**NTNU – Trondheim**
Norwegian University of
Science and Technology

# Quadratic Invariant-Preserving Runge-Kutta Methods for the Numerical Solution of Stochastic Differential Equations

Sverre Anmarkrud

Master of Science in Physics and Mathematics
Submission date: April 2015
Supervisor: Anne Kværnø, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

**Abstract**

Stochastic Runge-Kutta (SRK) and stochastic partitioned Runge-Kutta (SPRK) methods that conserve quadratic invariants when applied to stochastic ordinary differential equations are studied. Conditions that guarantee the conservation of quadratic invariants are found for the SRK and SPRK coefficients, and an existing order theory based on rooted trees is generalized to SPRK methods. Using the rooted tree theory it is shown that the majority of the order conditions of SRK and SPRK methods are dependent when the methods conserve quadratic invariants. Two new methods are constructed and their properties are verified numerically.

# Sammendrag

Stokastiske Runge-Kutta-metoder (SRK) og stokastiske partisjonerte Runge-Kutta-metoder (SPRK) som bevarer kvadratiske invarianter studeres. Det blir funnet betingelser på SRK og SPRK-koeffisientene som garanterer at kvadratiske invarianter blir bevart, og en eksisterende ordensteori basert på trær med røtter blir generalisert til SPRK-metoder. Ved å bruke ordensteorien blir det bevist at de fleste ordensbetingelsene til S(P)RK-metoder er avhengige av hverandre når metodene bevarer kvadratiske invarianter. To nye metoder blir konstruert og testet numerisk.

# Contents

# 1 Introduction

The subject of this thesis lies in the intersection of the two mathematical fields geometric numerical integration and stochastic differential equations. Geometric numerical integration is concerned with numerical methods that preserve the geometric properties of the exact flow of a differential equation. Examples of such geometric properties are reversibility, volume preservation, symplecticity[1] and invariants, though only the last two properties will make an appearance in this paper.

Informally, an invariant is a property of the system that is conserved over time. The total energy of a system and total linear and angular momentums are typical examples of invariants stemming from the laws of physics. We can also image a model of a chemical reaction where the total mass is preserved, or it may be true that all solutions lie on the surface of a sphere in 3-dimensional space. These are all examples of invariant quantities. The other geometric property, symplecticity, will be discussed at the end of chapter 2. While preserving the geometric properties of the flow can be important in its own right, it has also been shown that the structure preserving numerical methods allow for more accurate long-term integration than general purpose methods [7].

The goal has been to extend some of the theory from geometric numerical integration to stochastic differential equations (SDEs). These are differential equations that, informally, include some "randomness", or more accurately where one or more of the terms is a stochastic process. Typically, the random fluctuations are modelled by white noise which can be thought of as the derivative of Brownian motion, though other options such as jump processes are possible. Including stochastic effects allows us to create more accurate models for systems that exhibit similar random behaviour, with typical examples being the problem of option pricing from financial mathematics, the Langevin equations in statistical physics and the Kalman-Bucy filter. However, this comes at the expense of complexity. SDEs are more difficult to analyze and solve than ODEs, and in most cases analytical solutions cannot be found. Numerical methods then become valuable.

This paper is concerned with the creation of such methods, and we will proceed as follows. In chapters 2 and 3 we present material from geometric numerical integration and stochastic differential equations, respectively, and no new results are given. Rather, their purpose is to introduce the ideas that are either necessary or motivational for the rest of the paper. In chapter 2 we define invariants and introduce known, deterministic, numerical methods that conserve specific types of invariants. The numerical methods of the Runge-Kutta class are our tools of

[1] Sanz-Serna [13] uses the term *symplecticness* while Hairer, Lübich and Wanner [7] use *symplecticity*. We have chosen to go with the latter.

choice, both here and in the rest of the paper. At the end of the chapter we look briefly at Hamiltonian systems, which is a class of problems from Hamiltonian mechanics. Some of the numerical methods given in this chapter are, as we shall see, particularly suited for solving these systems, and some of the example problems at the end of this paper belong to this class.

Chapter 3 contains an overview of stochastic calculus, including the Wiener process, stochastic integrals and stochastic differential equations. We wrote a specialization project [1] covering much of the same material, and as such we have chosen to present an abridged version this time.

My own work begins in chapter 4, where we expand the concepts from chapter 2 into the stochastic setting of chapter 3. We find the necessary conditions on the stochastic Runge-Kutta (SRK) and stochastic partitioned Runge-Kutta (SPRK) coefficients that guarantee the conservation of quadratic invariants. While doing so we also find that no Itō SRK method of the general form studied in this paper can conserve quadratic invariants, and for the rest of the paper we consider Stratonovich SDEs.

Furthermore, we generalize an order theory based on rooted trees given in [6] to SPRK methods. We then study what happens to the rooted trees when we require the methods to conserve quadratic invariants, and we find that the majority of the trees lead to dependent order conditions. For a strong order 1.5 SPRK method with 1 stochastic process this reduces the number of order conditions from 140 to 10.

In chapter 5 we construct some new methods using the theory developed in chapter 4. Stochastic versions of both the 4th order Gauss method and the Störmer-Verlet methods are found, and their properties are verified numerically in chapter 6.

We note that some of the results in this paper were reached independently by Hong, Xu and Wang and published in a paper [8] during the course of this work. Primarily this concerns Theorems 4.4 and 4.5 that are stochastic extensions of theorems found in Hairer, Lubich and Wanner [7].

# 2 Deterministic theory

In this chapter we give a quick overview of relevant definitions and results from the theory of deterministic, ordinary differential equations (ODEs). For the sake of brevity we focus on material that we need in later chapters, and we do not intend to cover all the details. A complete presentation can be found in "Geometric Numerical Integration" by Hairer, Lubich and Wanner [7]. We will use results from there without further comments, although for the most relevant theorems we give a direct reference to the book (HLW).

The main purpose of this chapter is to define invariants and present some known numerical methods that conserve quadratic invariants for deterministic ODEs. When we later create new methods we will start with some of these deterministic methods and expand them to SRK methods. In fact, we will show that any SRK method that conserves quadratic invariants must be built on a deterministic RK method that conserves quadratic invariants. In the author's opinion, it therefore makes sense to start with methods that have already proven their worth, such as the Gauss and Lobatto methods we present in this chapter.

At the end of the chapter we briefly discuss Hamiltonian systems. This is a class of problems where the kind of structure preserving methods we discuss here have been used with great success, and some of the example problems at the end of this paper are stochastic Hamiltonian systems.

## 2.1 Invariants and flow

Our starting point is the system of ordinary differential equations (ODEs)

$$\dot{y} = f(y),\ y(t_0) = y_0 \tag{2.1}$$

where $y \in \mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}^d$ is a sufficiently smooth function.

The *flow* over time $t$ of (2.1) is a mapping $\phi_t$ defined by

$$\phi_t(y_0) = y(t) \quad \text{if } y(0) = y_0$$

for any point $y_0$ in phase space. The flow thus describes the evolution of the system as time passes given the initial condition, $y_0$.

**Definition 2.1.** (Invariant) *A function $I : \mathbb{R}^d \to \mathbb{R}$ is an* invariant *of (2.1) if*

$$I(y(t)) = I(y(t_0)) = Const$$

*for all solutions of (2.1).*

The next theorem follows immediately from the definition.

**Theorem 2.1.** *The function $I(y)$ is an invariant of (2.1) if and only if*

$$\nabla I(y) f(y) = 0, \quad \text{for all } y. \tag{2.2}$$

An invariant on the form $I(y) = c^\top y$ for some constant vector $c \in \mathbb{R}^d$ is called a *linear invariant* . Similarly, an invariant on the form $Q(y) = y^\top C y$ for some constant, symmetric, square matrix $C \in \mathbb{R}^{d\times d}$ is called a *quadratic invariant*. One of the main goals of this paper is to construct numerical methods that conserve quadratic invariants.

## 2.2 Runge-Kutta methods

All of the numerical methods considered in this paper belong to the class of Runge-Kutta (RK) methods or their counterparts, the stochastic Runge-Kutta (SRK) methods. Let us first recall the RK methods for the initial value problem (2.1).

An $s$-stage Runge-Kutta method for the numerical solution of (2.1) has the form

$$H_i = y_n + h\sum_{j=1}^{s} A_{ij} f(H_j), \quad i = 1, \dots, s \tag{2.3a}$$

$$y_{n+1} = y_n + h\sum_{i=1}^{s} b_i f(H_i). \tag{2.3b}$$

for a matrix $A \in \mathbb{R}^{s\times s}$, vector $b \in \mathbb{R}^s$ and a step size $h$. We assume the ODEs to be autonomous as we can always write ODEs on autonomous form by introducing new variables.

The RK methods are uniquely determined by their Runge-Kutta coefficients $A$ and $b$, and it's usually more convenient to present them as *Butcher tableaus*, as shown in Table 1 for the general method (2.3). We will use this convenient representation extensively.

All Runge-Kutta methods preserve linear invariants as for $I(y) = d^\top y$ where $d$ is a constant vector we have

$$I(y_{n+1}) = d^\top y_{n+1} = d^\top y_n + h\sum_{j=1}^{s} b_i d^\top f(H_i) = d^\top y_n = I(y_n),$$

where we have used that $d^\top f(y) = 0$ for all $y$ by Theorem 2.1.

For quadratic and higher order invariants we have the following theorems.

Table 1: Butcher tableau representation of the general $s$-stage Runge-Kutta method (2.3).

$$\begin{array}{ccc} A_{11} & \cdots & A_{1s} \\ \vdots & \vdots & \vdots \\ A_{s1} & \cdots & A_{ss} \\ \hline b_1 & \cdots & b_s \end{array}$$

**Theorem 2.2** (HLW IV Theorem 2.2)**.** *If the coefficients of a Runge-Kutta method satisfy*

$$b_i A_{ij} + b_j A_{ji} = b_i b_j, \qquad \text{for all } i, j = 1, \cdots, s,$$

*then it conserves quadratic invariants.*

**Theorem 2.3** (HLW IV Theorem 3.3)**.** *For $n \geq 3$, no Runge-Kutta method can conserve all polynomial invariants of degree $n$.*

## 2.3 Collocation methods

The conditions on the Runge-Kutta coefficients given in Theorem 2.2 is just one part of the puzzle. What further restrictions we should impose on the Runge-Kutta coefficients to create satisfactory methods is a whole subject by itself. Order of convergence is obviously important, and so is numerical stability. In this chapter we present some known methods that have good order of convergence and conserve quadratic invariants. We also briefly explain the theory behind them, but this is only tangentially related to our goals as our main interest is the methods themselves. For more details see chapter II.1 in HLW [7].

The idea behind collocation methods is to choose a polynomial of degree $s$ and require the derivative of the polynomial to equal the derivative of the solution of the differential equation at $s$ points.

**Definition 2.2.** *Let $c_1, c_2, \cdots, c_s$ be distinct real numbers (usually $0 \leq c_i \leq 1$). The* collocation polynomial *$u(t)$ is a polynomial of degree $s$ satisfying*

$$\begin{aligned} u(t_0) &= y_0, \\ \dot{u}(t_0 + c_i h) &= f(u(t_0 + c_i h)), \ i = 1, \cdots, s, \end{aligned} \tag{2.4}$$

*and the numerical solution of the* collocation method *is defined by $y_1 = u(t_0 + h)$.*

Clearly, the linear polynomial $u(t) = y_0+(t-t_0)k$ with $k = f(y_0+hc_1k)$ satisfies (2.4). Choosing, for example, $c_1 = 0$ and $c_1 = 1$ we recover the familiar explicit and implicit Euler methods. These methods can be written as Runge-Kutta methods, and the following theorem states that this is true for all collocation methods.

**Theorem 2.4** (HLW II Theorem 1.4)**.** *The collocation method of Definition 2.2 is equivalent to the s-stage Runge-Kutta method (2.3) with coefficients*

$$A_{ij} = \int_0^{c_i} \ell_j(\tau)d\tau, \quad b_i = \int_0^1 \ell_i(\tau)d\tau,$$

*where $\ell_i(\tau)$ is the Lagrange polynomial, $\ell_i(\tau) = \prod_{l\neq i}(\tau - c_l)/(c_i - c_l)$.*

Choosing $c_1, \cdots, c_s$ as the zeros of the $s$th shifted Legendre polynomial

$$\frac{d^s}{dx^s}\left(x^s(x-1)^s\right),$$

leads to the *Gauss methods*. It can be shown that these methods have order $2s$ and preserve quadratic invariants. The 4th and 6th order Gauss methods are shown in Tables 2 and 3, respectively.

Table 2: Gauss method of order 4.

$$\begin{array}{cc} \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline \frac{1}{2} & \frac{1}{2} \end{array}$$

Table 3: Gauss method of order 6.

$$\begin{array}{ccc} \frac{5}{36} & \frac{2}{9} - \frac{\sqrt{15}}{15} & \frac{5}{36} - \frac{\sqrt{15}}{30} \\ \frac{5}{36} + \frac{\sqrt{15}}{24} & \frac{2}{9} & \frac{5}{36} - \frac{\sqrt{15}}{24} \\ \frac{5}{36} + \frac{\sqrt{15}}{30} & \frac{2}{9} + \frac{\sqrt{15}}{15} & \frac{5}{36} \\ \hline \frac{5}{18} & \frac{4}{9} & \frac{5}{18} \end{array}$$

Choosing $c_1 = 0$, $c_s = 1$ and the letting the remaining nodes be the zeros of

$$\frac{d^{s-2}}{dx^{s-2}}\left(x^{s-1}(x-1)^{s-1}\right) \tag{2.5}$$

leads to the Lobatto IIIA methods. They have order $2s-2$, and together with the Lobatto IIIB methods they form a class of partitioned Runge-Kutta methods that we shall look at in later chapters. The second order Lobatto IIIA method is the trapezoidal rule. The 4th order Lobatto IIIA method is shown in Table 4.

Table 4: Lobatto IIIA method of order 4.

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

The idea behind collocation methods can be modified to include a larger class of methods.

**Definition 2.3.** *Let $c_2, \cdots, c_{s-1}$ be distinct real numbers (usually $0 \le c_i \le 1$), and let $b_1, b_s$ be two arbitrary real numbers. The corresponding* discontinuous collocation method *is then defined via a polynomial of degree $s-2$ satisfying*

$$\begin{aligned} u(t_0) &= y_0 - hb_1(\dot{u}(t_0) - f(u(t_0)) \\ \dot{u}(t_0 + c_i h) &= f(u(t_0 + c_i h)),\ i = 2, \cdots, s-1, \\ y_1 &= u(t_1) - hb_s(\dot{u}(t_1) - f(u(t_1))). \end{aligned}$$

If $b_1 = b_s = 0$ we see that this reduces to the collocation methods. Similarly to the collocation methods these can be written as Runge-Kutta methods.

**Theorem 2.5** (HLW II Theorem 1.8)**.** *The discontinuous collocation method of Definition 2.3 is equivalent to an s-stage Runge-Kutta method (2.3) with coefficients determined by $c_1 = 0$, $c_s = 1$, and*

$$\begin{aligned} A_{i1} = b_1, \quad A_{is} = 0, \quad i = 1, \cdots, s, \\ C(s-2) \text{ and } B(s-2), \end{aligned}$$

*with the conditions $C(q)$ and $B(p)$ defined as*

$$\begin{aligned} C(q): &\quad \sum_{j=1}^{s} A_{ij} c_j^{k-1} = \frac{c_i^k}{k},\ k = 1, \cdots, q, \forall i \\ B(p): &\quad \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k},\ k = 1, \cdots, p. \end{aligned}$$

Now choose $c_1 = 0$, $c_s = 1$, $b_1 \neq 0$, and $b_s \neq 0$. The class of Lobatto IIIB methods is found by choosing $c_2, \cdots, c_{s-1}$ as zeros of (2.5). Like the Lobatto IIIA methods they have order of convergence $2s - 2$. The second order Lobatto IIIA method is the trapezoidal rule, and the 4th order Lobatto IIIB method in shown in Table 5.

Table 5: Lobatto IIIB method of order 4.

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

## 2.4 Partitioned Runge-Kutta methods

Let us now consider differential equations in a partitioned form,

$$\dot{y} = f(y, z), \quad \dot{z} = g(y, z), \tag{2.6}$$

where $y$ and $z$ may be vectors, possibly of different dimensions.

Partitioned Runge-Kutta (PRK) methods are simply numerical schemes where, at each step, $y_{n+1}$ is found by one Runge-Kutta method and $z_{n+1}$ is found by a different Runge-Kutta method. This was originally proposed for problems with stiff and non-stiff parts, but with certain restrictions on the structure of the problem (2.6) they also allow the construction of explicit methods that conserve quadratic invariants on the form $Q(y, z) = y^\top Dz$, where $D$ is a constant matrix of appropriate dimension. In particular, the Hamiltonian systems introduced in the next chapter are of the form (2.6).

First let us define schemes of this type.

**Definition 2.4.** *Let $b_i, A_{ij}$ and $\hat{b}_i, \hat{A}_{ij}$ be the coefficients of two Runge-Kutta methods (2.3). A* partitioned Runge-Kutta method *for the solution of (2.6) is given by*

$$H_i = y_n + h\sum_{j=1}^{s} A_{ij} f(H_j, \hat{H}_j), \quad i = 1, \dots, s \tag{2.7a}$$

$$\hat{H}_i = z_n + h\sum_{j=1}^{s} \hat{A}_{ij} g(H_j, \hat{H}_j), \tag{2.7b}$$

$$y_{n+1} = y_n + h\sum_{i=1}^{s} b_i f(H_i, \hat{H}_i), \tag{2.7c}$$

$$z_{n+1} = z_n + h\sum_{i=1}^{s} \hat{b}_i g(H_i, \hat{H}_i). \tag{2.7d}$$

The corresponding Butcher tableau is shown in Table 6.

Table 6: Butcher tableau representation of a general PRK.

$$\begin{array}{ccc} A_{11} & \cdots & A_{1s} \\ \vdots & \vdots & \vdots \\ A_{s1} & \cdots & A_{ss} \\ \hline b_1 & \cdots & b_s \end{array} \qquad \begin{array}{ccc} \hat{A}_{11} & \cdots & \hat{A}_{1s} \\ \vdots & \vdots & \vdots \\ \hat{A}_{s1} & \cdots & \hat{A}_{ss} \\ \hline \hat{b}_1 & \cdots & \hat{b}_s \end{array}$$

**Theorem 2.6** (HLW IV Theorem 2.4)**.** *If the coefficients of a partitioned Runge-Kutta method (2.7d) satisfy*

$$b_i \hat{A}_{ij} + \hat{b}_j A_{ji} = b_i \hat{b}_j, \ \textit{for } i, j = 1, \dots, s, \tag{2.8}$$

$$b_i = \hat{b}_i, \ \textit{for } i = 1, \dots, s, \tag{2.9}$$

*then it conserves quadratic invariants of the form $Q(y, z) = y^\top D z$.*

*If the partitioned differential equation is of the special form $\dot{y} = f(z)$, $\dot{z} = g(y)$, then condition (2.8) alone implies that invariants of the form $Q(y, z) = y^\top D z$ are conserved.*

A simple example of a partitioned Runge-Kutta method that conserves quadratic invariants is the symplectic Euler method, which is simply the combination of the implicit and explicit Euler methods. The coefficients are $b_1 = 1$, $A_{11} = 1$, $\hat{b}_1 = 1$ and $\hat{A}_{11} = 0$ and the conditions in Theorem 2.6 are trivially satisfied.

The $s$-stage Lobatto IIIA and IIIB methods from the previous chapter can be combined to form *Lobatto IIIA-IIIB pairs*.

**Theorem 2.7** (HLW IV Theorem 2.3)**.** *The Lobatto IIIA-IIIB pair conserves all quadratic invariants of the form $Q(y, z) = y^\top D z$.*

**Theorem 2.8** (HLW II Theorem 2.2)**.** *The partitioned Runge-Kutta method composed of the s-stage Lobatto IIIA and the s-stage Lobatto IIIB method, is of order 2s-2.*

The pair of 2-stage methods leads to the *Störmer-Verlet* method, shown in Table 7. By the preceding theorems this is an order 2 PRK method that preserves quadratic invariants. Similarly we can construct an order 4 PRK method by combining the 3-stage Lobatto IIIA and IIIB methods, as shown in Table 8.

Table 7: Butcher tableau for the Störmer-Verlet method as a partitioned Runge-Kutta method.

$$\begin{array}{cc} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{cc} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} \end{array}$$

Table 8: Butcher tableau for the 3-stage Lobatto IIIA-IIIB pair.

$$\begin{array}{ccc} 0 & 0 & 0 \\ \frac{5}{24} & \frac{1}{3} & -\frac{1}{24} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \hline \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array} \qquad \begin{array}{ccc} \frac{1}{6} & -\frac{1}{6} & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{5}{6} & 0 \\ \hline \frac{1}{6} & \frac{2}{3} & \frac{1}{3} \end{array}$$

## 2.5 Symplecticity and Hamiltonian systems

Hamiltonian mechanics is a reformulation of classical mechanics and was first formulated by William Rowan Hamilton in 1834 [13]. It allows us to compute the dynamics of general mechanical systems by solving a system of differential equations derived from the *Hamiltonian*, which often represents the total mechanical energy of the system.

**Definition 2.5.** *Let $\Omega$ be a domain in the oriented Euclidean space $\mathbb{R}^{2d}$ of the points $(p, q) = (p_1, \cdots, p_d; q_1, \cdots, q_d)$. If $H$ is a sufficiently smooth real function*

*defined in* $\Omega$*, then the* Hamiltonian system *of differential equations with Hamiltonian* $H$ *is, by definition, given by*

$$\dot{p} = -H_q(p,q), \quad \dot{q} = H_p(p,q), \tag{2.10}$$

*where* $H_q$ *and* $H_p$ *are the vectors of partial derivatives.*

The variables $q_i$ and $p_i$ represent the generalized position coordinates and generalized momenta, respectively, for $i = 1, \ldots, d$. The integer $d$ is called the *number of degrees of freedom* and $\Omega$ is the *phase space*.

