STOCHASTIC DIFFERENTIAL EQUATIONS AND MODELS OF RANDOM PROCESSES

E. J. McSHANE
UNIVERSITY OF VIRGINIA

1. Description of a desirable model

Let us suppose that we are investigating a system whose state can be adequately specified by n real numbers $x^1, \cdots, x^n$. We shall suppose that by some acceptable scientific theory it is predicted that, in the absence of disturbances from outside the system, the x^i develop in time in accordance with certain differential equations,

$$(1.1) \qquad \dot{x}^i = g_0^i(t, x), \qquad\qquad i = 1, \cdots, n.$$

If there are disturbances or noises, $n^1(t), \cdots, n^r(t)$, the underlying theory of such systems will often permit us to conclude that

$$(1.2) \qquad \dot{x}^i = g_0^i(t, x) + \sum_{\rho=1}^{r} g_\rho^i(t, x) n^\rho(t), \qquad\qquad i = 1, \cdots, n,$$

where g_ρ^i is the sensitivity of the ith coordinate to the ρth noise. However in the underlying theory, equation (1.2) will usually have a limited domain of applicability; in particular, we could not usually retain confidence in the trustworthiness of (1.2) if the noise were unbounded. But for sufficiently well-behaved bounded noises we can rewrite (1.2) in the form

$$(1.3) \qquad dx^i = g_0^i(t, x)\, dt + \sum_\rho g_\rho^i(t, x)\, dz^\rho,$$

or

$$(1.4) \qquad x^i(t) = x_0^i + \int_a^t g_0^i[s, x(s)]\, ds + \sum_\rho \int_a^t g_\rho^i[s, x(s)]\, dz^\rho(s),$$

where

$$(1.5) \qquad z^\rho(t) = z^\rho(a) + \int_a^t n^\rho(s)\, ds;$$

with bounded n^ρ or Lipschitzian z^ρ, these are solvable by traditional methods, and (perhaps with still stronger requirements on the z^ρ) will describe the evolution of the system with as much certainty as the underlying scientific theory of such systems permits.

Usually however, we are interested, not in the response of the system to specified noises z^ρ, but in statistical properties of the responses of the system to random noises. As is well known, this causes a dilemma. The processes z^ρ most amenable to probabilistic study are martingales, especially the Wiener process and closely related processes. But these have almost surely non-Lipschitzian sample functions and lie outside the domain of applicability of the scientific theory that led to (1.4). The integrals with respect to z^ρ in (1.4) cannot even be interpreted as Riemann-Stieltjes or Lebesgue-Stieltjes integrals. Interpreting them as Itô integrals restores meaning to all terms in (1.4), but gives no ground for confidence that the solution (1.4) will continue to represent the time development of the system. It is a familiar fact that the uncritical use of (1.4) can lead to mismatches between system and model that are often considered paradoxical.

E. Wong and M. Zakai have made a major contribution [8], [9] to the removal of these "paradoxes." Suppose that we are studying a system which, for Lipschitzian disturbances z^ρ, is governed by (1.4) with $n = r = 1$. For notational simplicity we omit the superscripts on x^0, z^ρ, g_ρ^i, and so forth. Let z be a Brownian motion process on an interval $[a, b]$. Let π be a finite set of numbers $t_1, \cdots, t_{k+1}$ with

$$(1.6) \qquad a = t_1 < t_2 < \cdots < t_{k+1} = b,$$

and let Z be the process whose sample paths coincide with those of z at the t_j and are linear between them. Then the solutions X of (1.4) with Z in place of z, that is, the solutions of the ordinary equations

$$(1.7) \qquad X(t) = x_0 + \int_a^t g_0[s, X(s)] + \int_a^t g_1[s, X(s)] \, dZ(s),$$

are random variables; and as the mesh of π (that is, $\max [t_{j+1} - t_j]$) tends to 0, the X converge in quadratic mean to a limit x. But this limit is not the solution of (1.4), but of

$$(1.8) \qquad x(t) = x_0 + \int_a^t g_0[s, x(s)] \, ds + \int_a^t g_1[s, x(s)] \, dz(s)$$
$$+ \frac{1}{2} \int_a^t g_1[s, x(s)] g_{1,x}[s, x(s)] \, ds.$$

(Wong and Zakai have also established this for a more general class of disturbances than Brownian motion processes, see [9].)

These results of Wong and Zakai show us, at least in some important cases, how to model systems affected by noise. If for Lipschitzian disturbances the system evolves according to (1.4) (with subscripts and superscripts suppressed), then if the physically admissible Lipschitzian distances are idealized to Brownian

motion processes equation (1.4) should be replaced by (1.8). If this is done, the solution of (1.8) will be close in quadratic mean to the solutions of (1.4) for at least some Lipschitzian disturbances with finite dimensional distributions close to those of the Brownian motion idealization.

Nevertheless, it is at least inconvenient, as well as aesthetically unsatisfying, to have different equations for different types of disturbances. It would be preferable to have a theory of integration that would apply both to processes with Lipschitzian sample functions, to Brownian motion and to other martingales that have so often proved useful; and correspondingly, it would be preferable to have a method of modeling systems that is consistent with the basic model (1.4) when the disturbances are Lipschitzian and gives "nearly" the same result when a Lipschitzian disturbance is replaced by a martingale type idealization that is in some reasonable sense "close" to it. More specifically, we shall seek to replace (1.4) by another set of so called differential equations (really integral equations) with the following desirable properties.

(a) *Inclusiveness.* The integrals in the equations should be defined for some recognizable class of processes z^p, large enough to include all processes with Lipschitzian sample paths and also to include all Brownian motion and such modifications of Brownian motion as have been useful in applications.

(b) *Consistency.* For Lipschitzian disturbances, the solutions of the equations should coincide with the solutions of the equations (1.4) that are given to us (for smooth disturbances) by the scientific theory of the system.

(c) *Stability.* This property is not easy to describe precisely. Suppose that we have introduced some sort of topology in the space of random processes, so that the convergence of a sequence of processes $z_1, z_2 \cdots$ to a limit process z is meaningful and is in principle experimentally verifiable, with the customary allowance for experimental error. Then, under unexcessive restrictions, if processes $(z_j^1, \cdots, z_j^r)$ converge to $(z^1, \cdots, z^r)$, the solutions $(x_j^1, \cdots, x_j^n)$ corresponding to the z_j^p should also converge to the solutions $(x^1, \cdots, x^n)$ corresponding to the limit $(z^1, \cdots, z^r)$. As a special case, if $n = r = 1$, the solution of the equation when z is Brownian motion should coincide with the solution of Wong-Zakai equation (1.8).

In order to develop such a theory, we must define, for a class of processes with the inclusiveness property (a) the types of integrals needed in the equations; we must develop a calculus for these integrals that will permit us to study differential equations; we must specify the differential equations of our model; we must show that these differential equations are solvable; and we must show that their solutions possess the consistency property (b) and the stability property (c). The remainder of this paper is an outline of the steps in this program.

During the Sixth Berkeley Symposium, I had the pleasure and profit of several conversations with Professor Eugene Wong. In particular, the present version of Theorem 9.1 owes its existence to his tactfully expressed dissatisfaction with an earlier version in which weaker conclusions were drawn from stronger hypotheses.

2. Definition of the integral

To avoid repetition, we henceforth suppose that $(\Omega, \mathscr{A}, P)$ is a probability triple, that T is a set of real numbers, that $[a, b]$ is a closed interval contained in T, and also that

$$(2.1) \qquad f = (f(\tau, \omega): \tau \in T, \omega \in \Omega), \qquad Z^k = (z^k(t, \omega): t \in [a, b], \omega \in \Omega),$$

$k = 1, \cdots, q$, are real stochastic processes on T and on $[a, b]$, respectively. By a *partition* of $[a, b]$ (with evaluation points in T) we shall mean a finite set

$$(2.2) \qquad \Pi = (t_1, \cdots, t_{\ell+1}; \tau_1, \cdots, \tau_\ell)$$

of real numbers such that

$$(2.3) \qquad a = t_1 \leqq t_2 \leqq \cdots \leqq t_{\ell+1} = b$$

and $\tau_i \in T$, $i = 1, \cdots, \ell$. The t_i are called the division points of Π, and the τ_i the partition points of Π. (We usually omit the words "with evaluation points in T.")

Apart from notation, the partitions Π with $\tau_i = t_i$ were used one hundred and fifty years ago by Cauchy to define the integral of a continuous function; so we shall call them *Cauchy partitions*. Partitions with τ_i in $[t_i, t_{i+1}]$ for each i were used by Riemann, and we shall call them *Riemann partitions*. But for use with stochastic processes it proves highly advantageous to use partitions such that $t_i \geqq \tau_i$, $i = 1, \cdots, \ell$, and these we shall call *belated partitions*.

If Π is a Cauchy, or Riemann, or belated partition, with notation (2.2), we define

$$(2.4) \qquad \text{mesh } \Pi = \max \{t_{j+1} - \min \{t_j, \tau_j\}: j = 1, \cdots, \ell\}.$$

Corresponding to the processes (2.1) and the partition (2.2), we define the *Riemann sum* $S(\Pi; f, z^1, \cdots, z^q)$ to be the random variable (r.v.) whose value at ω (in Ω) is given by

$$(2.5) \qquad S(\Pi; f, z^1, \cdots, z^q)(\omega) = \sum_{i=1}^{\ell} \left\{ f(\tau_i; \omega) \prod_{k=1}^{g} [z^k(t_{i+1}\omega) - z^k(t_i, \omega)] \right\}.$$

We can now define the family of integrals that we shall use in our models.

DEFINITION 2.1. *The process f has a* belated integral *with respect to* $(z^1, \cdots, z^q)$ *over* $[a, b]$ *if*, Π *being restricted to the class of belated partitions of* $[a, b]$, *there is an r. v. J such that* $S(\Pi; f, z^1, \cdots, z^q)$ *converges in probability to J as mesh Π tends to 0. Every such J is called a* weak version *of the integral, and is denoted (possibly ambiguously) by*

$$(2.6) \qquad (w) \int_a^b f(t, \omega) \, dz^1(t, \omega), \cdots, dz^q(t, \omega).$$

Such a J is a strong version *of the integral, and is denoted by*

$$(2.7) \qquad \int_a^b f(t, \omega) \, dz^1(t, \omega), \cdots, dz^q(t, \omega),$$

if for each ω_0 in Ω such that the limit (*with notation* (2.2))

$$(2.8) \qquad \ell(\omega_0) = \lim_{\text{mesh } \Pi \to 0} f(t_i, \omega_0) \prod_{k=1}^{g} \left[z^k(t_{i+1}, \omega_0) - z^k(t_i, \omega_0) \right]$$

exists it is true that $J(\omega_0) = \ell(\omega_0)$.

As usual, we omit the ω when convenient. It is quite easy to show that if f is integrable with respect to $(z^1, \cdots, z^q)$, a strong version of the integral exists.

3. The stochastic model

If the sample functions of f are bounded and those of the z^k are Lipschitzian, there is no difficulty in proving that if $q > 1$, then

$$(3.1) \qquad \int_a^b f(t) \, dz^1(t) \cdots dz^q(t) = 0.$$

Suppose then that the functions f^i and g_ρ^i and the derivatives of the latter with respect to the x^i are continuous. By (3.1), if the sample functions x^i all satisfy (1.4) and the functions

$$(3.2) \qquad g_{\rho,\sigma}^i(x, t), \qquad i = 1, \cdots, n; \rho, \sigma = 1, \cdots, r; t \in T, x \in R^n$$

are continuous, then the integrals

$$(3.3) \qquad \sum_{\rho,\sigma} \int_a^t g_{\rho,\sigma}^i[x(s), s] \, dz^\rho(s) \, dz^\sigma(s)$$

exist and are zero for all i in $\{1, \cdots, n\}$ and t in T. Hence, the x^i also satisfy

$$(3.4) \qquad x^i(t) = x^i(a) + \int_a^t g_0^i[x(s), s] \, ds + \sum_\rho \int_a^t g_\rho^i[x(s), s] \, dz^\rho(s)$$
$$+ \tfrac{1}{2} \sum_{\rho,\sigma} \int_a^t g_{\rho,\sigma}^i[x(s), s] \, dz^\rho(s) \, dz^\sigma(s),$$

$i = 1, \cdots, n; a \leq t \leq b$, the integrals either being computed for each sample curve or understood as strict versions of belated integrals. No matter how we choose the (continuous) functions (3.2), we obtain the consistency property (b).

