Numerical Algorithms

Structure-preserving Runge-Kutta methods for Stochastic Hamiltonian equations with additive noise

--Manuscript Draft--

Manuscript Number:	NUMA-D-13-00147R2
Article Type:	Butcher
Keywords:	stochastic Hamiltonian problems; Runge-Kutta methods; symplecticity
Corresponding Author:	Pamela Burrage Brisbane, AUSTRALIA
First Author:	Pamela Burrage
Order of Authors:	Pamela Burrage
	Kevin Burrage
Abstract:	There has been considerable recent work on the development of energy conserving one-step methods that are not symplectic. Here we extend these ideas to stochastic Hamiltonian problems with additive noise and show that there are classes of Runge-Kutta methods that are very effective in preserving the expectation of the Hamiltonian, but care has to be taken in how the Wiener increments are sampled at each timestep. Some numerical simulations illustrate the performance of these methods.

Structure-preserving Runge-Kutta methods for Stochastic Hamiltonian equations with additive noise

Pamela M. Burrage*, Kevin Burrage*†

Abstract

There has been considerable recent work on the development of energy conserving one-step methods that are not symplectic. Here we extend these ideas to stochastic Hamiltonian problems with additive noise and show that there are classes of Runge-Kutta methods that are very effective in preserving the expectation of the Hamiltonian, but care has to be taken in how the Wiener increments are sampled at each timestep. Some numerical simulations illustrate the performance of these methods.

1 Introduction

One of the most important developments in recent years in the field of numerical solution of ordinary differential equations (ODEs) is that of geometric numerical integration theory. A numerical method is said to be geometric if it preserves one or more geometric properties of the system exactly, and of this new theory perhaps the concept of symplecticity is the most important. In the Introduction we briefly review some of these concepts.

When applied to the nonlinear multidimensional system of ODEs

$$y' = f(y), \quad y(0) = y_0, \quad y \in \mathbb{R}^d \tag{1}$$

an s-stage Runge-Kutta method takes the form

$$\begin{aligned} Y_i &= y_n + h\sum_{j=1}^{s} a_{ij}\, f(Y_j), \quad i = 1, \cdots, s \\ y_{n+1} &= y_n + h\sum_{j=1}^{s} b_j\, f(Y_j). \end{aligned} \tag{2}$$

This class of methods is often characterised by the so-called Butcher tableau

$$\begin{array}{c|c} c & A \\ \hline & b^\top \end{array}$$

*Mathematical Sciences School, Queensland University of Technology, Brisbane 4001, Australia.
†Department of Computer Science, University of Oxford, Oxford, OX13QD, U.K.

where $A = (a_{ij})_{i,j=1,\cdots,s}$, $b^\top = (b_1, \cdots, b_s)$, $c = Ae$, $e = (1, \cdots, 1)^\top$. If the matrix A is strictly lower triangular then the method is said to be explicit and is suitable for nonstiff problems, otherwise the method is implicit and may be suitable for stiff problems if the method has appropriate stability properties.

The system (1) is said to be dissipative if

$$\langle y - z, f(y) - f(z) \rangle \leq 0 \quad \forall y, z$$

and is said to be a Hamiltonian problem if, with $d = 2m$,

$$f(y) = J\nabla H(y), \quad y \in \mathbb{R}^{2m},$$

where

$$J = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix},$$

I_m is the Identity matrix and $H(y)$ is the Hamiltonian associated with the problem.

A dissipative problem has the property

$$\|y(t) - z(t)\| \leq \|y(s) - z(s)\|, \quad t > s$$

and so a Runge-Kutta method is said to be B-stable if two numerical solutions satisfy

$$\|y_{n+1} - z_{n+1}\| \leq \|y_n - z_n\|.$$

Burrage and Butcher [8] and Crouzeix [13] independently showed that an algebraic property, called algebraic stability, guarantees B-stability. Thus a Runge-Kutta method is said to be algebraically stable if $b_i > 0$, $i = 1, \cdots, s$ and if $M = BA + A^\top B - bb^\top$ is non-negative definite, where $B = diag(b_1, \cdots, b_s)$.

In the case that $M = 0$, then for Hamiltonian problems

$$\|y_{n+1} - z_{n+1}\| = \|y_n - z_n\|$$

and the method is said to be symplectic [22]. The concept of symplecticity has opened up a new field of study called structure-preserving methods [16].

The maximum order of an s-stage Runge-Kutta method is $2s$; these methods have stage order s and quadrature order $2s$, so that the weights (b_i) and abscissae (c_i) correspond to those associated with the Legendre polynomials that are orthogonal on $[0, 1]$. These methods can also be constructed through the technique known as collocation. These so-called Gauss methods of order $2s$ are known to be symplectic [6].

