STOCHASTIC EQUATIONS WITH DISCONTINUOUS DRIFT(¹)

BY

EDWARD D. CONWAY

Abstract. We study stochastic differential equations, $dx = a\,dt + \sigma\,d\beta$ where β denotes a Brownian motion. By relaxing the definition of solutions we are able to prove existence theorems assuming only that a is measurable, σ is continuous and that both grow linearly at infinity. Nondegeneracy is not assumed. The relaxed definition of solution is an extension of A. F. Filippov's definition in the deterministic case. When σ is constant we prove one-sided uniqueness and approximation theorems under the assumption that a satisfies a one-sided Lipschitz condition.

We consider the stochastic equation

$$dx = a\,dt + \sigma\,d\beta.$$

If the drift coefficient, a, and diffusion matrix, σ, are continuous on R^{d+1}, then solutions are known to exist [1]. If σ is assumed to be nondegenerate, i.e. $\sigma\sigma^*$ is positive definite, then existence of a solution can be proved assuming that a is bounded and measurable [9]. Moreover, consideration of the one-dimensional example $\sigma = 0$, $a(t, x) = -\text{sgn}(x)$ shows that continuity is necessary if degeneracy is allowed. However, in the deterministic case ($\sigma = 0$), A. F. Filippov [2] has shown that a natural and fruitful theory for discontinuous direction fields is possible if the definition of solution is relaxed. In this paper we bring Filippov's ideas into the stochastic context. This results in a fairly general existence theorem (*Theorem* 3). We were less successful in our treatment of uniqueness, however. That result (*Theorem* 4) is limited to the case of constant diffusion matrix. Our last result (*Theorem* 5) concerns approximation of relaxed solutions by solutions of non-degenerate equations.

It is a pleasure to acknowledge J. Goldstein with whom I have had several fruitful conversations. I also thank the referee for bringing the work of Prohorov to my attention.

1. **Existence of relaxed solutions.** We shall denote points of R^d by $x = (x_1, \ldots, x_d)$ and points of R^{d+1} by $(t, x) = (t, x_1, \ldots, x_d)$. $\langle x, y \rangle = x_1 y_1 + \cdots + x_d y_d$ denotes the inner product in R^d and $|x| = \sqrt{\langle x, d \rangle}$.

We shall be concerned with the stochastic equation

$$(1.1) \qquad x(t) = x_0 + \int_{t_0}^{t} a(s, x(s)) \, ds + \int_{t_0}^{t} \sigma(s, x(s)) \, dw(s)$$

for $t \in I = [t_0, T]$. a maps $I \times R^d$ into R^d and σ maps $I \times R^d$ into $d \times l$ matrices with components denoted by σ_{ij}. x_0 is an R^d-valued random variable on a probability space (Ω, Σ, P) and $\{w(t) : t \in I\}$ is the generic symbol for an l-dimensional Wiener process (mean $= 0$, variance parameter $= 1$) [1] (cf. also "standard Brownian motion process" in [3]). The stochastic integral in (1.1) is the R^d-valued process whose ith component is

$$\sum_{j=1}^{l} \int_{t_0}^{t} \sigma_{ij}(s, x(s)) \, dw_j(s)$$

where these integrals are the usual Itô integrals with respect to the (independent) component processes of $\{w(s) : s \in I\}$ [1].

DEFINITION. Let $W = \{w(t) : t \in I\}$ be a Wiener process over a probability space (Ω, Σ, P) and x_0 be a random variable over this space. A *solution of* (1.1) *relative to* W is a sample continuous process $\{x(t) : t \in I\}$ over (Ω, Σ, P) such that $P(A) = 1$ where A is the subset of Ω for which (1.1) holds for all $t \in I$.

Skorohod has shown in [1] that if a and σ are continuous and have at most linear growth at infinity then there is a solution relative to some Wiener process over the Lebesgue interval, $[0, 1]$, as sample space. In studying discontinuous coefficients we shall also consider the associated "mollified" equation

$$(1.2) \qquad x^{(n)}(t) = x_0 + \int_{t_0}^{t} a^{(n)}(s, x^{(n)}(s)) \, ds + \int_{t_0}^{t} \sigma^{(n)}(s, x^{(n)}(s)) \, d\beta(s)$$

where

$$a^{(n)}(t, \cdot) = a(t, \cdot) * \omega_{1/n}, \qquad \sigma^{(n)}(t, \cdot) = \sigma(t, \cdot) * \omega_{1/n}$$

and ω_ρ is the usual d-dimensional smoothing kernel of support radius ρ (cf. "regularization" in [4]).

