ANNALES DE L'I. H. P., SECTION B

THOMAS G. KURTZ
ÉTIENNE PARDOUX
PHILIP PROTTER

Stratonovich stochastic differential equations driven by general semimartingales

Annales de l'I. H. P., section B, tome 31, nº 2 (1995), p. 351-377
<http://www.numdam.org/item?id=AIHPB_1995__31_2_351_0>

Ann. Inst. Henri Poincaré,

Vol. 31, n° 2, 1995, p. 351-377 *Probabilités et Statistiques*

Stratonovich stochastic differential equations driven by general semimartingales

by

Thomas G. KURTZ (1)

Departments of Mathematics and Statistics,
University of Wisconsin, Madison, Wisconsin 53706, U.S.A.

Étienne PARDOUX

Mathématiques, U.R.A. 225, Université de Provence,
13331 Marseille Cedex 3, France

and

Philip PROTTER (2)

Departments of Mathematics and Statistics,
Purdue University, West Lafayette, IN 47907-1395, U.S.A.

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.

Key words: Stratonovich integrals, stochastic differential equations, reflection, semimartingale.

A.M.S. Classification : 60 H 10, 60 H 20, 60 J 25, 60 J 60.

(1) Supported in part by NSF grant #DMS-8901464.

(2) Supported in part by NSF grant #DMS-9103454.

RÉSUMÉ. – On considère des équations stochastiques différentielles où le « bruit » est une semimartingale quelconque (avec des sauts). On propose une interprétation 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'interprétation de Meyer-Stratonovich). De plus on établit les propriétés de Markov fortes, et on démontre un genre de convergence faible du 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 [16] or Protter [17].) 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:

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

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 [17], p. 62). The third term is a (possibly countable) sum of terms of order $(\Delta Z_s)^2$ and therefore converges absolutely (*see* Section 2). Were we to have interpreted (1.1) as a Stratonovich equation in the sense of the semimartingale Stratonovich integral as defined by Meyer [16] (*see* also Protter [17]), 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 6) 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 4.) The third (*see* Section 3) 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 [17], 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 ([13], [14]) in the case where Z has finitely many jumps on compact time intervals. The corresponding "Wong-Zakai" results were investigated by Kushner [12]. Recently Estrade [4] 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_j^i 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 $\left(\text{that is we write } a_i \text{ to denote } \sum_{i=1}^{d} a_i\right)$.

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 [17]. A k-dimensional semimartingale $Z = \{Z_t; t \geq 0\}$ is assumed given with $Z_0 = 0$. $[Z,Z] \equiv (([Z^j, Z^m]))$ will denote the matrix of covariations and $[Z,Z] = [Z,Z]^c + [Z,Z]^d$ denotes its decomposition into continuous and purely discontinuous parts. $[Z]$ will denote the scalar quadratic variation, that is, $[Z] = \sum_{j=1}^{k} [Z^j, Z^j]$, with $[Z]^c$ and $[Z]^d$ the corresponding continuous and purely discontinuous parts.

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

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

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 [17], p. 58). Let us finally define the notation $\varphi(f\Delta Z_s, x)$. Given $g \in C^1(\mathbb{R}^d;\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, the following equation:

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

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

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

If $\xi > 1$,

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

If $\xi \leq 1$, $\varphi(g,x)$ 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 $\mathcal{C}^2$, we have:

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

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

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

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 simplified a previous version of this paper. For a given vector of semimartingales Z, we define

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

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*

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

is also a solution of

$$X_t = X_0 + \int_0^t f(X_{s-})dZ_s + \frac{1}{2}\int_0^t f'f(X_{s-})d[Z,Z]_s^c + \int_0^t h(s,\cdot,X_{s-})d[Z]_s^d, \tag{2.5}$$

and conversely.

3. EXISTENCE, UNIQUENESS AND FLOWS OF THE EQUATION

One can study equation (2.1) directly (as the authors did during their preliminary efforts), but it is much more efficient to consider (2.5). We will call an operator F on processes *process Lipschitz* as defined in Protter ([17], 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

$$u \rightarrow \varphi(f\Delta Z_s, x, u) - \varphi(f\Delta Z_s, y, u):$$

to obtain

$$\begin{aligned}
&|\varphi(f\Delta Z_s,x,u) - \varphi(f\Delta Z_s,y,u) - x - y - (f(x)-f(y))\Delta Z_s| \\
&\quad\leq \frac{1}{2}|\{f'f(\varphi(f\Delta Z_s,x,\theta)) - f'f(\varphi(f\Delta Z_s,y,\theta))\}\Delta Z_s \Delta Z_s^t| \\
&\quad\leq c|\varphi(f\Delta Z_s,x,\theta) - \varphi(f\Delta Z_s,y,\theta)||\Delta Z_s|^2 \\
&\quad\leq c|x-y|e^{c|\Delta Z_s|}|\Delta Z_s|^2,
\end{aligned} \tag{3.1}$$

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

$$|h(s,w,x) - h(s,w,y)| \leq c|x-y|e^{c|\Delta Z_s|}, \tag{3.2}$$

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 [17].