We will find it convenient to write a Hamiltonian problem in the form

$$\begin{array}{rcl} q' & = & \nabla_p H(p, q) \\ p' & = & -\nabla_q H(p, q), \end{array} \tag{3}$$

where in many applications $H(y)$, $y = (q, p)^\top$ is often the total energy (kinetic energy plus potential energy) of the system. Since the Hamiltonian is a first integral of the system, it is constant along exact solutions. It is known that symplectic Runge-Kutta methods nearly conserve the Hamiltonian of (3) over exponentially long times [15]. However, Faou et al. [14] noted that certain non-symplectic methods can also be effective in terms of energy conservation. Using B-series and backward error analysis, they gave conditions on the coefficients of a method that guarantee the existence

of a first integral of a modified equation that is close to the Hamiltonian. If these conditions are satisfied up to a certain order, then the Hamiltonian is conserved up to this same order by the numerical method. They also showed that not all symmetric Runge-Kutta methods behave this way even if (3) is reversible under the reflection $p \leftrightarrow -p$. One of the consequences of this analysis is that the three stage collocation Lobatto IIIA method of order 4 has the property that for every Hamiltonian system the dominating error term in the numerical Hamiltonian has no drift. Thus the overall behaviour of the energy error of the Lobatto IIIA method is $O(h^4 + th^6)$, whereas the Lobatto IIIB method behaves as $O(th^4)$.

In a slightly different setting Quispel and McLaren [21] developed a new class of B-series methods that preserves energy for all canonical Hamiltonian vector fields. It is called the averaged vector field method and takes the form

$$y_{n+1} = y_n + h \int_0^1 f((1-s)y_n + sy_{n+1})\, ds. \tag{4}$$

In some sense method (4) could be viewed as a Runge-Kutta method with an infinite number of stages. We can write down finite stage approximations to (4) leading to, for example, the symmetric Kahan method [17], and when restricted to quadratic vector fields, it has the form

$$y_{n+1} = y_n + \frac{h}{2}\left(-f(y_n) + f(\frac{y_n + y_{n+1}}{2}) - f(y_{n+1})\right). \tag{5}$$

Celledoni et al. [11] show that when the vector field is Hamiltonian the map determined by discretisation (5) has both a conserved modified Hamiltonian and an invariant measure. Similarly, the method

$$y_{n+1} = y_n + \frac{h}{6}\left(f(y_n) + 4f(\frac{y_n + y_{n+1}}{2}) + f(y_{n+1})\right) \tag{6}$$

can be shown to preserve a quartic Hamiltonian. We will call this method the MQ method. We note in passing that Chartier et al. [12] have also shown that there are energy preserving B-series methods.

Very recently, Brugnano et al. [1, 2, 3, 4, 5] have developed new classes of Runge-Kutta methods, that they call Hamiltonian BVMs (HBVMs) that are energy preserving for canonical Hamiltonian systems. The idea is based on the discretisation of a local Fourier expansion of the given ODE problem in which different choices of the basis lead to different classes of methods. Thus Brugnano et al. [1, 2, 3, 4] have constructed classes of s-stage Runge-Kutta methods based on generalising the idea of collocating polynomials up to degree r. These are known as HBVM(s,r) and take the form (2) where

$$a_{ij} = b_j \sum_{l=0}^{r-1} P_l(c_j) \int_0^{c_i} P_l(\tau) d\tau, \quad i,j = 1, \cdots, s, \tag{7}$$

where $P_0, \cdots, P_{r-1}$ are the shifted Legendre polynomials orthonormal on $[0,1]$. If the quadrature order of the method is q, then an HBVM(s,r) has order $p = \min(q, 2r)$. Note that if $s = r$, and the nodes are placed at the Gauss points, then $q = 2s$, and these methods reduce to the Gauss methods of order $2s$.

HBVM(s,r) methods with quadrature order $2s$ based on Gaussian quadrature are A-stable with the stability region coinciding with the left half complex plane $\mathbb{C}^-$ and have the remarkable property that they are energy conserving for all polynomial Hamiltonians of degree not larger than $2s/r$. Furthermore, since these methods are of low rank r, the computational efficiency is closer to an r-stage Runge-Kutta method, rather than an s-stage method, apart for the additional $s-r$ function evaluations [1, 2, 3].

The outline of this paper is as follows. In section 2 we introduce our stochastic Hamiltonian problem which takes the form of (3) with additive noise terms introduced in the p equation. In section 3 we discuss various numerical approaches that have been introduced in an attempt to mimic the expectation of the Hamiltonian ($E(H(t))$) over time as accurately as possible. In section 4 we introduce new classes of Runge-Kutta methods that possess the same property. In particular we focus on the Kahan, MQ and three stage Lobatto IIIA methods with a particular implementation of the additive noise terms. Finally, in section 5, we compare the performance of the midpoint rule, MQ and Lobatto IIIA methods, on two test problems, where the Hamiltonian corresponds to the double well problem and the Hénon-Heiles problem. The paper concludes with some comments.

