NUMERICAL METHODS FOR SECOND-ORDER STOCHASTIC EQUATIONS

KEVIN BURRAGE,* IAN LENANE* AND GRANT LYTHE †

Abstract. We seek numerical methods for second-order stochastic differential equations that accurately reproduce the stationary distribution for all values of damping. A complete analysis is possible for linear second-order equations (damped harmonic oscillators with noise), where the statistics are Gaussian and can be calculated exactly in the continuous-time and discrete-time cases. A matrix equation is given for the stationary variances and correlation for methods using one Gaussian random variable per timestep. The only Runge-Kutta method with a nonsingular tableau matrix in the class that gives the exact steady state density for all values of damping is the implicit midpoint rule. Numerical experiments comparing the implicit midpoint rule with Heun and leapfrog methods suggest that the qualitative behavior is similar to the linear case for nonlinear equations with additive or multiplicative noise.

Key words. damped harmonic oscillators with noise, stationary distribution, stochastic Runge-Kutta methods, implicit midpoint rule, multiplicative noise.

AMS subject classifications. 60-08, 65c30.

1. Introduction. Newton's law states that acceleration is proportional to force. Consequently, second-order differential equations are common in scientific applications, in the guise of "Langevin", "Monte Carlo", "Molecular" or "Dissipative particle" dynamics [1, 2], and the study of methods for second-order ordinary differential equations is one of the most mature branches of numerical analysis [3]. The most exciting advances in recent decades have been the development of symplectic methods, capable of exactly preserving an energy-like quantity over very long times [4]. In the stochastic setting, the long-time dynamics of a typical physical system is governed by fluctuation-dissipation, so that the amount of time spent in different regions of phase space is controlled by the stationary density. The stationary density can have a relatively simple explicit expression even when the dynamics is highly nonlinear [5]. Numerical methods replace continuous-time with discrete-time dynamics, generating values at times $t_0, t_1, \ldots$. Usually $t_{n+1} - t_n$ is a fixed number Δt. The criterion for a good numerical method that will be examined in this work is that its discrete time dynamics has a steady-state density as close as possible to that of the continuous-time system.

The differential equations describing second-order systems contain a parameter known as damping. The steady-state density is independent of damping, but time-dependent quantities and the usefulness or otherwise of numerical algorithms are strongly dependent. As the damping tends to infinity the system becomes first order. The limit of zero damping, on the other hand, corresponds to Hamiltonian systems where symplectic methods can be applied. The aim in this paper is to devise methods capable of accurately reproducing the steady-state density *for all values of damping.*

We shall consider equations of the following form:

$$\ddot{x} = f(x) - \eta\dot{x} + \epsilon\xi(t), \tag{1.1}$$

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

*Advanced Computational Modelling Centre, Department of Mathematics, University of Queensland, Australia.

†Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, England

(SDE) (1.1) 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 $V(x)$ via

$$f(x) = -V'(x). \tag{1.2}$$

The amplitude of the random forcing, ϵ, is related to the temperature T and damping coefficient η by the fluctuation-dissipation relation

$$\epsilon^2 = 2\eta KT. \tag{1.3}$$

We can write (1.1) as a pair of first-order equations for $\mathbf{X}_t$ and $\mathbf{V}_t$, the position and velocity variables:

$$\begin{aligned} \mathrm{d}\mathbf{X}_t &= \mathbf{V}_t \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\eta \mathbf{V}_t \mathrm{d}t + f(\mathbf{X}_t)\mathrm{d}t + \epsilon \mathrm{d}\mathbf{W}_t, \end{aligned} \tag{1.4}$$

where $\mathbf{W}_t$ is a Wiener process satisfying $\left\langle \mathbf{W}_t \mathbf{W}_s \right\rangle = \min(t, s)$. The probability density at time t is $P(x, v; t)$, where

$$P(x, v; t) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{Prob}\left(\mathbf{X}_t < x, \mathbf{V}_t < v\right). \tag{1.5}$$

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

$$P_\infty(x, v) = \lim_{t\to\infty} P(x, v; t), \tag{1.6}$$

has the following analytical form:

$$P_\infty(x, v) = N \exp\left(-v^2/2KT - V(x)/KT\right). \tag{1.7}$$

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 this paper we examine how faithfully the stationary density is reproduced by standard timestepping methods for SDEs. These methods produce approximate values for position X_n and velocity V_n at discrete times $t_0, t_1, \ldots$, where $t_{n+1} - t_n = \Delta t$. We consider the evolution of X_n and V_n and their statistical properties as $t_n \to \infty$, and compare with the exact form (1.6). In particular, we consider $P^*(y, u; t_n)$, the discrete-time analogue of (1.5), and compare

$$P^*_\infty(y, u) = \lim_{t_n\to\infty} P^*(y, u; t_n) \tag{1.8}$$

with $P_\infty(y, u)$. In this work we instead base our analysis on linear second-order equations, where the statistics are Gaussian and completely characterized by three quantities: the mean-squares of the position and velocity variables, and the correlation between the position and velocity.

The Euler and Heun methods simply treat (1.4) as a pair of SDEs, without attempting to take advantage of their special structure due to their origin in a single second-order differential equation. We shall show that they perform reasonably well at intermediate values of damping, but fail at high damping, when the equations are

stiff, and low damping, when it is important to conserve energy-like quantities over long times.

The Verlet algorithm [1, 8, 9] produces numerical solutions of second order differential equations by updating the position variable without reference to the velocity variable. Extensions of this idea to Langevin equations [9, 10, 11, 15] have proved successful, and convergence properties may be analyzed using a second-order difference equations [11, 12, 13]. The implicit midpoint rule cannot be written as a difference equation but can be written as a matrix equation [13]. Time-dependent properties of numerical algorithms for the class of linear equations considered in this paper can be considered by means of a modified frequency [13] or by the method of modified equations [16].

Recently, analysis by Mannella [20] of the stationary distribution resulting from numerical timestepping methods, based on an expansion of the exponent in (1.7), lead to a proposed modification of the leapfrog method. We shall take this method as one of our examples and find that it provides a notable improvement on the standard leapfrog method.

A slightly different system has been studied by Melbo and Higham [21]. They considered a second-order system, but without a damping term. Rather than approaching a steady-state density, the sum of the mean squared velocity and position grows proportional to time. Partitioned methods are superior to the Euler method for this system.

In Section 2, we consider linear second-order equations, describing harmonic oscillators with noise and damping. The statistics for linear systems are Gaussian, described by a covariance matrix, and can be calculated exactly in the continuous time and discrete time cases. A matrix equation is derived for the stationary variances and correlations resulting from a large class of numerical methods. We then calculate the stationary variances produced by some well-known methods. All methods we shall consider use one realization per timestep of a Gaussian distribution with mean zero and variance Δt, denoted ΔW. This framework includes many methods with multiple intermediate steps and implicit methods (which are in fact explicit for linear equations). The Euler and Heun methods are shown to be unstable for sufficiently small and sufficiently large damping, while leapfrog methods are stable for arbitrarily small damping. Instability reveals itself in this context as infinite stationary means or correlations.

In Section 3 we seek "measure-exact" methods for linear second-order SDEs, that is methods that give the correct late-time mean square of the position and velocity and absence of correlation between them. In Section 4, we restrict consideration to Runge-Kutta methods and show that the unique measure-exact method is the implicit midpoint rule. In Section 5 we report on numerical experiments carried out on a nonlinear system; in Section 6 we report on numerical experiments carried out on the same system with multiplicative noise. In both cases the implicit midpoint rule is the most satisfactory over the full range of values of damping.