THEOREM 1. *Let* (i) *a and σ be Lebesgue measurable,*

(ii) *there be a function b in $L_p(I)$, $p > 4$, such that $|a(t, x)|^2 \leqq b^2(t)(1 + |x|^2)$ and $|\sigma_{ij}(t, x)|^2 \leqq b^2(t)(1 + |x|^2)$, $i = 1, \ldots, d; j = 1, \ldots, l$, for all $t \in I$ and all x,*

(iii) *x_0 be a random variable on (Ω, Σ, P) such that $E\{|x_0|^4\} < \infty$,*

(iv) *$B = \{\beta(t) : t \in I\}$ be any Wiener process over (Ω, Σ, P).*

Then the sequence of solutions relative to B of (1.2) *for $n = 1, 2, \ldots$ is tight in the space of continuous R^d-valued functions, $C(I)$.*

REMARKS. Conditions (i) and (ii) imply that $a^{(n)}$ and $\sigma^{(n)}$ are smooth functions of x for fixed t in I. This, together with (iii), assures the existence of unique solutions, $x^{(n)}$, of (1.2) [1, Chapter 3, Theorem 3]. For a discussion of tightness in $C(I)$ see [5].

We also note that without loss of generality we may assume the following inequalities to be valid:

$$|a^{(n)}(t, x)|^2 \leq b^2(t)(1+|x|^2),$$
(1.3)
$$|\sigma_{ij}^{(n)}(t, x)|^2 \leq b^2(t)(1+|x|^2)$$

for all $n = 1, 2, \ldots$.

LEMMA 1. *There is a constant C, independent of $n = 1, 2, \ldots$, such that*

$$E\{|x^{(n)}(t)|^4\} \leq C$$

for all $t \in I$.

Proof. For this proof we suppress the superscript (n). Let $g(t)$ be the indicator of the event $[t < \tau_N]$ where τ_N is the first hitting time relative to $\{x^{(n)}(s) : s \in I\}$ of the set $\{x \in R^d : |x| \geq N\}$. Then

$$g(t)|x(t)|^4 \leq c_1 \sum_i \left\{ |x_0|^4 + g(t)\left(\int_{t_0}^t g(s)a_i(s, x(s))\, ds \right)^4 \right.$$
$$\left. + g(t) \sum_j \left(\int_{t_0}^t g(s)\sigma_{ij}(s, x(s))\, d\beta(s) \right)^4 \right\}$$

where we have used the fact that $g(t) = g(t)g(s)$ for $t > s$ and basic properties of stochastic integrals. The constant c_1 depends only upon d and l. Having truncated we may now take expectations.

$$E\{g(t)|x(t)|^4\} \leq c_1 E|x_0|^4 + c_2 \int_{t_0}^t b^4(s)\, ds + c_2 \int_{t_0}^t b^4(s)E\{g(s)|x(s)|^4\}\, ds$$

where we have used Hölder's inequality, (1.3), and an estimate of the fourth moment of a stochastic integral as is found in Theorem 4, Chapter 2 of [1]. The constant c_2 depends upon d, l and $T - t_0$. We now apply a Gronwall-like result to conclude from the above estimate that

(1.4) $$E\{g(t)|x(t)|^4\} \leq (c_0 E\{|x_0|^4\} + 1) \exp\left\{ c_2 \int_{t_0}^t b^4(s)\, ds \right\}.$$

Since this is true for all $N > 0$ we see that Lemma 1 is proved.

LEMMA 2. *There are positive constants K and α, independent of n, such that*

(1.5) $$E\{|x^{(n)}(t_1) - x^{(n)}(t_2)|^4\} \leq K|t_1 - t_2|^{1+\alpha}$$

for all t_1, t_2 in I and all integers n.

