Stratonovich Stochastic Differential Equations Driven by General Semimartingales

by

Thomas G. Kurtz[1] and Etienne Pardoux
University of Wisconsin Université de Provence

Philip Protter[2]
Purdue University

Technical Report #91-22

Department of Statistics
Purdue University

March 1992

[1] Supported in part by NSF grant #DMS–8901464
[2] Supported in part by NSF grant #DMS–9103454

◇

STRATONOVICH STOCHASTIC DIFFERENTIAL EQUATIONS DRIVEN BY GENERAL SEMIMARTINGALES

by

Thomas G. Kurtz[1]
Department of Mathematics
and Statistics
University of Wisconsin
Madison, Wisconsin 53706, USA

and

Etienne Pardoux
Mathématiques, URA 225
Université de Provence
13331 Marseille Cedex 3
France

Philip Protter[2]
Department of Mathematics
and Statistics
Purdue University
West Lafayette, IN 47907-1395, USA

Abstract

We investigate stochastic differential equations driven by semimartingales with jumps. These are interpreted as Stratonovich type equations, with the "integrals" being of the kind introduced by S. Marcus, rather than the more well known type proposed by P. A. Meyer. We establish existence and uniqueness of solutions; we show the flows are diffeomorphisms when the coefficients are smooth (not the case for Meyer-Stratonovich differentials); we establish strong Markov properties; and we prove a "Wong-Zakai" type weak convergence result when the approximating differentials are smooth and continuous even though the limits are discontinuous.

Reśumé

On considère des equations stochastiques différentielles où le "bruit" est une semimartingale quelquonque (avee des sauts). On propose une interpretation des intégrales stochastiques du type "Stratonovich", mais du genre de celles introduites par S. Marcus, plutôt que du genre de celles de P.A. Meyer. On établit l'existence et l'unicité des solutions et on démontre que les flots sont des difféomorphismes quand les coefficients sont convenables (ce qui n'est pas le cas pour l'interpretation de Meyer–Stratonovich). De plus on établit les propriétés de Markov fortes, et on démontre un genre de convergence faible du

[1] Supported in part by NSF grant #DMS–8901464
[2] Supported in part by NSF grant #DMS–9103454

type "Wong–Zakai" quand les approximants sont réguliers et continus, même si les limites ne sont pas continues.

§1. Introduction

We investigate here a stochastic differential equation of "Stratonovich type", where the differential semimartingales Z can have jumps. We write the equation with the customary "circle" notation to indicate that it is not a standard Itô type semimartingale integral:

$$X_t = X_0 + \int_0^t f(X_s) \circ dZ_s. \tag{1.1}$$

The "integral" in the equation is a new type of Stratonovich stochastic integral with respect to a semimartingale Z with jumps. (Our integral is different from the one given by Meyer [13] or Protter [14].) Unfortunately we have been able to define our new integral only for integrands that are solutions of stochastic integral equations, and not for arbitrary integrands.

The equation (1.1) above is given the following meaning, for the case of scalar processes X, Z:

$$\begin{aligned} X_t = X_0 &+ \int_0^t f(X_{s-})dZ_s \\ &+ \frac{1}{2}\int_0^t f'f(X_s)d[Z,Z]_s^c \\ &\sum_{0<s\le t} \{\varphi(f\Delta Z_s, X_{s-}) - X_{s-} - f(X_{s-})\Delta Z_s\} \end{aligned} \tag{1.2}$$

where $\varphi(g,x)$ denotes the value at time $u=1$ of the solution of the following ordinary differential equation:

$$\frac{dy}{du}(u) = g(y(u)); \quad y(0) = x.$$

We also write $\varphi(g,x,u)$ to denote the solution at time u; thus $\varphi(g,x) = \varphi(g,x,1)$.

The first term on the right side of equation (1.2) is the standard Itô-semimartingale stochastic integral with respect to the semimartingale Z; the second term is a (semimartingale or) Stieltjes integral with respect to the increasing process $[Z,Z]^c$, where $[Z,Z]$ denotes the quadratic variation process of Z and $[Z,Z]^c$ denotes its path by path continuous part (see Protter [14, p. 62]). The third term is a (possibly countable) sum of terms of order $(\Delta Z_s)^2$ and therefore converges absolutely (see Section Two). Were we to have interpreted (1.1) as a Stratonovich equation in the sense of the semimartingale Stratonovich integral as defined by Meyer [13] (see also Protter [14]), the right side of (1.2) would have contained the first two terms only.

The inclusion of the third term on the right side of (1.2) has several beneficial consequences. The first (as we show in Section Six) is that the solution to (1.1) is the weak limit of the solutions to approximate equations where the driving semimartingales are *continuous* piecewise approximations of the driving semimartingale Z (a "Wong-Zakai" type of result). The second is that the solution remains on a manifold M whenever it starts there and the coefficients of the equation are vector fields over M. (This is proved in Section Four.) The third (see Section Three) is that the flows of the solution are diffeomorphisms when the coefficients are smooth. This last property does not hold in general for semimartingale nor Stratonovich-semimartingale stochastic differential equations, because (for example) the injectivity fails (see Protter [14, Chapter V, §10]).

We feel that the first consequence mentioned above, that of the "Wong-Zakai" property, is important from a modelling viewpoint, since a jump in the differential can be regarded as a mathematical idealization for a very rapid continuous change.

The idea to interpret equation (1.1) by (1.2) is not new. It was introduced by S. Marcus [10], [11] in the case where Z has finitely many jumps on compact time intervals. The corresponding "Wong-Zakai" results were investigated by Kushner [9]. Recently Estrade [3] has studied equations similar to (1.1) and (1.2) on Lie groups, and Cohen [2] has given an intrinsic language for stochastic differential equations on manifolds, which relates to section four of this article.

In this paper we prove existence and uniqueness of a solution of (1.2), we show the associated flow is a diffeomorphism of $\mathbb{R}^d$ in the vector case, we show the solution is a strong Markov process when the driving semimartingales Z are Lévy processes, and of course we establish "Wong-Zakai" type approximation results for weak convergence.

One notation caveat: the i^{th} component of a vector x will be denoted x^i; the j^{th} column vector of a matrix f will be denoted f_j, and hence f^i_j stands for the (i,j) term of the matrix f. Finally, when the meaning is clear, we use the convention of implicit summing over indices (that is we write a_i to denote $\sum_{i=1}^{d} a_i$).

§2. Discussion of the Equation

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a probability space equipped with a filtration $\{\mathcal{F}_t; t \geq 0\}$ of sub-σ-fields of $\mathcal{F}$. We assume the filtration satisfies the "usual hypotheses", i.e. it is right-continuous, and $\mathcal{F}_0$ contains all P-zero measure sets of $\mathcal{F}_\infty = \mathcal{F}$.

A process Z which has right continuous paths with left limits a.s. (known as *"càdlàg"*, after the French acronym) is called a *semimartingale* if it has a decomposition $Z = M + A$, where M is a càdlàg local martingale and A is an adapted, càdlàg process, whose paths are a.s. of finite variation on compacts. For all details of semimartingales the reader is referred to, for example, Protter [13]. A k-dimensional semimartingale $Z = \{Z_t; t \geq 0\}$ is assumed given with $Z_0 = 0$. Let $f \in C^1(\mathbb{R}^d; \mathbb{R}^{d\times k})$. Given an $\mathcal{F}_0$ measurable d-dimensional random vector X_0, we want to study an equation, which we write symbolically as:

$$X_t = X_0 + \int_0^t f(X_s) \circ dZ_s \tag{2.1}$$

and is to be understood as

$$\begin{cases} X_t = X_0 + \int_0^t f(X_{s-})dZ_s + \frac{1}{2}\int_0^t f'f(X_s)d[Z,Z]_s^c \\ \qquad + \sum_{0<s\leq t} \{\varphi(f\Delta Z_s, X_{s-}) - X_{s-} - f(X_{s-})\Delta Z_s\} \end{cases} \tag{2.2}$$

Let us explain the meaning of the three last terms on the right of (2.2).

$$\int_0^t f(X_{s-})dZ_s = \int_0^t f_j(X_{s-})dZ_s^j,$$

where the sum runs from $j = 1$ to $j = k$ (we use throughout the convention of summation of repeated indices), is the "Itô integral" of the predictable process $\{f(X_{t-})\}$ with respect to the semimartingale Z.

$$\int_0^t f'f(X_s)d[Z,Z]_s^c = \int_0^t \frac{\partial f_j}{\partial x^\ell}(X_s) f_m^\ell(X_s) d[Z^j, Z^m]_s^c \tag{2.3}$$

is a Stieltjes integral with respect to the continuous bounded variation processes $[Z^j, Z^m]^c$ which are the continuous parts of the quadratic covariation process (cf. Protter [14, p. 58]).

