On Numerical Schemes to Solve $dX_t = \sigma(X_t) \circ dW_t$ Under Commutative Noise*

Luis Jose Roman†

September, 2001

Abstract

First, we give a finite-difference scheme of global order h^2, in the mean-square sense, to solve numerically a class of stochastic differential equations (SDEs). This scheme is inspired on Runge-Kutta methods of order 4 in the case of deterministic ordinary differential equations (ODEs). It depends only on increments of the Wiener processes evaluated at the partition points and does not required evaluation of derivatives of the coefficients of the SDE. Second, we generalize this result to conclude that any numerical scheme of order $2n$ for ODEs becomes a scheme of order n for this class of SDEs. The significance of this paper is mainly appreciated when solving d-dimensional systems of SDEs which satisfy some restrictive conditions. Finally, some numerical examples are provided.

Key words and phrases. Stochastic differential equations, Stratonovich integrals, Runge-Kutta methods, Commutative noise, rate of convergence.

*1991 *Mathematics Subject Classification.* 34F05, 60H10, 65L06, 65L70.

†Department of Mathematics, University of California, Irvine, CA 92697-3875, USA. e-mail: roman@math.uci.edu.

1 Introduction

We first provide a motivation to this paper. Stochastic differential equations have been used to model systems subject to random influences and have become increasingly important in the analysis of a broad range of phenomena in natural sciences and economics. Many systems are described by differential equations where some of the parameters and/or the initial data are not known with complete certainty due to lack of information, uncertainty in the measurements or incomplete knowledge of the mechanisms themselves. In most cases we are concerned with temporal variation, and the state of the variable system at any particular instant t is described by an equation of the form:

$$dX_t = a(X_t)dt + \sum_{r=1}^{m} \sigma_r(X_t)dW_r(t) \tag{1}$$

where $a(x)$ and $\sigma_r(x)$ functions defined on $\mathbf{R}^d$ with values in $\mathbf{R}^d$ and the $W_{r'_s}(t)$ are one-dimensional independent Wiener processes. It has been of much interest to find ways of obtaining numerical approximations of (1), see [4], [5], [6], [7], [10], [13], [17], [19] and [20], for instance, but, the literature on that subject is much more extensive. The main idea of most of the methods currently being used are based on the stochastic Taylor expansion of (1), which involves derivatives of the functions $a(x)$, $\sigma_{r'_s}(x)$ and stochastic integrals. These present some complications in simulation and computations so it would be better to come up with other methods in which we do not have to deal neither with derivatives, or at least not many, nor stochastic integrals. On the other hand, typically, one of the main concerns of numerical analysts of deterministic differential equations have been to produce effective and accurate schemes where derivatives are avoided as much as possible, thus we can find a good source of candidates for our scheme in the literature of numerical analysis of deterministic ODEs, for example see [3], [23]. In particular, Runge-Kutta methods are well known and have been used for so many years already in solving numerically ordinary differential equations. The idea of the method is to replace derivatives of the solution of an ODE in its Taylor expansion by suitable iterative substitutions of the vector field. These "suitable" substitutions are found using the mean value theorem, the Taylor expansion of the solution and equating terms of the same order. Thus, it is reasonable to expect that similar methods can be translated into the stochastic case to solve stochastic differential equations and this is our main inspiration for the current paper. Runge-Kutta schemes for SDEs have been already introduced in [4], and [5], for a rather general class of SDEs. However, to achieve better order of accuracy they include additional random variables representing higher order stochastic integrals, and a cumbersome "rooted tree table" is introduced. Furthermore, in the best scenario they give, the best order of convergence obtained is $3/2$, and it is only attained in one dimensional cases. Even less interesting, they analyze the L_1-error whereas in most of the literature it seems to be of much interest the mean square error (L_2 sense) and the error of weak approximations. Here, we found that the classical Runge-Kutta scheme of order 4 for ODEs works perfectly as a scheme of order 2 in the mean square sense for certain type of SDEs, specifically, those of the form $dX_t = \sum_{r=1}^{m} \sigma_r(X_t) \circ dW_r(t)$, where "$\circ$" stands for stochastic integral in the Stratonovich sense, and satisfying the "commutativity noise condition", in other words $\Lambda_j \sigma_k = \Lambda_k \sigma_j, \ \ \forall 1 \leq j \leq k \leq m$. Moreover, this scheme does not use stochastic integrals. We will concentrate on approximations to (1) which have this last feature, in other words, depend only on evaluations of $W(t)$ at the points of the partition. However, for schemes depending only on evaluations of $W(t)$ at partitions points, it has been shown in [7] that, in general, no numerical method can guarantee accuracy along the trajectory, in the mean-square sense, of higher order than $O(h)$ in the 1-dimensional case, and in the multidimensional case, the best rate of convergence we can have is of the order $O(h^{1/2})$, unless a commutativity condition is satisfied, in which case order $O(h)$ can be achieved. Nevertheless, we prove in this paper that under commutativity noise and one more condition: no drift (or $a(x) = 0$), when written in the Stratonovich sense, any order of accuracy can be achieved. We emphasize, this result does not contradict the work on Cameron-Clark in [7], instead, it gives a smaller class of SDEs which can be approximated at any desired rate of convergence. Our approach here is to use the representation of solution of SDEs given by Doss in [9], to obtain and compare terms in the stochastic expansion of the solution. Doss' representation makes possible this proof in very simple and clear terms. Although, for this kind of SDEs, we can compute explicitly a closed form of the solution in the 1-dimensional case, the main value of this paper stands in its application to the multidimensional case. There, we do not have necessarily a closed form of the solution, thus numerical approximations are needed. We should mention that similar statement to our main result is given in [19], but no proofs or details of this fact are given. Also, in [6], they claim that this result is true without assuming the "no drift" condition (in the Stratonovich sense), therefore their claim is wrong as presented there. Finally, despite the various work in this subject, and the simplicity of the proof, the author is not aware at this time of any work where this observation has been proved other than in his previous works [21, 22].

2 Some Background from Numerical Approximation of SDEs

In order to derive an analyze numerical approximations to SDEs, we need a sort of “Taylor representation” for the solution of (1), and some results which grant us the convergence of of a given scheme as well as its order of accuracy. Indeed, we have all those tools and they are presented in the subsection below.

