Volume preserving numerical integrators for ordinary differential equations

Huiyan Xue

Dissertation for the degree of Philosophiae Doctor (PhD)

Department of Mathematics
University of Bergen

August 2013

Acknowledgements

I would like to thank my supervisors Prof. Antonella Zanna Munthe-Kaas and Prof. Robert McLachlan for their support and guidance throughout the period of my PhD study. With the support of the NFR grant, highly encouraged by Antonella, I have spent more than six months abroad, worked with many scholars, attended several international conferences, summer schools and workshops, which tremendously broadened my vision not only in mathematics but also many other aspects.

Thanks Prof. Robert McLachlan and Prof. Arieh Iserles for inviting me to visit Massey University and University of Cambridge where I had a wonderful time. The discussions with Olivier Verdier inspired me a lot. Thanks Nils Henry, Håkon, Geir, Brett and Eirik for reading earlier drafts and providing invaluable comments. Thanks Prof. Reinout Quispel for inviting me to the TRAGIC conference in Launceston, Australia. The travel in Australia in 2010 with the Munthe-Kaas family was highly appreciated. Thanks Elena, Brynjulf, Hans, Robert and Antonella for organizing the MAGIC conferences.

I am happy and thankful to have shared office with Hilde Kristine and Alexander who have been helpful and answered all kinds of questions. Thanks Morten and other participants for the orienteering trainings, competitions and gatherings afterwards. I am grateful to Atle for the coffee time at HIB. Many thanks to Kent, Qianqian, ChinChin, Nils Henry, Shengtian, Tineke and many other friends, colleagues and the staff of department of Mathematics, UiB. You have made my life in Norway better, brought me Norwegian culture and your kind help whenever I needed.

Thanks Chinese Students and Scholars Associate of Bergen for organizing many activities. I am indebted a lot to many Chinese friends in Bergen and my neighbors at Fantoft (13C) for supporting me in the last years. The kind encouragement from my former friends and teachers helped me a lot during my stay in Norway.

Special thanks to my parents and family in China. Their support makes me become better and encourages me to go further. This dissertation is dedicated to my parents and their generation who were forced to leave school too early.

Outline of the thesis

This dissertation is submitted as a partial fulfillment of the requirements for the Doctor of Philosophy (PhD) degree at the Faculty of Mathematics and Natural Sciences, University of Bergen, Norway.

This thesis is about geometric numerical integration (GNI), focusing on volume preserving numerical integrators for ordinary differential equations, more precisely, for divergence-free vector fields. The thesis is part of the project GeNuIn Applications, whose aim is design, analysis, and implementation of leading edge algorithms for the numerical integration of differential equations arising in mechanics and control applications.

The thesis is organized in two parts. **Part I** includes the background for understanding the papers included in **Part II**. There are four chapters in **Part I**. We give an overview of geometric numerical integration and application of volume preservation in Chapter 1. In Chapter 2, we provide some relevant mathematical definitions and background theory. In Chapter 3, we give an overview of the volume preserving numerical integrators. In Chapter 4, we give an introduction to the papers in **Part II**.

The four papers included in **Part II** are as follows:

Paper A: Xue, H. and Zanna A., Explicit volume-preserving splitting methods for polynomial divergence-free vector fields, *BIT Numerical Mathematics*, Volume 53, Issue 1, pp. 265-281, March 2013.

Paper B: Xue, H., High order volume preserving integrators for three kinds of divergence-free vector fields via commutators. *Preprint*, 2013.

Paper C: Xue, H. and Zanna A., Generating functions and volume preserving mappings, To appear: *Discrete and Continuous Dynamics Systems*, Volume 34, Number 3, pp. 1229-1249, March 2014. Publish online: *doi:10.3934/dcds.2014.34.1229*.

Paper D: Xue, H., Verdier O. and Zanna A., A study on volume preserving generating forms in $\mathbb{R}^3$. *Preprint*, 2013.

Contents

Part I

Background

Chapter 1

Introduction

> Motion is described by the differential equations, which are derived by the laws of physics. ... These laws are known as *symplectic* and *volume preservation*.
>
> p4, [14]

Summary: In this chapter, we introduce geometric numerical integration in Section 1.1, and in Section 1.2 we present the application of volume preserving integrators. The challenges and difficulties of constructing volume preserving numerical integrators are discussed in Section 1.3.

1.1 Geometric numerical integration

The goal of geometric numerical integration (GNI) is to design numerical integrators which preserve geometric structures, for instance, first integrals, phase space volume or symmetries of dynamical systems.

Let us consider the differential equation with initial value,

$$\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}), \;\; \mathbf{x}(0) = \mathbf{x}_0, \tag{1.1}$$

where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{a} : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{a}(\mathbf{x}) = [a_1(\mathbf{x}), \ldots, a_n(\mathbf{x})]^T$. The exact flow of this system, denoted as φ_τ, is defined by

$$\mathbf{x}(t+\tau) = \varphi_\tau(\mathbf{x}(t)), \text{ for any } t, \tau.$$

The flow has the continuous group property,

$$\varphi_{\tau_1} \circ \varphi_{\tau_2} = \varphi_{\tau_1+\tau_2}, \forall \tau_1, \tau_2 \in \mathbb{R}.$$

If we solve the differential equation from a given initial value $\mathbf{x}_0$ to time t_1, and then solve the equation from the resulting point $(\varphi_{t_1}(\mathbf{x}_0))$ forward to time t_2, the final result is the same as solving the equation with initial value $\mathbf{x}_0$ up to time t_1+t_2. In particular, if we set $\tau_2 = -\tau_1$, we have $\varphi_{\tau_1} \circ \varphi_{-\tau_1} = Id$, that is, $\varphi_{-\tau_1} = \varphi_{\tau_1}^{-1}$.

We can take the Taylor expansion for $\mathbf{x}(\tau)$ at 0:

$$\mathbf{x}(\tau) = \mathbf{x}(0) + \tau \frac{d\mathbf{x}}{dt}(0) + \frac{\tau^2}{2}\frac{d^2\mathbf{x}}{dt^2}(0) + \ldots$$

and from (1.1), we obtain

$$\frac{d\mathbf{x}}{dt} = \mathbf{a}(\mathbf{x}),$$
$$\frac{d^2\mathbf{x}}{dt^2} = (d\mathbf{a})\frac{d\mathbf{x}}{dt} = (d\mathbf{a})\mathbf{a}.$$

Therefore, for the flow φ_τ, we can obtain the Taylor expansion at $\mathbf{x}_0$,

$$\varphi_\tau(\mathbf{x}_0) = \mathbf{x}_0 + \tau\mathbf{a}(\mathbf{x}_0) + \frac{1}{2}\tau^2(d\mathbf{a}(\mathbf{x}_0))\mathbf{a}(\mathbf{x}_0) + \ldots \tag{1.2}$$

Denote by ψ_τ the approximation of the exact flow of the ODE system. We expect that the numerical integrator ψ_τ preserves as many as possible of the properties of the system. A standard method for the differential equation (*numerical integrator*) takes an initial value and provides a rule for computing approximations to the discrete trajectory. *Geometric numerical integrators* are numerical integrators that preserve some geometrical structures of the dynamical system. Since some of the underlying properties of the system have been preserved, the geometric numerical integrators behave favorably, especially when looking at long time periods. One of the most well-known examples illustrating the advantages of geometric numerical integrators is the calculation of the solar system by Wisdom et al. using a simple one-step, second order symplectic integrator. They ran a billion year simulation of the solar system with time step 7.5 days and obtained an energy error of only 2×10^{-11}. More details about Wisdom's example can be found in [14]. In [14], McLachlan and Quispel give some further examples on the advantages of preserving structures in dynamic systems. There are many references in the field of preserving the underlying structure of a dynamical system, see the standard reference book [5], and other references such as [9, 14–16].

1.2 Applications of volume preserving integrators

Motion is described by differential equations, and those equations are derived from the laws of physics. These laws are known as *symplecticity* and *volume preservation* [14]. The former law is for the Hamiltonian systems, and the latter one is for the divergence-free vector fields. Symplectic numerical integrators have been very popular in the last decades and hundreds of works have been done, see for instance, [5, 9] and references therein. Divergence-free vector fields occur naturally in incompressible fluid dynamics, and preservation of phase-space volume is also a crucial ingredient to reformulate hydrodynamics [10]. In fact, it is important to preserve the divergence-free structure since it gives rise to the existence of invariant tori [24]. In addition, preservation of volume by a numerical method for differential equations is a desirable property in the study of dynamical systems since the ODE system is volume preserving for any divergence-free vector field. Furthermore, volume preserving maps are simple and natural higher-dimensional generalization of the area-preserving maps which have been studied for many years.

We take the ABC flow to show that volume preserving integrators are crucial to

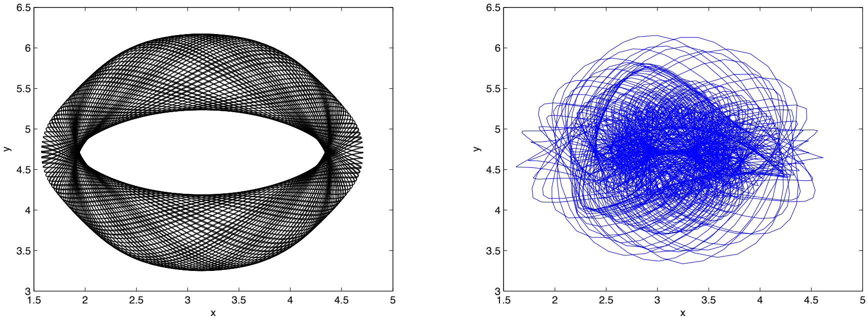

Figure 1.1: The orbit of ABC flow ($A = B = 1, C = 2$) with time step 0.1, integration time 750 and initial value [2,5,0]. The phase space $\mathbb{T}^3$ is viewed along the z-axis. Left: A second-order volume-preserving integrator. Right: MATLAB's ODE45 routine with the default setting.

preserve some structures of the flow for the divergence-free vector field,

$$\begin{aligned} \dot{x} &= A\sin z + C\cos y, \\ \dot{y} &= B\sin x + A\cos z, \\ \dot{z} &= C\sin y + B\cos x. \end{aligned} \tag{1.3}$$

The torus $\mathbb{T}^3$ is a periodicity box which we shall choose to be the cube $0 \leq x \leq 2\pi, 0 \leq y \leq 2\pi, 0 \leq z \leq 2\pi$. The ABC system has a mixture of quasi-periodic and chaotic orbits in $\mathbb{T}^3$. From Figure 1.1, one can see that the orbit obtained by a second order volume-preserving integrator (see Appendix A) lies on the torus. Its regular, quasi-periodic behavior is apparent, and the orbit obtained from the standard ODE45 routine is total chaos since the ODE45 uses adaptive time, which breaks the spatial symmetries.

1.3 Challenges in construction volume preserving numerical integrators

1.3.1 B-series methods cannot be volume preserving

Feng and Shang [4] showed that for a general, linear, divergence-free vector field of more than two dimension, there are no classical methods (e.g. Runge-Kutta methods) that analytically can preserve volume. Feng and Shang proved the result by showing that there exists no consistent approximation to the exponential function that maps the special linear Lie algebra $\mathfrak{sl}(n)$ ($tr(A) = 0$, A is the Jacobian of the flow) to the special linear group SL(n) ($det(A) = 1$), except for the exponential function itself. Exponential Runge-Kutta methods [6] can preserve volume for all linear divergence-free ODEs [23]. Unfortunately, a larger class of integrators characterized by being "B-Series" methods, including Runge-Kutta and Exponential Runge-Kutta methods, was proved to be not volume-preserving for all divergence-free ODEs [2, 8].

In [8], Iserles et al. showed that for any divergence-free differential equation $\nabla \cdot \mathbf{a}(\mathbf{x}) = 0$, if a B-series method is to be volume preserving, its modified differential equation $\tilde{\mathbf{a}}(\mathbf{x})$ must also be divergence-free. Then, the authors proved that for the B-series method, the modified equation must be the exact flow. As we know, partitioned Runge-Kutta methods are not B-series methods, hence it is reasonable to study volume preserving integrators based on them. Some notes on the study of the one-stage partitioned Runge-Kutta method for Hamiltonian systems and divergence-free vector fields in $\mathbb{R}^3$ can be found in Appendix C. Although the most useful partitioned Runge-Kutta methods in geometric integration are the Lobatto IIIA-IIIB methods, some of them are in fact splitting methods. In the next subsection we show that it is possible to construct volume preserving numerical integrators by splitting methods.

1.3.2 Difficulty of addressing the general space of divergence-free vector fields

One of the earliest volume preserving numerical methods is a splitting method proposed by Feng and Shang [4], decomposing the vector field into the sum of 2D Hamiltonian systems, which are then solved analytically or by a general symplectic method. Usually the splitting is not unique, one approach was shown in [4] and another was addressed in [14, 15].

Since it is hard to construct volume preserving integrators for generic divergence-free vector fields, many researchers recently turned to function spaces, for instance, the space of divergence-free polynomial and trigonometric fields, see references [13, 17, 20, 26, 30] and Paper A. The splitting method [4], the generating function approach [21] and the correction method [19] for general divergence-free vector fields are (often) implicit. However, explicit integrators via splitting methods for divergence-free polynomial vector fields were constructed by Xue and Zanna [26] (Paper A). For more general vector fields based on tensor product bases, explicit integrators were proposed by Zanna [30].

1.3.3 $n-1$ generating functions for a volume preserving map in n dimension

Generating function methods and generating forms have the property that they include the B-series methods as well as splitting methods as special cases. Therefore, it is reasonable to study generating functions and generating forms, see references [28] (Paper C) and [27] (Paper D).

Unlike the symplectic case in which a single function S (called *generating function*) can be found to describe entirely the dynamics of the mechanical system [5], the n-dimensional volume preserving problem is determined by $n-1$ functions, see details in [1]. Alternatively, $n-2$ differential forms (called *generating forms*) [10, 28], can be used to describe the volume preserving map on any contractible manifold [27, 28]. Therefore, the complexity of constructing volume preserving maps makes it very difficult to address the general n-dimensional case.

1.4 Prerequisites

Some prerequisites in geometric numerical integration are assumed throughout the text. Familiarity with the material in the classical book by E. Hairer, C. Lubich and G. Wanner [5] is recommended. Some knowledge about differential manifolds and dynamical systems is also needed. Some mathematical background is introduced in Chapter 2. In Chapter 3, we review the existing volume preserving integrators, including the methods presented in **Part II**.

Chapter 2

Background theory

> For any mathematical statement, if you say it is true or not, you should be able to prove it. 'I think it is true' is not a proof, it is a conjecture.
>
> A. Zanna Munthe-Kaas

Summary: In this chapter, we introduce some definitions and background theory. To start with, we give a brief description of differential manifolds, tangent spaces, vector fields and their flows in Section 2.1. In Sections 2.2 and 2.3, we give the basic notations, definitions and mathematical background for understanding [27, 28] (Paper C and D): differential forms, wedge products, exact and closed forms, volume forms, volume preservation and generating forms. In addition, in Section 2.4 we have a concise exposition of Lagrangian and Hamiltonian systems, and discrete Legendre transformations which relate the discrete Lagrangian systems and Hamiltonian systems and pave a background for part of [27].

2.1 Manifolds, tangent spaces and vector fields

In this section, we define *differential manifolds*, *tangent spaces*, *vector fields* and *flows*. This terminology is fundamental in differential geometry, dynamical systems and many other fields. The following text is mainly based on [7].

2.1.1 Differentiable manifolds

Manifolds can be seen as a generalization of vector spaces and intuitively one should think of an n-dimensional manifold as being a smooth domain which in a neighborhood of any point looks like $\mathbb{R}^n$ but typically looks different globally. We give the definition below.

Definition 1. *A differential manifold is a set $\mathcal{M}$ equipped with the following structures:*

1. A chart, which is a subset $U \subset \mathcal{M}$ with a bijective map $\chi : U \mapsto \chi(U) \subset \mathbb{R}^n$. A family of charts is called an atlas. Each point of $x \in \mathcal{M}$ is represented in at least one chart.

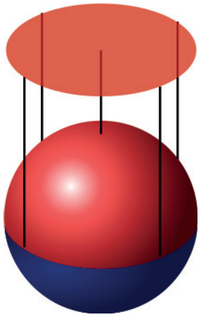

Figure 2.1: The chart maps the part of the sphere with positive z coordinate to a disc. The figure is from Wiki/Manifold.

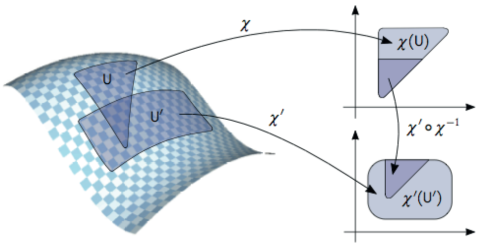

Figure 2.2: Two overlapping charts for a differential manifold. The map $\chi^{'} \circ \chi^{-1}$ is smooth. The figure is taken from [18].

2. *Every pair of charts* (U, χ) *and* $(U^{'}, \chi^{'})$ *of the atlas,* $\chi(U \cap U^{'})$ *and* $\chi^{'}(U \cap U^{'})$ *are open sets of* $\mathbb{R}^n$ *and the map*

$$\chi^{'} \circ \chi^{-1} : \chi(U \cap U^{'}) \mapsto \chi^{'}(U \cap U^{'})$$

is smooth.

In order to get a visual understanding of Definition 1, see Figure 2.2. See Figure 2.1 for the definition of a chart map.

