Projection and nested force-gradient methods for quantum field theories

Dissertation

zur Erlangung
des akademischen Grades
eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

der
Fakultät 4 – Mathematik und Naturwissenschaften
der
Bergischen Universität Wuppertal (BUW)
vorgelegt von

Dmitry Shcherbakov

geboren am 13.12.1987 in Wolgograd, Russland

betreut von: Prof. Dr. Michael Günther (BUW)
                Prof. Dr. Matthias Ehrhardt (BUW)

Wuppertal, 26. Juli 2017

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20170802-120731-0
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20170802-120731-0]

Acknowledgments

First and foremost I would like to express my deepest gratitude to Prof. Michael Günther and Prof. Matthias Ehrhardt for the continuous support of my doctoral study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me through all the time of my research and writing this thesis. I would like to thank my colleagues from the physics department Prof. Dr. Francesco Knechtli and Dr. Jacob Finkenrath for the fruitful discussions and useful meetings. They have played a very important role in my research and without their help it would be almost impossible to complete this thesis. I would like to thank Prof. Dr. Mike Peardon from Trinity College Dublin for his inspiring ideas and helpful supervision during my secondment in Ireland.

I also would like to thank all my colleagues from the chair of Applied Mathematics and Numerical Analysis for providing a great working atmosphere. I have learned a lot from them and the experience we shared together through all these years has truly made me a better person.

I would like to thank my friend Alexey Gutermann and his wonderful parents, who have been kind to me and have supported me from my first day in Germany. I owe a great debt to my friend Nemanja Bozovic and his lovely wife Milana for making my life brighter during my staying in this rainy city of Wuppertal. I also would like to thank all my friends from Russia and all over the world for always been their for me no matter how far distances were between us.

And last but not least I would like to thank my family. I could not have achieved anything in my life without their endless support, love and understanding.

Contents

List of Figures

1 MOTIVATION AND OVERVIEW

The need of developing structure-preserving algorithms for special classes of problems arose independently from very different areas of research as mechanics, astronomy, molecular dynamics, theoretical physics as well as from other areas of both applied and pure mathematics. Numerical methods that preserve geometric properties of the flow of a differential equation: symplectic integrators for Hamiltonian systems, symmetric integrators for reversible systems, methods preserving first integrals and numerical methods on manifolds etc. show that the conservation of geometric properties of the flow not only produces an improved qualitative behavior, but also allows for a more accurate long-time integration than with general-purpose methods [28].

For the Hybrid Monte Carlo algorithm (HMC) [16], often used to study the fundamental quantum field theory of quarks and gluons, quantum chromodynamics (QCD), on the lattice, one is interested in efficient numerical time integration schemes which preserve geometric properties of the flow and are optimal in terms of computational costs per trajectory for a given acceptance rate. High order numerical methods allow the use of larger step sizes, but demand a larger computational effort per step; low order schemes do not require such large computational costs per step, but need more steps per trajectory. So there is a need to balance these opposing effects.

Omelyan integration schemes [42] of a force-gradient type have proved to be an efficient choice, since it is easy to obtain higher order schemes that demand a small additional computational effort. These schemes use higher-order information from force-gradient terms to both increase the convergence of the method and decrease the size of the leading error coefficient. Other ideas to achieve better efficiency for numerical time integrators are given by multirate or nested approaches. These schemes do not increase the order but reduce the computational costs per path by recognizing the different dynamical time-scales generated by different parts of the action. Slow forces, which are usually expensive to evaluate, need only to be sampled at low frequency while fast forces which are usually cheap to evaluate need to be sampled at a high frequency. The important feature of both these class of numerical methods is that their construction guarantees the preservation of some geometrical properties of the flow. A natural way to inherit the advantages from both force-gradient type schemes and multirate approaches is to combine these two ideas.

A different approach to improve the efficiency of the HMC can be a possibility of neglecting the accept/reject step of the algorithm by using the idea of symmetric projection methods [25]. This class of numerical methods conserves the energy of a system exactly, which leads to a hundred percent acceptance rate of the HMC and hence there is no need for the accept/reject step in HMC. Due to the requirements for the numerical integrator in HMC to preserve a certain geometric properties the idea of projection methods has to be modified and analyzed accordingly.

This thesis is organized in the following way. First we introduce the basis concepts of

geometric numerical time integrators for ordinary differential equations in Chapter 2. We present some geometric properties in the context of Hamiltonian mechanics and then transfer these concepts to a structure-preserving approach for numerical methods together with an idea how to estimate a qualitative behavior of the numerical schemes. Examples of different geometric integrators, most commonly used in the integration on manifolds for both Abelian and non-Abelian cases, are presented.

Chapter 3 introduces the main application for the methods developed in this thesis. We present the main ideas how to treat quantum field theories on the lattice. We introduce basic concepts for the path integral formulation of quantum field theories together with classical examples of such theories. The HMC algorithm is shown as the main way to treat problems given in the path integral formulation. The role of structure-preserving numerical methods and the need to conserve geometric properties of the considered physical system is also explained.

The idea of the structure-preserving projection methods is presented in Chapter 4. We analyze the possibility to construct a projection method which would satisfy all the requirements needed for HMC simulations. We present another view on symmetric projection method and show our attempts to develop symplectic projection schemes. At the end the conclusion about structure-preserving projection methods is given based on both analytical and numerical results.

In Chapter 5 we give a short introduction to the two well-known ideas based on operator splitting, namely decomposition splitting and multirate splitting. The force-gradient extension of the decomposition approach is introduced. We analyze the structure-preserving properties of these numerical schemes as well as order conditions. Then we introduce and study a novel class of numerical integrators, the adapted nested force-gradient schemes, obtained via combination of both the force-gradient and the nested algorithm approaches in order to obtain more efficient numerical integrators.

In the end, we study the behavior of the adapted nested force-gradient schemes for an example of the n-body problem in order to learn more about their usefulness for lattice field theory calculations and confirm our analytical findings. We also show the derivation of a force-gradient term. Then we test these methods for the Schwinger model on the lattice, a well known benchmark problem to cope with the problems given by the non-Abelian setting in the HMC for lattice QCD. We derive the analytical basis of nested force-gradient type methods and demonstrate the advantage of the proposed approach, namely reduced computational costs compared with other numerical integration schemes in HMC.

Our main goal is to develop novel geometric numerical time integrators in order to improve the efficiency of the HMC algorithm in application to the problems of quantum field theories. This thesis is based on author's publications [51, 52].

2 GEOMETRIC NUMERICAL INTEGRATORS

The motion of various objects, whether we are talking about particles or planets, is described by differential equations, which are derived from the laws of physics. These equations define the current state of the system including all the physical laws relevant to the particular situation. Due to the usual complexity of the underlying physical problem, in order to obtain necessary information about the considered system one has to use numerical methods to find the solution of the differential equations, describing the given problem [24].

Standard numerical integration methods for simulating motion take an initial condition and move the objects in the direction specified by the differential equations. Unfortunately they completely ignore some of the hidden physical laws contained within the equations. Geometric integrators on the other hand obey these extra laws. Naturally, those physical laws have to be known if the integrator is going to obey it. The advantages of this structure-preserving approach may yield faster, simpler, more stable and/or more accurate methods as well as more robust schemes with quantitatively better results than standard methods. Also it must be mentioned that the requirement of structure preservation necessity restricts the choice of the method and may increase the cost. Therefore in each case the benefits and costs must be balanced [38].

In this chapter we briefly introduce the basic concept of geometric numerical time integrators for ordinary differential equations (ODEs), where we mainly focus on the Hamiltonian equations. We also recall some definitions of geometric properties of physical systems, which can be conserved by structure-preserving numerical methods. We show how to estimate the qualitative behavior of such integrators. Examples of the most commonly used geometric integrators are presented as well. Finally we give a short insight in the numerical treatment of differential equations on manifolds, more precisely Lie groups.

2.1 Geometric time integration of ODEs

Let us first introduce the standard formulation of numerical time integration of an initial value problem (IVP) for a system of ODEs to be solved. It is usually written in the following form

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \dot{y} = f(t, y), \quad 0 \le t \le T, \quad y(0) = y_0 \in \mathbb{R}^d, \tag{2.1}$$

which can describe a couple of real life physical problems.

It can easily happen that the solution $y(t)$ of the system (2.1) might be deduced from the given problem. Our main focus is related to its conservation laws, i.e. if $y(t)$ is a solution

of the IVP (2.1) which satisfies

$$g(t, y(t)) = 0 \tag{2.2}$$

for some t, then it satisfies the equation (2.2) for all t of interest. There are some other hidden apriori information, which can be unfortunately lost by the numerical solution obtained via numerical integration. Here we must emphasize that the solution $y(t)$ is determined completely by the IVP (2.1). Laws, such as equation (2.2), represent conclusions drawn about $y(t)$ but not conditions imposed on it [10].

For example, *linear conservation laws* arise, if for the right side f of the ODE (2.1) there exist a column vector x such that

$$x^\top f(t, y(t)) \equiv 0$$

identically in t and $y(t)$. This immediately implies from (2.1)

$$x^\top \dot{y} = x^\top f(t, y(t)) = 0$$

and thus it can be further represented by integration as

$$g(t, y(t)) = x^\top y(t) - x^\top y(0) = 0.$$

Such laws are associated with terms like *charge balance* or *conservation of mass*.

Nonlinear conservation laws can arise, for example, if there is a function $F(t, y)$ such as

$$\frac{\partial F}{\partial t}(t, y(t)) + \frac{\partial F}{\partial y}(t, y(t)) f(t, y(t)) = 0$$

identically, then for the total derivative we have

$$\frac{\mathrm{d}}{\mathrm{d}t} F(t, y(t)) = 0$$

for any solution $y(t)$ of the system (2.1). This yields by integration to the following condition

$$g(t, y(t)) = F(t, y(t)) - F(0, y(0)) = 0.$$

These type of laws are associated with the *conservation of energy* or the *conservation of angular momentum*.

Generally numerical time integrators do not preserve geometric properties of the underlying physical problem, even so, in principle, most numerical methods satisfy linear conservation laws [9]. But since it is very important to obtain a numerical solution, which possesses meaningful physical properties and exhibits a correct qualitative behavior, obviously there is a need in developing efficient structure-preserving numerical algorithms.

The research in geometric integration focuses on three main areas [4]:

1. development of new types of integrators, and integrators preserving new structures;

2. improvement of the efficiency of the integrators, by constructing methods tuned for special cases of systems;

3. and study of the behavior of the integrators, especially their long-time dynamics;

and the extent to which the phase portrait of the system is preserved.

2.2 Hamiltonian mechanics

In order to obtain a better insight in the world of geometric integration, we consider structure preserving properties of geometric integrators in the application to Hamiltonian mechanics. Namely, because of the fact that Hamiltonian dynamics has a physical interpretation that can provide useful intuitions and gives an example of a differential problem with an underlying structure, which encapsulates invariance and symmetries [10].

It is a very general and appealing underlying geometric framework with a good perspective for generalization to various disciplines, a strong connection between symmetries and conservation laws and a perspective for the application of statistical mechanics. While, for example, Lagrange's equation describes the motion of a particle as a single second-order ODE, Hamilton's equations describe the motion as a coupled system of two first-order ODEs [38].

One of the main advantages of Hamiltonian mechanics is that it exhibits a similar structure as quantum mechanics, the theory that describes the motion of particles at very tiny (subatomic) distance scales. An understanding of Hamiltonian mechanics thus provides a good introduction to the mathematics of quantum mechanics, which is proved to be a great test site for exposing advantages of geometric numerical integrators [9].

Hamiltonian dynamics operates on a d-dimensional position vector q, and a d-dimensional momentum vector p, such that the full state space has $2 \times d$ dimensions. The system is described by a function of q and p, known as the Hamiltonian $H(q, p)$. The partial derivatives of the Hamiltonian determine how q and p change with respect to time t, according to *Hamilton's equations of motion*:

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad\qquad \dot{p} = -\frac{\partial H}{\partial q}. \qquad (2.3)$$

For any time interval of duration ε, these equations define a mapping, $\mathcal{T}_\varepsilon$, from the state at any time t to the state at time $t + \varepsilon$. Here, the Hamiltonian function H, and hence the time mapping $\mathcal{T}_\varepsilon$, are assumed not to depend explicitly on t.

Alternatively, we can combine the vectors q and p into the vector $y = (q, p)$ with $2 \times d$ dimensions, and write Hamilton's equations (2.3) as

$$\frac{\mathrm{d}y}{\mathrm{d}t} = J^{-1} \nabla H(y),$$

where $\nabla H(\mathbf{y})$ is the gradient of the Hamiltonian H, and

$$J = \begin{pmatrix} 0_{d \times d} & I_{d \times d} \\ -I_{d \times d} & 0_{d \times d} \end{pmatrix} \qquad (2.4)$$

is a $2d \times 2d$ matrix whose blocks are defined above in terms of identity and zero matrices. In the literature the matrix J in (2.4) is called *standard symplectic matrix* [28].

We usually use separated Hamiltonian functions that can be written as

$$H(q, p) = T(p) + V(q), \tag{2.5}$$

where $V(q)$ represents the *potential energy* of the system and $T(p)$ its *kinetic energy* often defined as

$$T(p) = \frac{p^\top M^{-1} p}{2}. \tag{2.6}$$

Here, $M \in \mathbb{R}^{d \times d}$ denotes a symmetric, positive-definite *mass matrix*, which is typically diagonal, and is often a scalar multiple of the identity matrix. With these forms for decoupled H and T, Hamilton's equations of motion (2.3) can be written

$$\frac{\mathrm{d}q_j}{\mathrm{d}t} = \left[M^{-1} p\right]_j, \qquad\qquad \frac{\mathrm{d}p_j}{\mathrm{d}t} = -\frac{\partial V}{\partial q_j}, \qquad\qquad j = 1, \ldots, d. \tag{2.7}$$

Let us now take a look at some geometric properties of the solution to the given system (2.7).

Time-reversibility. First, Hamiltonian dynamics is reversible. The mapping $\mathcal{T}_\varepsilon$ from the state $(q(t), p(t))$ at time t to the state $(q(t + \varepsilon), p(t + \varepsilon))$ at time $t + \varepsilon$, is one-to-one, and hence has an inverse, $\mathcal{T}_{-\varepsilon}$. This inverse mapping is obtained by simply negating the time derivatives in equations (2.7). If the Hamiltonian has the form (2.5)

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

and $T(p) = T(-p)$, assign the quadratic form for the kinetic energy (2.6), the inverse mapping can also be obtained by negating p, applying $\mathcal{T}_\varepsilon$, and then negating p again, like it is shown in Figure 2.1.
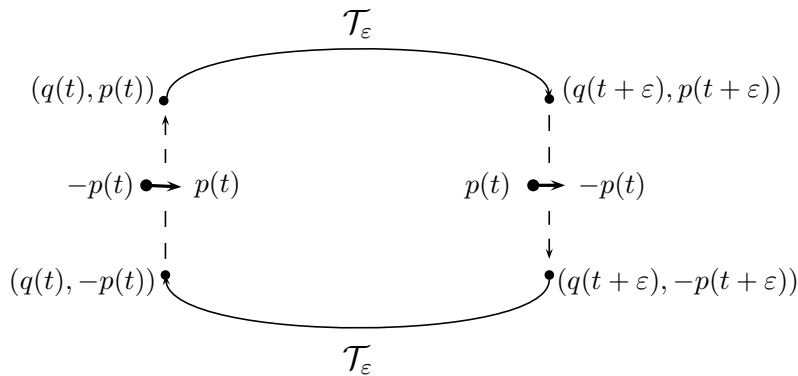
Figure 2.1: Time-reversibility.

Conservation of the Hamiltonian. A second property of the dynamics is that it keeps the Hamiltonian invariant (i.e. conserved). This is easily seen from equations (2.7) as follows:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_{j=1}^{d} \left[\frac{\mathrm{d}q_j}{\mathrm{d}t} \frac{\partial H}{\partial q_j} + \frac{\mathrm{d}p_j}{\mathrm{d}t} \frac{\partial H}{\partial p_j} \right] = \sum_{j=1}^{d} \left[\frac{\partial H}{\partial p_j} \frac{\partial H}{\partial q_j} - \frac{\partial H}{\partial q_j} \frac{\partial H}{\partial p_j} \right] = 0.$$

We will see later, however, that a numerical time integration can only make H approximately invariant, and hence we will not be able to achieve this conservation exactly. But

using numerical time integrators of higher order we can obtain a better conservation of the Hamiltonian H (2.5).

Volume-Preservation. A third fundamental property of Hamiltonian dynamics is that it preserves the volume (q, p) in space (a result known as Liouville's theorem). If we apply the mapping $\mathcal{T}_\varepsilon$ to the points in some region R of the (q, p) space, with volume $\mathcal{D}$, the image of R under $\mathcal{T}_\varepsilon$ will also have the same volume $\mathcal{D}$ as it is shown in Figure 2.2.
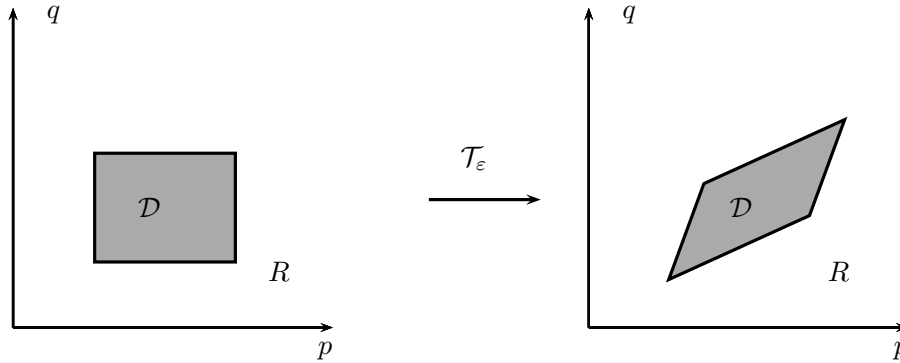
Figure 2.2: Symplecticity(volume preservation).

The preservation of volume by Hamiltonian dynamics can be proved in several ways. One is to note that the divergence of the vector field defined by equations (2.7) is zero, which can be seen as follows:

$$\sum_{j=1}^{d} \left[\frac{\partial}{\partial q_j} \frac{\mathrm{d}q_j}{\mathrm{d}t} + \frac{\partial}{\partial p_j} \frac{\mathrm{d}p_j}{\mathrm{d}t} \right] = \sum_{j=1}^{d} \left[\frac{\partial}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial}{\partial p_j} \frac{\partial H}{\partial q_j} \right] = \sum_{j=1}^{d} \left[\frac{\partial^2 H}{\partial q_j p_j} - \frac{\partial^2 H}{\partial p_j q_j} \right] = 0$$

A vector field with zero divergence is known to preserve volume [1].

Symplecticity. The volume preservation is also a consequence of Hamiltonian dynamics being symplectic. Letting $y = (q, p)$, and defining matrix J (2.4), the *symplecticity condition* is that the Jacobian matrix, B_ε, of the mapping $\mathcal{T}_\varepsilon$ satisfies

$$B_\varepsilon^\top J B_\varepsilon = J.$$

This implies a volume conservation, since $\det(B_\varepsilon^\top) \det(J) \det(B_\varepsilon) = \det(J)$ implies that $\det(B_\varepsilon)^2$ is one. When $d > 1$, the symplecticity condition is stronger than the volume preservation. Hamiltonian dynamics and the symplecticity condition can be generalized to the case when J is any matrix for which $J^\top = -J$ and $\det(J) \neq 0$.

It is important to mention that reversibility, preservation of volume, and symplecticity can be maintained exactly even when Hamiltonian dynamics is approximated, which is usually the case in practice. We will demonstrate this property in the next section.

2.3 Conservation of physical properties

In this section we introduce the structure preserving properties in terms of numerical time integrators. We again consider the IVP for the system of ODEs (2.1) $\dot{y} = f(t, y)$

with $y(0) = y_0$ and we study the question to what extent geometric properties of the exact flow $\varphi_t(y_0)$ can be preserved by a numerical approximation y_{n+1} obtained via time integration by some numerical scheme $\Phi_h(y_n)$ with the time step-size h. Here we present some definitions and theorems, which are very important for understanding the idea of transferring preservation laws to the world of numerical analysis.

2.3.1 Symmetry, time-reversibility

The first property we consider is a conservation of the symmetries, since it plays a crucial role in physical problems, which current state changes during the time. Therefore it will be beneficial to have such a numerical scheme.

Definition 1. *[28] A numerical method Φ_h is called symmetric, if it satisfies the following condition*

$$\Phi_h \circ \Phi_{-h} = id \quad \text{or equivalently} \quad \Phi_h = \Phi_{-h}^{-1}.$$

Hence a method $y_{n+1} = \Phi_h(y_n)$ is *symmetric* if exchanging $y_n \leftrightarrow y_{n+1}$ and $h \leftrightarrow -h$ leaves the method unaltered.

In order to define the time-reversibility we must first introduce few definitions.

Definition 2. *[28] Let ρ be an invertible linear transformation in the phase space of $\dot{y} = f(y)$. This differential equation and the vector field $f(y)$ are called ρ-reversible if*

$$\rho f(y) = -f(\rho y) \qquad\qquad \text{for all } \ y.$$

Definition 3. *[28] A map $\Phi(y)$ is called ρ-reversible if*

$$\rho \circ \Phi = \Phi^{-1} \circ \rho.$$

Theorem 1. *[28] If a numerical method Φ_h, applied to a ρ-reversible differential equation, satisfies*

$$\Phi_{-h} \circ \rho = \rho \circ \Phi_h^{-1},$$

then the numerical flow ϕ_h is a ρ-reversible map if and only if Φ_h is symmetric.

From this theorem we know that symmetry does not imply the ρ-reversibility of the numerical flow [28]. If Theorem 1 holds for $\rho = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$ then a numerical method Φ_h is time reversible.

Symmetric numerical schemes, applied to reversible systems, show an improved long-time behavior in comparison to the standard numerical integrators [28].

2.3.2 Volume-preservation, symplecticity

The second fundamental property is symplecticity of the numerical flow. For example, if $y(0)$ on some domain D possess certain properties, then $y(t)$ retain those properties after the transformation through φ_t. Hence, it is natural to look for numerical methods that share this property.

Definition 4. *[28] A numerical method Φ_h is called symplectic if the following condition holds*

$$\left(\frac{\partial y_{n+1}}{\partial y_n}\right)^{\top} J \left(\frac{\partial y_{n+1}}{\partial y_n}\right) = J,$$

where $\frac{\partial y_{n+1}}{\partial y_n}$ is the sensitivity matrix of the given scheme Φ_h with respect to the initial value and J is the skew-symmetric matrix (2.4).

Sometimes a weaker property can be enough to be satisfied by a numerical integrator.

Definition 5. *[28] A numerical method Φ_h is called area-preserving if the following condition holds*

$$\left|\det\left(\frac{\partial y_{n+1}}{\partial y_n}\right)\right| = 1.$$

It is straightforward to observe the following conclusion.

Corollary 1. *Symplecticity implies area-preservation.*

Symplectic integrators are known to have a very good energy behavior, meaning that the considered system does not loose or gain energy. For example Hamiltonian systems possess both time-reversibility and symplecticity properties and therefore require integrators to conserve both properties.

2.3.3 Energy conservation, convergence order

The convergence of a numerical integrator is related to the global accuracy, and bounds the global error after n time steps.

Definition 6. *[5] An integrator Φ_h is said to be convergent of order r if there exist an open set U and constant $\tilde{h} > 0$ such that*

$$\|\Phi_h^N(y_0) - \varphi_t(y_0)\| = \mathcal{O}(h^r) \quad for \quad h \leq \tilde{h}, \quad y_0 \in U.$$

Since it is almost never possible to obtain the exact flow $\varphi_t(y_0)$ of the system (2.1) we use in the next section the technique of backward error analysis, which helps us to estimate the qualitative performance of the numerical scheme via the amount of the conserved energy.

The third important feature of the Hamiltonian systems (2.7) is the invariance of the total energy, which is represented by the Hamiltonian H (2.5).

Theorem 2. *[28] Consider a Hamiltonian system with an analytic Hamiltonian function $H : D \mapsto \mathbb{R}$ (where $D \subset \mathbb{R}^{2d}$), and apply a symplectic numerical method Φ_h with step size h. If the numerical solution stays in the compact set $K \subset D$, then there exist h_0 and $N = N(h)$ (namely N equal to the largest integer satisfying $hN \geq h_0$) such that*

$$\tilde{H}(y_n) = \tilde{H}(y_0) + O(e^{\frac{-h_0}{2h}})$$
$$H(y_n) = H(y_0) + O(h^r)$$

over exponentially long time intervals $nh \leq e^{\frac{h_0}{2h}}$. The modified analytic Hamiltonian $\tilde{H}$ is exactly preserved by the numerical scheme Φ_h

$$\tilde{H}(q,p) = H(q,p) + hH_2(q,p) + h^2H_3(q,p) + \ldots, \qquad (2.8)$$

where the smooth functions $H_j : \mathbb{R}^{2d} \mapsto \mathbb{R}$ for $j = 2, 3, \ldots$, such that $f_j(y) = J^{-1}\nabla H_j(y)$, where defined by (2.9).

For non-symplectic methods we can obtain an estimate of the energy preservation by a computation similar to that of the proof of Theorem 2 which is given in [24]. And from a Lipschitz condition of the Hamiltonian H and from the standard local error estimate, we obtain $H(y_{n+1}) - H(\varphi_h(y_n)) = \mathcal{O}(h^{r+1})$. Since $H(\varphi_h(y_n)) = H(y_n)$, a summation of these terms leads to $H(y_n) - H(y_0) = O(th^r)$ for $t = nh$. This means that the energy does not stay invariant for non-symplectic methods.

From Theorem 2 we can conclude that numerical methods cannot be symplectic and conserve the total energy exactly. They can preserve the Hamiltonian H up to some degree r, which depends of the convergence order of the numerical scheme. Meaning, the higher the convergence order of the integrator the better it conserves the total energy.

2.4 Backward error analysis

The idea of backward error analysis is the following: for a given numerical integrator search for a modified differential equations, such that the exact solution of this modified equation approximates very well the numerical solution. An analysis of the modified differential equation then gives much insight into the numerical flow [27].

Unlike, a forward error analysis, which consists of the study of the errors $y_1 - \varphi_h(y_0)$ (*local error*) and $e_n = y_n - \varphi_{nh}(y_0)$ (*global error*) in the solution space, the idea of backward error analysis is to search for a modified differential equation $\dot{\tilde{y}} = f_h(\tilde{y})$ of the form

$$\dot{\tilde{y}} = f(\tilde{y}) + hf_2(\tilde{y}) + h^2f_3(\tilde{y}) + \ldots, \qquad (2.9)$$

such that $y_n = \tilde{y}(nh)$, and in studying the difference of the vector fields $f(y)$ and $f_h(y)$. This approach then gives much insight into the qualitative behavior of the numerical solution and into the global error $y_n - y(nh) = \tilde{y}(nh) - y(nh)$. We remark that the series (2.9) usually diverges and that one has to truncate it suitably. For the moment we content ourselves with a formal analysis without taking care of convergence issues. The idea of interpreting the numerical solution as the exact solution of a modified equation is common to many numerical analysts [28].

Theorem 3. *[28] Suppose that the method $y_{n+1} = \Phi_h(y_n)$ is of order r, i.e.,*

$$\Phi_h(y) = \varphi_h(y) + h^{r+1}\delta_{r+1}(y) + \mathcal{O}(h^{r+2}),$$

where $\varphi_t(y)$ denotes the exact flow of $\dot{y} = f(y)$, and $h^{r+1}\delta_{r+1}(y)$ the leading term of the local truncation error. The modified equation then satisfies

$$\dot{\tilde{y}} = f(\tilde{y}) + h^r f_{r+1}(\tilde{y}) + h^{r+1}f_{r+2}(\tilde{y}) + \cdots, \qquad \tilde{y}(0) = y_0$$

with $f_{r+1}(y) = \delta_{r+1}(y)$.

A first application of the modified equation is the existence of an asymptotic expansion of the global error. Indeed, by the nonlinear variation of constants formula, the difference between its solution $y(t)$ and the solution $\tilde{y}(t)$ of $\dot{\tilde{y}} = f_h(\tilde{y})$ satisfies

$$\tilde{y}(t) - y(t) = h^r e_r(t) + h^{r+1} e_{r+1}(t) + \ldots,$$

where $e_i(t)$ are terms of the global truncation error. Since $y_n = \tilde{y}(nh) + \mathcal{O}(h^n)$ for the solution of a truncated modified equation, this proves the existence of an asymptotic expansion in powers of h for the global error $y_n - y(nh)$.

Theorem 4. *[28](Modified Hamiltonians) If a symplectic method Φ_h is applied to a Hamiltonian system with a smooth Hamiltonian $H : \mathbb{R}^{2d} \mapsto \mathbb{R}$, then the modified equation (2.9) is also Hamiltonian. More precisely, there exist smooth functions $H_j : \mathbb{R}^{2d} \mapsto \mathbb{R}$ for $j = 2, 3, \ldots$ such that $f_j(y) = J^{-1} \nabla H_j(y)$.*

The result of the above theorem allows us to estimate the energy conservation of the symplectic numerical schemes via the difference between the Hamiltonian of the system and the Hamiltonian preserved by numerical method as we show in the next section.

Modified Hamiltonian functions.

The motivation for studying numerically the conservation properties of these *modified Hamiltonians* are multifaceted [54], e.g. numerical evidence for the existence of a Hamiltonian for a particular calculation, exposure of energy drifts caused by numerical instability, etc.. Skeel and Hardy [54] proposed a simple strategy for deriving highly accurate estimates for modified Hamiltonians. Since these modified Hamiltonians approximate well the true Hamiltonian, they are referred as "shadow" Hamiltonians $\tilde{H}$ (2.8), cf. [18]. The existence of these shadow Hamiltonians guarantees the boundedness of the error in the symplectic map, in fact we have $\tilde{H}(q, p, h) \to H(q, p)$ for $h \to 0$.

Conversely, if one starts from a given numerical solver then it is well known that any symplectic integrator different from the Hamiltonian flow itself does not preserve the Hamiltonian however a nearby system, the so-called *shadow Hamiltonian* $\tilde{H}$ is conserved. The energy computed from the shadow Hamiltonian of a symplectic integrators differs by $H(q, p) - \tilde{H}(q, p, h) = \mathcal{O}(h^r)$ from the true Hamiltonian [30], with r being the order of the integration scheme. Hence, computing the shadow Hamiltonian of a symplectic integrator is equivalent to determining the order of the integrator.

To compute a shadow Hamiltonian it is necessary to expand an exponential map to a *Hausdorff series*. Let $\hat{T}$ and $\hat{V}$ be a random (usually non-commuting) operators then the *Baker-Campbell-Hausdorff (BCH) formula* [32] reads

$$\ln(e^{\hat{T}} e^{\hat{V}}) = \sum_{n=1}^{\infty} c_n(T, V), \tag{2.10}$$

where the coefficients c_n are recursively determined from the relations $c_1 = T + V$ and

$$(n+1)c_{n+1} = \sum_{m=1}^{\lfloor n/2 \rfloor} \frac{B_{2m}}{(2m)!} \sum_{k_1,\ldots,k_{2m} \geq 1} \mathrm{adc}_{k_1} \ldots \mathrm{adc}_{k_{2m}}(T + V) - \frac{1}{2}(\mathrm{adc}_n)(T - V),$$

for $n \geq 0$, where $\mathrm{ad}\, a : b \mapsto [a, b]$ and B_n denote the Bernoulli numbers.

For a differential equation

$$\dot{y} = f^{[1]}(y) + f^{[2]}(y),$$

it is convenient to study the composition of the flows $\varphi_t^{[1]}$ and $\varphi_t^{[2]}$ of the systems

$$\dot{y} = f^{[1]}(y) \qquad\qquad \dot{y} = f^{[2]}(y), \qquad\qquad (2.11)$$

respectively.

