Explicit form and robustness of martingale representations

Jean Jacod*, Sylvie Méléard† and Philip Protter‡

ABSTRACT

Stochastic integral representation of martingales has been undergoing a renaissance due to questions motivated by Stochastic Finance theory. In the Brownian case one usually has formulas (of differing degrees of exactness) for the predictable integrands. We extend some of these to Markov cases where one does not necessarily have stochastic integral representation of all martingales. Moreover we study various convergence questions that arise naturally from (for example) approximations of "price processes" via Euler schemes for solutions of stochastic differential equations. We obtain general results of the following type: let U, U^n be random variables with decompositions:

$$U = \alpha + \int_0^\infty \xi_s dX_s + N_\infty$$

$$U^n = \alpha_n + \int_0^\infty \xi^n_s dX^n_s + N^n_\infty$$

where X, N, X^n, N^n are martingales. If $X^n \to X$ and $U^n \to U$, when and how does $\xi^n \to \xi$?

1 Introduction

1) Consider a sequence X^n of square-integrable martingales, which converge to another square-integrable martingale X: this convergence may hold in a strong sense (as in $\mathbb{L}^2$) and all the X^n's and X are on the same probability space, or it may hold in the weak sense (convergence in law) and each X^n is defined on its own probability space. Let also Φ be a bounded continuous functional (say, on the Skorokhod space of all right continuous functions with left limits), and set $U^n = \Phi(X^n)$ and $U = \Phi(X)$, so that U^n converges to U. Suppose in addition that we have the martingale representation property for each X^n and for X, so we can write U^n and U as stochastic integrals as follows:

$$U^n \;=\; \alpha_n + \int_0^\infty \xi^n_s dX^n_s, \qquad U \;=\; \alpha + \int_0^\infty \xi_s dX_s, \tag{1.1}$$

*Laboratoire de Probabilités, Université Paris 6, 4 place Jussieu, 75252 Paris, France (CNRS UMR 7599)

†Université Paris 10, MODAL'X, JE 421, 200 av. de la République, 92000 Nanterre, France

‡Mathematics and Statistics Departments, Purdue University, W.Lafayette, IN 47907-1395, USA. Supported in part by ONR grant # N00014-96-1-0262 and NSF grant # 9401109-INT

where α_n and α are random variables measurable w.r.t. the relevant initial σ-fields, and ξ^n and ξ are predictable processes. Then an important theoretical problem is to find whether the sequence ξ^n converges in law, for a suitable topology, to ξ.

This problem has also much practical relevance. For example in financial mathematics, suppose that X models the price of a stock, and U is a claim based upon this stock, and for simplicity the riskless bond has constant price 1. If the model is complete and with no arbitrage opportunity, then X is a martingale under the unique risk neutral equivalent measure, the price of the claim is the expectation $E(U) = \alpha$ under this measure, and we have the martingale representation property w.r.t. X: then the process ξ in (1.1) is the so-called hedging strategy. Now, for computational purposes we might want to take a discrete time approximation for X: e.g. a binomial approximation X^n which thus converges in law to X, or an Euler approximation X^n which thus converges strongly to X when this process is the solution of a stochastic differential equation. If one also has the martingale representation property for the discrete time models (as is the case for the binomial approximation), then it is important to know whether the "approximate" hedging strategies ξ^n do converge in some sense to ξ. Such questions have been touched upon in [6] for example.

The above brief description immediately gives rise to two kinds of problems. The first one comes from the fact that the martingale representation property quite often does not hold: it holds under reasonably general conditions when the basic martingale X is continuous, but it is usually lost as soon as X as jumps, and in particular in the discrete time setting (except for the binomial model).

The second problem is to find an adequate topology for which the ξ^n's might converge. This is not obvious, because these processes have *a priori* no regularity in time (they are predictable, but otherwise neither right continuous nor left continuous in general).

2) To begin with, let us consider the first problem described above. Let X be a locally square-integrable martingale on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ having $\mathcal{F} = \bigvee_t \mathcal{F}_t$, and U be a square-integrable random variable. Using the theory of "stable" subspaces generated by a martingale (see Dellacherie and Meyer [5], or [15] or [9] for this fact, as well as for all results on martingales and stochastic integrals), we have the decomposition

$$U \;=\; \alpha + \int_0^\infty \xi_s dX_s + N_\infty, \tag{1.2}$$

where $\alpha = E(U|\mathcal{F}_0)$ and N is a square-integrable martingale (i.e. a martingale such that $\sup_t |N_t|$ is square-integrable), orthogonal to X and ξ is a predictable process, and this decomposition is unique up to null sets: it comes in fact from the (unique) decomposition of the square-integrable martingale $M_t = E(U|\mathcal{F}_t)$ as a stochastic integral w.r.t. X, plus an orthogonal term. Recall also that two locally square-integrable martingales M and N are *orthogonal* if their product MN is a local martingale, and this is denoted by $M \perp N$.

Observe that α and N are defined uniquely up to a P-null set, while ξ is defined uniquely up to a null set w.r.t. the following measure

$$Q_X(d\omega, dt) \;=\; P(d\omega) d\langle X, X\rangle_t(\omega) \tag{1.3}$$

on $\Omega \times \mathbb{R}_+$. Here, $\langle X, X \rangle$ denotes the "angle" (or predictable) bracket. We will denote the process ξ by $\xi(X, U)$, which is square-integrable w.r.t. Q_X.

Section 2 of this paper is devoted to finding an "explicit" expression for the process ξ above: first in the discrete time setting, where it is very simple; next in some Markovian situations, when U has the form $U = f(Y_T)$ for a fixed time T and an underlying Markov process Y and X is a locally square-integrable martingale on this Markov process. We thus extend the well known Clark-Haussmann formula, usually given for Brownian motion, in two directions: the Brownian motion is replaced by a rather general Markov process, and we do not assume the martingale representation property. But of course we are limited to variables U of the form $U = f(Y_T)$ or more generally of the form $U = f(Y_{T_1}, \ldots, Y_{T_k})$ for fixed times $T_1 < \ldots < T_k$.

Let us come back to the financial interpretation of (1.2): if the martingale representation property w.r.t. X does not hold, the variable N_∞ in (1.2) is in general not equal to 0. We are in the incomplete model case, and the process ξ is shown to be a risk minimizing strategy for hedging the claim U: see Föllmer and Sondermann [7].

3) Let us now turn to convergence results. To get an idea of what to expect as far as convergence results are concerned, here is a trivial special case: we have a sequence U^n of random variables tending to a limit U in $\mathbb{L}^2(P)$, and a fixed locally square-integrable martingale X. Writing M^n, α^n, ξ^n and N^n for the terms associated with U^n and X in (1.2), the three variables $\alpha^n - \alpha$, $\int_0^\infty (\xi_s^n - \xi_s) dX_s$ and $N_\infty^n - N_\infty$ are orthogonal in $\mathbb{L}^2(P)$ and add up to $U^n - U$, so they all go to 0 in $\mathbb{L}^2(P)$. Since the expected value of $\left(\int_0^\infty \eta_s dX_s\right)^2$ is $Q_X(\eta^2)$, we deduce in particular that

$$U^n \to^{\mathbb{L}^2(P)} U \qquad \Rightarrow \qquad \xi(X, U^n) \to^{\mathbb{L}^2(Q_X)} \xi(X, U). \tag{1.4}$$

This leads us to consider first the case where all locally square-integrable martingales X^n and X are defined on the same space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ with $\mathcal{F} = \bigvee_t \mathcal{F}_t$. The simplest result one can state in this direction is as follows:

Theorem A *Assume that X^n and X are locally square-integrable martingales on a filtered space, such that $\langle X^n - X, X^n - X \rangle_t \to 0$ in probability for all $t \in \mathbb{R}_+$, and that U^n converges to U in $\mathbb{L}^2(P)$. Then ξ^n converges to ξ in Q_X-measure.*

We also give a series of other results, which are more difficult to state, and which mainly concern discrete time approximations of a given martingale X, of various kinds: stepwise approximations, or Euler schemes when X is the solution of a stochastic differential equation. All these results are proved in Section 3.

4) Section 4 is devoted to weak convergence results. First, we take advantage of the explicit results of Section 2 in the Markov case to show that if X^n is the solution of the equation $dX_t^n = g_n(X_{t-}^n) dZ_t^n$ and X is the solution of a similar equation with g and Z, where Z^n and Z are Lévy processes, and if $g_n \to g$ and Z^n converges in law to Z, then under some mild additional assumptions the processes ξ^n converge to ξ for a suitable topology, when $U^n = f(X^n_T)$ and $U = f(X_T)$ and f is a differentiable function (typically Z is a Brownian motion, but the Z^n's are not, so we have the martingale representation property w.r.t. X, but not w.r.t. X^n). We also give a discrete time version of this result.

Finally, we give an analogous convergence result when $U^n = \Phi(X^n)$ and $U = \Phi(X)$ for a continuous bounded function Φ on the Skorokhod space, when X is the solution of an equation as above with Z a Brownian motion, and the X^n's are discrete time solutions of difference equations converging to X. As an example, particularly relevant in financial applications, let us mention the case where

$$\bar{X}^n_{i+1} \;=\; \bar{X}^n_i + g(\bar{X}^n_i)Y^n_{i+1},$$

where for each n the $(Y^n_i)_{i\geq 1}$ are i.i.d. bounded variables, centered with variance $1/n$. Then the processes $X^n_t = \bar{X}^n_{[nt]}$ converge in law to the solution of $dX_t = g(X_t)dW_t$, where W is a Brownian motion, as soon as g is Lipschitz. In this situation, with $U^n = \Phi(X^n)$ and $U = \Phi(X)$ with Φ as above, the processes ξ^n (naturally defined as some sort of interpolations of the discrete time processes $\xi(\bar{X}^n, U^n)$) do converge in a suitable sense to ξ, in law.

2 Explicit representations of the integrand

In this section our aim is to give an "explicit" form for the integrand $\xi(X,U)$ in essentially two specific cases: one is the discrete-time case, with an extension to the discretization of a continuous-time process; the other is a Markov situation. It seems hopeless to obtain such an explicit form in general, but other cases are found in the literature, essentially on the Wiener space and using Malliavin calculus: see e.g. the book of Nualart [14] and the references therein.

Before starting we wish to make precise the various notions of (locally) square-integrable martingales used in this paper, since they play a crucial role. As said in the introduction, a process X given on a stochastic basis, either with discrete or with continuous time, is called a square-integrable martingale if it is a martingale and if the supremum of X over all time is square-integrable: then the limit X_∞ exists and is a square-integrable variable. X is called a locally square-integrable martingale if there is a sequence R_n of stopping times increasing to $+\infty$, such that the process X stopped at any R_n is a square-integrable martingale. In between, we say that X is a *martingale square-integrable on compacts* if the process X stopped at any finite deterministic time is a square-integrable martingale: for example the Wiener process is a martingale square-integrable on compacts in this sense.

2.1 The discrete-time case

In this subsection, time is discrete: we have the basis $(\Omega, \mathcal{F}, (\mathcal{F}_i)_{i\in I\!N}, P)$ with $\mathcal{F} = \bigvee \mathcal{F}_i$ and with a given locally square-integrable martingale X. We also have a square-integrable variable U. For any process Y we write $\Delta Y_i = Y_i - Y_{i-1}$. In this discrete-time case, (1.2) becomes

$$U \;=\; \alpha + \sum_{i=1}^{\infty} \xi_i \Delta X_i + N_\infty, \tag{2.1}$$

where the series converges in $I\!L^2$ and ξ_i is $\mathcal{F}_{i-1}$-measurable and N is a square-integrable martingale orthogonal to X. Here the orthogonality of X and N amounts to say that

$$E(\Delta X_i \Delta N_i | \mathcal{F}_{i-1}) \;=\; 0, \qquad \forall\, i \geq 1. \tag{2.2}$$

The above conditional expectation is to be understood in the generalized sense, since the variable $\Delta X_i \Delta N_i$ might be not integrable: it is however integrable on each $\mathcal{F}_{i-1}$-measurable set $\{R_n \geq i\}$ (where R_n is as above), while $\cup_n \{R_n \geq i\} = \Omega$. The same comment applies below.

Proposition 2.1 *Assume that X is a locally square-integrable martingale, and let $M_i = E(U|\mathcal{F}_i)$. Then a version of $\xi = \xi(X,U)$ is given by*

$$\xi_i \;=\; \frac{E(\Delta X_i U|\mathcal{F}_{i-1})}{E((\Delta X_i)^2|\mathcal{F}_{i-1})} \;=\; \frac{E(\Delta X_i \Delta M_i|\mathcal{F}_{i-1})}{E((\Delta X_i)^2|\mathcal{F}_{i-1})}. \tag{2.3}$$

Proof. By definition of M and by the property $E(\Delta X_i|\mathcal{F}_{i-1}) = 0$ (where again the conditional expectation is in the generalized sense), the last equality in (2.3) is obvious. Define ξ by (2.3). The measurability condition is obviously met. Set

$$\Delta N_i \;=\; E(U|\mathcal{F}_i) - E(U|\mathcal{F}_{i-1}) - \xi_i \Delta X_i \;=\; \Delta M_i - \xi_i \Delta X_i$$

and $N_i = \sum_{j=1}^{i} \Delta N_j$. Then N is a square-integrable martingale with (2.2). That (2.1) holds is then obvious. $\square$

2.2 Discretization in time

Here we have a basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ such that $\mathcal{F} = \bigvee_t \mathcal{F}_t$. We consider a square-integrable martingale X. We also consider a locally finite subdivision τ of $I\!R_+$, consisting of an increasing sequence $\tau = (T_i : i \in I\!N)$ of stopping times such that

$$T_0 = 0, \qquad T_i < \infty \;\;\Rightarrow\;\; T_i < T_{i+1}, \qquad \lim_i T_i = \infty \quad \text{a.s.} \tag{2.4}$$

The discretized process is then

$$\bar{X}_i \;=\; X_{T_i} \qquad i \in I\!N, \tag{2.5}$$

which makes sense even on the set $\{T_i = \infty\}$. Then the sequence $(\bar{X}_i)_{i\in I\!N}$ is a square-integrable martingale w.r.t. the discrete-time filtration $(\mathcal{F}_{T_i})_{i\in I\!N}$. If $U \in I\!L^2$, we then have the two decompositions (1.2) and (2.2), namely

$$U \;=\; \begin{cases} \alpha + \int_0^\infty \xi_s dX_s + N_\infty, \\ \alpha + \sum_{i=1}^\infty \bar{\xi}_i \Delta \bar{X}_i + \bar{N}_\infty, \end{cases} \tag{2.6}$$

where N is a square-integrable martingale w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ orthogonal to X, and $\bar{N}$ is a square-integrable martingale w.r.t. $(\mathcal{F}_{T_i})_{i\in I\!N}$ orthogonal to $\bar{X}$, and $\alpha = E(U|\mathcal{F}_0)$. Then it is natural to call the *discretized version* of the integrand ξ the following continuous-time process:

$$\xi'_t \;=\; \bar{\xi}_i \qquad \text{if } \; T_{i-1} < t \leq T_i, \quad i \geq 1. \tag{2.7}$$

In a sense, this process ξ' naturally occurs if we replace X by the discretized version along the subdivision τ.

