WEAK CONVERGENCE IN STOCHASTIC ANALYSIS

by

Philip Protter *
Purdue University

Technical Report #89-36

Department of Statistics
Purdue University

October 1989

* Supported in part by NSF Grant No. DMS–8805595

*

WEAK CONVERGENCE IN STOCHASTIC ANALYSIS

Philip Protter *
Mathematics and Statistics Departments
Purdue University
W. Lafayette, Indiana 47907 USA

In this article we discuss the weak convergence of stochastic integrals and of solutions of stochastic differential equations. The differentials are semimartingales, which have been shown by the Bichteler–Dellacherie theorem to be the most general reasonable stochastic differentials possible. See Protter [4] for all background on semimartingales, stochastic integration and differential equations, as well as for an exposition of the Bichteler–Dellacherie theorem. All results in this article were obtained jointly with Tom G. Kurtz, and proofs and details can be found in Kurtz–Protter [2].

Naively, one would like a theorem that if a sequence of semimartingales Y_n converges weakly to Y, and if predictable processes X_n converge weakly to X, then Y is a semimartingale and $\int X_n dY_n$ converges to $\int X dY$. However one immediately runs into the problem of interpreting the weak convergence of X_n to X, since one cannot use the reasonable path spaces $\mathcal{C}$ or $\mathbb{D}$ (continuous and right continuous, left–limited functions, respectively; we call the latter <u>cadlag</u> henceforth, after its French acronym). One approach to this problem has been taken by Pages [3], but it is not useful for applications. Here we follow the approach of Jakubowski, Mémin, and Pages [1] and restrict our attention to integrands X_n which are adapted, cadlag processes. We then seek hypotheses such that if (X_n, Y_n) are defined on a sequence of spaces $(\Omega_n, \mathcal{F}_n, P_n)$ with X_n adapted, cadlag, and Y_n semimartingales, and such that if $(X_n, Y_n) \Rightarrow (X, Y)$, then $\int X_n(s-)dY_n(s) \Rightarrow \int X(s-)dY(s)$, where the notation "$\Rightarrow$" denotes weak convergence of the distribution measures on path space.

We begin with a technical definition. Fix an integer m and let $\mathbb{D}_{\mathbb{R}^m}[0,\infty)$ denote the space of cadlag functions mapping $[0,\infty)$ to $\mathbb{R}^m$. Let x denote a generic element of $\mathbb{D}_{\mathbb{R}^m}[0,\infty)$ and define

* Supported in part by NSF Grant No. DMS–8805595

$h_\delta : [0,\infty) \to [0,\infty)$ by $h_\delta(r) = (1-\delta/r)^+$. Define $J_\delta \colon D_{\mathbb{R}^m}[0,\infty) \to D_{\mathbb{R}^m}[0,\infty)$ by

$$J_\delta(x)(t) = \sum_{s\leq t} h_\delta(|x(s)-x(s-)|)(x(s)-x(s-))$$

It is easy to check that $x \to J_\delta(x)$ and $x \to x - J_\delta(x)$ are continuous in the Skorohod topology, and further that if $(x_n, y_n) \to (x, y)$ in the Skorohod topology, then so also does

$$\int_o^\cdot x_n(s-)dJ_\delta(y_n)(s) \to \int_0^\cdot x(s-)dJ_\delta(y)(s).$$

Let us next recall the classical definition of a semimartingale. Let $(\mathcal{F}_t)_{t\geq 0}$ be an increasing sequence of σ–algebras. A cadlag, $\{\mathcal{F}_t\}$–adapted process Y is a semimartingale if it can be decomposed as $Y = M + A$ where M is an $\{\mathcal{F}_t\}$–local martingale and the sample paths of A have finite variation on bounded time intervals, that is, there exists a sequence of $\{\mathcal{F}_t\}$–stopping times, τ_k, such that $\tau_k \to \infty$ a.s and for each $k, M^{\tau_k} \equiv M(\cdot \wedge \tau_k)$ is a uniformly integrable martingale, and for every $t > 0$, $T_t(A) = \sup \Sigma|A(t_{i+1}) - A(t_i)| < \infty$ a.s (where the supremum is over partitions of $[0,t]$).