Let us introduce *a Lie derivative*, the differential operators of the following form

$$\mathrm{D}_i = \sum_j f_j^{[i]}(y) \frac{\partial}{\partial y_j},$$

which means that for some differentiable functions $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ we have

$$\mathrm{D}_i F(y) = F'(y) f^{[i]}(y).$$

It follows from the chain rule that, for the solutions $\varphi_t^{[i]}(y_0)$ of the system (2.11),

$$\frac{\mathrm{d}}{\mathrm{d}t} F\left(\varphi_t^{[i]}(y_0)\right) = (\mathrm{D}_i F)\left(\varphi_t^{[i]}(y_0)\right) \qquad\qquad (2.12)$$

and iteratively applying this operator we obtain

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} F\left(\varphi_t^{[i]}(y_0)\right) = (\mathrm{D}_i^k F)\left(\varphi_t^{[i]}(y_0)\right),$$

and Taylor series of $F\left(\varphi_t^{[i]}(y_0)\right)$ at $t = 0$ yield

$$F\left(\varphi_t^{[i]}(y_0)\right) = \sum_{k \geq 0} \frac{t^k}{k!} (\mathrm{D}_i^k F)(y_0) = \exp(t\mathrm{D}_i) F(y_0).$$

Setting $F(y) = \mathrm{Id}(y) = y$ to be the identity map, we can see that this is the Taylor series of the solution itself

$$\varphi_t^{[i]}(y_0) = \sum_{k \geq 0} \frac{t^k}{k!} (\mathrm{D}_i^k \, \mathrm{Id})(y_0) = \exp(t\mathrm{D}_i) \, \mathrm{Id}(y_0).$$

Lemma 1. *[22] Let $\varphi_s^{[2]}$ and $\varphi_t^{[1]}$ be the flows of the differential equations $\dot{y} = f^{[1]}(y)$ and $\dot{y} = f^{[2]}(y)$, respectively. For their composition we then have*

$$\left(\varphi_s^{[2]} \circ \varphi_t^{[1]}\right)(y_0) = \mathrm{e}^{sD_1} \, \mathrm{e}^{tD_2} \, \mathrm{Id}(y_0).$$

For example, in case of the Hamiltonian function (2.5) for the system (2.3) the standard

formulation of the Störmer-Verlet scheme (leap-frog method) yields

$$p_{n+\frac{1}{2}} = p_n - \frac{h}{2}\frac{\partial V}{\partial q}$$

$$q_{n+1} = q_n + h\frac{\partial T}{\partial p_{n+\frac{1}{2}}}$$

$$p_{n+1} = p_{n+\frac{1}{2}} - \frac{h}{2}\frac{\partial V}{\partial q_{n+1}}.$$

From Lemma 1 it follows that the above method can be rewritten in the following formulation

$$\mathrm{e}^{\frac{h}{2}\hat{V}}\,\mathrm{e}^{h\hat{T}}\,\mathrm{e}^{\frac{h}{2}\hat{V}},$$

where exponential operators represent shifts in momenta and coordinates of the standard formulation of the method. Then by using the BCH formula (2.10) we can obtain its shadow Hamiltonian, which is given by

$$\tilde{H} = H - \frac{h^2}{24}\Big(2\big[V,[T,V]\big] + \big[T,[T,V]\big]\Big) + \mathcal{O}(h^4)$$

and it is of the second order accuracy.

We later use the exponential operator formulation for the newly developed integrators since it turns the derivation of the modified Hamiltonians more clear and notation universal for the case of Abelian and non-Abelian structures as we demonstrate further though this thesis.

2.5 Numerical time integrators for ODEs

In this section we briefly define the main classes of numerical integrators that have been found to exhibit useful geometric properties.

2.5.1 Splitting and composition methods

With phase space $\mathcal{M}$ for the system

$$\dot{y} = f(y), \qquad\qquad y(0) = 0, \qquad\qquad y \in \mathcal{M},$$

splitting methods involve three steps [37]:

1. choosing a set of vector fields $f^{[i]}(y)$ such that $f = \sum_{i=1}^{n} f^{[i]}(y)$

2. integrating either exactly or approximately each $f^{[i]}(y)$ with $\varphi_h^{[i]}$

3. combining these solutions to yield an integrator for f.

The order can be increased by including more operators in a time step. For a splitting into two parts, $f(y) = f^{[1]}(y) + f^{[2]}(y)$, we have the general non symmetric composition

$$\varphi_{a_K h}^{[1]} \circ \varphi_{b_K h}^{[2]} \circ \cdots \circ \varphi_{a_1 h}^{[1]} \circ \varphi_{b_1 h}^{[2]} \circ \varphi_{a_0 h}^{[1]}. \tag{2.13}$$

By convention, we only count the evaluations of the flow of $\varphi_h^{[2]}$, and refer to equation (2.13) as an k-stage method. The number of stages and the coefficients a_k and b_k are to be chosen to ensure that the method has some order r. The method Φ_h is time symmetric if one of the following conditions holds $a_1 = 0$, $a_{k+1} = a_{K-k+1}$, $b_k = b_{K-k+1}$ or $a_k = a_{K-k+1}$, $b_k = b_{K-k}$ and $b_K = 0$ [42].

It is easy to find time-symmetric methods, e. g. if Φ_h is any method of order r, then $\Phi_{\frac{1}{2}h}\Phi_{-\frac{1}{2}h}$ is time-symmetric and of order at least $r + 2$ (if r is even) or at least $r + 1$ (if r is odd). In general, if Φ_h is an explicit method, then Φ_{-h} is implicit. However, if Φ is a composition of (explicitly given) flows, then Φ_{-h} is also explicit [37].

The advantages of splitting methods are that they are usually simple to implement, explicit, and can preserve a wide variety of structures. Their disadvantages are that, except in particular cases, higher-order methods are relatively expensive, and that the splitting may violate some special property, such as a symmetry, that one may want to preserve.

2.5.2 Runge-Kutta methods

Runge–Kutta (RK) methods are defined for systems with linear phase space $\mathbb{R}^n$. For the system

$$\dot{y} = f(y), \qquad\qquad y(0) = 0, \qquad\qquad y \in \mathbb{R}^n,$$

the *s-stage RK method* with parameters a_{ij}, b_i for $i, j = 1, \ldots, s$ is given by

$$k_i = f\left(y_k + h\sum_{j=1}^{s} a_{ij}k_j\right), \qquad\qquad y_{k+1} = y_k + h\sum_{j=1}^{s} b_j k_j. \qquad (2.14)$$

RK methods are "linear", that is, the map from vector field f to the map $y_k \mapsto y_{k+1}$ commutes with linear changes of variable $y_k \mapsto Ay$. (Alternatively, the method is independent of the basis of $\mathbb{R}^n$). This implies, for example, that all RK methods preserve all linear symmetries of the system. They are explicit if $a_{ij} = 0$ for $j \geq i$, although they cannot then be symplectic. The most useful RK methods in geometric integration are the Gaussian methods. They are implicit, A-stable, possess the maximum possible order for an s-stage RK method (namely $2s$), preserve all quadratic first integrals of f, and are symplectic for canonical Hamiltonian systems [10].

If we fix a basis in $\mathbb{R}^n$, writing $y = (y^1, y^2, \ldots, y^n)$, we can define the partitioned Runge–Kutta (PRK) methods, in which a different set of coefficients are used for each component of y

$$k_i^l = f^l\left(y_k^l + h\sum_{j=1}^{s} a_{ij}^l(k_j)\right), \qquad\qquad y_{k+1}^l = y_k^l + h\sum_{j=1}^{s} b_j^l k_j. \qquad (2.15)$$

Some PRK methods are splitting methods, but apart from these, the most useful PRK methods in geometric integration are the Lobatto IIIA-IIIB methods, of order $2s - 2$. These are symplectic for separated Hamiltonian systems and have two sets of coefficients,

one for the q and one for the p variables.

2.5.3 Projection methods

Some properties can be preserved by simply taking a step of an arbitrary method and then enforcing the property. Integrals and weak integrals (invariant manifolds) can be preserved by projecting onto the desired manifold at the end of a step or steps, typically using orthogonal projection with respect to a suitable metric. For example, energy-preserving integrators have been constructed using discrete gradient methods, a form of projection [25]. Although it is still used, projection is something of a last resort, as it typically destroys other properties of the method (such as symplecticity) and may not yield good long-time behavior. Reversibility is an exception, for R-reversibility can be imposed on the map Φ by using $\Phi R \Phi^{-1} R^{-1}$, where $R : M \to M$ is an arbitrary diffeomorphism of the phase space. Since this is a composition, it can preserve the group properties of Φ such as symplecticity. Symmetries are a partial exception; the composition $\Phi S \Phi S$ is not S-equivalent, but it is closer to S-equivalent than Φ is, when $S^2 = 1$.

2.5.4 Variational methods

Many ODEs and PDEs of mathematical physics are derived from variational principles with natural discrete analogs. For an ODE with Lagrangian $L(q, \dot{q})$ one can construct an approximate discrete Lagrangian $L_d(q_0, q_1)$ such as $L_d(q_0, q_1) = L_d(q_0, (q_1 - q_0)/\tau)$ from an integrator by requiring that the discrete action $\sum_{i=0}^{N} L_d(q_i, q_{i+1})$ be stationary with respect to all variations of the q_i, for $i = 1, \ldots, N - 1$, with fixed end-points q_0 and q_N. For regular Lagrangian, the integrator can be seen to be symplectic in a very natural way, and in fact the standard symplectic integrators such as leapfrog and the symplectic Runge–Kutta methods can be derived in this way [38]. The advantage of the discrete Lagrangian approach is that it acts as a guide in new situations. The Newmark method, well known in computational engineering, is variational, a fact that allowed the determination of the (nonstandard) symplectic form it preserves; singular Lagrangian can be treated; it suggests a natural treatment of holonomic (position) and nonholonomic (velocity) constraints and of non smooth problems involving collisions; and powerful "asynchronous" variational integrators can be constructed, which use different, even incommensurate time steps on different parts of the system. In these situations variational integrators appear to be natural, and to work extremely well in practice, even if the reason for their good performance (e.g., by preserving some geometric feature) is not yet fully understood [35].

2.5.5 Linear multistep methods

Defined on a linear space $M \in \mathbb{R}^n$ s-step method is

$$\sum_{j=0}^{s} \alpha_j y_{k+j} = \tau \sum_{j=0}^{s} \beta_j f(y_{k+j}). \tag{2.16}$$

Such methods define a map on the product space M^s, and can sometimes preserve a structure (e.g. a symplectic form) in this space. However, this will not usually give good long-time behavior of the sequence of points $y_k \in M$. Instead, one considers the so-called underlying one-step method $\Phi_h : M \rightarrow M$, which is defined such that the sequence of points $y_k := \Phi_h(y_0)$ satisfies the multistep formula. (It always exists, at least as a formal power series in τ). Often the dynamics of Φ dominates the long-term behavior of the multistep method. Recently it has been proved that the underlying one-step methods for a class of time-symmetric multistep methods for second-order problems $\ddot{y} = f(y)$ are conjugate to symplectic, which explains their near-conservation of energy over long times and their practical application in solar system dynamics [11].

2.6 Numerical integration on Lie groups

For certain applications numerical time integration of ODEs on manifold is required. There are a number of techniques to preserve geometric properties of the numerical flow and numerical integrators on Lie groups play a very important role. In this section we give a short insight for numerical treatment of ODEs defined on Lie groups.

First let $\mathcal{M} \in \mathbb{R}^d$ be a given manifold then the system of ODEs (2.1)

$$\dot{y} = f(y), \qquad\qquad y(0) = 0, \qquad\qquad y \in \mathbb{R}^n,$$

is a differential equation on the manifold $\mathcal{M}$, if $y_0 \in \mathcal{M}$ implies $y(t) \in \mathcal{M}$ for all t. This is equivalent to the requirement on the vector field that $f(y) \in \mathcal{T}_y\mathcal{M}$ for $y \in \mathcal{M}$, where $\mathcal{T}_y\mathcal{M}$ is the tangent space of $\mathcal{M}$ at the point $y \in \mathcal{M}$ and for a manifold given by [26]

$$\mathcal{M} = \{y \in \mathbb{R}^d | g(y) = 0\},$$

the tangent space takes the form

$$\mathcal{T}_y\mathcal{M} = \{v \in \mathbb{R}^d | g'(y)v = 0\}.$$

Then a Lie group is a differentiable manifold G, for which the product is a differentiable mapping $G \times G \rightarrow G$. The tangent space $\mathfrak{g} = T_I G$ at the identity I of a matrix Lie group G is closed under commutators of its elements. This makes $\mathfrak{g}$ an algebra, the Lie algebra of the Lie group G [28]. As an example of a Lie group, which can be widely used in quantum field theory, is the *special unitary group* $\mathrm{SU}(n)$ of $n \times n$ unitary matrices with a determinant equals to 1.

Next two lemmas introduce some important properties of Lie groups.

Lemma 2. *Let G be a matrix Lie group and let $\mathfrak{g} = T_I G$ be the tangent space at the identity. The Lie bracket (or commutator)*

$$[A, B] = AB - BA \tag{2.17}$$

defines an operation $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ which is bilinear, skew-symmetric and satisfies the Jacobi identity

$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0 \tag{2.18}$$

Lemma 3. *Consider a matrix Lie group G and its Lie algebra $\mathfrak{g}$. The matrix exponential is a map*

$$\exp : \mathfrak{g} \to G, \tag{2.19}$$

i. e., for $A \in \mathfrak{g}$ we have $\exp(A) \in G$.

Further in this section we present some general ideas for numerical treatment of the problems defined on the Lie groups.

2.6.1 Methods based on Magnus expansion

Let us consider the linear matrix differential equations

$$\dot{Y} = A(t)Y, \tag{2.20}$$

where the matrix A depends constantly on t. For the scalar case the solution of (2.20) with $Y(0) = Y_0$ is given by

$$Y(t) = \exp\left(\int_0^t A(\tau)\mathrm{d}\tau\right) Y_0. \tag{2.21}$$

Defining the inverse of the derivative of the matrix exponential by

$$\mathrm{d}\exp_\Omega^{-1}(\Lambda) = \sum_{k \geq 0} \frac{B_k}{k} \mathrm{ad}_\Omega^k(\Lambda), \tag{2.22}$$

where B_k are the Bernoulli numbers, and $\mathrm{ad}_\Omega(A)$ is the adjoint operator, we can formulate Theorem 5

Theorem 5. *The solution of the differential equation (2.20) can be written as $Y(t) = \exp(\Omega(t))Y_0$ with $\Omega(t)$ defined by*

$$\dot{\Omega} = d\exp_\Omega^{-1}(A(t)), \qquad \qquad \Omega(0) = 0 \tag{2.23}$$

As long as $\|\Omega(t)\| < \pi$, the convergence of the $d\exp_\Omega^{-1}$ expansion (2.22) is assured.

The so-called *Magnus expansion* is obtained via integrating (2.23) and yields

$$\begin{aligned}
\Omega(t) = \int_0^t A(\tau)\mathrm{d}\tau &- \frac{1}{2} \int_0^t \left[\int_0^\tau A(\sigma)\mathrm{d}\sigma, A(\tau)\right] \mathrm{d}\tau \\
&+ \frac{1}{4} \int_0^t \left[\int_0^\tau \left[\int_0^\sigma A(\mu)\mathrm{d}\mu, A(\sigma)\right] \mathrm{d}\sigma, A(\tau)\right] \mathrm{d}\tau \\
&+ \frac{1}{2} \int_0^t \left[\int_0^\tau A(\sigma)\mathrm{d}\sigma, \left[\int_0^\tau A(\mu)\mathrm{d}\mu, A(\tau)\right]\right] \mathrm{d}\tau + \dots.
\end{aligned} \tag{2.24}$$

The truncated series (2.24) give an excellent approximation to the solution of (2.20),

cf. [28].

In order to approximate (2.20) it is proposed in [29] to use $Y_{n+1} = \exp(h\Omega_n)Y_n$, where Ω_n is an approximation of $\Omega(h)$ given by (2.24) with $A(t_n + \tau)$ instead of $A(\tau)$ and the collocation approach yields to replace $A(t)$ with

$$\hat{A}(t) = \sum_{i=1}^{s} l_i(t)A(t_n + c_i h)$$

and to solve $\dot{Y} = \hat{A}(t)Y$ on $[t_n, t_n + h]$ by use of (2.24).

It is possible to construct methods of arbitrary high order with reduced number of commutators [26].

2.6.2 Crouch-Grossman methods

The idea of RK methods (2.14) applied to differential equations (2.20) on Lie groups has the disadvantage since for any $Y \in G$ and $Z \in G$ the update of the form $Y + haA(Z)Z$ is not general in the Lie group G. The replacement of the above update operation with $\exp(haA(Z))Y$ was proposed to resolve this disadvantage in [15].

Definition 7. *Let b_i, a_{ij} $(i, j = 1, \ldots, s)$ be real numbers. An explicit s-stage Crouch-Grossman method is given by*

$$Y^{(i)} = \exp(ha_{i,i-1}K_{i-1}) \ldots \exp(ha_{i1}K_1)Y_n, \qquad K_i = A(Y^i),$$
$$Y_{n+1} = \exp(hb_sK_s) \ldots \exp(hb_1K_1)Y_n.$$

The Crouch-Grossman methods yield the approximation Y_n which lie exactly on the manifold defined by a Lie group and the accuracy order is defined in Theorem 6

Theorem 6. *[29] Let $c_i = \sum_j a_{ij}$. A Crouch-Grossman method has order p $(p \leq 3)$ if the following order conditions are satisfied:*

$$order \ \ 1: \qquad \sum_i b_i = 1$$

$$order \ \ 2: \qquad \sum_i b_1 c_i = \frac{1}{2}$$

$$order \ \ 3: \qquad \sum_i b_i c_i^2 = \frac{1}{3}$$

$$\sum_{ij} b_i a_{ij} c_j = \frac{1}{6}$$

$$\sum_i b_i^2 c_i + 2\sum_{i<j} b_i c_i b_j = \frac{1}{3}.$$

We note that the construction of the high order Crouch-Grossman methods is very complicated [28].

2.6.3 Munthe-Kaas methods

The main purpose to develop this class of numerical methods was to develop a theory of RK methods in a coordinate free framework [28].

Let us consider the problem (2.20) with $A(Y) \in \mathfrak{g}$ for $Y \in G$. From Theorem 5 we know that the solution of (2.20) can be written as $Y(t) = \exp(\Omega(t))Y_0$, where $\Omega(t)$ is the solution of $\dot{\Omega} = \mathrm{d}\exp_{\Omega}^{-1}(A(Y(t)))$, $\Omega(0) = 0$.

Suitably truncating the series (2.24) and considering the differential equation

$$\dot{\Omega} = A\left(\exp(\Omega)Y_0\right) + \sum_{k \geq 0}^{q} \frac{B_k}{k} \operatorname{ad}_{\Omega}^k \left(A\left(\exp(\Omega)Y_0\right)\right), \qquad \Omega(0) = 0. \qquad (2.25)$$

yields the following algorithm.

Assume that Y_n lies in the Lie group G, then, the step $Y_n \mapsto Y_{n+1}$ can be defined as follows

- consider the differential equation (2.25) with Y_n instead of Y_0 and apply a RK method (2.14) to get an approximation $\Omega_1 \approx \Omega(h)$

- define the numerical solution by $Y_{n+1} = \exp(\Omega_1)Y_n$.

Important properties of the Munthe-Kass methods are given in the two following theorems.

Theorem 7. *[29] Let G be a matrix Lie group and $\mathfrak{g}$ its Lie algebra. If $A(Y) \in \mathfrak{g}$ for $Y \in G$ and if $Y_0 \in G$, then the numerical solution of the Lie group Munthe-Kass method lies in G, i. e., $Y_n \in G$ for all $n = 0, 1, 2, \ldots$.*

Theorem 8. *[29] If the RK method is of (classical) order p and if the truncation index in (2.25) satisfies $q \geq r - 2$, then the method of this type is of the order r.*

Every classical RK method defines a Munthe-Kaas method of the same order, unlike Crouch-Grossman methods, which need more stages for the same order [28].

Munthe-Kaas methods are in general not symmetric, even if the underlying Runge–Kutta method is symmetric [28], but symmetric versions of these methods have been developed [62]. As well as the symplectic versions based on the variational principles were recently presented in [7]. But so far there is no time-reversible and symplectic Lie group method.

It is important to mention that the projection methods are suited for the numerical treatment of equation of the type (2.20) since it is a special case of differential equation on the manifolds. As well as the splitting methods with a small modification (we introduce later) are capable to take care of problems on the Lie groups. In this thesis we use these two ideas as a benchmark of our research and to study our main application, which we present in the next chapter.

3 Chapter 3
QUANTUM FIELD THEORIES ON THE LATTICE

Quantum field theory is a set of ideas and tools that combines three of the major themes of modern physics: the quantum theory, the field concept and the principle of relativity. The theory underlies modern elementary particle physics and supplies essential tools to nuclear physics, atomic physics, condensed matter physics and astrophysics [45]. It summarizes the knowledge about the fundamental forces of electromagnetism, as well as the weak and strong interactions and it has been tested over an extremely wide range of length or energy scales. Quantum field theory includes a vast number of physical theories, such as quantum electrodynamics, which describes all phenomena of both electromagnetic and weak interactions, and quantum chromodynamics, which focus on the strong interactions between quarks and gluons [23].

Due to the author's involvement in the Marie Curie Initial Training Network STRONGnet on *Strong Interaction Supercomputing Training Network* and later in the project B5 within the SFB/Transregio 55 *Hadronenphysik mit Gitter-QCD*, the main goal of this thesis lies within the application to the theory of lattice quantum chromodynamics.

In this chapter we briefly introduce basic concepts of quantum field theory, show the quantization approach from classical field theory up to quantum fields. Also very shortly the idea of regularization path integral formulation of quantum field theory on the lattice is presented as well as couple examples of quantum field theories. Later we show one of the way to solve a path integral on the lattice, namely the hybrid Monte Carlo algorithm. Finally, we discuss the challenges for numerical integration and the reasons for studying structure preserving geometric numerical time integrators.

3.1 Basic concepts of quantum field theories

Quantum field theory in its function integral formulation follows from two classical concepts of physics: classical field theory and quantum mechanics. We do not consider here the complete path from point mechanics to quantum field theory, but rather just shortly introduce the basic ideas of the underlying theories for better understanding of the quantum field theory itself.

Classical field theory represent the generalization of point mechanics to n degrees of freedom in each space-time point $x = (t, \mathbf{x})$ and a system with infinite number of degrees of freedom is described by a field $\phi(x)$. The examples of fields are Abelian gauge fields $A_\mu \in \mathbb{R}$, describing the electromagnetic potential, where $\mu = 0, \ldots, 3$ and has $n = 4$ degrees of freedom with two of them corresponding to U(1) gauge group; and non-Abelian gauge fields A_μ^a, where $\mu = 0, \ldots, 3$, $a = 1, \ldots, 8$ and $n = 16$, represents a gluon field and

corresponds to gauge group SU(3).

Adding formalism of quantum mechanics we obtain the representation of probability amplitude for a particle to move from point x to point x' within the time interval T and $\hbar = 1$

$$\langle x' | \, \mathrm{e}^{-i\hat{H}T} \, | x \rangle, \tag{3.1}$$

where a state $|x\rangle$ of the position operator is a vector in Hilbert space and $\hat{H}$ represents the Hamilton operator. Here standard *bra-ket* notation $\langle | \rangle$ defines the scalar product.

The propagator given in (3.1) can be approximated as following Feynman path integral [39]

$$\langle x' | \, \mathrm{e}^{-i\hat{H}T} \, | x \rangle = \int Dx \, \mathrm{e}^{iS[x]} \,. \tag{3.2}$$

Here the functional measure Dx represents

$$Dx = \lim_{N \to \infty} \left(\frac{m}{2\pi i \epsilon} \right)^{N/2} dx_1 \cdots dx_{N-1} \tag{3.3}$$

and represents the integration over all possible paths from (t, x) to (t', x'). The functional formalism has been proved to be very useful in field theory since many results can be derived in a compact and easy way through formal manipulations of functional integrals [39].

Introducing the concept of Euclidean time $T \to -i\tau$ gives us a different representation for the path integral (3.2)

$$\langle x' | \, \mathrm{e}^{-\hat{H}\tau} \, | x \rangle = \int Dx \, \mathrm{e}^{-S_E[x]}, \tag{3.4}$$

where the functional measure Dx differs from one in (3.3) and given by

$$Dx = \lim_{N \to \infty} \left(\frac{m}{2\pi \epsilon} \right)^{N/2} dx_1 \cdots dx_N \tag{3.5}$$

and the Euclidean action S_E is related to the action S through

$$S = iS_E.$$

The path integral (3.4) is real now and has an integrand, which is damped for widely oscillating paths with a large Euclidean action S_E [39].

It is possible, according to quantum statistical mechanics, to extract thermal expectation values for some chosen operator $\hat{\mathbb{O}}$ using the path integral formulation (3.4) in the following way

$$\langle \hat{\mathbb{O}} \rangle = \frac{1}{Z} \int Dx \mathbb{O}(x) \, \mathrm{e}^{-S_E[x]} \,. \tag{3.6}$$

Here Z defines a partition function known from statistical mechanics

$$Z = \mathrm{Tr} \, \mathrm{e}^{-\beta \hat{H}} = \int Dx \, \mathrm{e}^{-S_E}, \tag{3.7}$$

where $\beta = T$ and the integral measure Dx is given by (3.5). Particularly it is interesting

to consider 2-point functions (correlators) of operator at different instances in Euclidean time

$$\langle \mathbb{O}(x(0)\mathbb{O}(x(t)))\rangle = \frac{1}{Z} \int Dx \mathbb{O}(x(0))\mathbb{O}(x(t)) \, e^{-S(x)}$$

especially connected 2-point functions presented by

$$\lim_{\beta \to 0} \left[\langle \mathbb{O}(x(0)\mathbb{O}(x(t)))\rangle - |\langle \mathbb{O}(x)\rangle|^2 \right] = |\langle 0|\hat{\mathbb{O}}|1\rangle|^2 \, e^{-(E_1 - E_0)t} \, . \tag{3.8}$$

The connected 2-point function (3.8) decays exponentially at large Euclidean time separations. By measuring this decay, we obtain in practice the energy gap $E_1 - E_0$, which represents the mass of the lightest particle involved [59].

Since field theories are represented as systems with infinitely many degrees of freedom with a certain number given per space point. Their quantization is a subtle issue because too naive approaches lead to divergent results. In order to avoid meaningless divergent results, quantum field theories must be regularized by introducing an ultraviolet cut-off. In order to properly define a quantum field theory one must also specify the integration measure $D\phi$ of the fields in a path integral (3.4) [59].

One approach is to expand the path integral (3.4) in powers of the coupling constant β. The resulting Feynman diagrams are then regularized order by order in the coupling. This perturbation approach to field theory has led to impressive results in weakly interacting theories. Still, even at weak coupling the perturbation approach to field theory is not entirely satisfactory. It is known that perturbation theory is only an asymptotic expansion. The sum of all orders is divergent and thus does not define the theory beyond perturbation theory. Even more important, for strongly coupled theories, like QCD at low energies, the perturbation regularization is completely useless [45].

The lattice regularization provides a clean way of doing this by replacing the space-time continuum with a discrete mesh of lattice points

$$Z_{reg} = \prod_{x \in lat} \int D\phi_{reg} e^{-S_E[\phi]}.$$

One should not regard the lattice as an approximation to the continuum theory. It rather provides a definition of a theory that is undefined directly in the continuum. Of course, in order to recover the continuum limit, the theory must be renormalized by letting the lattice spacing tend to zero while adjusting the bare coupling constants appropriately to obtain finite expectation values

$$\langle \mathbb{O} \rangle_{reg} = \frac{1}{Z_{reg}} \int D\phi_{reg} \mathbb{O}(\phi(x)) \, e^{-S_E[\phi]} \, .$$

The observable operators $\mathbb{O}$ can be operators that create or annihilate states, or operators that measure observables or combinations of all of these.

New quantization prescription, which is based on the above mentioned metric, states that it is no longer implemented by enforcing canonical commutation rules for the commutators, but instead by a path integral over classical field variables.

The steps involved in the quantization of some system can be formulated as in [19]:

1: Replace the continuum space time by an Euclidean lattice with a lattice constant a. The degrees of freedom are the classical field variables ϕ living on the lattice.

2: Discretize the Euclidean action $S_E[\phi]$ on the lattice such that in the limit $a \to 0$ the Euclidean continuum action is obtained.

3: The operators that appear in the Euclidean correlator to be studied are translated to functionals by replacing the field operators with the classical lattice field variables.

4: Compute the Euclidean correlation functions (3.8) by evaluating these functionals on some lattice field configuration, weighting them with the Boltzmann factor and integrating over all possible field configurations.

Once physical observables are calculated on the lattice one is interested in their values in the limit $a \to 0$, the so-called *continuum limit*. We will discuss this topic since it not a part of our research.

In the next section we consider the lattice regularization for different kinds of field variables and introduce examples of the quantum field theory as the main application of this work.

3.2 Regularization on the lattice

Lattice field theory is the only regularization scheme, which captures finite couplings for field interaction on a non-perturbative level. By the analogy with quantum mechanical path integral (3.2), defined as a limit of a finite-dimensional integral, which resulted from a discretization in time, the same procedure is applied to quantum field theory by considering the integral (3.4) as a limit of well defined integral over discretized Euclidean space-time [39].

Here we consider a hypercubic lattice Λ with a lattice spacing constant a given by

$$\Lambda = a\mathbb{Z}^4 = \left\{ x \mid \frac{x_\mu}{a} \in \mathbb{Z} \right\}. \tag{3.9}$$

One can consider ϕ^4 theory, which is one of the simplest field theories with interaction terms, which in certain cases does not actually requires regularization in order to show the most important features and techniques of lattice field theory.

For example, for the scalar fields $\phi(x)$, defined on the points $x \in \Lambda$ with the action $S[\phi(x)]$ the partition function Z is given by

$$Z_\Lambda = \int \prod_{x \in \Lambda} D\phi(x) e^{-S_E[\phi(x)]}$$

and observables $\langle \mathbb{O} \rangle_\Lambda$ can be evaluated numerically by means of Monte Carlo simulations.

We will discuss the Monte Carlo algorithm at the end of this chapter in more details but first we briefly introduce a regularization procedure from the continuum to the lattice for different field variables, which are necessary to introduce our main application.

3.2.1 Gauge fields

For simplicity let us consider a complex vector field $\phi(x)$ and some action S, which is invariant under transformations of the form

$$\phi(x) \rightarrow \phi' = \Omega^{-1}\phi(x), \qquad\qquad \Omega \in \mathrm{SU}(N), \qquad\qquad (3.10)$$

where Ω is a $N \times N$ matrix satisfying

$$\Omega^\dagger \Omega = 1, \qquad\qquad \det \Omega = 1. \qquad\qquad (3.11)$$

The first condition (3.10) defines a unitary group $\mathrm{U}(N)$ and the second one (3.11) restricts to a special unitary group $\mathrm{SU}(N)$. The transformation (3.10) is called a global gauge transformation.

The bigger class of transformations called local gauge transformations consist of the global gauge transformations (3.10)

$$\phi(x) \rightarrow \phi' = \Omega^{-1}(x)\phi(x), \qquad\qquad (3.12)$$

where $\Omega(x)$ changes with x.

In order to construct gauge invariant action one has to define so the called *covariant derivative* [19]

$$D_\mu(x) = \partial_\mu + iA_\mu(x), \qquad\qquad (3.13)$$

where $A_\mu(x) \in \mathrm{su}(N)$ is an element of the Lie algebra of $\mathrm{SU}(N)$ and called gauge field.

Then the field strength tensor $F_{\mu\nu}(x)$ corresponds to the curvature tensor and defined as a commutator

$$F_{\mu\nu}(x) = -i[D_\mu(x), D_\nu(x)] = \partial\mu A_\nu(x) - \partial_\nu A_\mu(x) + i[A_\mu(x), A_\nu(x)].$$

The fact that the field strength tensor $F_{\mu\nu}(x)$ is the commutator of two covariant derivatives implies that it inherits the transformation properties of (3.13), i.e., it transforms as

$$F_{\mu\nu}(x) \mapsto F'_{\mu\nu}(x) = \Omega(x)F_{\mu\nu}(x)\Omega(x)^\dagger$$

with Ω satisfying (3.11). The dynamics of the gauge field $A_\mu(x)$ is then introduced by means of the Yang-Mills action

$$S_{YM} = -\frac{1}{2g^2} \int \mathrm{d}^4x \, \mathrm{Tr} \, F_{\mu\nu} F_{\mu\nu}. \qquad\qquad (3.14)$$

For the regularization of the gauge field A_μ on the lattice, let n be a point on the lattice Λ and $n + a\hat{\mu}$ the neighboring point in the direction of the lattice axis $\mu = 1, 2, 3, 4$ with a lattice spacing a. The link $U_{x,\mu}$ on the straight path from x to $x + \mu$ is an element of a gauge group $\mathcal{G}(\mathrm{U}(N)$ or $\mathrm{SU}(N))$ and satisfies

$$U(y, n) = U^{-1}(n, y).$$

The collection of all links $\{U_\mu(n)\}$ represents the lattice gauge field [39].

In order to construct a gauge invariant action for the lattice gauge field the smallest closed loops on the lattice called *plaquettes* are used.

A plaquette $\mathbb{P}_\mu(n)$ consisting of four links and containing points

$$n, \qquad\qquad n + a\hat{\mu}, \qquad\qquad n + a\hat{\mu} + a\hat{\nu}, \qquad\qquad n + a\hat{\nu}$$

presented as a product of links in the following form

$$\mathbb{P}_\mu(n) = U_\mu(n)U_\nu(n + \hat{\mu})U_\mu^\dagger(n + \hat{\nu})U_\nu^\dagger(n), \tag{3.15}$$

where the link on opposite direction is defined by

$$U_{-\mu}(n) = U_\mu^\dagger(n - \hat{\mu})$$

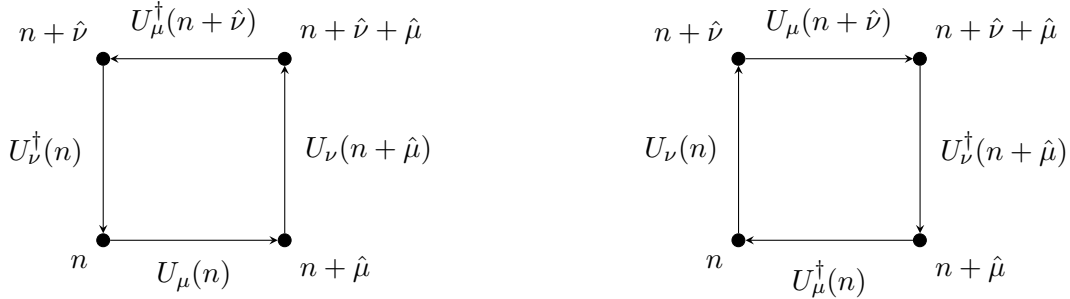and can be visualized as it is shown on Figure 3.1.