2.1 The One-step Approximation.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, let $\mathcal{F}_t$, $t_0 \leq t \leq t_0 + T$, be a nondecreasing family of σ-algebras of $\mathcal{F}$, and let $(W_r(t), \mathcal{F}_t)$, $r = 1, \ldots, m$, be an independent Wiener processes. Consider the *system of stochastic differential equations in the sense of Itô*:

$$dX_t = a(t, X_t)dt + \sum_{r=1}^{m} \sigma_r(t, X_t)dW_r(t) \tag{2}$$

where X, a, σ_r are vectors of dimension d. Assume that the functions $a(t, x)$ and $\sigma_r(t, x)$ are defined and continuous for $t \in [t_0, t_0 + T]$, $x \in \mathbf{R}^d$ and satisfy a Lipschitz condition: for all $t \in [t_0, t_0 + T]$, $x \in \mathbf{R}^d$, $y \in \mathbf{R}^d$ there is an inequality

$$|a(t, x) - a(t, y)| + \sum_{r=1}^{m} |\sigma_r(t, x) - \sigma_r(t, y)| \leq K|x - y|. \tag{3}$$

Here and below $|x|$ denotes the Euclidean norm of the vector x, and we denote by xy the scalar inner product of two vectors x and y. Let $(X(t), \mathcal{F}_t)$, $t_0 \leq t \leq t_0 + T$, be a solution of the system (2) with $\mathbf{E}|X(t_0)|^2 < \infty$. The one-step approximation $\overline{X}_{t,x}(t + h), t_0 \leq t \leq t + h \leq t_0 + T$, is defined as follows, and depends on x, t, h, and $\{W_1(\theta) - W_1(t), \ldots, W_m(\theta) - W_m(t) : t \leq \theta \leq t + h\}$:

$$\overline{X}_{t,x}(t + h) = x + \mathbf{A}(t, x, h; Ws, i = 1, \ldots, m, t \leq \theta \leq t + h) \tag{4}$$

where $\mathbf{A}$ is a function of t, x, h and $W_i(\theta) - W_i(t)$. An example of such $\mathbf{A}$ are given below. Using the one-step approximation we recurrently construct the approximations $(\overline{X}_k, \mathcal{F}_{t_k}), k = 0, \ldots, N, t_{k+1} - t_k = h_{k+1}, t_N = t_0 + T$:

$$\begin{aligned} \overline{X}_0 &= X_0 = X(t_0) \\ \overline{X}_{k+1} &= \overline{X}_{t_k, \overline{X}_k}(t_{k+1}) \\ &= \overline{X}_k + \mathbf{A}(t_k, \overline{X}_k, h_{k+1}; W_i(\theta) - W_i(t_k), i = 1, \ldots, m, t_k \le \theta \le t_{k+1}) \end{aligned} \tag{5}$$

In order to get a better understanding of the *One-step Approximation process* consider the following example:
Euler's Method:

$$\overline{X}_{t,x}(t, h) = x + a(t, x)h + \sum_{r=1}^{m} \sigma_r(t, x)\Delta_t W_r(h)$$

where $\Delta_t W_r(h) = W_r(t + h) - W_r(t)$. By (5), this approximation generates the scheme:

$$X_{k+1} = X_k + a_k h + \sum_{r=1}^{m} \sigma_{r_k} \Delta_k W_r(h) \tag{6}$$

where a_k, σ_{r_k} are the values of the coefficients a and σ_r at the point (t_k, X_k), and $\Delta_k W_r(h) = W_r(t_k + h) - W_r(t_k)$.

2.2 Expansion of the solution of a system of stochastic differential equations.

The reader interested in further details about the stochastic expansion of the solution of a SDE can find more explanations in [17, 29]. Here, we just provide the necessary terminology, tools, and results to understand our problem and some of the techniques we use. Let $X_{t,x}(s) = X(s)$ be the solution of the system (2), and let $f(t, x)$ be a sufficiently smooth (scalar or vector) function. By Itô's formula we have for $t_0 \le t \le \theta \le t_0 + T$:

$$f(\theta, X(\theta)) = f(t, x) + \sum_{r=1}^{m} \int_t^\theta \Lambda_r f(\theta_1, X(\theta_1)) dW_r(\theta_1) + \int_t^\theta L f(\theta_1, X(\theta_1)) d\theta_1 \tag{7}$$

where the operators $\Lambda_r, r = 1, \ldots, m$, and L are given by:

$$
\begin{aligned}
\Lambda_r &= (\sigma_r, \frac{\partial}{\partial x}), \\
L &= \frac{\partial}{\partial t} + (a, \frac{\partial}{\partial x}) + \frac{1}{2}\sum_{r=1}^{m}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma_r^i\sigma_r^j\frac{\partial^2}{\partial x^i \partial x^j}.
\end{aligned}
$$

Formula (7) is the analog of the Taylor Expansion of an ODE in the deterministic case.

Apply (7) to the functions $\Lambda_r f$ and Lf, and subsequently insert the expressions obtained for $\Lambda_r f(\theta, X(\theta))$ and $Lf(\theta, X(\theta))$ into (7). We find

$$
\begin{aligned}
f(s, X(s)) &= f + \sum_{r=1}^{m}\Lambda_r f \int_t^s dW_r(\theta) + Lf \int_t^s d\theta \\
&+ \sum_{r=1}^{m}\int_t^s \left(\sum_s^m \int_t^\theta \Lambda_s\Lambda_r f(\theta_1, X(\theta_1)) dW_s(\theta_1)\right) dW_r(\theta) \\
&+ \sum_{r=1}^{m}\int_t^s \left(\int_t^\theta L\Lambda_r f(\theta_1, X(\theta_1)) d\theta_1\right) dW_r(\theta) \\
&+ \sum_{r=1}^{m}\int_t^s \left(\int_t^\theta \Lambda_r L f(\theta_1, X(\theta_1)) dW_r(\theta_1)\right) d\theta \\
&+ \int_t^s \left(\int_t^\theta L^2 f(\theta_1, X(\theta_1)) d\theta_1\right) d\theta
\end{aligned} \tag{8}
$$

where, e.g., $\Lambda_r f$ and Lf are computed at (t, x). Continuing this way we obtain an expansion for $f(t+h, X(t+h))$. Such expansion is know as the Stochastic Taylor Expansion (STE, in short), or the Wagner-Platen expansion (see [29]). In a deterministic situation, this expansion is the classical Taylor expansion in powers of h, with remainder of integral type. However, in the stochastic case we observe, not only powers of h, but also, powers given by random variables of the form (they are independent of $\mathcal{F}_t$):

$$
I_{i_1,\ldots,i_j}(h) = \int_t^{t+h} dW_{i_j}(\theta) \int_t^\theta dW_{i_{j-1}}(\theta_1) \int_t^{\theta_1} \cdots \int_t^{\theta_{j-2}} dW_{i_1}(\theta_{j-1}) \tag{9}
$$

where $i_1, \ldots, i_j$ take values in the set $\{0, 1, \ldots, m\}$, and where $dW_0(\theta_r)$ is understood to mean $d\theta_r$.