2.1.2 Tangent spaces

The single most important property of a manifold is the existence of *tangents* to the manifold in any point $x \in \mathcal{M}$. One approach to define tangents is by differentiating a curve.

Definition 2. *Let* $\mathcal{M}$ *be an n-dimensional manifold and suppose that* $\gamma(t) \in \mathcal{M}$ *is a smooth curve such that* $\gamma(0) = x$*. A* tangent vector *at x is defined as*

$$v = \frac{d\gamma(t)}{dt}|_{t=0}.$$

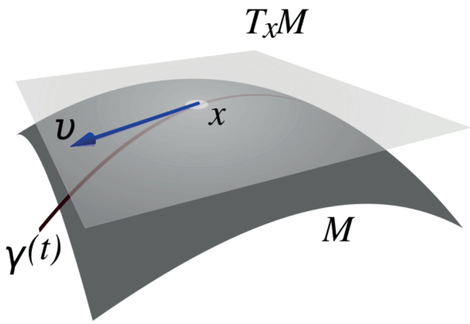

Figure 2.3: A tangent vector in the tangent space T_xM (from wiki/Tangent space)

The set of all tangents at x is called the *tangent space* at x and is denoted by $T_x\mathcal{M}$. The tangent space is an n-dimensional linear space, so addition and scalar multiplication of the tangent vectors are closed.

The collection of all tangent spaces at all points $x \in \mathcal{M}$ is called the *tangent bundle* of $\mathcal{M}$ and is denoted by $T\mathcal{M}$. $T\mathcal{M}$ is a $2n$-dimensional space, with elements (x,a) consisting of all possible points $x \in \mathcal{M}$ and all of its possible tangents a.

Definition 3. *The cotangent space of $\mathcal{M}$ at x denoted by $T_x^*\mathcal{M}$ is the dual space of $T_x\mathcal{M}$, that is, the space of linear forms on $T_x\mathcal{M}$.*

2.1.3 Vector fields and flows

Definition 4. *A vector field on $\mathcal{M}$ is a smooth function $F : \mathcal{M} \mapsto T\mathcal{M}$ such that $F(x) \in T\mathcal{M}|_x$ for all $x \in \mathcal{M}$. $\mathfrak{X}(\mathcal{M})$ denotes the collection of all vector fields on $\mathcal{M}$.*

Addition and scalar multiplication of *vector fields* are closed. If $F,G \in \mathfrak{X}(\mathcal{M})$, then $(F+G)(x) = F(x)+G(x) \in \mathfrak{X}(\mathcal{M})$ and $(bF)(x) = b(F(x)) \in \mathfrak{X}(\mathcal{M})$ for all real b.

Definition 5. *Let $F \in \mathfrak{X}(\mathcal{M})$. Given the differential equation on $\mathcal{M}$*

$$\mathbf{y}' = F(\mathbf{y}), \;\; \mathbf{y}(0) \in \mathcal{M}, \tag{2.1}$$

the flow *of F is the solution operator $\varphi_{t,F} : \mathcal{M} \mapsto \mathcal{M}$ such that*

$$\mathbf{y}(t) = \varphi_{t,F}(\mathbf{y}_0)$$

solves (2.1).

The *flow* of F is a smooth map which satisfies: 1) $\varphi_{o,x} = x$ for all $x \in \mathcal{M}$. 2) $\varphi_{s,\varphi_{t,x}} = \varphi_{s+t,x}$ for all $s,t \in \mathbb{R}$, $x \in \mathcal{M}$.

On a compact manifold there is a one-to-one correspondence between a vector field and its (global) flow. Going from flow to vector field is simple since we just need to differentiate. The other way around is much harder since we need to integrate the system. If the system cannot be integrated exactly, the goal is to approximate it by a numerical solution. If the system is divergence-free, a natural approach is to find a volume preserving (numerical) integrator.

2.2 Differential and volume forms

In this section, we introduce some basic definitions and notations which are used in [27, 28] (Papers C and D). We give basic definitions, properties and operators of differential forms, which are mostly based on Marsden and Ratiu [11]. To understand differential forms, we would like to quote the following description (p. 129, [11]).

> *The main idea of differential forms is to provide a generalization of the basic operations of vector calculus, div, grad and curl, and the integral theorems of Green, Gauss and Stokes to manifolds of arbitrary dimension.*

Definition 6 (Differential forms)**.** *Given a smooth manifold $\mathcal{M}$, a differential form ω of order k on $\mathcal{M}$ is a field of alternating k-linear maps. Define*

$$\omega_x : T_x\mathcal{M} \times \cdots \times T_x\mathcal{M} \mapsto \mathbb{R}$$

such that for all permutations σ of $\{1,\ldots,k\}$,

$$\forall (u_1,\ldots,u_k) \in (T_x\mathcal{M})^k,\ \omega_x(u_{\sigma(1)},\ldots,u_{\sigma(k)}) = \mathrm{sgn}(\sigma)\omega_x(u_1,\ldots,u_k),$$

where $\mathrm{sgn}(\sigma)$ *denotes the sign of σ and ω_x is linear with respect to each u_i, $i = 1,\ldots,k$.*

From Definition 6, we see that a differential k-form is multi-linear and skew,

$$\begin{aligned} \omega_x(v_1,\ldots,bv_j + cv'_j,\ldots,v_k) &= b\omega_x(v_1,\ldots,v_j,\ldots,v_k) \\ &\quad + c\omega_x(v_1,\ldots,v'_j,\ldots,v_k), \forall b,c \in \mathbb{R}, \end{aligned}$$

$$\omega_x(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = -\omega_x(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_k).$$

Let $(x^1,\ldots,x^n)$ denote the coordinates on $\mathcal{M}$, let

$$\{e_1,\ldots,e_n\} = \{\partial/\partial x^1,\ldots,\partial/\partial x^n\}$$

be the corresponding basis for the tangent space $T_x\mathcal{M}$ and let the dual basis for $T_x^*\mathcal{M}$ be

$$\{e^1,\ldots,e^n\} = \{dx^1,\ldots,dx^n\}.$$

Then, at each point $x \in \mathcal{M}$, a two-form can be written as

$$\omega_x(v_1,v_2) = \omega_{ij}v_1^i v_2^j,\ \text{ where } \omega_{ij}(x) = \omega_x(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}),$$

and more generally, a k-form can be defined similarly. More details can be found on page 130 in [11].

Definition 7 (The wedge product)**.** *The wedge product of a k-form ω and a l-form η on $\mathcal{M}$ is a $(k+l)$-form such that for all $x \in \mathcal{M}$ and for all $(u_1,\ldots,u_{k+l}) \in (T_x\mathcal{M})^{k+l}$,*

$$\begin{aligned} (\omega \wedge \eta)_x(u_1,\ldots,u_{k+l}) = \sum_{\sigma \in S_{k,l}} &\mathrm{sgn}(\sigma)\,\omega_x(u_{\sigma(1)},\ldots,u_{\sigma(k)}) \cdot \\ &\eta_x(u_{\sigma(k+1)},\ldots,u_{\sigma(k+l)}), \end{aligned}$$

where $S_{k,l}$ is the subset of permutations σ of $\{1,\ldots,k+l\}$, such that $\sigma(1) < \cdots < \sigma(k)$ and $\sigma(k+1) < \cdots < \sigma(k+l)$.

If ω and η are one-forms, then

$$(\omega \wedge \eta)(v_1, v_2) = \omega(v_1)\eta(v_2) - \omega(v_2)\eta(v_1).$$

If ω is a two-form and η is a one-form, then

$$(\omega \wedge \eta)(v_1, v_2, v_3) = \omega(v_1, v_2)\eta(v_3) + \omega(v_3, v_1)\eta(v_2) + \omega(v_2, v_3)\eta(v_1).$$

Proposition 1. *The wedge product has the following properties:*
1. It is associative: $\omega \wedge (\eta \wedge \gamma) = (\omega \wedge \eta) \wedge \gamma$.
2. It is bilinear in ω *and* η.
3. For any k-form ω *and any l-form* η, $\omega \wedge \eta$ *is anti-commutative:* $\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$.

Remark 1. *In terms of the dual basis* dx^i*, any k-form can be written locally as*

$$\omega = \sum \omega_{i_1,\dots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

where the sum is over all i_j *satisfying* $i_1 < \cdots < i_k$.

Definition 8 (Pullback)**.** *If* f *denotes a* C^1 *map from a smooth manifold* $\mathcal{M}$ *onto a smooth manifold* $\mathcal{N}$*, and* ω *is a differential form of order* k *on* $\mathcal{N}$*, then the pullback of* ω *by* f *at* x *is defined as*

$$\forall (u_1, \dots, u_k) \in (T_x\mathcal{M})^k, (f^*\omega)_x(u_1, \dots, u_k) = \omega_{f(x)}(df_x(u_1), \dots, df_x(u_k)),$$

where df_x *is the usual differential of* f *at* x.

Definition 9 (The exterior derivative)**.** *The exterior derivative* $d : \Lambda^k(\mathcal{M}) \mapsto \Lambda^{k+1}(\mathcal{M})$ *(*$\Lambda^k(\mathcal{M})$*: collection of all k-forms on* $\mathcal{M}$*) is a unique mapping such that:*

1. *If* ω *is a 0-form (i.e.* $\omega = f$*, where* f *is a function), then the one-form* df *is the differential of* f.
2. d *is linear, that is,* $d(c_1\omega + c_2\eta) = c_1 d\omega + c_2 d\eta$.
3. *If* ω *is a k-form and* η *is a l-form,* $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$.

If ω is given in canonical coordinates as $\omega = \sum_{i_1 < \cdots < i_k} \omega_{i_1,\dots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, then

$$d\omega = \sum_j \sum_{i_1 < \cdots < i_k} \frac{\partial \omega_{i_1,\dots,i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Proposition 2. *The exterior derivative is natural with respect to the pullback,* $d(f^*\omega) = f^* d\omega$.

Moreover, for any k-form ω, we have:

Proposition 3. *For any differential form* ω,

$$d^2\omega = d(d\omega) = 0.$$

Definition 10 (Exact forms and closed forms). *A differential k-form ω is exact if there exists a $(k-1)$-form ν such that $\omega = d\nu$. Any such ν is also called a primitive (or potential form) of ω. A differential form ω is closed if $d\omega = 0$.*

Thus, closed forms are the kernel of d, while exact forms are the image of d. Since $d^2 = 0$, any exact form is also closed. Is any closed form also exact? The following Poincaré lemma tells us when this is the case.

Theorem 1 (Poincaré lemma). *A closed form ($d\omega = 0$) is locally exact ($\omega = d\nu$), that is, there is a neighborhood U about each point on which $\omega = d\nu$. The statement is globally true on a contractible manifold. Furthermore, if $\mathcal{M}$ is simply connected, any closed form is also exact. In particular, if $\mathcal{M} = \mathbb{R}^n$, any closed form is exact.*

We define the Lie derivative on the differential form, then give the definition of the divergence of the vector field. It is shown that for any divergence-free vector field, its flow must be volume preserving.

Definition 11 (The Lie derivative). *Let ω be the k-form and let X be the vector field with flow φ_t. The Lie derivative of ω along X is given by*

$$L_X\omega = \lim_{t\to 0}\frac{1}{t}[\varphi_t^*\omega - \omega] = \frac{d}{dt}\varphi_t^*\omega|_{t=0}. \tag{2.2}$$

A *volume form* is a nowhere-vanishing (non-degenerate) n-form on an n-manifold $\mathcal{M}$. The *divergence* of the vector field X relative to $\nu \in \Omega^n(\mathcal{M})$ (a volume form) is

$$L_X(\nu) = (\nabla\cdot X)\nu = div_\nu(X).$$

Theorem 2. *Let φ_t be the exact flow of X. Then $div_\nu(X) = 0$ if and only if*

$$\varphi_t^*\nu = \nu, \tag{2.3}$$

i.e. the exact flow is volume preserving.

Proof. The proof is straightforward. From (2.2) and condition $div_\nu(X) = 0$, we obtain (2.3), and vice versa. □

2.3 Volume preservation

We consider the vector field (1.1) subject to the *divergence-free* (or *source-free*) condition

$$\nabla\cdot\mathbf{a} = \sum_{i=1}^{n}\partial_{x_i}a_i(\mathbf{x}) = 0. \tag{2.4}$$

Let $d\mathbf{a}$ be the derivative of the vector field $\mathbf{a}$, i.e. the matrix with elements $(\partial a_i/\partial x_j)$ at (i,j). Define $A = \partial\varphi_\tau/\partial\mathbf{x}$ to be the Jacobian of the flow $\varphi_\tau(\mathbf{x})$. The Jacobian of the flow φ_τ satisfies

$$\frac{dA}{dt} = (d\mathbf{a})A,\ \ A(0) = Id.$$

This implies that $\det(A(0)) = 1$. In fact,

$$\frac{d}{dt}det(A) = tr(d\mathbf{a})det(A) = 0,$$

since the divergence-free condition gives $\nabla \cdot d\mathbf{a} = tr(d\mathbf{a}) = 0$. So we have $\det(A) = 1$ for all time, that is, the determinant of the Jacobian of the flow is 1; the flow is volume preserving.

The goal of constructing a volume preserving numerical integrator is to find an integrator ψ_τ such that

$$det(\frac{\partial \psi_{\tau,i}}{\partial x_j}) = 1. \tag{2.5}$$

Alternatively, for a given volume preserving map $\mathbf{f} : \mathbf{x} \mapsto \mathbf{X}$, the Jacobian (determinant) is

$$J_{\mathbf{f}} = \left| \frac{\partial \mathbf{X}}{\partial \mathbf{x}} \right| = \left| \frac{\partial (X_1, X_2, \ldots, X_n)}{\partial (x_1, x_2, \ldots, x_n)} \right| = 1. \tag{2.6}$$

However, the above equations (2.5) and (2.6) are very hard to check. Considering that for any divergence-free vector field, its flow satisfies (2.3), we try to find a map $\mathbf{f}$ which also satisfies this condition. We have the following definitions.

Definition 12. *(*Volume preservation*). A volume form Ω on a manifold $\mathcal{M}$ is preserved by a C^1 map $\mathbf{f} : \mathcal{M} \mapsto \mathcal{M}$ if*

$$\mathbf{f}^*\Omega = \Omega. \tag{2.7}$$

We assume that the volume form Ω is exact, that is, there exists an $n-1$-form ν such that $\Omega = d\nu$. Then (2.7) becomes

$$\mathbf{f}^* d\nu - d\nu = 0$$

If $\mathcal{M} = \mathbb{R}^n$, through the Poincaré lemma, we have an $n-2$-form λ, such that $\mathbf{f}^*\nu - \nu = d\lambda$. We summarize this below.

Definition 13. *[10]. Let Ω be a volume form and ν a primitive, i.e. $\Omega = d\nu$. A diffeomorphism $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^n$ is ν-exact volume preserving if there exists an $n-2$ form λ on $\mathbb{R}^n$ such that*

$$\mathbf{f}^*\nu - \nu = d\lambda. \tag{2.8}$$

Primitives ν of a differential n-form are not uniquely determined. This motivates the generalization below.

Definition 14. *[10]. Suppose that $d\nu = d\tilde{\nu} = \Omega$ (volume form). A diffeomorphism $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^n$ is exact volume preserving with respect to $(\nu, \tilde{\nu})$ if*

$$\mathbf{f}^*\tilde{\nu} - \nu = d\lambda, \tag{2.9}$$

for an $n-2$ form λ. λ is called a generating form.

2.4 Discrete Lagrangians are generating functions

It is well-known that any n-dimensional divergence-free vector field can be split into $n-1$ 2D Hamiltonian systems [4]. Furthermore, the Hamiltonian systems can be described as Lagrangian systems via Legendre transformations. We study the discrete versions for the numerical integrators which require the definition of discrete Lagrangians and discrete Legendre transformations. One of the reasons why we study this background is that [12] showed that discrete Lagrangians are in fact generating functions and an objective of this thesis has been to use the generating functions and generating forms to study the volume preserving maps. Therefore, in this section, we follow the notations of Marsden and West [12] to discuss Lagrangian and Hamiltonian systems in order to understand part of [27] (Paper D). The readers who are familiar with this field can skip this subsection.

Consider a *configuration manifold* Q and a *Lagrangian* L: $TQ \mapsto \mathbb{R}$. For a given interval $[0,T]$, define the *path space*

$$\mathcal{C}(Q) = ([0,T],Q) = \{\mathbf{q} : [0,T] \mapsto Q | \mathbf{q} \textit{ is a } C^2 \textit{ curve}\},$$

and the *action map* $\mathfrak{G} : \mathcal{C}(Q) \mapsto \mathbb{R}$ as

$$\mathfrak{G}(\mathbf{q}) = \int_0^T L(\mathbf{q}(t), \dot{\mathbf{q}}(t))dt.$$

It can be proved that $\mathcal{C}(Q)$ is a smooth manifold and the action map $\mathfrak{G}$ is as smooth as the Lagrangian L [12].

Denote $T_q(\mathcal{C}(Q))$ as the tangent space of $\mathcal{C}(Q)$ at q which is the set of C^2 maps $v_q : [0,T] \mapsto TQ$. Let $\pi_Q : TQ \mapsto Q$ be the canonical projection such that $\pi_Q \circ v_q = q$. Define $\ddot{Q}$ (*second order submanifold*) to be the set of second order derivatives $\frac{d^2q}{dt^2}(0)$ of the curves $q : \mathbb{R} \mapsto Q$, which are elements of the form $((q,\dot{q}),(\dot{q},\ddot{q})) \in T(TQ)$.