An $\mathbb{R}^m$–valued process is an $\{\mathcal{F}_t\}$–semimartingale, if each component is a semimartingale. Let $\mathbb{M}^{km}$ denote the real–valued, $k \times m$ matrices. Throughout, $\int XdY$ will denote $\int X(s-)dY(s)$. Every semimartingale Y has associated to it an increasing, right continuous, adapted process $[Y]$, known as the quadratic variation process of Y. This process is often easy to calculate, or to estimate. Indeed one has the following facts concerning $[Y]$ (see [4] for proofs):

(1) let τ_n be a refining sequence of partitions of $[0,t]$ with $\lim_n$ mesh $(\tau_n) = 0$. Then

$$\lim_{n\to\infty} \sum_{t_i\in\tau_n} (Y_{t_i} - Y_{t_{i-1}})^2 = [Y]_t$$

with convergence in probability;

(2) $Y_t^2 - 2\int_0^t Y(s-)dY(s) = [Y]_t$, with the convention $Y(0-) = 0$;

(3) if $Y = M$ is a local martingale, and if $E\{[M]_t\} < \infty$, then M is a martingale and $E\{M_t^2\} = E\{[M]_t\}$;

(4) if M is a local martingale, then $M_t^2 - [M]_t$ is also a local martingale;

(5) if Y is a continuous semimartingale with (unique) decomposition $Y = M+A$, with $A_0 = 0$, then $[Y] = [M]$;

(6) if $Z = \int X(s-)dY(s)$, then $[Z] = \int X(s-)^2 d[Y]_s$;

(7) if Z is the solution of

$$dZ = \sum_{i=1}^k f_i(Z(s-))dY^i(s),$$

for semimartingales $Y_i,\ 1 \leq i \leq k$, then

$$[Z] = \sum_{i,j} \int f_i(Z(s-)) f_j(Z(s-)) d[Y^i, Y^j]_s,$$

where $[Y^i, Y^j] = \frac{1}{2}\{[Y^i + Y^j] - [Y^i] - [Y^j]\}$; in particular, if $(B^1, \ldots, B^k)$ is an $\mathbb{R}^k$–Brownian motion, and if

$$dZ = \sum_{i=1}^{k} f_i(Z(s-)) dB^i(s),$$

then $[Z]_t = \sum_{i=1}^{k} \int_0^t f_i(Z(s-))^2 ds$, since $[B^i, B^j] = 0$ for $i \neq j$, and $[B^i]_t = t$.

In view of the preceding discussion, the next theorem is particularly useful for applications, since its hypotheses are usually easy to verify. (Another version of this theorem, with essentially equivalent hypotheses which are in a more obtuse form, can be found in [1]). For a finite variation process A let $T_t(A) = \int_0^t |dA_s|$ be the random variable denoting the total variation of the paths of A on $[0, t]$.

THEOREM 1. For each n, let (X_n, Y_n) be an $\{\mathcal{F}_t^n\}$–adapted process with sample paths in $D_{\mathbb{M}^{km}} \times \mathbb{R}^m[0, \infty)$, and let Y_n be an $\{\mathcal{F}_t^n\}$–semimartingale. Fix $\delta > 0$ (allowing $\delta = \infty$), and define $Y_n^\delta = Y_n - J_\delta(Y_n)$. (Note that Y_n^δ will also be a semimartingale.) Let $Y_n^\delta = M_n^\delta + A_n^\delta$ be a decomposition of Y_n^δ into an $\{\mathcal{F}_t^n\}$–local martingale and a process with finite variation. Suppose

(*) For each $\alpha > 0$, there exist stopping times $\{\tau_n^\alpha\}$ such that $P\{\tau_n^\alpha \leq \alpha\} \leq \frac{1}{\alpha}$ and $\sup_n E[[M_n^\delta]_{t\wedge\tau_n^\alpha} + T_{t\wedge\tau_n^\alpha}(A_n^\delta)] < \infty$.

If $(X_n, Y_n) \Rightarrow (X, Y)$ in the Skorohod topology on $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0, \infty)$, then Y is a semimartingale with respect to a filtration to which X and Y are adapted, and $(X_n, Y_n, \int X_n dY_n) \Rightarrow (X, Y, \int X dY)$ in the Skorohod topology on $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^k}[0, \infty)$. If $(X_n, Y_n) \to (X, Y)$ in probability, then the triple converges in probability.

<u>Remark</u> For $c > 0$, define $\tau_n^c = \inf\{t\colon |M_n^\delta(t)| \vee |M_n^\delta(t-)| \geq c \text{ or } T_t(A_n^\delta) \geq c\}$. Suppose the following conditions hold.

(i) $\{T_t(A_n^\delta)\}$ is stochastically bounded for each $t > 0$.

(ii) For each $c > 0$, $\sup_n E[M_n^\delta(t \wedge \tau_n^c)^2 + T_{t\wedge\tau_n^c}(A_n^\delta)] < \infty$.