THEOREM 3.2. – *Let f and $f'f$ be globally Lipschitz. Then there exists a càdlàg solution to* (2.1), *it is unique, and it is a semimartingale.*

Proof. – Rewriting equation (2.1) in its equivalent form (2.5), we observe that (2.5) is a standard stochastic differential equation with semimartingale differentials Z, $[Z,Z]^c$, and $[Z]^d$, 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 the collection of càglàd (left continuous with right limits), adapted processes to itself. However the process $[Z,Z]_s^j$ 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 ([17], 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 [15], Theorem 34.7, p. 246 or Protter [17], 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* (2.1). *Moreover* $\limsup_{t \to T} ||X_t|| = \infty$ *a.s. on the event* $\{T < \infty\}$.

Remark. – A more general equation than (2.2) is the following

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

where J is a càdlàg, adapted process such that $[J,Z]^c$ exists (in the sense defined in Protter [17], 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

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

This equation can be shown, by a slight extension of Theorem 6.5, to be the natural limit of approximating equations of the form

$$X_t^h = J_t + \int_0^t f(X_s^h) dZ_s^h.$$

Existence and uniqueness of solutions for equation (3.3) follows as in Theorem 3.2. Note that if J is a semimartingale, then equation (3.3) can be put into the form of equation (2.2). We shall restrict ourselves to the case where J is a semimartingale in this paper.

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

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

The *flow* of the stochastic differential equation (3.4) is the function $x \longrightarrow X_t(x, \omega)$, which can be considered as a mapping from $\mathbb{R}^d \to \mathbb{R}^d$ for (t, ω) fixed, or as a mapping from $\mathbb{R}^d \to \mathcal{D}^d$, where $\mathcal{D}^d$ denotes the space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}^d$, 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 $x \to X(x, \omega)$ from $\mathbb{R}^d$ 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 (2.5). 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 ([17], p. 246). □

We henceforth consider the flow of equation (3.4) as a function from $\mathbb{R}^d$ to $\mathbb{R}^d$, for each fixed (t, ω). Let Ψ denote the flow: that is, $\Psi : \mathbb{R}^d \to \mathbb{R}^d$ is given by $\Psi_t(x) = X_t(x, \omega)$ 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 $||Z||_{\mathcal{H}^\infty} \leq \varepsilon$, 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, $1 \leq \alpha \leq m$. (*See* Theorem V.5, p. 192 of Protter [17].) The above observation allows us to first consider semimartingale differentials with small $\mathcal{H}^\infty$ norm.

Let $X_t^j(x)$ satisfy

$$X_t^j(x) = x + \int_0^t f(X_s) \circ dZ_s^{\alpha,j}$$

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

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

Define the *linkage operators* H^j by $H^j(x) = \varphi(f\Delta Z_{T_j}, x)$.

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 $\mathbb{R}^d$.*

Next we have the obvious result:

THEOREM 3.6. – *The solution of* (3.4) *is given by*

$$X_t(x) = X_t^{j+1}(x_{j+}), \qquad T_j \leq t < T_{j+1},$$

where

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

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* (3.4) *is a diffeomorphism if, for each* j,

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

is a diffeomorphism.

Proof. – By Theorem 3.6, the solution X of (3.4) 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 [17].

THEOREM 3.8. – *Let* f, $f'f$ *in* (3.4) *be* $\mathcal{C}^\infty$ *with all derivatives bounded. If* $||Z||_{\mathcal{H}^\infty} < \varepsilon$ *for* $\varepsilon > 0$ *sufficiently small, then the corresponding flow* Ψ *is a diffeomorphism of* $\mathbb{R}$.

Proof. – We rewrite the equation (3.4) in the form (2.5). Equation (2.5) 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 [17] to deduce the result. □

Combining Theorems 3.7 and 3.8, we have:

THEOREM 3.9. – *Let* f *in* (3.4) *be* $\mathcal{C}^\infty$ *with all derivatives of* f *and* $f'f$ *bounded. Then the flow* $\Psi : x \rightarrow X_t(x,\omega)$ *of the solution* X *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 (2.1). We will establish for $g \in C^1(\mathbb{R}^d;\mathbb{R}^k)$ that we can define an integral

$$\int_0^t g(X_s) \circ dZ_s = \int_0^t g_i(X_s) \circ dZ_s^i \tag{4.1}$$

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 [16] and Protter [17], 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.)

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:

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

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:

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

□

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 comment. First, 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:

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

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:*

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

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

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

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. □

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

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

i.e. $\{f_j(x), x \in M\}_{1\leq j\leq k}$ 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 does not immediately imply that X stays on M.