We again suppress the superscript (n).
Proof. From (1.2) and Hölder's inequality we derive

$$|x(t_1) - x(t_2)|^4 \leq c_3 \sum_i \left(\int_{t_1}^{t_2} a_i(s, x(s))\, ds \right)^4 + c_3 \sum_{i,j} \left(\int_{t_1}^{t_2} \sigma_{ij}(s, x(s))\, d\beta \right)^4$$

where c_3 depends upon d and l. From Lemma 1 we see that the expectations in question are finite. We may therefore proceed as in the proof of Lemma 1 to derive

$$E\{|x(t_1)-x(t_2)|^4\} \leqq c_4|t_1-t_2| \int_{t_1}^{t_2} b^4(s)[1+E\{|x(s)|^4\}] \, ds$$

where c_4 depends upon d, l and $T-t_0$. Using Lemma 1 we obtain

$$(1.6) \qquad E\{|x(t_1)-x(t_2)|^4\} \leqq c_5|t_1-t_2| \int_{t_1}^{t_2} b^4(s) \, ds$$

where again c_5 does not depend upon n. But since b is in L_p, $p>4$, we can apply Hölder's inequality to derive (1.5) with $\alpha = p/(p-4)$.

Now according to a theorem of Prohorov (Theorem 12.3 in [5]; original reference is [6]) condition (iii) and Lemma 2 imply the tightness in $C(I)$ of the sequence $\{x^{(n)}(t) : t \in I\}$.

By the "Lebesgue interval" is meant $[0, 1]$ endowed Lebesgue measure.

THEOREM 2. *With the same assumption as Theorem 1 there is a subsequence of the integers, J_1, and a sequence of sample continuous stochastic processes $\{w(t) : t \in I\}$, $\{w^{(n)}(t) : t \in I\}$, $\{y(t) : t \in I\}$, $\{y^{(n)}(t) : t \in I\}$, $n \in J_1$, having the Lebesgue interval as sample space and possessing the following properties:*

(A) $\{(y^{(n)}(t), w^{(n)}(t)) : t \in I\}$ *and* $\{(x^{(n)}(t), \beta(t)) : t \in I\}$ *have the same distribution considered as random elements in* $C(I) \times C(I)$.

(B) *w and $w^{(n)}$ are Brownian motion processes.*

(C) *For $n \in J_1$, $y^{(n)}$ is a solution of (1.2) relative to the Brownian motion process, $w^{(n)}$, i.e. almost surely,*

$$(1.7) \qquad y^{(n)}(t) = y^{(n)}(t_0) + \int_{t_0}^t a^{(n)}(s, y^{(n)}(s)) \, ds + \int_{t_0}^t \sigma^{(n)}(s, y^{(n)}(s)) \, dw^{(n)}(s)$$

for all $t \in I$.

(D) *With probability one, $y^{(n)}$ converges uniformly to y and $w^{(n)}$ converges uniformly to w as $n \to \infty$, $n \in J_1$.*

Proof. According to an important theorem due to Prohorov ([6]; Theorem 6.1 of [5]) tightness of the processes $x^{(n)}$, $n \geqq 1$, implies the relative compactness of their distributions in $C(I)$. It is also clear that the joint processes $\{(x^{(n)}(t), \beta(t)) : t \in I\}$ are tight hence their distributions in $C(I) \times C(I)$ also form a relatively compact sequence. Hence, there is a subsequence of the integers, J_1, such that the distributions of the joint processes converge weakly as $n \to \infty$, $n \in J_1$, to a probability measure on $C(I) \times C(I)$. From a theorem of Skorohod [7, Theorem 3.1.1] we can then conclude that there exist the processes (y, w) and $(y^{(n)}, w^{(n)})$ mentioned in the theorem such that (A) and (D) are satisfied and the distribution of (y, w) is the limiting probability measure referred to above. That (B) is true follows from the continuity of the processes and the fact that the finite-dimensional distributions

are those of Brownian motion. The proof of (C) is elementary and we shall not give the details.

It seems reasonable to refer to the process $\{y(t) : t \in I\}$ as a generalized solution of (1.1). This is a definition of solution via a "completion" of a family of approximating solutions. We emphasize that measurability and certain growth conditions were the only assumptions made concerning the drift and diffusion coefficients in proving the existence of solutions in this sense. If we assume now that the diffusion coefficient is continuous, we can say a bit more concerning this generalized solution. For this we shall need a few ideas from real variable theory.