Figure 3.1: Plaquette.

The action proposed for the pure gauge theory is called Wilson action and defined in terms of the plaquettes (3.15)

$$S[U] = \sum_{n;\mu,\nu} S_{n;\mu,\nu}(\mathbb{P}_{\mu,\nu}(n)),$$

with the plaquette term

$$S_{n;\mu,\nu} = \beta \left\{ 1 - \frac{1}{N} \operatorname{Re} \operatorname{Tr} U_{\mu,\nu}(n) \right\} \quad \text{for} \quad \mathrm{SU}(N).$$

There are other possibilities for defining gauge invariant actions, but the choice of Wilson action appears to be the simplest one [39].

3.2.2 Fermion fields

Theories with fermions have no immediate classical limit, and the definition of the path integral needs special care [59].

It requires an introduction of the anti-commuting variables η_i with $i \in 1, 2, \ldots, \mathbb{N}$ defined on a so-called Grassmann algebra, which is characterized by the following relations

$$\{\eta_i, \eta_j\} = \eta_i\eta_j + \eta_j\eta_i = 0,$$

such that any function $f(\eta)$ has a representation in terms of finite degree polynomials of the form

$$f(\eta) = f + \sum_i f_i \eta_i + \sum_{ij} f_{ij} \eta_i \eta_j + \sum_{ijk} f_{ijk} \eta_i \eta_j \eta_k + \cdots,$$

where the numbers $f_{ij\ldots l}$ are complex or real numbers and antisymmetric in $i, j, \cdots, l$. Then formal differentiation and integration procedures are also defined correspondingly for differentiation

$$\frac{\partial}{\partial \eta_i} \eta_i = 1, \quad \frac{\partial}{\partial \eta_i} \eta_i \eta_j = \eta_j, \quad \frac{\partial}{\partial \eta_j} \eta_i = -\eta_j,$$

and for integration

$$\int \mathrm{d}\eta_i = 0, \quad \int \mathrm{d}\eta_i \eta_i = 1, \quad \int \mathrm{d}\eta_i \mathrm{d}\eta_i \eta_i \eta_j = -1.$$

The fermion action S_F is presented by

$$S_F[\bar{\phi}\phi] = \int \mathrm{d}^d x \bar{\phi}(\gamma_\mu \partial_\mu + m)\phi, \tag{3.16}$$

with γ_μ are Euclidean Dirac matrices.

The Grassmann algebra is used to define fermion fields generated by independent Grassmann numbers ϕ_n and $\bar{\phi}_n$. The index n varies over all space-time points as well as over all spin, flavor and color indices. On the lattice the continuum field $\bar{\phi}, \phi$ is replaced by Grassmann variables $\phi_n, \bar{\phi}_n$.

There are different ways to discretize the fermion action (3.16). First we introduce the naive discretization of the fermion action (3.16).

Naive lattice fermions. The idea is to discretize the continuum derivative in (3.16) by a finite difference, such that

$$S_N[\bar{\phi}, \phi] = a^d \sum_{n,\mu} \frac{1}{2a} \left(\bar{\phi}_n \gamma_\mu \phi_{n+\hat{\mu}} - \bar{\phi}_{n+\hat{\mu}} \gamma_\mu \phi_n \right) + a^d \sum_n m \bar{\phi}_n \phi_n. \tag{3.17}$$

Unfortunately this approach does not lead to the correct continuum theory, due to the fermion doubling effect, which pose a severe problem in lattice field theory.

There are two basic types of lattice formulation, applied in most cases are the Wilson formulation and the Kogut Susskind staggered formulation [39]. In this thesis we stick to the Wilson fermion formulation as it is sufficient for achieving the main goal of our work.

Wilson fermions. This discretization is based on the naive idea with an introduction of an additional term the so called Wilson term, which gives the fermion doublers a mass of the order of the cut off while the physical fermion remain massless. The discretized action then looks like

$$S_W[\bar{\phi}, \phi] = S_N[\bar{\phi}, \phi] + a^d \sum_{n,\mu} \frac{1}{2a} \left(2\bar{\phi}_n \phi_n - \bar{\phi}_n \gamma_\mu \phi_{n+\hat{\mu}} - \bar{\phi}_{n+\hat{\mu}} \gamma_\mu \phi_n \right),$$

where S_N is a naive discretization given by (3.17).

Here we introduce two important examples of quantum field theories, namely quantum electrodynamics and quantum chromodynamics.

Quantum electrodynamics

Quantum electrodynamics (QED) is a quantum field theory of the electromagnetic force. It successfully describes electromagnetic interactions of electrons, muons etc. to a high precision [23]. As a part of the standard electroweak model QED describes all phenomena of both electromagnetic and weak interactions in the energy range up to 100 GeV. Lattice study of QED mainly motivated by desire to improve the theoretical understanding of the general mathematical properties of this type of quantum field theories.

For example, let us define the Wilson formulation of lattice action for QED.

The gauge field variables $U_\mu(n) \in \mathrm{U}(1)$ are located on the links of the lattice connecting two neighboring sites $(n, n + \hat{\mu})$ and can be written as

$$U_\mu(n) = \mathrm{e}^{ieA_\mu(n)} .$$

Here $A_\mu(n) \in \mathbb{R}$ corresponds to the continuum vector potential and e is the bare gauge coupling.

The fermion part of the QED lattice action with Wilson fermions is presented by

$$S_F = \sum_n \left\{ M(\bar{\phi}_n \phi_n) - K \sum_{\mu=\pm 1}^{\pm 4} (\bar{\phi}_{n+\hat{\mu}}[r + \gamma_\mu]U_\mu(n)\varphi_n) \right\} .$$

The fermion mass in lattice units (am) is given in [39]

$$am + 4r = \frac{M}{2K}.$$

Usually one chooses $K = \frac{1}{2}$ in lattice perturbation theory and $M = 1$ in numerical simulations. By adding N_f such actions one can describe QED with N_f fermion flavors. Abelian gauge fields are given by

$$S_G = \frac{1}{2e^2} \sum_n \sum_{\nu,\mu=1}^4 \left\{ 1 - U_\mu(n)U_\nu(n + \hat{\mu})U_\mu^\dagger(n + \hat{\nu})U_\nu^\dagger(n) \right\} ,$$

where the summation goes over all positively oriented plaquettes of the lattice. The whole lattice QED action is given by the combination of fermion and gauge part.

Despite the fact that QED does not require the lattice regularization since it gives very successful results in the framework of perturbation theory, it is still used to improve the theoretical understanding of the general mathematical properties of this type of quantum field theories [39].

Since QED does not require a large computational effort in comparison to more advance theories and shares the main properties of these systems it is often used as a test example for developing numerical techniques. Later we use a QED model for the numerical testing.

Quantum chromodynamics

Quantum chromodynamics (QCD) is the fundamental quantum field theory of quarks and gluons. QCD is believed by most physicists to be the correct theory of strong nuclear force [39]. Lattice QCD has become a standard tool in elementary particle physics [19].

QCD is a type of quantum field theory, a non-Abelian gauge theory with symmetry group SU(3). The QCD analogue of electric charge is a property called color. Gluons are the force carrier of the theory, like photons are for the electromagnetic force in quantum electrodynamics. The theory is an important part of the Standard Model of particle physics. A large body of experimental evidence for QCD has been gathered over the years [55].

The lattice action of QCD depends on the gauge group field $U_\mu(n) \in$ SU(3) on lattice links and Grassmann fields φ_{qn}, $\bar{\varphi}_{qn}$ on lattice sites. Here n represents a lattice point and $\mu = \pm 1, \pm 2, \pm 3, \pm 4$ are the directions of neighboring points $n + \hat{\mu}$ on the four-dimensional hypercubic lattice.

For example, let us define the Wilson formulation of lattice action for QCD.

The pure gauge lattice action is an sum over all plaquettes

$$S_g[U] = \beta \sum_{n,\mu} \left(1 - \frac{1}{3} \operatorname{Re} \operatorname{Tr} \mathbb{P}_\mu(n) \right),$$

where $\beta = 6g^{-2}$ is the coupling term and the plaquette is given by

$$\mathbb{P}_\mu(n) = U_\mu(n) U_\nu(n + \hat{\mu}) U_\mu^\dagger(n + \hat{\nu}) U_\nu^\dagger(n).$$

The quark fields $\phi_{qn} = \phi_{qn\alpha c}$ and $\bar{\phi}_{qn} = \bar{\phi}_{qn\alpha c}$ have flavor index $q = u, d, c, s, t, b$, Dirac spinor index $\alpha = 1, 2, 3, 4$ and SU(3) color index $c = 1, 2, 3$.

The Wilson action for QCD is

$$S[U, \phi, \bar{\phi}] = S_g[U] + S_q[U, \phi, \bar{\phi}],$$

where the quark field action is

$$S_q[U, \phi, \bar{\phi}] = \sum_{n,\mu} \left\{ (\bar{\phi}_{qn} \phi_{qn}) - K_q \sum_{\mu=\pm 1}^{\pm 4} (\bar{\phi}_{qn+\bar{\mu}}[r_q + \gamma_\mu] U_\mu(n) \phi_{qn}) \right\}.$$

Here the hopping parameter K_q and Wilson parameter r_q can depend on the flavor index q. In most cases the Wilson parameter is chosen to be $r_q = 1$. There are another representations of quark action for QCD such as an idea of staggered fermions [19], which we do not consider here.

The lattice QCD is a very perspective quantum field theory with a lot of room for improvement. Therefore we consider it as the main application of this thesis. In the next section we introduce algorithm to treat the path integral formulation of the quantum field theories regularized on the lattice.

3.3 Hybrid Monte Carlo algorithm

The main purpose of numerical simulations in lattice quantum field theory is to find an estimation of some observable $\mathbb{O}[\phi]$ of the field variables ϕ which is given by the path integral as

$$\langle \mathbb{O} \rangle = \frac{1}{Z} \int [\mathrm{d}\phi] \, \mathrm{e}^{-S[\phi]} \, \mathbb{O}[\phi], \qquad\qquad Z = \int [\mathrm{d}\phi] \, \mathrm{e}^{-S[\phi]}, \qquad (3.18)$$

which leads to systems with a high number of integration variables (around 10^4) [40].

So the only possibility would be to use Monte Carlo integration. It would mean to randomly generate field configurations ϕ in the space of field variables, but since the number of lattice point is very large only a small amount of the free energy density will contribute in the path integral.

Hence the importance sampling is required in the Monte Carlo integration in such way that the distribution of configurations follows the Boltzmann factor $\exp(-S[\phi])$ [16].

The task in numerical simulations is to generate samples consisting of large number N of configurations $\{[\phi_n], \quad n = 1, 2, \cdots, N\}$, in such a way that the distribution within a sample approximates the desired distribution is the canonical ensemble[16].

The sequence of field configuration $\{[\phi_n], \quad n = 1, 2, \cdots, N\}$ is obtained by repeatedly applying an algorithm, which updates a configuration $[\phi_n]$ to the next state $[\phi_{n+1}]$ with a given transition probability $P([\phi'] \leftarrow [\phi])$, which refers to a large number of independent updating.

The most useful technique for generating a sequence of configurations with the desired probability is to construct a Markov process. A Markov process is a stochastic procedure which generates a new configuration $[\phi']$ from its predecessor $[\phi]$ with probability $P_M([\phi'] \leftarrow [\phi])$. Any Markov process converges to a unique fixed point distribution P_S provided that it is ergodic and that it satisfies the detailed balance

$$P_S(\phi)P_M([\phi'] \leftarrow [\phi]) = P_S(\phi')P_M([\phi] \leftarrow [\phi']). \qquad (3.19)$$

The convenient way to construct a Markov process is to choose a new configuration ϕ' with probability $P_C([\phi'] \leftarrow [\phi])$ and then to either accept ϕ' with the probability $P_A([\phi'] \leftarrow [\phi])$ or to reject it and keep the old configuration ϕ.

The Metropolis algorithm [21] provides one choice of P_A to reach a detailed balance (3.19) for any probability P_C in the following generalized procedure

$$P_A([\phi'] \leftarrow [\phi]) = \min\left(1, \frac{P_S(\phi')P_C([\phi] \leftarrow [\phi'])}{P_S(\phi)P_C([\phi'] \leftarrow [\phi])}\right).$$

This approach is not very efficient because in most cases the new value of the action will be larger than the previous one and hence the proposed change would be not accepted [39]. However, if the updating process is defined by a Hamiltonian, then the classical dynamics equations will give new configurations with different field variables, but the value of a Hamiltonian is close to the original one [16].

The HMC algorithm can be briefly summarized in three steps [34]:

- The construction of samples, pairs of links $U \in \Omega = \mathrm{SU}(N)$ and momenta $P \in \hat{\Omega} = \mathrm{su}(N)$ according to the probability distribution $F_{\mathcal{V}}$ with probability density $\mathcal{V}(U,P)$ given by

$$\mathcal{V}(U,P) = \frac{(2\pi)^{N/2}}{Z_H} \exp(-S[U]) \frac{1}{(2\pi)^{N/2}} \exp(-\langle P,P \rangle / 2) = \frac{1}{Z_H} \exp(-H(U,P)),$$

  with $Z_H = \int_{\Omega \times \hat{\Omega}} \mathcal{V}(U,P) \mathrm{d}(U,P)$ and $H(U,P) = \langle P,P \rangle / 2 + S[U]$, where $\langle P,P \rangle = \sum_{x,\mu} \langle P_{x,\mu}, P_{x,\mu} \rangle$ is a natural scalar product.

- The transition from one configuration (U_0, P_0) to the next consists of a proposal step

$$(U_0, P_0) \to g(U_0, P_0) \tag{3.20}$$

  with mapping $g : \Omega \times \hat{\Omega} \to \Omega \times \hat{\Omega}$.

- The acceptance step, where the new configuration $g(U_0, P_0)$ is accepted with probability

$$\begin{aligned} P_A((U_0, P_0), g(U_0, P_0)) &= \min\left(1, \frac{\mathcal{V}(g(U_0, P_0))}{\mathcal{V}(U_0, P_0)}\right) \\ &= \min\left(1, \mathrm{e}^{-(H(g(U_0, P_0)) - H(U_0, P_0))}\right), \end{aligned} \tag{3.21}$$

  otherwise the old configuration (U_0, P_0) is kept as the next entry in the chain.

In order for process (3.21) to satisfy the detailed balance condition (3.19) exactly the dynamics (3.20) must be time-reversible and generate an area-preserving map on the phase space for any value of Δt as it is shown in details in [39]. It can be achieved by defining g to be the Hamiltonian flow and therefore Hamilton's equation of motion has to be solved. In practice equations of motion are integrated by numerical schemes approximately.

Due to the fact that the detailed balance condition (3.19) must be satisfied, it is required from numerical time integrators to preserve some geometric properties of the numerical flow such as time-reversibility and area-preservation.

Also since the acceptance rate of the Hybrid Monte Carlo (HMC) algorithm depends on $\Delta H = H(g(U_0, P_0)) - H(U_0, P_0)$ the higher convergence order of a numerical time integrator or the full conservation of energy $\Delta H = 0$ will be an advantage because in the first case we obtain higher acceptance for the new configurations and in the second case the acceptance rate is a hundred percent $(P_A = 1)$.

In the next two chapters we propose newly developed numerical time integrators, which would satisfy all the requirements to be suitable for the molecular dynamics step of the HMC. In the sequel we will show a detailed analysis.

4 PROJECTION METHODS

As we emphasized in the previous chapter, we are highly interested in the energy conservation property of the numerical time integrators, since it yields hundred percent acceptance rate in the accept/reject step of the HMC algorithm. There is a class of numerical methods we must pay attention to, so-called projection methods. The main advantage of this kind of integrators is to preserve the energy of the system completely. Unfortunately, as it was mentioned before, the numerical scheme for the molecular dynamics step of HMC algorithm has to be also time-reversible and symplectic. And while symmetric (time-reversible) projection methods have been presented in [25], projection integrators preserving both time-reversibility and symplecticity are unknown. In spite of the above mentioned properties symmetric projection methods seem to be a suitable candidate for an improvement of the HMC algorithm, namely to neglect the accept/reject step.

In this chapter, first we introduce the standard approach of projection methods. Then we investigate projection schemes, which allow to interpret both the perturbation and projection step as a one-step scheme with step size μ, applied to the underlying Hamiltonian $H(y)$. We introduce our attempts to construct a new class of projection schemes, which preserve the Hamiltonian and at the same time preserve the physical properties of time-reversibility and symplecticity of the Hamiltonian flow (4.3) via choosing the symmetric, time-reversible and symplectic one-step method Φ_h for the intermediate step of the classic projection algorithm (4.2), preserve these properties and obey the energy conservation properties by choosing suitable perturbation and projection steps of the method (4.2). After that we show our attempts to obtain high-order integrators via investigating modified Hamiltonians of the numerical integrators presented in previous sections, where we try to increase the order of conserved energy by eliminating the low order error terms of its shadow Hamiltonians. At last we show our last try to construct a structure-preserving modification of the classical projection approach and give a short conclusion, based on theoretical and numerical results.

4.1 Introduction in the projection methods

For $t \in [0, T]$, we consider the solution $y(t; y_0)$ of an initial value problem of ODEs in $\mathbb{R}^d$ defined by

$$\dot{y} = f(t, y), \qquad y(0) = y_0, \tag{4.1}$$

lying on a $(d - m)$ submanifold $\mathcal{M} = \{y : g(t, y) = 0\}$ defined by a function $g : \mathbb{R}^d \to \mathbb{R}^m$. Geometric numerical integration aims at determining a numerical approximate to the solution $y(t; y_0)$ fulfilling the constraint defined by $\mathcal{M}$ at the same time. One such approach is given by the projection methods [24], where one step reads as follows:

- compute $\tilde{y}_{n+1} = \Phi_h(y_n)$, where Φ_h is an arbitrary one-step method,

- project the perturbed initial values $\tilde{y}_{n+1}$ onto the manifold $\mathcal{M}$ to obtain the numerical solution $y_{n+1} \in \mathcal{M}$.

This procedure leads to solving the following system

$$
\begin{aligned}
\tilde{y}_{n+1} &= \Phi_h(y_n), & \text{compute with arbitrary } \Phi_h \\
y_{n+1} &= \tilde{y}_{n+1} + G(\tilde{y}_{n+1})^\top \lambda, & \text{project on } \mathcal{M} \\
0 &= g(y_{n+1}), & \text{conservation condition}
\end{aligned}
$$

where $G^\top(y) = g'(y)$ and λ is implicitly defined by $g(y_{n+1}) = 0$.

The projection does not deteriorate the convergence order of the method Φ, because the distance of $\tilde{y}_{n+1}$ to the manifold $\mathcal{M}$ is of the size of the local error.
Unfortunately this general approach lacks of two other geometric properties our numerical integrator must conserve. There exist the suitable symmetric modifications of the standard approach, which allow us to preserve the time-reversibility of the numerical flow.

This approach has been first generalized by Ascher and Reich [2] for Hamiltonian systems and later by Hairer [25] for general systems with constraints to obtain an overall symmetric scheme provided that the underlying one-step scheme is also symmetric:

$$
\begin{aligned}
\tilde{y}_n &= y_n + G(y_n)^\top \mu, & \text{perturbation step} \\
\tilde{y}_{n+1} &= \Phi_h(\tilde{y}_n), & \text{symmetric one-step method} \\
y_{n+1} &= \tilde{y}_{n+1} + G(y_{n+1})^\top \mu, & \text{projection step} \\
g(y_{n+1}) &= 0, & \text{energy preservation condition}
\end{aligned}
\tag{4.2}
$$

where $G(y_n)$ denotes the Jacobian of $g(y_n)$ and μ can be implicitly defined via the constraint $g(y_{n+1}) = 0$. The overall scheme $\Psi_h(y_n)$ is then composed of three different mappings:

$$
\Psi_h(y_n) := (BP)_\mu(\Phi_h(P_\mu(y_n))),
$$

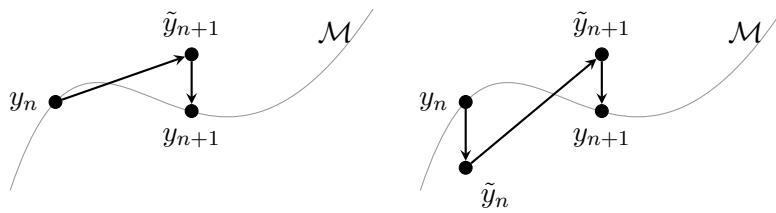with the projection step $(BP)_\mu$ and perturbation step P_μ as defined above in (4.2).

Figure 4.1: Standard Projection Method (left); Symmetric Projection Method (right)

An important special case is given by applying symmetric projection schemes to the classical Hamiltonian equation of motion (4.1) with

$$
f(y) = J^{-1} \nabla H(y), \qquad y = (q, p), \qquad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},
\tag{4.3}
$$

where $\nabla H(y) = H'(y)^\top$, $H(y)$ is a Hamiltonian function, q and p are the generalized coordinates and impulses of the system, and the matrix J denotes the standard symplectic matrix. In this case, the constraint to be preserved is the Hamiltonian itself: $g(y) := H(y(T)) - H(y_n)$.

So far, projection schemes are regarded as one-step schemes Ψ_h, starting from an initial value y_n, and computing a new approximation y_{n+1} with step size h. The parameter μ is only an internal variable fixed by the demand of fulfilling the constraint. Thus the solution Ψ_h depends on the initial value y_n only, with μ implicitly defined by y_n. If one wants to generalize such projection schemes to symplectic or volume-preserving schemes, one has to carefully check the sensitivity of Ψ_h with respect to y_n, and the dependence of μ on y_n has to be included. For such schemes, if existing, also the perturbed solutions satisfy the constraints.

However, when generalizing projection schemes to symplectic and/or volume-preserving schemes, the μ-dependence on y_n is not easy to handle. Taking another point of view, one can get rid of this problem: we regard projection schemes as a class of schemes Ψ_h^μ with step size h, parametrized by μ. Hereby, the optimal parameter μ_0 is defined by $H(y) = \text{const.}$. For one step under consideration, this $\mu = \mu_0$ is then constant, and we use the scheme $\Psi_h^{\mu_0}$. To check for symplecticity or volume preservation, we have to investigate the sensitivity of Ψ_{μ_0} with respect to the initial value with a fixed and constant μ_0. Obviously, the perturbed solutions will not fulfill the energy constraint, but we do not have to care about it.

Taking this point of view, we can easily generalize the symmetric projection scheme (4.2) to time-reversible and volume-preserving schemes: the perturbation and projection step in (4.2), if G evaluated at $J^{-1}y_n$ and $J^{-1}y_{n+1}$ instead of y_n and y_{n+1}, represent an explicit and implicit Euler step with step size μ, resp., applied to a Hamiltonian equation of motion (4.3) with Hamiltonian $\tilde{H}(y) := H(J^{-1}y)$ instead of $H(y)$. If we replace these steps by time-reversible and symplectic one-step schemes such as implicit midpoint rule or leap-frog, the overall scheme Ψ_h^μ, consisting of three one-step schemes with step sizes μ, h and μ applied to the Hamiltonians $\tilde{H}$, H and $\tilde{H}$, resp., inherits these properties for all feasible parameters μ, if the intermediate scheme Φ_h is symmetric, time-reversible and symplectic, too.

4.2 Another view on symmetric projection schemes

In our first approach, we replace the perturbation and projection step in the symmetric projection scheme (4.2) by a symmetric, time-reversible and symplectic one-step method $\tilde{\Phi}_\mu$ of order q, which performs one step of step size μ applied to the Hamiltonian equation of motion (4.3) with the Hamiltonian $\tilde{H}(y) := H(J^{-1}y)$ instead of $H(y)$.

The overall scheme $\Psi_h^\mu(y_n)$ is then composed of three different mappings:

$$\Psi_h^\mu(y_n) := \tilde{\Phi}_\mu(\Phi_h(\tilde{\Phi}_\mu(y_n))). \tag{4.4}$$

Ψ_h^μ is a composition of three symmetric, symplectic and time-reversible schemes with step sizes μ, h and μ, resp. In composition methods, one step size is the multiple of the others, and these properties are preserved. Though this is not the case for Ψ_h^μ, it inherits these

properties:

- *Symmetry:* $\Psi_h^\mu \circ \Psi_{-h}^{-\mu} = \mathrm{id}$.
  We have

$$
\begin{aligned}
\Psi_h^\mu \circ \Psi_{-h}^{-\mu} &= (\tilde{\Phi}_\mu \circ \Phi_h \circ \tilde{\Phi}_\mu) \circ (\tilde{\Phi}_{-\mu} \circ \Phi_{-h} \circ \tilde{\Phi}_{-\mu}) \\
&= (\tilde{\Phi}_\mu \circ \Phi_h \circ \tilde{\Phi}_\mu) \circ ((\tilde{\Phi}_\mu)^{-1} \circ (\Phi_h)^{-1} \circ (\tilde{\Phi}_\mu)^{-1}) \\
&= \mathrm{id}
\end{aligned}
$$

  due to the symmetry of the underlying methods.

- *Time reversibility:* $\Psi_h^\mu \circ \rho \circ \Psi_h^\mu \circ \rho = \mathrm{id}$ with $\rho = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$.
  This condition is equivalent to

$$
\begin{aligned}
\Psi_h^\mu \circ \rho \circ \Psi_h^\mu \circ \rho &= \tilde{\Phi}_\mu \circ \Phi_h \circ \tilde{\Phi}_\mu \circ \rho \circ \tilde{\Phi}_\mu \circ \Phi_h \circ \tilde{\Phi}_\mu \circ \rho \\
&= \tilde{\Phi}_\mu \circ \Phi_h \circ \underbrace{\tilde{\Phi}_\mu \circ \rho \circ \tilde{\Phi}_\mu \circ \rho}_{=\mathrm{id}} \circ \rho \circ \Phi_h \circ \tilde{\Phi}_\mu \circ \rho \\
&= \tilde{\Phi}_\mu \circ \underbrace{\Phi_h \circ \rho \circ \Phi_h \circ \rho}_{=\mathrm{id}} \circ \rho \circ \tilde{\Phi}_\mu \circ \rho \\
&= \mathrm{id}
\end{aligned}
$$

  due to the time-reversibility of the underlying methods.

- *Symplecticity:* $\left(\frac{\partial \Psi_h^\mu(y_n)}{\partial y_n}\right)^\top J \left(\frac{\partial \Psi_h^\mu(y_n)}{\partial y_n}\right) = J$.
  The symplecticity of Ψ_h^μ is a direct consequence of the symplecticity of the underlying methods:

$$
\begin{aligned}
\left(\frac{\partial \Psi_h^\mu}{\partial y_n}\right)^\top J \left(\frac{\partial \Psi_h^\mu}{\partial y_n}\right) &= \left(\frac{\partial \tilde{\Phi}_\mu}{\partial y_n} \frac{\partial \Phi_h}{\partial y_n} \frac{\partial \tilde{\Phi}_\mu}{\partial y_n}\right)^\top J \left(\frac{\partial \tilde{\Phi}_\mu}{\partial y_n} \frac{\partial \Phi_h}{\partial y_n} \frac{\partial \tilde{\Phi}_\mu}{\partial y_n}\right) \\
&= \left(\frac{\partial \tilde{\Phi}_\mu}{\partial y_n}\right)^\top \left(\frac{\partial \Phi_h}{\partial y_n}\right)^\top \underbrace{\left(\frac{\partial \tilde{\Phi}_\mu}{\partial y_n}\right)^\top J \frac{\partial \tilde{\Phi}_\mu}{\partial y_n}}_{=J} \frac{\partial \Phi_h}{\partial y_n} \frac{\partial \tilde{\Phi}_\mu}{\partial y_n} \\
&= \left(\frac{\partial \tilde{\Phi}_\mu}{\partial y_n}\right)^\top \underbrace{\left(\frac{\partial \Phi_h}{\partial y_n}\right)^\top J \frac{\partial \Phi_h}{\partial y_n}}_{=J} \frac{\partial \tilde{\Phi}_\mu}{\partial y_n} \\
&= \left(\frac{\partial \tilde{\Phi}_\mu}{\partial y_n}\right)^\top J \frac{\partial \tilde{\Phi}_\mu}{\partial y_n} \\
&= J.
\end{aligned}
$$

For a given fixed step size h, the optimal μ_{opt} to preserve the Hamiltonian is then fixed by the non-linear system

$$
F(h, y_{n+1}, \mu) = \begin{pmatrix} y_{n+1} - \Psi_h^\mu(y_n) \\ H(y_{n+1}) - H(y_n) \end{pmatrix}, \tag{4.5}
$$

which can be rewritten as scalar equation

$$\tilde{g}_h(\mu) := g(\Psi_h^\mu(y_n)) = H(\Psi_h^\mu(y_n)) - H(y_n) = 0, \tag{4.6}$$

if Ψ_h^μ is explicit.

With $F(0, y_n, 0) = 0$ and

$$\frac{\partial F}{\partial(y_{n+1}, \mu)}(0, y_n, 0) = \begin{pmatrix} I & -2\nabla H(y_n) \\ \nabla H(y_n)^\top & 0 \end{pmatrix} \text{ regular } \Leftrightarrow \nabla H(y_n) \neq 0,$$

the argument in [28] regarding the existence of the numerical solution also applies to our case: for sufficiently small step size h, the existence of the numerical approximation $\Psi_h^{\mu_{\text{opt}}}(y_n)$ with local order $\mathcal{O}(h^{r+1})$ is guaranteed, and $\mu = \mathcal{O}(h^{r+1})$, independent of the order k of Ψ_μ.

For non-separable Hamiltonians, all symplectic schemes with order larger than one are implicit, and one has to solve the nonlinear system (4.5) to obtain the approximate $\Psi_h^{\mu_{\text{opt}}}(y_n)$. However, for separable Hamiltonians, the leap-frog scheme is explicit, and one has to solve the scalar equation (4.6) only. The solvability of this scalar equations directly follows from the argument above, or directly via

$$\begin{aligned} \tilde{g}_h'(\mu) &= \nabla H((\Psi_h^\mu(y_n))^\top \cdot \frac{\partial \Psi_h^\mu(y_n)}{\partial \mu} \\ &= \nabla H((\Psi_h^\mu(y_n))^\top \left(\frac{\partial \Phi_h(\Phi_\mu(y_n)))}{\partial \Phi_\mu(y_n)} \nabla H(y_n) + \nabla H(\Phi_h(\Phi_\mu(y_n))) \right) + \\ &\quad + \mathcal{O}(h) + \mathcal{O}(\mu) \\ &= 2\|\nabla H(y_n)\|_2^2 + \mathcal{O}(h) + \mathcal{O}(\mu). \end{aligned}$$

4.3 The Structure-preserving approach

In our second approach, we replace the perturbation and projection step in the symmetric projection scheme (4.2) by a symmetric, time-reversible and symplectic one-step method $\hat{\Phi}_\mu$ of order k, which performs one step of step size μ applied to the original Hamiltonian equation of motion (4.3) with Hamiltonian $H(y)$. In contrast to the first approach, the right-hand sides of the perturbation and projection step contain linearly transformed Jacobian of the constraint, and not the Jacobian itself.

The overall scheme, depending on the parameter μ is then defined by the composition of three numerical integration schemes

$$y_{n+1} = \bar{\Psi}_h^\mu(y_n) := \hat{\Phi}_\mu(\Phi_h(\hat{\Phi}_\mu(y_n))), \tag{4.7}$$

with step size μ, h and μ, resp. Similar to the first approach, the overall numerical scheme $\Psi_\mu(y_n)$ inherits the properties of symmetry, time-reversibility and symplecticity from the three basis schemes.