2. Linear equation and matrix notation. If $f(x) = -gx$, then the stationary density is

$$P_\infty(x, v) = N \exp\left(-g\, x^2/2KT - v^2/2KT\right).$$

The distributions of the position and velocity variables are Gaussian with

$$\begin{aligned}\lim_{t\to\infty}\left\langle \mathbf{X}_t{}^2\right\rangle &= \frac{1}{g}KT\\ \lim_{t\to\infty}\left\langle \mathbf{V}_t^2\right\rangle &= KT\\ \lim_{t\to\infty}\left\langle \mathbf{X}_t\mathbf{V}_t\right\rangle &= 0.\end{aligned} \tag{2.1}$$

In matrix notation, the linear second-order SDE can be written as

$$\mathrm{d}\begin{pmatrix}\mathbf{X}_t\\ \mathbf{V}_t\end{pmatrix} = Q\begin{pmatrix}\mathbf{X}_t\\ \mathbf{V}_t\end{pmatrix}\mathrm{d}t + \epsilon\begin{pmatrix}0\\ 1\end{pmatrix}\mathrm{d}\mathbf{W}_t, \tag{2.2}$$

where

$$Q = \begin{pmatrix}0 & 1\\ -g & -\eta\end{pmatrix}. \tag{2.3}$$

Numerical methods, single or multistep, for the linear system can be written as

$$\begin{pmatrix}X_{n+1}\\ V_{n+1}\end{pmatrix} = R\begin{pmatrix}X_n\\ V_n\end{pmatrix} + \epsilon r\Delta W, \tag{2.4}$$

where

$$R = \begin{pmatrix}r_{11} & r_{12}\\ r_{21} & r_{22}\end{pmatrix} \qquad \text{and} \qquad r = \begin{pmatrix}r_1\\ r_2\end{pmatrix}. \tag{2.5}$$

Since the numerical evolution is also linear, $P_\infty^*(x,v)$ is also Gaussian. Let

$$\Sigma = \begin{pmatrix}\sigma_x^2 & \mu\\ \mu & \sigma_v^2\end{pmatrix}, \tag{2.6}$$

where

$$\begin{aligned}\sigma_x^2 &= \lim_{t_n\to\infty}\left\langle X_n^2\right\rangle\\ \sigma_v^2 &= \lim_{t_n\to\infty}\left\langle V_n^2\right\rangle\\ \mu &= \lim_{t_n\to\infty}\left\langle X_nV_n\right\rangle.\end{aligned} \tag{2.7}$$

Then

$$P_\infty^*(x,v) = \frac{1}{2\pi}|M^*|^{\frac{1}{2}}\exp\left(-\frac{1}{2}(x,v)M^*\begin{pmatrix}x\\ v\end{pmatrix}\right),$$

where $M^* = \Sigma^{-1}$.

The stationary density of a numerical method is defined as being unchanged by the transformation (2.4). Now, if $\begin{pmatrix}X_n\\ V_n\end{pmatrix}$ is Gaussian with mean zero and correlation matrix Σ, then $R\begin{pmatrix}X_n\\ V_n\end{pmatrix}$ is Gaussian with mean zero and correlation matrix $R\Sigma R^{\mathrm{T}}$. Thus the stationary correlation matrix that results from a method of the form (2.4) satisfies

$$\Sigma = R\Sigma R^{\mathrm{T}} + \epsilon^2 rr^{\mathrm{T}}\Delta t$$

or

$$R\Sigma R^{\mathrm{T}} = \Sigma - \epsilon^2 r r^{\mathrm{T}} \Delta t. \tag{2.8}$$

In the remainder of this Section, we shall use (2.8) to calculate the σ_x^2, μ, and σ_v^2 as a function of η and Δt for various numerical methods. We rewrite (2.8) in a form suitable for inversion:

$$\begin{pmatrix} r_{11}^2 - 1 & 2r_{11}r_{12} & r_{12}^2 \\ r_{11}r_{21} & r_{11}r_{22} + r_{12}r_{21} - 1 & r_{12}r_{22} \\ r_{21}^2 & 2r_{21}r_{22} & r_{22}^2 - 1 \end{pmatrix} \begin{pmatrix} \sigma_x^2 \\ \mu \\ \sigma_v^2 \end{pmatrix} = -\epsilon^2 \Delta t \begin{pmatrix} b_1^2 \\ b_1 b_2 \\ b_2^2 \end{pmatrix}, \tag{2.9}$$

and consider what happens for some well-known methods.

2.1. Forward Euler method. Under the Euler method, the position and velocity variables are updated as follows:

$$\begin{aligned} X_{n+1} &= X_n + V_n \, \Delta t \\ V_{n+1} &= V_n - \eta V_n \Delta t + f(X_n) \, \Delta t + \epsilon \Delta W. \end{aligned} \tag{2.10}$$

With the notation of (2.4), $R = R_{\mathrm{E}}$ and $r = r_{\mathrm{E}}$ where

$$R_{\mathrm{E}} = 1 + \Delta t Q = \begin{pmatrix} 1 & \Delta t \\ -g\Delta t & 1 - \eta \Delta t \end{pmatrix}, \tag{2.11}$$

$r_{\mathrm{E}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and the solution of (2.8) yields

$$\Sigma_{\mathrm{E}} = \frac{KT}{1 - \frac{g}{\eta}\Delta t} (2 - \eta \Delta t + \frac{1}{2} g \Delta t^2)^{-1} \begin{pmatrix} \frac{1}{g}(2 - \eta \Delta t + g \Delta t^2) & -\Delta t \\ -\Delta t & 2 \end{pmatrix}. \tag{2.12}$$

In Figure 2.1 we display the differences between the stationary variances obtained by applying the Euler method to the linear equation and the exact values (2.1). We conclude that the forward Euler method works best at intermediate values of η. It is unstable for $\eta < g\Delta t$ and for $\eta > 2\Delta t^{-1} + \frac{1}{2} g \Delta t$. In Figure 2.2, we plot the correlation between velocity and position, which is zero in the exact solution but is proportional to Δt under the Euler method.

2.2. The Heun method. Under the Heun method, intermediate values are obtained via the Euler method:

$$\begin{aligned} \hat{X} &= X_n + V_n \, \Delta t \\ \hat{V} &= V_n - \eta V_n \Delta t + f(X_n) \, \Delta t + \epsilon \, \Delta W. \end{aligned} \tag{2.13}$$

The update is

$$\begin{aligned} X_{n+1} &= X_n + \frac{1}{2}(V_n + \hat{V}) \, \Delta t \\ V_{n+1} &= V_n - \frac{1}{2}\eta (V_n + \hat{V}) \Delta t + \frac{1}{2}(f(X_n) + f(\hat{X})) \, \Delta t + \epsilon \, \Delta W. \end{aligned} \tag{2.14}$$

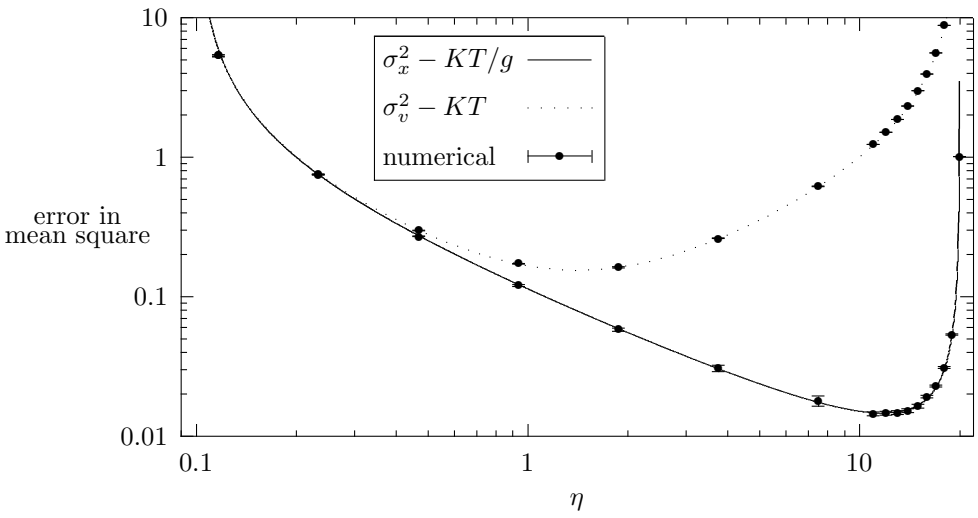

FIG. 2.1. *Stationary variances vs damping for the forward Euler method. The differences between the late-time mean squares and the exact values are plotted against η for fixed $\Delta t = 0.1$, $g = 1$ and $KT = 1.0$. The lines use* (2.12)*. Dots with error bars are obtained from numerical solution of the linear SDE.*

