SEL MATH SOV
VOL. 1, NO. 1, 1981

0272-9903/81/01/0019-43$10.10

On Strong Solutions of Stochastic Differential Equations*

A. K. Zvonkin and N. V. Krylov

1. Definitions, examples, discussion

1. Introduction. For the last two decades most authors writing on random processes have used the words "solution" of a stochastic differential equation, uniqueness of a "solution" etc., not realizing that the word "solution" was being used in different senses in different works. This has led to much confusion (and even outright errors in some cases), and only the introduction of the notion of "strong" and "weak" solutions as well as "strong" and "weak uniqueness" could resolve this problem. On examination of extant works it appears that the classical Itô theorem is the only result on the existence and uniqueness of strong solutions [see 6], and that most of the "stronger" versions and generalizations [see 21, 12, 23, and others] are actually valid only for "weak" solutions. The problem of the existence of strong solutions has had to be re-examined: some cases have been resolved, others appear to be very difficult to settle. The objective of our survey is to describe the situation at the present time.**

2. Notation

1. Let $\{\Omega, F, P\}$ be a probability space and R^n an n-dimensional Euclidean space equipped with an affine structure, i.e., each point $x \in R^n$ is associated with an n-dimensional space L_x^n having the coordinate origin at x (and a natural identification operation among all L_x^n, $x \in R^n$). A scalar product induced by the metric in R^n is fixed in L_x^n. Cartesian

*Originally published in Trudy shkoly-seminar po teorii sluchaynykh protsessov, part 2 (1975) Vilnius. Translated by A. B. Aries.

**The authors are thankful to A. N. Shiryayev and all the participants of the Seminar on the Theory of Stochastic Processes for their very useful discussions throughout.

coordinates are fixed in R^n. We consider the stochastic integral equation

$$x_t = x + \int_0^t b(s, x_s)\, ds + \int_0^t \sigma(s, x_s)\, dw_s, \tag{1}$$

where $x \in R^n$; $t \in [0, T]$; x_t, x_s are random variables with values in R^n; $b(t, x)$ is a vector field on R^n, that is, for all $(t, x) \in [0, T] \times R^n$ the vector $b(t, x) \in L_x^n$; for each $(t, x) \in [0, T] \in R^n$, $\sigma(t, x)$ is the matrix $\begin{pmatrix} \sigma_{11}, \ldots, \sigma_{1n} \\ \cdots\cdots\cdots \\ \sigma_{n1}, \ldots, \sigma_{nn} \end{pmatrix}$ of a linear mapping σ: $L_x^n \to L_x^n$; $w_t = \begin{bmatrix} w_t^{(1)} \\ \vdots \\ w_t^{(n)} \end{bmatrix} \in R^n$ is a standard n-dimensional Wiener process.

The second integral on the right-hand side of (1) is an Itô integral [see 17].

We shall often write Eq. (1) as a differential equation:

$$dx_t = b(t, x_t)\, dt + \sigma(t, x_t)\, dw_t,$$

$$x_0 = x. \tag{2}$$

2. If $\xi_t = \xi_t(\omega)$ is some random process on $\{\Omega, F, P\}$ then:
   ξ_t is the value of the process at time $t \in [0, T]$;
   ξ_0^t is the trajectory of the process on a time interval $[0, t]$;
   F_t^ξ is the σ-algebra of subsets of Ω generated by the sets $\{\omega: \xi_s(\omega) \in \Gamma\}$, where $s \leqslant t$, the Γ are Borel sets in the range of values of ξ_t; as a rule we shall assume that the σ-algebras F_t^ξ are complete in all probability measures on Ω;
   $H^\xi[0, t]$ is the family of F_t^ξ-measurable random variables;
   $H_2^\xi[0, t]$ is the family of square-summable random variables in $H^\xi[0, t]$.
3. For a mapping f: $R^n \to R^m$ we have [see 20, for example]:
   The derivative $f'(x)$ is the linear mapping $L_x^n \to L_y^m$ (where $y = f(x) \in R^m$) given by the $m \times n$ matrix $(\partial f_i / \partial x_j)$.
   The second derivative $f''(x)$ is a bilinear mapping $L_x^n \times L_x^n \to L_y^m$.

For a mapping f: $R^n \times R^m \to R^k$, $f(x, y) = z$, we denote by $L_{(x,y)}^n$ and $L_{(x,y)}^m$ the linear subspaces of the space $L_{(x,y)}^{n+m}$ that are parallel to R^n and R^m respectively. Then the partial derivative $f_x'(x, y)$ is a linear mapping $L_{(x,y)}^n \to L_z^k$.

The second partial derivative $f_{xx}''(x, y)$ is a bilinear mapping $f_{xx}''(x, y)$: $L_{(x,y)}^n \times L_{(x,y)}^n \to L_z^k$.

In similar way we define the derivatives $f_y'(x, y)$, $f_{yy}''(x, y)$, $f_{xy}''(x, y)$.

4. We introduce the symmetric matrix $a(t, x) = \sigma \cdot \sigma^*$ (here σ^* is the transpose of the matrix σ); it will play the role of a bilinear form on L_x^n, a: $L_x^n \times L_x^n \to R$).[1]

[1] It would be more correct to regard a as a bilinear form on the space $(L_x^n)^*$ dual to L_x^n; we have identified L_x^n and $(L_x^n)^*$ by fixing a scalar product in L_x^n.

Let $e_1,\dots,e_n$ be an arbitrary basis in L_x^n, and let $d_1,\dots,d_n$ be its orthogonal basis (i.e., $(e_i,d_j)=\sigma_{ij}$).[2]

We can determine the differential operator $L^{(x_t)}$ corresponding to the process x_t as follows. For $u:[0,T]\times R^n\to R^m$,

$$\mathcal{L}^{(x_t)}u(t,x)=u_t'+u_t'+u_x'\cdot b(t,x)+\frac{1}{2}\sum_{i,j=1}^{n}u_{xx}''(e_i,e_j)a(t,x)(d_i,d_j).^3 \quad (3)$$

5. It will be convenient to choose $|(a_{ij})|=\max_{ij}|a_{ij}|$ as the matrix norm:

$$|f'(x)|=\max_{i,j}\left|\frac{\partial f_i}{\partial x_j}(x)\right|,$$

$$|f''(x)|=\max_{i,j,k}\left|\frac{\partial^2 f_k}{\partial x_i\partial x_j}(x)\right|.$$

For a region $D\subset R^n$ the space $W_p^{1,2}([0,T]\times D)$, $p>1$, is the completion of the family of infinitely differentiable functions in the norm

$$|u(t,x)|_{W_p^{1,2}}=\|u(t,x)\|_{L_p}+\|u_t'(t,x)\|_{L_p}+\|u_{xx}''(t,x)\|_{L_p},$$

where

$$\|f(t,x)\|_{L_p}=\left(\int_{[0,T]\times D}|f(t,x)|^p\,dt\,dx\right)^{1/p}.$$

In a similar way we define the space $W_p^2(D)$ (dropping the term with derivative in t and integrating over D). For a more detailed discussion of the spaces W see [22], [15].

If the function $u(t,x)$ defined on $[0,T]\times R^n$ belongs to the space $W_p^{1,2}([0,T]\times D)$ for any bounded region $D\subset R^n$, we shall simply write $U\in W_p^{1,2}$. We introduce a similar notation for the spaces W_p^2, L_p.

3. *Definitions*

1. STRONG SOLUTION. Given is a probability space $\{\Omega,F,P\}$ with a Wiener process $\{w_t,F_t^w\}$. Also given are measurable functions $b_k:[0,T]\times R^n\to R$, $k=1,\dots,n$, $\sigma_{ij}:[0,T]\times R^n\to R$, $i,j=1,\dots,n$.

The problem is to construct a process x_t which is measurable with respect to the σ-algebra F_t^w for each t (i.e., $x_t\in H^w[0,t]$ for all $t\in[0,T]$) and such that (1) is satisfied for all $t\in[0,T]$ with probability one. The pair of processes $\{(x_t,w_t),F_t^w\}$ is said to be a *strong solution* of Eq. (1) (or Eq. (2)).

[2]Note that if the basis $e_1,\dots,e_n$ is orthonormal, $e_k=d_k$.

[3]The sum on the right-hand side of Eq. (3) does not depend on the choice of the basis (see [2]).

2. WEAK SOLUTION. We are given only the two functions $b_k(t, x)$ and $\sigma_{ij}(t, x)$. The problem is to construct a probability space $\{\Omega, F, P\}$ and a pair of processes $\{(x_t, w_t), F_t\}$ such that $\{w_t, F_t\}$ is a Wiener process and (1) holds for all $t \in [0, T]$ with probability one. Such a pair of processes $\{(x_t, w_t), F_t\}$ is said to be *a weak solution* of Eq. (1) (or Eq. (2)).

3. The formal difference between the two definitions above is that:

a) a strong solution and a weak solution are solutions to two different problems: in one case we need to construct only x_t (everything else is given), in the other case we need to construct $\Omega, F, P, F_t, w_t, x_t$;
b) (crucial!) if (x_t, w_t) is a weak solution, x_t need not be measurable with respect to the σ-algebra of F_t^w.

Actually, as we shall see from the results in §2, the difference between strong and weak solutions is much deeper and is not merely formal (i.e., due not only to measurability).

Remark. We shall permit the following "abuses" in our notation.

a) By a weak solution we mean not the pair of processes (x_t, w_t) but the process x_t itself.
b) We shall call the process x_t a strong solution if it turns out to be F_t^w-measurable (therefore, a weak solution may be strong).
c) We call a weak solution simply a "solution" of Eq. (1).

The next two notions of uniqueness are identical for both weak and strong solutions.

4. STRONG UNIQUENESS (uniqueness with respect to sample paths). If for any two solutions (x_t', w_t') and (x_t'', w_t'') on the same probability space, it follows (a.s.) from the equalities $x_0' = x_0''$ and $w_t' \equiv w_t''$ that

$$P\left\{\max_{t \in [0, T]} |x_t' - x_t''| > 0\right\} = 0, \tag{4}$$

then we say that the solution of Eq. (1) is *unique in the strong sense,* or *unique with respect to the sample paths*, or simply, pathwise unique.

5. WEAK UNIQUENESS (uniqueness in measure). If for any two solutions (x_t', w_t') and (x_t'', w_t'') of Eq. (1) all the finite-dimensional distributions coincide in all the pairs of processes mentioned above, we say that the solution of Eq. (1) is *unique in the weak sense*, or *unique in the sense of measure.*

4. *Examples*

Example 1. In this example we demonstrate a canonical technique for constructing weak solutions. We shall construct a solution of the equation

$$\begin{cases} dx_t = \alpha(t, x_0^t)\, dt + dw_t \\ x_0 = 0, \end{cases} \tag{5}$$

where $\alpha(t, x_0^t)$ is a bounded nonanticipatory (i.e., F_t^x-measurable for each t) functional.

Let us take an arbitrary Wiener process ξ_t and consider the new process

$$\eta_t = \xi_t - \int_0^t \alpha(s, \xi_0^s)\, ds.$$

It is well known (see [4] or [16]) that the process η_t also is a Wiener process, with respect, however, to a new measure $\tilde{P}$ which is absolutely continuous with respect to the initial measure P, with Radon-Nikodym derivative:

$$\exp\left\{ \int_0^T \alpha(t, \xi_0^t)\, d\xi_t - \frac{1}{2} \int_0^T \alpha^2(t, \xi_0^t)\, dt \right\}.$$

Taking this new measure $\tilde{P}$, we see that η_t is a Wiener process and

$$\xi_t = \int_0^t \alpha(s, \xi_0^s)\, ds + \eta_t;$$

i.e., if we write $x_t = \xi_t$, $w_t = \eta_t$, we get a solution of Eq. (5). Note that the same (initially Wiener) process ξ_t can be a solution of Eq. (5) for various $\alpha(t, x_0^t)$. In that case the measure $\tilde{P}$ and the Wiener process η_t change. First we construct the required process x_t. Further, we choose the measure $\tilde{P}$ and the Wiener process w_t (we choose them instead of specifying them in advance). As a result, instead of $x_t \in H^w[0, t]$, we have, conversely, $w_t \in H^x[0, t]$.

The following question arises. As we have shown, a solution exists with one Wiener process (the one we have constructed). If, however, we are given in advance another Wiener process, can it happen that a solution does not exist in this case? What is then the difference between these two Wiener processes?

This is a strange question, and it has not yet been answered. Obviously, the question itself needs to be more precise. We can only note that in order to *construct* x_t from w_t in some way, we need to express x_t in terms of w_t, using a measurable mapping. This is exactly what a weak solution cannot do.

Example 2 (H. Tanaka). This example shows that in fact there exist weak solutions that are not strong, and that uniqueness in measure does not imply uniqueness with respect to sample paths.

Consider the equation

$$\begin{cases} dx_t = \sigma(x_t)\, dw_t, \\ x_0 = 0, \end{cases} \tag{6}$$

where $\sigma(x) = \begin{cases} 1, & x \geqslant 0, \\ -1, & x < 0, \end{cases}$ (note that $\sigma^2(x) \equiv 1$).

a) *Weak uniqueness.* The martingale $x_t = \int_0^t \sigma(x_s)\, dw_s$, according to [17, §2.5], is a Wiener process with respect to a new time $\tau_t = \int_0^t \sigma^2(x_s)\, ds = t$, which coincides in this case with the old one. Therefore, any solution x_t has the same (namely, Wiener) finite-dimensional distributions.

b) *The existence of a weak solution.* Taking an arbitrary Wiener process for x_t, we construct $\tilde{w}_t = \int_0^t \sigma(x_s)\, dx_s$. The process $\tilde{w}$ is also a Wiener process (it is possible to make a time substitution as we did in (a)), so $d\tilde{w}_t = \sigma(x_t)\, dx_t$, yielding

$$\sigma(x_t)\, d\tilde{w}_t = \sigma(x_t)\sigma(x_t)\, dx_t = dx_t.$$

Therefore, $(x_t, \tilde{w}_t)$ is a (weak) solution of Eq. (6).

c) *The absence of strong uniqueness.* In addition to the solution $(x_t, \tilde{w}_t)$, Eq. (6) has at least one more solution, $(-x_t, \tilde{w}_t)$.

d) *The absence of a strong solution.*[4] No solution of Eq. (6) is strong; i.e., x_t is known to be *non-measurable* with respect to the σ-algebra of F_t^w. To make sure this is so, consider the random variable θ_t—the local time during which the process x_t sojourns at the zero point. Since the process x_t is a Wiener process, θ_t is representable in the form (see [17], §3.8)

$$\theta_t = x_t^+ - \int_0^t \chi_{[0,\infty)}(x_s)\, dx_s.$$

$$\left(\text{Here } x^+ = \max(x,0);\ x^- = -\min(x,0),\right.$$

$$\left.\chi_{[a,b]}(x) = \begin{cases} 1, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}.\right)$$

Due to the obvious symmetry,

$$\theta_t = x_t^- + \int_0^t \chi_{(-\infty,0]}(x_s)\, dx_s.$$

Adding the last two equalities, we get

$$2\theta_t = |x_t| - \int_0^t \operatorname{sgn} x_s\, dx_s = |x_t| - \int_0^t \operatorname{sgn} x_s \sigma(x_s)\, dw_s = |x_t| - w_t,$$

[4] Non-measurability of x_t with respect to F_t^w follows from the general theory developed in §2 (see Corollary 3 of Theorem 1). For this particular case, however, we can give a special proof.

yielding, by the definition of local time,

$$w_t = |x_t| - 2\theta_t = |x_t| - \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \chi_{[0,\varepsilon]}(|x_s|)\, ds,$$

i.e., w_t is measurable and can be expressed in terms of $|x_t|$, and $F_t^w \subset F_t^{|x|}$. Assuming that $x_t \in H^w[0, t]$, i.e., $F_t^x \subset F_t^w$, we arrive at the incompatible relation $F_t^x \subset F_t^{|x|}$.

