Stochastics, 1979, Vol. 3, pp. 105–125
0090-2234/79/0302-0105 $04.50/0

A Formal Approach to Stochastic Integration and Differential Equations

ARTHUR J. KRENER†
Department of Mathematics, University of California, Davis, California 95616

(*Received October 23, 1978; in final form May 10, 1979*)

Stochastic integrals are defined using a differential rule and the fundamental theorem of calculus. It is shown that such integrals lead to the solution of stochastic differential equations driven by a single Wiener process or square integral sample path continuous martingale.

1. INTRODUCTION

The importance of stochastic models has clearly been demonstrated in numerous applications in science and engineering. Frequently the models take the form of an initial value problem for an ordinary differential equation with stochastic driving terms.

$$dX = f(t, X)dt + \sum_{i=1}^{k} g_i(t, X)dW_i$$

$$X(0) = X^0. \tag{1.1}$$

On the intuitive level, the solution $X(t)$ is a stochastic process representing the ensemble of states of the system at time t, X^0 is the random initial condition, $f(x)$ the disturbance-free dynamics, the Wiener process $W_i(t)$ is the accumulated disturbance of type i to time t and $g_i(x)dW_i$ the infinitesimal effect of an incremental disturbance dW_i on a system in state x at time t.

†This research supported in part by The National Science Foundation under grant MC 575-05248 A01.

To make (1.1) mathematically precise is difficult because the sample paths of a Wiener process are of unbounded variation on any interval almost surely. An alternate formulation of (1.1) is the integral equation

$$X(t)=X^0+\int_0^t f(s, X(s))ds+\sum_{i=1}^{k} \int_0^t g_i(s, X(s))dW_i(s) \tag{1.2}$$

but the meaning of this requires a definition of the stochastic integrals.

Wiener defined stochastic integrals of the form

$$\int_0^1 h(t)dW(t) \tag{1.3}$$

where h is a square integrable function, Ito [2] and Stratonovich [5] extended this to integrals of the form

$$\int_0^1 Y(t)dW(t) \tag{1.4}$$

where $Y(t)$ is stochastic process satisfying certain conditions. Wiener and Ito took a Lebesque-like approach, i.e., first defining (1.3) and (1.4) for simple functions or processes and then extending it to a larger class by a limiting process. Stratonovich took a Riemannian approach defining (1.4) as the limit of Riemann sums where the integrand is evaluated at the midpoint of each subinterval. The Stratonovich integral is less general than the Ito integral and generally gives a different answer. A variation of Stratonovich is to evaluate the integrand in the Riemann sum at the point of each subinterval which splits it into pieces of proportion λ and $1-\lambda$. This is called the generalized Stratonovich or λ-integral, $\lambda=1/2$ is Stratonovich and $\lambda=0$ agrees with Ito when the former is defined. The Stratonovich integral can also be defined via a "correction term" added to the Ito integral. The correction involves the quadratic variation of integrand with the Wiener process. Details can be found in the Appendix of [12].

In this paper we take a formal approach; instead of defining the stochastic integral we shall define a family of chain rules indexed by λ, $0 \leqq \lambda \leqq 1$, for the composition of a C^1 function and a stochastic process. Each chain rule and the fundamental theorem of calculus fixes the definition of a stochastic integral,

$$\int_0^1 h(t, W(t))\, d_\lambda W \tag{1.5}$$

where $h(t, w)$ is a C^1 function. This integral is approximately as general as the corresponding λ-integral and agrees with it when both are defined.

The formal approach is in the spirit of Schwartzian distributions and generalized–stochastic processes [1]. It generalized the alternate definition of the Wiener integral (1.3) for C^1 functions $h(t)$ using integration by parts as given by Paley, Wiener and Zygmund [4]. Generalizations to integrals involving more than one noise fail because of a geometric obstruction.

In Section 3 we use the formal approach to define the solution of the stochastic differential equation (1.1) deriven by a single noise ($k=1$). This is done by showing that under suitable hypothesis on f and g there exist a unique random field $X(t,w)$ depending on two parameters t and w with C^1 sample functions almost surely. The solution of (1.1) is $X(t) = X(t,w(t))$. This is related to work of Lamperti [3], Sussmann [6] and Doss [11]. They have formally defined the solution of the stochastic differential equation (1.1) for $k=1$ directly, without having first defined a stochastic integral. We discuss their work in more detail in Section 3.

Although the formal approach to (1.1) is more restrictive in that it requires $k=1$, it has some advantages. For example it shows that the solution is of the form $X(t)=X(t,W(t))$ where $X(t,w)$ is a random field whose sample functions are C^1 almost surely. From the formal approach it is easy to see why the correction term of Wong and Zakai [9] must be added to (1.1) when $W(t)$ is approximated by a Lipschitz continuous process. The formal approach may prove useful in developing numerical schemes to simulate solutions to (1.1).

In Section 4 we extend the formal approach to differential equations driven by a square integrable sample path continuous martingale.

2. INTEGRATION

Let $W(t)$ be a Wiener process; we wish to define integrals of the form

$$\int_a^b h(t, W(t))dW(t) \tag{2.1}$$

when $h(t,w)$ is a real-valued function which is C^1 in t and w. To do so we shall use two axioms. The first is the fundamental theorem of calculus:

$$\int_a^b dH(t, W(t)) = H(t, W(t))]_a^b. \tag{2.2}$$

The second is the chain rule for differentials, actually a family of chain roles indexed by λ,

$$d_\lambda H(t, W(t)) = H_t(t, W(t))dt + H_w(t, W(t))d_\lambda W(t) + (1/2-\lambda)H_{ww}(t, W(t))dt \tag{2.3}$$

The subscripts t and w denote partial differentiation. We restrict (2.3) to functions $H(t,w)$ which are continuously differentiable with respect to t and twice continuously differentiable with respect to w. If $\lambda=1/2$, (2.3) is standard chain rule and if $\lambda=0$ it is a special case of the Ito differential rule.

We take (2.2) and (2.3) as primitive and use them to define (2.1). The generalized Stratonovich approach is to define (2.1) by Riemannian sums, take (2.2) as primitive and deduce (2.3) in the following fashion.