With the notation of (2.4), $R = R_{\rm H}$ and $r = r_{\rm H}$ where

$$R_{\rm H} = I + \Delta t Q + \frac{1}{2}\Delta t Q^2 = \begin{pmatrix} 1 - \frac{1}{2}g\Delta t^2 & \Delta t - \frac{1}{2}\eta\Delta t^2 \\ -g\Delta t + \frac{1}{2}\eta g\Delta t^2 & 1 - \eta\Delta t + \frac{1}{2}(-g+\eta^2)\Delta t^2 \end{pmatrix}$$

and

$$r_{\rm H} = \begin{pmatrix} \frac{1}{2}\Delta t \\ 1 - \frac{1}{2}\eta\Delta t \end{pmatrix}.$$

Then the stationary density is characterized by

$$\sigma_x^2 = \frac{KT}{g}\frac{4 - \eta\Delta t + \frac{1}{2}\eta^2\Delta t^2 - \frac{1}{8}(g\eta + \eta^3)\Delta t^3 + \frac{3}{16}g\eta^2\Delta t^4 - \frac{1}{8}g^2\eta\Delta t^5 + \frac{1}{32}g^3\Delta t^6}{\left(2 - \eta\Delta t + (\eta^2 - g)\Delta t^2 - \frac{1}{4}\eta g\Delta t^3 + g^2\Delta t^4/8\right)\left(2 - \eta\Delta t + g\Delta t^2 - g^2\Delta t^3/(2\eta)\right)}$$

$$\sigma_v^2 = KT\frac{(2 - \eta\Delta t)^2}{\left(2 - \eta\Delta t + (\eta^2 - g)\Delta t^2 - \frac{1}{4}\eta g\Delta t^3 + g^2\Delta t^4/8\right)\left(2 - \eta\Delta t + g\Delta t^2 - g^2\Delta t^3/(2\eta)\right)}$$

$$\mu = \frac{\Delta t^2 KT}{4}\frac{(2 - \eta\Delta t)(2\eta - g\Delta t)}{\left(2 - \eta\Delta t + (\eta^2 - g)\Delta t^2 - \frac{1}{4}\eta g\Delta t^3 + g^2\Delta t^4/8\right)\left(2 - \eta\Delta t + g\Delta t^2 - g^2\Delta t^3/(2\eta)\right)}.$$

Compared with the Euler method, the upper limit on η for stability is little changed. The lower limit is, however, proportional to Δt^3. As $\Delta t \to 0$, the Heun method is stable for

$$\frac{1}{4}g^2\Delta t^3 < \eta < 2\Delta t^{-1}.$$

The correlation between position and velocity is proportional to Δt^2 under the Heun method.

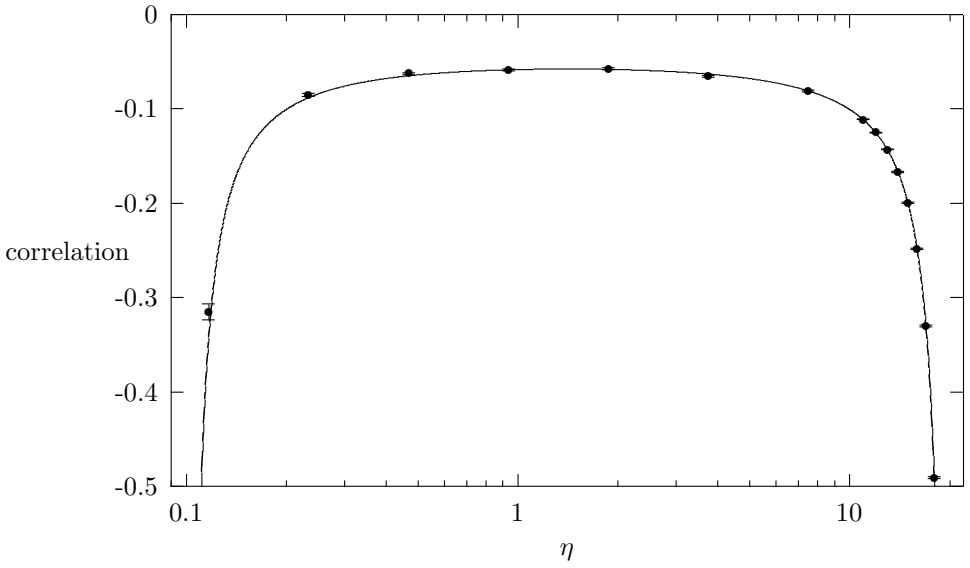

FIG. 2.2. *Stationary correlation vs damping for the forward Euler method. The late-time correlation, μ, obtained using the forward Euler method, is plotted against η for fixed $\Delta t = 0.1$, $g = 1$ and $KT = 1.0$. The line use* (2.12). *Dots with error bars are obtained from numerical solution of the linear SDE.*

2.3. The leapfrog method. Under leapfrog methods, velocity and position are updated successively rather than together. The simplest possibility is

$$\begin{aligned} \hat{X} &= X_n + \frac{1}{2} V_n \Delta t \\ V_{n+1} &= V_n - \eta V_n \Delta t + f(\hat{X})\, \Delta t + \epsilon\, \Delta W \\ X_{n+1} &= \hat{X} + \frac{1}{2} V_{n+1} \Delta t. \end{aligned} \tag{2.15}$$

Under this method, with the notation of (2.4), $R = R_{\mathrm{L}}$ and $r = r_{\mathrm{L}}$ where

$$R_{\mathrm{L}} = \begin{pmatrix} 1 - \frac{1}{2} g \Delta t^2 & \Delta t - \frac{1}{2} \eta \Delta t^2 - \frac{1}{4} g \Delta t^3 \\ -g \Delta t & 1 - \eta \Delta t - \frac{1}{2} g \Delta t^2 \end{pmatrix},$$

and

$$r_{\mathrm{L}} = \begin{pmatrix} \frac{1}{2} \Delta t \\ 1 \end{pmatrix}.$$

The leapfrog method maintains the independence of position and velocity and produces the exact stationary variance of the position variable. However the error in σ_v^2 is an increasing function of η:

$$\Sigma_{\mathrm{le}} = KT \begin{pmatrix} 1/g & 0 \\ 0 & (1 - \frac{1}{2} \eta \Delta t - \frac{1}{4} g \Delta t^2)^{-1} \end{pmatrix}. \tag{2.16}$$

