THE CAUCHY PROBLEM FOR DEGENERATE PARABOLIC EQUATIONS WITH DISCONTINUOUS DRIFT

BY

EDWARD D. CONWAY

ABSTRACT. The coefficient of the gradient is allowed to be discontinuous but is assumed to satisfy a "one-sided" Lipschitz condition. This condition insures the pathwise uniqueness of the underlying Markov process which in turn yields the existence of a unique stable generalized solution of the parabolic equation. If the data is Lipschitz continuous, then so is the solution.

Introduction. Consider the backward Cauchy problem,

$$u_t + \langle a, \nabla u \rangle + \langle D\nabla, \nabla u \rangle = bu + c \qquad \text{for } t < T,$$

$$u = \phi \qquad \text{for } t = T.$$

If D is nonnegative definite but not positive definite, i.e. if the equation is degenerate parabolic, then no matter how smooth ϕ might be, unless a is differentiable a continuous solution need not exist even locally. A simple example (discussed by Gelfand in [6]) is the one dimensional totally degenerate equation where $D \equiv 0$, $a(x) = \sqrt{|x|}$ and $b \equiv c \equiv 0$. However, in an earlier paper [1], I showed that there was a stable Lipschitzian solution to the first order problem ($D \equiv 0$) if a satisfies a "one-sided" Lipschitz condition,

$$\langle a(t,x) - a(t,y), x - y \rangle \leqslant k|x - y|^2.$$

This condition allows a to be discontinuous (e.g. in one dimension it is satisfied by any nonincreasing function of x), but it insures the uniqueness (in the direction of increasing t) of the underlying characteristic curves. This uniqueness, it turns out, is the crucial issue, not the smoothness. In this paper, I show that the same considerations are valid for the second order equation. The uniqueness in question is now the pathwise uniqueness of the underlying characteristic Markov process. A detailed statement of the results is given below but they include the following. Let D be C^2 and $O(|x|^2)$ at infinity. Let a satisfy the above condition and be $O(|x|)$. Let ϕ, b and c be bounded and continuous. Then there is a unique stable continuous generalized solution. By stable I mean that it is the uniform limit on compacta of the bounded solutions of uniformly parabolic equations which approximate the given equation. If b, c and ϕ satisfy a Lipschitz condition,

then so does u. The stability is important since there are unstable bounded Lipschitzian solutions with the same initial data (I gave an example in [1]). But now let us be a bit more detailed.

Write the equation in the form,

$$(1) \qquad u_t + a_j(t,x)u_{x_j} + \tfrac{1}{2}\sigma_{ik}(t,x)\sigma_{jk}(t,x)u_{x_ix_j} = b(t,x)u + c(t,x).$$

The terminal condition is

$$(2) \qquad\qquad\qquad u(T,x) = \phi(x).$$

Notation. $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$; $\langle x,y\rangle = x_jy_j = \sum x_jy_j$ (summation convention); σ is a $d \times n$ matrix-valued function which I will assume satisfies a local Lipschitz condition in the space variables. If D is C^2 then σ may be taken to be the symmetric square root of $2D$ (cf. [10]). The functions a, σ, b and ϕ are defined throughout $I \times \mathbf{R}^d$ where $I = [T_0, T]$.

I will approach (1)–(2) through an underlying Markov process. This is the natural generalization of the elementary study of the first order equation

$$(3) \qquad\qquad v_t + a_j(t,x)v_{x_j} = b(t,x)v + c(t,x)$$

based upon the theory of characteristics [5]. Basic to that theory is the ordinary differential equation

$$(4) \qquad\qquad\qquad dx = a(t,x)\,dt.$$

Let $z_{t,x}$ denote the *characteristic* through the point (t,x), i.e.