Now, it is clear that for each *one-step approximation* there is an *error* committed. In other words, given a finite interval, without losing generality, let's say $[0, T]$, and suppose we want to approximate the solution of (2) at time $t = T$ by successive *one-step approximations*. Since at each step we make an error, a question arises: How large is the error carried up to time $t = T$?. The answer to this question is given by the next theorem below, but, let's first get familiar with notation. As usual, $X_{t_k,X}(t)$ denotes the solution of (2) for $t_k \leq t \leq t_0 + T$ satisfying the following initial condition at $t = t_k$: $X(t_k) = X$. By $\overline{X}_{t_k,X}(t_i), t_i \geq t_k$, we denote the approximation of the solution at step i and such that $\overline{X}_k = X$,(where $\overline{X}_k = \overline{X}(t_k)$). For example,

$$\overline{X}_{t_0,X_0}(t_{k+1}) = \overline{X}_{t_k,\overline{X}_k}(t_{k+1}) = \overline{X}_{k+1} \tag{10}$$

For simplicity reasons we assume that $t_{k+1} - t_k = h = \frac{T}{N}$. The theorem we state below is the one we used to determine the order of accuracy of a given scheme. Basically, it tell us what is the order of a scheme, based on the terms of the stochastic expansion of the solution. In what follows $\Lambda_0 = L$. Let

$$\bar{\imath}_k = \begin{cases} 0, & i_k = 0 \\ 1, & i_k \neq 0 \end{cases}$$

Then we have the following:

Theorem 1 *([13],[17])*

Suppose that $\overline{X}_{t,x}(t+h)$ *includes all terms of the form* $\Lambda_{i_1} \dots \Lambda_{i_j} f I_{i_1,\dots,i_j}$*, where* $f \equiv x$*, up to order* m *inclusive.*
Suppose that all functions $\Lambda_{i_1} \dots \Lambda_{i_j} f(t,x)$*, where* $f \equiv x$*,* $\sum_{k=1}^{j}(2 - \bar{\imath}_k)/2 \leq m + 1$*, satisfy* $\left|\Lambda_{i_1} \dots \Lambda_{i_j} f(t,x)\right| \leq K(1 + |x|^2)^{1/2}$*. Then the mean-square order of accuracy of the method based on this approximation is equal to* m*.*
Suppose that $\overline{X}_{t,x}(t+h)$ *includes all terms of the form* $\Lambda_{i_1} \dots \Lambda_{i_j} f I_{i_1,\dots,i_j}$*, where* $f \equiv x$*, up to order* $m + 1/2$ *inclusive, as well as the term*
$L^m a \int_t^{t+h} d\theta \int_t^{\theta} d\theta_1 \dots \int_t^{\theta_{m-1}} d\theta_m = L^m a h^{m+1}/(m+1)!$*. Suppose that all functions* $\Lambda_{i_1} \dots \Lambda_{i_j} f(t,x)$*, where* $f \equiv x$*,* $\sum_{k=1}^{j}(2 - \bar{\imath}_k)/2 \leq m + 2$*, satisfy*
$\left|\Lambda_{i_1} \dots \Lambda_{i_j} f(t,x)\right| \leq K(1 + |x|^2)^{1/2}$ *. Then the mean-square order of accuracy of the method based on this approximation is equal to* $m + 1/2$*.*

3 Runge-Kutta type schemes for certain type of SDEs (1-dimensional case).

To begin, let us consider a simpler version of (2), $d = 1$ and $m = 1$. The case for $d = 1$ and arbitrary m follows easily from this. Also, for simplicity, assume that $a(t, x)$ and $\sigma(t, x)$ are not time dependent. In fact, we will assume enough smoothness on the coefficients so that $a(t, x)$ and $\sigma(t, x)$ can be sought as functions on $\mathbf{R}^{d+1}$ satisfying a Lipschitz condition. Thus, we study a stochastic differential equation of the form:

$$\begin{cases} dX_t = a(X_t)dt + \sigma(X_t)dW_t \\ X(0) = X_0 \end{cases} \tag{11}$$

where $t \in [0, T]$, $a(x)$ and $\sigma(x)$ are real valued functions defined on $\mathbf{R}$ and satisfy the following Lipschitz condition:

$$|a(x) - a(y)| + |\sigma(x) - \sigma(y)| \leq K|x - y| \tag{12}$$

and W_t is a standard Brownian motion.
Under condition (12) it is proved that equation (11) with initial condition X_0, such that $\mathbf{E}|X_0|^2 < \infty$, has an unique solution $X_{(\cdot)} \in L_2(\Omega \times [0, T])$ (see [8]).

3.1 RK4 is a scheme of order 2 for SDEs of the form: $dX_t = \sigma(X_t) \circ dW_t$

Our next goal is to prove that the classical Runge-Kutta method of order 4, (RK4, in short), used in deterministic ODEs, works as a scheme of order of accuracy equal to 2, to solve numerically certain SDEs. Unfortunately, we cannot have the same result in the general case (see [7]), but, we show that when $a(x) = \frac{1}{2}\sigma(x)\sigma'(x)$, i.e., for equations of the form $dX_t = \sigma(X_t) \circ dW_t$, where "$\circ$" means stochastic integral in the Stratonovich sense, we actually have a scheme with the desired accuracy, which does not involves derivatives of $\sigma(x)$ and it is very easy to implement.
Next, we give the STE of order 2, of the solution of (2):

$$X_{t+h} = X_t + a(X_t)h + \sigma(X_t)\Delta_h W_t + \frac{1}{2}\sigma(X_t)\sigma'(X_t)\left((\Delta_h W_t)^2 - h\right)$$