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 $x \in M$. The desired result then follows from Proposition 4.2, by using the same argument as for ODE's (*see*, for example, Hirsch [7], pp. 149-152).

PROPOSITION 4.3. – *Let M be a C^2 manifold without boundary embedded in $\mathbb{R}^d$, 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 is a Lévy process (*i.e.*, a process with stationary and independent increments), and if $f : \mathbb{R}^d \to \mathbb{R}^{d \times k}$ is Lipschitz, then the solution of

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

is strong Markov (*see* Protter [17], 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 [6]). We have the same Markov property for solutions with our Stratonovich-type differentials.

THEOREM 5.1. – *Let f and $f'f$ be globally Lipschitz, let Z be a Lévy process, and let X_0 be independent of Z. Then the solution X of*

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

is strong Markov.

Proof. – We rewrite equation (5.2) as in (2.5) as:

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

Note that $[Z,Z]_t^c = \alpha t$ for some constant α because Z is a Lévy process (*see*, e.g., Theorem V.33 of Protter [17], p. 239); thus $[Z,Z]^c$ is trivially also a Lévy process. One easily verifies that $[Z]^d$ 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 càglàd processes; however, the differential for h, $d[Z]_s^d$, 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 ([17], 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

Wong and Zakai [18] consider differentiable approximations of Brownian motion and show that the solutions of ordinary differential equations driven by these smooth approximants converge to the solution of an analogous Stratonovich-type stochastic differential equation driven by the Brownian motion, and not to the solution of the corresponding Itô-type equation. Their result has undergone many generalizations, culminating in Kurtz-Protter [9], where the Brownian differentials are replaced by general semimartingales. In Kurtz-Protter [9], 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 limiting equation is then of the type introduced above. This result 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 the case where Z is a one-dimensional semimartingale. A generalization to systems of equations driven by vector-valued semimartingales is possible with appropriate assumptions.

We define the approximating semimartingales 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 of finite variation on compacts. Moreover $\lim_{\substack{h\to 0\\ h>0}} Z_t^h = Z_{t-}$ a.s., each $t > 0$.

Let $f: \mathbb{R} \to \mathbb{R}$ and f' be bounded and Lipschitz continuous, and let X^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^h converges, in an appropriate sense, to the solution X of the equation

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

introduced above. Note that X^h is continuous while X may have discontinuities, so convergence in the Skorohod topology will not, in general, hold. The type of convergence we will establish is that studied in Kurtz [8]. In particular, we will show the existence of a sequence of time transformations τ_h for which $Y^h(t) = X^h_{\tau_h(t)}$ satisfies $(Y^h, \tau_h) \to (Y, \tau)$ in the compact uniform topology and $X(t) = Y_{\tau^{-1}(t)}$.

The new time scale we introduce includes the "fictitious time" during which the solution follows the vector field f to form the jump. Define

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

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

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

which is also strictly increasing and adapted, although not continuous. Note that

$$\lim_{\substack{h\to 0\\ h>0}} \gamma_h(t) = \gamma_0(t-).$$

For each h, the desired time change is given by the continuous inverse

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

Then for all $h \geq 0$, $\gamma_h^{-1}(t)$ is a stopping time for each t and γ_h^{-1} is continuous. For $h > 0$, γ_h^{-1} is strictly increasing and, hence, is the inverse of γ_h. Note that each discontinuity $\Delta Z(t)$ of Z corresponds to an interval of length $|\Delta Z(t)|^2$ on which γ_0^{-1} is constant. Note also that $\gamma_0^{-1} \circ \gamma_0(t) = t$ and $\gamma_0 \circ \gamma_0^{-1}(t) \geq t$ for all $t \geq 0$.

The time-changed driving process

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

is continuous and has paths of finite variation on compacts, since Z^h does. The time-changed solution

$$Y_t^h = X^h_{\gamma_h^{-1}(t)} \tag{6.5}$$

is then the unique solution of

$$Y_t^h = X_0 + \int_0^t f(Y_s^h) dV_s^h. \tag{6.6}$$

We next establish several preliminary results.

LEMMA 6.1. – *For each $h > 0$ and $t \geq 0$, $\gamma_h^{-1}(t) - h < \gamma_0^{-1}(t) < \gamma_h^{-1}(t)$, and hence,* $\lim_{h\to 0} \gamma_h^{-1}(t) = \gamma_0^{-1}(t)$, *uniformly in t.*

Proof. – The lemma follows from the observation that $\gamma_h(t) < \gamma_0(t) < \gamma_h(t+h)$. □

To identify the limit of the processes V^h 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) $V_t =$

$$\begin{cases} Z_{\gamma_0^{-1}(t)} & \text{if } \eta_1(t) = \eta_2(t) \\ Z_{\gamma_0^{-1}(t)}\left(\dfrac{t-\eta_1(t)}{\eta_2(t)-\eta_1(t)}\right) + Z_{\gamma_0^{-1}(t)-}\left(\dfrac{\eta_2(t)-t}{\eta_2(t)-\eta_1(t)}\right) & \text{if } \eta_1(t) \neq \eta_2(t) \end{cases}$$