Theorem 3. *[12]. Given a C^k Lagrangian L, $k > 2$, there exists a unique C^{k-2} mapping $D_{EL}L : \ddot{Q} \mapsto T^*Q$ and a unique C^{k-1} one-form Θ_L such that for all variations $\delta q \in T_q\mathcal{C}(Q)$ of the $q(t)$, we have*

$$< d\mathfrak{G}(q), \delta q >= \int_0^T D_{EL}L(\ddot{q}) \cdot \delta q dt + \Theta_L(\dot{q}) \cdot \hat{\delta} q|_0^T,$$

where $\hat{\delta}q(t) = ((q(t), \frac{\partial q}{\partial t}(t)),(\delta q(t), \frac{\partial \delta q}{\partial t}(t)))$. D_{EL} (called the Euler-Lagrange map*) has the coordinate form*

$$(D_{EL}L)_i = \frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}.$$

*The one-form Θ_L (*Lagrangian one-form*) has the coordinate expression*

$$\Theta_L = \frac{\partial L}{\partial \dot{q}^i} dq^i.$$

Proof. Using integration by parts. See details in [12]. □

The *Lagrangian vector field* $X_L : TQ \mapsto T(TQ)$ is a second order vector field on TQ which satisfies the following equation,

$$D_{EL} \circ X_L = 0,$$

and the *Lagrangian flow* $F_L : TQ \times \mathbb{R} \mapsto TQ$ is the flow of X_L. The *Lagrangian flow* is symplectic and its symplectic form is given in coordinates as

$$\Omega_L = d\Theta_L = \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j} dq^i \wedge dq^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} d\dot{q}^i \wedge dq^j.$$

A curve $\mathbf{q} \in \mathcal{C}(Q)$ is a solution of the *Euler-Lagrangian equations* if $\mathbf{q}$ satisfies the *Euler-Lagrangian equations*

$$\frac{\partial L}{\partial q^i}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}^i}(\mathbf{q}, \dot{\mathbf{q}})) = 0, \tag{2.10}$$

Example 1. *Given a Lagrangian* $L = \frac{1}{2}\dot{\mathbf{q}}^T M \dot{\mathbf{q}} - V(\mathbf{q})$ *where* M *is the mass matrix with elements* m_{ij} *at* (i, j)*, it is clear that*

$$\frac{\partial^2}{\partial \dot{q}^i \partial q^j} = 0, \quad \frac{\partial L^2}{\partial \dot{q}^i \partial \dot{q}^j} = m_{ij}.$$

The symplectic form in coordinates of the Lagrangian is given as

$$\Omega_L = M d\dot{\mathbf{q}} \wedge d\mathbf{q} = d\mathbf{p} \wedge d\mathbf{q},$$

where $\mathbf{p} = M\dot{\mathbf{q}}$.

Remark 2. [1] *If we define the path space as*

$$\mathcal{C}_{[c,b]}(Q) = \{C^2 \; curve \; [0, T] \mapsto Q \; with \; fixed \; end \; points \; q(0) = c, \; q(T) = b\},$$

the action map $\mathfrak{G}(\mathcal{C}_{[c,b]}(Q) \mapsto \mathbb{R})$ *is still defined as*

$$\mathfrak{G}(\mathbf{q}) = \int_0^T L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) dt,$$

then the trajectory of the mechanics system between two fixed points a,b is the one that minimizes the action function, i.e.

$$\mathbf{q} = argmin_{\mathbf{q}(0)=c, \mathbf{q}(T)=b} \int_0^T L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) dt,$$

which implies

$$0 = \delta \int_0^T L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) dt,$$

[1] Many books and references use path space with fixed points as in Remark 2. In this text, we introduce the path space without any fixed points, by involving the definition of the set of second order derivatives.

for all small variations of $\mathbf{q}$. *Using integration by parts, we have*

$$\begin{aligned} 0 &= \delta \int_0^T L(\mathbf{q}(t), \dot{\mathbf{q}}(t))dt \\ &= \int_0^T (\partial_{\mathbf{q}} L \delta \mathbf{q} + \partial_{\dot{\mathbf{q}}} L \delta \dot{\mathbf{q}})dt \\ &= \int_0^T (\partial_{\mathbf{q}} L - \frac{d}{dt}\partial_{\dot{\mathbf{q}}} L)\delta \mathbf{q} dt \\ &+ \partial_{\dot{\mathbf{q}}} L \delta_{\mathbf{q}}(T) - \partial_{\dot{\mathbf{q}}} L \delta_{\mathbf{q}}(0). \end{aligned}$$

The last two terms are zero since we fixed the end points, and the above equation is true for any $\mathbf{q}$. *Therefore, we obtain*

$$\partial_{\mathbf{q}} L - \frac{d}{dt}\partial_{\dot{\mathbf{q}}} L = 0,$$

i.e. the Euler-Lagrange equation.

To introduce the discrete Lagrangian L_d we define the *discrete state space* $Q \times Q$ instead of TQ. First of all, it is necessary to introduce the time step Δt, then take the L_d on this time step. Setting $\{t_k = k\Delta t | k = 0, \ldots, N\} \subset \mathbb{R}$, we can define the *Discrete Lagrangian* as a function $L_d : Q \times Q \mapsto R$ and the *discrete path space* as

$$\mathcal{C}_d(Q) = \mathcal{C}_d(\{t_k\}_{k=0}^N, Q) = \{q_d : \{t_k\}_{k=0}^N \mapsto Q\}.$$

Identifying a discrete trajectory $q_d \in \mathcal{C}_d(Q)$ with its image $q_d = \{q_k\}_{k=0}^N$ and $q_k = q_d(t_k)$, we can define the *discrete action map* $\mathfrak{G} : \mathcal{C}_d(Q) \mapsto \mathbb{R}$ as

$$\mathfrak{G}_d(q_d) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}).$$

Similarly to the continuous case, we can define the *discrete second-order submanifold* $\ddot{Q}_d$ with the form $((q_0, q_1), (q_1, q_2))$. Taking the derivative of the discrete action map as we did in the continuous case, we can easily obtain Theorem 1.3.1 [12].

The *discrete Euler-Lagrange map* has coordinate form

$$D_{DEL} L_d((q_{k-1}, q_k), (q_k, q_{k+1})) = D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}),$$

and the *discrete Lagrangian one-form* $\Theta^+_{L_d}$ and $\Theta^-_{L_d}$ in coordinates are

$$\begin{aligned} \Theta^+_{L_d}(q_0, q_1) &= D_2 L_d(q_0, q_1) dq_1 = \frac{\partial L_d}{\partial q_1^i} dq_1^i \\ \Theta^-_{L_d}(q_0, q_1) &= -D_1 L_d(q_0, q_1) dq_0 = -\frac{\partial L_d}{\partial q_0^i} dq_0^i. \end{aligned}$$

Observe that $dL_d = \Theta^+_{L_d} - \Theta^-_{L_d}$. From the property that $d^2 = 0$, this gives us

$$d\Theta^+_{L_d} = d\Theta^-_{L_d}.$$

The discrete Lagrangian maps are symplectic [12]. A discrete path $q_d \in \mathcal{C}_d(Q)$ is said to be the solution of *discrete Euler-Lagrange equations* if the points $\{q_k\}$ satisfy the *discrete Lagrange equations*

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0.$$

Example 2. *Given a Lagrangian $L = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$ and the discrete Lagrangian $L_d = \Delta t(\frac{1}{2}\frac{q_1-q_0}{\Delta t}M\frac{q_1-q_0}{\Delta t}) - \Delta t V(q_0)$, the one-form $\Theta^+_{L_d}(q_0,q_1)$ is*

$$\Theta^+_{L_d}(q_0,q_1) = \frac{\partial L_d}{\partial q_1}(q_0,q_1)dq_1 = M\frac{q_1-q_0}{\Delta t}dq_1.$$

Then

$$\Omega_d = d\Theta^+_{L_d}(q_0,q_1) = d(M\frac{q_1-q_0}{\Delta t}dq_1) = dp_1 \wedge dq_1$$

if we set $p_1 = M(\frac{q_1-q_0}{\Delta t})$

The *Hamiltonian* is a function $H : T^*Q \mapsto \mathbb{R}$, where T^*Q is defined as the *phase space* to be the cotangent bundle ($T^*Q = \{(p,d\phi)|p \in Q, d\phi \in T^*_pQ\}$). Take the local coordinates on T^*Q to be $(\mathbf{q},\mathbf{p})$ and the canonical one-form in coordinates as $\Theta(\mathbf{q},\mathbf{p}) = p_i dq^i$. The *canonical two-form* Ω on T^*Q is defined as

$$\Omega = -d\Theta = dq^i \wedge dp_i.$$

Given a Hamiltonian H, we can define the corresponding *Hamiltonian vector field* X_H to be the unique vector field on T^*Q satisfying

$$i_{X_H}\Omega = dH, \tag{2.11}$$

where i_{X_H} is the interior derivative defined by

$$(i_{X_H}\Omega)_x(v_2,\ldots,v_k) = \Omega_x(X_H(x),v_2,\ldots,v_k).$$

Taking $X_H = (X_q,X_p)$, (2.11) becomes

$$-X_{p_i}dq^i + X_{q^i}dp_i = \frac{\partial H}{\partial q^i}dq^i + \frac{\partial H}{\partial p_i}dp_i,$$

which gives the *Hamiltonian equations* for X_H,

$$\begin{aligned} X_{q^i}(\mathbf{q},\mathbf{p}) &= \frac{\partial H}{\partial p_i}(\mathbf{q},\mathbf{p}), \\ X_{p_i}(\mathbf{q},\mathbf{p}) &= -\frac{\partial H}{\partial p^i}(\mathbf{q},\mathbf{p}). \end{aligned}$$

The *Hamiltonian flow* $F_H : T^*Q \times \mathbb{R} \mapsto T^*Q$ is the flow of the Hamiltonian vector field X_H and for any fixed $t \in \mathbb{R}$, we can prove that the flow map $F^t_H : T^*Q \mapsto T^*Q$ is symplectic. By differentiating and using (2.2), we have

$$\frac{\partial (F^t_H)^*}{\partial t}\Omega|_0 = L_{X_H}\Omega,$$

and from Cartan's 'magic' formula, $L_{X_H}\Omega = di_{X_H}\Omega + i_{X_H}d\Omega = d^2H - i_{X_H}d^2\Theta = 0$ (using the fact that $d^2 = 0$).

In order to relate the Hamiltonian system and Lagrangian system, we define the *Legendre transform* $\mathbb{F}L : TQ \mapsto T^*Q$ (in coordinate form)

$$\mathbb{F}L : (\mathbf{q},\dot{\mathbf{q}}) \mapsto (\mathbf{q},\mathbf{p}) = (\mathbf{q},\frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q},\dot{\mathbf{q}})).$$

To relate the Lagrangian system and Hamiltonian system, define the *Legendre transformation (fibre derivative)* of the Hamiltonian to be the map $\mathbb{F}H : T^*Q \mapsto TQ$ (in coordinate form)

$$\mathbb{F}H : (\mathbf{q},\mathbf{p}) \mapsto (\mathbf{q},\dot{\mathbf{q}}) = (\mathbf{q}, \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q},\mathbf{p})).$$

The relation between the Hamiltonian and Lagrangian is

$$H(\mathbf{q},\mathbf{p}) = \mathbb{F}L(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} - L(\mathbf{q},\dot{\mathbf{q}}). \tag{2.12}$$

The *discrete Legendre transformations* $\mathbb{F}^+L_d, \mathbb{F}^-L_d : Q \times Q \mapsto T^*Q$

$$\begin{aligned} \mathbb{F}^+L_d &: (q_0,q_1) \mapsto (q_1,p_1) = (q_1, D_2L_d(q_0,q_1)) \\ \mathbb{F}^-L_d &: (q_0,q_1) \mapsto (q_0,p_0) = (q_0, -D_1L_d(q_0,q_1)) \end{aligned}$$

are defined to relate the discrete Lagrangian system and Hamiltonian system. The *discrete Lagrangian map* $F_{L_d} : Q \times Q \mapsto Q \times Q$ and the *discrete Hamiltonian map* $\tilde{F}_{L_d} : T^*Q \mapsto T^*Q$ by $\tilde{F}_{L_d} = F^{\pm}_{L_d} \circ F_{L_d} \circ (F^{\pm}_{L_d})^{-1}$ (see below for three expressions). We have the following theorem:

Theorem 4. *[12]. The following diagram commutes.*

$$\begin{tikzcd}
 & (q_0,q_1) \arrow[rr, "FL_d"] \arrow[dl, "\mathbb{F}^-L_d"'] \arrow[dr, "\mathbb{F}^+L_d"] & & (q_1,q_2) \arrow[dl, "\mathbb{F}^-L_d"'] \arrow[dr, "\mathbb{F}^+L_d"] & \\
(q_0,p_0) \arrow[rr, "\tilde{F}L_d"] & & (q_1,p_1) \arrow[rr, "\tilde{F}L_d"] & & (q_2,p_2)
\end{tikzcd}$$

Corollary 1.5.3 [12] shows that the three definitions of the Hamiltonian map

$$\begin{aligned} \tilde{F}_{L_d} &= F^+_{L_d} \circ F_{L_d} \circ (F^+_{L_d})^{-1} \\ \tilde{F}_{L_d} &= F^-_{L_d} \circ F_{L_d} \circ (F^-_{L_d})^{-1} \\ \tilde{F}_{L_d} &= F^+_{L_d} \circ (F^-_{L_d})^{-1} \end{aligned}$$

are equivalent and its coordinate expression is $\tilde{F}_{L_d} : (q_0,p_0) \mapsto (q_1,p_1)$, where

$$p_0 = -D_1L_d(q_0,q_1), \tag{2.13}$$
$$p_1 = D_2L_d(q_0,q_1). \tag{2.14}$$

The above expression is exactly the generating function Type I. of the symplectic case (see Chapter 3). It shows from one side that the discrete Lagrangian is a generating function, which paves the background and has been the inspiration for the main idea for part of [27] (Paper D).

Chapter 3

Volume preserving numerical integrators

> Divergence-free vector fields occur naturally in incompressible fluid dynamics, and preservation of phase-space volume is also a crucial ingredient in many if not all ergodic theorems.
>
> p336, [13]

Summary: In this chapter, we review some volume preserving numerical integrators. In Section 3.1, we begin with one of the earliest volume preserving integrators proposed by Feng and Shang [4]. Since it is difficult to address the general space of divergence-free vector fields, we switch to discussing (smaller) function spaces in Section 3.2, for instance vector fields of polynomials [26] (Paper A), trigonometric polynomials (special cases of the exponential bases or Fourier series) and mixed tensor product bases of monomials and exponentials [30]. In Section 3.3, we review the correction method by Quispel [19]. In Sections 3.4 and 3.5, Shang's [21, 22] generating function method and the methods based on generating forms [10, 27, 28] (including Papers C and D) are discussed.

3.1 Decomposition of divergence-free vector fields into 2D Hamiltonians

In [4], Feng and Shang proposed to decompose the n-dimensional divergence-free vector field into $n-1$ two-dimensional Hamiltonians, and then solved each Hamiltonian with a symplectic difference scheme. It was proved that this approach naturally gives a volume preserving difference scheme. We summarize below.

For any vector field (1.1) subject to the divergence-free condition (2.4), we can rewrite it in the form

$$
\begin{aligned}
\dot{x}_1 &= \partial_{x_2} F^{(1)}(x_1, x_2, \dots, x_n), \\
\dot{x}_2 &= -\partial_{x_1} F^{(1)}(x_1, x_2, \dots, x_n) + \partial_{x_3} F^{(2)}(x_1, x_2, \dots, x_n), \\
&\dots \\
\dot{x}_{n-1} &= -\partial_{x_{n-2}} F^{(n-2)}(x_1, x_2, \dots, x_n) + \partial_{x_n} F^{(n-1)}(x_1, x_2, \dots, x_n), \\
\dot{x}_n &= -\partial_{x_{n-1}} F^{(n-1)}(x_1, x_2, \dots, x_n),
\end{aligned}
\tag{3.1}
$$

where

$$
\begin{aligned}
F^{(1)} &= \int_0^{x_2} a_1 dx_2, \\
F^{(k)} &= \int_0^{x_{k+1}} (a_k + \frac{\partial F^{(k-1)}}{\partial x_{k-1}}) dx_{k+1}, \quad 2 \le k \le n-2, \\
F^{(n-1)} &= \int_0^{x_n} (a_{n-1} + \frac{\partial F^{(n-2)}}{\partial x_{n-2}}) dx_n - \int_0^{x_{n-1}} a_n|_{x_n=0} dx_{n-1}.
\end{aligned}
$$

The above representation of a divergence-free vector field is just one of many possibilities. For instance, McLachlan and Quispel proposed another approach in Appendix A in [16].