$$
\begin{aligned}
&+L(\sigma)\int_t^{t+h}(\theta-t)dW(\theta)+\Lambda(a)\int_t^{t+h}(W(\theta)-W(t))d\theta\\
&+\Lambda^2(\sigma)\int_t^{t+h}\int_t^{\theta}\int_t^{\theta_1}dW(\theta_2)dW(\theta_1)dW(\theta)\\
&+L(a)\frac{h^2}{2}+\Lambda^3(\sigma)\int_t^{t+h}\int_t^{\theta}\int_t^{\theta_1}\int_t^{\theta_2}dW(\theta_3)dW(\theta_2)dW(\theta_1)dW(\theta)\\
&+L\Lambda(\sigma)\int_t^{t+h}\int_t^{\theta}\int_t^{\theta_1}d\theta_2 dW(\theta_1)dW(\theta)\\
&+\Lambda L(\sigma)\int_t^{t+h}\int_t^{\theta}\int_t^{\theta_1}dW(\theta_2)d\theta_1 dW(\theta)\\
&+\Lambda^2(a)\int_t^{t+h}\int_t^{\theta}\int_t^{\theta_1}dW(\theta_2)dW(\theta_1)d\theta
\end{aligned}
\tag{13}
$$

where $L := a\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}$ and $\Lambda := \sigma\frac{\partial}{\partial x}$ and all coefficients are evaluated at the point X_t.
Now, two of the terms of order $3/2$, namely the factors of $L(\sigma)$ and $\Lambda(a)$ respectively, can be combined using the identity $\int_t^{t+h}(\theta - t)dW(\theta) = h\Delta_h W_t - \int_t^{t+h}(W(\theta) - W(t))d\theta$ (see [17]), to obtain instead:

$$
L(\sigma)h\Delta_h W_t + (\Lambda(a) - L(\sigma))\int_t^{t+h}(W(\theta) - W(t))d\theta \tag{14}
$$

Notice that it is impossible to obtain the factor "$\int_t^{t+h}(W(\theta) - W(t))d\theta$" by just iterating in the arguments of $a(\cdot)$ and $\sigma(\cdot)$, as we are supposed to do when implementing Runge-Kutta of order 4 for ODEs, into the stochastic case. In fact, by iterating $a(\cdot)$ and $\sigma(\cdot)$, we would get an expansion of the form $p(h, \Delta_h W_t)$, where $p(x, y)$ is a polynomial of two variables of degree≤ 4. Therefore, if we expect this scheme to give a solution to (11) we better assume that:

$$
\Lambda(a)(x) - L(\sigma)(x) = 0 \tag{15}
$$

which in our case becomes:

$$
\sigma(x)a'(x) - a(x)\sigma'(x) = \frac{1}{2}\sigma^2(x)\sigma''(x) \tag{16}
$$

whose general solution is given by:

$$a(x) = \frac{1}{2}\sigma(x)\sigma'(x) + c\sigma(x) \tag{17}$$

where c is any arbitrary constant. In particular, if $c = 0$, we have $a(x) = \frac{1}{2}\sigma(x)\sigma'(x)$ so (11) turns out to be of the form:

$$dX_t = \frac{1}{2}\sigma(X_t)\sigma'(X_t)dt + \sigma(X_t)dW_t = \sigma(X_t) \circ dW_t \tag{18}$$

Now, it is a straightforward computation to proof that if $\Lambda(a)(x) - L(\sigma)(x) = 0$, then the following identities hold:

$$\begin{aligned} L\Lambda(\sigma) = \Lambda L(\sigma) &= \Lambda^2(a) = \frac{1}{2}\Lambda^3(\sigma) \\ \Lambda(a) &= \frac{1}{2}\Lambda^2(\sigma) \\ L(a) &= \frac{1}{4}\Lambda^3(\sigma) \end{aligned} \tag{19}$$

Define

$$\mathcal{I}_{i_1,\ldots,i_j}(h) := \int_t^{t+h} dW_{i_j}(\theta) \int_t^{\theta} dW_{i_{j-1}}(\theta_1) \int_t^{\theta_1} \cdots \int_t^{\theta_{j-2}} dW_{1_1}(\theta_{j-1}) \tag{20}$$

where the indices $i_1, \ldots, i_j$ can take either the value 0 (then $dW_0(\theta) = d\theta$) or 1 (then $dW_1(\theta) = dW(\theta)$).
Then (13) can be written as:

$$\begin{aligned} X_{t+h} &= X_t + \sigma(X_t)\Delta_h W_t + \frac{1}{2}\sigma(X_t)\sigma_x(X_t)(\Delta_h W_t)^2 \\ &\quad + \frac{1}{2}\Lambda^2(\sigma)h\Delta_h W_t + \frac{1}{2}\Lambda^3(\sigma)\left(\mathcal{I}_{0,1,1}(h) + \mathcal{I}_{1,1,0}(h) + \mathcal{I}_{1,0,1}(h)\right) \\ &\quad + \Lambda^2(\sigma)\mathcal{I}_{1,1,1}(h) + \frac{1}{4}\Lambda^3(\sigma)\frac{h^2}{2} \\ &\quad + \Lambda^3(\sigma)\mathcal{I}_{1,1,1,1}(h) \end{aligned} \tag{21}$$

by (19).
Moreover, it is not hard to show that

$$\mathcal{I}_{0,1,1}(h) + \mathcal{I}_{1,1,0}(h) + \mathcal{I}_{1,0,1}(h) = h\left(\frac{(\Delta_h W_t)^2 - h}{2}\right) \tag{22}$$

Also, after many mistakes, corrections and very tedious computations of some of the stochastic integrals in (21), it yields

$$\begin{aligned}\mathcal{I}_{1,1,1}(h) &= \frac{1}{6}\left(W^3(t+h) - W^3(t)\right) \\ &- \frac{1}{2}\Delta_h W_t \left(W^2(t) + h + W\Delta_h W_t\right)\end{aligned}$$

and

$$\begin{aligned}\mathcal{I}_{1,1,1,1}(h) &= \frac{1}{24}\left(W^4(t+h) + W^4(t)\right) \\ &- \frac{1}{4}\left(W^2(t+h)h - W^2(t)W^2(t+h) + W^2(t)h\right) \\ &+ \frac{1}{2}W(t)W(t+h)h - \\ &+ \frac{1}{2}\left(\frac{W(t)W^3(t+h)}{3} - \frac{W(t+h)W^3(t)}{3} + \frac{h^2}{4}\right)\end{aligned}$$

In these previous calculations, identities obtained by using the Hermite polynomials evaluated at $(t, \Delta_h W(t))$ are used, as well as construction of stochastic integrals in a recursive way (see [17]).
Thus, substituting the last two expressions above into (21) and after some long but straightforward computations we have:

$$\begin{aligned}X_{t+h} &= X_t + \sigma(X_t)\Delta_h W_t + \Lambda\sigma(X_t)\frac{(\Delta_h W_t)^2}{2!} + \Lambda^2\sigma(X_t)\frac{(\Delta_h W_t)^3}{3!} \\ &+ \Lambda^3\sigma(X_t)\frac{(\Delta_h W_t)^4}{4!}\end{aligned} \tag{23}$$