$$z'_{t,x}(s) = a(s, z_{t,x}(s))$$

for all s in an interval containing t and $z_{t,x}(t) = x$. Under suitable conditions on the function a there is one and only one such absolutely continuous function and it is defined for all $s, t \leqslant s \leqslant T$. In such a case, if ϕ, b and c are smooth enough the unique solution of (3)–(2) is given by

$$(5) \qquad\qquad\qquad v(t,x) = H_t(z_{t,x})$$

where H_t is the functional with value

$$(6) \qquad H_t(\alpha) = \phi(\alpha(T))\exp\left(-\int_t^T b(s,\alpha(s))\,ds\right)$$
$$- \int_t^T c(\theta, \alpha(\theta))\exp\left(-\int_t^\theta b(s,\alpha(s))\,ds\right)\,d\theta$$

at any continuous function $\alpha : [t, T] \to \mathbf{R}^d$. In the classical theory these statements are justified when a, b, c and ϕ are continuously differentiable and when a is $O(|x|)$ as $|x| \to \infty$. In the earlier paper [1], I carried out this program under less stringent conditions. These results are contained as special cases in the

treatment given below of the second order problem (1)–(2), i.e. the results proved below contain those of [1] when $\sigma \equiv 0$.

In the second order case (1)–(2), we replace (4) by the stochastic characteristic equations

$$(7) \qquad\qquad dx = a(t, x)\, dt + \sigma(t, x)\, d\beta(t)$$

where $\{\beta(t) : t \in I\}$ is an n-dimensional Wiener process [9]. Let $\{\zeta_{t,x}(s) : t \leqslant s \leqslant T\}$ denote the solution of the Itô equation (7) with the property that

$$(8) \qquad\qquad \zeta_{t,x}(t) = x \quad \text{a.s.}$$

Such processes are known to exist uniquely under suitable assumptions on a and σ, e.g. if a and σ are continuously differentiable in x and grow no more than linearly at infinity (cf. [7] or [9]). If we now also assume sufficient smoothness and growth properties of ϕ, b and c then the unique solution (among functions with suitably restricted growth) of (1)–(2) is given by

$$(9) \qquad\qquad u(t, x) = E H_t(\zeta_{t,x}) = \int_\Omega H_t(\zeta_{t,x}(\omega)) P(d\omega)$$

where (Ω, Σ, P) is the underlying probability space supporting β and ζ. In other words the data ϕ affects u in the same way as in the case $\sigma \equiv 0$ but now the effects come from many different points on the line $t = T$ and are then averaged according to a particular measure upon the space of paths joining (t, x) to that line. This is well known; see, for example, [8]. (9) can actually be derived under suitable assumptions from Itô's chain rule; an example of this is given in [4].

Now, the relationship of (9) to (1)–(2) is usually discussed under two general classes of assumptions. If equation (1) is assumed *strictly parabolic*, then continuity and boundedness are all that is required of σ while a is assumed to be merely bounded and measurable. This is carried out for $b \equiv c \equiv 0$ by Stroock and Varadhan in their impressive paper [12]. When degeneracy is allowed (i.e. $\sigma x = 0$ for some $x \neq 0$) the representation has been proved (cf. [7], [8] and [13]) only under the assumption that a and σ are smooth enough to insure the pathwise uniqueness of solutions of (7). *But it is the uniqueness that is important, not the smoothness.* For in the discussion given below I prove that (9) is a highly regular and stable generalized solution even when a is discontinuous as long as it satisfies a dissipative-like condition (condition 2° below) which insures pathwise uniqueness of the characteristic processes. Since a need not be continuous, however, the characteristic processes are solutions of (7) in the relaxed sense introduced by the author in [2] and [3].

The detailed statements of results together with their proofs are given in the sections below.

1. The basic assumptions are as follows:

1° The functions a and σ are measurable and, for some constant M, satisfy

$$|a(t,x)|^2 + |\sigma(t,x)|^2 \leqslant M(1 + |x|^2)$$

throughout $I \times \mathbf{R}^d$.

2° For each $N > 0$ there is a number k_N such that

$$(10) \qquad\qquad \langle a(t,x) - a(t,y), x - y \rangle \leqslant k_N |x - y|^2$$

and

$$(11) \qquad\qquad |\sigma(t,x) - \sigma(t,y)|^2 \leqslant k_N |x - y|^2$$

for all t in I and $|x| \leqslant N$, $|y| \leqslant N$.