Our aim here is to compute ξ' in terms of ξ. This is simple, after recalling that the process ξ is square-integrable w.r.t. the finite measure Q_X defined by (1.3), and after introducing the σ-field $\mathcal{P}'$ on $\tilde{\Omega} = \Omega \times I\!R_+$ which is generated by the sets $D \times (T_i, T_{i+1}]$, where $i \in I\!N$ and $D \in \mathcal{F}_{T_i}$:

Proposition 2.2 *Assume that X is a square-integrable martingale. With the above notation we have $\xi' = Q_X(\xi|\mathcal{P}')$ (the conditional expectation of ξ w.r.t. $\mathcal{P}'$ for the finite measure Q_X).*

Proof. Set $A = \langle X, X\rangle$ and $B_t = \int_0^t \xi_s dA_s$. If $M_t = E(U|\mathcal{F}_t)$, then (1.2) yields $M_t = \alpha + \int_0^t \xi_s dX_s + N_t$, hence $\langle X, M\rangle = B + \langle X, N\rangle = B$ because X and N are orthogonal. So an application of Proposition 2.1 yields the following explicit form for $\bar{\xi}$:

$$\bar{\xi}_i \;=\; \frac{E(B_{T_i} - B_{T_{i-1}}|\mathcal{F}_{T_{i-1}})}{E(A_{T_i} - A_{T_{i-1}}|\mathcal{F}_{T_{i-1}})}. \tag{2.8}$$

For $D \in \mathcal{F}_{T_{i-1}}$ we have

$$Q_X(1_{D\times(T_{i-1},T_i]}\xi) \;=\; E(1_D(B_{T_i} - B_{T_{i-1}})) \;=\; E(1_D E(B_{T_i} - B_{T_{i-1}}|\mathcal{F}_{T_{i-1}})).$$

By (2.8) and (2.7) this is equal to

$$E(1_D\bar{\xi}_i(A_{T_i} - A_{T_{i-1}})) \;=\; Q_X(1_{D\times(T_{i-1},T_i]}\xi').$$

Since ξ' is obviously $\mathcal{P}'$-measurable, this implies the result. $\Box$

Remark 2.3 Exactly the same result (with the same proof) holds if we assume that X is a locally square-integrable martingale, such that each stopped process $(X_t^{T_i} = X_{T_i\wedge t})_{t\geq 0}$ is a square-integrable martingale. $\Box$

2.3 A Clark-Haussmann formula for Markov processes

In this subsection we give an alternative form of the Clark-Haussmann formula giving the integrand $\xi(X, U)$: see Nualart [14] for a general form of this formula.

The setting is as follows: we have a quasi-left continuous $I\!R^d$-valued strong Markov process Y on $(\Omega, \mathcal{F}, (\mathcal{F}_t), P_x)$, where P_x is the probability measure under which $Y_0 = x$ a.s., and we assume also that Y is a semimartingale under each P_x. Let μ be the jump measure of Y, and (B, C, ν) its characteristics: we refer for this to [9], and also to [3] for the following structural results, showing in particular that (B, C, ν) do not depend on the starting point: there exist a continuous increasing additive functional A, a Borel $I\!R^d$-valued function b, a Borel nonnegative symmetric $I\!R^d \otimes I\!R^d$-valued function c and a transition measure F from $\mathbb{R}^d$ into itself integrating $z \rightsquigarrow |z|^2 \wedge 1$ (the "modified Lévy measure"), such that

$$\left.\begin{array}{rcl} B_t &=& \int_0^t b(Y_{s-})dA_s, \\ C_t^{ij} &=& \int_0^t c(Y_{s-})^{ij}dA_s, \\ \nu(\omega, ds, dz) &=& dA_s(\omega)F(Y_{s-}(\omega), dz). \end{array}\right\} \tag{2.9}$$

We denote by (P_t) the transition semi-group of Y.

Now, we work under the measure P of the form $P = \int m(dx)P_x$ (so m is the law of Y_0). Denote by $\mathcal{D}_T$ the class of all Borel functions f such that $f(Y_T) \in \mathbb{L}^2(P)$ and that the function $(t, y) \rightsquigarrow P_t f(y)$ on $(0, \infty) \times \mathbb{R}^d$ is once differentiable in t and twice differentiable in y, with all partial derivatives being continuous.

Next, our basic locally square-integrable martingale X is of the form

$$X_t \;=\; X_0 + \int_0^t \gamma_s^{\mathrm{T}} dY_s^c + \int_0^t \int_{\mathbb{R}^d} \bar{\gamma}(s,z)(\mu - \nu)(ds, dz), \tag{2.10}$$

where Y^c denotes the continuous martingale part of Y, and "T" denotes the transpose, and $\gamma = (\gamma^i)_{1\le i\le d}$ and $\bar{\gamma}$ are predictable functions on $\Omega \times \mathbb{R}_+$ and $\Omega \times \mathbb{R}_+ \times \mathbb{R}^d$, such that for all $t < \infty$:

$$\int_0^t a_s dA_s \;<\; \infty, \qquad \text{where} \;\; a_s \;=\; \gamma_s^{\mathrm{T}} c(Y_{s-})\gamma_s + \int F(Y_{s-}, dz)\bar{\gamma}(s,z)^2. \tag{2.11}$$

Observe that under (2.11), X is well defined and is a locally square-integrable martingale under each P_x and P, with angle bracket $\langle X, X\rangle_t = \int_0^t a_s dA_s$.

Theorem 2.4 *Let $f \in \mathcal{D}_T$ and $U = f(Y_T)$ for a given $T \in \mathbb{R}_+$. Then a version of the process $\xi = \xi(X, U)$ is given for $s > T$ by $\xi_s = 0$ and for $s \le T$ by*

$$\xi_s = \frac{1}{a_s}\left(\gamma_s^{\mathrm{T}} c(Y_{s-})\nabla(P_{T-s}f)(Y_{s-}) + \int F(Y_{s-}, dz)\bar{\gamma}(s,z)(P_{T-s}f(Y_{s-}+z) - P_{T-s}f(Y_{s-}))\right). \tag{2.12}$$

We use here the traditional convention $\frac{0}{0} = 0$, since when $a_s = 0$ the numerator in the right side of (2.12) is also 0. Observe that the process ξ does not depend on the measure m in $P_m := P = \int m(dx)P_x$, as long as $f(Y_T)$ is in $\mathbb{L}^2(P_m)$.

Proof. Since U is $\mathcal{F}_T$-measurable, that $\xi_s = 0$ for $s > T$ is trivial. By the Markov property, $M_t = E(U|\mathcal{F}_t)$ is given for $t \le T$ by $M_t = g(t, Y_t)$, where $g(t,y) = P_{T-t}f(y)$. By hypothesis, g is once differentiable in t and twice differentiable in y with continuous partial derivatives. By Itô's formula,

$$M_t \;=\; \alpha + \int_0^t \frac{\partial}{\partial s} g(s, Y_{s-})ds + \frac{1}{2}\int_0^t \sum_{1\le i,j\le d} \frac{\partial^2}{\partial y^i \partial y^j} g(s, Y_s)c(Y_{s-})^{ij} dA_s$$

$$+\int_0^t \sum_{1\le i\le d} \frac{\partial}{\partial y^i} g(s,Y_{s-}) b(Y_{s-})^i dA_s + \int_0^t \sum_{1\le i\le d} \frac{\partial}{\partial y^i} g(s,Y_{s-}) dY_s^{i,c}$$

$$+\int_0^t \int_{I\!R^d} \sum_{1\le i\le d} \frac{\partial}{\partial y^i} g(s,Y_{s-}) z^i 1_{\{|z|\le 1\}} (\mu-\nu)(ds,dz)$$

$$+\int_0^t \int_{I\!R^d} \left(g(s,Y_{s-}+z) - g(s,Y_{s-}) - \sum_{1\le i\le d} \frac{\partial}{\partial y^i} g(s,Y_{s-}) z^i 1_{\{|z|\le 1\}} \right) \mu(ds,dz).$$

The first three integrals above are predictable process of finite variation. The last integral may be rewritten as the sum of the stochastic integral w.r.t. the measure martingale $\mu-\nu$, plus the integral w.r.t. ν, which again is a predictable process of finite variation. Since M is a martingale, the sum of all predictable processes of finite variation must equal 0, and after a simple transformation we get

$$M_t = \alpha + \int_0^t \sum_{1\le i\le d} \frac{\partial}{\partial y^i} g(s,Y_{s-}) dY_s^{i,c} + \int_0^t \int_{I\!R^d} (g(s,Y_{s-}+z) - g(s,Y_{s-}))(\mu-\nu)(ds,dz).$$

Then (1.2) and (2.10) give for $t \le T$:

$$N_t = \int_0^t \sum_{1\le i\le d} \left(\frac{\partial}{\partial y^i} g(s,Y_{s-}) - \xi_s \gamma_s^i \right) dY_s^{i,c}$$

$$+ \int_0^t \int_{I\!R^d} \left(g(s,Y_{s-}+z) - g(s,Y_{s-}) - \xi_s \bar\gamma(s,z) \right) (\mu-\nu)(ds,dz).$$

Then we get

$$\langle N,X\rangle_t = \int_0^t \left(\sum_{1\le i,j\le d} \left(\frac{\partial}{\partial y^i} g(s,Y_{s-}) - \xi_s \gamma_s^i \right) c(Y_{s-})^{ij} \gamma_s^j + \right.$$

$$\left. + \int_{I\!R^d} F(Y_{s-},dz) \left(g(s,Y_{s-}+z) - g(s,Y_{s-}) - \xi_s \bar\gamma(s,z) \right) \bar\gamma(s,z) \right) dA_s.$$

In view of (2.11), this becomes

$$\langle N,X\rangle_t = \int_0^t \left(-\xi_s a_s + \sum_{1\le i,j\le d} \frac{\partial}{\partial y^i} g(s,Y_{s-}) c(Y_{s-})^{ij} \gamma_s^j \right.$$

$$\left. + \int_{I\!R^d} F(Y_{s-},dz)(g(s,Y_{s-}+z) - g(s,Y_{s-}))\bar\gamma(s,z) \right) dA_s.$$

Since ξ is charaterized by the orthogonality of N and X, that is by $\langle N,X\rangle = 0$, a version of ξ is thus given by (2.12), and we have proved the claim. □

The class $\mathcal{D}_T$ of functions for which (2.12) holds is rather restrictive. It might be of interest to enlarge this class. To this effect, for each $x \in I\!R^d$ we introduce the set $\mathcal{D}'_T$ of all functions f for which there is a sequence $f_n \in \mathcal{D}_T$ (called an "approximating sequence") such that $f_n(Y_T) \to f(Y_T)$ in $I\!L^2(P)$.

The measure Q_X associated by (1.3) with X (and relative to P) is here $Q_X(d\omega, dt) = P(d\omega)dA_t(\omega)a_t(\omega)$.

Finally, let $\mathcal{D}''_T$ be the subset of all $f \in \mathcal{D}'_T$ such that $y \rightsquigarrow P_t f(y)$ is differentiable for $0 < t \leq T$, and for which there is an approximating sequence f_n in $\mathcal{D}_T$ such that for all $t \in (0,T]$ and $y \in I\!R^d$ we have

$$P_t f_n(y) \ \to \ P_t f(y), \qquad \frac{\partial}{\partial y^i} P_t f_n(y) \ \to \ \frac{\partial}{\partial y^i} P_t f(y), \tag{2.13}$$

and that for Q_X-almost all (ω, t) with $t \leq T$ we have

$$\int F(Y_{t-}(\omega), dz)\bar{\gamma}(\omega, t, z)(P_{T-t} f_n(Y_{t-}(\omega) + z) - P_{T-t} f_n(Y_{t-}(\omega)))$$
$$\to \int F(Y_{t-}(\omega), dz)\bar{\gamma}(\omega, t, z)(P_{T-t} f(Y_{t-}(\omega) + z) - P_{T-t} f(Y_{t-}(\omega))). \tag{2.14}$$

Observe that (2.13) implies (2.14) as soon as f, the f_n's and the $\frac{\partial}{\partial y^i} P_t f_n$'s are uniformly bounded for each t, by virtue of (2.11) and of the fact that $\int F(y, dz)(|z|^2 \wedge 1) < \infty$.