Let us finally define the notation $\varphi(f\Delta Z_s, x)$. Given $g \in C^1(\mathbf{R}^d;\mathbf{R}^d)$ and $x \in \mathbf{R}^d$, the following equation:

$$\begin{cases} \dfrac{dy}{du}(u) = g(y(u)) \\ \quad y(0) = x \end{cases}$$

has a unique maximal solution $\{\varphi(g,x,u); 0 \le u < \xi\}$ and

$$\overline{\lim_{u\uparrow\xi}}|\varphi(g,x,u)| = +\infty \text{ if } \xi < \infty$$

If $\xi > 1$,

$$\varphi(g,x) = \varphi(g,x,1)$$

If $\xi \le 1$, $\varphi(g,x)$ is undefined: the solution of (2.1) explodes at the corresponding jump time of Z. We shall be mainly concerned with the case where f is globally Lipschitz, in which case $\varphi(f\Delta Z_s, X_{s-})$ is always defined as a d-dimensional $\mathcal{F}_s$ measurable random vector (given that X_{s-} is $\mathcal{F}_s$-measurable).

For equation (2.2) to make sense we must show that the sum on the right side is absolutely convergent. This follows from Taylor's theorem: Since $u \to \varphi(f\Delta Z_s, x, u)$ is C^2, we have:

$$\begin{aligned}\varphi(f\Delta Z_s, x, 1) = x &+ f(x)\Delta Z_s \\ &+ \frac{1}{2} f'f(\varphi(f\Delta Z_s, x, \theta))\Delta Z_s \Delta Z_s^t\end{aligned}$$

for $\theta \in (0,1)$ which depends on (s,ω,x). Note that the notation used above is defined in equation (2.3). Thus

$$\begin{aligned}&\sum_{0<s\le t} |\varphi(f\Delta Z_s, X_{s-}) - X_{s-} - f(X_{s-})\Delta Z_s| \\ &\le \frac{1}{2} \sup_{\substack{s\le t \\ 0\le\theta\le 1}} |f'f(\varphi(f\Delta Z_s, X_{s-}, \theta)|(\sum_{0<s\le t} |\Delta Z_s|^2) \\ &\le K \sum_{0<s\le t} |\Delta Z_s|^2,\end{aligned}$$

which is a.s. finite since $K(\omega) < \infty$ and the sum of squares of the jumps of a semimartingale is always finite a.s.

The next observation allows us to use many of the results of the well developed theory of stochastic differential equations, and it has greatly simplied a previous version of this paper. Let $[Z,Z]^d$ be the jump component of the quadratic variation process $[Z,Z]$, analogous to $[Z,Z]^c$ being the continuous part. That is,

$$[Z,Z]^d_t = \sum_{0<s\leq t} |\Delta Z_s|^2$$

where Z is a given vector of semimartingales.

Again for a given vector of semimartingales Z, we define

$$h(s,\omega,x) = \frac{\varphi(\Delta Z_s f, x) - x - f(x)\Delta Z_s}{|\Delta Z_s|^2}.$$

where the ω comes from the terms $\Delta Z_s = \Delta Z_s(\omega)$. We have the following obvious result:

Lemma 2.1 For f and $f'f$ well defined and Lipschitz continuous, a solution X of equation (2.1), interpreted as a solution of

$$\begin{aligned} X_t = X_0 + &\int_0^t f(X_{s-})dZ_s + \\ &\frac{1}{2}\int_0^t f'f(X_s)d[Z,Z]^c_s \\ + &\sum_{0<s\leq t} \{\varphi(f\Delta Z_s, X_{s-}) - X_{s-} - f(X_{s-})\Delta Z_s\}, \end{aligned} \tag{2.4}$$

is also a solution of

$$\begin{aligned} X_t = X_0 + &\int_0^t f(X_{s-})dZ_s + \frac{1}{2}\int_0^t f'f(X_{s-})d[Z,Z]^c_s \\ + &\int_0^t h(s,\cdot,X_{s-})d[Z,Z]^d_s, \end{aligned} \tag{2.5}$$

and conversely.

§3. Existence, Uniqueness and Flows of the Equation

One can study the equation in question

$$X_t = X_0 + \int_0^t f(X_s) \circ dZ_s, \tag{3.1}$$

directly (as the authors did during their preliminary efforts), but it is much more efficient to realize that equation (3.1) can be rewritten as

$$X_t = X_0 + \int_0^t f(X_{s-})dZ_s + \frac{1}{2}\int_0^t f'f(X_s)d[Z,Z]_s^c + \int_0^t h(s,w,X_{s-})d[Z,Z]_s^d, \tag{3.2}$$

where $h(s,\omega,x) = \frac{\varphi(\Delta Z_s f,x)-x-f(x)\Delta Z_s}{|\Delta Z_s|^2}$, and $[Z,Z]_s^d = \sum_{0<s\leq t} \Delta Z_s^t \Delta Z_s$, the sum of the squares of the jumps.

We will call an operator F on processes *process Lipschitz* as defined in Protter [14, p. 195]) if (i) whenever $X^{T-} = Y^{T-}$, then $F(X)^{T-} = F(Y)^{T-}$ for any stopping time T; and (ii) $|F(X)_t - F(Y)_t| \leq K_t|X_t - Y_t|$, for an adapted process K.

Lemma 3.1 For f and $f'f$ Lipschitz continuous, the function $h(s,\omega,x)$ is process Lipschitz. If Z has bounded jumps, then h is random Lipschitz with a bounded Lipschitz constant.

Proof: To show the result we apply Taylor's theorem to the mapping

$$u \to \varphi(\Delta Z_s f, x, u) - \varphi(\Delta Z_s f, y, u):$$
$$\begin{aligned} &|\varphi(\Delta Z_s f, x, u) - \varphi(\Delta Z_s f, y, u) - x - y - \Delta Z_s(f(x) - f(y))| \\ &\leq \frac{1}{2}|\{f'f(\varphi(\Delta Z_s f, x, \theta)) - f'f(\varphi(\Delta Z_s f, y, \theta))\}\Delta Z_s \Delta Z_s^t| \\ &\leq c|\varphi(\Delta Z_s f, x, \theta) - \varphi(\Delta Z_s f, y, \theta)||\Delta Z_s|^2 \\ &\leq c|x-y|e^{c|\Delta Z_s|}|\Delta Z_s|^2, \end{aligned}$$

where the last inequality follows from Gronwall's lemma. This implies

$$|h(s,w,x) - h(s,w,y)| \leq c|x-y|e^{c|\Delta Z_s|},$$

and the result follows. □

Lemma 3.1 allows us to use the already well developed theory of stochastic differential equations as found in Chapter V of Protter [14].

Theorem 3.2 Let f and $f'f$ be globally Lipschitz. Then there exists a càdlàg solution to (3.1), it is unique, and it is a semimartingale.

Proof: Rewriting equation (3.1) in its equivalent form (3.2), we observe that (3.2) is a standard stochastic differential equation with semimartingale differentials Z, $[Z,Z]^c$, and $[Z,Z]^j$, and process Lipschitz coefficients. There is one technical problem: the coefficient $h(s,w,x)$ is not predictable for each fixed x, and does not map càglàd (left continuous with right limits), adapted processes to itself. However the process $[Z,Z]^j_s$ is an increasing, finite variation process, and since h is optionally measurable for each fixed x, this does not pose a problem. Thus we need only to apply a trivial extension of (for example) Theorem V.7 of Protter [14, p. 197] to deduce the result. □

We can weaken the globally Lipschitz hypotheses of Theorem 3.1 to locally Lipschitz, by standard techniques (see e.g., Métivier [12, Theorem 34.7, p. 246] or Protter [14, pp. 247–249]). We will call a function g *locally Lipschitz* if for any n there exists a constant c_n such that for all x, $y \in \mathbb{R}^d$ with $||x|| \leq n$, $||y|| \leq n$, $||g(x) - g(y)|| \leq c_n||x-y||$.

Corollary 3.3 Let f and $f'f$ be locally Lipschitz continuous. Then there exists a stopping time T, called the explosion time, and a càdlàg, adapted d-dimensional process $\{X_t, 0 \leq t < T\}$ that is the unique solution of equation (3.1). Moreover $\limsup_{t\to T} ||X_t|| = \infty$ a.s. on the event $\{T < \infty\}$.

