JOURNAL OF DIFFERENTIAL EQUATIONS 11, 156–168 (1972)

Stability and Existence of Diffusions with Discontinuous or Rapidly Growing Drift Terms*

H. J. KUSHNER

Center for Dynamical Systems, Divisions of Applied Mathematics and Engineering Brown University, Providence, Rhode Island 02912

Received March 8, 1971

1. INTRODUCTION

For a vector $x = \{x_i\}$ and matrix $\sigma = \{\sigma_{ij}\}$, define the Euclidean norms $|x|^2 = \sum_i x_i^2$, $|\sigma|^2 = \sum_{i,j} \sigma_{ij}^2$, resp. Consider the homogeneous[1] Itô stochastic differential equation

$$dx = f(x)\,dt + \sigma(x)\,dz, \qquad t \geqslant 0, \tag{1}$$

where $\sigma(\cdot)$ satisfies growth and Lipschitz conditions of the types[2]

$$|\sigma(x)|^2 \leqslant K(1 + |x|^2), \tag{2a}$$

$$|\sigma(x) - \sigma(y)| \leqslant K|x - y| \tag{2b}$$

and $z(t)$ is a normalized vector-valued Wiener process. If

$$|f(x)|^2 \leqslant K(1 + |x|^2), \tag{3a}$$

and

$$|f(x) - f(y)| \leqslant K|x - y|, \tag{3b}$$

then the Itô existence theory is applicable to (1) and the stability properties can be discussed [1]. If (3b) holds locally, but (3a) is violated, a "local" stability

* This research was supported in part by the National Science Foundation under grant GK 24485, in part by the National Aeronautics and Space Administration under grant NGL 40-002-015, and in part by the Air Force Office of Scientific Research under grant AF-AFOSR 693-67D.

[1] The homogeneity condition is not essential, except in Section 4.

[2] K and K_i always denote real numbers; their value may change from usage to usage.

property [1, Theorem 8, Chap. 2] ensures the existence of a solution to (1) for all $t \geqslant 0$.

Recent investigations [2–5] have studied an important class of Eqs. (1), where $f(\cdot)$ is allowed some discontinuities. Rewrite (1) in the form (x^1 and x^2 are vectors)

$$dx = \begin{pmatrix} dx^1 \\ dx^2 \end{pmatrix} = \begin{matrix} f^1(x)\, dt \\ f^2(x)\, dt + \hat{f}(x)\, dt + \hat{\sigma}(x)\, dz, \end{matrix} \tag{4}$$

where we assume that the f^i and $\hat{\sigma}$ satisfy (3) and (2), respectively, and $\hat{\sigma}(x)$ has a uniformly bounded inverse. (Thus $\hat{\sigma}^{-1}(x)$ satisfies (2), but $\hat{f}(\cdot)$ does not necessarily satisfy (3)). In the sequel, we prove existence, uniqueness, and other properties of (4), when neither (3a) nor (3b) necessarily holds, but a 'local' stability property obtains. We also treat the problems of asymptotic stability, the existence of a unique invariant measure, and the convergence of the measures of (1) to the invariant measure.

DIffusions of the type (4) occur frequently in control applications. Consider, for example, a "white noise"-driven n-th-order differential equation where $\hat{f}$ is a "bang-bang" control taking the values $\{+1, -1\}$, or which may be discontinuous on a smooth "switching curve" and tend to infinity in certain directions. Also models such as

$$dx = \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} = \begin{pmatrix} x_2\, dt \\ -(x_1 + x_1^3)\, dt + \sigma\, dz \end{pmatrix}$$

are sometimes used, and the existence and asymptotic character of the corresponding measures are of interest.

2. Mathematical Preliminaries

Assume:

(C1) f^i and $\hat{\sigma}$ satisfy (3) and (2), respectively, and $\hat{\sigma}^{-1}(x)$ is uniformly bounded. $\hat{f}(\cdot)$ is a vector-valued Borel function of x which is bounded in any compact set.

(C2) The $I + \hat{\sigma}$ process (5) has a transition density $p(x; t, y)$.

(C3) (A condition on the discontinuities of $\hat{f}$.) Let S_m denote a sphere of radius m, whose center is the origin. Let $N_\epsilon(A)$ denote an ϵ-neighborhood of the set A and $\mu(A)$ the Lebesgue measure of A. Suppose there is a (discontinuity) set D so that

$$\mu(N_\epsilon(D \cap S_m)) \to 0,$$

as $\epsilon \to 0$ for each $m < \infty$. For each $\epsilon' > 0$, let there be an $\epsilon > 0$, so that $|x - y| < \epsilon$ implies $|\hat{f}(x, t) - \hat{f}(y, t)| < \epsilon'$ uniformly in x in bounded regions, provided that $x \notin N_\epsilon(D)$.