Example 3. This example develops further the idea in Example 2. Consider the two-dimensional version of Example 2:

$$\begin{cases} dx_t = \sigma(x_t)\, dw_t \\ x_0 = 0, \end{cases} \tag{7}$$

where

$$\sigma(x) = \begin{bmatrix} \dfrac{x_1}{\sqrt{x_1^2 + x_2^2}} & \dfrac{-x_2}{\sqrt{x_1^2 + x_2^2}} \\ \dfrac{x_2}{\sqrt{x_1^2 + x_2^2}} & \dfrac{x_1}{\sqrt{x_1^2 + x_2^2}} \end{bmatrix} = \frac{1}{r} \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix}.$$

Note that the matrix $\sigma(x)$ is orthogonal for all x and infinitely differentiable everywhere except at the point $x = 0$, where the matrix $\sigma(x)$ is discontinuous.

a) We shall now construct a weak solution. To this end, we take for x_t an arbitrary Wiener process. And we construct the process w_t by the formula

$$w_t = \int_0^t \sigma^{-1}(x_s)\, dw_s.$$

The pair of processes (x_t, w_t) yields a solution since $dw_t = \sigma^{-1}(x_t)\, dx_t$, or $dx_t = \sigma(x_t)\, dw_t$.

That the process w_t thus constructed is a Wiener process follows from the orthogonality of the matrix σ: because $\sigma^{-1} = \sigma^*$, we have

$$\langle w_t^{(i)}, w_t^{(j)} \rangle = \int_0^t (\sigma^{-1}(\sigma^{-1})^*)_{ij}\, ds = \delta_j^i t,$$

since $\sigma^{-1}(\sigma^{-1})^* = E$ (the unit matrix) (for the definition of $\langle \cdot, \cdot \rangle$, see [16], Ch. 5, §1).

b) Similarly, for the process x_t we have

$$E\left[x_t^{(i)}, x_t^{(j)} \right] = \int_0^t (\sigma\sigma^*)_{ij}\, ds = \delta_j^i t;$$

i.e., any solution x_t of Eq. (7) has Wiener measure.

c) It is easy to verify that in addition to the solution (x_t, w_t) Eq. (7) has an entire family of solutions (Ax_t, w_t), where A is an arbitrary orthogonal matrix. For

$$d(Ax_t) = A\, dx_t = A\sigma(x_t)\, dw_t,$$

and it remains only to verify that $A\sigma(x)=\sigma(Ax)$, which is immediate.

d) The absence of a strong solution follows from Theorem 1, §2 (see Corollary 3).

Note that if we took an initial point $x_0 \neq 0$, the solution would be strong and pathwise unique (and would remain so until it leaves the region where $\sigma(x)$ is smooth, i.e., the region $R^2 \setminus \{0\}$; the solution, being a Wiener process, will, however, never leave this region; that is, the solution will never get to the zero point).

5. Discussion. In which cases is it sufficient to have a weak solution of Eq. (1), and in which cases is it necessary to have a strong solution?

The difference between strong and weak solutions is very similar to the difference between a random variable and its distribution: by constructing a strong solution we construct a particular function $x_t = x(t, w)$. Hence:

1. Weak solutions are adequate in those situations where the answers to the questions we are concerned with involve only the measure on the space of trajectories. Such questions include: the determination of various probabilities and mathematical expectations; problems related to the stability of processes and to the existence of invariant measures; the problems of absolute continuity of measures for various processes; probabilistic representation of solutions of partial differential equations.

However, there are situations where a diffusion process needs to be regarded as a particular family of trajectories. In such cases we must have a strong solution. We shall illustrate this by examples.

2. There is a well-known result, referred to as "a comparison theorem" (see [3]):

If x'_t and x''_t are solutions of the equations

$$dx'_t = b_1(t, x'_t)\,dt + dw_t \tag{8.1}$$

$$dx''_t = b_2(t, x''_t)\,dt + dw_t, \tag{8.2}$$

with $x'_0 = x''_0$ and $b_1(t, x) > b_2(t, x)$, then for all $t > 0$ we have $x'_t > x''_t$ (a.s.).

This result makes no sense for weak solutions, if only because the latter may be given on different probability spaces. Even if we construct Eqs. (8.1) and (8.2) on the same space, as in Example 1, choosing the processes x'_t and x''_t in advance without knowing the coefficients b_1 and b_2, we can have $x'_t(w) = x''_t(w)$, or $x'_t(w) = -x''_t(w)$, etc.; i.e., the theorem does not hold.

3. Another example occurs in the theory of diffusion control processes.[5] On the interval $[0, T]$ let there be given

$$\begin{cases} dx_t = u_t\,dt + dw_t, \\ x_0 = 0, \end{cases} \tag{9}$$

where we can choose the control u_t from the set $\{u: |u| \leqslant 1\}$. We assume that our aim is to deviate as little as possible from zero (i.e., maximize one of the functionals

$$M|x_T|, \tag{10.1}$$

$$Mx_T^2, \tag{10.2}$$

$$M\int_0^T x_t^2\,dt, \tag{10.3}$$

or maximize one of the functionals

$$M\{\tau_a \wedge T\}, \tag{10.4}$$

$$P\{|x_T| < a\}, \tag{10.5}$$

where $\tau_a = \inf\{t: |x_t| > a\}$, $\tau_a \wedge T = \min(\tau_a, T)$, etc.). It is natural in this situation to drive the process x_t as fast as possible to zero; i.e., to choose the control so that:

$$u_t = \begin{cases} -1, & x_t > 0, \\ 0, & x_t = 0, \\ 1, & x_t < 0. \end{cases}$$

Eq. (9) thus becomes

$$\begin{cases} dx_t = -\operatorname{sgn} x_t + dw_t, \\ x_0 = 0. \end{cases} \tag{11}$$

Note that the drift coefficient is discontinuous in this case, and hence the Itô theorem [6] cannot be applied to yield a strong solution.

The value of each of the functionals (10.1)–(10.5) remains the same regardless of whether we consider a weak or a strong solution of Eq. (11). However, the meaning of our actions is fundamentally different; we have stopped *controlling* the process x_t. In fact, the meaning of the control is that while observing an individual trajectory of x_t, we try to bring it back to zero as soon as this trajectory deviates from 0. But in constructing weak solutions of Eq. (9) under various strategies u_t, we do not in general change the process x_t; we keep all the trajectories of x_t unchanged. We change only the measure on the set of *all* the trajectories, although the behavior of the trajectories other than the one we are observing is of no concern to us.

[5] It is difficult to say in what way this discussion is related to engineering control problems (perhaps in no way). Our discussion is thus confined to mathematical theory.

Apparently, it is not possible to regard the mathematical theory of controlled diffusion processes as complete until sufficiently general results are obtained about the existence of strong solutions.

4. We now consider the problem of filtering an unobservable component of a two-dimensional process. Let (θ_t, ξ_t) be a process satisfying the equations

$$d\theta_t = a(t, \theta_t, \xi_t)\, dt + dw_t^1,$$
$$d\xi_t = A(t, \theta_t, \xi_t)\, dt + dw_t^2.$$

The problem is to estimate θ_t from the trajectory ξ_0^t. It is well known (see [16], Ch. 8, §3) that the best mean-square estimate $M(\theta_t | F_t^\xi)$ is representable in the form

$$M(\theta_t | F_t^\xi) = M(\theta_0 | F_0^\xi) + \int_0^t [\cdot]\, ds + \int_0^t [\cdot]\, d\bar{w}_s$$

(the specific form of the integrands is of no interest to us), where $\bar{w}_t$ is the Wiener process given by

$$\xi_t - \int_0^t M(A(s, \theta_s, \xi_s) | F_s^\xi)\, ds.$$

The construction shows that the process $\bar{w}_t$ is F_t^ξ-measureable, i.e., $F_t^{\bar{w}} \subset F_t^\xi$. It is well known that the case when the process $\bar{w}_t$ is an *innovation* process is particularly important, i.e., $F_t^{\bar{w}} = F_t^\xi$ (this means that no information has been lost in going from the process ξ_t to the process $\bar{w}_t$), and the equality $F_t^{\bar{w}} = F_t^\xi$ can be satisfied if and only if the equation

$$d\xi_t = \alpha(t, \xi_0^t)\, dt + d\bar{w}_t$$

has a strong solution, where

$$\alpha(t, \xi_0^t) = M(A(t, \theta_t, \xi_t) | F_t^\xi).$$

5. The result on strong uniqueness obtained in §3 leads to a corollary which looks paradoxical within the framework of the theory of ordinary differential equations. Consider the equation $\dot{x} = f(t, x)$, or more precisely,

$$x_t = x_0 + \int_0^t f(s, x_s)\, ds. \tag{12}$$

If $f(t, x)$ does not satisfy a Lipschitz condition in x, a solution of Eq. (12) may not be unique (for example, $\dot{x} = 2\sqrt{x}$, $x_0 = 0$; solutions of this equation will be $x(t) \equiv 0$, $x(t) = t^2$, and others). We add an arbitrary small stochastic part to the right-hand side of Eq. (12); i.e., we consider the equation

$$x_t = x_0 + \int_0^t f(s, x_s)\, ds + \varepsilon w_t. \tag{13}$$

It turns out that for (almost) any trajectory w_0^T of the process w_t the solution x_t of Eq. (13) is unique. It will be unique even for the well-known

example (see [5], Ch. II, §5) in which the equation $\dot{x}=f(t, x)$ exhibits nonuniqueness at each point x.

Note that this corollary cannot be derived from the results on weak uniqueness of solutions of Eq. (13). The classical Itô theorem [6] proves the strong uniqueness of a solution of Eq. (13) only in the case where the function $f(t, x)$ is Lipschitz in x, i.e. when Eq. (12) has a unique solution.

6. *A survey of some results on weak solutions.* Let the coefficients $b(t, x)$ and $\sigma(t, x)$ in Eq. (1) satisfy the following conditions.

A. Continuity of the matrix $a(t, x)$: the matrix $a(t, x)=\sigma\sigma^*$ is continuous in t and in x.

B. Uniform ellipticity of the operator $L^{(x_t)}$: there exists a number $\mu>0$ such that for any vector $e\in L_x^n$

$$a(t, x)(e, e)\geqslant \mu^2|e|^2. \tag{14}$$

C. Boundedness of the coefficients: there is a number c such that

$$|b(t, x)|+|\sigma(t, x)|\leqslant c. \tag{15}$$

Then the following assertions hold:

1. A (weak) solution of Eq. (1) exists and is unique in measure.
2. If solutions of Eq. (1) are considered on intervals $[t, T]\subset[0, T]$ for initial values $x_t=x\in R^n$, each solution is a (nonstationary) strictly Markov process.

   We denote the mathematical expectation with respect to the measure corresponding to this process by $M_{(t,x)}$, and for $t=0$ we denote it by M_x.
3. Let $D\subset R^n$ be a bounded region. Let τ be the time of first exit of the process x_t from the region D. And let a function $f(t, x)$ be given, $f\in L_p$ $([0, T]\times D)$, where $p\geqslant n+1$; we then have the estimate

$$M_x\int_0^{\tau\wedge T}|f(t, x_t)|\,dt\leqslant N\|f\|_{L_p}, \tag{16}$$

   where the constant N depends only on T, μ, c. If $f(t, x)=f(x)$ does not depend on t, and if $f\in L_p(D)$, the same estimate holds for $p\geqslant n$, N being independent of T.
4. Consider the process $y_t=u(t, x_t)$, where the function $u\in W_p^{1,2}$, $p\geqslant n+1$. We assert that the Itô differential rule holds for the function $u(t, x)$:

$$du(t, x_t)=\left[\mathcal{L}^{(x_t)}u(t, x_t)\right]dt+\left[u_x'(t, x_t)\sigma(t, x_t)\right]dw_t. \tag{17}$$

   We call the function $f(t, x)$ "slowly increasing" if it increases infinitely slower than $e^{k|x|^2}$ for any $k>0$. Consider the equation

$$\mathcal{L}^{(x_t)}u(t, x)+f(t, x)=0,$$
$$u(T, x)=\varphi(x), \tag{18}$$

where $u: [0, T] \times R^n \to R^m$; $f: [0, T] \times R^n \to R^m$; $\varphi: R^n \to R^m$, $f(t, x)$ and $f(x)$ are slowly increasing functions, and $f \in L_p, \varphi \in W_p^2$ for some $p > n + 2/2$. Then:

5. In the class of slowly increasing functions the solution $u(t, x)$ of Eq. (18) exists, is unique, and $u \in W_p^{1,2}$.
6. If $p \geqslant n + 1$, this solution has the probabilistic representation

$$u(t, x) = M_{(t,x)}\left(\int_t^T f(s, x_s)\, ds + \varphi(x_T) \right). \tag{1}$$

Assertions 1, 2 have been proved in [23]; assertion 5 has been proved in [15]; a propos of assertions 3, 4, 6, see [14].

From now on, unless stated otherwise, we shall assume conditions A, B, C to be satisfied, although some of the assertions proved hold under weaker restrictions.

2. Properties of strong solutions

1. *Basic assertions*

The main result of this section is:

Theorem 1. Let conditions A, B, C (§1.6) be satisfied. Then a solution of Eq. (1) is strong if and only if it is pathwise unique.

In other words, for given coefficients $b(t, x)$, $\sigma(t, x)$ we have the following

Alternative. Either any solution of Eq. (1) given on any probability space is strong and pathwise unique, or no solution on any probability space is strong, and moreover these solutions are not necessarily pathwise unique.