Corollary 2.5 *a) If $f \in \mathcal{D}'_T$ with the approximating sequence f_n, then a version of the process $\xi(X, f(Y_T))$ is the limit of $\xi(X, f_n(Y_T))$ in $I\!L^2(Q_X)$.*

b) If further $f \in \mathcal{D}''_T$, then a version of the process $\xi(X, f(Y_T))$ is given by (2.12) for $s \leq T$, and by 0 for $s > T$.

Proof. The claim a) readily follows from (1.4). Assume now that $f \in \mathcal{D}''_T$, with the approximating sequence f_n. Then if ξ' is given by (2.12), on the one hand $\xi(X, f_n(Y_T))_s(\omega) \to \xi'_s(\omega)$ for Q_X-almost all (ω, s), and on the other hand (1.4) holds: hence $\xi' = \xi(X, f(Y_T))$ Q_X-a.s., and we have b). $\square$

2.4 A particular case

Theorem 2.4 and its corollary are not quite satisfactory, because they give $\xi(X, U)$ for a variable U of the form $U = f(Y_T)$, while one would like to have it for $U = f(X_T)$. It becomes more satisfactory when X itself is Markov. We give in some detail a simple case of this situation, namely when X is the solution of the equation $dX = g(X_-)dZ$, where Z is a 1-dimensional Lévy process and g a smooth enough coefficient.

Since we wish X to be a locally square-integrable martingale, it is natural to assume first that the Lévy process Z, which is defined on some space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$, is a locally square-integrable martingale itself. This in fact implies that it is then a martingale square-integrable on compacts, and its characteristic function has the form

$$E\left(e^{iuZ_t}\right) \ = \ \exp t \left(-\frac{cu^2}{2} + \int F(dx)(e^{iux} - 1 - iux)\right), \tag{2.15}$$

where $c \geq 0$ and the Lévy measure F integrates x^2. We set

$$\tilde{c} \ = \ c + \int F(dz) z^2, \tag{2.16}$$

so $\langle Z, Z\rangle_t = \tilde{c}t$, and of course we assume that $\tilde{c} > 0$ (otherwise $Z = 0$ and what follows is empty). Next we have a continuously differentiable function g with bounded derivative g', and for any x we consider the solution X^x of the following stochastic differential equation:

$$X_t^x \;=\; x + \int_0^t g(X_{s-}^x)dZ_s \tag{2.17}$$

A classical argument (see (5.2) in the Appendix) yields that X^x is a square-integrable martingale over each finite interval $[0,T]$. Further, the solution of the following linear equation

$$X_t'^x \;=\; 1 + \int_0^t g'(X_{s-}^x)X_{s-}'^x dZ_s, \tag{2.18}$$

is also a square-integrable martingale over each finite interval $[0,T]$. So for each measurable function f with at most linear growth we can set

$$P_t f(x) \;=\; E(f(X_t^x)), \qquad Q_t f(x) \;=\; E(f(X_t^x)X_t'^x). \tag{2.19}$$

Observe that (P_t) is the semi-group of X^x, which is a Markov process. We then have:

Theorem 2.6 *Assume that g has a continuous and bounded derivative.*

a) *For any $T \in I\!\!R_+$ and any differentiable function f with bounded derivative f' the variable $f(X_T^x)$ is square-integrable, and a version of $\xi(X^x, f(X_T^x))$ is given by*

$$\xi_s(X^x, f(X_T^x)) \;=\; \eta(s, X_{s-}^x)1_{[0,T]}(s), \tag{2.20}$$

where

$$\eta(s,y) \;=\; Q_{T-s}f'(y) + \frac{1}{\tilde{c}}\int F(dz)z^2 \int_0^1 (Q_{T-s}f'(y+g(y)zu) - Q_{T-s}f'(y))du. \tag{2.21}$$

b) *The same holds when f is the difference of two convex functions, with a right derivative f'_r bounded and f' above replaced by f'_r, provided we have $P_t(y,.)$ has no atom for all $t \in (0,T]$, $y \in I\!\!R$.*

In the last claim one can of course replace the right derivative f'_r by the left derivative f'_l. The last condition is obviously satisfied when $P_t(x,.)$ has a density: this is the case when $c > 0$ as soon as g does not vanish (or, does not vanish in the set in which the process X^x takes its values). When $c = 0$, one can find conditions implying the existence of a density in e.g. [2].

Proof. 1) Our assumptions always imply that f and g have at most linear growth. Then the property $f(X_T^x) \in I\!\!L^2(P)$ follows from (5.2) in the Appendix.

2) We first prove the result under the three additional assumptions that Z has bounded jumps (which is equivalent to saying that F has compact support), that g is infinitely differentiable with bounded derivatives of all orders, and that f is twice continuously differentiable with f, f' and f'' bounded.

By virtue of Theorem 2.4 applied to $Y = X^x$, it suffices to prove that $f \in \mathcal{D}_T$ and that (2.12) reduces to (2.21) in our situation. The first property is proved in Lemma 5.1 of the Appendix. So it remains to identify (2.12) with (2.21). With $Y = X$ we have $b = 0$, $c(y) = cg(y)^2$, $A_t = t$ and $F(y,.)$ is the image of the measure F under the map $z \mapsto g(y)z$, while in (2.10) we must take $\gamma_s = 1$ and $\bar\gamma(s,z) = z$. So if $a(y) = cg(y)^2 + \int F(dz)z^2 g(y)^2$, (2.11) becomes $a_s = a(X_{s-})$, while $\nabla P_t f = Q_t f'$ by (5.3). Then (2.12) yields that we have (2.20) with η' instead of η, given by

$$\eta'(s,y) \;=\; \frac{g(y)^2 c Q_{T-s} f'(y) + \int F(dz) z g(y)(P_{T-s}f(y+g(y)z) - P_{T-s}f(y))}{g(y)^2(c + \int F(dz)z^2)}$$

if $g(y) \neq 0$, and $\eta'(s,y) = 0$ if $g(y) = 0$. By Taylor's formula and again the property $\nabla P_t f = Q_t f'$ and (2.16) we note that $\eta'(s,y)$ equals $\eta(s,y)$ as given by (2.21) when $g(y) \neq 0$. Finally the process $\xi(X^x, f(X^x_T))$ is unique up to a Q_{X^x}-null set, and the set $\{(\omega,t) : g(X^x_{t-}(\omega)) = 0\}$ is Q_{X^x}-negligible: hence (2.20) gives a version of $\xi(X^x, f(X^x_T))$.

3) Second, we prove the result when f and g are once continuously differentiable with bounded derivatives f' and g', and further f is bounded, and when Z is a locally square-integrable martingale.

First we replace Z by a sequence of Lévy processes Z^n obtained by truncating the jumps of Z of size bigger than n. More precisely, we may write

$$Z_t \;=\; Z^c_t + \int_0^t \int_{I\!R} z(\mu - \nu)(ds,dz),$$

where Z^c is the continuous martingale part of Z (of the form cW, where W is a Wiener process), μ is the jump measure of Z and $\nu(ds,dz) = ds \otimes F(dz)$. Then we set

$$Z^n_t \;=\; Z^c_t + \int_0^t \int_{I\!R} z 1_{\{|z| \le n\}} (\mu - \nu)(ds,dz).$$

We have $\langle Z, Z\rangle_t = \tilde c t$ where $\tilde c = c + \int F(dz) z^2$, and $\langle Z^n, Z^n\rangle_t = \tilde c_n t$ where $\tilde c_n = c + \int F(dz) z^2 1_{\{|z|\le n\}}$, and obviously

$$\langle Z - Z^n, Z - Z^n\rangle_t \;\to\; 0. \tag{2.22}$$

Next, let ϕ be an infinitely differentiable function with compact support and integral equal to 1. Then we replace g by $g_n(x) = \int n\phi(ny) g(x-y)dy$ and f by $f_n(x) = \int n\phi(ny) f(x-y)dy$. Thus g_n and f_n are infinitely differentiable with bounded derivatives of all orders, and there exists K such that for all n:

$$|g_n(0)| \;\le\; K, \quad |g'_n(x)| \;\le\; K, \quad |f_n(x)| \;\le\; K, \quad |f'_n(x)| \;\le\; K, \tag{2.23}$$

and moreover

$$g_n \;\to\; g, \quad g'_n \;\to\; g' \quad f_n \;\to\; f, \quad f'_n \;\to\; f' \quad \text{locally uniformly.} \tag{2.24}$$

Now, denote by $X^{n,x}$ and $X'^{n,x}$ the solutions of (2.17) and (2.18), with Z and g substituted with Z^n and g_n, and by Q^n_t the kernel associated by (2.19). Since Z^n, g_n and

f_n satisfy the conditions of Step 2, $\xi^n = \xi(X^{n,x}, f_n(X_T^{n,x}))$ is given by (2.20), with η_n given by

$$\eta_n(s,y) = Q^n_{T-s} f'_n(y) + \frac{1}{\tilde{c}_n} \int F(dz) z^2 1_{\{|z|\leq n\}} \int_0^1 (Q^n_{T-s} f'_n(y + g_n(y) z u) - Q^n_{T-s} f'_n(y)) du.$$

By virtue of (2.22) and (2.24), stability results for stochastic differential equations (see [10]) imply that (X^{n,y_n}, X'^{n,y_n}) converges locally uniformly (in time) in probability to (X^y, X'^y) for any sequence $y_n \to y$, and further $f_n(X_t^{n,y_n}) \to f(X_t^y)$ and $f'_n(X_t^{n,y_n}) \to f'(X_t^y)$ in probability by (2.24), hence also in $\mathbb{L}^p(P)$ for all p by (2.23). Thus applying $\tilde{c}_n \leq \tilde{c}$ and (5.2) of the Appendix and (2.19) we see that

$$Q^n_t f'_n(y_n) \;\to\; Q_t f'(y), \qquad |Q^n_t f'_n(y)| \;\leq\; K_t \tag{2.25}$$

for a constant K_t independent of n and x. If further we use the facts that $\tilde{c}_n \to \tilde{c}$ and that $\int F(dz) z^2 < \infty$, (2.24) and (2.25) allow us to deduce that if η is defined by (2.21),

$$y_n \;\to\; y \qquad \Rightarrow \qquad \eta_n(s, y_n) \;\to\; \eta(s,y). \tag{2.26}$$

Finally, we have

$$X_t^{n,x} - X_t^x \;=\; \int_0^t g_n(X_{s-}^{n,x}) d(Z^n - Z)_s + \int_0^t (g_n(X_{s-}^{n,x}) - g(X_{s-}^x)) dZ_s,$$

hence

$$\langle X^{n,x} - X^x, X^{n,x} - X^x \rangle_t \;\leq\; 2(\tilde{c} - \tilde{c}_n) \int_0^t g_n(X_s^{n,x})^2 ds + 2\tilde{c} \int_0^t (g_n(X_s^{n,x}) - g(X_s^x))^2 ds.$$

Using $\tilde{c}_n \to \tilde{c}$, (2.23), (2.24) and the fact that $X_s^{n,x} \to X_s^x$ uniformly in $s \in [0,t]$ in probability, we readily deduce that $\langle X^{n,x} - X^x, X^{n,x} - X^x \rangle_t \;\to\; 0$ in probability.

Since $f_n(X_T^{n,x}) \;\to\; f(X_t^x)$ in $\mathbb{L}^2(P)$, we are in a position to apply a result proved in the next section (not based upon the present theorem, of course), namely Theorem 3.3: this theorem asserts that ξ^n converges to $\xi(X^x, f(X_T^x))$ in Q_X-measure. Then, since $\xi_s^n = \eta_n(s, X_{s-}^{n,x}) 1_{[0,T)}(s)$ and since $X^{n,x}$ converges locally uniformly in time, in probability, to X^x, we deduce (2.20) from (2.26).

4) For (a) it remains to consider the case where f has a continuous bounded derivative but is not bounded itself. We can find a sequence (f_n) of bounded continuously differentiable functions such that $|f'_n(x)| \leq K$ for some constant K and $f_n(x) = f(x)$ for all $|x| \leq n$ and $|f_n| \leq |f|$. Then $\xi^n = \xi(X^x, f_n(X_T^x))$ is given by (2.20) with η_n given by (2.21), where f is substituted with f_n. That $f_n(X_T^x) \to f(X_T^x)$ in $\mathbb{L}^2(P)$ and that (2.25) holds with Q_t instead of Q_t^n are obvious by the previous estimates on f_n and f'_n and (5.2) of the Appendix. It follows as above that $\eta_n \to \eta$ pointwise, where η is given by (2.21). Then by (1.4) we have (2.20).

5) It remains to prove (b). We set $f_n(x) = \int n\phi(ny) f(x-y) dy$ as in Step 3. Then $f_n \to f$ locally uniformly, while $f'_n \to f'_r$ everywhere except on an at most countable set D, and we still have $|f'_n(x)| \leq K$ and $|f_n(0)| \leq K$ for some constant K. Therefore, exactly

as in Step 3, we have $f_n(X_t^y) \to f(X_t^y)$ and $f_n'(X_t^{y_n}) \to f_r'(X_t^y)$ (as $y_n \to y$) in $\mathbb{L}^2(P)$, at least when restricted to the set $\Omega_{y,t} = \{\omega : X_t^y(\omega) \notin D\}$. Now, if $P_t(y,.)$ has no atom, the set $\Omega_{y,t}$ is P-negligible: then we have (2.25) with Q_t instead of Q_t^n, and the rest of the proof follows as in Step 4. $\square$

Remark 2.7 When $c = 0$ and F is a finite measure, the process Z is a compensated compound Poisson process, and the situation is much simpler. One can show with the same methods used in Steps 3 or 4 above that the result hold without any differentiability condition. We need f to be continuous, and both f and g with linear growth, and (2.21) takes the following simple form (with $\frac{0}{0} = 0$):

$$\eta(s,y) \;=\; \frac{\int F(dz)z(P_{T-s}f(y+g(y)z) - P_{T-s}f(y))}{g(y)\int F(dz)z^2}. \tag{2.27}$$

Of course in this simple situation we could also write an "elementary" proof which looks like the proof of Proposition 2.1. This is no surprise, since the compound Poisson case has a "discrete" structure.