Assume (C1). Let Ω denote the sample space. We use the notation $(\Omega, z(t), \mathscr{B}_t, P)$ for the Wiener process on $[0, \infty]$, where $\mathscr{B}_t$ measures $z(s)$, $s \leqslant t$ and $z(r_2)$–$z(r_1)$ is independent of $\mathscr{B}_t$ for $t \leqslant r_1 \leqslant r_2$, and P is the measure on all the $\mathscr{B}_t$. We say that $z(t)$ is a Wiener process on $(\Omega, \mathscr{B}_t, P)$. Let $x(t)$ be the unique solution to the Itô Eq. (5)

$$dx = \begin{matrix} dx^1 = f^1(x)\, dt \\ dx^2 = f^2(x)\, dt + \hat{\sigma}(x)\, dz. \end{matrix} \tag{5}$$

We say that $x(t)$ is an Itô process with respect to $(\Omega, z(t), \mathscr{B}_t, P_x)$, where P_x denotes the probability given that $x(0) = x$ (and E_x denotes the corresponding expectation). E and P denote expectation and probability for functionals of $z(t)$. Define Ω_T as the sample space for $z(t)$, $t \leqslant T$. Suppose that

$$\int_0^T |\hat{\sigma}^{-1}(x(t))\, \hat{f}(x(t))|^2\, dt < \infty \qquad \text{w.p.1.} \tag{6}$$

(which is certainly true if $\hat{f}$ is bounded). Define

$$\zeta_0^T(\hat{f}) = \int_0^T \hat{\sigma}^{-1}(x(t))\, \hat{f}(x(t))\, dz(t) - \tfrac{1}{2} \int_0^T |\hat{\sigma}^{-1}(x(t))\, \hat{f}(x(t))|^2\, dt,$$

and suppose that

$$E_x \exp \zeta_0^T(\hat{f}) = 1 \tag{7}$$

((7) holds for all $T < \infty$ if $\hat{f}$ is bounded.) Then the set function $\hat{p}_x^T$ defined by[3]

$$\tilde{P}_x^T(A) = \int_A \exp \zeta_0^T(f) \cdot P(d\omega)$$

is a probability measure on the $\mathscr{B}_t$, $t \leqslant T$. The process $\tilde{z}(t)$, $t \leqslant T$,

$$\tilde{z}(t) = z(t) - \int_0^t \hat{\sigma}^{-1}(x(s))\, \hat{f}(x(s))\, ds$$

is a Wiener process on $(\Omega_T, \mathscr{B}_t, \tilde{P}_x^T)$, and the process

$$\begin{aligned} dx &= \begin{matrix} f^1(x)\, dt \\ f^2(x)\, dt + \hat{f}(x)\, dt + \hat{\sigma}(x)[dz - \hat{\sigma}^{-1}(x)\, \hat{f}(x)\, dt] \end{matrix} \\ &= \begin{matrix} f^1(x)\, dt \\ f^2(x)\, dt + \hat{f}(x)\, dt + \hat{\sigma}(x)\, d\tilde{z} \end{matrix} \end{aligned} \tag{8}$$

[3] The measure $\tilde{P}_x^T$ depends on the initial condition of (5), as does the Wiener process $\tilde{z}(t)$.

is an Itô process with respect to $(\Omega_T, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x^T)$. The construction was first done by Girsanov [4], and exploited by Benes [5], Rishel [2] and then Kushner [3], for several control problems. Note the sample space Ω_T, the σ-algebras $\mathscr{B}_t$, and the random variables $x(t)$ for the Wiener process $\tilde{z}(t)$, and Itô process $(\Omega_T, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x^T)$ are the same as those for the Wiener process $z(t)$ and Itô process (5), for $t \leqslant T$. Only the measures have been changed. The process (8) is constructed by a transformation of measures on the "nicer" process (5).

Assume that $\hat{f}$ is bounded and that (C1) *holds.*

The following facts (drawn from [2–4]) about (8) will be needed:

(01) [3, Theorem 5]. Assume, in addition, (C2–3). The multivariate distributions of (8) are continuous with respect to the initial condition $x(0)$ (in the sense that the characteristic functions are continuous in $x(0)$).

(02 [3, Theorem 2]. The solution to (8) is unique in the sense that any solutions to (8) have the same multivariate distributions.

(03) Assume, in addition, (C2–3). Then

$$\tilde{E}_x^T \sup_{t \geqslant s \geqslant 0} | x(s) - x |^2 \leqslant K_1 t(1 + | x |^2), \qquad t \leqslant T,$$

$$\tilde{E}_x^T \sup_{t \geqslant s \geqslant 0} | x(s) - x |^4 \leqslant K_1 t^2(1 + | x |^4), \qquad t \leqslant T,$$

where $\tilde{E}_x^T$ is the expectation given $x(0) = x$, and K_1 depends on the bound on $\hat{f}$. The proof of (03) is close to that of (27)–(28) of [3, Theorem 6].

