Low rank Runge-Kutta methods, symplecticity and stochastic Hamiltonian problems with additive noise

Kevin Burrage[*], Pamela M. Burrage[†]

Abstract

In this paper we extend the ideas of Brugnano, Iavernaro and Trigiante in their development of HBVM(s,r) methods to construct symplectic Runge-Kutta methods for all values of s and r with $s \geq r$. However, these methods do not see the dramatic performance improvement that HBVMs can attain. Nevertheless, in the case of additive stochastic Hamiltonian problems an extension of these ideas, which requires the simulation of an independent Wiener process at each stage of a Runge-Kutta method, leads to methods that have very favourable properties. These ideas are illustrated by some simple numerical tests for the modified midpoint rule.

1 Introduction

Runge-Kutta methods have been effective solvers of initial value ordinary differential equation systems for well over a hundred years. When applied to the problem

$$y' = f(y), \quad y(0) = y_0, \quad y \in \Re^N \tag{1}$$

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

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

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

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

[†]Electrical Engineering and Computer Science School, Queensland University of Technology, Brisbane 4001, Australia.

for nonstiff problems, otherwise the method is implicit and may be suitable for stiff problems if the method has appropriate stability properties.

Perhaps one of the most important developments in the last thirty or so years is the recognition that certain classes of implicit Runge-Kutta methods can satisfy algebraic properties that guarantee stable or structure-preserving properties for general classes of nonlinear problems. A system 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

$$f(y) = J \nabla H(y), \quad y \in \Re^N$$

where

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

where I_N 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 [10] 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 [15]. The concept of symplecticity has opened up a new field of study called structure-preserving methods [12].

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 [7].

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.

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

$$a_{ij} = b_j \sum_{l=0}^{r-1} P_l(c_j) \int_0^{c_i} P_l(\tau) d\tau, \quad i, j = 1, \cdots, s, \tag{3}$$

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 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 plane 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].

We will find it useful to write down the general classes of HBVM(s,2) and HBVM(s,3) methods, respectively. We note that

$$P_1(x) = 2\sqrt{3}(x - \frac{1}{2}), \quad P_2(x) = 6\sqrt{5}(x^2 - x + \frac{1}{6}).$$

Example 1
i) HBVM(s,2):

$$A = cb^T + 6(c^2 - c)(b^T C - \frac{1}{2}b^T), \tag{4}$$

ii) HBVM(s,3):

$$A = cb^T + 6(c^2 - c)(b^T C - \frac{1}{2}b^T) + (10c^3 - 15c^2 + 5c)(6b^T C^2 - 6b^T C + b^T), \tag{5}$$

where vector operations are component-wise and

$$C = diag(c_1, \cdots, c_s).$$

2 Low rank symplectic Runge-Kutta methods

The class of methods given in (4) and (5) have at most orders 4 and 6, respectively. However, they are only symplectic in the case of HBVM(s,2) if $s = 2$ and $c^2 - c + \frac{1}{6}e = 0$ and in the case of HBVM(s,3) if $s = 3$ and $c^3 - \frac{3}{2}c^2 + \frac{3}{5}c - \frac{1}{20}e = 0$. In these two cases the methods reduce to the two and three stage Gauss methods of orders four and six, respectively. In this section we will take a different view to that of Brugnano, Iavernaro and Trigiante [1, 2, 3] and view the HBVM methods as a particular class of low rank Runge-Kutta methods. But as a particular feature of these classes of s-stage, r-rank methods we will require that they be symplectic for all values of s and r, with $s \geq r$. This will constrain them to have a similar but different form to the HBVM(s,r) methods. We will denote this class of symplectic low rank, r, s-stage Runge-Kutta methods as SLIRK(s,r) methods.

In what follows we will often write the Runge-Kutta matrix as

$$A = \bar{A}B, \quad B = diag(b_1, \cdots, b_s).$$

Clearly the algebraic stability matrix can be written as

$$M = B(\bar{A} + \bar{A}^T - ee^T)B$$

and we will let

$$\bar{M} = \bar{A} + \bar{A}^T - ee^T. \tag{6}$$

We will construct low rank methods with

$$\bar{A} = \sum_{j=0}^{r-1} u_j w_j^T \tag{7}$$

such that $M = 0$, so that if $B > 0$ then these methods will be symplectic.

