MICHAEL KRAUS

Variational Integrators in Plasma Physics

$$\sum_{\substack{\square \\ \square \cap \partial \mathcal{U} \neq \varnothing}} \left(\sum_{\substack{a \\ \square^a \in \partial \mathcal{U}}} (j^1 \varphi)^* J^a(\square) \right) = 0$$

$$\sum_a \frac{\partial \mathcal{L}_\square}{\partial y^a} \big(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4} \big) = 0$$

$$\mathcal{A}_d[\varphi] = \sum_\square \mathcal{L}_d \circ j^1 \varphi(\square)$$

q_0

q_N
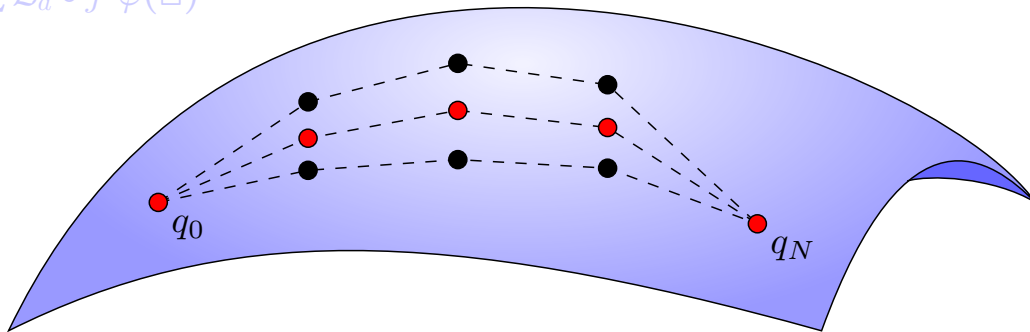
$$\Omega_{L_d} = \frac{\partial^2 L_d}{\partial q_k \, \partial q_{k+1}} (q_k, q_{k+1}) \, dq_k \wedge dq_{k+1}$$

$$\mathbf{d}\mathcal{A}_d \cdot V = \sum_{\substack{\square \\ \square \cap \mathrm{int}\,\mathcal{U} \neq \varnothing}} \mathbf{d}\mathcal{L}_\square \cdot V + \theta_{L_d}(\varphi) \cdot V$$

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

$$\sum_{\substack{\square \\ \square \cap \partial \mathcal{U} \neq \varnothing}} \left(\sum_{\substack{a \\ \square^a \in \partial \mathcal{U}}} \left[(j^1 \varphi)^* (V \lrcorner W \lrcorner \Omega^a_{L_d}) \right](\square) \right) = 0$$

TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Mathematik

Lehrstuhl für Numerische Methoden der Plasmaphysik

Variational Integrators in Plasma Physics

Michael Kraus

Vollständiger Abdruck der von der Fakultät für
Mathematik der Technischen Universität München
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigten Dissertation.

Vorsitzende: Univ.-Prof. Dr. Simone Warzel

Prüfer der Dissertation: 1. Univ.-Prof. Dr. Eric Sonnendrücker
                           2. Univ.-Prof. Dr. Oliver Junge
                           3. Prof. Philip J. Morrison, Ph.D.,
                               University of Texas at Austin/USA

Die Dissertation wurde am 03.04.2013 bei der
Technischen Universität München eingereicht
und durch die Fakultät für Mathematik
am 01.07.2013 angenommen.

Abstract

To a large extent, research in plasma physics is concerned with the description and analysis of energy and momentum transfer between different scales and different kinds of waves. In the numerical modelling of such phenomena it appears to be crucial to describe the transfer processes preserving the underlying conservation laws in order to prevent physically spurious solutions.

In this work, special numerical methods, so called variational integrators, are developed for several models of plasma physics. Special attention is given to conservation properties like conservation of energy and momentum.

By design, variational integrators are applicable to all systems that have a Lagrangian formulation. Usually, equations of motion are derived by Hamilton's action principle and then discretised. In the application of the variational integrator theory, the order of these steps is reversed. At first, the Lagrangian and the accompanying variational principle are discretised, such that discrete equations of motion can be obtained directly by applying the discrete variational principle to the discrete Lagrangian. The advantage of this approach is that the resulting discretisation automatically retains the conservation properties of the continuous system.

Following an overview of the geometric formulation of classical mechanics and field theory, which forms the basis of the variational integrator theory, variational integrators are introduced in a framework adapted to problems from plasma physics. The applicability of variational integrators is explored for several important models of plasma physics: particle dynamics (guiding centre dynamics), kinetic theory (the Vlasov-Poisson system) and fluid theory (magnetohydrodynamics).

These systems, with the exception of guiding centre dynamics, do not possess a Lagrangian formulation to which the variational integrator methodology is directly applicable. Therefore the theory is extended by linking it to Ibragimov's theory of integrating factors and adjoint equations. It allows us to find a Lagrangian for all ordinary and partial differential equations and systems thereof, for which the number of variables equals the number of equations. Consequently, the applicability of the variational integrators is extended to a much larger family of systems as envisaged in the original theory. The theory allows for the application of Noether's theorem to analyse the conservation properties of the system, both at the continuous and the discrete level.

In numerical examples, the conservation properties of the derived schemes are analysed. In case of guiding centre dynamics, momentum in the toroidal direction of a tokamak is preserved exactly. The particle energy exhibits an error, but the absolute value of this error stays constant during the entire simulation. Therefore numerical dissipation is absent. In case of the kinetic theory, the total number of particles, total linear momentum and total energy are preserved exactly, i.e., up to machine accuracy. In case of magnetohydrodynamics, the total energy, cross helicity and the divergence of the magnetic field are preserved up to machine precision.

These conservation properties not only make the numerical schemes more stable than those obtained by traditional discretisation methods, but they also reduce unphysical behaviour like spurious loss of energy or momentum, thereby increasing the trustworthiness of numerical simulations.

Zusammenfassung

Weite Teile der theoretischen Plasmaphysik beschäftigen sich mit der Beschreibung und Untersuchung des Transfers von Energie und Impuls zwischen verschiedenen Skalen und verschiedenen Arten von Wellen. In der numerischen Modellierung dieser Phänomene erscheint es entscheidend, die den Transferprozessen zugrunde liegenden Erhaltungsgesetze zu bewahren, um unphysikalische Lösungen zu unterbinden.

In der vorliegenden Arbeit werden spezielle numerische Verfahren, sogenannte Variationsintegratoren, für verschiedene Modelle der Plasmaphysik entwickelt. Ein besonderer Augenmerk liegt dabei auf der Erhaltung physikalischer Größen wie Impuls und Energie. Prinzipiell sind Variationsintegratoren auf alle Systeme anwendbar, die eine Lagrangesche Formulierung aufweisen. Die grundlegende Idee ist, dass man nicht, wie üblich, mit Hilfe des Hamiltonschen Variationsprinzips Bewegungsgleichungen ableitet und diese dann diskretisiert sondern die Abfolge dieser Schritte umkehrt. Man diskretisiert die Lagrange-Funktion und das Variationsprinzip und leitet damit direkt diskrete Bewegungsgleichungen ab. Diese Vorgehensweise hat den Vorteil, dass man Diskretisierungen erhält, die automatisch die Erhaltungseigenschaften des kontinuierlichen Systems bewahren.

Im Folgenden wird zuerst die Theorie der Variationsintegratoren entwickelt. Da diese auf einer geometrischen Formulierung der klassischen Mechanik und Feldtheorie beruht, werden deren Grundlagen ebenfalls dargestellt. Anschließend wird die Anwendbarkeit der Variationsintegratoren auf verschiedene Systeme der Plasmaphysik untersucht: auf Teilchendynamik (Guiding Centre Dynamik), die kinetische Theorie (das Vlasov-Poisson-System) und die Fluiddynamik (Magnetohydrodynamik).

Da diese Systeme, mit Ausnahme der Guiding Centre Dynamik, keine passende Lagrangesche Formulierung aufweisen, wird die Theorie der Variationsintegratoren durch Kopplung mit Ibragimovs Theorie der integrierenden Faktoren und adjungierten Gleichungen erweitert. Diese erlaubt es, eine Lagrange-Funktion für alle gewöhnlichen und partiellen Differentialgleichungen und Differentialgleichungssysteme zu finden, bei denen die Anzahl der Variablen der Anzahl der Gleichungen entspricht. Dadurch ist es möglich, die Anwendbarkeit der Variationsintegratoren auf eine viel größere Familie von Systemen zu erweitern, als dies in der ursprünglichen Theorie vorgesehen ist. Die Theorie ermöglicht die Anwendung des Noether-Theorems zur Untersuchung der Erhaltungseigenschaften des Systems, sowohl im Kontinuierlichen wie auch im Diskreten.

In numerischen Beispielen werden die Erhaltungseigenschaften der Variationsintegratoren untersucht. Im Falle der Guiding Centre Dynamik ist der Impuls in toroidaler Richtung im Tokamak exakt erhalten. Die Energie weist einen Fehler auf, dessen Absolutwert vom gewählten Zeitschritt abhängt, aber im Laufe einer Simulation konstant bleibt, d.h. es tritt keine numerische Dissipation auf. Im Falle der kinetischen Theorie werden die Gesamtzahl der Teilchen, der Gesamtimpuls und die Gesamtenergie exakt erhalten (d.h. bis auf Maschinengenauigkeit). Im Falle der Magnetohydrodynamik wird ebenfalls die Gesamtenergie und zusätzlich die Kreuzhelizität und die Divergenz des Magnetfeldes exakt erhalten.

Diese Erhaltungseigenschaften führen zu Verfahren mit verbesserter numerischer Stabilität im Vergleich zu Diskretisierungen durch traditionelle Verfahren. Mindestens ebenso wichtig ist aber, dass sie unphysikalisches Verhalten wie die numerische Dissipation von Energie oder Impuls reduzieren und dadurch die Glaubwürdigkeit numerischer Ergebnisse erhöhen.

Contents

Notation

$\mathcal{D}$	Differential operator
$\mathcal{A}$	Action
$\mathcal{A}_d$	Discrete action
H	Hamiltonian
L	Lagrangian
h	Particle Hamiltonian
L_d	Discrete Lagrangian
$\mathcal{H}$	Hamiltonian density
$\mathcal{L}$	Lagrangian density
Θ	Canonical one-form
Ω	Symplectic two-form
Θ_L	Lagrangian one-form
Ω_L	Lagrangian two-form
Θ_d	Discrete one-form
Ω_d	Discrete two-form
t	Time
q	Generalised coordinates
$\dot{q}$	Generalised velocities
p	Generalised (conjugate) momenta
E	Energy
$\mathcal{E}$	Energy functional
f	Distribution function
g	Ibragimov multiplier of f
ϕ	Electrostatic potential
ψ	Ibragimov multiplier of ϕ
Φ	Potential vorticity
F	Functional of the distribution function f
Z	L^2 norm of the distribution function f
p	Kinetic gas pressure
P	Generalised pressure
V	Velocity field
ω	Vorticity
ψ	Streaming function
E	Electric field
B	Magnetic field
J	Current
A	Magnetic vector potential
A^*	Modified vector potential
A	Electromagnetic 4-potential
$\mathbf{F}$	Electromagnetic field tensor
$\mathcal{J}$	Jacobian

$[\cdot,\cdot]$	Poisson brackets
$\{\cdot,\cdot\}$	Lie-Poisson brackets
$[\cdot,\cdot,\cdot]$	Nambu particle brackets
$\{\cdot,\cdot,\cdot\}$	Nambu field brackets
$\langle\cdot,\cdot\rangle$	Pairing
r	Minor tokamak radius
R	Major tokamak radius
q	Safety factor
μ	Magnetic moment
φ	Toroidal angle
ϑ	Poloidal angle
$\mathcal{M},\mathcal{N}$	Manifold
$\Omega^n(\mathcal{M})$	n forms on $\mathcal{M}$
$\mathcal{Q}$	Configuration manifold with coordinates (q)
$\mathsf{T}Q$	Tangent bundle of $\mathcal{Q}$ with coordinates (q,v), velocity phase manifold
T^*Q	Cotangent bundle of $\mathcal{Q}$ with coordinates (q,p)
$\mathcal{C}(\mathcal{Q})$	Set of parametrised curves in $\mathcal{Q}$
$\mathcal{C}(q_1,q_2,[t_1,t_2])$	Set of parametrised curves in $\mathcal{Q}$ that start at q_1, end at q_2, and are parametrised by values in the interval $[t_1,t_2]$
$\mathcal{C}_L(\mathcal{Q})$	Set of parametrised curves in $\mathcal{Q}$ that are solutions of the Euler-Lagrange equations
$\mathcal{C}(\mathcal{Y})$	Set of parametrised sections in $\mathcal{Y}$
$\mathcal{C}_L(\mathcal{Y})$	Set of parametrised sections in $\mathcal{Y}$ that are solutions of the Euler-Lagrange equations
$\mathcal{X}$	Base manifold (usually time, spacetime or phasespacetime)
$\mathcal{Y}$	Fibre bundle
$\mathcal{F}$	Typical fibre
$\mathcal{U}_\mathcal{X}$	Closed submanifold of the base manifold $\mathcal{X}$
$\mathcal{U}$	Parametrisation manifold of $\mathcal{U}_\mathcal{X}$
J^k	k'th jet bundle
j^k	k'th jet prolongation
φ	Section
V,X	Vector field
$\aleph$	General geometric object (scalar function, vector field, differential form)
ξ,η	Transformation map
ϵ,s,t	Group parameter
$\pounds$	Lie derivative
$\mathbf{d}$	Exterior derivative
$\wedge$	Wedge product
$\boldsymbol{\imath}$	Interior product
$\lrcorner$	Contraction
$\cong$	is isomorphic to
iff	if and only if
$\mathrm{int}(\mathcal{U})$	interior of $\mathcal{U}$
$\mathrm{cl}(\mathcal{U})$	closure of $\mathcal{U}$
$\partial\mathcal{U}$	boundary of $\mathcal{U}$

Local deviations from this list cannot be excluded. Some letters are defined twice, such as φ for the toroidal angle as well as for a general section. However, the current meaning should always be clear from the context in which the symbol is used. Some symbols with varying meaning (like for general scalars, scalar fields, vector fields, differential forms, etc.) are not listed here or may overlap with notation defined here.

Coordinates indices i, j, k run from 1 to n, indices μ, ν, σ run from 0 to n, where n is the dimension of the space. Indices a, b correspond the components of fields or sections.

Throughout the whole thesis, all maps are assumed to be smooth, while all manifolds are assumed to be smooth as well as oriented.

1. Introduction

Plasma physics is one of the most challenging fields in classical physics. Not only does it describe systems that consist of a vast number of particles, but these particles are charged and interact through the mean field they generate, leading to a collective behaviour and a tremendous complexity of the dynamics. This complexity repeatedly provides us with new imponderabilities, not anticipated before. To investigate the complicated behaviour inherent to any plasma system, pure theory is not sufficient. Too great is the complexity of nowadays' problems to solve them by pen and paper alone. Therefore computer simulations have become an essential part of plasma physics research, and have been for some time already.

With ever more powerful computers becoming available, ever larger simulations become feasible. Larger simulations in terms of the simulation domain but also in terms of the simulation time. Consider for example simulations of the entire plasma of ITER, a new experimental device under construction in France, which has a volume of about $840\,\mathrm{m}^3$. And think of simulations of an entire plasma discharge, which in the case of ITER might last up to 400 seconds, almost an eternity on the timescales of important plasma processes like small scale turbulence.

To be able to do such long simulations and still obtain accurate results, standard discretisation methods do not suffice. Most often they are based on the minimisation of local errors, but do not limit global error growth, thereby accumulating errors in each and every timestep, eventually leading to unphysical solutions.

Another example is turbulence, one of the large standing problems in classical physics and an ubiquitous topic in plasma physics. Its description involves the analysis of energy and momentum transfer processes between different scales and different kinds of waves. In the numerical modelling of such phenomena it appears to be crucial to describe these transfer processes while preserving the underlying conservation laws in order to prevent physically spurious solutions. It cannot be expected to describe an energy cascade correctly, if energy is numerically created or dissipated.