To obtain preservation of the Hamiltonian in addition, μ is fixed by the constraint

$$0 = g(y_{n+1}) := H(y_{n+1}) - H(y_n). \tag{4.8}$$

For a given fixed step size h, the optimal μ_{opt} to preserve the Hamiltonian is then fixed by the nonlinear system

$$0 = F(h, y_{n+1}, \mu) := \begin{pmatrix} y_{n+1} - \bar{\Psi}_h^\mu(y_n) \\ H(y_{n+1}) - H(y_n) \end{pmatrix}. \tag{4.9}$$

Since $F(0, y_n, 0) = 0$, existence of the numerical approximation is given for sufficiently small step sizes h by the implicit function theorem, if the Jacobian

$$\frac{\partial F}{\partial(y_{n+1}, \mu)}\bigg|_{(h, y_n, 0)} = \begin{pmatrix} I & -\frac{\partial \Phi_h(y_n)}{\partial y_n} J^{-1} \nabla H(y_n) + J^{-1} \nabla H(y_{n+1}) \\ \nabla H(y_{n+1})^\top & 0 \end{pmatrix}$$

is regular for $h \to 0$, which is equivalent to a non-vanishing Schur complement

$$\tilde{g}_h'(0) = \frac{\partial H(\Phi_h(y_n))}{\partial y_n} J^{-1} \nabla H(y_n) \tag{4.10}$$

with $\tilde{g}_h(\mu) := g(\Psi_h^\mu(y_n))$. The Schur complement depends only on the intermediate scheme Φ_h. For a scheme of order 2, the numerical approximate $\Phi_h(y_n)$ can be written by Taylor series expansion as

$$\Phi_h(y_n) = y_n + h J^{-1} \nabla H(y_n) + \frac{1}{2} h^2 J^{-1} \nabla^2 H(y_n) J^{-1} \nabla H(y_n) + h^3 \delta(y_n) + \mathcal{O}(h^4).$$

Inserting this expansion into (4.10), and expanding H around y_n, one gets

$$\tilde{g}_h'(0) = h^3 \frac{\partial S(y_n)}{\partial y_n} J^{-1} \nabla H(y_n) + \mathcal{O}(h^4)$$

with

$$S(y_n) := \nabla H^\top(y_n) \delta(y_n) + \frac{1}{6} \nabla^3 H(y_n)(J^{-1} \nabla H, J^{-1} \nabla H, J^{-1} \nabla H).$$

If we now replace the constraint $g(y) = H(y) - H(y_n)$ by

$$\bar{g}(y) = \frac{H(y) - H(y_n)}{h^3}, \tag{4.11}$$

the corresponding Schur complement reads

$$\bar{g}_0'(0) = \frac{\partial S(y_n)}{\partial y_n} J^{-1} \nabla H(y_n).$$

If this Schur complement is not zero, we can repeat the argument of the first approach: for sufficiently small step size h, the existence of the numerical approximation $\bar{\Psi}_h^{\mu\,\text{opt}}(y_n)$ with local order $\mathcal{O}(h^{r+1})$ is guaranteed, and $\mu = \mathcal{O}(h^{r+1})$, independent of the order k of Ψ_μ.

For the leap-frog (Störmer-Verlet) method, for example, one gets for a separable Hamiltonian

$$\delta(y_n) = \begin{pmatrix} \frac{1}{8} H_{ppp}(y_n)(H_q(y_n), H_q(y_n)) \\ -\frac{1}{4}(H_{qqq}(y_n)(H_p(y_n), H_p(y_n)) - H_{qq}(y_n) H_{pp}(y_n) H_q(y_n)) \end{pmatrix},$$

resulting in the Schur complement

$$S(y_n) = \frac{7}{24} H_{ppp}(H_q(y_n), H_q(y_n), H_q(y_n)) - \frac{1}{12} H_{qqq}(H_p(y_n), H_p(y_n), H_p(y_n))$$
$$+ \frac{1}{4} H_{qq}(y_n) H_{pp}(y_n)(H_q(y_n), H_p(y_n)).$$

Unfortunately despite all positive results, obtained, it is clear that the parameter μ depends on not only step-size h, but also on the initial values y_0, thus we cannot rely on the results obtained in previous sections, since we cannot guarantee conservation of the geometric properties in general with $\mu(h)$.
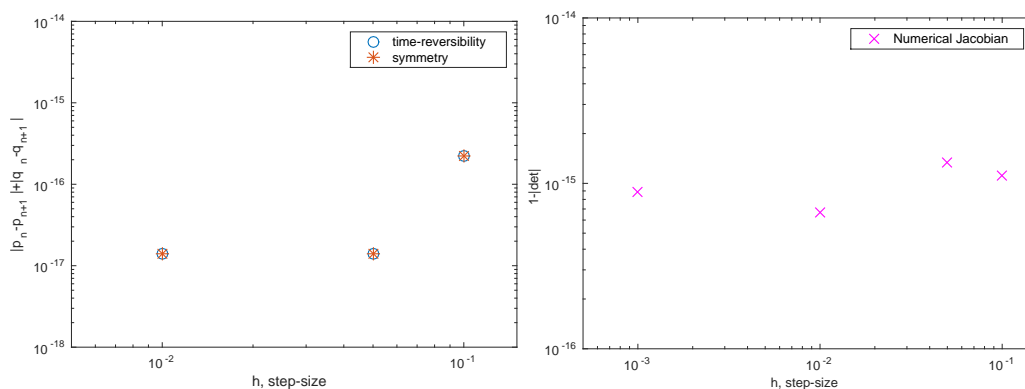

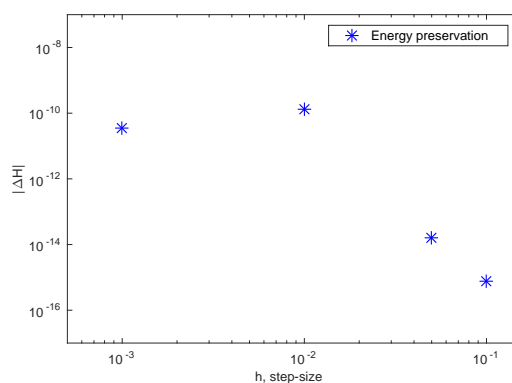Figure 4.2: Symmetry(time-reversibility) (left); Symplecticity (right).

Figure 4.3: Energy drift.

Figures 4.2–4.3 show numerically the conservation of geometrical properties by the method (4.7) with $\mu = \mu(y_0, h)$ applied to the simple Harmonic oscillator (SHO) given by the Hamiltonian function $H = \frac{1}{2}p^2 + \frac{1}{2}q^2$. We can clearly observe very promising results. Unfortunately we must realize that it would require a computation of $\mu(y_n, h)$ satisfying (4.8) on every time-step of the algorithm (4.7) and even in the case of SHO it results in extensive computational time.

4.4 Structure-preserving approach $\left(\mu = \mu(h)\right)$

Although the results obtained in the previous sections do not satisfy our needs, we still tend to believe that there are some possibilities to increase the convergence order using ideas, inspired by the projection method approach, namely by fixing $\mu = \mu(h)$ such that it would reduce the leading error term.

Our first attempt consists of the following composition of numerical methods

$$y_{n+1} = \Psi_h^\mu(y_n) := \hat{\Phi}_\mu(\Phi_h(\hat{\Phi}_\mu)), \tag{4.12}$$

where $\hat{\Phi}_\mu$ and Φ_h are Störmer-Verlet methods with the step-sizes h and μ respectively and find such $\mu = \mu(h)$, in order to apply the theory, developed in the previous sections, and possibly to increase the order of overall scheme.

To do so we rewrite our composition scheme (4.12) in the form of a symplectic mapping

$$\underbrace{e^{\frac{\mu}{2}\hat{V}}e^{\mu\hat{T}}e^{\frac{\mu}{2}\hat{V}}}_{\hat{H}_1}\underbrace{e^{\frac{h}{2}\hat{V}}e^{h\hat{T}}e^{\frac{h}{2}\hat{V}}}_{\hat{H}_2}\underbrace{e^{\frac{\mu}{2}\hat{V}}e^{\mu\hat{T}}e^{\frac{\mu}{2}\hat{V}}}_{\hat{H}_3}$$

apply the BCH formula (see Appendix A) and obtain the shadow Hamiltonian $\tilde{H} = H_1 + H_2 + H_3$ or

$$\tilde{H} = H + \frac{(2\mu^3 + h^3)}{24(2\mu + h)}\left(2[T[T, V]] + [V[T, V]]\right).$$

In order to increase the order of the following scheme the factor $\frac{(2\mu^3 + h^3)}{24(2\mu + h)}$ must be equal to 0, which is satisfied by the choice $\mu = -\frac{h}{\sqrt[3]{2}}$

The second attempt is to consider "alike" projection method of the following form

$$y_{n+1} = \Psi_h^\mu(y_n) := \hat{\Phi}_{\mu_1}(\hat{\Phi}_{\mu_2}(\Phi_h(\hat{\Phi}_{\mu_2}(\hat{\Phi}_{\mu_1})))), \tag{4.13}$$

where $\hat{\Phi}_\mu$ and Φ_h are Störmer-Verlet methods with the step-sizes h, μ_1 and μ_2 respectively. We chose two different step-sizes $\mu_1 = \mu_1(h)$ and $\mu_2 = \mu_2(h)$ in order to have an additional degree of freedom. For this purpose we apply the BCH formula on the symplectic map from the numerical composition scheme (4.13)

$$\underbrace{e^{\frac{\mu_1}{2}\hat{V}}e^{\mu_1\hat{T}}e^{\frac{\mu_1}{2}\hat{V}}}_{\hat{H}_1}\underbrace{e^{\frac{\mu_2}{2}\hat{V}}e^{\mu_2\hat{T}}e^{\frac{\mu_2}{2}\hat{V}}}_{\hat{H}_2}\underbrace{e^{\frac{h}{2}\hat{V}}e^{h\hat{T}}e^{\frac{h}{2}\hat{V}}}_{\hat{H}_3}\underbrace{e^{\frac{\mu_2}{2}\hat{V}}e^{\mu_2\hat{T}}e^{\frac{\mu_2}{2}\hat{V}}}_{\hat{H}_4}\underbrace{e^{\frac{\mu_1}{2}\hat{V}}e^{\mu_1\hat{T}}e^{\frac{\mu_1}{2}\hat{V}}}_{\hat{H}_5}$$

and obtain (see Appendix B) $\tilde{H} = H_1 + H_2 + H_3 + H_4 + H_5$ or

$$\tilde{H} = H + \frac{(2\mu_1^3 + 2\mu_2^3 + h^3)}{24(2\mu_1 + 2\mu_2 + h)}\left(2[T[T, V]] + [V[T, V]]\right).$$

Thus we can find μ_1 and μ_2 such that $\frac{(2\mu_1^3 + 2\mu_2^3 + h^3)}{24(2\mu_1 + 2\mu_2 + h)} = 0$.

Here we have to emphasize that unfortunately regardless of our investigations the result is unsatisfactory. Since the presented schemes are different derivations of the well-known idea of composition approach [28] and hence yields no original ideas, but give a different

representation of already existing numerical time integration schemes.

4.5 Linear projection methods

After obtaining rather disappointing results we have decided to reconsider our understanding of the idea of symmetric projection methods. Also taking into account that using the composition approach can dramatically increase the computation time, thus making the effort of getting the higher convergence order virtually worthless. We have come up with the idea of substituting the perturbation and the projection steps of the scheme (4.2) by some linear transformation.

It leads us to consider the following scheme

$$y_{n+1} = L^{-1}\Phi_h L(y_n), \tag{4.14}$$

where L is a linear operator.

As we can see the symmetry of the method is satisfied by the construction of the method. The time-reversibility can be shown by using the definition of time-reversibility

$$\rho \cdot L^{-1} \cdot \Phi_h \cdot L \cdot \rho \cdot L^{-1} \cdot \Phi_h \cdot L \overset{!}{=} \mathrm{id}\,.$$

Using the time-reversibility of the underlying method Φ_h and the fact that ρ is idempotent, we can replace Φ_h with $\rho \cdot \Phi_{-h}^{-1} \cdot \rho$ and get

$$\mathrm{id} \overset{!}{=} \rho \cdot L^{-1} \cdot \Phi_h \cdot (L \cdot \rho \cdot L^{-1}\rho) \cdot \Phi_{-h}^{-1} \cdot \rho \cdot L.$$

If we assume $L \cdot \rho \cdot L^{-1}\rho = \mathrm{id}$ the time-reversibility condition is satisfied, it holds only when L has a block-diagonal structure.

The volume-preservation is satisfied; it can be proved simply by applying the definition of the volume-preservation

$$1 = \det\left(\frac{\partial L^{-1}\Phi_h L(y_0)}{\partial y_0}\right) = |L^{-1}|\, \underbrace{\Phi_h'}_{=1}\, |L| = 1.$$

The open question is whether it possible to achieve the higher order of convergence for the method (4.14). To check it we write down the scheme (4.14) in the form of a symplectic mapping

$$e^{L^{-1}\frac{h}{2}V}e^{hT}e^{\frac{h}{2}VL}$$

and by applying the BCH formula (see Appendix C) we obtain

$$\begin{aligned}
\tilde{H} = {} & H + L^{-1}\frac{h}{2}V + hT + \frac{h}{2}VL - \frac{h^2}{4}[T, L^{-1}V] + \frac{h^2}{4}[T, VL] \\
& - \frac{h^3}{48}[L^{-1}V, [T, L^{-1}V]] + \frac{h^3}{24}[T, [T, L^{-1}V]] + \frac{h^3}{16}[VL, [T, L^{-1}V]] \\
& + \frac{h^3}{48}[L^{-1}V, [T, VL]] - \frac{h^3}{48}[VL, [T, VL]] + \frac{h^3}{24}[T, [T, VL]].
\end{aligned}$$

It is easy to see that L has to be equal to L^{-1} to sustain the second order of the underlying method, but that would mean, together with above conditions, L is the identity matrix. This indicates, our scheme (4.14) will not yield any advantages, in the case of $\mu = \mu(h)$. Only if we consider $\mu = \mu(y_0, h)$ for the numerical method (4.14) and $L = (I + A\mu)$, where A is block diagonal matrix, we can obtain rather interesting results. We illustrate the advantages of this approach using an example.

Let us consider a simple harmonic oscillator (SHO) given by

$$H = \frac{p^2}{2} + \frac{q^2}{2}$$

with the equation of motions

$$\dot{p} = -\frac{\partial H}{\partial q} = -q, \qquad\qquad \dot{q} = \frac{\partial H}{\partial p} = p. \qquad (4.15)$$

The full scheme will look like

$$\tilde{p}_n = (1 + \mu)p_n$$
$$\tilde{q}_n = q_n$$
$$\tilde{p}_{n+\frac{1}{2}} = \tilde{p}_n - \frac{h}{2}\tilde{q}_n$$
$$\tilde{q}_{n+1} = \tilde{q}_n + h\tilde{p}_{n+\frac{1}{2}}$$
$$\tilde{p}_{n+1} = \tilde{p}_{n+\frac{1}{2}} - \frac{h}{2}\tilde{q}_{n+1}$$
$$p_{n+1} = \frac{\tilde{p}_{n+1}}{1 + \mu} = \frac{(1 + \mu)p_n - \frac{hq_n}{2} - \frac{1}{2}h\left(q_n + h\left((1 + \mu)p_n - \frac{hq_n}{2}\right)\right)}{1 + \mu}$$
$$q_{n+1} = \tilde{q}_{n+1} = q_n + h\left((1 + \mu)p_n - \frac{hq_n}{2}\right).$$

The energy conservation condition reads

$$\frac{p_{n+1}^2}{2} + \frac{q_{n+1}^2}{2} - \frac{p_n^2}{2} - \frac{q_n^2}{2} = 0$$

and solving this equation with respect to μ we obtain

$$\mu_1 = \frac{-2p_n + hq_n}{2p_n} \qquad\qquad \mu_2 = \frac{-2hp_n - 4q_n + h^2q_n}{2hp_n}$$
$$\mu_3 = \frac{1}{2}\left(-2 - \sqrt{4 - h^2}\right), \qquad\qquad \mu_4 = \frac{1}{2}\left(-2 + \sqrt{4 - h^2}\right).$$

Since the time-reversibility and the symmetry of (4.14) are satisfied by the construction of the method and the symplecticity can be checked simply by computing the determinant of the Jacobian of the scheme (4.14)

$$\det\begin{pmatrix} \partial_{p_n}p_{n+1} & \partial_{q_n}p_{n+1} \\ \partial_{p_n}q_{n+1} & \partial_{q_n}q_{n+1} \end{pmatrix} = -1,$$

where the result is obtained by using the symbolic computation of *Mathematica*, and it satisfies the area-preservation condition.
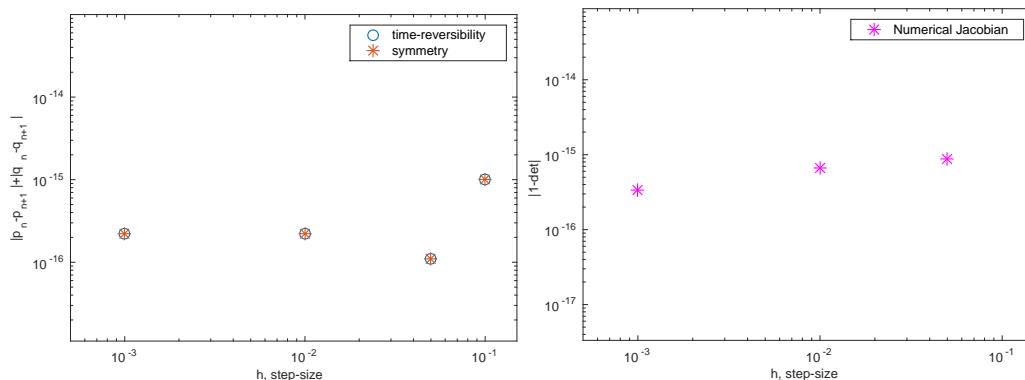

Figure 4.4: Symmetry(time-reversibility) (left); Symplecticity (right).
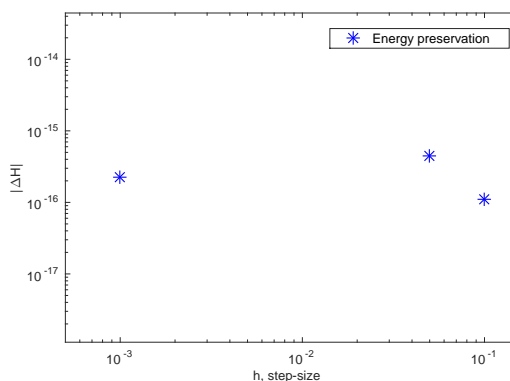

Figure 4.5: Energy drift.

Figures 4.4–4.5 show conservation of the geometric structure numerically. As we can observe all desired properties are preserved exactly (up to machine precision).

4.6 Conclusion

Despite very promising results of the previous section regarding the linear projection (4.14) we must admit that the outcome of this whole chapter is only semi-positive. The presented ideas work only for rather simple problems, where with an increase of degrees of freedom and use of more general (non-linear, non-separable) Hamiltonian function, the calculation of $\mu(y_0, h)$ on each time step of the algorithm would result in very high computational costs, minimizing the practicability of this schemes. For our main application, presented in Chapter 3, it plays a crucial role, since the constrains (4.8) are extremely complicated systems and they would require additional numerical treatment and hence lead to vast computational costs, which are not affordable. Summarizing everything written above we conclude that we have to choose the other direction of our research, namely developing higher order geometric numerical integrator and at the same time reducing the computational effort. In the next chapter we present our research in this direction.

5

Chapter 5

NESTED FORCE-GRADIENT METHODS

Previously we mentioned in Chapter 3 that for the HMC algorithm, used in QCD, one is interested in efficient numerical time integration schemes, which are optimal in terms of computational costs per trajectory for a given acceptance rate level. High order numerical methods allow the usage of larger step sizes, but ask for larger computational effort per step; low order schemes do not require vast computational costs per step, but need more steps per trajectory. So there is a need to balance these opposing effects.

Omelyan's numerical time integration schemes [42] of a force-gradient type have been shown to be an efficient choice, since it is easy to obtain a higher order for a scheme and it demands just a small additional computational effort. These methods use high order information of the force-gradient terms in order to increase the convergence order of the given numerical scheme and/or decrease the size of the leading error coefficient. Another idea to achieve the better efficiency of numerical time integrators is given by the multirate or nested approach. These schemes do not increase the order of the scheme, but reduce the computational costs per path by exploiting the multirate behavior given the different parts of the right-hand side: slow forces, which are usually expensive to evaluate, but need only to be sampled at low frequency; and fast forces, which are usually cheap to evaluate, but need to be sampled at a high frequency. The natural way to inherit the advantages from force-gradient type schemes and multirate approach would be to combine these two ideas [52].

In this chapter we first briefly present the idea of splitting decomposition schemes, which are actually a basis to both of the above mentioned approaches. Then we use this class of methods to introduce the force-gradient decomposition integrators and consider their structure preserving properties as well as order conditions for this type of numerical time integrators. We continue by presenting the idea of multirate time integration and use the advantages of these two classes of methods to combine both approaches in order to obtain a more efficient numerical scheme.

5.1 Splitting decomposition schemes

The main idea of this type of numerical time integrators is to decompose the vector field into integrable pieces and treat them separately [28]. Here we consider an arbitrary system $\dot{y} = f(t, y) \in \mathbb{R}^n$, and suppose that the vector field is *split* as

$$\dot{y} = f^{[1]}(t, y) + f^{[2]}(t, y),$$

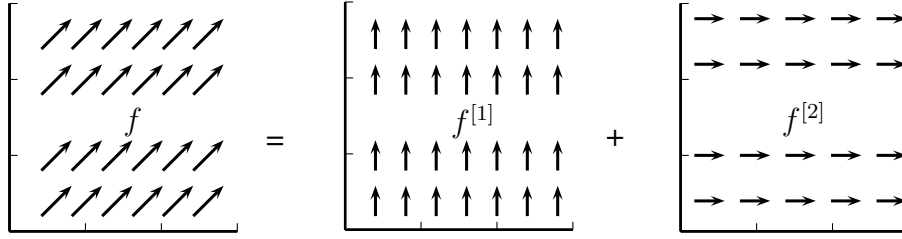and it can be visualized as in Figure 5.1.

Figure 5.1: Splitting of a vector field.

If $\varphi_t^{[1]}$ and $\varphi_t^{[2]}$ the exact flows, respectively, of the systems $\dot{y} = f^{[1]}(y)$ and $\dot{y} = f^{[2]}(t, y)$ can be calculated explicitly, then from a given initial value y_0 one can solve the first system in order to obtain a value $y_{1/2}$, and from this value integrate the second system to obtain y_1. In this way we have introduced the numerical methods

$$\Phi_h^* = \varphi_h^{[2]} \circ \varphi_h^{[1]}, \qquad\qquad \Phi_h = \varphi_h^{[1]} \circ \varphi_h^{[2]},$$

where the scheme Φ_h^* is adjoint of the scheme Φ_h, both of them yield first order approximation and usually called *Lie-Trotter splitting* [28]. The composition of the Lie-Trotter methods and its adjoint with halved step sized $\Phi_h^S = \Phi_{\frac{h}{2}} \circ \Phi_{\frac{h}{2}}^*$ results in a symmetric numerical scheme of the second order, which can be written as

$$\Phi_h^S = \varphi_{\frac{h}{2}}^{[1]} \circ \varphi_h^{[2]} \circ \varphi_{\frac{h}{2}}^{[1]}$$

and it is known as the *Strang splitting* [28].

The general splitting procedure can be formed by using arbitrary coefficients a_1, b_1, a_2, ..., a_n, b_n (where, eventually, a_1 or b_n, or both, are zero)

$$\Phi_h = \varphi_{b_n h}^{[2]} \circ \varphi_{a_n h}^{[1]} \circ \varphi_{b_{n-1} h}^{[2]} \circ \cdots \circ \varphi_{a_2 h}^{[1]} \circ \varphi_{b_1 h}^{[2]} \circ \varphi_{a_1 h}^{[1]} \tag{5.1}$$

and try to increase the order of the scheme by suitably determining the free coefficients. There have been several developments of the special cases of this class of methods (5.1) until the systematic study was proposed in [56]-[61].

Here we follow the approach presented in [61] to obtain the order conditions for splitting methods (5.1) and also to determine structure preserving properties such as time-reversibility and symplecticity, of these class of methods.

Since our main interest lies within a numerical time integration of canonical transformation $(q(0), p(0)) \rightarrow (q(\tau), p(\tau))$ on the molecular dynamic step in the HMC algorithm it restricts us to a certain type of problems for Hamiltonian systems of the form

$$H = T(p) + V(q). \tag{5.2}$$

Let us write down the Hamilton's equation of motion in the following form

$$\dot{\rho} = \{\rho, H(\rho)\}, \tag{5.3}$$

where $\rho = (q, p)$ and $\{\ ,\ \}$ stands for Poisson bracket. The above equation can be also

written as

$$\dot{\rho} = \hat{H}\rho, \tag{5.4}$$

where $\hat{H}$ is a differential operator $\hat{H}\rho := \{\rho, H(\rho)\}$ and $\{ \, , \, \}$ represents

$$\{f, g\} = \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p}.$$

Then the formal solution, or exact time evolution $\rho(t)$ from $t = 0$ to $t = \tau$ is given by

$$\rho(T) = \mathrm{e}^{\tau\hat{H}}\,\rho(0). \tag{5.5}$$

For the Hamiltonian of the form (5.2), the differential operator $\hat{H} = \hat{T} + \hat{V}$, the solution (5.5) becomes

$$\rho(T) = \mathrm{e}^{\tau(\hat{T}+\hat{V})}\,\rho(0). \tag{5.6}$$

Here we use Lemma 1 to write the method (5.1) as a product of exponentials and formally apply the Baker-Campbell-Hausdorff (BCH) formula (2.10) in order to obtain one exponential of a series in powers of step-size h. Finally, we compare this series with (5.6), which corresponds to the exact solution, which results, cf. [28]

$$\mathrm{e}^{(\hat{T}+\hat{V})h+\mathcal{O}(h^{r+1})} = \prod_{k=1}^{K} \mathrm{e}^{\hat{T}a_k h}\,\mathrm{e}^{\hat{V}b_k h}, \tag{5.7}$$

where the evolution operators $\mathrm{e}^{\hat{T}a_k h}$ and $\mathrm{e}^{\hat{V}b_k h}$ displace q and p forward in time with

$$q \to q + a_k h\hat{T}\rho \quad \text{and} \quad p \to p + b_k h\hat{V}\rho. \tag{5.8}$$

The coefficients a_k and b_k in evolution operators (5.7) have to be chosen in such way to obtain the highest possible order $r \geq 1$ for a given integer $K \geq 1$.

Before we comment on the choice of coefficients for the order conditions for our splitting methods (5.7) let us first have a look at the structure-preserving properties of this particular class of numerical schemes. Since it results a simplification, better understanding and an improvement of the original decomposition scheme (5.7).

The most important feature of this decomposition approach (5.7) is its *exact conservation of the symplectic map* of the numerical flow, because the separate shifts (5.15) of coordinate variables q and momenta variables p do not change the phase volume and decomposition (5.7) is just a product of the symplectic mappings [61].

Time-reversibility can be ensured by imposing two conditions, namely $a_1 = 0$, $a_{k+1} = a_{K-k+1}$ and $b_k = b_{K-k+1}$ as well as $a_k = a_{K-k+1}$ and $b_k = b_{K-k}$ with $b_K = 0$. Then the single-exponential propagator enters into the decomposition (5.7), providing automatically the required reversibility [43]. This symmetry also leads to elimination of all the even-order terms in the error function of the splitting integrators (5.7) setting $r = 2, 4, 6, 8, \ldots$, i.e.

$$\mathcal{O}\left(h^{r+1}\right) = \mathcal{O}_1 h + \mathcal{O}_3 h^3 + \mathcal{O}_5 h^5 + \cdots + \mathcal{O}_{r+1} h^{r+1} + \ldots. \tag{5.9}$$

For example we can represent the famous the Störmer-Verlet scheme (leap-frog method)

by the decomposition (5.7) by putting $K = 2$ with $a_1 = 0$, $b_1 = b_2 = 1/2$, and $a_2 = 1$

$$\mathrm{e}^{h\frac{\hat{V}}{2}} \, \mathrm{e}^{h\hat{T}} \, \mathrm{e}^{h\frac{\hat{V}}{2}} \,. \tag{5.10}$$

The fourth-order $(r = 4)$ numerical scheme from [17] is obtained via setting $K = 4$ with $a_1 = 0$, $a_2 = a_4 = \theta$, $a_3 = (1 - 2\theta)$, $b_1 = b_4 = \theta/2$, and $b = 2 = b_3 = (1 - \theta)/2$, where $\theta = 1/(2 - \sqrt[3]{2})$ and given by

$$\mathrm{e}^{\frac{\theta}{2}h\hat{V}} \, \mathrm{e}^{\theta h\hat{T}} \, \mathrm{e}^{\frac{1-\theta}{2}h\hat{V}} \, \mathrm{e}^{(1-2\theta)h\hat{T}} \, \mathrm{e}^{\frac{1-\theta}{2}h\hat{V}} \, \mathrm{e}^{\theta h\hat{T}} \, \mathrm{e}^{\frac{\theta}{2}h\hat{V}} \,. \tag{5.11}$$

Schemes of the sixth order $(r = 6)$ are derivable starting from $K = 8$ with numerical representation of time coefficients [42].

In Chapter 2 we introduced the concept of the modified Hamiltonian for symplectic integrators as the way to determine the convergence order of numerical schemes. Here we can apply BCH formula (2.10) the same way as in (5.9) and show that for the method (5.11) the shadow Hamiltonian has the following form

$$\tilde{H}_{FR} = H + \Bigg((32 - 20\sqrt[3]{4}) \Big[V, \Big[V, [V, [V, T]] \Big] \Big] + \ldots$$
$$+ (8 + 20\sqrt[3]{4}) \Big[T, \Big[T, [T, [V, T]] \Big] \Big] \Bigg) \frac{h^4}{34560} + \mathcal{O}(h^6).$$

The decomposition (5.7) has certain disadvantages. First and foremost with an increasing order of the integration from 6 to 8 and higher, the number $2K$ of unknowns a_k and b_k increases too fast. This makes it impossible to explicitly represent algorithms (5.7) for the case $r > 6$, since one cannot solve the same number of the resulting nonlinear equations with respect to a_p and b_p even using the capabilities of modern supercomputers [43]. Also it is impossible to derive a decomposition scheme (5.7) for $r > 2$ at any K with only positive time coefficients [56]. For example, the numerical scheme (5.11) has negative coefficients, namely, a_3, b_2 and b_3, and since schemes with negative time coefficients have a restricted region of application and are not acceptable for simulating non-equilibrium statistical mechanics, quantum statistics, stochastic dynamics, and other important processes. Moreover, for schemes expressed in terms of force evaluation only, the main term $\mathcal{O}(h^{r+1})$ of truncation uncertainties appears to be, as a rule, too big, resulting in a decrease in the efficiency of the computations [42].

5.2 Force-gradient decomposition method

The drawbacks, mentioned above, make it quite desirable to introduce a more general approach that is free of all disadvantages. At the same time, the proposed approach must be explicit, in order to lead to analytical propagations. And, in addition, the already known decomposition algorithms should be able to be derived from it as particular cases.

A so called force-gradient decomposition approach [42] follows from the standard decom-

position (5.7) via adding the force-gradient term $\mathcal{C}$, which results to the splitting

$$\mathrm{e}^{(\hat{T}+\hat{V})h+\mathcal{O}(h^{r+1})} = \prod_{k=1}^{K} \mathrm{e}^{\hat{T}a_k h}\, \mathrm{e}^{\hat{V}b_k h+\mathcal{C}c_k h^3}, \tag{5.12}$$

where $\mathcal{C} = [V, [T, V]]$ and $[\ ,\ ]$ denotes the commutator of two operators. Here the force-gradient term $\mathcal{C}$ arises from the third order leading error term $\mathcal{O}(h^3)$ from (5.9), which can be obtained via Taylor expansion and yields

$$\mathcal{O}(h^3) = \left(\frac{1}{12}[T, [T, V]] + \frac{1}{24}[V, [T, V]] \right) h^3 + \mathcal{O}(h^5). \tag{5.13}$$

Taking into account the explicit expressions for operators T and V it can be shown that one of the two third-order operators in (5.13), namely $[V, [T, V]]$, is relatively simple and, more importantly, it allows to be handled explicitly, contrary to the operator $[T, [T, V]]$ [43]. Therefore by choosing in the certain c_k one can reduce the leading error term.

For example, the force-gradient version of the leap-frog method (5.10)

$$\mathrm{e}^{\frac{h}{2}\hat{V}-\frac{h^3}{48}\hat{\mathcal{C}}}\, \mathrm{e}^{h\hat{T}}\, \mathrm{e}^{h\frac{\hat{V}}{2}+\frac{h^3}{48}\hat{\mathcal{C}}}, \tag{5.14}$$

has the shadow Hamiltonian

$$\tilde{H}_{LP_{fg}} = H - \frac{h^2}{12}[T, [T, V]] + \mathcal{O}(h^4),$$

which is smaller and therefore better preserves energy for the system in the comparison to the standard leap frog formulation, which shadow Hamiltonian yields

$$\tilde{H}_{LP} = H - \frac{h^2}{24}\Big(2[V, [T, V]] + [T, [T, V]]\Big) + \mathcal{O}(h^4).$$

Since the force-gradient term $\mathcal{C}$ not only affects the third order leading error term $\mathcal{O}(h^3)$ but also higher order error terms it is possible to obtain an improvement for higher order numerical schemes.

Let us first mention the conservation properties of the force-gradient approach (5.12). In the similar way as for methods presented in (5.7) the coefficients a_k, b_k and c_k in (5.12) have to be chosen in such way to obtain the highest possible order $r \geq 1$ for a given integer $K \geq 1$. The expression (5.12) represents the general form of the decomposition, while for $c_k \equiv 0$ the decomposition reduces to the standard non-gradient factorization (5.7).