We note that the existence of a weak solution and its weak uniqueness are ensured by conditions A, B, and C.

The following corollaries of Theorem 1 are actually the theorem itself in a more complete form.

Corollary 1. Let Eq. (1) have a strong solution on some probability space. Then:

(a) any weak solution of Eq. (1) is strong;
(b) on any probability space two strong solutions of Eq. (1) coincide for all t with probability 1.

Corollary 2. Let Eq. (1) have a weak solution and let it be pathwise unique. Then all solutions of Eq. (1) are strong.

Corollary 3. If on some probability space Eq. (1) has two solutions (not coinciding for all t with probability 1), then Eq. (1) has no strong solution on any probability space.

We note that Corollary 3 follows trivially from Corollaries 1 and 2. Hence we need only prove Corollaries 1 and 2.

(In 5 we shall formulate and prove a criterion for the existence of a strong solution.)

2. *Proof of Corollary 1(b)*

1. Denote by M the set of bounded nonrandom (Lebesgue) measurable functions $m(t)$ on $[0,T]$, $m:[0,T]\to R^n$. Let

$$p_t(m)=\exp\left\{\sum_{k=1}^{n}\left(\int_0^t m_k(s)\,dw_s^{(k)}-\frac{1}{2}\int_0^t m_k^2(s)\,ds\right\}\right. . \tag{19}$$

We assert that the class of random variables $\{p_T(m): m\in M\}$ is fundamental in the Hilbert space $H_2^w[0,T]$.

Lemma 2.1. If $\xi\in H_2^w[0,T]$, and if for every $m\in M$

$$M\xi p_T(m)=0,$$

then $\xi=0$ (a.s.).

Proof. This assertion is proved while proving Theorem 5.5 [16].

2. On the interval $[0,t]$ we consider the equation

$$\mathcal{L}^{(x_t)}u(s,x)+U_x'\sigma(s,x)m(s)=0,\qquad s\in[0,t]$$
$$u(t,x)=\varphi(x). \tag{20}$$

Lemma 2.2. Let the coefficients of the operator $L^{(x_t)}$ satisfy conditions A, B, C (§2.6), $m\in M$. And let $\varphi(x)$ be a slowly increasing function, $\varphi\in W_p^2$, $p\geqslant n+1$. Finally, let x_t be a (weak) solution of Eq. (1). Then

$$Mp_t(m)\varphi(x_t)=u(0,x). \tag{21}$$

Proof. Let $y_t=p_t(m)u(t,x_t)$. If y_t were a martingale, the relation

$$Mp_t(m)\varphi(x_t)=Mp_t(m)u(t,x_t)=Mp_0(m)u(0,x_0)=u(0,x)$$

would be satisfied, which is to be proved. We shall prove that y_t is indeed a martingale. To this end we apply the Itô formula to y_t (see §1.6.4): since

$$dp_t(m)=p_t(m)\sum_{k=1}^{n}m_k(t)\,dw_t^{(k)},$$

we have

$$dy_t=p_t(m)\left[\mathcal{L}^{(x_t)}u(t,x_t)\,dt+u_x'\sigma\,dw_t\right]+u(t,x_t)p_t(m)\sum_{k=1}^{n}m_k\,dw_t^{(k)}$$
$$+p_t(m)u_x'(t,x_t)\sigma(t,x_t)m(t)\,dt$$
$$=[\cdot]\,dw_t+p_t(m)\left\{\mathcal{L}^{(x_t)}u(t,x_t)+u_x'(t,x_t)\sigma(t,x_t)m(t)\right\}dt.$$

Note that by Eq. (20) the coefficient of dt is zero, which was to be proved.

3. We proceed to prove Corollary I(b). Let $\varphi(x)=x$ in Eq. (20) and let x_t', x_t'' be two solutions of Eq. (1). By Lemma 2.2 we have

$$Mp_t(m)x_t' = u(0,x) = Mp_t(m)x_t''. \tag{22}$$

If x_t' and x_t'' are F_t^w-measurable, i.e., lie in $H_2^w[0,t]$, by Lemma 2.1 it will follow from Eq. (22) that $x_t' = x_t''$ (a.s.).

It remains only to obtain from

$$\forall t \leqslant T \qquad P\{x_t' = x_t''\} = 1$$

that

$$P\left\{\sup_{t<T} |x_t' - x_t''| = 0\right\} = 1,$$

which can be proved in the usual way.

4. **Remark.** We wish to expand on exactly what we have proved. For the given set $\{\Omega, F, P; F_t, w_t\}$ let there exist several solutions of Eq. (1); among them there may be both strong solutions and weak (i.e., F_t^w-measurable and non-measurable) solutions. We have proved that there can be no more than one strong solution. However, we have not ruled out the possibility that besides this strong solution there may exist several other weak solutions differing from the strong solution as well as from each other. We shall disprove this possibility in (3).

3. *Proof of Corollary 1 (a)*

1. **Lemma 3.1.** Let a sequence ξ_n, $n=1,2,\ldots$ of elements in $H_2^w[0,t]$ and an element $\xi \in H_2^w[0,t]$ be bounded in norm by a common constant (i.e., $M|\xi_n|^2 \leqslant K\ \forall n$, and $M|\xi|^2 \leqslant K$), and for all $m \in M$ let

$$Mp_t(m)\xi_n \to Mp_t(m)\xi. \tag{23}$$

Then $\xi_n \to \xi$ weakly in $H_2^w[0,t]$.

Proof. Let ξ_{n_k} be a subsequence not converging weakly to ξ; since the sequence ξ_{nk} is bounded, it has a weakly convergent subsequence (we may assume that this subsequence coincides with the sequence ξ_{n_k}); i.e., $\xi_{n_k} \to \zeta$ weakly in $H_2^w[0,t]$, $\zeta \neq \xi$. By the definition of weak convergence, $Mp_t(m)\xi_{n_k} \to Mp_t(m)\zeta\ \forall m \in M$, which leads us, if we use Eq. (23), to $Mp_t(m)\xi = Mp_t(m)\zeta\ \forall m \in M$. Therefore, by Lemma 2.1, $\xi = \zeta$ (a.s.).

2. The next lemma is of interest in itself.

Lemma 3.2. Suppose we are given a sequence of functions $b^n(t,x)$, $\sigma^n(t,x)$, $n=1,2,\ldots$ such that:

(α) for each n, $b^n(t,x)$ and $\sigma^n(t,x)$ satisfy conditions A, B, C (§1.6), the constants C and μ being independent of n;

(β) for each n, $b^n(t,x)$ and $\sigma^n(t,x)$ satisfy a Lipschitz condition in x;

(γ) as $n\to\infty$, $b^n(t,x)\to b(t,x)$, $\sigma^n(t,x)\to\sigma(t,x)$ for almost all (t,x).

Let x_t be a strong solution of Eq. (1). Denote by x_t^n a solution of the equation

$$x_t^n = x + \int_0^t b^n(s,x_s^n)\,ds + \int_0^t \sigma^n(s,x_s^n)\,dw_s$$

given on the same probability space $\{\Omega, F, P\}$ with the same Wiener process w_t as the process x_t.[6]

We assert that $x_t^n \to x_t$ in the mean-square, uniformly in $t\in[0,T]$.

Remark. The converse is trivial: if $x_t^n\to x_t$ in the mean-square, x_t is F_t^w-measurable, being the limit of the sequence of F_t^w-measurable random variables x_t^n.

3. PROOF OF LEMMA 3.2.

(a) Denote by $u^n(s,x)$ the solution of the equation

$$\begin{cases} \mathcal{L}^n u^n(s,x) + (u^n)'_x \sigma^n(s,x) m(s) = 0, & s\in[0,t] \\ u(t,x)=\varphi(x). \end{cases}$$

It is well known from the theory of differential equations (see [15]) that

$$u^n(s,x)\to u(s,x), \tag{24}$$

where $u(s,x)$ is a solution of Eq. (20).

(b) By Lemma 2.2, it follows from Eq. (24) that

$$Mp_t(m)\varphi(x_t^n)\to Mp_t(m)\varphi(x_t). \tag{25}$$

(c) Setting $\varphi(x)=x$ in formula (25), we find that for any $m\in M$

$$Mp_t(m)x_t^n \to Mp_t(m)x_t. \tag{26}$$

Adding to the above the estimate

$$M|x_t^n|^2 \leqslant k \quad \forall n, \qquad M|x_t|^2 \leqslant k$$

(which easily follows from the boundedness of the coefficients of b^n, b, σ^n, σ), we find, by Lemma 3.1, that $x_t^n \to x_t$ weakly in $H_2^w[0,t]$[7].

[6] It is always possible to construct such an x_t^n, since $b^n(t,x)$ and $\sigma^n(t,x)$ are Lipschitz (see [6]).

[7] Here we essentially use the assumption that $x_t\in H_2^w[0,t]$, i.e., that x_t is a strong solution.

(d) We now show that the convergence $x_t^n \to x_t$ is actually strong. To this end it suffices to show that the norms of x_t^n converge to the norm of x_t; i.e., that $M|x_t^n|^2 \to M|x_t|^2$, which is a particular case of formula (25) for $\varphi(x)=|x|^2$, $m(t)\equiv 0$ (then $p_t(m)\equiv 1$). Uniformity in t on each finite interval of time follows from the similar uniformity in formula (24).

We have thus proved Lemma 3.2.

4. We shall next prove Corollary 1.(a); however, we do it here only for continuous $\sigma(t, x)$. Under the weaker requirement that $a(t, x)$ be continuous, a similar assertion follows from Theorem 2 (see Remark at the end of (5)).

Let (x_t, w_t) be a strong solution of Eq. (1). Let $(\tilde{x}_t, \tilde{w}_t)$ be another solution of Eq. (1) (possibly given on a different probability space). We need to show that the process $\tilde{x}_t$ is measurable with respect to $F_t^{\tilde{w}}$.

(a) Let $\psi_t = M\{\tilde{x}_t | F_t^{\tilde{w}}\}$; we need to show that $\psi_t = \tilde{x}_t$ (a.s.). By Jensen's inequality we have

$$|M\{\tilde{x}_t | F_t^{\tilde{w}}\}|^2 \leqslant M\{|\tilde{x}_t|^2 | F_t^{\tilde{w}}\}.$$

Taking the mathematical expectation on both sides of this inequality, we get $M|\psi_t|^2 \leqslant M|\tilde{x}_t|^2$. As is well known, equality is attainable only in case $\psi = \tilde{x}_t$ (a.s.). Hence it suffices to show that $M|\psi_t|^2 = M|\tilde{x}_t|^2$.

(b) We now take a sequence of functions $b^n(t, x)$ $\sigma^n(t, x)$, which are smooth in x and such that the assumptions of Lemma 3.2 can be satisfied for them. Further, we construct the processes x_t^n, $\tilde{x}_t^n$ as solutions of the equations

$$dx_t^n = b^n(t, x_t^n)\,dt + \sigma^n(t, x_t^n)\,dw_t, \quad x_0^n = x,$$

$$d\tilde{x}_t^n = b^n(t, \tilde{x}_t^n)\,dt + \sigma^n(t, \tilde{x}_t^n)\,d\tilde{w}_t, \quad \tilde{x}_0^n = x,$$

respectively. Note that:

1. The processes x_t^n and $\tilde{x}_t^n$ are strong solutions of the corresponding equations (since the coefficients of $b^n(t, x)$ and $\sigma^n(t, x)$ satisfy a Lipschitz condition in x).
2. The processes x_t^n converge to x_t in the mean-square since x_t is a strong solution (the convergence follows from Lemma 3.2).
3. The processes x_t and $\tilde{x}_t$ and, for any n, the processes x_t^n and $\tilde{x}_t^n$, have the same finite-dimensional distributions (this follows from Assertion 1, §6.1); moreover, for any n the n-vectors $(x_t^1, \ldots, x_t^n)$ and $(\tilde{x}_t^1, \ldots, \tilde{x}_t^n)$ have the same finite-dimensional distributions.

(c) By Eq. (26),

$$M\tilde{p}_t(m)x_t^n \to M\tilde{p}_t(m)x_t.$$

Since the random variables $\tilde{x}_t^n$, $\tilde{x}_t$, $\tilde{p}_t(m)$[8] have the same distributions, we get $M\tilde{p}_t(m)\tilde{x}_t^n \to M\tilde{p}_t(m)\tilde{x}_t$.

Taking the conditional mathematical expectation $M\{\cdot | F_t^{\tilde{w}}\}$, we have

$$M\tilde{p}_t(m)\tilde{x}_t^n \to M\tilde{p}_t(m)\psi_t; \tag{27}$$

since $\psi_t \in H_2^w[0, t]$, according to Lemma 3.1 the weak convergence $\tilde{x}_t^n \to \psi_t$ follows from Eq. (27).

(d) By the Banach-Saks theorem (see [1]) it is possible to derive from the sequence x_t^n a subsequence (which, we assume, coincides with the sequence x_t^n) whose arithmetic means converge to ψ_t strongly. Then

$$M|\psi_t|^2 = \lim_{n\to\infty} \frac{M|\tilde{x}_t^1 + \cdots \tilde{x}_t^n|^2}{n} = \lim_{n\to\infty} \frac{M|x_t' + \cdots x_t^n|^2}{n}$$
$$= M|x_t|^2 = M|\tilde{x}_t|^2, \qquad Q.E.D.$$

4. Proof of Corollary 2 (Yamada, Watanabe [24])

1. **Lemma 4.1.** Let a pair of processes $\{(x_t, w_t), F_t\}$ be a solution of Eq. (1). And let a pair of continuous processes $(\tilde{x}_t, \tilde{w}_t)$ have the same finite-dimensional distributions as the processes (x_t, w_t). Then $\{(\tilde{x}_t, \tilde{w}_t), F_t^{\tilde{w}, \tilde{x}}\}$ also yield a solution of Eq. (1).

Proof. Since the distributions coincide, we have

$$M\left(\tilde{x}_t - x - \int_0^t b(s, \tilde{x}_s)\, ds - \int_0^t \sigma(s, \tilde{x}_s)\, d\tilde{w}_s\right)^2$$
$$= M\left(x_t - x - \int_0^t b(s, x_s)\, ds - \int_0^t \sigma(s, x_s)\, dw_s\right)^2 = 0, \tag{28}$$

yielding

$$\tilde{x}_t = x + \int_0^t b(s, \tilde{x}_s)\, ds + \int_0^t \sigma(s, \tilde{x}_s)\, d\tilde{w}_s, \qquad Q.E.D.$$

2. The preceding argument is not quite valid. In fact, Eq. (28) contains a stochastic integral with respect to the Wiener process $\tilde{w}_t$. A necessary condition for this stochastic integral to be correctly defined is that the integrand be independent of the future of $\tilde{w}_t$.