(04) Assume, in addition, (C2). Then (8) has a transition density, which is any version of [2, Lemma 1] (boundedness of $\hat{f}$ is not required if (6)–(7) hold) for $t \leqslant T$

$$q(x; t, y) = \tilde{E}_x^T[\exp \zeta_0^t(\hat{f}) | \, x(t) = y] \, p(x; t, y).$$

Also $\hat{f}$ is not required to be bounded in (05).

(05) [4, Corollary to Lemma 3]. Let $g(\omega)$ be $\mathscr{B}_t$ measurable with $\tilde{E}_x^T | g(\omega)| < \infty$, and $t \leqslant T$. Then, for $s \leqslant t \leqslant T$, w.p.1,

$$\tilde{E}_x^T[g(\omega) \mid \mathscr{B}_s] = E[g(\omega) \exp \zeta_s^t(\hat{f}) | \, \mathscr{B}_s].$$

(The equation also holds if $\mathscr{B}_s$ is replaced by any sub σ-algebra of $\mathscr{B}_s$.)

Fix T, and define $\tilde{z}(t)$ and $\tilde{P}_x^T$ by the Girsanov transformation. *Write* $\tilde{z}(t)$ *as* $\tilde{z}^T(t)$. Suppose that (6)–(7) hold for a time $T_1 > T$, and define the corres ponding Ω_{T_1}, $\tilde{z}^{T_1}(t)$, $\tilde{P}_x^{T_1}$. Then $\tilde{z}^{T_1}(t) = \tilde{z}^T(t)$ for $t \leqslant T$, and on sets B of $\mathscr{B}_T$ we have $\tilde{P}_x{}^T(B) = \tilde{P}_x^{T_1}(B)$. This follows from (05) since (χ_B is the characteristic function of the set B)

$$\begin{aligned}\tilde{P}_x^{T_1}(B) &= E_x[E_x(\chi_B \exp \zeta_0^{T_1}(\hat{f}) \mid \mathscr{B}_T)] \\ &= E_x\chi_B \exp \zeta_0{}^T(\hat{f})[E_x(\exp \zeta_T^{T_1}(\hat{f}) \mid \mathscr{B}_T)] \\ &= E_x\chi_B \exp \zeta_0{}^T(\hat{f}) = \tilde{P}_x{}^T(B).\end{aligned}$$

Thus $\tilde{P}_x^{T_1}$ is an extension of $\tilde{P}_x{}^T$. If (6)–(7) hold for each $T < \infty$, we can replace Ω_T by Ω and define a unique measure $\tilde{P}_x$ on all the $\mathscr{B}_t$, $t < \infty$, which will be consistent with the $\tilde{P}_x{}^T$ on $\mathscr{B}_T$. Then $\tilde{z}(t)$ will be an Itô process with respect to $(\Omega, \mathscr{B}_t, \tilde{P}_x)$, and $(\Omega, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x)$ an Itô process (for all $t < \infty$). Both (6)–(7) hold for all $T < \infty$ if $\hat{f}$ is bounded. Let $\mathscr{B} = \bigcup_{t \geqslant 0} \mathscr{B}_t$.

3. Existence of a Solution to (8) for Unbounded $\hat{f}$

Let $V(x)$ denote a nonnegative twice continuously differentiable function which tends to infinity as $|x| \to \infty$. Define $Q_N = \{x : V(x) < N\}$ and define $\hat{f}^N(x) \equiv \hat{f}(x)$ for $x \in Q_N$ and $\hat{f}^N(x) = 0$, $x \notin Q_N$. Define

$$C_N{}^T = \{\omega : x(t) \in Q_N , t \in [0, T]\}.$$

Let $\tilde{\mathscr{L}}$ denote the differential generator of the process (8) and write $\tilde{\mathscr{L}}^N$ for the differential generator when $\hat{f}$ is replaced by $\hat{f}^N$ in (8). Theorem 1 uses a stability idea to prove existence for (8), for all $t < \infty$. Lemma 1 is used in Theorem 1.

Lemma 1. *Assume* (C1)–(C3) *and that* $\hat{f}(x)$ *is bounded. Then* $x(t)$, *the Itô process* (8) *on* $(\Omega, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x)$ *is a Markov process and a Feller process, hence a strong Markov process.*[4]

Proof. It is sufficient only to consider the process on an arbitrary finite interval $[0, T]$. The Markov property, under (C1) and for bounded $\hat{f}(x)$, is proved in the first part of Theorem 2, and we will not duplicate the proof here. To prove the Feller property it is sufficient to show that $G(x) \equiv \tilde{E}_x{}^T g(x(t))$ is continuous in x for any continuous function $g(x)$ with compact support.