The evolution operators $\mathrm{e}^{\hat{T}a_k h}$ and $\mathrm{e}^{\hat{V}b_k h+\mathcal{C}c_k h^3}$ in (5.12) displace p and q forward in time with

$$p \to p + b_k h\hat{V}(\rho) + c_k \mathcal{C}h^3 \qquad \text{and } q \to q + a_k h\hat{T}(\rho). \tag{5.15}$$

Symplecticity of the decomposition integration (5.12) is conserved (similar to the decomposition (5.7)), since the separate shifts (5.15) of positions and velocities do not change the phase volume. Here, it is important to mention that in the non-Abelian case, when, for example, the gauge fields U are elements of some Lie group, and the momenta P, elements of the corresponding Lie algebra, substituting the separable shifts (5.15) with

the following

$$
\begin{aligned}
\mathrm{e}^{\hat{T}h} : \quad U &\quad \rightarrow \quad U' = \exp(iPh)U \\
\mathrm{e}^{\hat{V}h} : \quad P &\quad \rightarrow \quad P' = P - ihF_V(U),
\end{aligned}
$$

results the symplectic mapping and therefore the conservation of the phase volume. It makes the formulation (5.12) universal for the Abelian case as well as for the case of the integration on the manifolds.

Time-reversibility with the addition of the coefficient c_k can be ensured by imposing two conditions, namely $a_1 = 0$, $a_{k+1} = a_{K-k+1}$, $b_k = b_{K-k+1}$, $c_k = c_{K-k+1}$, as well as $a_k = a_{K-k+1}$, $b_k = b_{K-k}$, $c_k = c_{K-k}$ with $b_K = 0$ and $c_K = 0$ [42].

As well as for the standard decomposition (5.7) the above symmetry will result to an automatic disappearing of all even-order terms in the error function (5.9) for the force-gradient version (5.12). The cancellation of the remaining error terms in (5.9) is provided by fulfilling a set of basic conditions for a_k, b_k, and c_k.

For example the assumption

$$
\sum_{k=1}^{K} a_k = \sum_{k=1}^{K} b_k = 1
$$

results in the elimination of the first order leading error term $\mathcal{O}_1(h)$ in (5.9) and the error function (5.9) can be then rewritten

$$
\mathcal{O}\left(h^{r+1}\right) = \mathcal{O}_3 h^3 + \mathcal{O}_5 h^5 + \cdots + \mathcal{O}_{r+1} h^{r+1} + \dots. \tag{5.16}
$$

In order to annihilate high order error terms in (5.16) first we need to explicitly derive these terms by expanding both sides of the decomposition (5.12) in Taylor series, which yields

$$
\begin{aligned}
\mathcal{O}_1 &= (\nu - 1)T + (\sigma - 1)V, \quad \mathcal{O}_3 = \alpha[T,[T,V]] + \beta[V,[T,V]] \\
\mathcal{O}_5 &= \gamma_1\,[T,[T,[T,[T,V]]]] + \gamma_2\,[T,[T,[V,[T,V]]]] + \gamma_3\,[V,[T,[T,[T,V]]]] \\
&\quad + \gamma_4\,[V,[B,[T,[T,V]]]],
\end{aligned} \tag{5.17}
$$

where coefficients ν, σ, α, β and γ_{1-4} are chosen in such a way to eliminate error terms (5.17) further we show the optimal algorithm proposed by Omelyan et al. [42] to define these coefficients. Here we also keep the error term $\mathcal{O}_1$ in mind for the generalization of the order conditions.

Omelyan et al. [42] proposed the following algorithm to derive the multipliers ν, σ, α, β and γ_{1-4}, arising in (5.17), which depend on the coefficients a_k, b_k, and c_k, where $k = 1, 2, \dots, K$. Since we are dealing with self-adjoint schemes here, the total number of single-exponential operators (stages) in (5.12) actually equals to $S = 2K - 1$, and taking into account the time-reversibility condition $a_1 = 0$ or $b_K = c_K = 0$, it only accepts odd values. Then the simplest way to obtain explicit expressions according to [43] consists of repeatedly choosing a central single-exponential operator and apply $K - 1$ times the following two types of symmetric transformations

$$\mathrm{e}^{\mathcal{W}^{(n+1)}} = \mathrm{e}^{Ta^{(n)}h}\, \mathrm{e}^{\mathcal{W}^{(n)}}\, \mathrm{e}^{Ta^{(n)}h}$$

$$\mathrm{e}^{\mathcal{W}^{(n+1)}} = \mathrm{e}^{Vb^{(n)}h+\mathcal{C}c^{(n)}h^3}\, \mathrm{e}^{\mathcal{W}^{(n)}}\, \mathrm{e}^{Vb^{(n)}h+\mathcal{C}c^{(n)}h^3}, \tag{5.18}$$

where

$$\mathcal{W} = (\nu T + \sigma V)h + \mathcal{O}_3 h^3 + \mathcal{O}_5 h^5 + \cdots + \dots \tag{5.19}$$

The result leads us to the factorization (5.12), where the time coefficients a_k, b_k, and c_k will be related to the quantities $a^{(n)}$, $b^{(n)}$, and $c^{(n)}$, respectively.

For the first kind of self-adjoint constraints (when $a_1 = 0$) with even P or the second one (when $b_P = c_P = 0$) with odd K, the central operator of the symmetric decompositions will be correspondingly $e^{\hat{T}a_{(K-2)/2+1}h}$ or $e^{\hat{T}a_{(K-1)/2+1}h}$. So that here we must put $\sigma^{(0)} = 0$, $\alpha^{(0)} = \beta^{(0)} = \gamma_{1-4}^0 = 0$ as well as either $\nu^{(0)} = a_{(K-2)/2+1}$ or $\nu^{(0)} = a_{(K-1)/2+1}$ on the very beginning ($n = 0$) of the recursive procedure (5.18). Moreover, we should start from the second line of (5.18) at $b^{(0)} = b_{(K-2)/2}$ and $c^{(0)} = c_{(K-2)/2}$ or $b^{(0)} = b_{(K-1)/2}$ and $c^{(0)} = c_{(K-1)/2}$ with further consecutive decreasing the index k by unity with increasing the number $n = 1, 2, \dots, K - 1$ at $a^{(n)} \equiv a_k$, $b^{(n)} \equiv b_k$, and $c^{(n)} \equiv c_k$ on both lines of transformation (5.18) [42].

For the first kind of constraints with odd K or the second kind with even K, the central operator is $e^{\hat{V}b_{(K-1)/2+1}h+\mathcal{C}c_{(K-1)/2+1}h^3}$ or $e^{\hat{V}b_{(K-2)/2+1}h+\mathcal{C}c_{(K-2)/2+1}h^3}$ that corresponds to $\sigma^{(0)} = b_{(K-1)/2+1}$ and $\beta^{(0)} = c_{(K-1)/2+1}$ or $\sigma^{(0)} = b_{(K-2)/2+1}$ and $\beta^{(0)} = c_{(K-2)/2+1}$, respectively, with $\nu^{(0)} = 0$ and $\alpha^{(0)} = \gamma_{1-4}^{(0)} = 0$. In this case, the procedure should be started from the first line of (5.18) at $a^{(0)} = a_{(K-1)/2+1}$ or $a^{(0)} = a_{(K-2)/2+1}$ with decreasing k at increasing n for $b^{(n)} \equiv b_k$, $c^{(n)} \equiv c_k$, and $a^{(n)} \equiv a_k$ [42].

The recursive relations between the multipliers ν, σ, α, β, and γ_{1-4} corresponding to the first line of (5.18) are:

$$\nu^{(n+1)} = \nu^{(n)} + 2a^{(n)}, \qquad \sigma^{(n+1)} = \sigma^{(n)},$$

$$\alpha^{(n+1)} = \alpha^{(n)} - \frac{a^{(n)}\sigma^{(n)}\left(a^{(n)} + \nu^{(n)}\right)}{6},$$

$$\beta^{(n+1)} = \beta^{(n)} - \frac{a^{(n)}\sigma^{(n)2}}{6},$$

$$\gamma_1^{(n+1)} = \gamma_1^{(n)} + \frac{a^{(n)}\left(a^{(n)} + \nu^{(n)}\right)\left(\left(7a^{(n)2} + 7a^{(n)}\nu^{(n)} + \nu^{(n)2}\right)\sigma^{(n)} - 60\alpha^{(n)}\right)}{360},$$

$$\gamma_2^{(n+1)} = \gamma_2^{(n)} + \frac{a^{(n)}\left(30\alpha^{(n)}\sigma^{(n)} - 30a^{(n)}\beta^{(n)} - 30\beta^{(n)}\nu^{(n)}\right)}{180}, \tag{5.20}$$

$$+ \frac{a^{(n)}\left(3a^{(n)2}\sigma^{(n)2} + 2a^{(n)}\nu^{(n)}\sigma^{(n)2} + \nu^{(n)2}\sigma^{(n)2}\right)}{180},$$

$$\gamma_3^{(n+1)} = \gamma_3^{(n)} + \frac{a^{(n)}\sigma^{(n)}\left(\left(8a^{(n)2} + 12a^{(n)}\nu^{(n)} + \nu^{(n)2}\right)\sigma^{(n)} - 120\alpha^{(n)}\right)}{360},$$

$$\gamma_4^{(n+1)} = \gamma_4^{(n)} + \frac{a^{(n)}\sigma^{(n)}\left(\left(6a^{(n)} + \nu^{(n)}\right)\sigma^{(n)2} - 60\beta^{(n)}\right)}{180}.$$

The relations between the multipliers ν, σ, α, β, and γ_{1-4} corresponding to the second line of (5.18) are:

$$
\begin{aligned}
&\nu^{(n+1)} = \nu^{(n)}, \quad \sigma^{(n+1)} = \sigma^{(n)} + 2b^{(n)}, \\
&\alpha^{(n+1)} = \alpha^{(n)} + \frac{b^{(n)}\nu^{(n)2}}{6}, \\
&\beta^{(n+1)} = \beta^{(n)} - \frac{12c^{(n)} + b^{(n)}\nu^{(n)}\left(b^{(n)} + \sigma^{(n)}\right)}{6}, \\
&\gamma_1^{(n+1)} = \gamma_1^{(n)} - \frac{b^{(n)}\nu^{(n)4}}{360}, \\
&\gamma_2^{(n+1)} = \gamma_2^{(n)} - \frac{\nu^{(n)}\left(60\alpha^{(n)}b^{(n)} - \nu^{(n)}\left(30c^{(n)} - b^{(n)}\nu^{(n)}\left(6b^{(n)} + \sigma^{(n)}\right)\right)\right)}{180}, \\
&\gamma_3^{(n+1)} = \gamma_3^{(n)} + \frac{b^{(n)}\nu^{(n)}\left(60\alpha^{(n)} + \nu^{(n)2}\left(4b^{(n)} - \sigma^{(n)}\right)\right)}{360}, \\
&\gamma_4^{(n+1)} = \gamma_4^{(n)} - \frac{30\alpha^{(n)}b^{(n)}\left(b^{(n)} + \sigma^{(n)}\right) - \nu^{(n)}\left(30\beta^{(n)}b^{(n)} + 60b^{(n)}c^{(n)}\right)}{180} \\
&\qquad\qquad - \frac{\nu^{(n)}\left(3b^{(n)3}\nu^{(n)} + 30c^{(n)}\sigma^{(n)} - 2b^{(n)2}\nu^{(n)}\sigma^{(n)} - b^{(n)}\nu^{(n)}\sigma^{(n)2}\right)}{180}.
\end{aligned}
\tag{5.21}
$$

Thus, the multipliers can readily be obtained at the end (i.e. after $K-1$ steps) of the recursive process [42].

As we mentioned above the first two multipliers are particularly simple and look as $\nu = \sum_{k=1}^{K} a_k$ and $\sigma = \sum_{k=1}^{K} b_k$ then setting $\nu = 1$ and $\sigma = 1$ cancels the first-order error leading term $\mathcal{O}_1 = 0$. The next multipliers α, β, γ_{1-4} in order to eliminate higher-order error terms, namely, $\mathcal{O}_3$ and $\mathcal{O}_5$ (5.16) have to be set to zero.

The required order conditions form a system of non-linear equations, which must be solved with respect to unknown time coefficients a_k, b_k, and c_k. The number m of the conditions as well as the minimal numbers $S_{\min}^{(f)} = 2m - 1$ and $S_{\min}^{(g)}$ of stages, which are necessary to obtain at least the same number m of unknowns within non-gradient (f) and gradient (g) self-adjoint decomposition schemes [42].

According to the idea proposed in [42] the minimal number of stages corresponding to a gradient scheme of an order r can be considerably smaller (especially for high r) than that of a non-gradient $(c_k \equiv 0)$ integrator of the same order, which plays the crucial role for the main application of this work, since we would like to find an efficient algorithm this property of the force-gradient decomposition methods is essential.

Here we consider the most efficient versions of the force-gradient approach (5.12) due to Omelyan [42] and present the most commonly used numerical time integrations schemes in molecular dynamics and particularly in molecular dynamics step of the HMC algorithm for QCD.

The second order numerical time integrators include the leap-frog method, a 3-stage composition scheme

$$
\Delta(h) = e^{h\frac{\hat{V}}{2}} e^{h\hat{T}} e^{h\frac{\hat{V}}{2}},
\tag{5.22}
$$

with the shadow Hamiltonian

$$\tilde{H}_{LF} = H - \frac{h^2}{24}\Big(2\big[V,[T,V]\big] + \big[T,[T,V]\big]\Big) + \mathcal{O}(h^4)$$

and a 5-stage extension widely used in QCD computations:

$$\Delta(h)_5 = \mathrm{e}^{\frac{1}{6}h\hat{V}}\,\mathrm{e}^{\frac{1}{2}h\hat{T}}\,\mathrm{e}^{\frac{2}{3}h\hat{V}}\,\mathrm{e}^{\frac{1}{2}h\hat{T}}\,\mathrm{e}^{\frac{1}{6}h\hat{V}}, \tag{5.23}$$

which conserves the shadow Hamiltonian [14]

$$\tilde{H}_{5S} = H - \big[V,[T,V]\big]\frac{h^2}{72} + \mathcal{O}(h^4).$$

The forth order numerical integrators are presented by a basic scheme is given by the 5-stage force-gradient scheme proposed by Omelyan et al [42]

$$\Delta(h)_{5C} = \mathrm{e}^{\frac{1}{6}h\hat{V}}\,\mathrm{e}^{\frac{1}{2}h\hat{T}}\,\mathrm{e}^{\frac{2}{3}h\hat{V}+\frac{1}{72}h^3\mathcal{C}}\,\mathrm{e}^{\frac{1}{2}h\hat{T}}\,\mathrm{e}^{\frac{1}{6}h\hat{V}}. \tag{5.24}$$

This scheme conserves the shadow Hamiltonian [31]

$$\tilde{H}_{5C} = H + \left(41\Big[V,\Big[V,\big[V,[V,T]\big]\Big]\Big] + \ldots \right.$$
$$\left. + 54\Big[T,\Big[T,\big[T,[V,T]\big]\Big]\Big]\right)\frac{h^4}{155520} + \mathcal{O}(h^6),$$

which gains two orders of accuracy in the comparison to its non-gradient (5.23) case, and an extension is given by the 11-stage decomposition approach [42], recently mentioned as the integrator in the open source code openQCD as one of the standard options [36]

$$\Delta(h)_{11} = \mathrm{e}^{\sigma h\hat{V}}\,\mathrm{e}^{\eta h\hat{T}}\,\mathrm{e}^{\lambda h\hat{V}}\,\mathrm{e}^{\theta h\hat{T}}\,\mathrm{e}^{(1-(\lambda+\sigma))h\hat{V}}\,\mathrm{e}^{(1-(\theta+\eta))h\hat{T}}\,\mathrm{e}^{(1-(\lambda+\sigma))h\hat{V}}\,\mathrm{e}^{\theta h\hat{T}}\,\mathrm{e}^{\lambda h\hat{V}}\,\mathrm{e}^{\eta h\hat{T}}\,\mathrm{e}^{\sigma h\hat{V}}, \tag{5.25}$$

where $\sigma,\theta,\lambda,\gamma_1\ldots\gamma_4$ and η are parameters from [42, eqn.(71)] with a shadow Hamiltonian of the following form

$$\tilde{H}_{11} = H + \left(\gamma_1\Big[V,\Big[V,\big[V,[V,T]\big]\Big]\Big] + \ldots \right.$$
$$\left. + \gamma_4\Big[T,\Big[T,\big[T,[V,T]\big]\Big]\Big]\right)\frac{h^4}{32440} + \mathcal{O}(h^6).$$

We use above mentioned methods as a benchmark for the comparison of the numerical results and introduce nested versions of these numerical schemes in the next section.

5.3 Multirate approach

In this section we consider the case, that many solution components of an ordinary differential equation vary slowly and only a few components exhibit fast dynamics. Here it is tempting to use large step sizes for the slow components and small step sizes for the fast ones. Such integrators, called multirate methods (commonly termed nested schemes in the physics community), were first formulated by Rice (1960) and Gear and Wells [20]. These

schemes do not increase the order of the scheme, but reduce the computational costs per path by exploiting the multirate behavior given in the different parts of the right-hand side: slow forces, which are usually expensive to evaluate, but need only to be sampled at a low frequency; and fast forces, which are usually cheap to evaluate, but need to be sampled at a high frequency. In the case of our primal application to the quantum field theories we have a chance to exploit this property to our advantage, since the general problem of quantum field theory has a slow and costly fermion part and fast and cheap gauge part as we showed in Chapter 3.

First we consider a system of ODEs and suppose that the vector field f is split into summands contributing to slow and fast dynamics as following

$$\dot{y} = f(t, y), \quad \dot{y} = f^{[\text{slow}]}(t, y) + f^{[\text{fast}]}(t, y), \quad y \in \mathbb{R}^n, \tag{5.26}$$

where $f^{[\text{slow}]}(t, y)$ is more expensive to evaluate than $f^{[\text{fast}]}(t, y)$. Multirate methods can often be cast into this framework by collecting in $f^{[\text{slow}]}(t, y)$ those components of $f(t, y)$ which produce slow dynamics and in $f^{[\text{fast}]}(t, y)$ the remaining components [20].

For a given $N \geq 1$ and for the ODE (5.26) a multiple time stepping method is obtained from

$$\Phi_{\frac{h}{2}}^{[\text{slow}]} \circ \left(\Phi_{\frac{h}{M}}^{[\text{fast}]} \right)^M \circ \Phi_{\frac{h}{2}}^{[\text{slow}]} \tag{5.27}$$

where $\Phi_h^{[\text{slow}]}$ and $\Phi_h^{[\text{fast}]}$ are numerical integrators consistent with $\dot{y} = f^{[\text{slow}]}(t, y)$, respectively and $\dot{y} = f^{[\text{fast}]}(t, y)$. The method (5.27) is already stated in symmetric form . It is often called *the impulse method*, because the slow part $f^{[\text{slow}]}(t, y)$ of the vector field is used - impulse-like - only at the beginning and at the end of the step, whereas the many small sub-steps in between are concerned solely through integrating the fast system $\dot{y} = f^{[\text{fast}]}(t, y)$ [47]. Multirate time stepping approach can be visualized as it is shown on Figure 5.2.
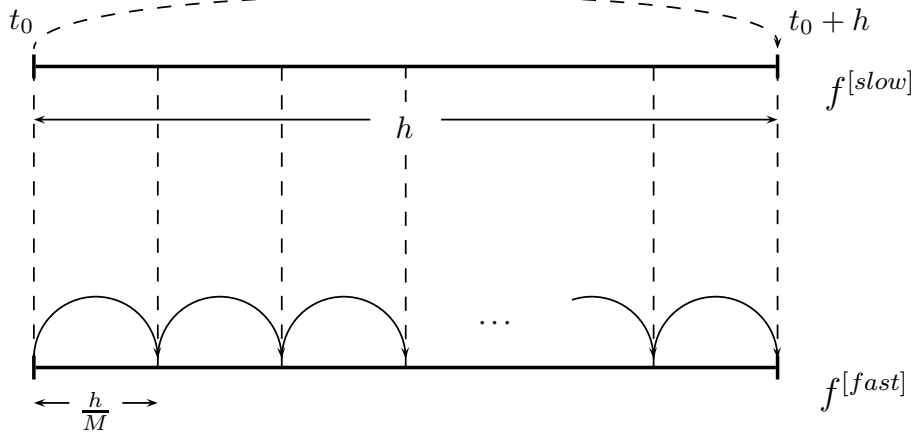

Figure 5.2: Multirate time stepping structure.

Structure preserving properties of fast-slow splitting multirate algorithms (5.26) can be easily checked via using the representation of the multirate numerical schemes by the composition approach, which results the following Lemma 4

Lemma 4. *Let* $\Phi_h^{[slow]}$ *be an arbitrary method of order 1, and* $\Phi_h^{[fast]}$ *a symmetric method*

of order 2. Then, the multiple time stepping algorithm (5.27) is symmetric and of order 2.

If $f^{[slow]}(t, y)$ and $f^{[fast]}(t, y)$ are Hamiltonian and if $\Phi_h^{[slow]}$ and $\Phi_h^{[fast]}$ are both symplectic, then the multiple time stepping method is also symplectic.

The most important application of multiple time stepping is in Hamiltonian systems with a separable Hamiltonian

$$H(p, q) = T(p) + V(q), \quad V(q) = V_1(q) + V_2(q), \tag{5.28}$$

where T represents the kinetic part, V_1 is the potential energy of the small (fast) scale part of the system and V_2 corresponds to the potential energy of the large (slow) scale part.

For example, using the notation of the standard decomposition approach (5.7) we choose the following integrator to compute the inner part $H = T + V_1$

$$\Delta(h)_M = \left[e^{\frac{h}{2M} \hat{V}_1} e^{\frac{h}{M} \hat{T}} e^{\frac{h}{2M} \hat{V}_1} \right]^M,$$

and therefore define $\Delta(h)$ a nested integrator to solve the split problem (5.28), it yields

$$\hat{\Delta}(h) = e^{\frac{h}{2} \hat{V}_2} \Delta(h)_M e^{\frac{h}{2} \hat{V}_2}. \tag{5.29}$$

The obtained method, called *nested leap-frog*, conserves the shadow Hamiltonian [53]

$$\tilde{H}_M = H + \left(-\frac{1}{24} [V_2, [V_2, T]] + \frac{1}{12} [V_1, [V_2, T]] + \frac{1}{12} [T, [V_2, T]] \right.$$
$$\left. + \frac{1}{M^2} \left(-\frac{1}{24} [V_1, [V_1, T]] + \frac{1}{12} [T, [V_1, T]] \right) \right) h^2 + \mathcal{O}(h^4).$$

Notice that for $M = 1$ the method (5.29) reduces to the Störmer-Verlet scheme applied to the Hamiltonian system with $H(q, p)$. This is a consequence of the fact that $\varphi_t^{[fast]} \circ \varphi_t^{[slow]} = \varphi_t^V$ is the exact flow of the Hamiltonian system corresponding to $V(q)$ of (5.28) [6].

5.4 Nested force-gradient schemes

Our main idea in this thesis is to combine the idea of force-gradient decomposition methods (5.12) to improve the energy preservation of symplectic schemes applied to Hamiltonian systems with multiple time splitting techniques for multi-time scales problems, using the fact that the potential of the general quantum field problems has different parts with strongly varying dynamics, thereby reducing the computational costs.

In order to get a better insight into the multi-rate formulation of the force-gradient approach, we start this section with a short introduction into the nested integration idea proposed by Sexton and Weingarten [49].

The basis for multirate schemes is to split the action into different parts according to the

activity level:

$$H(U, P) = \frac{\langle P, P \rangle}{2} + \sum_{m=0}^{M} S_k[U, \phi] \tag{5.30}$$

If the actions of the Hamiltonian (5.30) are ordered such that computational costs are increasing, while at the same time the strength of the forces is decreasing, a multi-rate integration based on the leap-frog scheme with macro step size h_0 and micro step size $h_1 := h_0/m_1$ due to Sexton and Weingarten [49] proceeds as follows:

$$V_H(h_0) = V_{H_2}(\frac{h_0}{2}) \, (V_{H_1}(h_1))^{m_1} \, V_{H_2}(\frac{h_0}{2})$$

with $H_1(U, P) := \langle P, P \rangle/2 + \sum_{m=0}^{M-1} S_k[U, \phi]$ and $H_2(U, P) = S_M[U, \phi]$. This scheme can be nested, by introducing a next finer step size $h_2 := h_1/m_2$ and further splitting H_1 into

$$H_1(U, P) = H_{11}(U, P) + H_{12}(U, P)$$

$H_{11}(U, P) := \langle P, P \rangle/2 + \sum_{m=0}^{M-2} S_k[U, \phi]$ and $H_{12}(U, P) = S_{M-1}[U, \phi]$ in order to replace $V_{H_1}(h_1)$ above by

$$V_{H_1}(h_1) = V_{H_{12}}(\frac{h_1}{2}) \, (V_{H_{11}}(h_2))^{m_2} \, V_{H_{12}}(\frac{h_1}{2})$$

This procedure can be applied recursively [58] to obtain M different step size ratios at the end, corresponding to the activity levels of the M actions $S_k[U, \phi]$.

This approach can be easily generalized to base schemes of higher order such as force-gradient schemes. It can be easily shown that by construction the overall multirate scheme inherits the properties of the underlying base scheme: symmetry, time-reversibility, volume preservation and order (according to the slowest time step h). Nested force-gradient schemes in QCD have been tested by [14]; applications to Domain Wall fermions are discussed in [3].

A nested version of a standard force-gradient integrator (5.24), which probably has been used in [3] reads

$$\hat{\Delta}_{5C}(h) = e^{\frac{1}{6}h\hat{V}_2} \, \Delta_2\left(\frac{h}{2}\right)_M e^{\frac{2}{3}h\hat{V}_2 + \frac{1}{72}h^3 \mathcal{C}_f} \, \Delta_2\left(\frac{h}{2}\right)_M e^{\frac{1}{6}h\hat{V}_2}, \tag{5.31}$$

where

$$\Delta_2(h)_M = \left(e^{\frac{1}{6M}h\hat{V}_1} \, e^{\frac{1}{2M}h\hat{T}} \, e^{\frac{2}{3M}h\hat{V}_1 + \frac{1}{72}\left(\frac{h}{M}\right)^3 \mathcal{C}_g} \, e^{\frac{1}{2M}h\hat{T}} \, e^{\frac{1}{6M}h\hat{V}_1} \right)^M,$$

with $\mathcal{C}_g = [V_1, [T, V_1]]$ and $\mathcal{C}_f = [V_2, [T, V_2]]$. It is important to notice that, apart of implementation of the integrator (5.31), there was no detailed analysis made for the integrator properties.

Despite the fact that the idea of nested force-gradient methods is of a great interest, in general it remained relatively unknown and there is not that many research been published on this matter.

One serious drawback of this approach is that the same base scheme is used at all levels, i.e., high computational costs of high order schemes appear at all levels. In order to overcome this limitation it is natural to suggest using different schemes of different order at the different levels without loosing the effective high order of the overall multirate

scheme. For the latter, we could make use of the idea of including appropriate force-gradient information. For this, we restrict the system (5.30) to two time levels, as the ultimate goal will be to develop nested force-gradient schemes for lattice QCD, where the gauge field plays the role of the fast action and cheap part, and the fermion field plays the role of the slow action and expensive part, i.e., only two time scales are involved.

If one uses a second-order scheme for the fast action with time step h/M and a fourth-order scheme for the slow action with time step h, the error of the overall multirate scheme will be of order $\mathcal{O}(h^2) + \mathcal{O}\left(\left(\frac{h}{M}\right)^2\right) + \mathcal{O}(h^4)$ the use of force-gradient information only at the slowest level it is possible to cancel the leading error term of order $\mathcal{O}(h^2)$. As $M \geq \frac{1}{h}$ usually holds in the multirate setting, the overall order is then given by the leading error term of order $\mathcal{O}(h^4)$, i.e., the scheme has an effective order of four. In the next section we perform the proof of this statement.

5.5 Adapted nested force-gradient schemes

In order to overcome the previously mentioned drawbacks of a standard nested force-gradient integrator (5.31) we propose the following scheme introduced in [51]

$$\tilde{\Delta}_{5C}(h) = \mathrm{e}^{\frac{1}{6}h\hat{V}_2}\,\Delta_1\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{2}{3}h\hat{V}_2+\frac{1}{72}h^3\mathcal{C}_f}\,\Delta_1\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{1}{6}h\hat{V}_2}. \tag{5.32}$$

To summarize, the adapted scheme (5.32) differs from the original one (5.31) in two perspectives:

- The force gradient scheme for the fast action is replaced by a leap-frog scheme.

- Only the part $[V_2, [T, V_2]]$ of the full force gradient $[V_1 + V_2, [T, V_1 + V_2]]$ is needed to gain the effective order of four.

Geometric properties of the numerical flow such as *time-reversibility* and *symplecticity* preserved by the construction, since the scheme (5.32) can be presented as the composition of the time-reversible and symplectic integrators and hence also possesses these properties.

Now it is crucial to show that the proposed integrator has the effective order of four. Let us start by introducing following *alike 5-stage nested integrator*

$$\Delta(h) = \mathrm{e}^{\lambda h\hat{V}_2}\,\Delta\left(\frac{h}{2}\right)_M \mathrm{e}^{(1-2\lambda)h\hat{V}_2}\,\Delta\left(\frac{h}{2}\right)_M \mathrm{e}^{\lambda h\hat{V}_2}, \tag{5.33}$$

where

$$\Delta\left(\frac{h}{2}\right)_M = \left[\mathrm{e}^{\frac{h}{4M}\hat{V}_1}\,\mathrm{e}^{\frac{h}{2M}\hat{T}}\,\mathrm{e}^{\frac{h}{4M}\hat{V}_1}\right]^M$$

and analyzing the energy conservation of this scheme. The method (5.33) was chosen to be a base of our proposed integrator (5.32), because it has an optimal number of stages, necessary for increasing its order.

We have to find the shadow Hamiltonian for the integrator (5.33). In order to do so, we use the BCH formula (2.10). To simplify this task we consider the limit of the integrator given in (5.33), as M tends to infinity.