For a function $a: I \times R^d \to R^d$ we shall define its R^d-*essential extension*, A, as the set-valued mapping defined on $I \times R^d$ with values

$$A(t, x) = \bigcap_{\delta > 0} A^\delta(t, x)$$

where

$$A^\delta(t, x) = \bigcap_N \text{co} \{a(t, y) : |y - x| \leqq \delta, y \notin N\},$$

N ranging over all sets of zero d-dimensional Lebesgue measure and co $\{ \ \}$ denoting the closed convex hull of $\{ \ \}$. All operations and relations are to be interpreted memberwise, e.g. $\langle A(t, x), v \rangle$ denotes the set of real numbers $\langle \alpha, v \rangle$ for $\alpha \in A(t, x)$; $A(t, x) \leqq B(t, x)$ (for $d = 1$) means that $\alpha \leqq \beta$ for all $\alpha \in A(t, x)$ and $\beta \in B(t, x)$.

DEFINITION. A process $\{x(t) : t \in I\}$ is said to be a *relaxed solution* of (1.1) relative to a Wiener process $\{w(t) : t \in I\}$ if with probability one the following conditions are satisfied:

(a) $x(t_0) = x_0$.

(b) For all $v \in R^d$ and for all $t_1, t_2 \in I$, $t_1 < t_2$,

$$(1.8) \quad \langle x(t_1) - x(t_2), v \rangle \leqq \int_{t_1}^{t_2} \sup \langle A(s, x(s)), v \rangle \, ds + \left\langle \int_{t_1}^{t_2} \sigma(s, x(s)) \, dw(s), v \right\rangle.$$

The following propositions are easily checked. In both propositions we assume that a and σ are Lebesgue measurable and that, for x confined to any compact set, a and σ are dominated by functions which are integrable on I. We refer to this assumption as *Condition* L.

PROPOSITION 1. *If a relaxed solution is almost surely bounded on I then it is almost surely continuous on I.*

PROPOSITION 2. *If for all t in I, $a(t, \cdot)$ is continuous in R^d, then a relaxed solution which is almost surely continuous is a solution in the ordinary sense.*

When we say that there exists a solution (resp. relaxed solution) of (1.1) on I with a given distribution function F_0, as initial distribution, we mean that on some probability space there is a Wiener process W and a random variable x_0 having F_0 as its distribution function relative to which there is a solution (resp. relaxed solution) of (1.1) on I.

THEOREM 3. *Let* (i), (ii) *and* (iii) *be as in Theorem* 1. *Let* σ *be continuous in x for each fixed t in I. Then there is a relaxed solution of* (1.1) *whose initial distribution is that of x_0. In fact, each process $\{y(t) : t \in I\}$ of Theorem* 2 *is such a relaxed solution.*

Proof. For $t_1 - t_2$ we see that

$$
E \left| \int_{t_1}^{t_2} \sigma^{(n)}(s, y^{(n)}(s))\, dw^{(n)}(s) - \int_{t_1}^{t_2} \sigma(s, y(s))\, dw(s) \right|^2
$$

$$
(1.9) \qquad \leq E \int_{t_1}^{t_2} |\sigma^{(n)}(s, y^{(n)}(s)) - \sigma(s, y(s))|^2\, ds
$$

$$
+ E \left| \int_{t_1}^{t_2} \sigma(s, y(s))\, dw^{(n)}(s) - \int_{t_1}^{t_2} \sigma(s, y(s))\, dw(s) \right|^2
$$

by the isometric property of the Itô integral. The first term on the right $\to 0$ because of the continuity of σ in the space variables and the growth condition (ii). That the second term $\to 0$ is an easily verified fact concerning Itô integrals.

Hence there is a subsequence J_2 of J_1 such that with probability one

$$
(1.10) \qquad \int_{t_1}^{t_2} \sigma^{(n)}(s, y^{(n)}(s))\, dw^{(n)}(s) \to \int_{t_1}^{t_2} \sigma(s, y(s))\, dw(s)
$$

for an everywhere dense countable set of values of t_1, t_2 as $n \to \infty$, $n \in J_2$.