Let $\bar{P}_i(c)$ be the Gauss-Legendre polynomials of degree i, with

$$\bar{P}_1(c) = c - \tfrac{1}{2}, \quad \bar{P}_2(c) = 6c^2 - 6c + 1, \quad \bar{P}_3(c) = 10c^3 - 15c^2 + 6c - \tfrac{1}{2}.$$

Then it is easily shown that if

$$\begin{aligned} u_0 &= c, \quad u_i = -\bar{P}_{i-1}(c) + \bar{P}_{i+1}(c),\ i = 1, \cdots, r-2, \quad u_{r-1} = -\bar{P}_{r-2}(c) \\ w_i &= \bar{P}_i(c),\ i = 0, \cdots, r-1 \end{aligned} \tag{8}$$

then $\bar{M} = 0$. A SLIRK(s, s) with $B(s)$ and $\bar{P}_s(c) = 0$, reduces to HBVM(s, s), namely the symplectic s-stage Gauss methods of order $2s$. It is easy to show

Theorem 1. For $t = 1, \cdots, r-1$ quadrature order $r+t-1$ implies a SLIRK$(r+t-1, r)$ method is order $\min(r+t-1, 2t+1)$ with stage order t. In addition, quadrature order $2r$ implies order $2r$.

SLIRK(s, r) methods can be implemented in an efficient manner, as is the case of HBVMs [5]. The case $r = 2$ is easily generalised to $r > 2$. In the case $r = 2$

$$A = cb^T - e(b^T C - \frac{1}{2} b^T)$$

and so

$$a_{ij} = b_j(c_i - c_j + \frac{1}{2}), \quad \forall i, j = 1, \cdots, s.$$

Hence

$$\begin{aligned} Y_1 - y_n &= h \sum_{j=1}^{s} b_j (c_1 - c_j + \tfrac{1}{2}) f(Y_j) \\ Y_k - y_n &= h \sum_{j=1}^{s} b_j (c_k - c_j + \tfrac{1}{2}) f(Y_j),\ k = 2, \cdots, s \end{aligned}$$

and subtracting gives

$$\begin{aligned} Y_k - Y_1 &= (c_k - c_1)\, h \sum_{j=1}^{s} b_j f(Y_j) \\ &= (c_k - c_1)(y_{n+1} - y_n). \end{aligned}$$

Let $z_{n+1} = y_{n+1} - y_n$, then

$$Y_k = Y_1 + (c_k - c_1) z_{n+1}, \quad k = 2, \cdots, s$$

and so the method reduces to

$$\begin{aligned} Y_1 &= y_n + h\sum_{j=1}^{s} b_j(c_1 - c_j + \frac{1}{2})f(Y_1 + (c_j - c_1)z_{n+1}) \\ z_{n+1} &= h\sum_{j=1}^{s} b_j f(Y_1 + (c_j - c_1)z_{n+1}) \\ y_{n+1} &= z_{n+1} + y_n. \end{aligned}$$

It is easily seen that, apart from additional function evaluations, as long as the quadrature order three condition $B(3)$ holds then the effective implementation is that of a two stage method with structure

$$\begin{pmatrix} c_1 & -c_1^2 + \frac{1}{2}c_1 - \frac{1}{12} \\ 1 & \frac{1}{2} - c_1 \end{pmatrix}$$

which has the same characteristic polynomial, $\lambda^2 - \frac{1}{2}\lambda + \frac{1}{12}$, as the two stage Gauss method. This is easily seen by computing the Jacobian of the nonlinear system involving Y_1 and z_{n+1}.

The basis of this section has been to show that there are other classes of low rank methods and we have imposed symplecticity for all r and s. However, these methods do not appear to have the same effect as the class of HBVMs, as for given r they improve little on the accuracy of the Hamiltonian as s increases. However because they are symplectic, they do preserve exactly the angular momentum for all values of s and r. We now turn to the case of efficient methods for stochastic Hamiltonian problems with additive noise.