Lemma 4.2. Let $\tilde{B} \in F_t^{\tilde{x}}$. Then $p\{\tilde{B} | F_T^{\tilde{w}}\} = p\{\tilde{B} | F_t^{\tilde{w}}\}$.

[8] To get $\tilde{p}_t(m)$ we need to substitute w_t for $\tilde{w}_t$ in Eq. (19).

Proof. It suffices to take the event $\tilde{B}$ to be $\tilde{B}=\{w: \tilde{x}_{t_1} \leqslant c_1, \ldots, \tilde{x}_{t_n} \leqslant c\}$, $t_1, \ldots, t_n \leqslant t$ and to show that the latter does not depend on an event $\tilde{A}$ of the form

$$\tilde{A}=\{w: \tilde{w}_{u_2}-\tilde{w}_{u_1} \leqslant d_1, \ldots \tilde{w}_{u_{n+1}} \leqslant d_n\}, t \leqslant u_1, \ldots u_{n+1} \leqslant T.$$

This independence follows from the independence of the events A and B constructed in the same way as the events $\tilde{A}$ and $\tilde{B}$ but now for the processes x_t, w_t.

3. The definition of strong uniqueness given in §1.3.4 can be interpreted as follows: for a fixed trajectory w_0^T of the Wiener process w_t the random variables x_t' and x_t'' in the plane of the variables (x', x'') are distributed on the diagonal $x'=x''$. Our next objective is to show that this distribution is in fact concentrated at a single point; i.e., for a fixed Wiener trajectory the random variables x_t'' and x_t'' are no longer random and can be expressed in terms of a functional of this trajectory.

Let $C_{[0,T]}$ be the space of continuous functions mapping $[0, T]$ into R^n. Let (x_t', w_t') and (x_t'', w_t'') be two solutions of Eq. (1). Denote by P' and P'' the pertinent measures in $(C_{[0,T]} \times C_{[0,T]}, B_T \times B_T)$. By [13] there exist regular conditional distributions $P'(\cdot|w_0^T)$ and $P''(\cdot|w_0^T)$.

We now consider a probability space $\{\Omega, F, P\}$ for which

(a) $$\Omega=C_{[0,T]} \times C_{[0,T]} \times C_{[0,T]};$$

write the elements $\omega \in \Omega$ as $\omega=(x', x'', w)$, where x', x'', w are continuous functions on $[0, T]$ with values in R^n;

(b) $F=B_T \times B_T \times B_T$;

(c) the measure P is:

$$P(dx'\, dx''\, dw)=P'(dx'|w)P''(dx''|w)W(dw),$$

where W is a Wiener measure on $C_{[0,T]}$.

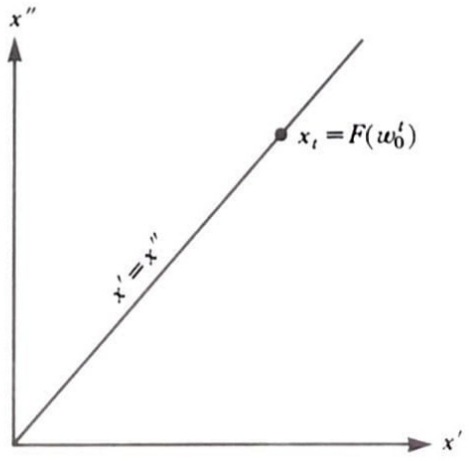

Figure 1

We define the random processes $x'_t(\omega)=x'(t)$, $x''_t(\omega)=x''(t)$, $w_t(\omega)=w(t)$. Then, by Lemma 4.1, the pair of processes (x'_t, w_t) and (x''_t, w_t) satisfy Eq. (1).

4. We proceed to prove Corollary 2.

(a) **Proof.** We have constructed the probability space $\{\Omega, F, P\}$ and two solutions (x'_t, w_t) and (x''_t, w_t) of Eq. (1) having the same Wiener process w_t, $x'_0=x''_0=x$. Hence, by our assumption on the strong uniqueness of solution of Eq. (1), we have:

$$P\{\forall t\in[0,T] \quad x'_t=x''_t\}=1, \tag{29}$$

which implies that each of the conditional measures $P'(\cdot|w)$ and $P''(\cdot|w)$ is almost surely concentrated at a single point.

(b) Fix $t\in[0,T]$ and $w=w_0^T\in C_{[0,T]}$. Then the measure Q_T on $R^n\times R^n$ induced by the pair (x'_t, x''_t) has the following properties:

1) The measure Q_t is concentrated on the hyperplane $\{x'=x''\}$ (i.e., $Q_t\{x'=x''\}=1$; this follows from Eq. (29));
2) The measure Q_t is the product of two measures: $Q_t=P'_tP''_t$, where P'_t and P''_t are measures induced by the processes x'_t and x''_t, respectively (this follows from the construction of $\{\Omega, F, P\}$ and the fact that we have fixed w_0^T).

(c) We have (we omit the subscript t):

$$0=\int_{R^n\times R^n}|x'-x''|^2\,dQ=\int_{R^n\times R^n}|x'-x''|^2\,dP'\,dP''$$

$$=\int_{R^n}\left(\int_{R^n}|x'-x''|^2\,dP''\right)dP', \tag{30}$$

where for the inner integral:

$$\int_{R^n}|x'-x''|^2\,dP''=\int_{R^n}\left[|x'|^2+|x''|^2-2(x',x'')\right]dP''$$

$$=|x'|^2+M|x''|^2-2(x',Mx'')>|x'|^2+|Mx''|^2-2(x',Mx'')=|x'-\bar{x}|^2,$$

where $\bar{x}=Mx''$ is a nonrandom vector on R^n.

Continuing the chain of equalities (30), we get

$$0>\int_{R^n}|x'-\bar{x}|^2\,dP'>0,$$

or,

$$\int_{R^n}|x'-\bar{x}|^2\,dP'=0,$$

yielding $x'=\bar{x}P'$(a.s.). We can argue in a similar way for x''.

(d) Thus for (almost) any $w \in C_{[0,T]}$ there is $\bar{x} \in R^n$ such that with probability 1 $x_{t'} = \bar{x}$, $x_{t''} = \bar{x}$. In other words, there exists a mapping $F_t: C_{[0,T]} \to R^n$ such that if x_t is a solution of Eq. (1), then $x_{t_0} = F_t(w_0^T)$ (a.s.). It remains to prove that the mapping F_t is F_t^w-measurable. But by the definition of regular conditional probability and by Lemma 4.2, for any Borel set $\Gamma \subset R^n$ the function $f(w) = P\{x_t \in \Gamma | w\}$ is F_t^w-measurable; at the same time, this function is given by

$$P\{x_t \in \Gamma | w\} = \begin{cases} 1, & \text{if } F_t(w) \in \Gamma \\ 0, & \text{if } F_t(w) \notin \Gamma. \end{cases}$$

Therefore,

$$\{w \in C_{[0,T]}: F_t(w) \in \Gamma\} = \{w: f(w) = 1\} \in F_t^w, \qquad Q.E.D.$$

We have thus proved Corollary 2 and thereby Theorem 1.

5. **Corollary 2′.** If for some set $\{\Omega, F, P; w_t, F_t^w\}$ there exists a strong solution x_t of Eq. (1), then there exists a strong solution $\tilde{x}_t$ for any set $\{\tilde{\Omega}, \tilde{F}, \tilde{P}, \tilde{w}_t, F_t^{\tilde{w}}\}$.

Proof. Let $\tilde{x}_t = F_t(\tilde{w}_0^T)$. A pair of processes $(\tilde{x}_t, \tilde{w}_t)$ obviously generates the same measure as (x_t, w_t) in $C_{[0,T]} \times C_{[0,T]}$. It remains only to use Lemma 4.1 and Corollary 1.a.

Corollary 2″. Pathwise uniqueness implies uniqueness in measure.

Proof. The proof is obvious, due to the representation $x_t = F_t(w)$.

5. *Criteria for existence of strong solutions*

1. The arguments in §2.5 are based on investigating more closely the structure of the Hilbert space $H_2^w[0,T]$ introduced in §2.2. For the sake of simplicity we confine ourselves to considering the one-dimensional case, although all the arguments can be repeated for the multivariate case.

We have already noted (see Lemma 2.1) that the random variables $P_t^{(m)}$, $m \in M$, form a fundamental set in $H_2^w[0,T]$. Using these vectors we can construct an orthonormal basis in $H_2^w[0,T]$ as follows:

In the Hilbert space $L_2[0,T]$ choose an orthonormal basis $\{m_n(t)\}$ consisting of uniformly bounded functions (in particular, all $m_n(t) \in M$). Let $z = (z_1, z_2, \dots)$, where

$$\sum_{k=1}^{\infty} z_k^2 \leqslant C < \infty, \qquad \sum_{k=1}^{\infty} |z_k| \leqslant C < \infty. \tag{31}$$

Then the function

$$m(t) = \sum_{k=1}^{\infty} z_k m_k(t) \in L_2[0,T] \cap M.$$

And let $p_T(m) = p_T(z)$.

For a set of indices $I=(i_1,\dots,i_n)$, where the i_n are nonnegative integers, we write $|I|=i_1+\cdots+i_n$,

$$I!=i_1!\dots i_n!$$

$$D_z^I=\frac{\partial^{|I|}}{\partial z_1^{i_1}\dots\partial z_n^{i_n}}$$

(if $I=(0,\dots,0)$, D_z^I is the identity operator[9]). Next consider the vectors from $H_2^w[0,T]$ defined by the formula

$$\xi_I=\frac{1}{\sqrt{I!}}D_z^I(p_T(z))\Big|_{z=0}.$$

2. **Lemma 5.2** The set of random variables ξ_I forms an orthonormal basis in the space $H_2^w[0,T]$.

Proof.

(a) *Orthonormality.* Let $I=(i_1,\dots,i_n)$, $J=(j_1,\dots,j_n)$. We have

$$M_{\xi_I\xi_J}=M\left(\frac{1}{\sqrt{I!}}D_{z'}^I p_T(z')\right)\left(\frac{1}{\sqrt{J!}}D_{z''}^J p_T(z'')\right)\Big|_{z'=z''=0}$$

$$=\frac{1}{\sqrt{I!}}\frac{1}{\sqrt{J!}}M(D_{z'}^I D_{z''}^I(p_T(z')p_T(z'')))|_{z'=z''=0}$$

$$=\frac{1}{\sqrt{I!}}\frac{1}{\sqrt{J!}}D_{z'}^I D_{z''}^J(Mp_T(z')p_T(z''))|_{z'=z''=0}$$

$$=\frac{1}{\sqrt{I!}}\frac{1}{\sqrt{J!}}D_{z'}^I D_{z''}^J(Mp_T(z'+z'')e^{z'z''})|_{z'=z''=0}. \tag{32}$$

(Here $z'z''=\sum_{k=1}^{\infty}z_k'z_k''$; the verification of the equality $P_T(z')P_T(z'')=P_T(z'+z'')e^{z'z''}$ is trivial.)

Remark. While going from the second row to the third in the chain (32) we interchanged the order of differentiation and mathematical expectation. This is legitimate if the derivatives are integrable (over w) uniformly in z', z'', which can be fulfilled if each z_n', z_n'' is bounded in advance by a fixed number, which is guaranteed by condition (31).

Recalling that

$$Mp_T(z)=Mp_T\left(\sum_{k=1}^{\infty}z_k m_k(t)\right)=Mp_T(m)=1$$

[9]We shall not distinguish here between the two sets of indices $I=(i_1,\dots,i_n)$ and $J=(j_1,\dots,j_m)$, $m>n$, if $i_k=j_k$ for $k\leqslant n$, $j_k=0$ for $k>n$.

if only $m(t)=\sum_{k=1}^{\infty} z_k m_k(t)\in M$, we proceed as follows:

$$M\xi_I\xi_J = \frac{1}{\sqrt{I!}}\frac{1}{\sqrt{J!}} D_{z'}^I D_{z''}^I e^{z'z''}\big|_{z'=z''=0}$$

$$= \frac{1}{\sqrt{I!}}\frac{1}{\sqrt{J!}} \prod_{k=1}^{n} \frac{\partial^{i_k+j_k}}{(\partial z_k')^{i_k}(\partial z_k'')^{j_k}} (e^{z_k' z_k''})\big|_{z'=z''=0}$$

$$= \prod_{k=1}^{n} \delta_{i_k j_k}, \qquad Q.E.D.$$

(b) *Totality.* Let $\eta\in H_2^w[0,T]$, and $M\eta\xi_I=0$ for all I. We shall show that $\eta=0$ (a.s.).

Since $p_T(\sum_{k=1}^{n} z_k m_k)$ is an analytic function of the variables $z_1,\ldots,z_n$, expanding this function as a Taylor series in z we have

$$p_T\left(\sum_{k=1}^{n} z_k m_k\right) = \sum_{i_1,\ldots,i_n=0}^{\infty} \frac{D_z^{(i_1,\ldots,i_n)} p_T\left(\sum z_k m_k\right)}{i_1!\ldots i_n!} z_1^{i_1}\ldots z_n^{i_n}$$

$$= \sum_{i_1,\ldots,i_n=0}^{\infty} \frac{z_1^{i_1}\ldots z_n^{i_n}}{\sqrt{i_1!\ldots i_n!}} \xi_{(i_1,\ldots,i_n)}.$$

Therefore

$$M\eta p_T\left(\sum_{k=1}^{n} z_k m_k\right) = \sum_{i_1,\ldots,i_n=0}^{\infty} \frac{z_1^{i_1}\ldots z_n^{i_n}}{\sqrt{i_1!\ldots i_n!}} M\eta\xi_{(i_1,\ldots,i_n)} = 0.$$

It is easy to derive from this that $M\eta p_T(m)=0$ for all $m\in L_2[0,T]\cap M$. By Lemma 2.1 we finally get $\eta=0$ a.s., thus proving the lemma.