Now let $v \in R^d$ and let δ be any positive number. Then from (1.7) we see that, for $n > \delta^{-1}$, with probability one we have

$$
\langle y^{(n)}(t_2) - y^{(n)}(t_1), v \rangle
$$

$$
\leq \int_{t_1}^{t_2} \sup \langle A^\delta(s, y^{(n)}(s)), v \rangle\, ds + \left\langle \int_{t_1}^{t_2} \sigma^{(n)}(s, y^{(n)}(s))\, dw^{(n)}(s), v \right\rangle
$$

for all $t_1 \leq t_2$. But from the almost sure uniform convergence of $y^{(n)}$ to y we see that

$$
\limsup_{n \in J_2} \int_{t_1}^{t_2} \sup \langle A^\delta(s, y^{(n)}(s)), v \rangle\, ds \leq \int_{t_1}^{t_2} \sup \langle A^{2\delta}(s, y(s)), v \rangle\, ds.
$$

Since this last is true for all $\delta > 0$ we see that, with probability one,

$$
\langle y(t_2) - y(t_1), v \rangle
$$

$$
\leq \int_{t_1}^{t_2} \sup \langle A(s, y(s)), v \rangle\, ds + \limsup \left\langle \int_{t_1}^{t_2} \sigma^{(n)}(s, y^{(n)}(s))\, dw^{(n)}(s), v \right\rangle
$$

where the lim sup is taken over $n \in J_2$. But then from (1.10) we see that with probability one,

$$
(1.11) \quad \langle y(t_2) - y(t_1), v \rangle \leq \int_{t_1}^{t_2} \sup \langle A(s, y(s)), v \rangle\, ds + \left\langle \int_{t_1}^{t_2} \sigma(s, y(s))\, dw(s), v \right\rangle
$$

for all t_1, $t_2 \geqq t_1$ in an everywhere dense countable set. But then the sample continuity of the processes figuring in (1.10) allows us to conclude that it is satisfied for all t_1, $t_2 \geqq t_1$. This proves the theorem.

2. One-sided uniqueness and approximation of relaxed solutions. We now restrict our attention to the case of a constant diffusion matrix, σ. A relaxed solution in this case is a process $\{x(t) : t \in I\}$ such that, with probability one,

$$(2.1) \qquad \langle x(t_1) - x(t_2), v \rangle \leqq \int_{t_2}^{t_1} \sup \langle A(s, x(s)), v \rangle \, ds + \langle \sigma w(t_1) - \sigma w(t_2), v \rangle$$

for all $v \in R^d$ and all $t_1 \geqq t_2$ in I. A second characterization of relaxed solution is given by

PROPOSITION 3. *A necessary and sufficient condition that a process* $\{x(t) : t \in I\}$ *which is almost surely bounded on* I *be a relaxed solution of*

$$(2.2) \qquad\qquad\qquad dx = a \, dt + \sigma \, dw$$

is that, with probability one, the process $\{z(t) = x(t) - \sigma w(t), t \in I\}$ *have absolutely continuous paths and*

$$(2.3) \qquad\qquad\qquad dz(t)/dt \in A(t, x(t))$$

for almost all t *in* I.

Proof. The sufficiency is clear. On the other hand if x is a relaxed solution then it follows from (2.1) that

$$|\langle z(t_1) - z(t_2), v \rangle| \leqq \int_{t_2}^{t_1} \sup |\langle A(s, x(s)), v \rangle| \, ds$$

so that absolute continuity follows from the boundedness of the sample paths and Condition L.

We can use (2.1) to conclude

$$\langle dz(t)/dt, v \rangle \leqq \sup \langle A(t, x(t)), v \rangle$$

for all $v \in R^d$. Relation (2.3) follows from this since a compact convex set is the intersection of all half-spaces containing it.

The basic condition which we shall impose on the drift coefficient is the following:

Condition M. *For each* $N > 0$ *there is a nonnegative function* k_N, *integrable on* I, *such that*

$$(2.4) \qquad\qquad \langle a(t, x) - a(t, y), x - y \rangle \leqq k_N(t) |x - y|^2$$

if $|x| \leqq N$ *and* $|y| \leqq N$.

It is an easy exercise to prove

PROPOSITION 4. *If* a *satisfies Conditions* L *and* M *then*

$$\langle A(t, x) - A(t, x), x - y \rangle \leqq k_N(t) |x - y|^2, \qquad |x| \leqq N, \quad |y| \leqq N.$$

We remind the reader that this is a set inequality the sense of which is that $\langle x'-y', x-y \rangle \leqq k_N(t)|x-y|^2$ for all $x' \in A(t, x)$ and $y' \in A(t, y)$.