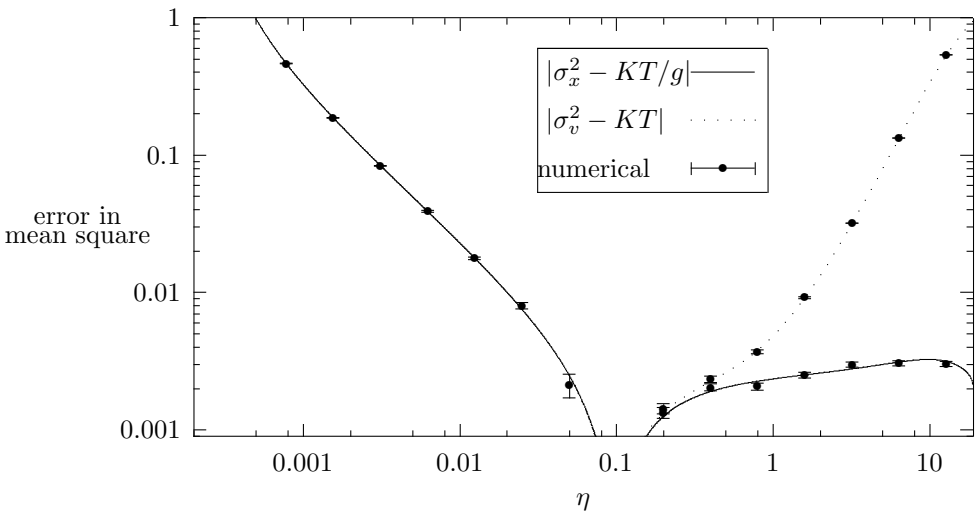

FIG. 2.3. *Stationary variances vs damping for the Heun method. The differences between the late-time mean squares and the exact values are plotted against η for fixed $\Delta t = 0.1$, $g = 1$ and $KT = 1.0$. The lines use* (2.15)*. Dots with error bars are obtained from numerical solution of the linear SDE.*

Mannella [20] has proposed the following modification of the leapfrog method:

$$\begin{aligned} \hat{X} &= X_n + \frac{1}{2} V_n \Delta t \\ V_{n+1} &= c_1 \left(c_2 V_n + f(\hat{X})\, \Delta t + \epsilon\, \Delta W \right) \\ X_{n+1} &= \hat{X} + \frac{1}{2} V_{n+1} \Delta t, \end{aligned} \tag{2.17}$$

where $c_1 = 1 - \frac{1}{2}\eta\Delta t$ and $c_2 = (1 + \frac{1}{2}\eta\Delta t)^{-1}$. The corresponding quantities are

$$R_{\mathrm{M}} = \begin{pmatrix} 1 - c_2 \frac{1}{2} g \Delta t^2 & \frac{1}{2}\Delta t \left(1 + c_1 c_2 - \frac{1}{2} c_2 \Delta t^2\right) \\ -c_2 g \Delta t & c_1 c_2 - \frac{1}{2} c_2 g \Delta t^2 \end{pmatrix},$$

$$r_{\mathrm{L}} = c_2 \begin{pmatrix} \frac{1}{2}\Delta t \\ 1 \end{pmatrix},$$

and

$$\Sigma_{\mathrm{M}} = KT \begin{pmatrix} 1/g & 0 \\ 0 & (1 - \frac{1}{2}\eta\Delta t - \frac{1}{4} g \Delta t^2)^{-1} \end{pmatrix}. \tag{2.18}$$

The only error, in σ_v^2, is independent of η.

2.4. The BBK method. Under the BBK method [10], the position variable is updated without reference to the velocity variable. The two most recent values of the

position variable need to be retained and the update is

$$X_{n+1} = X_n + \frac{1-\frac{1}{2}\eta\Delta t}{1+\frac{1}{2}\eta\Delta t}(X_n - X_{n-1}) + \frac{\Delta t}{1+\frac{1}{2}\eta\Delta t}\left(f(X_n) + \epsilon\,\Delta W\right). \tag{2.19}$$

With the definition of velocity $V_n\Delta t = X_n - X_{n-1}$, proposed in [11], the BBK method can be written in the form (2.4) with [14]

$$R_{\mathrm{B}} = \begin{pmatrix} 1-\frac{g\Delta t^2}{1+\frac{1}{2}\eta\Delta t} & \frac{1-\frac{1}{2}\eta\Delta t}{1+\frac{1}{2}\eta\Delta t}\Delta t \\ -\frac{g\Delta t}{1+\frac{1}{2}\eta\Delta t} & \frac{1-\frac{1}{2}\eta\Delta t}{1+\frac{1}{2}\eta\Delta t} \end{pmatrix},$$

and

$$r_{\mathrm{B}} = \frac{1}{1+\frac{1}{2}\eta\Delta t}\begin{pmatrix}\Delta t \\ 1\end{pmatrix}.$$

Whatever the definition of velocity, we find [11, 13, 16]

$$\sigma_x^2 = \frac{KT}{g}\frac{1}{1-\frac{1}{4}g\Delta t^2}. \tag{2.20}$$

However σ_v^2 and μ do depend on the definition of velocity. Note that, with the velocity definition $2V_n\Delta t = X_{n+1} - X_{n-1}$, the BBK update cannot be put into the form (2.4).

2.5. The implicit midpoint method. Under the implicit midpoint method, intermediate values are obtained via the implicit procedure

$$\hat{X} = X_n + \frac{1}{2}\hat{V}\,\Delta t$$

$$\hat{V} = V_n - \frac{1}{2}\eta\hat{V}\Delta t + \frac{1}{2}f(\hat{X})\,\Delta t + \epsilon\,\Delta W. \tag{2.21}$$

The update is

$$X_{n+1} = X_n + \hat{V}\,\Delta t$$

$$V_{n+1} = V_n - \eta\hat{V}\Delta t + f(\hat{X})\,\Delta t + \epsilon\,\Delta W. \tag{2.22}$$

With the notation of (2.4), $R = R_{\mathrm{I}}$ and $r = r_{\mathrm{I}}$ where

$$R_{\mathrm{I}} = \left(I - \frac{1}{2}\Delta t Q\right)^{-1}\left(I + \frac{1}{2}\Delta t Q\right) = \begin{pmatrix} 1-\frac{1}{2}g\Delta t^2 & \alpha^{-1}\Delta t \\ -\alpha^{-1}\Delta t & 2\alpha^{-1}-1 \end{pmatrix},$$

and

$$r_{\mathrm{I}} = \alpha^{-1}\begin{pmatrix}\frac{1}{2}\Delta t \\ 1\end{pmatrix},$$

where $\alpha = 1 + \frac{1}{2}\eta\Delta t + \frac{1}{4}g\Delta t^2$, then

$$\Sigma_{\mathrm{im}} = KT\begin{pmatrix} 1/g & 0 \\ 0 & 1\end{pmatrix}. \tag{2.23}$$

Since R_{I} and r_{I} satisfy (2.8), the implicit midpoint method gives the exact stationary variances for all values of η. Only an algebraic error in Appendix A of [13] prevented this observation being made by Mishra and Schlick. The behavior of the implicit midpoint method is explored further in Section 4.

3. "Measure-Exact" algorithms. The exact covariance matrix is.

$$\Sigma_{\rm e} = KT \begin{pmatrix} g^{-1} & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.1}$$

We want to find R and r that satisfy (2.8) with $\Sigma = \Sigma_e$, that is

$$R \begin{pmatrix} g^{-1} & 0 \\ 0 & 1 \end{pmatrix} R^{\rm T} = \begin{pmatrix} g^{-1} & 0 \\ 0 & 1 \end{pmatrix} - 2\eta r r^{\rm T} \Delta t. \tag{3.2}$$

We shall assume Σ is positive definite and let Δt is be chosen small enough to ensure

$$G = \Sigma - r r^{\rm T} \Delta t$$

is also positive definite. Matrices $\Sigma^{\frac{1}{2}}$ and $G^{\frac{1}{2}}$ can be constructed to satisfy

$$G = G^{\frac{1}{2}} \left(G^{\frac{1}{2}}\right)^{\rm T},$$
$$\Sigma = \Sigma^{\frac{1}{2}} \left(\Sigma^{\frac{1}{2}}\right)^{\rm T}.$$

Let $\Sigma^{-\frac{1}{2}}$ and $G^{-\frac{1}{2}}$ be the inverses of $\Sigma^{\frac{1}{2}}$ and $G^{\frac{1}{2}}$. Condition (2.8) can then be rewritten as

$$P P^{\rm T} = I,$$

where $P = G^{-\frac{1}{2}} R \Sigma^{\frac{1}{2}}$ and I is the identity matrix. Any $n \times n$ orthogonal matrix P generates a solution of (3.2) as

$$R = G^{\frac{1}{2}} P \Sigma^{-\frac{1}{2}}. \tag{3.3}$$

A real 2×2 orthogonal matrix can be written as

$$P = \begin{pmatrix} (1-a^2)^{1/2} & a \\ -a & (1-a^2)^{1/2} \end{pmatrix},$$

for some $|a| \leq 1$. This form is convenient because we can take $a \propto \Delta t$; P reduces to the identity matrix when $a = 0$.