2 Stochastic Hamiltonian problems with additive noise

Very recently, there has been considerable focus on the dynamics of stochastic Hamiltonian problems driven by additive Wiener noise. Mattingly et al. [19] studied the dynamics of gradient systems of the form

$$dy = -\nabla F(y)dt + \Sigma dW,\ y \in \mathbb{R}^d,\ W \in \mathbb{R}^m, \tag{8}$$

where $W(t)$ is a vector of independent Wiener processes and the columns of $\Sigma \in \mathbb{R}^{d\times m}$ are linearly independent. Under appropriate conditions on F including $F \geq 0$, $F(a) \to \infty$ as $|a| \to \infty$, (8) has a unique invariant measure and sharp bounds on the behaviour of $E(F(y(t))^l)$, $l > 0$, are possible. They also showed that the Euler-Maruyama method does not preserve ergodicity but that the implicit Euler method will, under appropriate conditions on F.

In a slightly more general setting, Soize [24] has considered non-linear Hamiltonian dissipative systems excited by white noise, $\xi(t)$, of the form

$$y'' + \epsilon f(H)y' + \nabla V(y) = S\,\xi(t),$$

which can be written as

$$\begin{aligned} dq &= p\,dt \\ dp &= (-\epsilon f(H)\,p - \nabla V(q))\,dt + S\,dW, \end{aligned}$$

where the Hamiltonian is given by

$$H(q,p) = \frac{1}{2}p^2 + V(q).$$

Soize showed there exists a unique solution that tends asymptotically as $t \to \infty$ to a stationary process whose stationary probability density function satisfies

$$p_S = C_0 \, \exp\left(-\frac{2\epsilon}{S^2} \int_0^H f(x) dx \right).$$

Such classes of problems arise in statistical mechanics as a Langevin formulation in which a particle is considered to be moving under a potential V with a frictional term $\epsilon f(H)p$. For higher dimensional systems the theory is more complicated, but results are known for specific cases [24].

Consider now the stochastic Hamiltonian system with additive noise, namely

$$\begin{aligned} dq &= \nabla_p H \, dt \\ dp &= -\nabla_q H \, dt + \Sigma \, dW, \end{aligned} \tag{9}$$

where $q, p \in \mathbb{R}^m$ ($d = 2m$) and Σ is a diagonal matrix given by $\Sigma = diag(\epsilon_1, \cdots, \epsilon_m)$ and $W \in \mathbb{R}^m$ is a vector of independent Wiener processes.

Using Itô's Lemma [18], a Stochastic Differential Equation (SDE) can be written for which the Hamiltonian is a solution. Thus, given the multidimensional Wiener noise Itô SDE,

$$dy = f(t, y) \, dt + g(t, y) \, dW, \quad y(0) = y_0, \quad y \in \mathbb{R}^{2m}$$

the SDE for $U = H(t, y)$ is given by

$$dU = \left(\frac{\partial U}{\partial t} + \sum_{i=1}^{2m} f_i \frac{\partial U}{\partial y_i} + \frac{1}{2} tr(g g^\top \nabla^2 U) \right) dt + \sum_{i=1}^{m} g_i \frac{\partial U}{\partial y_{m+i}} \, dW_i,$$

where the g_i are the columns of g. Therefore, from (9), with $U = H(q, p)$ and

$$\begin{aligned} f &= (\nabla_p H, -\nabla_q H)^\top \\ \nabla U &= (\nabla_q H, \nabla_p H)^\top \\ \nabla^2 U &= \begin{pmatrix} \nabla_{qq} H & \nabla_{qp} H \\ \nabla_{pq} H & \nabla_{pp} H \end{pmatrix}, \end{aligned}$$

and $g_i = \epsilon_i (0, e_i)^\top$, where 0 is the zero vector of length m and e_i is the ith basis vector of $\mathbb{R}^m$, we find

$$dH = (\frac{1}{2} \sum_{i=1}^{m} \epsilon_i^2 \nabla_{pp}^{ii} H) \, dt + \sum_{i=1}^{m} \epsilon_i \nabla_p^i H \, dW_i,$$

where $\nabla_p^i H$ is the ith component of $\nabla_p H$ and $\nabla_{pp}^{ii} H$ is the ith diagonal element of $\nabla_{pp} H$.