Since $\sup_{t\leq\alpha}|M_n^\delta(t)| = \sup_{t\leq\alpha}|Y_n^\delta(t) - A_n^\delta(t)| \leq \sup_{t\leq\alpha}|Y_n(t)| + T_\alpha(A_n^\delta)$ is stochastically bounded in n for each α, there exists c_α so that $P\{\tau_n^{c_\alpha} \leq \alpha\} \leq \frac{1}{\alpha}$. In addition $E[[M_n^\delta]_{t\wedge\tau_n^{c_\alpha}}] = E[(M_n^\delta(t \wedge \tau_n^{c_\alpha})^2]$, and (*) is satisfied with $\tau_n^\alpha = \tau_n^{c_\alpha}$.

For $\delta < \infty$, (ii) above will usually be immediate since the discontinuities of Y_n^δ are bounded in magnitude by δ (making Y_n^δ a special semimartingale) and there will exist a decomposition with the discontinuities of each term bounded by 2δ.

We next turn our attention to stochastic differential equations. We wish to consider systems of equations of the form $(1 \leq i \leq k)$

$$X_t^i = U_t^i + \sum_{j=1}^{m} \int_0^t F_j^i(X, s-) dY_s^j$$

where U^i are cadlag, adapted processes, and F_j^i are coefficients which are functionals, allowing dependence on the past. Such equations arise in applications such as stochastic control theory. Natural hypotheses on a sequence F_n of coefficients (and a limiting coefficient F) might be:

(i) For each compact subset $\mathcal{X} \subset D_{\mathbb{R}^k}[0,\infty)$ and $t > 0$, $\sup_{x\in\mathcal{X}} \sup_{s\leq t} |F_n(x,s) - F(x,s)| \to 0$.

(**) (ii) For $\{x_n\}$ and x in $D_{\mathbb{R}^k}[0,\infty)$ and each $t > 0$, $\sup_{s\leq t}|x_n(s) - x(s)| \to 0$ implies $\sup_{s\leq t}|F(x_n, s) - F(x,s)| \to 0$.

(iii) For each compact subset $\mathcal{X} \subset D_{\mathbb{R}^k}[0,\infty)$ and $t > 0$, there exists a continuous function $\omega: [0,\infty) \to [0,\infty)$ with $\omega(0) = 0$ such that for all $\lambda \in \Lambda^1$, $\sup_{x\in\mathcal{X}} \sup_{s\leq t} |F(xo\lambda, s) - F(x, \lambda(s))| \leq \omega(\gamma(\lambda))$.

The above conditions imply that if (x_n, y_n) converges to (x,y) in the Skorohod topology, then so also $(x_n, y_n, F_n(x_n))$ converges to $(x, y, F(x))$. Nevertheless, for technical reasons, we need a slight strengthening of (**), involving properties of F_n and F under transformations of the time scale. Let $T_1[0,\infty)$ denote the collection of nondecreasing mappings λ of $[0,\infty)$ onto $[0,\infty)$ (in particular $\lambda(0) = 0$) such that $\lambda(t+h) - \lambda(t) \leq h$ for all $t,\ h \geq 0$. Let i denote the identity map $i(s) = s$. We will assume that there exist mappings $G_n,\ G: D_{\mathbb{R}^k}[0,\infty) \times T_1[0,\infty) \to D_{\mathbb{M}^{km}}[0,\infty)$ such that $F_n(x)o\lambda = G_n(xo\lambda, \lambda)$ and $F(x)o\lambda = G(xo\lambda, \lambda)$ for $(x,\lambda) \in D_{\mathbb{R}^k}[0,\infty) \times T_1[0,\infty)$. We further make the following assumptions on G_n and G:

(i) For each compact subset $\mathcal{X} \subset D_{\mathbb{R}^k}[0,\infty) \times T_1[0,\infty)$ and $t > 0$, $\sup_{(x,\lambda)\in\mathcal{X}} \sup_{s\leq t} |G_n(x,\lambda,s) - G(x,\lambda,s)| \to 0$.

(* * *)

(ii) For $\{(x_n, \lambda_n)\} \in D_{\mathbb{R}^k}[0,\infty) \times T_1[0,\infty)$, $\sup_{s\leq t}|x_n(s) - x(s)| \to 0$ and $\sup_{s\leq t}|\lambda_n(s) - \lambda(s)| \to 0$ for each $t > 0$ implies $\sup_{s\leq t}|G(x_n, \lambda_n, s) - G(x, \lambda, s)| \to 0$.

Note that (* * *) implies (**). Let M^{km} denote $k \times m$ real valued matrices.