Lastly, consider magnetic reconnection, a problem that will be dealt with in some more detail later on. It describes how magnetic field lines open up and reconnect in certain physical situations, resulting in a change of the topology of the magnetic field. Even under ideal conditions, that do not feature reconnection processes (i.e., the magnetic field line topology is fixed), most numerical schemes find reconnection events due to numerical dissipation. If such methods are used to model real reconnection processes, one can never be absolutely certain to which extent the results are due to physical effects and to which extent they are just numerical artefacts.

1.1. Geometric Discretisation

To overcome these problems, the global structure of the equations, namely their geometry, has to be taken into account in the course of discretisation. Following Christiansen et al. [31], a geometric structure is a global property, that can be defined independently of

particular coordinate representations of the differential equations at hand (see also Budd and Piggott [22]). Examples for such structures encompass topology, like magnetic field line topology, conservation laws and symmetries, such as conservation of energy which arises through the invariance of a system under infinitesimal time translations, constraints like the divergence of the magnetic field which has to vanish, or identities like those from vector calculus and their generalisations from differential geometry.

The preservation of such geometric properties on the discrete level can have crucial influence on the quality of a simulation. It affects stability and global error growth, reduces numerical artefacts, like spurious loss of energy or momentum, and thereby reduces the likelihood of inaccurate and unphysical behaviour. In difficult cases, simulations often only become possible by using geometric discretisation methods. This is especially true for long time simulations, where the unlimited growth of global errors, like in the energy of the system, can lead to numerical instabilities or at least physically wrong results.

1.2. Symplecticity

A geometric concept that plays an important role throughout this work is symplecticity. For a one-dimensional Hamiltonian system, the symplectic structure amounts to a skew-symmetric matrix that is a measure for phasespace area. Computing the product of this matrix with two phasespace vectors yields the area of the parallelogram spanned by the two vectors. For autonomous Hamiltonian systems, this area is always preserved. Consequently the Hamiltonian flow is called *symplectic*. Maintaining this preservation of area on the discrete level has important consequences for the resulting integrators like very good energy behaviour (for more details see Sanz-Serna and Calvo [111], Leimkuhler and Reich [73] and Hairer et al. [52]).

In more than one dimension, the conserved quantity is the sum of the areas of the parallelograms that result by projecting the two phasespace vectors to the coordinate axes. As a consequence, phasespace volume is preserved under symplectic maps leading to other conservation laws like conservation of the total number of particles in a system. In the framework of partial differential equations, the concept of symplecticity is generalised to multisymplecticity. Simply put, a multisymplectic map is symplectic with respect to both space and time.

It is noteworthy that a symplectic structure can also be defined on the Lagrangian side, indicating that the class of systems endowed with a symplectic structure is larger than the class of Hamiltonian systems. A fact that was already known to Lagrange and the details of which will be explained in chapter two.

1.3. Variational Integrators

One special geometric discretisation method is represented by variational integrators. They can be applied to any equation or system of equations that can be derived by means of a variational principle. The general idea is simple. It can be described as discretising the theory instead of discretising the equations. Part of the development of variational integrators was therefore the development of discrete counterparts of classical mechanics and classical field theory. Although those are not complete counterparts, they are sufficient to derive geometric integration schemes and analyse their properties with respect to the observance of conservation laws.

In order to derive variational integrators one first has to discretise the basic constituents of the variational principle, the Lagrangian and the action integral. One has to approximate the particle positions or fields and their derivatives and select a quadrature rule. Then a discrete variational principle is applied to the resulting discrete action, directly leading to discrete equations of motion (Euler-Lagrange equations). There are several advantages of this method compared with a direct discretisation of the continuous Euler-Lagrange equations. The obtained integrators preserve a discrete analogue of the symplectic or multisymplectic form (for finite or infinite dimensional systems, respectively). This implies conservation of phasespace volume and a very good energy behaviour. In general, the energy is not preserved exactly, but it exhibits an oscillating behaviour about a fixed value. Consequently, the energy is not constant. But what is important is that the energy is not dissipated or growing unphysically, instead its error is bounded. Furthermore, variational integrators conserve discrete momenta, that is conserved quantities corresponding to a symmetry of the system, practically exactly (up to machine precision).

Quite often, one can recover existing methods that are well known for their good conservative properties via a discrete variational principle. The Newmark scheme, Störmer–Verlet, or symplectic Runge-Kutta methods are examples.

1.4. Outline and Contributions

In chapter two, an introduction to the geometric formulation of classical mechanics and classical field theory is given, together with an overview of the most important differential geometric tools. The theory of variational principles is reviewed in a geometric setting. While the material presented in this section is not original, the presentation is detailed and self-contained. It should be accessible to non-specialists, applied mathematicians and theoretical physicists alike.

The theory of Ibragimov is presented. It allows us to find a variational formulation for certain systems that naturally do not have such a formulation. Furthermore, the Noether theorem can be applied in this framework to study symmetries and link them with conservation laws. The combination of Ibragimov's theory with the discrete variational principle is a very important result of this work, as it allows to derive variational integrators for a much larger class of systems than had been foreseen in the original theory.

In chapter three, the basic theory and methodology of variational integrators is presented, both for finite dimensional systems (e.g., particle mechanics) and infinite dimensional systems (e.g., field theories). Proofs for the discrete conservation properties are given or at least sketched, including a discrete version of Noether's theorem. The chapter is closed by an example, namely the advection equation. Ibragimov's theory is used to construct a Lagrangian and consecutively a variational integrator is derived. Both the continuous and the discrete Noether theorem are applied to obtain conservation laws for that equation. Again, the general theory of variational integrators is not original and largely influenced by Marsden and coworkers [84, 65, 83, 76], but the presentation is adapted to our framework. As opposed to this, the application of variational integrators to Ibragimov's extended Lagrangians is proposed here for the first time. Its role for this work is crucial, since for most plasma physics problems a natural variational formulation in terms of Eulerian coordinates has not yet been found.

Chapters four to six explore the applicability of variational integrators to different systems which are important in plasma physics. Three classes of problems are distinguished: particle dynamics, kinetic theory, and plasma fluid theory.

In chapter four, several variational integrators are derived for the motion of particles in a non-uniform magnetic field. Specifically, the motion of centres of the helical trajectory of a particle (guiding centre) is considered, extending previous work in several aspects: different discretisations of the Lagrangian are explored and dynamics in higher dimensions is considered. We find that the variational integrators obtained here show excellent long-time behaviour, describing particle orbits correctly after millions of characteristic times and hundreds of millions of timesteps, while for standard methods like a fourth order Runge-Kutta scheme large deviations from the correct orbit are observed. The chapter concludes with the sketch of a possible application of the derived integrators in particle-in-cell codes. Here, it is possible to employ a variational principle for the combined system of particles and fields, leading to schemes that consistently respect the conservation properties of the complete system.

In chapter five, variational integrators for the Vlasov-Poisson system in one-dimension are derived. This is a typical test bed for kinetic problems, e.g., it has recently been employed to test the conservation properties of new schemes such as discontinuous Galerkin methods [11, 12, 10, 53, 29, 30]. Considering a one dimensional problem (one space plus one velocity coordinate) reduces the computational burden, while retaining the qualitative physical behaviour, including phase-mixing and collective effects. One of the integrators for this system shows extraordinary conservation properties, preserving the total particle number, the total energy, total linear momentum and the L^2 norm exactly, i.e., up to machine accuracy. Problems only arise if the grid resolution is insufficient to resolve small scale structures in the distribution function. To treat such cases, a velocity space collision operator is introduced. It dissipates the L^2 norm and removes subgrid modes, while retaining the conservation of the total particle number, energy and momentum. Furthermore, a linear integrator is derived, which is computationally less demanding but keeps the conservation properties intact, albeit with less accuracy. Energy for example is not preserved to machine precision, but oscillating about some fixed value, as is typical for symplectic integrators. Still, no numerical dissipation is present. The derived integrators are then applied to different standard benchmark cases like Landau damping, the twostream instability and the Jeans instability.

In chapter six, a variational discretisation of magnetohydrodynamics is obtained. The resulting integrator has similarly astonishing properties as the one for the Vlasov-Poisson system, namely exact conservation of the total energy and cross helicity. Here, a staggered grid approach has to be taken to avoid unphysical oscillations in the velocity and pressure fields, a typical problem in incompressible fluid dynamics. The integrator is applied to a range of quite different examples like Alfvén waves, which appear to travel virtually forever through the computational domain, the passive advection of a magnetic loop by the velocity field, the emergence of current sheaths in the turbulent setting of a Orszag-Tang vortex, and several current sheath models as they are used in reconnection studies.

In the appendix, semi-discretisation strategies based on variational integrators or closely related methods are sketched. In appendix A, a variational-spectral method for the vorticity equation and for the Vlasov-Poisson system is derived. Here, the spatial dimensions are transformed into Fourier space and only time or time and velocity are treated variationally. In appendix B, discretisations of Poisson brackets and various generalisations thereof are considered. Here, only phasespace is discretised but not time. The derivations of these discretisations share many similarities with the derivation of variational integrators. Therefore it is not surprising that the resulting schemes are found to be similar.

2. Geometric Mechanics and Field Theory

"Physicists have had a long-lasting love affair with the idea of generating physical laws by setting the derivative of some functional to zero. This is called an action principle. The most famous action principle is Hamilton's principle, which produces Lagrange's equations of mechanics upon variation." - Philip Morrison [89]

In this chapter, a short overview of the geometric formulation of classical mechanics and field theory is given. It is important to understand some of the geometric underpinnings of the treated systems to appreciate the presented geometric discretisation methods that aim at preserving exactly these structures. Unfortunately, the geometric point of view is seldom treated in lectures on classical mechanics at university, nor are they common knowledge in the plasma physics community. The elegance and beauty of the geometric formulation will certainly appeal to the reader yet unfamiliar with it.

We begin with the presentation of the geometric setting, i.e., some basic notions about manifolds, differential forms and fibre bundles. Hereafter, the formulation of Lagrangian mechanics and field theory is presented, at first in an analytic and then in a geometric way, applying the utilities introduced in the first section. In this context, the theory of Ibragimov is reviewed. It allows us to find a Lagrangian for any ordinary or partial differential equation or any system of differential equations where the number of equations equals the number of dependent variables, which is usually the case in physical systems. Hence it allows us to find extended Lagrangian formulations for systems that do not posses a classical Lagrangian as it is often the case in plasma physics.

Some emphasis is put on Noether's theorem, which connects symmetries and conservation laws. Beginning with an analytic description of point transformations and one-parameter groups, the Noether theorem is presented for particles, fields and extended Lagrangian formulations according to Ibragimov. Hereafter, a geometric formulation of Noether's theorem is developed using the notion of momentum maps.

This treatment restrains itself mostly to the Lagrangian side, nevertheless connections with Hamiltonian mechanics and field theory are drawn to compare some results with discretisation methods developed on that side and to outline some alternative strategies. Central to these ideas are various kinds of brackets, namely the classical Poisson brackets and their generalisations in form of Nambu, Lie-Poisson and Dirac brackets.

If the reader is interested in more detailed treatments he can find some recommendations below. There are lots of classical as well as modern introductions to differential geometry and exterior calculus. The more recent ones include Lovett [79], Torres del Castillo [130] and Epstein [37] on the physics oriented side and Lee [72] and Tu [131] on the math oriented side. Some classics are Schutz [116], Burke [23] and Abraham et al. [2]. A short and nevertheless comprehensive overview of differential forms are the lecture notes by Sjamaar [119] which are freely available on the internet. Good introductions can quite often also be found in general relativity textbooks, e.g., Ryder [105], Hobson et al. [54]

and Carroll [24].

A basic introduction to the geometric formulation of classical mechanics is Jose and Saletan [61]. More advanced treatments are Holm et al. [57] and Marsden and Ratiu [82]. Some classics that are still very useful today are Arnold [8], Abraham and Marsden [1], Sudarshan and Mukunda [125] and Saletan and Cromer [107]. Freely available lectures by Holm [55], Marsden [81] and Ratiu [104] can be found on the internet. Two review papers that focus on problems related to fluid dynamics but also have an introductory character are Morrison [89] and Salmon [108].

The theory of jet bundles is introduced in the monographs of Krasil'shchik and Vinogradov [66], Olver [95] and Saunders [113]. Its application to classical mechanics and field theories is explained in Gotay et al. [49], Marsden et al. [84], Aldaya and De Azcarraga [3], Echeverria-Enríquez et al. [35, 36], Giachetta et al. [47, 45, 46], Sardanashvily [112] and references therein. The survey articles by Saunders [114] and Krupka [67] are also instructive.

This chapter is largely influenced by all of the aforementioned references. It makes no claim of originality, except for presenting the material in a mostly self-contained and coherent way. However, the following presentation, especially the sections after the geometric introduction, should be easier accessible, as the level of detail and explanation often surpasses that of the original works which are quite challenging at times.

2.1. Geometric Foundations

This section tries to give a short overview of the geometric foundations underlying the theory presented below. All of the geometric tools that are used later on should be covered, the only exception being basic Lie group theory.

2.1.1. Smooth Manifolds

Probably the simplest definition of a manifold $\mathcal{M}$ is a set of points that can be labelled by coordinates. Locally, manifolds look like the Euclidean space. Globally, however, they might have a much more complicated structure. Therefore it is often not possible to define a global coordinate system on a manifold (think of a circle or a sphere) and one has to find coordinate patches (charts) that together cover the whole manifold. A *chart* (local coordinate system) is a pair $(\mathcal{U}, \phi)$, where $\mathcal{U}$ is an open subset of $\mathcal{M}$ and ϕ is a one-to-one map from $\mathcal{U}$ onto some open subset of $\mathbb{R}^n$

$$\phi(p) = \big(x^1(p), x^2(p), ..., x^n(p)\big), \qquad\qquad p \in \mathcal{U}. \tag{2.1}$$

Hence, a chart labels each point p in $\mathcal{U}$ by n real numbers. If more than one chart is necessary to cover the whole manifold, most likely some points will lie in the domain of more than one chart. In that case we demand that there exists a transition map as follows. If $(\mathcal{U}_1, \phi_1)$ and $(\mathcal{U}_2, \phi_2)$ are two coordinate patches overlapping in $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$, we request that $\psi = \phi_2 \circ \phi_1^{-1}$ is smooth. In that case, the charts ϕ_1 and ϕ_2 are said to be compatible. The set of compatible charts that covers all of a manifold is called an *atlas*. Consider as an example the configuration space $\mathcal{Q}$ of a mechanical system. The definition of a coordinate chart on $\mathcal{Q}$ amounts to a choice of generalised coordinates. Fortunately, for the cases considered in this work, it is always possible to find a global coordinate patch, thereby avoiding the subtleties arising from having more than one coordinate patch.

Most manifolds in mathematical physics are smooth manifolds, continuous and infinitely often differentiable. We will always assume that this is the case. Furthermore, we shall assume that all of the considered manifolds are orientable.

In the remainder of this section we will consider some intrinsic objects and operations that can be defined on manifolds and will be used in the subsequent treatment.

2.1.2. Vector Fields

On trivial manifolds, i.e., such manifolds that can be identified with a linear vector space like $\mathbb{R}^n$, the definition of vector fields is straight forward. Indeed, it is customary to identify points p of the space with the corresponding vector $x = (x^1, ..., x^n)$, leading to the usual notion of vectors. General manifolds, however, are not necessarily linear, so vectors cannot be defined by the usual means. The simplest geometric way to describe a vector V at a point p on a nontrivial manifold is intuitively as the tangent to a parametrised curve $c(t)$ in $\mathcal{M}$, satisfying $c(0) = p$.

Tangent Vectors

A parametrised curve $c(t)$ in $\mathcal{M}$ is a smooth map from some interval $\mathcal{I} \subseteq \mathbb{R}$ to the manifold $\mathcal{M}$

$$c : \mathcal{I} \to \mathcal{M}. \tag{2.2}$$

If coordinates on $\mathcal{M}$ are denoted (x^μ), this can be explicitly written as

$$c : t \mapsto x^\mu(t). \tag{2.3}$$

Without loss of generality assume that $\mathcal{I}$ contains the point $0 \in \mathbb{R}$ and that $c(0) = p$. Consider the directional derivative of a function $f : \mathcal{M} \to \mathbb{R}$ along the curve c, that is

$$V_p(f) = \frac{d}{dt}\Big[f \circ c(t)\Big]\bigg|_{t=0}. \tag{2.4}$$

For the trivial case, $\mathcal{M} = \mathbb{R}^n$, this corresponds to

$$V_p(f) = \frac{dx^\mu}{dt}\frac{\partial f}{\partial x^\mu} \qquad \text{with} \qquad \frac{dx^\mu}{dt} \equiv V_p^\mu. \tag{2.5}$$

V_p^μ are the components of the tangent vector of $c(t)$ at $t = 0$. In the general case, this is used as a definition. As (2.5) is fully general and independent of f, the vector V_p can be written as

$$V_p = V_p^\mu \partial_\mu \qquad \text{where} \qquad \partial_\mu \equiv \frac{\partial}{\partial x^\mu}. \tag{2.6}$$

Vectors on a manifold therefore correspond to first order differential operators. ∂_μ are the local basis in which the vector components are expressed.