Remark 2.8 We have considered above the "homogeneous" situation, where the coefficient g does not depend on time. Similar formulas would obviously hold when the coefficient depends on time, and also when the process Z is a non-homogeneous process with independent increments.

Remark 2.9 Let f be a function on $\mathbb{R}^k$ (say, with linear growth) and $0 < T_1 < \ldots < T_k$ be deterministic times, and $\xi = \xi(X^x, U)$ where $U = f(X_{T_1}^x, \ldots, X_{T_k}^x)$. Then the martingale $M_t = E(U|\mathcal{F}_t)$ when $T_{i-1} \le t \le T_i$ is $M_t = \int P_{T_i - t}(X_t^x, dy) f_{X_{T_1}^x, \ldots X_{T_{i-1}}^x}(y)$, where

$$f_{x_1,\ldots x_{i-1}}(y) = \int P_{T_{i+1}-T_i}(y, dx_{i+1}) \ldots P_{T_k - T_{k-1}}(x_{k-1}, dx_k) f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots x_k).$$

By iteration of the previous result we then get that

$$\xi_s \;=\; \sum_{i=1}^k \eta_{i;X_{T_1}^x,\ldots X_{T_{i-1}}^x}(s, X_{s-}^x) 1_{(T_{i-1},T_i]},$$

where

$$\eta_{i;x_1,\ldots x_{i-1}}(s,y) \;=\; Q_{T_i - s} f'_{x_1,\ldots x_{i-1}}(y)$$
$$+\frac{1}{\tilde{c}} \int F(dz) z^2 \int_0^1 (Q_{T_i-s} f'_{x_1,\ldots x_{i-1}}(y + g(y)zu) - Q_{T_i - s} f'_{x_1,\ldots x_{i-1}}(y)) du.$$

Of course we need some smoothness conditions of f to do that: that f is continuously differentiable with all partial derivatives bounded is enough, in which case we need to reproduce the proof of the previous theorem.

We *do not have* an explicit form for $\xi(X,U)$ when U is a function of the whole path of X^x over $[0,T]$. But the variables of the form above are dense into the set of square-integrable variables, measurable w.r.t. the σ-field $\sigma(X_s^x : s \le T)$. This is to be compared to the Clark-Haussmann formula in the Wiener case, see e.g. Nualart [14]: in this case one has an "explicit" form for $\xi(X^x,U)$ for variables U that are smooth in the Malliavin sense (and thus include the variables $f(X^x_{T_1},\ldots)$ for smooth f's). But this approach is limited to the Wiener space, and the explicit form involves the not so explicit Malliavin derivatives and predictable projections of such derivatives.

3 Strong convergence results

3.1 Discretization of a process

Here we consider the situation of Section 2.2, and we look at what happens when we have a sequence of subdivisions whose meshes go to 0. More precisely, we have a square-integrable martingale X on a basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},P)$ with $\mathcal{F}=\bigvee_t \mathcal{F}_t$, and for each n a subdivision $\tau_n=(T(n,i):i\in I\!N)$ satisfying (2.4). The sequence (τ_n) satisfies

$$\sup_{i\geq 1}(T(n,i)\wedge t-T(n,i-1)\wedge t) \quad\to\quad 0 \quad \text{a.s. as } n\to\infty. \tag{3.1}$$

Then set $\bar{X}^n_i=X_{T(n,i)}$: for each n the sequence $(\bar{X}^n_i)_{i\in I\!N}$ is a square-integrable martingale w.r.t. $(\mathcal{F}_{T(n,i)})_{i\in I\!N}$. Finally, let $U\in I\!L^2(P)$ be fixed. We have the first decomposition (2.6), and the second one for each n with the process $\bar{\xi}^n$, and we associate with $\bar{\xi}^n$ and τ_n the process ξ'^n by (2.7). Then we have:

Theorem 3.1 *Under (3.1), and if X is a square-integrable martingale and $U\in I\!L^2(P)$, the functions ξ'^n tend to ξ in $I\!L^2(Q_X)$.*

Proof. For each n we endow the space $\tilde{\Omega}$ with the σ-field $\mathcal{P}'_n$ generated by the sets $D\times(T(n,i-1),T(n,i)]:=\{(\omega,t):\omega\in D, T(n,i-1)(\omega)<t\leq T(n,i)(\omega)\}$, where $i\geq 1$ and $D\in\mathcal{F}_{T(n,i-1)}$. By virtue of Proposition 2.2, we have $\xi'^n=Q_X(\xi|\mathcal{P}'_n)$. Recall that here Q_X is a finite measure.

The sequence (ξ'^n) is bounded in $I\!L^2(Q_X)$, and thus is in a compact set for the weak topology in $I\!L^2(Q_X)$. So there exists a subsequence, again denoted by ξ'^n for simplicity, which converges weakly to a variable ξ' in $I\!L^2(Q_X)$.

Let us first show that $\xi'=\xi$ Q_X-a.s. Take $\eta=1_{D\times]s,t]}$, where D is $\mathcal{F}_s$-measurable. Then $Q_X(\xi'^n\eta)\to Q_X(\xi'\eta)$. Consider the two stopping times $S_n=\inf\{T(n,i):i\in I\!N, T(n,i)\geq s\}$ and $T_n=\inf\{T(n,i):i\in I\!N, T(n,i)\geq t\}$. Then

$$\left.\begin{aligned} Q_X(\xi'^n\eta) &= Q_X(\xi'^n 1_{D\times(S_n,T_n]})+Q_X(\xi'^n 1_{D\times(s,S_n]})-Q_X(\xi'^n 1_{D\times(t,T_n]}) \\ Q_X(\xi\eta) &= Q_X(\xi 1_{D\times(S_n,T_n]})+Q_X(\xi 1_{D\times(s,S_n]})-Q_X(\xi 1_{D\times(t,T_n]}). \end{aligned}\right\} \tag{3.2}$$

If $A(n,s,\varepsilon)=\{S_n>s+\varepsilon\}$ we have

$$Q_X(D\times(s,S_n]) \;\leq\; Q_X(A(n,s,\varepsilon)\times I\!R_+)+Q_X(\Omega\times(s,s+\varepsilon]).$$

Since $Q_X(.\times I\!R_+)$ is absolutely continuous w.r.t. P and since (3.1) implies $P(A(n,s,\varepsilon))\to 0$ as $n\to\infty$, we also have $Q_X(A(n,s,\varepsilon)\times I\!R_+)\to 0$ as $n\to\infty$ for all $\varepsilon>0$. On the other hand, $Q_X(\Omega\times(s,s+\varepsilon])\to 0$ as $\varepsilon\to 0$ because Q_X is a finite measure, so we deduce that $Q_X(D\times(s,S_n])\to 0$. The variables ξ'^n being Q_X-uniformly integrable, we deduce that $Q_X(\xi'^n 1_{D\times(s,S_n]})\to 0$, and similarly $Q_X(\xi'^n 1_{D\times(t,T_n]})\to 0$, and $Q_X(\xi 1_{D\times(s,S_n]})\to 0$ and $Q_X(\xi 1_{D\times(t,T_n]})\to 0$. Since further $D\times(S_n,T_n]\in\mathcal{P}'_n$, we get $Q_X(\xi'^n 1_{D\times(S_n,T_n]}) = Q_X(\xi 1_{D\times(S_n,T_n]})$. It then follows from (3.2) and from the fact that $Q_X(\xi'^n\eta)\to Q_X(\xi'\eta)$, that $Q_X(\xi'\eta)=Q_X(\xi\eta)$. Then by a monotone class argument, this relation holds for all bounded predictable η, which yields $\xi=\xi'$ Q_X-a.s. (recall that ξ and all ξ'^n, hence ξ' as well, are predictable).

In particular, $Q_X((\xi'^n)^2)=Q_X(\xi'^n\xi)$ tends to $Q_X(\xi^2)$. Now, if a sequence ξ'^n in $\mathbb{L}^2(Q_X)$ converges weakly to ξ and the norms of ξ'^n converge to the norm of ξ, we have indeed strong convergence. Thus the $\mathbb{L}^2(Q_X)$-convergence of the original sequence ξ'^n to ξ follows. □

Remark 3.2 When the subdivisions (τ_n) are finer and finer, the sequence of σ-fields $\mathcal{P}'_n$ is increasing, hence the fact that $\xi'^n = Q_X(\xi|\mathcal{P}'_n)$ implies that the sequence ξ'^n is a square-integrable martingale and the convergence to ξ readily follows from the fact that $\mathcal{P}'_n$ increases to $\mathcal{P}$ up to P-null sets (by (3.1)). □

3.2 A general continuous time convergence theorem

We have again a stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},P)$ with $\mathcal{F}=\bigvee\mathcal{F}_t$, supporting locally square-integrable martingales X and X^n and square-integrable variables U^n and U. We have the following (unique) decompositions, as in (1.2):

$$\left.\begin{array}{rcl} U &=& \alpha+\int_0^\infty \xi_s dX_s + N_\infty, \\ U^n &=& \alpha^n+\int_0^\infty \xi^n_s dX^n_s + N^n_\infty, \end{array}\right\} \qquad (3.3)$$

where $\alpha=E(U|\mathcal{F}_0)$ and $\alpha^n=E(U^n|\mathcal{F}_0)$ and N (resp. N^n) is a square-integrable martingale orthogonal to X (resp. to X^n).

Below, we consider again the measure Q_X associated with X by (1.3). It is not necessarily finite, so we recall that $\xi^n\to^{Q_X}\xi$ means that $\xi^n\to\xi$ in R-measure for one (hence for all) finite measure R equivalent to Q_X. Our main result is the following:

Theorem 3.3 *Assume that $U^n\to U$ in $\mathbb{L}^2(P)$ and that X and X^n are locally square-integrable martingales satisfying*

$$\langle X^n-X,X^n-X\rangle_t \;\to\; 0 \qquad \textit{in probability for all } t>0. \qquad (3.4)$$

Then $\xi^n\to\xi$ in Q_X measure. (We denote $\xi^n\to^{Q_X}\xi$).

Proof. 1) To begin with, we introduce the following orthogonal decompositions for the locally square-integrable martingales X^n and the square-integrable martingales N^n (recall (3.3)); below the processes L^n are locally square-integrable martingales and T^n are square-integrable martingales (recall also that the orthogonality between local martingales is denoted by $\perp$):

$$X^n_t \;=\; X^n_0+\int_0^t \gamma^n_s dX_s + L^n_t, \qquad L^n\perp X, \qquad (3.5)$$

$$N^n_t \;=\; \int_0^t \lambda^n_s dX_s + T^n_t, \qquad T^n \perp X. \tag{3.6}$$

In what follows we prove a bit more than is strictly necessary for the present theorem, but the following facts will also be used in the subsequent results. The orthogonality of X^n and N^n yields

$$\int_0^t \gamma^n_s \lambda^n_s d\langle X,X\rangle_s + \langle L^n,T^n\rangle_t \;=\; 0, \qquad \forall t, \;\; \text{a.s.} \tag{3.7}$$

We also have

$$\langle X^n - X, X^n - X\rangle_t \;=\; \int_0^t (\gamma^n_s - 1)^2 d\langle X,X\rangle_s + \langle L^n,L^n\rangle_t, \tag{3.8}$$

$$Q_X((\xi^n\gamma^n + \lambda^n - \xi)^2) \;\leq\; E((U^n - U)^2) \;\to\; 0, \tag{3.9}$$

$$\left.\begin{array}{rcl} E(\langle T^n,T^n\rangle_\infty) & \leq & E(\langle N^n,N^n\rangle_\infty) \\ E(\int_0^\infty (\xi^n_s)^2 d\langle L^n,L^n\rangle_s) & \leq & E((\int_0^\infty \xi^n_s dX^n_s)^2) \end{array}\right\} \;\leq\; E((U^n)^2) \;\leq\; K \tag{3.10}$$

for some constant K, and where we have used that $U^n \to U$ in $\mathbb{L}^2(P)$ for the last two properties.

2) After these preliminaries, we can go the proof of our claim. First, we can write the (pathwise) Lebesgue decomposition of the process $\langle L^n,T^n\rangle$, which is of locally bounded variation, w.r.t. the increasing process $\langle X,X\rangle$ as $\langle L^n,T^n\rangle_t = \int_0^t \rho^n_s d\langle X,X\rangle_s + A^n_t$, where A^n is a function of locally bounded variation which is singular w.r.t. to $\langle X,X\rangle$. Then (3.7) yields

$$\gamma^n\lambda^n + \rho^n \;=\; 0 \qquad Q_X - \text{a.s.} \tag{3.11}$$

But it is well known by Kunita-Watanabe inequality that the variation of the process $\langle L^n,T^n\rangle$ over $[0,t]$ is smaller than or equal to $\sqrt{\langle L^n,L^n\rangle_t}\sqrt{\langle T^n,T^n\rangle_t}$, while by the above Lebesgue decomposition it is bigger than $\int_0^t |\rho^n_s| d\langle X,X\rangle_s$. Then we readily deduce from (3.4), (3.8) and (3.10) that $\int_0^t |\rho^n_s| d\langle X,X\rangle_s \to^P\; 0$ for all t, so in view of (3.11) we get

$$\gamma^n\lambda^n \;\to^{Q_X}\; 0. \tag{3.12}$$

Next, (3.4) and (3.8) on the one hand, (3.9) on the other hand, give us:

$$\xi^n\gamma^n + \lambda^n \;\to^{Q_X}\; \xi, \qquad \gamma^n \;\to^{Q_X}\; 1. \tag{3.13}$$

Now, combining (3.12) and (3.13) readily gives us $\xi^n \to^{Q_X} \xi$. $\square$

Associated with this theorem, we have a result about the rate of convergence:

Theorem 3.4 *Assume that $U^n \to U$ in $\mathbb{L}^2(P)$ and that X and X^n are locally square-integrable martingales satisfying (3.4). Assume further that there is a sequence (a_n) in $\mathbb{R}_+$ going to $+\infty$ such that the sequence $(a_n(U^n - U) : n \geq 1)$ is bounded in $\mathbb{L}^2(P)$ and that for each t the sequence of variables $(a_n^2\langle X^n - X, X^n - X\rangle_t : n \in \mathbb{N})$ is uniformly tight. Then the sequence $(a_n(\xi^n - \xi) : n \in \mathbb{N})$ is uniformly tight with respect to any finite measure equivalent to Q_X.*

Proof. 1) We choose a finite measure R equivalent to Q_X. Let us first recall that if (u^n) is a sequence of processes such that for all t the sequence of random variables $(\int_0^t |u_s^n| d\langle X,X\rangle_s : n\geq 1)$ is tight, then the sequence (u^n) is R-tight.