3° The real valued functions b and c are continuous on $I \times \mathbf{R}^d$; both $-b$ and $|c|$ are bounded there.

4° The function ϕ is bounded and continuous on $\mathbf{R}^d$.

Let $\zeta_{t,x} = \{\zeta_{t,x}(s) : t \leqslant s \leqslant T\}$ denote the unique *relaxed solution* of (7) starting at (t,x), i.e. satisfying (8). A complete discussion of these processes including proofs of existence and uniqueness under conditions 1° and 2° is given in [2] and [3]. Since the paths of $\zeta_{t,x}$ are continuous, we see that 3° and 4° insure that

$$(9) \qquad\qquad u(t,x) = EH_t(\zeta_{t,x})$$

is well defined and in fact bounded throughout $I \times \mathbf{R}^d$. We shall refer to this function u as a generalized solution of (1)–(2). The first and most generally valid fact justifying this term is that it is the uniform (on compacta) limit of bounded solutions of nondegenerate equations whose coefficients converge to those of (1). To be precise consider the sequence of strictly parabolic equations

$$(12) \qquad\qquad u_t + a_j^{(m)} u_{x_j} + \tfrac{1}{2}\sigma_{ik}^{(m)} \sigma_{jk}^{(m)} u_{x_i x_j} = bu + c$$

where

A1. For all $m > 0$, $a^{(m)}$ and $\sigma^{(m)}$ are C^∞ and satisfy 1° with the constant M independent of m. For each m, $\sigma^{(m)}\sigma^{(m)*}$ is positive definite.

A2. $a^{(m)} \to a$ in $L^2_{\mathrm{loc}}(I \times \mathbf{R}^d)$ as $m \to \infty$.

A3. For each t in I, $\sigma^{(m)}(t, \cdot) \to \sigma(t, \cdot)$ uniformly on compact sets of $\mathbf{R}^d$ in such a way that, for all $N > 0$,

$$\int_I \max_{|x| \leqslant N} |\sigma^{(m)}(t,x) - \sigma(t,x)|^2 \, dt \to 0 \quad \text{as } m \to \infty.$$

Note that the choice of convergence in A2 and A3 allows a to be discontinuous and σ to be discontinuous in t.

Theorem 1. *Let conditions 1° through 4° and A1 through A3 be satisfied. Let $u^{(m)}$ be the unique bounded solution of (12) in $I \times \mathbf{R}^d$ such that $u^{(m)}(T,x) = \phi(x)$. Let u be as in (9). Then the sequence $\{u^{(m)}\}$ converges uniformly on compact sets to u.*

I wish to emphasize that the sequence $\{u^{(m)}\}$ *converges* and is not merely compact, i.e. u is a *stable* generalized solution.

Proof. Let $\zeta_{t,x}^{(m)} = \{\zeta_{t,x}^{(m)}(s) : t \leqslant s \leqslant T\}$ be the unique solution of the Itô equation

$$(13) \qquad\qquad dx = a^{(m)} dt + \sigma^{(m)} d\beta$$

satisfying

$$\zeta_{t,x}^{(m)}(t) = x \quad \text{a.s.}$$

That A1 is sufficient for the existence and uniqueness of these processes is proved in [9]. Since $u^{(m)}$ is the bounded solution we see (cf. [8] or [13]) that

$$(14) \qquad\qquad u^{(m)}(t, x) = E\{H_t(\zeta_{t,x}^{(m)})\}.$$

Now because of 3° and the fact that ϕ is bounded we see that for any continuous $\alpha : [t, T] \to \mathbf{R}^d$, $|H_t(\alpha)| \leqslant K_0$ where K_0 may be chosen independently of α and t. Thus, if $E_N(t, x)$ denotes the indicator random variable of the event,

$$|\zeta_{t,x}(s)| + |\zeta_{t,x}^{(m)}(s)| \leqslant N \quad \text{for all } s, t \leqslant s \leqslant T,$$

then

$$|u(t, x) - u^{(m)}(t, x)| \leqslant E(|H_t(\zeta_{t,x}) - H_t(\zeta_{t,x}^{(m)})|E_N(t, x)) + 2K_0 P(E_N(t, x) = 0).$$