3. As a corollary we have the following expansion: if $\eta\in H_2^w[0,T]$, then

$$\eta = \sum_T \xi_I(\eta,\xi_I) = \sum_I \xi_I M\eta \frac{1}{\sqrt{I!}} D_z^I p_T(z)\big|_{z=0}$$

$$= \sum_I \frac{1}{\sqrt{I!}} D_z^I(M\eta p_T(z))\big|_{z=0}\xi_I. \tag{33}$$

If η is F_t^w-measurable, because of the martingale property of $p_t(z)$ we may replace $M\eta p_T(z)$ by $M\eta p_t(z)$ in the right-hand side of Eq. (33). Furthermore, if $\eta=\eta(w)$ is an arbitrary random variable, a similar equality holds for $M\{\eta|F_t^w\}$:

$$M\{\eta|F_t^w\} = \sum \frac{1}{\sqrt{I!}} \xi_I D_z^I(M(M\{\eta|F_t^w\}p_t(z)))\big|_{z=0}$$

$$= \sum \frac{1}{\sqrt{I!}} D_z^I(M\eta p_t(z))\big|_{z=0}\xi_I \tag{34}$$

(convergence of all the series is in the mean-square). By Parseval's

theorem we get

$$M[M\{\eta|F_t^w\}]^2=\sum_I \frac{1}{I!}\left[D_z^I(M\eta p_t(z))\right]^2\Big|_{z=0}. \tag{35}$$

Finally we put $\eta=\varphi(x_t)$, in (34), (35), where x_t is an arbitrary (weak) solution of Eq. (1).

We denote by $u(s,x,z)$ a solution of Eq. (20) in which $m(s)=\sum_{k=1}^{\infty}z_k m_k(s)$.

Using Lemma 2.2, we put

$$M\eta p_t(z)=M\varphi(x_t)p_t(m)=u(0,x,z),$$

yielding

$$M\left[M\{\varphi(x_t)|F_t^w\}\right]^2=\sum_I \frac{1}{I!}\left[D_z^I u(0,x,z)\right]^2\Big|_{z=0}. \tag{36}$$

4. Now we have everything ready for us to formulate a necessary and sufficient condition for a strong solution of Eq. (1) to exist.

Theorem 2. Let $0\leqslant s\leqslant t\leqslant T$, $x\in R^1$, $z=(z_1,\dots,z_n,\dots)$.

Let $u(s,x,z)$ and $v(s,x)$ be solutions of the equations

$$\mathcal{L}^{x_t}u(s,x)+u'_x\sigma(s,x)m(s)=0, \text{ where } m(s)=\sum_1^{\infty}z_k m_k, \tag{37}$$

$$u(t,x)=x.$$

$$\begin{cases}\mathcal{L}^{(x_t)}v(s,x)=0,\\ v(t,x)=x^2,\end{cases} \tag{38}$$

respectively. And let $u(0,x,z)=u(x,z)$, $v(0,x)=v(x)$.

Eq. (1) has a strong solution for $t\leqslant T$ if and only if the identity

$$v(x)=\sum_I \frac{1}{I!}\left[D_z^I u(x,z)\right]^2\Big|_{z=0} \tag{39}$$

holds.

Remark. Although the foregoing necessary and sufficient condition is hard to verify, it has the advantage of being formulated in terms of the operator $L^{(x_t)}$ and Eqs. (37), (38). Thus the presence of a strong solution depends only on the coefficients $b(t,x)$ and $\sigma(t,x)$ and is not connected with an appropriate or inappropriate choice of σ-algebras, probability spaces, etc.

5. PROOF OF THEOREM 2.

(a) Setting $\varphi(x)=x$ in Eq. (36) we get

$$M[M(x_t|F_t^w)]^2=\sum_I \frac{1}{I!}\left[D_z^I u(x,z)\right]^2\Big|_{z=0}. \tag{40}$$

(b) *Necessity.* If x_t is a strong solution of Eq. (1), in the left-hand side of Eq. (40) we have $M\{x_t|F_t^w\}=x_t$ and $Mx_t^2=v(x)$, which thus yields Eq. (39).

(c) *Sufficiency.* Let Eq. (39) be satisfied. Since the right-hand sides in Eq. (39) and Eq. (40) coincide, their left-hand sides will coincide as well; i.e.,

$$Mx_t^2=v(x)=M\big[\,M\{x_t|F_t^w\}\,\big]^2.$$

Then we have (as in 3.4(a)) $x_t=M\{x_t|F_t^w\}$. *Q.E.D.*

6. **Remark.** This provides one more proof of Corollary 1.a (see 1). In fact, let Eq. (1) have a strong solution on some probability space. Then Eq. (39) is satisfied by this solution. Due to uniqueness in measure, Eq. (39) will also hold for any other solution. This means that this new solution is also strong.

3. Theorems on existence and uniqueness

1. Basic assertions. In this section we prove the following theorem on existence and uniqueness of strong solutions.

Theorem 3. Let the coefficients $b(t,x)$ and $\sigma(t,x)$ of Eq. (1) satisfy assumptions A, B, C (§1.6). Then the following conditions are sufficient for a strong solution to exist and to be pathwise unique.

1) The n-dimensional case: diffusion satisfies a Lipschitz condition in x. In addition both coefficients satisfy the Dini condition:

$$|\sigma(t,x_2)-\sigma(t,x_1)|\leqslant k|x_2-x_1| \tag{41}$$

$$\int_0^\varepsilon \frac{\rho(r)}{r}\,dr<\infty, \tag{42}$$

where $\rho(r)$ is the modulus of continuity of the functions $b(t,x)$, $\sigma(t,x)$ (with respect to a pair of arguments).

2) The one-dimensional case: diffusion satisfies a Hölder condition with exponent $\alpha\geqslant\frac{1}{2}$:

$$|\sigma(t,x_2)-\sigma(t,x)|\leqslant k|x_2-x_1|^\alpha, \qquad \alpha\geqslant\tfrac{1}{2} \tag{43}$$

and has bounded measurable drift.

3) The one-dimensional case, with $b(t,x)=b(x)$, $\sigma(t,x)=\sigma(x)$: the diffusion coefficient has constant sign and has bounded variation on each finite interval

$$\sigma(x)\geqslant\mu>0 \tag{44}$$

$$\operatorname{var}_{.[-N,N]}\sigma(x)=\sup_{-N=x_0<\cdots<x_n=N}\sum|\sigma(x_{k+1})-\sigma(x_k)|<\infty \tag{45}$$

with bounded measurable drift.

Remark.

a) By Theorem 1, we need only prove either the existence of a strong solution or the pathwise uniqueness of a (weak) solution. We prefer proving pathwise uniqueness.

b) Obviously it suffices to prove pathwise uniqueness for an arbitrarily small interval $[0, \varepsilon]$. Moreover, in that case we may replace in Eqs. (1) and (2) the initial condition $x_0 = x$ by $x_0 = \xi$, where ξ is a random variable. Splitting a long interval $[0, T]$ into small parts, we get uniqueness on the entire interval $[0, T]$. In the future we assume $T > 0$ to be sufficiently small (although this will be necessary only in §2).

2. A drift-eliminating transformation

1. **Lemma. 2.1** Let $b(t, x)$, $\sigma(t, x)$ satisfy conditions A, B, C (§1.6), let $T > 0$ be sufficiently small, and let the function $u: [0, T] \times R^n \to R^n$ be a solution of the equation

$$\begin{aligned} \mathcal{L}^{(x_t)} u(t, x) &= 0 \qquad t \in [0, T] \\ u(T, x) &= x. \end{aligned} \tag{46}$$

Then the function $u(t, x)$ has the following properties.

1) For any fixed $t \in [0, T]$ the function $u(t, x) = u_t(x)$ is a one-to-one mapping $u_t: R^n \to R^n$ onto R^n.

   We denote the inverse of $u_t(x)$ by $v_t(y) = v(t, y)$.

2) The functions $u(t, x)$ and $v(t, y)$ belong to $W_p^{1,2}$ for all $p < \infty$; in particular, the derivatives u_x' and v_y' satisfy a Hölder condition (with respect to t and x) with exponent $\alpha < 1$.

3) The derivatives $u_x'(t, x)$ and $v_y'(t, y)$ are bounded. Hence, the mappings $u(t, x)$ and $v(t, y)$ are quasi-isometric for each t: there is a constant $m > 0$ such that

$$m \leqslant \frac{u(t, x_2) - u(t, x_1)}{|x_2 - x_1|} = \frac{y_2 - y_1}{|v(t, y_2) - v(t, y_1)|} \leqslant \frac{1}{m}.$$

2. The detailed proof of Lemma 2.1 and an investigation of the properties of the mappings $u(t, x)$, $v(t, y)$ can be found in [24]. Here we outline only the basic steps in the proof.

(a) From the theory of parabolic differential equations (see [15] for instance), it follows that a solution of Eq. (46) exists, is unique and belongs to $W_p^{1,2}$ for any $p < \infty$.

(b) From embedding theorems (for instance, see [15], Lemma II.3.3) it follows that if $u \in W_p^{1,2}$ for $p > n + 2$, all the derivatives $(\partial u_i / \partial x_j)$ (t, x) satisfy a Hölder condition with respect to t and x. In that

case we can choose any Hölder exponent in the range $0<\alpha<1-\frac{(n+2)}{p}$; since p can be arbitrarily large, we can choose α as close to 1 as desired; the Hölder constant k for the domain $[0,T]\times D$ (here the region D is bounded) is to be estimated in terms of the norm of the function $\varphi(x)=x$ in the space $W_p^2(E)$, where E is a specially constructed sphere, $D\subseteq E$. However, due to the linearity of the operator $L^{(x_1)}$ the constant k should not change if we add an arbitrary constant to $\varphi(x)$, or, equivalently, if we move the sphere E to the coordinate origin.

Therefore, the derivatives $(\partial u_i/\partial x_j)(t,x)$ satisfy a Hölder condition with respect to t uniformly in x.

(c) Note that for $t=T$

$$\frac{\partial u_i}{\partial x_j}(T,x)=\frac{\partial \varphi_i}{\partial x_j}(x)=\left\{\begin{array}{ll} 1, & i=j \\ 0, & i\neq j \end{array}\right\};$$

i.e., $|u_x'(T,x)|=1$. Hence it follows from the Hölder property of the derivatives that for all $t\in[0,T]$ we have the estimate

$$|u_x'(t,x)|\leqslant \text{constant}. \tag{47}$$

(d) Note that for $t=T$,

$$\det u_x'(T,x)=\det \varphi'(x)=1,$$

and $\det u_x'(t,x)$ satisfies a Hölder condition with respect to t since it has been obtained by multiplication and addition operations from functions satisfying a Hölder condition. Hence if T is sufficiently small we have the estimate

$$\det' u_x''(t,x)\geqslant \text{const}>0. \tag{48}$$

(e) The estimates (47) and (48) imply the estimate

$$|\lambda_{\min}(t,x)|\geqslant \text{const}>0, \tag{49}$$

where $\lambda_{\min}(t,x)$ is an eigenvalue of minimum value of the matrix $u_x'(t,x)$; this yields

$$|u(t,x_2)-u(t,x_1)|\geqslant \text{const}|x_2-x_1|. \tag{50}$$

(f) Condition (50), as is shown in [9] (Corollary of Theorem II), is sufficient for the global invertibility of the mapping $u_t(x)$. We note that this fact is not trivial for dimension $n>1$. In particular, one inequality (48) is in general not sufficient for global invertibility of $u_t(x)$, although it is sufficient for local invertibility.

(g) Assertion 1 of the lemma has been proved. Assertion 3 follows from the estimates (47), (50). It remains only to prove that the inverse function $v(t,y)$ belongs to $W_p^{1,2}$. We can prove this by taking the limit of the smooth functions $u_n(t,x)$ and their inverses $v_n(t,y)$. The lemma is proved.

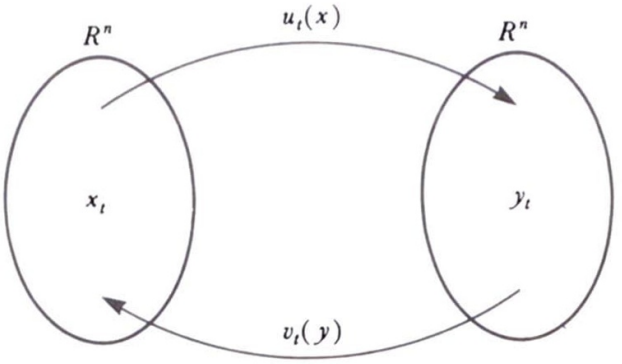

Figure 2

3. The mappings $u(t, x)$ and $v(t, y)$ constructed have a very useful property. Consider two processes x_t and y_t that can be obtained from each other according to the formula

$$\begin{aligned} y_t &= u(t, x_t), \\ x_t &= v(t, y_t). \end{aligned} \tag{51}$$

Let x_t satisfy the equation

$$\begin{cases} dx_t = b(t, x_t)\,dt + \sigma(t, x_t)\,dw_t \\ x_0 = \xi. \end{cases} \tag{52}$$

By the Itô formula (see §1, assertion 6.4) we then have

$$dy_t = \mathcal{L}^{(x_t)} u(t, x_t)\,dt + u'_x(t, x_t) \cdot \sigma(t, x_t)\,dw_t.$$

But $\mathcal{L}^{(x_t)} u(t, x) = 0$ by construction. Hence

$$dy_t = u'_x(t, x_t) \cdot \sigma(t, x_t)\,dw_t.$$

To convert the last equality into an equation for y_t we replace x_t by $v(t, y_t)$. We then find that y_t satisfies the equation

$$\begin{cases} dy_t = s(t, y_t)\,dw_t, \\ y_0 = \zeta, \end{cases} \tag{53}$$

where $\zeta = u(0, \xi)$; the matrix $s(t, y)$ is seen to be

$$s(t, y) = u'_x(t, v(t, y)) \cdot \sigma(t, v(t, y)). \tag{54}$$

We can formulate now a trivial lemma which essentially simplifies the proof of Theorem 3.

Lemma 2.3. A solution of Eq. (52) is pathwise unique if and only if Eq. (53) has a pathwise unique solution.

Proof. The proof is obvious.

4. **Lemma 2.4** The matrix $s(t, y)$ given by formula (54) satisfies conditions A, B, C (§1.6).

Proof.

(a) Continuity of the matrix $s \cdot s^* = u'_x \cdot \sigma \cdot \sigma^* \cdot (u'_x)^* = u'_x \cdot a \cdot (u'_x)^*$ follows from the continuity of the function $v(t, y)$ with respect to (t, y) and the continuity of the matrices $a(t, x)$, $u'_x(t, x)$, $(u'_x)^*(t, x)$ with respect to (t, x).

(b) Uniform positive definiteness of the matrix $s \cdot s^*$ follows from the uniform positive definiteness of the matrix $a(t, x)$ and the uniform nonsingularity of the matrix $u'_x(t, x)$.

(c) Boundedness of the matrix $s(t, y)$ follows from the boundedness of the matrices $\sigma(t, x)$ and $u'_x(t, x)$.

3. Proof of the first assertion of Theorem 3

1. **Lemma 3.1.** If $b(t, x)$ and $\sigma(t, x)$ satisfy Eqs. (42) and (41), the matrix $s(t, y)$ given by formula (54) satisfies a Lipschitz condition with respect to y.