Tangent Bundle

The tangent vector V_p is an element of the local *tangent space* $\mathsf{T}_p\mathcal{M}$, where $\mathsf{T}_p\mathcal{M}$ is the set of all tangent vectors (i.e., all possible directional derivatives) to $\mathcal{M}$ at p and has the same dimension as $\mathcal{M}$. That means, $\mathsf{T}_p\mathcal{M}$ can be obtained by considering the tangents to all possible curves passing through that point. Coordinates x^μ on $\mathcal{M}$ induce a basis ∂_μ on $\mathsf{T}_p\mathcal{M}$. Therefore (∂_μ) build a *natural coordinate system* on $\mathsf{T}_p\mathcal{M}$. One possible way of obtaining $\mathsf{T}_p\mathcal{M}$ is to consider all possible curves c through p and evaluate (2.6) for each single one.

An important consequence of vectors at different points p of the manifold being elements of different vector spaces $\mathsf{T}_p\mathcal{M}$ is that they cannot be added or subtracted. This is only possible for vectors at the same point p, i.e., vectors which are elements of the same tangent space $\mathsf{T}_p\mathcal{M}$.

Collecting all the $\mathsf{T}_p\mathcal{M}$, for each point $p \in \mathcal{M}$, into one single object (which is a disjoint union) yields the *tangent bundle*

$$\mathsf{T}\mathcal{M} = \bigcup_{p \in \mathcal{M}} \mathsf{T}_p\mathcal{M}. \tag{2.7}$$

It is the set of all tangent vectors at all points of $\mathcal{M}$ and has the structure of a differentiable manifold. More details on the tangent bundle will be presented in section 2.1.4 on fibre bundles.

A vector field V on a manifold $\mathcal{M}$ is a function that assigns a vector $V_p \in \mathsf{T}_p\mathcal{M}$ to each point $p \in \mathcal{M}$. Consequently, all vector fields V on $\mathcal{M}$ lie in $\mathsf{T}\mathcal{M}$. A vector field V is therefore a smooth, linear map

$$V : \mathcal{M} \to \mathsf{T}\mathcal{M}. \tag{2.8}$$

The tangent space $\mathsf{T}_p\mathcal{M}$ at each point $p \in \mathcal{M}$ is a real vector space. Therefore two vector fields V and W may be added or multiplied by a scalar field $f : \mathcal{M} \to \mathbb{R}$ as follows

$$(V + W)(p) = V(p) + W(p), \qquad (fV)(p) = f(p)\,V(p), \qquad p \in \mathcal{M}. \tag{2.9}$$

In the physics literature, the V^μ are usually referred to as contravariant components of the vector field V.

2.1.3. Integral Curves and Flows

We will now study a topic that is essential in the geometric formulation of the action principle and the study of symmetries. It is based on the observation that vector fields induce, at least locally, a family of transformations of the manifold onto itself.

A one-parameter family of transformations is a differentiable map

$$\varphi : \mathcal{M} \times \mathbb{R} \to \mathcal{M} \tag{2.10}$$

that depends on a real parameter. Therefore they are called one-parameter-groups of transformations. They map points of $\mathcal{M}$ onto different points of $\mathcal{M}$

$$\varphi : p \mapsto \varphi(p, t) \qquad \text{with} \qquad p \in \mathcal{M}, \quad t \in \mathbb{R}, \tag{2.11}$$

such that

$$\varphi(p, 0) = p \quad \text{and} \quad \varphi\big(\varphi(p, s), t\big) = \varphi(p, s + t) \quad \text{for all} \quad p \in \mathcal{M}, \quad s, t \in \mathbb{R}. \tag{2.12}$$

Thus, upon defining $\varphi_t(p) \equiv \varphi(p,t)$ we can write

$$\varphi_{s+t}(p) = \varphi_s \circ \varphi_t = \varphi_t \circ \varphi_s \qquad \text{and} \qquad \varphi_0 = \text{id} \,. \tag{2.13}$$

As

$$\varphi_t \circ \varphi_{-t} = \varphi_{-t} \circ \varphi_t = \varphi_0 = \text{id} \tag{2.14}$$

each map φ_t has an inverse $\varphi_t^{-1} = \varphi_{-t}$ that is also differentiable. Therefore, each φ_t is a diffeomorphism of $\mathcal{M}$ onto itself, and the set of transformations $\{\varphi_t \mid t \in \mathbb{R}\}$ is a group of diffeomorphisms of $\mathcal{M}$ onto itself.

Each one-parameter-group of transformations φ on $\mathcal{M}$ determines a family of curves in $\mathcal{M}$ (referred to as the *orbits* of the group). The map

$$\varphi_p : \mathbb{R} \to \mathcal{M} \tag{2.15}$$

given by

$$\varphi_p(t) = \varphi(p,t) \tag{2.16}$$

is a differentiable curve in $\mathcal{M}$ for each $p \in \mathcal{M}$. The vector field tangent to these curves generated by the one-parameter-group of transformations

$$V = \frac{d}{dt} \varphi_t \bigg|_{t=0} \tag{2.17}$$

is called the *infinitesimal generator* of φ. Since $\varphi_p(0) = \varphi(p,0) = p$ the tangent vector to the curve φ_p belongs to $\mathsf{T}_p\mathcal{M}$. The curves φ_p are integral curves of V.

2.1.4. Fibre Bundles

Reconsider the construction of the tangent bundle $\mathsf{T}\mathcal{M}$ of a manifold $\mathcal{M}$ from section 2.1.2. $\mathsf{T}\mathcal{M}$ was built by attaching to each point $p \in \mathcal{M}$ the tangent vector space $\mathsf{T}_p\mathcal{M}$ at that point (2.7). The resulting object is generally referred to as a *fibre bundle* with the vector spaces $\mathsf{T}_p\mathcal{M}$ being the *fibres* that are attached to each point p of the *base space* $\mathcal{M}$. See figure 2.1 for a pictorial view of the tangent bundle of the circle $\mathbb{S}^1$.
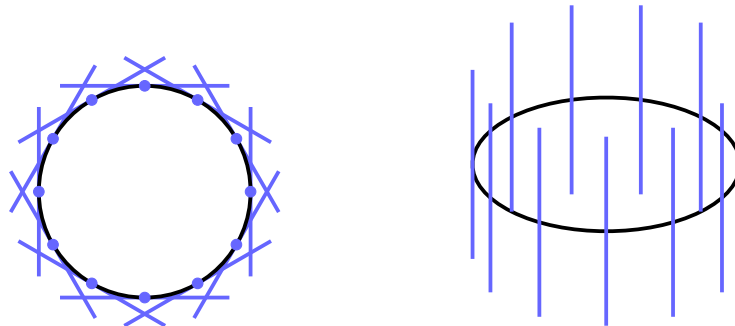
Fig. 2.1.: Left: Tangent bundle of the circle $\mathbb{S}^1$. Each of the blue lines attached to a point of the circle depicts a fibre of the tangent bundle $\mathsf{T}\mathbb{S}^1$. Right: To avoid spurious intersections, the same tangent bundle $\mathsf{T}\mathbb{S}^1$ is drawn with the fibres parallel to each other. Now the circle is depicted in the horizontal plane and the fibres $\mathsf{T}_p\mathbb{S}^1$ are vertical lines.

In general, a fibre bundle is characterised by four quantities, the base space $\mathcal{X}$, the total space $\mathcal{Y}$, a typical fibre $\mathcal{F}$ and a projection π. Fibre bundles are particular manifolds with the property of being decomposable into fibres. The points of a single fibre are related to one another while points of different fibres are not. This is formalised by defining a *projection map.*

Projections

Consider a fibre bundle $\mathcal{Y}$ over $\mathcal{X}$, a point x in the base space $\mathcal{X}$, and the fibre $\mathcal{Y}_x \cong \mathcal{F}$ at that point. The *natural* (or *canonical*) *projection* π maps each element y of $\mathcal{Y}_x$ to the point x, the fibre is attached to, i.e.,

$$\pi_\mathcal{X} : \mathcal{Y} \to \mathcal{X} \tag{2.18}$$

or in coordinates

$$\pi_\mathcal{X} : (x, y) \mapsto x. \tag{2.19}$$

To avoid confusion, e.g., if more than one projection appears in a treatment, we sometimes also write $\pi_{\mathcal{X}\mathcal{Y}}$ to denote both the source and the target space explicitly.

Sections

A section φ of a fibre bundle $\mathcal{Y}$ over $\mathcal{X}$ is a map that assigns to each point x in the base manifold a point in the fibre bundle

$$\varphi : \mathcal{X} \to \mathcal{Y} \qquad \text{such that} \qquad \pi_{\mathcal{X}\mathcal{Y}} \circ \varphi = \mathrm{id}_\mathcal{X}, \tag{2.20}$$

or in coordinates

$$\varphi : x \mapsto (x, y). \tag{2.21}$$

Sections $\varphi : \mathcal{X} \to \mathcal{Y}$ are also denoted as $\varphi \in \Gamma(\pi_{\mathcal{X}\mathcal{Y}})$, where $\Gamma(\pi_{\mathcal{X}\mathcal{Y}})$ denotes the set of all sections in the fibre bundle $\mathcal{Y}$ over $\mathcal{X}$, characterised by the projection $\pi_{\mathcal{X}\mathcal{Y}}$. Similarly, the fibre bundle itself is often just denoted as $\pi_{\mathcal{X}\mathcal{Y}}$. If the model space $\mathcal{F}$ for fibres is a linear space, then $\Gamma(\pi_{\mathcal{X}\mathcal{Y}})$ has a natural linear space structure as well.

In this framework, vector fields are sections of $\mathsf{T}\mathcal{M}$. Later on, we will introduce jet bundles that are very practical when dealing with field theories. These are fibre bundles defined over spacetime that contain fields and their derivatives up to a given order.
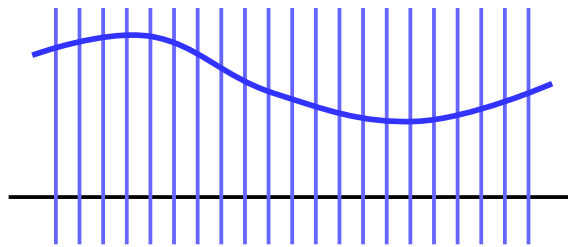
Fig. 2.2.: A section in the tangent bundle $\mathsf{T}\mathbb{R}$ of the real line.

Vertical Lifts

Some fibre bundles $\pi_{\mathcal{X}\mathcal{Y}}$, e.g., tangent bundles or jet bundles, have a natural way of *lifting* a curve in $\mathcal{X}$ by a section y. The tangent lift

$$\mathsf{T}f : \mathsf{T}\mathcal{M} \to \mathsf{T}\mathcal{N} \tag{2.22}$$

of a function

$$f : \mathcal{M} \to \mathcal{N} \tag{2.23}$$

is defined by

$$V_p \mapsto Df\, V_p. \tag{2.24}$$

The lift of a curve $c(t)$ in $\mathcal{X}$ to the tangent bundle $\mathsf{T}\mathcal{X}$ is given as

$$\mathsf{T}c(t) : t \mapsto \left(x^\mu(t), \frac{dx^\mu}{dt}(t) \right). \tag{2.25}$$

Obviously, not every curve in the tangent bundle is the lift of a curve in the base space. Those curves that are lifted ones are called *holonomic*. If φ_t is a one-parameter group of diffeomorphisms on the basespace $\mathcal{X}$, then $\mathsf{T}\varphi_t$ will be a one-parameter group of diffeomorphisms on the tangent bundle $\mathsf{T}\mathcal{X}$. Thus, if V is the infinitesimal generator of φ_t, its tangent lift to $\mathsf{T}(\mathsf{T}\mathcal{X})$ is given as

$$\mathsf{T}V = \frac{d}{dt}\,\mathsf{T}\varphi_t \bigg|_{t=0} \tag{2.26}$$

in coordinates

$$\mathsf{T}V : (x^\mu, V^\mu) \mapsto \left(\left(x^\mu, \frac{dx^\mu}{dt} \right), \left(\frac{dx^\mu}{dt}, \frac{d^2x^\mu}{dt^2} \right) \right). \tag{2.27}$$

The space of such vector fields over $\mathsf{T}\mathcal{X}$ plays an important role in the next section on Lagrangian dynamics.

2.1.5. Differential Forms

"Very loosely speaking, differential forms are whatever appears under an integral sign."
<div align="right">- Loring Tu [131]</div>

Somewhat less loosely speaking, they allow us to define integrands over manifolds, thereby providing a natural way to integrate over curved spaces. Just as vector fields, they are intrinsic objects associated to any manifold, and in fact their simplest instance, differential one-forms, are the dual concept to vector fields (which are first order differential operators).

Differential One-Forms

A differential one-form α (hereafter just referred to as one-form) on a manifold $\mathcal{M}$ assigns an element of the dual space $\mathsf{T}_p^*\mathcal{M}$ of the tangent space $\mathsf{T}_p\mathcal{M}$ to each point $p \in \mathcal{M}$. It is therefore a smooth, linear map

$$\alpha : \mathcal{M} \to \mathsf{T}^*\mathcal{M}. \tag{2.28}$$

Each such one-form takes values in the *cotangent space* $\mathsf{T}_p^*\mathcal{M}$ at that point, where $\mathsf{T}_p^*\mathcal{M}$ contains all dual vectors (covariant vectors) at that point p and has the same dimension as $\mathcal{M}$. Collecting all the $\mathsf{T}_p^*\mathcal{M}$ for each point $p \in \mathcal{M}$ into one single object gives the *cotangent bundle*

$$\mathsf{T}^*\mathcal{M} = \bigcup_{p \in \mathcal{M}} \mathsf{T}_p^*\mathcal{M}, \tag{2.29}$$

such that the one-form α is a section of $\mathsf{T}^*\mathcal{M}$. That way, $\mathsf{T}^*\mathcal{M}$ is also the dual to $\mathsf{T}\mathcal{M}$. The space $\mathsf{T}_p^*\mathcal{M}$ is a vector space, such that two forms α and β of the same order may be added or multiplied by a scalar field $f : \mathcal{M} \to \mathbb{R}$

$$(\alpha + \beta)(p) = \alpha(p) + \beta(p), \qquad (f\beta)(p) = f(p)\,\beta(p), \qquad p \in \mathcal{M}. \tag{2.30}$$

The simplest geometric way to describe a one-form α is as the differential of a function f on $\mathcal{M}$. Assign to each point p of the curve local coordinates $x^\mu(t)$ and a scalar function $f(x^\mu)$ on $\mathcal{M}$. The differential of the function f at p is

$$df = \frac{\partial f}{\partial x^\mu}\,dx^\mu. \tag{2.31}$$

The first expression corresponds to the components of the gradient of f, and (dx^μ) forms a local basis, dual to the basis (∂_μ) on $\mathsf{T}\mathcal{M}$. However, not all one-forms are differentials of a function. In general, one-forms are written as

$$\alpha = \alpha_\mu\,dx^\mu, \tag{2.32}$$

where (dx^μ) is the aforementioned basis on $\mathsf{T}^*\mathcal{M}$, defined by letting a basis one-form act on a basis vector, i.e.,

$$dx^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \delta_\nu^\mu, \tag{2.33}$$

such that the result of a general one-form α acting on a general vector v is given by

$$\alpha(v) = \alpha_\mu\,v^\nu\,\delta_\nu^\mu = \alpha_\mu\,v^\mu. \tag{2.34}$$

A one-form α is a linear functional, such that it acts on a linear combination of vectors $v, w \in \mathsf{T}\mathcal{M}$ with scalars $a, b \in \mathbb{R}$ as

$$\alpha(av + bw) = a\,\alpha(v) + b\,\alpha(w). \tag{2.35}$$

In the physics literature, the α_μ are usually referred to as covariant components of the covector field α. Strictly speaking, one-forms can only be identified with covector fields if the underlying manifold is endowed with a metric, which defines a canonical isomorphism of $\mathsf{T}_p^*\mathcal{M}$ and $\mathsf{T}_p\mathcal{M}$, thus identifying vectors and there duals. In physical applications this is indeed most often the case. To change between vectors and one-forms, the $\flat$ and $\sharp$ operators can be defined like

$$V^\flat = V_i dx^i \qquad \text{and} \qquad \alpha_\sharp = \alpha^i \partial_i. \tag{2.36}$$

The flat operator $\flat$ returns the one-form corresponding to a vector field, and the sharp operator $\sharp$ returns the vector field corresponding to a one-form. Therefore, the action is the same as in music, but with respect to indices instead of notes.