3 Stochastic Hamiltonian problems with additive noise

Newton's Second Law of Motion relates force to acceleration and so second order differential equations arise frequently in scientific applications. These differential equations often contain a parameter known as damping. The stationary density is independent of damping, but dynamical quantities, and the usefulness, or otherwise, of numerical algorithms, are strongly dependent on it. In the infinite-damping limit, the system becomes first order. The limit of zero damping, on the other hand, corresponds to Hamiltonian systems where symplectic methods can be applied, which are capable of preserving geometric or energy-like properties over long time periods. Thus it is important to derive methods capable of accurately reproducing the stationary density *for all positive values of damping.*

We shall consider equations of the following form:

$$x'' = f(x) - \eta s^2(x)x' + \epsilon s(x)\xi(t), \tag{9}$$

where $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$ and the damping parameter is denoted η. Angled brackets denote mean over realizations. The second-order stochastic differential equation (SDE) (9) describes the position of a particle subject to deterministic forcing $f(x)$ and random forcing $\xi(t)$. The deterministic forcing is related to the potential function $E(x)$ via $f(x) = -E'(x)$. Thus the particle is said to be moving

noisily in a potential well. The amplitude of the random forcing, ϵ, is related to the temperature T and damping coefficient η by the fluctuation-dissipation relation [11] $\epsilon^2 = 2\eta KT$. This is the so-called Einstein relationship.

We can write (9) as a pair of first-order equations for X and V, the position and velocity variables:

$$\begin{aligned} dX &= V\,dt \\ dV &= -\eta s^2(X)V dt + f(X)dt + \epsilon s(X)dW(t), \end{aligned} \tag{10}$$

where $W(t)$ is a Wiener process satisfying $\langle W(t)W(s)\rangle = \min(t,s)$. If $s(x)$ is not a constant, the noise amplitude is a function of position and the equation is commonly said to have "multiplicative noise". However, because the coefficient of dW in the SDE for V is a function of X only, there is no difference between the Itô and Stratonovich forms [11] of (10). The probability density at time t is $P(x,v;t)$, where

$$P(x,v;t) = \frac{d}{dx}\frac{d}{dv}\mathrm{Prob}\left(X(t) < x, V(t) < v\right). \tag{11}$$

The stationary density, $P_\infty(x,v)$, defined as

$$P_\infty(x,v) = \lim_{t\to\infty} P(x,v;t), \tag{12}$$

has the following analytical form, independent of $\eta > 0$ and $s(x)$ [11]:

$$P_\infty(x,v) = N\exp\left(-v^2/2KT - E(x)/KT\right). \tag{13}$$

Thus the late-time statistics of the velocity are Gaussian and uncorrelated with the position. It is notable that the stationary density has a closed tractable form for many nonlinear functions $f(x)$ when analytical study of the full evolution is not possible.

In a slightly more general setting, Soize [17] has shown that for non-linear Hamiltonian dissipative systems excited by white noise of the form

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

which can be written as

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

with

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

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

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

If $f(H) = 1$ then

$$p_S = C_0\,\exp\left(-\frac{2\epsilon}{S^2}H(q,p)\right),$$

while for the non-linear problem

$$y'' + \epsilon\left(1 - (y')^2 - y^2\right) y' + T_0 y = S\xi(t),$$

$$p_S = C_0 \exp\left(-\frac{2\epsilon}{S^2} H(1-H)\right), \quad H = \frac{1}{2}p^2 + \frac{1}{2}T_0 q^2.$$

These classes of problems arise in statistical mechanics as a Langevin formulation in which a particle is 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 [17].

Consider now the separable Hamiltonian problem with additive noise, namely

$$\begin{aligned} dq &= H_p\, dt \\ dp &= -H_q\, dt + \sigma\, dW \\ H(q,p) &= \frac{1}{2}p^2 + V(q). \end{aligned} \tag{14}$$

Using Itô's Lemma [13], an SDE can be written for which the Hamiltonian is a solution.