Along the solution curves of (2.10) the Hamiltonian is an invariant. This follows immediately from Theorem 2.1 and (2.10),

$$\frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial q}\right) + \frac{\partial H}{\partial q}\left(\frac{\partial H}{\partial p}\right) = 0.$$

If the Hamiltonian has the form

$$H(p,q) = T(p) + V(q)$$

it is called *separable*, and in a mechanical system $T(p)$ and $V(q)$ would represent the kinetic and potential energy, respectively. For example, the Hamiltonian of an *N-body* system is given by

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{m_i}p_i^T p_i + \sum_{i=2}^{N}\sum_{j=1}^{i-1} V_{ij}(||q_i - q_j||),$$

where $V_{ij}(r)$ are potential functions and $m_i$ is the mass of body $i$.

The Hamiltonian could also be non-autonomous (i.e. depend on time), but in this paper we restrict ourselves to the autonomous case.

We note that the Hamiltonian is typically not a quadratic invariant and thus not conserved by the methods in this paper. What is conserved, however, is the *symplecticity* of the flow of Hamiltonian systems. For $d = 1$ symplecticity of the flow means that the flow, $\phi_t$, is *area-preserving* in the sense that for a bounded subdomain $\Sigma$ for which $\phi_t(\Sigma)$ is defined, $\Sigma$ and $\phi_t(\Sigma)$ have the same oriented area. In higher dimensions we consider the sum of the oriented areas of the projections of the two-dimensional surface $\Sigma$ onto the coordinate planes $(p_i, q_i)$. Denote this sum by $m(\Sigma)$. If $m(\Sigma) = m(\phi_t(\Sigma))$ then the transformation $\phi_t$ is called symplectic.

Sanz-Serna [13] notes that the area-preserving property means that asymptotically stable equilibria and limit cycles cannot occur, the Poincaré recurrence holds, and "in fact, all properties specific to the Hamiltonian dynamics can be derived from the preservation of area property. This is no surprise because the

area-preserving character of the flow [...] actually holds *only* for Hamiltonian systems".

It can be shown [2] that for Runge-Kutta methods, the conservation of quadratic invariants implies that the method preserves the symplecticity of the Hamiltonian system.

# 3 Stochastic theory

We covered, among other things, basic probability theory, properties of Brownian motion and the construction of the Itō integral in a specialization project [1], and we have chosen not to repeat the presentation in detail. For example, we will take the existence of stochastic integrals for granted, and definitions of filtrations and martingales are assumed to be known. We will, however, define systems of stochastic differential equations and discuss the differences between Itō and Stratonovich integrals. The corresponding material in the specialization project was based on Øksendal's book "Stochastic Differential Equations" [12] which serves as the primary source for this chapter as well.

In the last section we focus on multiple stochastic integrals. They are frequently encountered when developing stochastic numerical methods, and defining a convenient notation is well worth the effort.

## 3.1 Preliminaries

We assume we are working with a complete probability space $(\Omega, \mathcal{F}, P)$ where $\Omega$ is a sample space, $\mathcal{F}$ is a $\sigma$-algebra and $P$ is a probability measure. A *stochastic process* is a parametrized collection of random variables

$$\{X(t)\}_{t\in T},$$

where in this paper the parameter space $T$ is the half-line $[0, \infty)$. For each "experiment" $\omega \in \Omega$ the function

$$t \to X(t, \omega)$$

is called a *path* of $X(t)$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Borel-measurable function and $\mu_X$ the distribution of $X$. If $\int_\Omega |f(X(\omega))| dP(\omega) < \infty$ then the number

$$\mathrm{E}[f(X)] := \int_\Omega f(X(\omega)) dP(\omega) = \int_{\mathbb{R}^n} f(x) d\mu_X(x)$$

is called the *expectation* of $f(X)$.

**Definition 3.1. (Brownian motion)** *A real-valued, scalar stochastic process $W(t)$ is called a* Brownian motion *or* Wiener process *if*

1. *$W(t) - W(s)$ is normally distributed with mean 0 and variance $t - s$ for all $t \geq s \geq 0$,*
2. *for all times $0 < t_1 < t_2 < \cdots < t_n$, the random variables $W(t_1), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ are independent.*

*If* $W(0) = 0$ *then* $W$ *is called* standard Brownian motion.

The Brownian sample paths have unbounded variation and are nowhere differentiable with probability 1 [10]. Using Kolmogorov's continuuity theorem it can be shown [12] that there exists a continuous version of Brownian motion, and we will assume that we are always working with a continuous version.

## 3.2 Itō and Stratonovich integrals

Let us first consider scalar stochastic differential equations (SDEs) on the form,

$$\begin{aligned} dX(t) &= f(t, X(t))dt + g(t, X(t))dW(t), \\ X(t_0) &= x_0, \end{aligned} \tag{3.1}$$

where $W(t)$ is a Brownian motion. The coefficients $f(t, X(t))$ and $g(t, X(t))$ are often referred to as the *drift* and *diffusion* coefficients, respectively.

Brownian motion is nowhere differentiable [10], so the differential form (3.1) has to be interpreted as an abbreviation for the stochastic integral equation

$$X(t) = X(t_0) + \int_{t_0}^{t} f(s, X(s))ds + \int_{t_0}^{t} g(s, X(s))dW(s) \tag{3.2}$$

since the differentials have no meaning on their own.

Let us first look at a specific case with $g(s, X(s)) = W(s)$,

$$\int_0^T W(s)dW(s).$$

We can approximate this integral by Riemann sums,

$$R := \sum_{k\geq 0} W(t_k^*)(W(t_{k+1}) - W(t_k))$$

where $t_k^* = t_k + \lambda(t_{k+1} - t_k)$ and $\lambda \in [0, 1]$. Then

$$\begin{aligned} \mathrm{E}\left[\int_0^T W(s)dW(s)\right] &= \mathrm{E}\left[\sum_{k\geq 0} W(t_k + \lambda(t_{k+1} - t_k))(W(t_{k+1}) - W(t_k))\right] \\ &= \sum_{k\geq 0} (t_k + \lambda(t_{k+1} - t_k) - t_k) = \lambda T, \end{aligned}$$

where we used the property $\mathrm{E}[W(t)W(s)] = \min(s, t)$.

Unlike the case for the Riemann-Stieltjes integral, the value of the stochastic integral depends on the choice of point $t_k^*$. Two choices have proven particularly useful:

1. Taking the left end point, corresponding to $\lambda = 0$, leads to the *Itō integral*, denoted by
$$\int_S^T g(t, X(t)))dW(t).$$
2. Taking the mid point ($\lambda = \frac{1}{2}$) corresponds to the *Stratonovich integral*, denoted by
$$\int_S^T g(t, X(t)) \circ dW(t).$$

Both integrals are used extensively, and the choice ultimately depends on the modelling problem. Stratonovich calculus has the advantage that the chain rule holds, while an advantage of the Itō integral is that it does not "look into the future". Specifically the Itō integral is a martingale, and also obeys certain simple formulas

$$E\left[\int_S^T g(t, X(t))dW(t)\right] = 0,$$
$$E\left[\left\|\int_S^T g(t, X(t))dW(t)\right\|^2\right] = \int_S^T E\left[||g(t, X(t))||^2\right] dt,$$

that are important in, for example, financial mathematics [3]. The latter is known as the Itō isometry.

It can be shown [9] that the solutions of a scalar Stratonovich SDE

$$dX(t) = f(t, X(t))dt + g(t, X(t)) \circ dW(t)$$

also satisfy an Itō SDE with the same diffusion coefficient, $g(t, x)$, but with a modified drift coefficient,

$$\bar{f}(t, x) = f(t, x) + \frac{1}{2}g(t, x)\frac{\partial g}{\partial x}(t, x)$$

This allows us to, given sufficiently smooth coefficient functions, work with whatever integral is most convenient.

## 3.3 Stochastic differential equations

In the previous section we looked at scalar SDEs with a single stochastic process to emphasize how the choice of midpoint leads to different stochastic integrals. We would now like to define a more general class of vector SDEs with multiple stochastic processes, and we first introduce a class of functions that serve as coefficient functions in this definition. As mentioned at the beginning of the chapter we assume concepts such as martingales and filtrations are known, alternatively we refer to Øksendal [12].

**Definition 3.2.** *Let $\mathcal{W}_{\mathcal{H}}(S,T)$ be the class of processes $f(t,\omega) \in \mathbb{R}$ satisfying*

$$f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}$$

*such that*

1. *$(t,\omega) \to f(t,\omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable, were $\mathcal{B}$ denotes the Borel $\sigma$-algebra on $[0,\infty)$.*
2. *There exists an increasing family of $\sigma$-algebras $\mathcal{H}_t; t \geq 0$ such that*
   a) *$W(t)$ is a martingale with respect to $\mathcal{H}_t$ and*
   b) *$f_t$ is $\mathcal{H}_t$-adapted*
3. *$P\left[\int_S^T f(s,\omega)^2 ds < \infty\right] = 1.$*

This choice of coefficient functions allows us to define the general class of multi-dimensional Itō processes with multiple stochastic processes we will use in the rest of the paper.

**Definition 3.3.** *Let $W(t) = (W_1(t), \ldots, W_m(t))$ be an $m$-dimensional Wiener process and for each process $g_{ij}(t,\omega) \in \mathcal{W}_{\mathcal{H}}$ for all $i = 1, \ldots, d$ and $j = 1, \ldots, m$ we assume*

$$P\left[\int_0^t g_{ij}(s,\omega)^2 ds < \infty \text{ for all } t \geq 0\right] = 1.$$

*We also assume that $f_i(s,\omega)$ is $\mathcal{H}_t$-adapted for all $i = 1, \ldots, d$ and*

$$P\left[\int_0^t |f_i(s,\omega)| ds < \infty \text{ for all } t \geq 0\right] = 1.$$

*Then the following is a $d$-*dimensional Itō process*,*

$$dX = f dt + g dW \tag{3.3}$$

*where*

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_d \end{pmatrix}, \quad g = \begin{pmatrix} g_{11} \ldots g_{1m} \\ \vdots \\ g_{d1} \ldots g_{dm} \end{pmatrix}, \quad dW = \begin{pmatrix} dW_1 \\ \vdots \\ dW_m \end{pmatrix}.$$

The definition also applies to Stratonovich integrals. It can be shown [9] that the Stratonovich SDE corresponding to the Itō SDE (3.3) is

$$dX = \underline{f}dt + g \circ dW \tag{3.4}$$

where the modified drift coefficient is defined componentwise by

$$\underline{f_i}(t,x) = f_i(t,x) - \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{m} g_{j,k}(t,x)\frac{\partial g_{j,k}}{\partial x_j}(t,x). \tag{3.5}$$

The next theorem is called the *Itō formula* or *Itō's Lemma* and is a key result in stochastic calculus, serving the role of the chain rule for Itō processes.

**Theorem 3.1** (Øksendal Theorem 4.2.1)**.** *Let*

$$dX = fdt + gdW$$

*be a d-dimensional Itō process as above. Let* $h(t,x) = (h_1(t,x), \ldots, h_p(t,x))$ *be a* $C^2$ *map from* $[0,\infty) \times \mathbb{R}^d$ *into* $\mathbb{R}^p$*. Then the process*

$$Y = h(t,X)$$

*is again an Itō process, whose component number k,* $Y_k$*, is given by*

$$dY_k = \frac{\partial h_k}{\partial t}(t,X)dt + \sum_i \frac{\partial h_k}{\partial x_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 h_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j$$

*where* $dW_i dW_j = \delta_{ij}dt$*,* $dW_i dt = dt dW_i = 0$*.*

When doing numerical calculations it is important to know that an exact solutions exists. We conclude this section with the following existence and uniqueness result.

**Theorem 3.2** (Øksendal Theorem 5.2.1)**. (Existence and uniqueness theorem for stochastic differential equations)**
*Let* $T > 0$ *and* $f(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n, g(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ *be measurable functions satisfying*

$$|f(t,x)| + |g(t,x)| \leq C(1+|x|);\ x \in \mathbb{R}^n, t \in [0,T]$$

*for some constant C, (where* $|g|^2 = \sum |g_{ij}|^2$*) and such that*

$$|f(t,x) - f(t,y)| + |g(t,x) - g(t,y)| \leq D|x-y|;\ x,y \in \mathbb{R}^n, t \in [0,T]$$

*for some constant D. Let Z be a random variable which is independent of the σ-algebra* $\mathcal{F}_\infty^{(m)}$ *generated by* $W(s,\cdot)$, $s \geq 0$ *and such that*

$$\mathrm{E}[|Z|^2] < \infty.$$

*Then the stochastic differential equation*

$$dX(t) = f(t, X(t))dt + g(t, X(t))dW(t),\ 0 \leq t \leq T, X(0) = Z$$

*has a unique t-continuous solution* $X(t,\omega)$ *with the property that* $X(t,\omega)$ *is adapted to the filtration* $\mathcal{F}_t^Z$ *generated by* $Z$ *and* $W(s,\cdot); s \leq t$ *and*

$$\mathrm{E}\left[\int_0^T |X(t)|^2 dt\right] < \infty. \tag{3.6}$$

## 3.4 Multiple stochastic integrals

To simplify the presentation of multiple stochastic integrals in subsequent chapters we will borrow a convenient notation from Kloeden and Platen (KP) [9].

We define a *multi-index* as a row vector

$$\alpha = (j_1, \cdots, j_l)$$

where

$$j_i \in \{0, 1, \cdots, m\}$$

for $i \in \{1, 2, \cdots, l\}$ and with *length* $l := l(\alpha) \in \{1, 2, \cdots\}$. Denote by $\mathcal{M}$ the set of all such multi-indices, i.e.

$$\mathcal{M} = \{(j_1, \cdots, j_l) : j_i \in \{0, 1, \cdots, m\}, i \in \{1, \cdots, l\}, \text{ for } l = 1, 2, \cdots\} \cup \{v\}.$$

Here we have denoted the multi-index of length zero by $v$.

For $\alpha \in \mathcal{M}$ with $l(\alpha) \geq 1$ we define $\alpha-$ as the multi-index obtained by removing the last component of $\alpha$. We also introduce the notation that $dW_0 = ds$. Now we define *multiple Itō integrals* $I_\alpha$ recursively by

$$I_\alpha := \begin{cases} 1 & \text{for } l = 0, \\ \int_0^t I_{\alpha-}\, dW_{j_l}(t) & \text{for } l \geq 1. \end{cases}$$

A very useful relation for our purposes exists between integrals of this type.

**Proposition 3.1** (KP Proposition 5.2.3)**.** *Let* $j_1, \ldots, j_l \in \{0, 1, \ldots, m\}$ *and* $\alpha = (j_1, \ldots, j_l) \in \mathcal{M}$ *where* $l = 1, 2, 3, \ldots$. *Denote by* $\mathcal{I}_A$ *the indicator function of the set* $A$. *Then*

$$W_j(t)I_\alpha = \sum_{i=0}^{l} I_{(j_1,\ldots,j_i,j,j_{i+1},\ldots,j_l)} + \sum_{i=1}^{l} \mathcal{I}_{\{j_i=j\neq 0\}} I_{(j_1,\ldots,j_{i-1},0,j_{j+1},\ldots,j_l)}. \tag{3.7}$$

**Corollary 3.1.** *Suppose that* $\alpha = (j_1, \cdots, j_l)$ *with* $j_1 = \cdots = j_l = j \in \{0, \cdots, m\}$ *where* $l \geq 2$. *Then for* $t \geq 0$

$$I_\alpha = \begin{cases} \frac{1}{l!}t^l & \text{for } j = 0, \\ \frac{1}{l}\left(W_j(t)I_{\alpha-} - tI_{(\alpha-)-}\right) & \text{for } j \geq 1. \end{cases} \tag{3.8}$$

For *multiple Stratonovich integrals* we have a similar definition,

$$J_\alpha := \begin{cases} 1 & \text{for } l = 0, \\ \int_0^t J_{\alpha-} \circ dW_{j_l}(t) & \text{for } l \geq 1. \end{cases}$$

The fact that the chain rule holds in Stratonovich calculus leads to simpler relationships between the Stratonovich integrals than the corresponding results for Itō integrals.

**Proposition 3.2** (KP Proposition 5.2.10)**.** *Let* $j_1, \ldots, j_l \in \{0, 1, \ldots, m\}$ *and* $\alpha = (j_1, \ldots, j_l) \in \mathcal{M}$ *where* $l = 1, 2, 3, \ldots$. *Then*

$$W_j(t)J_\alpha = \sum_{i=0}^{l} J_{(j_1,\ldots,j_i,j,j_{i+1},\ldots,j_l)}. \tag{3.9}$$

**Corollary 3.2.** *Suppose that* $\alpha = (j_1, \cdots, j_l)$ *with* $j_1 = \ldots = j_l = j \in \{0, \ldots, m\}$ *where* $l \geq 0$. *Then for* $t \geq 0$

$$J_\alpha = \frac{1}{l!}(J_{(j)})^l. \tag{3.10}$$

# 4 Numerical methods

In this chapter we will consider numerical methods for the solution of $d$-dimensional stochastic differential equations with $m$ stochastic processes,

$$X(t) = x_0 + \sum_{l=0}^{m} \int_{t_0}^{t} g_l(X(s)) \star dW_l(s), \tag{4.1}$$

where $g_l : \mathbb{R}^d \to \mathbb{R}^d$. This is a slight reformulation of (3.3). We use the notation $\star dW_0(s) = ds$ and for $l \geq 0$ we interpret $\star dW_l(s)$ in either the Itō, $\star dW_l(s) = dW_l(s)$, or Stratonovich, $\star dW_l(s) = \circ dW_l(s)$, sense. Many of the following results do not depend on the choice of stochastic integral.

The main theoretical result of this chapter is the generalization of the theory of stochastic B-series in Debrabant and Kværnø [6] to stochastic, partitioned Runge-Kutta (SPRK) methods. We also develop conditions for conservation of quadratic invariants for both stochastic Runge-Kutta (SRK) methods and SPRK methods, and show how the rooted trees in the order theory become "non-rooted" when quadratic invariants are conserved. This greatly reduces the number of independent order conditions. However, requiring the conservation of quadratic invariants is in itself a strong requirement.

## 4.1 Assumptions and notation

In the following we use the discretization $I^h = t_0, t_1, \ldots, t_N$ with $t_0 < t_1 < \cdots < t_N = T$ on the interval $I = [t_0, T]$. We will always assume a uniform step size that we denote by $h = t_{n+1} - t_n$ for $n = 0, 1, \ldots, N-1$. For our numerical methods we write $Y(t_n) = Y_n$.

Furthermore, we shall assume that the coefficients $g_l$ satisfy the conditions of the Existence and Uniqueness Theorem 3.2 for Itō SDEs. For Stratonovich SDEs we must additionally require that the vectors $g_l' g_l$ satisfy a Lipschitz and a linear growth condition so the theorem can be applied with the modified drift coefficient (3.5).

## 4.2 Convergence

When constructing a numerical method it is useful to have some measure of efficiency of the method. For deterministic methods this is done through the order of convergence, but in the stochastic setting there are several types of convergence that make sense. To create an efficient method we must first ask ourselves what we want our method to approximate, which in turn depends on the problem we wish to model. In the literature there are two major types of convergence that dominate.

*Strong* convergence where the individual sample paths are approximated, and *weak* convergence where only the corresponding distributions are approximated. It is easier to implement numerical methods constructed with respect to weak convergence, and much computation time can be saved if we realize that our problem does not need pathwise approximation.

However, in this paper we will only concern ourselves with strong convergence and the development of strong methods. Let us first give a precise definition [6].

**Definition 4.1.** *A time discrete approximation $Y = (Y_t)_{t\in I^h}$ converges strongly, respectively, in the mean square with order $p$ to $X$ as $h \to 0$ at time $t \in I^h$ if there exists a constant $C$ and a finite $\delta_0 > 0$ such that*

$$\mathrm{E}||Y(t) - X(t)|| \leq Ch^p, \text{ respectively, } \sqrt{\mathrm{E}||Y(t) - X(t)||^2} \leq Ch^p$$

*holds for each $h \in ]0, \delta_0[$.*

The order $p$ can be fractional since the root mean square order of the Wiener process is $h^{\frac{1}{2}}$. While we are interested in strong convergence, it is often easier to work with mean square convergence. This is unproblematic since by Jensen's inequality we have

$$(\mathrm{E}||Y(t) - X(t)||)^2 \leq \mathrm{E}[||Y(t) - X(t)||^2],$$

which shows that mean square convergence implies strong convergence of the same order. We also see that if $X(t)$ and $Y(t)$ are deterministic this reduces to the familiar, deterministic definition of convergence.

Definition 4.1 concerns global order, i.e. order at the end of the simulation after a number of steps. When analyzing numerical methods it is usually more convenient to find the error of a single step. Let $le^m(h;t,x)$ and $le^{ms}(h;t,x)$ be the mean and mean square local error of the method starting at the point $(t,x)$ with respect to the step size $h$,

$$le^m(h;t,x) = \mathrm{E}(Y(t+h) - X(t+h)|Y(t) = X(t) = x), \tag{4.2}$$

$$le^{ms}(h;t,x) = \sqrt{\mathrm{E}((Y(t+h) - X(t+h))^2|Y(t) = X(t) = x)}. \tag{4.3}$$

The following theorem from Milstein [11] relates local order to global order of accuracy.