Proof. As is shown in [8], if $b(t, x)$ and $\sigma(t, x)$ satisfy the Dini condition (42), then the second derivative $u''_{xx}(t, x)$ is bounded. Therefore $u'_x(t, x)$ satisfies a Lipschitz condition with respect to x, and the product $u'_x(t, x) \cdot \sigma(t, x)$ satisfies a Lipschitz condition with respect to x (since $\sigma(t, x)$ satisfies (41)). If we substitute x for $x = v(t, y)$, we find that $s(t, y)$ is Lipschitz with respect to y because $v(t, y)$ is Lipschitz with respect to y (see Assertion 3 of Lemma 2.1).

2. The validity of the first assertion of Theorem 3 follows in an obvious manner from the Itô theorem on uniqueness [6] and from Lemma 2.3.

4. Proof of the second assertion of Theorem 3

1. **Lemma 4.1** If $x \in R^1$ and $\sigma(t, x)$ satisfies condition (43), $s(t, y)$ satisfies a Hölder condition with respect to y, with the same exponent α.

Proof. The proof follows trivially from (43) and the second assertion of Lemma 3, if we recall that a product of two functions satisfying the Hölder condition with exponents α_1, α_2 respectively, satisfies the Hölder condition with exponent $\alpha = \min\{\alpha_1, \alpha_2\}$.

2. **Lemma 4.2.** (Yamada, Watanabe [24]). Let the coefficient $s(t, y)$ in Eq. (53) satisfy the condition

$$|s(t, y_2) - s(t, y_1)| \leq \rho(|y_2 - y_1|), \tag{55}$$

where $\rho(r)$ is a nonnegative increasing function such that

$$\int_0^{\varepsilon} \frac{1}{\rho^2(r)}\,dr = \infty. \tag{56}$$

Then a solution of Eq. (53) is pathwise unique.

Remark. The functions $\rho(r)=c\cdot|r|^{\alpha}$, $\alpha \geqslant \frac{1}{2}$, the functions $c\sqrt{r|\ln r|}$, $c\cdot\sqrt{r|\ln r|\cdot|\ln|\ln r||}$, etc. satisfy condition (56). But the functions $c\cdot|r|^{\frac{1}{2}+\varepsilon}$, $c\sqrt{r|\ln r|^{1+\varepsilon}}$, $c\sqrt{r|\ln r|\cdot|\ln|\ln r||^{1+\varepsilon}}$, etc., do not satisfy condition (56).

3. Proof of Lemma 4.2

(a) Let $a_0=1$, $a_k \downarrow 0$,

$$\int_{a_n}^{a_{n-1}} \frac{1}{\rho^2(r)}\,dr = k;$$

then

$$\int_{a_k}^{a_{k-1}} \frac{2}{k\rho^2(r)}\,dr = 2.$$

We construct the functions $\varphi_k(r)$ in the following ways (see Figure 3):

$$\varphi_k''(r) = \begin{cases} 0 & \text{for } r \leqslant a_k,\ r \geqslant a_{k-1}, \\ \text{between 0 and } 2/k\rho^2(r) & \text{for } a_k \leqslant r \leqslant a_{k-1}. \end{cases}$$

Here $\varphi_k''(r)$ is continuous, and the shaded region under it is equal to 1 (the region under $2/k\rho^2(r)$ on an interval $[a_k, a_{k-1}]$ is equal to 2). Further,

$$\varphi_k'(r) = \begin{cases} 0 & \text{for } r \leqslant a_k,\ r \geqslant a_{k-1}, \\ \text{between 0 and 1} & \text{for } a_k \leqslant r \leqslant a_{k-1}. \end{cases}$$

Let $\varphi_k(0)=0$. And continue $\varphi_k(r)$ to $(-\infty,0]$ by making it even. It is clear that the functions $\varphi_k(r)$ are twice continuously differentiable, and $\varphi_k(r)\uparrow|r|$ since $|r| \geqslant \varphi_k(r) \geqslant |r|-a_{k-1}$.

(b) Let y_t' and y_t'' be two solutions of Eq. (53) with the same Wiener process w_t and the same initial value ξ. Then

$$y_t' - y_t'' = \int_0^t (s(t, y_t') - s(t, y_t''))\,dw_t.$$

By the Itô formula,

$$\varphi_k(y_t' - y_t'') = \int_0^t \varphi_k'(y_t' - y_t'')[s(t, y_t') - s(t, y_t'')]\,dw_t + \int_0^t \tfrac{1}{2}\varphi_k''(y_t' - y_t'')[s(t, y_t') - s(t, y_t'')]^2\,dt = \mathfrak{I}_1 + \mathfrak{I}_2.$$

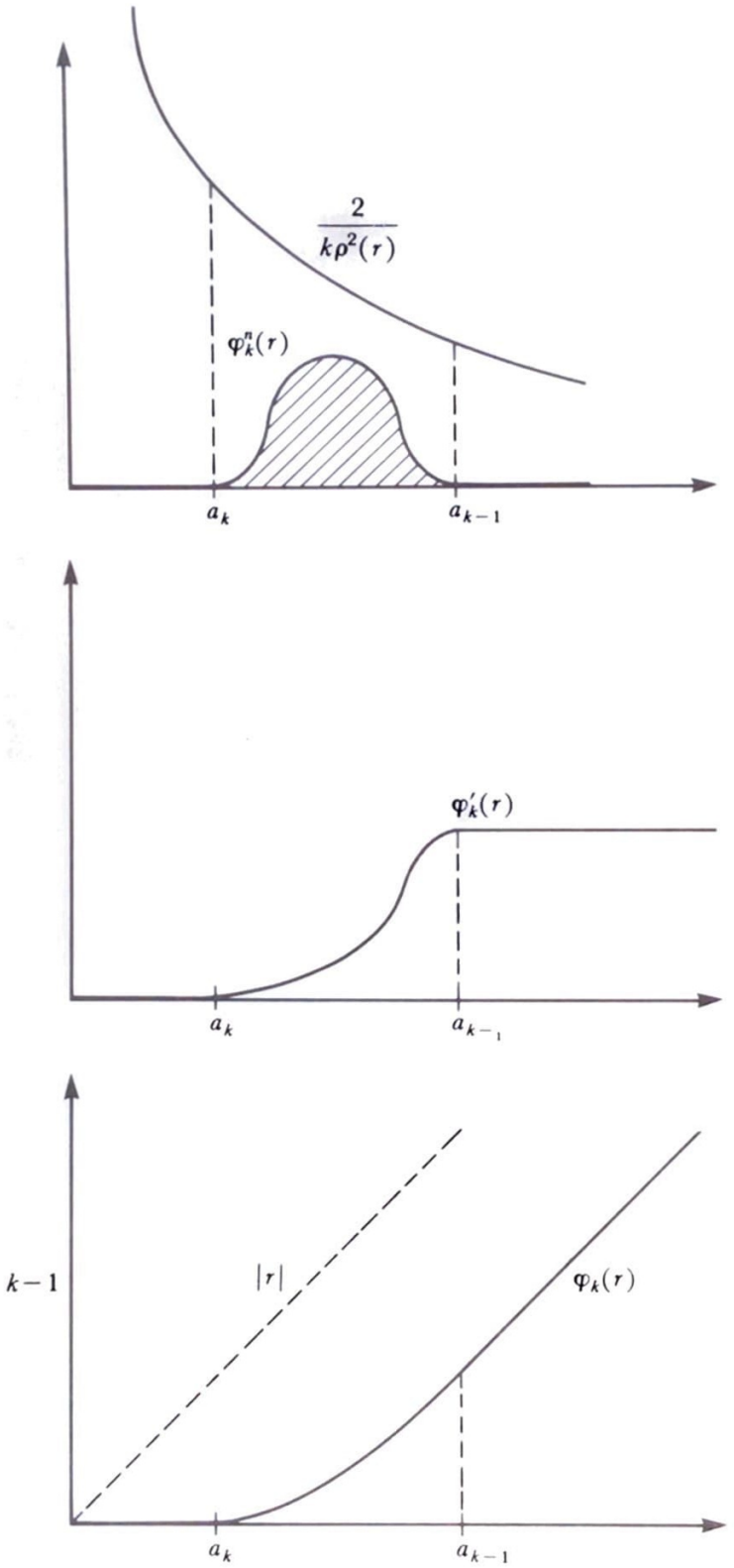
$\frac{2}{k\rho^2(r)}$
$\varphi_k^n(r)$
a_k
a_{k-1}
$\varphi_k'(r)$
a_k
a_{k-1}
$k-1$
$|r|$
$\varphi_k(r)$
a_k
a_{k-1}

Figure 3

(c) We have

$$M\vartheta_1 = 0$$

$$|\vartheta_2| \leqslant \tfrac{1}{2}\int_0^t |\varphi_k''(y_t' - y_t'')| \cdot |s(t, y_t') - s(t, y_t'')|^2 \, dt$$

$$\leqslant \tfrac{1}{2}t \cdot \max_{a_k < |r| < a_{k-1}} \varphi_k''(r)\rho^2(|r|) \leqslant \tfrac{1}{2} \cdot t \cdot \frac{2}{k} \to 0 \text{ as } k \to \infty.$$

We get $M\varphi_k(y_t' - y_t'') \to 0$ as $k \to \infty$, yielding $M|y_t' - y_t''| = 0$, *Q.E.D.*

4. **Remark.** In proving Lemma 4.2 we never used the nonsingularity of $s(t, y)$. Therefore, the assertion of the lemma remains valid for the degenerate equations (53).

5. *Proof of the third assertion of Theorem 3* (See [18])

1. **Remark.** Many discontinuous $\sigma(x)$ satisfy conditions (44), (45) of the third assertion of Theorem 3. Hence we cannot formally refer to either the assertions in §1.6 or Lemma 2.1, §3. Actually, all the assertions formulated above remain valid. To convince ourselves that this is so, we need, instead of the parabolic operator

$$\frac{\partial}{\partial t} + b(x)\frac{\partial}{\partial x} + \tfrac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2},$$

the elliptic operator

$$b(x)\frac{\partial}{\partial x} + \tfrac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}.$$

The properties of this elliptic operator do not change if we divide it by $\sigma^2(x)$; i.e., we go over the operator

$$\frac{b(x)}{\sigma^2(x)} \cdot \frac{\partial}{\partial x} + \frac{1}{2}\frac{\partial^2}{\partial x^2}.$$

In particular, in Lemma 2.1 we may take, instead of a solution of Eq. (46), the function

$$u(x) = \int_0^x \exp\left\{-2\int_0^y \frac{b(z)}{\sigma^2(z)}\,dz\right\} dy, \tag{57}$$

which satisfies all the needed properties.

2. **Lemma 5.2.** If $x \in R^1$, $\sigma(x)$ satisfies conditions (44), (45), then $s(y)$ is of bounded variation on each finite interval.

Proof. The proof follows trivially from the remark that u_x' in the formulas (54), (57) is of bounded variation.

3. **Lemma 5.3.** Let (ξ_t, v_t), $t\in[0,T]$ be a pair of continuous real random processes such that:

1) ξ_t is a martingale, $\xi_0=0$;
2) v_t is a process with a bounded (in the mean) variation:
$$\mathrm{M}\,\mathrm{var}_{[0,T]} v(t,\omega)<\infty, \tag{58}$$
where $\mathrm{var}_{[0,T]} v(t,\omega)$ is the total variation (in t) of the function $v(t)$ for fixed ω;
3) There exist constants $C>c>0$ such that on the set $\{(t,\omega)\colon \xi_t(\omega)>0\}$ the inequality
$$c\xi_t(\omega)\leqslant v_t(\omega)\leqslant C\xi_t(\omega) \tag{59}$$
holds.

Then $\xi_t=0$ with probability one for all $t\in[0,T]$.

Proof.

(a) Let $n(z,\omega)$ be the number of points $t\in[0,T]$ such that $v_t(\omega)=z$ (the function $n(z,\omega)$ has the possible values $0,1,2,\ldots,\infty$). By the Banach theorem (see [19]):
$$\mathrm{var}_{[0,T]} v_t(\omega)=\int_{-\infty}^{\infty} n(z,\omega)\,dz. \tag{60}$$

(b) Consider an interval $[a,b]=\left[\frac{1}{C}z,\frac{1}{c}z\right]$. Let $m(z,\omega)$ be the number of down-crossings of this interval by the martingale ξ_t. We see that
$$n(z,\omega)\geqslant m(z,\omega) \tag{61}$$
(since whenever $\xi_t=(1/c)z$, $v_t\geqslant z$, and when $\xi_t=(1/C)z$, $v_t\leqslant z$).

(c) We now estimate from below the mean of $m(t,\omega)$ for $t>0$. To this end consider the Markov times
$$\tau_0=0,$$
$$\tau_{2k+1}=\inf\{t\geqslant\tau_{2k}\colon \xi_t>b\}\wedge T$$
$$T_{2k+2}=\inf\{t\geqslant\tau_{2k+1}\colon \xi_t<a\}\wedge T$$
$$k=0,1,2,\ldots$$
We see from Figure 4 that
$$\sum_{k=1}^{\infty}(\xi_{\tau_{2k-1}}-\xi_{\tau_{2k}})=m(z,\omega)\cdot(b-a)+\zeta, \tag{62}$$
where $\zeta=0$ or $\zeta=(1/c)z-\xi_T$ depending on whether the number n of the last $\tau_n<T$ is odd or even.