Integrating gives

$$H(t) = H(t_0) + \frac{1}{2} \sum_{i=1}^{m} \epsilon_i^2 \int_{t_0}^{t} \nabla_{pp}^{ii} H \, ds + \sum_{i=1}^{m} \epsilon_i \int_{t_0}^{t} \nabla_p^i H \, dW_i(s)$$

and using the expectation property of Itô integrals this gives

$$E(H(t)) = E(H(t_0)) + \frac{1}{2} \sum_{i=1}^{m} \epsilon_i^2 \int_{t_0}^{t} E(\nabla_{pp}^{ii} H) \, ds. \tag{10}$$

Now if the Hamiltonian is separable of the form

$$H(q,p) = \frac{1}{2}\sum_{j=1}^{m} p_j^2 + V(q)$$

then (10) becomes

$$E(H(t)) = E(H(t_0)) + \frac{1}{2}\sum_{j=1}^{m} \epsilon_j^2\,(t - t_0) \tag{11}$$

and the expected value of the Hamiltonian grows linearly with time.

3 Numerical methods for additive noise problems

In the case of additive noise problems, we could ask a numerical method to perform well in terms of the stationary distribution function and the evolution of the mean of the underlying Hamiltonian. Schurz [23], for example, showed that the implicit midpoint rule has the same stationary distribution as a multidimensional Ornstein-Uhlenbeck process. Burrage et al. [9] extended this analysis and showed that for linear second order equations with additive white noise and damping, only the implicit midpoint rule is measure-exact, that is, preserves the position, velocity and absence of correlation at equilibrium. Not only that, the implicit midpoint rule seems to be very effective on nonlinear second order equations with damping, although the leapfrog (Verlet) method can also be effective and has the advantage of being explicit. Burrage and Lythe [10] extended these ideas by constructing additional classes of s-stage Runge-Kutta methods that can preserve the correlation matrix to $O(h^s)$.

In terms of considering the evolution of the mean of the Hamiltonian, Melbo and Higham [20] showed that partitioned approaches can be effective. Burrage and Burrage [7] explored under what conditions a Runge-Kutta method can preserve property (11) for a separable Hamiltonian problem driven by additive noise.

The analysis was given for the linear, additive noise problem

$$dy = Qy\,dt + \epsilon\, r\, dW, \tag{12}$$

where

$$Q = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad r = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

corresponding to the Hamiltonian $H = \frac{1}{2}(p^2 + q^2)$.

To illustrate these ideas, first consider the general additive noise problem

$$dY = f(Y)\,dt + \epsilon r\, dW, \tag{13}$$

where Y and r (constant vector) are $2m \times 1$ column vectors, ϵ is a scalar value and $W(t)$ is a scalar Wiener process. If only one Wiener increment $W(t_n + h) - W(t_n) = \Delta W_n \sim N(0,h)$ is sampled per step then an s-stage Runge-Kutta method is given

by

$$\begin{aligned} Y_i &= y_n + h\sum_{j=1}^{s} a_{ij} f(Y_j) + \epsilon\, c_i\, r\, \Delta W_n \\ y_{n+1} &= y_n + h\sum_{j=1}^{s} b_j f(Y_j) + \epsilon\, r\, \Delta W_n. \end{aligned} \tag{14}$$

Applying Runge-Kutta method (14) to (12) gives

$$y_{n+1} = R(hQ)\, y_n + \epsilon\, \Delta W_n\, S(hQ)\, r,$$

where for a scalar z

$$R(z) = 1 + z\, b^\top (I_s - Az)^{-1}\, e, \quad S(z) = 1 + z b^\top (I_s - Az)^{-1}\, c.$$

It is trivial to show that

$$S(z) = \frac{1}{z}\left(R(z) - 1\right).$$

This can be easily seen by noting that with $c = Ae$ and $b^\top e$=1 then

$$\begin{aligned} S(z) &= 1 - b^\top (I_s - Az)^{-1}(-zAe) \\ &= 1 - b^\top (I_s - Az)^{-1}(I_s - Az - I_s)e \\ &= 1 - b^\top e + (zb^\top (I_s - Az)^{-1}e)/z \\ &= (R(z) - 1)/z. \end{aligned}$$

For a system

$$R(hQ) = I_2 + h(b^\top \otimes Q)\left(I_s \otimes I_2 - h(A \otimes Q)\right)^{-1}(e \otimes I_2)$$

and

$$S(hQ) = (hQ)^{-1}\left(R(hQ) - I_2\right),$$

where $e = (1, \cdots, 1)^\top \in \mathbb{R}^s$.