**Theorem 4.1.** *Suppose the one-step approximation $Y(t+h)$ has order of accuracy $p_1$ for the mathematical expectation of the deviation and order of accuracy $p_2$ for the mean-square deviation; more precisely, for arbitrary $t_0 \leq t \leq t_0 + T - h$, $x \in \mathbb{R}^n$ the following inequalities hold:*

$$|le^m(h;t,x)| \leq K(1 + |x|^2)^{1/2}h^{p_1},$$
$$le^{ms}(h;t,x) \leq K(1 + |x|^2)^{1/2}h^{p_2}$$

*Also, let*

$$p_2 \geq \frac{1}{2}, \quad p_1 \geq p_2 + \frac{1}{2}.$$

*Then for any $N$ and $k = 0, 1, \ldots, N$ the following inequality holds:*

$$\left[\mathrm{E}|Y(t_k) - X(t_k)|^2\right]^{1/2} \leq K(1 + |X_0|^2)^{1/2} h^{p_2 - 1/2},$$

*i.e. the order of accuracy of the method constructed using the one-step approximation $Y(t+h)$ is $p = p_2 - \frac{1}{2}$.*

## 4.3 Invariants for SDEs

We note that this and all subsequent chapters constitute our own work. We begin by expanding some of the theory in chapter 2 to the stochastic setting of chapter 3, and the following is an extension of Theorem 2.1 to SDEs.

**Theorem 4.2.** *The function $I(y)$ is an invariant of (4.1) if and only if*

$$\nabla I(y) \cdot g_l(y) = 0, \quad l = 0, 1, \ldots, m, \ \forall y \in \mathbb{R}^d \tag{4.4}$$

*Proof.* We use Definition 2.1 of invariants and take the derivative with respect to $t$,

$$\begin{aligned} &Const. = I(x(t)) \\ \implies &0 = \frac{dI(x(t))}{dt} = \nabla I(x(t)) \frac{dx}{dt} \\ \implies &0 = \nabla I(x(t)) dx \end{aligned}$$

We then insert the differential form of the SDE (4.1) to obtain

$$0 = \nabla I(X(t)) dX(t) = \nabla I(X(t)) \left( \sum_{l=0}^{m} g_l(X(t)) \star dW_l(t) \right).$$

Because $W_i$ and $W_j$ are independent for all $i, j = 0, \cdots, m$ and $i \neq j$ we must have

$$\nabla I(X(t)) g_l(X(t)) = 0, \quad l = 0, 1, \ldots, m.$$

Now assume $\nabla I(X(t)) g_l(X(t)) \neq 0$ for one or more $l$. Again using the independence of the $W_i$'s there must be a $t$ such that

$$0 \neq \nabla I(X(t)) \left( \sum_{l=0}^{m} g_l(X(t)) \star dW_l(t) \right) = \nabla I(X(t)) dX(t).$$

and we have $I(X(t)) \neq Const.$ which contradicts the assumption that $I(X(t))$ is an invariant. □

**Corollary 4.1.** *If $I(y)$ is an invariant of the ODE*

$$dy = f(y)dt$$

*then $I(y)$ is an invariant of the SDE*

$$dX(t) = f(X(t))dt + \alpha f(X(t)) \star dW(t)$$

*for $\alpha \in \mathbb{R}$.*

## 4.4 Stochastic Runge-Kutta methods

The class of Stochastic Runge-Kutta (SRK) methods is a fairly straightforward generalization of Runge-Kutta methods. The main principle remains the same. A number of intermediate steps are calculated and the next step is a weighted average of the intermediate steps.

We note that, as in the deterministic case, we can always write SDEs on autonomous form by introducing new variables. Thus we will limit ourselves to the autonomous case, but the results also apply to non-autonomous SDEs.

For the solution of SDEs on the form (4.1) we define $s$-stage SRK methods

$$Y_{n+1} = Y_n + \sum_{l=0}^{m} \sum_{j=1}^{s} \gamma_j^{(l)} g_l(H_j), \tag{4.5a}$$

for $n = 0, 1, \ldots, N-1$, $t_n \in I^h$, initial condition $Y_0 = x_0$ and with stage values

$$H_i = Y_n + \sum_{l=0}^{m} \sum_{j=1}^{s} Z_{ij}^{(l)} g_l(H_j). \tag{4.5b}$$

For $m = 0$ this is almost the definition of a deterministic Runge-Kutta method (2.3), except the matrix $Z^{(0)}$ and vector $\gamma^{(0)}$ are now functions of $h$. For $m > 0$ we must include stochastic processes to approximate the stochastic integrals, which is the motivation behind the following definition from [6].

Denote by $\Xi$ a set of families of measurable mappings,

$$\Xi := \{\{\phi(h)\}_{h\geq 0} : \phi(h) : \Omega \to \mathbb{R} \text{ is } \mathcal{A} - \mathcal{B}\text{-measurable } \forall h \geq 0\}, \tag{4.6}$$

and by $\Xi_0$ the subset of $\Xi$ defined by

$$\Xi_0 := \{\{\phi(h)\}_{h\geq 0} \in \Xi : \phi(0) \equiv 0\}.$$

We then have

$$\gamma^{(l)} \in \Xi_0^s, \quad Z^{(l)} \in \Xi_0^{s\times s}$$

for $l = 0, 1, \dots, m$.

It is sometimes more convenient to write (4.5) on matrix form

$$Y_{n+1} = Y_n + \sum_{l=0}^{m} \left( {\gamma^{(l)}}^\top \otimes I_d \right) g_l(H),$$

with stage values

$$H = \mathbb{1}_s \otimes Y_n + \sum_{r=0}^{m} \left( Z^{(r)} \otimes I_d \right) g_r(H),$$

where $\mathbb{1}_s = (1, \dots, 1)^\top \in \mathbb{R}^s$ and

$$g_l(H) = \left( g_l(H_1)^\top, \dots, g_l(H_s)^\top \right)^\top,$$

for $l = 0, \dots, m$.

The Butcher tableau is extended in an obvious way, as shown in Table 9. Of course, the RK coefficients are now functions of the step size and stochastic integrals, but for simple methods where the meaning is obvious we will leave out the step size and stochastic integrals when writing the Butcher tableau.

Table 9: Butcher tableau representation of a general SRK (4.5) with $s$ stages and $m$ stochastic processes.

| | | | | | | |
|---|---|---|---|---|---|---|
| $Z_{11}^{(0)}$ | $\cdots$ | $Z_{1s}^{(0)}$ | | $\hat{Z}_{11}^{(m)}$ | $\cdots$ | $\hat{Z}_{1s}^{(m)}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\cdots$ | $\vdots$ | $\cdots$ | $\vdots$ |
| $Z_{s1}^{(0)}$ | $\cdots$ | $Z_{ss}^{(0)}$ | | $\hat{Z}_{s1}^{(m)}$ | $\cdots$ | $\hat{Z}_{ss}^{(m)}$ |
| $\gamma_1^{(0)}$ | $\cdots$ | $\gamma_s^{(0)}$ | $\cdots$ | $\hat{\gamma}_1^{(m)}$ | $\cdots$ | $\hat{\gamma}_s^{(m)}$ |

Next we will find conditions of the SRK coefficients that guarantee the conservation of linear and quadratic invariants. The first result is trivial.

**Theorem 4.3.** *All stochastic Runge-Kutta methods preserve linear invariants.*

*Proof.* Let $I(y) = d^\top y$ where $d$ is a constant vector be an invariant. We then have

$$I(Y_{n+1}) = d^\top Y_{n+1} = d^\top Y_n + \sum_{l=0}^{m} \sum_{i=1}^{s} \gamma_i^{(l)} d^\top g_l(H_i) = d^\top Y_n = I(Y_n),$$

where we have used that $\sum_{l=0}^{m} d^\top g_l(y) = 0$ for all $y$ by Theorem 4.2. □

Let us look at quadratic invariants, $I(y) = y^\top Cy$ where $C$ is a constant square matrix of appropriate dimension.

**Theorem 4.4.** *If a SRK-method (4.5) satisfies* $\gamma_i^{(l)} Z_{ij}^{(k)} + \gamma_j^{(k)} Z_{ji}^{(l)} = \gamma_i^{(l)} \gamma_j^{(k)}$ *for all* $i, j = 1, \dots, s$ *and* $k, l = 0, \dots, m$ *then it conserves quadratic invariants,* $I(y) = y^\top Cy$.

*Proof.* That $I(y)$ is an invariant implies that $\sum_{l=0}^m (g_l(y)^\top Cy + y^\top C g_l(y)) \star dW_l = 0$ for all $y$. By the independence of the Wiener processes this implies $g_l(y)^\top Cy + y^\top C g_l(y) = 0$ for $l = 0, 1, \dots, m$.

$$
\begin{aligned}
I(Y_{n+1}) = Y_{n+1}^\top C Y_{n+1} = Y_n^\top C Y_n &+ \sum_{j=1}^s \left[ Y_n^\top C \left( \sum_{k=0}^m \gamma_j^{(k)} g_k(H_j) \right) \right] \\
&+ \sum_{i=1}^s \left[ \left( \sum_{l=0}^m \gamma_i^{(l)} g_l(H_i) \right)^\top C Y_n \right] \\
&+ \sum_{i=1}^s \left( \sum_{l=0}^m \gamma_i^{(l)} g_l(H_i) \right)^\top C \sum_{j=1}^s \left( \sum_{k=0}^m \gamma_j^{(k)} g_k(H_j) \right)
\end{aligned}
$$

We can use (4.5b) to eliminate $Y_n$.

$$
\begin{aligned}
Y_{n+1}^\top C Y_{n+1} = Y_n^\top C Y_n &+ \sum_{j=1}^s \left[ \left( H_j - \sum_{l=0}^m \sum_{i=1}^s Z_{ji}^{(k)} g_l(H_i) \right)^\top C \left( \sum_{k=0}^m \gamma_j^{(k)} g_k(H_j) \right) \right] \\
&+ \sum_{i=1}^s \left[ \left( \sum_{l=0}^m \gamma_i^{(l)} g_l(H_i) \right)^\top C \left( H_i - \sum_{k=0}^m \sum_{j=1}^s Z_{ij}^{(k)} g_k(H_j) \right) \right] \\
&+ \sum_{i=1}^s \left( \sum_{l=0}^m \gamma_i^{(l)} g_l(H_i) \right)^\top C \sum_{j=1}^s \left( \sum_{k=0}^m \gamma_j^{(k)} g_k(H_j) \right) \\
&= Y_n^\top C Y_n + \sum_{l=0}^m \sum_{i=1}^s \left( \gamma_i^{(l)} \left( H_i^\top C g_l(H_i) + g_l(H_i)^\top C H_i \right) \right) \\
&- \sum_{i=1}^s \sum_{j=1}^s \sum_{k=0}^m \sum_{l=0}^m (\gamma_i^{(l)} Z_{ij}^{(k)} + \gamma_j^{(k)} Z_{ji}^{(l)} - \gamma_i^{(l)} \gamma_j^{(k)}) g_l(H_i)^\top C g_k(H_j)
\end{aligned}
$$

□

## 4.5 Stochastic partitioned Runge-Kutta methods

We now assume that the SDE (4.1) is partitioned into

$$dY = \sum_{l=0}^{m} g_l(Y_t, Z_t) \star dW_l, \tag{4.7a}$$

$$dZ = \sum_{l=0}^{m} \hat{g}_l(Y_t, Z_t) \star dW_l. \tag{4.7b}$$

The dimensions of $Y$ and $Z$ are not required to be equal.

Solving for $Y$ and $Z$ with different SRK methods leads to *stochastic partitioned Runge-Kutta* (SPRK) methods on the form

$$H_i = Y_n + \sum_{l=0}^{m} \sum_{j=1}^{s} Z_{ij}^{(l)} g_l(H_j, \hat{H}_j), \tag{4.8a}$$

$$\hat{H}_i = Z_n + \sum_{l=0}^{m} \sum_{j=1}^{s} \hat{Z}_{ij}^{(l)} \hat{g}_l(H_j, \hat{H}_j), \tag{4.8b}$$

$$Y_{n+1} = Y_n + \sum_{l=0}^{m} \sum_{i=1}^{s} \gamma_i^{(l)} g_l(H_j, \hat{H}_j), \tag{4.8c}$$

$$Z_{n+1} = Z_n + \sum_{l=0}^{m} \sum_{i=1}^{s} \hat{\gamma}_i^{(l)} \hat{g}_l(H_j, \hat{H}_j), \tag{4.8d}$$

for $i = 1, \ldots, s$.

We also assume that the system has a quadratic invariant $I(y, z) = y^\top D z$ for a matrix $D$ of the appropriate dimension, and we want to find conditions on the coefficients that allow (4.8) to conserve this quadratic invariant, i.e. such that $I(Y_{n+1}, Z_{n+1}) = I(Y_n, Z_n)$.

**Theorem 4.5.** *If the coefficients of the stochastic partitioned Runge-Kutta method (4.8) satisfy*

$$\hat{\gamma}_i^{(l)} = \gamma_i^{(l)} \quad \forall i = 1, \ldots, s \text{ and } \forall l = 0, 1, \ldots, m, \tag{4.9}$$

$$\gamma_i^{(l)} \hat{\gamma}_j^{(k)} = \hat{\gamma}_j^{(k)} Z_{ji}^{(l)} + \gamma_i^{(l)} \hat{Z}_{ij}^{(k)} \quad \forall i, j = 1, \ldots, s \text{ and } \forall k, l = 0, 1, \ldots, m \tag{4.10}$$

*then it preserves quadratic invariants on the form $I(y, z) = y^\top D z$.*

*If (4.7) is on the special form*

$$dY = \sum_{l=0}^{m} g_l(Z_t) \star dW_l,$$

$$dZ = \sum_{l=0}^{m} \hat{g}_l(Y_t) \star dW_l,$$

*then only condition (4.10) is required.*

*Proof.* That $I(y)$ is an invariant implies that $dI(y,z) = \sum_{l=0}^{m}(g_l(y,z)Dz + yD\hat{g}_l(y,z))\star dW_l = 0$, and by the independence of the Wiener processes we have $g_l(y,z)Dz + yD\hat{g}_l(y,z) = 0$ for $l = 0, 1, \ldots, m$ and for all $y$ and $z$. Inserting (4.8c) and (4.8d) into $I(Y_{n+1}, Z_{n+1})$ we get

$$\begin{aligned} I(Y_{n+1},Z_{n+1}) = Y_{n+1}^\top DZ_{n+1} &= Y_n^\top CZ_n + \sum_{j=1}^{s}\left[Y_n^\top D\left(\sum_{k=0}^{m}\hat{\gamma}_j^{(k)}\hat{g}_k(H_j,\hat{H}_j)\right)\right] \\ &+ \sum_{i=1}^{s}\left[\left(\sum_{l=0}^{m}\gamma_i^{(l)}g_l(H_i,\hat{H}_i)\right)^\top DZ_n\right] \\ &+ \sum_{i=1}^{s}\left(\sum_{l=0}^{m}\gamma_i^{(l)}g_l(H_i,\hat{H}_i)\right)^\top D\sum_{j=1}^{s}\left(\sum_{k=0}^{m}\hat{\gamma}_j^{(k)}\hat{g}_k(H_j,\hat{H}_j)\right) \end{aligned}$$

From (4.8a) and (4.8b) we have $Y_n = H_i - \sum_{l=0}^{m}\sum_{j=1}^{s} Z_{ij}^{(l)} g_l(H_j,\hat{H}_j)$ and a similar expression for $Z_n$, and we insert these on the right hand side,

$$\begin{aligned} Y_{n+1}^\top DZ_{n+1} &= Y_n^\top DZ_n \\ &+ \sum_{j=1}^{s}\left[\left(H_j - \sum_{l=0}^{m}\sum_{i=1}^{s} Z_{ji}^{(l)} g_l(H_i,\hat{H}_i)\right)^\top D\left(\sum_{k=0}^{m}\hat{\gamma}_j^{(k)}\hat{g}_k(H_j,\hat{H}_j)\right)\right] \\ &+ \sum_{i=1}^{s}\left[\left(\sum_{l=0}^{m}\gamma_i^{(l)}g_l(H_i,\hat{H}_i)\right)^\top D\left(\hat{H}_i - \sum_{k=0}^{m}\sum_{j=1}^{s}\hat{Z}_{ij}^{(k)}\hat{g}_k(H_j,\hat{H}_j)\right)\right] \\ &+ \sum_{i=1}^{s}\left(\sum_{l=0}^{m}\gamma_i^{(l)}g_l(H_i,\hat{H}_i)\right)^\top D\sum_{j=1}^{s}\left(\sum_{k=0}^{m}\hat{\gamma}_j^{(k)}\hat{g}_k(H_j,\hat{H}_j)\right) \\ &= Y_n^\top DZ_n + \sum_{i=1}^{s}\sum_{l=0}^{m}\left[\hat{\gamma}_i^{(l)}H_i^\top D\hat{g}_l(H_i,\hat{H}_i) + \gamma_i^{(l)}g_l(H_i,\hat{H}_i)^\top D\hat{H}_i\right] \\ &+ \sum_{i=1}^{s}\sum_{j=1}^{s}\sum_{k=0}^{m}\sum_{l=0}^{m}\left[\left(\gamma_i^{(l)}\hat{\gamma}_j^{(k)} - \hat{\gamma}_j^{(k)}Z_{ji}^{(l)} - \gamma_i^{(l)}\hat{Z}_{ij}^{(k)}\right)g_l(H_i,\hat{H}_i)^\top D\hat{g}_k(H_j,\hat{H}_j)\right] \end{aligned}$$

If $\hat{\gamma}_i^{(l)} = \gamma_i^{(l)}$ then

$$\begin{aligned} &\sum_{i=1}^{s}\sum_{l=0}^{m}\left[\hat{\gamma}_i^{(l)}H_i^\top D\hat{g}_l(H_i,\hat{H}_i) + \gamma_i^{(l)}g_l(H_i,\hat{H}_i)^\top D\hat{H}_i\right] = \\ &\sum_{i=1}^{s}\sum_{l=0}^{m}\gamma_i^{(l)}\left[H_i^\top D\hat{g}_l(H_i,\hat{H}_i) + g_l(H_i,\hat{H}_i)^\top D\hat{H}_i\right] = 0, \end{aligned}$$

where we have used that $y^\top Dg_(y,z) + \hat{g}(y,z)^\top Dz = 0$ for all $y$ and $z$ by the definition of the invariant. Condition (4.10) then follows.

Proving the special case only requires us to note that the assumption $g_l(z)Dz + yD\hat{g}_l(y) = 0$ for all $y$ and $z$ implies that $g_l(z)Dz = -yD\hat{g}_l(y) = Const.$ Because this is true for $y = 0$ and $z = 0$ we must have $g_l(z)Dz = -yD\hat{g}_l(y) = 0$, and condition (4.9) is no longer required. □

## 4.6 Stochastic B-series for partitioned Runge-Kutta methods

The theory of B-series and rooted trees developed by J.C. Butcher [5] for the numerical analysis of deterministic ODEs has been extended, by various authors [6], to stochastic Stratonovich and Itō SDEs for both weak and strong convergence. Debrabant and Kværnø [6] note that the construction of the B-series does not depend on the choice of Itō or Stratonovich integrals, or strong or weak convergence, and they give a unifying theory for the construction of stochastic B-series. The work in this section is based on their approach, and our aim is to slightly generalize the theory to more easily find order conditions for SPRK methods. We do this in a general setting of $q$ partitions, but the primary concern in this paper is SPRK methods on the form (4.8) where there are only 2 partitions.

We thus consider a system of $q$ stochastic processes with $m$ diffusion terms,

$$\begin{aligned}
X^{(1)}(h) &= x_0^{(1)} + \sum_{l=0}^{m} \int_0^h g_{1,l}(X^{(1)}(s), X^{(2)}(s), \dots, X^{(q)}(s)) \star dW_l(s), \\
X^{(2)}(h) &= x_0^{(2)} + \sum_{l=0}^{m} \int_0^h g_{2,l}(X^{(1)}(s), X^{(2)}(s), \dots, X^{(q)}(s)) \star dW_l(s), \\
&\vdots \\
X^{(q)}(h) &= x_0^{(q)} + \sum_{l=0}^{m} \int_0^h g_{q,l}(X^{(1)}(s), X^{(2)}(s), \dots, X^{(q)}(s)) \star dW_l(s).
\end{aligned} \tag{4.11}$$

Denote the dimension of $X^{(k)}$ by $d_k$. The deterministic term is represented by $l = 0$, such that $dW_0(s) = ds$, and the integral w.r.t. the Wiener process is interpreted as either an Itō integral, $\star dW_l(s) = dW_l(s)$, or a Stratonovich integral, $\star dW_l(s) = \circ dW_l(s)$. We also define the vector of initial values, $x_0 = [x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(q)}]$.

Our first goal is to find B-series representations of (4.11), and we begin by assuming $X^{(k)}(h)$ can be written as a B-series $B^{(k)}(\phi, x_0; h)$,

$$B^{(k)}(\phi, x_0; h) = \sum_{\tau \in T_k} \alpha(\tau) \cdot \phi(\tau)(h) \cdot F(\tau)(x_0),$$

where we define the mapping $\phi : T \to \Xi$ as

$$\phi(\emptyset)(h) \equiv 1, \ \phi(\tau)(0) \equiv 0, \ \forall \tau \in T \setminus \{\emptyset\}, \ k = 1, \dots, q.$$

The terms $\alpha(\tau)$ are combinatoric terms, while $\phi(\tau)(h)$ are stochastic integrals and $F(\tau)(x_0)$ are the elementary differentials. We note that $\Xi$ was defined in (4.6). If $\phi : T \to \Xi^s$ then $B(\phi, x_0; h) = [B(\phi_1, x_0; h), \ldots, B(\phi_s, x_0; h)]^\top$.

Our definitions for trees and elementary differentials are similar to those in [6], but in the case of SPRK methods it is also necessary to keep track of the individual partitions.

**Definition 4.2. (Trees)** *The set of shaped, rooted trees*

$$T = \{\emptyset\} \cup T_1 \cup T_2 \cup \cdots \cup T_q$$

*where*

$$\begin{aligned} T_1 &= \{\emptyset\} \cup T_{1,0} \cup T_{1,1} \cup \cdots \cup T_{1,m} \\ T_2 &= \{\emptyset\} \cup T_{2,0} \cup T_{2,1} \cup \cdots \cup T_{2,m} \\ &\vdots \\ T_q &= \{\emptyset\} \cup T_{q,0} \cup T_{q,1} \cup \cdots \cup T_{q,m} \end{aligned}$$

*is recursively defined as follows:*

*(a) The graph* $\bullet_{k,l} = [\emptyset]_{k,l}$ *with only one vertex of shape* $k$ *and color* $l$ *belongs to* $T_{k,l}$. *Let* $\tau = [\tau_1, \tau_2, \ldots, \tau_\kappa]_{k,l}$ *be the tree formed by joining the subtrees* $\tau_1, \tau_2, \ldots, \tau_\kappa$ *each by a single branch to a common root of shape* $k$ *and color* $l$.

