume without loss of generality that h is differentiable such that $h'(t)=O(t^{-1}h(t))$. From Ito's formula

$$dH(|X|)=-\frac{h(|X|)}{|X|^2}\left(\langle X, b(X)\rangle+\frac{1}{2}\operatorname{trace} a(X)-\left(1-\frac{|X|}{2}\frac{h'(|X|)}{h(|X|)}\right)\frac{X^T a(X)X}{|X|^2}\right)dt$$
$$-\frac{h(|X|)}{|X|^2}\sum_{i,j} X_i\,\sigma_{ij}(X)\,dW_j.$$

(If $H(|x|)$ is not smooth enough at $x=0$, modify it suitably in a neighbourhood of 0. This has no effect on our conclusions.) Since $\langle x, b(x)\rangle\sim|x|\cdot|b(x)|$ for large $|x|$, from (A 3) with a suitable $c>0$

$$H(|X_t|)\leqq c-\frac{1}{2}\int_0^t h(|X_s|)\cdot|b(X_s)|\frac{ds}{|X_s|}+S_t$$
$$\leqq c+S_t,$$

where S_t denotes a stochastic integral. Now with probability 1 either $\liminf S_t=-\infty$ or S_t has a finite limit. (For a proof recall that any stochastic integral is, up to a change of time, a Brownian motion.) The first alternative is excluded here in view of the above inequality, since $H(|X_t|)\to 0$ a.s., as $t\to\infty$. Thus, S_t has a.s. a finite limit, and, using the above inequality a second time, we conclude that

$$\int_0^\infty h(|X_t|)\cdot|b(X_t)|\frac{dt}{|X_t|}<\infty \quad \text{a.s.} \tag{2.1}$$

(iii) Let us denote $P_{ij}(x)=\delta_{ij}-\frac{x_i x_j}{|x|^2}$, where δ_{ij} is Kronecker's symbol. $P(x)=(P_{ij}(x))$ is the projection matrix onto the subspace of dimension $d-1$, which is perpendicular to the vector x. Some calculations, involving Ito's formula, yield

$$d\frac{X}{|X|}=\frac{b_0(X)}{|X|}dt+\frac{1}{|X|}\tau(X)\,dW + \frac{X}{2|X|^3}\left(3\frac{X^T a(X)X}{|X|^2}-\operatorname{trace} a(X)\right)dt+\frac{c(X)}{2|X|^2}dt. \tag{2.2}$$

Here $\tau(x)=P(x)\,\sigma(x)$, $c(x)$ is the vector with components

$$c_i(x)=e_i^T\,a(x)\,e_i+\frac{x^T a(x)x}{|x|^2}-\left(e_i+\frac{x}{|x|}\right)^T a(x)\left(e_i+\frac{x}{|x|}\right),$$

and e_i is the unit vector $(\delta_{1i},\ldots,\delta_{di})^T$. The i-th component of the stochastic integral in (2.2) has the quadratic variational process

$$\int_0^t \tilde{a}_{ii}(X_s)\frac{ds}{|X_s|^2},$$

where $\tilde{a}(x)=\tau(x)\,\tau(x)^T=P(x)\,a(x)\,P(x)^T$. Since $P(x)$ is a projector, the maximal eigenvalue of $\tilde{a}(x)$ is not bigger than that of $a(x)$, therefore, using (A3), $\tilde{a}_{ii}(x)=e_i^T\,\tilde{a}(x)\,e_i$ is of order $O(|b(x)|\cdot|x|\cdot h(|x|))$. In view of (2.1)

$$\int_0^\infty \tilde{a}_{ii}(X_s)\frac{ds}{|X_s|^2}<\infty \quad \text{a.s.}$$

which implies the a.s. convergence of the stochastic integrals in (2.2). Furthermore, using (A2) and (A3), the length of the drift vectors in (2.2) are of order $O(|b(x)|\cdot|x|\cdot h(|x|))$, and, using (2.1) again, we get the a.s. convergence of all integrals on the right hand side of (2.2). Therefore $X_t/|X_t|$ converges a.s. □

C. *Rate of Divergence*

We continue to study (X_t) under the assumptions of the last paragraph. At which rate does $|X_t|$ diverge? Let us assume additionally

(A4) *There is a positive function $f(t)$, $t>0$, such that for all $|x|\geqq 1$*

$$\langle x, b(x)\rangle \geqq |x|\,f(|x|)>0.$$

Define the real function u_t, $t\geqq 0$, by

$$du_t=f(u_t)\,dt, \qquad u_0=1.$$

From $dx_t = b(x_t)\,dt$ it follows $d|x_t| \geqq f(|x_t|)\,dt$, consequently, if $|x_0| \geqq 1$, $|x_t| \geqq u_t$ for all $t \geqq 0$. In the low-noise situation this essentially remains true for X_t, too.

Proposition 2. *Assume* (A1)–(A4) *and let* $f(t)/t$ *be ultimately decreasing. Then with probability* 1

$$\liminf_{t\to\infty} \frac{|X_t|}{u_t} > 0.$$

If $f(t) = o(t)$,

$$\liminf_{t\to\infty} \frac{|X_t|}{u_t} \geqq 1.$$

If we revers the inequality in (A4), the analog results for $\limsup |X_t|/u_t$ are valid.

Proof. (i) From Ito's formula

$$d \log |X| = \frac{1}{|X|^2} \langle X, b(X)\rangle\, dt + \frac{1}{|X|^2} \sum_{i,j} X_i\, \sigma_{ij}(X)\, dW_j + \frac{1}{2|X|^2} \left(\operatorname{trace} a(X) - 2\frac{X^T a(X) X}{|X|^2}\right) dt. \tag{2.3}$$

Now from (A3)

$$\int_0^t X_s^T\, a(X_s)\, X_s \frac{ds}{|X_s|^4} = O\left(\int_0^t h(|X_s|) \cdot |b(X_s)| \frac{ds}{|X_s|}\right).$$

(2.1) gives the convergence of the right hand side. The left hand side is the quadratic variational process of the stochastic integral in (2.3). This integral is therefore a.s. convergent, as $t \to \infty$. Similarly from (A3) and (2.1) the last integral on the right hand side of (2.3) converges a.s., as $t \to \infty$. Thus from (2.3)

$$\log |X_t| = \int_0^t \langle X_s, b(X_s)\rangle \frac{ds}{|X_s|^2} + M_t,$$

where M_t is a.s. convergent. Using (A4), for any $s \leqq t$

$$\log \frac{|X_t|}{u_t} - \log \frac{|X_s|}{u_s} \geqq \int_s^t \left(\frac{f(|X_s|)}{|X_s|} - \frac{f(u_s)}{u_s}\right) ds + M_t - M_s. \tag{2.4}$$