Higher Order Differential Forms

A differential two-form ω is a function

$$\omega : \mathcal{M} \to \Omega^2(\mathcal{M}) \tag{2.37}$$

where $\Omega^2(\mathcal{M})$ is the space of two-forms on $\mathcal{M}$. It is generally written[1]

$$\omega = \frac{1}{2}\,\omega_{\mu\nu}\,dx^\mu \wedge dx^\nu. \tag{2.38}$$

where $\wedge$ denotes the wedge product, which is defined in the next paragraph. Two-forms are antisymmetric, such that

$$\omega_{\mu\nu} = -\omega_{\nu\mu}. \tag{2.39}$$

A similar result as (2.35) holds also for a two-form ω, which is a bilinear functional, acting on vectors $v, w, z \in \mathsf{T}\mathcal{M}$ with scalars $a, b \in \mathbb{R}$ as

$$\omega(av + bw, z) = a\,\omega(v, z) + b\,\omega(w, z) \quad \text{and} \quad \omega(v, aw + bz) = a\,\omega(v, w) + b\,\omega(v, z). \tag{2.40}$$

Differential n-forms (differential forms of order n) are completely covariant, totally anti-symmetric tensors. A n-form θ is a function

$$\theta : \mathcal{M} \to \Omega^n(\mathcal{M}), \tag{2.41}$$

where $\Omega^n(\mathcal{M})$ is the space of n forms on $\mathcal{M}$. As a consequence of the antisymmetry property (2.39), the highest order forms that can exist on a manifold $\mathcal{M}$ of dimension m are of order m. The generalisation of (2.40) to higher order forms is straight forward.

Examples: Differential Forms in a Three-Dimensional Manifold

1-Form:

$$A = A_\mu\,dx^\mu = A_1\,dx^1 + A_2\,dx^2 + A_3\,dx^3 \tag{2.42}$$

2-Form:

$$F = \frac{1}{2}\,F_{\mu\nu}\,dx^\mu \wedge dx^\nu = F_{12}\,dx^1 \wedge dx^2 + F_{23}\,dx^2 \wedge dx^3 + F_{31}\,dx^3 \wedge dx^1 \tag{2.43}$$

3-Form:

$$\Omega = \frac{1}{3!}\,\Omega_{\mu\nu\sigma}\,dx^\mu \wedge dx^\nu \wedge dx^\sigma = \Omega_{123}\,dx^1 \wedge dx^2 \wedge dx^3 \tag{2.44}$$

---

[1] The factor of 1/2 originates from the fact that in the sum over indices all contributions are taken into account twice. Whether it is written or not depends on notational convention.

Wedge Product

The wedge product takes a p-form ξ and a q-form η and returns a $(p+q)$-form

$$\xi \wedge \eta \, (v_1, v_2, ..., v_{p+q}) = \frac{1}{(p+q)!} \sum_{\sigma \in S_{p+q}} \mathrm{sgn}(\sigma) \, \xi(v_{\sigma_1}, v_{\sigma_2}, ..., v_{\sigma_p}) \, \eta(v_{\sigma_{p+1}}, v_{\sigma_{p+2}}, ..., v_{\sigma_{p+q}}),$$

$$(2.45)$$

where $(\sigma_1, ..., \sigma_p, \sigma_{p+1}, ..., \sigma_{p+q})$ is an element of S_{p+q}, the group of all permutations of the numbers $\{1, 2, ..., p+q\}$, and $\mathrm{sgn}(\sigma)$ is the sign of the permutation, i.e.,

$$\mathrm{sgn}(\sigma) = \begin{cases} 1 & \text{odd permutation,} \\ 0 & \text{even permutation,} \end{cases} \tag{2.46}$$

such that

$$\xi \wedge \eta = (-1)^{pq} \eta \wedge \xi. \tag{2.47}$$

It is associative,

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma), \tag{2.48}$$

and bilinear,

$$\begin{aligned} (a\,\alpha + b\,\beta) \wedge \gamma &= a\,(\alpha \wedge \gamma) + b\,(\beta \wedge \gamma), \\ \alpha \wedge (b\,\beta + c\,\gamma) &= b\,(\alpha \wedge \beta) + c\,(\alpha \wedge \gamma), \end{aligned} \tag{2.49}$$

but in general not commutative. Due to the antisymmetry property (2.39), the wedge product of a basis form with itself vanishes,

$$dx^\mu \wedge dx^\mu = 0. \tag{2.50}$$

For that reason there can be no forms of higher order than the dimensionality of the space they are defined on.

Examples: Wedge Products of Differential Forms

Consider the examples from above, again defined on a three-dimensional manifold,

$$A = A_1 \, dx^1 + A_2 \, dx^2 + A_3 \, dx^3, \tag{2.51a}$$
$$F = F_{12} \, dx^1 \wedge dx^2 + F_{23} \, dx^2 \wedge dx^3 + F_{31} \, dx^3 \wedge dx^1, \tag{2.51b}$$
$$\Omega = \Omega_{123} \, dx^1 \wedge dx^2 \wedge dx^3. \tag{2.51c}$$

The wedge product of A with itself is

$$\begin{aligned} A \wedge A &= A_1 A_2 \, dx^1 \wedge dx^2 + A_2 A_3 \, dx^2 \wedge dx^3 + A_3 A_1 \, dx^3 \wedge dx^1 \\ &\quad + A_2 A_1 \, dx^2 \wedge dx^1 + A_3 A_2 \, dx^3 \wedge dx^2 + A_1 A_3 \, dx^1 \wedge dx^3 = 0, \end{aligned} \tag{2.52}$$

which is obvious as for one-forms $A \wedge A = -A \wedge A$ (2.47). The wedge product of A and F is

$$A \wedge F = (A_1 F_{23} + A_2 F_{31} + A_3 F_{12}) \, dx^1 \wedge dx^2 \wedge dx^3. \tag{2.53}$$

Let us try to compute the wedge product of Ω with the basis forms dx^μ, that is

$$\Omega \wedge dx^1 = \Omega_{123}\, dx^1 \wedge dx^2 \wedge dx^3 \wedge dx^1 = \Omega_{123}\, \underbrace{dx^1 \wedge dx^1}_{=0} \wedge dx^2 \wedge dx^3 = 0, \tag{2.54a}$$

$$\Omega \wedge dx^2 = \Omega_{123}\, dx^1 \wedge dx^2 \wedge dx^3 \wedge dx^2 = -\Omega_{123}\, dx^1 \wedge \underbrace{dx^2 \wedge dx^2}_{=0} \wedge dx^3 = 0, \tag{2.54b}$$

$$\Omega \wedge dx^3 = \Omega_{123}\, dx^1 \wedge dx^2 \wedge \underbrace{dx^3 \wedge dx^3}_{=0} = 0. \tag{2.54c}$$

We see that all of these vanish, which is no surprise as Ω is a form of maximum order.

Interior Product

One-forms α are linear functionals that map vector fields to functions

$$\alpha : \mathsf{T}\mathcal{M} \to \mathbb{R}. \tag{2.55}$$

In general, n-forms θ are n-linear functionals, mapping n vector fields to functions

$$\theta : \bigotimes_n \mathsf{T}\mathcal{M} \to \mathbb{R}. \tag{2.56}$$

The *interior product* of a vector field v and a one-form α is defined as their *contraction*, denoted

$$\boldsymbol{\imath}_v \alpha = v \lrcorner \alpha = \langle \alpha, v \rangle. \tag{2.57}$$

The interior product $\boldsymbol{\imath}_v$ of some n-form θ and a vector field v yields a $(n-1)$-form $\boldsymbol{\imath}_v\theta$,

$$\boldsymbol{\imath}_v \theta (\underbrace{\ldots\ldots}_{n \text{ slots}}) = \theta(\overbrace{v, \underbrace{\ldots\ldots}_{n-1 \text{ slots}}}^{n \text{ slots}}). \tag{2.58}$$

It can therefore be regarded as a map

$$\boldsymbol{\imath}_v : \Omega^n(\mathcal{M}) \to \Omega^{n-1}(\mathcal{M}) \tag{2.59}$$

or component-wise

$$\boldsymbol{\imath}_v : \theta^{j_1 j_2 \ldots j_p} \mapsto \theta_{k\, j_2 \ldots j_p}\, v^k. \tag{2.60}$$

The interior product of a vector field v and a scalar function f is zero by definition. $\boldsymbol{\imath}_v$ is an anti-derivation: for a p-form ξ and a q-form η

$$\boldsymbol{\imath}_v(\xi \wedge \eta) = (\boldsymbol{\imath}_v \xi) \wedge \eta + (-1)^p \xi \wedge (\boldsymbol{\imath}_v \eta). \tag{2.61}$$

For example, for $v = v^\sigma \partial_\sigma$ and the two-form $dx^\mu \wedge dx^\nu$, we have

$$\boldsymbol{\imath}_v(dx^\mu \wedge dx^\nu) = (\boldsymbol{\imath}_v dx^\mu)\, dx^\nu - dx^\mu\, (\boldsymbol{\imath}_v dx^\nu) = v^\mu\, dx^\nu - v^\nu\, dx^\mu. \tag{2.62}$$

Exterior Derivative

The exterior derivative **d** maps n-forms into $(n + 1)$-forms

$$\mathbf{d} : \Omega^n(\mathcal{M}) \to \Omega^{n+1}(\mathcal{M}), \tag{2.63}$$

thus taking functions (which are considered zero-forms) to one-forms, one-forms to two-forms, and so on. It is axiomatically defined as follows. If $f : \mathcal{M} \to \mathbb{R}$ is a function (zero-form), than $\mathbf{d}f$ is the ordinary differential

$$\mathbf{d}f(v) = vf = v^\mu \partial_\mu f, \tag{2.64}$$

equivalently

$$\mathbf{d}f = (\partial_\mu f) \, dx^\mu. \tag{2.65}$$

d is an anti-derivation, i.e., if ξ is a p-form and η a q-form, than

$$\mathbf{d}(\xi \wedge \eta) = \mathbf{d}\xi \wedge \eta + (-1)^p \, \xi \wedge \mathbf{d}\eta. \tag{2.66}$$

When applied twice, the exterior derivative vanishes, i.e., $\mathbf{d}^2 = 0$ or $\mathbf{d}(\mathbf{d}\alpha) = 0$ for any n-form θ. The exterior derivative **d** is linear, such that for every $a \in \mathbb{R}$

$$\mathbf{d}(a\alpha) = a \, \mathbf{d}\alpha \qquad \text{and} \qquad \mathbf{d}(\alpha + \beta) = \mathbf{d}\alpha + \mathbf{d}\beta. \tag{2.67}$$

The vanishing of the exterior derivative when applied twice, $\mathbf{d}^2 = 0$, leads to the notion of closed and exact forms. A n-form θ is *closed* if $\mathbf{d}\theta = 0$. A n-form θ is *exact* if $\theta = \mathbf{d}\eta$ for a $(n-1)$-form η. An exact form is always closed, but a closed form is not necessarily exact.

Example: Electromagnetic Field

The natural description of the magnetic potential A is as a one-form

$$A = A_1 \, dx^1 + A_2 \, dx^2 + A_3 \, dx^3. \tag{2.68}$$

The exterior derivative of A,

$$\mathbf{d}A = \left(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2}\right) dx^1 \wedge dx^2 + \left(\frac{\partial A_1}{\partial x^3} - \frac{\partial A_3}{\partial x^1}\right) dx^3 \wedge dx^1 + \left(\frac{\partial A_3}{\partial x^2} - \frac{\partial A_2}{\partial x^3}\right) dx^2 \wedge dx^3$$

$$= \frac{1}{2} \left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) dx^\mu \wedge dx^\nu \tag{2.69}$$

$$\equiv \frac{1}{2} F_{\mu\nu} \, dx^\mu \wedge dx^\nu, \tag{2.70}$$

yields the magnetic field tensor

$$F = F_{12} \, dx^1 \wedge dx^2 + F_{23} \, dx^2 \wedge dx^3 + F_{31} \, dx^3 \wedge dx^1. \tag{2.71}$$

Interestingly, $\mathbf{d}A$ looks like a curl, and indeed, the components of F correspond to the components of the magnetic field $B = \nabla \times A$,

$$F = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}. \tag{2.72}$$

> Therefore the natural representation of the magnetic field is a two-form. The exterior derivative of F,
>
> $$\mathbf{d}F = \left(\frac{\partial F_{23}}{\partial x^1} + \frac{\partial F_{31}}{\partial x^2} + \frac{\partial F_{12}}{\partial x^3}\right) dx^1 \wedge dx^2 \wedge dx^3, \tag{2.73}$$
>
> does of course vanish as $\mathbf{d}F = \mathbf{d}^2 A = 0$. Interestingly, $\mathbf{d}F$ looks like a divergence and indeed, it corresponds to $\nabla \cdot B = 0$.

These examples show that on a three-dimensional manifold, the exterior derivative corresponds to the operators from vector calculus. The exterior derivative of a zero-form corresponds to the gradient, the exterior derivative of a one-form corresponds to the curl, and the exterior derivative of a two-form corresponds to the divergence. On manifolds of dimension other than three, the exterior derivative provides a generalisation of these operators.

2.1.6. Pullback

In all considerations of this subsection, φ is regarded as a diffeomorphism

$$\varphi : \mathcal{M} \to \mathcal{N}. \tag{2.74}$$

The pullback of φ allows us to *pull back* geometric objects from the target manifold $\mathcal{N}$ to the source manifold $\mathcal{M}$. This includes functions, vector fields and differential forms. The pullback of a scalar field $f : \mathcal{N} \to \mathbb{R}$ by φ is given by composition

$$\varphi^* f = f \circ \varphi. \tag{2.75}$$

The result is a scalar field $\varphi^* f : \mathcal{M} \to \mathbb{R}$. The pullback of a n-form $\omega \in \Omega^n(\mathcal{N})$ by φ is a n-form $\varphi^* \omega \in \Omega^n(\mathcal{M})$, defined point-wise by

$$\left(\varphi^* \omega\right)_p(v_1, ..., v_n) = \omega_{\varphi(p)}\left(D\varphi(p) \cdot v_1, ..., D\varphi(p) \cdot v_n\right), \qquad p \in \mathcal{N}. \tag{2.76}$$

While $\varphi^* \omega$ is acting on vectors $v_i \in \mathsf{T}_p\mathcal{M}$, ω is acting on vectors $D\varphi(p) \cdot v_i \in \mathsf{T}_{\varphi(p)}\mathcal{N}$. The following diagram should help to clarify this.

$$
\begin{array}{ccc}
\mathsf{T}\mathcal{M} & \xrightarrow{\ \mathsf{T}\varphi\ } & \mathsf{T}\mathcal{N} \\
\downarrow & & \downarrow \\
\mathcal{M} & \xrightarrow{\ \varphi\ } & \mathcal{N} \\
\in & & \in \\
\varphi^*\omega & & \omega
\end{array}
$$

The pullback of a wedge product is the wedge product of the pullback

$$\varphi^*(\alpha \wedge \beta) = (\varphi^* \alpha) \wedge (\varphi^* \beta). \tag{2.77}$$

The pullback of an exterior derivative is the exterior derivative of the pullback

$$\varphi^*(\mathbf{d}\omega) = \mathbf{d}(\varphi^* \omega) \tag{2.78}$$

where ω is any differential form.