<u>Examples</u> Let $g: \mathbb{R}^k \times [0,\infty) \to \mathbb{M}^{km}$ and $h: [0,\infty) \to [0,\infty)$ be continuous. The following functions satisfy (**) and have a representation in terms of a G satisfying (* * *):

a) $F(x,t) = g(x(t), t)$

b) $F(x,t) = \int_0^t h(t-s) g(x(s), s) ds$

For $k = m = 1$

c) $F(x,t) = \sup_{s\le t} h(t-s)g(x(s),s)$

d) $F(x,t) = \sup_{s\le t} h(t-s)g(x(s)-x(s-),s)$

Using vector and matrix notation, for $n = 1,2,\dots$ let $F_n: D_{\mathbb{R}^k}[0,\infty) \to D_{\mathbb{M}^{km}}[0,\infty)$, let U_n and Y_n be processes with sample paths in $D_{\mathbb{R}^k}[0,\infty)$ and $D_{\mathbb{R}^m}[0,\infty)$ respectively, adapted to a filtration $\{\mathcal{F}_t^n\}$. Suppose Y_n is a semimartingale and that F_n is nonanticipating in the sense that $F_n(x,t) = F_n(x^t,t)$ for all $t \ge 0$ and $x \in D_{\mathbb{R}_k}[0,\infty)$, where $x^t(\cdot) = x(\cdot \wedge t)$. Let X_n be adapted to $\{\mathcal{F}_t^n\}$ and satisfy

$$(\dagger) \qquad X_n(t) = U_n(t) + \int_0^t F_n(X_n, s-)dY_n(s),$$

and let X satisfy the limiting equation:

$$(\dagger\dagger) \qquad X(t) = U(t) + \int_0^t F(X, s-)dY(s).$$

We say that (X,τ) is a local solution of $(\dagger\dagger)$ above if there exists a filtration $\{\mathcal{F}_t\}$ to which X, U and Y are adapted, Y is an $\{\mathcal{F}_t\}$–semimartingale, τ is an $\{\mathcal{F}_t\}$–stopping time, and

$$X(t\wedge\tau) = U(t\wedge\tau) + \int_0^{t\wedge\tau} F(X, s-)dY(s).$$

We say that local uniqueness holds for $(\dagger\dagger)$ if any two local solutions (X_1,τ_1), (X_2,τ_2) satisfy $X_1(t) = X_2(t)$, $t \le \tau_1 \wedge \tau_2$, a.s. See Protter [4], Chapter V, for sufficient conditions for uniqueness.

The next theorem can be found, together with a proof, in Kurtz–Protter [2]. It improves upon results of Slominski [5].

THEOREM 2. Suppose that (U_n, X_n, Y_n) satisfies $(\dagger)$, $(U_n, Y_n) \Rightarrow (U, Y)$ in the Skorohod topology and that $\{Y_n\}$ satisfies $(*)$ for some $0 < \delta \le \infty$. Assume that $\{F_n\}$ and F have representations in terms of $\{G_n\}$ and G satisfying $(***)$. For $b > 0$, define $\eta_n^b = \inf\{t: |F_n(X_n,t)| \vee |F_n(X_n,t-)| \ge b\}$ and let X_n^b denote the solution of

$$X_n^b(t) = U_n(t) + \int_0^t 1_{[0,\eta_n^b)}(s-)F_n(X_n^b, s-)dY_n$$

that agrees with X_n on $[0,\eta_n^b)$. Then $\{(U_n, X_n^b, Y_n)\}$ is relatively compact and any limit point, (U, X^b, Y), gives a local solution (X^b,τ) of $(\dagger\dagger)$ with $\tau = \eta^c \equiv \inf\{t: |F(X^b,t)| \vee |F(X^b,t-)| \ge c\}$ for any $c < b$. If there exists a global solution X of $(\dagger\dagger)$ and local uniqueness holds, then $(U_n, X_n, Y_n) \Rightarrow (U, X, Y)$.

Applications of Theorems 1 and 2 to statistics, filtering theory, economics (finance theory), and M. Emery's martingale structure equation are given in Section 3 of Kurtz–Protter [2].

REFERENCES

1. Jakubowski, A., Mémin, J. and Pages, G. (1989). Convergence en loi des suites d'intégrales stochastiques sur l'espace D^1 de Skorokhod. *Probab. Th. Rel. Fields* **81**, 111–137.

2. Kurtz, T.G., Protter, P. (1989). Weak limit theorems for stochastic integrals and stochastic differential equations. To appear.

3. Pages, G. (1986). Un théorème de convergence fonctionnelle pour les intégrales stochastiques. Séminaire de Proba. XX, 572–611. Springer Lecture Notes in Math. 1204.

4. Protter, Philip (1990). *Stochastic Integration and Differential Equations: A New Approach.* Springer–Verlag, Berlin, Heidelberg, New York (in press: to appear February 1990).

5. Slominski, Leszek (1989). Stability of strong solutions of stochastic differential equations. *Stochastic Processes Appl.* **31**, 173-202.