*(b) If* $\tau_1, \tau_2, \ldots, \tau_\kappa \in T$, *then* $\tau = [\tau_1, \tau_2, \ldots, \tau_\kappa]_{k,l} \in T_{k,l}$.

**Definition 4.3. (Elementary differentials)** *For a tree* $\tau \in T$ *the elementary differential is a mapping* $F(\tau)$: $\mathbb{R}^{qd} \to \mathbb{R}^d$ *defined recursively by*

*(a)* $F(\emptyset)(x_0) = x_0^{(k)}$, $\emptyset \in T_k$[2],

*(b)* $F(\bullet_{k,l})(x_0) = g_{k,l}(x_0)$,

*(c) If* $\tau = [\tau_1, \tau_2, \ldots, \tau_\kappa]_{k,l} \in T_{k,l}$, *then*

$$F(\tau)(x_0) = g_{k,l}^{(\kappa)}(x_0)(F(\tau_1)(x_0), F(\tau_2)(x_0), \ldots, F(\tau_\kappa)(x_0)).$$

Fundamental for this work is the following lemma which says that if $Y^{(k)}(h)$ can be written as a B-series, then $f(Y^{(1)}(h), \ldots, f(Y^{(q)}(h)))$ can also be written as a B-series. This is a trivial extension of the lemma found in [6].

[2]The meaning of this notation is that when we take the sum of all tress in $T_k$ we define $F(\emptyset)(x_0) = x_0^{(k)}$.

**Lemma 4.1.** *If $Y^{(k)}(h) = B^{(k)}(\phi, x_0; h)$ are some B-series and $f \in C^\infty(\mathbb{R}^{d\times q}, \mathbb{R}^d)$, then $f(Y^{(1)}(h), \dots, f(Y^{(q)}(h)))$ can be written as formal series of the form*

$$f(Y^{(1)}(h), \dots, Y^{(q)}(h)) = \sum_{u\in U_f} \beta(u)\cdot\psi_\phi(u)(h)\cdot G(u)(x_0), \tag{4.12}$$

*where $U_f$ is a set of trees derived from $T$, by*

*(a)* $[\emptyset]_f \in U_f$, *and if* $\tau_1, \tau_2, \dots, \tau_\kappa \in T$, *then* $[\tau_1, \tau_2, \dots, \tau_\kappa]_f \in U_f$.

*(b)* $G([\emptyset]_f)(x_0) = f(x_0)$ *and*
$G(u = [\tau_1, \tau_2, \dots, \tau_\kappa]_f)(x_0) = f^{(\kappa)}(x_0)(F(\tau_1)(x_0), \dots, \tau_\kappa(x_0))$.

*(c)* $\beta([\emptyset]_f) = 1$ *and* $\beta(u = [\tau_1, \tau_2, \dots, \tau_\kappa]_f) = \frac{1}{r_1! r_2! \dots r_q!}\prod_{j=1}^{\kappa}\alpha(\tau_j)$, *where* $r_1, r_2, \dots, r_q$ *count equal trees among* $\tau_1, \tau_2, \dots, \tau_\kappa$.

*(d)* $\psi_\phi([\emptyset]_f)(h) \equiv 1$ *and*
$\psi_\phi(u = [\tau_1, \tau_2, \dots, \tau_\kappa]_f)(h) = \phi(\tau_1)(h) \odot \phi(\tau_2) \odot \dots \odot \phi(\tau_\kappa)(h)$.

In the theorem above we have used the notation $\odot$ for the Hadamard-product[3].

**Definition 4.4.** *The order of a tree $\tau \in T \cup U_f$ is defined by*

$$\rho(\emptyset) = 0, \quad \rho([\tau_1, \dots, \tau_\kappa]_f) = \sum_{i=1}^{\kappa}\rho(\tau_i),$$
$$\rho(\tau = [\tau_1, \dots, \tau_\kappa]_{k,l}) = \sum_{i=1}^{\kappa}\rho(\tau_i) + \begin{cases} 1 \text{ for } l = 0, \\ \frac{1}{2} \text{ otherwise} \end{cases}$$

If we apply Lemma 4.1 to the functions $g_{k,l}$ in (4.11) we get

$$g_{k,l}(X^{(1)}(h), \dots, X^{(q)}(h)) = \sum_{u\in U_{g_{k,l}}} \beta(u)\cdot\psi_\phi(u)(h)\cdot G(u)(x_0).$$

Clearly, by the definitions of trees, $T_{k,l}$, and elementary differentials, $F(\tau)(x_0)$, we can write this as

$$g_{k,l}(X^{(1)}(h), \dots, X^{(q)}(h)) = \sum_{\tau\in T_{k,l}} \alpha(\tau)\cdot\phi'_{k,l}(\tau)(h)\cdot F(\tau)(x_0), \tag{4.13}$$

if we also define $\alpha(\tau)$ as

$$\alpha(\emptyset) = 1,\ \alpha(\bullet_{k,l}) = 1,\ \alpha(\tau = [\tau_1, \dots, \tau_\kappa]_{k,l}) = \frac{1}{r_1! r_2 \dots r_q!}\prod_{j=1}^{\kappa}\alpha(\tau_j),$$

[3]The Hadamard-product is also known as the Schur-product or element-by-element product. We use the notation $\odot$ instead of the usual $\circ$ to avoid confusion with the Stratonovich integral.

where $r_1, r_2, \ldots, r_q$ count equal trees among $\tau_1, \tau_2, \ldots, \tau_\kappa$. From Lemma 4.1 we also have

$$\phi'_{k,l}(\tau)(h) = \begin{cases} 1, & \text{if } \tau = \bullet_{k,l}, \\ \prod_{i=1}^{\kappa} \phi(\tau_i)(h), & \text{if } \tau = [\tau_1, \ldots, \tau_\kappa]_{k,l} \in T_{k,l}. \end{cases}$$

We now write the exact solutions (4.11) as B-series and use (4.13) to obtain

$$\sum_{\tau \in T_k} \alpha(\tau) \cdot \phi(\tau)(h) \cdot F(\tau)(x_0) = x_0^{(k)} + \sum_{l=0}^{m} \sum_{\tau \in T_{k,l}} \alpha(\tau) \cdot \int_0^h \phi'_{k,l}(\tau)(s) \star dW_l(s) F(\tau)(x_0).$$

Comparing term by term we see that

$$\phi(\emptyset)(h) \equiv 1, \text{ and } \phi(\tau)(h) = \int_0^h \phi'_{k,l}(\tau)(s) \star dW_l(s)$$
$$\text{for } \tau \in T_{k,l},\ l = 0, 1, \ldots, m, k = 1, \ldots, q.$$

With induction on $\tau$ we have proven the following Theorem:

**Theorem 4.6.** *The exact solutions of (4.11) can be written as B-series $B_k(\phi, x_0; h)$ with*

$$\phi(\emptyset)(h) \equiv 1, \quad \phi(\bullet_{k,l})(h) = W_l(h),$$
$$\phi(\tau = [\tau_1, \tau_2, \ldots, \tau_\kappa]_{k,l})(h) = \int_0^h \prod_{j=1}^{\kappa} \phi(\tau_j(s)) \star dW_l(s),$$
$$\tau \in T_{k,l},\ k = 1, \ldots, q,\ l = 0, 1, \ldots, m.$$

We must also find a similar result for the numerical solutions. For the solution of (4.11) we use the $s$-stage SPRK method

$$H_k = \mathbb{1}_s \otimes Y_{k,n} + \sum_{r=0}^{m} \left( Z^{(k,r)} \otimes I_d \right) g_{k,r}(H_1, \ldots, H_q) \tag{4.14}$$

$$Y_{k,n+1} = Y_{k,n} + \sum_{l=0}^{m} \left( (\gamma^{(k,l)})^\top \otimes I_d \right) g_{k,l}(H_1, \ldots, H_q) \tag{4.15}$$

for $k = 1, \cdots, q$.

**Theorem 4.7.** *If the coefficients $Z^{(k,l)} \in \Xi_0^{s \times s}$ and $\gamma^{(k,l)} \in \Xi_0^s$, then the numerical solutions $Y_{k,1}$ as well as the stage values can be written in terms of B-series*

$$H_k = B_k(\Psi, x_0; h), \quad Y_{k,1} = B_k(\Phi, x_0; h)$$

*for all $k$, with*

$$\Psi(\emptyset)(h) \equiv \mathbb{1}_s, \quad \Psi(\bullet_{k,r})(h) = Z^{(k,r)}\mathbb{1}_s, \tag{4.16a}$$
$$\Psi(\tau = [\tau_1, \dots, \tau_\kappa]_{k,r})(h) = Z^{(k,r)} \left(\Psi(\tau_1)(h) \odot \cdots \odot \Psi(\tau_\kappa)(h)\right) \tag{4.16b}$$

*and*

$$\Phi(\emptyset)(h) \equiv 1, \quad \Phi(\bullet_{k,l})(h) = \gamma^{(k,l)}\mathbb{1}_s \tag{4.17a}$$
$$\Phi(\tau = [\tau_1, \dots, \tau_\kappa]_{k,l})(h) = \gamma^{(k,l)} \left(\Psi(\tau_1)(h) \odot \cdots \odot \Psi(\tau_\kappa)(h)\right) \tag{4.17b}$$

*Proof.* Write $H_k$ as B-series,

$$H_k = \sum_{\tau \in T_k} \alpha(\tau) \left(\Psi(\tau)(h) \otimes \mathbb{1}_{d_k}\right) \odot (\mathbb{1}_s \otimes F(\tau)(x_0)),$$

for $k = 1, \cdots, q$. Inserted into (4.14) and using (4.13) this gives

$$H_k = \mathbb{1}_s \otimes x_{k,0} + \sum_{r=0}^{m} \sum_{\tau \in T_{k,r}} \alpha(\tau) \left(\left(Z^{(k,r)} \cdot \Psi'_{k,r}(\tau)(h)\right) \otimes \mathbb{1}_{d_k}\right) \odot (\mathbb{1}_s \otimes F(\tau)(x_0))$$

A term by term comparison yields (4.16). The proof of (4.17) is similar. □

For the mean, (4.2), and mean square, (4.3), local error we then find by Theorems 4.6 and 4.7

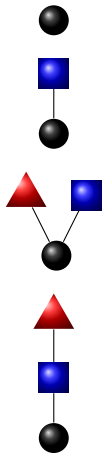
$$le^m(h; t, x) = \sum_{k=1}^{q} \sum_{\tau \in T_k} \alpha(\tau) \cdot E(\Phi(\tau)(h) - \phi(\tau)(h)) \cdot F(\tau)(x),$$
$$le^{ms}(h; t, x) = \sqrt{\mathrm{E}\left(\sum_{k=1}^{q} \sum_{\tau \in T_k} \alpha(\tau) \cdot (\Phi(\tau)(h) - \phi(\tau)(h)) \cdot F(\tau)(x)\right)^2}.$$

Finally, applying Theorem 4.1 from Milstein we find the conditions for mean square global order $p$,

$$\Phi(\tau)(h) = \phi(\tau)(h) + \mathcal{O}(h^{p+\frac{1}{2}}), \tag{4.18a}$$
$$E[\Phi(\tau)(h)] = E[\phi(\tau)(h)] + \mathcal{O}(h^{p+1}). \tag{4.18b}$$

In Table 10 we have listed the 4 different tree shapes allowed by trees with up to 3 nodes. This includes all the trees needed for methods of order 1.

Table 10: Trees, $\tau$, with of to 3 nodes. The B-series $B(\phi, x_0; h)$ and $B(\Phi, x_0; h)$ correspond to the exact and numerical solution, respectively.

| $\tau$ | $\phi$ | $\Phi$ |
|---|---|---|
|  | $\int_0^h dW_{l_1}(s)$ | $\gamma^{(k_1,l_1)}\mathbb{1}_s$ |
| | $\int_0^h \int_0^{s_1} dW_{l_2}(s_2)dW_{l_1}(s_1)$ | $\gamma^{(k_1,l_1)} Z^{(k_2,l_2)}\mathbb{1}_s$ |
| | $\int_0^h \int_0^{s_1} dW_{l_1}(s_2) \int_0^{s_1} dW_{l_2}(s_2)dW_{l_1}(s_1)$ | $\gamma^{(k_1,l_1)} \left(Z^{(k_2,l_2)}\mathbb{1}_s\right) \odot \left(Z^{(k_3,l_3)}\mathbb{1}_s\right)$ |
| | $\int_0^h \int_0^{s_1} \int_0^{s_2} dW_{l_3}(s_3)dW_{l_2}(s_2)dW_{l_1}(s_1)$ | $\gamma^{(k_1,l_1)} \left(Z^{(k_2,l_2)} Z^{(k_3,l_3)}\mathbb{1}_s\right)$ |

## 4.7 Dependent order conditions

While the order theory based on B-series makes it trivial to find the order conditions, the number of order conditions for SPRK methods becomes quite large even for relatively modest order. We want to explore which order conditions are automatically fulfilled or depend on each other when the conditions for preserving quadractic invariants in Theorem 4.5 are fulfilled.

In 1991 Sanz-Serna and Abia [14] found these dependent trees in the deterministic case, and in particular found that the order conditions can be written in terms of *non-rooted* trees. Inspired by an approach in chapter IV.7 of [7] we will prove this for SPRKs on the form (4.8), which corresponds to a system (4.11) with two partitions, $q = 2$, in the order theory. Let us first write the conditions in Theorem 4.7 in terms of the $i$'th elements of the vectors $\Phi(\tau)(h)$ and $\Psi(\tau)(h)$,

$$\begin{aligned} \Psi_i(\emptyset)(h) &\equiv 1, \quad \Psi_i(\bullet_{k,l})(h) = \sum_{j=1}^{s} Z_{ij}^{(k,l)}, \\ \Psi_i(\tau = [\tau_1, \ldots, \tau_\kappa]_{k,l})(h) &= \sum_{j=1}^{s} Z_{ij}^{(k,l)} \prod_{k=1}^{\kappa} \Psi_j(\tau_k)(h), \end{aligned} \tag{4.19}$$

and

$$\Phi_i(\emptyset)(h) \equiv 1, \quad \Phi_i(\bullet_{k,l})(h) = \sum_{i=1}^{s} \gamma_i^{(k,l)},$$
$$\Phi_i(\tau = [\tau_1, \ldots, \tau_\kappa]_{k,l})(h) = \sum_{i=1}^{s} \gamma_i^{(k,l)} \prod_{k=1}^{\kappa} \Psi_i(\tau_k)(h). \tag{4.20}$$

Furthermore, let

$$u = [u_1, \ldots, u_{\kappa_1}]_{k_1,l_1}, \quad v = [v_1, \ldots, v_{\kappa_2}]_{k_2,l_2},$$

be two trees. The *Butcher product* is defined by

$$u \circ v = [u_1, \ldots, u_{\kappa_1}, v]_{k_1,l_1}.$$

Multiply both sides of (4.10) in Theorem 4.5 by $\prod_{k_1=1}^{\kappa_1} \Psi_i(u_{k_1})(h) \prod_{k_2=1}^{\kappa_2} \Psi_j(v_{k_2})(h)$ and sum over all $i, j = 1, \ldots, s$,

$$\left(\sum_i \gamma_i^{(l)} \prod_{k_1=1}^{\kappa_1} \Psi_i(u_{k_1})(h)\right) \left(\sum_j \hat{\gamma}_j^{(k)} \prod_{k_2=1}^{\kappa_2} \Psi_j(v_{k_2})(h)\right) =$$
$$\sum_{i,j} \left(\hat{\gamma}_j^{(k)} Z_{ji}^{(l)} + \gamma_i^{(l)} \hat{Z}_{ij}^{(k)}\right) \prod_{k_1=1}^{\kappa_1} \Psi_i(u_{k_1})(h) \prod_{k_2=1}^{\kappa_2} \Psi_j(v_{k_2})(h)$$

Using (4.19) and (4.20) we get the following lemma.

**Lemma 4.2.** *If a SPRK method on the form (4.8) preserves quadratic invariants, i.e. satisfies the conditions in Theorem 4.5, then for all $u \in T_{k_1,l_1}$ and $v \in T_{k_2,l_2}$*

$$\Phi(u)(h) \cdot \Phi(v)(h) = \Phi(u \circ v)(h) + \Phi(v \circ u)(h)$$

*for all $k_1, k_2 = 1, 2$ and $l_1, l_2 = 0, 1, \ldots, m$.*

A corresponding result can be obtained for the exact solution.

**Lemma 4.3.** *For Stratonovich SDEs on the form (4.7) we have for all $u \in T_{k_1,l_1}$ and $v \in T_{k_2,l_2}$*

$$\phi(u)(h) \cdot \phi(v)(h) = \phi(u \circ v)(h) + \phi(v \circ u)(h)$$

*for all $k_1, k_2 = 1, 2$ and $l_1, l_2 = 0, 1, \ldots, m$.*

*Proof.* Write $X(t) = (X^{(1)}(t), X^{(2)}(t))$ where

$$X^{(1)}(t) = \int_0^t g_{1,l_1}(X(t)) \circ dW_{l_1}(s),$$
$$X^{(2)}(t) = \int_0^t g_{2,l_2}(X(t)) \circ dW_{l_2}(s),$$

are 1-dimensional stochastic processes. We are working with Stratonovich calculus so the chain rule applies [9],

$$\frac{d(x_1(t)x_2(t))}{dt} = x_2(t)\frac{dx_1(t)}{dt} + x_1(t)\frac{dx_2(t)}{dt}.$$

Inserting the stochastic processes $X^{(1)}(t)$ and $X^{(2)}(t)$ on differential form and integrating both sides we find

$$X^{(1)}(t)X^{(2)}(t) = X^{(1)}(0)X^{(2)}(0) + \int_0^t X^{(2)}(s)g_{1,l_1}(X(s)) \circ dW_{l_1}(s) + \int_0^t X^{(1)}(s)g_{2,l_2}(X(s)) \circ dW_{l_2}(s). \tag{4.21}$$

Let $u = [u_1, \ldots, u_{\kappa_1}]_{k_1,l_1} \in T_{k_1,l_1}$ and $v = [v_1, \ldots, v_{\kappa_2}]_{k_2,l_2} \in T_{k_2,l_2}$. Let $X_1(t) = \phi(u)(t)$ and $X_2(t) = \phi(v)(t)$ into (4.21) we obtain,

$$\begin{aligned}\phi(u)(t)\phi(v)(t) &= \int_0^t \phi(v)(s) \prod_{i=1}^{\kappa_1} \phi(u_i)(s) dW_{l_1}(s) \\ &\quad + \int_0^t \phi(u)(s) \prod_{i=1}^{\kappa_2} \phi(v_i)(s) dW_{l_2}(s) \\ &= \phi(u \circ v)(t) + \phi(v \circ u)(t),\end{aligned}$$

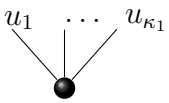
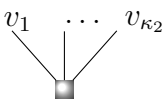
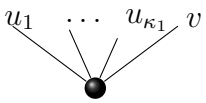
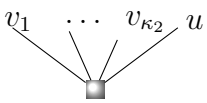
where we have used Theorem 4.6 and the fact that $\phi(\tau)(0) = 0$ by definition. □

As illustrated in Table 11, when we take the Butcher product of two arbitrary trees $u = [u_1, \ldots, u_{\kappa_1}]_{k_1,l_1}$ and $v = [v_1, \ldots, v_{\kappa_2}]_{k_2,l_2}$ we connect the root nodes of $u$ and $v$. The roots of $u \circ v$ and $v \circ u$ are the roots of $u$ and $v$, respectively. In other words, the trees $u \circ v$ and $v \circ u$ have the same structure, but differ only in the choice of root node. From Lemmas 4.2 and 4.3 we see that the order conditions obtained from the trees $u \circ v$ and $v \circ u$ are dependent.

In Table 12 we give some examples of trees that lead to dependent order conditions. Clearly, if either the order condition for $u \circ v$ or $v \circ u$ and the order conditions for the lower order trees are satisfied, then order condition for $v \circ u$ or $u \circ v$, respectively, is also satisfied.

We can turn this observation into a general result, but we will first prove the following lemma.

Table 11: Butcher products of trees $u = [u_1, \cdots, u_{\kappa_1}]_{k_1,l_1}$ and $v = [v_1, \ldots, v_{\kappa_2}]_{k_2,l_2}$.

| $u$ | $v$ | $u \circ v$ | $v \circ u$ |
|---|---|---|---|
|  |  |  |  |



**Lemma 4.4.** *If $\gamma^{(l)} = \hat{\gamma}^{(l)}$, i.e condition (4.9) in Theorem 4.5 is satisfied then two trees that only differ by the shape of the root node have equal order conditions.*

*Proof.* This follows directly from Theorems 4.6 and 4.7 by applying the condition $\gamma^{(l)} = \hat{\gamma}^{(l)}$. □

Let $|\tau|$ denote the number of nodes in the tree $\tau \in T$. The next theorem shows that when the method conserves quadratic invariants the trees become "non-rooted", and, under certain conditions, also "shapeless".