With Σ given by (3.1), we can choose

$$\Sigma^{\frac{1}{2}} = (KT)^{1/2} \begin{pmatrix} g^{-1/2} & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.4}$$

So

$$\begin{aligned} G &= \Sigma - r r^{\rm T} \Delta t \\ &= \begin{pmatrix} KT/g - r_1^2 \Delta t & -r_1 r_2 \Delta t \\ -r_1 r_2 \Delta t & KT - r_2^2 \Delta t \end{pmatrix}, \end{aligned}$$

and

$$G^{\frac{1}{2}} = \begin{pmatrix} \left(KT/g - r_1^2 \Delta t\right)^{\frac{1}{2}} & 0 \\ \frac{-r_1 r_2 \Delta t}{\left(KT/g - r_1^2 \Delta t\right)^{\frac{1}{2}}} & \left(KT - r_2^2 \Delta t + \frac{(r_1 r_2 \Delta t)^2}{KT/g - r_1^2 \Delta t}\right)^{\frac{1}{2}} \end{pmatrix}. \tag{3.5}$$

Any measure-exact method, using one Gaussian random variable per timestep, can be obtained from (3.3) with some choice of a, r_1 and r_2. For example, we may obtain a measure-exact Euler-like method by choosing $r = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. That is,

$$G^{1/2} = (KT)^{1/2} \begin{pmatrix} g^{-1/2} & 0 \\ 0 & (1-2\eta\Delta t)^{1/2} \end{pmatrix}$$

and

$$R = G^{1/2} P \Sigma^{-1/2} = \begin{pmatrix} (1-a^2)^{1/2} & a g^{-1/2} \\ -a g^{1/2}(1-2\eta\Delta t)^{1/2} & (1-a^2)^{1/2}(1-2\eta\Delta t)^{1/2} \end{pmatrix}.$$

If we use the one remaining degree of freedom by choosing a such that the Euler method is obtained as $\Delta t \to 0$,

$$a = g^{1/2}\Delta t,$$

then

$$R = \begin{pmatrix} (1-g\Delta t^2)^{1/2} & \Delta t \\ -g\Delta t(1-2\eta\Delta t)^{1/2} & (1-g\Delta t^2)^{1/2}(1-2\eta\Delta t)^{1/2} \end{pmatrix}. \tag{3.6}$$

Notice that the difference between (3.6) and (2.11) is Order Δt^2. This modified Euler method reproduces the *exact stationary density* as long as $2\eta\Delta t < 1$. As a next step, we may impose agreement of R with $\exp(\Delta t Q)$ to order Δt^2:

$$R = I + \Delta t Q + \frac{1}{2}\Delta t^2 Q^2 + \cdots. \tag{3.7}$$

Then $a = g^{\frac{1}{2}}\Delta t - \frac{1}{2} g^{\frac{1}{2}} \eta \Delta t^2$, $r_1 = \frac{1}{2}\Delta t + \cdots$ and $r_2 = 1 - \eta\frac{1}{2}\Delta t + \cdots$. Methods constructed in this way will have a stationary density differing from the exact density by some power of Δt.

In the next Section we return to the question of solving the condition (3.2) exactly, yielding an R matrix and corresponding numerical method that reproduces the exact stationary density.

4. Runge-Kutta methods. Consider the m-dimensional additive noise SDE [7]

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}t + \epsilon H \mathrm{d}\mathbf{B}_t, \tag{4.1}$$

where $\mathbf{Y}_t$ and $\mathbf{B}_t$ are $m \times 1$ column vectors, the entries of $\mathbf{B}_t$ are independent Wiener processes and H is an $m \times m$ matrix with constant entries. 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:

$$y_{n+1} = y_n + \sum_{j=1}^{s} b_j f(Y_i)\Delta t + \epsilon H \Delta B_n, \tag{4.2}$$

where $\sum_j b_j = 1$. The intermediate values satisfy

$$Y_i = y_n + \sum_{j=1}^{s} a_{ij} f(Y_j)\Delta t + \epsilon c_i H \Delta B_n. \tag{4.3}$$

Let $Y = (Y_1, Y_2, \ldots, Y_s)^{\mathrm{T}}$, $c = (c_1, c_2, \ldots, c_s)^{\mathrm{T}}$, $e = (1, 1, \ldots, 1)^{\mathrm{T}}$ and A be the $s \times s$ matrix whose entries are the coefficients in (4.3). For the linear equation (2.2) with $m = 2$, $H = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and $f(y) = Qy$,

$$Y = e \otimes y_n + A \otimes QY\Delta t + \epsilon c \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Delta W. \tag{4.4}$$

If $b = (b_1, b_2, \ldots, b_s)^{\mathrm{T}}$, we can write

$$\begin{aligned} y_{n+1} &= y_n + b^{\mathrm{T}} Y \otimes Q\Delta t + \epsilon \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Delta W \\ &= y_n + b^{\mathrm{T}} \otimes Q \left(I_s \otimes I_m - A \otimes Q\Delta t\right)^{-1} \left(e \otimes y_n + \epsilon c \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Delta W\right) \\ &\qquad + \epsilon \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Delta W \\ &= \left(I_m + b^{\mathrm{T}} \otimes Q \left(I_s \otimes I_m - A \otimes Q\Delta t\right)^{-1} e\right) y_n \\ &\qquad + \epsilon \left(I_m + b^{\mathrm{T}} \otimes Q \left(I_s \otimes I_m - A \otimes Q\Delta t\right)^{-1} c\right) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Delta W. \end{aligned} \tag{4.5}$$

In the notation of (2.4), with $r = R_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$,

$$y_{n+1} = R(\Delta tQ) y_n + R_1(\Delta tQ) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \epsilon \Delta W, \tag{4.6}$$

where, for scalar z,

$$R(z) = 1 + b^{\mathrm{T}} z (I_s - Az)^{-1} e, \tag{4.7}$$

and

$$R_1(z) = 1 + b^{\mathrm{T}} z (I_s - Az)^{-1} c. \tag{4.8}$$

Let us now examine the condition for exact stationary measure. We shall take $g = 1$, which can always be achieved by rescaling time, so

$$\Sigma_{\mathrm{e}} = KT \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The equation to be satisfied is now

$$RR^{\mathrm{T}} - I + 2\eta \Delta t R_1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} R_1^{\mathrm{T}} = 0. \tag{4.9}$$

Let

$$R(\Delta tQ) = I + \sum_{i=1}^{\infty} \alpha_i (\Delta tQ)^i \tag{4.10}$$

and

$$R_1(\Delta tQ) = I + \sum_{i=1}^{\infty} \beta_i (\Delta tQ)^i. \tag{4.11}$$

Then (4.9) can be expanded in powers of Δt as

$$\begin{aligned}
&\Delta t\left(\alpha_1(Q+Q^{\mathrm{T}})+2\eta\begin{pmatrix}0&0\\0&1\end{pmatrix}\right)\\
&+\Delta t^2\left(\alpha_1^2QQ^{\mathrm{T}}+\alpha_2(Q^2+(Q^{\mathrm{T}})^2)+2\eta\beta_1(Q\begin{pmatrix}0&0\\0&1\end{pmatrix}+\begin{pmatrix}0&0\\0&1\end{pmatrix}Q^{\mathrm{T}})\right)\\
&+\Delta t^3\Bigg(\alpha_3(Q^3+(Q^{\mathrm{T}})^3)+\alpha_1\alpha_2(Q^2Q^{\mathrm{T}}+Q(Q^2)^{\mathrm{T}})\\
&\qquad+2\eta(\beta_2Q^2\begin{pmatrix}0&0\\0&1\end{pmatrix}+\beta_2\begin{pmatrix}0&0\\0&1\end{pmatrix}(Q^{\mathrm{T}})^2+\beta_1^2Q\begin{pmatrix}0&0\\0&1\end{pmatrix}Q^{\mathrm{T}})\Bigg)+\cdots=0.
\end{aligned}\tag{4.12}$$