Applying this to (3.8) and (3.9) multiplied by a_n^2 gives that

$$\text{the two sequences} \quad a_n(\gamma^n-1), \quad a_n(\xi^n\gamma^n+\lambda^n-\xi) \quad \text{are } R\text{-tight.} \tag{3.14}$$

We also deduce from (3.8) that for each t the sequence $(\langle L^n,L^n\rangle_t : n\geq 1)$ is tight. Exactly as in the last step of the previous proof, we deduce that the sequences $(\int_0^t a_n|\rho_s^n| d\langle X,X\rangle_s : n\geq 1)$ are tight, hence the sequence $(a_n\rho^n)$ is R-tight. In view of (3.11) we deduce that

$$\text{the sequence} \quad a_n\gamma^n\lambda^n \quad \text{is } R\text{-tight.} \tag{3.15}$$

Now, we can write

$$a_n(\xi^n-\xi) \;=\; a_n\xi^n(1-\gamma^n)(1+\gamma^n)+a_n\gamma^n(\gamma^n\xi^n+\lambda^n-\xi)-a_n\gamma^n\lambda^n+a_n\xi(\gamma^n-1).$$

We also know that the sequences (γ^n) and (ξ^n) are R-tight (by (3.13) and the previous theorem). Then the result readily follows from (3.14) and (3.15). □

3.3 A discrete version of Section (3.2)

Now we consider for each n a subdivision $\tau_n=(T(n,i) : i\in I\!N)$ of stopping times on the basis $(\Omega,\mathcal{F},(\mathcal{F}_t),P)$ with $\mathcal{F}=\bigvee \mathcal{F}_t$, satisfying (2.4), and we suppose that the sequence (τ_n) satisfies (3.1). For each n we have a square-integrable martingale X^n and a square-integrable variable U^n. Analoguous to (2.5), we set $\bar{X}_i^n = X_{T(n,i)}^n$. As in (2.6) we have (3.3), as well as the decomposition

$$U^n \;=\; \alpha^n+\sum_{i=1}^{\infty}\bar{\xi}_i^n\Delta\bar{X}_i^n+\bar{N}_\infty^n. \tag{3.16}$$

Then, as in (2.7) we set

$$\xi_t'^n \;=\; \bar{\xi}_i^n \qquad \text{if} \quad T(n,i-1)<t\leq T(n,i). \tag{3.17}$$

Theorem 3.5 *Assume that $U^n\to U$ in $I\!L^2(P)$ and that X^n and X are square-integrable martingales and that*

$$E(\langle X^n-X,X^n-X\rangle_\infty) \;\to\; 0. \tag{3.18}$$

Then the sequence ξ'^n converges to $\xi=\xi(X,U)$ in Q_X-measure.

Proof. In view of Proposition 2.2 we have $\xi'^n=Q_{X^n}(\xi^n|\mathcal{P}_n')$, where $\mathcal{P}_n'$ is the σ-field on $\tilde{\Omega}$ defined in the proof of Theorem 3.1. Let also $\mathcal{P}$ be the predictable σ-field on $\tilde{\Omega}$. We consider the decomposition (3.5) for X^n.

We can find a probability measure R on $(\tilde{\Omega}, \mathcal{P})$ which dominates all the finite measures Q_X and Q_{L^n}, and such that $Q_X \leq aR$ for some constant a. We can thus find nonnegative R-integrable and predictable functions V, V^n such that $V \leq a$ and

$$Q_X \;=\; V \bullet R, \qquad Q_{L^n} \;=\; V^n \bullet R.$$

Then we have $Q_{X^n} = W^n \bullet R$, with $W^n = (\gamma^n)^2 V + V^n$.

Now, (3.18) and (3.8), then (3.9), then (3.10), yield

$$\gamma^n \;\to^{\mathbb{L}^2(Q_X)}\; 1, \qquad Q_{L_n}(1) \;\to\; 0, \tag{3.19}$$

$$\xi^n \gamma^n + \lambda^n \;\to^{\mathbb{L}^2(Q_X)}\; \xi, \tag{3.20}$$

$$Q_{T^n}(1) \;\leq\; K, \qquad Q_{L_n}((\xi^n)^2) \;\leq\; K. \tag{3.21}$$

Furthermore we get $Q_X(|\rho^n|) \leq \sqrt{Q_{L^n}(1)}\sqrt{Q_{T^n}(1)}$, exactly as in the proof of Theorem 3.3, and in view of (3.19), (3.21) and (3.11), we obtain

$$\gamma^n \lambda^n \;\to^{\mathbb{L}^1(Q_X)}\; 0. \tag{3.22}$$

Then (3.19) and (3.20) yield that $(\gamma^n)^2 \xi^n + \gamma^n \lambda^n \to \xi$ in $\mathbb{L}^1(Q_X)$, hence also $(\gamma^n)^2 \xi^n \to \xi$ in $\mathbb{L}^1(Q_X)$ by (3.22). Since V is bounded, we readily deduce that $(\gamma^n)^2 V \xi^n \to V\xi$ and $(\gamma^n)^2 V \to V$ in $\mathbb{L}^1(R)$ (use (3.19) again for the later). Furthermore $V^n \to 0$ in $\mathbb{L}^1(R)$ by (3.19), while we have $R(V^n|\xi^n|) \leq \sqrt{R(V^n)}\sqrt{R(V^n(\xi^n)^2)}$, which goes to 0 by (3.19) and (3.21): then $V^n \xi^n \to 0$ in $\mathbb{L}^1(R)$. Putting all these results together yields

$$W^n \;\to^{\mathbb{L}^1(R)}\; V, \qquad W^n \xi^n \;\to^{\mathbb{L}^1(R)}\; V\xi.$$

It readily follows that

$$R(W^n|\mathcal{P}'_n) - R(V|\mathcal{P}'_n) \;\to^{\mathbb{L}^1(R)}\; 0, \qquad R(\xi^n W^n|\mathcal{P}'_n) - R(V\xi|\mathcal{P}'_n) \;\to^{\mathbb{L}^1(R)}\; 0. \tag{3.23}$$

On the other hand, Bayes' rule yields

$$\xi'^n \;=\; Q_{X^n}(\xi^n|\mathcal{P}'_n) \;=\; \frac{R(\xi^n W^n|\mathcal{P}'_n)}{R(W^n|\mathcal{P}'_n)}. \tag{3.24}$$

Now let us apply the proof of Theorem 3.1 to R instead of Q_X: first with V instead of ξ, which, since V is bounded, yields $R(V|\mathcal{P}'_n) \to V$ in $\mathbb{L}^2(R)$. Next with ξV instead of ξ, which, since $V \leq a$ and thus $R((\xi V)^2) \leq aR((\xi)^2 V) = aQ_X((\xi)^2) < \infty$, yields $R(\xi V|\mathcal{P}'_n) \to \xi V$ in $\mathbb{L}^2(R)$. Combining this with (3.23) yields

$$R(W^n|\mathcal{P}'_n) \;\to\; V, \quad R(\xi^n W^n|\mathcal{P}'_n) \;\to\; \xi V \qquad \text{in } \mathbb{L}^1(R), \text{ hence also in } Q_X\text{-measure}.$$

Since further we have $V > 0$ Q_X-a.s., it follows from (3.24) that $\xi'^n \to \xi$ in Q_X-measure, and we are done $\square$

In the previous theorem, we would like to replace (3.18) by (3.4), with X and X^n being only locally square-integrable martingales. But we have been unable to prove such a result under "reasonable" conditions.

3.4 Application to the Euler scheme

We apply the previous results to the Euler approximation scheme for a stochastic differential equation. The setting, similar to that of Subsection 2.4, is as follows: we have a locally square-integrable martingale Z on a space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ with $\mathcal{F} = \bigvee \mathcal{F}_t$ and a locally Lipschitz continuous function with linear growth g, and X is the (unique) solution of the following stochastic equation (where X_0 is a given $\mathcal{F}_0$-measurable square-integrable variable):

$$X_t \;=\; X_0 + \int_0^t g(X_{s-})dZ_s. \tag{3.25}$$

In comparison with Subsection 2.4, we relax the assumptions on g and Z and allow an arbitrary initial condition X_0. We also consider subdivisions $\tau_n = (T(n,i) : i \in I\!\!N)$ of stopping times satisfying (2.4), such that (3.1) holds. With $\phi_0^n = 0$ and $\phi_t^n = T(n, i-1)$ for $T(n, i-1) < t \leq T(n, i)$, we have the "continuous" Euler approximation at stage n, which is the solution of

$$X_t^n \;=\; X_0 + \int_0^t g(X_{\phi_s^n}^n)dZ_s. \tag{3.26}$$

Let U and U^n be square-integrable variables such that $U^n \to U$ in $I\!\!L^2$: typically $U = f(X_t)$ and $U^n = f(X_t^n)$ for some t, where f is a bounded continuous function. In this case, since by a well known result (see e.g. [11]) X^n goes in probability to X, locally uniformly in time, we do indeed have $U^n \to U$ in $I\!\!L^2$.

Note that X and X^n are locally square-integrable martingales. Recall also (1.3). Then as a corollary of Theorem 3.3 we get:

Theorem 3.6 *Let $U^n \to U$ in $I\!\!L^2(P)$, and let $\xi = \xi(X, U)$ and $\xi^n = \xi(X^n, U^n)$. Then $\xi^n \to^{Q_X} \xi$.*

Proof. It is enough to prove that (3.4) holds. Note

$$\begin{aligned}\langle X^n - X, X^n - X\rangle_t \;&=\; \int_0^t (g(X_{\phi_s^n}^n) - g(X_{s-}))^2 d\langle Z, Z\rangle_s \\ &\leq\; 2\int_0^t (g(X_{\phi_s^n}^n) - g(X_{\phi_s^n}))^2 d\langle Z, Z\rangle_s + 2\int_0^t (g(X_{\phi_s^n}) - g(X_{s-}))^2 d\langle Z, Z\rangle_s.\end{aligned}$$

We have already mentioned that X^n goes to X uniformly in time, in P-measure. Thus the sequence $\sup_{s\leq t}(|X_s| + |X_s^n|)$ is bounded in probability and, since g is continuous and locally bounded, it follows that the first term in the right side of the above inequality goes to 0 in probability for each t. On the other hand $\phi_s^n \to s$ and $\phi_s^n < s$ for all $s > 0$: thus for all ω and all $s > 0$ we have $X_{\phi_s^n}(\omega) \to X_{s-}(\omega)$. Thus, by the continuity of g again, the second term in the right side of the above inequality goes to 0 for all ω: hence (3.4) is proved. $\square$

Let us now pass to the "discrete" Euler approximation:

$$\bar{X}_i^n \;=\; X_{T(n,i)}^n. \tag{3.27}$$

Here we have some problems of integrability, because in order to apply the previous results we need each $\bar{X}^n$ to be a discrete-time locally square-integrable martingale. On the other hand we do not wish to assume that X, X^n and Z are square-integrable up to infinity.

In order to resolve this problem, we suppose that Z is a martingale square-integrable on compacts, and also that there is a constant K such that for all i, n:

$$T(n,i) - T(n,i-1) \;\leq\; K. \tag{3.28}$$

Then the process Z stopped at any time $T(n,i)$ (which is bounded by (3.28)), is a square-integrable martingale, and $\bar{Z}_i^n = Z_{T(n,i)}$ is a martingale square-integrable on compacts w.r.t. $(\mathcal{F}_{T(n,i)})_{i\geq 0}$. Due to the linear growth of g, and similarly to (5.2) of the Appendix, one also checks easily that $\bar{X}_i^n = X^n_{T(n,i)}$ is also a martingale square-integrable on compacts w.r.t. the filtration $(\mathcal{F}_{T(n,i)})_{i\geq 0}$.

As soon as U^n is square-integrable, analogous to (3.16), we may thus write

$$U^n \;=\; \begin{cases} \alpha^n + \sum_{i=1}^\infty \bar{\xi}_i^n \Delta \bar{X}_i^n + \bar{N}_\infty^n \\ \alpha^n + \sum_{i=1}^\infty \bar{\chi}_i^n \Delta \bar{Z}_i^n + \bar{N}'^n_\infty \end{cases} \tag{3.29}$$

where $\bar{N}^n$ (resp. $\bar{N}'^n$) is a square-integrable martingale w.r.t. $(\mathcal{F}_{T(n,i)})_{i\geq 0}$, null at 0 and orthogonal to the discrete time locally square-integrable martingale $(\bar{X}_i^n)_{i\geq 0}$ (resp. $(\bar{Z}_i^n)_{i\geq 0}$), and $\bar{\xi}_i^n$ and $\bar{\chi}_i^n$ are $\mathcal{F}_{T(n,i-1)}$-measurable. Further, we set

$$\left.\begin{array}{rcl} \xi'^n_t &=& \bar{\xi}_i^n \\ \chi'^n_t &=& \bar{\chi}_i^n \end{array}\right\} \qquad \text{if} \quad T(n,i-1) < t \leq T(n,i). \tag{3.30}$$

Recall that $\xi = \xi(X,U)$, and set $\chi = \xi(Z,U)$. Then:

Theorem 3.7 *Assume (3.28) and that Z is a martingale square-integrable on compacts. If the variables U^n and U are $\mathcal{F}_T$-measurable for some $T \in \mathbb{R}_+$ and satisfy $U^n \to U$ in $\mathbb{L}^2(P)$, we have $\xi'^n \to^{Q_X} \xi$ and $\chi'^n \to^{Q_Z} \chi$.*

Proof. 1) Take $T' = T + K$, where K occurs in (3.28). Then the processes χ^n are the same if we replace Z by the stopped process $Z^{T'}$ in (3.29), and also $\chi = \xi(Z^{T'}, U)$. So we can assume that $Z = Z^T$ is square-integrable. Applying Theorem 3.5 with $X^n = X = Z$ then yields that $\chi'^n \to \chi$ in Q_Z-measure.