(ii) Assume now that $\liminf |X_t|/u_t = 0$. Then there are numbers $s_1 < t_1 < s_2 < \dots$ going to infinity, such that $|X_{t_n}|/u_{t_n} \leqq 1/2\,|X_{s_n}|/u_{s_n}$ and $|X_t|/u_t \leqq 1$ for $s_n \leqq t \leqq t_n$. Since $f(t)/t$ is decreasing, we deduce from (2.4)

$$\log \tfrac{1}{2} \geqq M_{t_n} - M_{s_n}.$$

Letting $n \to \infty$ we see that M_t cannot converge at infinity. Thus $\liminf |X_t|/u_t > 0$ almost surely, which is our first assertion. By the same kind of reasoning one can show that $\lim|X_t|/u_t$ exists almost surely on the event $\{\liminf|X_t|/u_t<1\}$. This is useful for the proof of the second assertion. Thus assume now that $f(t)=o(t)$. If $\lim_t|X_t|/u_t<(1-\varepsilon)$ for some $\varepsilon>0$, in view of Lemma A2 in the appendix

$$\int_0^\infty \left(\frac{f(|X_s|)}{|X_s|}-\frac{f(u_s)}{u_s}\right) ds=\infty.$$

Applying (2.4) with $s=0$, $t=\infty$, we get

$$\log(1-\varepsilon)-\log\frac{|X_0|}{u_0}\geqq \int_0^\infty \left(\frac{f(|X_s|)}{|X_s|}-\frac{f(u_s)}{u_s}\right) ds+M_\infty-M_0,$$

thus $M_t\to-\infty$ for $t\to\infty$. Consequently the event $\{\liminf|X_t|/u_t<1\}$ has zero probability and the second assertion is proven. □

D. On the Stratonovitch Solution

Here we show that in the low-noise case our results carry over to Stratonovitch solutions of the equations under consideration. We study processes (Y_t) which satisfy the equation

$$dY_t=b(Y_t)\,dt+\sigma(Y_t)\circ dW_t.$$

The $\circ$ indicates that we are now considering the symmetric differential. (For a discussion of the symmetric differential we refer the reader to the monograph of Ikeda/Watanabe (1981), p. 100.) We need an additional assumption

(A 5) *For all i, j, k, as* $|x|\to\infty$

$$(D_i\,\sigma_{jk}(x))^2=O\left(\frac{h(|x|)}{|x|}\,|b(x)|\right).$$

Note that, since $\sigma_{jk}(x)^2\leqq a_{jj}(x)=e_j^T\,a(x)\,e_j$, from (A 3)

$$\sigma_{jk}(x)^2=O(|b(x)|\cdot h(|x|)\cdot|x|), \tag{2.5}$$

in fact (2.5) and (A 3) are equivalent. Thus (A 3) and (A 5), though not equivalent, are related.

Proposition 3. *Under* (A 1)–(A 3) *and* (A 5), *as* $t\to\infty$, *almost surely* $|Y_t|\to\infty$ *and* $Y_t/|Y_t|$ *converges to a finite limit.*

Proof. This is much along the lines of the proof of Proposition 1.

(i) The equation for Y_t may be rewritten as the Ito equation

$$dY=b(Y)\,dt+\sigma(Y)\,dW+\tfrac{1}{2}d\sigma(Y)\,dW$$

more precisely

$$dY_i = b_i(Y)\,dt + \sum_j \sigma_{ij}(Y)\,dW_j + \tfrac{1}{2}\sum_{k,j} D_k\,\sigma_{ij}(Y)\,\sigma_{kj}(Y)\,dt.$$

In view of (A 5) and (2.5) the additional drift term is of order $O(h(|y|)\cdot|b(y)|)$, thus, since $h(t)=o(1)$, of smaller order than the drift vector $b(y)$. Consequently the transience criterion of Has'minskii and Battacharya applies as in the proof of Proposition 1, and it follows $|Y_t| \to \infty$ a.s.

(ii) Let $H(t)$ as in the proof of Proposition 1. For the symmetric differential the chain rule holds, therefore

$$\begin{aligned} dH(|Y|) &= -\frac{h(|Y|)}{|Y|^2}\Big(\langle Y, b(Y)\rangle\,dt + \sum_{i,j} Y_i\,\sigma_{ij}(Y)\circ dW_j\Big) \\ &= -\frac{h(|Y|)}{|Y|^2}\Big(\langle Y, b(Y)\rangle\,dt + \sum_{i,j} Y_i\,\sigma_{ij}(Y)\,dW_j\Big) \\ &\quad -\frac{1}{2}\sum_{i,j,k} D_k\left(\frac{h(|y|)}{|y|^2}\,y_i\,\sigma_{ij}(y)\right)(Y)\,\sigma_{kj}(Y)\,dt. \end{aligned} \tag{2.6}$$

Using $h'(t)=O(t^{-1}h(t))$, (A 5) and (2.5) it is not difficult to show that

$$D_k\left(\frac{h(|y|)}{|y|^2}\,y_i\,\sigma_{ij}(y)\right)\sigma_{kj}(y) = \mathrm{O}\left(\frac{h(|y|)^2}{|y|}\,|b(y)|\right),$$

while from (A 1) and (A 2)

$$\frac{h(|y|)}{|y|^2}\langle y, b(y)\rangle \sim \frac{h(|y|)}{|y|}\,|b(y)|.$$

Recalling $h(t)=o(1)$ we see that the additional drift term in (2.6) is small for large $|y|$ compared to the first term. Now a repetition of the arguments leading to (2.1) leads to

$$\int_0^\infty |b(Y_t)|\cdot h(|Y_t|)\,\frac{dt}{|Y_t|} < \infty \quad \text{a.s.}$$

(iii) Using the chain rule we obtain for the angular process

$$d\frac{Y}{|Y|} = \frac{b_0(Y)}{|Y|}\,dt + \frac{1}{|Y|}\,\tau(Y)\circ dW,$$

which for the i-th component translates into the Ito equation

$$d\frac{Y_i}{|Y|} = \frac{\langle e_i, b_0(Y)\rangle}{|Y|}\,dt + \frac{1}{|Y|}\sum_j \tau_{ij}(Y)\,dW_j + \frac{1}{2}\sum_{j,k}\left(D_k\frac{\tau_{ij}(y)}{|y|}\right)(Y)\,\tau_{kj}(Y)\,dt.$$

τ and e_i are defined as in the proof of Proposition 1. Now it is not difficult to check, using (A 5) and (2.5) that the last term on the right hand side is of order $h(|Y|)\cdot|b(Y)|\cdot|Y|^{-1}\,dt$. Therefore we may conclude as in the proof of Proposition 1 that all integrals on the right hand side of the above equation converge, and $Y_t/|Y_t|^{-1}$ converges a.s. Details are left to the reader. □

In much the same way the conclusion of Proposition 2 carries over to Y_t under the assumption of (A 1)–(A 5).

III. The High-noise Case

In this section we argue that the drift component has no essential effect on the behaviour of X_t, if there is enough noise in the system. Our approach consists in the construction of new coordinates deviating insignificantly from the original ones such that the system has no longer any drift in the new scale. Let us develop our ideas in the one-dimensional case, which from a mathematical point of view is pretty trivial.