Note that $\eta_1(t) = \eta_2(t)$ unless Z has a discontinuity at $\gamma_0^{-1}(t)$ and that γ_0^{-1} is constant on $[\eta_1(t), \eta_2(t)]$. V is the semimartingale Z time–changed according to γ_0^{-1}, except when Z jumps. At the jump-times of Z, we add "imaginary" time intervals $[\eta_1(t), \eta_2(t)]$ of length $|\Delta Z|^2$. During these intervals V is defined by linear interpolation over the discontinuity of Z.

Note that if $\sum_{0<s\leq t} |\Delta Z_s| < \infty$ a.s., for each $t > 0$, then it is clear that V 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 need not be a semimartingale. In all cases, however, V is a continuous process adapted to the filtration $\mathcal{G}_t = \mathcal{F}_{\gamma_0^{-1}(t)}$.

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

Proof. – We need to show that $t_h \geq 0$ and $t_h \to t$ imply $V^h_{t_h} \to V_t$. If Z is continuous at $\gamma_0^{-1}(t)$, that is, if $\eta_1(t) = \eta_2(t)$, the limit will hold by Lemma 6.1. Assume that Z has a discontinuity at $\gamma_0^{-1}(t)$ or equivalently, that $\eta_1(t) \neq \eta_2(t)$. By (6.1), along a subsequence satisfying $\gamma_h^{-1}(t_h) \leq \gamma_0^{-1}(t)$, $V^h_{t_h} \to V_t = Z_{\gamma_0^{-1}(t)-}$. In particular, $V^h_{\gamma_h \circ \gamma_0^{-1}(t)} \to Z_{\gamma_0^{-1}(t)-}$ and $\gamma_h \circ \gamma_0^{-1}(t) \to \eta_1(t)$. Along a subsequence satisfying $\gamma_h^{-1}(t_h) - h \geq \gamma_0^{-1}(t)$, $V^h_{t_h} \to V_t = Z_{\gamma_0^{-1}(t)}$. Note also, that $\gamma_h(\gamma_0^{-1}(t) + h) \to \eta_2(t)$.

Observe that

$$\frac{d}{du}\gamma_h^{-1}(u) = \frac{h}{[Z]^d_{\gamma_h^{-1}(u)} - [Z]^d_{\gamma_h^{-1}(u)-h} + h} \tag{6.8}$$

and

$$\frac{d}{du}V_u^h = \frac{Z_{\gamma_h^{-1}(u)} - Z_{\gamma_h^{-1}(u)-h}}{[Z]^d_{\gamma_h^{-1}(u)} - [Z]^d_{\gamma_h^{-1}(u)-h} + h}. \tag{6.9}$$

It follows that

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

uniformly in u satisfying $\gamma_0^{-1}(t) < \gamma_h^{-1}(u) < \gamma_0^{-1}(t) + h$ which is the derivative of V_t in $[\eta_1(t), \eta_2(t)]$. Consequently, along any subsequence satisfying $\gamma_0^{-1}(t) < \gamma_h^{-1}(t_h) < \gamma_0^{-1}(t) + h$, $(t_h - \gamma_h \circ \gamma_0^{-1}(t)) \to (t - \eta_1(t))$ and hence

$$\lim_{h\to 0} V_{t_h}^h = Z_{\gamma_0^{-1}(t)-} + (t - \eta_1(t))\frac{Z_{\gamma_0^{-1}(t)} - Z_{\gamma_0^{-1}(t)-}}{\eta_2(t) - \eta_1(t)} = V_t.$$

□

Before continuing we need to introduce a concept from Kurtz-Protter [10].

DEFINITION. – For each n let Z^n be a semimartingale with respect to a filtration $\{\mathcal{F}_t^n\}$, and suppose that Z^n converges in distribution in the Skorohod topology to a process Z. Then the sequence $(Z^n, \{\mathcal{F}_t^n\})$ is said to be *good* if Z is a semimartingale and for any H^n, càdlàg and adapted to $\{\mathcal{F}_t^n\}$, such that (H^n, Z^n) converges in distribution to (H, Z) in the Skorohod topology, $\int_0^\cdot H_{s-}^n dZ_s^n$ converges in distribution in the Skorohod topology to $\int_0^\cdot H_{s-} dZ_s$.