2.1.7. Lie Derivative

As already pointed out, on a manifold $\mathcal{M}$ it is generally not possible to add or subtract vectors at different points $p \in \mathcal{M}$, as those live in different vector spaces. This causes a problem as for differentiation in the usual sense, one needs to do exactly that. Nevertheless, there are ways to define derivatives on manifolds. Otherwise, they would not be fun to deal with. The arguably most important derivative on manifolds is the Lie derivative. It describes how a geometric object (a function, a vector, a form) changes when it is dragged along some vector field X. It will be defined below in two different approaches that have been shown to be equivalent [57, 82].

Algebraic Definition

The Lie derivative $\pounds_X$ along a vector field $X \equiv X^\mu \partial_\mu$ is defined through its action on a scalar function f,

$$\pounds_X f \equiv X f = X^\mu f_{,\mu}, \tag{2.79}$$

a vector field $Y \equiv Y^\nu \partial_\nu$,

$$\pounds_X Y \equiv XY - YX = \left(X^\mu (\partial_\mu Y^\nu) - Y^\mu (\partial_\mu X^\nu) \right) \partial_\nu, \tag{2.80}$$

where in the second identity, the products XY and YX are viewed as composition of differential operators. The Lie derivative of anything else is defined through the requirement that it is a derivative, i.e., that it fulfils Leibniz' rule. The Lie derivative of differential forms can be obtained by following this rule. The result is a very beautiful relation called Cartan's magic formula or *infinitesimal homotopy relation*

$$\pounds_X \omega = \mathbf{d} \left(\boldsymbol{\imath}_X \omega \right) + \boldsymbol{\imath}_X \left(\mathbf{d} \omega \right) \qquad \text{(\textbf{Cartan's Magic Formula})}. \tag{2.81}$$

Dynamical Definition

An alternative definition of the Lie derivative along a vector field X with flow φ_t is given by

$$\pounds_X \aleph = \left. \frac{d}{dt} \varphi_t^* \aleph \right|_{t=0}, \tag{2.82}$$

where $\aleph$ can now be a scalar function, a vector field or a differential form. So this definition, referred to as dynamical definition of the Lie derivative, is formally the same for all geometric entities (see for example Marsden and Ratiu [82] or Holm et al. [57]).

Properties

The Lie derivative does not change the tensorial character of the object it is acting on, e.g., a scalar stays a scalar, a vector stays a vector, a one-form stays a one-form, and so on. It commutes with the exterior derivative,

$$\pounds_X \mathbf{d}\omega = \mathbf{d} \left(\pounds_X \omega \right), \tag{2.83}$$

and if $\varphi : \mathcal{M} \to \mathcal{N}$ is a diffeomorphism, the pullback of the Lie derivative by φ is

$$\varphi^* \pounds_X \omega = \pounds_{\varphi^* X} (\varphi^* \omega). \tag{2.84}$$

Now we should be well equipped to approach the geometric formulation of Lagrangian dynamics.

2.2. Lagrangian Dynamics

At the age of 19, Lagrange found a solution to the long-standing isoperimetric problem[2,3]. As it turned out, more important than the answer to this special problem was Lagrange's solution method, which lead to what we now call the *Euler-Lagrange equations*

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0, \tag{2.85}$$

where $L(q, \dot{q})$ is the Lagrangian function, which often corresponds to the kinetic energy minus the potential energy, q are generalised coordinates and $\dot{q}$ generalised velocities. The great advantage of this formulation is that it is completely covariant. It does not depend on a specific coordinate representation.

Some years later, at the age of 18, Hamilton found that these equations can be derived by the *principle of stationary action*[4]. It states that, considering all possible trajectories $q(t)$ a system could follow to get from state a to state b, the following integral, called the action,

$$\mathcal{A}[q] = \int_a^b L\big(q(t), \dot{q}(t)\big) \, dt, \tag{2.86}$$

is stationary for the actual physical trajectory $q(t)$. This means that the variation of $\mathcal{A}$,

$$\delta \mathcal{A}[q] = \delta \int_a^b L\big(q(t), \dot{q}(t)\big) \, dt, \tag{2.87}$$

vanishes for the trajectory $q(t)$ that is actually taken by the system, i.e., the trajectory $q(t)$ that fulfils the Euler-Lagrange equations (2.85).

Starting from this description, the generalisation of classical mechanics to field theories is mostly straight forward. The Lagrangian generalises from a function of position q and velocity $\dot{q}$ to a function of the coordinates x^μ (independent variables; most often spacetime), the fields $\varphi^a(x)$ (dependent variables) and their derivatives with respect to the

---

[2] Historical notes according to Holm [55].

[3] The isoperimetric problem asks, among all closed surfaces of a given fixed perimeter in the plane, which curve maximises the area that it encloses? Lagrange sent his solution to this problem to Euler in 1755.

[4] Quite often Hamilton's principle is called "principle of least action", which is misleading. In fact, the action does not need to take a minimum but just a critical point. For the derivation of the equations of motion, it doesn't make a difference if the critical point is a minimum, a maximum or a saddle point. Admittedly, most often it is indeed a minimum, but there are counter-examples as well (e.g., under certain conditions the action of the harmonic oscillator takes neither a minimum nor a maximum).

coordinates. The envisaged applications are all first order theories, i.e., their Lagrangians depend only on first order derivatives of the fields, $\varphi_\mu^a = \partial\varphi^a/\partial x^\mu$, and are thus of the form

$$L = L\big(x^\mu, \varphi^a(x), \varphi_\mu^a(x)\big). \tag{2.88}$$

In classical field theory, the Lagrangian density $\mathcal{L}$ is often preferred over the Lagrangian function L, as it allows for more general notation. The connection between the two,

$$\mathcal{L} = L\,\omega, \tag{2.89}$$

is drawn by the volume form ω of the base space (e.g., $\omega = dt \wedge dx \wedge dy \wedge dz$ for spacetime). Finally, the action becomes an integral not only over time but over the whole base space $\mathcal{X}$,

$$\mathcal{A} = \int_{\mathcal{X}} \mathcal{L}(x^\mu, \varphi^a, \varphi_\mu^a). \tag{2.90}$$

In the case of spacetime this is

$$\mathcal{A} = \int_{\mathcal{X}} L\big(x^\mu, \varphi^a, \varphi_\mu^a\big)\, dt\, dx\, dy\, dz. \tag{2.91}$$

In this section, a thorough derivation of the Euler-Lagrange equations based on Hamilton's action principle is presented. At first from an analytic point of view, utilising the notion of one-parameter families of transformations, thereby staying close to what is usually taught at university classes in classical mechanics. This is followed by a presentation from a geometric point of view, namely on tangent and cotangent bundles, which is a much more natural description of the problem. After a short comment on the phasespace Lagrangian, a popular object in the description of reduced kinetic theories in plasma physics, the theory on jet bundles is outlined. It allows to unify the theory for finite dimensional and infinite dimensional systems in one single framework and has many advantages over the tangent bundle theory. Nevertheless, we also include the tangent bundle theory, which has the two-fold purpose of a gentle introduction to abstract formalism and a useful tool for problems where the full fledged framework of jet bundles is not needed (e.g., in large parts of our treatment of particle dynamics). Finally, it is shown how to find what we call extended Lagrangians for systems that do not posses a classical Lagrangian, as it is often the case in plasma physics. The geometric point of view as it is stressed in this section is essential in the derivation of the variational integrators and the analysis of symmetries and conservation properties.

2.2.1. Hamilton's Action Principle

Hamilton's action principle answers the following question: of all possible paths $q(t)$ a system could choose to go from $q(t_1)$ to $q(t_2)$, which is the one it actually picks?[5] Inserting a given path $q(t)$ into the Lagrangian $L(q(t), \dot{q}(t))$, it becomes a function of time only. This time dependency is different for all the possible $q(t)$, such that the integral

$$\int_{t_1}^{t_2} L\big(q(t), \dot{q}(t)\big)\, dt \tag{2.92}$$

---

[5] The following derivation follows along the lines of Saletan and Cromer [107] and Jose and Saletan [61].
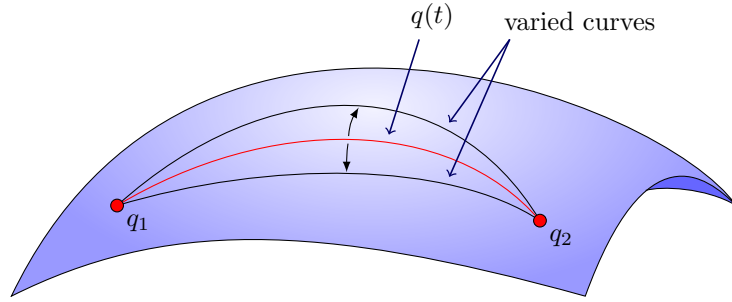
Fig. 2.3.: Variations of the trajectory $q(t)$.

takes different values for different paths $q(t)$. Hamilton's principle states that this integral takes a critical point for the physical path.

To make calculations tractable let us not consider all possible paths from $q(t_1)$ to $q(t_2)$ but a family in which each path is determined by a parameter ϵ. This family shall contain the actual, physical path for $\epsilon = 0$. Each path is a function of time, labelled by ϵ,

$$q_\epsilon(t) \equiv q(t, \epsilon) \qquad \text{with fixed } \epsilon. \tag{2.93}$$

The function $q_\epsilon(t)$ shall be differentiable for both, t at fixed ϵ and ϵ at fixed t, such that mixed partial derivatives can be exchanged

$$\frac{\partial^2 q_\epsilon}{\partial t \, \partial \epsilon} = \frac{\partial^2 q_\epsilon}{\partial \epsilon \, \partial t}. \tag{2.94}$$

All paths shall start at $q(t_1)$ and end at $q(t_2)$, such that

$$\begin{aligned} q_\epsilon(t_1) &= q_0(t_1) = q(t_1) \equiv q_1 \\ q_\epsilon(t_2) &= q_0(t_2) = q(t_2) \equiv q_2 \end{aligned} \tag{2.95}$$

or

$$\frac{\partial q_\epsilon}{\partial \epsilon}(t_1) = \frac{\partial q_\epsilon}{\partial \epsilon}(t_2) = 0. \tag{2.96}$$

One example of such a family, often considered exclusively in classical mechanics textbooks, is

$$q_\epsilon(t) = q(t) + \epsilon \, \delta q. \tag{2.97}$$

This, however, requires that $q(t)$ takes values in a linear space, an assumption that generally cannot be made in the geometric framework on manifolds. Therefore we consider more general transformations of the form (2.93). The action integral is given by

$$\mathcal{A}[q_\epsilon] = \int_{t_1}^{t_2} L\big(q_\epsilon(t), \dot{q}_\epsilon(t)\big) \, dt, \tag{2.98}$$

and has different values for different ϵ. Hamilton's principle of stationary action states that for each one-parameter family q_ϵ that fulfils the above conditions (2.94 - 2.96), q is a critical point of the action iff

$$\frac{d}{d\epsilon}\mathcal{A}[q_\epsilon]\bigg|_{\epsilon=0} = \left[\frac{d}{d\epsilon}\int_{t_1}^{t_2} L\big(q_\epsilon(t), \dot{q}_\epsilon(t)\big)\, dt\right]\bigg|_{\epsilon=0} = 0 \qquad \textbf{(Hamilton's Action Principle)}.$$

$$(2.99)$$

This means that the time integral of the Lagrangian becomes stationary for the actual, physical motion. As t_1 and t_2 are arbitrary, this is a general statement of Newton's second law. The integration limits are fixed, so that the derivative can be moved into the integral, such that

$$\frac{d}{d\epsilon}\mathcal{A}[q_\epsilon]\bigg|_{\epsilon=0} = \int_{t_1}^{t_2} \frac{dL}{d\epsilon}\big(q_\epsilon(t), \dot{q}_\epsilon(t)\big)\bigg|_{\epsilon=0} dt = \int_{t_1}^{t_2}\left[\frac{\partial L}{\partial q}\frac{\partial q_\epsilon}{\partial \epsilon} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}_\epsilon}{\partial \epsilon}\right]\bigg|_{\epsilon=0} dt. \qquad (2.100)$$

The derivative $\dot{q}_\epsilon \equiv dq_\epsilon/dt$ is the generalised velocity along a particular trajectory that is labelled by a specific value of ϵ. This time derivative is taken for fixed ϵ, so it should better be denoted $\partial q_\epsilon/\partial t$. However, the important point is that the derivatives with respect to time t and with respect to ϵ can be exchanged as in (2.94) such that integration by parts can be performed

$$\frac{d}{d\epsilon}\mathcal{A}[q_\epsilon]\bigg|_{\epsilon=0} = \int_{t_1}^{t_2}\left[\frac{\partial L}{\partial q}\frac{\partial q_\epsilon}{\partial \epsilon} + \frac{\partial L}{\partial \dot{q}}\frac{\partial}{\partial t}\frac{\partial q_\epsilon}{\partial \epsilon}\right]\bigg|_{\epsilon=0} dt \qquad (2.101)$$

$$= \int_{t_1}^{t_2}\left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right]\frac{\partial q_\epsilon}{\partial \epsilon}\bigg|_{\epsilon=0} dt + \int_{t_1}^{t_2}\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}}\frac{\partial q_\epsilon}{\partial \epsilon}\right]\bigg|_{\epsilon=0} dt. \qquad (2.102)$$

The second integral vanished according to (2.96),

$$\left[\frac{\partial L}{\partial \dot{q}}\frac{\partial q_\epsilon}{\partial \epsilon}\right]\bigg|_{t_1}^{t_2} = 0, \qquad (2.103)$$

such that the first integral has to vanish as well. At that, it has to vanish for each and every path. And as the $\partial q_\epsilon/\partial \epsilon$ are arbitrary functions of time (only restricted in that they have to vanish at the endpoints), the expression in square brackets has to vanish[6]. Of course, this expression corresponds to the Euler-Lagrange equations

$$\frac{\partial L}{\partial q}(q, \dot{q}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) = 0 \qquad \textbf{(Euler-Lagrange Equations)}. \qquad (2.104)$$

The usual notation is retained by defining

$$\delta \equiv \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \qquad (2.105)$$

and calling δ an infinitesimal variation. The previous derivation summarises

$$\delta\mathcal{A} = \int_{t_1}^{t_2}\delta L\big(q(t), \dot{q}(t)\big)\, dt = \int_{t_1}^{t_2}\left[\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q}\right] dt = \int_{t_1}^{t_2}\left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right]\delta q\, dt = 0. \quad (2.106)$$

---

[6] For a deeper discussion of this point see Jose and Saletan [61], section 3.1, or Gelfand and Fomin [44].

By the above arguments it is clear that the variation of the time derivative of q equals the time derivative of the variation of q, i.e.,

$$\delta \dot{q} = \frac{d}{dt} \delta q, \tag{2.107}$$

an important point that is often obfuscated by oversimplification.

Hamilton's Action Principle for Fields

In the infinite dimensional case (field theory), the Lagrangian can in principle depend on the fields, their derivatives, and also the coordinates. The latter is however not the case in the envisaged applications, hence for simplicity it is not considered here. The theory on jet bundles includes this case without further complications.