[4] If $\tilde{E}_x g(x(t))$ is continuous in x for each continuous bounded real-valued function $g(x)$, then the process is said to be a Feller process. A Markov process which is a right continuous Feller process is a strong Markov process [6, Theorem 3.10].

Let $g(x)$ be such a function. For any $\epsilon > 0$ there are finitely many rectangles (open or closed or partly open) $A_{\epsilon 1}, \ldots, A_{\epsilon n_\epsilon}$ and points $x_{\epsilon i} \in A_{\epsilon i}$ so that

$$\sup_x | g(x) - g_\epsilon(x)| < \epsilon,$$

where

$$g_\epsilon(x) = \sum_{i=1}^{n_\epsilon} g(x_{\epsilon i}) I_{A_{\epsilon i}}(x).$$

By (01) and (04), the Itô process (8) on $(\Omega_T, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x^T)$ has a density, and the $\tilde{P}_x^T\{x(t) \in A_{\epsilon i}\}$ are continuous in x, for any rectangle (with closed, open, or partly open boundary). Thus $G_\epsilon(x) = \tilde{E}_x^T g_\epsilon(x(t))$ is continuous in x, and since $G_\epsilon(x)$ is within ϵ of $G(x)$, the Lemma is proved.

THEOREM 1. *Assume* (C1)–(C3) *and the conditions on* $V(x)$ *given above Lemma 1. Let* $\tilde{\mathscr{L}} V(x) \leqslant 0$ *for all* x *not in some* Q_a, $a < \infty$. *Then*

$$E_x \exp \zeta_0^T(\hat{f}) = 1 \tag{9}$$

for all $T < \infty$, *and*

$$\tilde{z}(t) = z(t) - \int_0^t \hat{\sigma}^{-1}(x(s)) \hat{f}(x(s))\, ds$$

is a Wiener process, for all $t < \infty$ *with respect to* $(\Omega, \mathscr{B}_t, \tilde{P}_x)$. *The solution to* (8) *exists for all* $t < \infty$. *It is an Itô process with respect to* $(\Omega, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x)$, *and it is unique* (*in the sense that the multivariate distributions of any two solutions are equal*).

Remark. Let $f(y), \sigma(y)$ satisfy (3), (2) locally, and let $\mathscr{L}_1$ denote the differential generator, with coefficients determined by $f(y), \sigma(y)$. If $V(x)$ and $\mathscr{L}_1 V(x)$ have the properties required in Theorem 1, then the proof can be altered to yield existence and uniqueness for the process

$$dy = f(y)\, dt + \sigma(y)\, dz.$$

Proof. Let $\hat{f}^N$ replace $\hat{f}$ in (8), where $N > a$. Let $\tilde{P}_x^{N,T}$ denote the transformed measure with $\tilde{P}_x^{N,T}(A) = \int_A \exp \zeta_0^T(\hat{f}^N)\, dP$ and $\tilde{P}_x^N$ the extension of the $\tilde{P}_x^{N,T}$ to the σ-algebra $\mathscr{B}$ on Ω. Write the Wiener process corresponding to $\tilde{P}_x^N$ as $\tilde{z}^N(t)$ (instead of $\tilde{z}(t)$). Then (8) is an Itô process with respect to $(\Omega, \tilde{z}^N(t), \mathscr{B}_t, \tilde{P}_x^N)$. By virtue of (03)[5] (for $x = x(0)$)

$$\tilde{P}_x^N\{ \sup_{t \geqslant s \geqslant 0} | x(s) - x | \geqslant \epsilon > 0\} \to 0 \tag{10}$$

[5] For the proof of (03), (C1)–(C3) are used.

as $t \to 0$, uniformly for x in compact regions. Also $\tilde{\mathscr{L}}^N V(x) \leqslant 0$ in[6] $Q_N - Q_a - \partial Q_a \equiv Q_{N,a}$. Let τ denote the first exit[7] time of the path $x(t)$ from $Q_N - Q_a - \partial Q_a$, and $t \cap \tau \equiv \min(t, \tau)$. Then, by Itô's Lemma, $\tilde{E}_x^N V(x(t \cap \tau)) - V(x) \leqslant 0$ for $x \in Q_N - Q_a$. Since

$$\tilde{E}_x^N V(x(t \cap \tau)) - a \geqslant (N - a)\, \tilde{P}_x^N\{x(s) \text{ hits } \partial Q_N \text{ before } \partial Q_a$$
$$\text{and leaves } Q_{N,a} \text{ in } [0, t]\},$$

we can conclude that

$$\tilde{P}_x^N\{x^N(t) \text{ hits } \partial Q_N \text{ before } \partial Q_a \text{ and leaves } Q_{N,a} \text{ in } [0, T]\} \leqslant \frac{V(x) - a}{N - a} \equiv \epsilon_3 . \tag{11}$$