Define a family of difference operators

$$\Delta_\lambda f(t)=f(t+(1-\lambda)\Delta t)-f(t-\lambda\Delta t)$$

for $0\leqq\lambda\leqq 1$. If $\lambda=0$ (1/2 or 1) Δ_λ is the forward (centered or backward) difference operator. If $\Delta t=(b-a)/n$ and $t_i=a+i\Delta t$

$$H(t,W(t))]_a^b=\lim_{\Delta t\to 0}\sum_{i=1}^n \Delta_\lambda H(t_i,W(t)).$$

Expanding $\Delta_\lambda H(t,W(t))$ in a Taylor series based at t_i we obtain

$$\Delta_\lambda H(t_i,W(t_i))=H_t(t_i,W(t_i))\Delta t+H_w(t_i,W(t_i))\Delta_\lambda W(t_i)$$
$$+(\tfrac{1}{2}-\lambda)H_{ww}(t_i,W(t_i))(\Delta_\lambda W(t_i))^2+o(\Delta t)+o(\Delta_\lambda W(t_i))^2 \qquad (2.4)$$

Then

$$\sum_{i=1}^n H_t(t_i,W(t_i))\Delta t$$

converges almost surely to the sample path Riemann integral

$$\textstyle\int_a^b H_t(t,W(t))dt,$$

$$\sum_{i=1}^n H_w(t_i,W(t_i))\Delta_\lambda W$$

converges by definition to the generalized Stratonovich integral

$$\textstyle\int_a^b H_w(t,W(t))d_\lambda W,$$

and it can be shown (McKean [8] pp. 34–35) that if Δt goes to zero fast enough then

$$\sum_{i=1}^n H_{ww}(t_i,W(t_i))(\Delta_\lambda W(t_i))^2$$

has the same limit almost surely as

$$\sum_{i=1}^n H_{ww}(t_i,W(t_i))\Delta t$$

which is of course the sample path Riemann integral

$$\int_a^b H_{ww}(t, W(t))dt.$$

The sum

$$\sum_{i=1}^{n} o(\Delta t) + o(\Delta_\lambda W(t_i))^2$$

goes to zero almost surely.

To define (2.1) where $h(t,w)$ is C^1 we choose a function $H(t,w)$ such that

$$H_w(t,w) = h(t,w).$$

From (2.2) and (2.3) we obtain

$$\int_a^b h(t, W(t))d_\lambda W(t) = H(t, W(t))]_a^b - \int_a^b H_t(t, W(t))dt + (\lambda - 1/2)\int_a^b H_{ww}(t, W(t))dt \qquad (2.5)$$

The integrands on the right are sample path continuous almost surely and hence Riemann integrable.

This is generalization of Wiener integration as defined by Paley, Wiener and Zygmund [4]. For C^1 functions $h(t)$ they defined

$$\int_a^b h(t)dW(t) = h(t)W(t)]_a^b - \int_a^b W(t)dh(t).$$

The principal weakness of the formal approach (2.5) is that it does not appear to generalize to integrals involving two Wiener processes, i.e. of the form

$$\int_a^b h(t, W_1(t), W_2(t))dW_1(t) \qquad (2.6)$$

Similar problems have been reported by Lamperti, Sussmann and Doss regarding the formal solution of the stochastic differential equation (1.1) when $k > 1$.

The obstruction is geometric in character. Suppose we try to integrate (2.10) as before, then we must find a function $H(t, w_1, w_2)$ such that

$$H_{w_1} = h \quad \text{and} \quad H_{w_2} = 0.$$

This is not always possible. If h does not depend on t this can be put another way. Any one form $h(w)dw$ in one independent variable is closed hence exact on simply connected domains but a one form $h(w_1, w_2)dw_1$ in two variables need not be closed. For example, consider $w_2 dw_1$.

3. DIFFERENTIAL EQUATIONS

Consider the stochastic differential equation

$$dX = f(t, X)dt + g(t, X)dW$$
$$X(0) = X^0 \tag{3.1}$$

or in its integral formulation

$$X(t) = X^0 + \int_0^t f(s, X(s))ds + \int_0^t g(s, X(s))d_\lambda W. \tag{3.2}$$

The solution will depend on λ; for $\lambda = 0$ we expect the Ito solution and for $\lambda = 1/2$ the Stratonovich solution. The appropriate choice of λ depends on the heuristics of the model. For example if (3.1) is the infinitesimal limit of a difference equation

$$\Delta_\lambda X = f(t, X)\Delta t + g(t, X)\Delta_\lambda W. \tag{3.3}$$

then the appropriate λ is obvious. For a further discussion of this point, see Turelli [7].

Lamperti [3] and Sussmann [6] obtained the Ito and Stratonovich solutions of (3.1) essentially by making a change of coordinates in (t, x)-space which takes the coefficient of $d_\lambda W(t)$ to a constant function. The former used the Ito differential rule ((2.3) with $\lambda = 0$) and the latter the standard chain rule ((2.3) with $\lambda = 1/2$) to compute the effect of the coordinate change on (3.1). It is interesting to note that the same change of coordinates works for all λ.

In the new coordinates the second integrand of (3.2) is constant so the integral follows from the fundamental theorem of calculus. This allows one to demonstrate by Picard iteration the existence and uniqueness of the solutions of (3.1) for almost every sample path under suitable regularity conditions on f and g, namely that f and the partial derivatives of g are locally Lipschitz continuous in x. This technique cannot be used to solve

$$dX = f(t, X)dt + g_1(t, X)dW_1 + g_2(t, X)dW_2 \tag{3.4}$$

because regardless of which chain rule is employed, there does not always exist a change of coordinates which makes the coefficients of both dW_1 and dW_2 constant. If g_1 and g_2 are commuting vector fields then such a change of coordinates can be found. Commuting vector fields are geometrically dual to closed one forms.

Doss used a similar approach, noticing that if $z(t)$ is a solution of

$$\dot{z}=g(t,z)$$

then $z(t)=z(W(t))$ is a solution of

$$dz=g(t,z)dW.$$

This solves (3.1) if $f=0$ if $f\neq 0$ then one can make a change of coordinates in (t,x) space which transforms (3.1) to a differential equation with $f=0$.

We would like to demonstrate the existence and uniqueness of solutions to (3.1) using the formal integral of Section 2.