But from 1° and A1 it follows that there is a constant K_1 which is independent of m, x, s and t in I such that

$$E|\zeta_{t,x}(s)|^2 + E|\zeta_{t,x}^{(m)}(s)|^2 \leqslant K_1(1 + |x|^2).$$

(This is well known; Lemma 1 of [2] gives a quick proof.) Thus for any compact set $Q \subset I \times \mathbf{R}^d$ and for any $\varepsilon > 0$ there is an N such that, for all $m > 0$,

$$|u(t, x) - u^{(m)}(t, x)| \leqslant E(|H_t(\zeta_{t,x}) - H_t(\zeta_{t,x}^{(m)})|E_N(t, x)) + \varepsilon,$$

for all $(t, x) \in Q$. But now from the continuity of b, c and ϕ it follows that there is a $\delta = \delta(\varepsilon, N)$ such that

$$|H_t(\alpha) - H_t(\gamma)| < \varepsilon$$

if $|\alpha(s)| + |\gamma(s)| \leqslant N$ and $|\alpha(s) - \gamma(s)| < \delta$ for all $s \in [t, T]$. Hence for all (t, x) in Q and all $n > 0$ we have

$$|u(t, x) - u^{(m)}(t, x)| \leqslant 2K_0 P(A_\delta(t, x)) + 2\varepsilon$$

where $A_\delta(t, x)$ is the event where $|\zeta_{t,x}(s) - \zeta_{t,x}^{(m)}(s)| > \delta$ for some s in $[t, T]$. But from Theorem 2 of [3] it follows that $P(A_\delta) \to 0$ as $m \to \infty$. An examination of the proof of that theorem (cf. especially equation (9) of [3]) shows that the convergence is uniform for (t, x) in Q. This completes the proof of Theorem 1.

The previous theorem shows that u is continuous and is stable under regularization, i.e. the sequence $a^{(m)}$ can be obtained from a by "mollifying" (convolution with smooth kernel) and $\sigma^{(m)}$ can be obtained by mollifying and adding $m^{-1} \cdot$ identity matrix to D. But u is also stable with respect to arbitrary perturbation of the coefficients as long as we stay within the class of drift and diffusion coefficients satisfying 1° and 2°, i.e. if, for each m, $a^{(m)}$ and $\sigma^{(m)}$ satisfy 1° and 2° with M and k_N independent of m and if they converge to a and σ in the sense of A2 and A3 then $u^{(m)} \to u$ uniformly on compacta where $u^{(m)}$ is given by (14) but the $\zeta^{(m)}$ are now relaxed processes. This follows from Theorem 3 of [3] just as the previous theorem followed from Theorem 2 of that paper.

2. **Regularity of the solution.** The following lemma provides the key estimate for this section.

Lemma 1. *Let* (i) *1° and 2° be satisfied,*
(ii) $|\sigma(t, x)|^2 \leqslant M(1 + |x|^2)^\mu$ *in* $I \times \mathbf{R}^d$ *where* $0 \leqslant \mu \leqslant 1$,
(iii) $k_N = O(N^{2\nu})$ *as* $N \to \infty$ *where* $\nu < 1$.
Then (a) *if* $\mu + \nu < 1$ *we have*

$$(15) \qquad E|\zeta_{t,x}(s) - \zeta_{t',x'}(s)| \leqslant Q_R|x - x'| + S_R|t - t'|^{1/2}$$

for all x, x' *such that* $|x| \leqslant R$, $|x'| \leqslant R$ *and all* $s, t \vee t' \leqslant s \leqslant T$. *The constants Q_R and S_R depend upon R, $|I|$, d, n, M, μ and ν.*

(b) *If $\nu = 0$ then μ can equal 1 and Q_R may be chosen independently of R.*

Remarks. Part (b) is well known for smooth a and σ. Part (a) seems to be new even in the smooth case. It is certainly not unexpected, however.