But soon we shall show that the belated integrals can be defined for a class of processes large enough to possess the inclusiveness property (a). When this larger class of z^ρ is permitted, the integrals in (3.3) no longer all vanish, and the stability property (c) does not hold for all choices of functions (3.2). In fact, it is far from clear that it will hold for *any* such functions. We make the choice

$$(3.5) \qquad g_{\rho,\sigma}^i(x, t) = \tfrac{1}{2} \sum_{j=1}^n g_{\rho,x}^i j(x, t) g_\sigma^j(x, t), \qquad i = 1, \cdots, n.$$

This is our selection principle. We do *not* consider that we have added a

correction term (3.5) to the "standard" equation (1.4). Rather, from the aggregate of all equations (3.4), we have selected the one specified by (3.5) instead of the simplest looking one with all functions (3.2) equal to 0. All the equations (3.4) are equally in accord with the underlying theory that gave us equations (1.4), assuming as before that this theory has been established only for Lipschitzian z^ρ. But setting the functions (3.2) equal to 0 gives us merely typographical simplicity, while (as we shall ultimately show) the choice (3.5) gives us, at least under some restrictions, the much more important virtue of stability.

4. Principal existence theorem for the belated integral

Throughout this paper, note T will denote a set of real numbers and $[a, b]$ a closed interval contained in T. Moreover, the symbol $F_\tau (\tau \in T)$ will always denote a σ-subalgebra of $\mathscr{A}$. and we shall always assume *if* τ *and* σ *are in* T *and* $\sigma \leq \tau$, *then* $F_\sigma \subseteq F_\tau$.

For the sake of brevity, if x is a process defined on some subset D_x of T. we shall use the expression "x is F. measurable" to mean "for each t in D_x, $(x(t, \cdot)$ is F_t measurable." Furthermore, to avoid complicated typography we use $F(\tau)$ to denote F_τ whenever convenient; in particular, we write $F(t_j)$ instead of writing the t_j as a subscript to the F.

The processes $(z^1, \cdots, z^r)$ that play the principal role in our theory are those processes on $[a, b]$ that satisfy the following conditions.

CONDITION 4.1.   *Each* $z^\rho (\rho = 1, \cdots, r)$ *is* F. *measurable, and there exist positive numbers* K *and* δ *and a positive integer* q *such that if* $\rho \in \{1, \cdots, r\}$ *and* $a \leq s \leq t \leq b$ *and* $t - s < \delta$. *then a.s.*

$$(4.1) \quad \begin{aligned} \left| E\left(\left[z^\rho(t) - z^\rho(s) \right] \middle| F_s \right) \right| &\leq K(t - s), \\ E\left(\left[z^\rho(t) - z^\rho(s) \right]^{2k} \middle| F_s \right) &\leq K(t - s), \qquad k = 1, \cdots, q. \end{aligned}$$

If x *is a vector in* R^n, *we define* $|x| = \left[\Sigma_1^n (x^i)^2 \right]^{1/2}$; *if* $(x(\omega); \omega \in \Omega)$ *is an* n *vector valued r.v., we define* $\|x\| = E(|x|^2)^{1/2}$ *whenever this expectation exists.*

At this stage we observe that the existence of the integral in Definition 2.1 can be proved, for $q = 1$, under much weaker hypotheses than Condition 4.1; for $q \geq 2$ considerable weakening is also possible, though not as much as for $q = 1$. But the gain in generality is bought at a high price in simplicity. With Condition 4.1, we already have as much inclusiveness as was asked for in (a) and Definition 2.1 is only slightly more complicated than the standard definition of the Riemann integral. Greater inclusiveness would require introducing concepts and methods too sophisticated for some potential users, and would not justify its cost.

The next lemma is an essential element of several later proofs. Its proof differs only trivially from that of Lemma 1 in [5].

LEMMA 4.1.   *Let* $F_1, \cdots, F_m$ *be* σ-*subalgebras of* $\mathscr{A}$ *with* $F_1 \subseteq F_2 \subseteq \cdots \subseteq F_m$. *Let* $u_1, \cdots, u_m$ *and* $\Delta_1, \cdots, \Delta_m$ *be r.v. with finite second moments such that for*

each k in $\{1, \cdots, m\}$, all u_j with $j \leq k$ and all Δ_j with $j < k$ are F_k measurable. Let C_j, D_j for $j = 1, \cdots, m$, be numbers such that a.s.

$$(4.2) \qquad |E(\Delta_j | F_j)| \leq C_j, \; E(\Delta_j^2 | F_j) \leq D_j.$$

Then

$$(4.3) \qquad \left\| \sum_{j=1}^{m} u_j \Delta_j \right\| \leq 2 \sum_{j=1}^{m} C_j \|u_j\| + \left\{ \sum_{j=1}^{m} D_j \|u_j\|^2 \right\}^{1/2}.$$

It is convenient to state a frequently used corollary.

COROLLARY 4.1.   *Let Condition 4.1 be satisfied, and let* Π *(with notation (2.2)) be a partition of* $[a, b]$. *For each j in $\{1, \cdots, \ell\}$, let u_j be an $F[t_j]$ measurable r.v. with finite second moment. Then*

$$(4.4) \qquad \left\| \sum_{j=1}^{\ell} u_j \prod_{k=1}^{q} [z^k(t_{j+1}) - z^k(t_j)] \right\| \leq B \left\{ \sum_{j=1}^{\ell} \|u_j\|^2 (t_{j+1} - t_j) \right\}^{1/2},$$

where $B = 2K(b - a)^{1/2} + K^{1/2}$.

PROOF.   We define

$$(4.5) \qquad \begin{aligned} \Delta_j &= \prod_{k=1}^{q} [2^k(t_{j+1}) - z^k(t_j)], \\ C_j &= D_j = K(t_{j+1} - t_j). \end{aligned}$$

The hypotheses of Lemma 4.1 are satisfied, so

$$(4.6) \qquad \left\| \sum_{j=1}^{\ell} u_j \Delta_j \right\| \leq 2K \sum_{j=1}^{\ell} \{ \|u_j\| (t_{j+1} - t_j) \}^{1/2} \{ (t_{j+1} - t_j) \}^{1/2}$$
$$+ \left\{ \sum_{j=1}^{\ell} \|u_j\|^2 K(t_{j+1} - t_j) \right\}^{1/2}.$$

Applying the Cauchy-Buniakowsky-Schwarz inequality to the first sum in the right member yields the desired conclusion.

We can now state and prove an existence theorem of particular importance in the rest of this paper. In this theorem the integrand is assumed to have the following rather strong continuity property:

(d) *f is bounded in L_2 norm on T, and is continuous in L_2 norm at almost all points of $[a, b]$.*

For such integrands we have the following theorem.

THEOREM 4.1.   *Let $z^1, \cdots, z^q$ satisfy Condition 4.1. Let $(f(\tau): \tau \in T)$ be F. measurable and satisfy* (d). *Then for every subinterval $[c, e]$ of $[a, b]$, f has a (belated stochastic) integral over $[c, e]$, and this integral has an F_e measurable version. Moreover, for every $\varepsilon > 0$ there is a δ' with $0 < \delta' < \delta$ such that, for all subintervals $[c, e]$ of $[a, b]$ and all belated partitions of $[c, e]$ with mesh $\Pi < \delta'$*

it is true that

$$(4.7) \qquad \left\| S(\Pi; f, z^1, \cdots, z^q) - \int_c^e f(t) \, dz^1(t) \cdots dz^q(t) \right\| < \varepsilon.$$

PROOF. Consider the case in which f is continuous in L_2 norm on T. If Π' and Π'' are belated partitions

$$(4.8) \qquad \begin{aligned} \Pi' &= (t_1, \cdots, t_{\ell+1}; \tau'_1, \cdots, \tau'_\ell), \\ \Pi'' &= (t_1, \cdots, t_{\ell+1}; \tau''_1, \cdots, \tau''_\ell), \end{aligned}$$

of $[c, e]$ with the same division points,

$$(4.9) \qquad S(\Pi'; f, z^1, \cdots, z^q) - S(\Pi''; f, z^1, \cdots, z^q)$$

$$= \sum_{j=1}^{\ell} [f(\tau'_j) - f(\tau''_j)] \prod_{k=1}^{q} [z^k(t_{j+1}) - z^k(t_j)].$$

By Corollary 4.1,

$$(4.10) \qquad \left\| S(\Pi'; f, z^1, \cdots, z^q) - S(\Pi''; f, z^1, \cdots, z^q) \right\|$$

$$\leq B(e - c)^{1/2} \max_j \left\| f(\tau'_j) - f(\tau''_j) \right\|.$$

For $q = 1$, the restriction to Π' and Π'' with the same division points is easily removed. Given two partitions

$$(4.11) \qquad \begin{aligned} \Pi' &= (t'_1, \cdots, t'_{\gamma+1}; \tau'_1, \cdots, \tau'_\gamma), \\ \Pi'' &= (t''_1, \cdots, t''_{n+1}; \tau''_1, \cdots, \tau''_n) \end{aligned}$$

of $[c, e]$, we say that the latter is obtained from the former by *adjunction of division points* if each t'_i is one of the t''_j, and if $[t''_j, t''_{j+1}] \subseteq [t'_j, t'_{j+1}]$ then $\tau''_j = \tau'_i$. In this case it is easily seen that

$$(4.12) \qquad S(\Pi'; f, z^1) = S(\Pi''; f, z^1).$$

If Π' and Π'' are any two belated partitions of $[c, e]$, in computing their Riemann sums (for $q = 1$) there is thus no loss of generality in supposing that Π' and Π'' have the same division points, so

$$(4.13) \quad \left\| S(\Pi''; f, z^1) - S(\Pi''; f, z^1) \right\| \leq B(e - c)^{1/2} \max_i \left\| f(\tau'_i) - f(\tau''_i) \right\|,$$

and this can be made arbitrarily small by restricting Π' and Π'' to have small mesh. So the Riemann sums converge in L_2 norm and mesh Π tends to 0, and Theorem 4.1 holds if $q = 1$ and f is continuous in L_2 norm. The latter restriction can be removed by much the same devices as in the case of the ordinary Riemann integral.

If $q > 1$ and Π'' is obtained from Π' by adjunction of division points, the analogue of (4.12) fails. For example, if $q = 2$ and each $[t_j, t_{j+1}]$ of Π' contains

in its interior either a single division point of Π'' (which we then call s_i) or no such point (in which case we define s_i to be t_i), we readily calculate

$$(4.14) \qquad S(\Pi', f, z^1, z^2) - S(\Pi''; f, z^1, z^2)$$

$$= \sum_{j=1}^{\ell} f(\tau_j)\{[z^1(s_j) - z^1(t_j)][z^2(t_{j+1}) - z^2(s_j)]$$
$$+ [z^2(s_j) - z^2(t_j)][z^1(t_{j+1}) - z^1(s_j)]\},$$

which is not in general 0. However, we can find a useful estimate of its norm. Define $\mu(\Pi') = \max_j[t_{j+1} - t_j]$. Since we have assumed Condition 4.1 holds, we find

$$(4.15) \qquad |E([z^1(s_j) - z^1(t_j)][z^2(t_{j+1}) - z^2(s_j)]\,|\,F(t_j))|$$
$$= |E(z^1(s_j) - z^1(t_j)E([z^2(t_{j+1}) - z^2(s_j)]\,|\,F(s_j))\,|\,F(t_j))|$$
$$\leqq K(t_{j+1} - s_j)E(|z^1(s_j) - z^1(t_j)|\,|\,F(t_j))$$
$$\leqq K(t_{j+1} - s_j)[E([z^1(s_j) - z^1(t_j)]^2\,|\,F(t_j))]^{1/2}$$
$$\leqq K^{3/2}\mu(\Pi')^{1/2}(t_{j+1} - t_j).$$

The same estimate holds with z^1 and z^2 interchanged. Similarly,

$$(4.16) \qquad E([z^1(s_j) - z^1(t_j)]^2[z^2(t_{j+1}) - z^2(s_j)]^2\,|\,F(t_j))$$
$$\leqq K^2\mu(\Pi')(t_{j+1} - t_j),$$

and likewise with z^1 and z^2 interchanged. We now apply Lemma 4.1 to each of the two sums in (4.14); by (4.15) and (4.16), we obtain

$$(4.17) \qquad \|S(\Pi'; f, z^1, z^2) - S(\Pi''; f, z^1, z^2)\| \leqq C[\mu(\Pi')]^{1/2},$$

where

$$(4.18) \qquad C = 2K \sup \{\|f(\tau)\|: \tau \in T\}[2K^{1/2}(b - a) + (b - a)^{1/2}].$$

If $q > 2$, the estimate (4.17) (with a different c) is still valid, the proof is not essentially different but the details are more tedious.