The system (3.1) is split into $n-1$ small systems. Each of the systems, for instance, the ith,

$$
\begin{aligned}
\dot{x}_1 &= 0 \\
&\dots \\
\dot{x}_{i-1} &= 0 \\
\dot{x}_i &= \frac{\partial F^{(i)}}{\partial x_{i+1}}, \\
\dot{x}_{i+1} &= -\frac{\partial F^{(i)}}{\partial x_i}, \\
\dot{x}_{i+2} &= 0, \\
&\dots \\
\dot{x}_n &= 0,
\end{aligned}
$$

corresponds to a two-dimensional Hamiltonian system

$$
\begin{aligned}
\frac{dx_i}{dt} &= \frac{\partial H_i}{\partial x_{i+1}} \\
\frac{dx_{i+1}}{dt} &= -\frac{\partial H_i}{\partial x_i}
\end{aligned}
$$

with Hamiltonian $H_i := F^{(i)}$, treating x_j, $j \neq i, i+1$, as fixed (frozen) parameters. One can either solve the 2D-Hamiltonian system exactly or use any symplectic integrator ψ_i. Although ψ_i is not symplectic in the whole space $\mathbb{R}^n$, it is volume preserving. The proof and details can be found in [4].

3.2 Splitting methods for divergence-free polynomials and tensor product bases vector fields

In this section, we consider vector fields in function spaces, for instance, the space of monomials, exponentials and mixed tensor products of these. We begin with the polynomial divergence-free vector field.

An arbitrary polynomial vector field $\mathbf{f}^{\mathbf{vf}}(\mathbf{x})$ can always be decomposed as

$$\mathbf{f}^{\mathbf{vf}}(\mathbf{x}) = \mathbf{f}^{\mathrm{diag}}(\mathbf{x}) + \mathbf{f}^{\mathrm{offdiag}}(\mathbf{x}).$$

We denote $\mathbf{f}^{\mathrm{diag}}$ and $\mathbf{f}^{\mathrm{offdiag}}(\mathbf{x})$ as the *diagonal part* (component-wise, $f_i^{\mathrm{diag}}(\mathbf{x})$ is the collection of terms in $f_i(\mathbf{x})$ that depend on x_i, i.e. $\partial_{x_i} f_i^{\mathrm{diag}}(\mathbf{x}), \neq 0$) and the *off-diagonal* (not diagonal) part of the vector field $\mathbf{f}^{\mathbf{vf}}$, respectively. The *off-diagonal* part can be solved by the technique in [13]. The *diagonal part* of (1.1) subject to the divergence-free condition (2.4) can be split into even smaller divergence-free vector fields associated to a monomial basis $\mathbf{x}^{\mathbf{j}} = x_1^{j_1} \dots x_n^{j_n}$, which has the form

$$\dot{x}_i = \alpha_i x_i \mathbf{x}^{\mathbf{j}}, \qquad i = 1, \dots, n, \qquad \alpha_v = (\alpha_1, \dots, \alpha_n)^T \tag{3.2}$$

with its divergence-free condition

$$\alpha_v^T (\mathbf{j} + \mathbf{1}) = 0, \;\; \mathbf{1} = (1, \dots, 1)^T. \tag{3.3}$$

Each of these divergence-free vector fields (3.2) is called an *elementary divergence-free vector field* (EDFVF). Each EDFVF can be solved analytically [26], which allows for explicit volume preserving methods for arbitrary polynomial divergence-free vector fields, including those with negative degree. Furthermore, the methods can be extended to the case in which the vector field can be expanded as a linear combination of monomials, with arbitrary (real) exponents.

Explicit volume preserving splitting methods for divergence-free ODEs by tensor product bases decompositions, which include the polynomial vector fields and trigonometric vector fields [20] as special cases, was treated in [30], where Zanna considered the EDFVF associated to tensor basis elements $\phi_{\mathbf{j}}(\mathbf{x})$. Assume that the basis function $\phi_{\mathbf{j}}(\mathbf{x})$ has the form,

$$\phi_{\mathbf{j}}(\mathbf{x}) = \phi_{j_1}(x_1) \dots \phi_{j_n}(x_n)$$
$$j_l \in \mathfrak{J}_l, \; \mathbf{j} = (j_1, \dots, j_n) \in \mathfrak{J} = \mathfrak{J}_1 \times \dots \times \mathfrak{J}_n.$$

The choice $\phi_{j_l} = x_l^{j_l}$ and $\phi_{\mathbf{j}} = \mathbf{x}^{\mathbf{j}}$ corresponds to the case of monomial basis for polynomials, which was just discussed, see also [26]. In [30], Zanna also presented the possibility to treat the mixed tensor product bases of the polynomial and exponential case.

The splitting method for the polynomial case including negative degree [26], the exponential case and mixed tensor product of the monomial and exponential basis [30] is an explicit first order method and second order method by symmetrization. Higher order methods can be obtained by composition, for instance using the Yoshida technique [29]. In [25] (Paper B), we discuss how to obtain high order methods via commutators for three kinds of divergence-free vector fields: the monomial basis, exponential basis and tensor product of the monomial and exponential basis.

3.3 Quispel's method

In 1995, Quispel [19] presented a general method for constructing volume preserving integrators up to n dimensions. Consider a map $\mathbf{f}:\mathbb{R}^3 \mapsto \mathbb{R}^3$ defined implicitly by the equations

$$\begin{aligned} x_1 &= f_1(X_1,x_2,x_3), \\ X_2 &= f_2(X_1,x_2,x_3), \\ X_3 &= f_3(X_1,X_2,x_3). \end{aligned} \tag{3.4}$$

If we take the derivatives with respect to x_1,x_2,x_3 for each equation in (3.4), then we obtain

$$\begin{aligned} 1 &= \partial_1 f_1 \tfrac{\partial X_1}{\partial x_1}, \\ 0 &= \partial_1 f_1 \tfrac{\partial X_1}{\partial x_2} + \partial_2 f_1, \\ 0 &= \partial_1 \tfrac{\partial X_1}{\partial x_3} + \partial_3 f_1, \end{aligned} \tag{3.5}$$

$$\begin{aligned} \tfrac{\partial X_2}{\partial x_1} &= \partial_1 f_2 \tfrac{\partial X_1}{\partial x_1}, \\ \tfrac{\partial X_2}{\partial x_2} &= \partial_1 f_2 \tfrac{\partial X_1}{\partial x_2} + \partial_2 f_2, \\ \tfrac{\partial X_2}{\partial x_3} &= \partial_1 f_2 \tfrac{\partial X_1}{\partial x_3} + \partial_3 f_2, \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} \tfrac{\partial X_3}{\partial x_1} &= \partial_1 f_3 \tfrac{\partial X_1}{\partial x_1} + \partial_2 f_3 \tfrac{\partial X_2}{\partial x_1}, \\ \tfrac{\partial X_3}{\partial x_2} &= \partial_1 f_3 \tfrac{\partial X_1}{\partial x_2} + \partial_2 f_3 \tfrac{\partial X_2}{\partial x_2}, \\ \tfrac{\partial X_3}{\partial x_3} &= \partial_1 f_3 \tfrac{\partial X_1}{\partial x_3} + \partial_2 f_3 \tfrac{\partial X_2}{\partial x_3} + \partial_3 f_3. \end{aligned} \tag{3.7}$$

From (3.5), (3.6) and (3.7), we have the matrix equation

$$\begin{pmatrix} \partial_1 f_1 & 0 & 0 \\ \partial_1 f_2 & -1 & 0 \\ \partial_1 f_3 & \partial_2 f_3 & -1 \end{pmatrix} \begin{pmatrix} \frac{\partial X_1}{\partial x_1} & \frac{\partial X_1}{\partial x_2} & \frac{\partial X_1}{\partial x_3} \\ \frac{\partial X_2}{\partial x_1} & \frac{\partial X_2}{\partial x_2} & \frac{\partial X_2}{\partial x_3} \\ \frac{\partial X_3}{\partial x_1} & \frac{\partial X_3}{\partial x_2} & \frac{\partial X_3}{\partial x_3} \end{pmatrix} = \begin{pmatrix} 1 & -\partial_2 f_1 & -\partial_3 f_1 \\ 0 & -\partial_2 f_2 & -\partial_2 f_3 \\ 0 & 0 & -\partial_3 f_3 \end{pmatrix}.$$

Now, the Jacobian determinant

$$\begin{vmatrix} \frac{\partial X_1}{\partial x_1} & \frac{\partial X_1}{\partial x_2} & \frac{\partial X_1}{\partial x_3} \\ \frac{\partial X_2}{\partial x_1} & \frac{\partial X_2}{\partial x_2} & \frac{\partial X_2}{\partial x_3} \\ \frac{\partial X_3}{\partial x_1} & \frac{\partial X_3}{\partial x_2} & \frac{\partial X_3}{\partial x_3} \end{vmatrix} = 1$$

implies that

$$\begin{vmatrix} \partial_1 f_1 & 0 & 0 \\ \partial_1 f_2 & -1 & 0 \\ \partial_1 f_3 & \partial_2 f_3 & -1 \end{vmatrix} = \begin{vmatrix} 1 & -\partial_2 f_1 & -\partial_3 f_1 \\ 0 & -\partial_2 f_2 & -\partial_2 f_3 \\ 0 & 0 & -\partial_3 f_3 \end{vmatrix},$$

that is,

$$\partial_1 f_1(X_1,x_2,x_3) = \partial_2 f_2(X_1,x_2,x_3)\partial_3 f_3(X_1,X_2,x_3). \tag{3.8}$$

The reverse statement is also true, that is, if (3.8) is satisfied, then (3.3) is true. We can see that (3.4) determines f_1 as a function of functions f_2 and f_3 for a volume preserving map. Therefore, (3.4) is volume preserving if and only if (3.8) is satisfied. Consider the differential equation

$$\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{3.9}$$

where $\mathbf{x} \in \mathbb{R}^3$ and $\mathbf{a} : \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{a}(\mathbf{x}) = [a_1(\mathbf{x}), a_2(\mathbf{x}), a_3(\mathbf{x})]^T$, subject to the divergence-free condition

$$\frac{\partial a_1}{\partial x_1} + \frac{\partial a_2}{\partial x_2} + \frac{\partial a_3}{\partial x_3} = 0. \tag{3.10}$$

Consider the first order explicit Euler method to solve (3.9). We obtain

$$\begin{aligned} X_1 &= x_1 + \Delta t a_1(x_1, x_2, x_3), \\ X_2 &= x_2 + \Delta t a_2(x_1, x_2, x_3), \\ X_3 &= x_3 + \Delta t a_3(x_1, x_2, x_3). \end{aligned}$$

To obtain the first order scheme for X_2 and X_3, one can rewrite the second and the third equations according to the map (3.4),

$$\begin{aligned} X_2 &= x_2 + \Delta t a_2(X_1, x_2, x_3) = f_2(X_1, x_2, x_3), \\ X_3 &= x_3 + \Delta t a_3(X_1, X_2, x_3) = f_3(X_1, X_2, x_3). \end{aligned}$$

From (3.8), integrating both sides, we have

$$\begin{aligned} x_1 &= \int^{X_1} \partial_2 f_2(X_1, x_2, x_3) \partial_3 f_3(X_1, X_2, x_3) dX_1 \\ &= \int^{X_1} (1 + \Delta t \frac{\partial a_2}{\partial x_2})(1 + \Delta t \frac{\partial a_3}{\partial x_3}) dX_1 \\ &= X_1 - \Delta t a_1(X_1, x_2, x_3) + \int_{const}^{X_1} [\Delta t \frac{\partial a_3}{\partial x_3}(X_1, X_2, x_3) - \Delta t \frac{\partial a_3}{\partial x_3}(X_1, x_2, x_3) \\ &+ \Delta t^2 \frac{\partial a_2}{\partial x_2}(X_1, x_2, x_3) \frac{\partial a_3}{\partial x_3}(x_1, x_2 + \Delta t a_2(X_1, x_2, x_3), x_3)] dX_1, \end{aligned}$$

where $const = \Delta t a_1(const, x_2, x_3)$ and the third equality is true because of (3.10). Hence, we have

$$X_1 = x_1 + \Delta t a_1(X_1, x_2, x_3) - f_{correct}, \tag{3.11}$$

$f_{correct}$ is a correction form to get the volume preserving integrators corresponding to the divergence free condition,

$$\begin{aligned} f_{correct} &= \int_{const}^{X_1} [\Delta t \frac{\partial a_3}{\partial x_3}(X_1, x_2 + \Delta t a_2(X_1, x_2, x_3), x_3) - \Delta t \frac{\partial a_3}{\partial x_3}(X_1, x_2, x_3) \\ &+ \Delta t^2 \frac{\partial a_2}{\partial x_2}(X_1, x_2, x_3) \frac{\partial a_3}{\partial x_3}(x_1, x_2 + \Delta t a_2(X_1, x_2, x_3), x_3)] dX_1. \end{aligned}$$

So, we obtain the method with the correction term,

$$\begin{aligned} X_1 &= x_1 + \Delta t a_1(X_1, x_2, x_3) - f_{correct}(X_1, x_2, x_3), \\ X_2 &= x_2 + \Delta t a_2(X_1, x_2, x_3), \\ X_3 &= x_3 + \Delta t a_3(X_1, X_2, x_3). \end{aligned} \tag{3.12}$$

This method is called the *correction method* [14]. By construction, we can see that (3.12) is volume preserving.

3.4 Shang's generating function approach

In this subsection, we show the generating function approach for volume preserving mappings proposed by Shang [21, 22]. We consider the linear transformation

$$\begin{pmatrix} \tilde{\mathbf{w}} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{x} \end{pmatrix}. \tag{3.13}$$

Theorem 5. *[3, 21]. Let $\alpha = \begin{pmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{pmatrix} \in GL(2n)$, and denote $\alpha^{-1} = \begin{pmatrix} A^\alpha & B^\alpha \\ C^\alpha & D^\alpha \end{pmatrix}$. Let $\mathbf{g} : \mathbb{R}^n \mapsto \mathbb{R}^n$ be a differentiable volume preserving mapping satisfying the transversality condition*

$$\left| C_\alpha \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}) + D_\alpha \right| \neq 0, \tag{3.14}$$

in a neighborhood of some $\mathbf{x}_0 \in \mathbb{R}^n$. Then there exists a differentiable mapping $\mathbf{f}_\alpha(\mathbf{w}) = \mathbf{f}_{\alpha,\mathbf{g}}(f_{1,\alpha}(\mathbf{w}), f_{2,\alpha}(\mathbf{w}), \ldots, f_{n,\alpha}(\mathbf{w}))^T$ satisfying the condition

$$\left| \frac{\partial \mathbf{f}_\alpha}{\partial \mathbf{w}} C_\alpha - A_\alpha \right| = \left| B_\alpha - \frac{\partial \mathbf{f}_\alpha}{\partial \mathbf{w}}(\mathbf{w}) D_\alpha \right| \neq 0, \tag{3.15}$$

in a neighborhood of the point $\mathbf{w}_0 = C_0 \mathbf{g}(\mathbf{x}_0) + D_\alpha \mathbf{x}_0$ in $\mathbb{R}^n$ such that the mapping $\mathbf{X} = \mathbf{g}(\mathbf{x})$ can be constructed from $\mathbf{f}_\alpha = \mathbf{f}_{\alpha,\mathbf{g}}$ by the relation

$$A_\alpha \mathbf{X} + B_\alpha \mathbf{x} = \mathbf{f}_\alpha(C_\alpha \mathbf{X} + D_\alpha \mathbf{x}), \tag{3.16}$$

in a neighborhood of the point $\mathbf{x}_0$ in R^n.

Remark 3. *Theorem 5 constructs a differential volume preserving map $\mathbf{X} = \mathbf{g}(\mathbf{x})$. Through a transformation, $\tilde{\mathbf{w}}$ and $\mathbf{w}$ are defined by (3.13), that is, $\tilde{\mathbf{w}} = A_\alpha \mathbf{X} + B_\alpha \mathbf{x}$ and $\mathbf{w} = C_\alpha \mathbf{X} + D_\alpha \mathbf{x}$. The inverse transformation map gives us $\mathbf{X} = A^\alpha \tilde{\mathbf{w}} + B^\alpha \mathbf{w}$ and $\mathbf{x} = C^\alpha \tilde{\mathbf{w}} + D^\alpha \mathbf{w}$. Define another map such that $\tilde{\mathbf{w}} = \mathbf{f}_\alpha(\mathbf{w})$, which has the form*

$$A_\alpha \mathbf{X} + B_\alpha \mathbf{x} = \mathbf{f}_\alpha(C_\alpha \mathbf{X} + D_\alpha \mathbf{x}).$$

Since the map $\mathbf{g} : \mathbf{x} \mapsto \mathbf{X}$ is volume preserving, then

$$\left| \frac{\partial \mathbf{X}}{\partial \mathbf{x}} \right| = 1. \tag{3.17}$$

We take derivatives of both sides of (3.16) with respect to $\mathbf{x}$, and get

$$A_\alpha \frac{\partial \mathbf{X}}{\partial \mathbf{x}} + B_\alpha = \frac{\partial \mathbf{f}_\alpha}{\partial \mathbf{w}} (C_\alpha \frac{\partial \mathbf{X}}{\partial \mathbf{x}} + D_\alpha). \tag{3.18}$$

The above equation leads to a transversality condition

$$\left| C_\alpha \frac{\partial \mathbf{X}}{\partial \mathbf{x}} + D_\alpha \right| = \left| C_\alpha \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}) + D_\alpha \right| \neq 0$$

in a neighborhood of the point $\mathbf{x}_0 = C^{\alpha}\mathbf{f}(\mathbf{w}_0) + D^{\alpha}\mathbf{w}_0$. *Then solving with respect to the term* $\frac{\partial \mathbf{X}}{\partial \mathbf{x}}$ *from (3.18), taking determinants of both sides and using condition (3.17), we obtain (3.15).*

Theorem 6. *[21]. Let* α *and* α^{-1} *be defined above. Then* $\mathbf{f}_{\alpha}(\mathbf{w},t)$*, the generating mapping of type* α *of the time-dependent diffeomorphism* $\mathbf{X} = \mathbf{g}(\mathbf{x},t) = \mathbf{g}_{\alpha}^{t}(M_0\mathbf{x})$*, is well-defined. Moreover, it also satisfies the following first order partial differential equation*

$$\frac{\partial \mathbf{f}_{\alpha}}{\partial t} = (A_{\alpha} - \frac{\partial \mathbf{f}_{\alpha}}{\partial \mathbf{w}} C_{\alpha})\mathbf{a}(A^{\alpha}\mathbf{f}_{\alpha} + B^{\alpha}\mathbf{w}), \tag{3.19}$$

with initial condition

$$\mathbf{f}_{\alpha}(\mathbf{w},0) = (A_{\alpha}M_0 + B_{\alpha})(C_{\alpha}M_0 + D_{\alpha})^{-1}\mathbf{w}, \tag{3.20}$$

where $\mathbf{a}$ *is the divergence free vector field (1.1) and* $|C_{\alpha}M_0 + D_{\alpha}| \neq 0$ *for some* $M_0 \in GL(n)$.