With the restriction to Lagrangians that are only functions of the fields and their first derivatives, the derivation of Hamilton's action principle is not any more complicated than in the finite dimensional case (particle mechanics). Without loss of generality, consider as base space $\mathcal{X}$ only time plus one spatial dimension (t, x) and a theory of just one scalar field $\varphi(t, x)$. The Lagrangian density is thus a function

$$\mathcal{L} = \mathcal{L}\big(\varphi(t, x), \varphi_t(t, x), \varphi_x(t, x)\big) \tag{2.108}$$

and the action is given by

$$\mathcal{A}[\varphi] = \int_{\mathcal{X}} \mathcal{L}(\varphi, \varphi_t, \varphi_x) = \int_{\mathcal{X}} L(\varphi, \varphi_t, \varphi_x) \, dt \, dx. \tag{2.109}$$

For instructive reasons, all considerations in this section will be taken with respect to the Lagrangian function L. Consider a family of variations φ^ϵ of φ that is defined by

$$\varphi^\epsilon(t, x) = \varphi(t, x, \epsilon) \qquad \text{with} \qquad \varphi^0 = \varphi(t, x). \tag{2.110}$$

The variation of the action can be expressed as

$$\frac{d}{d\epsilon} \mathcal{A}[\varphi^\epsilon]\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\bigg[\int_{\mathcal{X}} L\big(\varphi^\epsilon(t, x), \varphi_t^\epsilon(t, x), \varphi_x^\epsilon(t, x)\big) \, dt \, dx\bigg]\bigg|_{\epsilon=0} \tag{2.111}$$

and Hamilton's principle of stationary action states that φ is a critical point of the action iff (2.111) vanishes for all variations φ^ϵ of φ. The differentiation is carried out under the integral to give

$$\frac{d}{d\epsilon} \mathcal{A}[\varphi^\epsilon]\bigg|_{\epsilon=0} = \int_{\mathcal{X}} \bigg[\frac{dL}{d\epsilon}\big(\varphi^\epsilon, \varphi_t^\epsilon, \varphi_x^\epsilon\big)\bigg]\bigg|_{\epsilon=0} dt \, dx \tag{2.112}$$

$$= \int_{\mathcal{X}} \bigg[\frac{\partial L}{\partial \varphi}\frac{\partial \varphi^\epsilon}{\partial \epsilon} + \frac{\partial L}{\partial \varphi_t}\frac{\partial \varphi_t^\epsilon}{\partial \epsilon} + \frac{\partial L}{\partial \varphi_x}\frac{\partial \varphi_x^\epsilon}{\partial \epsilon}\bigg]\bigg|_{\epsilon=0} dt \, dx. \tag{2.113}$$

The second and third term are integrated by parts with respect to t and x, respectively,

$$\frac{d}{d\epsilon} \mathcal{A}[\varphi^\epsilon]\bigg|_{\epsilon=0} = \int_{\mathcal{X}} \bigg[\frac{\partial L}{\partial \varphi} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \varphi_t} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \varphi_x}\bigg]\frac{\partial \varphi^\epsilon}{\partial \epsilon}\bigg|_{\epsilon=0} dt \, dx, \tag{2.114}$$

where it is assumed that the fields vanish at infinity such that the boundary terms vanish. Except for this restriction, the functions $\partial_\epsilon \varphi^\epsilon$ are arbitrary, such that the variation of the action vanishes, iff the expression in square brackets vanishes, which is what now leads to the Euler-Lagrange field equations in one spatial dimension,

$$\frac{\partial L}{\partial \varphi} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \varphi_t} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \varphi_x} = 0 \qquad \text{(\textbf{Euler-Lagrange Field Equations})}. \qquad (2.115)$$

As expected, there were no surprises and the derivation was very similar to the one of the finite dimensional case, presented in section 2.2.1. The usual notation is retained by identifying

$$\delta \varphi = \frac{d}{d\epsilon} \varphi^\epsilon \bigg|_{\epsilon=0} \qquad (2.116)$$

such that the variation of the action reads

$$\delta \mathcal{A}[\varphi] = \int_{\mathcal{X}} \delta L\big(\varphi^\epsilon, \varphi_t^\epsilon, \varphi_x^\epsilon\big)\, dt\, dx \qquad (2.117)$$

and the usual manipulations amount to

$$\delta \mathcal{A}[\varphi] = \int_{\mathcal{X}} \left[\frac{\partial L}{\partial \varphi}\delta\varphi + \frac{\partial L}{\partial \varphi_t}\delta\varphi_t + \frac{\partial L}{\partial \varphi_x}\delta\varphi_x\right] dt\, dx = \int_{\mathcal{X}} \left[\frac{\partial L}{\partial \varphi} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \varphi_t} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \varphi_x}\right]\delta\varphi\, dt\, dx$$
$$(2.118)$$

with the term in square brackets corresponding to the Euler-Lagrange field equations (2.115).

2.2.2. Dynamics on the Tangent Bundle

In introductory textbooks on classical mechanics, the Lagrangian is often defined as a function on the configuration space $\mathcal{Q}$ with coordinates q, which for example might be the three-dimensional euclidean space $\mathbb{E}^3$. In that setting, the velocities $\dot{q}$ and accelerations $\ddot{q}$ correspond to the first and second time derivative of q, and if q is a vector in $\mathbb{E}^3$, they are as well. The Euler-Lagrange equations are second order differential equations.

This setting, however, does not seem natural. The Lagrangian is defined with respect to q and its first time derivative $\dot{q}$. So, strictly speaking, L is not a function on $\mathcal{Q}$ but on a larger space.

Let us take a step back and ask what determines the state of a system. It is not just the position q of all its constituents, but also their respective velocities $\dot{q}$. So the state of a system corresponds to a point in a state space labelled by $(q, \dot{q})$[7]. When the system evolves in time, both q and $\dot{q}$ change, consequently the evolution of both, q and $\dot{q}$, has to be computed, not just the evolution of the coordinates q. It is therefore natural to define the Lagrangian on exactly this space of states. This point of view has many advantages. The obvious one is that the Euler-Lagrange equations become first order differential equations for q and $\dot{q}$. The consequence of this first order nature of the equations is a separation of the trajectories in state space. There is only one trajectory passing through each point in state space, allowing for the construction of phase portraits. These are visual

---

[7] This state space is also called *velocity phasespace* in analogy to the *phasespace* in Hamiltonian dynamics.

solutions of the dynamical equations and often useful in the analysis of a dynamical systems' behaviour.

In the next step, these ideas are translated into the geometric language of manifolds. The configuration space is regarded as a smooth manifold, still denoted $\mathcal{Q}$ and called the *configuration manifold*, with points labelled by q. The velocity phasespace corresponds to the tangent bundle $\mathsf{T}\mathcal{Q}$ of that configuration manifold $\mathcal{Q}$, called the *velocity phase manifold*, with elements labelled by (q, v)[8]. The Lagrangian therefore comes naturally as a map

$$L : \mathsf{T}\mathcal{Q} \to \mathbb{R}. \tag{2.119}$$

The advantage of this point of view might not be apparent if one just thinks in term of Euclidean spaces. It will become clearer considering a particle whose motion is constrained to the two-dimensional surface of a sphere $\mathbb{S}^2$. The velocity vector of a particle moving in $\mathbb{E}^3$ is also a vector in $\mathbb{E}^3$. In $\mathbb{S}^2$, however, the velocity vector of a particle is tangent to the sphere. It is not contained in the sphere, but reaches out into the $\mathbb{E}^3$ in which the sphere is embedded. So to describe the particle motion on $\mathbb{S}^2$, one has to consider an embedding space. It is not possible to do that only by means of $\mathbb{S}^2$ alone. In this example, there might not be much of an issue, but in other dynamical systems the embedding space might not be so easily found and if it can be found might not have any physical meaning.[9]

Of course, $\mathsf{T}\mathcal{Q}$ is also a space embedding $\mathcal{Q}$, but in contrast to $\mathbb{S}^2$ and $\mathbb{E}^3$, there is an intrinsic relation between $\mathcal{Q}$ and $\mathsf{T}\mathcal{Q}$, given by the tangent lift as discussed in section 2.1.4, i.e., the tangent bundle $\mathsf{T}\mathcal{Q}$ is obtained from $\mathcal{Q}$ by attaching to each point $q \in \mathcal{Q}$ the tangent space $\mathsf{T}_q\mathcal{Q}$ at that point. The linear space $\mathsf{T}_q\mathcal{Q}$ contains all possible velocities at q, which are of course tangent to $\mathcal{Q}$ at that point.

The Lagrangian maps points (q, v) of $\mathsf{T}\mathcal{Q}$ to the real numbers $\mathbb{R}$. The resulting values are completely independent from the coordinates on $\mathsf{T}\mathcal{Q}$. That way, the description of the dynamics is intrinsic, independent on any particular choice of coordinate systems. Restricting the Lagrangian to solutions of the Euler-Lagrange equations (2.85), it becomes a function of $(q, \dot{q})$, as before. In the applications considered later on, it will always be possible to find a global coordinate system for the configuration manifold $\mathcal{Q}$ and its tangent bundle $\mathsf{T}\mathcal{Q}$. This allows to circumvent the issues arising in the case, when the path $q(t)$ and its deformations $q_\epsilon(t)$ are not located in a single coordinate patch[10].

Hamilton's Action Principle on the Tangent Bundle

In this section, a derivation of Hamilton's action principle in a geometric framework is presented[11]. Consider the space of paths $\mathcal{C}(\mathcal{Q})$ that connect two points in $\mathcal{Q}$,

$$\mathcal{C}(\mathcal{Q}) = \left\{ c : \mathcal{I} \to \mathcal{Q} \,\middle|\, \mathcal{I} \subset \mathbb{R} \text{ smooth and bounded} \right\}. \tag{2.120}$$

---

[8] At this point, a comment is in order. In the literature, points of $\mathsf{T}\mathcal{Q}$ are often labelled $(q, \dot{q})$. This notation implies that all curves in $\mathsf{T}\mathcal{Q}$ are the tangent lift of some curve $q(t)$ in $\mathcal{Q}$. That is of course not the case! The previous statement is only true for physical trajectories, i.e., solutions $(q, \dot{q})$ of the Euler-Lagrange equations (2.85). But there exist much more curves in $\mathsf{T}\mathcal{Q}$ for which $v \neq \dot{q}$.

[9] See Jose and Saletan [61], section 2.4, for a more detailed discussion. This example is taken from there.

[10] For a discussion of these issues see Holm et al. [57], section 4.1.

[11] The following derivation follows along the lines of Marsden and Ratiu [82] and Marsden and West [83]. For proofs of some of the statements have a look at [82], section 8.1.

Fixing two points q_1 and q_2 in $\mathcal{Q}$ as well as an interval $[t_1, t_2]$, the path space from q_1 to q_2 is defined as

$$\mathcal{C}(q_1, q_2, [t_1, t_2]) = \big\{ c : [t_1, t_2] \to \mathcal{Q} \,\big|\, c(t_1) = q_1, c(t_2) = q_2 \big\} \subset \mathcal{C}(\mathcal{Q}). \tag{2.121}$$

Elements c of $\mathcal{C}(q_1, q_2, [t_1, t_2])$ are maps that relate points q in configuration space $\mathcal{Q}$ to points t in the time interval $[t_1, t_2]$, whereby the first and last points, $c(t_1)$ and $c(t_2)$, take fixed values, q_1 and q_2, respectively. Consequently, the action can be written as a map $\mathcal{A} : \mathcal{C}(q_1, q_2, [t_1, t_2]) \to \mathbb{R}$ assigning real values to each path c,

$$\mathcal{A}[c] = \int_{t_1}^{t_2} L\big(c(t), \dot{c}(t)\big) \, dt. \tag{2.122}$$

If L is the Lagrangian on $\mathsf{T}\mathcal{Q}$, and c is a path $c : [t_1, t_2] \to \mathcal{Q}$ that connects $q_1 = c(t_1)$ with $q_2 = c(t_2)$, Hamilton's principle of stationary action states that c obeys the Euler-Lagrange equations (2.85), iff c is a critical point of the function $\mathcal{A} : \mathcal{C}(q_1, q_2, [t_1, t_2]) \to \mathbb{R}$, that is $\mathcal{A}$ is stationary for c or $\delta\mathcal{A}[c] = 0$. Stationarity of $\mathcal{A}[c]$ means that $\mathcal{A}[c]$ does not change under infinitesimal variations of the path c. Such infinitesimal variations of c live in the tangent space $\mathsf{T}_c\mathcal{C}(q_1, q_2, [t_1, t_2])$ of $\mathcal{C}(q_1, q_2, [t_1, t_2])$ at c. The tangent vector to some path $c_\epsilon \in \mathcal{C}(q_1, q_2, [t_1, t_2])$ is given by

$$V(t) = \frac{d}{d\epsilon} c_\epsilon(t) \Big|_{\epsilon=0}. \tag{2.123}$$

For each fixed t, c_ϵ is a curve in $\mathcal{Q}$ through the point $c(t)$, such that $V(t)$ is a tangent vector to $\mathcal{Q}$ based at $c(t)$, i.e., $V(t) \in \mathsf{T}_{c(t)}\mathcal{Q}$ and thus $\pi_\mathcal{Q} \circ V = c$, where $\pi_\mathcal{Q}$ is the canonical projection $\pi_\mathcal{Q} : \mathsf{T}\mathcal{Q} \to \mathcal{Q}$. From the restrictions $c_\epsilon(t_1) = q_1$ and $c_\epsilon(t_2) = q_2$ follows that $V(t_1) = 0$ and $V(t_2) = 0$, but otherwise V is an arbitrary function. To summarise, the *infinitesimal variation* of a path $c : [t_1, t_2] \to \mathcal{Q}$, is the set of maps

$$V : [t_1, t_2] \to \mathsf{T}\mathcal{Q} \qquad \text{for which} \qquad \pi_\mathcal{Q} \circ V = c \qquad \text{and} \qquad V(t_1) = V(t_2) = 0. \tag{2.124}$$

V is called an infinitesimal variation of the path c with fixed endpoints and naturally denoted $V = \delta c$. With the chain rule one obtains

$$\delta\mathcal{A}[c] = \frac{d}{d\epsilon} \mathcal{A}[c_\epsilon] \Big|_{\epsilon=0} = \frac{\partial\mathcal{A}}{\partial c}[c] \cdot \frac{dc_\epsilon}{d\epsilon} \Big|_{\epsilon=0} = \mathbf{d}\mathcal{A}[c] \cdot V, \tag{2.125}$$

where $\mathbf{d}\mathcal{A}[c]$ and $V = dc_\epsilon/d\epsilon|_{\epsilon=0}$ are regarded as elements of the cotangent and tangent spaces $\mathsf{T}_c^*\mathcal{C}$ and $\mathsf{T}_c\mathcal{C}$ on the manifold $\mathcal{C}$, respectively. Therefore the variation of the action can be formulated as

$$\mathbf{d}\mathcal{A}[c] \cdot V = \frac{d}{d\epsilon} \left[\int_{t_1}^{t_2} L\big(c_\epsilon(t), \dot{c}_\epsilon(t)\big) \, dt \right] \Big|_{\epsilon=0}, \tag{2.126}$$

where L is a function of the tangent lift of c_ϵ. Computation of the derivative under the integral,

$$\mathbf{d}\mathcal{A}[c] \cdot V = \int_{t_1}^{t_2} \frac{d}{d\epsilon} \left[L\big(c_\epsilon(t), \dot{c}_\epsilon(t)\big) \right] \Big|_{\epsilon=0} dt = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} \cdot \frac{dc_\epsilon}{d\epsilon} \Big|_{\epsilon=0} + \frac{\partial L}{\partial v} \cdot \frac{d\dot{c}_\epsilon}{d\epsilon} \Big|_{\epsilon=0} \right] dt, \tag{2.127}$$

leads to

$$\mathbf{d}\mathcal{A}[c] \cdot V = \int\limits_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} \cdot V + \frac{\partial L}{\partial v} \cdot \dot{V} \right] dt = \int\limits_{t_1}^{t_2} \mathbf{d}L\big(c(t), \dot{c}(t)\big) \cdot \mathsf{T}V(t)\, dt, \tag{2.128}$$

with $(V, \dot{V})$ the coordinates of the tangent lift of V,

$$\mathsf{T}V : (q, V) \mapsto \big((q, \dot{q}), (V, \dot{V})\big). \tag{2.129}$$

The second term can be integrated by parts, leading to

$$\mathbf{d}\mathcal{A}[c] \cdot V = \int\limits_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial v} \right] \cdot V\, dt + \left[\frac{\partial L}{\partial v} \cdot V \right]_{t_1}^{t_2}, \tag{2.130}$$

where the second term vanishes as V vanishes on both ends of the trajectory (2.124), such that

$$\mathbf{d}\mathcal{A}[c] \cdot V = \int\limits_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial v} \right] \cdot V\, dt = \int\limits_{t_1}^{t_2} D_{\mathrm{EL}}L\big(\ddot{c}(t)\big) \cdot V(t)\, dt, \tag{2.131}$$

where

$$D_{\mathrm{EL}}L(c) : \ddot{\mathcal{Q}} \to \mathsf{T}^*\mathcal{Q} \tag{2.132}$$

is an one-form valued function, defining the Euler-Lagrange operator, and $\ddot{\mathcal{Q}}$ is a submanifold of $\mathsf{T}(\mathsf{T}\mathcal{Q})$, such that

$$\ddot{\mathcal{Q}} = \big\{ w \in \mathsf{T}(\mathsf{T}\mathcal{Q}) \,\big|\, \mathsf{T}\pi_{\mathcal{Q}}(w) = \pi_{\mathsf{T}\mathcal{Q}}(w) \big\} \subset \mathsf{T}(\mathsf{T}\mathcal{Q}). \tag{2.133}$$

In other words, $\ddot{\mathcal{Q}}$ is the set of second derivatives $\ddot{c}(0)$ of curves $c : \mathcal{I} \to \mathcal{Q}$, which are of the form $\big((q, \dot{q}), (\dot{q}, \ddot{q})\big) \in \mathsf{T}(\mathsf{T}\mathcal{Q})$. The requirement $\delta\mathcal{A}[c] = 0$ is equivalent to $\mathbf{d}\mathcal{A}[c] \cdot V = 0$ for all $V \in \mathsf{T}_c\mathcal{C}(q_1, q_2, [t_1, t_2])$ as well as to the Euler-Lagrange equations (2.85), that now are rewritten $D_{\mathrm{EL}}L\big(c(t)\big) = 0$, as V is arbitrary, except for it has to vanish at the end points of the trajectory. The covariance of the Euler-Lagrange equation, first observed in the original work of Lagrange, is obtained here as a natural consequence of the geometric framework.