A. The One-dimensional Case

Let us give an instance where the drift term has negligible influence. Consider the real-valued diffusion (Z_t) given by

$$dZ_t=\tau(Z_t)\,dW_z. \tag{3.1}$$

Let $g(z)$ be some smooth increasing function such that

$$g'(z)\to 1, \qquad \text{as } |z|\to\infty.$$

Then $X_t=g(Z_t)$ obeys the equation

$$dX=b(X)\,dt+\sigma(X)\,dW,$$

where

$$b(g(z))=\tfrac{1}{2}g''(z)\,\tau(z)^2,$$
$$\sigma(g(z))=g'(z)\,\tau(z).$$

Since $g(z)\sim z$ for large $|z|$, the long-term behaviour of X_t and Z_t are very much alike. On the other hand $b(x)$ depends highly on the local properties of $g(x)$ via its second derivative. Thus the drift term contains no relevant information about the behaviour of X_t. For instance, it makes little sense to try to compare X_t with the solution of the equation $dx_t=b(x_t)\,dt$, as we did in the low-noise case.

Which are the diffusions X_t which can be obtained from (3.1) by means of a transformation $g(z)$ as above? Note that, since $g'(z)\to 1$, typically $g''(z)=o(|z|^{-1})$, as $|z|\to\infty$. Then, letting $x=g(z)$,

$$b(x)=o(|z|^{-1}\tau^2(z))=o(|g(z)|^{-1}\sigma^2(x))=o(|x|^{-1}a(x)),$$

which means that we are just in the high-noise case. Not every diffusion belonging to the high-noise class can be obtained from a diffusion of form (3.1) in the described manner, however, we have the following result.

Proposition 3.1. *Suppose that* $\int_{-\infty}^{\infty} \frac{|b(x)|}{a(x)} dx < \infty$. *Then there is a monotone, differentiable function* $u(x)$ *such that*

i) $u'(x) \to 1$, *as* $|x| \to \infty$,

ii) *for* $Z_t = u(X_t)$

$$dZ_t = \tau(Z_t)\, dW_t, \qquad \text{if } |Z_t| \geqq 1,$$

where $\tau(u(x)) = u'(x)\,\sigma(x)$.

For the proof choose any function $u(x)$ such that

$$u'(x) = \exp\left(\int_x^{\infty} \frac{2b(y)}{a(y)} dy\right), \qquad \text{if } x \geqq 1,$$

$$u'(x) = \exp\left(-\int_{-\infty}^{x} \frac{2b(y)}{a(y)} dy\right), \qquad \text{if } x \leqq -1.$$

Remarks

1) We can work out our argument further by adding a time change to the scale change. Let

$$\rho_t = \int_0^t \frac{\sigma^2(Z_s)}{\tau^2(Z_s)} ds,$$

$$\tilde{Z}_t = Z_{\rho_t},$$

where Z_t and $\tau(x)$ are as in the above proposition. Then

$$d\tilde{Z}_t = \sigma(\tilde{Z}_t)\, dB_t, \qquad \text{if } |\tilde{Z}_t| \geqq 1,$$

where B_t is a suitable standard Wiener process. If Z_t has a stationary limiting distribution then by the ergodic theorem almost surely

$$\rho_t / t \to \int_{-\infty}^{\infty} \frac{\sigma^2(x)}{\tau^2(x)} \psi(x)\, dx,$$

where ψ denotes the stationary density. To treat the remaining case let us assume that

$$\frac{\sigma(x)}{\sigma(y)} \to 1, \quad \text{if } |x|, |y| \to \infty \quad \text{such that } \frac{|x|}{|y|} \to 1.$$

From the definition of $\tau(x)$ we obtain $\sigma(x)/\tau(x)\to 1$, as $|x|\to\infty$. Furthermore almost surely

$$\int_0^t \frac{\sigma^2(Z_s)}{\tau^2(Z_s)}\chi_{\{|Z_s|\leq\gamma\}}\,ds=o(t)$$

for any $\gamma>0$, since we are in the case where Z_t has no limiting distribution. It follows $\rho_t\sim t$ a.s. Such a time change obviously has a negligible influence on the behaviour of Z_t. Altogether, by minor changes of time and space we have achieved that the drift coefficient vanishes outside of a compact set, while the diffusion term remains unchanged.

2) It is well-known that for any one-dimensional diffusion X_t there is a monotone function $u(x)$ such that $u(X_t)$ has no drift component. Just define

$$u'(x)=c\exp\left(-\int_0^x \frac{2b(y)}{a(y)}\,dy\right).$$

Of course $u(x)$ and x, $u'(x)$ and 1 may be far apart from each other so that our argumentation is not valid in general. – Note that $u'(x)\to 1$, as $x\to\pm\infty$, can only be achieved, if $\int_{-\infty}^{\infty}(b(y)/a(y))\,dy=0$. Only in this case we can attain that the stochastic equation formulated in Proposition 3.1 holds on the whole real line.

3) For the (Stratonovitch) solution Y_t of the equation

$$dY=b(Y)\,dt+\sigma(Y)\circ dW$$

Proposition 3.1 is not valid. Introduce the function

$$H(x)=\int_0^x \frac{dy}{\sigma(y)}$$

and assume $H(\infty)=\infty$, $H(-\infty)=-\infty$. By means of the chain rule

$$dH(Y)=\frac{b(Y)}{\sigma(Y)}\,dt+dW,$$

or, letting $U_t=H(Y_t)$, $c(H(x))=\dfrac{b(x)}{\sigma(x)}$,

$$dU=c(U)\,dt+dW.$$

For this equation the Stratonovitch- and Ito-equation coincide. By a substitution

$$\int_{-\infty}^{\infty}|c(x)|\,dx=\int_{-\infty}^{\infty}\frac{|b(x)|}{a(x)}\,dx.$$

Thus, if $\int |b(x)|\, a(x)^{-1}\, dx < \infty$, we may apply the conclusion of Proposition 3.1 to $U_t = H(Y_t)$. Again the drift component in the equation for U_t is negligible and, since $H(x)$ depends only on $\sigma(x)$, this is true for the drift term in the above Stratonovitch equation, too. The conclusion is similar as for the Ito equation, however, unlike circumstances in the low-noise case, the Stratonovitch and Ito solutions in general behave differently.

B. The Higher-dimensional Case

We extend the type of result given in Proposition 3.1. Our main additional assumption is the following uniform ellipticity property of the diffusion matrix:

(3.2) $$\lambda = \limsup_{|x| \to \infty} \lambda_{\max}(x)/\lambda_{\min}(x) < \infty.$$

We are looking for a bijection $u = u(x)$ of the $\mathbb{R}^d$ onto itself such that the diffusion $u(X_t)$ has no drift term outside some compact set. Let

$$L = \tfrac{1}{2} \sum_{i,j} a_{ij}(x)\, D_{ij} + \sum_i b_i(x)\, D_i$$

and denote by $Du(x)$ the Jacobian of u at point x. Then from Ito's formula

$$du(X) = Lu(X)\, dt + Du(X) \cdot \sigma(X)\, dW,$$

(the dot denotes matrix multiplication). We are thus loocking for bijections $u: \mathbb{R}^d \to \mathbb{R}^d$ such that

(S) (i) $|u(x) - x| = o(|x|)$ *and* $|Du(x) - E| = o(1)$, *as* $|x| \to \infty$, (*E denotes the identity matrix*),

(ii) *There is a* $C > 0$ *such that*

$$Lu(x) = 0, \quad \text{if } |x| \geqq C.$$

We conjecture that in the high-noise case such a transformation of the coordinates can mostly be found.