If this equation is to be satisfied for all Δt then each coefficient in the expansion must be zero. It is easily shown that this can hold if and only if

$$\alpha_i=(1/2)^i \tag{4.13}$$

and

$$\beta_i=\alpha_{i+1}. \tag{4.14}$$

Thus the family of Runge-Kutta methods that gives the exact stationary density is characterized by

$$R(\Delta tQ)=\left(I-\frac{1}{2}\Delta tQ\right)^{-1}\left(I+\frac{1}{2}\Delta tQ\right) \tag{4.15}$$

and

$$R_1(\Delta tQ)=(\Delta tQ)^{-1}(R(\Delta tQ)-1). \tag{4.16}$$

Explicitly

$$r^{\mathrm{T}}=\Delta t^{-1}\begin{pmatrix}-r_{22}-\eta r_{12}\\ r_{12}\end{pmatrix}.$$

For example, the Euler method has $R=\Delta tQ$ and $r^{\mathrm{T}}=\begin{pmatrix}0&1\end{pmatrix}$. While Runge-Kutta methods with more than one stage can be constructed that have the stability function given by the above, they all have a singular tableau matrix. *The unique Runge-Kutta method,with a nonsingular tableau matrix, extended with a single Gaussian random variable per time step, that preserves the exact stationary density of the linear equation for all values of damping is thus the implicit midpoint method.* For the linear equation, the method is explicit and implemented as follows. We first generate $\hat{X}$ and $\hat{V}$:

$$\hat{V}=\alpha^{-1}\left(V_n-gX_n\frac{1}{2}\Delta t+\frac{1}{2}\epsilon\,\Delta W\right) \tag{4.17}$$

$$\hat{X}=X_n+\hat{V}\,\frac{1}{2}\Delta t, \tag{4.18}$$

where $\alpha=1+\frac{1}{2}\eta\Delta t+\frac{1}{4}g\Delta t^2$. Then

$$X_{n+1}=X_n+\hat{V}\,\Delta t \tag{4.19}$$

$$V_{n+1}=V_n-\eta\hat{V}\Delta t-g\hat{X}\,\Delta t+\epsilon\Delta W. \tag{4.20}$$

The steps (4.17)–(4.20) can be carried out in the order given.

5. Double-well systems. The excellent properties of the implicit midpoint and leapfrog methods for linear second-order SDEs prompt us to investigate its accuracy for nonlinear equations. In this Section we take as our example the double-well system (1.4), with

$$V(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4. \tag{5.1}$$

The steady-state density is given explicitly by (1.7). The statistics of the position variable is non-Gaussian, but the exact value of σ_x^2 can be evaluated to arbitrary accuracy by a numerical integration.

In general, using an implicit method on a nonlinear equation requires an iterative procedure at each timestep. However, the structure of second-order systems makes a very simple iteration possible. In the case of (1.4) we proceed as follows. Firstly, the intermediate value $\hat{X}$ is generated by repeated evaluation of

$$\hat{X} = X_n + (1 + \frac{1}{2}\eta\Delta t)^{-1}\frac{1}{2}\Delta t\left(V_n + \frac{1}{2}\Delta t f(\hat{X}) + \frac{1}{2}\epsilon\Delta W\right), \tag{5.2}$$

by fixed-point iteration with the starting value $\hat{X} = X_n$. The rest of the algorithm is explicit:

$$\begin{aligned}
\hat{V} &= V_n + \frac{1}{2}\Delta t\hat{X} \\
X_{n+1} &= X_n + \hat{V}\,\Delta t \\
V_{n+1} &= V_n - \eta\hat{V}\Delta t + f(\hat{X})\Delta t + \epsilon\Delta W.
\end{aligned}$$

In practice, only a handful of evaluations of (5.2) are necessary, so the method can be viewed as an explicit predictor-corrector approach.

In Figure 5.1 we display results obtained at $KT = 0.1$ using the Heun, leapfrog and implicit midpoint methods, as a function of η for $\Delta t = 0.1$. The Heun method is seen to be inaccurate both at large and small values of damping, while the leapfrog method is accurate for small damping but inaccurate for large damping. The implicit midpoint method is accurate (but no longer exact) for all values of damping. Not shown in this figure is Mannella's modification of the leapfrog method, that performs almost as well as the implicit midpoint method; it has a small error in σ_x^2 and σ_v^2 at all values of η.

In Figure 5.2 we display results obtained at $KT = 0.1$ using the Heun, leapfrog and implicit midpoint methods, as a function of Δt for $\eta = 1$. Both leapfrog and implicit midpoint methods maintain the independence of position and velocity for all values of Δt. The implicit midpoint method also gives the exact value of the late-time mean square of the velocity variable. All of the methods have an error in σ_x^2 proportional to Δt^2.

6. Multiplicative noise system. If the noise amplitude is a function of position then it is commonly termed "multiplicative". If the damping coefficient is multiplied by an appropriate function of position:

$$\ddot{x} = f(x) - \eta s^2(x)\dot{x} + \epsilon s(x)\xi(t), \tag{6.1}$$

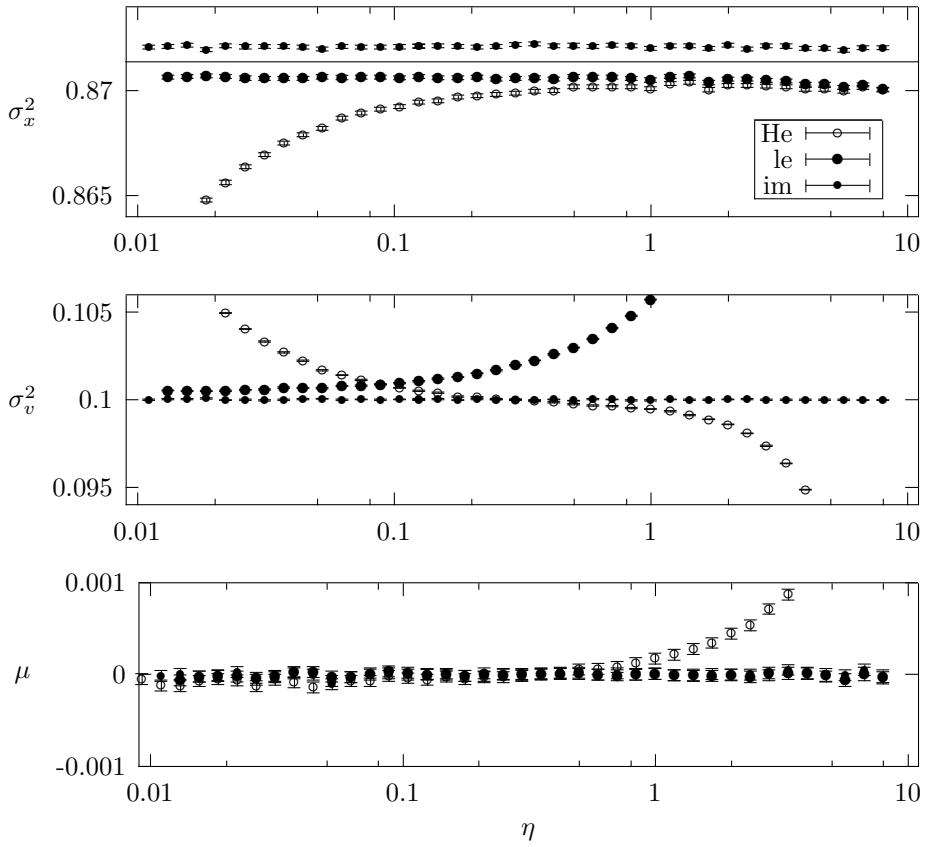

FIG. 5.1. *Mean squares and correlation versus η: double well system. Results obtained at $KT = 0.1$ using the Heun (circles), leapfrog (larger filled circles) and implicit midpoint (small filled circles) methods with $\Delta t = 0.1$. The top graph shows the late-time mean square of the position variable; the exact value is shown as a solid line. The middle graph shows the late-time mean square of the velocity variable; the exact value is* 0.1*. The lower graph shows the correlation between the position and velocity variables.*