Remark: A more general equation than (3.1) is the following

$$
\begin{aligned}
X_t = J_t &+ \int_0^t f(X_{s-})dZ_s + \frac{1}{2}\int_0^t f'f(X_s)d[Z,Z]^c_s \\
&+ \frac{1}{2}\int_0^t f'(X_s)d[J,Z]^c_s \\
&+ \sum_{0<s\leq t;\Delta Z_s\neq 0} \{\varphi(\tilde{f}\Delta\tilde{Z}_s, X_{s-}) - X_{s-} - \tilde{f}(X_{s-})\Delta\tilde{Z}_s\}
\end{aligned}
\tag{3.3}
$$

where J is a càdlàg, adapted process such that $[J,Z]^c$ exists (in the sense defined in Protter [14, p. 215]), and moreover $\sum_{0<s\leq t,\Delta Z_s\neq 0} |\Delta J_s|^2 < \infty$, each $t \geq 0$. Also, $\tilde{f}(x) \in \mathbb{R}^{d\times(k+d)}$ is defined as

$$\tilde{f}(x) = [f(x) \vdots I] \text{ and } \tilde{Z}_t = \begin{pmatrix} Z_t \\ J_t \end{pmatrix}.$$

We can prove an existence and uniqueness result for equation (3.3). Note that if J is a semimartingale, then equation (3.3) can be put into the form of equation (3.1). We shall restrict ourselves to the case where J is a semimartingale in this paper.

Letting the initial condition be $x \in \mathbf{R}^d$, we can write $X(t, \omega, x)$ for the solution

$$X_t = x + \int_0^t f(X_s) \circ dZ_s. \tag{3.4}$$

The *flow* of the stochastic differential equation (3.4) is the function $x \longrightarrow X(t, \omega, x)$, which can be considered as a mapping from $\mathbf{R}^d \to \mathbf{R}^d$ for (t, ω) fixed, or as a mapping from $\mathbf{R}^d \to \mathcal{D}^d$, where $\mathcal{D}^d$ denotes the space of càdlàg functions from $\mathbf{R}_+$ to $\mathbf{R}^d$, equipped with the topology of uniform convergence on compacts, for ω fixed.

Theorem 3.4 Let f and $f'f$ be globally Lipschitz. Then the flow $x \to X(\cdot, \omega, x)$ from $\mathbf{R}^d$ to $\mathcal{D}^d$ is continuous in the topology of uniform convergence on compacts.

Proof: We can express equation (3.4) in the equivalent form (3.2). Since f and $f'f$ are globally Lipschitz and h is process Lipschitz, Theorem 3.4 is a special case of Theorem V.37 in Protter [14, p. 246]. □

We henceforth consider the flow of equation (3.4) as a function from $\mathbf{R}^d$ to $\mathbf{R}^d$, for each fixed (t, ω). Let Ψ denote the flow: that is, $\Psi : \mathbf{R}^d \to \mathbf{R}^d$ is given by $\Psi_t(x) = X(x, t, \omega)$ for fixed (t, ω), where X is the solution of equation (3.4).

For a semimartingale Z with $Z_0 = 0$, let $Z = N + A$ be a decomposition into a local martingale N and an adapted, càdlàg process A with paths of finite variation on compacts, and $N_0 = A_0 = 0$. For $1 \leq p \leq \infty$ define

$$j_p(N, A) = ||[N, N]_\infty^{1/2} + \int_0^\infty |dA_s|||_{L^p},$$

where $|| \cdot ||_{L^p}$ denotes the L^p norm with respect to the underlying probability measure P, and $\int_0^\infty |dA_s|$ denotes the total variation of the paths of A, ω by ω. Next define

$$||Z||_{\mathcal{H}^p} = \inf_{Z=N+A} j_p(N, A),$$

where the infimum is taken over all decompositions $Z = N + A$. We will be especially interested in the $\mathcal{H}^\infty$ norm. Note that if $||Z||_{\mathcal{H}^\infty} \le \varepsilon$, then the jumps of each component of Z are bounded by ε.

For a given $\varepsilon > 0$, and $Z = (Z^1, \ldots, Z^m)$, we can find stopping times $0 = T_0 < T_1 < T_2 < \ldots$ tending a.s. to ∞ such that

$$Z^{\alpha,j} = (Z^\alpha)^{T_j -} - (Z^\alpha)^{T_{j-1}}$$

has an $\mathcal{H}^\infty$ norm less than ε, $1 \le \alpha \le m$. (See Theorem V.5, p. 192 of Protter [14].) The above observation allows us to first consider semimartingale differentials with small $\mathcal{H}^\infty$ norm.

Let $X_t^j(x)$ denote the solution of

$$X_t = x + \int_0^t f(X_s) \circ dZ_s \tag{3.1}$$

where the equation has driving semimartingales $Z^{\alpha,j} = (Z^\alpha)^{T_j -} - (Z^\alpha)^{T_{j-1}}$. Outside of the interval (T_{j-1}, T_j) the solution is:

$$X_t^j(x) = \begin{cases} x & \text{for } t < T_{j-1} \\ X_{T_j -}^j & \text{for } t \ge T_j \end{cases}$$

We next define the *linkage operators* H^j : let $y = y(t,x)$ be the solution of

$$\frac{dy}{dt} = f(y(t))\Delta Z_{T_j};\ y(0) = x$$

and define $H^j(x) = y(1,x)$.

The next lemma is classical:

Lemma 3.5 Let f be $\mathcal{C}^\infty$ with all derivatives bounded. Then H^j is a.s. a $\mathcal{C}^\infty$ diffeomorphism of $\mathbf{R}^d$.

Next we have the obvious result:

Theorem 3.6 The solution X of

(3.1) $$X_t = x + \int_0^t f(X_s) \circ dZ_s$$

is equal to, for $T_j \leq t < T_{j+1}$:

$$X_t(x) = X_t^{j+1}(x_{j+}),$$

where

$$x_{0+} = x$$
$$x_{1-} = X^1_{T_1-}(x), x_{1+} = H^1(x_1-)$$
$$\vdots$$
$$x_{j-} = X^j_{T_j-}(x_{(j-1)+}); x_{j+} = H^j(x_{j-}).$$

Theorem 3.7 Let f be $\mathcal{C}^\infty$ with all derivatives bounded. The flow $\Psi: \ x \longrightarrow X_t(x,\omega)$ of the solution X of

(3.1) $$X_t = x + \int_0^t f(X_s) \circ dZ_s$$

is a diffeomorphism if the collection

$$x \longrightarrow X_t^j(x,\omega)$$

are diffeomorphisms.

Proof: By Theorem 3.6, the solution X of (3.1) can be constructed by composition of the functions X^j and the linkage operators H^j. But the linkage operators are diffeomorphisms by Lemma 3.5, and since the composition of diffeomorphisms is a diffeomorphism, the theorem is proved. □

To show the functions $x \longrightarrow X_t^j(x,\omega)$ are diffeomorphisms we are able to use the results of Section 10 of Chapter V of Protter [14].

Theorem 3.8 Let f, $f'f$ be $\mathcal{C}^\infty$ with all derivatives bounded, and $Z = (Z^1, \ldots, Z^m)$ be semimartingales, $Z_0 = 0$, and let X be the solution of

(3.1) $$X_t = x + \int_0^t f(X_s) \circ dZ_s.$$

The flow Ψ is a diffeomorphism of $\mathbb{R}^d$ if $||Z||_{\mathcal{H}^\infty} < \varepsilon$ for $\varepsilon > 0$ sufficiently small.

Proof: We rewrite the equation (3.1) in the form (3.2). Equation (3.2) is in the classical form with process Lipschitz, smooth coefficients. We then invoke Hadamard's theorem (Theorem 59, p. 275), along with Theorem 62 (p. 279) and Theorem 64 (p. 281) of Protter [14] to deduce the result. □

Combining Theorems 3.7 and 3.8, we have:

Theorem 3.9 Let f be $\mathcal{C}^\infty$ with all derivatives of f and $f'f$ bounded. Let $Z = (Z^1, \ldots, Z^m)$ be a vector of semimartingales. Then the flow $\Psi : x \to X_t(x, \omega)$ of the solution X of

$$X_t = x + \int_0^t f(X_s) \circ dZ_s$$

is a diffeomorphism of $\mathbb{R}^d$.