Proof. We first prove part (a). Let $x(s) = \zeta_{t,x}(s)$ and $y(s) = \zeta_{t',x'}(s)$. If $t \geqslant t'$ then from the definition of relaxed solution and Itô's lemma (see equation (5) of [3]) we have

$$
\begin{aligned}
|x(s) - y(s)|^2 = {} & |x(t) - y(t)|^2 + 2 \int_t^s \langle x(\theta) - y(\theta), \Delta(\theta) \rangle \, d\theta \\
(16) \qquad & + \sum_{i,j} \int_t^s |\sigma_{ij}(\theta, x(\theta)) - \sigma_{ij}(\theta, y(\theta))|^2 \, d\theta \\
& + 2 \int_t^s \langle x(\theta) - y(\theta), [\sigma(\theta, x(\theta)) - \sigma(\theta, y(\theta))] \, d\beta(\theta) \rangle
\end{aligned}
$$

where $\Delta(\theta)$ is equal to $a(\theta, x(\theta)) - a(\theta, y(\theta))$ if a is continuous in x but in any case

$$\langle x(\theta) - y(\theta), \Delta(\theta) \rangle \leqslant k_N|x(\theta) - y(\theta)|^2$$

if $|x(\theta)| \leqslant N$ and $|y(\theta)| \leqslant N$. Now let τ_N be the first time either $|x(\theta)| = N$ or $|y(\theta)| = N$. If in (16) we replace t by $t \wedge \tau_N$ and s by $s \wedge \tau_N$ then the Itô integral has mean zero because $|\sigma(t, x)| = O(|x|)$ and x and y have bounded second moments. Hence,

$$
\begin{aligned}
(17) \quad E|x(s \wedge \tau_N) - y(s \wedge \tau_N)|^2 &\leqslant E|x(t \wedge \tau_N) - y(t \wedge \tau_N)|^2 \\
&+ (nd + 2)E \int_{t \wedge \tau_N}^{s \wedge \tau_N} k_N |x(\theta) - y(\theta)|^2 \, d\theta.
\end{aligned}
$$

If we now let $g_N(\theta)$ be the indicator function of $[\theta \leqslant \tau_N]$ this last estimate implies

$$
\begin{aligned}
E\{|x(s) - y(s)|^2 g_N(s)\} &\leqslant E\{|x(t) - y(t)|^2 g_N(t)\} \\
&+ (nd + 2)k_N \int_t^s E\{|x(\theta) - y(\theta)|^2 g_N \theta\} d\theta,
\end{aligned}
$$

so that from Gronwall's lemma we have

$$
\begin{aligned}
(18) \quad E\{|x(s) - y(s)|^2 g_N(s)\} &\leqslant E\{|x(t) - y(t)|^2 g_N(t)\} e^{(nd+2)k_N(s-t)} \\
&\leqslant E\{|x(t) - y(t)|^2\} e^{(nd+2)k_N(s-t)}.
\end{aligned}
$$

Now let $J_N(s)$ be the indicator function of the event

$$
\left[N - 1 < \sup_{\theta \leqslant s} |x(\theta)| + \sup_{\theta \leqslant s} |y(\theta)| \leqslant N \right].
$$

We have

$$
\begin{aligned}
E|x(s) - y(s)| &= \sum_{N=1}^{\infty} E\{|x(s) - y(s)|J_N(s)\} \\
&= \sum_N E\{|x(s) - y(s)|g_N J_N(s)\} \qquad (J_N = 1 \Rightarrow g_N = 1) \\
&\leqslant \sum_N (E|x(s) - y(s)|^2 g_N(s))^{1/2} (E(J_N(s)))^{1/2},
\end{aligned}
$$

so that from (18) we have

$$
\begin{aligned}
(19) \quad E|x(s) - y(s)| &\leqslant (E|x(t) - y(t)|^2)^{1/2} \\
&\cdot \sum_N \exp((nd + 2)k_N t/2)(P[J_N(s) = 1])^{1/2}.
\end{aligned}
$$