Remark 4. *In Theorem 6, we consider the time dependent volume preserving map* $\mathbf{X} = \mathbf{g}(\mathbf{x},t)$*. Moreover, we define a time dependent map* $\mathbf{f}_{\alpha}(\mathbf{w},t)$ *corresponding to the transformation (3.13). We then have a time dependent relation*

$$A_{\alpha}\mathbf{X} + B_{\alpha}\mathbf{x} = \mathbf{f}_{\alpha}(C_{\alpha}\mathbf{X} + D_{\alpha}\mathbf{x},t). \tag{3.21}$$

To obtain (3.19), we take derivatives with respect to t in (3.21).

As we discussed in Theorem 5 and Remark 3, condition (3.15) implies that $\left|\frac{\partial \mathbf{X}}{\partial \mathbf{x}}\right| = 1$. But it is very hard to check (3.15) with an arbitrary matrix α. Therefore, the Hadamard matrix $\alpha = \alpha_{(1,1)}$ was considered by Shang [21, 22] to reduce the complexity of the problem[1]. Here $\alpha_{(1,1)} = \begin{pmatrix} I_n - E_{11} & E_{11} \\ E_{11} & I_n - E_{11} \end{pmatrix}$, and E_{ij} is a matrix with all zero entries except 1 at the i-th row and j-th column. According to Theorem 6, one could have the following corollary.

Corollary 1. *[21]. Assume that the dynamical system is divergence-free, that is,* $\mathbf{a}$ *is a divergence-free vector field. Then the phase flow* $\mathbf{X} = g_{\mathbf{a}}^{t}(\mathbf{x})$ *can be generated from the* $n-1$ *functions* $f_{2,\alpha}, \ldots, f_{n,\alpha}$ *by the relation*

$$\begin{cases} x_1 &= f_{1,\alpha}(X_1, x_2, \ldots, x_n, t), \\ X_k &= f_{k,\alpha}(X_1, x_2, \ldots, x_n, t), \\ k &= 2, \ldots, n, \end{cases} \tag{3.22}$$

where $f_{2,\alpha}, \ldots, f_{n,\alpha}$ *are solutions of the Cauchy problems*

$$\begin{cases} \frac{\partial f_{1,\alpha}}{\partial t} = -a_1(w_1, f_2, \ldots, f_n)\frac{\partial f_{1,\alpha}}{\partial w_1}, \\ \frac{\partial f_{k,\alpha}}{\partial t} = a_k(w_1, f_{2,\alpha}, \ldots, f_{n,\alpha}) - a_1(w_1, f_{2,\alpha}, \ldots, f_{n,\alpha})\frac{\partial f_{k,\alpha}}{\partial w_1}, \\ f_{k,\alpha}(w_1, \ldots, w_n, 0) = w_k, \\ k = 2, \ldots, n. \end{cases} \tag{3.23}$$

[1] Although the linear transformation was considered in both articles [22, 28], Xue and Zanna [28] use simplifying condition Lemma 3.3. Through this process, they derive new volume preserving integrators with parameters, see more details in [28].

Remark 5. *The Hadamard matrix $\alpha_{1,1}$ gives us the transformation*

$$\begin{array}{rclcrcl} \tilde{w}_1 &=& x_1, & \quad & w_1 &=& X_1, \\ \tilde{w}_2 &=& X_2, & & w_2 &=& x_2, \\ & \cdots & & & & \cdots & \\ \tilde{w}_n &=& X_n, & & w_n &=& x_n. \end{array}$$

Then $\tilde{\mathbf{w}} = \mathbf{f}(\mathbf{w})$ implies (3.22). Substituting the Hadamard matrix into (3.19) gives us (3.23). Taking $M_0 = I_n$, the initial condition becomes $\mathbf{f}_\alpha(\mathbf{w}, 0) = \mathbf{w}$. From Theorem 5, we know that $\mathbf{f}$ with initial condition satisfies

$$\frac{\partial f_{1,\alpha}}{\partial w_1} = \left| \frac{\partial(f_{2,\alpha}, \ldots, f_{n,\alpha})}{\partial(w_2, \ldots, w_n)} \right|.$$

Therefore, we have

$$f_{1,\alpha}(\mathbf{w}, t) = w_1 - \int_0^t a_1(w_1, f_{2,\alpha}(\mathbf{w}, \tau), \ldots, f_{n,\alpha}(\mathbf{w}, t)) \left| \frac{\partial(f_{2,\alpha}, \ldots, f_{n,\alpha})}{\partial(w_2, \ldots, w_n)} \right| (\mathbf{w}, \tau) d\tau. \tag{3.24}$$

In order to solve the equations of the Cauchy problems for $f_{k,\alpha}, k = 2, \ldots, n$, assume that the generating map $\mathbf{f}_\alpha(\mathbf{w}, t)$ depends analytically on w and t in some neighborhood of $\mathbb{R}^n$ and for small $|t|$. One could expand the mapping as a power series

$$\mathbf{f}_\alpha(\mathbf{w}, t) = \sum_{k=0}^{\infty} \mathbf{f}_\alpha^{(k)}(\mathbf{w}) t^k. \tag{3.25}$$

In [22], Shang gave an explicit formula of $\mathbf{f}_\alpha^{(k)}$ for all k and constructed the volume preserving numerical scheme for divergence-free systems. Here we just show the truncation up to first order in Theorem 7.

Theorem 7. *[22]. Let the vector field $\mathbf{a}(\mathbf{z})$ depend analytically on $\mathbf{z}$. Then the solution of the Cauchy problem $\mathbf{f}_\alpha(\mathbf{w}, t)$ is expressible as a convergent power series in t for sufficiently small $|t|$ (see (3.25)), with recursively determined coefficients (only up to first order)*

$$\mathbf{f}_\alpha^{(0)}(\mathbf{w}) = N_0 \mathbf{w}, \; N_0 = (A_\alpha + B_\alpha)(C_\alpha + D_\alpha)^{-1}, \tag{3.26}$$

$$\mathbf{f}_\alpha^{(1)}(\mathbf{w}) = L_0 \mathbf{a}(E_0 \mathbf{w}), \; E_0 = (C_\alpha + D_\alpha)^{-1}, \; L_0 = A_\alpha - N_0 C_\alpha. \tag{3.27}$$

Proof. We use (3.25) and differentiate with respect to $\mathbf{w}$ and t respectively. Substituting the corresponding terms into (3.19) and comparing both sides, until first order terms, we obtain the theorem. □

In order to obtain numerical schemes, Shang defined the truncation of $\mathbf{f}_\alpha(\mathbf{w}, t)$ as

$$\phi_i^{(m)} = \sum_{k=0}^{m} f_{i,\alpha}^{(k)}(\mathbf{w}) t^k, \;\; i = 2, \ldots, n,$$

$$\psi_1^{(m)} = \sum_{k=0}^{m} f_{1,\alpha}^{(k)}(\mathbf{w}) t^k.$$

For some fixed value, for instance 0,

$$\phi_1^{(m)}(\mathbf{w},t) = \psi_1^{(m)}(0,w_2,\ldots,w_n,t) + \int_0^{w_1} \left| \frac{\partial(\phi_2^{(m)},\ldots,\phi_n^{(m)})}{\partial(w_2,\ldots,w_n)} \right| (\xi,w_2,\ldots,w_n)d\xi.$$

Remark 6. *Theorem 7 tell us how to solve $f_{k,\alpha}, k=2,n$. Together with equation (3.24), we obtain the whole picture of $\mathbf{f}_\alpha$ (until first order).*

Now, one could easily obtain the first order volume preserving numerical scheme in the following theorem. Moreover, we show that the first order numerical scheme satisfies $\wedge^n \mathbf{X} - \wedge^n \mathbf{x} = 0$, where $\wedge^n = dX_1 \wedge dX_2 \wedge \cdots \wedge dX_n$.

Theorem 8. [2] *Shang's first order numerical scheme [22] is*

$$\begin{aligned} x_1^k &= \phi_1^{(1)}(x_1^{k+1},x_2^k,\ldots,x_n^k,\Delta t), \\ x_i^{k+1} &= \phi_i^{(1)}(x_1^{k+1},x_2^k,\ldots,x_n^k,\Delta t), i=2,\ldots,n, \end{aligned}$$

where

$$\begin{aligned} \phi_1^{(1)}(\mathbf{w},\Delta t) &= -\Delta t a_1(0,w_2,\ldots,w_n) \\ &+ \int_0^{w_1} \begin{vmatrix} 1+\Delta t\frac{\partial a_2}{\partial w_2} & \Delta t\frac{\partial a_2}{\partial w_3} & \ldots & \Delta t\frac{\partial a_2}{\partial w_n} \\ \Delta t\frac{\partial a_3}{\partial w_2} & 1+\Delta t\frac{\partial a_3}{\partial w_3} & \ldots & \Delta t\frac{\partial a_3}{\partial w_n} \\ \ldots & \ldots & \ldots & \ldots \\ \Delta t\frac{\partial a_n}{\partial w_2} & \Delta t\frac{\partial a_n}{\partial w_3} & \ldots & 1+\Delta t\frac{\partial a_n}{\partial w_n} \end{vmatrix} (\xi,w_2,\ldots,w_n)d\xi, \\ \phi_i^{(1)} &= w_i + \Delta t a_i(\mathbf{w}), i=2,\ldots,n. \end{aligned}$$

Observe that the above first order scheme is volume preserving by construction, i.e.

$$dx_1^{k+1} \wedge dx_2^{k+1} \wedge \cdots \wedge dx_n^{k+1} = dx_1^k \wedge dx_2^k \wedge \cdots \wedge dx_n^k. \tag{3.28}$$

Proof. Set

$$p = (p_{i,j}) = \begin{pmatrix} 1+\Delta t\frac{\partial a_2}{\partial w_2} & \Delta t\frac{\partial a_2}{\partial w_3} & \ldots & \Delta t\frac{\partial a_2}{\partial w_n} \\ \Delta t\frac{\partial a_3}{\partial w_2} & 1+\Delta t\frac{\partial a_3}{\partial w_3} & \ldots & \Delta t\frac{\partial a_3}{\partial w_n} \\ \ldots & \ldots & \ldots & \ldots \\ \Delta t\frac{\partial a_n}{\partial w_2} & \Delta t\frac{\partial a_n}{\partial w_3} & \ldots & 1+\Delta t\frac{\partial a_n}{\partial w_n} \end{pmatrix}, i,j=1,\ldots,n-1.$$

Define

$$\delta_{i_1,i_2,\ldots,i_{n-1}} = \begin{cases} 1 & if\ \{i_1,\ldots,i_{n-1}\}\ is\ even, \\ -1 & if\ \{i_1,\ldots,i_{n-1}\}\ is\ odd, \end{cases}$$

where $\{i_1,\ldots,i_{n-1}\}$ is the permutations $\{1,2,\ldots,n-1\}$. Observe that

$$\sum_{i_1,i_2,\ldots,i_{n-1}=1}^{n-1} \delta_{i_1,i_2,\ldots,i_{n-1}} p_{1,i_1}\ldots,p_{n-1,i_{n-1}} = \det p.$$

[2] Identify $x_i^{k+1} := X_i$ and $x_i^k := x_i$ for $i=1,2\ldots,n$

Differentiating on both sides of the numerical scheme, we obtain

$$\begin{aligned}
dx_1^{k+1} &= (\det p)^{-1}(dx_1^k + \Delta t \frac{\partial a_1}{\partial w_2} dx_2^k + \cdots + \Delta t \frac{\partial a_1}{\partial w_n} dx_n^k),\\
dx_2^{k+1} &= dx_2^k + \Delta t \frac{\partial a_2}{\partial w_1} dx_1^{k+1} + \Delta t \frac{\partial a_2}{\partial w_2} dx_2^k + \cdots + \Delta t \frac{\partial a_2}{\partial w_n} dx_n^k,\\
\cdots & \\
dx_n^{k+1} &= dx_n^k + \Delta t \frac{\partial a_n}{\partial w_1} dx_1^{k+1} + \Delta t \frac{\partial a_n}{\partial w_2} dx_2^k + \cdots + \Delta t \frac{\partial a_n}{\partial w_n} dx_n^k.
\end{aligned}$$

Substituting $dx_2^{k+1}, \ldots dx_n^{k+1}$ into $dx_1^{k+1} \wedge dx_2^{k+1} \wedge \cdots \wedge dx_n^{k+1}$, we obtain

$$\begin{aligned}
dx_1^{k+1} \quad \wedge \cdots \wedge \quad & dx_n^{k+1} = dx_1^{k+1} \wedge \left[dx_2^k + \Delta t \frac{\partial a_2}{\partial w_1} dx_1^{k+1} + \Delta t \frac{\partial a_2}{\partial w_2} dx_2^k + \cdots + \Delta t \frac{\partial a_2}{\partial w_n} dx_n^k\right]\\
\wedge \cdots \wedge \quad & \left[dx_n^k + \Delta t \frac{\partial a_n}{\partial w_1} dx_1^{k+1} + \Delta t \frac{\partial a_n}{\partial w_2} dx_2^k + \cdots + \Delta t \frac{\partial a_n}{\partial w_n} dx_n^k\right]\\
= \quad & \sum_{i_1,\ldots,i_{n-1}=1}^{n-1} \delta_{i_1,i_2,\ldots,i_{n-1}} p_{1,i_1} \ldots, p_{n-1,i_{n-1}} dx_1^{k+1} \wedge dx_2^k \wedge \cdots \wedge dx_n^k\\
= \quad & \sum_{i_1,\ldots,i_{n-1}=1}^{n-1} \delta_{i_1,i_2,\ldots,i_{n-1}} p_{1,i_1} \ldots, p_{n-1,i_{n-1}} (\det p)^{-1} dx_1^k \wedge dx_2^k \wedge \cdots \wedge dx_n^k\\
= \quad & dx_1^k \wedge dx_2^k \wedge \cdots \wedge dx_n^k.
\end{aligned}$$

□

3.5 Generating functions and generating forms

In [27, 28] (Papers C and D), Xue et al. present a systematic study on the generating functions (forms) approach for volume preserving mappings, where Definition 14 was used.

3.5.1 The symplectic case

Symplecticity implies volume preservation for Hamiltonian systems. In this subsection, we start with the symplectic case. Two choices of differential one-forms $\nu, \tilde{\nu}$ ($\mathbf{p}^T d\mathbf{q}$ and $-\mathbf{q}^T d\mathbf{p}$) for the canonical symplectic form $\omega = d\mathbf{p} \wedge d\mathbf{q}$ give us four possible combinations. According to Definition 14, for instance, taking $\nu = \tilde{\nu} = \mathbf{p}^T d\mathbf{q}$ (top left in Table 3.1) we have

$$\mathbf{f}^* \tilde{\nu} - \nu = \mathbf{P}^T d\mathbf{Q} - \mathbf{p}^T d\mathbf{q} = dS(\mathbf{p}, \mathbf{q}, \mathbf{P}, \mathbf{Q}).$$

Comparing both sides of the above equation, we have $S_{\mathbf{p}} = 0$ and $S_{\mathbf{P}} = 0$, which implies that the generating function S depends only on the variables $\mathbf{q}$ and $\mathbf{Q}$, that is, $S = S(\mathbf{q}, \mathbf{Q})$. Moreover, we have the following relations:

$$\begin{aligned}
\mathbf{P} &= S_{\mathbf{Q}}(\mathbf{q}, \mathbf{Q}),\\
\mathbf{p} &= -S_{\mathbf{q}}(\mathbf{q}, \mathbf{Q}).
\end{aligned} \tag{3.29}$$

Assuming that $\mathbf{p}$ and $\mathbf{q}$ are known, the second equation of (3.29) generates $\mathbf{Q}$ if $\frac{\partial \mathbf{p}}{\partial \mathbf{Q}} \neq 0$, and this is called the *twist equation*. A similar procedure gives us all other cases in Table 3.1.