§4. A Change of Variable Formula and Manifold-Valued solutions

One could argue that even the Stratonovich integral for Brownian motion should not be called an "integral", since it does not satisfy a minimally acceptable "dominated convergence theorem", as does — for example — the semimartingale "Itô-type" integral. However our "integral" is even less of an integral than the Meyer-Stratonovich integral, since it is only defined for integrands which are solutions of stochastic differential equations.

Nevertheless there are circumstances under which we can establish a change of variables formula. Let X denote a solution of our equation:

$$X_t = X_0 + \int_0^t f(X_s) \circ dZ_s. \tag{4.1}$$

We will establish for $g \in \mathcal{C}^1(\mathbb{R}^d; \mathbb{R}^k)$ that we can define an integral

$$\int_0^t g(X_s) \circ dZ_s = \int_0^t g_i(X_s) \circ dZ_s^i$$

for $t \geq 0$, which we call the Stratonovich integral of $g(X)$ with respect to Z. (Note that this definition is *not* consistent with that of Meyer [13] and Protter [14], when Z

has jumps; however it agrees with the integral originally proposed by Stratonovich for Brownian motion. Also all generalizations of the Stratonovich integral agree when Z is continuous.)

Throughout this section $\{Z_t\}$ shall denote a given k dimensional semimartingale. For $d \in \mathbb{N}$ and $f \in C^1(\mathbb{R}^d;\mathbb{R}^{dk})$, we shall say that the d-dimensional process X belongs to $\mathcal{E}^d(Z,f)$ if there exists a d-dimensional $\mathcal{F}_0$-measurable random vector X_0 such that:

$$X_t = X_0 + \int_0^t f(X_{s-})dZ_s + \frac{1}{2}\int_0^t f'f(X_s)d[Z,Z]_s^c \\ + \sum_{0<s\le t} \{\varphi(f\Delta Z_s, X_{s-}) - X_{s-} - f(X_{s-})\Delta Z_s\}$$

Definition 4.1 Let $d \in \mathbb{N}$, $X \in \mathcal{E}^d(Z,f)$ and $g \in C^1(\mathbb{R}^d;\mathbb{R}^k)$. We define the Stratonovich integral of the row vector $g(X)$ with respect to Z as follows:

$$\int_0^t g(X_s)\circ dZ_s = \int_0^t g(X_{s-})dZ_s \\ + \frac{1}{2}\,\mathrm{Tr}\int_0^t g'(X_s)d[Z,Z]_s^c f(X_s)^t \\ + \sum_{0<s\le t}\left(\int_0^1 \{g(\varphi(f\Delta Z_s, X_{s-}, u) - g(X_{s-})\}du\right)\Delta Z_s$$

□

Let us now comment on that definition. The first two terms on the right side of the above formula should be clear from the usual definition of Stratonovich integrals. However, the last term merits some comments. First of all, note that each term in the sum is of the order of $|\Delta Z_s|^2$, so that the sum converges. Furthermore that expression tells us that:

$$\Delta\left(\int_0^\cdot g(X_s)\circ dZ_s\right)_t = \left(\int_0^1 g(\varphi(\Delta Z f, X_{t-}, u))du\right)\Delta Z_t.$$

This formula can be interpreted as follows. At each jump time of Z, we open a unit length interval of "fictitious time", over which the integrand varies continuously from $g(X_{t-})$ to $g(X_t)$, and the jump of the integral equals the jump of the driving semimartingale multiplied by the mean of $g(x)$ along the curve joining X_{t-} to X_t.

We can now state and prove the associated change of variable formula:

Proposition 4.2 Let $d \in \mathbb{N}$, $f \in C^1(\mathbb{R}^d; \mathbb{R}^{dk})$, $X \in \mathcal{E}^d(Z, f)$, and $\Psi \in C^2(\mathbb{R}^d)$. We then have:

$$\Psi(X_t) = \Psi(X_0) + \int_0^t \Psi'(X_s) f(X_s) \circ dZ_s, t \geq 0$$

Proof: We know that X is a semimartingale and that:

$$\begin{aligned} dX_t &= f(X_{t-})dZ_t + \frac{1}{2} f' f(X_t) d[Z, Z]_t^c + \\ &\quad + \sum_{0<s\leq t} \{\varphi(\Delta Z_s f, X_{s-}) - X_{s-} - f(X_{s-})\Delta Z_s\}; \\ d[X, X]_t^c &= f(X_{t-}) d[Z, Z]_t^c f(X_{t-})^t; \\ \Delta X_t &= \varphi(\Delta Z_t f, X_{t-}) - X_{t-}. \end{aligned}$$

We plug these expressions into the Itô formula:

$$\begin{aligned} \Psi(X_t) = \Psi(X_0) &+ \int_0^t \Psi'(X_{s-}) dX_s + \frac{1}{2} \operatorname{Tr} \int_0^t \Psi''(X_s) d[X, X]_s^c + \\ &+ \sum_{0<s\leq t} (\Psi(X_{s-} + \Delta X_s) - \Psi(X_{s-}) - \Psi'(X_{s-})\Delta X_s) \end{aligned}$$

It is then easy to check that this expression coincides with

$$\Psi(X_0) + \int_0^t \Psi'(X_s) f(X_s) \circ dZ_s,$$

with the help of Definition 4.1. □

Let X denote the unique solution of equation

$$X_t = X_0 + \int_0^t f(X_s) \circ dZ_s, \quad t \geq 0$$

where $f \in C^1(\mathbb{R}^d; \mathbb{R}^k)$, f and $f'f$ being locally Lipschitz.

Let now M be a C^2 manifold without boundary embedded in $\mathbb{R}^d$, and assume that

$$f_j(x) \in T_x M, \quad x \in M, \quad 1 \leq j \leq k$$

i.e. $\{f_j(x), x \in M\}_{1\le j\le k}$ are vector fields over M. It is then intuitively clear that, starting on M, the solution X should stay on M. Indeed, between jumps, it obeys a continuous Stratonovich differential equation, and

$$x \longrightarrow \varphi(f\Delta Z_s, x)$$

maps M onto M. However, since there can be infinitely many jumps in a compact time interval, the above argument is not sufficient.

Suppose that the dimension of M is $\ell < d$. Locally, one can find a C^2 chart φ s.t. $\varphi_1(x), \ldots, \varphi_\ell(x)$ are coordinates on M, and $\varphi_{\ell+1}(x) = \ldots = \varphi_d(x) = 0$ if and only if $x \in M$. It then follows from Proposition 4.2, by using the same argument as for ODE's (see, for example, Hirsch [5], pp. 149–152):

Proposition 4.3 Let M be a C^2 manifold without boundary embedded in $\mathbf{R}^d$, and suppose that $\{f_j(x); x \in M\}_{1\le j\le k}$ are vector fields over M. Then $P(X_0 \in M) = 1$ implies that $P(X_t \in M,\ t \ge 0) = 1$. □

§5. Strong Markov Property

In the usual theory of stochastic differential equations, if $Z = (Z^1, \ldots, Z^m)$ is a vector of Lévy processes (i.e., processes with stationary and independent increments), and if $f: \mathbf{R}^d \longrightarrow \mathbf{R}^{d\times m}$ is Lipschitz, then the solution

$$X_t = X_0 + \int_0^t f(X_{s-})dZ_s \tag{5.1}$$

is strong Markov (see Protter [14, p. 238]). Recently the converse has been shown: Suppose f never vanishes and let X^x denote the solution with initial condition $X_0 = x$. If the processes X^x are time homogeneous Markov with the same transition semigroup for all x, then Z is a Lévy process (see Jacod-Protter [4]). We have the same Markov property for solutions with our Stratonovich-type differentials.

Theorem 5.1 Let f and $f'f$ be globally Lipschitz, and $Z = (Z^1, \ldots, Z^m)$ be a vector of Lévy processes. Then the solution X of

$$X_t = X_0 + \int_0^t f(X_s) \circ dZ_s \tag{5.2}$$

is strong Markov if X_0 is independent of $(Z_t)_{t\geq 0}$.