On the other hand, we perform the RK4 scheme as follows:

$$\begin{aligned}k_1 &= \sigma(x(t))\Delta_h W_t, \\ k_2 &= \sigma(x(t) + \frac{1}{2}k_1)\Delta_h W_t, \\ k_3 &= \sigma(x(t) + \frac{1}{2}k_2)\Delta_h W_t, \\ k_4 &= \sigma(x(t) + k_3)\Delta_h W_t, \\ x(t+h) &= x(t) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).\end{aligned} \tag{24}$$

Then, we compute the expansion of (24) of order 4 in $\Delta_h W_t$ using Maple V. If the reader feels insecure of Maple, the author has to say that doing the expansion by hand could be one of the most challenging and longest computations from which any learning or useful information is gained other than just applications of the mean value theorem over and over. Besides, the chances of making mistakes are enormous and it is a very tedious task.

After getting the expansion of (24), we observe that, in fact, it generates all terms of the stochastic Taylor expansion of X_t up to order 2, so by theorem 2, we conclude that RK4 has mean square order of accuracy equal to 2. Once again, we notice that the order of accuracy is reduced by half compared with Runge-Kutta of order 4 in the case of deterministic ordinary differential equations. Therefore, our result can be briefly described by:

Runge-Kutta of order 4 for ODEs becomes a scheme of order 2 for SDEs of the form: $dX_t = \sigma(X_t) \circ dW_t$, $X_t \in \mathbf{R}$.

3.2 Multidimensional Case

So far we have worked only in the case when X_t and W_t are one dimensional stochastic processes, and $a(x)$,$\sigma(x)$ real valued functions. In this section we consider the following equation:

$$dX_t = a(X_t)dt + \sigma(X_t)dW_t \tag{25}$$

where X_t is a d-dimensional stochastic process, $a : \mathbf{R}^d \mapsto \mathbf{R}^d$, $\sigma : \mathbf{R}^d \mapsto \mathcal{L}(\mathbf{R}^m, \mathbf{R}^d)$ (the space of linear functionals from $\mathbf{R}^m$ to $\mathbf{R}^d$), and W_t is a m-dimensional Wiener processes.

We also observe that (25) can be written as:

$$dX_t = a(X_t)dt + \sum_{r=1}^{m} \sigma_r(X_t)dW_r(t) \tag{26}$$

where

$$\sigma := \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \dots & \sigma_{dm} \end{pmatrix}$$

and

$$\sigma_r := \begin{pmatrix} \sigma_{1r} \\ \sigma_{2r} \\ \vdots \\ \sigma_{dr} \end{pmatrix} \quad W_t := \begin{pmatrix} W_1(t) \\ W_2(t) \\ \vdots \\ W_m(t) \end{pmatrix}$$

The $W_{i's}$ are m-independent real valued Wiener processes.

3.2.1 Main Result

We begin by giving the notation used in this part. For $u \in \mathbf{R}^d$ we denote by $u^{(i)}$, or u^{Δ}, any of its components. In particular, if $h(x,y) = (h^1(x,y), \ldots, h^d(x,y))$ then $h^{\Delta}(x,y)$ represents an arbitrary $h^i(x,y)$ for $1 \leq i \leq d$. We consider equation (25) as in the previous subsection, but now we require two more conditions. Thus, let us assume the following:

H1) $a(X)$ is Lipschitz.

H2) σ is in $\mathcal{C}^2(\mathbf{R}^d)$.

H3) *Commutative Noise*:

$$\sum_{i=1}^{d} \sigma_{ik} \frac{\partial \sigma_{lj}}{\partial x_i} = \sum_{i=1}^{d} \sigma_{ij} \frac{\partial \sigma_{lk}}{\partial x_i} \tag{27}$$

$\forall j,k,l: \quad 1 \leq j \leq m; 1 \leq k \leq m; 1 \leq l \leq d$, and where σ_{ij} denotes the (i,j) component of σ.

Also, consider the system of ODEs:

$$\begin{cases} \dfrac{\partial h^{(i)}}{\partial y_j}(x,y) = \sigma_{ij}(h(x,y)), \quad 1 \leq i \leq d, 1 \leq j \leq m \\ h(x,0) = x \end{cases} \tag{28}$$

Remark 1 *From (28), we can see that it might be very difficult, or sometimes impossible, to get a closed form of the function h. Further, even numerical approximations of h could be a difficult and cumbersome task.*

Remark 2 *H3 is equivalent to $\Lambda_j \sigma_k = \Lambda_k \sigma_j \;\; \forall 1 \leq k, j \leq m$.*

Further, let $D_t \in \mathbf{R}^d$ be the solution of the following system of ODEs:

$$\begin{cases} D_0 = X_0 \\ \dfrac{d}{dt} D_t^{\Delta} = \displaystyle\sum_{i=1}^{d} \left\{ \frac{\partial h^{\Delta}}{\partial x_i}(h(D_t, W_t), -W_t) \right\} \\ \qquad \times \left\{ a^i(h(D_t, W_t)) - \frac{1}{2} \displaystyle\sum_{r=1}^{d} \sum_{j=1}^{m} \frac{\partial \sigma_{ij}}{\partial x_r}(h(D_t, W_t)) \sigma_{rj}(h(D_t, W_t)) \right\} \end{cases} \tag{29}$$

where equation (29) is understood pathwise. Then, under conditions $H1$, $H2$ and $H3$, it is proved in [27] the following:

- (28) has unique solution $h(\cdot, \cdot) : \mathbf{R}^d \times \mathbf{R}^m \to \mathbf{R}^d$ and

$$\frac{\partial h}{\partial \beta}(\alpha, \beta) = \sigma\left(h(\alpha, \beta)\right) \tag{30}$$

  furthermore,

- $X_t = h(D_t, W_t)$.