But

$$
\begin{aligned}
(20) \quad P(J_N(s) = 1) &\leqslant P\left(\sup_I |x(\theta)| > \frac{N-1}{2} \right) + P\left(\sup_I |y(\theta)| > \frac{N-1}{2} \right) \\
&\leqslant Q_1 N \exp(-Q_2 N^{2-2\mu}).
\end{aligned}
$$

This is because $|a(t, x)| = O(|x|)$ and $|\sigma(t, x)|^2 = O(|x|^{2\mu})$; (20) follows easily from the estimate on line three from the bottom, p. 93 of [9]. The constants Q_1 and Q_2

depend upon d, n, M, $|I|$ and R. From (20) and the fact that $k_N = O(N^{2\nu})$, $\nu + \mu < 1$, it follows that the series in (19) converges so that

$$(21) \qquad E|x(s) - y(s)| \leqslant Q_3(E|x(t) - y(t)|^2)^{1/2}.$$

But

$$(22) \qquad \begin{aligned} E|x(t) - y(t)|^2 &= E|x - \zeta_{t',x'}(t)|^2 \\ &\leqslant 2|x - x'|^2 + 2E|x' - \zeta_{t',x'}(t)|^2. \end{aligned}$$

But because of 1° it follows that

$$(23) \qquad E|\zeta_{t',x'}(t) - x'|^4 \leqslant \text{constant } (1 + |x'|^4)|t - t'|^2.$$

This is well known for solutions of Itô equations (cf. [7, p. 398]). The constant depends upon M, $|t - t'|$, d and n. That (23) holds for relaxed solutions under condition 2° follows from the fact that relaxed solutions are the uniform limit in probability of solutions in the usual sense having the same growth properties, i.e. satisfying condition 1° with the same M. This is proved in [3, Theorem 2]. The estimate (15) now follows from (21), (22) and (23). This completes the proof of part (a).

Now if $\nu = 0$ we may assume k_N to be independent of N. Then from (16) we immediately derive

$$E|x(s) - y(s)|^2 \leqslant E|x(t) - y(t)|^2 + (2 + nd) \int_t^s kE|x(\theta) - y(\theta)|^2 \, d\theta$$

so that

$$(24) \qquad E|x(s) - y(s)|^2 \leqslant e^{(2+nd)k(s-t)} E|x(t) - y(t)|^2.$$

This estimate together with (22) and (23) establishes the truth of part (b).

We now impose a stronger condition on b, c and ϕ.

5° For each $N > 0$ there is a number K_N such that

$$|b(t,x) - b(t,y)| + |c(t,x) - c(t,y)| + |\phi(x) - \phi(y)| \leqslant K_N|x - y|$$

for all t in I if $|x| \leqslant N$ and $|y| \leqslant N$.

Theorem 2. *Let* (i) 1° *through* 5° *be satisfied,*
(ii) $|\sigma(t,x)|^2 \leqslant M(1 + |x|^2)^\mu$, $0 \leqslant \mu \leqslant 1$,
(iii) $k_N = O(N^{2\nu})$, $N \to \infty$, $\nu < 1$,
(iv) $\log K_N = O(N^{2\lambda})$, $N \to \infty$, $\lambda < 1$.
Then (a) *if* $\mu + \nu < 1$ *and* $\mu + \lambda < 1$ *we have*

$$(25) \qquad |u(t,x) - u(t',y)| \leqslant P_R|x - x'| + L_R|t - t'|^{1/2}$$

for all t, t' in I if $|x| \leqslant R$ and $|x'| \leqslant R$. P_R and L_R depend upon R, d, n, $|I|$, M, μ, ν and λ. On the other hand,