A necessary and sufficient condition for a convergent sequence of semimartingales Z^n to be good was obtained in Kurtz-Protter [9,10]. Let $h_\delta(r) = (1 - \delta/r)^+$, and define $J_\delta\colon D[0,\infty) \to D[0,\infty)$ by

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

Let

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

Then $Z^{n,\delta}$ has jumps bounded by δ. 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 for "goodness" of the sequence $\{Z_n\}$ is that for each n, there exist such decompositions satisfying

($\star$) For each $\alpha > 0$, there exist stopping times T_n^α such that

$$P(T_n^\alpha \leq \alpha) \leq 1/\alpha$$

and

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

Note that if $Z^n = Z$ for each n, then the sequence is good. Furthermore, since $\gamma_h^{-1}(t) \leq t$ for all t and h, the sequence $\{(Z_{\gamma_h^{-1}}, \{\mathcal{F}_{\gamma_h^{-1}(t)}\}), h > 0\}$ is good. If $\{V^h\}$ were good, then we could apply Theorem 5.4 of Kurtz and Protter [9] to conclude that the solution of (6.5) converges. Unfortunately, not only is $\{V^h\}$ not in general a good sequence, the limit V is not in general a semimartingale. To address this problem, we first define

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

and rewrite (6.5) as

$$Y_t^h = X_0 + \int_0^t f(Y_s^h) dZ_{\gamma_h^{-1}(s)} + \int_0^t f(Y_s^h) dU_s^h \tag{6.11}$$

Note that by Lemma 6.2 and the fact that $\gamma_h^{-1}(t)$ converges to $\gamma_0^{-1}(t)$ from above, we have

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

In addition, the convergence of U^h to U is in the Skorohod topology. Following the general approach to Wong-Zakai-type theorems taken in Kurtz and Protter [9], we integrate the last term by parts to obtain

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

Since Y^h is continuous and of finite variation and f is $\mathcal{C}^1$, $[f(Y^h),\ U^h] = 0$. Consequently,

$$\begin{aligned} \int_0^t f(Y_s^h) dU_s^h &= f(Y_t^h) U_t^h - \int_0^t U_s^h f'(Y_s^h) dY_s^h \\ &= f(Y_t^h) U_t^h - \int_0^t f'(Y_s^h) f(Y_s^h) U_s^h dV_s^h \end{aligned} \tag{6.12}$$

With the last term in (6.12) in mind, we prove the following lemma.

LEMMA 6.3. – *The sequence of semimartingales* $A_t^h = \int_0^t U_s^h dV_s^h$ *is good, and*

$$\int_0^t U_s^h dV_s^h \to A_t^0 \equiv \frac{1}{2}\{(V_t - Z_{\gamma_0^{-1}(t)})^2 - [Z]_{\gamma_0^{-1}(t)}\}$$

where the convergence is in probability in the Skorohod topology.

Proof. – A_t^h can be written

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

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

and substituting $u = \gamma_h^{-1}(s)$, we have [using (6.8)]

$$(6.13) \quad \int_0^t U_s^h dV_s^h = \int_0^{\gamma_h^{-1}(t)} \frac{\left(\frac{1}{h}\int_{u-h}^{u} Z_r dr - Z_u)(Z_u - Z_{u-h}\right)}{h} du.$$

It is easy to see that if a sequence of semimartingales $(Z^n)_{n\geq 1}$ defined on the same space and converging in probability (not just in distribution) 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 (*see* Dellacherie and Meyer [3], p. 251), 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.

To verify goodness of the sequence in (6.13), we estimate the total variation.

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

$$\leq 2E\left\{\int_0^T \frac{\sup_{u-h\leq s\leq u}|Z_s - Z_{u-h}|^2}{h} du\right\}$$

$$\leq 4E\left\{\int_0^T \frac{\sup_{u-h\leq s\leq u}|M_s - M_{u-h}|^2}{h} du\right\}$$

$$+ 4E\left\{\int_0^T \frac{(|A|_u - |A|_{u-h})^2}{h} du\right\}$$

$$\leq \frac{16}{h}E\left\{\int_0^T ([M]_u - [M]_{u-h})du\right\}$$

$$+ \frac{4}{h}E\left\{\int_0^T (|A|_u^2 - |A|_{u-h}^2)du\right\}$$

$$\leq 16E\{[M]_T\} + 8E\{|A|_T^2\}$$

giving a bound on the expected total variation that is independent of h. We conclude that $(\star)$ is satisfied, and hence we have goodness.

To identify the limit of $\int U_s^h dV_s^h$, we use integration by parts to obtain

$$\int_0^t U_s^h dV_s^h = \int_0^t U_{s-}^h dU_s^h + \int_0^t U_{s-}^h dZ_{\gamma_h^{-1}(s)}$$

$$= \frac{1}{2}((U_t^h)^2 - [Z]_{\gamma_h^{-1}(t)}) + \int_0^t U_{s-}^h dZ_{\gamma_h^{-1}(s)}.$$

By the definition of U^h and Lemma 6.2, the right side converges to

$$\frac{1}{2}\{(V_t - Z_{\gamma_0^{-1}(t)})^2 - [Z]_{\gamma_0^{-1}(t)}\} + \int_0^t U_{s-} dZ_{\gamma_0^{-1}(s)}. \tag{6.14}$$

Since U vanishes off of the intervals on which γ_0^{-1} is constant, the last term in (6.14) vanishes, completing the proof of the lemma. □

Finally, we need to show that the first term on the right of (6.12) is relatively compact.