Solutions to (3.1) using this integral exist and are unique under precisely the same conditions as for Stratonovich, Lamperti and Sussmann integrals. The Ito integral requires the weaker hypothesis that f and g be locally Lipschitz.

THEOREM 1 *Let f, $\partial g/\partial t$ and $\partial g/\partial x$ be locally Lipschitz continuous in x then for any $\lambda\in[0,1]$ there exists a random field $X(t,w)$ with C^1 sample functions almost surely and a random variable $T>0$ such that $X(t)=X(t,W(t))$ satisfies (in the sense of (2.3)) the integral equation (3.2) almost surely for $0\leqq t\leqq T$. Moreover the solution $X(t)$ is unique in that any two solutions agree almost surely on their common domain of definition.*

The theorem is local in character; solutions exist and are unique for small time periods. Of course they can be continued by the standard techniques and if suitable growth conditions are assumed then solutions will exist for all t.

The solution is constructed independently for each sample point $\omega\in\Omega$. We use a lower case symbol to denote the sample point evaluation of a random variable or process denoted by the corresponding upper case symbol. For example,

$$w(t)=W(t,\omega)$$

$$x(t,w)=X(t,w,\omega)$$

$$x^0=X^0(\omega).$$

Without loss of generality we take $w(0)=0$.

As with most existence and uniqueness theorems for ordinary differential equations the proof of the above depends upon the contraction mapping principle. Let $C^1(\delta,\varepsilon)$ be the space of n vector valued functions $y(t,w)$ which are continuously differentiable for $0\leqq t\leqq\delta$ and $|w|\leqq\varepsilon$. We endow $C^1(\delta,\varepsilon)$ with the topology of uniform convergence of the function and its first partials. For fixed $w(t)$ let F be the map from $C^1(\delta,\varepsilon)$ to itself defined by

$$F(y)(t,w)=x^0+\int_0^t f(s,y)(s,w(s)))ds$$
$$+H(t,w)-\int_0^t H_t(s,w(s))ds+(\lambda-1/2)\int_0^t h_w(s,w(s))ds \qquad (3.5)$$

where

$$h(t,w)=g(t,y(t,w)) \qquad (3.6a)$$

$$H(t,w)=\int_0^w h(t,v)dv \qquad (3.6b)$$

and subscripts denote partial differentiation.

Notice that

$$F(y)(t,w(t)) = x^0+\int_0^t f(s,y(s,w(s)))ds+\int_0^t g(s,y(s,w(s)))d_\lambda w. \qquad (3.7)$$

The following lemmas demonstrate that F is a local contraction, from this the theorem follows. We use $\|\cdot\|$ to denote a norm on $\mathbb{R}^n$ and also the associated operator norm on $\mathbb{R}^{n\times n}$.

LEMMA 1 *Let L, C be chosen so that for all y^1, y^2, $\|y^i-x^0\|\leqq C$ and $0\leqq t\leqq\varepsilon$, the following inequalities hold*

$$\|f(t,y^1)\|,\|g(t,y^1)\|,\left\|\frac{\partial g}{\partial t}(t,y^1)\right\|,\left\|\frac{\partial g}{\partial t}(t,y^1)\right\|\leqq L \qquad (3.8a)$$

and

$$\|f(t,y^1)-f(t,y^2)\|,\|g(t,y^1)-g(t,y^2)\|,\left\|\frac{\partial g}{\partial t}(t,y^1)-\frac{\partial g}{\partial t}(t,y^2)\right\|,$$
$$\left\|\frac{\partial g}{\partial x}(t,y^1)-\frac{\partial g}{\partial x}(t,y^2)\right\|\leqq L\|y^1-y^2\|. \qquad (3.8b)$$

There exist sufficiently small δ, $\varepsilon>0$ so that if $y(t,w)\in C^1$ (δ,ε) and

$$\|y(t,w)-x^0\|\leqq C \qquad (3.9a)$$

$$\left\|\frac{\partial y}{\partial t}(t,w)\right\|\leqq 1+2L+2L^2 \qquad (3.9b)$$

$$\left\|\frac{\partial y}{\partial w}(t,w)\right\|\leqq L \qquad (3.9c)$$

for $0\leqq t\leqq\delta$ and $|w|\leqq\varepsilon$ then $F(y)(t,w)\in C^1$ (δ,ε) and satisfies (3.9) also for $0\leqq t\leqq\delta$ and $|w|\leqq\varepsilon$.

Proof For any choice of $\varepsilon>0$ we will always choose $\delta>0$ so that $|w(t)| \leqq \varepsilon$ for $0 \leqq t \leqq \delta$.

From the definition of F

$$\|F(y)(t,w)-x^0\| \leqq \|\textstyle\int_0^t f(s,y(s,w(s)))ds\| + \|H(t,w)\| + \|\int_0^t H_t(s,w(s))ds\| + \frac{1}{2}\|\int_0^t h_w(s,w(s))ds\|$$

We estimate the first two terms,

$$\|\textstyle\int_0^t f(s,y(s,w(s)))ds\| \leqq Lt \leqq L\delta,$$

$$\|H(t,w)\| = \|\textstyle\int_0^w g(t,y(t,v))dv\| \leqq L|w| \leqq L\delta.$$

As for the third term,

$$\begin{aligned}\|H_t(t,w)\| &= \|\textstyle\int_0^w h_t(t,v)dv\| \\ &= \left\| \int_0^w \frac{\partial g}{\partial t}(t,y(t,v)) + \frac{\partial g}{\partial x}(t,y(t,v))\frac{\partial y}{\partial t}(t,v)dv \right\| \\ &\leqq |w|(L+L(1+2L+2L^2)) \leqq 2\varepsilon(L+L^2+L^3)\end{aligned} \tag{3.10}$$

So

$$\|\textstyle\int_0^t H_t(s,w(s))ds\| \leqq 2\delta\varepsilon(L+L^2+L^3)$$

Similarly

$$\|h_w(t,w)\| = \left\|\frac{\partial g}{\partial x}(t,y(t,w))\frac{\partial y}{\partial w}(t,w)\right\| \leqq L^2 \tag{3.11}$$

so

$$\|\textstyle\int_0^t h_w(s,w(s))ds\| \leqq tL^2 \leqq \delta L^2$$

Hence if δ, ε are sufficiently small

$$\|F(y)(t,w)-x^0\| \leqq C.$$

Using the definition of F

$$\left\|\frac{\partial F(y)}{\partial t}(t,w)\right\| \leqq \|f(t,y(t,w(t)))\| + \|H_t(t,w)\| + \|H_t(t,w(t))\| + \tfrac{1}{2}\|h_w(t,w(t))\|$$