$\mathbf{f}^*\tilde{\nu} \downarrow \quad \nu \rightarrow$	$\mathbf{p}^T d\mathbf{q}$	$-\mathbf{q}^T d\mathbf{p}$
$\mathbf{P}^T d\mathbf{Q}$	I. $S(\mathbf{q},\mathbf{Q})$ $\mathbf{P} = \partial_{\mathbf{Q}} S(\mathbf{q},\mathbf{Q})$ $\mathbf{p} = -\partial_{\mathbf{q}} S(\mathbf{q},\mathbf{Q})$ $\frac{\partial \mathbf{p}}{\partial \mathbf{Q}} \neq 0$	II. $S(\mathbf{p},\mathbf{Q})$ $\mathbf{P} = \partial_{\mathbf{Q}} S(\mathbf{p},\mathbf{Q})$ $\mathbf{q} = \partial_{\mathbf{p}} S(\mathbf{p},\mathbf{Q})$ $\frac{\partial \mathbf{q}}{\partial \mathbf{Q}} \neq 0$
$-\mathbf{Q}^T d\mathbf{P}$	III. $S(\mathbf{q},\mathbf{P})$ $\mathbf{Q} = \partial_{\mathbf{P}} S(\mathbf{q},\mathbf{P})$ $\mathbf{p} = \partial_{\mathbf{q}} S(\mathbf{q},\mathbf{P})$ $\frac{\partial \mathbf{p}}{\partial \mathbf{P}} \neq 0$	IV. $S(\mathbf{p},\mathbf{P})$ $\mathbf{Q} = \partial_{\mathbf{P}} S(\mathbf{p},\mathbf{P})$ $\mathbf{q} = -\partial_{\mathbf{p}} S(\mathbf{p},\mathbf{P})$ $\frac{\partial \mathbf{q}}{\partial \mathbf{P}} \neq 0$

Table 3.1: The four classical types of generating functions for the canonical symplectic form $\omega = d\mathbf{p} \wedge d\mathbf{q}$.

We consider the Hamiltonian system $H(\mathbf{p},\mathbf{q})$[3],

$$
\begin{aligned}
\dot{\mathbf{p}} &= -\nabla_{\mathbf{q}} H(\mathbf{p},\mathbf{q}), \\
\dot{\mathbf{q}} &= \nabla_{\mathbf{p}} H(\mathbf{p},\mathbf{q}).
\end{aligned} \tag{3.30}
$$

If we choose $S(\mathbf{p},\mathbf{Q}) = \mathbf{pQ} - S^2(\mathbf{p},\mathbf{Q})$, the equations of Type II. (top right corner of Table 3.1) becomes

$$
\begin{aligned}
\mathbf{P} &= \mathbf{p} - S^2_{\mathbf{Q}}(\mathbf{p},\mathbf{Q}), \\
\mathbf{Q} &= \mathbf{q} + S^2_{\mathbf{p}}(\mathbf{p},\mathbf{Q}).
\end{aligned}
$$

It is easy to see that by choosing $S^2 = \Delta t H$, identifying $\mathbf{P} = \mathbf{p}^{n+1}$, $\mathbf{p} = \mathbf{p}^n$, $\mathbf{Q} = \mathbf{q}^{n+1}$ and $\mathbf{q} = \mathbf{q}^n$, we obtain the Symplectic Euler-B method. For Type III., by setting

$$
S(\mathbf{q},\mathbf{P}) = \mathbf{qP} + S^3(\mathbf{q},\mathbf{P}),
$$

choosing $S^3 = \Delta t H$, identifying $\mathbf{P} = \mathbf{p}^{n+1}$, $\mathbf{p} = \mathbf{p}^n$, $\mathbf{Q} = \mathbf{q}^{n+1}$ and $\mathbf{q} = \mathbf{q}^n$, we obtain the Symplectic Euler-A method.

By change of variables through a linear transformation, one can get the Symplectic Theta method through Type II. or Type III.. Details can be found in [28].

Theorem 9. *[28]. For any choice of θ, the scheme*

$$
\begin{aligned}
\mathbf{P} &= \mathbf{p} + \partial_2 S(\theta\mathbf{P} + (1-\theta)\mathbf{p}, (1-\theta)\mathbf{Q} + \theta\mathbf{q}) \\
\mathbf{Q} &= \mathbf{q} - \partial_1 S(\theta\mathbf{P} + (1-\theta)\mathbf{p}, (1-\theta)\mathbf{Q} + \theta\mathbf{q})
\end{aligned} \tag{3.31}
$$

is symplectic. Moreover, letting $S = -\Delta t H$, the choices $\theta = \frac{1}{2}, 1, 0$ yield the Implicit Midpoint Rule method, the Symplectic Euler-A method and Symplectic Euler-B method, respectively.

[3] We follow the notation in [3]

The conjugate method of (3.31) is given in Appendix B, where the second order method is constructed by symmetrization. Unlike the Theta method (see Remark 8 in Appendix B), which is symplectic if and only if $\theta = \frac{1}{2}$, the Symplectic Theta method works for all values of θ. Numerical experiments by solving the Harmonic oscillator problem [9] can be found in Figure B.1 in Appendix B. Furthermore, it is shown in Appendix C that the Symplectic Theta scheme (3.31) is a one-stage partitioned Runge-Kutta method. Moreover, we have the following theorem, see also more discussions in Appendix C.

Theorem 10. *For any Hamiltonian system $H(\mathbf{p},\mathbf{q})$, if we apply one-stage partitioned RK method formed of $(\theta_1, 1)$ to the variable $\mathbf{p}$ and $(\theta_2, 1)$ to the variable $\mathbf{q}$, the numerical scheme*

$$\begin{aligned}\mathbf{P} &= \mathbf{p} - \partial_2 H(\theta_1\mathbf{P} + (1-\theta_1)\mathbf{p}, (1-\theta_2)\mathbf{q} + \theta_2\mathbf{Q})\\ \mathbf{Q} &= \mathbf{q} + \partial_1 H(\theta_1\mathbf{P} + (1-\theta_1)\mathbf{p}, (1-\theta_2)\mathbf{q} + \theta_2\mathbf{Q})\end{aligned} \tag{3.32}$$

is volume preserving if and only if $\theta_1 = 1 - \theta_2$.

One can solve Type I. and IV. by the technique in [12], see also Section 2.4 for more details. Given a Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}}, t) = L(q_1, \dots, q_d, \dot{q}_1, \dots, \dot{q}_d, t)$ for the configuration space $\mathbb{R}^d$ with coordinates $q_i, i = 1, \dots, d$, if we take the discrete Lagrangian $L_d(q_k, q_{k+1}, \Delta t)$ ($\mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$) as the generating function for Type I., we obtain

$$\begin{aligned}p_k &= -D_1 L_d(q_k, q_{k+1}, \Delta t),\\ p_{k+1} &= D_2 L_d(q_k, q_{k+1}, \Delta t),\end{aligned} \tag{3.33}$$

while changing the position of $\mathbf{p}$ with $\mathbf{q}$ and $\mathbf{P}$ with $\mathbf{Q}$, we obtain the discrete scheme for Type IV.. When we choose different forms of the discrete Lagrangians L_d, equations (3.33) cover the Symplectic Euler-A, Symplectic Euler-B and Implicit Midpoint Rule methods.

Proposition 4. *In (3.33), when we set the discrete Lagrangian as*

$$L_d^0(q_k, q_{k+1}, \Delta t) = \Delta t L(q_k, \frac{q_{k+1} - q_k}{\Delta t}),$$

we obtain the Symplectic Euler-A method; when set the discrete Lagrangian as

$$L_d^1(q_k, q_{k+1}, \Delta t) = \Delta t L(q_{k+1}, \frac{q_{k+1} - q_k}{\Delta t}),$$

we obtain the Symplectic Euler-B method; when set the discrete Lagrangian as

$$L_d^{\frac{1}{2}}(q_k, q_{k+1}, \Delta t) = \Delta t L(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{\Delta t}),$$

we obtain the Implicit Midpoint Rule.

3.5.2 The volume case in $\mathbb{R}^3$

We start with one-forms $\nu = \tilde{\nu} = x_3 dx_1 \wedge dx_2$. This is denoted as the (123,123) case. Let (123,123) A-C denote the top left case of Table 3.2, and similarly (123,123) B-C denote the top right case and so on.

(123,123)	Adx_1	Bdx_2
CdX_1	$\lambda = A(x_1,x_2,X_1)dx_1 + C(x_1,X_1,X_2)dX_1$ $x_3 = \partial_{x_2} A$ $\partial_{X_1} A = \partial_{x_1} C$ $X_3 = -\partial_{X_2} C$ $\frac{\partial X_1}{\partial x_3} \neq 0, \quad \frac{\partial x_1}{\partial X_3} \neq 0$	$\lambda = B(x_1,x_2,X_1)dx_2 + C(x_2,X_1,X_2)dX_1$ $x_3 = -\partial_{x_1} B$ $\partial_{X_1} B = \partial_{x_2} C$ $X_3 = -\partial_{X_2} C$ $\frac{\partial X_1}{\partial x_3} \neq 0, \quad \frac{\partial x_2}{\partial X_3} \neq 0$
DdX_2	$\lambda = A(x_1,x_2,X_2)dx_1 + D(x_1,X_1,X_2)dX_2$ $x_3 = \partial_{x_2} A$ $\partial_{X_2} A = \partial_{x_1} D$ $X_3 = \partial_{X_1} D$ $\frac{\partial X_2}{\partial x_3} \neq 0, \quad \frac{\partial x_1}{\partial X_3} \neq 0$	$\lambda = B(x_1,x_2,X_2)dx_2 + D(x_2,X_1,X_2)dX_2$ $x_3 = -\partial_{x_1} B$ $\partial_{X_2} B = \partial_{x_2} D$ $X_3 = \partial_{X_1} D$ $\frac{\partial X_2}{\partial x_3} \neq 0, \quad \frac{\partial x_2}{\partial X_3} \neq 0$

Table 3.2: The four basic types of generating 1-forms λ for $\nu = \tilde{\nu} = x_3 dx_1 \wedge dx_2$, adapted from [10]. All the other tables are obtained by applying cyclic even permutations to the variables (x_1,x_2,x_3) and (X_1,X_2,X_3).

We take the (123,123) A-C case for instance to deduce the generating form λ. For this case, choosing $\nu = \tilde{\nu} = x_3 dx_1 \wedge dx_2$, from Definition 14, we obtain

$$\mathbf{f}^* \tilde{\nu} - \nu = d(A(x_1,x_2,x_3,X_1,X_2,X_3)dx_1 + C(x_1,x_2,x_3,X_1,X_2,X_3)dX_1).$$

Noticing that A and C depend only on the variables x_1,x_2,X_1 and x_1,X_1,X_2, respectively, we have

$$X_3 dX_1 \wedge dX_2 - x_3 dx_1 \wedge dx_2 = C_{X_2} dX_2 \wedge dX_2 + (C_{x_1} - A_{X_1}) dx_1 \wedge dX_1 + A_{x_2} dx_2 \wedge dx_1.$$

Comparing both sides, we have the following equations:

$$\begin{aligned} x_3 &= \partial_{x_2} A(x_1,x_2,X_1), \\ \partial_{X_1} A &= \partial_{x_1} C, \\ X_3 &= -\partial_{X_2} C(x_1,X_1,x_2). \end{aligned}$$

In order to solve the second equation (the compatibility equation), the *twist conditions* $\frac{\partial X_1}{\partial x_3} \neq 0, \quad \frac{\partial x_1}{\partial X_3} \neq 0$ are needed. We obtain the other cases in Table 3.2 similarly.

Remark 7. *In $\mathbb{R}^3$, ν can be chosen from $x_3 dx_1 \wedge dx_2, x_2 dx_3 \wedge dx_1$ and $x_1 dx_2 \wedge dx_3$, and so can $\tilde{\nu}$. For each choice of $(\nu, \tilde{\nu})$, there are four cases, like Table 3.2. Overall, there are thirty-six generating one-forms in $\mathbb{R}^3$.*

According to Remark 7, we have all together 9 tables like Table 3.2 for thirty-six generating one-forms in $\mathbb{R}^3$. The other tables are obtained by applying cyclic even permutations to the variables (x_1, x_2, x_3) and (X_1, X_2, X_3). For instance, we take $\nu = x_2 dx_3 \wedge dx_1$ and $\tilde{\nu} = x_1 dx_2 \wedge dx_3$. We then obtain a table of the (312,231) case by changing the variables from $(x_1, x_2, x_3) \mapsto (x_3, x_1, x_2)$ (use 312 to denote the permutation for lowercase variables, see the top labels in Figure 3.1) and $(X_1, X_2, X_3) \mapsto (X_2, X_3, X_1)$ (use 231 to denote the permutation for uppercase variables, see the left labels in Figure 3.1). Putting all the 9 tables together, we obtain the Figure 3.1.

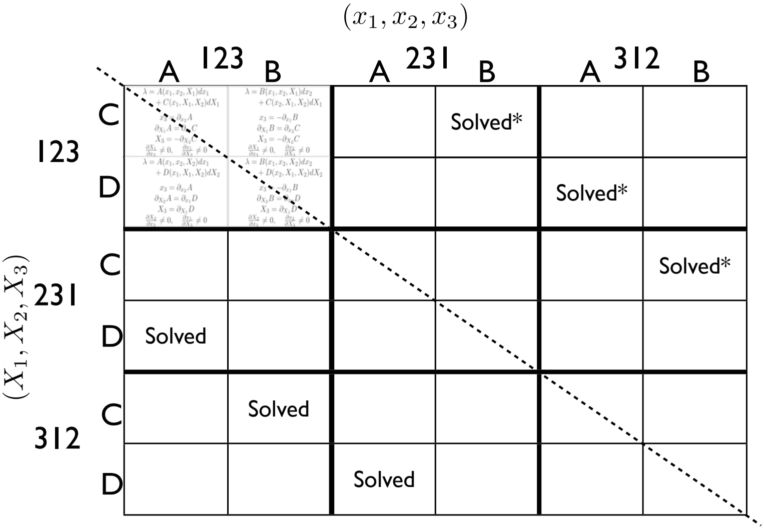

Figure 3.1: The overall view of the thirty-six cases of generating one-forms for volume preserving maps in $\mathbb{R}^3$. The top left 4×4 box is the same as in Table 3.2. The labels 123,231,312 correspond to permutations of variables. The table has diagonal symmetry corresponding to the *adjoint* map. The 'solved' cases are those identified in [28]. The table is taken from [27].

In [28], Xue and Zanna give explicit formulas for generating functions for the (231,123) B-C case, (312,231) B-C case and (312,123) A-D case and their conjugate cases (marked 'solved' in Figure 3.1). Here, we demonstrate the (231,123) B-C case as an example. For a divergence-free vector field[4] written as

$$\begin{aligned} \dot{x}_1 &= \partial_{x_2} F^{(1)}(x_1, x_2, x_3), \\ \dot{x}_2 &= -\partial_{x_1} F^{(1)}(x_1, x_2, x_3) + \partial_{x_3} F^{(2)}(x_1, x_2, x_3), \\ \dot{x}_3 &= -\partial_{x_2} F^{(2)}(x_1, x_2, x_3). \end{aligned} \tag{3.34}$$

the one-form is $\lambda = B(X_1, x_2, x_3)dx_3 + C(X_1, X_2, x_3)dX_1$ and B, C satisfy

$$\begin{aligned} x_1 &= -\partial_{x_2} B(X_1, x_2, x_3), \\ \partial_{X_1} B &= \partial_{x_3} C, \\ X_3 &= -\partial_{X_2} C(X_1, X_2, x_3), \end{aligned} \tag{3.35}$$

[4]See Remark 5 in [28] for more details about different representations of the vector field (1.1) subject to the divergence-free condition (2.4).

with the twist conditions

$$\frac{\partial X_1}{\partial x_1} \neq 0, \quad \frac{\partial x_3}{\partial X_3} \neq 0.$$

Assuming that $\frac{\partial X_1}{\partial x_1} = 1 + O(\Delta t)$, $\frac{\partial x_3}{\partial X_3} = 1 + O(\Delta t)$, choosing $B = -X_1 x_2 + \Delta t F^{(1)}(X_1, x_2, x_3)$ and $C = -X_2 x_3 + \Delta t F^{(2)}(X_1, X_2, x_3)$, and following the procedure in [28], we obtain the first order volume preserving scheme

$$\begin{aligned}
X_1 &= x_1 + \Delta t \partial_{x_2} F^{(1)}(X_1, x_2, x_3), \\
X_2 &= x_2 - \Delta t \partial_{X_1} F^{(1)}(X_1, x_2, x_3) + \Delta t \partial_{x_3} F^{(2)}(X_1, X_2, x_3), \\
X_3 &= x_3 - \Delta t \partial_{X_2} F^{(2)}(X_1, X_2, x_3).
\end{aligned} \tag{3.36}$$

In [27], Xue, Verdier and Zanna analyze the remaining volume preserving forms for $\mathbb{R}^3$. We identify cases which are geometrically different in the sense that they cannot be obtained from others by permutations of the variables. It is shown that there are only six geometrically different generating one-forms in $\mathbb{R}^3$, and they correspond to the first column in Figure 3.1. The six geometrically different cases are named Type SE+SE, Type DL+DL, Type DL+SE, Type SE+DL, Type S1 and S2. The six cases marked 'solved' in Figure 3.1 and identified in [28] belong to Type SE+SE.

Type S1 and S2 are novel cases, and their general understanding is still under progress. The remaining four cases are solved by splitting the 3D divergence-free vector field into two 2D Hamiltonian systems, while the Hamiltonian systems can be understood by looking into the corresponding Lagrangian systems. One can use the Symplectic Euler (SE) method to solve the Hamiltonian system and discrete Lagrangian (DL) to solve the Lagrangian system.

As an example, we show how the (231,123) A-C case solves a divergence-free vector field under the proper representation. From Figure 3.1, we can see that the (231,123) A-C case (Type SE+DL) is the adjoint 'map' of the (123,231) A-C case (Type DL+SE). Here, we refer to [27], where the geometric structure of the first column of Figure 3.1 was discussed, and by considering the symmetry, we do not distinguish Type SE+DL and Type DL+SE.