We shall repeatedly use the following procedure.

PROCEDURE 4.1. *Given a partition* $\Pi = (t_1, \cdots, t_{\ell+1}; \tau_1, \cdots, \tau_\ell)$, *we adjoin to* Π *as new division points the midpoints of all those intervals* $[t_j, t_{j+1}]$ *such that* $t_{j+1} - t_j \geqq \frac{1}{2}\mu(\Pi)$.

We form a sequence of partitions

$$(4.19) \qquad \Pi_0' = \Pi', \Pi_1', \Pi_2', \cdots, \Pi_\alpha',$$

each formed from the preceding by applying Procedure 4.1. We carry it to a large enough α so that no interval of the original Π' remains unsubdivided; then each interval of Π_α' will have length at least $\frac{1}{2}\mu(\Pi_\alpha')$. Since

$$(4.20) \qquad \mu(\Pi_k') = 2^{-k}\mu(\Pi'),$$

we may also suppose that $\mu(\Pi'_\alpha)$ is less than half the length of the smallest interval in Π''. Next, starting with Π'', we form the sequence

$$(4.21) \qquad\qquad \Pi''_0 = \Pi'', \Pi''_1, \Pi''_2, \cdots, \Pi''_\beta$$

by repeated application of Procedure 4.1. We can and do choose β so that

$$(4.22) \qquad\qquad 2^{-1/2}\mu(\Pi'_\alpha) \leqq \mu(\Pi''_\beta) \leqq 2^{1/2}\mu(\Pi'_\alpha);$$

this is possible by the analogue of (4.20). By (4.17) and (4.20),

$$(4.23) \qquad \left\| S(\Pi'_\alpha; f, z^1, z^2) - S(\Pi'; f, z^1, z^2) \right\|$$
$$\leqq \sum_{n=0}^{\alpha-1} C[\mu(\Pi'_h)]^{1/2}$$
$$\leqq (2 + 2^{1/2})C[\mu(\Pi')]^{1/2}.$$

Similarly,

$$(4.24) \qquad \left\| S(\Pi''_\beta; f, z^1, z^2) - S(\Pi''; f, z^1, z^2) \right\|$$
$$\leqq (2 + 2^{1/2})C[\mu(\Pi'')]^{1/2}.$$

Every interval in Π'_α has length at least $\tfrac{1}{2}\mu(\Pi'_\alpha)$, and likewise for Π''_β. So every interval in Π'_α has length at least $\mu(\Pi''_\beta)/2^{3/2}$, and vice versa. Thus, each interval of Π'_α contains at most three division points of Π''_β. We can adjoin these to Π'_α in two stages, obtaining a partition Π'_* such that $\big($by (4.17)$\big)$

$$(4.25) \qquad \left\| S(\Pi'_*, f, z^1, z^2) - S(\Pi'_\alpha; f, z^1, z^2) \right\|$$
$$\leqq 2C[\mu(\Pi')]^{1/2}.$$

Similarly, we can adjoin the division points of Π''_ρ to Π''_β in at most two stages, obtaining a partition Π''_* such that

$$(4.26) \qquad \left\| S(\Pi''_*; f, z^1, z^2) - S(\Pi''_\beta; f, z^1, z^2) \right\|$$
$$\leqq 2C[\mu(\Pi'')]^{1/2}.$$

Now Π'_* and Π''_* have the same division points. By (4.10), (4.19), (4.20), (4.21), and (4.22), for L_2 continuous f, the Riemann sums for Π', for Π'_*, for Π''_* and for Π'' have differences (in the order named) that have L_2 norms which are arbitrarily small if mesh Π' and mesh Π'' are small. This implies that $S(\Pi; f, z^1, \cdots, z^q)$ converges in L_2 norm as mesh Π tends to 0, and the integral exists. The uniform closeness of Riemann sum to integral follows from the fact that all estimates of L_2 norms were uniformly valid; and we could have used F_e everywhere in place of $\mathscr{A}$ without changing anything, which would give us an F_e measurable integral.

5. An estimate, and a second existence theorem

Suppose that the hypotheses of Theorem 4.1 are satisfied, and that Π (with notation (2.2)) is a belated partition of a subinterval $[c, e]$ of $[a, b]$. By Corollary 4.1,

$$(5.1) \qquad \|S(\Pi; f, z^1, \cdots, z^q)\| \leqq B \left\{ \sum_{j=1}^{\ell} \|f(\tau_j)\|^2 (t_{j+1} - t_j) \right\}^{1/2}.$$

But by the special case of Theorem 4.1 in which Ω contains a single point and $q = 1$ and $z(t) = t$, the belated integral of $\|f\|$ with respect to t exists. Since Cauchy partitions are both Riemann partitions and belated partitions, the belated integral of $\|f\|$ is its Riemann integral, and by letting mesh Π tend to 0 we obtain from (5.1)

$$(5.2) \qquad \left\| \int_c^e f(t)\, dz^1(t) \cdots dz^q(t) \right\| \leqq B \left\{ \int_c^e \|f(t)\|^2\, dt \right\}^{1/2}.$$

If T is an interval, and f is F. measurable and $(f(\tau, \omega): \omega \in t, \omega \in \Omega)$ is $dt\, dP$ measurable on $T \times \Omega$, and

$$(5.3) \qquad \int_T E[f(\tau)^2]\, d\tau < \infty,$$

it is possible to find (as in [1], p. 440) a sequence of bounded processes $f_1, f_2, \cdots$, satisfying the hypotheses of Theorem 4.1 such that

$$(5.4) \qquad \lim_{n \to \infty} \int_a^b E(|f_n - f|^2)\, d\tau = 0;$$

in fact, by a slight modification of the construction in [1] we may choose f_n that are continuous in L_2 norm. Then by (5.2) the integrals

$$(5.5) \qquad \int_c^e f_n(t)\, dz^1(\tau) \cdots dz^q(\tau), \qquad n = 1, 2, 3, \cdots,$$

form a Cauchy sequence in $L_2(\Omega, P)$, and hence have a limit in that space. We can accept this limit as the definition of the integral of f with respect to $(z^1, \cdots, z^q)$, thus extending the class of integrable functions so as to have the same sort of closure properties as the Itô integral. Such properties are valuable in many investigations. But in this paper we have no need of them, so we pursue this no farther.

In Definition 2.1, we used the concept of convergence in probability. In Theorem 4.1, we obtained more: the Riemann sums converged in L_2 norm. There is an intermediate kind of convergence that is sometimes encountered, that we shall call *uniform convergence in near L_2 norm*. We define it in the setting of functions of partitions, although it evidently can be applied to more general limit processes. (In the definition, 1_A denotes the indicator function of the set A.)

DEFINITION 5.1. *Assume that to each subinterval* $[c, e]$ *of* $[a, b]$ *and to each belated partition* Π *of* $[c, e]$ *there corresponds an r.v.* $x(\Pi, [c, e])$ *and an r.v.* $x_0([c, e])$. *Then* $x(\Pi, [c, e])$ *converges to* $x_0([c, e])$ *in near* L_2 *norm, uniformly on subintervals* $[c, e]$ *of* $[a, b]$, *if to each positive* ε *there corresponds a positive* δ *and a subset* A *of* Ω *with* $P(A) > 1 - \varepsilon$ *such that for all* $[c, e] \subseteq [a, b]$ *and all* Π *with mesh* $\Pi < \delta$,

$$(5.6) \qquad \|1_A[x(\Pi, [c, e]) - x_0([c, e])]\| < \varepsilon.$$

Clearly this implies uniform convergence in the metric of convergence in probability, and is implied by uniform convergence in L_2 norm.

Many processes f possess the following important property:

(e) f *is separable, and with probability* 1 *the sample function* $[f(\tau, \omega): \tau \in T]$ *is bounded.* (Note that the bound is not assumed to be independent of ω.)

For integrands f with this property, we can prove the following theorem.

THEOREM 5.1. *Let Condition* 4.1 *be satisfied. Let* $[f(\tau): \tau \in T]$ *satisfy* (e), *and be* F. *measurable, and be continuous in probability at almost all points of* $[a, b]$. *Then for every subinterval* $[c, e]$ *of* $[a, b]$, f *has a belated integral with respect to* $(z^1, \cdots, z^q)$ *over* $[c, e]$, *and this integral has an* F_e *measurable version. Moreover, the Riemann sums* $S(\Pi; f, z^1, \cdots, z^q)$ *corresponding to belated partitions* Π *of* $[c, e]$ *converge to the integral over* $[c, e]$ *uniformly in near* L_2 *norm as mesh* $\Pi \to 0$.

PROOF. Let $S \subseteq T$ be a separate set for f. There is a subset Λ of Ω with $P\Lambda = 0$ such that for every open interval I and every ω in $\Omega - \Lambda$, the functions $[f(\tau, \omega): \tau \in I \cap T]$ and $[f(\tau, \omega): \tau \in I \cap S]$ have equal suprema and equal infima. Let ε be positive. For each positive N we define $A_N(\tau)$ $(\tau \in T)$ to be the set of all ω in Ω such that $|f(s, \omega)| \leq N$ for $s = \tau$ and for all $s < \tau$ in S. This is F_τ measurable, and $P[A_N(\tau)]$ is nonincreasing. By (e), we can choose N large enough so that

$$(5.7) \qquad P[A_N(b)] > 1 - \varepsilon.$$

Let $\phi_N(\tau, \cdot)$ be the indicator function of $A_N(\tau)$. Then by definition of A_N, $f\phi_N$ is bounded. It is also fairly obviously F_τ measurable at each τ in T, and it is continuous in probability (hence, being bounded, it is continuous in L_2 norm) except on the union of the null set of discontinuities of f and the countable set of discontinuities of $P(A_N(\tau))$. So, by Theorem 4.1 for every $[c, e] \subseteq [a, b]$ the Riemann sums

$$(5.8) \qquad S(\Pi; f\phi_N, z^1, \cdots, z^q)$$

converge as mesh $\Pi \to 0$ to the integral of $f\phi_N$ over $[c, e]$, uniformly with respect to $[c, e]$. But the sums (5.8) coincide on $A_N(b) - \Lambda$ with

$$(5.9) \qquad S(\Pi; f, z^1, \cdots, z^2).$$

From this and (5.7), it follows readily that the sums (5.9) converge in near L_2 norm to a limit, which is by definition the integral of f over $[c, e]$; and the convergence is uniform with respect to $[c, e]$.

6. Examples

Suppose first that z^1 and z^q are both the same Wiener process w. Then if $a \leqq s \leqq t \leqq b, w(t) - w(s)$ is independent of F_s. Let Π (with the usual notation (2.2)) be a belated partition, and define

$$(6.1) \qquad \Delta_j = [w(t_{j+1}) - w(t_j)]^2 - (t_{j+1} - t_j), \qquad j = 1, \cdots, \ell.$$

Then $E(\Delta_j | F(t_j)) = 0, E(\Delta_j^2 | F(t_j)) = 2(t_{j-1} - t_j)^2$.

If f satisfies the hypotheses of Theorem 4.1 by Lemma 4.1,

$$(6.2) \qquad \left\| \sum_{j=1}^{\ell} f(t_j) \Delta_j \right\| \leqq \left\{ \sum_{j=1}^{\ell} 2\|f(t_j)\|^2 (t_{j+1} - t_j)^2 \right\}^{1/2},$$

which tends to 0 with mesh Π. By Theorem 4.1, $S(\Pi; f, z^1, z^2)$ has a limit as mesh $\Pi \to 0$; by (6.1) and (6.2), $S(\Pi; f, t)$ has the same limit. So if f satisfies the hypotheses of Theorem 4.1, we have

$$(6.3) \qquad \int_a^b f(t)(dw)^2 = \int_a^b f(t)\, dt.$$

This also holds if f satisfies the hypotheses of Theorem 5.1.

The next lemma is useful because it often permits us to discard integrals with several dz^ρ. It applies to disturbances that satisfy the following condition.

CONDITION 6.1. *To each positive ε there corresponds a positive δ and a set $A \subseteq \Omega$ with $P(A) > 1 - \varepsilon$ such that if $a \leqq s \leqq t \leqq b$ and $t - s < \delta$,*

$$(6.4) \qquad |z^\rho(t, \omega) - z^\rho(s, \omega)| < \varepsilon(t - s)^{1/3}$$

for all ω in A.