Thus

$$E[y_{n+1}^\top\, y_{n+1}] = E[y_n^\top\, R^\top(hQ)\, R(hQ)\, y_n] + \epsilon^2 h r^\top S^\top(hQ) S(hQ)\, r.$$

Since $Q^\top = -Q$, and letting

$$P(hQ) = R(-hQ)R(hQ),$$

then

$$E[y_{n+1}^\top\, y_{n+1}] = E[y_n^\top P(hQ) y_n] + \epsilon^2 h r^\top (Qh)^{-2}\left(P(hQ) - R(-hQ) - R(hQ) + I_2\right) r.$$

In the case of a symmetric Runge-Kutta method of order two or more, $P = I_2$ and as $E[H_n] = \frac{1}{2}E[y_n^\top y_n]$ then

$$E[H_{n+1}] = E[H_n] + \frac{\epsilon^2}{2}\, h r^\top (hQ)^{-2}\left(2I_2 - (R(-hQ) + R(hQ))\right) r.$$

Thus Burrage and Burrage [7] concluded it is not possible for any Runge-Kutta method, even one that is symmetric, to preserve property (11) given only one Wiener increment per step.

Burrage and Burrage [7] then propose a new formulation of a Runge-Kutta method that involves using $s+1$ samples of a Wiener process per step. This idea is in line with the ideas of Brugnano et al. [1, 2, 3].

Thus let

$$z_i = W(t_n + c_i h) - W(t_n + c_{i-1} h) = \sqrt{c_i - c_{i-1}}\, \Delta W_i, \quad i = 1, \cdots, s+1,$$

where $c_0 = 0$, $c_{s+1} = 1$ and ΔW_i are $s+1$ independent samples of $N(0, h)$. Thus the Runge-Kutta formulation for (13) is

$$\begin{aligned} Y_i &= y_n + h \sum_{j=1}^{s} a_{ij} f(Y_j) + \epsilon \sum_{j=1}^{i} z_j\, r \\ y_{n+1} &= y_n + h \sum_{j=1}^{s} b_j f(Y_j) + \epsilon \sum_{j=1}^{s+1} z_j\, r. \end{aligned} \tag{15}$$

Let

$$Z = (z_1, \cdots, z_s)^\top$$

and note that

$$\left(z_1, z_1 + z_2, \cdots, \sum_{j=1}^{s} z_j \right)^\top = VZ,$$

where V is the $s \times s$ matrix whose lower triangular component has ones, and the strictly upper triangular component has zeros. Applying this method to the linear test equation gives

$$y_{n+1} = R(hQ)\, y_n + \epsilon\, S(hQ)\, r$$

where

$$S(hQ) = z_{s+1} I_2 + e^\top Z I_2 + h(b^\top \otimes Q)\, (I_s \otimes I_2 - h(A \otimes Q))^{-1}\, (VZ \otimes I_2).$$

For example, in formulation (15), the implicit midpoint rule is

$$\begin{aligned} Y &= y_n + \frac{h}{2} f(Y) + \frac{1}{\sqrt{2}} \Delta W_1\, \epsilon\, r \\ y_{n+1} &= y_n + h\, f(Y) + \frac{1}{\sqrt{2}} (\Delta W_1 + \Delta W_2)\, \epsilon\, r \end{aligned} \tag{16}$$

and so

$$\begin{aligned} R(hQ) &= (I_2 - \frac{h}{2} Q)^{-1}\, (I_2 + \frac{h}{2} Q) \\ S(hQ) &= z_2 I_2 + R(hQ)\, z_1. \end{aligned} \tag{17}$$

Thus

$$E(y_{n+1}^\top\, y_{n+1}) = E(y_n^\top R^\top(hQ)\, R(hQ)\, y_n) + \epsilon^2\, r^\top\, E(S^\top(hQ)\, S(hQ))\, r. \tag{18}$$

With $Q^\top = -Q$ then the midpoint rule has the property

$$R^\top(hQ)\, R(hQ) = I_2$$

$$E(S^\top(hQ)\, S(hQ)) = E(z_1^2 + z_2^2)\, I_2 = h\, I_2,$$

and so

$$E(H_{n+1}) = E(H_n) + \frac{\epsilon^2}{2}\, h. \tag{19}$$

Thus the implicit midpoint rule (16) with two Wiener samples per step preserves (11) when the Hamiltonian is given by $\frac{1}{2}(p^2 + q^2)$.

In the next section we address the question as to whether there are other classes of Runge-Kutta methods that possess this property of the implicit midpoint method.