LEMMA 6.4. – *The sequence* $\{f(Y^h)U^h\}$ *is relatively compact in the sense of convergence in distribution in the Skorohod topology.*

Proof. – To show relative compactness, it is sufficient to show that every subsequence has a further subsequence that converges. Let $\{\tau_i\}$ be the times at which Z has a jump. The boundedness of f and f', the "goodness" of $\{Z_{\gamma_h^{-1}}\}$ and $\{A^h\}$, and the fact that U^h converges, ensure that $\sup_{t\leq T} Y_t^h$ is stochastically bounded. For any sequence $\{h_n\}$, $h_n \to 0$, there will be a further subsequence along which $(Y^h_{\gamma_h(\tau_1)}, Y^h_{\gamma_h(\tau_2)}, \ldots)$ converges in distribution in $\mathbb{R}^\infty$. Denote the limit by $(Y_1, Y_2, \ldots)$. For $\gamma_h(\tau_i) \leq s < \gamma_h(\tau_i + h)$

$$|\dot{V}_s^h| \leq \frac{1}{|\Delta Z_{\tau_i}|^2} \sup_{\tau_i - h\leq u,v\leq \tau_i+h} |Z_u - Z_v|$$

and

$$\sup_{\gamma_h(\tau_i)\leq s<\gamma_h(\tau_i+h)} \left| Y_s^h - Y^h_{\gamma_h(\tau_i)} - \int_{\gamma_h(\tau_i)}^s f(Y_u^h)\frac{\Delta Z_{\tau_i}}{|\Delta Z_{\tau_i}|^2}du \right| \to 0.$$

It follows that

$$\sup_{\gamma_h(\tau_i)\leq s<\gamma_h(\tau_i+h)} \left| Y_s^h - \varphi\left(f\Delta Z_{\tau_i}, Y^h_{\gamma_h(\tau_i)}, \frac{s-\gamma_h(\tau_i)}{|\Delta Z_{\tau_i}|^2}\right)\right| \to 0,$$

where φ is defined in Section 2. Let $\Gamma_h = \bigcup_{i=1}^{\infty} [\gamma_h(\tau_i), \gamma_h(\tau_i + h))$. Then for each $T > 0$, $\lim_{h\to 0} \sup_{t\in[0,T]-\Gamma_h} |U_t^h| = 0$. Noting that $\gamma_h(\tau_i) \to \gamma_0(\tau_i-) = \eta_1(\gamma_0(\tau_i))$ and $\gamma_h(\tau_i + h) \to \gamma_0(\tau_i) = \eta_2(\gamma_0(\tau_i))$, it follows that along the subsequence, $f(Y^h)U^h \Rightarrow R$, where

$$(6.15)\quad R(t) = \sum_i I_{[\gamma_0(\tau_i-),\gamma_0(\tau_i))}(t) f\left(\varphi\left(f\Delta Z_{\tau_i}, Y_i, \frac{t-\gamma_0(\tau_i-)}{|\Delta Z_{\tau_i}|^2}\right)\right)U_t$$

completing the proof of the lemma. □

We can now apply Theorem 5.4 of Kurtz and Protter [9] to conclude that $\{Y^h\}$ is relatively compact in the sense of convergence in distribution in the Skorohod topology and that any limit point must satisfy

$$(6.16)\qquad Y_t = X_0 + R(t,Y) + \int_0^t f(Y_s)dZ_{\gamma_0^{-1}(s)} - \int_0^t f'f(Y_s)dA_s^0$$

where $R(t, Y)$ is given by (6.15) and A^0 is defined in Lemma 6.3.

Substituting for A^0 in (6.16) and writing $[Z] = [Z]^d + [Z]^c$, we obtain

$$(6.17)\quad \begin{aligned} Y_t =& X_0 + R(t,Y) + \int_0^t f(Y_s)dZ_{\gamma_0^{-1}(s)} + \frac{1}{2}\int_0^t f'f(Y_s)d[Z]^c_{\gamma_0^{-1}(s)} \\ & - \frac{1}{2}\int_0^t f'f(Y_s)d(U_s^2 - [Z]^d_{\gamma_0^{-1}(s)}). \end{aligned}$$

Recall that $U_t = 0$ unless $\eta_1(t) \neq \eta_2(t)$ and that if $\eta_1(t) \neq \eta_2(t)$,

$$(6.18)\qquad U_t = -\Delta Z_{\gamma_0^{-1}(t)} \frac{\eta_2(t) - t}{\eta_2(t) - \eta_1(t)}.$$