2) In (3.29) we may write, in view of (3.26) and (3.27):

$$U^n \;=\; \alpha^n + \sum_{i=1}^\infty \bar{\xi}_i^n g(\bar{X}_{i-1}^n) \Delta \bar{Z}_i^n + \bar{N}_\infty^n,$$

$$U^n \;=\; \alpha^n + \sum_{i=1}^\infty \bar{\chi}_i^n 1_{\{g(\bar{X}_{i-1}^n)\neq 0\}} \Delta \bar{Z}_i^n + \sum_{i=1}^\infty \bar{\chi}_i^n 1_{\{g(\bar{X}_{i-1}^n) = 0\}} \Delta \bar{Z}_i^n + \bar{N}'^n_\infty.$$

The last three terms above are orthogonal martingales, and thus by identification with the previous expression we get that a.s.:

$$\bar{\xi}_i^n g(\bar{X}_{i-1}^n) \;=\; \bar{\chi}_i^n 1_{\{g(\bar{X}_{i-1}^n)\neq 0\}}. \tag{3.31}$$

This yields

$$\xi'^n_s g(X^n_{\phi^n_s}) \;=\; \chi'^n_s 1_{\{g(X^n_{\phi^n_s})\neq 0\}} \qquad Q_Z\text{-a.s.} \tag{3.32}$$

A similar argument shows that

$$\xi_s g(X_{s-}) \;=\; \chi_s 1_{\{g(X_{s-})\neq 0\}} \qquad Q_Z\text{-a.s.} \tag{3.33}$$

As seen in the proof of Theorem 3.6, $g(X^n_{\phi^n_s}) \to g(X_{s-})$ in probability for all s. Then one deduces from the fact that $\chi'^n \to \chi$ in Q_Z-measure and from (3.32) and (3.33) that $\xi'^n \to \xi$ in Q_Z-measure on the set $\{(\omega,t) : |g(X_{t-}(\omega)| \geq \varepsilon\}$, for every $\varepsilon > 0$. Hence the same convergence holds also on the set $A = \{(\omega,t) : g(X_{t-}(\omega)) \neq 0\}$, and since Q_X is absolutely continuous w.r.t. Q_Z and does not charge the complement of A, we deduce that $\xi'^n \to \xi$ in Q_X-measure. □

Remark 3.8 The same proof as above would also work for Theorem 3.6: we have $\chi^n = \xi(Z, U^n) \to \chi = \xi(Z, U)$ by (1.4), while the relation (3.32) holds between χ^n and ξ^n.

4 Weak convergence results

In this section we consider the weak convergence of integrands: we have a sequence X^n of locally square-integrable martingales, each defined on its own probability space $(\Omega^n, \mathcal{F}^n, (\mathcal{F}^n_t), P^n)$, and for each n a square-integrable variable U^n on the relevant space. The aim is to prove that if (X^n, U^n) converges in law to (X, U), with X a locally square-integrable martingale and U a square-integrable variable on the space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, then $(X^n, \xi(X^n, U^n))$ converges in law to $(X, \xi(X, U))$ in some sense.

It seems impossible to solve such a general problem, so we will concentrate on some particular cases.

4.1 Application of the Clark-Haussmann formula

Here we consider a sequence of processes of the form studied in Subsection 2.4. More precisely, we have Z, g and X^x as in this subsection, given on $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. For each n, we also have a Lévy process Z^n which is a martingale square-integrable on compacts on a space $(\Omega^n, \mathcal{F}^n, (\mathcal{F}^n_t), P^n)$, satisfying (2.15) with c_n and F_n, and as before, we assume that the numbers

$$\tilde{c} \;=\; c + \int F(dz) z^2, \qquad \tilde{c}_n \;=\; c_n + \int F_n(dz) z^2$$

are finite and strictly positive.

Then we have differentiable functions g_n, and we consider the equations (2.17) and (2.18) w.r.t. Z^n and g_n, and whose solutions are denoted by $X^{n,x}$ and $X'^{n,x}$. We make the following assumptions. First on g_n and g:

$$|g_n(0)| \;\leq\; K, \qquad |g'_n(x)| \;\leq\; K, \qquad |g'_n(x) - g'_n(y)| \;\leq\; K|x-y|, \tag{4.1}$$

$$g_n \;\to\; g, \qquad g'_n \;\to\; g' \qquad \text{pointwise.} \tag{4.2}$$

Next on Z^n and Z: we basically assume that Z^n converges in law to Z, plus a slightly stronger assumption which is reminiscent of the Lindeberg condition; more precisely we assume that

$$\left.\begin{array}{l} \tilde{c}_n \to \tilde{c}, \qquad \int F_n(dz)h(z) \to \int F(dz)h(z) \\ \text{for } h \text{ continuous, bounded and vanishing in a neighborhood of } 0, \end{array}\right\} \tag{4.3}$$

$$A(x) \;=\; \sup_n \int F_n(dz)z^2 1_{\{|z|\geq x\}} \;\to\; 0 \quad \text{as } x \to \infty. \tag{4.4}$$

These two conditions imply that the second convergence in (4.3) also holds when h is continuous, and $h(x) = O(x^2)$ at infinity, and $h(x) = o(x^2)$ at 0. They imply the convergence in law of Z^n to Z (see e.g. [9]).

Then we can state:

Theorem 4.1 *Assume (4.1), (4.2), (4.3) and (4.4). Let f be a differentiable function with a bounded and Lipschitz derivative and $T > 0$. The processes $\xi = \xi(X^x, f(X^x_T))$ and $\xi^n = \xi(X^{n,x}, f(X^{n,x}_T))$ have versions which are left continuous with right limits, and if we set $\xi(+)_s = \lim_{t\downarrow s, t>s} \xi_t$ and $\xi(+)^n_s = \lim_{t\downarrow s, t>s} \xi^n_t$, the processes $(X^{n,x}, \xi^n(+))$ converge in law for the Skorokhod topology on $\mathbb{R}^2$ to $(X^x, \xi(+))$.*

Proof. 1) A version of ξ is given by (2.20), with η given by (2.21). We wish to prove here that this version is left continuous with right limits. We can rewrite η as

$$\left.\begin{array}{rcl} k(s,y,z) &=& \int_0^1 (Q_{T-s}f'(y+uzg(y)) - Q_{T-s}f'(y))du, \\ \eta(s,y) &=& Q_{T-s}f'(y) + \frac{1}{\tilde{c}} \int F(dz)z^2 k(s,y,z). \end{array}\right\} \tag{4.5}$$

In view of (5.9) of the appendix and of the properties of f, we have for $0 \leq s \leq t \leq T$:

$$\left.\begin{array}{rcl} |k(s,y,z)| &\leq& C, \\ |k(s,y,z) - k(t,y,z)| &\leq& C(1+|y|(1+|z|))\sqrt{t-s} \end{array}\right\} \tag{4.6}$$

for a constant C. Now, (4.3) and (4.4) yield that $\int F(dz)z^2 1_{\{|z|\geq x\}} \leq A(x)$, so the above estimates and (5.9) again yield that for all $N > 1 \vee T^{-1/4}$ and for two other constants C', C'':

$$|\eta(s,y) - \eta(t,y)| \;\leq\; C'(1+N|y|)\sqrt{t-s} + CA(N) \;\leq\; C''(1+|y|)(t-s)^{1/4} + A((t-s)^{-1/4}) \tag{4.7}$$

(take $N = (t-s)^{-1/4}$ to get the last estimate). On the other hand, as in the proof of Theorem 2.6 we have (2.25) with Q_t instead of Q^n_t, hence it is clear from (4.5) and (4.7) and another application of (5.9) that $(s,y) \mapsto \eta(s,y)$ is continuous: hence (2.20) readily yields that ξ is left continuous with right limits.

2) Similarly, for each n we associate with Z^n, F_n, $\tilde{c}_n$, g_n the functions k_n and η_n given by (4.5). Exactly as before, we obtain that ξ^n, as given by (2.20) with η_n and $X^{n,x}$ instead of η and X^x, is left continuous with right limits. Moreover, in view of the Appendix and

of (4.1), (4.3) and (4.4), it is clear that the estimates (4.6) and (4.7) hold for all k_n and η_n with constants C, C' independent of n.

3) Now we apply again the stability results of [10]: by (4.1), (4.2) and (4.3), for any sequence $y_n \to y$, the processes (X^{n,y_n}, X'^{n,y_n}) converge in law to (X^y, X'^y), and further the estimate (5.2) of the Appendix yields that each sequence $(X_t'^{n,y_n})_{n\geq 1}$ is uniformly integrable. Hence if Q_t^n is associated with $(X^{n,x}, X'^{n,x})$ by (2.19) we readily deduce that (2.25) holds.

We will deduce that if $y_n \to y$ and $s_n \to s$ we have

$$\eta_n(s_n, y_n) \ \to \ \eta(s,y). \tag{4.8}$$

Indeed, by (2.25) and (5.9) of the Appendix, we have $Q^n_{T-s_n} f'(y_n) \to Q_{T-s} f'(y)$, hence also $k_n(s_n, y_n, z_n) \to k(s,y,z)$ as soon as $z_n \to z$ because of (4.2). Hence for (4.8) it remains to prove that if $h_n(z) = k_n(s_n, y_n, z)$ and $h(z) = k(s,y,z)$

$$\frac{1}{\tilde{c}_n} \int F_n(dz) z^2 h_n(z) \ \to \ \frac{1}{\tilde{c}} \int F(dz) z^2 h(z), \tag{4.9}$$

knowing that $h_n(z_n) \to h(z)$ if $z_n \to z$ and h is continuous and $|h_n| \leq C$ for a constant C. Now, consider the probability measures $G_n(dz) = \frac{1}{\tilde{c}_n}(F_n(dz)z^2 + c_n\varepsilon_0(dz))$ and $G(dz) = \frac{1}{\tilde{c}}(F(dz)z^2 + c\varepsilon_0(dz)$. Since $h_n(0) = h(0) = 0$, (4.9) reads as $G_n(h_n) \to G(h)$. Furthermore (4.3) and (4.4) imply that G_n converges weakly to G.

By the Skorokhod representation theorem we can find random variables V_n, V on a suitable probability space, such that V_n and V have laws G_n and G, and that $V_n \to V$ everywhere. Then $G_n(h_n) = E(h_n(V_n))$ and $G(h) = E(h(V))$, and the fact that $h_n(z_n) \to h(z)$ if $z_n \to z$ yields that $h_n(V_n) \to h(V)$ everywhere. Since further $|h_n| \leq C$, it follows that $G_n(h_n) \to G(h)$: hence (4.9) and (4.8) are proved.

4) Observe that $\xi(+)^n_s = \eta_n(s, X_s^{n,x}) 1_{[0,T)}(s)$ and $\xi(+)_s = \eta(s, X_s^x) 1_{[0,T)}(s)$. Further, (4.8) implies that $\eta_n \to \eta$ locally uniformly. Since $X^{n,x}$ converges in law to X^x and since X^x has no fixed time of discontinuity, an application of the continuous mapping theorem yields that $(X^{n,x}, \xi(+)^n)$ converges in law for the Skorokhod topology to $(X^x, \xi(+))$. □

Remark 4.2 Suppose now that f is a continuously differentiable function on $\mathbb{R}^k$ with all partial derivatives bounded and Lipschitz, and let $0 < T_1 < \ldots < T_k$. Set $\xi = \xi(X^x, f(X^x_{T_1}, \ldots, X^x_{T_k}))$ and $\xi^n = \xi(X^{n,x}, f(X^{n,x}_{T_1}, \ldots, X^{n,x}_{T_k}))$, as in Remark 2.9. Then the statement of Theorem 4.1 holds, with exactly the same proof.

4.2 A discrete time version

Here we consider a "discrete time" version of the previous results. The setting is as follows, and will also be the same in the next subsection.

For each n we have a sequence $(Y_i^n)_{i\geq 1}$ of i.i.d. variables on a given space $(\Omega^n, \mathcal{F}^n, P^n)$, with

$$E^n(Y_i^n) \ = \ 0, \qquad E^n((Y_i^n)^2) \ = \ \frac{1}{n}, \qquad E^n((Y_i^n)^4) \ \leq \ \frac{\varepsilon_n}{n}, \tag{4.10}$$

where $\varepsilon_n \to 0$. These conditions imply that the partial sums processes

$$Z_t^n = \sum_{i=1}^{[nt]} Y_i^n \tag{4.11}$$

converge weakly to a standard Wiener process $Z = W$, defined on a (possibly different) filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. We also have a function g on $I\!R$ which is differentiable with a bounded Lipschitz derivative, and we consider the difference equation

$$\bar{X}_0^{n,x} = x, \qquad \bar{X}_i^{n,x} = \bar{X}_{i-1}^{n,x} + g(\bar{X}_{i-1}^{n,x}) Y_i^n, \tag{4.12}$$

whose solution is a square-integrable martingale w.r.t. the discrete-time filtration $\mathcal{F}_i^n = \sigma(Y_j^n : j \leq i)$. We also consider the associated continuous-time martingale w.r.t. the filtration $(\mathcal{F}_{[nt]}^n)_{t\geq 0}$:

$$X_t^{n,x} = \bar{X}_{[nt]}^{n,x}. \tag{4.13}$$

This process $X^{n,x}$ can be viewed as the solution of the stochastic differential equation

$$X_t^{n,x} = x + \int_0^t g(X_{s-}^{n,x}) dZ_s^n, \tag{4.14}$$

and by stability theorems (see [10]) it converges weakly to the unique strong solution of the following equation:

$$X_t^x = x + \int_0^t g(X_s^x) dZ_s. \tag{4.15}$$

We even have that the pair $(Z^n, X^{n,x})$ weakly converges to (Z, X^x). Further $X^{n,x}$ and $\bar{X}^{n,x}$ are also related by (2.5) with $T_i = i/n$, and $X^{n,x}$ is a locally square-integrable martingale.