**Theorem 4.8.** *Assume the coefficients of a SPRK method on the form (4.8) satisfy the conditions in Theorem 4.5, i.e. the method conserves quadratic invariants. Let $\tau \in T$ be a tree with $|\tau| = n$ nodes for some $n \in \{1, 2, \ldots\}$, and assume the order condition (4.18) for $\tau$ is satisfied. Furthermore, assume the order conditions are satisfied for all trees with $n - 1$ or less nodes. Then the order conditions for all trees that only differ from $\tau$ by choice of the root node are also satisfied.*

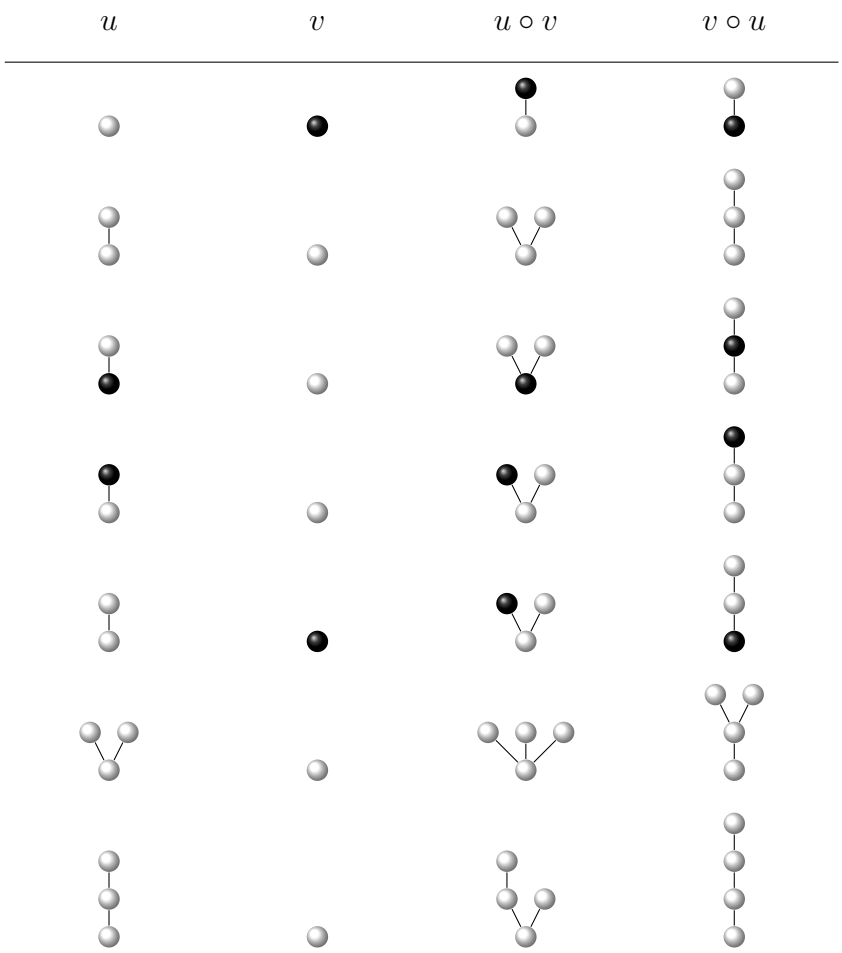
*Furthermore, if condition (4.9) in Theorem 4.5 is satisfied then the order conditions for all trees that can be obtained from $\tau$ by changing the shape of the nodes and/or choosing a different node as root node are satisfied.*

*Proof.* The proof for $|\tau| < 2$ is trivial. Let $\tau = [\tau_1, \cdots, \tau_\kappa]_{k,l}$ be an arbitrary tree with $|\tau| \geq 2$ and denote the root node by $\tau^0$. Choose subtrees $v_i = \tau_i$ and $u_i = [\tau_1, \cdots, \tau_{i-1}, \tau_{i+1}, \cdots, \tau_\kappa]_{k,l}$ for $i \in \{1, \cdots, \kappa\}$, denote the root nodes of the subtrees $\tau_i$ by $\tau_i^0$ and let $m(\tau)$ be the number of trees $\tau_i$, i.e. $m(\tau) = \kappa$. We note that $\{\tau_1^0, \cdots, \tau_\kappa^0\}$ form the set of all nodes adjacent to $\tau^0$. Furthermore, let $\tau$ have $|\tau| = n$ nodes, and assume that the order conditions for all trees with less than $n$ nodes are satisfied.

Now define the transformation $\pi_i : T \to T$ such that $\pi_i(\tau)$ is the tree obtained from $\tau$ by choosing the node $\tau_i^0$ as root. Then the order conditions obtained from $\tau = u_i \circ v_i$ and $\pi_i(\tau) = v_i \circ u_i$ are dependent by Lemmas 4.2 and 4.3, and if one of them is satisfied the others are automatically satisfied.

We have proven that all trees that can be obtained from one tree $\tau$ by choosing as root node any node adjacent to $\tau^0$ lead to dependent order conditions. Because this is true for all $\tau \in T$ it is also true for all $\pi_i(\tau)$, i.e. the order conditions

Table 12: Dependent trees.




obtained from $\pi_i(\tau)$ and $\pi_j(\pi_i(\tau))$, $j = 1, \ldots, m(\pi_i(\tau))$ are dependent. By repeated application of this argument we have proven the first part of the lemma.

To prove the second part we first note that if $\gamma^{(l)} = \hat{\gamma}^{(l)}$ then by Lemma 4.4 the order conditions of two trees that only differ by the shape of the root are the same. Let $\tau$ be a tree with $n$ nodes and let $v$ be an arbitrary tree that differs from $\tau$ by choice of the root node and/or the shape of the nodes. We will prove that if the order conditions of $\tau$ are satisfied, then so are the order conditions of $v$.

We can assume that $v$ has the same root node as $\tau$, because by assumption $v$ can be obtained from $\tau$ by choosing a different root node and/or changing the shape of the nodes, and we have already proven that choosing a different root node leads to dependent order conditions. So let $\tau$ and $v$ differ only by the shape of the

nodes. Let $S = \{n_1, n_2, \dots, n_r\}$ be the set of nodes that differ in shape from the nodes of $v$, and denote the root node of $\tau$ by $n_0$. Let $\tau^*$ be the tree obtained from $\tau$ by choosing $n_1$ as the root node. We have already proven that the order condition for $\tau^*$ is then satisfied. Furthermore, $n_1$ is now the root node, and by changing the shape of $n_1$ we obtain a new tree, $\tau^*_{n_1}$. By Lemma 4.4 the order condition for $\tau^*_{n_1}$ is automatically satisfied. By choosing the original root, $n_0$, as the root of $\tau^*_{n_1}$ we get a new tree, $\tau_{n_1}$, and the order condition of $\tau_{n_1}$ is satisfied. What we have obtained is a tree, $\tau_{n_1}$, that only differs from $\tau$ by the shape of the node $n_1$, and we have shown that if the order condition for $\tau$ is satisfied, then so is the order conditions for $\tau_{n_1}$

Repeating this process, starting from $\tau_{n_1}$, we can change the shape of all the nodes in $S$ until we obtain $v$. □

We will give an example of how Theorem 4.8 can be used. Assume we want to construct an order 1.5 SPRK method that preserves quadratic invariants and that both conditions (4.9) and (4.10) are satisfied. The method will solve SODEs of the form (4.7), which corresponds to a system with 2 partitions in the order theory (4.11). We will also assume a single stochastic process, $m = 1$, with the usual representation of black deterministic nodes and white stochastic nodes.

Strong order 1.5 requires all trees up to order 2, and we list all trees where all nodes are of shape 1 in Table 13. Every node can have one of two shapes, so for every tree in Table 13 with $n$ nodes there are $2^n$ trees for a total of 140 trees. However, by Theorem 4.8 we know that it is enough to consider only one shape. In Table 14 we list the trees that remain after applying Theorem 4.8. We see that we are left with 10 trees that cannot be obtained from each other by a change of root node. This is a dramatic reduction from the intial 140 trees.

Table 13: All trees of order 2 or less with nodes of shape 1.

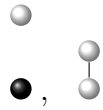
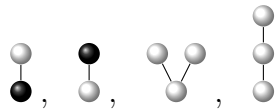
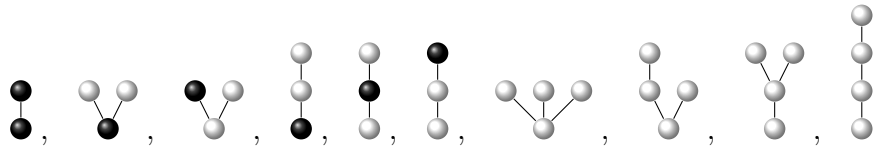
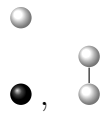
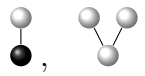
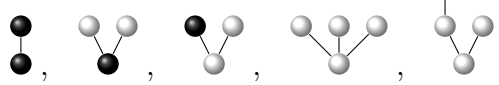
| $\rho(\tau)$ | $\tau$ |
|---|---|
| $\frac{1}{2}$ |  |
| 1 |  |
| $\frac{3}{2}$ |  |
| 2 |  |

Table 14: Non-superfluous trees of order 2 or less.

| $\rho(\tau)$ | $\tau$ |
|---|---|
| $\frac{1}{2}$ |  |
| 1 |  |
| $\frac{3}{2}$ |  |
| 2 |  |

# 5 Constructing new methods

In the previous chapter we developed the order theory and found conditions for conserving quadratic invariants. We now turn to the problem of putting it all together and creating new methods. It turns out that any SRK method, partitioned or not, that conserves quadratic invariants must be based on a deterministic RK method that does the same. This is simply a consequence of the conditions on the Runge-Kutta coefficients for the deterministic methods being a subset of the conditions for the stochastic methods. In the case of partitioned methods this follows from the conditions in Theorem 2.6 being a subset of the conditions in Theorem 4.5. For this reason we find it natural to base any new methods on known, deterministic RK methods that conserve quadratic invariants.

## 5.1 Itō stochastic Runge-Kutta methods

We look at Itō SRK methods (4.5) with two stages and a single stochastic process. The goal is to find coefficients that satisfy both the order conditions (4.18) for strong order 1 and the conditions for conserving quadratic invariants in Theorem 4.4. From the latter theorem we find the following conditions,

$$\gamma_1^{(0)} Z_{11}^{(0)} = \frac{1}{2}(\gamma_1^{(0)})^2, \tag{5.1a}$$
$$\gamma_1^{(1)} Z_{11}^{(1)} = \frac{1}{2}(\gamma_1^{(1)})^2, \tag{5.1b}$$
$$\gamma_2^{(0)} Z_{22}^{(0)} = \frac{1}{2}(\gamma_2^{(0)})^2, \tag{5.1c}$$
$$\gamma_2^{(1)} Z_{22}^{(1)} = \frac{1}{2}(\gamma_2^{(1)})^2, \tag{5.1d}$$
$$\gamma_1^{(0)} Z_{12}^{(0)} + \gamma_2^{(0)} Z_{21}^{(0)} = \gamma_1^{(0)}\gamma_2^{(0)}, \tag{5.1e}$$
$$\gamma_1^{(1)} Z_{12}^{(1)} + \gamma_2^{(1)} Z_{21}^{(1)} = \gamma_1^{(1)}\gamma_2^{(1)}, \tag{5.1f}$$
$$\gamma_1^{(0)} Z_{11}^{(1)} + \gamma_1^{(1)} Z_{11}^{(0)} = \gamma_1^{(0)}\gamma_1^{(1)}, \tag{5.1g}$$
$$\gamma_1^{(0)} Z_{12}^{(1)} + \gamma_2^{(1)} Z_{21}^{(0)} = \gamma_1^{(0)}\gamma_2^{(1)}, \tag{5.1h}$$
$$\gamma_1^{(1)} Z_{12}^{(0)} + \gamma_2^{(0)} Z_{21}^{(1)} = \gamma_1^{(1)}\gamma_2^{(0)}, \tag{5.1i}$$
$$\gamma_2^{(0)} Z_{22}^{(1)} + \gamma_2^{(1)} Z_{22}^{(0)} = \gamma_2^{(0)}\gamma_2^{(1)}. \tag{5.1j}$$

From (4.18) we find the order conditions for Itō strong order 1,

$$\gamma_1^{(0)} + \gamma_2^{(0)} = h, \tag{5.2a}$$
$$\gamma_1^{(1)} + \gamma_2^{(1)} = I_{(1)}, \tag{5.2b}$$
$$\gamma_1^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)}) + \gamma_2^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)}) = I_{(1,1)}, \tag{5.2c}$$
$$\mathrm{E}\left[\gamma_1^{(1)}(Z_{11}^{(0)} + Z_{12}^{(0)}) + \gamma_2^{(1)}(Z_{21}^{(0)} + Z_{22}^{(0)})\right] = 0, \tag{5.2d}$$
$$\mathrm{E}\left[\gamma_1^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)})^2 + \gamma_2^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)})^2\right] = 0, \tag{5.2e}$$
$$\mathrm{E}\left[\gamma_1^{(0)}(Z_{11}^{(1)} + Z_{12}^{(1)}) + \gamma_2^{(0)}(Z_{21}^{(1)} + Z_{22}^{(1)})\right] = 0, \tag{5.2f}$$
$$\begin{aligned}\mathrm{E}\Big[&\gamma_1^{(1)} Z_{11}^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)}) + \gamma_2^{(1)} Z_{21}^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)}) \\ &+\gamma_1^{(1)} Z_{12}^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)}) + \gamma_2^{(1)} Z_{22}^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)})\Big] = 0.\end{aligned} \tag{5.2g}$$

However, we find that there are no 2-stage, strong order 1 Itō SRK methods of the form (4.5) that conserve quadratic invariants. We can show this by substituting the conditions (5.1b), (5.1f) and (5.1d) into order condition (5.2c),

$$\begin{aligned}&\frac{1}{2}\left(\gamma_1^{(1)}\right)^2 + \gamma_1^{(1)}\gamma_2^{(1)} - \gamma_2^{(1)} Z_{21}^{(1)} + \gamma_2^{(1)} Z_{21}^{(1)} + \frac{1}{2}\left(\gamma_2^{(1)}\right)^2 = \\ &\frac{1}{2}(\gamma_1^{(1)} + \gamma_2^{(1)})^2 = \frac{1}{2} I_{(1)}^2 \neq I_{(1,1)}\end{aligned}$$

This result can be extended to $s$-stage methods.

**Theorem 5.1.** *There are no Itō stochastic Runge-Kutta methods on the general form (4.5) of strong order 1 or greater that preserve quadratic invariants.*

*Proof.* Order conditions corresponding to (5.2b) and (5.2c) are, respectively,

$$\sum_{i=1}^{s} \gamma_i^{(1)} = I_{(1)},$$
$$\sum_{i=1}^{s} \gamma_i^{(1)} \sum_{j=1}^{s} Z_{ij}^{(1)} = I_{(1,1)}.$$

We can write the latter condition as

$$\sum_{i=1}^{s} \gamma_i^{(1)} Z_{ii}^{(1)} + \sum_{i=2}^{s} \sum_{j=1}^{i-1} \left(\gamma_i^{(1)} Z_{ij}^{(1)} + \gamma_j^{(1)} Z_{ji}^{(1)}\right) = I_{(1,1)}$$

From Theorem 4.4 we have the relations $\gamma_i^{(1)} Z_{ij}^{(1)} = \gamma_i^{(1)}\gamma_j^{(1)} - \gamma_j^{(1)} Z_{ji}^{(1)}$ and $\gamma_i^{(1)} Z_{ii}^{(1)} = \frac{1}{2}\left(\gamma_i^{(1)}\right)^2$, and we obtain

$$\sum_{i}^{s} \left(\gamma_i^{(1)}\right)^2 + \sum_{i=2}^{s} \sum_{j=1}^{i-1} \left(\gamma_i^{(1)}\gamma_j^{(1)}\right) = \frac{1}{2}\left(\sum_{i=1}^{s} \gamma_i^{(1)}\right)^2 = \frac{1}{2} I_{(1)}^2 \neq I_{(1,1)}$$

□

However, this result suggests that there may be Stratonovich methods that conserve quadratic invariants, because in the Stratonovich case $\frac{1}{2}J_{(1)}^2 = J_{(1,1)}$, as required.

## 5.2 A Stratonovich stochastic Gauss method

From (4.18) we find that the order conditions for strong order 1 Stratonovich SRK methods are

$$\gamma_1^{(0)} + \gamma_2^{(0)} = h \tag{5.3a}$$

$$\gamma_1^{(1)} + \gamma_2^{(1)} = J_{(1)} \tag{5.3b}$$

$$\gamma_1^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)}) + \gamma_2^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)}) = J_{(1,1)} = \frac{1}{2}J_{(1)}^2 \tag{5.3c}$$

$$\mathrm{E}\left[\gamma_1^{(1)}(Z_{11}^{(0)} + Z_{12}^{(0)}) + \gamma_2^{(1)}(Z_{21}^{(0)} + Z_{22}^{(0)})\right] = 0 \tag{5.3d}$$

$$\mathrm{E}\left[\gamma_1^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)})^2 + \gamma_2^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)})^2\right] = 0 \tag{5.3e}$$

$$\mathrm{E}\left[\gamma_1^{(0)}(Z_{11}^{(1)} + Z_{12}^{(1)}) + \gamma_2^{(0)}(Z_{21}^{(1)} + Z_{22}^{(1)})\right] = 0 \tag{5.3f}$$

$$\begin{aligned}\mathrm{E}\Big[&\gamma_1^{(1)}Z_{11}^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)}) + \gamma_2^{(1)}Z_{21}^{(1)}(Z_{11}^{(1)} + Z_{12}^{(1)}) \\ &+\gamma_1^{(1)}Z_{12}^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)}) + \gamma_2^{(1)}Z_{22}^{(1)}(Z_{21}^{(1)} + Z_{22}^{(1)})\Big] = 0\end{aligned} \tag{5.3g}$$

From the proof of Theorem 4.4 we see that (5.3c) is automatically fulfilled if the conditions for conserving quadratic invariants (5.1) are satisfied.

We now turn to the problem of finding a SRK method of strong order 1 that conserves quadratic invariants. The conditions in Theorem 4.4 with $k, l = 0$ are exactly the same as those found for deterministic RK methods, and for this reason we find it convenient to use a known method from the deterministic theory as our starting point. We will choose the 4th order Gauss method introduced in Chapter 2.3 with the coefficients

$$\begin{array}{cc} \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline \frac{1}{2} & \frac{1}{2} \end{array}$$

The deterministic conditions (5.1a), (5.1c) and (5.1e) are now obviously satisfied. Next we look at the conditions that link the stochastic and deterministic

coefficients together, (5.1g)-(5.1j),

$$\begin{aligned}
\frac{1}{2}Z_{11}^{(1)} + \gamma_1^{(1)}\frac{1}{4} = \frac{1}{2}\gamma_1^{(1)} &\implies Z_{11}^{(1)} = \frac{1}{2}\gamma_1^{(1)} \\
\frac{1}{2}Z_{12}^{(1)} + \gamma_2^{(1)}\left(\frac{1}{4} + \frac{\sqrt{3}}{6}\right) = \frac{1}{2}\gamma_2^{(1)} &\implies Z_{12}^{(1)} = \left(\frac{1}{2} - \frac{1}{\sqrt{3}}\right)\gamma_2^{(1)} \\
\gamma_1^{(1)}\left(\frac{1}{4} - \frac{\sqrt{3}}{6}\right) + \frac{1}{2}Z_{21}^{(1)} = \frac{1}{4}\gamma_1^{(1)} &\implies Z_{21}^{(1)} = \left(\frac{1}{2} + \frac{1}{\sqrt{3}}\right)\gamma_1^{(1)} \\
\frac{1}{2}Z_{22}^{(1)} + \gamma_2^{(1)}\frac{1}{4} = \frac{1}{2}\gamma_2^{(1)} &\implies Z_{22}^{(1)} = \frac{1}{2}\gamma_2^{(1)}
\end{aligned}$$

With these coefficients the final 3 conditions are trivially satisfied.

When solving the order conditions we will assume that $\gamma_i^{(1)} = \beta_i J_{(1)}$ with $\beta_i \in \mathbb{R}$ for $i = 1, 2$. This means that $\mathrm{E}[\gamma_i^{(1)}] = \mathrm{E}[Z_{ij}^{(1)}] = 0$ for $i, j = 1, 2$, and conditions (5.3d) and (5.3f) are immediately satisfied. From (5.3b) we have $\beta_1 = 1 - \beta_2$, and from the proof of Theorem 5.1 we already know that (5.3c) is satisfied. Condition (5.3e) becomes

$$\mathrm{E}\left[\beta_1 J_{(1)}^3 \left(\frac{1}{2}\beta_1 + \frac{1}{2}\beta_2 - \frac{1}{\sqrt{3}}\beta_2\right)^2 + \beta_2 J_{(1)}^3 \left(\frac{1}{2}\beta_1 + \frac{1}{\sqrt{3}}\beta_1 + \frac{1}{2}\beta_2\right)^2\right] = 0,$$

and (5.3g) similarly takes the form $f(\beta_1, \beta_2)\mathrm{E}[J_{(1)}^3] = 0$. We gather the results in a theorem.

**Theorem 5.2.** *Let $\gamma^{(l)} = \beta^{(l)} W_l(h)$, $Z^{(l)} = A^{(l)} W_l(h)$ for $\beta^{(l)} \in \mathbb{R}^s$, $A^{(l)} \in \mathbb{R}^{s\times s}$ and $l = 0, 1$. The Stratonovich SRK method with coefficients*

$$\begin{array}{cc|cc}
\frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} & \frac{a}{2} & \left(\frac{1}{2} - \frac{1}{\sqrt{3}}\right)(1-a) \\
\frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} & \left(\frac{1}{2} + \frac{1}{\sqrt{3}}\right)a & \frac{1}{2}(1-a) \\
\hline
\frac{1}{2} & \frac{1}{2} & a & 1-a
\end{array}$$

*for some $a \in \mathbb{R}$ has strong order of convergence 1 and conserves quadratic invariants of the form $Q(y) = y^\top C y$.*

While $a$ can be chosen freely we will limit ourselves to $|a| \leq 1$. Whether or not large values of $|a|$ can lead to other problems, such a poor numerical stability, was not studied.

## 5.3 A stochastic Störmer-Verlet method

Using the Störmer-Verlet method as a starting point we create a 2-stage, strong order 1 SPRK method for a partitioned system of Stratonovich SDEs with a 1-dimensional Wiener process. This corresponds to $s = 2$, $q = 2$, $m = 1$ in the order theory developed in chapter 4.6. It turns out that there is only one such method, which is just a trivial extension of the original method.