Note that $-\frac{1}{2}(U^2 - [Z]^d)$ is absolutely continuous and nondecreasing, and its derivative is

$$\sum_i I_{[\gamma_0(\tau_i-),\gamma_0(\tau_i))}(t)\frac{\gamma_0(\tau_i) - t}{|\Delta Z_{\tau_i}|^2}.$$

With τ_i as in (6.15) and $Y_i = Y_{\gamma_0(\tau_i-)}$, we have for $t \in [\gamma_0(\tau_i-), \gamma_0(\tau_i))$

$$(6.19)\quad \begin{aligned} Y_t =& Y_i + f\left(\varphi\left(f\Delta Z_{\tau_i}, Y_i, \frac{t - \gamma_0(\tau_i-)}{|\Delta Z_{\tau_i}|^2}\right)\right)U_t + f(Y_i)\Delta Z_{\tau_i} \\ & + \int_{\gamma_0(\tau_i-)}^t f'f(Y_s)\frac{\gamma_0(\tau_i) - s}{|\Delta Z_{\tau_i}|^2}ds \end{aligned}$$

Observe that the solution of (6.19) on the interval $[\gamma_0(\tau_i-), \gamma_0(\tau_i))$ is unique given Y_i, and differentiating, it is easy to check that $Y_t = \varphi\left(f\Delta Z_{\tau_i}, Y_i, \frac{t - \gamma_0(\tau_i-)}{|\Delta Z_{\tau_i}|^2}\right)$. It follows that

$$(6.20)\quad \begin{aligned} Y_t =& X_0 + \sum_i I_{[\gamma_0(\tau_i-),\infty)}(t) \\ & \times \left(\varphi\left(f\Delta Z_{\tau_i}, Y_{\gamma_0(\tau_i-)}, \frac{t \wedge \gamma_0(\tau_i) - \gamma_0(\tau_i-)}{|\Delta Z_{\tau_i}|^2}\right)\right. \\ & \left. - Y_{\gamma_0(\tau_i-)} - f(Y_{\gamma_0(\tau_i-)})\Delta Z_{\tau_i}\right) \\ & + \int_0^t f(Y_s)dZ_{\gamma_0^{-1}(s)} + \frac{1}{2}\int_0^t f'f(Y_s)d[Z]^c_{\gamma_0^{-1}(s)} \end{aligned}$$

THEOREM 6.5. – *Let Z be a semimartingale, and let f, f', and $f'f$ be globally Lipschitz. For $h > 0$, define $Z_t^h = \frac{1}{h}\int_{t-h}^t Z_s ds$, and let X^h and X satisfy*

$$\text{(6.21)} \qquad \begin{cases} X_t^h = X_0 + \int_0^t f(X_s^h)dZ_s^h \\ X_t = X_0 + \int_0^t f(X_s) \circ dZ_s. \end{cases}$$

Let γ_h^{-1} and γ_0^{-1} be defined as above and define $Y_t^h = X^h_{\gamma_h^{-1}(t)}$. Then Y^h converges in probability in the compact uniform topology to a process Y such that $X_t = Y_{\gamma_0(t)}$. For all but countably many $t > 0$, $X_t^h \to X_t$ in probability.

Proof. – Assume that f and f' are bounded and globally Lipschitz. (The boundedness assumption can be removed by a localization argument.) As noted above, Y^h converges in distribution to the solution of (6.17) and, equivalently, (6.20). Note however that we have strong local uniqueness of the solution of (6.20). This can be strengthened to convergence in probability as follows (here we follow Corollary 5.6 of Kurtz and Protter [9]). (Since Y^h and Y are continuous we need not bother with the Skorohod topology.) Let F be a bounded continuous function from $C(\mathbb{R}_+; \mathbb{R})$ to $\mathbb{R}$, and let G be one mapping $C(\mathbb{R}_+, \mathbb{R}^2)$ to $\mathbb{R}$. Recall:

$$\begin{aligned} Y_t^h &= X_0 + \int_0^t f(Y_s^h)dV_s^h \\ &= X_0 + \int_0^t f(Y_s^h)dZ_{\gamma_h^{-1}}(s) + \int_0^t f(Y_s^h)dU_s^h \\ &= X_0 + f(Y_t^h)U_t^h + \int_0^t f(Y_s^h)dZ_{\gamma_h^{-1}}(s) + \int_0^t f'(Y_s^h)f(Y_s^h)dA_s^h. \end{aligned}$$

We have seen that $(Y^h, Z_{\gamma_h^{-1}},\ A^h,\ f(Y^h)U^h)$ converges in distribution to $(Y,\ Z_{\gamma_0^{-1}},\ A^0,\ R(Y))$, therefore

$$\lim_{h\to 0} E\{F(Y^h,\ f(Y^h)U^h)G(Z_{\gamma_h^{-1}},\ A^h)] = E\{F(Y,R)G(Z_{\gamma_0^{-1}},\ A^0)\}.$$