Conjecture. If there is an $\varepsilon > 0$ *such that*

$$|x|^{1+\varepsilon} |b(x)| = O(\lambda_{\min}(x))$$

(*and if* $a(x)$ *and* $b(x)$ *are sufficiently smooth*), *then there exists a bijection* u *such that* (S) *is satisfied.*

In this section we develop two results in this direction. Note that, if we multiply both $a(x)$ and $b(x)$ with some positive scalar $c(x)$, then the high-noise condition formulated in our conjecture is not affected. Also the problem (S) is not changed since the operator L just turns into $c(x)L$. (Probabilistically this corresponds to a random time change of X_t.) In particular, choosing for instance $1/c(x)=\lambda_{\min}(x)$ and recalling (3.2) we see that it is no loss of generality to require strong ellipticity of L:

(3.3) *There are numbers* $0<c_1<c_2<\infty$ *such that for all* x

$$c_1 \leqq \lambda_{\min}(x) \leqq \lambda_{\max}(x) \leqq c_2.$$

We need a certain amount of smoothness of the coefficients of L.

(3.4) *As* $r\to\infty$ $\sup\limits_{|y|,|z|\geqq r} \dfrac{|a_{ij}(y)-a_{ij}(z)|+|b_i(y)-b_i(z)|}{|y-z|}=o(1).$

Theorem 3.2. *Let* (3.3) *and* (3.4) *be satisfied. Let*

$$\alpha=2, \qquad \text{if } \lambda<d-1,$$
$$\alpha=1+\lambda(d-1), \qquad \text{if } \lambda\geqq d-1.$$

If for some $\varepsilon>0$,

(3.5) $$|b(x)|=O(|x|^{-\alpha-\varepsilon}),$$

as $|x|\to\infty$, *then there is a bijection u such that* (S) *holds.*

In this theorem always $\alpha\geqq 1$. Since $\lambda\geqq 1$, always $\alpha\geqq 2$, if the dimension $d\geqq 2$. The next result shows that it is not too much to hope that the conclusion of the theorem above remains valid if (3.5) holds just for $\alpha=1$.

Theorem 3.3. *Let* (3.3) *be satisfied. Assume that there is a positive definite matrix* $(a_{ij}(\infty))$ *and some* $\varepsilon>0$ *such that*

(3.6) $$|b(x)|=O(|x|^{-1-\varepsilon}),$$

(3.7) $$|a_{ij}(x)-a_{ij}(\infty)|=O(|x|^{-\varepsilon}),$$

as $|x|\to\infty$. *Also assume that, as* $r\to\infty$,

(3.8) $$\sup_{|y|,|z|\geqq r}\frac{|b_i(y)-b_i(z)|}{|y-z|}=O(r^{-2-\varepsilon}),$$

(3.9) $$\sup_{|y|,|z|\geqq r}\frac{|a_{ij}(y)-a_{ij}(z)|}{|y-z|}=O(r^{-1-\varepsilon}).$$

Then there is a bijection u such that (S) *holds.*

In the sequel we always assume $d\geqq 2$, the case $d=1$ being treated in Propositions 3.1. In the proof of these theorems we are mainly concerned with the function $v(x)=u(x)-x$. Note that (S) translates into

(S′) (i) $|v(x)|=o(|x|)$ *and* $|Dv(x)|=o(1)$, *as* $|x|\to\infty$.

(ii) $Lv(x)=-b(x)$, *if* $|x|\geqq C>0$.

The following lemmas deal with the solvability of this kind of Dirichlet-problem. We start with an auxiliary result.

Lemma 3.4. *Let $0<C<s$ and define*

$$\tau_s=\inf\{t\colon |X_t|=C \text{ or } |X_t|=s\}.$$

Under the assumptions of Theorem 3.2, if C is large enough, there is a $D>0$ such that for all $s>C$, $C\leqq|x|\leqq s$

$$E_x \int_0^{\tau_s} |X_t|^{-\alpha-\varepsilon} dt \leqq D,$$

where α is as in Theorem 3.2.

Proof. For every $s>C$ let $f(t)=f_s(t)$, $t\geqq C$, be the function given by

$$f(s)=f(C)=0 \tag{3.10}$$

and

$$f'(t)=\begin{cases} t^{1-\alpha-\varepsilon}\left|1-\left(\dfrac{a}{t}\right)^{\lambda(d-1)+1-\alpha-\varepsilon/2}\right|, & \text{if } C\leqq t\leqq a,\\ -t^{1-\alpha-\varepsilon}\left|1-\left(\dfrac{a}{t}\right)^{(d-1)/\lambda+1-\alpha-3\varepsilon/2}\right|, & \text{if } a\leqq t,\end{cases}$$

where $C<a<s$ is uniquely determined by (3.10). Note that $f(t)$ attains its maximum in a. Now

$$Lf(|x|)=\frac{f''(|x|)}{2|x|^2}x^T a(x)x+\frac{f'(|x|)}{2|x|}(\operatorname{trace} a(x)-|x|^{-2}x^T a(x)x+2\langle x, b(x)\rangle).$$

We estimate this function from above. To this end note that

$$\operatorname{trace} a(x)\leqq(d-1)\,\lambda_{\max}(x)+|x|^{-2}x^T a(x)x,$$

since $\lambda_{\min}(x)\leqq|x|^{-2}x^T a(x)x$. In view of (3.3) $|x|^{-2}x^T a(x)x\geqq c_1$, thus from (3.2) and (3.5), if C is large enough

$$\operatorname{trace} a(x)-|x|^{-2}x^T a(x)x+2\langle x, b(x)\rangle\leqq\left(\lambda(d-1)+\frac{\varepsilon}{2}\right)|x|^{-2}x^T a(x)x,$$

$$\text{if } |x|\geqq C.$$

If $C\leqq|x|\leqq a$, this entails

$$\begin{aligned} Lf(|x|)&\leqq\frac{1}{2}|x|^{-2}x^T a(x)x\left(f''(|x|)+(\lambda(d-1)+\varepsilon/2)\frac{f'(|x|)}{|x|}\right)\\ &\leqq-\frac{1}{2}c_1|\lambda(d-1)+1-\alpha-\varepsilon/2|\cdot|x|^{-\alpha-\varepsilon}.\end{aligned}$$