We will show that for each $\epsilon > 0$, there is an $N < \infty$ so that

$$\tilde{P}_x^N\{C_N^T\} \geqslant 1 - \epsilon. \tag{12}$$

Fix $a_1 > a$. There is a $\delta_0 > 0$ so that

$$\min_{\tilde{y} \in \partial Q_a, y \in \partial Q_{a_1}} | \tilde{y} - y | \geqslant \delta_0 .$$

Let $A \subset C_N^T$. Then, since $\hat{f}^N(x(t)) = \hat{f}^M(x(t))$ on $[0, T]$ for $M \geqslant N$ and $\omega \in C_N^T$, we have

$$\tilde{P}_x^M(A) = E_x \exp \zeta_0^T(\hat{f}^M)\, \chi_A = E_x \exp \zeta_0^T(\hat{f}^N)\, \chi_A = \tilde{P}_x^N(A). \tag{13}$$

Let δ_2 be an arbitrary real number. (03) implies that, for any $\delta_1 < \delta_2$,

$$\sup_{y \in \partial Q_a} \tilde{P}_y^N\{ \max_{\delta_1 \geqslant t \geqslant 0} | x(t) - y | \geqslant \delta_0\} \leqslant K_2 \frac{\delta_1^2}{\delta_0^4} = \epsilon_2 .$$

But (13) implies that the constant K_2 depends only on the number a_1 and does not depend on N, for $N > a_1$. Thus, we can assume that K_2 does not depend on N.

Let n be an arbitrary integer for which $T/n \leqslant \delta_2$. Define δ_1 by $T/n = \delta_1$. Let G_N^T denote the event that $x(t)$ goes to ∂Q_a before ∂Q_N (*or never leaves* $Q_{N,a}$), then takes no less time than δ_1 to reach ∂Q_{a_1}, then returns to ∂Q_a *no fewer than* $n - 1$ additional times and after each return takes no less than δ_1 to reach ∂Q_{a_1} before leaving Q_N for the first time. Since $C_N^T \supset G_N^T$ and

[6] ∂Q_N is the boundary of the set Q_N.

[7] If τ is undefined for some path, set $\tau = +\infty$. Note that the exit time $\tau(\omega)$ (as a path function) for $x(t)$ and $\tilde{x}^N(t)$ are the same; but their distributions may differ.

$G_N^T \in \mathscr{B}$, we have $\tilde{P}_x^N\{C_N^T\} \geqslant \tilde{P}_x^N\{G_N^T\}$. Also the inequality $\tilde{P}_x^N\{G_N^T\} \geqslant 1 - n(\epsilon_1 + \epsilon_2) - \epsilon_3$, where

$$\epsilon_1 = \max_{y \in \partial Q_{a_1}} \tilde{P}_y^N\{x(t) \text{ reaches } \partial Q_N \text{ before } \partial Q_a\} \leqslant \frac{(a_1 - a)}{(N - a)}$$

follow from the fact that the Itô process (8) on $(\Omega_T, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x^{T,N})$, with $\tilde{z}(t)$ and $f(x)$ replaced by $\tilde{z}^N(t)$ and $\hat{f}^N(x)$, is a Feller (hence a strong Markov) process (Lemma 1). Thus, using $\delta_1 = T/n$,

$$\tilde{P}_x^N\{G_N^T\} \geqslant 1 - n\left(\frac{a_1 - a}{N - a} + \frac{K_2 T^2}{n^2 \delta_0^4}\right) - \frac{V(x) - a}{N - a}$$

and N and n can be chosen so that $\tilde{P}_x^N\{G_N^T\} \geqslant 1 - \epsilon$.

$$1 \geqslant \tilde{P}_x^T(\Omega_T) \equiv E_x \exp \zeta_0^T(\hat{f}) \geqslant E_x \exp \zeta_0^T(\hat{f}^N)\, \chi_{C_N^T} \geqslant 1 - \epsilon.$$

Since ϵ is arbitrary, (9) holds, $\tilde{z}(t)$, $t \leqslant T$, is a Brownian motion with respect to $(\Omega_T, \mathscr{B}_t, \tilde{P}_x^T)$ and $x(t)$, $t \leqslant T$, an Itô process with respect to $(\Omega_T, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x^T)$. Furthermore, since T is arbitrary, we can replace $t \leqslant T$ by $t < \infty$ and $\tilde{P}_x^T$ and Ω_T by $\tilde{P}_x$ and Ω.