(b) *if $\nu = 0$ and $\lambda = 0$ then P_R may be chosen independently of R, i.e. for all t in I and all x, x' in $\mathbf{R}^d$,*

$$(26) \qquad |u(t, x) - u(t, x')| \leqslant P|x - x'|.$$

Proof. Because of 3°, 4° and 5° we see that if $t' \leqslant t \leqslant T$

$$(27) \quad |H_t(\alpha) - H_{t'}(\gamma)| \leqslant K_N \overline{P} \left\{ |\alpha(T) - \gamma(T)| + \int_t^T |\alpha(s) - \gamma(s)| \, ds \right\} + \overline{P}|t - t'|$$

for all curves $\alpha : [t, T] \to \mathbf{R}^d$, $\gamma : [t', T] \to \mathbf{R}^d$ provided that $|\alpha(s)| \leqslant N$ and $\gamma(s)| \leqslant N$ for all $s \leqslant T$. $\overline{P}$ does not depend upon N. Hence with J_N and g_N as in the proof of Lemma 1 we have

$$|u(t, x) - u(t', x')| \leqslant E|H_t(\zeta_{t,x}) - H_{t'}(\zeta_{t',x'})|$$

$$= \sum_N E(J_N(T)|H_t(\zeta_{t,x}) - H_{t'}(\zeta_{t',x'})|)$$

$$\leqslant \overline{P} \sum_N K_N E(J_N(T)|\zeta_{t,x}(T) - \zeta_{t',x'}(T)|)$$

$$+ \overline{P} \sum_N K_N \int_t^T E(J_N(T)|\zeta_{t,x}(s) - \zeta_{t',x'}(s)|) \, ds + \overline{P}|t - t'|$$

$$\leqslant \overline{P}|t - t'| + \overline{P} \sum_N K_N (E|\zeta_{t,x}(T) - \zeta_{t',x'}(T)|^2 g_N(T))^{1/2} \, (P[J_N = 1])^{1/2}$$

$$+ \overline{P} \sum_N K_N \int_t^T (E|\zeta_{t,x}(s) - \zeta_{t',x'}(s)|^2 g_N(s))^{1/2} \, ds (P[J_N = 1])^{1/2}$$

where in the last estimate we have used Hölder's inequality and the fact that $J_N(T) = 1 \Rightarrow g_N(s) = 1$ for all $s \leqslant T$. From (18) we may now conclude that

$$(28) \quad \begin{aligned} |u(t, x) - u(t', x')| &\leqslant \overline{P}|t - t'| + \overline{P}(1 + T - t)(E|\zeta_{t,x}(t) - \zeta_{t',x'}(t)|^2)^{1/2} \\ &\quad \cdot \sum_N K_N (P[J_N = 1])^{1/2} \exp(\tfrac{1}{2} k_N (2 + nd)(T - t)). \end{aligned}$$

Now the general term in this series is (cf. proof of Lemma 1) bounded by

$$N \exp(-Q_2 N^{2-2\mu} + \text{const} \cdot N^{2\lambda} + \text{const} \cdot N^{2\nu}).$$

Since $2 - 2\mu$ exceeds 2λ and 2ν this series converges. Also

$$E|\zeta_{t,x}(t) - \zeta_{t',x'}(t)|^2 \leqslant 2|x - x'| + 2E|x' - \zeta_{t',x'}(t)|^2$$

so that from (23) and (28) we obtain (25). This proves (a). If now $\nu = 0$ and $\lambda = 0$ then (27) is valid for all curves α and γ and K_N is independent of N. Thus

$$|u(t, x) - u(t, x')| = E|H_t(\zeta_{t,x}) - H_t(\zeta_{t,x'})|$$

$$\leqslant K\overline{P}E|\zeta_{t,x}(T) - \zeta_{t,x}(t)| + K\overline{P} \int_t^T E|\zeta_{t,x}(s) - \zeta_{t,x'}(s)|\,ds.$$

Part (b) of Lemma 1 now yields (26). This completes the proof of Theorem 2.