Proof: As in section three, we rewrite equation (5.2) as:

$$
\begin{aligned}
(5.3) \qquad X_t = X_0 + \int_0^t f(X_{s-})dZ_s \\
+ \frac{1}{2}\int_0^t f'f(X_s)d[Z,Z]_s^c \\
+ \int_0^t h(s,\cdot,X_{s-})d[Z,Z]_s^d
\end{aligned}
$$

Note that $[Z,Z]_t^c = \alpha t$ for some constant α because Z is a vector of Lévy processes (see, e.g., Theorem V.33 of Protter [14, p. 239]); thus $[Z,Z]^c$ is trivially also a Lévy process. One easily verifies that $[Z,Z]^j$ is also a Lévy process. Thus equation (5.3) falls within the "classical" province, where the equation is driven by Lévy semimartingales. The coefficients f and $f'f$ are Lipschitz, and h is process Lipschitz. There is one technical point: for fixed x, $h(s,\omega,x)$ is not predictable; it is optional. Moreover for fixed x it does not map càglàd (left continuous with right limits) processes into itself. However the differential for h, $d[Z,Z]_s^j$, is an increasing, finite variation process, and thus the established theory trivially extends to this case.

Adopting the framework of Çinlar-Jacod-Protter-Sharpe [1], we note that the coefficients f, $f'f$, and h are homogeneous in the sense of [1]; see page 214. (The coefficients f and $f'f$, being deterministic, are of course trivially homogeneous.)

The result now follows by a straightforward combination of the technique used to prove Theorem V.32 of Protter [14, p. 288] (where the coefficients are non-random), and the technique used to prove Theorem 8.11 of Çinlar-Jacod-Protter-Sharpe [1, p. 215], where the coefficients are homogeneous. □

§6. "Wong–Zakai" Type Approximations by Continuous Differentials

In 1965 Wong and Zakai [15] approximated the paths of Brownian motion with smooth processes, and they showed that the solutions of the corresponding ordinary differential equations converged to the solution of a Stratonovich–type stochastic differential equation, and not to the solution of an Itô–type equation. Their result has undergone many generalizations, culminating in Kurtz–Protter [7], where the Brownian differentials are replaced by general semimartingales. In Kurtz–Protter [7], however, and in all other treatments involving semimartingales with jumps, the approximating differentials must also have jumps, since convergence is in the Skorohod topology; and the limit of continuous approximants in either the uniform or Skorohod topologies must be continuous. Here we approximate the general semimartingale differentials with *continuous* approximants, even though the original semimartingale differentials may have jumps. The result is that the limiting equation is that of our new type of Stratonovich integral. This gives a justification for the use of our integral when one is modelling very sudden, sharp changes in an essentially continuous system.

For simplicity we consider here the case where Z is a one dimensional, given semimartingale. A generalization to systems of equations driven by vectors of semimartingales is possible with appropriate assumptions.

For the given (and fixed) semimartingale, we define approximations by:

$$Z_t^h = \frac{1}{h}\int_{t-h}^t Z_s ds \tag{6.1}$$

for $h > 0$. Then Z^h is adapted, continuous, and it is of finite variation on compacts. Moreover $\lim_{\substack{h\to 0\\ h>0}} Z_t^h = Z_{t-}$ a.s., each $t > 0$.

For given and fixed $f: \mathbf{R} \to \mathbf{R}$ that is continuously differentiable, we let X_t^h denote the unique solution of:

$$X_t^h = X_0 + \int_0^t f(X_s^h) dZ_s^h. \tag{6.2}$$

We want to show that X_t^h *converges to* X_t in probability, each $t > 0$, where

$$X_t = X_0 + \int_0^t f(X_s) \circ dZ_s. \tag{6.3}$$

We will use a method involving changes of time. Our new time scale will allow us to introduce the "fictitious time" where the solution follows the vector field f to form the jump. Let for $t > 0$:

$$[Z,Z]_t^d = \sum_{0<s\leq t} (\Delta Z_s)^2 < \infty \text{ a.s.,}$$

and $[Z,Z]^c = [Z,Z] - [Z,Z]^d$. The process $[Z,Z]^d$ is the purely discontinuous part of the quadratic variation process $[Z,Z]$ of Z. We define:

$$\gamma_h(t) = \frac{1}{h}\int_{t-h}^{t}([Z,Z]_s^d + s)ds. \tag{6.3}$$

Then γ_h is strictly increasing (since $[Z,Z]^d$ is increasing), continuous, and adapted. We also define

$$\gamma_0(t) = [Z,Z]_t^d + t,$$

which is also strictly increasing and adapted, although not continuous. Note that $\lim_{\substack{h\to 0\\ h>0}} \gamma_h(t) = \gamma_0(t-)$ a.s.

We define the continuous inverses

$$\gamma_h^{-1}(t) = \inf\{u > 0 : \gamma_h(u) > t\}.$$

Then $\gamma_h^{-1}(t)$ is a stopping time for each t, it is continuous, and it is strictly increasing for each $h > 0$. We next define

$$V_t^h = Z_{\gamma_h^{-1}(t)}^h. \tag{6.4}$$

The process V^h is continuous and it has paths of finite variation on compacts, since Z^h does. We next define Y^h to be the unique solution of

$$Y_t^h = X_0 + \int_0^t f(Y_s^h)dV_s^h. \tag{6.5}$$

Note that

$$Y_t^h = X_{\gamma_h^{-1}(t)}^h \tag{6.6}$$

by Lebesgue's change of time formula, where X^h is defined in (6.2). We next establish several preliminary results.

Lemma 6.1 $\lim_{\substack{h\to 0\\ h>0}} \gamma_h^{-1}(t) = \gamma_0^{-1}(t)$, uniformly.

Proof: Since $\gamma_h(t) < \gamma_0(t) < \gamma_h(t+h)$, it follows that $\gamma_h^{-1}(t) - h < \gamma_0^{-1}(t) < \gamma_h^{-1}(t)$, and the uniform convergence follows as well. □

We next want to find the limit of the processes V^h. Towards this end define

$$\eta_1(t) = \sup\{s: \gamma_0^{-1}(s) < \gamma_0^{-1}(t)\}$$
$$\eta_2(t) = \inf\{u: \gamma_0^{-1}(u) > \gamma_0^{-1}(t)\},$$

and

$$(6.7)\qquad V_t^0 = \begin{cases} Z^0_{\gamma_0^{-1}(t)} & \text{if } \eta_1(t) = \eta_2(t) \\ Z_{\gamma_0^{-1}(t)}\left(\frac{t-\eta_1(t)}{\eta_2(t)-\eta_1(t)}\right) + Z_{\gamma_0^{-1}(t)-}\left(\frac{\eta_2(t)-t}{\eta_2(t)-\eta_1(t)}\right) & \text{if } \eta_1(t) \neq \eta_2(t). \end{cases}$$

Note that $\eta_1(t) \leq \eta_2(t)$ always and that $\gamma_0^{-1}(s)$ is constant on $[\eta_1(t), \eta_2(t)]$. We can intuitively interpret V^0 as the semimartingale Z time–changed according to γ_0^{-1}, except when Z jumps. When Z jumps we add "imaginary" time intervals $[\eta_1(t), \eta_2(t)]$, of length of order $(\Delta Z)^2$; we add them at the times of the jumps of Z. During these intervals Z is made continuous by linear interpolation. Note that if $\sum_{0<s\leq t} |\Delta Z_s| < \infty$ a.s., each $t > 0$, then it is clear that V^0 can be interpreted as a semimartingale. However since it is possible to have $\sum_{0<s\leq t} |\Delta Z_s| = \infty$ a.s., every $t > 0$, these linear interpolations can have infinite length even on compact time intervals, and V^0 need not be a semimartingale. In all cases however V^0 is a continuous process adapted to the filtration $\mathcal{G}_t = \mathcal{F}_{\gamma_0^{-1}(t)}$, because $\lim_{h\to 0} \gamma_0(\gamma_0^{-1}(t) - h) = \eta_1(t)$ and $\lim_{h\to 0} \gamma_0(\gamma_0^{-1}(t) + h)) = \eta_2(t)$.

Lemma 6.2 $\lim_{\substack{h\to 0\\ h>0}} V^h = V^0$, uniformly on bounded intervals.

Proof: Note that $V^h = Z^h_{\gamma_h^{-1}(t)}$ as defined in (6.4), and V^0 is given by (6.7). By Lemma 6.1 we have $\lim_{h\to 0} \gamma_h^{-1}(t) = \gamma_0^{-1}(t)$, hence if $\eta_1(t) = \eta_2(t)$, we have $\lim_{h\to 0} V_t^h = V_t^0$, since $\gamma_0^{-1}(t)$ is a continuity point of Z. Thus we assume $\eta_1(t) \neq \eta_2(t)$.