We obtain the following representation of (5.33)

$$\Delta(h) = e^{\lambda h \hat{V}_2} \, e^{\frac{h}{2}(\hat{V}_1 + T)} \, e^{(1-2\lambda)h\hat{V}_2} \, e^{\frac{h}{2}(\hat{V}_1 + \hat{T})} \, e^{\lambda h \hat{V}_2} \, . \tag{5.34}$$

Theorem 9 (Shadow Hamiltonian of (5.34)). *The shadow Hamiltonian of the nested multirate integrator* (5.34) *is given by*

$$\tilde{H} = H + \Big(\frac{-1 + 6\lambda - 6\lambda^2}{12} \big[V_2, [T, V_2]\big]$$
$$+ \frac{-1 + 6\lambda}{24} \big[V_1, [T, V_2]\big] + \frac{-1 + 6\lambda}{24} \big[T, [T, V_2]\big] \Big) h^2 + \mathcal{O}(h^4). \tag{5.35}$$

Proof: We apply the BCH formula to the first two evolution operators

$$X = \ln \left(e^{\lambda h \hat{V}_2} \, e^{h \frac{\hat{V}_1 + \hat{T}}{2}} \right) = c_1 h + c_2 h^2 + c_3 h^3 + \mathcal{O}(h^5),$$

where

$$c_1 = \lambda V_2 + \frac{V_1 + T}{2},$$
$$2c_2 = \frac{B_2}{2!} \mathrm{ad} c_1 \left(\lambda V_2 + \frac{V_1 + T}{2} \right) - \frac{1}{2} \mathrm{ad} c_1 \left(\lambda V_2 - \frac{V_1 + T}{2} \right),$$
$$c_2 = -\frac{1}{4} \left[\lambda V_2 + \frac{V_1 + T}{2}, \ \lambda V_2 - \frac{V_1 + T}{2} \right] = -\frac{\lambda}{4}[V_1, V_2] - \frac{\lambda}{4}[T, V_2],$$
$$3c_3 = \frac{B_3}{3!} \mathrm{ad} c_1 \mathrm{ad} c_1 \left(\lambda V_2 + \frac{V_1 + T}{2} \right) - \frac{1}{2} \mathrm{ad} c_2 \left(\lambda V_2 - \frac{V_1 + T}{2} \right),$$
$$c_3 = -\frac{1}{6} \left[-\frac{\lambda}{4}[V_1, V_2] - \frac{\lambda}{4}[T, V_2], \ \lambda V_2 - \frac{V_1 + T}{2} \right]$$
$$= -\frac{\lambda^2}{24} \big[V_2, [V_1, V_2]\big] - \frac{\lambda^2}{24} \big[V_2, [T, V_2]\big] + \frac{\lambda}{48} \big[V_1, [V_1, V_2]\big]$$
$$+ \frac{\lambda}{48} \big[V_1, [T, V_2]\big] + \frac{\lambda}{48} \big[T, [V_1, V_2]\big] + \frac{\lambda}{48} \big[T, [T, V_2]\big].$$

Then we have the result for our first two operators

$$X = \left(\lambda V_2 + \frac{V_1 + T}{2} \right) h + \left(-\frac{\lambda}{4}[V_1, V_2] - \frac{\lambda}{4}[T, V_2] \right) h^2$$
$$+ \left(-\frac{\lambda^2}{24} \big[V_2, [V_1, V_2]\big] - \frac{\lambda^2}{24} \big[T, [T, V_2]\big] + \frac{\lambda}{48} \big[V_1, [V_1, V_2]\big] \right.$$
$$\left. + \frac{\lambda}{48} \big[V_1, [T, V_2]\big] + \frac{\lambda}{48} \big[T, [V_1, V_2]\big] + \frac{\lambda}{48} \big[T, [T, V_2]\big] \right) h^3 + \mathcal{O}(h^5).$$

The next step is to apply the BCH formula on the following operators

$$Y = \ln \left(e^X \, e^{(1-2\lambda)h\hat{V}_2} \right) = c_1 + c_2 + c_3 + \mathcal{O}(h^5)$$

and coefficients

$$c_1 = X + (1 - 2\lambda)hV_2 = \left((1 - \lambda)V_2 + \frac{V_1 + T}{2}\right)h + (\ldots)h^2 + (\ldots)h^3,$$

$$2c_2 = \frac{B_2}{2!}\mathrm{ad}c_1\big(X + (1 - 2\lambda)hV_2\big) - \frac{1}{2}\mathrm{ad}c_1\big(X - (1 - 2\lambda)hV_2\big),$$

$$c_2 = -\frac{1}{4}\big[X + (1 - 2\lambda)hV_2, \ X - (1 - 2\lambda)hV_2\big]$$

$$= \left(\frac{1 - 2\lambda}{4}[V_1, V_2] + \frac{1 - 2\lambda}{4}[T, \ V_2]\right)h^2$$

$$+ \left(\frac{\lambda - 2\lambda^2}{8}[V_2, [V_1, V_2]] + \frac{\lambda - 2\lambda^2}{8}[V_2, [T, V_2]]\right)h^3$$

$$3c_3 = \frac{B_3}{3!}\mathrm{ad}c_1\mathrm{ad}c_1\big(X + (1 - 2\lambda)hV_2\big) - \frac{1}{2}\mathrm{ad}c_2\big(X - (1 - 2\lambda)hV_2\big),$$

$$c_3 = -\frac{1}{6}\big[c_2, \ X - (1 - 2\lambda)hV_2\big]$$

$$= \left(-\frac{(1 - 2\lambda)(1 - 3\lambda)}{24}[V_2, [V_1, V_2]] - \frac{(1 - 2\lambda)(1 - 3\lambda)}{24}[V_2, [T, V_2]]\right.$$

$$+ \frac{(1 - 2\lambda)}{48}[V_1, [V_1, V_2]] + \frac{(1 - 2\lambda)}{48}[V_1, [T, V_2]] + \frac{(1 - 2\lambda)}{48}[T, [V_1, V_2]]$$

$$\left. + \frac{(1 - 2\lambda)}{48}[T, [T, V_2]]\right)h^3,$$

and we obtain the following expansion

$$Y = \left((1 - \lambda)V_2 + \frac{V_1 + T}{2}\right)h + \left(\frac{1 - 3\lambda}{4}[V_1, V_2] + \frac{1 - 3\lambda}{4}[T, V_2]\right)h^2$$

$$+ \left(\frac{-1 + 8\lambda - 13\lambda^2}{24}[V_2, [V_1, V_2]] + \frac{-1 + 8\lambda - 13\lambda^2}{24}[T, [T, V_2]]\right.$$

$$+ \frac{1 - \lambda}{48}[V_1, [V_1, V_2]] + \frac{1 - \lambda}{48}[V_1, [T, V_2]] + \frac{1 - \lambda}{48}[T, [V_1, V_2]]$$

$$\left. + \frac{1 - \lambda}{48}[T, [T, V_2]]\right)h^3 + \mathcal{O}(h^5),$$

The next step would be to repeat the previous procedures to find

$$Z = \ln\left(e^Y e^{h\frac{\hat{V}_1 + \hat{T}}{2}}\right) = c_1 + c_2 + c_3 + \mathcal{O}(h^5).$$

Using the BCH formula we obtain

$$
c_1 = Y + h\frac{V_1 + T}{2} = ((1-\lambda)V_2 + V_1 + T)h + (\ldots)h^2 + (\ldots)h^3,
$$

$$
2c_2 = \frac{B_2}{2!}\mathrm{ad}c_1\left(Y + h\frac{V_1 + T}{2}\right) - \frac{1}{2}\mathrm{ad}c_1\left(Y - h\frac{V_1 + T}{2}\right),
$$

$$
c_2 = -\frac{1}{4}\left[Y + h\frac{V_1 + T}{2},\; Y - h\frac{V_1 + T}{2}\right]
$$

$$
= \left(-\frac{1-\lambda}{4}[V_1,\; V_2] - \frac{1-\lambda}{4}[T,\; V_2]\right)h^2
$$

$$
+ \left(-\frac{1-3\lambda}{16}\big[V_1, [V_1, V_2]\big] - \frac{1-3\lambda}{16}\big[V_1, [T, V_2]\big]\right.
$$

$$
\left. -\frac{1-3\lambda}{16}\big[T, [V_1, V_2]\big] - \frac{1-3\lambda}{16}\big[T, [T, V_2]\big]\right)h^3,
$$

$$
3c_3 = \frac{B_3}{3!}\mathrm{ad}c_1\mathrm{ad}c_1\left(Y + h\frac{V_1 + T}{2}\right) - \frac{1}{2}\mathrm{ad}c_2\left(Y - h\frac{V_1 + T}{2}\right)
$$

$$
c_3 = -\frac{1}{6}\left[c_2,\; Y - h\frac{V_1 + T}{2}\right],
$$

$$
= \left(-\frac{(1-\lambda)^2}{24}\big[V_2, [V_1, V_2]\big] - \frac{(1-\lambda)^2}{24}\big[V_2, [T, V_2]\big]\right)h^3.
$$

Therefore we obtain

$$
Z = \big((1-\lambda)V_2 + V_1 + T\big)h + \left(\frac{-\lambda}{2}[V_1, V_2] + \frac{-2\lambda}{2}[T, V_2]\right)h^2
$$

$$
+ \left(\frac{-1 + 5\lambda - 7\lambda^2}{12}\big[V_2, [V_1, V_2]\big] + \frac{-1 + 5\lambda - 7\lambda^2}{12}\big[T, [T, V_2]\big]\right.
$$

$$
+ \frac{-1 + 4\lambda}{48}\big[V_1, [V_1, V_2]\big] + \frac{-1 + 4\lambda}{48}\big[V_1, [T, V_2]\big]
$$

$$
\left. + \frac{-1 + 4\lambda}{48}\big[T, [V_1, V_2]\big] + \frac{-1 + 4\lambda}{48}\big[T, [T, V_2]\big]\right)h^3 + O(h^4).
$$

Applying the BCH formula for a last time we obtain the shadow Hamiltonian

$$
\tilde{H} = \ln\left(e^Z\, e^{\lambda\hat{V}_2 h}\right) = c_1 + c_2 + c_3 + \mathcal{O}(h^5),
$$

with the coefficients

$$
\begin{aligned}
c_1 &= Z + \lambda V_2 h = (V_2 + V_1 + T)h + (\ldots)h^2 + (\ldots)h^3, \\
2c_2 &= \frac{B_2}{2!}\mathrm{ad}c_1\left(Z + \lambda V_2 h\right) - \frac{1}{2}\mathrm{ad}c_1\left(Z - \lambda V_2 h\right), \\
c_2 &= -\frac{1}{4}\left[Z + \lambda V_2 h,\ Z - \lambda V_2 h\right] \\
&= \left(\frac{\lambda}{2}[V_1,\ V_2] + \frac{\lambda}{2}[T,\ V_2]\right)h^2 + \left(\frac{\lambda^2}{4}\left[V_2,[V_1,V_2]\right] + \frac{\lambda^2}{4}\left[V_2,[T,V_2]\right]\right)h^3, \\
3c_3 &= \frac{B_3}{3!}\mathrm{ad}c_1\mathrm{ad}c_1\left(Z + \lambda V_2 h\right) - \frac{1}{2}\mathrm{ad}c_2\left(Z - \lambda V_2 h\right), \\
c_3 &= -\frac{1}{6}\left[c_2,\ Z - \lambda V_2 h\right] \\
&= \left(\frac{\lambda(1-2\lambda)^2}{12}\left[V_2,[V_1,V_2]\right] + \frac{\lambda(1-2\lambda)^2}{12}\left[V_2,[T,V_2]\right] + \frac{\lambda}{12}\left[V_1,[V_1,V_2]\right]\right. \\
&\quad \left. + \frac{\lambda}{12}\left[V_1,[T,V_2]\right] + \frac{\lambda}{12}\left[T,[V_1,V_2]\right] + \frac{\lambda}{12}\left[T,[T,V_2]\right]\right)h^3,
\end{aligned}
$$

and finally

$$
\begin{aligned}
\tilde{H} = H &+ \left(\frac{-1+6\lambda-6\lambda^2}{12}\left[V_2,[V_1,V_2]\right] + \frac{-1+6\lambda}{24}\left[V_1,[T,V_2]\right]\right. \\
&+ \frac{-1+6\lambda-6\lambda^2}{12}\left[V_2,[T,V_2]\right] + \frac{-1+6\lambda}{24}\left[V_1,[V_1,V_2]\right] \\
&\left. + \frac{-1+6\lambda}{24}\left[T,[V_1,V_2]\right] + \frac{-1+6\lambda}{24}\left[T,[T,V_2]\right]\right)h^2 + \mathcal{O}(h^4).
\end{aligned}
$$

Finally, taking into account that $[V_1,V_2] = 0$, hence $\left[V_2,[V_1,V_2]\right]$, $\left[V_1,[V_1,V_2]\right]$ and $\left[T,[V_1,V_2]\right]$ are equal to zero, we obtain (5.35). $\qquad\square$

We can eliminate a couple of terms by choosing $\lambda = 1/6$, thus

$$
\tilde{H} = H - \frac{1}{72}\left[V_2,[T,V_2]\right]h^2 + \mathcal{O}(h^4).
$$

We would like to increase the order of the method (5.33) by adding the force-gradient term, but first we consider the force-gradient itself. Due to the splitting (5.28) it can be represented as

$$
\begin{aligned}
\mathcal{C} = \left[V,[T,V]\right] &= \left[V_2 + V_1,[T,V_2 + V_1]\right] \\
&= \left[V_2,[T,V_2]\right] + \left[V_1,[T,V_1]\right] + \left[V_1,[T,V_2]\right] + \left[V_2,[T,V_1]\right].
\end{aligned}
$$

Then we can tune the original algorithm (5.33) by adding the first term of the force gradient $\left[V_2,[T,V_2]\right]$ and neglect the last three terms:

$$
\Delta(h) = \mathrm{e}^{\frac{1}{6}h\hat{V}_2}\,\Delta\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{2}{3}h\hat{V}_2 + \frac{1}{72}h^3\left[V_2,[T,V_2]\right]}\,\Delta\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{1}{6}h\hat{V}_2}, \tag{5.36}
$$

which preserves the fourth-order accurate shadow Hamiltonian

$$
\tilde{H} = H + \mathcal{O}(h^4).
$$

So now we can analyze the shadow Hamiltonian for the nested method (5.33) for the general case when the parameter M does not tend to infinity.

First, let us show that the evolution operator $\Delta(h)_M$ can also be presented as

$$\Delta\left(\frac{h}{2}\right)_M = e^{\frac{h}{2}(V_1+T)+\frac{h^3}{M^2}\left(\frac{1}{192}[V_1,\,[T,\,V_1]]+\frac{1}{96}[T,\,[T,\,V_1]]\right)+\mathcal{O}(h^5)}.$$

It is derived by applying the BCH formula on the original propagator $\Delta(h)_M$. Therefore the proposed scheme (5.33) can be rewritten as

$$\Delta(h) = e^{\lambda h V_2}\, e^{\frac{h}{2}(V_1+T)+\frac{h^3}{M^2}\ldots\mathcal{O}(h^5)}\, e^{(1-2\lambda)hV_2}\, e^{\frac{h}{2}(V_1+T)+\frac{h^3}{M^2}\ldots+\mathcal{O}(h^5)}\, e^{\lambda h V_2}.$$

To find the shadow Hamiltonian we apply the BCH formula to the first two evolution operators

$$X = \ln\left(e^{\lambda h V_2}\, e^{h\frac{V_1+T}{2}+\frac{h^3}{M^2}(\ldots)}\right) = c_1 h + c_2 h^2 + c_3 h^3 + \mathcal{O}(h^5),$$

where

$$c_1 = \lambda V_2 + \frac{V_1+T}{2} + \frac{h^3}{M^2}(\ldots) + \mathcal{O}(h^5)$$

$$2c_2 = \frac{B_2}{2!}\mathrm{ad}c_1\left(\lambda V_2 + \frac{V_1+T}{2}\right) - \frac{1}{2}\mathrm{ad}c_1\left(\lambda V_2 - \frac{V_1+T}{2}\right)$$

$$c_2 = -\frac{1}{4}\left[\lambda V_2 + \frac{V_1+T}{2},\ \lambda V_2 - \frac{V_1+T}{2}\right] = -\frac{\lambda}{4}[T, V_2],$$

$$3c_3 = \frac{B_3}{3!}\mathrm{ad}c_1\mathrm{ad}c_1\left(\lambda V_2 + \frac{V_1+T}{2}\right) - \frac{1}{2}\mathrm{ad}c_2\left(\lambda V_2 - \frac{V_1+T}{2}\right)$$

$$c_3 = -\frac{1}{6}\left[-\frac{\lambda}{4}[T, V_2],\ \lambda V_2 - \frac{V_1+T}{2}\right]$$

$$= -\frac{\lambda^2}{24}[V_2,\,[T,\,V_2]] + \frac{\lambda}{48}[V_1,\,[T,\,V_2]] + \frac{\lambda}{48}[T,\,[T,\,V_2]].$$

Then for the first two operators we have

$$X = h\left(\lambda V_2 + \frac{V_1+T}{2}\right) - \frac{\lambda h^2}{4}[T, V_2] - \frac{\lambda^2 h^3}{24}[V_2,\,[T,\,V_2]]$$

$$+ \frac{\lambda h^3}{48}[V_1,\,[T,\,V_2]] + \frac{\lambda h^3}{48}[T,\,[T,\,V_2]] + \frac{h^3}{M^2}(\ldots) + \mathcal{O}(h^5).$$

Next we apply the BHC formula to

$$Y = \ln\left(e^X\, e^{(1-2\lambda)hV_2}\right) = c_1 h + c_2 h^2 + c_3 h^3 + \mathcal{O}(h^5)$$

with coefficients

$$c_1 = X + (1 - 2\lambda)V_2 = (1 - \lambda)V_2 + \frac{V_1 + T}{2} + \ldots + \mathcal{O}(h^5)$$

$$2c_2 = \frac{B_2}{2!}\mathrm{ad}c_1\,(X + (1 - 2\lambda)V_2) - \frac{1}{2}\mathrm{ad}c_1\,(X - (1 - 2\lambda)V_2)$$

$$c_2 = -\frac{1}{4}[X + (1 - 2\lambda)V_2,\ X - (1 - 2\lambda)V_2] = \frac{1 - 2\lambda}{4}[T, V_2] + \frac{\lambda - 2\lambda^2}{8}[V_2, [T, V_2]]$$

$$3c_3 = \frac{B_3}{3!}\mathrm{ad}c_1\mathrm{ad}c_1\,(X + (1 - 2\lambda)V_2) - \frac{1}{2}\mathrm{ad}c_2\,(X - (1 - 2\lambda)V_2)$$

$$c_3 = -\frac{1}{6}[c_2, X - (1 - 2\lambda)V_2] = \frac{(1 - 2\lambda)(-1 + 3\lambda)}{24}[V_2, [T, V_2]]$$

$$+ \frac{(1 - 2\lambda)}{48}[V_1, [T, V_2]] + \frac{(1 - 2\lambda)}{48}[T, [T, V_2]]$$

and we obtain the following expansion

$$Y = \left((1 - \lambda)V_2 + \frac{V_1 + T}{2}\right)h + \frac{(1 - 3\lambda)h^2}{4}[T, V_2] + \frac{(-1 + 8\lambda - 13\lambda^2)h^3}{24}[V_2, [T, V_2]]$$

$$+ \frac{(1 - \lambda)h^3}{48}[V_1, [T, V_2]] + \frac{(1 - \lambda)h^3}{48}[T,\ [T,\ V_2]] + \frac{h^3}{M^2}(\ldots) + \mathcal{O}(h^5).$$

Then by repeating the previous procedures we obtain

$$Z = \ln\left(\mathrm{e}^Y\,\mathrm{e}^{+\frac{h^3}{M^2}\ldots + \mathcal{O}(h^5)}\right) = c_1 h + c_2 h^2 + c_3 h^3 + \mathcal{O}(h^5).$$

Using the BCH formula yields

$$c_1 = Y + \frac{V_1 + T}{2} = (1 - \lambda)V_2 + V_1 + T + \ldots + \mathcal{O}(h^5)$$

$$2c_2 = \frac{B_2}{2!}\mathrm{ad}c_1\left(Y + \frac{V_1 + T}{2}\right) - \frac{1}{2}\mathrm{ad}c_1\left(Y - \frac{V_1 + T}{2}\right)$$

$$c_2 = -\frac{1}{4}\left[Y + \frac{V_1 + T}{2}, Y - \frac{V_1 + T}{2}\right] = -\frac{1 - \lambda}{4}[T, V_2]$$

$$- \frac{1 - 3\lambda}{16}[V_1, [T, V_2]] - \frac{1 - 3\lambda}{16}[T, [T, V_2]]$$

$$3c_3 = \frac{B_3}{3!}\mathrm{ad}c_1\mathrm{ad}c_1\left(Y + \frac{V_1 + T}{2}\right) - \frac{1}{2}\mathrm{ad}c_2\left(Y - \frac{V_1 + T}{2}\right)$$

$$c_3 = -\frac{1}{6}\left[c_2,\ Y - \frac{V_1 + T}{2}\right] = -\frac{(1 - \lambda)^2}{24}[V_2, [T, V_2]]$$

and hence we obtain

$$Z = ((1 - \lambda)V_2 + V_1 + T)\,h + \frac{-\lambda h^2}{2}[T, V_2] + \frac{(-1 + 5\lambda - 7\lambda^2)h^3}{12}[V_2, [T, V_2]]$$

$$+ \frac{(-1 + 4\lambda)h^3}{48}[V_1, [T, V_2]] + \frac{(-1 + 4\lambda)h^3}{48}[T, [T, V_2]] + \frac{2h^3}{M^2}(\ldots) + \mathcal{O}(h^5).$$

Then using the BCH formula for the last time we obtain the shadow Hamiltonian

$$\tilde{H} = \ln\left(\mathrm{e}^Z\,\mathrm{e}^{\lambda V_2}\right) = c_1 h + c_2 h^2 + c_3 h^3 + \mathcal{O}(h^5)$$

with coefficients

$$c_1 = Z + \lambda V_2 = (V_2 + V_1 + T)h + \ldots + \mathcal{O}(h^5)$$

$$2c_2 = \frac{B_2}{2!}\mathrm{ad}c_1(Z + \lambda V_2) - \frac{1}{2}\mathrm{ad}c_1(Z - \lambda V_2)$$

$$c_2 = -\frac{1}{4}[Z + \lambda V_2, Z - \lambda V_2] = \frac{\lambda}{2}[T, V_2] + \frac{\lambda^2}{4}[V_2, [T, V_2]]$$

$$3c_3 = \frac{B_3}{3!}\mathrm{ad}c_1\mathrm{ad}c_1(Z + \lambda V_2) - \frac{1}{2}\mathrm{ad}c_2(Z - \lambda V_2)$$

$$c_3 = -\frac{1}{6}[c_2, Z - \lambda V_2]$$

$$= \frac{\lambda(1 - 2\lambda)}{12}[V_2,\ [T,\ V_2]] + \frac{\lambda}{12}[V_1,\ [T,\ V_2]] + \frac{\lambda}{12}[T,\ [T,\ V_2]],$$

which finally yields

$$\tilde{H} = V_2 + V_1 + T + \frac{(-1 + 6\lambda - 6\lambda^2)h^2}{12}[V_2, [T, V_2]] + \frac{(-1 + 6\lambda)h^2}{24}[V_1, [T, V_2]]$$

$$+ \frac{(-1 + 6\lambda)h^2}{24}[T,\ [T,\ V_2]] + \frac{h^2}{M^2}\left(\frac{1}{96}[V_1, [T, V_1]] + \frac{1}{48}[T, [T, V_1]]\right) + \mathcal{O}(h^4).$$

Similar to the case when $M \to \infty$ a choice $\lambda = 1/6$ eliminates the error terms $[V_1, [T, V_2]]$ and $[T,\ [T,\ V_2]]$; by adding the force-gradient term as for the numerical scheme (5.36) we obtain

$$\tilde{H} = H + \frac{h^2}{M^2}\left(\frac{1}{96}[V_1, [T, V_1]] + \frac{1}{48}[T, [T, V_1]]\right) + \mathcal{O}(h^4),$$

which results to an order of the method (5.32) to be $\mathcal{O}\left(\left(\frac{h}{M}\right)^2\right) + \mathcal{O}(h^4)$ but since in the multirate setting the parameter M usually satisfies the relation $M \geq \frac{1}{h}$, the overall order is then given by the leading error term of order $\mathcal{O}(h^4)$, i.e., the numerical scheme (5.32) has an effective order of four. We show the numerical proof in the next chapter.

In this chapter we have introduced a new decomposition scheme for Hamiltonian systems, which combines the idea of the force-gradient time-reversible and symplectic integrators and the splitting approach of nested algorithms. The new method (5.32) is fourth-order accurate. We also introduced some of the most commonly used numerical time integrators for the molecular dynamics such as for example leap-frog (5.22), its nested version (5.29), standard 5 stage decomposition integrator (5.23) and its nested version (5.33), which are second order; (5.24) and (5.25), which have the fourth order of convergence. Hence these are competitors to the proposed method (5.32).

We do not consider integrators of the order higher than 4, since the computational costs are too high. The schemes of the same convergence order differ from each other by the number of stages (updates of momenta and links per time step). Usually methods with more stages have smaller leading error coefficients and therefore have better accuracy, but higher computational costs. We would like to determine which integrator would represent the best compromise between high accuracy and computational efficiency.

In the next chapter we test our proposed numerical integrator in the comparison to the above mentioned standard numerical methods for the molecular dynamics step of the HMC algorithm. In order to analyze the behavior of the integrators we apply them to two different test problems, which are introduced next.

6 NUMERICAL RESULTS

Chapter 6

We compare the behavior of numerical time integration schemes currently used for HMC [42] with the nested force-gradient integrator [14] and the adapted version introduced in [51]. We investigate the computational costs needed to perform numerical calculations, as well as the effort required to achieve a satisfactory acceptance rate during the HMC evolution. Our goal is to find a numerical scheme for the HMC algorithm which would provide a sufficiently high acceptance rate while not drastically increasing the simulation time.

In this chapter we present results of numerical experiments, which were designed to test the new idea of the adapted nested force-gradient method. Since it was previously mentioned that all the considered numerical schemes preserve geometric properties of the numerical flow of the system of ODEs here we concentrate on the convergence order of the integrators and their computational efficiency.

First we test most commonly used geometric integrators on the classical n-body problem, then we apply these integrators to QED's two-dimensional Schwinger model and discuss our obtained results.

6.1 N-body problem

In order to obtain a better insight in dealing with the force-gradient (5.12) and multirate (5.27) numerical integration schemes we first apply certain numerical methods for solving the well known problem *n-body problem* of astrophysics.

This problem has been chosen because it can model the multirate time scaling for the multirate methods to be used. Also the case of the n-body problem we chose to consider is fairly simple (3-body problem) and may not yield significant advantages for the nested type integrators (the difference between fast and slow part of the action is too small), but it can still demonstrate important features of the proposed nested force-gradient schemes.

We recall that for classical mechanical systems, the equation of motion can be written as

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \{\rho, H\} \equiv \mathcal{L}(t)\rho(t), \tag{6.1}$$

where ρ is the set of phase variables, $\{\ ,\ \}$ denotes the Poisson bracket, H represents the Hamiltonian function, and $\mathcal{L}$ denotes the Liouville operator.

For the case of N particles, located in a spatially inhomogeneous time-dependent external field $u(q_i, t)$ and interacting through the pair-wise potential $\varphi(q_{ij}) \equiv \varphi(|q_i - q_j|)$, the

Hamiltonian reads

$$H = \sum_{i=1}^{N} \frac{m_i v_i^2}{2} + \frac{1}{2} \sum_{i \neq j}^{N} \varphi(q_{ij}) + \sum_{i=1}^{N} u(q_i, \ t) \equiv T(v) + V(q, t). \tag{6.2}$$

Here q_i represents the position of particle i $(i = 1, 2, \ldots, N)$ moving with velocity $v_i = \mathrm{d}q_i/\mathrm{d}t$ and carrying the mass m_i, so that T and V are the total kinetic and potential energies, respectively. Then $\rho = \{q_i, \ v_i\} \equiv \{q, \ v\}$, and the *Liouville operator* of the system takes the form

$$\mathcal{L}(t) = \sum_{i=1}^{N} \left(v_i \cdot \frac{\partial}{\partial q_i} + \frac{f_i(t)}{m_i} \cdot \frac{\partial}{\partial v_i} \right), \tag{6.3}$$

where

$$f_i(t) = - \sum_{j(j \neq i)}^{N} \frac{\varphi'(q_{ij})(q_i - q_j)}{q_{ij}} - \frac{\partial u(q_i, t)}{\partial q_i}$$

are forces acting on the particles due to their interactions.

If the initial configuration $\rho(0)$ is specified, the unique solution to the problem of equation (6.1) can be presented by the *time propagator operator* as

$$\rho(t) = \left[\mathrm{e}^{(\mathcal{D} + \mathcal{L})h} \right]^l \rho(0), \tag{6.4}$$

where h is a temporal step size and $l = t/h$ the total number of steps. $\mathcal{D} = \overleftarrow{\partial}/\partial t$ denotes the time derivative operator acting on the left of time-dependent functions. We consider in our paper $\mathcal{L}$, which does not depend explicitly on time, thus we set $\mathcal{D} = 0$.

In case of many-particle systems $(N > 2)$ the time propagator cannot be computed exactly even in the absence of time dependent potentials. Hence one has to apply numerical integration methods such as decomposition schemes, which both preserve the physical properties of the Hamiltonian system (6.4) (symplecticity, time reversibility) and are computationally efficient [42].

As we presented in the previous chapter the force-gradient schemes are based on the fact that the total propagator in (6.4) can be split in the following way:

$$\mathrm{e}^{(\hat{T} + \hat{V})h + \mathcal{O}(h^{K+1})} = \prod_{p=1}^{P} \mathrm{e}^{\hat{T} a_p h} \, \mathrm{e}^{\hat{V} b_p h + \mathcal{C} c_p h^3},$$

where $\mathcal{C} = [V, [T, V]]$ and $[\ , \ ]$ denotes the commutator of two operators and for the case of the Hamiltonian function (6.2) it yields

$$\mathcal{C} \equiv \big[V, [T, V] \big] = \sum_{i=1}^{N} \frac{g_i}{m_i} \cdot \frac{\partial}{\partial v},$$

where

$$g_{i\alpha} = 2 \sum_{j\beta} \frac{f_{j\beta}}{m_j} \frac{\partial f_{i\alpha}}{\partial q_{j\beta}},$$

α and β denote the Cartesian components of the vectors. The force-gradient evaluations

$\partial f_{i\alpha}/\partial q_{j\beta}$ can be explicitly represented taking into account that

$$f_{i\alpha} = m_i w_{i\alpha} - \frac{\partial u(q_i, t)}{\partial q_{i\alpha}},$$

where

$$w_{i\alpha} = -\frac{1}{m_i} \sum_{j(j \neq i)} \varphi'(q_{ij}) \frac{(q_{i\alpha} - q_{j\alpha})}{q_{ij}}$$

is the inter-particle part of the acceleration. The result is

$$g_i = -2 \sum_{j(j \neq i)}^{N} \left[(w_i - w_j)\frac{\varphi'_{ij}}{q_{ij}} + \frac{(q_i - q_j)}{q_{ij}^3} \left(q_{ij}\varphi''_{ij} - \varphi'_{ij} \right) \left(q_i \cdot (w_i - w_j) \right) \right] + h_i, \qquad (6.5)$$

where

$$h_{i\alpha} = \frac{2}{m_i} \sum_{\beta} \frac{\partial u}{\partial q_{i\beta}} \frac{\partial^2 u}{\partial q_{i\alpha}\partial q_{i\beta}}.$$

For this type of problem (6.1) we test well-known, commonly used second order integration schemes in molecular dynamics such as

- the leap-frog method, a 3-stage composition scheme:

$$\Delta(h) = e^{h\frac{\hat{V}}{2}} e^{h\hat{T}} e^{h\frac{\hat{V}}{2}}, \qquad (6.6)$$

- and a 5-stage extension widely used in QCD computations:

$$\Delta_5(h) = e^{\frac{1}{6}h\hat{V}} e^{\frac{1}{2}h\hat{T}} e^{\frac{2}{3}h\hat{V}} e^{\frac{1}{2}h\hat{T}} e^{\frac{1}{6}h\hat{V}}. \qquad (6.7)$$

We also consider the nested versions of the leap-frog method (6.6)

$$\hat{\Delta}(h) = e^{\frac{h}{2}\hat{V}_2} \Delta_1(h)_M e^{\frac{h}{2}\hat{V}_2}, \qquad (6.8)$$

where the inner cheaper system $H = T[P] + V_1[U]$ is solved by

$$\Delta_1(h)_M = \left(e^{\frac{h}{2M}\hat{V}_1} e^{\frac{h}{M}\hat{T}} e^{\frac{h}{2M}\hat{V}_1} \right)^M,$$

with M being a number of iterations for the fast part of the action; and similar 5-stage decomposition scheme (6.7), nested version of which has been recently introduced:

$$\hat{\Delta}_5(h) = e^{\frac{1}{6}h\hat{V}_2} \Delta_1\left(\frac{h}{2}\right)_M e^{\frac{2}{3}h\hat{V}_2} \Delta_1\left(\frac{h}{2}\right)_M e^{\frac{1}{6}h\hat{V}_2}. \qquad (6.9)$$

We test force gradient schemes (5.12) such as 5-stage force-gradient scheme proposed by Omelyan et al [42]

$$\Delta_{5C}(h) = e^{\frac{1}{6}h\hat{V}} e^{\frac{1}{2}h\hat{T}} e^{\frac{2}{3}h\hat{V} - \frac{1}{72}h^3\mathcal{C}} e^{\frac{1}{2}h\hat{T}} e^{\frac{1}{6}h\hat{V}} \qquad (6.10)$$

and the adapted nested force-gradient scheme introduced in [51]

$$\tilde{\Delta}_{5C}(h) = \mathrm{e}^{\frac{1}{6}h\hat{V}_2} \, \Delta_1\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{2}{3}h\hat{V}_2+\frac{1}{72}h^3\mathcal{C}_f} \, \Delta_1\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{1}{6}h\hat{V}_2}. \tag{6.11}$$

We also compare the extension given by the 11-stage decomposition [42], recently implemented as the integrator in the open source code openQCD as one of the standard options [36]

$$\Delta_{11}(h) = \mathrm{e}^{\sigma h\hat{V}} \, \mathrm{e}^{\eta h\hat{T}} \, \mathrm{e}^{\lambda h\hat{V}} \, \mathrm{e}^{\theta h\hat{T}} \, \mathrm{e}^{(1-2(\lambda+\sigma))\frac{h}{2}\hat{V}} \, \mathrm{e}^{(1-2(\theta+\eta))h\hat{T}} \, \mathrm{e}^{(1-2(\lambda+\sigma))\frac{h}{2}\hat{V}} \, \mathrm{e}^{\theta h\hat{T}} \, \mathrm{e}^{\lambda h\hat{V}} \, \mathrm{e}^{\eta h\hat{T}} \, \mathrm{e}^{\sigma h\hat{V}}, \tag{6.12}$$

where σ, θ, λ and η are parameters from equation (71) in Ref. [42].

The evolution operators $\mathrm{e}^{\hat{T}a_k h}$ and $\mathrm{e}^{\hat{V}b_k h+\mathcal{C}c_k h^3}$ displace v and q forward in time with

$$v \to v + b_k h\hat{V}(\rho) + c_k \mathcal{C}h^3 \qquad \text{and} \quad q \to q + a_k h\hat{T}(\rho). \tag{6.13}$$

We consider the three body problem, the *Sun-Earth-Moon problem* [28], which is a particular case of a *n-body problem* (6.2). The given system has the energy

$$E = \sum_{i=0}^{2} \frac{m_i v_i^2}{2} - G\sum_{i=1}^{2}\sum_{j=0}^{i-1} \frac{m_i m_j}{q_{ij}},$$

where $q_{ij} = \|q_i - q_j\|$, m_0, m_1 and m_2 represent the masses of the Sun, the Earth and the Moon, respectively and G is the gravitational constant. The physical parameters for this system are presented in Table 6.1.