Thus, given the multidimensional Wiener noise Itô SDE,

$$dy = f(t,y)\, dt + g(t,y)\, dW, \quad y(0) = y_0$$

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

$$dU = \left(\frac{\partial U}{\partial t} + \sum f_i \frac{\partial U}{\partial y_i} + \frac{1}{2} tr(gg^T \nabla^2 U)\right) dt + \sum g_i \frac{\partial U}{\partial y_i}\, dW_i.$$

Therefore, from (14), with $U = H(q,p)$ and

$$\begin{aligned} f &= (H_p, -H_q)^T \\ \nabla U &= (H_q, H_p)^T \\ \nabla^2 U &= \begin{pmatrix} H_{qq} & H_{qp} \\ H_{pq} & H_{pp} \end{pmatrix}, \end{aligned}$$

we find

$$dH = (\frac{1}{2}\sigma^2 H_{pp})\, dt + \sigma H_p\, dW,$$

and with H of the form in (14) then

$$dH = \frac{\sigma^2}{2}\, dt + \sigma\, p\, dW.$$

Integrating this last equation gives

$$H(t) = H(t_0) + \frac{\sigma^2}{2}\,(t - t_0) + \sigma \int_{t_0}^{t} p(s)\, dW(s)$$

and using the expectation property of Itô integrals this gives

$$E(H(t)) = E(H(t_0)) + \frac{\sigma^2}{2}\,(t - t_0). \tag{15}$$

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

4 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 [16], 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. In terms of considering the evolution of the mean of the Hamiltonian, Melbo and Higham [14] showed that partitioned approaches can be effective.

We shall illustrate these ideas by considering the evolution of the additive noise problem, with no friction and $H = \frac{1}{2}(p^2 + q^2)$. We can write this problem as

$$dy = Qy + \epsilon\, r\, dW, \tag{16}$$

where

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

First consider the general additive noise problem

$$dY = f(Y)\, dt + \epsilon r\, dW,$$

where Y and r (constant vector) are $m \times 1$ column vectors, and $W(t)$ is a scalar Wiener process. Let the numerically-generated approximations be denoted by column vectors y_n. Under an s-stage Runge-Kutta method, y_{n+1} is obtained from y_n as a weighted sum of s evaluations of the function f at intermediate values Y_i:

$$y_{n+1} = y_n + \sum_{j=1}^{s} b_j f(Y_j)\, h + \epsilon\, r\, \Delta W_n, \tag{17}$$

where $\sum_j b_j = 1$. The Wiener increment ΔW_n is sampled from a Gaussian distribution with mean zero and variance h. The column vectors of intermediate values satisfy

$$Y_i = y_n + h \sum_{j=1}^{s} a_{ij} f(Y_j) + \epsilon\, c_i\, r\, \Delta W_n.$$

Applying our Runge-Kutta method (17) to (16) we find

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

where

$$R(z) = 1 + z\, b^T (I - Az)^{-1}\, e$$

and

$$S(hQ) = (hQ)^{-1}\, (R(hQ) - I).$$

Thus

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

Since $Q^T = -Q$, and letting

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

then

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

In the case of a symmetric Runge-Kutta method of order two or more, $P = I$ and so

$$E[H_{n+1}] = E[H_n] + \epsilon^2\, h + \epsilon^2\, \alpha_4\, h^3 + O(h^5),$$

where α_4 is the coefficient of the z^4 term in the expansion of $R(z)$. If the method is order 4 then $\alpha_4 = \frac{1}{24}$ and the local error in the mean of the Hamiltonian is $O(h^3)$. But if $\alpha_4 = 0$ (so the method is order 2) then the local error is $O(h^5)$. This can be achieved by a method whose stability function is

$$R(z) = \frac{1 + \frac{1}{2}z + \frac{1}{8}z^2}{1 - \frac{1}{2}z + \frac{1}{8}z^2}.$$

On the other hand the stationary measure-exact implicit midpoint rule has a local error of $\frac{1}{8}h^3$ in the mean of the Hamiltonian. So there is a trade-off between these two metrics (order of preservation of the mean of the Hamiltonian and stationary measure exactness) when designing appropriate numerical methods.