Next observe that

$$(6.8)\qquad \frac{d}{dt}\gamma_h^{-1}(t) = \frac{h}{[Z,Z]^d_{\gamma_h^{-1}(t)} - [Z,Z]^d_{\gamma_h^{-1}(t)-h} + h}$$

and so

$$(6.9)\qquad \frac{d}{dt}V_t^h = \frac{Z_{\gamma_h^{-1}(t)} - Z_{\gamma_h^{-1}(t)-h}}{[Z,Z]^d_{\gamma_h^{-1}(t)} - [Z,Z]^d_{\gamma_h^{-1}(t)-h} + h}.$$

Therefore we have

$$\lim_{h\to 0}\left(\frac{d}{dt}V_t^h\right) = \frac{Z_{\gamma_0^{-1}(t)} - Z_{\gamma_0^{-1}(t)-}}{\eta_2(t)-\eta_1(t)},$$

which is the derivative of V_t^0 in $[\eta_1(t), \eta_2(t)]$, and which implies that $\lim_{h\to 0} V_t^h = V_t^0$ when $\eta_1(t) \neq \eta_2(t)$. Since it is easy to see that for $\varepsilon > 0$, there exists a $\delta > 0$ such that for $|s-t| < \delta$, $\sup_{h\leq h_0} |V_s^h - V_t^h| < \varepsilon$, we deduce the uniform convergence. □

Before continuing we need to introduce a concept from Kurtz–Protter [8].

Definition: A sequence of semimartingales Z^n is said to be *good in probability* if whenever (H^n, Z^n) converges in probability to (H, Z) in the Skorohod topology, where H^n, H are càdlàg, adapted; then Z is a semimartingale and $\int_0^\cdot H_{s-}^n dZ_s^n$ converges in probability in the Skorohod topology to $\int_0^\cdot H_{s-} dZ_s$.

A necessary and sufficient condition for a sequence of semimartingales Z^n to be good was obtained in Kurtz–Protter [7,8]: let $h_\delta(r) = (1-\delta/r)^+$, and $J_\delta\colon D[0,\infty) \to D[0,\infty)$ by

$$J_\delta(x)(t) = \sum_{0<s\leq t} h_\delta(|\Delta x_s|)\Delta x_s$$

and

$$Z^{n,\delta} = Z^n - J_\delta(Z^n).$$

Then $Z^{n,\delta}$ has jumps bounded by δ, and let

$$Z^{n,\delta} = M^{n,\delta} + A^{n,\delta}$$

be any decomposition of $Z^{n,\delta}$ into a local martingale $M^{n,\delta}$ and an adapted, càdlàg, finite variation process $A^{n,\delta}$. The condition is that there exists such a decomposition such that:

(*) For each $\alpha > 0$, there exist stopping times T_n^α such that $P(T_n^\alpha \leq \alpha) \leq 1/\alpha$ and

$$\sup_n E\left\{[M^{n,\delta}, M^{n,\delta}]_{t\wedge T_n^\alpha} + \int_0^{t\wedge T_n^\alpha} |dA_s^{n,\delta}|\right\} < \infty.$$

We next define

$$U_t^h = V_t^h - Z_{\gamma_0^{-1}(t)}. \tag{6.10}$$

Note that by Lemma 6.2 we have

$$\lim_{h\to 0} U_t^h = V_t^0 - Z_{\gamma_0^{-1}(t)}.$$

Lemma 6.3 The sequence of semimartingales $\int_0^t U_s^h dU_s^h$ is good in probability, and moreover

$$\int_0^t U_s^h dU_s^h \Rightarrow \frac{1}{2}\left\{(V_t^0 - Z_{\gamma_0^{-1}(t)})^2 - [Z,Z]_{\gamma_0^{-1}(t)}\right\}.$$

Proof: By definition

$$\int_0^t U_s^h dU_s^h = \int_0^t (V_s^h - Z_{\gamma_0^{-1}(s)})d(V_s^h - Z_{\gamma_0^{-1}(s)}),$$

and since $(Z_{\gamma_0^{-1}(s)})_{s\geq 0}$ is fixed, it is enough to study

$$\int_0^t (V_s^h - Z_{\gamma_0^{-1}(s)})dV_s^h.$$

Using our expression (6.9) for the derivative of V^h, we have

$$\int_0^t (V_s^h - Z_{\gamma_0^{-1}(s)})\frac{Z_{\gamma_h^{-1}(s)} - Z_{\gamma_h^{-1}(s)-h}}{[Z,Z]^d_{\gamma_h^{-1}(s)} - [Z,Z]^d_{\gamma_h^{-1}(s)-h} + h}\, ds$$

$$= \int_0^t \frac{\left(\frac{1}{h}\int_{\gamma_h^{-1}(s)-h}^{\gamma_h^{-1}(s)} Z_r dr - Z_{\gamma_0^{-1}(s)}\right)\left(Z_{\gamma_h^{-1}(s)} - Z_{\gamma_h^{-1}(s)-h}\right)}{[Z,Z]^d_{\gamma_h^{-1}(s)} - [Z,Z]^d_{\gamma_h^{-1}(s)-h} + h}\, ds$$

and letting $u = \gamma_h^{-1}(s)$, we have (using (6.8)):

$$= \int_0^{\gamma_h^{-1}(t)} \frac{(\frac{1}{h}\int_{u-h}^u Z_r dr - Z_{\gamma_0^{-1}(\gamma_h(u))})(Z_u - Z_{u-h})}{h}\, du.$$

It is easy to see that if a sequence of semimartingales $(Z^n)_{n\geq 1}$ defined on the same space is good for one probability measure P, then it is also good for any other probability Q equivalent to P, because if $Q << P$, then convergence in P–probability implies convergence in Q–probability. Thus without loss of generality, by changing to an equivalent probability measure if necessary, we can assume that Z is in $\mathcal{H}^2$; that is, Z has a canonical decomposition $Z = M + A$, where $E\left\{[M,M]_t + \int_0^t |dA_s|\right\}^2 < \infty$, for any finite time t. By stopping, we can further assume that for some $T < \infty$, $Z_t = Z_{t\wedge T}$. Therefore

$$E\left\{\int_0^{\gamma_h^{-1}(T)} \frac{|(\frac{1}{h}\int_{u-h}^u Z_s ds - Z_{\gamma_0^{-1}(\gamma_h(u))})(Z_u - Z_{u-h})|}{h}\, du\right\}$$

$$(6.11) \leq E\left\{\int_0^\infty \frac{(\frac{1}{h}\int_{u-h}^u Z_s ds - Z_{\gamma_0^{-1}(\gamma_h(u))})^2}{h}\, du\right\}^{1/2} E\left\{\int_0^\infty \frac{(Z_u - Z_{u-h})^2}{h}\, du\right\}^{1/2}$$

$$\leq E\left\{\int_0^\infty \frac{\sup_{u-h\leq s,t\leq u}(Z_s - Z_t)^2}{h}\, du\right\}^{1/2} E\left\{\int_0^\infty \frac{(Z_u - Z_{u-h})^2}{h}\, du\right\}^{1/2}$$

where we have used that $u - h \leq \gamma_0^{-1}(\gamma_h(u)) \leq u$, which follows from the definition of $\gamma_0^{-1}(t)$ together with the observation that $\gamma_h(u) > \gamma_0(u-h)$. Consider the first term on the right above. Using Doob's maximal quadratic inequality for martingales:

$$E\left\{\int_0^\infty \frac{\sup_{u-h\leq s,t\leq u}(Z_s - Z_t)^2}{h}\, du\right\}^{1/2}$$

$$\leq 2\int_0^\infty \frac{1}{h} E\left\{\sup_{u-h\leq s\leq u}(M_s - M_{s-h})^2\right\}\, du$$

$$+ 2\int_0^\infty \frac{1}{h} E\{(|A|_u - |A|_{u-h})^2\}\, du \quad (\text{where } |A|_u = \int_0^u |dA_s|),$$

$$\leq 8\int_0^\infty \frac{1}{h} E\{[M,M]_u - [M,M]_{u-h}\}\, du + 2\int_0^\infty \frac{1}{h} E\{(|A|_u - |A|_{u-h})^2\}\, du,$$

$$\leq 8\int_0^\infty \frac{1}{h} E\{[Z,Z]_u - [Z,Z]_{u-h}\}\, du + 2\int_0^\infty \frac{1}{h} E\left\{\left(\int_{u-h}^u |dA_s|\right)^2\right\}\, du,$$

with the last inequality holding because $E\{[M,M]_t\} \leq E\{[Z,Z]_t\}$ for the canonical decomposition (see Protter [14, p. 136]) and analogous inequalities hold for the increments. However since $E\{[Z,Z]_t\}$ and $E\{(\int_0^t |dA_s|)^2\}$ are increasing functions, their derivatives exist a.e. (dt). Therefore the first term on the right side of (6.11) is bounded in h; analogously we have the second term bounded in h. We conclude that $(*)$ is satisfied; hence we have goodness.