The process (8) is unique in the following sense. Suppose that both $x^i(t)$, $i = 1, 2$ satisfy (8). Let $x^{i,N}(t)$ denote the processes which result when $\hat{f}^N$ replaces $\hat{f}$. Suppose that if $x^{i,N}(t) \in Q_N$ for all $t \in [0, T]$, then $x^i(t)$ coincides with $x^{i,N}(t)$ on $[0, T]$. Then the uniqueness of the $x^{i,N}(t)$ (in the sense of multivariate distributions) and the fact that $\tilde{P}_x^N\{C_N^T\} = \tilde{P}_x^M\{C_N^T\} \geqslant 1 - \epsilon$ for $M > N$ (the $\tilde{P}_x^N$ do not depend on i) imply uniqueness of the $x^i(t)$ in the sense of multivariate distributions. Q.E.D.

Remark. Lemma 7 of [4] would appear to yield existence for a large class of unbounded $\hat{f}$. But an examination of the proof shows that its content is the following. Let processes (5) and (8) exist with respect to some Wiener process, with (5) being unique, and $\int_0^T |\,\hat{\sigma}^{-1}(x(t))\,\hat{f}(x(t))|^2\, dt < \infty$ w.p.1, where $x(t)$ is the solution to (5). Under some minor subsidiary condition, it is proved that

$$E_x \exp \zeta_0^T(\hat{f}) = 1,$$

where the expectation corresponds to (5). Then (8) can be obtained by a Girsanov transformation from (5). But both the square integrability property and existence for (8) must be established first. But these properties are essentially the desired result.

3. Markov Properties of (8)

In Theorem 2, we will use the condition (C4). In each compact x set, there is an $\alpha > 1$ and $M < \infty$ so that

$$\int p^{\alpha}(x; t, y) \leqslant M < \infty.$$

Theorem 2. *Assume* (C1)–(C3) *and the conditions on* V *and* $\mathscr{L}V$ *of Theorem* 1. *Then the process* (8) *is a strong Markov process.*

If (C4) *holds, for some* $\alpha > 1$, (8) *is a strong Feller process.*

Proof. The terminology of Theorem 1 will be used. By Theorem 1, the process is defined on the time interval $[0, \infty)$, and has continuous paths w.p.l.

First, we prove that (8) is a Markov process. Let $\mathscr{B}_t^x \subset \mathscr{B}_t$ measure $x(s)$, $s \leqslant t$. Define the transition function $\tilde{P}_x(x; t, A) = \tilde{P}_x\{x(t) \in A\}$. Since the right term of

$$\tilde{P}_x\{x(t) \in A\} = E_x \chi_{\{x(t)\in A\}} \exp \zeta_0^t(\hat{f})$$

is a Borel measurable function of x, so is $\tilde{P}(x; t, A)$ for each $A \in \mathscr{B}_t^x$. *Now assume that* $\hat{f}^N$ *replaces* $\hat{f}$. The Chapman–Kolmogorov equation holds since, by (05) and the fact that (5) is a Markov process,

$$\begin{aligned}\tilde{E}_x^N[\chi_{\{x(t+s)\in A\}} \mid \mathscr{B}_s^x] &= E_x[\chi_{\{x(t+s)\in A\}} \exp \zeta_s^{s+t}(\hat{f}^N) \mid \mathscr{B}_s^x] \\ &= E_{x(s)}[\chi_{\{x(t)\in A\}} \exp \zeta_0^t(\hat{f}^N)] = \tilde{P}^N(x(s); t, A)\end{aligned}$$

w.p.1. Thus by the definition Dynkin of [6, Chap. 3], $x^N(t)$ (the Itô process on $(\Omega, \tilde{z}^N(t), \mathscr{B}_t, \tilde{P}_x^N)$ corresponding to the use of $\hat{f}^N$) is a Markov process.

The σ-algebras $\mathscr{B}_t^x$ also measure (8). The measure $\tilde{P}_x$ for the unbounded $\hat{f}$, has the correct conditioning properties since, by (05) and the dominated convergence theorem,

$$\begin{aligned}&\tilde{E}_x[\chi_{\{x(t+s)\in A\}}\chi_{C_{t+s}^N} \mid \mathscr{B}_s^x] \\ &\quad = E_x[\chi_{\{x(t+s)\in A\}}\chi_{C_{t+s}^N} \exp \zeta_s^{t+s}(\hat{f}) \mid \mathscr{B}_s^x] \\ &\quad \to E_x[\chi_{\{x(t+s)\in A\}} \exp \zeta_s^{t+s}(\hat{f}) \mid \mathscr{B}_s^x] \\ &\quad = E_{x(s)}[\chi_{x\{(t)\in A\}} \exp \zeta_0^t(\hat{f})] \\ &\quad = \tilde{P}(x(s); t, A)\end{aligned}$$

w.p.1. Then, by the definition [6, Chapter 3], (8) is a Markov process.

Since (8) is a Feller process, it is also a strong Markov process [6, Theorem 3.10]. The proof is Lemma 1. The proof of the stronger "strong" Feller[8] property will be given next, under the additional condition (C4). Let (C4) hold.