From (3.8a) the first term is bounded by L, from (3.10) the next two are each bounded by $2\varepsilon(L+L^2+L^3)$ and from (3.11) the last is bounded by $1/2L^2$. Hence if $\varepsilon \leqq 1/(4L)$ then

$$\left\|\frac{\partial F(y)}{\partial t}(t, w)\right\| \leqq 1+2L+2L^2$$

Finally

$$\left\|\frac{\partial F(y)}{\partial w}(t, w)\right\| = \|H_w(t, w)\| = \|h(t, w)\|$$

$$= \|g(t, y(t, w))\| \leqq L. \qquad \text{Q.E.D.}$$

LEMMA 2 *Let L, C be chosen so that* (3.8) *holds. There exist sufficiently small δ, $\varepsilon > 0$ so that if $y^i(t, w) \in C^1(\delta, \varepsilon)$, satisfies* (3.9) *and*

$$\|y^1(t, w) - y^2(t, w)\| \leqq K(t+|w|) \tag{3.12a}$$

$$\left\|\frac{\partial y^1}{\partial t}(t, w) - \frac{\partial y^2}{\partial t}(t, w)\right\| \leqq K \tag{3.12b}$$

$$\left\|\frac{\partial y^1}{\partial w}(t, w) - \frac{\partial y^2}{\partial w}(t, w)\right\| \leqq K/(2L) \tag{3.12c}$$

for $0 \leqq t \leqq \delta$ and $|w| \leqq \varepsilon$ then $F(y^i)(t, w) \in C^1(\delta, \varepsilon)$, satisfies (3.9) *and*

$$\|F(y^1)(t, w) - F(y^2)(t, w)\| \leqq K(t+|w|)/2 \tag{3.13a}$$

$$\left\|\frac{\partial F(y^1)}{\partial t}(t, w) - \frac{\partial F(y^2)}{\partial t}(t, w)\right\| \leqq K/2 \tag{3.13b}$$

$$\left\|\frac{\partial F(y^1)}{\partial w}(t, w) - \frac{\partial F(y^2)}{\partial w}(t, w)\right\| \leqq K/(4L) \tag{3.13c}$$

for $0 \leqq t \leqq \delta$ and $|w| \leqq \varepsilon$.

Proof The first two assertions follow from Lemma 1, as for the last,

$$\|F(y^1)(t, w) - F(y^2)(t, w)\| \leqq \left\|\textstyle\int_0^t f(s, y^1(s, w(s))) - f(s, y^2(s, w(s)))ds\right\|$$
$$+ \|H^1(t, w) - H^2(t, w)\| + \left\|\textstyle\int_0^t H_t^1(s, w(s)) - H_t^2(s, w(s))ds\right\|$$
$$+ \tfrac{1}{2}\left\|\textstyle\int_0^t h_w^1(s, w(s)) - h_w^2(s, w(s))ds\right\|$$

where

$$h^i(t, w) = g(t, y^i(t, w))$$
$$H^i(t, w) = \textstyle\int_0^w h^i(t, v)dv.$$

By the Lipschitz continuity of f,

$$\left\|\int_0^t f(s, y^1(s, w(s))) - f(s, y^2(s, w(s)))ds\right\|$$
$$\leqq L\int_0^t \|y^1(s, w(s)) - y^2(s, w(s))\|ds$$
$$\leqq LK\int_0^t (\varepsilon+\delta)dt \leqq (\varepsilon+\delta)LKt \leqq (\varepsilon+\delta)LK(t+|w|).$$

Now

$$\|H^1(t, w) - H^2(t, w)\| = \left\|\int_0^w h^1(t, v) - h^2(t, v)dv\right\|$$
$$\leqq L\int_0^{|w|} \|y^1(t, v) - y^2(t, v)\|dv$$
$$\leqq LK\int_0^{|w|} (\varepsilon+\delta)dv \leqq (\varepsilon+\delta)LK|w| \leqq (\varepsilon+\delta)LK(t+|w|).$$

As for the third term we first bound

$$\|h_t^1(t, w) - h_t^2(t, w)\|$$
$$\leqq \left\|\frac{\partial g}{\partial x}(t, y^1)\frac{\partial y^1}{\partial t} - \frac{\partial g}{\partial x}(t, y^2)\frac{\partial y^2}{\partial t}\right\| + \left\|\frac{\partial g}{\partial t}(t, y^1) - \frac{\partial g}{\partial t}(t, y^2)\right\|$$
$$\leqq \left\|\frac{\partial g}{\partial x}(t, y^1) - \frac{\partial g}{\partial x}(t, y^2)\right\| \left\|\frac{\partial y^1}{\partial t}\right\|$$
$$+ \left\|\frac{\partial g}{\partial x}(t, y^2)\right\| \left\|\frac{\partial y^1}{\partial t} - \frac{\partial y^2}{\partial t}\right\| + L\|y^1 - y^2\| \leqq PK$$

where

$$P = 2(\varepsilon+\delta)(L+L^2+L^3)+L$$

So

$$\|H_t^1(t, w) - H_t^2(t, w)\| \leqq \int_0^{|w|}\|h_t^1(t, v) - h_t^2(t, v)\|dv$$
$$\leqq \varepsilon PK \leqq (\varepsilon+\delta)PK \qquad (3.14)$$

and

$$\left\|\int_0^t H_t^1(s, w(s)) - H_t^2(s, w(s))ds\right\| \leqq (\varepsilon+\delta)PK(t+|w|)$$

The estimate of the last term is simlar,

$$\|h_w^1(t, w) - h_w^2(t, w)\| \leqq \left\|\frac{\partial g}{\partial x}(t, y^1) - \frac{\partial g}{\partial x}(t, y^2)\right\| \left\|\frac{\partial y^1}{\partial w}\right\|$$
$$+ \left\|\frac{\partial g}{\partial x}(t, y^2)\right\| \left\|\frac{\partial y^1}{\partial w} - \frac{\partial y^2}{\partial w}\right\| \leqq (\varepsilon+\delta)KL^2 + K/2. \qquad (3.15)$$