Similarly check that for $|x|\geqq a$

$$\operatorname{trace} a(x)-|x|^{-2}x^T a(x)x+2\langle x, b(x)\rangle\geqq((d-1)/\lambda-\varepsilon/2)|x|^{-2}x^T a(x)x,$$

therefore we obtain for $|x| \geqq a$ the estimate

$$Lf(|x|) \leqq \frac{1}{2}|x|^{-2} x^T a(x) x \left(f''(|x|) + ((d-1)/\lambda - \varepsilon/2) \frac{f'(|x|)}{|x|} \right)$$
$$\leqq -\frac{1}{2} c_1 |(d-1)/\lambda + 1 - \alpha - 3\varepsilon/2| \cdot |x|^{-\alpha-\varepsilon}.$$

Altogether, if $\varepsilon > 0$ is so small that the factors on the right hand side in both estimates of $Lf(|x|)$ do not vanish there is a $c > 0$ such that

$$|x|^{-\alpha-\varepsilon} \leqq -cLf(|x|), \quad \text{if } |x| \geqq C.$$

By means of Ito's formula

$$df(|X|) = Lf(|X|)\,dt + d(\text{martingale}),$$

and from the optional sampling theorem and (3.10), if $C \leqq |x| \leqq s$,

$$E_x \int_0^{\tau_s} |X_t|^{-\alpha-\varepsilon} dt \leqq -cE_x \int_0^{\tau_s} Lf(|X_t|)\,dt$$
$$= cE_x(f(|X_0|) - f(|X_{\tau_s}|))$$
$$= cf(|x|) \leqq cf(a).$$

For the proof of our assertion it remains to show that $f(a)$ is bounded uniformly in s. If now $\alpha = 1 + \lambda(d-1)$, we obtain by means of a partial integration and (3.10)

$$f(a) = \int_C^a f'(t)\,dt = \int_C^a t^{-\lambda(d-1)-\varepsilon/2}(t^{-\varepsilon/2} - a^{-\varepsilon/2})\,dt$$
$$= \frac{\varepsilon/2}{\lambda(d-1) - 1 + \varepsilon/2} \int_C^a (C^{-\lambda(d-1)-\varepsilon/2+1} - t^{-\lambda(d-1)-\varepsilon/2+1})\, t^{-1-\varepsilon/2}\,dt$$
$$\leqq \frac{1}{\lambda(d-1) - 1 + \varepsilon/2} C^{-\lambda(d-1)-\varepsilon/2+1} \cdot C^{-\varepsilon/2}.$$

If on the other hand $\lambda < d-1$, $\alpha = 2$ and $\varepsilon > 0$ is so small that $(d-1)/\lambda + 1 - \alpha - 3\varepsilon/2 > 0$, then in view of (3.10)

$$f(a) = -\int_a^s f'(t)\,dt \leqq \int_a^s t^{1-\alpha-\varepsilon} dt \leqq \int_C^\infty t^{-1-\varepsilon} dt,$$

and the lemma is proven. □

Lemma 3.5. *Let $g: \mathbb{R}^d \to \mathbb{R}$ be such that for some $\varepsilon > 0$*

$$g(x) = O(|x|^{-\alpha-\varepsilon}),$$

as $|x| \to \infty$, *where* α *is as in Theorem* 3.2, *and*

$$\sup_{|y|,|z| \geqq r} \frac{|g(y)-g(z)|}{|y-z|} = o(1),$$

as $r \to \infty$. *Then, under the assumptions of Theorem* 3.2, *if* $C>0$ *is large enough, there is a function* φ *such that*

i) $L\varphi(x) = -g(x)$, *if* $|x| \geqq C$,
ii) $|\varphi(x)| \leqq C$, *if* $|x| \geqq C$,
iii) *for all* $1 \leqq i \leqq d$ $|D_i \varphi(x)| = o(1)$, *as* $|x| \to \infty$.

Proof. Without loss of generality let $g(x) \geqq 0$ for all x. Let $0<C<s$ and τ_s as in Lemma 3.4. As is well-known

$$\varphi_s(x) = E_x \int_0^{\tau_s} g(X_t)\, dt$$

solves the equation

$$L\varphi_s(x) = -g(x) \tag{3.11}$$

on the domain $C<|x|<s$. In view of (3.5) and Lemma 3.4 $\varphi_s(x)$ is uniformly bounded in $s>C$ and $C \leqq |x| \leqq s$. Furthermore, since $g(x) \geqq 0$, $\varphi_s(x)$ is increasing in s. Thus, letting $s \to \infty$, $\varphi_s(x) \uparrow \varphi(x)$, where $\varphi(x)$ is a bounded function on the domain $|x| \geqq C$. It is well-known that φ solves (3.11), too. For the reader's convenience we give a short probabilistic proof. Choose x such that $|x|>C$ and define $A_x = \{y\colon |y-x|<|x|-C\}$. Let $\bar{\varphi}$ be the unique continuous function on $A_x \cup \partial A_x$ such that

$$L\bar{\varphi}(y) = -g(y) \quad \text{for all } y \in A_x$$

and $\bar{\varphi} \equiv \varphi$ on ∂A_x. Then $L(\varphi_s - \bar{\varphi}) = 0$ on A_x. Letting $\rho = \inf\{t\colon X_t \in \partial A_x\}$, for any $y \in A_x$

$$\varphi_s(y) - \bar{\varphi}(y) = E_y(\varphi_s(X_\rho) - \varphi(X_\rho)).$$

If $s \to \infty$, the right hand side converges to zero, thus $\varphi \equiv \bar{\varphi}$ on A_x and assertion i) of the lemma follows, whereas ii) is valid by construction. In order to prove iii) we utilize Schauder's a priori bounds for partial differential equations. Let $0<r(t) \leqq t/2$ be some function going to infinity and let B_x be the ball of radius $r(|x|)$ around x. Using the notation of Theorem 6.2, p. 85 Gilbarg/Trudinger (1977), in view of (3.4)

$$|a_{ij}|^{(0)}_{0,1;B_x} \leqq r(|x|) \sup_{y,z \in B_x} \frac{|a_{ij}(y)-a_{ij}(z)|}{|y-z|} = o(1),$$

as $|x| \to \infty$, if only $r(t)$ goes to infinity slowly enough. Similarly

$$|b_i|^{(1)}_{0,1;B_x} \leqq r(|x|)^2 \sup_{y,z \in B_x} \frac{|b_i(y)-b_i(z)|}{|y-z|} = o(1).$$

Therefore the a priori bound given in the cited theorem (with $\alpha=1$, $\Omega=B_x$) holds uniformly over all balls B_x such that $|x|\geqq 1$. In particular there is a $c>0$, such that for $|x|\geqq 1$

$$r(|x|)|D_i\,\varphi(x)|\leqq c\left(\sup_{y\in B_x}|\varphi(y)|+r(|x|)^2\sup_{y\in B_x}|g(y)|+r(|x|)^3\sup_{y,z\in B_x}\frac{|g(y)-g(z)|}{|y-z|}\right).$$

By the assumptions of the lemma and since φ is bounded, the right hand side is a bounded function of x, if $r(t)$ is going to infinity slowly enough, and assertion iii) is proven. □

Next we treat the case of the Laplacian.