However there is another representation of a Runge-Kutta method for solving (16) 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]. Consider the general stochastic differential equation

$$dy = f(y)\, dt + g(y)\, dW, \quad y(0) = y_0.$$

In integral form this can be written as

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(y(s))\, ds + \int_{t_n}^{t_{n+1}} g(y(s))\, dW(s),$$

where the stochastic integral is considered in either Itô or Stratonovich form. In the case of additive noise, see (16), these two forms are the same, and

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(y(s))\, ds + \epsilon\, r \int_{t_n}^{t_{n+1}} dW(s).$$

In particular for offstep points $t_n + c_i h$

$$y(t_n + c_i h) = y(t_n) + \int_{t_n}^{t_n + c_i h} f(y(s))\, ds + \epsilon\, r \int_{t_n}^{t_n + c_i h} dW(s). \tag{18}$$

Since the Y_i should represent an approximation to (18) then a different formulation for a Runge-Kutta method is

$$\begin{aligned} Y_i &= y_n + h \sum_{j=1}^{s} a_{ij} f(Y_j) + \epsilon \left(W(t_n + c_i h) - W(t_n)\right) r, \quad i = 1, \cdots, s \\ y_{n+1} &= y_n + h \sum_{j=1}^{s} b_j f(Y_j) + \epsilon \left(W(t_n + h) - W(t_n)\right) r. \end{aligned} \tag{19}$$

However this formula does not take into account that when sampling $W(t_n + c_i h) - W(t_n)$ we should consider this as the sum of Wiener increments

$$\sum_{j=1}^{i} (W(t_n + c_j h) - W(t_n + c_{j-1} h)).$$

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 is

$$\begin{aligned} Y_i &= y_n + h \sum_{j=1}^{s} a_{ij} f(Y_j) + \epsilon \sum_{j=1}^{i} z_j\, r \\ y_{n+1} &= y_n + h \sum_{j=1}^{s} b_j f(Y_j) + \epsilon \sum_{j=1}^{s+1} z_j\, r. \end{aligned} \tag{20}$$

Let

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

and note that

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

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

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

where

$$S(z) = z_{s+1} + e^T\, Z + z\, b^T\, (I - Az)^{-1}\, V\, Z.$$

For example, in formulation (20), the implicit midpoint rule is

$$\begin{aligned} Y &= y_n + \frac{h}{2} f(Y) + \frac{1}{\sqrt{2}} \Delta W_1\, \epsilon\, r \\ y_{n+1} &= y_n + h\, f(Y) + \frac{1}{\sqrt{2}} (\Delta W_1 + \Delta W_2)\, \epsilon\, r \end{aligned} \tag{21}$$

and

$$\begin{aligned} R(hQ) &= (I - \frac{h}{2} Q)^{-1}\, (I + \frac{h}{2} Q) \\ S(hQ) &= z_2 I + R(hQ)\, z_1. \end{aligned} \tag{22}$$

In this formulation we note that

$$E[y_{n+1}] = R(hQ)\, E[y_n]$$

and that

$$E[y_{n+1}^T\, y_{n+1}] = E[y_n^T R^T(hQ)\, R(hQ)\, y_n] + \epsilon^2\, r^T\, E[S^T(hQ)\, S(hQ)]\, r. \tag{23}$$

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

$$R^T(hQ)\, R(hQ) = I$$

$$E[S^T(hQ)\, S(hQ)] = E[z_1^2 + z_2^2]\, I = h\, I.$$

Thus

$$E[H_{n+1}] = E[H_n] + \frac{\epsilon^2}{2}\, h \tag{24}$$

and we have the following result:

Theorem 2. The implicit midpoint rule (21) preserves the expectation of the Hamiltonian $\frac{1}{2}(p^2 + q^2)$ exactly, that is it preserves (15).