So

$$\tfrac{1}{2}\left\|\int_0^t h_w^1(s, w(s)) - h_w^2(s, w(s))ds\right\| \leqq (\tfrac{1}{2}(\varepsilon+\delta)KL^2 + K/4)(t+|w|)$$

Therefore (3.13a) holds if δ, ε are chosen so that

$$|w(t)| \leqq \varepsilon \quad \text{for} \quad 0 \leqq t \leqq \delta \quad \text{and} \quad (\varepsilon+\delta)[2L+P+L^2/2] \leqq 1/4$$

Turning to (3.13b)

$$\left\|\frac{\partial F}{\partial t}(y^1)(t, w) - \frac{\partial F}{\partial t}(y^2)(t, w)\right\|$$

$$\leqq \|f(t, y^1(t, w(t))) - f(t, y^2(t, w(t)))\|$$

$$+ \|H_t^1(t, w) - H_t^2(t, w)\| + \|H_t^1(t, w(t)) - H_t^2(t, w(t))\|$$

$$+ \tfrac{1}{2}\|h_w^1(t, w(t)) - h_w^2(t, w(t))\|$$

The first is bounded by

$$(\varepsilon+\delta)LK.$$

From (3.14) the second and third are bounded by

$$(\varepsilon+\delta)PK$$

and from (3.15) the last is bounded by

$$\tfrac{1}{2}(\varepsilon+\delta)KL^2 + K/4.$$

Hence (3.13b) holds if

$$(\varepsilon+\delta)[L+2P+L^2/2] \leqq 1/4.$$

The last (3.13c) is straightforward,

$$\left\|\frac{\partial F(y^1)}{\partial w}(t, w) - \frac{\partial F(y^2)}{\partial w}(t, w)\right\| = \|h^1(t, w) - h^2(t, w)\|$$

$$= \|g(t, y^1(t, w)) - g(t, y^2(t, w))\| \leqq (\varepsilon+\delta)LK$$

so we must choose

$$(\varepsilon+\delta)L \leqq 1/(4L)$$

for it to hold. Q.E.D.

Proof of theorem We start with uniqueness, suppose $y^i(t,w) \in C^1(\delta,\varepsilon)$ and both satisfy $y^i(t,w) = F(y^i)(t,w)$. We choose K, L, C, ε, $\delta > 0$ such that (3.8), (3.9), and (3.12) hold. By repeated application of Lemma 2, for $0 \leqq t \leqq \delta$, $|w| \leqq \varepsilon$, and for all $k \geqq 1$

$$\|y^1(t,w) - y^2(t,w)\| \leqq K(t + |w|)/2^k$$

hence $y^1(t,w) = y^2(t,w)$.

To show existence we use the Picard iterates

$$x^0(t,w) = x^0$$

$$x^{k+1}(t,w) = F(x^k)(t,w)$$

We choose K, L, C, ε, δ such that (3.8) and (3.9) and (3.12) hold for $x^0(t,w)$ and $x^1(t,w)$. By repeated application of Lemma 2 for $0 \leqq t \leqq \delta$, $|w| \leqq \varepsilon$ and all $k \geqq 1$

$$\|x^{k+1}(t,w) - x^k(t,w)\| \leqq K(t + |w|)/2^k$$

$$\left\|\frac{\partial x^{k+1}}{\partial t}(t,w) - \frac{\partial x^k}{\partial t}(t,w)\right\| \leqq K/2^k$$

$$\left\|\frac{\partial x^{k+1}}{\partial w}(t,w) - \frac{\partial x^k}{\partial t}(t,w)\right\| \leqq K/(2^{k+1}L).$$

Therefore $x^k(t,w)$ converges to some $x(t,w) \in C^1(\varepsilon,\delta)$ given by

$$x(t,w) = x^0 + \sum_{k=0}^{\infty} (x^{k+1}(t,w) - x^k(t,w))$$

since

$$x^{n+1}(t,w) = x^0 + \sum_{k=0}^{n} (x^{k+1}(t,w) - x^k(t,w)).$$

From the above we see that

$$\|x(t,w) - x^{n+1}(t,w)\| \leqq \sum_{k=n+1}^{\infty} \|x^{k+1}(t,w) - x^k(t,w)\|$$

$$\leqq \sum_{k=n+1}^{\infty} K(t + |w|)/2^k \leqq K(t + |w|)/2^n$$

and similarly

$$\left\|\frac{\partial x}{\partial t}(t,w)-\frac{\partial x^{n+1}}{\partial t}(t,w)\right\|\leqq K/2^{n}$$

$$\left\|\frac{\partial x}{\partial w}(t,w)-\frac{\partial x^{n+1}}{\partial w}(t,w)\right\|\leqq K/(2^{n+2}L).$$

Applying lemma 2 we obtain

$$\|F(x)(t,w)-x^{n+2}(t,w)\|\leqq K(t+|w|)/2^{n+1}$$

$$\left\|\frac{\partial F(x)}{\partial t}(t,w)-\frac{\partial x^{n+2}}{\partial t}(t,w)\right\|\leqq K/2^{n+1}$$

$$\left\|\frac{\partial F(x)}{\partial w}(t,w)-\frac{\partial x^{n+2}}{\partial w}(t,w)\right\|\leqq K/(2^{n+2}L).$$

Therefore $x(t,w)$ is a fixed point of F and hence the desired solution. Q.E.D.

4. FORMAL MARTINGALE CALCULUS

In this section we generalize the formal calculus to square integrable martingales with continuous sample paths. We refer the reader to Kunita and Watanabe [10] for the appropriate definitions and for the development of the martingale calculus along the lines of the Ito calculus. In particular we shall use the Kunita–Watanabe differential rule [10, Thm. 2.2] and the fundamental theorem of calculus to define our formal integrals. Since it is very similar to the above, except from proving the analogs of Lemmas 1 and 2, we shall only sketch the details.