For the (231,123) A-C case, we have[5]

$$\begin{aligned}
x_1 &= \partial_{x_3} A, \\
\partial_{X_1} A &= \partial_{x_2} C, \\
X_3 &= -\partial_{X_2} C,
\end{aligned} \tag{3.37}$$

where $\lambda = A(X_1, x_2, x_3) dx_2 + C(X_1, X_2, x_2) dX_1$ and the twist conditions are $\frac{\partial X_1}{\partial x_1} \neq 0$, $\frac{\partial x_2}{\partial X_3} \neq 0$. In this case, we consider the divergence-free vector field

$$\begin{aligned}
\dot{x}_1 &= -\partial_{x_3} F^{(1)}(x_1, x_2, x_3), \\
\dot{x}_2 &= \partial_{x_3} F^{(2)}(x_1, x_2, x_3), \\
\dot{x}_3 &= \partial_{x_1} F^{(1)}(x_1, x_2, x_3) - \partial_{x_2} F^{(2)}(x_1, x_2, x_3).
\end{aligned} \tag{3.38}$$

[5] In [27], all three-term equations, for instance (3.35) and (3.37), can be written in a unified manner by defining two bijections $\sum$ and σ.

From one of the twist conditions, $\frac{\partial X_1}{\partial x_1} \neq 0$, we assume that $\frac{\partial X_1}{\partial x_1} = 1 + O(\Delta t)$. Then taking $A(X_1, x_2, x_3) = x_3 X_1 + \Delta t F^{(1)}(X_1, x_2, x_3)$, we obtain

$$\begin{aligned} X_1 &= x_1 - \Delta t \partial_{x_3} F^{(1)}(X_1, x_2, x_3), \\ \partial_{X_1} A &= x_3 + \Delta t \partial_{X_1} F^{(1)}(X_1, x_2, x_3). \end{aligned}$$

For the next step, take $L = X_3 \dot{x_2} - F^{(2)}(X_1, x_2, X_3)$ and $C = -\Delta t L_d(X_1, x_2, X_3, (X_2 - x_2)/\Delta t)$, and obtain

$$\begin{aligned} X_2 &= x_2 + \Delta t \partial_{x_3} F^{(2)}(X_1, x_2, X_3), \\ \partial_{x_2} C &= X_3 + \Delta t \partial_{x_2} F^{(2)}(X_1, x_2, X_3). \end{aligned}$$

Overall, for the (231,123) A-C Case, we obtain the numerical scheme

$$\begin{aligned} X_1 &= x_1 - \Delta t \partial_{x_3} F^{(1)}(X_1, x_2, x_3), \\ X_2 &= x_2 + \Delta t \partial_{x_3} F^{(2)}(X_1, x_2, X_3), \\ X_3 &= x_3 + \Delta t \partial_{X_1} F^{(1)}(X_1, x_2, x_3) - \Delta t \partial_{X_2} F^{(2)}(X_1, x_2, X_3). \end{aligned} \tag{3.39}$$

Chapter 4

Introduction to the papers

In Paper A we derive splitting methods for polynomial divergence-free vector fields. In Paper B we construct high order methods for three kinds of divergence-free vector fields via commutators. In Papers C and D we get a slightly different view of the volume preserving maps through the generating functions and generating forms approach.

Paper A: *Huiyan Xue and Antonella Zanna. Explicit volume-preserving splitting methods for polynomial divergence-free vector fields. BIT Numerical Mathematics, Volume 53, Issue 1, pp. 265-281, March 2013.*

In this paper, we study volume preserving integrators for polynomial divergence-free vector fields. First, we decompose the polynomial divergence-vector field into a *diagonal* and an *off-diagonal* part. The *off-diagonal part* can be treated using the techniques in [13]. For the *diagonal part*, the main idea is to decompose the divergence polynomial by means of an appropriate basis for polynomials: the monomial basis. For each monomial basis function, the split fields are identified by collecting the appropriate terms in the vector field so that each split vector field is volume preserving. Associated to the monomial basis $\mathbf{x}^{\mathbf{j}} = x_1^{j_1} \dots x_n^{j_n}$, we have the form

$$\dot{x}_i = \alpha_i x_i \mathbf{x}^{\mathbf{j}},\ \alpha_v = (\alpha_1, \dots, \alpha_n)^T,$$

with the divergence-free condition

$$\alpha_v(\mathbf{j} + \mathbf{1}) = 0,\ \mathbf{1} = (1, \dots, 1)^T.$$

We show that each split field can be integrated exactly by analytical methods. Thus, the composition yields a volume preserving numerical method. The procedure can be extended to negative degrees.

My contribution to this paper was to implement all the numerical examples described in the paper and to write part of the paper, mainly Section 6. Furthermore, I contributed to the discussion of the numerical schemes and presented the results in the international conferences: The Tasmanian Rigorous Analysis and Geometric Integration Conference 2010 (Launceston, Australia) and Foundations of Computational Mathematics conference 2011 (Budapest, Hungary).

Paper B: *Huiyan Xue. High order volume preserving integrators for some divergence-free vector fields via commutators. Preprint, 2013.*

In this paper, we consider three kinds of divergence-free vector fields: the monomial basis, the exponential basis and the tensor product of the monomial basis and exponential basis. We find that for any two *elementary divergence-free vector fields* (EDFVFs) A and B, their commutator is still a divergence-free vector field of the same kind, i.e. $\nabla \cdot [A,B] = 0$. Using the BCH form, we have the fourth order method constructed by

$$S_4(\tau) = e^{\frac{\Delta t^3}{48}[A,[A,B]]} e^{-\frac{\Delta t^3}{24}[B,[B,A]]} e^{\frac{\Delta t}{2}A} e^{\Delta t B} e^{\frac{\Delta t}{2}A} e^{-\frac{\Delta t^3}{24}[B,[B,A]]} e^{\frac{\Delta t^3}{48}[A,[A,B]]},$$

which is obtained using the leapfrog second order method. A more accurate fourth order method is also derived by increasing the number of stages of the second order schemes. Moreover, we consider the ordering of the EDFVFs and their multi-commutators in order to reduce the error of the schemes, illustrating by numerical tests that the strategies in [31] work well.

Paper C: *Huiyan Xue and Antonella Zanna. Generating functions and volume preserving mappings. To appear: Discrete and Continuous Dynamics System-Series A, Volume 34, Number 3, pp. 1229-1249, March 2014. Publish online* doi:10.3934/dcds.2014.34.1229.

In this paper, we study generating forms and functions for volume-preserving mappings in $\mathbb{R}^n$. We derive some parametric classes of volume preserving numerical schemes for divergence-free vector fields. In passing, by extension of the Poincaré generating function and a change of variables, we obtain a symplectic equivalent of the theta-method for differential equations,

$$\begin{aligned}\mathbf{P} &= \mathbf{p} + \partial_2 S(\theta\mathbf{P} + (1-\theta)\mathbf{p}, (1-\theta)\mathbf{Q} + \theta\mathbf{q}),\\ \mathbf{Q} &= \mathbf{q} - \partial_1 S(\theta\mathbf{P} + (1-\theta)\mathbf{p}, (1-\theta)\mathbf{Q} + \theta\mathbf{q}),\end{aligned}$$

which includes the Implicit Midpoint Rule and Symplectic Euler A and B methods as special cases.

Based on feedback from the co-author, I wrote most of the paper. I contributed part of the ideas and was involved in most of the development. Moreover, I presented the results at the MAGIC workshop in Geilo, 2013.

Paper D: *Huiyan Xue. Olivier Verdier and Antonella Zanna, A study on volume preserving generating forms in* $\mathbb{R}^3$. *Preprint, 2013.*

For the volume case in $\mathbb{R}^3$, there are 36 one-forms that generate the volume preserving map. In [28], Xue and Zanna show how to construct numerical integrators from some of them. In this paper, we study the remaining cases. We notice that among the thirty-six cases, only six are geometrically different. Moreover, we present how to use these six cases to study the volume preserving schemes for divergence-free vector fields in $\mathbb{R}^3$ under proper representations.

For this paper I took part in the discussion with the co-authors, provided part of the idea and was involved in most of the development of the paper. Furthermore, I did the calculations for the main results and wrote a large part of the paper based on feedback from the co-authors.

Appendix A

The ODE $\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x})$ is divergence-free if

$$\nabla \cdot \mathbf{a} = \sum_{i=1}^{n} \frac{\partial a_i}{\partial x_i} = 0,$$

and is diagonal-free if

$$\frac{\partial a_i}{\partial x_i} = 0, \text{ for all } i.$$

Lemma 1. *[16] Any diagonal-free vector field is divergence-free. Each component*

$$V_i : \dot{x}_i = a_i(\mathbf{x}), \; \dot{x}_j = 0, \; j \neq i,$$

is integrated exactly by Euler method. Overall, combining each component gives us a volume preserving method.

The ABC system (1.3) is an example of a diagonal-free vector field. According to Lemma 1, we have the following splitting:

$$V_1: \begin{array}{rcl} \dot{x} &=& A\sin z + C\cos y, \\ \dot{y} &=& 0, \\ \dot{z} &=& 0, \end{array} \quad V_2: \begin{array}{rcl} \dot{x} &=& 0, \\ \dot{y} &=& B\sin x + A\cos z, \\ \dot{z} &=& 0, \end{array} V_3: \begin{array}{rcl} \dot{x} &=& 0, \\ \dot{y} &=& 0, \\ \dot{z} &=& C\sin y + B\cos x. \end{array}$$

We solve each vector field V_i $(i = 1,2,3)$ by the Euler method, denoted by $\phi_{V_i,h}$. In Figure 1.1, we use the second order method obtained by the composition $\phi_{V_1,h/2} \circ \phi_{V_2,h/2} \circ \phi_{V_3,h} \circ \phi_{V_2,h/2} \circ \phi_{V_1,h/2}$. The MATLAB code is as below.

```
A=1;
B=1;
C=2;
h=0.1;
T=750;
N=T/h;
Xnew=zeros(1,3);
Xold=[2,5,0];
xnewall=zeros(N,3);
for i=1:N
   %V1
   Xnew(1)=Xold(1)+(A*sin(Xold(3))+C*cos(Xold(2)))*h/2;
```

```
    Xold(1)=Xnew(1);
    %V2
    Xnew(2)=Xold(2)+(B*sin(Xold(1))+A*cos(Xold(3)))*h/2;
    Xold(2)=Xnew(2);
    %V3
    Xnew(3)=Xold(3)+(C*sin(Xold(2))+B*cos(Xold(1)))*h;
    Xold(3)=Xnew(3);
    %V2
    Xnew(2)=Xold(2)+(B*sin(Xold(1))+A*cos(Xold(3)))*h/2;
    Xold(2)=Xnew(2);
    %V1
    Xnew(1)=Xold(1)+(A*sin(Xold(3))+C*cos(Xold(2)))*h/2;
    Xold(1)=Xnew(1);
    xnewall(i,:)=Xold;
end
plot(wrapTo2pi(xnewall(:,1)),wrapTo2pi(xnewall(:,2)),'k');
```

Appendix B

In this appendix, we give the conjugate Symplectic Theta method and construct the second order method by symmetrization. Moreover, we test the Symplectic Theta method by a numerical example.

Lemma 2. *[Conjugate Symplectic Theta method] The scheme*

$$\begin{aligned}\mathbf{P} &= \mathbf{p} + \partial_2 S(\theta\mathbf{p} + (1-\theta)\mathbf{P}, (1-\theta)\mathbf{q} + \theta\mathbf{Q}),\\ \mathbf{Q} &= \mathbf{q} - \partial_1 S(\theta\mathbf{p} + (1-\theta)\mathbf{P}, (1-\theta)\mathbf{q} + \theta\mathbf{Q}),\end{aligned} \tag{B.1}$$

is the conjugate method of the scheme (3.31) and is symplectic. Setting $S = -\Delta t H$, we have the second order method

$$\begin{aligned}\bar{\mathbf{p}} &= \mathbf{p} - \Delta t/2\partial_2 H(\theta\bar{\mathbf{p}} + (1-\theta)\mathbf{p}, (1-\theta)\bar{\mathbf{q}} + \theta\mathbf{q}),\\ \bar{\mathbf{q}} &= \mathbf{q} + \Delta t/2\partial_1 H((\theta\bar{\mathbf{p}} + (1-\theta)\mathbf{p}, (1-\theta)\bar{\mathbf{q}} + \theta\mathbf{q}),\\ \mathbf{Q} &= \bar{\mathbf{q}} + \Delta t/2\partial_1 H((1-\theta)\mathbf{P} + \theta\bar{\mathbf{p}}, \theta\mathbf{Q} + (1-\theta)\bar{\mathbf{q}}),\\ \mathbf{P} &= \bar{\mathbf{p}} - \Delta t/2\partial_2 H((1-\theta)\mathbf{P} + \theta\bar{\mathbf{p}}, \theta\mathbf{Q} + (1-\theta)\bar{\mathbf{q}}).\end{aligned} \tag{B.2}$$

Proof. Denote the Symplectic Theta method by $\Psi_{\Delta t}$. Then the conjugate method $\Psi^*_{\Delta t}$ is obtained by exchanging $\mathbf{p} \mapsto \mathbf{P}$, $\mathbf{q} \mapsto \mathbf{Q}$, and substituting $-\Delta t$ for Δt. The scheme (B.2) is obtained by $\Psi^*_{\Delta t/2} \circ \Psi_{\Delta t/2}$, which is obviously of second order. □

Remark 8. *The Theta method*

$$z_{n+1} = z_n + \Delta t f(\theta z_n + (1-\theta) z_{n+1}), \tag{B.3}$$

for the ordinary differential equation $\frac{d}{dt}z = f(z)$, is symplectic and of second order if $\theta = \frac{1}{2}$. Otherwise, it is not symplectic and only of first order.

Proof. Let us consider the one-dimensional ordinary differential equation.

$$\dot{z} = f(z).$$

Integrating both sides of the above equation on a small interval $[t, t+\Delta t]$, we obtain

$$z(t+\Delta t) - z(t) = \int_0^{\Delta t} f(z(t+\tau))d\tau.$$

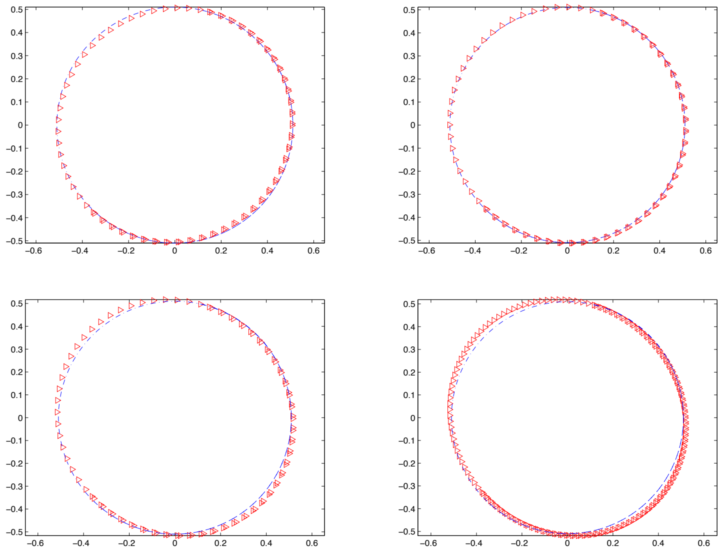

Figure B.1: Blue: reference figure (`ode45` with implementation `RelTol=1e-10`). Red: (1,1) $\theta = 0.9, \Delta t = 0.1$, (1,2) $\theta = 0.3, \Delta t = 0.1$, (2,1) $\theta = 0.1, \Delta t = 0.1$, (2,2) $\theta = -1.01, \Delta t = 0.05$. For all experiments, the initial value is $[0.1, 0.5]'$ and ω is 1.

The Theta method gives the discrete form as

$$\int_0^{\Delta t} f(z(t+\tau))d\tau = \Delta t f(\theta z(t) + (1-\theta)z(t+\Delta t)) + h.o.t.$$

Now, we need to find high order terms,

$$\begin{aligned}
\int_0^{\Delta t} f(z(t+\tau))d\tau &= \Delta t f(\theta z(t) + (1-\theta)z(t+\Delta t)) + \int_0^{\Delta t} \{f'(\theta z(t) \\
&+ (1-\theta)z(t+\tau))(\theta z(t) + (1-\theta)z(t+\Delta t) - z(t+\tau)\}d\tau \\
&= \Delta t f(\theta z(t) + (1-\theta)z(t+\Delta t)) + (\frac{1}{2} - \theta)f'(z(t))\dot{z}(t)\Delta t^2 + O(\Delta t^3) \\
&= \Delta t f(\theta z(t) + (1-\theta)z(t+\Delta t)) + (\frac{1}{2} - \theta)(f(z))'\Delta t^2 + O(\Delta t^3).
\end{aligned}$$

The Theta method is of first order . But when $\theta = \frac{1}{2}$, the method turns into Implicit Midpoint Rule method, which is of second order. If and only if $\theta = \frac{1}{2}$, the Theta method is symplectic, see Proposition 1 in [3]. □

Different to the Theta method, the Symplectic Theta method is symplectic for all θ and for any Hamiltonian system. The Symplectic Theta method is only of second order when $\theta = \frac{1}{2}$. Now, we would like to test the Symplectic Theta method by considering the harmonic oscillator $H = \frac{1}{2}p^2\omega^2 + \frac{1}{2}q^2$,

$$\begin{aligned}
\dot{q} &= H_p = p\omega^2, \\
\dot{p} &= -H_q = -q.
\end{aligned}$$

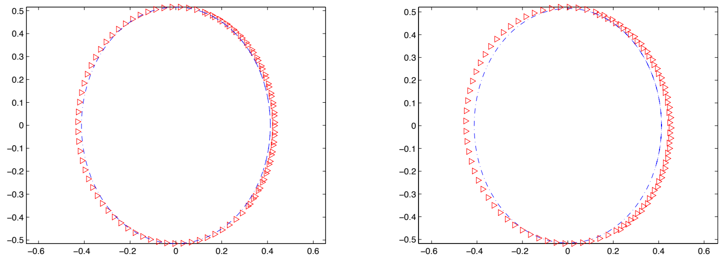

Figure B.2: Blue: reference figure (ode45 with implementation RelTol=1e-10). Red: (1,1) second order scheme $\theta = 0.3, \Delta t = 0.1$, (1,2) Symplectic Theta method first order $\theta = 0.3, \Delta t = 0.1$. For all experiments, the initial value is $[0.1, 0.5]'$ and ω is 0.8.