Phasespace Lagrangian

In plasma physics, there exists another notation that enjoys a certain prevalence, namely that of the *phasespace Lagrangian* [77]. In this formulation, the Lagrangian is not defined as a function on the tangent bundle $\mathsf{T}\mathcal{Q}$ of the configuration space $\mathcal{Q}$, but instead the tangent bundle takes the role of the configuration space, such that the Lagrangian is defined on $\mathsf{T}(\mathsf{T}\mathcal{Q})$.

In practice, the tangent bundle structure of the configuration space $\mathcal{Z} \cong \mathsf{T}\mathcal{Q}$ is neglected, and the Lagrangian is defined as a function on $\mathsf{T}\mathcal{Z} \cong \mathsf{T}(\mathsf{T}\mathcal{Q})$. Everything else then follows in a straight forward way.

2.2.3. Dynamics on the Jet Bundle

Another view is offered by employing jet bundle theory.[12] Rewriting the theory in this framework might at first sight seem to unnecessarily complicate things, but its great advantage is that it offers a concise notation that readily generalises to the case of field theories. That way, jet bundle theory offers a general formulation of the variational problem that applies to finite as well as infinite dimensional systems. Besides, this formulation is fully covariant, the analysis of symmetries with Noether's theorem is simpler, and it resembles the discrete setting quite nicely.

The idea of a jet is to combine the independent variables (coordinates), the dependent variables (trajectories, fields) and their partial derivatives up to a given order in one single geometric object. Jets provide a coordinate-free description of differential equations, which is especially useful in the theory of partial differential equations, where they allow us to represent an infinite-dimensional space of maps by sections of a finite dimensional space of jets, thereby avoiding the intricacies of infinite dimensional manifolds.

Consider a function $\varphi(x)$. It establishes a correspondence between each value $x \in \mathcal{X}$ and another value $\varphi(x) \in \mathcal{F}$. This second value $\varphi(x)$ can be considered as a point in the fibre $\mathcal{F}$ above x. So it seems natural to construct a fibre bundle $\mathcal{Y}$ over the base manifold $\mathcal{X}$ with fibres corresponding to $\mathcal{F}$. In other words, $\mathcal{Y}$ is obtained by attaching a fibre $\mathcal{F}$ to each point $x \in \mathcal{X}$, such that the fibres of $\mathcal{Y}$ contain all possible values of functions $\varphi(x)$, which can therefore be considered as sections φ in the bundle $\mathcal{Y}$,

$$\varphi : \mathcal{X} \to \mathcal{Y} \qquad \text{with} \qquad \pi_{\mathcal{X}\mathcal{Y}} \circ \varphi = \mathrm{id}_{\mathcal{X}}, \tag{2.134}$$

where $\pi_{\mathcal{X}\mathcal{Y}}$ is the canonical projection

$$\pi_{\mathcal{X}\mathcal{Y}} : \mathcal{Y} \to \mathcal{X}. \tag{2.135}$$

If (x^μ, y^a) are coordinates on $\mathcal{Y}$, a section φ is a map $x \mapsto (x^\mu, \varphi^a(x))$, where we denote by φ^a the vertical components of φ, i.e., the fibre coordinates of $\varphi(x)$. In this setting, the equivalent to the tangent bundle is the first jet bundle $\mathrm{J}^1\mathcal{Y}$, which contains the first order partial derivatives of each section $\varphi \in \mathcal{Y}$. In the same way, the kth jet bundle $\mathrm{J}^k\mathcal{Y}$ of $\mathcal{Y}$ is the space that contains the partial derivatives of each section $\varphi \in \mathcal{Y}$ up to order k. However, in what follows only the first order jet bundle $\mathrm{J}^1\mathcal{Y}$ is needed, so all considerations are restricted to that case.

Coordinates on $\mathrm{J}^1\mathcal{Y}$ are (x^μ, y^a, v^a_μ), where x^μ are the coordinates of the base manifold $\mathcal{X}$, y^a are the values of fields at x, and v^a_μ are all possible values of the partial derivative of y^a with respect to x^μ. $\mathrm{J}^1\mathcal{Y}$ has two natural projections. It can be viewed as a fibre bundle over $\mathcal{X}$ with the source projection

$$\pi_{\mathcal{X},\mathrm{J}^1\mathcal{Y}} : \mathrm{J}^1\mathcal{Y} \to \mathcal{X} \tag{2.136}$$

as well as a fibre bundle over $\mathcal{Y}$ with the target projection

$$\pi_{\mathcal{Y},\mathrm{J}^1\mathcal{Y}} : \mathrm{J}^1\mathcal{Y} \to \mathcal{Y}. \tag{2.137}$$

---

[12] The derivations of this section follow along the lines of Gotay et al. [49], Marsden et al. [84, 85], Kouranbaeva and Shkoller [65], Kouranbaeva [64] and West [135].

$$
\begin{array}{c}
\mathcal{Y} \\
\end{array}
$$
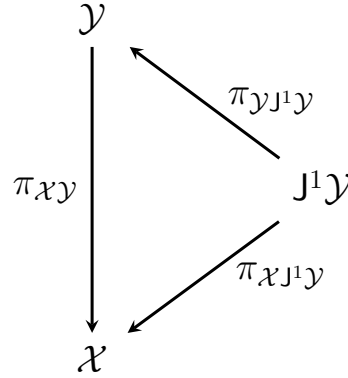
The first point of view is especially important. Consider a section $\varphi : \mathcal{X} \to \mathcal{Y}$ of $\pi_{\mathcal{X}\mathcal{Y}}$. Its tangent map $\mathsf{T}_x\varphi$ at $x \in \mathcal{X}$ is represented by the matrix $\partial\varphi^a(x)/\partial x^\mu = \varphi_\mu^a(x)$ and thus can be identified with an element of $\mathsf{J}^1_{\varphi(x)}\mathcal{Y}$. The map $x \mapsto \mathsf{T}_x\varphi$ is therefore a section of $\pi_{\mathcal{X},\mathsf{J}^1\mathcal{Y}}$, i.e., a section of $\mathsf{J}^1\mathcal{Y}$ regarded as a bundle over $\mathcal{X}$. This section, denoted $j^1\varphi$, is called the *first jet prolongation*[13] (also *canonical prolongation*) of a section $\varphi(x)$,

$$
j^1\varphi : \mathcal{X} \to \mathsf{J}^1\mathcal{Y} \qquad \text{in coordinates} \qquad j^1\varphi : x \mapsto \big(x^\mu, \varphi^a(x), \varphi_\mu^a(x)\big). \tag{2.138}
$$

Such sections $j^1\varphi$ of $\mathsf{J}^1\mathcal{Y}$ that correspond to the canonical prolongation[14] of a section $\varphi \in \mathcal{Y}$ are called holonomic. For them v_μ^a can be identified with $\varphi_\mu^a = \partial\varphi^a/\partial x^\mu$.

In this setting, a section $j^1\varphi$ of $\pi_{\mathcal{X},\mathsf{J}^1\mathcal{Y}}$ generalises the notion of a trajectory and a field. The *Lagrangian density* $\mathcal{L}$ is a n-form on the jet bundle $\mathsf{J}^1\mathcal{Y}$,

$$
\mathcal{L} : \mathsf{J}^1\mathcal{Y} \to \Omega^n(\mathcal{X}), \tag{2.139}
$$

where $\Omega^n(\mathcal{X})$ denotes the n-forms on $\mathcal{X}$. The Lagrangian L is a function on the jet bundle $\mathsf{J}^1\mathcal{Y}$,

$$
L : \mathsf{J}^1\mathcal{Y} \to \mathbb{R}. \tag{2.140}
$$

The connection between the two is drawn by the volume form ω of the base manifold $\mathcal{X}$,

$$
\mathcal{L} = L\omega. \tag{2.141}
$$

Here, n is the dimension of the base space $\mathcal{X}$ and $\omega = dx^1 \wedge dx^2 \wedge ... \wedge dx^n$, e.g., for $\mathcal{X}$ corresponding to spacetime we have $\omega = dt \wedge dx \wedge dy \wedge dz$.

Hamilton's Action Principle on the Jet Bundle

In the framework of jet bundles, the action is given as the integral of the pullback of the Lagrangian density $\mathcal{L}$ with the first jet prolongation $j^1\varphi$ of a section $\varphi : \mathcal{X} \to \mathcal{Y}$

$$
\mathcal{A}[\varphi] = \int_{\mathcal{X}} (j^1\varphi)^* \mathcal{L}. \tag{2.142}
$$

---

[13] The jet prolongation can be seen as producing a coordinate-free Taylor expansion to first (in general kth) order, as the jet bundle $\mathsf{J}^1\mathcal{Y}$ contains all functions that have the same Taylor series up to the first term.

[14] Not all sections of $\mathsf{J}^1\mathcal{Y}$ are prolongations of a section $\varphi \in \mathcal{Y}$.

As $\mathcal{L} = L\,\omega$ and L is a smooth function, the following expressions are equivalent

$$(j^1\varphi)^*\mathcal{L} = L(j^1\varphi)\,\omega. \tag{2.143}$$

Writing the action (2.142) with respect to the last expression and in coordinates

$$\mathcal{A}[\varphi] = \int_{\mathcal{X}} L(j^1\varphi)\,\omega = \int_{\mathcal{X}} L(x^\mu, \varphi^a, \varphi^a_\mu)\,\omega \tag{2.144}$$

establishes a correspondence between (2.142) and previous formulation (2.122). Hamilton's principle states that φ is a critical point of the action iff

$$\frac{d}{d\epsilon}\mathcal{A}[\varphi_\epsilon]\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\left[\int_{\mathcal{X}} (j^1\varphi_\epsilon)^*\mathcal{L}\right]\bigg|_{\epsilon=0} = 0 \tag{2.145}$$

for all variations φ_ϵ of φ. These variations are defined as a composition of the trajectory φ and the vertical transformation η_ϵ of the underlying fibre bundle, namely,

$$\varphi_\epsilon = \eta_\epsilon \circ \varphi = \eta_\epsilon(\varphi) \tag{2.146}$$

such that the variational vector field $V : \mathcal{X} \to \mathsf{T}\mathcal{Y}$ is defined as

$$V = \frac{d}{d\epsilon}(\eta_\epsilon \circ \varphi)\bigg|_{\epsilon=0} = \frac{d\eta_\epsilon}{d\epsilon}(\varphi)\bigg|_{\epsilon=0}, \tag{2.147}$$

or explicitly

$$V : x \mapsto \left(\big(x, \varphi^a(x)\big), \big(0, V^a_\eta\big)\right) \tag{2.148}$$

where V_η is the generating vector field of the transformation η_ϵ with components

$$V^a_\eta = \frac{d}{d\epsilon}\eta^a_\epsilon\bigg|_{\epsilon=0}, \tag{2.149}$$

η^a_ϵ being the y^a component of η_ϵ. For the moment, we are considering only vertical transformations as that is sufficient for the derivation of the Euler-Lagrange equations, but the Euler-Lagrange equations are also obtained for general variations, not necessarily of the form $\eta_\epsilon \circ \varphi$.

The flow map η_ϵ can be interpreted as dragging the path φ along V_η through the configuration space. From now on we drop the η index on the field components of the generating vector field. As we do not consider transformations in the coordinates this is no origin of confusion. The jet prolongation of V to $\mathsf{J}^1\mathcal{Y}$ is given by

$$j^1V = \frac{d}{d\epsilon}j^1(\eta_\epsilon \circ \varphi)\bigg|_{\epsilon=0} \tag{2.150}$$

or in coordinates

$$j^1V : x \mapsto \left(\big(x^\mu, \varphi^a(x), \varphi^a_\nu(x)\big), \big(0, V^a, V^a_\nu + V^a_b\,\varphi^b_\nu\big)\right). \tag{2.151}$$

With this and

$$j^1\varphi_\epsilon = j^1(\eta_\epsilon \circ \varphi) = j^1\eta_\epsilon \circ j^1\varphi \tag{2.152}$$

such that

$$(j^1\varphi_\epsilon)^*\mathcal{L} = (j^1\varphi)^*(j^1\eta_\epsilon)^*\mathcal{L} \tag{2.153}$$

the action principle (2.145) can be rewritten as

$$\frac{d}{d\epsilon}\mathcal{A}[\varphi_\epsilon]\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\bigg[\int_{\mathcal{X}} (j^1\varphi)^*(j^1\eta_\epsilon)^*\mathcal{L}\bigg]\bigg|_{\epsilon=0} = 0. \tag{2.154}$$

With the dynamical definition of the Lie derivative (2.82),

$$\pounds_{j^1V}\mathcal{L} = \frac{d}{d\epsilon}\big[(j^1\eta_\epsilon)^*\mathcal{L}\big]\bigg|_{\epsilon=0}, \qquad \text{and} \qquad \mathbf{d}\mathcal{A}\cdot V = \frac{d}{d\epsilon}\mathcal{A}[\varphi_\epsilon]\bigg|_{\epsilon=0}, \tag{2.155}$$

(2.154) becomes a beautiful, general, geometric formulation of Hamilton's action principle

$$\mathbf{d}\mathcal{A}[\varphi]\cdot V = \int_{\mathcal{X}} (j^1\varphi)^*(\pounds_{j^1V}\mathcal{L}) = 0 \qquad (\textbf{Hamilton's Action Principle}). \tag{2.156}$$

This form of the action principle has several advantages. First and most importantly, this equation is the very same for particles as well as for fields. Second, it makes explicit the use of the jet prolongation of the trajectory φ to $j^1\varphi$ and the variational vector field V to j^1V, whereas in the tangent bundle formulation, the tangent lift of φ and V is not explicit in the notation.

Last but not least, it is not too difficult to generalise from variations in the configuration space to variations in the full jet space. Thereby not only considering vertical variations, but general variations that might have both horizontal and vertical components.

Coming back to the derivation of the Euler-Lagrange equations, Cartan's magic formula

$$\pounds_{j^1V}\mathcal{L} = i_{j^1V}\mathbf{d}\mathcal{L} + \mathbf{d}(i_{j^1V}\mathcal{L}) \tag{2.157}$$

needs to be employed to give

$$\mathbf{d}\mathcal{A}[\varphi]\cdot V = \int_{\mathcal{X}} (j^1\varphi)^*(i_{j^1V}\mathbf{d}\mathcal{L}) + \int_{\mathcal{X}} \mathbf{d}\big((j^1\varphi)^*(i_{j^1V}\mathcal{L})\big) = 0 \tag{2.158}$$

where in the second integral we used that the pullback and the exterior derivative commute. The second integral vanishes due to Stokes' theorem and the assumption that the variations of φ vanish at the boundary $\partial\mathcal{X}$. Therefore, what is left is just

$$\mathbf{d}\mathcal{A}[\varphi]\cdot V = \int_{\mathcal{X}} (j^1\varphi)^*(i_{j^1V}\mathbf{d}\mathcal{L}) = 0. \tag{2.159}$$

This expression will be the basis for deriving the actual Euler-Lagrange equations in the jet bundle framework for both, particle mechanics and field theory, below.

Classical Mechanics on Jet Bundles

In classical mechanics, the base manifold is just time, $\mathcal{X} = \mathbb{R}$, with coordinates t. $\mathcal{Y}$ is a fibre bundle over time, with the fibres $\mathcal{Y}_t$ corresponding to the configuration space $\mathcal{Q}$, elements labelled by q and coordinates (t, q), i.e., time and the generalised coordinates.