**Theorem 5.3.** *Let $\gamma^{(k,l)} = \beta^{(k,l)} W_l(h)$, $Z^{(k,l)} = A^{(k,l)} W_l(h)$ for $\beta^{(k,l)} \in \mathbb{R}^s$, $A^{(k,l)} \in \mathbb{R}^{s\times s}$, $k = 1, 2$ and $l = 0, 1$. The Stratonovich SPRK method with coefficients*

$$\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{cc|cc} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}$$

*has strong order of convergence 1 and conserves quadratic invariants of the form $Q(y, z) = y^\top D z$.*

*Proof.* We will choose the Störmer-Verlet scheme, shown in Table 15 as a Butcher tableau, as a starting point.

Table 15: Butcher tableau for the Störmer-Verlet method as a partitioned Runge-Kutta method.

$$\begin{array}{cc} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{cc} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} \end{array}$$

Using Theorems 4.6 and 4.7 we can write the order conditions

$$J_{(l)} = \gamma^{(k,l)}\mathbb{1}_2, \quad l = 0, 1, \ k = 1, 2 \tag{5.4a}$$
$$J_{(1,1)} = \gamma^{(k_1,1)} Z^{(k_2,1)}\mathbb{1}_2, \quad k_1, k_2 = 1, 2 \tag{5.4b}$$
$$0 = \mathrm{E}\left[\gamma^{(k_1,0)} Z^{(k_2,1)}\mathbb{1}_2\right], \quad k_1, k_2 = 1, 2 \tag{5.4c}$$
$$0 = \mathrm{E}\left[\gamma^{(k_1,1)} Z^{(k_2,0)}\mathbb{1}_2,\right] \quad k_1, k_2 = 1, 2 \tag{5.4d}$$
$$0 = \mathrm{E}\left[\gamma^{(k_1,1)} \left(Z^{(k_2,1)}\mathbb{1}_2\right)\left(Z^{(k_3,1)}\mathbb{1}_2\right)\right], \ k_1, k_2, k_3 = 1, 2 \tag{5.4e}$$
$$0 = \mathrm{E}\left[\gamma^{(k_1,1)} \left(Z^{(k_2,1)} Z^{(k_3,1)}\mathbb{1}_2\right)\right], \ k_1, k_2, k_3 = 1, 2 \tag{5.4f}$$

We also want our method to conserve quadratic invariants, and from Theorem 4.5 we have the conditions

$$\gamma_i^{(1,l)} = \gamma_i^{(2,l)} \quad i = 1, 2 \text{ and } l = 0, 1, \tag{5.5a}$$
$$\gamma_i^{(1,l)}\gamma_j^{(2,r)} = \gamma_j^{(2,r)} Z_{ji}^{(1,l)} + \gamma_i^{(1,l)} Z_{ij}^{(2,r)}, \quad i, j = 1, 2 \text{ and } r, l = 0, 1. \tag{5.5b}$$

The purely deterministic conditions are, of course, satisfied, but we must determine the vectors $\gamma^{(k,1)}$ and matrices $Z^{(k,1)}$ for $k = 1, 2$. We will assume that $\gamma_i^{(k,1)} = b_i^{(k,1)} J_{(1)}$ where $b_i^{(k,1)} \in \mathbb{R}$ for $k = 1, 2$ and $i = 1, 2$. Similarly we assume that $Z_{ij}^{(k,1)} = A_{ij}^{(k,1)} J_{(1)}$ where $A_{ij}^{(k,1)} \in \mathbb{R}$ for $k = 1, 2$ and $i, j = 1, 2$. Order conditions (5.4c) and (5.4d) are then trivially satisfied as they become $C \cdot \mathrm{E}\left[J_{(1)}\right] = 0$ for some real number $C$. Similarly, conditions (5.4e) and (5.4f) become $C \cdot \mathrm{E}\left[J_{(1)}^3\right] = 0$ and are also trivially satisfied. From (5.5a) we have $\gamma_i^{(1,1)} = \gamma_i^{(2,1)}$ for $i = 1, 2$, and from (5.4a) we get $\gamma_2^{(1,1)} = J_{(1)} - \gamma_1^{(1,1)}$.

From (5.5b) we get 12 conditions when we ignore the 4 purely deterministic conditions. After using $\gamma^{(2,1)} = \gamma^{(1,1)}$, inserting the coefficients from the Störmer-

Verlet scheme and simplifying we get

$$\begin{aligned}
0 &= Z_{11}^{(1,1)},\\
\gamma_1^{(1,1)} &= Z_{11}^{(2,1)},\\
\gamma_1^{(1,1)} &= Z_{21}^{(1,1)},\\
0 &= Z_{12}^{(2,1)},\\
0 &= Z_{12}^{(1,1)},\\
\gamma_1^{(1,1)} &= Z_{21}^{(2,1)},\\
\gamma_2^{(1,1)} &= Z_{22}^{(1,1)},\\
0 &= Z_{22}^{(2,1)},\\
\gamma_1^{(1,1)}\gamma_1^{(1,1)} &= \gamma_1^{(1,1)} Z_{11}^{(1,1)} + \gamma_1^{(1,1)} Z_{11}^{(2,1)},\\
\gamma_1^{(1,1)}\gamma_2^{(1,1)} &= \gamma_2^{(1,1)} Z_{21}^{(1,1)} + \gamma_1^{(1,1)} Z_{12}^{(2,1)},\\
\gamma_2^{(1,1)}\gamma_1^{(1,1)} &= \gamma_1^{(1,1)} Z_{12}^{(1,1)} + \gamma_2^{(1,1)} Z_{21}^{(2,1)},\\
\gamma_2^{(1,1)}\gamma_2^{(1,1)} &= \gamma_2^{(1,1)} Z_{22}^{(1,1)} + \gamma_2^{(1,1)} Z_{22}^{(2,1)}.
\end{aligned}$$

The last 4 conditions above are automatically satisfied by the first 8.

Finally, from (5.4b) we get 2 conditions after noting that $\gamma^{(1,1)} = \gamma^{(2,1)}$,

$$\begin{aligned}
\frac{1}{2}J_{(1)}^2 &= (J_{(1)} - \gamma_1^{(1,1)})(\gamma_1^{(1,1)} + \gamma_2^{(1,1)}) = J_{(1)}^2 - J_{(1)}\gamma_1^{(1,1)},\\
\frac{1}{2}J_{(1)}^2 &= \gamma_1^{(1,1)}\gamma_1^{(1,1)} + (J_{(1)} - \gamma_1^{(1,1)})(\gamma_1^{(1,1)}).
\end{aligned}$$

The first condition implies that $\gamma_1^{(1,1)} = \frac{1}{2}J_{(1)}$ and the second condition becomes

$$\frac{1}{2}J_{(1)}^2 = \frac{1}{4}J_{(1)}^2 + (J_{(1)} - \frac{1}{2}J_{(1)})(\frac{1}{2}J_{(1)}) = \frac{1}{2}J_{(1)}^2.$$

□

The deterministic Störmer-Verlet method allows an explicit formulation for systems of the form $\dot{y} = f(z)$, $\dot{z} = g(y)$. A similar explicit formulation can be found in the stochastic case. Applying the stochastic Störmer-Verlet method to a system of the form

$$Y(t) = y_0 + \int_0^t f_1(Z(s))ds + \int_0^t g_1(Z(s)) \circ dW(s) \tag{5.6a}$$

$$Z(t) = z_0 + \int_0^t f_2(Y(s))ds + \int_0^t g_2(Y(s)) \circ dW(s) \tag{5.6b}$$

we get a scheme

$$\begin{aligned} Z_{n+1/2} &= Z_n + \frac{1}{2}hf_2(Y_n) + \frac{1}{2}J_{(1)}g_2(Y_n) \\ Y_{n+1} &= Y_n + hf_1(Z_{n+1/2}) + J_{(1)}g_1(Z_{n+1/2}) \\ Z_{n+1} &= Z_n + \frac{1}{2}h\left(f_2(Y_n) + f_2(Y_{n+1})\right) + \frac{1}{2}J_{(1)}\left(g_2(Y_n) + g_2(Y_{n+1})\right) \end{aligned} \tag{5.7}$$

Clearly, (5.7) is an explicit scheme.

## 5.4 Extending known deterministic methods

The results in the previous subsections suggest a more general result.

**Theorem 5.4.** *Let $J_{(0)} = h$, $\gamma^{(k,0)} = b^{(k)}J_{(0)}$ and $Z^{(k)} = A^{(k,0)}J_{(0)}$ for $k = 1, 2$ be the RK coefficients of a deterministic PRK scheme of at least order 1 that conserves quadratic invariants $Q(y, z) = y^\top Dz$. Then the Stratonovich SPRK method with $\gamma^{(k,l)} = b^{(k)}J_{(l)}$ and $Z^{(k,l)} = A^{(k)}J_{(l)}$ for $k = 1, 2$ and $l = 0, 1$ is of strong order 1 and conserves $Q(y, z)$.*

*Proof.* From Theorem 4.5 we find the conditions for preserving quadratic invariants,

$$\begin{aligned} (b_i^{(1)} - b_i^{(2)})J_{(l)} = 0 \quad \forall i = 1, \dots, s \text{ and } \forall l = 0, 1, \\ (b_i^{(1)}b_j^{(2)} - b_j^{(2)}A_{ji}^{(1)} - b_i^{(1)}A_{ij}^{(2)})J_{(l)}J_{(k)} = 0 \quad \forall i, j = 1, \dots, s \text{ and } \forall k, l = 0, 1 \end{aligned}$$

where we have written $\gamma^{(m,l)} = b^{(m)}J_{(l)}$ and $Z^{(m,l)} = A^{(m)}J_{(l)}$. By assumption this is true for the deterministic method, i.e. for $J_{(l)} = J_{(k)} = h$, and then clearly also true for all $k, l = 0, 1$. The SPRK then conserves quadratic invariants.

We can write the order conditions as

$$J_{(l)} = b^{(k)}\mathbb{1}_s J_{(l)}, \quad l = 0, 1,\ k = 1, 2 \tag{5.8a}$$

$$J_{(1,1)} = \gamma^{(k_1,1)}Z^{(k_2,1)}\mathbb{1}_s, \quad k_1, k_2 = 1, 2 \tag{5.8b}$$

$$0 = b^{k_1}A^{(k_2)}\mathbb{1}_s h\mathrm{E}\left[J_{(1)}\right], \quad k_1, k_2 = 1, 2 \tag{5.8c}$$

$$0 = b^{k_1}A^{(k_2)}\mathbb{1}_s \mathrm{E}\left[J_{(1)}\right] h, \quad k_1, k_2 = 1, 2 \tag{5.8d}$$

$$0 = b^{(k_1)}\left(A^{(k_2)}\mathbb{1}_s\right)\left(A^{(k_3)}\mathbb{1}_s\right)\mathrm{E}\left[J_{(1)}^3\right],\ k_1, k_2, k_3 = 1, 2 \tag{5.8e}$$

$$0 = b^{(k_1)}\left(A^{(k_2)}A^{(k_3)}\mathbb{1}_s\right)\mathrm{E}\left[J_{(1)}^3\right],\ k_1, k_2, k_3 = 1, 2. \tag{5.8f}$$

Condition (5.8a) is satisfied for $l = 0$ by assumption, and thus clearly also for $l = 1$. The latter 4 conditions are trivially satisfied because odd powers of the Wiener process have expected value 0. From the proof of Theorem 5.1 we know that (5.8b) is satisfied if the method conserves quadratic invariants. □

# 6 Numerical results

In this chapter we run some numerical tests using the stochastic Gauss and stochastic Störmer-Verlet methods we found in chapter 5. We numerically verify that they are of strong order 1 and conserve quadratic invariants.

## 6.1 Example 1: A non-linear test equation

To numerically test the order of convergence of the stochastic Gauss method we apply it to the test problem

$$dX(t) = \left(\frac{1}{2}X(t) + \sqrt{X(t)^2+1}\right)dt + \sqrt{X(t)^2+1}dW(t),\ X(0) = 0.$$

This is an Itō SDE taken from Kloeden and Platen [9] with exact solution $X(t) = \sinh(t + W(t))$. Because we use a Stratonovich method we must first find the corresponding Stratonovich SDE with the same solution. By the conversion formula (3.5) we find the modified drift coefficient

$$\begin{aligned}\bar{f}(t,x) &= f(t,x) - \frac{1}{2}g(t,x)\frac{\partial g(t,x)}{\partial x} \\ &= \left(\frac{1}{2}x + \sqrt{x^2+1}\right) - \frac{1}{2}\sqrt{x^2+1}\frac{x}{\sqrt{x^2+1}} = \sqrt{x^2+1},\end{aligned}$$

and the Stratonovich SDE is given by

$$dX(t) = \sqrt{X(t)^2+1}dt + \sqrt{X(t)^2+1}\circ dW(t),\ X(0) = 0. \tag{6.1}$$

We calculate 10,000 sample paths of (6.1) on the time interval $I = [0, 0.1]$ with step sizes $h = 0.1 \cdot 2^{-k}$ for $k = 5, \ldots, 9$ and compare it with the exact solution. The result is shown in Figure 1.

## 6.2 Example 2: Synchrotron oscillations

For the stochastic Störmer-Verlet method we look at set of equations from particle physics [15]. A storage ring is a particular type of cyclic particle accelerator where the kinetic energy is kept constant. Assume a reference particle travels with fixed energy along the design orbit of the storage ring. We consider another particle under the influence of fluctating electromagnetic fields, and this particle will perform stochastically perturbed oscillations with regards to the reference particle. The energy or phase oscillations with respect to the reference particle are called synchrotron oscillations.

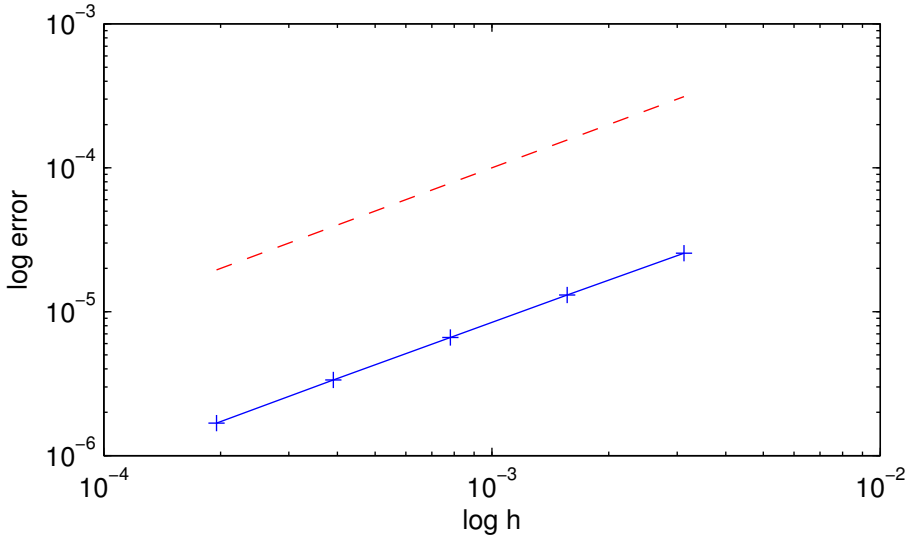



Figure 1: The order of convergence of the stochastic Gauss method. The dashed line is a reference line with slope 1.

The equations are derived from a stochastically perturbed Hamilitonian function of a pendulum [15]

$$H(p,x,t) = \frac{p^2}{2} - \omega^2(1+\lambda_A\xi_A(t))\cos(x) + \omega^2\sin(x)\lambda_{Ph}\xi_{Ph}(t). \tag{6.2}$$

Here $x$ is the longitudinal phase difference between the particle and the reference particle travelling along the design orbit, and $p$ is proportional to the energy deviation of the particle from the reference particle. The stochastic terms $\lambda_A\xi_A(t)$ and $\lambda_{Ph}\xi_{Ph}(t)$ are called amplitude noise and phase noise respectively.

We will consider the case of pure phase noise, $\lambda_A = 0$, and we write $\xi_{Ph}(t) = W(t)$. The equations of motion are then

$$dp = -\omega^2\sin(x)dt - \lambda_{Ph}\omega^2\cos(x)\circ dW(t), \quad dx = pdt, \tag{6.3}$$

and we solve them using the stochastic Störmer-Verlet scheme shown in Table 5.3.

Because the system is on the special form (5.6) we can use the explicit formulation (5.7). For the phase noise (6.3) this is simply

$$\begin{aligned} X_{n+1/2} &= X_n + \frac{1}{2}hP_n, \\ P_{n+1} &= P_n - h\omega^2\sin(X_{n+1/2}) - J_{(1)}\lambda_{Ph}\omega^2\cos(X_{n+1/2}), \\ X_{n+1} &= X_n + \frac{1}{2}h\left(P_n + P_{n+1}\right). \end{aligned}$$

We calculate 10,000 sample paths on the interval $I = [0, 0.1]$ for each of the step sizes $h = 0.1 \cdot 2^{-k}$ for $k = 8, 9, \ldots, 12$. The parameters are set to $\omega = 40\pi$ and $\lambda_{Ph} = 10^{-4}$, and we use initial values $X(0) = 0$ and $P(0) = 0$. Lacking an exact solution we apply the Stratonovich Heun scheme [4] with the step size $h = 0.1 \cdot 2^{-18}$ as a reference method. The order plot is shown in Figure 2

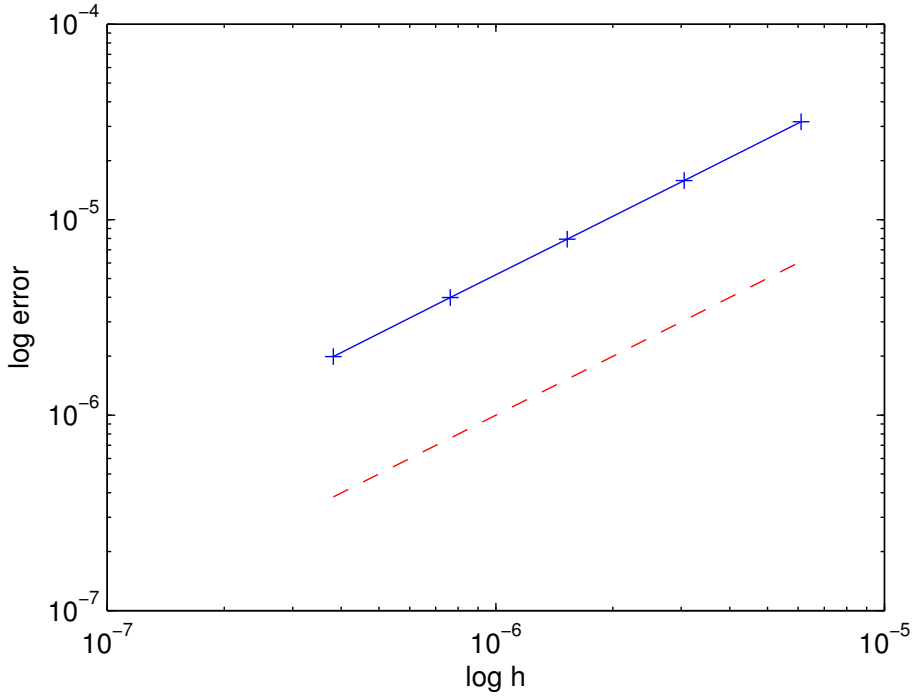



Figure 2: The order of convergence of the stochastic Störmer-Verlet method. The dashed line is a reference line with slope 1.

In the deterministic theory it has been proven [7] that structure preserving methods allow for more accurate long-term integration, and we would like to see how the stochastic Störmer-Verlet behaves over large time intervals. For each method we calculate the average of 1,000 sample paths, use a step size $h = 0.1$ and integrate to the time $T = 10,000$. In Figure 3 we show the error in the Hamiltonian as a function of time for the Stratonovich Heun and stochastic Störmer-Verlet schemes, respectively. Of course, the Hamiltonian is not a quadratic invariant and we do not expect the Störmer-Verlet method to conserve it, but in the deterministic theory it has been shown (see chapter IX in [7]) that symplectic methods[4] conserve a modified Hamiltonian. Figure 3 b) suggests that this might also be the case for

[4]As mentioned in chapter 2.5, all RK methods that conserve quadratic invariants are symplectic when applied to Hamiltonian systems.

the stochastic Störmer-Verlet method.

## 6.3 Example 3: A rigid body problem

The rigid body problem is based on a deterministic problem in [7]. Let $X = (X_1, X_2, X_3)^\top$ and let $W$ be a 1-dimensional Wiener process. The motion in 3 dimensions of a free rigid body with center of mass at the origin is described by the equations

$$dX = A(X)Xdt + g(X) \circ dW \tag{6.4}$$

with

$$A(X) = \begin{pmatrix} 0 & X_3/I_3 & -X_2/I_2 \\ -X_3/I_3 & 0 & X_1/I_1 \\ X_2/I_2 & -X_1/I_1 & 0 \end{pmatrix}$$

and some function $g(X)$. In the following we will use the initial value $x_0 = (\cos(1.1), 0, \sin(1.1))^\top$ and the principal moments of inertia $I_1 = 2$, $I_2 = 1$ and $I_3 = 2/3$. Furthermore, all methods use the step size $h = 0.1$ and end time $T = 10$.

For $g(x) = 0$ this reduces to the deterministic problem described in [7] with invariants $I(X) = X_1^2 + X_2^2 + X_3^2$ and

$$H(X) = \frac{1}{2}\left(\frac{X_1^2}{I_1} + \frac{X_2^2}{I_2} + \frac{X_3}{I_3}\right).$$

The solutions thus lie on the intersection between the sphere given by $I(X)$ and the ellipsoid given by $H(X)$. Figure 4 shows the result using the 4th order Gauss method. The blue ellipsoid corresponds to the invariant $H(X)$ while the gold sphere corresponds to the invariant $I(X)$. The solution is shown as white pearls, one for each step. As expected, the solution lies on the intersection between the sphere and ellipsoid, indicating that both invariants are conserved.