Now we let $T > 0$ and f be a differential function with a bounded and Lipschitz derivative. Then $U^n = f(X_T^{n,x})$ is square-integrable. We can consider the decomposition (3.16), which gives $\bar{\xi}_i^n$, and we associate ξ'^n as in (3.17) with $T(n,i) = i/n$. On the other hand $U = f(X_T^x)$ is also square-integrable, and we set $\xi = \xi(X^x, U)$.

Here again, by construction ξ'^n is left continuous with right limits, and we set $\xi(+)_s'^n = \lim_{t\downarrow s, t>s} \xi_s'^n$. On the other hand, the version of ξ given by Theorem 2.6 is not only left continuous, but even continuous except at time T: this is because the function η of (2.21) is continuous, and the process X^x also is continuous: then the process $\xi_s = \eta(s, X_s^x) 1_{[0,T)}(s)$ is another version of ξ, which is right continous with left limits (and also continuous except at T) and differs from the first version at time T only.

Theorem 4.3 *Assume (4.10), (4.11), (4.12), (4.14) and (4.15) with g differentiable with a bounded and Lipschitz derivative. Let f be a differentiable function with a bounded and Lipschitz derivative and let $T > 0$. Then the processes $(X^{n,x}, \xi(+)'^n)$ converge in law for the Skorokhod topology on $I\!R^2$ to (X^x, ξ).*

Proof. The explicit form of ξ is given by (2.20), with η taking the simple form $\eta(s,y) = Q_{T-s} f'(y)$. Now, if $P_t^n f(x) = E(f(X_t^{n,x}))$, we readily deduce from Proposition 2.1 and from (4.10) and (4.12) and (4.13) that a version of $\bar{\xi}_i^n$ is given by

$$\xi_i^n = \begin{cases} \frac{n}{g(\bar{X}_{i-1}^{n,x})} \int \mu_i^n(dy) y P_{T-\frac{i}{n}}^n f(\bar{X}_{i-1}^{n,x} + g(\bar{X}_{i-1}^{n,x}) y) & \text{if } \frac{i}{n} < T, \quad g(\bar{X}_{i-1}^{n,x}) \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where μ_i^n denotes the law of Y_i^n. In view of the properties of g, we readily deduce by induction on i that $x \mapsto \bar{X}_i^{n,x}$ is differentiable (for all ω), hence $x \mapsto X_t^{n,x}$ is also differentiable and its derivative satisfies

$$X_t'^{n,x} = 1 + \int_0^t g'(X_{s-}^{n,x}) X_{s-}'^{n,x} dZ_s^n,$$

and we set $Q_t^n f(x) = E(f(X_t^{n,x}) X_t'^{n,x})$. Then by virtue of (4.10) and of the properties of g again, one easily checks that $Q_t^n f'(y)$ is bounded in (n,y) and continuous in y, and that $\frac{\partial}{\partial y} P_t^n f(y) = Q_t^n f'(y)$; since further the Y_i^n's are centered, we get

$$\int \mu_i^n(dy) y P_{T-\frac{i}{n}}^n f(\bar{X}_{i-1}^{n,x} + g(\bar{X}_{i-1}^{n,x})y) = \frac{g(\bar{X}_{i-1}^{n,x})}{n} Q_{T-\frac{i}{n}}^n f'(\bar{X}_{i-1}^{n,x}) + \varepsilon_i^n,$$

where $\sup_i |\varepsilon_i^n| \to 0$. Therefore if $\phi_n(s) = i/n$ when $i/n < s \leq (i+1)/n$, we deduce that a version of $\xi(+)'^n$ is given by

$$\xi(+)_s'^n = Q_{T-\phi_n(s)}^n f'(X_{\phi_n(s)}^{n,x}) 1_{[0,\phi_n(T)]}(s) + \xi_s''^n,$$

where $\sup_s |\xi_s''^n| \to 0$. By the same argument as in Theorem 4.1 one has $Q_{s_n}^n f'(y_n) \to Q_s f'(y)$ when $s_n \to s$ and $y_n \to y$. Since $X^{n,x}$ converges in law to X^x, the result then follows as in Theorem 4.1 again. □

Remark 4.4 Exactly as in Remark 4.2, the same result holds when instead of $f(X_T^x)$ and $f(X_T^{n,x})$ we consider the variables $f(X_{T_1}^x, \ldots, X_{T_k}^x)$ and $f(X_{T_1}^{n,x}, \ldots, X_{T_k}^{n,x})$, where f is a continuously differentiable function on $\mathbb{R}^k$ with all partial derivatives bounded and Lipschitz, and $0 < T_1 < \ldots < T_k$.

4.3 Another discrete time version

Here we consider exactly the same setting as in the previous subsection: we have (4.10), (4.11), (4.12), (4.13), (4.14) and (4.15).

The only two differences are that we only assume g to be locally Lipschitz with at most linear growth, and that we will prove a convergence theorem for more general variables than $f(X_T^x)$, but in a much weaker sense.

More precisely, we consider a function Φ on the Skorohod space $\mathbb{D}$ of all right continuous with left limits functions on $\mathbb{R}_+$, which is bounded, continuous for the local uniform topology, and measurable w.r.t. the σ-field $\mathcal{D}_T$ generated by the coordinates on $\mathbb{D}$ up to some time $T > 0$ (recall that if Φ is continuous for the Skorokhod topology, it is a fortiori continuous for the local uniform topology). Then we take $U = \Phi(X^x)$ and $U^n = \Phi(X^{n,x})$.

For each n we can write the decomposition (3.29) for U^n, and define the continuous time processes ξ'^n and χ'^n by

$$\left.\begin{array}{rcl} \xi_t'^n &=& \bar{\xi}_i^n \\ \chi_t'^n &=& \bar{\chi}_i^n \end{array}\right\} \quad \text{if} \quad \frac{i-1}{n} < t \leq \frac{i}{n}. \tag{4.16}$$

Set also $\xi = \xi(X^x, U)$ and $\chi = \xi(Z, U)$.

To finish with our preliminaries, we need to introduce the topology w.r.t. which our processes will converge. We write $I\!\!F$ for the space of all Borel functions on $I\!\!R_+$, and Λ for the set of all bijective increasing maps from $I\!\!R_+$ into itself (= the set of continuous time-changes). We define the probability measure ρ on $I\!\!R_+$ by $\rho(dt) = e^{-t}dt$, and denote by Δ a distance metrizing the convergence in ρ-measure. Then for $x, y \in I\!\!F$ and with Id denoting the identity map on $I\!\!R_+$, we set

$$d_0(x,y) \;=\; \inf_{\lambda\in\Lambda}(\Delta(\lambda, Id) + \Delta(x\circ\lambda, y)), \qquad d(x,y) \;=\; d_0(x,y) + d_0(y,x). \tag{4.17}$$

This defines clearly a distance on $I\!\!F$, and a sequence x_n converges to x for this topology iff there is a sequence λ_n of time changes converging locally uniformly to Id and such that $\Delta(x_n\circ\lambda_n, y) \to 0$. This type of convergence is a weakening of convergence in Lebesgue measure, studied by many authors in the context of processes (see e.g. Grinblat [8], Cremers and Kadelka [4] or Meyer and Zheng [12]).

Finally, we endow the product $I\!\!D \times I\!\!D \times I\!\!F$ with the product of the local uniform topology on $I\!\!D$ and the topology induced on $I\!\!F$ by the distance d in (4.17). Then we have:

Theorem 4.5 *Assume (4.10), (4.11), (4.12), (4.14) and (4.15) with g locally Lipschitz with at most linear growth. The processes $(X^{n,x}, Z^n, \chi'^n)$ converge in law to (X^x, Z, χ) in the product space $I\!\!D \times I\!\!D \times I\!\!F$ with the above topology.*

If further the function $s \mapsto g(X^x_s)$ does not vanish the processes $(X^{n,x}, Z^n, \xi'^n)$ converge in law to (X^x, Z, ξ) in the same space.

Proof. 1) The idea of the proof is to embed in the Skorohod sense the random walk in the Wiener process.

Our basic space here will be $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ on which the Wiener process Z is defined, as well as the solution X^x of (4.15). By Skorohod embedding (see e.g. Skorokhod [16] or Azéma and Yor [1]), for each n we can find an increasing sequence $(T(n,i))_{i\geq 0}$ of stopping times with $T(n,0) = 0$ and such that if $S(n,i) = T(n,i) - T(n,i-1)$, the variables $(S(n,i), Z_{T(n,i)} - Z_{T(n,i-1)})_{i\geq 1}$ are independent and $Z_{T(n,i)} - Z_{T(n,i-1)}$ has the same law as Y^n_i, and further (compare to (4.10))

$$E(S(n,i)) \;=\; \frac{1}{n}, \qquad E(S(n,i)^2) \;\leq\; \frac{4\varepsilon_n}{n}. \tag{4.18}$$

In other words, since we are interested in convergence in law only and since thus the concrete realization of the variables Y^n_i does not matter, we can and will assume that $Y^n_i = Z_{T(n,i)} - Z_{T(n,i-1)}$. Then the process Z^n of (4.11) becomes $Z^n_t = Z_{T(n,[nt])}$. The solutions of (4.12), (4.13) and (4.15) are all defined on the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$, w.r.t. the same Z, as well as $U = \Phi(X^x)$ and $U^n = \Phi(X^{n,x})$, and thus also $\bar{\xi}^n_i$, $\bar{\chi}^n_i$, ξ'^n, χ'^n, ξ and χ.

2) Set $\Lambda^n_t = T(n,[nt])$ and $\phi^n_t = T(n,i-1)$ if $T(n,i-1) < t \leq T(n,i)$. Note that in (4.16) the time discretization is along the sequences i/n, while with the above representation of the Y^n_i's it is rather related to the sequences $T(n,i)$. This leads us to consider the equation

$$V^{n,x}_t \;=\; x + \int_0^t g(V^{n,x}_{\phi^n_s})dZ_s, \tag{4.19}$$

which is the Euler approximation of (4.15) along the $T(n,i)$'s. Note that $V^{n,x}_{T(n,i)} = \bar{X}^{n,x}_i$ (see (4.12)), hence

$$X^{n,x}_t \;=\; V^{n,x}_{\Lambda^n_t}. \tag{4.20}$$

Similarly, we will replace χ'^n by

$$\psi'^n_t \;=\; \bar{\chi}^n_i \qquad \text{if } \; T(n,i-1) < t \leq T(n,i), \tag{4.21}$$

so if Λ'^n is such that $\Lambda'^n_{i/n} = T(n,i)$ and is linear on each interval $(\frac{i-1}{n}, \frac{i}{n})$, we have

$$\chi'^n_t \;=\; \psi'^n_{\Lambda'^n_t}. \tag{4.22}$$

In the sequel we can assume without loss of generality that T is an integer. If $t \leq T$ we have $\Lambda^n_t \leq T(n,nT)$, so (4.20) implies that $U^n = \Phi(X^{n,x})$ is $\mathcal{F}_{T(n,nT)}$-measurable: it follows from (3.29) that $\bar{\chi}^n_i = 0$ for $i > nT$, and $\psi'^n_t = 0$ for $t \geq T(n,nT)$. Similarly, U is $\mathcal{F}_T$-measurable and we have $\chi_t = 0$ for $t > T$. Therefore for defining the process ψ'^n we can use the stopped process $H^n_t = Z_{t\wedge T(n,nT)}$, and for the process χ we can use the stopped process $H_t = Z_{t\wedge T}$.

Therefore, ψ'^n and χ are associated with H^n and H exactly as ξ'^n and ξ are associated with X^n and X in Theorem 3.5. So we will deduce from this theorem that

$$\psi'^n \;\to\; \chi \qquad \text{in } Q_H\text{-measure}, \tag{4.23}$$

provided we prove that

$$E(|U^n - U|^2) \;\to\; 0, \qquad E(\langle H^n - H, H^n - H\rangle_\infty) \;\to\; 0. \tag{4.24}$$

3) Recalling (4.18) and the independence of the $S(n,i)$'s for $i \geq 1$, we have that the mean of $T(n,[nt])$ is $\frac{[nt]}{n}$, and its variance is smaller than $4t\varepsilon_n$: therefore we have $T(n,[nt]) \to t$ in $\mathbb{L}^2(P)$. Therefore

$$\Lambda^n_t \;\to\; t, \quad \Lambda'^n_t \;\to\; t \qquad \text{locally uniformly in } t \text{ in } \mathbb{L}^2(P). \tag{4.25}$$

As already mentionned, $V^{n,x}$ converges locally uniformly in probability to X^x, and the limit X^x is continuous: so (4.20) and (4.25) imply that $X^{n,x}$ also converges locally uniformly in probability to X^x. Since Φ is bounded and continuous for the local uniform topology we have the first half of (4.24). As for the second half, since $\langle Z, Z\rangle_t = t$, it amounts to $E(|T - T(n,[nt])|) \to 0$: this is again a consequence of (4.25), hence (4.24) and (4.23) holds. Furthermore, since $\chi_t = 0$ for $t > T$ and $\psi'^n_t = 0$ for $t > T(n,nT)$, and since $Q_Z(d\omega, dt) = P(d\omega) \otimes dt$, we readily deduce from (4.23) and (4.25) that we even have

$$\psi'^n \;\to\; \chi \qquad \text{in } Q_Z\text{-measure}. \tag{4.26}$$

Now $Z^n \to Z$ locally uniformly for all ω, and $X^{n,x} \to X^x$ locally uniformly in probability as seen above. Finally, (4.25) and (4.22) implies that $d(\chi'^n, \chi) \to 0$ in probability, where d is defined in (4.17). That is, $(Z^n, X^{n,x}, \chi'^n)$ converges in probability to (Z, X^x, χ) in $\mathbb{D} \times \mathbb{D} \times \mathbb{F}$ for the desired topology, and the first claim is proved.