Supposing that (8) is a strong Feller process if $\hat{f}$ is bounded, we show that it is also a strong Feller process for unbounded $\hat{f}$. Let $g(\cdot)$ be bounded and measurable. Then $\tilde{E}_x^N g(x(t)) \equiv G^N(x)$ is continuous in x, for $t > 0$. Write $G(x) = \tilde{E}_x g(x(t))$. Then

$$| G(x) - G^N(x)| \leqslant \max_x | g(x)| \cdot [\tilde{P}_x\{\Omega - C_N{}^T\} + \tilde{P}_x{}^N\{\Omega - C_N{}^T\}] \to 0 \quad \text{as} \quad N \to \infty$$

uniformly in any compact x set. Thus, $G(x)$, being the uniform limit of continuous functions is continuous.

Finally, *suppose $\hat{f}$ is bounded and* (C4) *holds*. Reproducing an argument of Rishel [2], we show that for each compact x set there is a $\beta > 1$ and $M < \infty$ so that (q is the density of (8); see (04))

$$\int q^\beta(x; t, y)\, dy \leqslant M_1 < \infty. \tag{14}$$

Define $r(x; t, y) \equiv \tilde{E}_x[\exp \zeta_0{}^t(\hat{f}) \mid x(t) = y]$. Let $m^{-1} + n^{-1} = 1$, and note that, for any $\rho > 1$ and compact x set, there is an $N_\rho < \infty$ so that $\tilde{E}_x \exp \rho\zeta_0{}^t(\hat{f}) \leqslant N_\rho$ [4, Lemma 1]. Let $\beta > \beta_1$, $\beta > 1$. By Holder's inequality,

$$\begin{aligned}\int & p^\beta(x; t, y)\, r^\beta(x; t, y) \\ &= \int p^{\beta_1}(x; t, y)\, r^\beta(x; t, y)\, p^{\beta-\beta_1}(x; t, y)\, dy \\ &\leqslant \left[\int p^{\beta_1 n}(x; t, y)\, r^{\beta n}(x; t, y)\, dy\right]^{1/n} \left[\int p^{(\beta-\beta_1)m}(x; t, y)\, dy\right]^{1/m}.\end{aligned}$$

We can choose $\beta > 1$, $\beta > \beta_1$, m, n and $\rho > 1$ so that $(\beta - \beta_1)\, m = \alpha$, $\beta n = \rho$, $\beta_1 n = 1$, which, together with (C4), proves (14). Equation (14) implies that, as x varies in any compact set, the family $q(x; t, y)$ of functions of y is uniformly integrable. This, together with the continuity (in x) of $\tilde{P}(x; t, (-\infty, b))$ for any vector b (recall that there is a density) implies that $\tilde{P}(x; t, A)$ is continuous in x for any Borel set A, which implies, in turn, the strong Feller property. For more detail, note that the boundary of any

[8] $x(t)$ is a strong Feller process if $E_x f(x(t))$ is continuous in x for any bounded Borel function $f(x)$ and $t > 0$.

rectangle in the range space of $x(t)$ has zero probability, and that $\tilde{P}(x; t, A)$ is continuous in x on the algebra of sets which are sums of rectangles (open, closed or neither) by (01). Let $\tilde{P}(x; t, A_j)$ be continuous in x for a collection of sets A_j, which increase monotonically to A

$$\tilde{P}(x; t, A) = \int_{A_j} q(x; t, y)\, dy + \int_{A-A_j} q(x; t, y)\, dy.$$

The second integral goes to zero as $j \to \infty$ uniformly in x in any compact set, by the uniform integrability of $q(x; t, y)$. Since the first integral is continuous, so is the uniform limit $\tilde{P}(x; t, A)$. Thus $\tilde{P}\,(x; t, A)$ is continuous in x on the least σ-algebra containing sums of rectangles, the Borel field. Q.E.D.

4. The Invariant Measure, and the Asymptotic Properties of the Measures of (8)

In [8], under the conditions (D1)–(D5), Khasminskii proved the existence of a unique σ-finite invariant measure for a process $x(t)$ with a stationary transition function $\tilde{P}(x; t, A)$ under the conditions (D1–5).

(D1) For any ϵ neighborhood $N_\epsilon(x)$ of x, $1 - \tilde{P}(x; t, N_\epsilon(x)) = o(t)$ uniformly in x in any compact set.

(D2) The process is a strong Markov and strong Feller process.

(D3) $\tilde{P}(x; t, U) > 0$ for all open sets U and $t > 0$.

(D4) The paths are continuous w.p.1.

(D5) The process is recurrent. (There is some compact set K and a random time $\tau < \infty$ w.p.1. so that $x(\tau) \in K$ w.p.1., for each initial condition.)