The equations of motion (6.1) yield

$$\begin{aligned}
\frac{\mathrm{d}q_0}{\mathrm{d}t} &= v_0, & \frac{\mathrm{d}v_0}{\mathrm{d}t} &= -m_1 G\frac{q_0-q_1}{q_{01}^3} - m_2 G\frac{q_0-q_2}{q_{02}^3}, \\
\frac{\mathrm{d}q_1}{\mathrm{d}t} &= v_1, & \frac{\mathrm{d}v_1}{\mathrm{d}t} &= -m_0 G\frac{q_1-q_0}{q_{10}^3} - m_2 G\frac{q_1-q_2}{q_{12}^3}, \\
\frac{\mathrm{d}q_2}{\mathrm{d}t} &= v_2, & \frac{\mathrm{d}v_2}{\mathrm{d}t} &= -m_0 G\frac{q_2-q_0}{q_{20}^3} - m_1 G\frac{q_2-q_1}{q_{21}^3}.
\end{aligned} \tag{6.14}$$

Gravitational constant,	G	6.67384×10^{-11}, m^3/kg s	0.2662 AU^3/SU mo
Mass of the Sun,	m_0	1.9891×10^{30}, kg	1 SU
Mass of the Earth,	m_1	5.9736×10^{24}, kg	3×10^{-6} SU
Mass of the Moon,	m_2	7.3477×10^{22}, kg	0.0369×10^{-6} SU
Position of the Sun,	q_0	$(0,0)$, m	$(0,0)$, AU
Position of the Earth,	q_1	$(0, 1.52098 \times 10^{11})$, m	$(0, 1.0167138)$, AU
Position of the Moon,	q_2	$(0, 1.52504 \times 10^{11})$, m	$(0, 1.0191138)$, AU
Velocity of the Sun,	v_0	$(0,0)$, m/s	$(0,0)$, AU/mo
Velocity of the Earth,	v_1	$(0, 29.78 \times 10^3)$, m/s	$(0, 0.5160)$, AU/mo
Velocity of the Moon,	v_2	$(0, 30.802 \times 10^3)$, m/s	$(0, 0.5337)$, AU/mo

Table 6.1: Physical parameters of the Sun-Earth-Moon problem.

The force-gradient terms can be obtained from (6.5) for this case, using the external field potential $u(q_{ij}) = 0$ and the pair-wise potentials

$$\varphi(q_{ij}) = -G\frac{m_i m_j}{q_{ij}},$$

respectively for each interaction, with the *fast* V_1 and the *slow* V_2 potentials are given

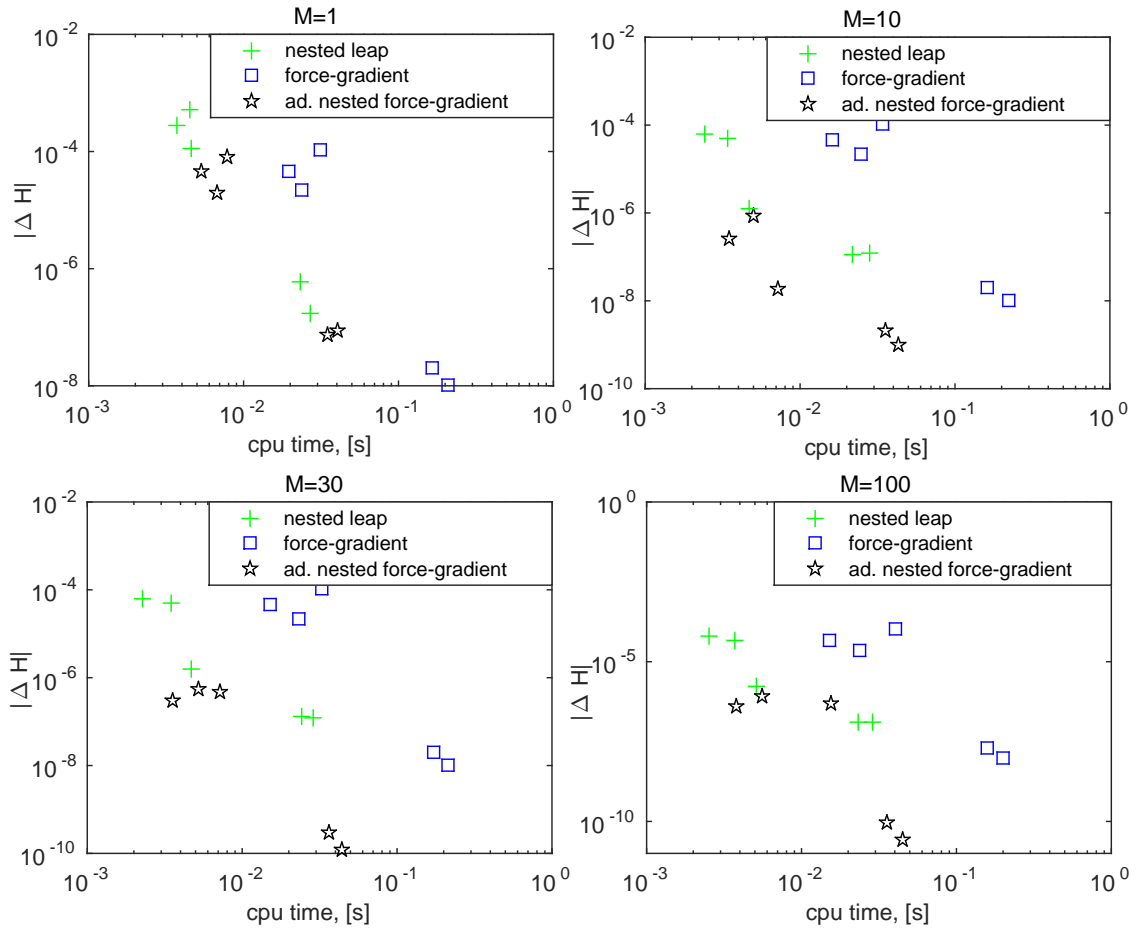$$V_1 = -G\frac{m_1 m_2}{q_{12}}, \qquad\qquad V_2 = -G\frac{m_0 m_1}{q_{01}} - G\frac{m_0 m_2}{q_{02}}.$$

Figure 6.1: Computational costs for different numerical methods(Sun-Moon-Earth problem).

Figure 6.1 presents the CPU time, required for the three different integrators against the achieved accuracy. Here we scale the time needed for the computation of the fast part by a factor of 0.001, since we assume that the computation of the fast scale functions is very cheap compared to the slow scale function evaluations. We can see that in general our nested force-gradient method (6.11) requires less CPU time and performs more accurate than the standard schemes, presented in Figure 6.1.
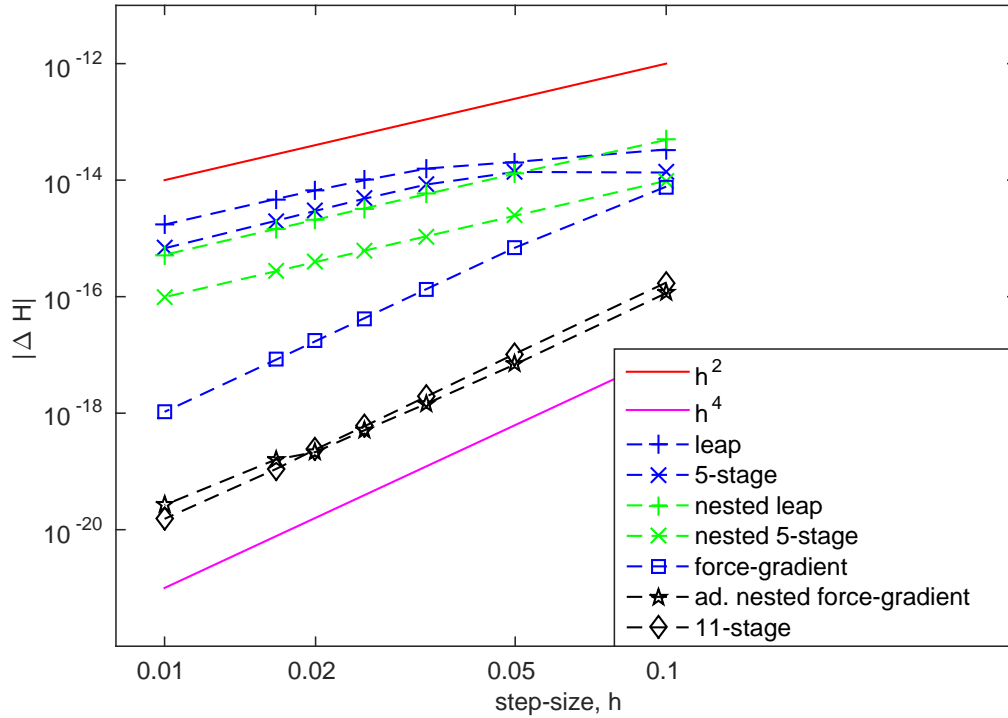
Figure 6.2: Absolute error for different numerical integrators (Sun-Moon-Earth problem).

Figure 6.2 presents a comparison between the standard numerical algorithms (6.6)-(6.7), nested approaches (6.8)–(6.9), the force-gradient (6.10) and our combined method (6.11), it shows the absolute error versus the step-size. The proposed integrator (6.11) with $M = 30$, which combines nested and force-gradient ideas, yields a better energy conservation even compared to 11-stage numerical scheme (6.12). These numerical results correspond to our analytical observations.

Thus we can argue that, if the evaluation of the *fast* function is significantly cheaper than the *slow* function, computational costs decrease. This is exactly the case found in our long-term goal applications in lattice quantum chromodynamics (LQCD), where the action can be split into two parts: the gauge action (whose force evaluations are cheap) and the fermion action (expensive). We have introduced a new decomposition scheme for Hamiltonian systems, which combines the idea of the force-gradient time-reversible and symplectic integrators and the splitting approach of nested algorithms. The new method of (5.36) is fourth-order accurate. Compared to other fourth-order schemes, the leading error coefficient is smaller and computational costs are lower.

6.2 Schwinger model

Previously, we studied the behavior of the adapted nested force-gradient scheme in the Abelian case on the example of the n-body problem [51]. Numerical tests have verified the potential of these schemes with increasing multirate potential, nested versions of force-gradient schemes outperform the unnested versions. This promising result suggested to perform the step from the Abelian to the non-Abelian setting of lattice QCD.

Due to the complexity of the lattice QCD problem requiring the huge computational effort

it was natural to concentrate on the intermediate step from Abelian to the non-Abelian case. We chose a well known problem of quantum electrodynamics (QED), the two-dimensional Schwinger model, since it suits the best as a test case for new concepts and ideas for more computational demanding problems [13]. As a lattice quantum field theory, it has many of the properties of more sophisticated models such as QCD, for example the numerical cost is still dominated by the fermion part of the action. The fact that this model, with far fewer degrees of freedom, does not require such large computational effort makes it the perfect choice for our testing purposes.

We apply the numerical integration schemes (6.6)–(6.11) to the two-dimensional Schwinger model, where the discrete updates (6.13) are given in the following form

$$
\begin{aligned}
\mathrm{e}^{\hat{T}h} : \quad U &\rightarrow U' = \exp(iPh)U \\
\mathrm{e}^{\hat{V}h} : \quad P &\rightarrow P' = P - ihF_V(U),
\end{aligned}
$$

where the gauge fields U are elements of some Lie group, and the momenta P, elements of the corresponding Lie algebra.

As well as the modification of the force-gradient method (6.10) proposed in [60], where the force-gradient term $\mathcal{C}$ is approximated via a Taylor expansion and a nested version of (6.10), which has been used in [3], which reads

$$
\hat{\Delta}_{5C}(h) = \mathrm{e}^{\frac{1}{6}h\hat{V}_2} \, \Delta_2\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{2}{3}h\hat{V}_2 + \frac{1}{72}h^3\mathcal{C}_f} \, \Delta_2\left(\frac{h}{2}\right)_M \mathrm{e}^{\frac{1}{6}h\hat{V}_2}, \tag{6.15}
$$

where

$$
\Delta_2(h)_M = \left(\mathrm{e}^{\frac{1}{6M}h\hat{V}_1} \, \mathrm{e}^{\frac{1}{2M}h\hat{T}} \, \mathrm{e}^{\frac{2}{3M}h\hat{V}_1 + \frac{1}{72}\left(\frac{h}{M}\right)^3 \mathcal{C}_g} \, \mathrm{e}^{\frac{1}{2M}h\hat{T}} \, \mathrm{e}^{\frac{1}{6M}h\hat{V}_1} \right)^M,
$$

with $\mathcal{C}_g = \{V_1, \{T, V_1\}\}$ and $\mathcal{C}_f = \{V_2, \{T, V_2\}\}$.

The most challenging task from the theoretical point of view is to derive the force-gradient term $\mathcal{C}$. First we introduce the Schwinger model and explain how to obtain the force-gradient term. The two-dimensional Schwinger model is defined by the following Hamiltonian function

$$
H = \frac{1}{2} \sum_{n=1,\mu=1}^{V,2} p_{n,\mu}^2 + S_{full}[U] = \frac{1}{2} \sum_{n=1,\mu=1}^{V,2} p_{n,\mu}^2 + S_G[U] + S_F[U]. \tag{6.16}
$$

with $V = L \times T$ the volume of the lattice. Unlike QCD, where $U \in SU(3)$ and $p_{n,\mu} \in su(3)$, for this QED problem (6.16), the links U are the elements of the Lie group $U(1)$ and the momenta $p_{n,\mu}$ belong to $\mathbb{R}$, which represents the Lie algebra of the group $U(1)$. This makes this test example (6.16) very cheap in terms of the computational time. This together with the fact that the Schwinger model also shares many of the features of QCD simulations, makes the Schwinger model an excellent test example when considering numerical integrators: a fast dynamics given by the computationally cheap gauge part $S_G[U]$ of the action demanding small step sizes, and a slow dynamics given by the computationally expensive fermion part $S_F[U]$ allowing large step sizes.

The pure gauge part of the action S_G sums up over all plaquettes $\mathbb{P}(n)$ in the two-

dimensional lattice with

$$\mathbb{P}(n) = U_1(n)U_2(n+\hat{1})U_1^\dagger(n+\hat{2})U_2^\dagger(n),$$

and is given by

$$S_G = \beta \sum_{n=1}^{V} \left(1 - \operatorname{Re} \mathbb{P}(n)\right). \tag{6.17}$$

The links U can be written in the form $U_\mu(n) = e^{iq_\mu(n)} \in \mathrm{U}(1)$ and connect the sites n and $n + \hat{\mu}$ on the lattice; $q_\mu(n) \in [-\pi, \pi]$, $\mu, \nu \in \{x, t\}$ are respectively space and time directions and β is a coupling constant. Note that from now on we will set the lattice spacing $a = 1$.

The fermion part of the action S_F is given by

$$S_F = \eta^\dagger \left(D^\dagger D\right)^{-1} \eta, \tag{6.18}$$

where η is a complex pseudofermion field. Here, D denotes the Wilson–Dirac operator given by

$$D_{n,m} = (2 + m_0)\delta_{n,m} - \frac{1}{2} \sum_{\mu=1}^{2} \left((1 - \sigma_\mu)U_\mu(n)\delta_{n,m-\hat{\mu}} + (1 + \sigma_\mu)U_\mu^\dagger(n - \hat{\mu})\delta_{n,m+\hat{\mu}}\right),$$

where σ_μ are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

m_0 is the mass parameter and the Kronecker delta $\delta_{n,m}$ acts on the pseudofermion field by

$$\sum_{m=1}^{V} \delta_{n,m}\eta(m) = \eta(n)$$

with $\eta(n)$ the pseudofermion field, a vector in the two-dimensional spinor space taking values at each lattice point n.

In order to proceed with the numerical integration we need to obtain the force F and the force gradient term $\mathcal{C}$. The force term $F(n, \mu)$ with respect to the link $U_\mu(n)$ is given by the first derivative of the action S_{full} and can be written as

$$F(n, \mu) = F_{S_G}(n, \mu) + F_{S_F}(n, \mu) = \frac{\partial S_G}{\partial q_\mu(n)} + \frac{\partial S_F}{\partial q_\mu(n)}. \tag{6.19}$$

Since the numerical schemes (6.9)–(6.11) use the multi-rate approach, the shifts in the momenta updates are split on F_{S_G} and F_{S_F} and we can consider them separately. The force terms F_{S_G} and F_{S_F} are obtained by differentiation over $U(1)$ group elements, which for the Schwinger model is the standard differentiation.

The force associated with link $U_\mu(n)$ from the gauge action is given by

$$\beta g(n, \mu) := F_{S_G}(n, \mu) = \beta \operatorname{Im} \left(\mathbb{P}(n) - \mathbb{P}(n - \hat{\nu})\right)\big|_{\mu \neq \nu}. \tag{6.20}$$

The force term of the fermion part is given by

$$f(n,\mu) := F_{S_F} = -\operatorname{Im}\left[\chi^\dagger(n)(1-\sigma_\mu)U_\mu(n)\xi(n+\hat\mu) - \chi^\dagger(n+\hat\mu)(1+\sigma_\mu)U_\mu^\dagger(n)\xi(n)\right] \qquad (6.21)$$

where vectors χ and ξ are given

$$\chi = D^{\dagger^{-1}}\eta, \qquad\qquad\qquad \xi = D^{-1}D^{\dagger^{-1}}\eta. \qquad (6.22)$$

For the numerical methods (6.10) and (6.15) we need to find the force gradient term $\mathcal{C}(n,\mu)$ with respect to the link $U_\mu(n)$. In case of the Schwinger model (6.16) this term reads

$$\mathcal{C}(n,\mu) = 2\sum_{m=1,\nu=1}^{V,2} \frac{\partial S_{full}}{\partial q_\nu(m)} \frac{\partial^2 S_{full}}{\partial q_\nu(m)\partial q_\mu(n)}. \qquad (6.23)$$

For simplicity we decompose the force gradient term (6.23) in four parts

$$\mathcal{C}_{GG} = 2\sum_{m=1,\nu=1}^{V,2} \frac{\partial S_G}{\partial q_\nu(m)} \frac{\partial^2 S_G}{\partial q_\nu(m)\partial q_\mu(n)}, \quad \mathcal{C}_{FG} = 2\sum_{m=1,\nu=1}^{V,2} \frac{\partial S_F}{\partial q_\nu(m)} \frac{\partial^2 S_G}{\partial q_\nu(m)\partial q_\mu(n)},$$

$$\mathcal{C}_{GF} = 2\sum_{m=1,\nu=1}^{V,2} \frac{\partial S_G}{\partial q_\nu(m)} \frac{\partial^2 S_F}{\partial q_\nu(m)\partial q_\mu(n)}, \quad \mathcal{C}_{FF} = 2\sum_{m=1,\nu=1}^{V,2} \frac{\partial S_F}{\partial q_\nu(m)} \frac{\partial^2 S_F}{\partial q_\nu(m)\partial q_\mu(n)}. \qquad (6.24)$$

This decomposition is also useful since the numerical integrator (6.15) only uses the term $\mathcal{C}_{FF}$ by construction. As shown in [51], to obtain the fourth order convergent scheme (6.15) from the second order convergent method (6.9) we must eliminate the leading error term, which is exactly represented by $\mathcal{C}_{FF}$. For completeness we discuss all 4 parts below.

The $\mathcal{C}_{GG}$ part of the force-gradient term is

$$\mathcal{C}_{GG} = 2\beta^2 \left[\operatorname{Im}(4\mathbb{P}_1(n,\mu) - \mathbb{P}_2(n,\mu) - \mathbb{P}_3(n,\mu) - \mathbb{P}_4(n,\mu) - \mathbb{P}_5(n,\mu))\operatorname{Re}(\mathbb{P}_1(n,\mu)) \right.$$
$$\left. - \operatorname{Im}(4\mathbb{P}_2(n,\mu) - \mathbb{P}_1(n,\mu) - \mathbb{P}_6(n,\mu) - \mathbb{P}_7(n,\mu) - \mathbb{P}_8(n,\mu))\operatorname{Re}(\mathbb{P}_2(n,\mu))\right]$$

with the set of plaquettes

$$\mathbb{P}_1(n,\mu) = U_\mu(n)U_\nu(n+\hat\mu)U_\mu^\dagger(n+\hat\nu)U_\nu^\dagger(n),$$
$$\mathbb{P}_2(n,\mu) = U_\mu(n-\hat\nu)U_\nu(n-\hat\nu+\hat\mu)U_\mu^\dagger(n)U_\nu^\dagger(n-\hat\nu),$$
$$\mathbb{P}_3(n,\mu) = U_\mu(n+\hat\mu)U_\nu(n+2\hat\mu)U_\mu^\dagger(n+\hat\nu+\hat\mu)U_\nu^\dagger(n+\hat\mu),$$
$$\mathbb{P}_4(n,\mu) = U_\mu(n+\hat\nu)U_\nu(n+\hat\mu+\hat\nu)U_\mu^\dagger(n+2\hat\nu)U_\nu^\dagger(n+\hat\nu),$$
$$\mathbb{P}_5(n,\mu) = U_\mu(n-\hat\mu)U_\nu(n)U_\mu^\dagger(n+\hat\nu-\hat\mu)U_\nu^\dagger(n-\hat\mu),$$
$$\mathbb{P}_6(n,\mu) = U_\mu(n-\hat\mu-\hat\nu)U_\nu(n-\hat\nu)U_\mu^\dagger(n-\hat\mu)U_\nu^\dagger(n-\hat\mu-\hat\nu),$$
$$\mathbb{P}_7(n,\mu) = U_\mu(n-2\hat\nu)U_\nu(n-2\hat\nu+\hat\mu)U_\mu^\dagger(n-\hat\nu)U_\nu^\dagger(n-2\hat\nu),$$
$$\mathbb{P}_8(n,\mu) = U_\mu(n-\hat\nu+\hat\mu)U_\nu(n-\hat\nu+2\hat\mu)U_\mu^\dagger(n+\hat\mu)U_\nu^\dagger(n-\hat\nu+\hat\mu).$$

Then by using the vectors $f(n,\mu)$ defined in (6.21) we obtain the $\mathcal{C}_{FG}$ piece of the force-gradient term given by

$$\mathcal{C}_{FG}(n,\mu) = 2\beta\left[(f(n,\mu) + f(n+\hat\mu,\nu) - f(n+\hat\nu,\mu) - f(n,\nu))\operatorname{Re}(\mathbb{P}_1) \right.$$
$$\left. + (f(n,\mu) - f(n+\hat\mu-\hat\nu,\nu) - f(n-\hat\nu,\mu) + f(n-\hat\nu,\nu))\operatorname{Re}(\mathbb{P}_2)\right].$$

The second derivative of the fermion action is

$$\frac{\partial^2 S_F}{\partial q_\nu(m)\partial q_\mu(n)} =$$

$$2\operatorname{Re}\chi^\dagger\left[\frac{\partial D}{\partial q_\nu(m)}D^{-1}\frac{\partial D}{\partial q_\mu(n)} + \frac{\partial D}{\partial q_\mu(n)}D^{-1}\frac{\partial D}{\partial q_\nu(m)} - \frac{\partial^2 D}{\partial q_\nu(m)\partial q_\mu(n)}\right]\xi+$$

$$2\operatorname{Re}\chi^\dagger\frac{\partial D}{\partial q_\mu(n)}(D^\dagger D)^{-1}\frac{\partial D^\dagger}{\partial q_\nu(m)}\chi\,,$$

$$= 2\operatorname{Re}\left[z_{1,m,\nu}^\dagger\frac{\partial D}{\partial q_\mu(n)}\xi + \chi^\dagger\frac{\partial D}{\partial q_\mu(n)}D^{-1}w_{2,m,\nu} - \chi^\dagger\frac{\partial^2 D}{\partial q_\nu(m)\partial q_\mu(n)}\xi + \chi^\dagger\frac{\partial D}{\partial q_\mu(n)}D^{-1}z_{1,m,\nu}\right]$$

$$= 2\operatorname{Re}\left[z_{1,m,\nu}^\dagger w_{2,n,\mu} + w_{1,n,\mu}^\dagger z_{2,m,\nu} - \chi^\dagger\frac{\partial^2 D}{\partial q_\nu(m)\partial q_\mu(n)}\xi\right]$$

in terms of the vectors χ and ξ defined in (6.22). Now the fields $z_{1,m,\nu}$ and $z_{2,m,\nu}$ are given by

$$z_{1,m,\nu} := D^{\dagger^{-1}}\frac{\partial D^\dagger}{\partial q_\nu(m)}\chi = D^{\dagger^{-1}}w_{1,m,\nu}$$

$$z_{2,m,\nu} := D^{-1}(\frac{\partial D}{\partial q_\nu(m)}\xi + z_{1,m,\nu}) = D^{-1}(w_{2,m,\nu} + z_{1,m,\nu})$$

with

$$w_{1,m,\nu}(n) := \sum_{n'}\frac{\partial D_{n,n'}^\dagger}{\partial q_\nu(m)}\chi(n') = \delta_{n,m+\hat\nu}\frac{i}{2}(1-\sigma_\nu)U_\nu^\dagger(m)\chi(m) - \delta_{n,m}\frac{i}{2}(1+\sigma_\nu)U_\nu(m)\chi(m+\hat\nu)\,,$$

$$w_{2,m,\nu}(n) := \sum_{n'}\frac{\partial D_{n,n'}}{\partial q_\nu(m)}\xi(n') = -\delta_{n,m}\frac{i}{2}(1-\sigma_\nu)U_\nu(m)\xi(m+\hat\nu) + \delta_{n,m+\hat\nu}\frac{i}{2}(1+\sigma_\nu)U_\nu^\dagger(m)\xi(m)\,.$$

In order to calculate C_{GF} and C_{FF} it is possible to perform the summation of $\sum_{m,\nu}$ before the inversions of D and $D^\dagger$ to get z_1 and z_2 which save $\mathcal{O}(V)$ additional inversions for the force gradient terms. It follows for the force gradient term C_{FF}

$$\mathcal{C}_{FF}(n,\mu) = 4\operatorname{Re}\left[Z_1^\dagger w_{2,n,\mu} + w_{1,n,\mu}^\dagger Z_2 - \chi^\dagger\frac{\partial^2 D}{\partial q_\mu(n)\partial q_\mu(n)}\xi\cdot f(n,\mu)\right] \tag{6.25}$$

with

$$Z_1 := D^{\dagger^{-1}}\sum_{m=1,\nu=1}^{V,2}(w_{1,m,\nu}\cdot f(m,\nu))\,,$$

$$Z_2 := D^{-1}\left(\sum_{m=1,\nu=1}^{V,2}[w_{2,m,\nu}\cdot f(m,\nu)] + Z_1\right)\,. \tag{6.26}$$

The expression for C_{GF} can be obtained from the one for C_{FF} by replacing in (6.25) and (6.26) the vector f with βg defined in (6.20).

It is important to mention that the computationally most demanding part of the numerical integration of the Schwinger model and quantum field theory in general is the inverse of the Dirac operator D^{-1}. Every momenta update, which includes fermion action (6.21) requires 2 inversions of the Dirac operator, the addition of the force-gradient term $\mathcal{C}$ requires 4 more inversions. Therefore leap-frog based methods (6.6) and (6.8) need 4 computations of D^{-1} per time step; schemes (6.7) and (6.9) 6 times; force-gradient based methods 8 for (6.15) and (6.11), 10 for (6.10) and the 11 stage method (6.12) has 12 inversions of the Dirac

operator. Since we use the multi-rate approach for schemes (6.9), (6.15) and (6.11), which leads generally to fewer macro time steps needed than for the standard schemes, we expect the integrator (6.11) will be the most efficient choice among the methods considered. In the next section we present numerical tests of this prediction.

We consider a 32 by 32 lattice with a coupling constant $\beta = 1.0$, scaling variable $z = 0.2$ and mass $m_0 = -0.231367$, simulate it on the trajectory $\tau = 2.0$ with 200 configurations. The parameters were taken from [13] for the better physical representation and also in order to increase the impact of the fermion part of the action.
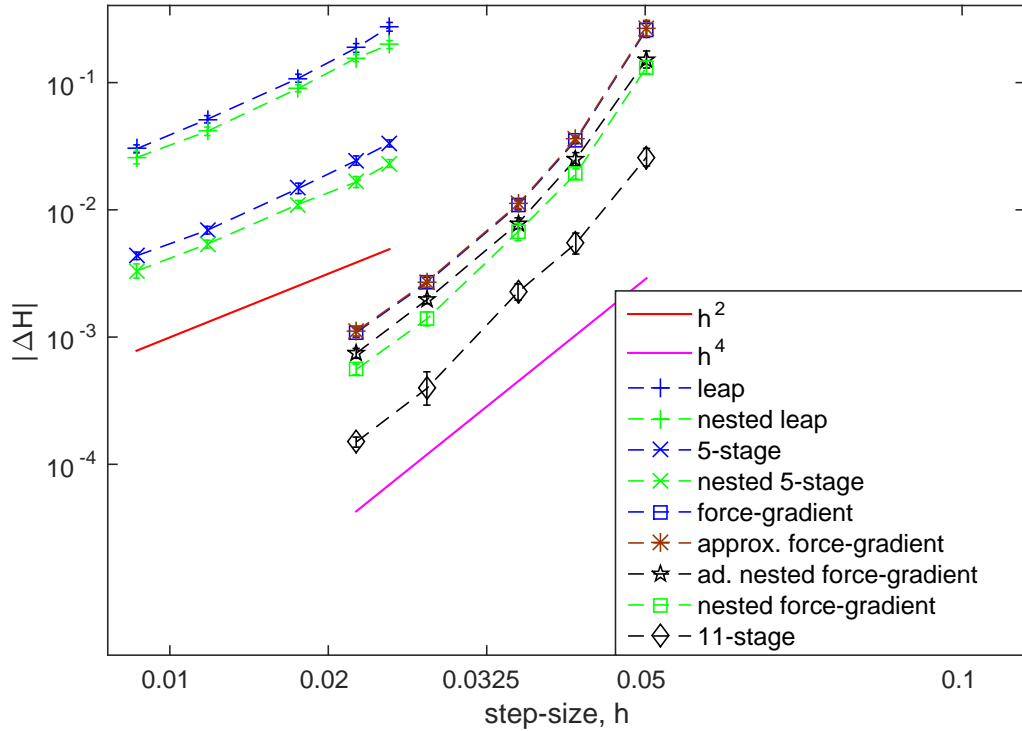

Figure 6.3: Absolute error for different numerical integrators (Schwinger model).

Figure 6.3 presents the comparison between the numerical integrators (6.6) – (6.15). It shows the absolute error versus the step-size. Here the multi-rate schemes (6.8), (6.9) and (6.15) are using four times smaller step-size for the gauge action (6.17). The numerical tests correspond to our analytical observations with the nested integrators perform better than their standard versions.

Figure 6.4 presents the CPU time, required for the proposed integrators (6.7)-(6.11), versus the achieved accuracy. We can observe that the nested force-gradient method (6.15) and adapted nested force-gradient method (6.11) show much better results in terms of a computational efficiency than the integrators (6.9) and (6.10); and even compared to the 11 stage scheme (6.12). Here we can see that the modification of (6.10) proposed in [60] also performs better than its original version. It shows almost similar computational costs as nested versions of the force-gradient approach (6.15)-(6.11).
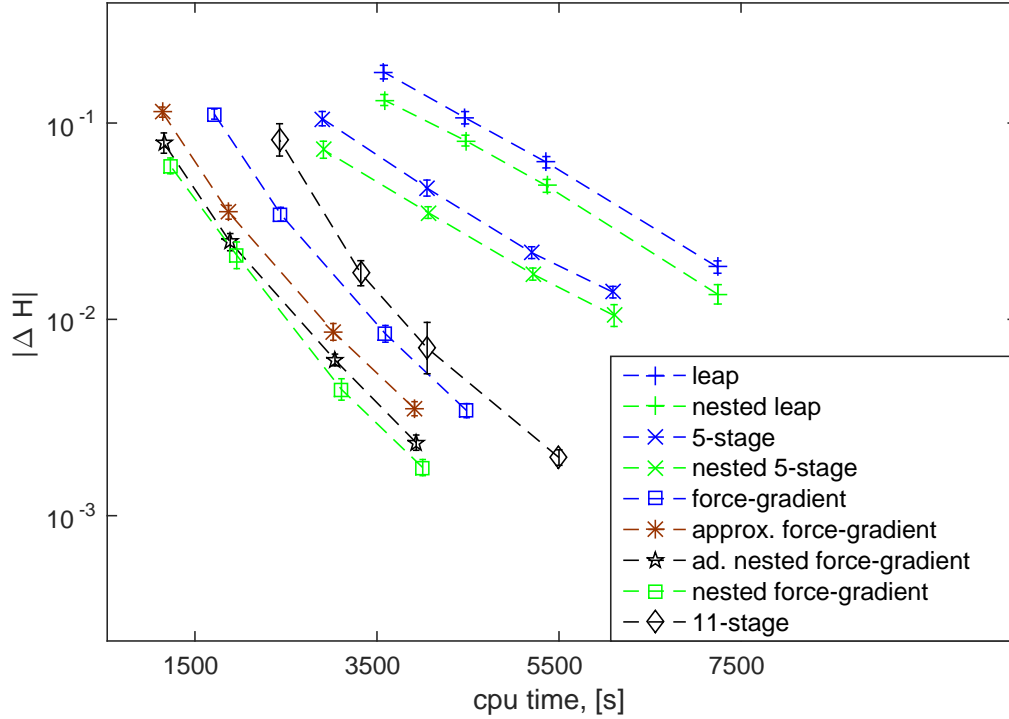
Figure 6.4: Computational costs for different numerical methods (Schwinger model).