Lemma 3.6. *Let $g\colon \mathbb{R}^d\to\mathbb{R}$ be Lipschitz such that for some $0<\varepsilon<1$, as $|x|\to\infty$,*

$$g(x)=O(|x|^{-1-\varepsilon}) \tag{3.12}$$

and, as $r\to\infty$,

$$\sup_{|y|,|z|\geqq r}\frac{|g(y)-g(z)|}{|y-z|}=O(r^{-2-\varepsilon}). \tag{3.13}$$

Let w_d denote the volume of the unit ball in $\mathbb{R}^d$ and define for $x, y\in\mathbb{R}^d$

$$\Gamma(x,y)=\begin{cases}-(4\pi)^{-1}(\log|y-x|-\log|y|), & \text{if } d=2,\\ (2d(d-1)\,w_d)^{-1}(|y-x|^{2-d}-|y|^{2-d}), & \text{if } d\geqq 3.\end{cases}$$

Then

$$\varphi(x)=\int_{\mathbb{R}^d}\Gamma(x,y)\,g(y)\,dy$$

is well-defined. Furthermore

i) $\frac{1}{2}\Delta\varphi=-g$,
ii) $\varphi(x)=O(|x|^{1-\varepsilon})$,
iii) $D_i\,\varphi(x)=O(|x|^{-\varepsilon})$,
iv) $D_{ij}\,\varphi(x)=O(|x|^{-1-\varepsilon})$,
v) $\displaystyle\sup_{|y|,|z|\geqq r}\frac{|D_{ij}\,\varphi(y)-D_{ij}\,\varphi(z)|}{|y-z|}=O(r^{-2-\varepsilon})$.

Proof. Let $B_x=\{y\in\mathbb{R}^d\colon |y|<\frac{1}{2}|x|\}$, $C_x=\{y\colon |y-x|<\frac{1}{2}|x|\}$. If $y\notin B_x\cup C_x$, some simple geometric considerations yield $|y|/3\leqq|y-x|\leqq 3|y|$. By the mean value theorem

$$||y-x|^{2-d}-|y|^{2-d}|\leqq(d-2)\,3^{d-1}|x||y|^{1-d}$$

and

$$|\log|y-x|-\log|y||\leqq 3|x||y|.$$

Thus for suitable $c_1, c_2 > 0$, by means of (3.12)

$$\int_{\mathbb{R}^d - B_x \cup C_x} |\Gamma(x,y)\, g(y)|\, dy \leqq c_1 |x| \int_{\mathbb{R}^d - B_x} |y|^{-d-\varepsilon} dy = c_2 |x| \int_{|x|/2}^{\infty} t^{-1-\varepsilon} dt$$
$$= O(|x|^{1-\varepsilon}).$$

Next, since $0 < \varepsilon < 1$, by means of (3.12)

$$\int_{B_x} |(|y-x|^{2-d} - |y|^{2-d})\, g(y)|\, dy \leqq c_1 \int_{B_x} |y|^{2-d} |g(y)|\, dy \leqq c_2 \int_0^{|x|} t^{-\varepsilon} dt = O(|x|^{1-\varepsilon}).$$

In the case $d=2$, since $|y| \leqq |y-x| \leqq 2|x|$ for $y \in B_x$,

$$\int_{B_x} |(\log|y-x| - \log|y|)\, g(y)|\, dy = \int_{B_x} \log \frac{|y-x|}{|y|} |g(y)|\, dy$$
$$\leqq c_1 \int_{B_x} |y|^{-1-\varepsilon} \log \frac{2|x|}{|y|}\, dy \leqq c_2 \int_0^{|x|} t^{-\varepsilon} \log \frac{|x|}{t}\, dt = O(|x|^{1-\varepsilon}).$$

Similarly one estimates $\int_{C_x} |\Gamma(x,y)\, g(y)|\, dy$. Altogether

$$\int |\Gamma(x,y)\, g(y)|\, dy < \infty,$$

thus $\varphi(x)$ is well-defined. Also ii) follows from our estimates. Next define

$$\varphi_s(x) = \int_{B_x} \Gamma(x,y)\, g(y)\, dy.$$

If $d \geqq 3$

$$\varphi_s(x) = (2d(d-1)\, w_d)^{-1} \int_{B_s} |y-x|^{2-d} g(y)\, dy + D_s,$$

where D_s is a constant depending only on s. By classical results from potential theory (Lemma 4.2 in Gilbarg/Trudinger (1977)) $\Delta \varphi_s(x) = -g(x)$ for all $x \in B_s$, further from the estimates above

$$|\varphi_s(x) - \varphi(x)| \leqq \int_{\mathbb{R}^d - B_s} |\Gamma(x,y)\, g(y)|\, dy \leqq c|x| \int_{s/2}^{\infty} t^{-1-\varepsilon} dt.$$

Letting $s \to \infty$, $\varphi_s \to \varphi$ pointwise, uniformly on compact sets. A theorem of Harnack (Theorem 2.9 in Gilbarg/Trudinger (1977)) implies that $\Delta \varphi = -g$, and i) follows. The case $d=2$ is treated similarly. To prove iii)–v) we use again Theorem 6.2 of Gilbarg/Trudinger (1977) (with $\Omega = C_x$ and $\alpha = 1$). In particular we obtain from (3.12) and (3.13)

$$|x|\,|D_i\,\varphi(x)| + |x|^2\,|D_{ij}\,\varphi(x)| + |x|^3 \sup_{|y-x|,|z-x| \leqq |x|/4} \frac{|D_{ij}\,\varphi(y) - D_{ij}\,\varphi(z)|}{|y-z|}$$

$$\leqq c\left(\sup_{y\in C_x} |\varphi(y)| + |x|^2 \sup_{y\in C_x} |g(y)| + |x|^3 \sup_{y,z\in C_x} \frac{|g(y)-g(z)|}{|y-z|}\right) = O(|x|^{1-\varepsilon}),$$

and iii)–v) follows. □

Lemma 3.7. *Let* (3.6)–(3.9) *be satisfied for some* $\varepsilon > 0$, *also* (3.12) *and* (3.13), *where* $g(x)$ *is as in Lemma* 3.6. *Then there is a* $C > 0$ *and a function* φ *such that*

i) $L\varphi(x) = -g(x)$, *if* $|x| \geqq C$,
ii) $\varphi(x) = o(|x|)$,
iii) $D_i\,\varphi(x) = o(1)$.