To identify the limit of $\int U_s^h dU_s^h$, we use integration by parts:

$$\int_0^t U_s^h dU_s^h = \frac{1}{2}((U_t^h)^2 - [Z,Z]_{\gamma_0^{-1}(t)}),$$

and by the definition of U^h (see (6.10)),

we have
$$\lim_{h\to 0}(U_t^h)^2 = \lim_{h\to 0}(V_t^h - Z_{\gamma_0^{-1}(t)})^2$$
$$= (V_t^0 - Z_{\gamma_0^{-1}(t)})^2$$

by Lemma 6.2. □

For the remaining results we need to use the concept of the topology of convergence of measure, popularized by Meyer and Zheng. For relevant definitions and results we refer the reader to Kurtz [6].

Lemma 6.4 Using the Meyer–Zheng topology of convergence in measure, the sequence $\int_0^t f(Y_s^h)dV_s^h$ is relatively compact, for convergence in probability.

Proof: By the definition of $U_t^h = V_t^h - Z_{\gamma_0^{-1}(t)}$, we have

$$\int_0^t f(Y_s^h)dV_s^h = \int_0^t f(Y_s^h)dZ_{\gamma_0^{-1}(s)} + \int_0^t f(Y_s^h)dU_s^h;$$

since $Z_{\gamma_0^{-1}(s)}$ is a fixed semimartingale it easily follows from Theorem 5.8 in Kurtz [6] that $\int f(Y_s^h)dZ_{\gamma_0^{-1}(s)}$ is relatively compact. Thus, it suffices to show that $\int_0^t f(Y_s^h)dU_s^h$ is relatively compact. Using integration by parts, we have

$$\int_0^t f(Y_s^h)dU_s^h = f(Y_t^h)U_t^h - \int_0^t U_s^h f'(Y_s^h)dY_s^h$$
$$- [f(Y^h),\ U^h]_t.$$

However Y^h is continuous and of finite variation, and since f is C^1, so also is $f(Y^h)$; therefore $[f(Y^h),\ U^h] = 0$. Hence

$$\int_0^t f(Y_s^h)dU_s^h = f(Y_t^h)U_t^h - \int_0^t U_s^h f'(Y_s^h)dY_s^h, \tag{6.12}$$

and considering the last term in (6.12) we have:

$$\int_0^t U_s^h f'(Y_s^h)dY_s^h = \int_0^t U_s^h f'f(Y_s^h)dZ_{\gamma_0^{-1}(s)} + \int_0^t U_s^h f'f(Y_s^h)dU_s^h$$
$$= \int_0^t U_s^h f'f(Y_s^h)dZ_{\gamma_0^{-1}(s)} + \int_0^t f'f(Y_s^h)U_s^h dU_s^h,$$

and therefore, since $U^h dU^h$ is good, we have relative compactness of this term.

Next consider the first term on the right side of (6.12). We have

$$f(Y_t^h)U_t^h = f(Y_t^h)(V_t^h - V_t^0) + f(Y_t^h)(V_t^0 - Z_{\gamma_0^{-1}(t)}).$$

Now $f(Y_t^h)(V_t^h - V_t^0)$ converges to zero and hence it is relatively compact. It remains to consider $f(Y_t^h)(V_t^0 - Z_{\gamma_0^{-1}(t)})$. Let $U_t = V_t^0 - Z_{\gamma_0^{-1}(t)}$. Note that $U_t = 0$ unless $\eta_1(t) \neq \eta_2(t)$. If $\eta_1(t) \neq \eta_2(t)$, then

$$U_t = V_t^0 - Z_{\gamma_0^{-1}(t)} = Z_{\gamma_0^{-1}(t)} \frac{t - \eta_1(t)}{\eta_2(t) - \eta_1(t)} + Z_{\gamma_0^{-1}(t)-} \frac{\eta_2(t) - t}{\eta_2(t) - \eta_1(t)} - Z_{\gamma_0^{-1}(t)},$$

and replacing $-Z_{\gamma_0^{-1}(t)}$ with $-Z_{\gamma_0^{-1}(t)}\left(\frac{\eta_2(t)-t}{\eta_2(t)-\eta_1(t)} + \frac{t-\eta_1(t)}{\eta_2(t)-\eta_1(t)}\right)$, this becomes

$$U_t = -\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - t}{\eta_2(t) - \eta_1(t)}. \tag{6.13}$$

Therefore if s, t are such that $|s - t| < \delta$:

$$\begin{aligned}
&|f(Y_t^h)U_t - f(Y_s^h)U_s| \\
&\leq |f(Y_t^h)\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - t}{\eta_2(t) - \eta_1(t)} 1_{[\eta_1(t),\eta_2(t))}(t) \\
&\quad - f(Y_s^h)\Delta Z_{\gamma_0^{-1}(s)} \frac{\eta_2(s) - s}{\eta_2(s) - \eta_1(s)} 1_{[\eta_1(s),\eta_2(s))}(s)| \\
&\leq |f(Y_t^h)\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - t}{\eta_2(t) - \eta_1(t)} 1_{[\eta_1(t),\eta_2(t))}(t) \\
&\quad - f(Y_t^h)\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - s}{\eta_2(t) - \eta_1(t)} 1_{[\eta_1(t),\eta_2(t))}(s)1_{[\eta_1(t),\eta_2(t))}(t)| \\
&\quad + |f(Y_t^h)\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - s}{\eta_2(t) - \eta_1(t)} 1_{[\eta_1(t),\eta_2(t))}(t)1_{[\eta_1(t),\eta_2(t))}(s) \\
&\quad - f(Y_s^h)\Delta Z_{\gamma_0^{-1}(s)} \frac{\eta_2(s) - s}{\eta_2(s) - \eta_1(s)} 1_{[\eta_1(s),\eta_2(s))}(s)| \\
&= I_h(t,s) + J_h(t,s). \qquad (6.14)
\end{aligned}$$

For an interval I and function α let $w(\alpha; I) = \sup_{s,t\in I} |\alpha(s) - \alpha(t)|$, and let

$$w_N(\alpha, \theta) = \sup\{w(\alpha; [t, t+\theta]) : 0 \leq t \leq t+\theta \leq N\},$$

for $\theta > 0$ and N an integer. Then by the Arzela–Ascoli theorem a subset A of $\mathcal{C}(\mathbf{R})$ is relatively compact locally uniformly if (i) $\sup_{\alpha\in A} |\alpha(0)| < \infty$, and (ii) for all N, $\lim_{\substack{\theta\to 0\\ \theta>0}} \sup_{\alpha\in A} w_N(\alpha, \theta) = 0$. We will use this criterion on $I_h(t,s)$ and $J_h(t,s)$. Consider for example $I_h(t,s)$ of (6.14) above. For our given δ, we can choose a partition $\{t_i\}$ such that $t_{i+1} - t_i > \delta$ and

$$\max_i w(Z(\omega); [t_i, t_{i+1})) < \varepsilon. \tag{6.15}$$

Then

$$\begin{aligned}|I_h(t,s)| &\leq \frac{|\Delta Z_{\gamma_0^{-1}(t)} f(Y_t^h)}{\eta_2(t) - \eta_1(t)} 1_{[\eta_1(t),\eta_2(t))}(t) |\eta_2(t) - t - (\eta_2(t) - s) 1_{[\eta_1(t),\eta_2(t))}(s)| \\ &< C\varepsilon,\end{aligned}$$

where the last inequality follows since the jumps are bounded by the choice of the partition in (6.15). Analogous arguments yield that the Arzela–Ascoli criterion holds for $J_h(t,s)$ as well. □

Theorem 6.5 Let Z be a given semimartingale, $Z_t^h = \frac{1}{h}\int_{t-h}^t Z_s ds$ for $h > 0$, and

$$\begin{aligned} X_t^h &= X_0 + \int_0^t f(X_s^h) dZ_s^h \\ X_t &= X_0 + \int_0^t f(X_s) \circ dZ_s. \end{aligned} \tag{6.16}$$

Then $\lim_{h\to 0} X_t^h = X_t$, each $t > 0$, with convergence in probability, except possibly for a countable set of t's.