For example, by a well-known theorem of Kolmogorov (see Neveu [7], p. 97) z^ρ satisfies Condition 6.1 if there is a constant K such that

$$(6.5) \qquad E([z^\rho(t) - z^\rho(s)]^8) \leqq K(t - s)^4, \qquad a \leqq s \leqq t \leqq b.$$

THEOREM 6.1. *Let the hypotheses of Theorem 4.1 or Theorem 5.1 be satisfied. If $q \geqq 3$, and $z^1, \cdots, z^q$, satisfy Lemma 6.1, then*

$$(6.6) \qquad \int_c^e f(t)\, dz^1(t) \cdots dz^q(t) = 0.$$

PROOF. Suppose first that $|f(\tau, \omega)|$ has an upper bound N on $T \times \Omega$. Let ε be positive, and let δ and A serve for all z^k in Condition 6.1. If Π (with notation (2.2)) is a belated partition with mesh $\Pi < \min(1, \delta)$, for all ω in A we have

$$(6.7) \qquad \left| \sum_{j=1}^{\rho} f(\tau_j, \omega) \prod_{k=1}^{q} (z^k(t_{j+1}, \omega) - z^k(t_j, \omega)) \right| \leqq N\varepsilon^q(b - \mathrm{a}).$$

So the Riemann sums converge in near L_2 norm to 0, and the integral is 0.

If f satisfies the hypotheses of Theorem 4.1, for each positive N, we define

(6.8) $f_N(\tau, \omega) = \begin{cases} f(\tau, \omega) & \text{if } -N \leqq f(\tau, \omega) \leqq N, \\ N & \text{if } f(\tau, \omega) > N, \\ -N & \text{if } f(\tau, \omega) < -N. \end{cases}$

By the proof just completed, the integral of f_N is 0 for all N. So by (5.2),

(6.9) $\left\| \int_c^e f(t) \, dz^1(t) \cdots dz^q(t) \right\| \leqq B \left\{ \int_c^e \|f(t) - f_N(t)\|^2 \, dt \right\}^{1/2}.$

The right member tends to 0 as $N \to \infty$, so the left member is 0.

If the hypotheses of Theorem 5.1 are satisfied, with the notation of that theorem, $f\phi_N$ has integral 0 for all N, so the integral of f is 0.

There are other useful sets of conditions that eliminate integrals, but we will confine ourselves to two simple cases.

THEOREM 6.2.   *If the hypotheses of Theorem 4.1 or of Theorem 5.1 hold with* $q \geqq 2$, *and* $z^q(t) = t$, *then* $\int_a^b f(t) \, dz^1 \cdots dz^q = 0$.

PROOF.   With Π as in (2.2), define

(6.10) $\Delta_j^k = z^k(t_{j+1}) - z^k(t_j).$

Then

(6.11) $\left| E\left(\prod_{k=1}^q \Delta_j^k \,\middle|\, F_{t_j} \right) \right| = (t_{j+1} - t_j) \left| E\left(\prod_{k=1}^{q-1} \Delta_j^k \,\middle|\, F_{t_j} \right) \right|,$

(6.12) $E\left(\left[\prod_{k=1}^q \Delta_j^k \right]^2 \,\middle|\, F_{t_j} \right) = (t_{j+1} - t_j)^2 E\left(\left[\prod_{k=1}^{q-1} \Delta_j^k \right]^2 \,\middle|\, F_{t_j} \right).$

If $q = 2$, the right members of (6.11) and (6.12) do not exceed $K(t_{j+1} - t_j)^2$ and $K(t_{j+1} - t_j)^3$, respectively, by Condition 4.1; so, under the hypotheses of Theorem 4.1, Lemma 4.1 assures us that $\|S(\Pi; f, z^1, \cdots, z^q)\|$ tends to 0 with mesh Π. If the hypotheses of Theorem 5.1 hold, the conclusion is established by the use of the functions ϕ_N of Theorem 5.1.

If $q > 2$, let r and s be integers at most $\frac{1}{2}q$ with $r + s = q - 1$. Then

(6.13) $E\left(\prod_{k=1}^{q-1} \Delta_j^k \,\middle|\, F_{t_j} \right) \leqq \left\{ E\left(\left[\prod_{k=1}^r \Delta_j^k \right]^2 \,\middle|\, F_{t_j} \right) \right\}^{1/2} \left\{ E\left(\left[\prod_{k=r+1}^{q-1} \Delta_j^k \right]^2 \,\middle|\, F_{t_j} \right) \right\}^{1/2}.$

Since $\left[\Pi_1^r \Delta_j^k \right]^2 \leqq \Sigma_1^r \left[\Delta_j^k \right]^{2r}$ and $2^r \leqq q$, by Condition 4.1, the first factor in the right member of (6.13) does not exceed a constant multiple of $(t_{j+1} - t_j)^{1/2}$. The same is true of the second factor, so the left member of (6.11) does not exceed a multiple of $(t_{j+1} - t_j)^2$. Likewise the left member of (6.12) does not exceed a multiple of $(t_{j+1} - t_j)^3$.

The rest of the proof is as for $q = 2$.

THEOREM 6.3. *Let z^1 and z^2 be processes such that if $a \leq s \leq t \leq b$, then $z^1(t) - z^1(s)$ and $z^2(t) - z^2(s)$ are conditionally independent as conditioned by F_s. Let the hypotheses of Theorem 4.1 hold. Then*

$$(6.14) \qquad \int_a^b f(t) \, dz^1(t) \, dz^2(t) = 0.$$

PROOF. We use the same notation as in the preceding proof. Then

$$(6.15) \qquad |E(\Delta_j, z^1 \Delta_j z^2 \,|\, F_s(t_j))| = |E(\Delta, z^1 \,|\, F(t_j)) E(\Delta_j z^2 \,|\, F(t_j))|$$
$$\leq K^2 (t_{j+1} - t_j)^2,$$

$$(6.16) \qquad E([\Delta_j z^1 \Delta_j z^2]^2 \,|\, F(t_j)) = E([\Delta_j z^1]^2 \,|\, F(t_j)) E([\Delta_j z^2]^2 \,|\, F(t_j))$$
$$\leq K^2 (t_{j+1} - t_j)^2.$$

By Lemma 4.1, $\|S(\Pi; f, z^1, z^2)\|$ tends to 0 with mesh Π.

7. Existence theorem for a functional equation

If the hypotheses of Theorem 4.1 are satisfied and we define a process F on $[a, b]$ by setting

$$(7.1) \qquad F(t) = \int_a^t f(s) \, dz^1(s) \cdots dz^q(s), \qquad\qquad t \in [a, b],$$

we know by Theorem 4.1 that $F(t)$ has finite second moment and can be chosen F_* measurable. By (5.2), we know that it satisfies a Hölder condition of exponent $1/2$ in L_2 norm. Processes with these properties occur often enough in succeeding pages to justify giving them a name.

DEFINITION 7.1. *Let $H_{1/2}(T, F_*)$ be the class of all (real or vector valued) processes x on T such that for all t in T, $x(t)$ is F_t measurable and $E(|x(t)|^2) < \infty$, and there is a number H^* such that if s and t are in T,*

$$(7.2) \qquad \|x(t) - x(s)\| \leq H^*(t - s)^{1/2}.$$

COROLLARY 7.1. *If the hypotheses of Theorem 4.1 are satisfied and F is defined by (7.1), F belongs to $H_{1/2}([a, b], F_*)$.*

Instead of restricting ourselves to stochastic "differential equations" such as (3.3), we shall discuss a class of functional equations

$$(7.3) \qquad x^i(\tau) = y^i(\tau), \qquad\qquad \tau \in T, \tau \leq a,$$

$$(7.4) \qquad x^i(t) = y^i(t) + \int_a^t g_0^i(s, x(s)) \, ds$$

$$+ \sum_{\mathscr{C}} \int_a^t g_{\rho, \sigma, \cdots, \phi}^i(s, x(s)) \, dz^\rho(s) \, dz^\sigma(s) \cdots dz^\phi(s),$$

$$i = 1, \cdots, n; a \leq t \leq b.$$

where the letters denote members of $\{1, \cdots, r\}$, and $\mathscr{C}$ is a finite set of finite ordered sequences $(\rho, \sigma, \cdots, \phi)$ of members of $\{1, \cdots, r\}$. The functions $g_0^i, g_{\rho, \sigma, \cdots, \phi}^i$ will be called coefficients. We shall make the following assumptions.

ASSUMPTION 7.1.   *The class $\mathscr{P}$ is a linear class of n-vector valued processes on T that contains $H_{1/2}(T, F_*)$, and is closed under uniform convergence in L_2 norm.*

ASSUMPTION 7.2.   *Each coefficient g is defined on $T \times \mathscr{P}$, and for fixed x in $\mathscr{P}$, $g(\cdot, x)$ is bounded in L_2 norm on T and is continuous in L_2 norm at almost all points of $[a, b]$.*

ASSUMPTION 7.3.   *If F is a σ-subalgebra of $\mathscr{A}$, and $t \in T$, and x is a process in $\mathscr{P}$ such that $x(\tau)$ is F measurable for all $\tau \leq t$ in T, then $g(t, x)$ is also F measurable.*

For Theorem 7.1 it would be adequate to choose $H_{1/2}(t, F_*)$ for $\mathscr{P}$. However, in the case of differential equations a little more latitude is convenient. Suppose that G_0^i and $G_{\rho, \sigma, \cdots, \phi}^i$ are functions on $T \times \mathbf{R}^n$ such that, for a certain subset N_0 of T with Lebesgue measure 0 and a certain positive L, it is true that G_0^i and $G_{\rho, \sigma, \cdots, \phi}^i$ are continuous in all variables at all points (t, x) with $t \in T - N_0$ and $x \in \mathbf{R}^n$, and for all t in T and x_1, x_2 in $\mathbf{R}^n$

$$(7.5) \qquad |G_0^i(t, x_1) - G_0^i(t, x_2)| \leq L|x_1 - x_2|.$$

and likewise for the $G_{\rho, \sigma, \cdots, \phi}^i$. Then for all processes x with finite second moments we can define

$$(7.6) \qquad g_0^i(t, x) = G_0^i(t, x(t)),$$

$$(7.7) \qquad g_{\rho, \sigma, \cdots, \phi}^i(t, x) = G_{\rho, \sigma, \cdots, \phi}^i(t, x(t)), \qquad\qquad t \in T,$$

and Assumption 7.3 is satisfied. To attain Assumptions 7.1 and 7.2 also, we can make the following assumption.

ASSUMPTION 7.4.   *$\mathscr{P}$ is the class of all processes bounded in L_2 norm on T and continuous in L_2 norm at almost all points of $[a, b]$.*

We can simplify notation a little by defining $z^0(t) = t$ for all real t, and adjoining the one element sequence (0) to $\mathscr{C}$. With this understanding, equations (7.3) and (7.4) take the notationally simpler form

$$(7.8) \qquad x^i(\tau) = y^i(\tau), \qquad\qquad \tau \in T, \tau \leq a,$$

$$(7.9) \qquad x^i(t) = y^i(t) + \sum_{\mathscr{C}} \int_a^t g_{\rho, \sigma, \cdots, \phi}^i(s, x(s)) \, dz^\rho(s) \cdots dz^\phi(s),$$

$$t = 1, \cdots, n; a \leq t \leq b.$$

THEOREM 7.1.   *Let the coefficients in (7.8) and (7.9) satisfy Assumptions 7.2 and 7.3, and let the z^ρ satisfy Condition 4.1. Assume also that there exists a positive L such that if x_1 and x_2 are in $\mathscr{P}$ and $t \in [a, b]$, for each coefficient g in (7.8) and (7.9) it is true that*

$$(7.10) \qquad \|g(t, x_1) - g(t, x_2)\| \leq L \sup \{\|x_1(s) - x_2(s)\| : s \in T, s \leq t\}.$$

Let y belong to $\mathscr{P}$ and be $F_{\cdot}$ measurable. Then there is an $F_{\cdot}$ measurable process $x(\cdot)$ in $\mathscr{P}$ such that $x^i(\tau) = y^i(\tau)$, $\tau \in T$, $\tau \leq a$, and (7.8) holds for $a \leq t \leq b$. If $y(\cdot) \in H_{1/2}(T, F_{\cdot})$, so does $x(\cdot)$. Moreover, if x_1 is any $F_{\cdot}$ measurable process satisfying (7.8) and (7.9) then $P[x_1(t) = x(t)] = 1$ for all t in T.

PROOF. Hypothesis (7.10) guarantees that the coefficients are nonanticipative; if x_1 and x_2 belong to $\mathscr{P}$ and $x_1(\tau) = x_2(\tau)$ if $\tau \in T$ and $\tau \leq t$, then $g(t, x_1) = g(t, x_2)$. If x is defined only on the part of T in $(-\infty, t]$ and has an extension $\tilde{x}$ to T that belongs to $\mathscr{P}$, by (7.10) all such extensions $\tilde{x}$ give the same value to $g(t, x)$. To simplify notation, we shall define $g(t, x)$ to mean that common value.