4) For the second claim, we observe that, exactly as in the proof of Theorem 3.7, the relations (3.31) and (3.33) hold, and thus also

$$\xi'^n_s g(X^{n,x}_{s-}) \;=\; \chi'^n_s 1_{\{g(X^{n,x}_{s-})\neq 0\}} \qquad Q_Z\text{-a.s.} \tag{4.27}$$

We have also seen that $X^{n,x} \to X^x$ locally uniformly in probability. So the second claim readily follows from the first one and from (3.33) and (4.27). □

Remark 4.6 The second claim is not very satisfactory, since it assumes that $g(X^x)$ does not vanish. If $S = \inf(t : g(X^x_t) = 0)$ is not everywhere infinite, then X^x is constant after S, and in the above proof we have the convergence of ξ'^n to ξ on the set $[0, S]$, but not necessarily on (S, ∞): when we go back to the original sequence Y^n_i and the original processes $X^{n,x}$, defined on possibly different spaces, one can no longer compare ξ'^n and ξ "pathwise", and the convergence in law "in restriction to $[0, S]$" makes no sense.

This is in contrast with Theorem 4.3, in which we obtained the convergence in law without restriction. Another difference with this theorem is that here the convergence of χ'^n and ξ'^n is in a much weaker sense, because the limiting processes χ and ξ are no longer left continuous with right limits. □

Remark 4.7 When the variables Y^n_i are $\mathcal{N}(0, \frac{1}{\sqrt{n}})$, the embedding in the previous proof is trivially realized with $T(n,i) = i/n$. Then there is no time-change involved: the convergence in law takes place in $\mathbb{D} \times \mathbb{D} \times \mathbb{F}$, with $\mathbb{F}$ endowed with the topology of convergence in Lebesgue measure. It does not seem to be true in general. □

Remark 4.8 The conditions (4.10) are far too strong for this result. In fact, Theorem 4.5 remains valid if the Y^n_i's have

$$E^n(Y^n_i|\mathcal{F}^n_{i-1}) \;=\; 0, \qquad E^n((Y^n_i)^2|\mathcal{F}^n_{i-1}) \;=\; \delta^n_i, \qquad E^n((Y^n_i)^4|\mathcal{F}^n_{i-1}) \;\leq\; \frac{\varepsilon_n}{n}, \tag{4.28}$$

where $\mathcal{F}^n_i = \sigma(Y^n_j : j \leq i)$ and $\varepsilon_n \to 0$ (the ε_n are constants) and the variables δ^n_i satisfy $\sum_{i=1}^{[nt]} \delta^n_i \to t$ in law for each t. The proof is almost the same: observing that δ^n_i is a function $h^n_i(Y^n_1, \ldots, Y^n_{i-1})$, the only difference is that the first equality in (4.18) is replaced by $E(S(n,i)|\mathcal{F}_{T(n,i-1)}) = h^n_i(S(n,1), \ldots, S(n,i-1))$ (using also the fact that embedding a random variable depending measurably on a parameter gives rise to a stopping time depending also measurably on this parameter, as is the case in the construction of Azema and Yor [1]). □

Remark 4.9 One could perhaps also consider the case of i.i.d. variables Y^n_i (or more generally triangular arrays of martingale increments, as in Remark 4.8) such that the processes Z^n of (4.11) converge in law to a Lévy process Z: this would probably require the embedding technique of Monroe [13], but we have not tried to do this. □

5 Appendix: some complements on stochastic differential equations

Here we gather some results about Equation (2.17). First, assume that Z is a Lévy process and a locally square-integrable martingale, so that $\langle Z, Z\rangle_t = \tilde{c}t$ for some $\tilde{c} > 0$. We are also given a differentiable function g, such that

$$|g(0)| \ \leq \ K, \qquad |g'(x)| \ \leq \ K. \tag{5.1}$$

In this case both (2.17) and the linear equation (2.18) have unique (strong) solutions. We have the following estimates, which rely upon Gronwall's Lemma and the property $\langle Z,Z\rangle_t = \tilde{c}t$:

$$E(\sup_{s\leq t}|X^x_s|^2) \ \leq \ (2x^2+1)\exp(4K^2\tilde{c}t), \qquad E(\sup_{s\leq t}|X'^x_s|^2) \ \leq \ 2\exp(2K^2\tilde{c}t). \tag{5.2}$$

Second, we prove the following lemma, which is less well known than the previous results:

Lemma 5.1 *Assume that the Lévy process Z above has bounded jumps, and that the coefficient g is infinitely differentiable with bounded derivatives of all order, and define P_t and Q_t by (2.19). Then for every twice continuously differentiable function f which is bounded as well as its two first derivatives, the function $(t,x)\mapsto P_tf(x)$ is twice differentiable in x and once differentiable in t, and all the partial derivatives are continuous in (t,x), and*

$$\frac{\partial}{\partial x}P_tf(x) \ = \ Q_tf'(x). \tag{5.3}$$

Proof. 1) In addition to (2.17) and (2.18), consider also the linear equation

$$X''^x_t \ = \ \int_0^t \left(g''(X^x_{s-})(X'^x_{s-})^2 + g'(X'^x_{s-})X''^x_{s-}\right)dZ_s. \tag{5.4}$$

Then we have the following properties, to be proved below:

$$\text{the maps } x\mapsto X^x_t,\ X'^x_t,\ X''^x_t \text{ are differentiable in } \mathbb{L}^2(P), \tag{5.5}$$

$$\text{the derivatives of } x\mapsto X^x_t,\ X'^x_t \text{ are } X'^x_t \text{ and } X''^x_t \text{ respectively,} \tag{5.6}$$

$$\text{the variables } (|X^x_t|^2, |X'^x_t|^2, |X''^x_t|^2)_{t\in[0,T]} \text{ are uniformly integrable.} \tag{5.7}$$

We readily deduce from these properties that (5.3) holds and moreover

$$\frac{\partial^2}{\partial x^2}P_tf(x) \ = \ E(f''(X^x_t)(X'^x_t)^2 + f'(X^x_t)X''^x_t) \tag{5.8}$$

hold. Further, the processes X^x, X'^x and X''^x are continuous in time, in probability: hence (5.6) and (5.7), together with (2.19) and (5.8), readily imply that $P_tf(x)$ and its two first derivatives in x are continuous in (t,x).

Moreover it is well known that f belongs to the domain of the infinitesimal generator $\mathcal{A}$ of (P_t), and

$$\mathcal{A}f(x) \ = \ \frac{c}{2}g(x)^2f''(x) + \int F(dz)(f(x+g(x)z) - f(x) - f'(x)g(x)z).$$

Hence $\frac{\partial}{\partial t}P_tf(x) = P_t\mathcal{A}f(x)$ exists and is continuous in (t,x), because $\mathcal{A}f$ is bounded and continuous.

2) It remains to prove that (5.5), (5.6) and (5.7) hold. For this, we will apply some results of Chapter 5 of [2]. The continuous martingale part of Z is $Z^c = cW$ for some $c \geq 0$ and a Wiener process W; let μ be the jump measure of Z, whose compensator is $\nu(dt, dx) = dt \otimes F(dx)$, where F is the Lévy measure of Z, which by hypothesis has compact support. The function $\eta(x) = x$ is thus in $\mathbb{L}^p(F)$ for all $p \geq 2$. Then we set for $\lambda \in \mathbb{R}$ and $y = (y_1, y_2, y_3) \in \mathbb{R}^3$:

$$H_t^{\lambda,1}(\omega) \;=\; x + \lambda, \qquad H_t^{\lambda,2}(\omega) \;=\; 1, \qquad H^{\lambda,3} \;=\; 0,$$

$$A^\lambda \;=\; 0,$$

$$B^{\lambda,1}(y,\omega,t) = g(y_1), \quad B^{\lambda,2}(y,\omega,t) = g'(y_1)y_2, \quad B^{\lambda,3}(y,\omega,t) = g''(y_1)(y_2)^2 + g'(y_1)y_3,$$

$$C^\lambda(y,\omega,t,z) \;=\; B^\lambda(y,\omega,t)z.$$

Then the set of the three equations (2.17), (2.18) and (5.4) for $x+\lambda$ instead of x reduces to Equation (5-22) of [2], with $Y^\lambda = (X^{x+\lambda}, X'^{x+\lambda}, X''^{x+\lambda})$. The assumptions of Theorem 5-24 of [2] are then obviously satisfied, and this theorem states that Y_t^λ is differentiable in all $\mathbb{L}^p(P)$ in λ, at $\lambda = 0$, and that the derivative is obtained by formal differentiation of the equation giving Y^λ, and that all powers of the derivatives $|Y_t'^0|^p$ are uniformly integrable when t runs through any finite interval: these properties imply (5.5), (5.6) and (5.7). □

Third, we give an estimate as in (5.2) for the kernel Q_t:

Lemma 5.2 *Assume (5.1). If f is a bounded Lipschitz function, for all $s \leq t$ we have with $C = \sup(|f(x)|, \frac{|f(x)-f(y)|}{|x-y|} : x, y \in \mathbb{R}, x \neq y)$:*

$$\left.\begin{array}{l} |Q_t f(x)| \;\leq\; 2C\; e^{K^2\tilde{c}t}, \\ |Q_t f(x) - Q_s f(x)| \;\leq\; 8KC\sqrt{1+x^2}\; e^{3K^2\tilde{c}t}\;\sqrt{t-s}. \end{array}\right\} \tag{5.9}$$

Proof. The first estimate in (5.9) follows from (5.2). For $s < t$ we have:

$$X_t^x - X_s^x \;=\; \int_s^t g(X_{r-}^x)dZ_r,$$

$$X_t'^x - X_s'^x \;=\; \int_s^t g'(X_{r-}^x)X_{r-}'^x dZ_r.$$

Hence by (5.1) and (5.2) we readily get

$$E(|X_t^x - X_s^x|^2) \;\leq\; 4K^2(1+x^2)e^{4K^2\tilde{c}t}(t-s),$$

$$E(|X_t'^x - X_s'^x|^2) \;\leq\; 2K^2 e^{2K^2\tilde{c}t}(t-s).$$

Now we write

$$Q_t f(x) - Q_s f(x) \;=\; E(f(X_t^x)(X_t'^x - X_s'^x)) + E((f(X_t^x) - f(X_s^x))X_s'^x),$$

and the second estimate in (5.9) follows from what precedes and from (5.2) again. □

References

[1] Azéma J. and Yor M.: Une solution simple au problème de Skorokhod. Sém. Proba. XIII, Lect. Notes in Math. **721**, 90-115, Springer Verlag: Berlin, 1979.

[2] Bichteler K., Gravereaux J-B. and Jacod J.: *Malliavin calculus for processes with jumps*, Gordon and Breach, London, 1987.

[3] Cinlar E., Jacod J., Protter P. and Sharpe M.: Semimartingales and Markov processes, *Z. Wahrsch. Verw. Geb.* **52**, 161-220, 1980.

[4] Cremers H. and Kadelka D.: On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in $\mathbb{L}_p^E$, *Stoch. Processes Appl.*, **21**, 305-317, 1986.

[5] Dellacherie C. and Meyer P.A.: *Probabilités et potentiel, Théorie des martingales*, Hermann: Paris, 1980.

[6] Duffie D. and Protter P.: The boundary between discrete and continuous time finance: weak convergence of the financial gain process, *J. Math. Finance*, **2**, 1-15, 1992.

[7] Föllmer H. and Sondermann D.: Hedging of non redundant contingent claims. In *Contributions to Mathematical Economics*, eds W. Hildebrand and A. Mas-Colell, 205-223, 1986.

[8] Grinblat L.S.: A limit theorem for measurable random processes and its applications, *Proceed. A.M.S.*, **61**, 371-376, 1976.

[9] Jacod J. and Shiryaev A.N.: *Limit theorems for stochastic processes*, Springer-Verlag: Berlin, 1987.

[10] Kurtz T.G. and Protter P.: Weak limit theorems for stochastic integrals and stochastic differential equations, *Annals Probab.*, **19**, 1035-1070, 1991.

[11] Kurtz T.G. and Protter P.: Wong Zakai corrections, random evolutions, and simulation schemes for SDE's, *Stochastic Analysis*, ed. Mayer-Wolf, Merzbach and Schwarz, AcademicPress: New York, 1991, 331-146.

[12] Meyer P.A. and Zheng W.A.: Tightness criteria for laws of semi-martingales, *Ann. Inst. H. Poincaré*, **20**, 371-376, 1984.

[13] Monroe I.: Processes that can be embedded in Brownian motion, *Annals Probab.*, **6**, 42-56, 1978.

[14] Nualart, D.: *Malliavin Calculus and Related Topics*, Springer Verlag: Berlin, 1995.

[15] Protter P.: *Stochastic Integration and Differential Equations*, Springer Verlag: Berlin, 1990.

[16] Skorokhod A.V.: *Studies in the theory of random processes*, Addison-Wesley: Reading, 1965.