If we define

$$R_i(z) = 1 + z\, b^T\, (I - Az)^{-1} \sum_{j=i}^{s} e_j \quad i = 1, \cdots, s+1,$$

where e_j is the jth unit vector and, in particular,

$$R_1(z) = R(z), \quad R_{s+1}(z) = 1$$

then from (23)

$$E[y_{n+1}^T\, y_{n+1}] = E[y_n^T\, R(-hQ)\, R(hQ)\, y_n] + \epsilon^2\, r^T\, E[S(-hQ)\, S(hQ)]\, r. \tag{25}$$

But from the definition of $S(hQ)$ and $R_i(hQ)$ it is easily seen that

$$S(hQ) = \sum_{i=1}^{s+1} z_i\, R_i(hQ).$$

Furthermore, since the z_i are independent and satisfy $E[z_i^2] = (c_i - c_{i-1})h$ then

$$E[S(-hQ)\, S(hQ)] = h \sum_{i=1}^{s+1} (c_i - c_{i-1})\, R_i(-hQ)\, R_i(hQ)$$

and so from (23)

$$E[y_{n+1}^T\, y_{n+1}] = E[y_n^T\, R(-hQ)\, R(hQ)\, y_n] + \epsilon^2\, h\, r^T \sum_{i=1}^{s+1} (c_i - c_{i-1})\, R_i(-hQ)\, R_i(hQ)\, r. \tag{26}$$

Hence we can use higher order methods in the mean and still obtain good properties in preserving the mean of the Hamiltonian.

We can achieve this best by assuming the underlying method has the form for the stability function

$$R(z) = \frac{P(-z)}{P(z)}.$$

Then from (26)

$$E[H_{n+1}] = E[H_n] + \frac{\epsilon^2}{2} C\, h^{l+1}$$

where

$$C = \sum_{i=1}^{s+1} (c_i - c_{i-1})\, D_i\, r^T\, Q^l\, r,$$

where the integer l and the error constants D_i are defined by the expansion

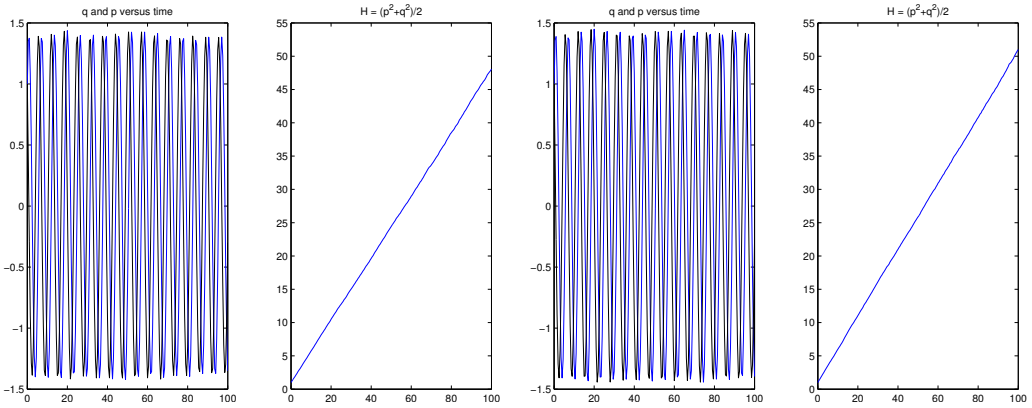
$$R_i(-hQ)\, R_i(hQ) = I + D_i\, h^l + O(h^{l+1}).$$

5 Simulations and Conclusions

We now present some numerical results, demonstrating the behaviour of the Hamiltonian in both a linear and non-linear setting.

Firstly we consider the linear SDE given by (16), and apply the implicit midpoint rule (using a single noise representation in the numerical method (17)). We then compare the results with those obtained from the alternative formulation (21). The following figure (Fig 1) shows the average solution vector and the growth of the Hamiltonian over time, where the integration is carried out for 200 steps with step size 0.5, noise term $\epsilon = 1.0$, and initial value $y_0 = (\ 1 \quad 1\ )^T$. The simulations were averaged over 10000 trajectories. In the single noise formulation, $\max(H(t)) = 48.067804$, while for multi-noise, $\max(H(t)) = 50.995163$. The expected value is 51 at $t = 100$.