Table 6.2 shows the number of inversions of the Dirac operator D, which is needed to reach 90% acceptance rate of the HMC. Since D^{-1} is the most computationally demanding part it is important to see how many of these inversions are required per each trajectory. From Table 6.2 it easy to see that the adapted nested force-gradient method (6.11) and nested force-gradient method (6.15) need the least number of D^{-1} evaluations per trajectory to reach the chosen acceptance rate $\approx 90\%$.

Integrator:	step size h	M	D^{-1} per step	D^{-1} per trajectory
5 stage method	0.0294	-	6	420
nested 5 stage method	0.0286	700	6	408
5 stage force-gradient	0.0550	-	10	370
approx.force-gradient [60]	0.0540	-	8	290
nested force-gradient	0.0560	450	8	285
adapted nested force-gradient	0.0560	450	8	285
11 stage method	0.0625	-	12	384

Table 6.2: Step-sizes and number of inversions of D per step and per trajectory for acceptance rate of 90%.

In this chapter we have presented the nested force-gradient approach (6.15) and its adapted version (6.11) applied to a model problem in quantum field theory, the two-dimensional Schwinger model. The derivation of the force-gradient terms was given and the Schwinger model was introduced. Nested force-gradient schemes seem to be an optimal choice with relatively high convergence order and low computational effort.

It also would be possible to improve the algorithm (6.11) by measuring the Poisson brackets of the shadow Hamiltonian of the proposed integrator and then tuning the set of optimal parameters, e. g. micro and macro step sizes.

We can also claim that methods (6.15) and (6.11) have a potential to perform even better with respect to the computational effort in the case of lattice QCD problems, since the impact of the fermion action (6.18) and the computational time to obtain the inversion of the Dirac operator D is much more significant.

In the last part of this thesis we summarize the numerical results obtained in this Chapter, draw the conclusion based on these results and give a short outlook on future research directions.

7
Chapter 7
SUMMARY AND OUTLOOK

In this thesis the general idea of geometric numerical time integration for ordinary differential equations was introduced. We presented this idea for the case of Hamiltonian mechanics as a basis of our main area of application for this work. We showed in detail the preservation of physical properties of a system such as time-reversibility, symplecticity and energy conservation for the numerical methods. We provided a brief introduction to the backward error analysis, which included the concept of shadow Hamiltonians. The main classes of geometric numerical time integrators were given as examples to treat the numerical problems on Abelian and non-Abelian manifolds.

Next we presented the main application of thesis quantum field theories where we gave a short summary of a quantization process. The basic concepts of the path integral formulation for quantum field theories were introduced as well as its regularization on the lattice. Two different kinds of field variables were presented in order to show a few important examples of quantum field theories such as QED and QCD. The role of geometric numerical time integrators were explained as the main way to treat the physical problems in path integral formulation is the HMC algorithms and numerical methods to be used in the molecular dynamics step of HMC have to conserve certain properties of quantum field theories.

The first attempt to improve the efficiency of HMC simulations were made by proposing the idea to use projection methods in order to eliminate the accept/reject step of the HMC algorithm, since this class of numerical methods conserves energy of a physical system exactly. We introduced both standard and symmetric variants of projection methods and presented an alternative view on symmetric projection schemes as well as a structure preserving approach for projection methods. We analyzed both analytically and numerically the conservation of geometric properties of the flow. Unfortunately we had to conclude that even so we were able to reach certain success in developing symplectic projection schemes for our main application it would require a huge computational effort and hence to use proposed methods is not advantageous.

We introduced another option to increase the efficiency of numerical time integrators for the HMC algorithm. We combined two operator splitting approaches, namely, force-gradient decomposition methods and multirate stepping strategy in order to gain a higher order of accuracy by adding the force-gradient term and reduce the computational cost by treating the slow and the fast part of the action of a system with not only different step size (smaller for the fast part and bigger for the slow one) but also applying different numerical schemes to each of these parts accordingly. The introduction into the force-gradient decomposition approach and the multirate time stepping methods was presented together with the analysis for the structure-preserving properties of these numerical scheme. The resulting adapted nested force-gradient approach satisfies the main goal of this thesis. We tested these methods first analytically to show that they satisfy the necessary structure-preserving properties necessary to be used in the HMC simulations and then numerically

to prove that they are more efficient than the other numerical integrators which are used for solving path integrals of quantum field theories.

The numerical tests were designed, first, to investigate the convergence order of the proposed adapted nested force-gradient decomposition schemes; then to derive the force-gradient term by applying them to the rather simple case of the n-body problem (Sun-Moon-Earth problem); and at last to estimate the efficiency of the proposed scheme via scaling the computational costs for the fast part of the action by small factor. The positive outcome of the first numerical test allowed us to carry on and we considered the 2-dimensional Schwinger model of quantum electrodynamics to showcase the advantages of the new integrator. The derivation of the force-gradient term was presented since it is the most challenging part of the implementation process and also to show the difference with the standard force-gradient and nested force-gradient schemes. Numerical results shows that we can claim that the proposed adapted nested force-gradient methods seem to be an optimal choice with relatively high convergence order and low computational effort.

The numerical results obtained by testing adapted force-gradient numerical integrators on the Schwinger model were very promising and showed that the adapted force-gradient integrators have a large potential for the ultimate goal of Lattice QCD applications, where two efficiency factors will act in favor of nested schemes: the multirate factor of lattice QCD is much higher than in the Schwinger model case; the number of inversions of the Dirac-Wilson operator is much lower for five stage than for eleven stage schemes.

The future work can be summarized as following

- Apply the adapted nested force-gradient schemes to the Lattice QCD applications and generalize this approach to more than two multirate time levels recursively, allowing to exploit additional sources of multirate potential. Since for the adapted nested force-gradient scheme we use cheaper integrators for the fast part of the action comparing to the standard nested force-gradient approach.

- Based on a first test implementation of the adapted nested force gradient integrators, the range of feasibility of this new scheme has to be tested against a series of competitors for a range of possible applications. For lattice QCD, the choice of the best integrator should be based on a cost metric introduced by Clark et al. [14], which scales the CPU time needed for a trajectory with the acceptance rate. If the time integration error (measured as the deviation from the Hamiltonian) is log-normal distributed, the latter will be given by the complementary error function, evaluated at one eighth of the variance σ^2 of the time integration error [33]. The variance will depend not only on the step size h, but also on the volume V, the mass m and the gauge coupling β, expecting a dependence of the type

$$\sigma^2(h, V, m, \beta) = A(\beta) \frac{V^b h^c}{(m - m_{\text{crit}})^\alpha},$$

where the unknown parameters can be detected by numerical experiments. For every setting, an optimal step size can be derived according to the cost metric above.

- To make the best use of the adapted nested force gradient integrators in Lattice QCD, one must implement the scheme, together with a proposal for an optimal choice of the step size according to the cost metric above, in the openQCD package

(see http://luscher.web.cern.ch/luscher/openQCD/) of Martin Lüscher and Stefan Schaefer. This choice is based on our previous experience in adding a software package for mass reweighing to the openQCD package.

We must also mention that there are possibilities to improve the proposed methods, by measuring the Poisson brackets of the shadow Hamiltonian of the adapted nested force-gradient scheme and then tuning the set of optimal parameters, e.g. micro and macro step sizes or/and to approximate the force-gradient term as it is proposed in [60].

Therefore we conclude that nested force-gradient methods have a huge potential to be used in Lattice QCD application, although it requires some additional work. Nevertheless we reached our main goal to develop a new geometric numerical time integrator for the molecular dynamics step of the HMC algorithm, which performs better with respect to accuracy and computational effort than most of numerical schemes generally used and still offers some space for improvements.

A Shadow Hamiltonian for the projection methods with $\mu = \mu(h)$

Let us consider the following scheme

$$\underbrace{e^{\frac{\mu}{2}\hat{V}}e^{\mu\hat{T}}e^{\frac{\mu}{2}\hat{V}}}_{\hat{H}_1}\underbrace{e^{\frac{h}{2}\hat{V}}e^{h\hat{T}}e^{\frac{h}{2}\hat{V}}}_{\hat{H}_2}\underbrace{e^{\frac{\mu}{2}\hat{V}}e^{\mu\hat{T}}e^{\frac{\mu}{2}\hat{V}}}_{\hat{H}_3}.$$

We know that it is the composition of three methods each of them preserves the corresponding shadow Hamiltonians

$$H_{1,3} = (T+V)\mu + \frac{\mu^3}{24}\left(2[T[T,V]] + [V[T,V]]\right)$$

$$H_2 = (T+V)h + \frac{h^3}{24}\left(2[T[T,V]] + [V[T,V]]\right).$$

Therefore, in order to find the shadow Hamiltonian for our method, we need to apply the BCH formula on

$$e^{\hat{H}_1}e^{\hat{H}_2}e^{\hat{H}_3}$$

First we consider the first part

$$e^{\hat{H}_1}e^{\hat{H}_2}$$

which yields

$$c_1 = H_1 + H_2,$$
$$c_2 = -\frac{1}{4}[H_1 + H_2, H_1 - H_2] = -\frac{1}{4}[(T+V)h + (T+V)\mu, (T+V)h - (T+V)\mu]$$
$$= -\frac{1}{4}\left((T+V)(T+V)\mu h + (T+V)(T+V)\mu^2 - (T+V)(T+V)h^2 - (T+V)(T+V)\mu h\right.$$
$$\left. -(T+V)(T+V)\mu h - (T+V)(T+V)\mu^2 + (T+V)(T+V)\mu h + (T+V)(T+V)h^2\right)$$
$$c_3 = -\frac{1}{6}[0, H_1 - H_2] = 0.$$

Here we skip the interaction of h^3 terms since we do not consider the error terms higher than h^3 and obtain

$$e^{\hat{H}_1}e^{\hat{H}_2} = e^{\hat{H}_1+\hat{H}_2}.$$

Then apply it for the second part

$$e^{\hat{H}_1+\hat{H}_2}e^{\hat{H}_3},$$

$$c_1 = H_1 + H_2 + H_3,$$

$$c_2 = -\frac{1}{4}[H_1 + H_2 + H_3, H_1 + H_2 - H_3] = -\frac{1}{4}[(T+V)2\mu + (T+V)h, (T+V)h]$$

$$= -\frac{1}{4}\left((T+V)(T+V)2\mu h + (T+V)(T+V)h^2 - (T+V)(T+V)2\mu h - (T+V)(T+V)h^2\right)$$

$$= 0$$

$$c_3 = -\frac{1}{6}[0, H_1 + H_2 - H_3] = 0.$$

and obtain

$$\tilde{H} = H_1 + H_2 + H_3$$

or

$$\tilde{H} = (2\mu + h)(T + V) + \frac{(2\mu^3 + h^3)}{24}\left(2[T[T,V]] + [V[T,V]]\right).$$

B Shadow Hamiltonian for the projection methods with $\mu_1 = \mu_1(h)$ and $\mu_2 = \mu_2(h)$

We consider the following integrator

$$\underbrace{e^{\frac{\mu_1}{2}\hat{V}}e^{\mu_1\hat{T}}e^{\frac{\mu_1}{2}\hat{V}}}_{\hat{H}_1}\underbrace{e^{\frac{\mu_2}{2}\hat{V}}e^{\mu_2\hat{T}}e^{\frac{\mu_2}{2}\hat{V}}}_{\hat{H}_2}\underbrace{e^{\frac{h}{2}\hat{V}}e^{h\hat{T}}e^{\frac{h}{2}\hat{V}}}_{\hat{H}_3}\underbrace{e^{\frac{\mu_2}{2}\hat{V}}e^{\mu_2\hat{T}}e^{\frac{\mu_2}{2}\hat{V}}}_{\hat{H}_4}\underbrace{e^{\frac{\mu_1}{2}\hat{V}}e^{\mu_1\hat{T}}e^{\frac{\mu_1}{2}\hat{V}}}_{\hat{H}_5}.$$

We know that it is the composition of three methods; each of them preserves the corresponding shadow Hamiltonians

$$H_{1,5} = (T+V)\mu_1 + \frac{\mu_1^3}{24}\left(2[T[T,V]] + [V[T,V]]\right)$$

$$H_{2,4} = (T+V)\mu_2 + \frac{\mu_2^3}{24}\left(2[T[T,V]] + [V[T,V]]\right)$$

$$H_3 = (T+V)h + \frac{h^3}{24}\left(2[T[T,V]] + [V[T,V]]\right).$$

Therefore, in order to find the general SH for our method, we need to apply the BCH formula on

$$e^{\hat{H}_1}e^{\hat{H}_2}e^{\hat{H}_3}e^{\hat{H}_4}e^{\hat{H}_5}.$$

Applying on the first part

$$e^{\hat{H}_1}e^{\hat{H}_2},$$

$c_1 = H_1 + H_2,$

$c_2 = -\frac{1}{4}[H_1 + H_2, H_1 - H_2] = -\frac{1}{4}[(T+V)\mu_1 + (T+V)\mu_2, (T+V)\mu_1 - (T+V)\mu_2]$

$\quad = -\frac{1}{4}\left((T+V)(T+V)\mu_2\mu_1 + (T+V)(T+V)\mu_1^2 - (T+V)(T+V)\mu_2^2 - (T+V)(T+V)\mu_1\mu_2\right.$

$\quad\quad \left. -(T+V)(T+V)\mu_1\mu_2 - (T+V)(T+V)\mu_1^2 + (T+V)(T+V)\mu_1\mu_2 + (T+V)(T+V)\mu_2^2\right)$

$c_3 = -\frac{1}{6}[0, H_1 - H_2] = 0$

Here we skip the interaction of $\mathcal{O}(\mu^3)$ terms since we do not consider the error terms higher than $\mathcal{O}(\mu^3)$ and obtain

$$e^{\hat{H}_1}e^{\hat{H}_2} = e^{\hat{H}_1 + \hat{H}_2}$$

Then we apply it for the second part

$$e^{\hat{H}_1 + \hat{H}_2}e^{\hat{H}_3}$$

$$c_1 = H_1 + H_2 + H_3,$$

$$c_2 = -\frac{1}{4}[H_1 + H_2 + H_3, H_1 + H_2 - H_3]$$

$$= -\frac{1}{4}[(T+V)\mu_1 + (T+V)\mu_2 + (T+V)h, (T+V)\mu_1 + (T+V)\mu_2 - (T+V)h]$$

$$= -\frac{1}{4}\left((T+V)(T+V)\mu_1^2 + (T+V)(T+V)\mu_1\mu_2 + (T+V)(T+V)\mu_1 h\right.$$

$$+ (T+V)(T+V)\mu_1\mu_2 + (T+V)(T+V)\mu_2^2 + (T+V)(T+V)\mu_2 h$$

$$- (T+V)(T+V)\mu_1 h - (T+V)(T+V)\mu_2 h - (T+V)(T+V)h^2$$

$$- (T+V)(T+V)\mu_1^2 - (T+V)(T+V)\mu_1\mu_2 + (T+V)(T+V)\mu_1 h$$

$$- (T+V)(T+V)\mu_1\mu_2 - (T+V)(T+V)\mu_2^2 + (T+V)(T+V)\mu_2 h$$

$$\left. (T+V)(T+V)\mu_1 h + (T+V)(T+V)\mu_2 h + (T+V)(T+V)h^2\right) = 0$$

$$c_3 = -\frac{1}{6}[0, H_1 + H_2 - H_3] = 0$$

and obtain

$$e^{\hat{H}_1}e^{\hat{H}_2}e^{\hat{H}_3} = e^{\hat{H}_1+\hat{H}_2+\hat{H}_3}$$

and then we apply BCH to

$$e^{\hat{H}_1+\hat{H}_2+\hat{H}_3}e^{\hat{H}_4}$$

which yields

$$c_1 = H_1 + H_2 + H_3 + H_4,$$

$$c_2 = -\frac{1}{4}[H_1 + H_2 + H_3 + H_4, H_1 + H_3]$$

$$= -\frac{1}{4}[(T+V)\mu_1 + 2(T+V)\mu_2 + (T+V)h, (T+V)\mu_1 + (T+V)h]$$

$$= -\frac{1}{4}\left((T+V)(T+V)\mu_1^2 + 2(T+V)(T+V)\mu_1\mu_2 + (T+V)(T+V)\mu_1 h\right.$$

$$+ (T+V)(T+V)\mu_1 h + 2(T+V)(T+V)\mu_2 h + (T+V)(T+V)h^2$$

$$- (T+V)(T+V)\mu_1^2 - 2(T+V)(T+V)\mu_1\mu_2 - (T+V)(T+V)\mu_1 h$$

$$\left. - (T+V)(T+V)\mu_1 h - 2(T+V)(T+V)\mu_2 h - (T+V)(T+V)h^2\right) = 0$$

$$c_3 = -\frac{1}{6}[0, H_1 + H_3] = 0$$

and therefore

$$e^{\hat{H}_1}e^{\hat{H}_2}e^{\hat{H}_3}e^{\hat{H}_4} = e^{\hat{H}_1+\hat{H}_2+\hat{H}_3+\hat{H}_4}.$$

At the end applying BCH formula on

$$e^{\hat{H}_1+\hat{H}_2+\hat{H}_3+\hat{H}_4}e^{\hat{H}_5}$$

$$c_1 = H_1 + H_2 + H_3 + H_4 + H_5,$$

$$c_2 = -\frac{1}{4}[H_1 + H_2 + H_3 + H_4 + H_5, H_2 + H_3 + H_4]$$

$$= -\frac{1}{4}[2(T+V)\mu_1 + 2(T+V)\mu_2 + (T+V)h, 2(T+V)\mu_2 + (T+V)h]$$

$$= -\frac{1}{4}\left(4(T+V)(T+V)\mu_2^2 + 4(T+V)(T+V)\mu_1\mu_2 + 2(T+V)(T+V)\mu_2 h\right.$$

$$+2(T+V)(T+V)\mu_1 h + 2(T+V)(T+V)\mu_2 h + (T+V)(T+V)h^2$$

$$-4(T+V)(T+V)\mu_2^2 - 4(T+V)(T+V)\mu_1\mu_2 - 2(T+V)(T+V)\mu_2 h$$

$$\left.-2(T+V)(T+V)\mu_1 h - 2(T+V)(T+V)\mu_2 h - (T+V)(T+V)h^2\right) = 0$$

$$c_3 = -\frac{1}{6}[0, H_2 + H_3 + H_4] = 0$$

and obtain

$$\tilde{H} = H_1 + H_2 + H_3 + H_4 + H_5$$

or

$$\tilde{H} = (2\mu_1 + 2\mu_2 + h)(T+V) + \frac{(2\mu_2^3 + 2\mu_2^3 + h^3)}{24}\left(2[T[T,V]] + [V[T,V]]\right).$$

C Shadow Hamiltonian for linear projection methods with $\mu = \mu(h)$

Let us write the method in the form of symplectic mapping

$$e^{\mathcal{M}^{-1}\frac{h}{2}\hat{V}}e^{h\hat{T}}e^{\frac{h}{2}\hat{V}}\mathcal{M}.$$

Applying BCH formula on the first part

$$e^{M^{-1}\frac{h}{2}\hat{V}}e^{h\hat{T}}$$

$$c_1 = \mathcal{M}^{-1}\frac{h}{2}V + hT$$

$$c_2 = -\frac{1}{4}[\mathcal{M}^{-1}\frac{h}{2}V + hT, \mathcal{M}^{-1}\frac{h}{2}V - hT] = -\frac{h^2}{4}[T, \mathcal{M}^{-1}V]$$

$$c_3 = -\frac{1}{6}[-\frac{h^2}{4}[T, \mathcal{M}^{-1}V], \mathcal{M}^{-1}\frac{h}{2}V - hT] = -\frac{h^3}{48}[\mathcal{M}^{-1}V, [T, \mathcal{M}^{-1}V]] + \frac{h^3}{24}[T, [T, \mathcal{M}^{-1}V]],$$

which yields

$$e^{\mathcal{M}^{-1}\frac{h}{2}\hat{V}}e^{h\hat{T}} = e^{\mathcal{M}^{-1}\frac{h}{2}\hat{V}+h\hat{T}-\frac{h^2}{4}[T,\mathcal{M}^{-1}V]-\frac{h^3}{48}[\mathcal{M}^{-1}V,[T,\mathcal{M}^{-1}V]]+\frac{h^3}{24}[T,[T,\mathcal{M}^{-1}V]]}$$

and then apply it for the second part

$$e^{\mathcal{M}^{-1}\frac{h}{2}\hat{V}+h\hat{T}\frac{h^2}{4}[T,\mathcal{M}^{-1}V]\cdots}e^{\frac{h}{2}\hat{V}}\mathcal{M}$$

$$c_1 = \mathcal{M}^{-1}\frac{h}{2}V + hT + \frac{h}{2}V\mathcal{M} - \frac{h^2}{4}[T, \mathcal{M}^{-1}V] - \frac{h^3}{48}[\mathcal{M}^{-1}V, [T, \mathcal{M}^{-1}V]] + \frac{h^3}{24}[T, [T, \mathcal{M}^{-1}V]]$$

$$c_2 = -\frac{1}{4}[\mathcal{M}^{-1}\frac{h}{2}V + hT + \frac{h}{2}V\mathcal{M} - \frac{h^2}{4}[T, \mathcal{M}^{-1}V] - \cdots, \mathcal{M}^{-1}\frac{h}{2}V + hT - \frac{h}{2}V\mathcal{M}$$

$$- \frac{h^2}{4}[T, \mathcal{M}^{-1}V] - \cdots] = \frac{h^2}{4}[T, V\mathcal{M}] + \frac{h^3}{16}[V\mathcal{M}, [T, \mathcal{M}^{-1}V]]$$

$$c_3 = -\frac{1}{6}[-\frac{h^2}{4}[T, V\mathcal{M}] + \frac{h^3}{16}[V\mathcal{M}, [T, \mathcal{M}^{-1}V]], \mathcal{M}^{-1}\frac{h}{2}V + hT - \frac{h}{2}V\mathcal{M} - \frac{h^2}{4}[T, \mathcal{M}^{-1}V] - \cdots]$$

$$= \frac{h^3}{48}[\mathcal{M}^{-1}V, [T, V\mathcal{M}]] - \frac{h^3}{48}[V\mathcal{M}, [T, V\mathcal{M}]] + \frac{h^3}{24}[T, [T, V\mathcal{M}]].$$

The result is following

$$
\begin{aligned}
\tilde{H} = {}& \mathcal{M}^{-1}\frac{h}{2}V + hT + \frac{h}{2}V\mathcal{M} - \frac{h^2}{4}[T, \mathcal{M}^{-1}V] + \frac{h^2}{4}[T, V\mathcal{M}] \\
& - \frac{h^3}{48}[\mathcal{M}^{-1}V, [T, \mathcal{M}^{-1}V]] + \frac{h^3}{24}[T, [T, \mathcal{M}^{-1}V]] + \frac{h^3}{16}[V\mathcal{M}, [T, \mathcal{M}^{-1}V]] \\
& + \frac{h^3}{48}[\mathcal{M}^{-1}V, [T, V\mathcal{M}]] - \frac{h^3}{48}[V\mathcal{M}, [T, V\mathcal{M}]] + \frac{h^3}{24}[T, [T, V\mathcal{M}]].
\end{aligned}
$$

Bibliography

[1] V. I. Arnold, Mathematical Methods of Classical Mechanics, Springer-Verlag New York 1989.

[2] U. Ascher, S. Reich, On some difficulties in integrating highly oscillatory Hamiltonian systems, in: Proceedings of Computational Molecular Dynamics, Springer Lecture Notes in Computational Science and Engineering, Vol. 4, 1998, pp. 281–296.

[3] A. Bazavov et al., Chiral transition and $U(1)_A$ symmetry restoration from lattice QCD using domain wall fermions, Phys. Rev. D 86 (201), 094503.

[4] S. Blanes, C. J. Budd, Adaptive Geometric Integrators for Hamiltonian Problems with Approximate Scale Invariance, J. Sci. Comput., Vol. 26, (2005), pp. 1089–1113.

[5] S. Blanes, F. Casas, Raising the order of geometric numerical integrators by composition and extrapolation, Numerical Algorithms, Springer, (2005) 38 pp 305–326.

[6] S. Blanes, F. Calas, A. Murua, Splitting and composition methods in the numerical integration of differential equations, Bol. Soc. Esp. Mat. Apl. 45(2008) pp. 89–145.

[7] G. Bogfjellmo, H. Marthinsen, High-Order Symplectic Partitioned Lie Group Methods, Foundations of Computational Mathematics 16.2 (2016) pp. 493–530.

[8] E. Borici, C. Joó, A. Frommer, Numerical methods in QCD, Springer, Berlin, 2002.

[9] C. J. Budd, A. Iserles, Geometric Integration: Numerical Solution of Differential Equations on Manifolds, Phil. Trans. R. Soc. Lond., Volume 357, 1999.

[10] C. J. Budd, M. D. Piggott, Geometric Integration and its Applications, Handbook of Numerical Analysis, vol. XI, Elsevier Science, 2003.

[11] J. C. Butcher, Numerical Methods for Ordinary Differential Equations, Wiley, 2003.

[12] E. Celledoni, H. Marthinsen, B. Owren, An introduction to Lie group integrators-basics, new developments and application, J. Comput. Phys., (2012).

[13] N. Christian, K. Jansen, K. Nagai, B. Pollakowski, Scaling test of fermion actions in the Schwinger model, Nuclear Physics B 739 (2006), pp. 60–84.

[14] M.A. Clark, B. Joó, A.D. Kennedy, P.J. Silva, Better HMC integrators for dynamical simulations, PoS 323(2010).

[15] P. E. Crouch, R. Grossman, Numerical integration of ordinary differential equations on manifolds, J. Nonlinear Sci., 3 (1993), pp. 1–33.

[16] S. Duane, A.D. Kennedy, B.J. Pendleton, D. Roweth, Hybrid Monte Carlo, Phys.

Lett. B195 (1987), pp. 216–222.

[17] E. Forest, R. D. Ruth, Fourth-order symplectic integration, Physica D 43 105 (1990).

[18] J. Gans, D. Shalloway, Shadow mass and the relationship between velocity and momentum in symplectic numerical integration, Phys. Rev. E (3), 61 (2000), pp. 4587–4592.

[19] C. Gattringer, C. B. Lang, Quantum Chromodynamics on the Lattice: An Introductory Presentation, Lect. Notes Phys. 788, Springer, Berlin Heidelberg, 2010, DOI 10.1007/978-3-642-01850-3.

[20] C. W. Gear, D. R. Wells, Multirate linear multistep methods, BIT 24(1984) pp 484–502.

[21] W. R. Gilks, S. Richardson, D. J. Spiegelhalter, Markov Chain Monte Carlo in Practice, Chapman & Hall (1995), pp. 1-19.

[22] W. Gröbner, Die Liereihen und ihre Anwendungen, VEB Deutscher Verlag der Wissenschaften, Berlin 1960.

[23] R. Gupta, Introduction to Lattice QCD, arXiv:hep-lat/9807028(1998).

[24] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems, 2nd ed., Springer Series in Comput. Math., Vol. 14, Springer-Verlag, Berlin, 1996.

[25] E. Hairer, Symmetric Projection Methods for Differential Equations on Manifolds, BIT Numerical Mathematics Vol. 40 (2000), pp. 726–734.

[26] E. Hairer, Geometric Integration of Ordinary Differential Equations on Manifolds, BIT (2001), Vol. 41, No. 5, pp. 996–1007.

[27] E. Hairer, Global modified Hamiltonian for constrained symplectic integrators, Numerical Mathematics Vol. 95 (2002), pp. 325–336.

[28] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, Springer, Berlin, 2006.

[29] A. Iserles, H. Z. Munthe-Kass, S. P. Norsett, A. Zanna, Lie-group methods, Acta Numerica, (2004) pp.1–148.

[30] A.D. Kennedy, M.A. Clark, Speeding up HMC with better integrators, PoS 038(2007).

[31] A. D. Kennedy, M. A. Clark, P. J. Silva, Force Gradient Integrators, PoS 021(2009).

[32] A.D. Kennedy, P.J. Silva, M.A. Clark, Shadow Hamiltonians, Poisson Brackets and Gauge Theories, Phys. Rev. D 87 (2013), 034511.

[33] F. Knechtli, U. Wolf, Dynamical fermions as a global correction, Nucl. Phys. B663 (2003) pp. 3–32.

[34] F. Knechtli, M. Günther, M. Peardon, Lattice Quantum Chromodynamics: Practical Essentials Springer (2016).

[35] A. Lew, J. E. Marsden, M. West, M. Ortiz, Variational time integrators, Int. J. Numer. Methods Eng., 60 (2004) pp.153–212.

[36] M. Lüscher, S. Schäfer, Lattice QCD with open boundary conditions and twisted-mass reweighting, Comput. Phys. Commun. 184(2013), pp 519–528.

[37] R. I. McLachlan, G. R. W. Quispel, Splitting methods, Acta Numerica (2002), pp. 341–434.

[38] R. I. McLachlan, G. R. W. Quispel, Geometric Integrators for ODEs, J. Phys. A 39(19) (2006), pp. 5251-5286.

[39] I. Montvay, G. Münster, Quantum Fields on a Lattice, Cambridge University Press: Cambridge, 1994.

[40] C. Morningstar, The Monte Carlo method in Quantum Field Theory, arXiv:hep-lat/0702020(2007).

[41] H. Munthe-Kaas, Runge-Kutta methods on Lie groups, BIT, 38 (1995), pp. 92–111.

[42] I.P. Omelyan, I.M. Mryglod, R. Folk, Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics, Comput. Phys. Commun. 151 (2003), pp. 272–314.

[43] I.P. Omelyan, I.M. Mryglod, R. Folk, Algorithm for Molecular Dynamics Simulations of Spin Liquids, Phys. Rev. Lett. 86 (2001) 898.

[44] I.P. Omelyan, Advanced gradient like methods for rigid-body molecular dynamics, J. Chem. Phys. 127(2003), 044102.

[45] M. E. Peskin, D. V. Schroeder, An Introduction to Quantum Field Theory, Westview Press, 1995 ISBN: 9780201503975, 0201503972.

[46] J. M. Sanz-Serna, Symplectic Methods, Encyclopedia of Applied and Computational Mathematics, Springer (2013).

[47] V. Savcenco, W. Hundsdorfer, G. J. Verwer, A multirate time stepping strategy for stiff ordinary differential equations, BIT Numerical Mathematics 47(2007) pp. 137–155.

[48] J.S. Schwinger, Gauge Invariance and Mass. 2, Phys. Rev. 128 (1962), pp. 2425–2429.

[49] J.C. Sexton, D.H. Weingarten, Hamiltonian evolution for the hybrid Monte Carlo algorithm, Nucl. Phys. B380 (1992), pp. 665–678.

[50] D. Shcherbakov, M. Ehrhardt, Multistep Methods for Lattice QCD Simulations, PoS (Lattice 2011), pp. 327–333.

[51] D. Shcherbakov, M. Ehrhardt, M. Günther, M. Peardon, Force-gradient nested multirate methods for Hamiltonian systems, Comput. Phys. Commun. 187(2015), pp 91–97.

[52] D. Shcherbakov, M. Ehrhardt, J. Finkenrath, M. Günther, F. Knechtli M. Peardon, Adapted nested force-gradient integrators: the Schwinger model case, Commun. in

Comput. Phy. 21(4) (2017), pp 1141–1153.

[53] P.J. Silva, A.D. Kennedy, M.A. Clark, Tuning HMC using Poisson brackets, PoS 041(2008).

[54] R.D. Skeel, D.J. Hardy, Practical Construction of Modified Hamiltonians, SIAM J. Sci. Comput., 23 (2001), pp. 1172–1188.

[55] J. Smit, Introduction to quantum fields on a lattice: a robust mate, Cambridge Lecture Notes in Physics, 2002, ISBN 0521890519.

[56] M. Suzuki, General theory of higher-order decomposition of exponential operators and symplectic integrators, Phys. Lett. A 165 (1992) 387.

[57] C.R. Sweet, S.S. Hampton, R.D. Skeel, J.A. Izaguirre, A separable shadow Hamiltonian hybrid Monte Carlo method, J. Chem. Phys. 131, 174106 (2009) (7 pages)

[58] C. Urbach, K. Jansen, A. Shindler, U. Wenger, HMC algorithm with multiple time scale integration and mass preconditioning, Comput. Phys. Commun. 174 (2006), pp. 87–98.

[59] U.J. Wiese, An Introduction to Lattice Field Theory, Lecture Material (2009).

[60] H. Yin, R. D. Mawhinney, Improving DWF Simulations: the Force Gradient Integrator and the Möbius Accelerated DWF Solver, PoS (Lattice 2011) 051.

[61] H. Yoshida, Construction of higher order symplectic integrators, Phys. Lett. A 150 (1990) 262.

[62] A. Zanna, K. Engø, H. Z. Munthe-Kaas, Adjoint and selfadjoint Lie group methods, BIT 2001 Vol. 41 No. 2 pp. 395–421.