They also showed, in [9] and [27], that if $\overline{D}_k$ is a scheme approximating the solution of (29) then $\overline{X}_k = h(\overline{D}_k, W_t)$ is an approximation of X_t at time t_k. We use this result to prove our main theorem:

Theorem 2 *Assume* **H1**, **H3**, *and* $\sigma_r \in C^{2n+1}(\mathbf{R}^d)$. *Besides, suppose that*

$$a(x) = \frac{1}{2} \sum_{r=1}^{m} \Lambda_r \sigma_r(x), \quad \forall x \tag{31}$$

where Λ_r *is the differential operator given by* $\Lambda_r = \left(\sigma_r, \frac{\partial}{\partial x}\right)$ *and finally, as in Theorem 2, assume that* $|\Lambda_{i_1} \ldots \Lambda_{i_j} f(t, x)| \leq K\left(1 + |x|^2\right)^{1/2}$ *where* $f \equiv x$ *and* $\sum_{k=1}^{j}(2 - \bar{\imath})/2 \leq n + 1$. *Then any numerical scheme of order* $2n$ *to approximate solutions of ODEs becomes a scheme of order* n *in the mean-square sense when approximating SDEs satisfying all the assumptions above.*

Remark 3 *By a numerical scheme of order* $2n$ *we mean that there exists an operator* $\Phi_h^{2n} : \mathcal{C}^{2n+1}\left(\mathbf{R}^d\right) \to \mathcal{C}\left(\mathbf{R}^d\right)$ *such that* $\Phi_h^{2n}(f)(x)$ *gives the multidimensional Taylor expansion of* f *at* x, *of order* $2n$, *in powers of* h, *and* $\overline{X}_{k+1} := \Phi_h^{2n}(f)(\overline{X}_k)$. *Also, the vectors* $\alpha = (\alpha_1, \ldots, \alpha_q)$, *whose components are nonnegative integers* α_k *are called multi-indices. We define for a multi-index* α, $|\alpha| := \alpha_1 + \ldots + \alpha_q$.

Proof:

In [9], it is proved that under these conditions there is a solution $h(x,y)$ to equation (28). Further, (31) implies that:

$$a^i(x) = \frac{1}{2}\sum_{r=1}^{d}\sum_{j=1}^{m}\frac{\partial\sigma_{ij}}{\partial x_r}\sigma_{rj}(x), \quad \forall i = 1,\ldots,d. \tag{32}$$

Hence the solution of the ODE (29) is given by $D_t = X_0$, for all t and moreover,

$$X_t = h(X_0, W_t) \tag{33}$$

By assumption we have that $\sigma \in \mathcal{C}^{2n+1}\left(\mathbf{R}^d\right)$ then applying Taylor's formula we have:

$$\begin{aligned} X_{t+h} - X_t &= h(X_0, W_{t+h}) - h(X_0, W_t) \\ &= \sum_{k=1}^{2n}\left(\sum_{|\alpha|=k} a_{\alpha_1\ldots\alpha_q}\mathcal{I}_1^{\alpha_1}(h)\ldots\mathcal{I}_m^{\alpha_q}(h)\right) + \rho^{2n} \end{aligned} \tag{34}$$

where $a_{\alpha_1\ldots\alpha_q} := \frac{\partial^{|\alpha|}h}{\partial\beta^{\alpha_1}\ldots\partial\beta^{\alpha_q}}$, evaluated at (X_0, W_t), $\alpha_i \in \{0,\ldots,2n\}$, for all $i \in \{0,\ldots,q\}$, and ρ^{2n} is a random variable defined by

$$\begin{aligned} \rho^{2n} &= \sum_{|\alpha|=2n+1}\int_0^1 \frac{\partial^{2n+1}h}{\partial\beta^{\alpha_1}\ldots\partial\beta^{\alpha_q}}((1-s)W_t + sW_{t+h}) \\ &\quad .(\Delta_h W_1)^{\alpha_1}\ldots(\Delta_h W_q)^{\alpha_q}s^{2n+1}ds \end{aligned} \tag{35}$$

which is just the reminder of the expansion. Moreover, by (30) and the growth condition in the statement, it yields

$$|\rho^{2n}| \leq K_\rho\left(1 + |W_t|^2 + |W_{t+h}|^2\right)\sum_{|\alpha|=2n+1}\mathcal{I}_1^{\alpha_1}(h)\ldots\mathcal{I}_m^{\alpha_q}(h) \tag{36}$$

where $\mathcal{I}_s(h)$ is defined as in section 2.2.
Since $\{W_1(t+h) - W_1(t), W_2(t+h) - W_2(t),\ldots,W_m(t+h) - W_m(t)\}$ are independent and each is a Gaussian process with mean zero and variance h it is easy to verify that:

$$\begin{aligned} \mathbf{E}|\rho^{2n}| &= 0 \\ \left(\mathbf{E}(\rho^{2n})^2\right)^{1/2} &\leq K(q,T,n)h^{n+1/2} \end{aligned} \tag{37}$$

The first expression follows easily since $|\alpha| = 2n+1$, and then at least one α_i must be an odd nonnegative integer. The second is obtained using the definition of the Euclidean norm, the fact that $t \in [0,T]$, and the order of each of the stochastic integrals $\mathcal{I}_s(h)$ (see [17]).
Notice that (34) shows that the only information we would need in order to approximate X_{t+h} are the increments of the Wiener processes at the partition points, i.e, $\Delta_h W_i(t_k)$. Therefore, any scheme which produces all the terms of the Taylor expansion of h up to order $2n$ will give the expansion of X_{t+h} in powers of $(W_1(t+h) - W_1(t))^{\alpha_1} \ldots (W_q(t+h) - W_q(t))^{\alpha_q}$, where $|\alpha| \le 2n$. Define $\overline{X}_{t_{k+1}} := \overline{X}_{t_k} + \Phi^{2n}_{\Delta_h W_{t_k}}(h)(W_{t_k})$ where $\Phi^{2n}_h(f)(x)$ is defined as in the remark above. Then $\overline{X}_{t+h}$ contains all the terms of the form $\Lambda_{i_1} \ldots \Lambda_{i_j} f I_{i_1,\ldots,i_j}$, where $f \equiv x$, up to order n (by definition of$\Phi^{2n}_h(f)(x)$), and by assumption we have $\sum_{k=1}^{j}(2-\bar{\imath})/2 \le n+1$, therefore by theorem 2 the scheme generated by using the one-step approximation process given above is of mean-square order n. This ends the proof.

$\square$

We make the remark that Doss' representation is what makes possible this proof in such a trivial terms. Now, in particular if we want to implement RK4 (used in the 1-dimensional case) for multidimensional equations, we need to know whether this method works for systems of first order equations. Indeed, it does work and some details can be found in [11] where it is proved that the method converges and still it is of order 4 under sufficiently smooth conditions on the vector field. Therefore, we can expect in the d-dimensional case that the classical Runge-Kutta method of order 4 also works for systems of SDEs, in other words it gives the Taylor expansions in terms of $\Delta W_t(h)$ up to the fourth power so that the method itself will be of order 2 as in the 1-dimensional case.