Figure 1: single noise (left) and multi-noise (right), for linear SDE (16)



The second example is a non-linear SDE of the form

$$dy = f(y)\,dt + \epsilon \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW$$

where

$$f(y) = \begin{pmatrix} p \\ -\nabla R(q) \end{pmatrix} = \begin{pmatrix} p \\ q - q^3 \end{pmatrix}$$

with Hamiltonian

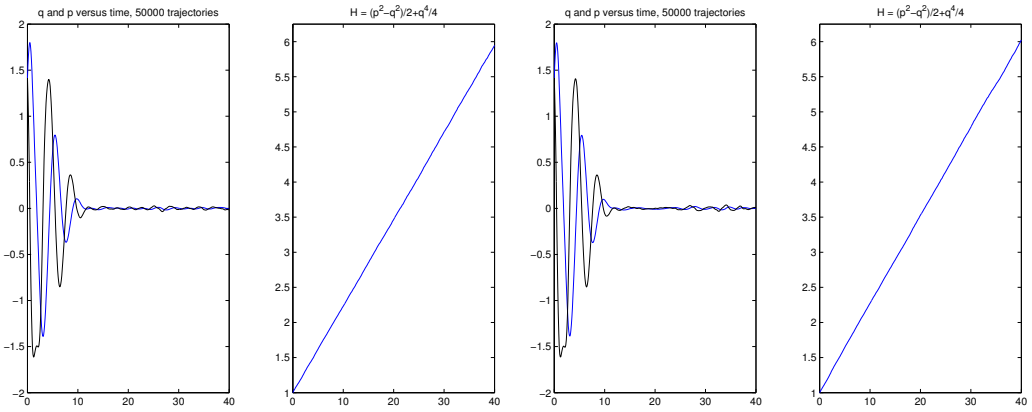
$$H = \frac{1}{2}p^2 + R(q) = \frac{1}{2}p^2 + \frac{1}{4}q^4 - \frac{1}{2}q^2.$$

This is also referred to as the Double Well problem. It is expected that the Hamiltonian grows over time according to

$$E[H(t)] = H(t_0) + \frac{1}{2}\epsilon^2\, t.$$

Again the single and multi-noise formulations are compared and due to the non-linear nature of the problem, we carried out 50000 simulations, over 400 steps with step size 0.1. The noise term was $\epsilon = 0.5$, and with initial value $y_0 = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \end{pmatrix}$, we would expect H to grow to $H(0) + \frac{1}{2}\epsilon^2\, t = 6$ at $t = 40$. For the single noise formulation, $\max(H(t)) = 5.946045$, while for the multi-noise approach, $\max(H(t)) = 6.022872$. Figure 2 shows the average solution y and the Hamiltonian over time.

Figure 2: single noise (left) and multi-noise (right) for the double well problem



In conclusion, we see that some of the ideas that lead to the class of HBVM methods, introduced by Brugnano et al. can be applied to the class of additive noise Hamiltonian problems. This leads to the generation of $s+1$ Wiener processes per step for an s-stage Runge-Kutta method. In the case of the implicit midpoint rule this leads to an implementation with two independent Wiener processes per step and such methods preserve the mean of the Hamiltonian exactly in the case of linear problems.

6 Acknowledgements

The authors would like to thank Luigi Brugnano for many interesting discussions when he visited Oxford in April 2011. The authors would like to pay their respects to Donato Trigiante, a fine mathematician and a fine man, and to recall happy times talking about life in his home in Pisa.

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*, in press doi:10.1016/j.amc.2011.03.022.

[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*, accepted for publication, doi:10.1016/j.amc.2012.01.074.

[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. (1978): High order algebraically stable Runge-Kutta methods, *BIT* **18**, 373–383.

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

[8] Burrage, K. and Butcher, J.C. (1979): Stability criteria for implicit Runge-Kutta methods. *SIAM Journal on Numerical Analysis*, **16**:4657.

[9] Burrage, K., Lenane, I. and Lythe, G. (2007): Numerical methods for second-order stochastic differential equations, *SIAM Journal on Scientific Computing*, **29**(1):245264.

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

[11] Gardiner, C.W. (2004): *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer, 3rd edition.

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

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

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

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

[16] Schurz H. (1999): The invariance of asymptotic laws of linear stochastic systems under discretization. *Z. Angew Math. Mech.*, **6**:375382.

[17] 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.