THEOREM 4. *Let a satisfy Conditions L and M. Let x and y be two relaxed solutions of (2.2) relative to the same Wiener process. Then if $P\{x(t_0) = y(t_0)\} = 1$ then*

$$P\{x(t) = y(t) : t_0 \leqq t \leqq T\} = 1.$$

Proof. Let $\Delta(t) = x(t) - y(t)$ for t in I. Then since σ is constant we see from Proposition 3 that $\{\Delta(t) : t \in I\}$ has absolutely continuous paths (a.s.) and

$$\dot{\Delta}(t) \in A(t, x(t)) - A(t, y(t)).$$

Therefore, for any $N > 0$,

$$\tfrac{1}{2} d|\Delta(t)^2|/dt \leqq \langle \dot{\Delta}(t), \Delta(t) \rangle$$
$$\leqq \sup \langle A(t, x(t)) - A(t, y(t)), x(t) - y(t) \rangle \leqq k_N(t)|\Delta(t)|^2$$

for all $t < \tau_N = $ smaller of the first hitting times for x and y of the set $\{x : |x| \geqq N\}$. (We define $\tau_N = -\infty$ if $|x(t_0)|$ or $|y(t_0)|$ exceeds N.) Integrating this simple inequality results in

$$|\Delta(t)|^2 \leqq |\Delta(t_0)|^2 \exp\left(2 \int_{t_0}^t k_N(s)\, ds\right)$$

for $t_0 \leqq t < \tau_N$. Since $\Delta(t_0) = 0$ almost surely, the truth of the theorem follows immediately since $P\{\tau_N = -\infty\} \to 0$ as $N \to \infty$ (sample continuity of relaxed solutions!).

THEOREM 5. *Let (i) $c^{(n)} : R^{d+1} \to R^d$ satisfy Condition L and $|c^{(n)}(t, x)|^2 \leqq b^2(t)(1 + |x|^2)$ for all $n \geqq 1$ where b is as in Theorem 1;*
   (ii) $a : R^{d+1} \to R^d$ satisfy Conditions L, M and $|a(t, x)|^2 \leqq b^2(t)(1 + |x|^2)$;
   (iii) $c^{(n)}$ converge to a in $L^2_{\text{loc}}(R^{d+1})$ as $n \to \infty$;
   (iv) $s^{(n)}, \sigma$ be $d \times d$ matrices, $s^{(n)} \to \sigma$ as $n \to \infty$;
   (v) each $s^{(n)}$ be positive definite, i.e. $\langle \xi, s^{(n)}\xi \rangle > 0$ for $\xi \neq 0$ in R^d;
   (vi) $y^{(n)}$ be a solution of

(2.5) $$dy^{(n)} = c^{(n)}\, dt + s^{(n)}\, dw$$

where w is a d-dimensional Wiener process;
   (vii) x be the unique solution of (2.2) relative to the same Wiener process;
   (viii) $y^{(n)}(t_0)$ converge in probability to $x(t_0)$;
   (ix) $E\{|x(t_0)|^2\}$ and $E\{|y^{(n)}(t_0)|^2\}$, $n \geqq 1$, be bounded by a constant K.
Then for all t in I, $y^{(n)}(t)$ converges in probability to $x(t)$.

Proof of Theorem 5. We first note that for any $\varepsilon > 0$ and any $N > 0$

$$P\{|y^{(n)}(t) - x(t)| > \varepsilon\} < E\{|x(t) - y^{(n)}(t)|^2 \psi_N^{(n)}\}/\varepsilon^2 + P\{\psi_N^{(n)} = 0\}$$

where $\psi_N^{(n)}$ is the indicator function of $\{\sup_{t \in I} |y^{(n)}(t)| + \sup_{t \in I} |x(t)| < N\}$. But

$$P\{\psi_N^{(n)} = 0\} \leqq P\left\{\sup_t |x(t)| > N/2\right\} + P\left\{\sup_t |y^{(n)}(t)| \geqq N/2\right\}$$

and since $c^{(n)}$ and a satisfy the same growth condition (i.e. (i) and (ii)) it easily follows (as in Lemma 1) that $E|y^{(n)}(t)|^2$ and $E|x(t)|^2$ are bounded by a constant k_0 depending on b, $T-t_0$, $|\sigma|$ and k but independent of n. Therefore there is an $N > 0$ such that $P\{\psi_N^{(n)} = 0\}$ is less than $\varepsilon/2$. For this N, which we fix, we have

$$(2.6) \qquad P\{|y^{(n)}(t) - x(t)| > \varepsilon\} < E\{|x(t) - y^{(n)}(t)|\psi_N^{(n)}\}/\varepsilon^2 + \varepsilon/2.$$