4 Geometric properties of additive stochastic Runge-Kutta methods with multiple Wiener processes

In section 1 we introduced three new methods that are not symplectic but perform well on deterministic separable Hamiltonian problems. Kahan's method (5) can be written in Runge-Kutta tableau form as

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{4} & 1 & -\frac{1}{4} \\ 1 & -\frac{1}{2} & 2 & -\frac{1}{2} \\ \hline & -\frac{1}{2} & 2 & -\frac{1}{2} \end{array} \tag{20}$$

The MQ method (6) can be written as

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{12} & \frac{1}{3} & \frac{1}{12} \\ 1 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array} \tag{21}$$

and the Lobatto IIIA method as

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{5}{24} & \frac{1}{3} & -\frac{1}{24} \\ 1 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array} \tag{22}$$

Remarks

1. The stability function of the MQ method (21) is

$$R(hQ) = (I_2 - \frac{h}{2}Q)^{-1}\, (I_2 + \frac{h}{2}Q), \tag{23}$$

which is the same as that of the implicit midpoint method, while for the Lobatto IIIA method (22) it is

$$R(hQ) = (I_2 - \frac{h}{2}Q + \frac{h^2}{12}Q^2)^{-1}\, (I_2 + \frac{h}{2}Q + \frac{h^2}{12}Q^2), \tag{24}$$

since the Lobatto IIIA method is order 4.

2. Both (21) and (22) can be written in a concise manner as

$$y_{n+1} = y_n + \frac{h}{6}\left(f(y_n) + 4f\left(\frac{y_n + y_{n+1}}{2} + \frac{\theta h}{8}(f(y_n) - f(y_{n+1}))\right) + f(y_{n+1})\right), \tag{25}$$

where $\theta = 0$ in the case of (21) and $\theta = 1$ in the case of (22).

3. (20) has quadrature order 2 but (21) and (22) have quadrature order 4.

Because of Remark 3, we will focus on methods (21) and (22) and we need to see how we can add the Wiener processes into these formulations in order to preserve the mean of the Hamiltonian property (11).

Before doing so, let us revisit the two Wiener process implementation of the Implicit Midpoint rule, namely

$$\begin{aligned} Y &= y_n + \frac{h}{2}f(Y) + \frac{1}{\sqrt{2}}z_1\epsilon r \\ y_{n+1} &= y_n + hf(Y) + \frac{1}{\sqrt{2}}(z_1 + z_2)\epsilon r. \end{aligned} \tag{26}$$

Now let

$$\begin{aligned} \bar{y}_n &= y_n + \frac{1}{\sqrt{2}}z_1\epsilon r \\ y_{n+1} &= \bar{y}_{n+1} + \frac{1}{\sqrt{2}}z_2\epsilon r \end{aligned} \tag{27}$$

and

$$\bar{y}_{n+1} = \bar{y}_n + h\, f(\frac{\bar{y}_n + \bar{y}_{n+1}}{2}), \tag{28}$$

then it is easily seen that the method given by (27) and (28) is equivalent to (26).

Now all three methods (20), (21) and (22) have the same structure: they have the first row in the Runge-Kutta tableau set to zero, they are stiffly accurate (the last row of the RK tableau is equal to the update row), they have symmetric weights, and they have an intermediate approximation at the midpoint of the interval. Thus they only need two Wiener increment evaluations per step just as in the case of the Implicit Midpoint rule. Consequently, formulations (27) and (28) give us the key to introducing the Wiener increments into these formulations.

Focusing on (21) and (22), and their characterisation via (25), leads to the formulation

$$\begin{aligned} \bar{y}_n &= y_n + \frac{1}{\sqrt{2}}z_1\epsilon r \\ \bar{y}_{n+1} &= \bar{y}_n + \frac{h}{6}\left(f(\bar{y}_n) + 4f\left(\frac{\bar{y}_n + \bar{y}_{n+1}}{2} + \frac{\theta h}{8}(f(\bar{y}_n) - f(\bar{y}_{n+1}))\right) + f(\bar{y}_{n+1})\right) \\ y_{n+1} &= \bar{y}_{n+1} + \frac{1}{\sqrt{2}}z_2\epsilon r. \end{aligned} \tag{29}$$

Applying this formulation to the linear test problem (12) we find

$$\bar{y}_{n+1} = R(hQ)\bar{y}_n$$

where $R(hQ)$ is either the order 2 (23) or order 4 (24) symmetric Padé approximation. Using the independence of z_1 and z_2 and the fact that $R(-hQ)R(hQ) = I_2$ in both cases then it is easily seen that

$$E(y_{n+1}^\top y_{n+1}) = E(y_n^\top y_n) + \frac{1}{2}\epsilon^2 r^\top r(E(z_1^2) + E(z_2^2))$$

and hence

$$E(H_{n+1}) = E(H_n) + \frac{\epsilon^2}{2}h,$$

as required.

In the next section we will compare the performance of the Implicit Midpoint rule, the MQ method and the Lobatto IIIA method on the double well and Hénon-Heiles additive noise problems.