Replace ζ by a larger value:
$$\zeta\leqslant\zeta'=\begin{cases}\zeta(\omega) & \text{if } \xi_T\geqslant b \text{ or } \zeta=0\\ b-a & \text{if } \xi_T<b,\end{cases}$$

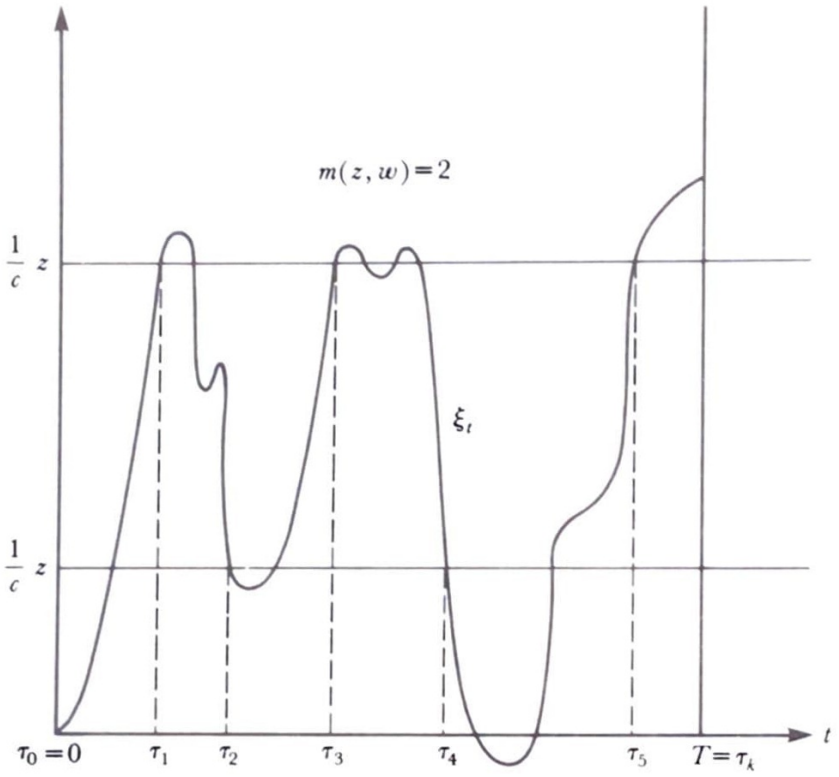

Figure 4

and take the mathematical expectation on both sides of Eq. (62). We get

$$0 \leqslant (b-a)(Mm(z,\omega)+1) - M\max\{\xi_T - b, 0\},$$

yielding

$$Mm(z,\omega) \geqslant \frac{M\max\{\xi_T - b, 0\}}{b-a} - 1 = \frac{M\max\left\{\xi_T - \frac{1}{c}z, 0\right\}}{\left(\frac{1}{c} - \frac{1}{C}\right)z} - 1.$$

(d) If $p\{\xi_T \neq 0\} > 0$, then $p\{\xi_T > 0\} > 0$. Therefore, there is $\delta > 0$ such that for all $z \in (0, \delta)$, $p\{\xi_T > (1/c)z\} > 0$; i.e., there is $\varepsilon > 0$ such that for all $z \in (0, \delta)$ we have the inequality

$$M\max\left\{\xi_T - \frac{1}{c}z, 0\right\} \geqslant \varepsilon,$$

whence

$$Mm(z,\omega) \geqslant \frac{\varepsilon}{kz} - 1, \tag{63}$$

where $k = (1/c) - (1/C)$.

(e) Combining inequalities (58), (60), (61), and (63), we get

$$\infty > M\mathrm{var}_{[0,T]}\, v_t(\omega) = \int_{-\infty}^{\infty} Mn(z,\omega)\, dz$$

$$\geqslant \int_0^{\delta} Mm(z,\omega)\, dz \geqslant \int_0^{\delta}\left(\frac{\varepsilon}{kz} - 1\right) dz = \infty.$$

The resulting contradiction shows that $P\{\xi_T = 0\} = 1$, yielding $P\{\forall t \leqslant T.\ \xi_t = 0\} = 1$. The lemma is proved.

4. **Lemma 5.4.** Let the coefficients $s(y)$ in Eq. (53) satisfy the following conditions:

$$0 < k \leqslant s(y) \leqslant K, \tag{64}$$

$$\operatorname{var}_{[-N,N]} s(y) < \infty \tag{65}$$

on any interval $[-N, N]$. Then the solution of Eq. (53) is pathwise unique.

Proof.

(a) Let y_t' and y_t'' be two solutions of Eq. (53), with the same Wiener process w_t and the same initial value ζ. Let

$$\tau = \inf\{t : |y_t'| = N \text{ or } |y_t''| = N\}.$$

It is clear that when $N \to \infty$, $\tau = \tau(N) \to \infty$. From now on, in order not to complicate the notation we shall write t instead of the Markov time $t \wedge \tau$. We consider our problem on the interval $[-N, N]$.

(b) To apply Lemma 5.3. we need to consider the pair of processes

$$\xi_t = \int_0^t [s(y_t') - s(y_t'')]\, dw_t, \tag{66}$$

$$v_t = h(y_t') - h(y_t''), \tag{67}$$

where $h(y) = \int_0^y \frac{1}{s(z)} \cdot dz$.

We see that ξ_t is a martingale, $\xi_0 = 0$. Next we estimate the variation of the process v_t.

(c) The function $h(y)$ is not twice differentiable. Hence it is not possible to apply the Itô formula immediately to the function $h(y)$. Consider a sequence of functions $s_n(y)$ such that:

$$s_n(y) \to s(y) \quad \text{in } L_1[-N, N]$$

$$s_n(y) \in C^1(R) \qquad \forall n,$$

$$k \leqslant s_n(y) \leqslant K \qquad \forall n,$$

$$\|s_n'(y)\|_{L_1[-N,N]} \leqslant \operatorname{var}_{[-N,N]} s(y) \qquad \forall n.$$

And let

$$h_n(y) = \int_0^y \frac{1}{s_n(z)}\, dz.$$

By the Itô formula,

$$h_n(y_t) = h_n(y_0) + \int_0^t \frac{s(y_t)}{s_n(y_t)}\, dw_t - \frac{1}{2} \cdot \int_0^t \frac{s_n'(y_t)}{s_n^2(y_t)} s^2(y_t)\, dt. \tag{68}$$

(d) Since $h_n(y)\to h(y)$ as $n\to\infty$ uniformly on $[-N, N]$, $h_n(y_t)\to h(y_t)$ (a.s.) and $h_n(y_0)\to h(y_0)$ (a.s.).

Further,

$$M\left|w_t - \int_0^t \frac{s(y_t)}{s_n(y_t)}\,dw_t\right|^2 = M\int_0^t \left(1 - \frac{s(y_t)}{s_n(y_t)}\right)^2 dt$$

$$\leqslant K_1 \left\| 1 - \frac{s(y)}{s_n(y)} \right\|^2_{L_2[-N,N]} \to 0 \quad \text{as } n\to\infty.$$

(Here the inequality follows from the estimate (16), see §1.6.3. We have convergence to zero due to the convergence $s_n(y)\to s(y)$ and the boundedness of $s_n(y)$.) Therefore, the stochastic integral in the right side of equality (68) converges as $n\to\infty$ (in the mean-square; because the interval $[-N, N]$ is bounded, it converges almost surely) to the Wiener process w_t.

Therefore, the last term in Eq. (68) converges almost surely to some process U_t. As a result we have for $h(y)$

$$h(y_t) = h(y_0) + w_t + U_t.$$

(e) Further, we estimate the variation of U_t. We have

$$M\,\mathrm{var}_{[0,T]} U_t \leqslant \frac{1}{2} M\int_0^T \left|\frac{s_n'(y_t)}{s_n^2(y_t)} s^2(y)\right| dt \leqslant K_2 \left\| \frac{s_n'(y)}{s_n^2(y)} s^2(y)\right\|_{L_1(-N,N)}$$

$$\leqslant K_3 \| s_n'(y)\|_{L_1[-N,N]} \leqslant K_3 \cdot \mathrm{var}_{[-N,N]}\, s(y) < \infty.$$

Therefore, the process

$$v_t = h(y_t') - h(y_t'') = U_t' - U_t''$$

also has a bounded (in the mean) variation.

(f) By the inequalities

$$\frac{1}{k} \leqslant h'(y) = \frac{1}{s(y)} \leqslant \frac{1}{k},$$

we have for $y_2 \geqslant y_1$:

$$\frac{1}{k}(y_2 - y_1) \leqslant h(y_2) - h(y_1) \leqslant \frac{1}{k}(y_2 - y_1),$$

$$\frac{1}{k}\xi_t \leqslant v_t \leqslant \frac{1}{k}\xi_t.$$

Then, by Lemma 6.3., $P\{\xi_t = 0\} = 1$, thus completing the proof of Theorem 3.

4. Additional results

In this section we state, though without proof, some additional results that strengthen and generalize the results obtained in the preceding sections. Not all the results below will be fully justified. Still it is possible to see how to go about proving them.

1. Representation of a strong solution in terms of Itô multiple integrals

1. It is known (see [7]) that any F_t^w-measurable random variable is representable as a series of Itô multiple integrals of nonrandom functions. Here we shall write a specific representation for the random variable x_t, where x_t is a strong solution of Eq. (1).

2. Denote by $T_{s,t}^m\varphi(x)$ the value of the solution $u(s,x)$ of the equation

$$\begin{cases} \dfrac{\partial u}{\partial s}+\dfrac{1}{2}\sigma^2(s,x)\dfrac{\partial^2 u}{\partial x^2}+[b(s,x)+m(s)\sigma(s,x)]\dfrac{\partial u}{\partial x}=0, \\ u(t,x)=\varphi(x) \end{cases} \tag{69}$$

at the point (s,x).

For $m(s)\equiv 0$ write $T_{s,t}\varphi(x)$. Let us rewrite Eq. (69):

$$\mathcal{L}^{(x_t)}u(s,x)=-m(s)\sigma(s,x)\frac{\partial u}{\partial x}(s,x). \tag{70}$$

Recalling how the solution of a nonhomogeneous parabolic equation may be expressed in terms of the solution of a homogeneous parabolic equation, we have:

$$T_{s,t}^m\varphi(x)\equiv u(s,x)=T_{s,t}\varphi(x)+\int_s^t T_{s,r}m(r)\cdot\sigma(r,\cdot)\frac{\partial}{\partial x}u(r,\cdot)(x)\,dr. \tag{71}$$

Let

$$\sigma(r,x)\frac{\partial}{\partial x}u(r,x)=\mathcal{D}u(r,x).$$

We shall solve Eq. (71) by an iteration method. We get

$$T_{s,t}^m\varphi(x)=T_{s,t}\varphi(x)+\sum_{n=1}^{\infty}\int\dots\int_{s<r_1<\dots r_n<t} T_{s,r_1}\mathcal{D}T_{r_1,r_2}\mathcal{D}$$
$$\times\dots\mathcal{D}T_{rn,t}\varphi(x)m(r_1)\dots m(r_n)\,dr_1\dots dr_n.$$

We also note that for any bounded nonrandom function $h(r_1,\dots,r_n)$ by the Girsanov theorem [4] we find that

$$\begin{aligned} Mp_t(m)&\int\dots\int_{s<r_1\dots<r_n<t} h(r_1,\dots r_n)\,dw_{r_1}\dots dw_{r_n} \\ &=M\int\dots\int_{s<r_1\dots<r_n<t} h(r_1,\dots r_n)\,(dw_{r_1}+m(r_1)\,dr_1)\dots(dw_{r_n}+m(r_n)\,dr_n) \\ &=\int\dots\int_{s<r_1\dots<r_n<t} h(r_1,\dots r_n)m(r_1)\,dr_1\dots m(r_n)\,dr_n. \end{aligned}$$

Therefore

$$T_{s,t}^m\varphi(x)=Mp_t(m)\Bigg[T_{s,t}\varphi(x)+\sum_{n=1}^{\infty}\int\dots\int_{s<r_1\dots<r_n<t}$$
$$\times T_{s,r_1}\mathcal{D}\dots\mathcal{D}T_{r_n,t}\varphi(x)\,dw_{r_1}\dots dw_{r_n}\Bigg]. \tag{72}$$

On the other hand, by Lemma 2.2, §2,

$$T^m_{0,t}\varphi(x) = M p_t(m)\cdot\varphi(x_t) = M\big[\, p_t(m)\cdot M\{\varphi(x_t)|F^w_t\}\,\big]. \tag{73}$$

By Lemma 2.1, §2, it follows from Eqs. (72) and (73) that for any (strong or weak) solution of Eq. (1)

$$M\{\varphi(x_t)|F^w_t\} = T_{0,t}\varphi(x) + \sum_{n=1}^{\infty}\int\cdots\int_{0<r_1<\cdots r_n<t} T_{0,r_1}\mathfrak{D}\ldots\mathfrak{D}T_{r_n,t}\varphi(x)\,dw_{r_1}\ldots dw_{r_n}. \tag{74}$$

Then, if x_t is a strong solution, and if $h(x)\equiv x$, we have

$$x_t = T_{0,t}h(x) + \sum_{n=1}^{\infty}\int\cdots\int_{0<r_1<\cdots r_n<t} T_{0,r_1}\mathfrak{D}\ldots\mathfrak{D}T_{r_n,t}h(x)\,dw_{r_1}\ldots dw_{r_n}, \tag{75}$$

yielding an "explicit" solution of Eq. (1).

3. Note that the terms in the right-hand side of formula (74) are orthogonal. Hence

$$M\big[\, M\{\varphi(x_t)|F^w_t\}\,\big]^2 = [\,T_{0,t}\varphi(x)\,]^2 + \sum_{n=1}^{\infty}\int\cdots\int_{0<r_1\cdots<r_n<t} \times\big[\, T_{0,r_1}\mathfrak{D}\ldots\mathfrak{D}T_{r_n,t}\varphi(x)\,\big]^2 dr_1\ldots dr_n. \tag{76}$$

This yields a representation for $M[\,M\{\varphi(x_t)|F^w_t\}\,]^2$ different from that in Eq. (36). By the way, it might be possible to prove that the right-hand sides coincide in Eq. (36) and Eq. (76) respectively without invoking probability theory.

Eq. (76) implies another criterion for the existence of a strong solution. The condition $Mx_t^2 = M[\,M\{x_t|F^w_t\}\,]^2$, necessary and sufficient for a solution x_t to be strong, can now be written as:

$$T_{0,t}h^2(x) = (T_{0,t}h(x))^2 + \sum_{n=1}^{\infty}\int\cdots\int_{0<r_1\cdots<r_n<t}\big[\, T_{0,r_1}\mathfrak{D}\ldots\mathfrak{D}T_{r_n,t}h(x)\,\big]^2 dr_1\ldots dr_n,$$

where $h(x)\equiv x$.

2. *The drift-eliminating transformation.* The mutually inverse transformations $u(t, x)$ and $v(t, y)$ constructed in §3.2 may appear useful not only in constructing strong solutions but in other situations as well. We show in what directions it is possible to weaken the restrictions on the coefficients.

1) Note first that a similar property holds not just for small T but for any $T<\infty$. In that case, though, to estimate the Jacobian $\det u'_x$ from below it does not suffice to invoke the embedding theorem—more refined probabilistic methods are needed.

2) The boundedness of the coefficients $b(t, x)$ and $\sigma(t, x)$ guarantees that the solution $u(t, x)$ of Eq. (46) lies in $W_p^{1,2}$ for all $p<\infty$. However, as one sees from the proof of Lemma 2.1, §3, it suffices to find a solution in any $W_p^{1,2}$ with $p>n+2$. Therefore, we may require only that the drift coefficient $b(t, x)$ belong locally to some L_p with $p>n+2$ (see [15], Theorem IV, 9.1).
3) It is possible to make only such restrictions on the behavior of the coefficients $b(t, x)$ and $\sigma(t, x)$ at infinity that guarantee the "absence of explosions" (see [17] and [10] for a criterion for the absence of explosions). For example, it suffices to require only that the coefficients grow linearly.
4) Evidently, the invertibility of $u(t, x)$ for each t will hold whenever the measure generated by the process x_t is equivalent to the measure generated by a process with diffusion $\sigma(t, x)$ and zero drift.