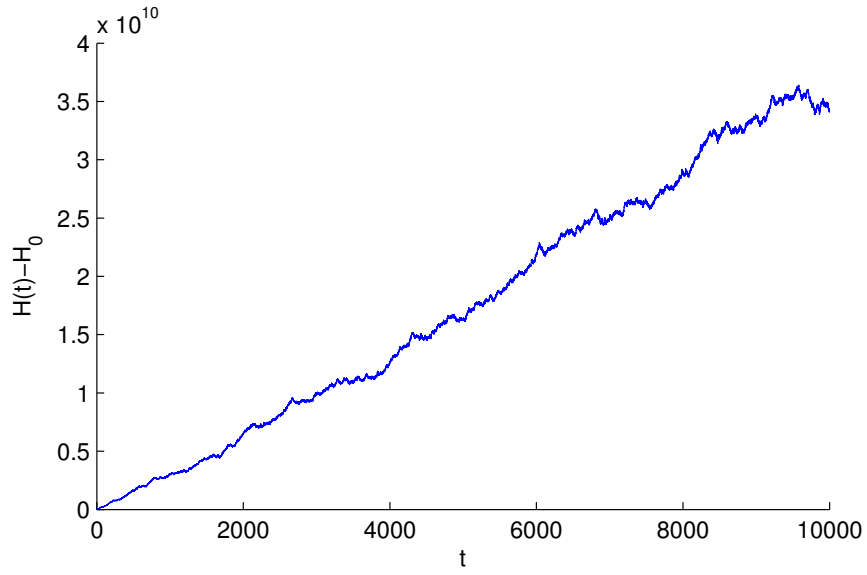
By Corollary 4.1 the choice $g(X) = A(X)X$ leads to an SDE with the same two invariants. The result of solving this with the stochastic Gauss method is shown in Figure 5, and, as we can see, the method conserves both invariants.

We can also choose $g(X)$ such that the system has only one of the two invariants. With $g(X) = (g_1(X), g_2(X), g_3(X))^\top$ we find

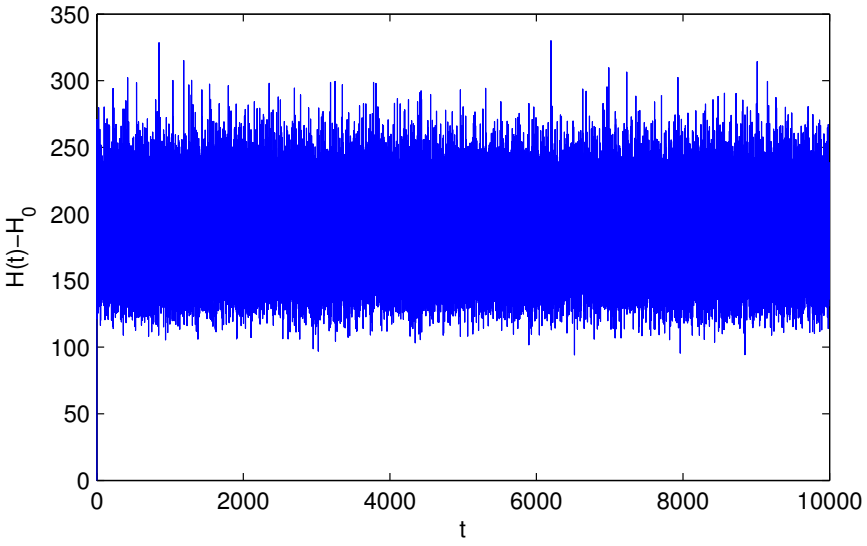
$$\nabla I(X)g(X) = 2\left(X_1 g_1(X) + X_2 g_2(X) + X_3 g_3(X)\right)$$

and

$$\nabla H(X)g(X) = \frac{1}{2}\left(X_1 g_1(X) + 2X_2 g_2(X) + \frac{4}{3}X_3 g_3(X)\right).$$




(a)




(b)

Figure 3: The error in the Hamiltonian as a function of time is shown for the synchrotron problem solved with a) the Stratonovich Heun method and b) the stochastic Störmer-Verlet method. For both methods we show an average of 1,000 sample paths, using a step size of 0.1 and integrating to $T = 10,000$. The stochastic Störmer-Verlet scheme shows much better long term behaviour when applied to this sample problem.

Choosing, for example, $g(X) = (X_2, -X_1, 0)^\top$ we see that $I(X)$ remains an invariant while $H(X)$ does not. This situation is shown in Figure 6 where the solution remains on the sphere given by $I(X)$, as expected.

As a comparison we apply the Euler-Maruyama method to the problem with $g(X) = A(X)X$. The system has two invariants, but as we can see in Figure 7, neither of them are conserved by this method.

## 6.4 Example 4: The outer solar system

We look at a Hamiltonian system describing the motion of the five outer planets relative to the Sun. This is an $N$-body problem taken from chapter I.2.4 in [7], and it has the form

$$H(p,q) = \frac{1}{2}\sum_{i=0}^{5}\frac{1}{m_i}p_i^\top p_i - G\sum_{i=1}^{5}\sum_{j=0}^{i-1}\frac{m_i m_j}{||q_i - q_j||}, \tag{6.5}$$

where $p$ and $q$ are supervectors composed of the vectors $p_i, q_i \in \mathbb{R}^3$ for $i = 0, 1, \ldots, 5$. The vectors $p_i$ and $q_i$ represent the moment and position, respectively, of planet $i$ in 3-dimensional space. Furthermore, $G$ is the gravitational constant, $m_i$ is the mass of planet $i$ and the norm is the Euclidean norm. The Sun and the inner planets are taken together and are represented as $i = 0$.

We add a stochastic term to the Hamiltonian equations such that the moment and position vectors become, respectively,

$$dq = H_p(p,q)dt + \alpha H_p(p,q)dW_t, \tag{6.6b}$$
$$dp = -H_q(p,q)dt - \alpha H_q(p,q)dW_t, \tag{6.6a}$$

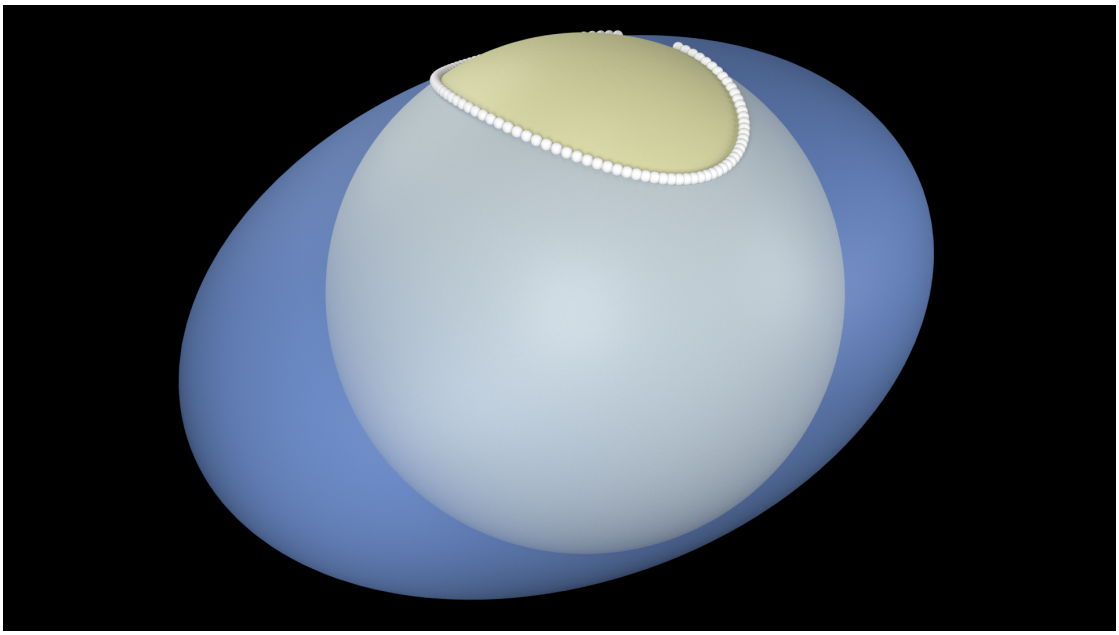
where $H_q = \nabla_q H = (\partial H/\partial q)^\top$ and $H_p = \nabla H = (\partial H/\partial p)^\top$ are the column vectors of partial derivatives and $\alpha \in \mathbb{R}$.

From (6.5) and (6.6) we get the equations of motion

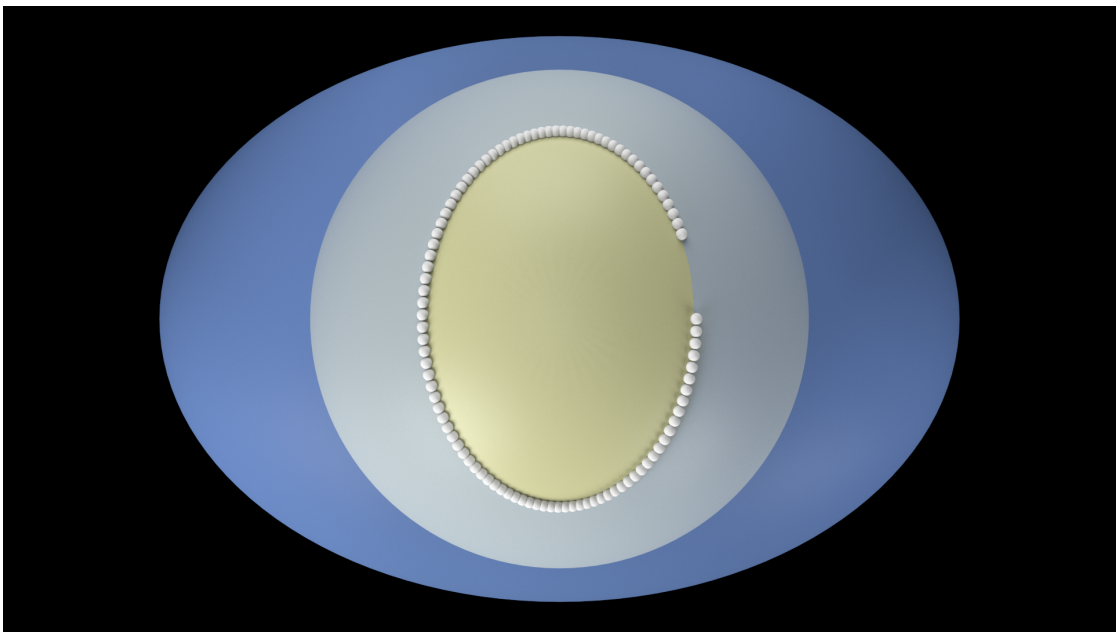
$$dq_i = \frac{1}{m_i}p_i(dt + \alpha dW_t), \tag{6.7a}$$
$$dp_i = -G\sum_{j=0, j\neq i}^{5}\frac{m_i m_j(q_i - q_j)}{||q_i - q_j||^3}(dt + \alpha dW_t). \tag{6.7b}$$

The Hamiltonian (6.5) is an invariant of this system because $H'(p,q) = (\partial H/\partial p, \partial H/\partial q)$ and

$$\frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial q}dt - \alpha\frac{\partial H}{\partial q}dW_t\right)^\top + \frac{\partial H}{\partial q}\left(\frac{\partial H}{\partial p}dt + \alpha\frac{\partial H}{\partial p}dW_t\right)^\top = 0.$$

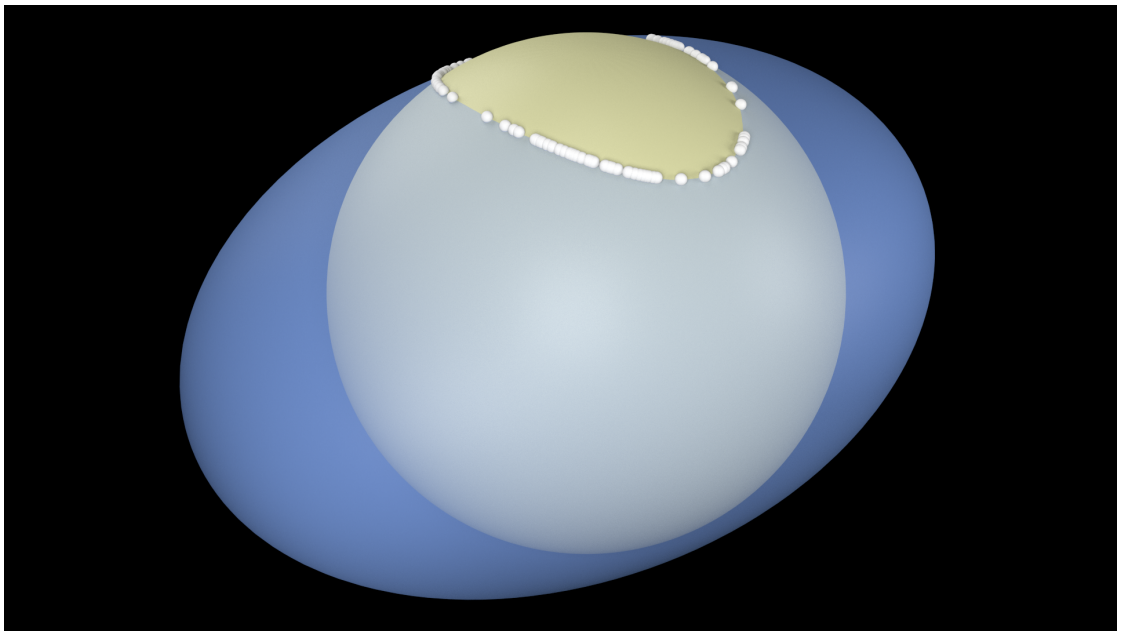


(a)

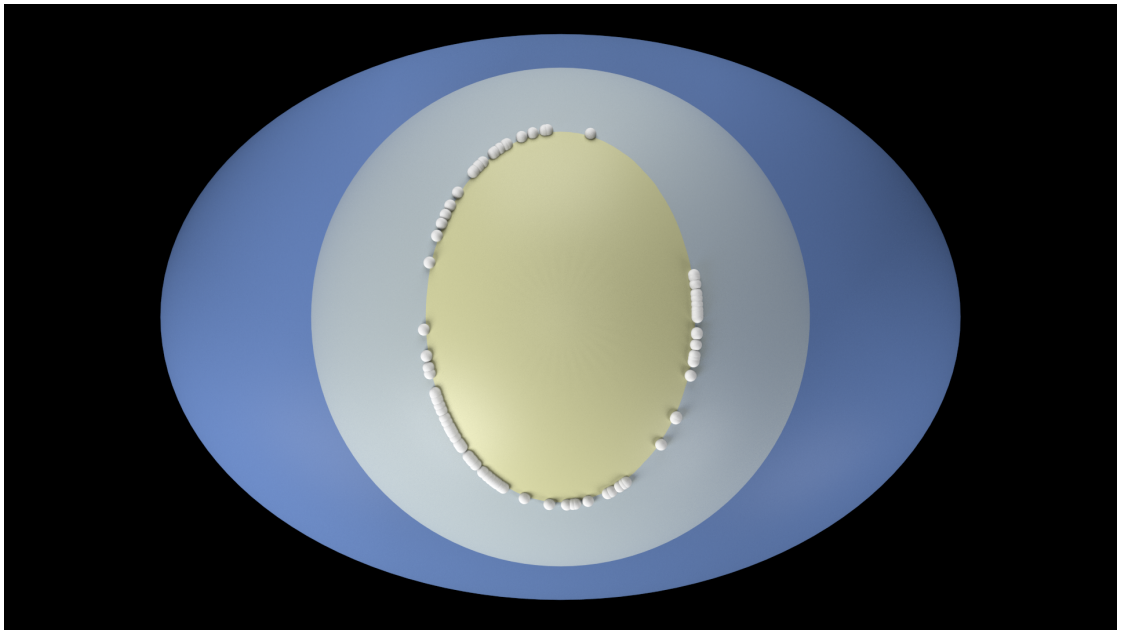


(b)

Figure 4: The rigid body problem with $g(X) = 0$ solved with the deterministic 4th order Gauss method. Both invariants are conserved.

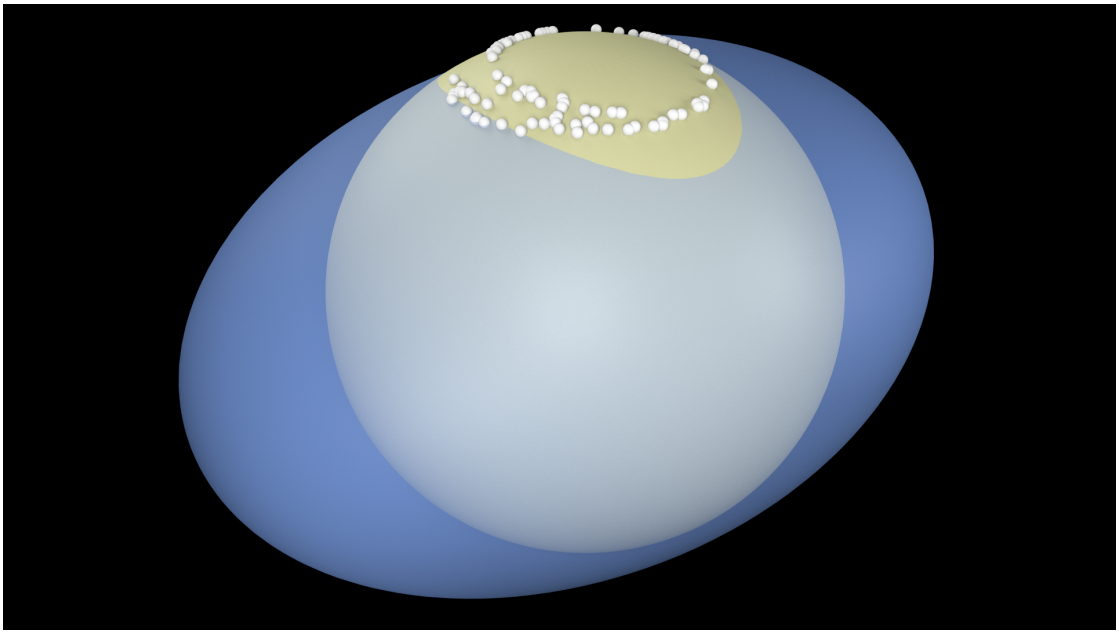


(a)

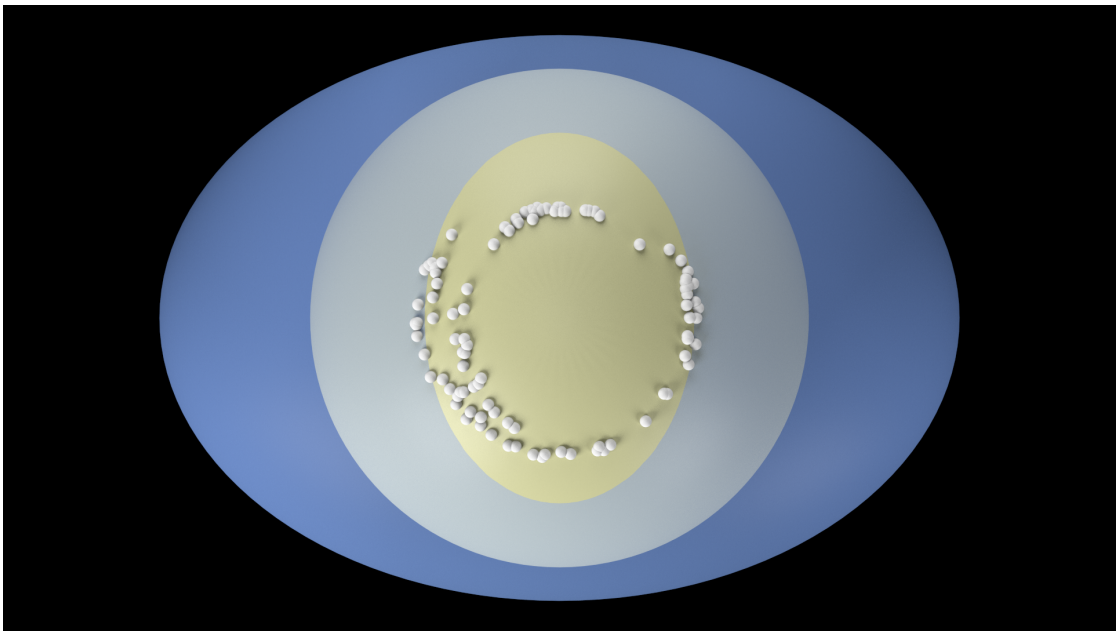


(b)

Figure 5: The rigid body problem with $g(X) = A(X)X$ and solved with the stochastic Gauss method. Both invariants are conserved.