then the stationary density (1.7) is independent of $s(x)$ [5]. In SDE notation (6.1) is a pair of equations:

$$\mathrm{d}\mathbf{X}_t = \mathbf{V}_t \mathrm{d}t$$

$$\mathrm{d}\mathbf{V}_t = -\eta s^2(\mathbf{X}_t)\mathbf{V}_t \mathrm{d}t + f(\mathbf{X}_t)\mathrm{d}t + \epsilon s(\mathbf{X}_t)\mathrm{d}\mathbf{W}_t. \tag{6.2}$$

Because the coefficient of $\mathrm{d}\mathbf{W}_t$ in the SDE for $\mathbf{V}_t$ is a function of $\mathbf{X}_t$ only, there is no difference between the Ito and Stratonovich [5] forms of (6.2). Note that, for systems of SDEs where there is a difference between Ito and Stratonovich forms, the implicit midpoint rule will converge to the Stratonovich form. However, it will only have strong order 0.5 for non-commutative SDEs with more than one noise term, in which case other approaches are needed to obtain a strong order of 1 [7].

We have performed numerical experiments with $s(x) = x$. Results for the leapfrog method are not shown because the method does not converge to the exact result for this system. In Figure 6.1 we display results obtained at $KT = 0.1$ using the Heun method, Mannella's modification of the leapfrog method and the implicit midpoint

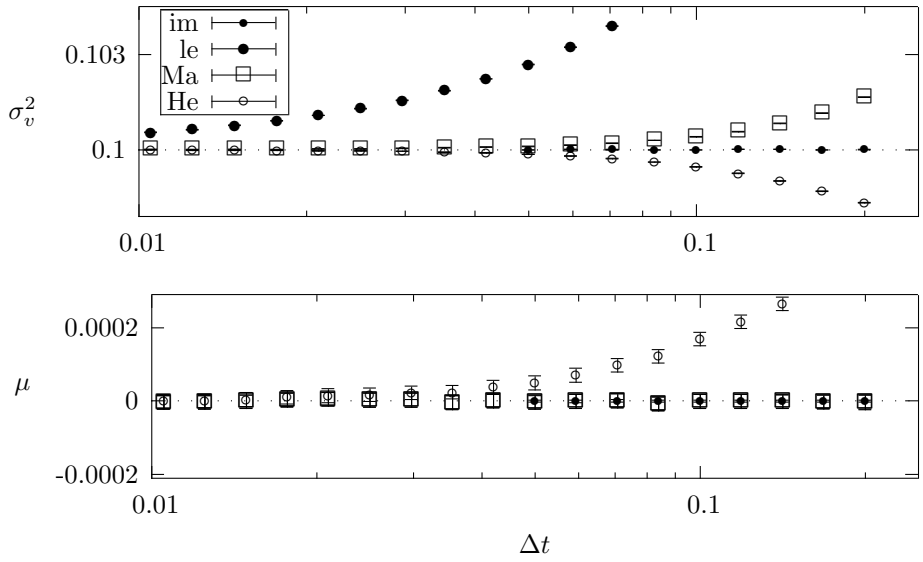

FIG. 5.2. *Mean square and correlation versus Δt: double well system. Results obtained at $KT = 0.1$ using the Heun method (circles), leapfrog method (larger filled circles) and Mannella's modification (squares), and the implicit midpoint method (small filled circles) with $\eta = 1$. The top graph shows the late-time mean square of the velocity variable; the exact value is* 0.1. *The lower graph shows the correlation between the position and velocity variables.*

method, as a function of η for $\Delta t = 0.1$. In terms of the error in σ_x^2, Mannella's method performs best at all values of η. The implicit midpoint method, however, is the only one that appears to give the exact value of σ_v^2 at all values of η. In Figure 6.2 we display numerical results with $KT = 0.1$ and $\eta = 1$, as a function of Δt. Mannella's method is the only one that has second-order convergence in σ_x^2, but the implicit midpoint method is more accurate for μ and σ_v^2.

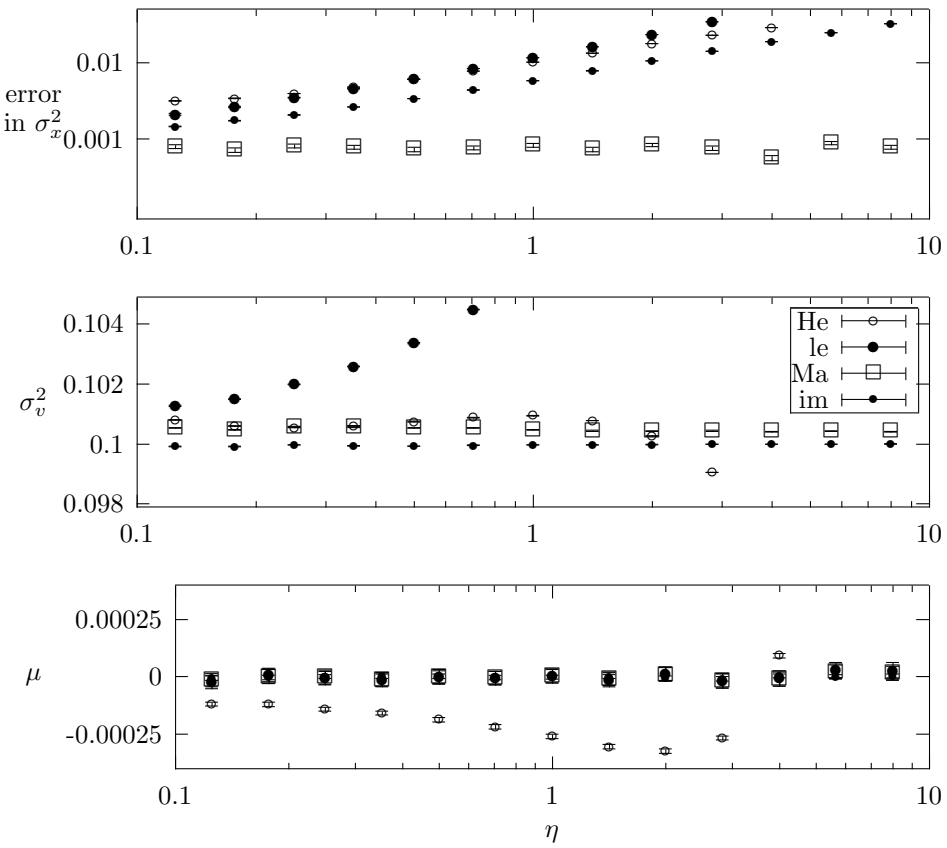

FIG. 6.1. *Mean squares and correlation versus η: double well system with multiplicative noise. Results obtained at $KT = 0.1$ and $\Delta t = 0.1$ using the Heun method (circles), the leapfrog method (larger filled circles), Mannella's modification of the leapfrog method (squares) and the implicit midpoint method (small filled circles). The top graph shows the error in the late-time mean square of the position variable. The middle graph shows the late-time mean square of the velocity variable; the exact value is* 0.1. *The lower graph shows the correlation between the position and velocity variables.*