Proof: Recalling the definitions (6.4), (6.5) and (6.6), let $V_t^h = Z^h_{\gamma_h^{-1}(t)}$, and

$$Y_t^h = X_0 + \int_0^t f(Y_s^h) dV_s^h$$

and

$$Y_t^h = X^h_{\gamma_h^{-1}(t)}.$$

Consider the terms $\int_0^t f(Y_s^h)dV_s^h$:

$$\int_0^t f(Y_s^h)dV_s^h = \int_0^t f(Y_s^h)dZ_{\gamma_0^{-1}(s)} + \int_0^t f(Y_s^h)dU_s^h$$

(where U^h is defined in (6.10)). Consider

$$\int_0^t f(Y_s^h)dU_s^h = f(Y_t^h)U_t^h - \int_0^t U_s^h f'f(Y_s^h)dZ_{\gamma_0^{-1}(s)} - \int_0^t U_s^h f'f(Y_s^h)dU_s^h.$$

However $\int_0^t U_s^h f'f(Y_s^h)dU_s^h = \int_0^t f'f(Y_s^h)U_s^h dU_s^h$, and since $U_s^h dU_s^h$ is good, using Lemma 6.4, Theorem 2.2 of Kurtz–Protter [7], and the relative compactness in the topology of convergence in measure of $\int_0^t f(Y_s^h)dV_s^h$, we conclude that any limit point of the sequence (Y^h) satisfies

$$Y_t = X_0 + \int_0^t f(Y_s)dZ_{\gamma_0^{-1}(s)} + f(Y_t)U_t - \int_0^t U_s f'f(Y_s)dZ_{\gamma_0^{-1}(s)} - \frac{1}{2}\int_0^t f'f(Y_s)d(U_s^2 - [Z,Z]_{\gamma_0^{-1}(s)}). \tag{6.17}$$

(Note that Y is continuous, and that U^2 is a semimartingale by the goodness of $U^h dU^h$ even though U need not be: the function $f(x) = \sqrt{x}$ is not the difference of two convexes.)

Let us study equation (6.17). Note that U_t is zero unless $\eta_1(t) \neq \eta_2(t)$; in this case the jump has size $-\Delta Z_{\gamma_0^{-1}(t)}$. Therefore

$$-\int_0^t U_s f'f(Y_s)dZ_{\gamma_0^{-1}(s)} = \sum_{0<s\leq t} f'f(Y_s)(\Delta Z_{\gamma_0^{-1}(s)})^2,$$

and so equation (6.17) simplifies to:

$$Y_t = X_0 + \int_0^t f(Y_s)dZ_{\gamma_0^{-1}(s)} + f(Y_t)U_t - \frac{1}{2}\int_0^t f'f(Y_s)dU_s^2 + \frac{1}{2}\int_0^t f'f(Y_s)d[Z,Z]^c_{\gamma_0^{-1}(s)}. \tag{6.18}$$

Further analyzing the process U, we recall from (6.13) that

$$U_t = -\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - t}{\eta_2(t) - \eta_1(t)} 1_{[\eta_1(t),\eta_2(t))}(t).$$

Therefore when $t \in [\eta_1(t), \eta_2(t))$, we have:

$$(6.19) \qquad Y_t = Y_{\eta_1(t)} + f(Y_{\eta_1(t)})\Delta Z_{\gamma_0^{-1}(t)} - f(Y_{\eta_1(t)})\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - t}{\eta_2(t) - \eta_1(t)}$$
$$+ \int_{[\eta_1(t),t)} f'f(Y_s)(\Delta Z_{\gamma_0^{-1}(t)})^2 \frac{\eta_2(t) - s}{(\eta_2(t) - \eta_1(t))^2} ds.$$

Since $f(Y_{\eta_1(t)})\Delta Z_{\gamma_0^{-1}(t)} - \int_{\eta_1(t)}^{\eta_2(t)} \frac{f(Y_s)\Delta Z_{\gamma_0^{-1}(t)}}{\eta_2(t)-\eta_1(t)} ds$ is of order $(\Delta Z_{\gamma_0^{-1}(t)})^2$, we can sum the terms (6.19) and then (6.18) becomes

$$(6.20)\, Y_t = X_0 + \int_0^t f(Y_s)dZ_{\gamma_0^{-1}(s)} + \frac{1}{2}\int_0^t f'f(Y_s)d[Z,Z]^c_{\gamma_0^{-1}(s)}$$
$$+ \sum_{0<s\leq\gamma_0^{-1}(\eta_2(t))} \left\{ \int_{\eta_1(\gamma_0(s))}^{\eta_2(\gamma_0(s))} \frac{f(Y_r)\Delta Z_s}{\eta_2(\gamma_0(s)) - \eta_1(\gamma_0(s))} dr - f(Y_{\eta_1(\gamma_0(s))})\Delta Z_s \right\}.$$

Since $(Y^h, V^h) = (X^h_{\gamma_h^{-1}}, Z^h_{\gamma_h^{-1}})$ is relatively compact and converges to (Y, V), so also is (X^h, Z^h) relatively compact in the topology of convergence in measure. Moreover $\gamma_h^{-1}(\gamma_h(t)) = t$ since γ_h is continuous and strictly increasing; therefore $(Y^h_{\gamma_h}, V^h_{\gamma_h})$ converges to $(Y_{\gamma_0}, V_{\gamma_0})$, and applying this to (6.20) yields, except for possibly countably many t:

$$Y_{\gamma_0(t)} = X_0 + \int_0^{\gamma_0(t)} f(X_s)dZ_{\gamma_0^{-1}(s)} + \frac{1}{2}\int_0^{\gamma_0(t)} f'f(Y_s)d[Z,Z]^c_{\gamma_0^{-1}(s)}$$
$$+ \sum_{0<s\leq\gamma_0^{-1}(\eta_2(\gamma_0(t)))} \left\{ \int_{\eta_1(\gamma_0(s))}^{\eta_2(\gamma_0(s))} \frac{f(Y_r)\Delta Z_s}{\eta_2(\gamma_0(s)) - \eta_1(\gamma_0(s))} dr - f(Y_{\eta_1(\gamma_0(s))})\Delta Z_s \right\}$$

and applying a change of variables formula shows that X satisfies equation (6.16). □

REFERENCES

1. E. Çinlar, J. Jacod, P. Protter, and M. Sharpe, Semimartingales and Markov Processes, *Z. für Wahrscheinlichkeitstheorie* **54** (1980), 161–220.

2. S. Cohen, Geometrie differentielle stochastique avec sauts, *C.R. Acad. Sci. Paris* **314** (1992), 767–770.

3. A. Estrade, Exponentielle stochastique et intégrale multiplicative discontinue, to appear in *Ann. Inst. H. Poincaré* (1992).

4. J. Jacod and P. Protter, Une remarque sur les equations différentielles stochastiques à solutions Markoviennes, in *Séminaire de Probabilités* **XXV**, Springer Lect. Notes in Math. 1485 (1991), 138–139.

5. M.W. Hirsch, *Differential Topology*, Springer–Verlag, New York Berlin 1976.

6. T. G. Kurtz, Random time changes and convergence in distribution under the Meyer-Zheng conditions, *Annals of Probability* **19** (1991), 1010–1034.

7. T. G. Kurtz and P. Protter, Weak limit theorems for stochastic integrals and stochastic differential equations, *Annals of Probability* **19** (1991), 1035–1070.

8. T. G. Kurtz and P. Protter, Characterizing the weak convergence of stochastic integrals, in *Stochastic Analysis* (M. Barlow and N. Bingham, eds.) (1991), 255–259.

9. H. J. Kushner, Jump-diffusion approximations for ordinary differential equations with wide-band random right hand sides, *SIAM J. Control* **17** (1979), 729–744.

10. S. I. Marcus, Modeling and analysis of stochastic differential equations driven by point processes, *IEEE Transactions on Information Theory*, **24** (1978), 164–172.

11. S. I. Marcus, Modeling and approximation of stochastic differential equations driven by semimartingales, *Stochastics*, **4** (1981), 223–245.

12. M. Métivier, *Semimartingales: a course on stochastic processes*, de Gruyter, Berlin New York, 1982.

13. P. A. Meyer, Un cours sur les intégrales stochastiques, Séminaire Proba. X, Lecture Notes in Mathematics, vol. 511, pp. 246–400, Springer-Verlag, Berlin New York, 1976.

14. P. Protter, *Stochastic integration and differential equations: a new approach*, Springer-Verlag, Berlin New York, 1990.

15. E. Wong and M. Zakai, On the convergence of ordinary integrals to stochastic integrals, *Ann. Math. Statist.* **36** (1965), 1560–1564.