We use Picard's method. We define $x_0 = y$, and then successively

$$(7.11) \qquad\qquad\qquad x_{k+1}^i(\tau) = y^i(\tau), \qquad\qquad \tau \in T, \tau \leq a,$$

$$(7.12) \qquad x_{k+1}^i(\tau) = y^i(\tau) + \sum_{\mathscr{C}} \int_a^\tau g_{\rho,\sigma,\cdots,\phi}^i(s, x_k)\, dz^\rho\, dz^\sigma \cdots dz^\phi.$$

By hypothesis x_0 is in $\mathscr{P}$. If we assume x_k in that class, the integrands in (7.11) and (7.12) satisfy the hypotheses of Theorem 4.1 by Corollary 7.1, x_{k+1}^i belongs to $H_{1/2}(T, F_{\cdot})$. Thus, (7.11) and (7.12) define x_k for $k = 0, 1, 2, 3, \cdots$. Define, for every process x on T and every t in T,

$$(7.13) \qquad\qquad N(t, x) = \sup \{\|x(\tau)\| \mid \tau \in T, \tau \leq t\}.$$

If $g_{\rho,\sigma,\cdots,\phi}^i$ is one of the coefficients in (7.8) and (7.9) and $k \geq 1$, by (5.2) and hypothesis (7.10),

$$(7.14) \quad \left\| \int_a^t \{g_{\rho,\sigma,\cdots,\phi}^i(s, x_k) - g_{\rho,\sigma,\cdots,\phi}^i(s, x_{k+1})\}\, dz^\rho\, dz^\sigma \cdots dz^\phi \right\|$$

$$\leq B \left\{ \int_a^t \left\| g_{\rho,\sigma,\cdots,\phi}^i(s, x_k) - g_{\rho,\sigma,\cdots,\phi}^i(s, x_{k-1}) \right\|^2 ds \right\}^{1/2}$$

$$\leq B \left\{ \int_a^t L^2 N(s, x_k - x_{k-1})^2\, ds \right\}^{1/2}.$$

Let B_0 be the product of n, B, L^2 and the number of sequences in the set $\mathscr{C}$. By (7.11), (7.12), and (7.14),

$$(7.15) \qquad \|x_{k+1}(t) - x_k(t)\| \leq B_0 \left\{ \int_a^t N(s, x_k - x_{k-1})\, ds \right\}^{1/2}.$$

Since this estimate is still valid if we replace t in the left member by any smaller member of T.

$$(7.16) \qquad N(t, x_{k+1} - x_k)^2 \leq B_0^2 \left\{ \int_a^t N(s, x_k - x_{k-1})^2\, ds \right\}.$$

We can now prove by induction (with $x_{-1} \equiv 0$)

$$(7.17) \qquad N(t, x_k - x_{k-1})^2 \leq \{\sup \|y(t)\|^2\} B_0^{2k} \frac{(t-a)^k}{k!}, \qquad k = 0, 1, 2, \cdots,$$

For $k = 0$, this is simply the statement $N(t, y)^2 \leq \sup \|y(t)\|^2$. If (7.17) holds for k, by (7.16),

$$(7.18) \qquad N(t, x_{k+1} - x_k)^2 \leq B_0^2 \left\{ \int_a^t (\sup \|y\|^2) \left[\frac{B_0^{2k}}{k!} \right] (s-a)^k \, ds \right\}$$

$$= \sup \|y\|^2 \left[\frac{B_0^{2k+2}}{(k+1)!} \right] (t-a)^{k+1}.$$

so (7.17) holds for all nonnegative integers k. It follows at once that the sums

$$(7.19) \qquad x_k = \sum_{k=0}^{h} (x_k - x_{k-1})$$

converge uniformly in L_2 norm to a limit, which we call x. This limit belongs to $H_{1/2}(T, F.)$, and by (7.11) and (7.12) it satisfies (7.8) and (7.9).

If x' and x'' both satisfy (7.8) and (7.9) and are $F.$ measurable, just as we proved (7.16) we can prove

$$(7.20) \qquad N(t, x' - x'')^2 \leq B_0^2 \left\{ \int_a^t N(s, x' - x'')^2 \, ds \right\}^{1/2}.$$

The only solution of this is $N(s, x' - x'') = 0$, which completes the proof.

8. Cauchy-Maruyama approximations

G. Maruyama [4] has extended the well-known Cauchy (or Euler) method of constructing polygonal approximate solutions, proceeding successively from each vertex to the next, to the stochastic differential equations (1.3). It is easy to extend this procedure still further to equations of the form of (7.8) and (7.9). Given any Cauchy partition

$$(8.1) \qquad \Pi = (t_1, \cdots, t_{\ell+1}; t_1, \cdots, t_\ell)$$

of $[a, b]$, we first define $\bar{x}(\tau) = y(\tau)$ for all τ in T with $\tau \leq a$. Then $\bar{x}$ having been defined for all τ in T with $\tau \leq t_j$, we define it on $(t_j, t_{j+1}]$ by setting

$$(8.2) \qquad \bar{x}^i(\tau) = \bar{x}^i(t_j) + y^i(t) - y^i(t_j)$$
$$+ \sum_{\mathscr{C}} g_{\rho,\sigma,\cdots,\phi}^i(t_j, \bar{x}) [z^\rho(t) - z^\rho(t_j)] \cdots [z^\phi(t) - z^\phi(t_j)].$$

(Notice that the coefficients are defined, even though $\bar{x}$ has been defined only up to t_j, by the first paragraph of the proof of Theorem 7.1.)

We can prove that under the hypotheses of Theorem 7.1 these Cauchy-Maruyama functions x converge to the solution x of (7.8) and (7.9) uniformly

in L_2 norm as mesh $\Pi \to 0$. But in Sections 9 and 11, we shall need a different approximation, in which we shall permit a small departure from equality in (8.2).

We shall suppose that $\mathscr{P}$ is defined by Assumption 7.4, and we shall adopt the abbreviations

(8.3) $\quad \Delta_j t = t_{j+1} - t_j, \; \Delta_j y = y(t_{j+1}) - y(t_j), \; \Delta_j z^\rho = z^\rho(t_{j+1}) - z^\rho(t_j).$

Suppose now that to each Cauchy partition Π (with notation (8.1)) then corresponds a process $\bar{x}$ with the following properties:

(f) $\bar{x}(\tau) = y(\tau), \tau \in T, \tau \leqq a$;

(g) *to each positive ε there corresponds a positive δ such that, if mesh $\Pi < \delta$*

(8.4) $\quad \| \bar{x}(t) - \bar{x}(t_j) - y(t) + y(t_j)$
$$- \Sigma \, g^i_{\rho,\sigma,\cdots,\phi}(t_j, \bar{x})[z^\rho(t) - z^\rho(t_j)] \cdots [z^\phi(t) - z^\rho(t_j)] \|$$
$$\leqq \varepsilon\big(1 + \sup \{\|\bar{x}(\tau)\| : \tau \in T, \tau \leqq t\}\big)(t_j \leqq t < t_{j+1}),$$

(8.5) $\quad \| \bar{x}(t_{j+1}) - \bar{x}(t_j) - \Delta_j y - \Sigma \, g_{\rho,\sigma,\cdots,\phi}(t_j, \bar{x})\Delta_j z^\rho \cdots \Delta_j z^\rho \|$
$$\leqq \varepsilon\big(1 + \sup \{\|\bar{x}(\tau)\| : \tau \in T, \tau \leqq t_{j+1}\}\big)\Delta_j t ;$$

(h) *if $\tau \in T$ and $\tau \leqq t_j$, $\bar{x}(\tau)$ is $F[t_j]$ measurable.*

(The Cauchy-Maruyama functions (8.3) clearly satisfy these requirements.) We can then prove the following theorem.

THEOREM 8.1. *Let the hypotheses of Theorem 7.1 hold. Assume that to each Cauchy partition Π of $[a, b]$ there corresponds a process $\bar{x}$ in $\mathscr{P}$ such that* (f), (g), *and* (h) *hold. Then as mesh Π tends to zero, $\bar{x}$ converges in L_2 norm, uniformly on T, to the solution x of* (7.8) *and* (7.9).

PROOF. By Theorem 4.1, the solution x of (7.8) and (7.9) exists, and we can and do choose it to be F, measurable. Let Π be a Cauchy partition, with notation (2.1); let t be a point of $[a, b]$; and define

(8.6) $\quad t_k = $ largest number in set $\{t_1, \cdots, t_\ell\} \cap (-\infty, t],$

(8.7) $\quad N(t) = \sup \{\|\bar{x}(\tau) - x(\tau)\| : \tau \in T, \tau \leqq t_k\}.$

(8.8) $\quad x^i(t) = y^i(t) + \sum_{j=1}^{k-1} g^i_{\rho,\sigma,\cdots,\phi}(t_j, x)\Delta_j z^\rho \cdots \Delta_j z$
$$+ \Sigma \, g^i_{\rho,\cdots,\phi}(t_k, x)[z^\rho(t) - z^\rho(t_k)] \cdots [z^\phi(t) - z^\phi(t_k)],$$

(For $\tau \in t, \tau \leqq a$ we take $X(\tau) = y(\tau)$.)

Let $M - 1$ be an upper bound for $\|x(\tau)\|$ on T, and let ε be any number such that $0 < 2(1 + b - a)\varepsilon < 1$. By Theorem 4.1, there is a positive δ_1 such that if mesh $\Pi < \delta_1$,

(8.9) $\quad \|X(\tau) - x(\tau)\| < \varepsilon(\tau \in T).$

With δ as in (g), we let Π be any belated partition such that

$$(8.10) \qquad\qquad \text{mesh } \Pi < \min \{\delta, \delta_1\},$$

Then, with t_k defined by (8.6),

$$(8.11) \quad \bar{x}^i(t) - x^i(t)$$
$$= X^i(t) - x^i(t)$$
$$+ \sum_{j=1}^{k-1} \{\bar{x}^i(t_{j+1}) - \bar{x}^i(t_j) - \Delta_j y^i - \sum_{\mathscr{C}} g^i_{\rho,\cdots,\phi}(t_j, \bar{x}) \Delta_j z^\rho \cdots \Delta_j z^\phi\}$$
$$+ \bar{x}^i(t) - \bar{x}^i(t_k) - [y^i(t) - y^i(t_k)]$$
$$- \sum_{\mathscr{C}} g^i_{\rho,\cdots,\phi}(t_k, \bar{x})[z^\rho(t) - z^\rho(t_k)] \cdots [z^\phi(t) - z^\phi(t_k)]$$
$$+ \sum_{j=1}^{k-1} \sum_{\mathscr{C}} [g^i_{\rho,\cdots,\phi}(t_j, \bar{x}) - g^i_{\rho,\cdots,\phi}(t_j, x)] \Delta_j z^\rho \cdots \Delta_j z^\phi$$
$$+ \sum_{\mathscr{C}} [g^i_{\rho,\cdots,\phi}(t_k, \bar{x}) - g^i_{\rho,\cdots,\phi}(t_k, x)][z^\rho(t) - z^\rho(t_k)] \cdots$$
$$[z^\phi(t) - z^\rho(t_k)].$$

By hypothesis (7.10) of Theorem 7.1 with (5.1),

$$(8.12) \quad \left\| \sum_{j=1}^{k-1} [g^i_{\rho,\cdots,\phi}(t_j, \bar{x}) - g^i_{\rho,\cdots,\phi}(t_j, x)] \Delta_j z^\rho \cdots \Delta_j z^\phi \right.$$

$$\left. + [g^i_{\rho,\cdots,\phi}(t_k, \bar{x}) - g^i_{\rho,\cdots,\phi}(t_k, x)][z^\rho(t) - z^\rho(t_k)] \cdots [z^\phi(t) - z^\phi(t_k)] \right\|$$

$$\leqq B\left\{ \sum_{j=1}^{k-1} L^2(N(t_j))^2 \Delta_j t + L^2(N(t_k))^2 [t - t_k] \right\}^{1/2}$$

$$= BL\left\{ \int_a^t N(s)^2 \, ds \right\}^{1/2}.$$

Since $M - 1$ is an upper bound for $\|x\|$, by (8.7), $\|\bar{x}(t)\| \leq M - 1 + N(t)$, and the right members of (8.4) and (8.5) are at most $\varepsilon(M + N(t))$, $\varepsilon(M + N(t_{j+1}))$, respectively. So if C is the number of members of the set $\mathscr{C}$, from (g), (8.9), (8.10), (8.11), and (8.12), we deduce

$$(8.13) \quad \|\bar{x}(t) - x(t)\|$$
$$\leqq \varepsilon + \varepsilon\{M + N(t_k)\}(t_k - a) + \varepsilon\{M + N(t)\} + CBL\left\{ \int_a^t N(s)^2 \, ds \right\}^{1/2}.$$