Based on the first order Symplectic Theta method in Theorem 10, we have[1]

$$\begin{aligned} q_{n+1} &= q_n + \Delta t \omega^2 (\theta p_{n+1} + (1-\theta) p_n), \\ p_{n+1} &= p_n - \Delta t (\theta q_n + (1-\theta) q_{n+1}). \end{aligned}$$

We can write this in matrix form,

$$\begin{pmatrix} q_{n+1} \\ p_{n+1} \end{pmatrix} = \frac{1}{1+\Delta t^2 \omega^2 \theta (1-\theta)} \begin{pmatrix} 1+\Delta t^2 \omega^2 \theta^2 & \Delta t \omega^2 \\ -\Delta t & 1-\Delta t^2 \omega^2 (1-\theta)^2 \end{pmatrix} \begin{pmatrix} q_n \\ p_n \end{pmatrix}.$$

From Lemma 2, we can obtain the second order scheme as

$$\begin{aligned} \begin{pmatrix} q_{n+1} \\ p_{n+1} \end{pmatrix} &= \frac{1}{(1+\Delta t^2 \omega^2 \theta (1-\theta)/4)^2} \begin{pmatrix} 1+\Delta t^2 \omega^2 \theta^2/4 & \Delta t \omega^2/2 \\ -\Delta t/2 & 1-\Delta t^2 \omega^2 (1-\theta)^2/4 \end{pmatrix} \\ &\quad \begin{pmatrix} 1-\Delta t^2 \omega^2 \theta^2/4 & -\Delta t/2 \\ \Delta t \omega^2/2 & 1-\Delta t^2 \omega^2 (1-\theta)^2/4 \end{pmatrix} \begin{pmatrix} q_n \\ p_n \end{pmatrix}. \end{aligned}$$

From the experiments, we can see that the Symplectic Theta method behaves quite well for all kinds of choices of θ (for instance, θ is chosen from 0.1, 0.3, -0.1 and -1.01, respectively, from Figure B.1). As we know, the explicit Euler and implicit Euler methods cannot preserve the symplectic property for the Harmonic oscillator. More details can be found on page 27 of [9]. The second order method behaves better than the first order method when ω is not 1, see Figure B.2.

[1] Here, we use subscripts $n+1$ and n instead of big and small variables.

Appendix C

Chartier et al. [2] studied the additively split system

$$\dot{y}(x) = \sum_{\nu=1}^{N} f^{[\nu]}(y) \tag{C.1}$$

for the divergence-free vector field f, while considering the splitting such that $f^{[\nu]}$ is also divergence-free, i.e. $div(f^{[\nu]})(y) = 0$ for $\nu = 1, \ldots, N$.

For the above splitting, we study the case based on coordinate partitioning, which implies that $f^{[\nu]}$ may not be divergence-free. Without loss of generality, we consider the divergence-free vector field in $\mathbb{R}^3$

$$\begin{aligned}
\dot{p} &= f(p,q,r),\\
\dot{q} &= g(p,q,r),\\
\dot{r} &= \phi(p,q,r),
\end{aligned} \tag{C.2}$$

with the divergence-free condition

$$f_p(p,q,r) + g_q(p,q,r) + \phi_r(p,q,r) = 0. \tag{C.3}$$

First, we consider one-stage partitioned Runge-Kutta method,

$$\begin{aligned}
p_1 &= p_0 + \Delta t f(P,Q,R),\\
q_1 &= q_0 + \Delta t g(P,Q,R),\\
r_1 &= r_0 + \Delta t \phi(P,Q,R),
\end{aligned} \tag{C.4}$$

where

$$\begin{aligned}
P &= p_0 + \Delta t \alpha_1 f(P,Q,R),\\
Q &= q_0 + \Delta t \alpha_2 g(P,Q,R),\\
R &= r_0 + \Delta t \alpha_3 \phi(P,Q,R).
\end{aligned} \tag{C.5}$$

If we set

$$M = \begin{pmatrix} f_p & f_q & f_r \\ g_p & g_q & g_r \\ \phi_p & \phi_q & \phi_r \end{pmatrix}$$

and

$$N = \begin{pmatrix} \alpha_1 & & \\ & \alpha_2 & \\ & & \alpha_3 \end{pmatrix} M = \alpha M,$$

differentiating both sides of (C.4) and (C.5), we obtain

$$\begin{pmatrix} dp_1 \\ dq_1 \\ dr_1 \end{pmatrix} = [I_3 + \Delta t M (I_3 - \Delta t \alpha M)^{-1}] \begin{pmatrix} dp_0 \\ dq_0 \\ dr_0 \end{pmatrix}.$$

The volume preserving condition requires that

$$\det(I_3 + \Delta t M (I_3 - h\alpha M)^{-1}) = 1,$$

which implies

$$\det(I_3 - \Delta t \alpha M) = \det(I_3 - \Delta t(\alpha - I_3)M). \tag{C.6}$$

One can see that for any divergence-free vector field in $\mathbb{R}^3$, (C.6) is a sufficient and necessary condition for the one-stage partitioned Runge-Kutta method to be volume preserving. It is easy to see that for any divergence-free vector field, matrix M has the property trace$(M) = 0$ due to the divergence-free condition. However, one cannot find a matrix α satisfying (C.6) for any divergence-free vector field (M is unknown). This implies that there is no one-stage partitioned Runge-Kutta method for any divergence-free vector field in $\mathbb{R}^3$.

For any given divergence-free vector field and time step (M and Δt are known), there is always a solution (relation). However, the solution generally depends on the vector field M itself and the time step Δt. For some special divergence-free vector fields, one can find the relation which does not depend on the vector field or time step, see Example 3.

Example 3. *We consider the three-cycle system [2]*

$$\begin{aligned} \dot{p} &= f(q), \\ \dot{q} &= g(r), \\ \dot{r} &= \phi(p). \end{aligned}$$

The volume preserving condition (C.6) for this system becomes

$$1 - \Delta t^3 \alpha_1 \alpha_2 \alpha_3 f'(q) g'(r) \phi'(p) = 1 - \Delta t^3 (\alpha_1 - 1)(\alpha_2 - 1)(\alpha_3 - 1) f'(q) g'(r) \phi'(p),$$

which gives us

$$(\alpha_1 - 1)(\alpha_2 - 1)(\alpha_3 - 1) = \alpha_1 \alpha_2 \alpha_3. \tag{C.7}$$

Remark 9. *(C.7) covers the results of Theorem 4.21 in [2], that is, a one-stage partitioned Runge-Kutta method formed of (α_1,1), (α_2,1) and (α_3,1) is volume preserving for the three-cycle system if and only if (C.7) is satisfied. Furthermore, even for the three-cycle system which is diagonal-free and divergence-free, no one-stage Runge-Kutta scheme can be volume preserving.*

Proof. Note that (C.7) is a sufficient and necessary condition for the three-cycle system. For the Runge-Kutta scheme, we have that $\alpha_1 = \alpha_2 = \alpha_3$. Then equation (C.7) becomes $3\alpha_1^2 - 3\alpha_1 + 1 = 0$, which has no real solutions. □

As a simpler case, we consider $\mathbb{R}^2$. By a similar procedure, we have the volume preserving condition for a one-stage partitioned Runge-Kutta method

$$\det(I_2 - \Delta t \alpha_{22} M_{22}) = \det(I_2 - \Delta t(\alpha_{22} - I_2) M_{22}), \tag{C.8}$$

where the matrices α_{22} and M_{22} are the first two rows and the first two columns of matrices α and M, respectively.

Theorem 11. *The partitioned Runge-Kutta method formed of $(\alpha_1, 1)$ and $(\alpha_2, 1)$ is volume preserving for any Hamiltonian system if and only if $\alpha_1 = 1 - \alpha_2$. The Symplectic Theta method in Theorem 9 in Chapter 3 can be interpreted as a one-stage partitioned Runge-Kutta method formed of $(\theta, 1)$ and $(1-\theta, 1)$.*

Proof. By using (C.8), we have

$$\begin{vmatrix} 1 - \Delta t \alpha_1 f_p & -\Delta t \alpha_1 f_q \\ -\Delta t \alpha_2 g_p & 1 - \Delta t \alpha_2 g_q \end{vmatrix} = \begin{vmatrix} 1 - \Delta t(\alpha_1 - 1) f_p & -\Delta t \alpha_1 f_q \\ -\Delta t \alpha_2 g_p & 1 - \Delta t(\alpha_2 - 1) g_q \end{vmatrix}.$$

Comparing both sides of the above equation and using $f_p + g_q = 0$, we have

$$\alpha_1 \alpha_2 = (\alpha_1 - 1)(\alpha_2 - 1).$$

Similarly, if we have $\alpha_1 = 1 - \alpha_2$, we obtain that $dp_1 \wedge dq_1 = dp_0 \wedge dq_0$, which gives the volume preserving property.

The scheme of partitioned Runge-Kutta method with $(\theta, 1)$ and $(1-\theta, 1)$ can be written as

$$\begin{aligned} p_1 &= p_0 + \Delta t f(p_0 + \theta \Delta t f(P,Q), q_0 + (1-\theta)\Delta t g(P,Q)), \\ q_1 &= q_0 + \Delta t g(p_0 + \theta \Delta t f(P,Q), q_0 + (1-\theta)\Delta t g(P,Q)), \end{aligned} \tag{C.9}$$

and

$$\begin{aligned} p_1 &= p_0 + \Delta t f(P,Q), \\ q_1 &= q_0 + \Delta t g(P,Q). \end{aligned}$$

Then, (C.9) becomes

$$\begin{aligned} p_1 &= p_0 + \Delta t f(\theta p_1 + (1-\theta) p_0, (1-\theta) q_1 + \theta q_0), \\ q_1 &= q_0 + \Delta t g(\theta p_1 + (1-\theta) p_0, (1-\theta) q_1 + \theta q_0). \end{aligned}$$

Considering the relation between f and g, i.e. $f_p + g_q = 0$, we can define a smooth function H such that

$$f = H_q, \;\; g = -H_p.$$

Substituting this for the $\mathbb{R}^2$ case, we obtain exactly the Symplectic Theta method in Theorem 9. □

Bibliography

[1] M.M. Carroll. A representation theorem for volume-preserving transformations. *J. International Journal of Non-Linear Mechanics*, Volume 39, Issue 2, Pages 219-224, 2004. 1.3.3

[2] P. Chartier and A. Murua. Preserving first integrals and volume forms of additively split systems. *IMA Journal of Numerical Analysis*, 27(2):381, 2007. 1.3.1, C, 3, 9

[3] K. Feng. Difference schemes for Hamiltonian formalism and symplectic geometry. *J. Comput. Math*, Volume 4, Issue 3, Pages 279-289, 1986. 5, 3, B

[4] K. Feng and Z.-J Shang. Volume-preserving algorithms for source-free dynamical systems. *Numericshe Mathematik*, Volume 71, Number 4, 451-463, 1995. 1.3.1, 1.3.2, 2.4, 3, 3.1, 3.1

[5] E. Hairer, C. Lubich and G. Wanner. Geometric numerical integration: structure-preserving algorithms for ordinary differential equations. *Spring*, second edition, 2006. 1.1, 1.2, 1.3.3, 1.4

[6] M. Hochbruck, C. Lubich and H. Selhofer. Exponential integrators for large systems of differential equations, *SIAM J. Sci. Comout.*, Volume 19, Pages 1552-1574, 1998. 1.3.1

[7] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett and A. Zanna. Lie-group methods. *Acta Numerica*, Volume 9, Pages 215-365, 2000. 2.1

[8] A. Iserles, G.R.W. Quispel and P.S.P. Tse. B-series methods cannot be volume-reserving. *BIT Numerical Mathematics*, Volume 47, Pages 351–378, 2007. 1.3.1

[9] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2004. 1.1, 1.2, 3.5.1, B

[10] H.E. Lomeli and J.D. Meiss. Generating forms for exact volume preserving maps. *Discrete and continuous dynamical systems series s*, Number 2, Pages 361-377, 2009. 1.2, 1.3.3, 13, 14, 3, 3.2

[11] J.E. Marsden and T.S. Radiu. Introduction to mechanics and symmetry, *Springer*, second edition, 1999. 2.2, 2.2

[12] J.E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, Volume 10, Pages 357–514, 2001. 2.4, 3, 2.4, 2.4, 4, 2.4, 3.5.1

[13] R.I. McLachlan, H.Z. Munthe-Kaas, G.R.W. Quispel, and A. Zanna. Explicit volume-preserving splitting methods for linear and quadratic divergence-free vector fields. *Found. Comput. Math.*, 8(3):335–355, 2008. 1.3.2, 3, 3.2, 4

[14] R.I. McLachlan, G.R.W. Quispel. Six lectures on geometric integration, *Foundation of Computational Mathematics* (R. Devore, A. Iserles and E. Süli, eds), Cambridge University Press, Cambridge, pp. 155-210. 1, 1.1, 1.2, 1.3.2, 3.3

[15] R.I. McLachlan, G.R.W. Quispel. Geometric integrators for ODES, *J. Phys. A: Math. Gen*, volume 39, pages 5251-5285, 2006. 1.1, 1.3.2

[16] R.I. McLachlan, G.R.W. Quispel, Splitting methods. *Acta Numerica*, Volume 11, Pages 341-434, 2002. 1.1, 3.1, 1

[17] R.I. McLachlan and G.R.W. Quispel. Explicit geometric integration of polynomial vector fields. *BIT*, 44:515–538, 2004. 1.3.2

[18] K. Modin. Adaptive geometric numerical integration of mechanics systems. *PhD thesis*, Lund University, Lund, Sweden, 2009. 2.2

[19] G.R.W. Quispel. Volume-preserving integrators. *Physics Letters A*, Volume 206, Pages 26-30, 1995. 1.3.2, 3, 3.3

[20] G.R.W. Quispel and D.I. McLaren. Explicit volume-preserving and symplectic integrators for trigonometric polynomial flows. *J. Comput. Phys.*, 186(1):308–316, 2003. 1.3.2, 3.2

[21] Z.-J Shang. Generating functions for volume preserving mapping with applications I: Basic theory. *China/Korea Joint Seminar: Dynamical systems and their application* 1.3.2, 3, 3.4, 5, 6, 3.4, 1

[22] Z.-J Shang. Construction of volume-preserving difference schemes for source-free systems via generating function. *Journal of computational mathematics*, Vol.12, No.3, Pages 265-272, 1994. 3, 3.4, 3.4, 1, 3.4, 7, 8

[23] P.S.P. Tse. Geometric Numerical Integration: On the Numerical Preservation of Multiple Geometric Properties for Ordinary Differential Equations, *PhD thesis*, La Trobe University, Melbourne, Australia, 2007. 1.3.1

[24] Z. Xia. Existence of invariant tori in volume-preserving diffeomorphisms, *Ergodic Theory and Dynamical Systems*, Volume 12, Issue 03, pages 621-631, 1992. 1.2

[25] H. Xue. High order volume preserving integrators for three kinds of divergence-free vector fields via commutators, *Preprint*, 2013. 3.2

[26] H. Xue and A. Zanna. Explicit volume-preserving splitting methods for polynomial divergence-free vector fields, *BIT Numer Math*, Volume 53, pages 265-281, 2013. 1.3.2, 3, 3.2

[27] H. Xue, O. Verdier and A. Zanna. A study on volume preserving generating forms in $\mathbb{R}^3$, *Preprint*, 2013. 1.3.3, 2, 2.2, 2.4, 2.4, 3, 3.5, 3.1, 3.5.2, 5

[28] H. Xue and A. Zanna. Generating functions and volume preserving mappings, To appear: *Discrete and Continuous Dynamical Systems*, Volume 34, Number 3, pages 1229-1249, March 2014. 1.3.3, 2, 2.2, 3, 1, 3.5, 3.5.1, 9, 3.1, 3.5.2, 4, 3.5.2, 3.5.2, 4

[29] H. Yoshida. Construction of higher order symplectic integrators, *Physics Letters A*, 150:262–268, 1990. 3.2

[30] A. Zanna. Explicit Volume-Preserving splitting methods for divergence-free ODEs by tensor product bases decompositions, *Submitted*, 2012. 1.3.2, 3, 3.2

[31] A. Zanna. The Euler equation of quasi-geostrophic fluids and volume preserving numerical methods, *Technical Report 1205.1947*, arXiv, 2012. 4

Part II

Included papers