Let $M(t)$ be a square integrable martingale with continuous sample paths almost surely and $\langle M\rangle(t)$ be its quadratic variation. The sample paths of $\langle M\rangle(t)$ are continuous and monotone increasing almost surely. For convenience we take $M(0)=\langle M\rangle(0)=0$.

Let $h(t,m)$ be absolutely continuous in t uniformly in m and continuously differentiable in m, i.e., there exists $K>0$ such that

$$\int_a^b\|h_t(t,m)\|dt<K$$

for all m and $h_m(t,m)$ is continuous for all t and m. A particular case of the Kunita—Watanabe rule is

$$dH(t, M(t)) = H_t(t, M(t))dt + H_m(t, M(t))dM(t) + \tfrac{1}{2}H_{mm}(t, M(t))d\langle M\rangle(t) \qquad (4.1)$$

This is a generalization of (2.3) for $W(t)$ is a square integrable sample path continuous martingale with quadratic variation $\langle W\rangle(t) = t$. For simplicity we have not considered different values of λ as in (2.3) although such a development is possible.

The fundamental theorem of calculus and (4.1) motivate the definition of martingale integral

$$\int_a^b h(t, M(t))dt = H(t, M(t))]_a^b - \int_a^b H_t(t, M(t))dt - \tfrac{1}{2}\int_a^b H_{ww}(t, W(t))d\langle M\rangle(t). \qquad (4.2)$$

where $h = H_m$. The integrals on the right are to be interpreted in the Lebesque and Riemann–Stieltjes senses.

Consider the stochastic differential equation

$$dX = f(t, X)dt + g(t, X)dM$$

$$X(0) = X^0. \qquad (4.3)$$

The integral formulation is

$$X(t) = X^0 + \int_0^t f(s, X(s))ds + \int_0^t g(s, X(s))dM(s). \qquad (4.4)$$

We have the following.

THEOREM 2 *Let f, $\partial g/\partial t$ and $\partial g/\partial x$ be locally Lipschitz continuous in x then there exists a random variable $T > 0$ and a random field $X(t,m)$ whose sample functions are absolutely continuous in t uniformly in m and continuously differentiable in m such that $X(t) = X(t, M(t)$ satisfies (in the sense of (4.2)) the integral equation (4.4) almost surely for $0 \leqq t \leqq T$. Moreover the solution is unique in that any two solutions agree almost surely on their common domain of definition.*

As before the theorem is local in character, under suitable hypothesis a global result can be proved.

The random field $X(t,m)$ is constructed independently for each sample point $\omega \in \Omega$. For fixed ω, we let

$$m(t) = M(t, \omega)$$

and

$$\mu(t) = \langle M \rangle (t, \omega)$$

$$x(t) = X(t, m, \omega).$$

The proof is similar to that of Theorem 1, for suitable $y(t, m)$ we define

$$F(y)(t,m) = x^0 + \int_0^t f(s, {}^y(s, m(s)))ds$$
$$+ H(t,m) - \int_0^t H_t(s, m(s))ds - \tfrac{1}{2}\int_0^t h_m(s, m(s))d\mu(s) \qquad (4.5)$$

where

$$h(t,m) = g(t, y(t,m))$$

$$H(t,m) = \int_0^m h(t,v)dv.$$

The desired solution $x(t,m)$ is a fixed point of F. It exists and is unique because F is a local contraction, as shown by the following lemmas.

Let $D^1(\delta, \varepsilon)$ be the space of all n vector valued function $y(t,m)$ which are absolutely continuous with respect to t uniformly in $|m| \leqq \varepsilon$ and continuously differentiable with respect to m for $0 \leqq t \leqq \delta$ and $|m| \leqq \varepsilon$. More precisely $y(t,m) \in D^1(\delta, \varepsilon)$ if there exists $K > 0$

$$\int_0^\delta \left\| \frac{\partial y}{\partial t}(s,m) \right\| ds < K$$

for every $|m| \leqq \varepsilon$ and

$$(t,m) \mapsto \frac{\partial y}{\partial m}(t,m)$$

is continuous for $0 \leqq t \leqq \delta$ and $|m| \leqq \varepsilon$.

LEMMA 3 *Let L, C be chosen so that* (3.8) *holds and $D > L + L^2$. There exist δ and ε sufficiently small so that if $y(t,m) \in D^1(\delta, \varepsilon)$ and*

$$\|y(t,m) - x^0\| \leqq C \qquad (4.6a)$$

$$\int_0^t \left\| \frac{\partial y}{\partial t}(s,m) \right\| ds \leqq D(t + \mu(t)) \qquad (4.6b)$$

$$\left\| \frac{\partial y}{\partial m}(t,m) \right\| \leqq L \qquad (4.6c)$$

for $0 \leqq t \leqq \delta$ and $|m| \leqq \varepsilon$ then $F(y)(t,m) \in D^1(\delta, \varepsilon)$ and satisfies (4.6) *for $0 \leqq t \leqq \delta$ and $|w| \leqq \varepsilon$.*

Proof For any choice of $\varepsilon > 0$ we will always choose $\delta > 0$ so that $|m(t)|$ and $\mu(t) \leqq \varepsilon$ for $0 \leqq t \leqq \delta$.

From the definition (4.5) of F

$$\|F(y)(t,m) - x^0\| \leqq \int_0^t \|f(s, y(s, m(s)))\| ds$$
$$+ \|H(t,m)\| + \int_0^t \|H_t(s, m(s))\| ds + \tfrac{1}{2}\int_0^t \|h_m(s, m(s))\| d\mu(s).$$

Now

$$\int_0^t \|f(s, y(s, m(s)))\| ds \leqq Lt \leqq L\delta \tag{4.7}$$

and

$$\|H(t,m)\| = \left\|\int_0^m g(t, y(t,v)) dv\right\| \leqq L|m| \leqq L\varepsilon.$$

As for the third term, by Fubini's Theorem

$$\int_0^t \|H_t(s, m(s))\| ds = \int_0^t \left\|\int_0^{m(s)} h_t(s,v) dv\right\| ds$$
$$\leqq \int_0^\varepsilon \int_0^t \left\|\frac{\partial g}{\partial t}(s, y(s,v)) + \frac{\partial g}{\partial x}(s, y(s,v)) \frac{\partial y}{\partial t}(s,v)\right\| dv ds$$
$$\leqq \int_0^\varepsilon Lt + L(t + \mu(t)) dv \leqq 2L\varepsilon(t + \mu(t)) \leqq 2L\varepsilon(\delta + \varepsilon) \tag{4.8}$$

The last term is easily bounded

$$\int_0^t \|h_m(s, m(s))\| d\mu(s)$$
$$\leqq \int_0^t \left\|\frac{\partial g}{\partial x}(y(s, m(s))) \frac{\partial y}{\partial m}(s, m(s))\right\| d\mu(s)$$
$$\leqq L^2 \mu(t) \leqq L^2 \varepsilon, \tag{4.9}$$

and so if δ and ε are sufficiently small (4.6a) holds.