In [9], Kushner applied the result in [8] to obtain a sufficient condition for the convergence of the measures of a class of diffusions to a unique invariant measure. Theorem 3 includes the prior result as a special case. Zakai [10] has treated the invariant measure problem for a class of diffusions satisfying (2)–(3), using a general method of Benes [11]. A similar problem is treated in Elliot [12]. Elliot's method involves a condition on a Lie algebra generated by certain functions of the diffusion coefficients, which is hard to check in special cases. The result of Benes [11] (concerning only existence of an invariant measure) uses the condition that $\lim_{|x|\to\infty} P(x; t, K) \to 0$ for all compact sets K. This would not always hold under our conditions. For example, the solution to $\dot{x} + x^3 = 0$, reaches $x = 1$ in a time that is bounded as $x(0) \to \infty$, and we would expect a similar result for $dx = -x^3\, dt + \sigma\, dz$.

Theorem 3. *Assume* (C1)–(C4), *and the conditions on* $V(\cdot)$ *in Theorem* 1, *except let $\tilde{\mathscr{L}}V(x) \leqslant -\epsilon < 0$ outside of Q_a. Let* (5) *have a nowhere zero density for each initial condition x. Then* (8) *has a unique invariant measure $Q(\cdot)$ and $\tilde{P}(x; t, A) \to Q(A)$ as $t \to \infty$ for any x. Both $\tilde{P}(x; t, A)$ and $Q(A)$ have nowhere zero densities.*

Remark. Theorem 3 only deals with invariant measures, but almost all of stability results in [1] can be carried over to the problem with discontinuous drift terms.

Proof. The second inequality of (03) implies (D1) for bounded $\hat{f}$, and, hence, for the processes $x^N(t)$. But, if (D1) holds for each $x^N(t)$, it holds for (8). (D2) is proved in Theorem 2. Since $\tilde{E}_x[\exp \zeta_0^t(\hat{f}) \mid x(t) = y] > 0$ w.p.1. and $p(x; t, y) > 0$ for all y by assumption, $q(x; t, y)$ (the density for $\tilde{P}(x; t, A)$) is positive for almost all y (Lebesgue measure). This implies (D3). (D4) is a consequence of Theorem 1. (D5) is a consequence of $\tilde{\mathscr{L}}V(x) \leqslant -\epsilon < 0$ for all large x. (See Theorem 4 in [9]). Indeed, the average time to leave the set $Q_N - Q_a - \partial Q_a$ (for $x(0) = x$) is bounded above by $(V(x) - a)/\epsilon < \infty$. This together with (11) gives (D5). Thus all (D1–5) hold.

$Q(A)$ satisfies

$$Q(A) = \int Q(dx)\, \tilde{P}(x; t, A)$$
$$= \int_A du \int Q(dx)\, q(x; t, u).$$

Thus $Q(A) > 0$ for all sets A of positive Lebesgue measure and has density $\int Q(dx)\, q(x; t, u)$, which must be positive almost everywhere.

For a process with a transition density and a unique invariant measure $Q(\cdot)$ with a nowhere zero density, Doob [7, Theorem 5] proves that $\tilde{P}(x; t, A) \to Q(A)$ as $t \to \infty$ for any x. Q.E.D.

References

1. H. J. Kushner, "Stochastic Stability and Control," Academic Press, New York, 1967.
2. R. W. Rishel, Weak solutions of the partial differential equations of dynamic programming, to appear in *SIAM J. Control.*
3. H. J. Kushner, Approximations to and local properties of diffusions with discontinuous controls, to appear in *SIAM J. Control.*
4. I. V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, *Theor. Probability Appl.* **5** (1960), 285–301 (translation).
5. V. E. Benes, The existence of optimal strategies based on specified information for a class of stochastic decision problems, *SIAM J. Control,* 1970.
6. E. B. Dynkin, "Markov Processes," Springer, Berlin, 1965 (translation from 1962 Russian Edition).

7. J. L. Doob, Asymptotic properties of Markov transition probabilities, *Trans. Amer. Math. Soc.* **3** (1948), 393–421.
8. R. Z. Khasminskii, Ergodic properties of recurrent diffusion processes and stabilization of the solution to the Cauchy problem for parabolic equations, *Theor. Probability Appl.* **9** (1960), 179–196.
9. H. J. Kushner, The cauchy problem for a class of degenerate parabolic equations, and asymptotic properties of related diffusion processes, *J. Differential Equations* **6** (1969), 209–231.
10. M. Zakai, A Lyapunov criteria for the existence of stationary probability distributions for systems perturbed by noise, *SIAM J. Control* **7** (1969), 390–397.
11. V. E. Benes, Finite regular invariant measures for Feller processes, *J. Appl. Probability* **5** (1968).
12. D. L. Elliot, "Behavior of controllable nonlinear systems driven by white noise," Second Hawaiian Conference on System Science, January 1969, Dissertation, Department of Engineering, University of California, Los Angeles, Calif., 1969.