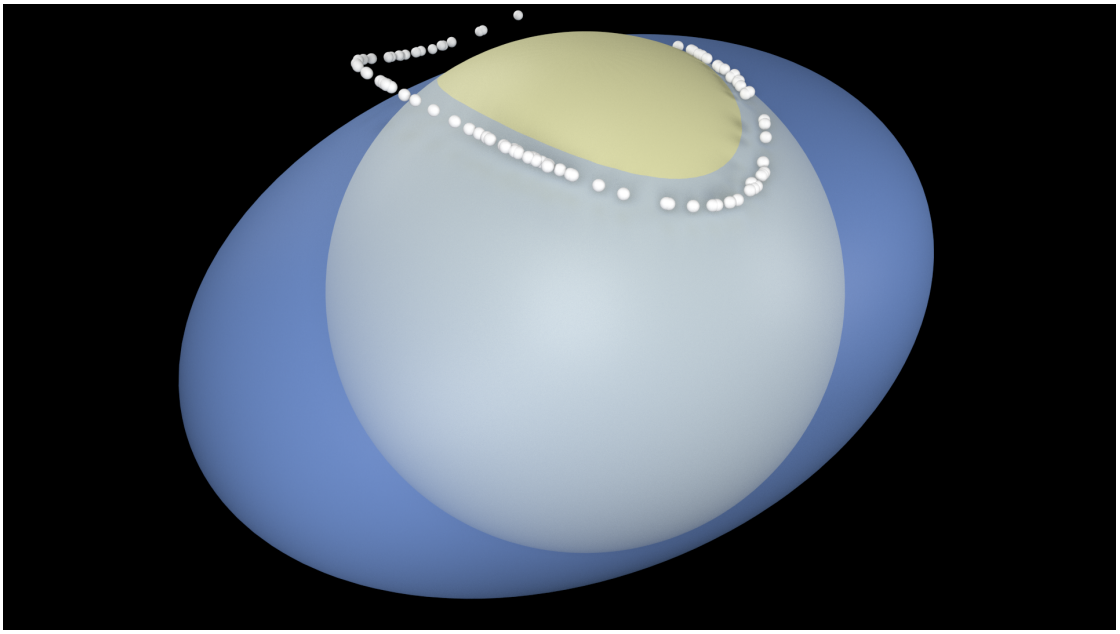


(a)

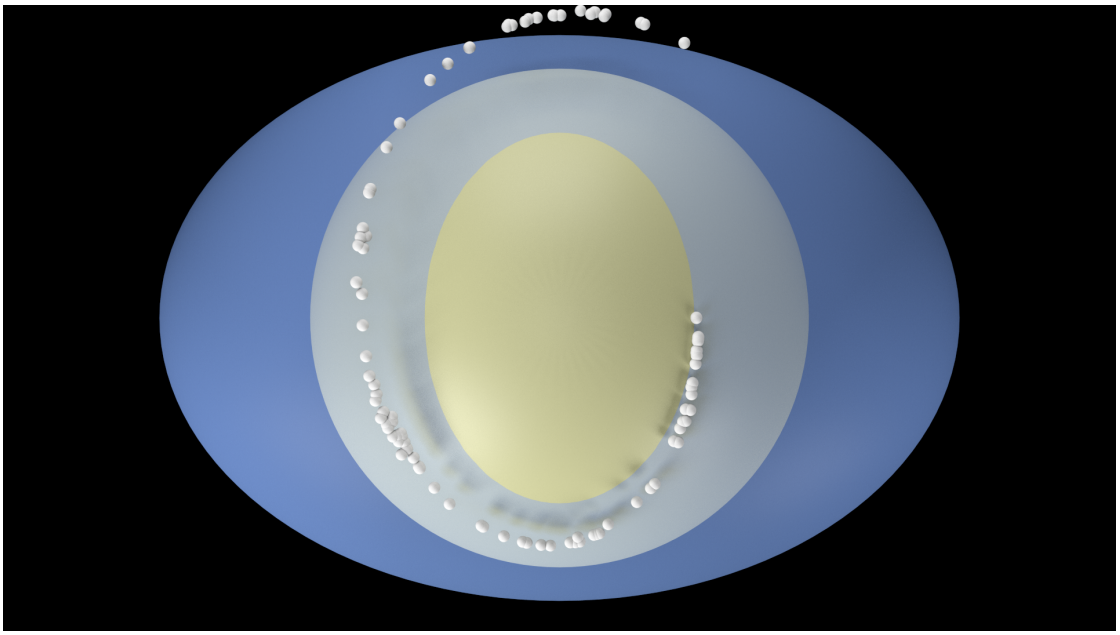


(b)

Figure 6: The rigid body problem with $g(X) = (X_2, -X_1, 0)^\top$ and solved with the stochastic Gauss method. The system has only one invariant in this case, drawn as a gold sphere, and we see that it is conserved by the method.



(a)



(b)

Figure 7: The rigid body problem with $g(X) = A(X)X$ and solved by the Euler-Maruyama method. The system has two invariants, but neither of them are conserved.

Similarly to the deterministic case we can show that the total linear momentum, $P = \sum_{i=0}^{5} p_i$, and angular momentum $L = \sum_{i=0}^{5} q_i \times p_i$ are invariants of (6.6). For the total linear momentum we have

$$\begin{aligned} dP &= -G \sum_{i=1}^{5} \sum_{j=0, j\neq i}^{5} \frac{m_i m_j (q_i - q_j)}{||q_i - q_j||^3} (dt + \alpha dW_t) \\ &= -G \sum_{i=1}^{5} \sum_{j=0}^{i-1} \frac{m_i m_j \left((q_i - q_j) + (q_j - q_i)\right)}{||q_i - q_j||^3} (dt + \alpha dW_t) = 0, \end{aligned}$$

and for the angular momentum we have

$$\begin{aligned} \frac{dL}{dt} &= \sum_{i=0}^{5} \frac{dq_i}{dt} \times p_i + \sum_{i=0}^{5} q_i \times \frac{dp_i}{dt} \\ dL &= \sum_{i=0}^{5} (p_i (dt + \alpha dW_t)) \times p_i + \sum_{i=1}^{5} q_i \times \left( -G \sum_{j=0, j\neq i}^{5} \frac{m_i m_j (q_i - q_j)}{||q_i - q_j||^3} (dt + \alpha dW_t) \right) \\ &= \sum_{i=0}^{5} (p_i \times p_i) (dt + \alpha dW_t) \\ &\quad -G \sum_{i=1}^{5} \sum_{j=0, j\neq i}^{5} \frac{m_i m_j}{||q_i - q_j||^3} (q_i \times ((q_i - q_j)) (dt + \alpha dW_t)) \\ &= -G \sum_{i=1}^{5} \sum_{j=0, j\neq i}^{5} \frac{m_i m_j}{||q_i - q_j||^3} (q_i \times (q_i - q_j)) (dt + \alpha dW_t) \\ &= -G \sum_{i=1}^{5} \sum_{j=0}^{i-1} \frac{m_i m_j}{||q_i - q_j||^3} ((q_i \times -q_j) + (q_j \times -q_i)) (dt + \alpha dW_t) \\ &= 0. \end{aligned}$$

We note that linear momentum and angular momentum are 3-dimensional vectors, and for both vectors we then have 3 invariants, one for each dimension.

We solve (6.7) using the stochastic Störmer-Verlet method. For the values of the gravitational constant, masses of the planets, initial positions and initial momenta used in this problem we refer to chapter I.2.4 in Hairer, Lubich and Wanner [7]. We use the step size $h = 200$, $\alpha = 1$ and integrate to time $T = 1,000,000$. The result is shown in Figure 8. We note that the stochastic effect is not readily apparent in this plot.

When studying the invariants we integrate to time $T = 20,000,000$ to better observe their longterm behaviour. In Figure 9 we show the error of the Hamiltonian. While $H$ is an invariant it is not a quadratic invariant, and thus not conserved by the stochastic Störmer-Verlet method.

The errors in the total angular momentum and total linear momentum appear to be conserved, however, as seen in Figure 10. While there is some error, it is on the scale of $10^{-17}$ for all the invariants. We attribute this to numerical imprecision.

For comparison we implemented the Stratonovich Heun method. The Hamiltonian is shown in Figure 11. The stochastic effect on the error is noticeably smaller, but it appears that that system is gradually losing energy, i.e. that $H$ is decreasing.

We know that the total linear momentum is conserved as the Stratonovich Heun method belongs to the SRK class which conserves all linear invariants. The angular momentum is a quadratic invariant and is not conserved, as shown in Figure 12.

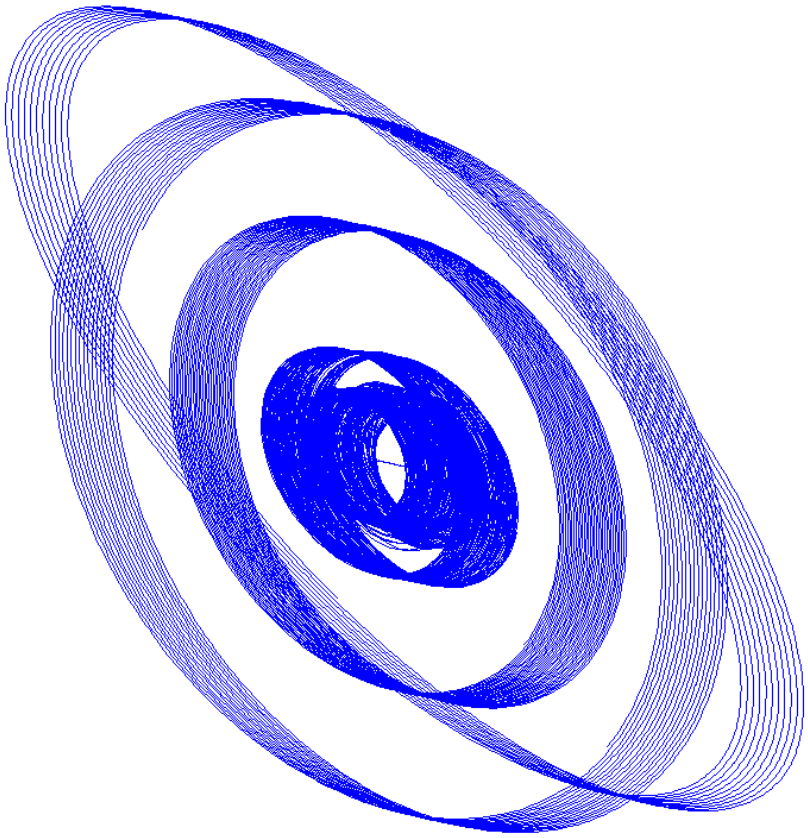


Figure 8: The motion of the 5 outer planets as described by the outer solar system problem in example 4. The system is solved with the stochastic Störmer-Verlet method using a step size of $h = 200$ and the end time $T = 1,000,000$.

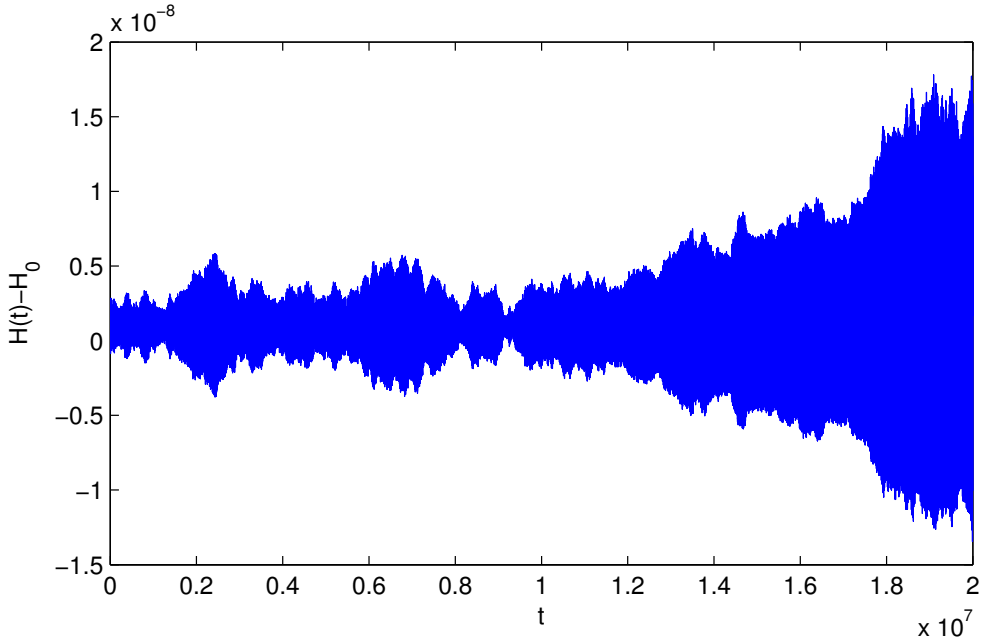



Figure 9: The error in the Hamiltonian as a function of time when applying the stochastic Störmer-Verlet scheme to the outer solar system problem in example 4. We see that the Hamiltonian is not conserved, as expected.

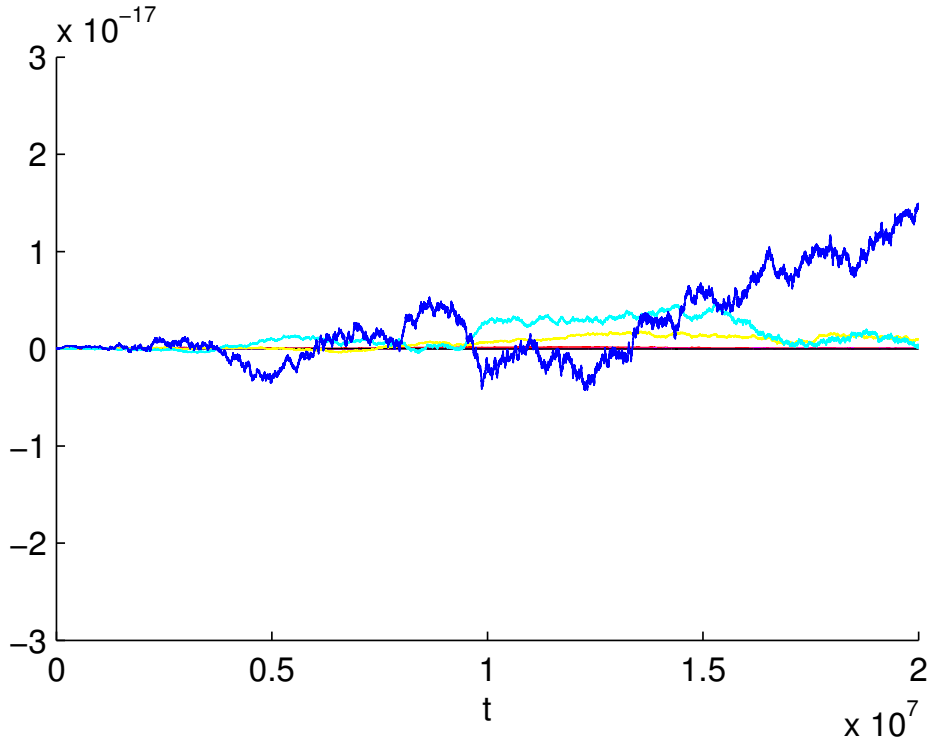



Figure 10: The errors in the invariants as functions of time when applying the stochastic Störmer-Verlet scheme to the outer solar system problem in example 4. The errors in the 3 dimensions of the total linear momentum are $P_1$ (red), $P_2$ (magenta) and $P_3$ (black). The errors for the total angular momentum are $L_1$ (yellow), $L_2$ (cyan) and $L_3$ (blue). We attribute these very small errors to numerical imprecision.

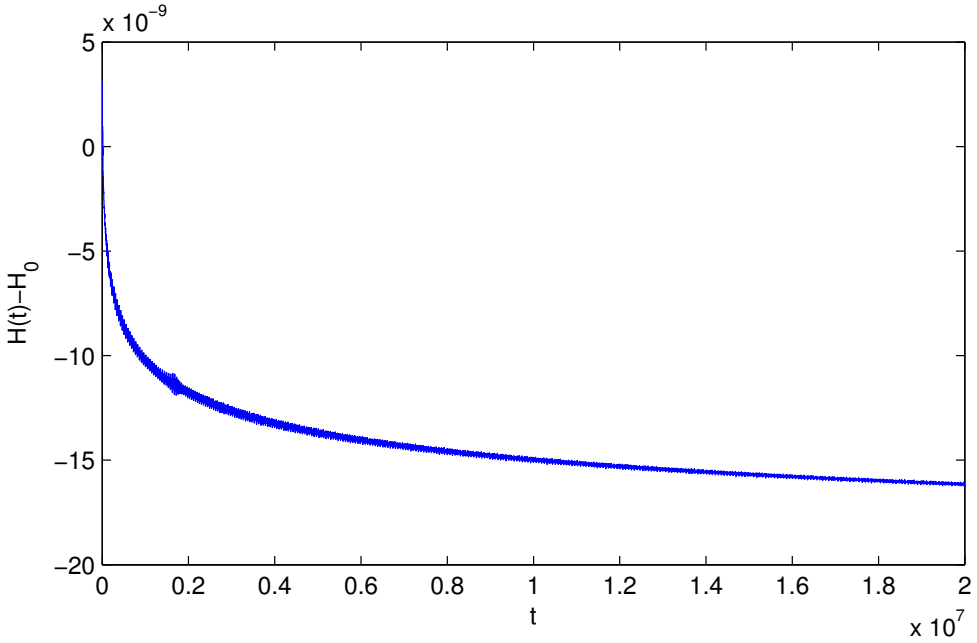



Figure 11: The error in the Hamiltonian as a function of time when applying the Stratonovich Heun scheme to the outer solar system problem in example 4. We see that $H$ is not conserved by the Stratonovich Heun method.

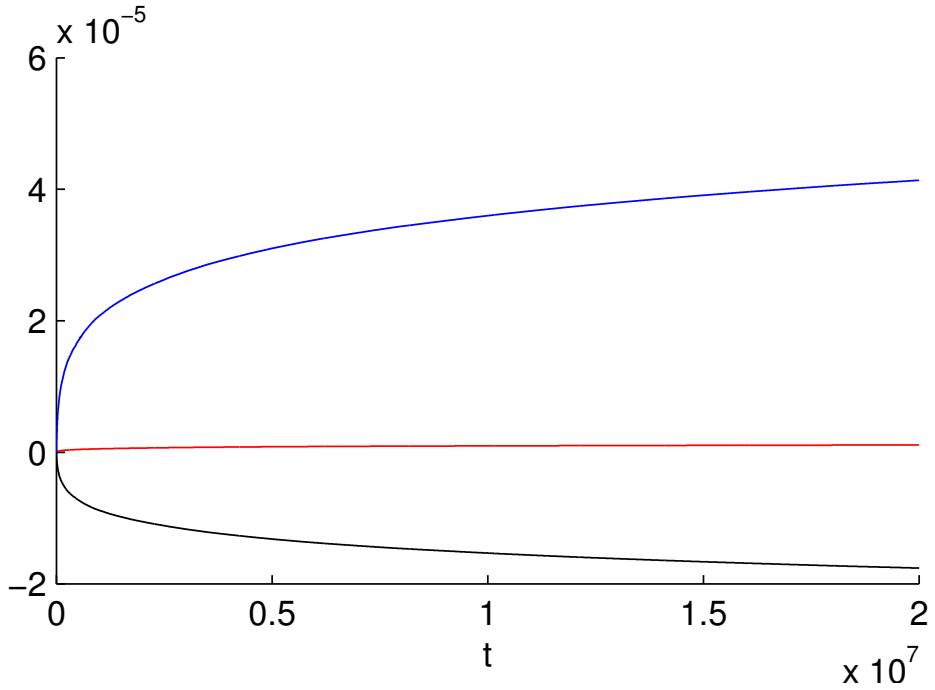



Figure 12: The errors in the 3 dimensions of the total angular momentum as a function of time when applying the Stratonovich Heun scheme to the outer solar system problem in example 4. Here shown as $P_1$ (red), $P_2$ (black) and $P_3$ (blue). We see that these invariants are not conserved.

# 7 Final remarks

## 7.1 Summary

We have extended some results from geometric numerical integration to stochastic differential equations. For stochastic Runge-Kutta (SRK) and stochastic partitioned Runge-Kutta (SPRK) methods we found necessary conditions for the conservation of quadratic invariants, and we also showed that the Itō SRK methods of the form we considered in this paper cannot conserve such invariants. We then expanded the order theory based on rooted trees to cover SPRK methods, and we showed how the majority of the order conditions are dependent if the SPRK method conserves quadratic invariants. Together, these results build the framework that allows us to construct new SRK and SPRK methods.

We then constructed some new methods of strong order 1 based on existing, deterministic methods. A general result was also found that lets us expand any known RK method of at least order 2 that conserves quadratic invariants to an order 1 Stratonovich SRK method that does the same. Finally, we ran numerical tests where we verified the order and that the methods indeed conserve quadratic invariants.

## 7.2 Future work

There are predominantly two sets of problems I would have liked to have explored further, but could not find the time for during the work on this thesis.

- Use the theoretical framework developed in this thesis to construct higher order methods and methods that allow multiple stochastic processes.
- Explore the benefits of these methods when applied to stochastic Hamiltonian systems. A good starting point would be expanding the results in chapter IX of [7] to a stochastic setting.

# References

[1] S. Anmarkrud. Investigation of mean-square stability properties of a class of first order stochastic Runge-Kutta methods. *Specialization project*, 2014.

[2] P.B. Bochev and C. Scovel. On quadratic invariants and symplectic structure. *BIT Numerical Mathematics*, 34(3):337–345, 1994.

[3] K. Burrage and P.M. Burrage. Order conditions of stochastic Runge–Kutta methods by b–series. *SIAM Journal on Numerical Analysis*, 38(5):1626–1646, 2000.

[4] K. Burrage, P.M. Burrage, and T. Tian. Numerical methods for strong solutions of stochastic differential equations: An overview. *Royal Society of London*, 460:373–402, 2004.

[5] J.C. Butcher. Coefficients for the study of Runge-Kutta integration processes. *Journal of the Australian Mathematical Society*, 3:185–201, 5 1963.

[6] K. Debrabant and A. Kværnø. B-series analysis of stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values. *SIAM J. NUMER. ANAL.*, 47:181–203, 2008.

[7] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer, 2006.

[8] J. Hong, D. Xu, and P. Wang. Preservation of quadratic invariants of stochastic differential equations via Runge–Kutta methods. *Applied Numerical Mathematics*, 87(0):38 – 52, 2015.

[9] P.E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Verlag, 1992.

[10] T. Mikosch. *Elementary stochastic calculus*. World Scientific, 2011.

[11] G.N. Milstein. *Numerical integration of stochastic differential equations*. Environmental Science and Technology. Kluwer Academic Publishers, 1995.

[12] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[13] J.M. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. *Acta Numerica*, pages 243–286, 1991.

[14] J.M. Sanz-Serna and L. Abia. Order conditions for canonical Runge-Kutta schemes. *SIAM J. NUMER. ANAL.*, 28:1081–1096, 1991.

[15] M. Seeßelberg, H.P. Breuer, H. Mais, Petruccione, and J. Honerkamp. Simulation of one-dimensional noisy Hamiltonian systems and their application to particle storage rings. *Zeitschrift für Physik C Particles and Fields*, 62(1):63–73, 1994.