Proof. Without loss let $0 < \varepsilon < 1$. Applying a linear transformation to $\mathbb{R}^d$, we may assume that $a_{ij}(\infty) = \delta_{ij}$ in (3.7) (δ_{ij} is Kronecker's symbol). Let φ_1 be the solution of the equation $\frac{1}{2}\Delta\,\varphi_1 = -g$, as given by Lemma 3.6 and let

$$g_1(x) = \tfrac{1}{2}\sum_{i,j} (a_{ij}(x) - \delta_{ij})\, D_{ij}\,\varphi_1(x) + \sum_i b_i(x)\, D_i\,\varphi_1(x).$$

In view of Lemma 3.6 iii)–v) and (3.6)–(3.9)

$$g_1(x) = O(|x|^{-1-2\varepsilon}),$$

$$\sup_{|y|,|z| \geqq r} \frac{|g_1(y) - g_1(z)|}{|y-z|} = O(r^{-2-2\varepsilon}),$$

as well as

$$\varphi_1(x) = o(|x|), \qquad D_i\,\varphi_1(x) = o(1).$$

Iterating this procedure we define φ_k and g_k such that

$$\tfrac{1}{2}\Delta\,\varphi_k = -g_{k-1},$$

$$g_k = (L - \tfrac{1}{2}\Delta)\,\varphi_k,$$

with $g_0 = g$, thus

$$L\varphi_k = g_k - g_{k-1},$$

$$g_k(x) = O(|x|^{-1-(k+1)\varepsilon}),$$

$$\sup_{|y|,|z| \geqq r} \frac{|g_k(y) - g_k(z)|}{|y-z|} = O(r^{-2-(k+1)\varepsilon}),$$

$$\varphi_k(x) = o(|x|), \qquad D_i\,\varphi_k(x) = o(1).$$

In view of Lemma 3.6 we can continue this construction as long as $1 \leqq k \leqq n$, where n is such that $n\varepsilon < 1 < (n+1)\varepsilon$. (By choosing an irrational ε we exclude the possibility $n\varepsilon = 1$.) In particular

$$g_n(x) = O(|x|^{-2-\delta})$$

where $\delta = (n+1)\varepsilon - 1 > 0$. Now note that, since $a_{ij}(\infty) = \delta_{ij}$, λ, as given in (3.2), is equal to 1. Applying Lemma 3.5 with $\alpha = 2$ we see that there is a function φ_{n+1} such that

$$L\varphi_{n+1}(x) = -g_n(x) \quad \text{for } |x| \geqq C > 0,$$
$$|\varphi_{n+1}(x)| \leqq C, \quad D_i\,\varphi_{n+1}(x) = o(1).$$

Now define $\varphi = \varphi_1 + \ldots + \varphi_{n+1}$. Then for $|x| \geqq C$

$$L\varphi(x) = \sum_{k=1}^{n} (g_k - g_{k-1})(x) - g_n(x) = -g(x).$$

By construction $\varphi(x) = o(|x|)$ and $D_i\,\varphi(x) = o(1)$. □

Lemma 3.8. *Let $\varphi(x)$, $|x| \geqq 1$, be a realvalued C^2-function such that*

$$\varphi(x) = o(|x|), \quad D_i\,\varphi(x) = o(1),$$

as $|x| \to \infty$, and let $\delta > 0$. Then there exists a $c \geqq 1$ and a C^2-function $\psi\colon \mathbb{R}^d \to \mathbb{R}$ such that

$$\psi(x) = \varphi(x), \quad \text{if } |x| \geqq c,$$
$$|D_i\,\psi(x)| \leqq \delta, \quad \text{for all } x.$$

Proof. We construct ψ by mollifying φ. Define $\varphi(x) = 0$ for all $|x| < 1$. Choose a C^2-function $M(x)$ such that $M(x) = 0$ for all $|x| \geqq 1/4$ and

$$\int M(x)\,dx = 1.$$

Let $\rho\colon [0, 1] \to [0, 1]$ be a C^2-function such that

$$\rho(t) = 1 \quad \text{for all } 0 \leqq t \leqq 1/2,$$
$$0 < \rho(t) \leqq 1 \quad \text{for all } 1/2 \leqq t < 1,$$
$$\rho(1) = \rho'(1) = \rho''(1) = 0.$$

Let $c > 4$. Now define

$$\psi(x) = \int \varphi(x - c\,y\,\rho(|x|/c))\,M(y)\,dy$$
$$= \rho(|x|/c)^{-d} \int \varphi(cz)\,M\left(\frac{x/c - z}{\rho(|x|/c)}\right) dz, \quad \text{if } |x| < c,$$
$$\psi(x) = \varphi(x), \quad \text{if } |x| \geqq c.$$

If now $x_v \to x$, $|x_v| < c$, $|x| = c$, then

$$\psi(x_v) \to \int \varphi(x)\, M(y)\, dy = \varphi(x),$$

since $\rho(1) = 0$. Consequently ψ is continuous everywhere. From the second integral representation it is clear that $\psi(x)$ is a C^2-function in the domain $|x| < c$. Let us estimate the partial derivatives. First we consider the case $|x| \leqq c/2$. Then $\rho(|x|/c) = 1$, therefore from the second integral representation

$$D_i\, \psi(x) = \int \frac{1}{c}\, \varphi(c z)\, D_i\, M(x/c - z)\, dz.$$

Taking absolute values we obtain

$$|D_i\, \psi(x)| \leqq \sup_{|y| \leqq c} \frac{1}{c} |\varphi(y)| \sup_y |D_i\, M(y)| \int_{|z| \leqq 1/4} dz.$$

Since by assumption $\varphi(x) = o(|x|)$,

$$|D_i\, \psi(x)| \leqq \delta \qquad \text{for all } |x| \leqq c/2,$$

if only c is large enough.

If $c/2 \leqq |x| < c$, we utilize the first integral representation for ψ. Since $|y| \leqq 1/4$ entails $|x - c\rho(|x|/c)\, y| \geqq c/4 > 1$, we obtain

$$D_i\, \psi(x) = \sum_k \int M(y)\, D_k\, \varphi(x - c\, y\, \rho(|x|/c)) \left(\delta_{ik} - \rho'(|x|/c) \frac{x_i}{|x|}\, y_k\right) dy,$$

consequently

$$\sup_{c/2 \leqq |x| < c} |D_i\, \psi(x)| \leqq d \sup_{\substack{1 \leqq k \leqq d \\ |y| \geqq c/4}} |D_k\, \varphi(y)| \sup_{0 \leqq t \leqq 1} (1 + |\rho'(t)|) \int M(y)\, dy \leqq \delta,$$

if c is large enough, since $D_k\, \varphi(y) = o(1)$ by assumption. Also, if x_v is a sequence such that $x_v \to x$, $|x_v| < c$, $|x| = c$, then, since $\rho(1) = \rho'(1) = 0$

$$D_i\, \psi(x_v) \to \sum_k \int M(y)\, D_k\, \varphi(x)\, \delta_{ik}\, dy = D_i\, \varphi(x).$$

Thus $D_i \psi$ is continuous in the whole Euclidean space. Also it is clear by construction that $\sup |D_i \psi| \leqq \delta$, if only c was chosen large enough. It remains to show that $D_{ij}\, \psi(x_v) \to D_{ij}\, \varphi(x)$, where $x_v \to x$ as above. The proof is along the same lines, differentiating the first integral representation of ψ and using $\rho(1) = \rho'(1) = \rho''(1) = 0$. We leave the details to the reader. It follows that $D_{ij} \psi$ is continuous everywhere, thus ψ is a C^2-function. □