The right member is a nondecreasing function of t, so (8.13) remains valid if we replace t_k in the right member by t and then replace t by any larger number,

or equivalently replace t in the left member by any smaller number. So

$$(8.14) \qquad N(t) \leqq \varepsilon[1 + M(1 + b - a)] + \varepsilon(1 + b - a)N(t)$$
$$+ CBL\left\{\int_a^t N(s)^2 \, ds\right\}^{1/2}.$$

By the fact we have chosen ε such that $0 < 2(1 + b - a)\varepsilon < 1$, this implies

$$(8.15) \qquad N(t) \leqq 2\varepsilon[1 + M(1 + b - a)] + 2CBL\left\{\int_a^t N(s)^2 \, ds\right\}^{1/2}$$

To condense notation, we write

$$(8.16) \qquad \begin{aligned} P &= 2[1 + M(1 + b - a)], \\ Q &= 2CBL. \end{aligned}$$

Then from (8.15) we can deduce that

$$(8.17) \qquad N(t) \leqq 2\varepsilon P \exp\{(2P^2Q^2[t - a])(a \leqq t \leqq b)\},$$

for (8.17) holds at $t = a$. If it does not hold everywhere in $[a, b]$, there is a first point t_0 at which it fails. Then it holds on $[a, t_0]$, so by (8.15),

$$(8.18) \qquad N(t_0) \leqq \varepsilon P + Q\left[\int_a^t 4\varepsilon^2 P^2 \exp\{4P^2Q^2[s - a]\}) \, ds\right]^{1/2}$$
$$= \varepsilon P + \varepsilon[\exp\{4P^2Q^2[t_0 - a]\} - 1]^{1/2}$$
$$< 2\varepsilon P \exp\{2P^2Q^2[t_0 - a]\},$$

contradicting the assumption that (8.17) fails at t_0. Since for every positive ε, (8.17) holds whenever (8.10) does, $\|\bar{x}(t) - x(t)\|$ converges uniformly to 0 as mesh $\Pi \to 0$, which completes the proof.

The only use made of the enlarged class defined by Assumption 7.4 was to guarantee that the coefficients $g^i_{\rho,\ldots,\phi}(t_j, \bar{x})$ are defined. If the $\bar{x}$ are the Cauchy-Maruyama functions defined by (8.2), they are in $H_{1/2}[T, F_{\cdot}]$, and we can use this for our class $\mathscr{P}$, abandoning Assumption 7.4.

9. Stochastic differential equations and related ordinary equations

We shall now revert back to stochastic differential equations like those in Sections 1 to 3, in which the coefficients g^i_ρ, and so forth, are functions of t and $x(t)$ and independent of earlier values of $x(\tau)$. We suppose that these have the properties ascribed to the coefficients G^i_0, and so forth, as stated after Assumption 7.3; but we use g instead of G. Moreover, we take T to be the same as $[a, b]$, and $y(t)$ is simply an initial value x_0, which is an F_a measurable r.v. For such equations we shall show that, with the definition in (3.5), equations of the form (3.4) have the stability property that for a rather large class of processes Z^ρ

interpolated in the z^ρ and having piecewise smooth sample paths, the solutions of (3.3) with Z^ρ tend to those with z^ρ, uniformly in near L_2 norm.

To avoid inordinately long formulae, we change the notation somewhat. We define

(9.1)
$$x^0(t) = t, \qquad z^0(t) = t, \qquad\qquad\qquad -\infty < t < \infty$$
$$g^0_0(x) = 1, \qquad g^0_1(x) = \cdots = g^0_r(x) = 0, \qquad\qquad x \in \mathbf{R}^{n+1}.$$

The variables α, β will always have range $\{0, \cdots, n\}$, and ρ, σ, τ will have range $\{0, \cdots, r\}$. A summation sign such as Σ_σ or Σ_ρ will denote the sum over the whole range of that variable. Also, Π will always denote a Cauchy partition with notation (8.1), and t_j will denote a division point of Π. An equation such as $u^i_\rho = v^i_\rho$ will always be understood to hold for all i, ρ in the range of those variables, unless some other range is expressly specified.

For all x^0 in $[a, b]$ and $(x^1, \cdots, x^n)$ in $\mathbf{R}^n$ we define

(9.2)
$$g^i_{\rho,\sigma}(x) = \sum_\alpha \frac{\partial g^i_\rho(x)}{\partial x^\alpha} g^\alpha_\sigma(x),$$

provided that the indicated derivatives exist. Equations (1.3) now take the form

(9.3)
$$x^i(t) = x^i_0 + \sum_\rho \int_a^t g^i_\rho(x(s)) \, dz^\sigma,$$

where the initial value x_0 is always assumed to be an F_a measurable r.v. The analogue of (3.4), with (3.5), is

(9.4)
$$x^i(t) = x^i_0 + \sum_\rho \int_a^t g^i_\rho(x(s)) \, dz^\rho(s) + \tfrac{1}{2} \sum_\rho \int_{\sigma a}^t g^i_{\rho,\sigma}(x(s)) \, dz^\rho(s) \, dz^\sigma(s).$$

This is not identical with (3.4), for the last sum contains terms with $\rho = 0$ or $\sigma = 0$, and (3.4) does not. However, by Theorem 6.2, all such integrals vanish for all processes z^ρ that we shall consider. Furthermore, even if ρ and σ are positive, the definition (9.2) contains a term (with $\alpha = 0$) which is lacking in (3.5). But by (9.1) this term is 0. So the solutions of (9.4) are the same as those of (3.4) with (3.5) for all processes z^ρ that we shall permit.

In Theorem 7.1, we assumed that the coefficients were Lipschitzian in $x(\cdot)$, and merely almost everywhere continuous in t. To simplify proofs, we now replace this by a somewhat unnecessarily strong substitute:

ASSUMPTION 9.1. *The functions g^i are continuously differentiable on the set of x with $a \leqq x^0 \leqq b$; and there is a positive L such that, if x and x'' are both in that set,*

(9.5)
$$|g^i_\rho(x') - g^i_\rho(x'')| \leqq L|x' - x''|,$$
$$|g^i_{\rho,\sigma}(x') - g^i_{\rho,\sigma}(x'')| \leqq L|x' - x''|.$$

Instead of restricting ourselves to linear interpolation as mentioned in Section 1, we shall permit certain other kinds. Let

$$(9.6) \qquad \phi_\rho(t): 0 \leqq t \leqq 1, \qquad \rho = 0, 1, \cdots, r,$$

be Lipschitzian functions such that

$$(9.7) \qquad \phi_\rho(0) = 0, \qquad \phi_\rho(1) = 1$$

and

$$(9.8) \qquad \phi_0(t) = t, \qquad 0 \leqq t \leqq 1.$$

Then for each z^ρ and each Cauchy partition Π, we define functions Z^ρ by setting

$$(9.9) \qquad Z^\rho(t, \omega) = z^\rho(t_j, \omega)$$

$$+ \phi_\rho\left(\frac{t - t_j}{t_{j+1} - t_j}\right)[z^\rho(t_{j+1}) - z^\rho(t_j)], \qquad t_j \leqq t \leqq t_{j+1}.$$

In particular,

$$(9.10) \qquad Z^0(t) = t, \qquad a \leqq t \leqq b.$$

We define

$$(9.11) \qquad J_{\rho,\sigma} = \int_0^1 [1 - \phi_\rho(s)]\dot{\phi}_\sigma(s)\, ds.$$

Our principal stability theorem, which we now state, overlaps considerably with the results of Wong and Zakai ([8], [9]). Although the present methods are different, Theorem 9.1 obviously owes its existence to those previous results. Besides this, the present version of Theorem 9.1 replaces an earlier version with stronger hypotheses because Professor Wong pointed out the desirability of improvement.

THEOREM 9.1. *Let Assumption 9.1 hold, and let the z^ρ satisfy Condition 6.1. Let $\phi_0, \cdots, \phi_r$ have the properties described above. Assume that for each ρ and σ in $\{0, 1, \cdots, r\}$, either:*

(i) *to each $\varepsilon > 0$ corresponds a $\delta > 0$ such that if $a \leqq s \leqq t \leqq b$ and $t - s < \delta$ then a.s.,*

$$(9.12) \qquad \begin{aligned} |E([z^\rho(t) - z^\rho(s)][z^\sigma(t) - z^\sigma(s)] \,|\, F_s)| &\leqq \varepsilon(t - s), \\ E([z^\rho(t) - z^\rho(s)]^2[z^\sigma(t) - z^\sigma(s)]^2 \,|\, F_s) &\leqq \varepsilon(t - s), \end{aligned}$$

or else

(ii) $J_{\rho,\sigma} = 1/2$.

Then, as mesh $\Pi \to 0$, the solution X of

$$(9.13) \qquad X^i(t) = x_0^i + \sum_\rho \int_a^t g_\rho^i(X(s))\dot{Z}^\rho(s)\, ds, \qquad a \leqq t \leqq b$$

converges uniformly in near L_2 norm on $[a, b]$ to the solution x of

$$(9.14) \qquad x(t) = x_0 + \sum_\rho \int_a^t g_\rho^i(x(s)) \, dz^\rho(s) + \tfrac{1}{2} \sum_{\rho, \sigma} \int_a^t g_{\rho, \sigma}^i(x(s)) \, dz^\rho(s) \, dz^\sigma(s),$$

$$a \leqq t \leqq b.$$

PROOF. Observe that if $\rho = \sigma$, condition (ii) is satisfied, while if $\rho = 0$ or $\sigma = 0$ condition (i) holds.

The solution x of (9.14) also satisfies

$$(9.15) \qquad x^i(t) = x_0^i + \sum_\rho \int_a^t g_\rho^i(x(s)) \, dz^\rho(s) + \sum_{\rho, \sigma} J_{\rho, \sigma} \int_a^t g_{\rho, \sigma}^i(x(s)) \, dz^\rho(s) \, dz^\sigma(s),$$

$$a \leqq t \leqq b,$$

since those integrals in (9.15) with coefficients $J_{\rho, \sigma} \neq 1/2$ all vanish by (5.2).

We again define $\Delta_j t$ and $\Delta_j z^\rho$ by (8.3). Let ε be positive, and let δ and A correspond to ε for all the z^ρ as in Condition 6.1. Let Π be a Cauchy partition with mesh $\Pi < \delta$. For each k in $\{1, \cdots, \ell\}$, we define A_k to be the set of ω in Ω such that the inequalities

$$(9.16) \qquad |\Delta_j z^\rho(\omega)| \leqq \varepsilon(\Delta_j t)^{1/3}, \qquad\qquad \rho = 1, \cdots, r$$

all hold for $j = 1, \cdots, k$; then $A_\ell \supseteqq A$. Corresponding to Π, we now define a process $\bar{x}$ as follows. First, $\bar{x}(a) = x_0$. Then, $\bar{x}(t_j, \omega)$ having been defined, we define

$$(9.17) \qquad\qquad \bar{x}^i(t_{j+1}, \omega) = X^i(t_{j+1}, \omega) \qquad \text{if} \quad \omega \in A_j,$$

$$(9.18) \qquad \bar{x}(t_{j+1}, \omega) = \bar{x}^i(t_j, \omega) + \sum_\rho g_\rho^i(\bar{x}(t_j, \omega)) \Delta_j z^\rho$$

$$+ \sum_{\rho, \sigma} J_{\rho, \sigma} g_{\rho, \sigma}^i(\bar{x}(t_j, \omega)) \Delta_j z^\rho \Delta_j z^\sigma \qquad \text{if} \quad \omega \in \Omega - A_j.$$

In either case we define

$$(9.19) \qquad\qquad \bar{x}^i(t, \omega) = \bar{x}^i(t_j, \omega) \qquad\qquad t_j \leqq t < t_{j+1}.$$

The set A_j defined by (9.16) is $F[t_{j+1}]$ measurable, and $Z^\rho(t)$ is a linear function of $z^\rho(t_j)$ and $z^\rho(t_{j+1})$ for $t_j \leqq t \leqq t_{j+1}$; so by (9.13) $X(t_{j+1})$ is a continuous function of the $z^\rho(t_h)$ for $h = 1, \cdots, j + 1$, and is $F[t_{j+1}]$ measurable. So by (9.17) and (9.18), $\bar{x}(t_{j+1})$ is $F[t_{j+1}]$ measurable, and hypothesis (h) is satisfied. So is (f); and (8.4) follows readily from (9.19) and Condition 4.1.