5 Simulations and Conclusions

In both our test problems, we perform 50000 simulations for varying values of the stepsize and the additive noise on the interval [0,40], and present the numerical error in the mean of the Hamiltonian at $T = 40$.

Example 1.

Our first example is the double well potential with Hamiltonian

$$H = \frac{1}{2}p^2 + \frac{1}{4}q^4 - \frac{1}{2}q^2.$$

We take $(q_0, p_0) = (\sqrt{2}, \sqrt{2})^\top$, and take $\epsilon = 0$, 0.01, 0.1, 0.5 and compute $E(H(t))$ on the interval $[0, 40]$. In this case (11) gives

$$E(H(40)) = 1 + 20\epsilon^2 = 1,\ 1.002,\ 1.2,\ 6.$$

Numerical results are given in Table 1.

ϵ	h	MQ	L	IM
0	0.05	9.5(-15)	7.9(-7)	4.3(-4)
0	0.2	5.6(-14)	2.1(-4)	6.1(-3)
0.001	0.1	8.4(-6)	2.1(-5)	1.6(-3)
0.01	0.1	1.5(-5)	4.7(-6)	1.2(-4)
0.01	0.2	4.3(-5)	1.5(-4)	6.0(-3)
0.5	0.05	1.5(-2)	8.0(-3)	1.2(-2)
0.5	0.2	1.6(-2)	6.5(-3)	6.9(-2)

Table 1: Errors in $E(H(40))$ for the three methods

Example 2.

Our second problem is the Hénon-Heiles problem with Hamiltonian

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(q_1^2 + q_2^2) + \alpha(q_1 q_2^2 - \frac{1}{3}q_1^3).$$

We take $(q_0, p_0) = (\sqrt{3}, 1, 1, 1)$ and $\alpha = \frac{1}{16}$. In this case (with noise $\epsilon_1 = \epsilon_2 = \epsilon$), (11) gives

$$E(H(T)) = 3 + \epsilon^2 T.$$

Numerical results are given in Table 2.

ϵ	h	MQ	L	IM
0	0.1	3.1(-15)	3.4(-8)	6.9(-5)
0.001	0.1	4.1(-5)	4.0(-5)	9.8(-5)
0.01	0.1	4.0(-4)	3.9(-4)	2.1(-3)
0.01	0.2	1.9(-4)	2.2(-4)	1.9(-3)
0.2	0.1	2.6(-3)	2.4(-3)	7.9(-1)
0.2	0.2	1.3(-2)	1.3(-2)	7.6(-1)

Table 2: Errors in $E(H(40))$ for the three methods

Discussion

In the case of the deterministic formulations of problems 1 and 2, the MQ method preserves the Hamiltonian almost exactly, while the Lobatto method and the Implicit Midpoint rule show an $O(h^4)$ and $O(h^2)$ behaviour, respectively. In the case of very small additive noise this property of the MQ method is lost, but it still performs better than the other two methods and substantially better than the Implicit Midpoint rule. As the additive noise increases further then the MQ and Lobatto methods seem to perform about the same, and both perform substantially better than the Implicit Midpoint rule. Trying to elucidate a dependence on the stepsize is a little harder and that is due to the presence of the Monte Carlo error and the fact that as the additive noise component increases, the error due to this drowns the stepsize dependence.

Implementation

The formulation described in (29) for both the MQ and Lobatto IIIA methods is elegant since it merely needs the same solution technique for the nonlinear equations to be solved at each time step as the deterministic case. This could be either via a fixed point iteration or Newton's method. We prefer the latter and in this case we need to solve a set of linear systems per step of the form

$$Ay = b$$

where

$$A = I - \frac{h}{2} J_n + \theta \frac{h^2}{12} J_n^2$$

where J_n is the Jacobian of f evaluated at y_n, and $\theta = 0$ or 1, respectively.

Conclusions

We have shown that we can take certain classes of non-symplectic Runge-Kutta methods that are very effective in terms of energy preservation over time and by introducing additive noise terms in an appropriate manner, demonstrate that these

methods will preserve very accurately the expectation of the Hamiltonian for additive noise problems. This work extends the analysis given in Burrage and Burrage [7] for the implicit midpoint rule. We only focussed on three-stage methods, as in these cases only two Wiener increments per stage were needed. However, we expect the ideas introduced here can be extended to higher stage, and hence higher order, methods, but at the cost of simulating more Wiener increments per step. Our theoretical numerical analysis of these methods was based on the stochastic linear Hamiltonian problem, corresponding to the Hamiltonian $\frac{1}{2}(q^2+p^2)$, but we believe the deterministic analysis based on B-series and backward error analysis could be extended to additive noise nonlinear Hamiltonian problems. All of these are considerations for future work.