3. The first order equation. Let us now specialize to the case $\sigma \equiv 0$. Equation, (1) now becomes

$$(29) \qquad\qquad\qquad u_t + a_j u_{x_j} = bu + c.$$

The characteristic equation becomes

$$(30) \qquad\qquad\qquad dx/dt = a(t, x).$$

$\zeta_{t,x}$ is now a degenerate process consisting of a single absolutely continuous curve. It satisfies (30) in the sense of Filippov [1], and is uniquely determined on the interval $[t, T]$. u is now given by

$$(31) \qquad\qquad\qquad u(t, x) = H_t(\zeta_{t,x})$$

and we have the following slight improvement over [1].

Theorem 3. *Let* $1°$ *through* $5°$ *be satisfied but let* $\sigma \equiv 0$. *Then*

(a) *u is locally Lipschitz continuous in both t and x and hence is differentiable almost everywhere in* $I \times \mathbf{R}^d$.

(b) *u satisfies* (29) *almost everywhere in* $I \times \mathbf{R}^d$.

Proof. (a) follows from a specialization of the proof of Theorem 2. The differentiability is of course Rademacher's theorem [11]. Part (b) is proved in my earlier paper [1].

We again emphasize that u is stable under smoothing of the coefficients. But it also follows from our work in §1 that u is the uniform limit on compacta of solutions of parabolic equations.

REFERENCES

1. E. D. Conway, *Generalized solutions of linear differential equations with discontinuous coefficients and the uniqueness question for multidimensional quasilinear conservation laws,* J. Math. Anal. Appl. **18** (1967), 238–251. MR **34** #6293.

2. ———, *Stochastic equations with discontinuous drift,* Trans. Amer. Math. Soc. **157** (1971), 235–245. MR **43** #1286.

3. ———, *Stochastic equations with discontinuous drift.* II, Indiana Univ. J. Math. **22** (1972), 91–99.

4. ———, *On the total variation of solutions of parabolic equations,* Indiana Univ. J. Math. **21** (1971), 493–503.

5. R. Courant, *Methods of mathematical physics*. Vol. II: *Partial differential equations*, Interscience, New York, 1962. MR **25** #4216.

6. I. M. Gel'fand, *Some problems in the theory of quasilinear equations*, Uspehi Mat. Nauk **14** (1959), no. 2 (86), 87–158; English transl., Amer. Math. Soc. Transl. (2) **29** (1963), 295–381. MR **22** #1736; MR **27** #3921.

7. I. I. Gihman and A. V. Skorohod, *Introduction to the theory of random processes*, "Nauka", Moscow, 1965; English transl., Saunders, Philadelphia, Pa., 1969. MR **33** #6689; MR **40** #923.

8. R. Z. Has'minskiĭ, *The averaging principle for parabolic and elliptic differential equations and Markov processes with small diffusion*, Teor. Verojatnost. i Primenen. **8** (1963), 3–25. (Russian) MR **28** #4253.

9. H. P. McKean, Jr., *Stochastic integrals*, Probability and Math. Statist., no. 5, Academic Press, New York, 1969. MR **40** #947.

10. R. S. Phillips and L. Sarason, *Elliptic-parabolic equations of the second order*, J. Math. Mech. **17** (1967/68), 891–917. MR **36** #2942.

11. J. B. Serrin, Jr., *On the differentiability of functions of several variables*, Arch. Rational Mech. Anal. **7** (1961), 359–372. MR **25** #3131.

12. D. W. Stroock and S. R. S. Varadhan, *Diffusion processes with continuous coefficients*, Comm. Pure Appl. Math. **22** (1969), 345–400, 479–530. MR **40** #6641; #8130.

13.———, *On degenerate elliptic-parabolic operators of second order and their associated diffusions*, Comm. Pure Appl. Math. **25** (1972), 651–713.

14. M. Zakai, *Some moment inequalities for stochastic integrals and for solutions of stochastic differential equations*, Israel J. Math. **5** (1967), 170–176. MR **37** #991.

DEPARTMENT OF MATHEMATICS, TULANE UNIVERSITY, NEW ORLEANS, LOUISIANA 70118