If $\omega \in A_j$, from (9.17) we obtain by integration by parts in (9.13),

$$(9.20) \quad \bar{x}^i(t_{j+1}) - \bar{x}^i(t_j) - \sum_\rho g_\rho^i(\bar{x}(t_j)) \Delta_j z^\rho$$

$$- \sum_{\rho, \sigma} g_{\rho, \sigma}^i(\bar{x}(t_j)) \int_{t_j}^{t_{j+1}} [Z^\rho(t_{j+1}) - Z^\rho(s)] \dot{Z}^\sigma(s) \, ds$$

$$= \sum_{\rho, \sigma} \int_{t_j}^{t_{j+1}} [g_{\rho, \sigma}^i[X(s)] - g_{\rho, \sigma}^i[\bar{x}(t_j)]][Z^\rho(t_{j+1}) - Z^\rho(s)] \dot{Z}^\sigma(s) \, ds.$$

For the rest of this proof C_1, C_2, and so forth, will denote positive numbers whose values are determined by the numbers n, r, K, L, $g_\rho^i(0)$ and $\sup |\dot{\phi}_\rho(t)|$; we omit the easy but uninspiring computation of the expressions for the C_i.

For t in $[t_j, t_{j+1}]$ and ω in A_j, we define

$$(9.21) \qquad N(t, \omega) = \sup \{|X(\tau, \omega) - \bar{x}(t_j, \omega)| : t_j \leq \tau \leq t\}.$$

Then by Assumption 9.1,

$$
\begin{aligned}
(9.22) \qquad |g_\rho^i(X(t, \omega))| &\leq |g_\rho^i(0)| + |g_\rho^i(\bar{x}(t_j, \omega)) - g_\rho^i(0)| \\
&\quad + |g_\rho^i(X(t, \omega)) - g_\rho^i(\bar{x}(t_j, \omega))| \\
&\leq C_1 + L|\bar{x}(t_j, \omega)| + LN(t, \omega).
\end{aligned}
$$

Hence by (9.17) and (9.16),

$$
\begin{aligned}
(9.23) \quad |X^i(t, \omega) - \bar{x}^i(t_j, \omega)| &\leq \sum_\rho \int_{t_j}^t |g_\rho^i(X(s))| \|\dot{Z}^\rho(s)| \, ds \\
&\leq \varepsilon[C_2 + C_3|\bar{x}(t_j, \omega)| + C_4 N(t, \omega)](\Delta_j t)^{1/3}.
\end{aligned}
$$

This remains valid if in the left member we replace t by any number τ in $[t_j, t]$, so

$$(9.24) \qquad N(t, \omega) \leq \varepsilon[C_5 + C_6|\bar{x}(t_j, \omega)| + C_7 N(t, \omega)](\Delta_j t)^{1/3}.$$

Since C_7 does not depend on ε, we may and shall restrict our attention to ε such that

$$(9.25) \qquad 0 < \varepsilon < \frac{1}{2C_7}(b - a)^{-1/3}.$$

Then from (9.24), we obtain

$$(9.26) \qquad N(t, \omega) \leq 2\varepsilon[C_5 + C_6|\bar{x}(t_j, \omega)|](\Delta_j t)^{1/3}.$$

From this, with (9.20) and (9.21),

$$
\begin{aligned}
(9.27) \quad &\left|\bar{x}^i(t_{j+1}) - \bar{x}^i(t_j) - \sum_\rho g_\rho^i(\bar{x}(t_j))\Delta_j z^\rho - \sum_{\rho,\sigma} J_{\rho,\sigma} g_{\rho,\sigma}^i(\bar{x}(t_j))\Delta_j z^\rho \Delta_j z^\sigma\right| \\
&\leq (C_8 + C_9|\bar{x}^i(t_j)|)\varepsilon\Delta_j t.
\end{aligned}
$$

If $\omega \in \Omega - A_j$, this is trivial; the left member of (9.27) is 0 by definition.

By (9.27),

$$
\begin{aligned}
(9.28) \quad &\left\|\bar{x}(t_{j+1}) - \bar{x}(t_j) - \sum_\rho g_\rho^i(\bar{x}(t_j))\Delta_j z^\rho - \sum_{\rho,\sigma} J_{\rho,\sigma} g_{\rho,\sigma}^i(\bar{x}(t_j))\Delta_j z^\rho \Delta_j z^\sigma\right\| \\
&\leq \|C_8 \varepsilon \Delta t\| + \|C_9 \varepsilon \Delta t \bar{x}(t_j)\|;
\end{aligned}
$$

so (8.5) is satisfied, and by Theorem 8.1, $\bar{x}$ converges to x uniformly in L_2 norm as mesh $\Pi \to 0$.

If $\omega \in A_j$, $\bar{x}(t_j)$ was defined to be $X(t_j)$, and by (9.21) and (9.26),

$$(9.29) \qquad |X^i(t) - \bar{x}^i(t)| \leq 2\varepsilon[C_5 + C_6|\bar{x}^i(t)|](\Delta_j t)^{1/3}.$$

Since $\bar{x}$ converges uniformly in L_2 norm to x, its L_2 norm is bounded, and (9.29) implies that if mesh Π is small,

$$(9.30) \qquad \left\{ \int_A |X(t, \omega) - \bar{x}(t, \omega)|^2 P(d\omega) \right\}^{1/2} \leq C_{10} \varepsilon (\text{mesh } \Pi)^{1/3}.$$

This, with the uniform convergence of $\bar{x}$ to x in L_2 norm, shows that X tends to x uniformly in near L_2 norm as mesh Π tends to 0.

10. Stability, and its limitations

Theorem 9.1 can be regarded as a statement about stability of the solutions of (9.14). Since the last set of stochastic integrals in (9.14) vanish for Lipschitzian z^ρ and z^σ, the theorem informs us that, on the family of disturbances z^ρ consisting of one process satisfying Condition 4.1 and Lemma 6.1 together with all processes interpolated in the z^ρ in accordance with Theorem 9.1, the solutions of (9.14) depend, in a stable or continuous manner, on the disturbances. By estimating the closeness of all approximations we could extend this to a larger collection of z^ρ, all satisfying Condition 4.1 and Condition 6.1 with the same constants, together with all disturbances interpolated in them as in Theorem 9.1.

It would be desirable to permit another kind of interpolation often encountered in applications, in which the z^ρ are approximated by functions Z^ρ that coincide with z^ρ at evenly spaced t_j and have derivatives whose Fourier transforms vanish outside some finite interval. Theorem 9.1 gives us a feeble substitute for this. Let Ψ be infinitely differentiable and nondecreasing on $(-\infty, \infty)$, with $\Psi(t) = 0$ if $t \leq 0$ and $\Psi(t) = 1$ if $t \geq 1$. We choose

$$(10.1) \qquad \phi_0(t) = t, \qquad \phi_\rho(t) = \Psi(t) \qquad\qquad \rho = 1, \cdots, r,$$

and for each Π, we interpolate Z^ρ in z^ρ using the ϕ_ρ and extend Z^ρ by constancy on $(-\infty, a]$ and $[b, \infty)$. Then the Z^ρ have Fourier transforms that tend to 0 at $\pm\infty$ faster than any negative power of the independent variable, and the Z^ρ satisfy the requirements of Theorem 9.1.

For the case $n = r = 1$, with z^1 a Wiener process, Wong and Zakai [9] have proved a theorem that shows the possibility of using Z^ρ whose derivatives have bounded spectra. Omitting superscripts i and ρ, let $Z_1, Z_2, \cdots$, be a sequence of processes on $[a, b]$ such that with probability 1, $Z_n(t, \omega)$ tends pointwise to $z(t, \omega)$ and has a bounded piecewise continuous derivative, and such that there are finite valued processes n_0, k such that a.s.

$$(10.2) \qquad\qquad |Z_n(t, \omega)| \leq k(\omega), \qquad a \leq t \leq b, \text{ if } n \geq n_0(\omega).$$

Wong and Zakai then showed that, under essentially the same hypotheses on g_1 and g_{11} as in Theorem 9.1 the solutions X of (9.13) converge almost surely to the solution x of (9.14) (in which we can replace $dzdz$ by dt, by (6.3)).

This theorem does not extend to the case $n = r = 2$, even when z^1 and z^2 are independent Brownian motions, as the following example shows. Consider the stochastic differential equations

$$(10.3) \qquad x^1(t) = \int_0^t dz^1(s), \qquad x^2(t) = \int_0^t x^1(s) \, dz^2(s), \qquad 0 \leqq t \leqq 1,$$

in which z^1 and z^2 are independent standard Wiener processes. Let ψ_1 and ψ_2 be infinitely differentiable nondecreasing functions on $(-\infty, \infty)$ such that

$$(10.4) \qquad \psi(t) = \begin{cases} 0 & \text{if} \quad t \leqq 0, \\ 1 & \text{if} \quad t \geqq 1/2. \end{cases}$$

$$(10.5) \qquad \psi(t) = \begin{cases} 0 & \text{if} \quad t \leqq 1/2, \\ 1 & \text{if} \quad t \geqq 1. \end{cases}$$

Then

$$(10.6) \qquad \begin{aligned} \int_0^1 \left[1 - \psi_1(s)\right] \dot{\psi}_2(s) \, ds &= 0, \\ \int_0^1 \left[1 - \psi_2(s)\right] \dot{\psi}_1(s) \, ds &= 1. \end{aligned}$$

Given a Cauchy partition Π, with the usual notation (8.1), for $\rho = 1, 2$ and for $t_j \leqq t \leqq t_{j+1}$, we define

$$(10.7) \qquad \begin{aligned} Z^\rho(t, \omega) &= z^\rho(t, \omega) + \psi_\rho\left(\frac{t - t_j}{\Delta_j t}\right) \Delta_j z^\rho(\omega), \\ \tilde{Z}^\rho(t, \omega) &= z^\rho(t, \omega) + \psi_{3-\rho}\left(\frac{t - t_j}{\Delta_j t}\right) \Delta_j z^\rho(\omega), \end{aligned}$$

if

$$(10.8) \qquad \Delta_j z^1(\omega) \Delta_j z^2(\omega) \geqq 0,$$

and we define $Z^\rho(t, \omega)$ and $\tilde{Z}^\rho(t, \omega)$ by (10.7) with the right members interchanged if (10.8) is false. We extend Z^ρ and $\tilde{Z}^\rho$ by constancy on $(-\infty, a]$ and on $[b, \infty)$. Then Z^ρ and $\tilde{Z}^\rho$, $\rho = 1, 2$, are infinitely differentiable, have the same bounds as z^ρ, and with probability 1 converge to z^ρ uniformly on $[a, b]$. By (9.20) and (10.6), for the corresponding solutions of (10.3) we have

$$(10.9) \qquad \begin{aligned} X^1(t_{j+1}) - X(t_j) &= \Delta_j z^1, \\ X^2(t_{j+1}) - X^2(t_j) &= X^1(t_j)\Delta_j z^2 + (\Delta_j z^1 \Delta_j z^2)^+. \end{aligned}$$

and

$$(10.10) \qquad \begin{aligned} \tilde{X}^1(t_{j+1}) - \tilde{X}^1(t_j) &= \Delta_j z^1, \\ \tilde{X}^2(t_{j+1}) - \tilde{X}^2(t_j) &= \tilde{X}^1(t_j)\Delta_j z^2 - (\Delta_j z^1 \Delta_j z^2)^-. \end{aligned}$$

Define

$$(10.11) \qquad\qquad \xi^\rho = X^\rho - \tilde{X}^\rho, \qquad\qquad \rho = 1, 2;$$

then from (10.9) and (10.10)

$$(10.12) \qquad\qquad \xi^1(t_{j+1}) - \xi^1(t_j) = 0,$$

$$(10.13) \qquad\qquad \xi^2(t_{j+1}) - \xi^2(t_j) = \xi^1(t_j) + \left| \Delta_j z^1 \Delta_j z^2 \right|.$$

From (10.12), $\xi^1(t_j) = 0$ for all j, so by (10.13)

$$(10.14) \qquad\qquad \xi^2(1) = \sum_j \left| \Delta_j z^1 \Delta_j z^2 \right|.$$

Since the $\Delta_j z^1$ and $\Delta_j z^2$ are independent normal r.v., $\Delta_j z^\rho$ having mean 0 and variance $\Delta_j t$, we readily compute

$$(10.15) \qquad\qquad E\big(\xi^2(1)\big) = \sum_j \frac{2\Delta_j t}{\pi} = \frac{2}{\pi},$$

and

$$(10.16) \qquad\qquad \operatorname{Var} \xi^2(1) \leqq E\Big(\sum_j \big[\Delta_j z^1 \big]^2 \big[\Delta_j z^2 \big]^2\Big)$$

$$= \sum_j (\Delta_j t)^2,$$

which tends to 0 with mesh Π. So $X^2(1) - \tilde{X}^2(1)$ tends in L_2 norm to $2/\pi$ as mesh Π tends to 0, and it is impossible that $X^2(1)$ and $\tilde{X}^2(1)$ both tend to the same limit $x^2(1)$, a.s., or even in probability.