Acknowledgements

The authors would like to thank Luigi Brugnano and Grant Lythe for various discussions on issues related to this submission.

References

[1] Brugnano, L., Iavernaro, F. and Trigiante, D. (2009): Analysis of Hamiltonian Boundary Value Methods (HBVMs): a class of energy-preserving Runge-Kutta methods for the numerical solution of polynomial Hamiltonian dynamical systems, submitted (arXiv:0909.5659).

[2] Brugnano, L., Iavernaro, F. and Trigiante, D. (2010): Hamiltonian boundary value methods (energy preserving discrete line integral methods), *J. of Numer. Anal. Industr. and Appl. Math.*, **5**, 1-2, 17–37 (arXiv:0910.3621).

[3] Brugnano, L., Iavernaro, F. and Trigiante, D. (2012): The lack of continuity and the role of infinite and infinitesimal in numerical methods for ODEs: the Case of Symplecticity, *Applied Mathematics and Computation*, **218**, 8053–8063.

[4] Brugnano, L., Iavernaro, F. and Trigiante, D. (2012): A simple framework for the derivation and analysis of effective classes of one-step methods for ODEs, *Applied Mathematics and Computation*, **218**, 8475–8485.

[5] Brugnano, L., Iavernaro, F. and Trigiante, D. (2011): A note on the efficient implementation of Hamiltonian BVMs, *Journal of Computational and Applied Mathematics*, **236**, 375-383.

[6] Burrage, K. (1995): *Parallel and Sequential Methods for Ordinary Differential Equations*, Oxford University Press.

[7] Burrage, K. and Burrage, P.M. (2012): Low rank Runge-Kutta methods, symplecticity and stochastic Hamiltonian problems with additive noise, *Journal of Computational and Applied Mathematics* **236**, 3920–3930.

[8] Burrage, K. and Butcher, J.C. (1979): Stability criteria for implicit Runge-Kutta methods. *SIAM Journal on Numerical Analysis*, **16**, 46–57.

[9] Burrage, K., Lenane, I. and Lythe, G. (2007): Numerical methods for second-order stochastic differential equations, *SIAM Journal on Scientific Computing*, **29**(1), 245–264.

[10] Burrage, K. and Lythe, G. (2009): Accurate stationary densities with partitioned numerical methods for stochastic differential equations, *SIAM J. on Numerical Analysis* **47**, 1601–1618.

[11] Celledoni, E., McLachlan,R., Owren, B. and Quispel, G.R.W. (2013): Geometric properties of Kahan's method, *J. Phys. A: Math. Theor.* **46**, 025201.

[12] Chartier, P., Faou, E; and Murua, A. (2006): An algebraic approach to invariant preserving integrators: the case of quadratic and Hamiltonian invariants, *Numer. Math.* **103**, 575–590.

[13] Crouzeix, M. (1979): Sur la B-stabilité des méthodes de Runge-Kutta, *Numer. Math.* **32**, 75–82.

[14] Faou, E., Hairer, E. and Pham, T-L. (2004): Energy conservation with non-symplectic methods: examples and counter-examples *BIT Numer. Math.* **44**, 699–709.

[15] Hairer, E. and Lubich, Ch. (1997): The life span of backward error analysis for numerical integrators, *Numer. Math.* **76**, 441–462.

[16] Hairer, E., Lubich, Ch. and Wanner, G. (2006): *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer, Berlin.

[17] Kahan, W. and Li, R.C. (1997): Unconventional schemes for a class of ordinary differential equations - with applications to the Korteweg-de Vries equation, *J. Comput. Phys.* **134**, 316–331.

[18] Kloeden, P.E. and Platen, E. (1999): *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.

[19] Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002): Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, *Stoch. Proc. and their Appl.* **101** (2), 185–232.

[20] Melbo, A.H.S. and Higham, D.J. (2004): Numerical simulation of a linear oscillator with additive noise. *Applied Numerical Mathematics* **51**, 89.

[21] Quispel, G.R.W. and McLaren, D.I. (2008): A new class of energy-preserving numerical integration methods, *J. Phys. A: Math. Theor.* **41**, 045206.

[22] Sanz-Serna, J.M. and Calvo, M.P. (1994): *Numerical Hamiltonian Problems*, Chapman & Hall, London.

[23] Schurz H. (1999): The invariance of asymptotic laws of linear stochastic systems under discretization. *Z. Angew Math. Mech.* **6**, 375–382.

[24] Soize, C. (1995): Exact steady-state solution of FKP equation in higher dimension for a class of nonlinear Hamiltonian dissipative dynamical systems excited by Gaussian white noise, *Probabilistic Methods in Applied Physics*, eds. P. Krée and W. Wedig, Lecture Notes in Physics **451**, 284–309.