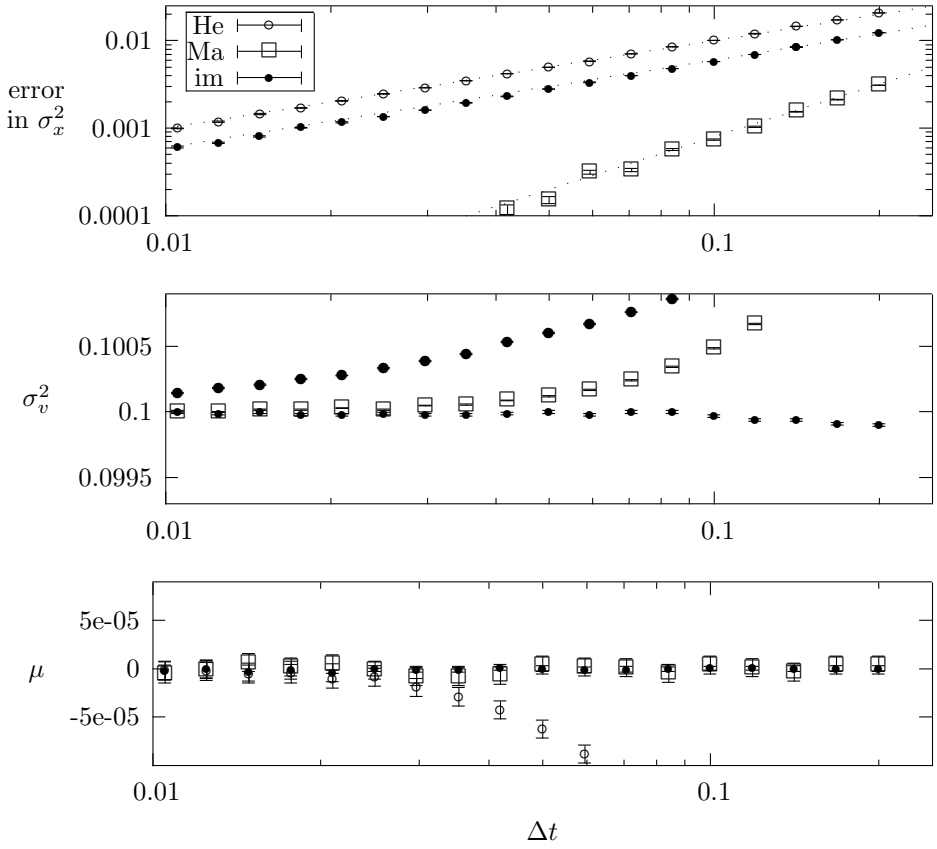

FIG. 6.2. *Mean square and correlation versus Δt: double well system with multiplicative noise. Results obtained at $KT = 0.1$ and $\eta = 1$, using the Heun method (circles), Mannella's modification of the leapfrog method (squares) and the implicit midpoint method (small filled circles). The top graph shows the error in the late-time mean square of the position variable; the dotted lines are $0.1\Delta t$, $0.06\Delta t$, and $0.08\Delta t^2$. The middle graph shows the late-time mean square of the velocity variable; the exact value is 0.1. The lower graph shows the correlation between the position and velocity variables.*

7. Discussion. Exact calculations for linear equations are of interest in their own right and because the qualitative form of the error as a function of Δt and η carries over to the nonlinear equations used as examples here. The implicit midpoint rule is the only Runge-Kutta method with a nonsingular tableau matrix that reproduces the exact stationary distribution for the linear equation. It also performs well when applied to second-order nonlinear equations and is not difficult to implement using a predictor-corrector approach. However, leapfrog type methods have the advantage of being fully explicit even for nonlinear equations. Applied to our linear equations, the simplest leapfrog method gives the exact stationary variance of the position variable and maintains the independence of position and velocity. The remaining error, in the variance of the velocity variable, is reduced under Mannella's modification of the leapfrog method. The qualitative picture is very similar for the nonlinear double-well system: the implicit midpoint method and Mannella's methods are the best, the former being superior in the error in σ_v^2 and the latter being superior in the error in σ_x^2.

In the deterministic case when solving separable Hamiltonian problems, explicit partitioned Runge-Kutta methods can be constructed that are symplectic. We aim to extend the idea in a stochastic setting to construct explicit partitioned methods that approximate the stationary correlation matrix with high-order accuracy.

REFERENCES

[1] M.P. Allen and D.J. Tildesley, *Computer simulation of Liquids* Oxford University Press, Oxford, 1997.

[2] Tony Shardlow, *Splitting for Dissipative particle dynamics* SIAM J. Sci. Comput. **24** 1267 (2003).

[3] E. Hairer, S.P. Norsett and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff problems* Springer, Berlin, 1993.

[4] Ernst Hairer, Christian Lubich and Gerhard Wanner. *Geometric numerical integration : structure-preserving algorithms for ordinary differential equations* Springer, Berlin, 2002.

[5] W. Gardiner, *Handbook of stochastic methods* Springer, Berlin, 1990.

[6] Peter E. Kloeden and Eckhard Platen, *Numerical Solution of Stochastic Differential Equations* Springer, Berlin, 1992.

[7] Kevin Burrage and Pamela Burrage, *Order conditions of stochastic Runge-Kutta methods by B-series* SIAM J. Num. Anal. **38, 5** 1626-1646 (2000).

[8] Loup Verlet, *Computer experiments on Classical Fluids. I. Thermodynamical properties of Lennard-Jones molecules* Physical Review **159** 98 (1967).

[9] W.F. van Gunsteren and H.J.C. Berendsen, *Algorithms for Brownian dynamics* Mol. Phys. **45** 637 (1982).

[10] Axel Brünger, Charles L. Brooks III and Martin Karplus, *Stochastic boundary conditions for molecular dynamics simulations of ST2 water* Chem. Phys. Lett. **105** 495 (1984).

[11] Richard W. Pastor, Bernard P. Brooks and Attila Szabo, *An analysis of the accuracy of Langevin and molecular dynamics algorithms* Mol. Phys. **65** 1409-1419 (1988).

[12] Guihua Zhang and Tamar Schlick, *Implicit discretization schemes for Langevin dynamics.* Mol. Phys. **84** 1077 (1995).

[13] Bima Mishra and Tamar Schlick, *The notion of error in Langevin dynamics. I. Linear analysis* J. Chem. Phys. **105** 299 (1996).

[14] Wei Wang, *Analysis of a few numerical integration methods for the Langevin equation* MSc. thesis, University of Illinois at Urbana-Champaign (2001).

[15] Jesus A. Izaguirre and Robert D. Skeel, *An impulse integrator for Langevin dynamics* Mol. Phys. **100** 3885-3891 (2002).

[16] Wei Wang and Robert D. Skeel, *Analysis of a few numerical integration methods for the Langevin equation* Mol. Phys. **101** 2149 (2003).

[17] Eli Hershkovitz, *A fourth-order numerical integrator for stochastic Langevin equations* J. Chem. Phys. **108** 9253 (1998).

[18] Kenneth S. Miller, *Multidimensional Gaussian Distributions* Wiley, New York, 1964.

[19] Riccardo Mannella, *Numerical tricks for studying large rare fluctuations* J. de Physique IV **8** 241 (1998).
[20] Riccardo Mannella, *Quasisymplectic integrators for stochastic differential equations* Physical Review E **69** 041107-1 (2004).
[21] Aslaug H. Strommen Melbo and Desmond J. Higham, *Numerical simulation of a linear oscillator with additive noise* Applied Numerical Mathematics **51** 89 (2004).
[22] W. Rümelin *Numerical treatment of stochastic differential equations* SIAM J. Numer. Anal. **19** 604 (1982).