We can also make similar changes in Theorem 3.

3. *Conditions for pathwise uniqueness*

1. Let $b(t, x)\equiv 0$ in Eq. (1). In [24] conditions for strong uniqueness are given, formulated in terms of the modulus of continuity of $\sigma(t, x)$ in x. They are:

Let

$$|\sigma(t, x)-\sigma(t, y)|\leqslant\rho(|x-y|),$$

where the function $\rho(r)$ is defined for $r\in[0, \infty)$, $\rho(0)=0$. The following conditions suffice for strong uniqueness of a solution of the equation

$$\begin{cases} dx_t=\sigma(t, x_t)\,dw_t, \\ x_0=x; \end{cases} \tag{77}$$

(a) $x\in R^1$; it is then required that $\forall\varepsilon>0$

$$\int_0^\varepsilon \frac{1}{\rho^2(r)}\,dr=\infty. \tag{78}$$

The functions $\rho(r)=c\sqrt{r}$; $c\sqrt{r}\,|\ln r|^{\frac{1}{2}+\varepsilon}$; $c\sqrt{r|\ln r|}\cdot|\ln|\ln r||^{\frac{1}{2}+\varepsilon}$, etc. satisfy condition (a). If we denote by M_ρ the class of functions satisfying condition (a), and by M_α the class of functions satisfying the Hölder condition with indicator α, $0<\alpha\leqslant 1$, then the following inclusions are strict

$$M_{\frac{1}{2}}\subset M_\rho\subset\bigcup_{\alpha<\frac{1}{2}} M_\alpha.$$

The corresponding uniqueness theorem was actually proved in §3, see Lemma 2.2, §3.

(b) $x\in R^2$; in this case it is required that the function $G(z)=z^3e^{2/z}\rho^2(e^{-1/z})$ be upward convex and that $\forall\varepsilon>0$

$$\int_0^\varepsilon \frac{r|\ln r|}{\rho^2(r)}dr=\infty. \tag{79}$$

Examples of functions satisfying the condition in (b) are: $\rho(r)=cr$; $cr|\ln r|$; $cr|\ln r|\sqrt{|\ln|\ln r||}$; $cr|\ln r|\sqrt{|\ln|\ln r||}\ \sqrt{|\ln|\ln|\ln r|||}$.

(c) $x\in R^n$, $n\geqslant 3$; in this case it is required that the function $\rho^2(r)/r$ be upward convex and that $\forall\varepsilon>0$,

$$\int_0^\varepsilon \frac{r}{\rho^2(r)}dr=\infty. \tag{80}$$

Examples of functions satisfying the conditions in (c) are: $\rho(r)=cr$; $cr\sqrt{|\ln r|}$; $cr\sqrt{|\ln r|\cdot|\ln|\ln r||}$.

In cases (b) and (c) the following strict inclusions hold:

$$M_1\subset M_\rho\subset\bigcap_{\alpha<1}M_\alpha.$$

Here we give no proofs of assertions (b) and (c) because first, they are cumbersome, second, they follow in general the proof of Lemma 4.2, §3, third, they weaken a Lipschitz condition only "very, very little" (as will be shown, they do not weaken the Lipschitz condition enough that we can reject the Dini condition in Theorem 3.1 without changing the proof).

2. Conditions (a), (b), (c) are unimprovable (in terms of the modulus of continuity), as is illustrated by the following example.

Example (Yamada, Watanabe [24]). We consider the case $n\geqslant 3$. Let the function $\rho(r)$ be upward convex, and let

$$\int_0^\varepsilon \frac{r}{\rho^2(r)}dr<\infty.$$

Further, let $\sigma_{ij}(t,x)=\delta_{ij}\cdot\rho(|x|)$ (then $|\sigma_{ij}(t,y)-\sigma_{ij}(t,x)|\leqslant|\rho(|y|)-\rho(|x|)|\leqslant\rho(|y-x|)$).

Considering the equation

$$\begin{cases} dx_t=\sigma(x_t)\,dw_t, \\ x_0=0, \end{cases} \tag{81}$$

we can see that this equation always (for any Wiener process) has the solution $x_t\equiv 0$. Let us construct one more, nonzero solution of Eq. (81).

Let (w_t, F_t) be a Wiener process. Consider the function

$$f(t)=f(t,\omega)=\int_0^t \frac{1}{\rho^2(|w_s|)}\,ds;$$

we will show that $Mf(t,\omega)<\infty$. In fact,

$$Mf(t,\omega)=\int_0^t\left(M\frac{1}{\rho^2(|w_s|)}\right)ds=k\int_0^t\left(\int_0^\infty \frac{1}{\rho^2(r)}\,\frac{1}{(2\pi s)^{n/2}}\,e^{-r^2/2s}\cdot r^{n-1}\,dr\right)ds$$

$$=k\int_0^\infty \frac{r^{n-1}}{\rho^2(r)}\left(\int_0^t \frac{1}{(2\pi)^{n/2}s^{n/2}}\,e^{-r^2/2s}\,ds\right)dr$$

$$\leqslant \text{const.}\left(1+\int_0^\varepsilon \frac{r}{\rho^2(r)}\,dr\right)<\infty.$$

Here we have used the fact that for small r the integral

$$\int_0^t \frac{1}{s^{n/2}}\,e^{-r^2/2s}\,ds$$

has the order of $1/r^{n-2}$.

We change the time-scale using the function $f(t,\omega)$: consider the function $g(h)=g(h,\omega)$, the inverse of the function $f(t,\omega)$ for each ω:

$$g(h)=\min\{t\colon f(t)=h\},$$

and consider the process $x_h=w_{g(h)}$. For this process we have

$$\langle x_n^i, x_n^j\rangle=\langle w^i_{g(h)}, w^j_{g(h)}\rangle=\delta_{ij}g(h)=\delta_{ij}\int_0^h \rho^2(|x_s|)\,ds,$$

which can be satisfied because $g(0)=0$, and

$$g'(h)=1/f'(g(h))=\left(\rho^2|w_{g(h)}|\right)=\left(\rho^2|x_h|\right),$$

yielding

$$\langle x_n^i, x_n^j\rangle=\delta_{ij}\int_0^h \sigma_{ij}^2(x_s)\,ds.$$

In a similar way we can show that the process $\tilde{w}_h=\int_0^h \sigma^{-1}(x_s)\,dx_s$ is Wiener. Therefore, we have two solutions of Eq. (81): $(0,\tilde{w}_n)$ and $(x_h,\tilde{w}_h)$.

Remark 1. In dimensions $n=1,2$, examples can be constructed in the same way; the inequality $Mf(t,\omega)<\infty$ can be obtained when conditions (78), (79) are violated.

Remark 2. In the example constructed above we used essentially the singularity of the diffusion coefficient $\sigma(x)$; in Theorem 3 the diffusion is nonsingular. However, we do not know how to make use of this nonsingularity to prove strong uniqueness.

3. If the drift-coefficient satisfies the Dini condition, the matrix $u_s'(t,x)$, as was mentioned, satisfies a Lipschitz condition in x. Hence the continu-

ity modulus of $s(t, y)$ (see Eq. (54)) changes, compared to the continuity modulus of $\sigma(t, x)$, by no more than a constant number of times. Hence, if the drift satisfies the Dini condition, it is possible to replace the conditions imposed on the diffusion coefficient in Theorem 3 by conditions (a), (b), (c) formulated in 3.1.

In the one-dimensional case, it suffices that $u'_x(t, x)$ satisfy the Hölder condition with exponent $\alpha \geqslant \frac{1}{2}$. Hence, instead of measurability and boundedness of $b(t, x)$ we may require that $b(t, x)$ be measurable and belong locally to L_p, where $p > 6$. If the coefficients do not depend on t, it suffices that $p > 3$.

It is possible to require that conditions A, B, C be satisfied not in the entire space but only in each finite region, with constants probably depending on the region—we then have strong uniqueness.

4. Regretfully, we have not yet succeeded in deleting the Dini condition from Theorem 3.1. Indeed, as the examples in [13] show, for $n > 2$ even if $\sigma(t, x)$ is a unit matrix and $b(t, x)$ is continuous, u''_{xx} may still turn out to be unbounded. In that case the unboundedness may be such that the new diffusion coefficient $s(t, y) = u'_x(t, v(t, y))$ will have a continuity modulus in y, which was prohibited by Example 3.2.

4. *Drift depending on the past.* For the equation

$$\begin{cases} dx_t = \alpha(t, x_0^t)\, dt + dw_t, \\ x_0 = x, \end{cases} \tag{82}$$

where $\alpha(t, x_0^t)$ is a measurable nonanticipatory bounded functional, we formulate two sufficient conditions for strong uniqueness (each condition is sufficient):

1) A functional Lipschitz condition:

$$|\alpha(t, x_0^t) - \alpha(t, y_0^t)| \leqslant k \|x_0^t - y_0^t\|, \tag{83}$$

where the norm is taken in the space $C_{[0,t]}$.

2) Delay: there is $\varepsilon > 0$ such that for all t the random variable $\alpha(t, x_0^t)$ is measurable with respect to the σ-algebra $F^x_{t-\varepsilon}$ (that is, $\alpha(t, x_0^t) = \alpha(t, x_0^{t-\varepsilon})$).

The proof of uniqueness under condition 1 coincides with the proof of the Itô theorem [1]. To prove that a solution is F_t^w-measurable under condition 2 we need to split the interval $[0, T]$ into intervals of length ε.

5. *Three unsolved problems.* In concluding we state three problems which have not yet been solved but which seem most interesting.

Problem 1. Let the diffusion coefficient $\sigma(t, x)$ be continuous with respect to both arguments, bounded, and uniformly nonsingular. The

problem is to prove that the equation

$$\begin{cases} dx_t = \sigma(t, x_t)\, dw_t, \\ x_0 = x \end{cases} \tag{84}$$

has a strong solution.

Problem 2. Let the drift coefficient $b(t, x)$ be bounded and measurable. The problem is to prove that the equation

$$\begin{cases} dx_t = b(t, x_t)\, dt + dw_t, \\ x_0 = x \end{cases} \tag{85}$$

has a strong solution.[10]

Problem 3. Let $x \in R^1$ be one-dimensional. And let $\alpha(t, x_0^t)$ be a measurable bounded nonanticipatory functional. The problem is to prove that the equation

$$\begin{cases} dx_t = \alpha(t, x_0^t)\, dt + dw_t, \\ x_0 = x \end{cases} \tag{86}$$

has a strong solution.[11]

Remark.

a) Problem 3 is known as the "innovation problem". Its solution would essentially simplify the proofs of many results in filtering theory (see [16]).

b) Problem 2 follows from Problem 1. In the one-dimensional case Problem 2 has been solved (Theorem 3.2). In the multivariate case, to solve Problem 2 it would suffice to prove strong uniqueness for Eq. (84) with the coefficient $\sigma(t, x) \in \bigcap_{\alpha < 1} M_\alpha$, where M_α is the class of Hölder functions in x with exponent α.

c) Problem 1 has been formulated implicitly in Strook and Varadhan [23].

References

1. N. Dunford and J. T. Schwartz, Linear Operators, I, General Theory, New York, Interscience Publishers, 1958.
2. I. M. Gelfand, Lectures on Linear Algebra, New York, Interscience Publishers, 1961.
3. I. I. Gikhman and A. V. Shorokhod, Stochastic Differential Equations, New York, Springer Verlag, 1972. Originally published by Naukova Dumka, Kiev, 1968.

[10] This problem is solved positively by A. Yu. Veretennikov in Matem. Sbornik 111 (1980), 434–452.

[11] This problem is solved negatively by Cirelson, cf. [16].

4. I. V. Girsanov, *On transforming a certain class of stochastic processes by absolutely continuous substitutions of measures*, Theory of probability and its applications 5 (1960), 285–300.
5. Ph. Hartman, Ordinary differential equations, New York, Wiley, 1964.
6. K. Itô, On stochastic differential equations, New York, American Mathematical Society, 1951.
7. _____, *Multiple Wiener Integrals*, J. Math. Soc. Japan 3 (1951), 157–169.
8. M. D. Ivanovich, *On the nature of continuity of solutions of second-order parabolic equations*, Vestnik MGU, Ser. Mat. 4 (1966), 31–41.
9. F. John, *On quasi-isometric mapping, I*, Comm. Pure Appl. Math. 21 (1968), 77–110.
10. P. Z. Khasminsky, Stability of systems of differential equations under random perturbations of their parameters, Moscow, Nauka, 1969.
11. S. N. Kruzhkov, *On estimates of higher derivatives for solutions of elliptic and parabolic equations with continuous coefficients*, Matem. Zam. 2 (1967), 549–560.
12. N. V. Krylov, *On Itô's stochastic integral equations*, Theory of probability and its applications 14 (1969), 330–336.
13. _____, *The regularity of conditional probabilities for stochastic processes*, Theory of probability and its applications 18 (1973), 150–153.
14. _____, Diffusion Control Processes, New York, Springer Verlag, 1980.
15. O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Uraltzeva, Linear and quasi-linear equations of parabolic type, Providence, American Mathematical Society, 1968. Originally published by Nauka, Moscow, 1967.
16. R. S. Liptser and A. N. Shiryayev, Statistics of Random Processes, I, II, New York, Springer-Verlag, 1977-1978. Originally published by Nauka, Moscow, 1974.
17. H. P. McKean, Stochastic Integrals, New York, Academic Press, 1969.
18. S. Nakao, *On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations*, Osaka J. Math. 9 (1972), 513–518.
19. S. Saks, Theory of the integral, New York, G. E. Stechert & Co., 1937.
20. L. Schwartz, Analysis, Vol. I, Moscow, Mir, 1972.
21. A. V. Skorokhod, *On existence and uniqueness of solutions of stochastic diffusion equations*, Sibirsky Matem. J., 2 (1961), 129–137.
22. S. L. Sobolev, Applications of functional analysis in mathematical physics, Providence, American Mathematical Society, 1963. Originally published by LGU, Leningrad, 1950.
23. D. W. Strook and S. R. S. Varadhan, *Diffusion processes with continuous coefficients, I*, Comm. Pure Appl. Math. 22 (1963), 345–400.
24. T. Yamada, S. Watanabe, *On the uniqueness of solutions of stochastic differential equations, I, II*, J. Math. Kyoto Univ 11 (1971), 155–167; 11 (1971), 553–563.
25. A. K. Zvonkin, *Transformation of a phase space of a drift eliminating diffusion process*, Matem. Sbornik, 93 (1974), 129–149.