Now let $z(t) = x(t) - y + (s - \sigma)(w(t) - w(t_0))$ for t in I where for the moment we have suppressed superscripts. From Proposition 3 we conclude that $\{z(t) : t \in I\}$ has absolutely continuous paths and that

$$(2.7) \qquad \dot{z}(t) \in A(t, x(t)) - c(t, y)$$

for almost all t in I. Obviously then $\{z(t)\psi_N : t \in I\}$ also has absolutely continuous sample paths and

$$(d/dt)\psi_N z(t) = \psi_N \dot{z}(t) \in \psi_N A(t, x(t)) - c(t, y(t)).$$

Now

$$(2.8) \qquad \begin{aligned} \langle \psi_N \dot{z}(t), z(t) \rangle &= \langle \psi_N \dot{z}(t), x(t) - y(t) \rangle + \langle \psi_N \dot{z}(t), (s - \sigma)(w(t) - w(t_0)) \rangle \\ &= Q_1 + Q_2, \end{aligned}$$

where from (2.7) we see that

$$\begin{aligned} Q_1 &\leqq \sup \langle A(t, x(t)) - c(t, y(t)), x(t) - y(t) \rangle \psi_N \\ &\leqq \sup \langle A(t, x(t)) - a(t, y(t)), x(t) - y(t) \rangle \psi_N \\ &\quad + \langle a(t, y(t)) - c(t, y(t)), x(t) - y(t) \rangle \psi_N \end{aligned}$$

the supremum being taken over the elements of A. Now using Condition M and Proposition 4 we have

$$\begin{aligned} Q_1 &\leqq k_N(t)|x(t) - y(t)|^2 \psi_N + 2N\psi_N|a(t, y(t)) - c(t, y(t))| \\ &\leqq 2k_N(t)|z(t)|^2 \psi_N + 2k_N(t)|s - \sigma|^2|w(t) - w(t_0)|^2 \\ &\quad + 2N\psi_N|a(t, y(t)) - c(t, y(t))|. \end{aligned}$$

Hence,

$$\begin{aligned} E\{Q_1\} &\leqq 2k_N(t)E\{|z(t)|^2\psi_N\} + 2k_N(t)|s - \sigma|^2(t - t_0) \\ &\quad + 2NE\{\psi_N|a(t, y(t)) - c(t, y(t))|\}. \end{aligned}$$

But,

$$E\{\psi_N|a(t, y(t)) - c(t, y(t))|\} \leqq \int_{|\xi| \leqq N} |a(t, \xi) - c(t, \xi)|p(t, \xi)\, d\xi$$

where $p(t, \xi)$ is the density function of the distribution $P\{y(t) < \xi\}$. It is here that we make use of the nondegeneracy of the processes y (condition (iv)). We must

have that the distribution of y be absolutely continuous relative to Lebesgue measure in R^d. We therefore have

$$E\{Q_1\} \leqq 2k_N(t)E\{|z(t)|^2\psi_N\} + 2k_N(t)|s-\sigma|^2(t-t_0) + 2N^2\|a-c\|_N$$

where $\|\cdot\|_N$ denotes the L_2 norm over $I \times \{\xi : |\xi| \leqq N\}$.