The convergence in probability of $(Z_{\gamma_h^{-1}},\ A^h)$ then implies

$$\text{(6.22)} \qquad \begin{aligned} &\lim_{h\to 0} E\{F(Y^h,\ f(Y^h)U^h)G(Z_{\gamma_0^{-1}},\ A^0)\} \\ &\quad = E\{F(Y,R)G(Z_{\gamma_0^{-1}},\ A^0)\}. \end{aligned}$$

Then L^1 approximation of measurable functions by continuous functions implies that (6.22) holds for all bounded, measurable G. It then follows by strong local uniqueness that there exists a bounded measurable G such that $F(Y,R) = G(Z_{\gamma_0^{-1}},\ A^0)$ a.s.

Therefore we can write simply

$$\lim_{h\to 0} E\{(F(Y^h, f(Y^h)U^h) - F(Y,R))^2] \\ = \lim_{h\to 0}(E\{F(Y^h, f(Y^h)U^h)^2\} \\ - 2E\{F(Y^h, f(Y^h)U^h)F(Y,R)\} + E\{F(Y,R)^2\}) = 0$$

and convergence in probability for Y^h follows.

Since Y^h and Y are continuous, convergence in the Skorohod topology is equivalent to convergence in the compact uniform topology. By this convergence and the continuity of Y, $X_t^h = Y^h_{\gamma_h(t)} \to Y_{\gamma_0(t)} = X_t$ at every point t at which γ_0 is continuous. □

ACKNOWLEDGMENT

We wish to thank Jean Jacod for several suggestions that greatly improved the paper, especially Section 6.

REFERENCES

[1] E. Çinlar, J. Jacod, P. Protter and M. Sharpe, Semimartingales and Markov Processes, *Z. für Wahrscheinlichkeitstheorie*, Vol. **54**, 1980, pp. 161-220.
[2] S. Cohen, Géométrie différentielle stochastique avec sauts, *C. R. Acad. Sci. Paris*, Vol. **314**, 1992, pp. 767-770.
[3] C. Dellacherie and P. A. Meyer, *Probabilities and Potential B*, North Holland, Amsterdam, 1982.
[4] A. Estrade, Exponentielle stochastique et intégrale multiplicative discontinue, to appear in *Ann. Inst. H. Poincaré*, 1992.
[5] T. Fujiwara, Stochastic differential equations of jump type on manifolds and Lévy flows, *J. Math. Kyoto Univ.*, Vol. **31**, 1991, pp. 99-119.
[6] J. Jacod and P. Protter, Une remarque sur les équations différentielles stochastiques à solutions Markoviennes, in *Séminaire de Probabilités XXV, Springer Lect. Notes in Math.*, **1485**, 1991, pp. 138-139.
[7] M. W. Hirsch, *Differential Topology*, Springer-Verlag, New York, Berlin, 1976.
[8] T. G. Kurtz, Random time changes and convergence in distribution under the Meyer-Zheng conditions, *Annals of Probability*, Vol. **19**, 1991, pp. 1010-1034.
[9] T. G. Kurtz and P. Protter, Weak limit theorems for stochastic integrals and stochastic differential equations, *Annals of Probability*, Vol. **19**, 1991, pp. 1035-1070.
[10] T. G. Kurtz and P. Protter, Characterizing the weak convergence of stochastic integrals, in *Stochastic Analysis*, M. Barlow and N. Bingham, Eds., 1991, pp. 255-259.

[11] T. G. KURTZ and P. PROTTER, Wong-Zakai corrections, random evolutions, and simulation schemes for SDE's, *Stochastic Analysis: Liber Amicorum for Moshe Zakai*, Academic Press, San Diego, 1991, pp. 331-346.

[12] H. J. KUSHNER, Jump-diffusion approximations for ordinary differential equations with wide-band random right hand sides, *SIAM J. Control*, Vol. **17**, 1979, pp. 729-744.

[13] S. I. MARCUS, Modeling and analysis of stochastic differential equations driven by point processes, *IEEE Transactions on Information Theory*, Vol. **24**, 1978, pp. 164-172.

[14] S. I. MARCUS, Modeling and approximation of stochastic differential equations driven by semimartingales, *Stochastics*, Vol. **4**, 1981, pp. 223-245.

[15] M. MÉTIVIER, *Semimartingales: a course on stochastic processes*, de Gruyter, Berlin, New York, 1982.

[16] 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.

[17] P. PROTTER, *Stochastic integration and differential equations: a new approach*, Springer-Verlag, Berlin, New York, 1990.

[18] E. WONG and M. ZAKAI, On the convergence of ordinary integrals to stochastic integrals, *Ann. Math. Statist.*, Vol. **36**, 1965, pp. 1560-1564.