Remark 4 *As we said before, the importance of this paper arises mainly when we applied to solve systems of SDEs. For example, consider de following system:*

$$\begin{cases} dx(t) = (y^2(t) - x^2(t))dW_1(t) + 2x(t)y(t)dW_2(t); \quad x(0) = x_0 \\ dy(t) = -2x(t)y(t)dW_1(t) + (y^2(t) - x^2(t))dW_2(t); \quad y(0) = y_0 \end{cases} \tag{38}$$

where $(W_1(t), W_2(t))$ is a 2-dimensional Wiener process, and $\mathbf{E}|x_0|^2 < \infty$, $\mathbf{E}|y_0|^2 < \infty$. We know that solution of (38) exists for some time, in other words, there exists a $t' < T$, and a stochastic process $(x(t), y(t))$, for $t < t'$, such that equation (38) holds for all $t < t'$. This system can be re-written as

$$d \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y^2(t) - x^2(t) & 2x(t)y(t) \\ -2x(t)y(t) & y^2(t) - x^2(t) \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \tag{39}$$
$$= \begin{pmatrix} y^2(t) - x^2(t) \\ -2x(t)y(t) \end{pmatrix} dW_1(t) + \begin{pmatrix} 2x(t)y(t) \\ y^2(t) - x^2(t) \end{pmatrix} dW_2(t)$$

Straightforward computations reveals that $\sigma_1(x, y)$, and $\sigma_2(x, y)$, satisfy $\Lambda_1\sigma_2 = \Lambda_2\sigma_1$, where $\sigma_1(x, y) := (y^2 - x^2, -2xy)^T$, and $\sigma_2(x, y) := (2xy, y^2 - x^2)^T$, in other words, σ_1 and σ_2 satisfy the commutativity condition, and that $\sum_{r=1}^{2} \Lambda_r\sigma_r = 0$, hence the Itô's system above has the same representation when the stochastic integrals are written in the Stratonovich sense, therefore, even though it is not easy or clear how to get a closed solution of (39), we know, by theorem (2), that the Runge-Kutta scheme of order 4 for deterministic ODEs, provide a scheme of order 2 to approximate numerically the solution of (39).

4 Numerical Examples

In this section we give three different examples where the Milhstein scheme of order 2 and the Runge-Kutta scheme developed in this paper are compared against each other. Again, for all the examples below we have a closed form of the solution, thus, indeed, numerical approximations are not needed. Nevertheless, our intention is to compare the Runge-Kutta scheme with the stochastic Taylor expansion or order 2, and we observed that, in fact, Runge-Kutta seems to perform better. Also, these same examples were taken in [4, 5] and their solutions were approximated using more complicated techniques than the one we present here. The programs were elaborated in Matlab and they are accessible upon request. For each experiment it is necessary to simulate many trajectories of the Wiener process and for simplicity we take always $M = 100$, where M stands for the number of different realizations of the Wiener process. Of course, the programs are made in such a way that users could choose M as they wish. Two different errors are computed at each trial. The first one, named e_1 is just the L_2 error on $[0, T] \times \Omega$ between the exact and the approximate solution, we refer to it as the path-wise error. The second, named e_2, is the mean-square deviation error (see [17]), between the exact solution X_t and the approximate solution $\overline{X}_t$ at $t = T$, where T is the end point of the interval $[0, T]$. We estimate e_2 by taking the average of the errors committed for each path of the Wiener process. In other words,

$e_2 = \left(\frac{1}{MN} \sum_{i=1}^{M} |X_T^i - \overline{X}_T^i|^2 \right)^{1/2}$, where the superscript i indicates the error corresponding to the ith path of the Wiener process. Finally, h stands always for the size-step of the discretization.

Example 1.

The first of the examples we show is the following equation:

$$\begin{cases} dX_t = a^2 X_t (1 + X_t^2) dt + a(1 + X_t^2) dW_t = a(1 + X_t^2) \circ dW_t \\ X(0) = 0 \end{cases}$$

where a is any real number.

The exact solution is $X_t = \tan(aW_t + \arctan X_0)$ and the results are the following:

Schemes	errors	a=.5, T=.2 h=.1	a=.5, T=.2 h=.01	a=.2, T=1 h=.1	a=.2, T=1 h=.01
STE of order 2	e_1	3.4×10^{-9}	$O(10^{-10})$	1.7×10^{-9}	$O(10^{-10})$
	e_2	1.2×10^{-4}	9.8×10^{-6}	6.9×10^{-5}	7×10^{-6}
Runge-Kutta of order 2	e_1	$O(10^{-10})$	$O(10^{-10})$	$O(10^{-10})$	$O(10^{-10})$
	e_2	6.6×10^{-7}	1.2×10^{-8}	5.8×10^{-7}	8.8×10^{-9}

There we observe some small difference in the path-wise error, but, Runge-Kutta shows considerably better results when estimating e_2.

Example 2.

The next example we will consider is the equation:

$$\begin{cases} dX_t = -\frac{1}{2} a^2 X_t dt - a\sqrt{1 - X_t^2} dW_t = -a\sqrt{1 - X_t^2} \circ dW_t \\ X(0) = 0 \end{cases}$$

where again a is any real number.

The exact solution is given by $X_t = \cos(aW_t + \arccos X_0)$ and we obtain:

Schemes	errors	a=.5, T=.2 h=.1	a=.5, T=.2 h=.01	a=.2, T=1 h=.1	a=.2, T=1 h=.01
STE of order2	e_1	$O(10^{-10})$	$O(10^{-10})$	$O(10^{-10})$	$O(10^{-10})$
	e_2	7.2×10^{-7}	1.4×10^{-8}	6.8×10^{-7}	1×10^{-8}
Runge-Kutta of order 2	e_1	$O(10^{-10})$	$O(10^{-10})$	$O(10^{-10})$	$O(10^{-10})$
	e_2	3.7×10^{-7}	6.2×10^{-9}	3.4×10^{-7}	4.9×10^{-9}

In this case, we do not see much difference in the accuracy of both schemes, in fact, they are of the same order for $h = .1$, but Runge-Kutta seems to be of one degree better for e_2 when $h = .01$. However, in order to implement the Milhstein scheme many different derivatives were needed, thus, its implementation is more tedious and longer than using Runge-Kutta.