Turning to (4.6b)

$$\int_0^t \left\|\frac{\partial F(y)}{\partial t}(s,m)\right\| ds \leqq \int_0^t \|f(s, y(s, m(s)))\| ds$$
$$+ \int_0^t \|H_t(s,m)\| ds + \int_0^t \|H_t(s, m(s))\| ds$$
$$+ \tfrac{1}{2}\int_0^t \|h_m(s, m(s))\| d\mu(s).$$

The first, third and fourth terms are estimated in (4.7), (4.8), and (4.9); the second term can be estimated similar to (4.8). Therefore

$$\int_0^t \left\|\frac{\partial F(y)}{\partial t}(s,m)\right\| ds \leqq Lt + 4L\varepsilon(t+\mu(t)) + L^2\mu(t)$$

which is less than $P(t+\mu(t))$ if ε is sufficiently small.

The last estimate (4.6c) is straightforward,

$$\left\|\frac{\partial F(y)}{\partial m}(t,m)\right\| = \|H_m(t,m)\| = \|h(t,m)\| = \|g(t,y(t,m))\| \leqq L$$

LEMMA 4 *Let L, C be chosen so that (3.8) holds and $D > L + L^2$. There exist δ and ε sufficiently small so that if $y^i(t,m) \in D^1(\delta,\varepsilon)$ satisfies (4.6) and*

$$\|y^1(t,m) - y^2(t,m)\| \leqq K(t+|m|-\mu(t)) \tag{4.10a}$$

$$\int_0^t \left\|\frac{\partial y^1}{\partial t}(s,m) - \frac{\partial y^2}{\partial t}(s,m)\right\| ds \leqq K(t+\mu(t)) \tag{4.10b}$$

$$\left\|\frac{\partial y^1}{\partial m}(t,m) - \frac{\partial y^2}{\partial m}(t,m)\right\| \leqq K/2L \tag{4.10c}$$

for $0 \leqq t \leqq \delta$ and $|m| \leqq \varepsilon$ then $F(y^i)(t,m) \in D^1(\delta,\varepsilon)$, satisfies (4.6) and

$$\|F(y^1)(t,m) - F(y^2)(t,m) \leqq K(t+|m|+\mu(t))/2 \tag{4.11a}$$

$$\int_0^t \left\|\frac{\partial F(y^1)}{\partial t}(s,m) - \frac{\partial F(y^2)}{\partial t}(s,m)\right\| ds \leqq K(t+\mu(t))/2 \tag{4.11b}$$

$$\left\|\frac{\partial F(y^1)}{\partial m}(s,m) - \frac{\partial F(y^2)}{\partial m}(s,m)\right\| \leqq K/(4L) \tag{4.11c}$$

for $0 \leqq t \leqq \delta$ and $|m| \leqq \varepsilon$.

Proof The first two assertions follow from Lemma 3, as for the last,

$$\begin{aligned}\|F(y^1)(t,m) - F(y^2)(t,m)\| &\\ \leqq \textstyle\int_0^t \|f(s,y^1(s,m(s))) - f(s,y^2(s,m(s)))\| ds&\\ + \|H^1(t,m) - H^2(t,m)\|&\\ + \textstyle\int_0^t \|H_t^1(s,m(s)) - H_t^2(s,m(s))\| ds&\\ + \tfrac{1}{2}\textstyle\int_0^t \|h_m^1(s,m(s)) - h_m^2(s,m(s))\| d\mu(s)&\end{aligned}$$

where

$$h^i(t,m) = g(t, y^i(t,m))$$
$$H^i(t,m) = \int_0^m h^i(t,v)dv$$

By Lipschitz continuity

$$\int_0^t \|f(s, y^1(s, m(s))) - f(s, y^2(s, m(s)))\| ds$$
$$\leqq L\int_0^t \|y^1(s, m(s)) - y^2(s, m(s))\| ds$$
$$\leqq KL\int_0^t \delta + 2\varepsilon\, dt \leqq KL(\delta + 2\varepsilon)(t + |m| + \mu(t)) \qquad (4.12)$$

Next

$$\|H^1(t,m) - H^2(t,m)\| \leqq \int_0^{|m|} \|h^1(t,v) - h^2(t,v)\| dv$$
$$\leqq \int_0^{|m|} L\|y^1(t,v) - y^2(t,v)\| dv$$
$$\leqq KL\int_0^{|m|} \delta + 2\varepsilon\, dv \leqq KL(\delta + 2\varepsilon)(t + |m| + \mu(t))$$

as for the third

$$\int_0^t \|H_t^1(s, m(s)) - H_t^2(s, m(s))\| ds$$
$$= \int_0^t \left\| \int_0^{m(s)} h_t^1(s,v) - h_t^2(s,v) dv \right\| ds$$
$$\leqq \int_0^\varepsilon \int_0^t \|h_t^1(s,v) - h_t^2(s,v)\| ds dv$$

Now

$$\|h_t^1(t,m) - h_t^2(t,m)\|$$
$$\leqq \left\| \frac{\partial g}{\partial t}(t, y^1(t,m)) - \frac{\partial g}{\partial t}(t, y^2(t,m)) \right\|$$
$$+ \left\| \frac{\partial g}{\partial x}(t, y^1(t,m)) \right\| \left\| \frac{\partial y^1}{\partial t} - \frac{\partial y^2}{\partial t} \right.$$
$$+ \left\| \frac{\partial y^2}{\partial t} \right\| \left\| \frac{\partial g}{\partial x}(t, y^1(t,m)) - \frac{\partial g}{\partial x}(t, y^2(t,m)) \right\|$$
$$\leqq L\|y^1 - y^2\| + L\left\| \frac{\partial y^1}{\partial t} - \frac{\partial y^2}{\partial t} \right\|$$
$$+ L\left\| \frac{\partial y^2}{\partial t} \right\| \|y^1 - y^2\|$$
$$\leqq L(\delta + 2\varepsilon) + L\left\| \frac{\partial y^1}{\partial t} - \frac{\partial y^2}{\partial t} \right\| + L(\delta + 2\varepsilon)\left\| \frac{\partial y^2}{\partial t} \right\|$$