Proof of Theorem 3.2 and 3.3. In view of Lemma 3.5, 3.7 and 3.8 there is for any $\delta>0$ a $C>0$ and a C^2-mapping $v=(v_1, \ldots, v_d)$: $\mathbb{R}^d \to \mathbb{R}^d$ such that

$$\begin{aligned} Lv(x) &= -b(x), \quad \text{if } |x| \geqq C, \\ |v(x)| &= o(|x|), \\ D_i v_j(x) &= o(1), \\ |D_i v_j(x)| &\leqq \delta \quad \text{for all } x. \end{aligned}$$

Define $u(x)=x+v(x)$. Then $u(x)$ satisfies (S). It remains to show that u is a bijection. From the mean value theorem, if $\delta>0$ is small enough

$$|v(x)-v(y)| \leqq \sum_{i,j} \sup_z |D_i v_j(z)| |x-y| \leqq d^2 \delta |x-y| < |x-y|.$$

Consequently $u(x) \neq u(y)$ for all $x \neq y$ such that u is injectiv. Next choose $\delta>0$ so small that $\det(Du(x)) = \det(E+Dv(x) \neq 0$ for all x. Then u is a regular function at any point x, consequently u is an open mapping. In particular, range(u) is open. Next let $y \in \overline{\text{range}(u)}$, $y_v \in \text{range}(u)$ such that $y_v \to y$, and x_v such that $u(x_v)=y_v$. The sequence (x_v) has to be bounded since otherwise $|y_v| \to \infty$ along some subsequence. Without loss $x_v \to x$. Since u is continuous, $y=u(x)$, i.e. $y \in \text{range}(u)$. We have shown that range(u) is both open and closed in $\mathbb{R}^d$, consequently range$(u)=\mathbb{R}^d$. The proof is finished. □

Appendix

We prove two auxiliary results.

Lemma A1. *Let $h(t)$, $t \geqq 0$, be a non-negative decreasing function such that $\int_1^\infty h(t) \frac{dt}{t} < \infty$. Then there exists a function $k(t)$ such that*

i) *$k(t)$ is non-negative and decreasing,*
ii) *$k(t) \geqq h(t)$ for all $t \geqq 0$,*
iii) $\int_1^\infty k(t) \frac{dt}{t} < \infty$,
iv) *$k(t)$ is continuously differentiable and $k'(t) = O(t^{-1} k(t))$, as $t \to \infty$.*

Proof. Define $j(t) = t^{-1} \int_0^t h(s)\, ds$, $t>0$. Since $h(t)$ is decreasing, also $j(t)$ is decreasing and $j(t) \geqq h(t)$. By a partial integration

$$\int_1^\infty j(t) \frac{dt}{t} = j(1) - j(\infty) + \int_1^\infty h(t) \frac{dt}{t} < \infty.$$

Now define $k(t)=t^{-1}\int_0^t j(s)\,ds$. By repeating our arguments we see that i)–iii) are satisfied. Moreover, $j(t)$ is continuous, thus $k(t)$ is continuously differentiable, and, since $j(t)\leqq k(t)$

$$-\frac{k(t)}{t}\leqq\frac{j(t)}{t}-\frac{k(t)}{t}=k'(t)\leqq 0.$$

Thus iv) holds. □

Lemma A2. *Let $f(t)$, $t\geqq 1$, be a positive function such that $f(t)=o(t)$ and $f(t)/t$ is decreasing. Let u_t, $t\geqq 0$, be the solution of the equation*

$$du_t=f(u_t)\,dt, \qquad u_0=1.$$

Let $\varepsilon>0$. Then

$$\int_0^\infty\left(\frac{f((1-\varepsilon)\,u_t)}{(1-\varepsilon)\,u_t}-\frac{f(u_t)}{u_t}\right)dt=\infty.$$

Proof. Define $g(s)=e^{-s}f(e^s)$. Substituting $u_t=e^s$ in the above integral, we realize that our assertion is equivalent to

$$\int_0^\infty\left(\frac{g(s-c)}{g(s)}-1\right)ds=\infty, \tag{4.1}$$

where $c=-\log(1-\varepsilon)>0$. Note that $g(s)$ is decreasing and going to zero by assumption. We distinguish two cases:

i) Since $g(s)=o(1)$,

$$\int_0^\infty \log\frac{g(s-c)}{g(s)}\,ds=\int_{-c}^{0}\log g(s)\,ds-\lim_{n\to\infty}\int_n^{n+c}\log g(s)\,ds=\infty.$$

If now $g(s-c/2)/g(s)\to 1$, as $s\to\infty$, consequently $g(s-c)/g(s)\to 1$, (4.1) follows.

ii) Now assume that there are numbers $t_1<t_2<\dots$ going to infinity such that $g(t_v-c/2)/g(t_v)\geqq 1+\delta$ for some $\delta>0$. Since $g(s)$ is decreasing, $g(s-c)/g(s)\geqq 1+\delta$ for $t_v\leqq s\leqq t_v+c/2$, and (4.1) is trivially satisfied. □

Acknowledgements. A major part of this paper emerged during a stay of the author at the University of Wisconsin, Madison. The author likes to express his gratitude to Peter Ney and Tom Kurtz for their great hospitality. – This project was partly supported by the DFG.

References

1. Arnold, L., Oeljeklaus, E., Paradoux, E.: Almost sure and moment stability for linear Ito equations. Lect. Notes Math. vol. 1186, Berlin Heidelberg New York: Springer 1986
2. Bhattacharya, R.N.: Criteria for recurrence and existence of invariant measures for multidimensional diffusions. Ann. Probab. **6**, 541–553 (1978)
3. Clark, C.R.: Asymptotic properties of some multidimensional diffusions. Ann. Probab. **15**, 985–1008 (1987)

4. Cranston, M.: Invariant σ-Fields for a class of diffusions. Z. Wahrscheinlichkeitstheor. Verw. Geb. **65**, 161–180 (1983)
5. Durrett, R.: Brownian motion and martingales in analysis. Belmont: Wadsworth 1984
6. Gilbarg, D., Trudinger, N.S.: Elliptic partial differential equations of second order. Berlin Heidelberg New York: Springer 1977
7. Has'minskii, R.Z.: Stochastic stability of differential equations. Rockville: Alphen 1980
8. Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes. Amsterdam: North Holland 1981
9. Keller, G., Kersting, G., Rösler, U.: On the asymptotic behaviour of solutions of stochastic differential equations. Z. Wahrscheinlichkeitstheor. Verw. Geb. **68**, 163–189 (1984)
10. Pinsky, R.: Recurrence, transience and bounded harmonic functions for diffusions in the plane. Ann. Probab. **15**, 954–984 (1987)

Received June 14, 1988; in revised form March 3, 1989