Example 3.
Finally, our third example is the general linear SDE:

$$\begin{cases} dX_t = aX_t dt + bX_t dW_t = (a - \frac{1}{2}b^2)X_t dt + bX_t \circ dW_t \\ X(0) = 1 \end{cases}$$

where a and b are any real numbers.
The exact solution is given by $X_t = X_0 \exp\left((a - \frac{1}{2}b^2)W_t + bW_t\right)$.
Of course, for most choices of a and b, drift and diffusion terms respectively, this equation will not satisfy the condition $a = \frac{1}{2}\sigma\sigma'$, however, we observe that the Runge-Kutta scheme performs much better than Milhstein's scheme (i.e. STE of order 1).

Schemes	errors	a=.2, b=1 h=.1	a=.2, b=1 h=.01	a=1, b=1 h=.1	a=1, b=1 h=.01
STE of order 1	e_1	3.3×10^{-5}	7.4×10^{-8}	1.2×10^{-3}	1.8×10^{-6}
	e_2	6.7×10^{-3}	4.4×10^{-4}	5.9×10^{-2}	2.1×10^{-3}
Runge-Kutta of order 2	e_1	4.1×10^{-6}	8×10^{-10}	2.1×10^{-5}	3.3×10^{-9}
	e_2	2.2×10^{-3}	4.4×10^{-5}	8.4×10^{-3}	1.2×10^{-4}

In the opinion of the author, this is a very interesting example because even though the equation is not of the type studied in this paper, still we see how Runge-Kutta performs way better that the STE scheme of order 1. In fact, empirically, it seems to behave as a scheme of order 2 even though there is not theory that supports this fact.

References

[1] Allan Bensoussan, R.Glowinski, and A.Rascanu. *Approximation of the Zakai Equation By the Splitting Up Method.* SIAM journal, Control and Optimization, Vol.28, No 6, 1990.

[2] Allan Bensoussan, R.Glowinski, and A.Rascanu.it Approximation Of Some Stochastic Differential Equations by The Splitting Up Method. Appl. Math. Optim, 25, 1992.

[3] Burden, Richard and Faires J.Douglas. *Numerical Analysis.* 5th edition, PWS-Kent Publishing Company, 1993.

[4] Burrage, K. and Burrage, P.M.. *High Strong Order Explicit Runge-Kutta Methods For Stochastic Ordinary Differential Equations.* Appl.Num.Math., 22, 1996, p.81-101.

[5] Burrage, K. and Burrage, P.M.. *General Order Conditions For Stochastic Runge-Kutta Methods For Both Commuting And Non-Commuting Stochastic Ordinary Equations.* Appl.Num.Math., 28, 1998, p.161-177..

[6] Castell, Fabienne and Jessica Gaines. *The Ordinary Differential Equation Approach To Asymptotically Efficient Schemes For Solution Of Stochastic Differential Equations*, Ann. Inst. Henri Poincare, Vol.32, 1996, p.231-250.

[7] J.M.C Clark, R.J. Cameron *The Maximum Rate of Convergence Of Discrete Approximations For Stochastic Differential Equations*, Lectures Notes in Control and Information Sciences, 25, 1978.

[8] N, Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, Second edition, North-Holland Publishing Company, 1989.

[9] Doss, Halim. *Liens entre équations différentielles stochastiques et ordinaires*, Ann. Inst. Henri Poincaré,Vol. XIII, no 2, 1977, p.99-125.

[10] J.G.Gaines. *Numerical Experiments with S(P)DE's*,

[11] Gear, C.William. *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall Inc, NJ, 1971.

[12] Gikhman, I. and Skorokhod, A. *Introduction à la théorie des processus aléatoires*, Ed. Mir. 1980.

[13] Kloeden, P.E. and Platen, E. *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1992.

[14] Kloeden, P.E, Platen, E. and Schurz, H. *Numerical Solution of SDE through Computer Experiments*,1997.

[15] Krylov N.V. *Introduction to the Theory of Diffusion Processes*, American Mathematical Society, 1995.

[16] Francois LeGland. *Splitting-Up Approximation For SPDE's and SDE's with Application To Nonlinear Filtering*.

[17] Milshtein G.N., *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers, 1995.

[18] Milshtein G.N., it A Method Of Second-Order Accuracy Integration Of Stochastic Differential Equations 1976.

[19] Nigel.J.Newton. *Asymptotically Efficient Runge-Kutta Methods For A Class Of Itô And Stratonovich Equations*. SIAM J. Appl. Math, Vol.51, No.2, pp. 542-567.

[20] Nigel.J.Newton. *An Asymptotically Efficient Difference Formula For Solving Stochastic Differential Equations*. Stochastics, Vol.19, pp. 175-206, 1986.

[21] Roman, Luis J. *A Runge-Kutta Type Scheme To Solve* $dX_t = \sigma(X_t) \circ dW_t$, *Under Commutative Noise*. Institute For Mathematics And Its Applications, University of Minnesota, Preprint Series # 1658, 1999.

[22] Roman, Luis J. *On Numerical Solutions Of Stochastic Differential Equations*. PhD Thesis, University of Minnesota, 2000.

[23] Strikwerda, John C. *Finite Difference Schemes and Partial Differential Equations*. Wadsworth and Brooks, 1989.

[24] Talay, Denis. *Efficient Numerical Schemes For The Approximation Of Expectations Of Functionals Of The Solution Of A S.D.E, And Applications*,Lecture Notes in Control and Information Sciences, Filtering and Control of Random Processes, 61, Proceedings of the E.N.S.T.-C.N.E.T. Colloquium Paris, France, February 23-24, 1983.

[25] Talay, Denis. *Discrétisation D'une Équation Différentielle Stochastique Et Calcul Approché D'Epérances De Fonctionnelles De La Slution*, Mathematical Modeling and Numerical Analysis, Vol. 20, n^o 1, 1986, p. 141 à 179.

[26] Talay, Denis. *Simulation of Stochastic Differential Systems*,Probabilistic Methods in Applied Physics, P. Krée, W. Wedig (Eds.), Springer, 1995.

[27] Talay, Denis. *Resolution Trajectorielle et Analyse Numerique des Equations Differentielles Stochastiques*, Stochastics, 1983, Vol.9, pp.275-306.

[28] Talay, Denis and Tubaro, Luciani.*Expansion of the global error for numerical solving Stochastic Differential Equations*, Stochastic Analysis and Applications, 8(4), 483-509, 1990.

[29] Wolfgang Wagner and Eckhard Platen. *Approximation Of Itô Integral Equations*. Als Preprint gedruckt, Berlin, 1978.