So

$$\begin{aligned}
\textstyle\int_0^t \|H_t^1(s,m(s)) - H_t^2(s,m(s))\|ds \\
&\leqq KL\varepsilon(\delta+2\varepsilon)t + L\textstyle\int_0^\varepsilon K(t+\mu(t))dv \\
&\quad + KL(\delta+2\varepsilon)\int_0^\varepsilon\int_0^t \left\|\frac{\partial y^2}{\partial t}(s,v)\right\|dsdv \\
&\leqq KL\varepsilon(\delta+2\varepsilon)t + KL\varepsilon(t+\mu(t)) \\
&\quad + KLD\varepsilon(\delta+2\delta)(t+\mu(t)) \qquad (4.13)
\end{aligned}$$

Finally

$$\begin{aligned}
\textstyle\int_0^t \|h_m^1(s,m(s)) - h_m^2(s,m(s))\|d\mu(s) \\
&\leqq \int_0^t \left\|\frac{\partial g}{\partial x}(s,y^1(s,m(s)))\right\| \left\|\frac{\partial y^1}{\partial m}(s,m(s)) - \frac{\partial y^2}{\partial m}(s,m(s))\right\| \\
&\quad + \left\|\frac{\partial y^2}{\partial m}(s,m(s))\right\| \left\|\frac{\partial g}{\partial x}(s,y^1(s,m(s))) - \frac{\partial g}{\partial x}(s,y^2(s,m(s)))\right\|d\mu(s) \\
&\leqq \textstyle\int_0^t K/2 + L^2\|y^1(s,m(s)) - y^2(s,m(s))\|d\mu(s) \\
&\leqq (K/2 + KL^2(\delta+2\varepsilon))\mu(t). \qquad (4.14)
\end{aligned}$$

Hence (4.11a) holds if δ and ε are chosen small enough so that

$$(2L + LD\varepsilon + L^2/2)(\delta+2\varepsilon) \leqq 1/4.$$

As for (4.11b)

$$\begin{aligned}
\int_0^t \left\|\frac{\partial F(y^1)}{\partial t}(s,m(s)) - \frac{\partial F(y^2)}{\partial t}(s,m(s))\right\|ds \\
&\leqq \textstyle\int_0^t \|f(s,y^1(s,m(s))) - f(s,y^2(s,m(s)))\|ds \\
&\quad + \textstyle\int_0^t \|H_t^1(s,m) - H_t^2(s,m)\|ds + \int_0^t \|H_t^1(s,m(s)) - H_t^2(s,m(s))\|ds \\
&\quad + \tfrac{1}{2}\textstyle\int_0^t \|h_m^1(s,m(s)) - h_m^2(s,m(s))\|d\mu(s)
\end{aligned}$$

The first, third and fourth terms are estimated above (4.12), (4.13), and (4.14) and the second can be estimated like (4.13). Hence

$$\begin{aligned}
\int_0^t \left\|\frac{F(y^1)}{\partial t}(s,m(s)) - \frac{\partial F(y^2)}{\partial t}(s,m(s))\right\|ds \\
&\leqq (K/4 + (KL + 2KLD\varepsilon + KL^2/2)(\delta+2\varepsilon))(t+\mu(t)),
\end{aligned}$$

so (4.11b) holds if

$$(L+2LD\varepsilon+L^2/2)(\delta+2\varepsilon)\leqq 1/4$$

Finally

$$\begin{aligned}\left\|\frac{\partial F(y^1)}{\partial m}(t,m)-\frac{\partial F(y^2)}{\partial m}(t,m)\right\| &= \left\|h^1(t,m)-h^2(t,m)\right\| \\ &= \left\|g(t,y^1(t,m))-g(t,y^2(t,m))\right\|\leqq KL(\delta+2\varepsilon)\end{aligned}$$

so for small δ and ε, (4.11c) follows. Q.E.D.

References

[1] I. N. Gelfand and N. Ya. Valenkin, *Generalized Functions*, Vol. 4, Academic Press, N.Y., 1964.
[2] K. Ito, The stochastic integral, *Proc. Imperial Acad.* Tokyo 20 (1944) pp. 519–524.
[3] J. Lamperti, A simple construction of certain diffusion processes, *J. Math. Kyoto* **4** (1964) pp. 161–170.
[4] R. E. A. C. Paley, N. Wiener and A. Zygmund, A note on random functions, *Math. Z.* **37** (1959), pp. 647–668.
[5] R. L. Stratonovich, *Conditional Markov Processes and their Application to the Theory of Optimal Control*, American Elsevier, N.Y., 1968.
[6] H. J. Sussmann, An interpretation of stochastic differential equations as ordinary differential equations which depend on the sample point, *Bul., A.M.S.* **83** (1977) pp. 296–298.
[7] M. Turelli, Random environments and stochastic calculus, *Theoretical Population Biology* **12** (1977) pp. 140–178.
[8] H. P. McKean, *Stochastic Integrals*, Academic Press, N.Y. 1969.
[9] E. Wong and M. Zakai, On the convergence of ordinary integrals to stochastic integrals, *Annals Math. Stat.* **36** (1965) pp. 1560–1564.
[10] H. Kunita and S. Watanabe, On square-integrable martingales, *Nagoya Math. J.* **30** (1967) pp. 209–245.
[11] H. Doss, Liens entre équations différentielles stochastiques et ordinaires, *Ann. Inst. Henri Poincaré* **13** (1977) pp. 99–125.
[12] M. H. A. Davis, Martingale integrals and stochastic calculus, in *Communication Systems and Random Process Theory*, J. K. Skwirzynski (Ed.), Sisthoff and Nordhoff, Holland, 1978.