On the other hand, using Cauchy's inequality we obtain

$$E\{Q_2\} \leqq (E\{|\dot z(t)|^2\psi_N\})^{1/2}(E\{|s-\sigma|^2|w(t)-w(t_0)|^2\psi_N\})^{1/2}$$

while from (2.7) and (i) we see that

$$|\dot z(t)|^2\psi_N \leqq 2b^2(t)(1+|N|^2).$$

Hence

$$E\{Q_2\} \leqq b(t)|s - \sigma|(1 + N^2)^{1/2}(2|t - t_0|)^{1/2}.$$

Using these estimates in (2.8) we obtain

$$E\{\tfrac{1}{2}(d/dt)|z(t)|^2\psi_N\} < 2k_N(t)E\{|z(t)|^2\psi_N\} + 2N^2\|a-c\|_N + \tfrac{1}{2}|s-\sigma|g_N(t)$$

where

$$g_N(t) = 4k_N(t)(t-t_0)|s-\sigma| + b(t)(1+N^2)^{1/2}(8|t-t_0|)^{1/2}.$$

But using the Fubini theorem, (2.9), (2.6) and the dominated convergence theorem it is easy to see that

$$\tfrac{1}{2}(d/dt)E\{|z(t)|^2\psi_N\} = E\{\tfrac{1}{2}(d/dt)|z(t)|^2\psi_N\}.$$

To this we apply a standard result on differential inequalities [8, p. 26] to conclude that for $t \geqq t_0$ (note the one-sidedness!)

$$E\{|z(t)|^2\psi_N\} \leqq E\{|z(t_0)|^2\psi_N\}k_1 + \|a-c\|_Nk_2 + |s-\sigma|k_3,$$

where

$$(2.9) \qquad \begin{aligned} k_1 &= 4\int_I k_N(n)\,dn, \\ k_2 &= 4N^2k_1|T-t_0|, \\ k_3 &= k_1\int_I g_N(n)\,dn. \end{aligned}$$

But $E\{|x(t)-y(t)|^2\psi_N\} \leqq 2E\{|z(t)|^2\psi_N\} + 2|s-\sigma|^2E\{|w(t)-w(t_0)|^2\}$ and $z(t_0) = x(t_0) - y(t_0)$ so that

$$(2.10) \qquad E\{|x(t)-y(t)|^2\psi_N\} \leqq E\{|x(t_0)-y(t_0)|^2\psi_N\}k_4 + \|a-c\|_Nk_5 + |s-\sigma|k_6$$

where $k_4 = 2k_1$, $k_5 = 2k_2$ and $k_6 = 2k_3 + 4(|\sigma|^2+1)|T-t_0|$. The important observation is of course that $k_1, \ldots, k_6$ are independent of n. This being the case we see that there is an n_0 such that

$$E\{|x(t)-y^{(n)}(t)|^2\psi_N^{(n)}\} < \tfrac{1}{2}\varepsilon^3$$

for all $n > n_0$. Hence, from (2.6) we see that

$$P\{|x(t) - y^{(n)}(t)| > \varepsilon\} < \varepsilon$$

for all $n > n_0$. Since ε was arbitrary this proves Theorem 5.

REFERENCES

1. A. V. Skorohod, *Studies in the theory of random processes*, Izdat. Kiev. Univ., Kiev, 1961; English transl., Addison-Wesley, Reading, Mass., 1965. MR **32** #3082a, b.

2. A. F. Filippov, *Differential equations with discontinuous right-hand side*, Mat. Sb. **51** (93) (1960), 99–128; English transl., Amer. Math. Soc. Transl. (2) **42** (1964), 199–231. MR **22** #4846.

3. K. Itô and H. P. McKean, Jr., *Diffusion processes and their sample paths*, Die Grundlehren der math. Wissenschaften, Band 125, Academic Press, New York; Springer-Verlag, Berlin, 1965. MR **33** #8031.

4. K. Yosida, *Functional analysis*, Die Grundlehren der math. Wissenschaften, Band 123, Academic Press, New York; Springer-Verlag, Berlin, 1965. MR **31** #5054.

5. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR **38** #1718.

6. Ju. V. Prohorov, *Convergence of random processes and limit theorems*, Teor. Verojatnost. i Primenen. **1** (1956), 157–214.

7. A. V. Skorohod, *Limit theorems for stochastic processes*, Teor. Verojatnost. i Primenen. **1** (1956), 289–319 = Theor. Probability Appl. **1** (1956), 261–289. MR **18**, 943.

8. P. Hartman, *Ordinary differential equations*, Wiley, New York, 1964. MR **30** #1270.

9. D. Stroock and S. R. S. Varadhan, *Diffusion processes with continuous coefficients*. I, Comm. Pure Appl. Math. **22** (1969), 345–400.

TULANE UNIVERSITY,
NEW ORLEANS, LOUISIANA 70118