The example shows the inherent limitations on stability of models. With such a simple system as (10.3), when mesh Π is small, the results of linear interpolation in z^ρ and of the interpolation (10.9) in z^ρ will have differences that are uniformly arbitrarily small for almost all ω. Yet the solutions X of the ordinary equations (10.4) and (10.5) corresponding to those two practically indistinguishable disturbances will not be arbitrarily close to each other in L_2 norm. Hence, no "selection principle" can possibly provide a model that is consistent, inclusive enough to include Lipschitzian processes and Brownian motions, and so thoroughly stable as to yield practically indistinguishable solutions corresponding to practically indistinguishable disturbances. The limited stability described in Theorem 9.1 may be about as much as we can attain.

Perhaps we are studying the problem from the wrong end. As mentioned in Section 1, if we wish to stay in the domain of trustworthiness of classical scientific theories, we should hold to Lipschitzian disturbances. Idealizations to martingales are made for mathematical convenience, and they depart from the Lipschitzian case so far that no martingale can have a.s. Lipschitzian sample paths unless the sample paths are a.s. constant (see Fisk, [3]).

In Theorem 9.1, and in the theorems of Wong and Zakai, the idealization is the starting point, and it is approximated by the Lipschitzian Z^ρ. Since it is the

Lipschitzian case that is presented to us by the outside world, it would seem more significant to find how well we can approximate the processes of classical theory by our idealizations, rather than the reverse. But this would appear to be a difficult undertaking.

11. A Runge-Kutta type of approximation

The Cauchy (or Euler) polygons are useful in the theory of ordinary differential equations, but for computation they are much inferior to the Runge-Kutta approximations. As adapted to equations (9.3), this method can be described thus. Given a partition Π (with the usual notation (8.1)), we define $y(a) = x_0$, and then define $y(t_2)$, $\cdots$ successively as follows. From $y(t_j)$ we first compute, as in (8.2), the value of

$$(11.1) \qquad y^i(t_j) + \sum_\rho g_\rho^i[y(t_j)]\Delta_j z^\rho.$$

But instead of using the g_ρ^i corresponding to these sums as coefficients for the next step, as in the Cauchy-Maruyama process, we average them with the $g_\rho^i(y(t_j))$ to furnish a second approximation to the values of the g_ρ^i for use in estimating $y(t_{j+1})$. Thus, we have

$$(11.2) \qquad y(t_{j+1}) = y(t_j) + \tfrac{1}{2}\sum_\rho \left[g_\rho^i(y(t_j))\Delta_j z^\rho\right]$$

$$+ \tfrac{1}{2}\sum_\rho g_\rho^i[y(t_j) + \sum_\sigma g_\sigma(y(t_j))\Delta_j z^\sigma]\Delta_j z^\rho.$$

The values of y at points interior to intervals $[t_k, t_{k+1}]$ are of secondary interest; we could, for example, define them by linear interpolation.

The preservation of a formula or an algorithm is a much less basic stability property than that discussed in the preceding section. Nevertheless, it is to some extent significant, as well as computationally convenient, that if we try to approximate solutions of (9.3) by the Runge-Kutta method for processes satisfying the hypotheses of Theorem 9.1 and one more continuity condition Assumption 11.1, the approximations will converge, not to the solution of (9.3), but to the solution of (9.14). This in a sense gives added recommendation to our "selection principle." But besides this, it permits us to use a well-known computation procedure to approximate the solution of (9.14), whether the z^ρ are Lipschitzian or are martingales or any other processes satisfying the hypotheses of Theorem 9.1, without having to interpolate to find the Z^ρ and without having to solve equations (9.13).

Equation (11.2) may be regarded as the first step in the iterative solution of

$$(11.3) \qquad y(t_{j+1}) = y(t_j) + \tfrac{1}{2}\sum_\rho \{g_\rho^i[y(t_j)] + g_\rho^i[y(t_{j+1})]\}\Delta_j z^\rho.$$

To guarantee the convergence of such an iterative process, it is desirable and usual to make assumptions that guarantee that the successive corrections form a diminishing sequence. One such assumption, for the present problem, is the following.

ASSUMPTION 11.1. *There are positive numbers* δ_1, L_1 *such that if* x_1 *and* x_2 *are points of* $\mathbf{R}^{n+1}$ *with* x_1^0 *in* $[a, b]$ *and* $|x_2 - x_1| \leqq \delta_1(1 + |x_1|)$, *then*

$$(11.4) \qquad |g_{\rho, x^{\alpha}}^i(x_2) - g_{\rho, x^{\alpha}}^i(x_1)| \leqq \frac{L_1 |x_1 - x_2|}{1 + |x_1 - x_2|}.$$

This rather strong uniform continuity requirement will be further discussed after proving the next theorem.

THEOREM 11.1. *Let the* z^ρ *and* g_ρ^i *satisfy the hypotheses of Theorem 9.1, and also satisfy Assumption 11.1. For each Cauchy partition* Π *of* $[a, b]$, *let* y *be the process determined by the Runge-Kutta process* (11.2), *with linear interpolation between the division points of* Π. *Then as mesh* $\Pi \to 0$, y *converges on* $[a, b]$ *uniformly in near* L_2 *norm to the solution* x *of* (9.14).

PROOF. Let ε be positive, and let δ and A correspond to ε for all the z^ρ as in Condition 6.1. We define the sets $A_1, \cdots, A_\ell$ as in the sentence containing (9.16) and we define a process $\bar{x}$ corresponding to Π as follows. First, $\bar{x}(a) = x_0$. Next, $\bar{x}(t_j)$ having been defined, we define $\bar{x}(t_{j+1})$ by

$$(11.5) \qquad \bar{x}(t_{j+1}, \omega) = y(t_{j+1}, \omega) \qquad \text{if} \quad \omega \, \varepsilon \, A_j,$$

$$(11.6) \qquad \bar{x}(t_{j+1}, \omega) = \bar{x}(t_j, \omega) + \sum_\rho g_\rho^i(\bar{x}(t_j, \omega))\Delta_j z^\rho$$
$$+ \tfrac{1}{2} \sum_{\rho, \sigma} g_{\rho, \sigma}^i(\bar{x}(t_j, \omega))\Delta_j z^\rho \Delta_j z^\sigma$$

$$\text{if} \quad \omega \in \Omega - A_j.$$

Finally, we define

$$(11.7) \qquad \bar{x}^i(t, \omega) = \bar{x}^i(t_j, \omega), \qquad\qquad t_j \leqq t < t_{j+1}.$$

It is easy to verify that these $\bar{x}$ satisfy conditions (f), (h), and (8.4). For ω in A_j, by applying the theorem of the mean to (11.2), we find

$$(11.8) \qquad y^i(t_{j+1}) = y^i(t_j) + \tfrac{1}{2} \sum_\rho g_\rho^i(y(t_j))\Delta_j z^\rho$$
$$+ \tfrac{1}{2} \sum_\rho \{g_\rho^i(y(t_j)) + \sum_\sigma g_{\rho, \sigma}^i(\eta)\Delta_j z^\sigma\}\Delta_j z^\rho,$$

where

$$(11.9) \qquad \eta^i = y^i(t_j) + \theta_j^i(\omega) \sum_\tau g_\tau^i(y(t_j))\Delta_j z^\tau$$

for some $\theta_j^i(\omega)$ in $(0, 1)$. As in Section 9, we use C_1, C_2, and so forth, to denote numbers determined by the data of the problem. Then, by Assumption 11.1, (11.9) implies

(11.10) $$\left| g_{\rho,\sigma}^i(\eta) - g_{\rho,\sigma}^i(y(t_j)) \right| \leqq \left(C_1 + C_2 |y(t_j)| \right) \varepsilon (\Delta_1 t)^{1/3}.$$

From this and (11.8), since $\omega \in A_j$,

(11.11) $$\left| \bar{x}^i(t_{j+1}) - \bar{x}^i(t_j) - \sum_\rho g_\rho^i(\bar{x}(t_j)) \Delta_j z^\rho - \tfrac{1}{2} \sum_{\rho,\sigma} g_{\rho,\sigma}^i(\bar{x}(t_j)) \Delta_j z^\rho \Delta_j z^\sigma \right|$$
$$\leqq \left(C_3 + C_4 |\bar{x}(t_j)| \right) \varepsilon^3 \Delta_j t.$$

This also holds if $\omega \in \Omega - A_j$, since then the left member is 0 by definition. By taking the expectation of the square of the left member of (11.11), we find, as in (9.28), that (8.5) also is satisfied.

Now by Theorem 8.1, $\bar{x}$ converges to the solution x of (9.14) uniformly in L_2 norm, as mesh $\Pi \to 0$. Since x is continuous in L_2 norm and the $\bar{x}(t_j)$ are uniformly close in L_2 norm to $x(t_j)$ at all division points of Π if mesh Π is small, it is easy to see that if we modify $\bar{x}$ by retaining its values at the t_j but interpolating linearly between them, the modified process also converges to x uniformly in L_2 norm. But this modified process coincides with the Runge-Kutta process y for all ω in A, and so the proof is complete.

Assumption 11.1 is strong, but from an experimental or computational point of view it can be tolerated. Ordinarily there will be some bound B on the norms of the x that interest us. In an experiment, points with $|x| > B$ will make the points too far away to be involved in the process under investigation; in computation, B could be a bound on the numbers within the machine's capacity. If we replace the $g_\rho^i(x)$ by other functions $G_\rho^i(x)$ that satisfy Assumption 11.1 and coincide with $g_\rho^i(x)$ whenever $|x| \leq B$, the solution of (9.14) with coefficients G_ρ^i will coincide with the solution of (9.13) as written unless the solution somewhere has norm greater than B: and unless the probability of this is negligibly small, we face worse troubles than the mere nonconvergence of the Runge-Kutta procedure.

Professor H. Rubin informs me that Dr. Donald Fisk, in his doctoral dissertation at Michigan State University, has defined and studied a stochastic integral which is the limit of "trapezoidal rule" approximations, the values of the integrand at the beginning and end of each interval $[t_j, t_{j+1}]$ being averaged. (Professor Rubin also furnished reference [2].) Existence theorems are established for integrals $\int f dz$ in which z is a quasi-martingale (see Fisk, [3]). I have not had the opportunity of seeing this dissertation, but it is evident that the application of Fisk's integral to differential equations (1.3) must be closely related to the procedure described in Theorem 11.1, and even more closely related to the process mentioned in (11.3).

REFERENCES

[1] J. L. Doob, *Stochastic Processes*, New York, Wiley, 1953.
[2] D. L. Fisk, "Quasi-Martingales and stochastic integrals," Technical Report No. 1, Dept. Statistics, Michigan State University, 1963.
[3] ———, "Quasi-Martingales," *Trans. Amer. Math. Soc.*, Vol. 120 (1965), pp. 369–389.
[4] G. Maruyama, "Continuous Markov processes and stochastic equations," *Rend. Circ. Mat. Palermo*, Vol. 4 (1955), pp. 48–90.
[5] E. J. McShane, "Stochastic integrals and stochastic functional equations," *SIAM J. Appl. Math.*, Vol. 17 (1969), pp. 287–306.
[6] ———, "Toward a stochastic calculus," *Proc. Nat. Acad. Sci., U.S.A.*, Vol. 63 (1969), Part I, pp. 275–280; Part II, pp. 1084–1087.
[7] J. Neveu, *Mathematical Foundations of the Calculus of Probability* (translated by A. Feinstein), San Francisco, Holden-Day, 1965.
[8] E. Wong and M. Zakai, "On the relation between ordinary and stochastic differential equations," *Internat. J. Engrg. Sci.*, Vol. 3 (1965), pp. 213–229.
[9] ———, "On the convergence of ordinary integrals to stochastic integrals," *Ann. Math. Statist.*, Vol. 36 (1965), pp. 1560–1564.