The fibres $J^1_{(t,q)}\mathcal{Y}$ of the first jet bundle $\mathsf{J}^1\mathcal{Y}$ contain the time derivatives of all sections $c : \mathcal{X} \to \mathcal{Y}$. Its coordinates are (t, q, v). The jet prolongation is given in coordinates by

$$j^1 c : (t, c) \mapsto (t, c, \dot{c}). \tag{2.160}$$

Observe that $\mathsf{J}^1\mathcal{Y}$ can be identified with (is isomorphic to) $\mathbb{R} \times \mathsf{T}\mathcal{Q}$, sections of $\mathcal{Y}$ correspond to trajectories $q(t)$ in $\mathcal{Q}$, sections of $\mathsf{J}^1\mathcal{Y}$ to trajectories $(q(t), v(t))$ in $\mathsf{T}\mathcal{Q}$, and that the jet prolongation is analogous to the tangent lift, such that holonomic sections $j^1 c$ of $\mathsf{J}^1\mathcal{Y}$ are solutions $(q(t), \dot{q}(t))$ of the Euler-Lagrange equations. In classical mechanics the volume form is just $\omega = dt$. This implies that the Lagrangian is a function

$$L : \mathsf{J}^1\mathcal{Y} \to \mathbb{R}. \tag{2.161}$$

The coordinate expression of $\mathcal{L}$ is

$$\mathcal{L} = L(q, v) \, dt, \tag{2.162}$$

where we are considering a time-independent Lagrangian L. Starting from (2.159), compute the exterior derivative, contract with $j^1 V$ and do the usual partial integration

$$\mathbf{d}\mathcal{A}[c] \cdot V = \int_\mathcal{X} \left[\frac{\partial L}{\partial q}(j^1 c) \cdot V + \frac{\partial L}{\partial v}(j^1 c) \cdot \dot{V} \right] dt = \int_\mathcal{X} \left[\frac{\partial L}{\partial q}(j^1 c) - \frac{d}{dt}\frac{\partial L}{\partial v}(j^1 c) \right] \cdot V \, dt = 0. \tag{2.163}$$

The usual arguments then again yield the Euler-Lagrange equations

$$\frac{\partial L}{\partial q}(j^1 c) - \frac{d}{dt}\frac{\partial L}{\partial v}(j^1 c) = 0. \tag{2.164}$$

Note that in the jet bundle framework, the case of an explicit time dependency of the Lagrangian is automatically included.

Field Theory on Jet Bundles

In field theory, the base manifold $\mathcal{X}$ is usually identified with spacetime. Its points are denoted x and its coordinates are (t, x, y, z) abbreviated as x^ν. $\mathcal{Y}$ is thus a fibre bundle over spacetime with coordinates (x^ν, y^a), where y^a are the different fields or field components of the theory, and the first jet bundle $\mathsf{J}^1\mathcal{Y}$ has coordinates (x^ν, y^a, v^a_ν). Hence, the Lagrangian density is a function $\mathcal{L}(x^\nu, y^a, v^a_\nu)$.

Considering a field theory of a (possibly vector valued) field $\varphi(x) : \mathcal{X} \to \mathcal{Y}$, defined over spacetime, one can directly start from the action principle as formulated in equation (2.156)

$$\mathbf{d}\mathcal{A}[\varphi] \cdot V = \int_\mathcal{X} (j^1\varphi)^* (\pounds_{j^1 V}\mathcal{L}) = 0, \tag{2.165}$$

as all considerations that lead to this equation were completely general. All the hard work of section 2.2.3 is paying off now. The coordinate expressions of φ and V and their jet prolongations $j^1\varphi$ and $j^1 V$ are

$$\varphi : x \mapsto (x^\mu, \varphi^a), \qquad\qquad j^1\varphi : x \mapsto (x^\mu, \varphi^a, \varphi^a_\mu), \tag{2.166a}$$

$$V : x \mapsto ((x^\mu, \varphi^a), (0, V^a)), \qquad j^1 V : x \mapsto ((x^\mu, \varphi^a, \varphi^a_\mu), (0, V^a, V^a_\mu)). \tag{2.166b}$$

Inserting this into (2.165) and making the exterior derivative, the contraction and the pullback explicit gives

$$\mathbf{d}\mathcal{A}[\varphi] \cdot V = \int_{\mathcal{X}} \left[\frac{\partial L}{\partial y^a}(j^1\varphi) \, V^a + \frac{\partial L}{\partial v_\nu^a}(j^1\varphi) \, V_\nu^a \right] \omega \tag{2.167}$$

$$= \int_{\mathcal{X}} \left[\frac{\partial L}{\partial y^a}(j^1\varphi) - \frac{\partial}{\partial x^\nu} \frac{\partial L}{\partial v_\nu^a}(j^1\varphi) \right] V^a \, \omega \tag{2.168}$$

which leads to the Euler-Lagrange field equations for a theory of a field φ on spacetime

$$\frac{\partial L}{\partial y^a}(j^1\varphi) - \frac{\partial}{\partial x^\nu} \frac{\partial L}{\partial v_\nu^a}(j^1\varphi) = 0 \qquad \textbf{(Euler-Lagrange Field Equations)}. \tag{2.169}$$

2.2.4. Variational Route to the Cartan Form

In this section we want to describe a variational derivation of the Cartan form, one of the two fundamental geometric structures of classical mechanics and classical field theories (the other one being the (multi)symplectic form, covered in the next section).

In most treatments, the Cartan form and the multisymplectic form are constructed by using the Legendre transformation to pull back the canonical forms from the Hamiltonian side (cotangent bundle) to the Lagrangian side (tangent bundle). However, it has been shown by Marsden et al. [84] that the Cartan form arises naturally in the boundary term of the variation of the action in Hamilton's action principle, thus allowing to obtain these structures while staying on the Lagrangian side, entirely. The advantage of this approach is the possibility of a geometric treatment of theories for which a Hamiltonian cannot be defined. This is especially important in the light of extended Lagrangians as they will be introduced in section 2.2.7.

After a short look at the Cartan one-form in the tangent bundle setting, which is restricted to autonomous systems of classical mechanics[15], we will generalise the derivation to jet bundles and obtain an expression of the Cartan form that is valid for autonomous as well as non-autonomous systems of classical mechanics and also field theories.

Lagrangian One- and Two-Form

Besides leading to the equations of motion, the variational principle provides a direct and natural way to derive the fundamental geometric structures of classical mechanics[16]. For this derivation, the boundary conditions $\delta q(t_1) = \delta q(t_2) = 0$ are removed, while the time interval is kept fixed. Thus the variational principle reads

$$\mathbf{d}\mathcal{A}[q(t)] \cdot \delta q(t) = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial v} \right] \cdot \delta q \, dt + \left[\frac{\partial L}{\partial v} \cdot \delta q \right]_{t_1}^{t_2} \tag{2.170}$$

---

[15] It is possible to derive the Cartan one-form for non-autonomous systems of classical mechanics in the tangent bundle framework (see Marsden and West [83], section 4.2), but it is not much clearer than the more general derivation on jet bundles.

[16] The following derivation follows along the lines of Marsden et al. [84], section 2, and Marsden and Ratiu [82], section 8.2.

where the variations δq do not vanish at the boundary point, so that the last term on the right hand side does not vanish. This last term corresponds to a linear pairing of the function $\partial L/\partial v$, which is a function of $(q, \dot{q})$, with the tangent vector δq. This term can be regarded as a one-form on $\mathsf{T}\mathcal{Q}$ [17], referred to as the *Lagrangian one-form* or *Cartan one-form*,

$$\Theta_L = \frac{\partial L}{\partial v} \, dq. \tag{2.171}$$

This means that the Lagrangian one-form Θ_L is the boundary term of the functional derivative of the action, if the boundary is varied. The negative of the exterior derivative of the Lagrangian one-form gives the *Lagrangian two-form*, also referred to as the *symplectic two-form*

$$\Omega_L \equiv -\mathbf{d}\Theta_L, \tag{2.172}$$

given in coordinates by

$$\Omega_L = \frac{\partial^2 L}{\partial q^i \, \partial v^j} \, dq^i \wedge dq^j + \frac{\partial^2 L}{\partial v^i \, \partial v^j} \, dv^i \wedge dq^j. \tag{2.173}$$

For details on the connection between the Lagrangian one-form Θ_L on $\mathsf{T}\mathcal{Q}$ and the canonical one-form Θ on $\mathsf{T}^*\mathcal{Q}$ as well as between the Lagrangian two-form Ω_L on $\mathsf{T}\mathcal{Q}$ and the canonical symplectic two-form Ω on $\mathsf{T}^*\mathcal{Q}$ the reader is referred to Marsden and Ratiu [82].

Cartan Form and Multisymplectic Form

To derive the Cartan form in a general setting that applies to classical mechanics as well as to field theories, the action principle on the jet bundle has to be generalised a bit further[18]. Before, only vertical variations of the action were considered, thereby implicitly restricting the treatment to Lagrangians, that are not explicitly time dependent in the case of particles, or do not explicitly depend on the coordinates in the case of fields, respectively. But in order to obtain the correct Cartan form in the general case also horizontal variations need to be considered. Whereas the Euler-Lagrange equations obtained by considering vertical or arbitrary variations are the same, the Cartan form is missing one term if only vertical variations are accounted for.

Allowing also for horizontal variations brings some complications. A transformation η acting on a section $\varphi : \mathcal{U}_\mathcal{X} \to \mathcal{Y}$, defined over a bounded domain $\mathcal{U}_\mathcal{X} \subset \mathcal{X}$,

$$\eta : \big(x, \varphi(x)\big) \mapsto \big(\eta_\mathcal{X}(x), \eta_\mathcal{Y}(x, \varphi(x))\big) \tag{2.174}$$

changes not only the section φ to $\eta \circ \varphi$ but also the base space from $\mathcal{U}_\mathcal{X}$ to $\eta_\mathcal{X}(\mathcal{U}_\mathcal{X})$. We explain now how to get around this issue.

Consider a smooth manifold $\mathcal{U}$ with smooth closed boundary $\partial\mathcal{U}$. $\mathcal{U}$ shall be a parametrisation of the space $\mathcal{U}_\mathcal{X} \subset \mathcal{X}$ on which the physical sections are defined. This is similar to

---

[17] One could be tempted to regard $\partial L/\partial \dot{q}$ as a one-form on $\mathcal{Q}$ as it only has a component in dq. The same way δq could be regarded as a tangent vector on $\mathcal{Q}$. However, $\partial L/\partial \dot{q}$ is a function of $(q, \dot{q})$ and therefore clearly a function on $\mathsf{T}\mathcal{Q}$. δq can also be replaced with a more general vector $\delta \hat{q} \in \mathsf{T}(\mathsf{T}\mathcal{Q})$ that has non-vanishing components $(\delta q, \delta v)$.

[18] The following derivation follows along the lines of Marsden et al. [84, 85], Kouranbaeva and Shkoller [65], Kouranbaeva [64], chapter 4, and West [135], chapter 5,.

the previous case, where a space of curves $\mathcal{C}(\mathcal{Q})$ was defined (2.120), such that elements of $\mathcal{C}(\mathcal{Q})$ correspond to parametrisations of the physical trajectories. Thus, in total analogy we define the set of smooth maps

$$\mathcal{C}(\mathcal{Y}) = \left\{ \phi : \mathcal{U} \to \mathcal{Y} \mid \pi_{\mathcal{X}\mathcal{Y}} \circ \phi : \mathcal{U} \to \mathcal{X} \text{ is an embedding} \right\} \tag{2.175}$$

in coordinates

$$\phi : u \mapsto \left(x^\mu(u), \phi^a(u) \right) \qquad \text{with} \qquad x(u) \text{ an embedding,} \tag{2.176}$$

and $\phi^a(u)$ are the fibre coordinates of $\phi(u)$. Points in $\mathcal{X}$ and $\mathcal{U}$ are denoted x and u, respectively, and their coordinates are denoted x^μ and u^μ, respectively. For each $\phi \in \mathcal{C}(\mathcal{Y})$ define

$$\phi_{\mathcal{X}} \equiv \pi_{\mathcal{X}\mathcal{Y}} \circ \phi \qquad \text{and} \qquad \mathcal{U}_{\mathcal{X}} \equiv \pi_{\mathcal{X}\mathcal{Y}} \circ \phi(\mathcal{U}) \qquad \text{such that} \qquad \phi_{\mathcal{X}} : \mathcal{U} \to \mathcal{U}_{\mathcal{X}} \tag{2.177}$$

in coordinates

$$\phi_{\mathcal{X}} : u \mapsto x^\mu(u). \tag{2.178}$$

Since $\phi_{\mathcal{X}}$ is assumed to be an embedding, $\mathcal{U}_{\mathcal{X}}$ is a submanifold of $\mathcal{X}$ that has a smooth closed boundary, just like $\mathcal{U}$. It is the physical space on which the fields and trajectories, i.e., the physical sections, are defined. Closed boundaries are necessary as the term we are interested in of the variational principle, the one that yields the Cartan form, is the boundary term that arises from the partial integration.

The map $\phi_{\mathcal{X}}$ is a diffeomorphism between $\mathcal{U}$ and $\mathcal{U}_{\mathcal{X}}$. It maps between the physical space and its parametrisation, such that the composition

$$\varphi = \phi \circ \phi_{\mathcal{X}}^{-1} \tag{2.179}$$

corresponds to a parametrisation of the sections that are physical fields or trajectories. These physical sections, defined on $\mathcal{U}_{\mathcal{X}}$, can be seen as sections of the fibre bundle represented by $\pi_{\mathcal{U}_{\mathcal{X}},\mathcal{Y}}$. These are maps

$$\varphi : \mathcal{U}_{\mathcal{X}} \to \mathcal{Y} \qquad \text{with} \qquad \pi_{\mathcal{U}_{\mathcal{X}},\mathcal{Y}} \circ \varphi = \pi_{\mathcal{X}\mathcal{Y}} \circ \varphi = \mathrm{id} \tag{2.180}$$

in coordinates

$$\varphi : x \mapsto \left(x^\mu, \varphi^a(x) \right). \tag{2.181}$$

The fields that are varied in the action principle are the ϕ. Horizontal variations of the φ would change the base space $\mathcal{U}_{\mathcal{X}}$ on which the fields are defined. However, a transformation $\mathcal{Y} \to \mathcal{Y}$ acts naturally on the space $\mathcal{C}$, mapping $\mathcal{C}$ into itself, even though $\mathcal{U}_{\mathcal{X}}$ is not mapped into itself.
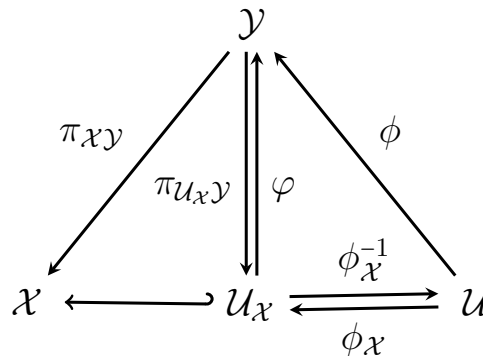
To clarify some of the previous statements, consider the analogous derivation on tangent bundles from section 2.2.2. We defined $\mathcal{C}(\mathcal{Q})$ as the space of trajectories that connect two points in $\mathcal{Q}$. These trajectories $c \in \mathcal{C}(\mathcal{Q})$ were considered as maps

$$c : \mathcal{I} \to \mathcal{Q} \qquad \text{with} \qquad \mathcal{I} \subset \mathbb{R} \text{ smooth and bounded.} \tag{2.182}$$

In the jet bundle framework they correspond to elements of $\mathcal{C}(\mathcal{Y})$

$$c : \mathcal{U} \to \mathcal{Y}. \tag{2.183}$$

In the general case, the parameter space $\mathcal{U}$ does not have just one dimension but as many as the base space $\mathcal{X}$. So the correspondence of the previous and the current notation is

$$\mathcal{I} \leftrightarrow \mathcal{U}, \qquad \mathcal{Q} \leftrightarrow \mathcal{Y}, \qquad \mathcal{C}(\mathcal{Q}) \leftrightarrow \mathcal{C}(\mathcal{Y}), \qquad c \leftrightarrow \phi, \qquad q \leftrightarrow \varphi.$$

Going back to the general theory, the tangent space to $\mathcal{C}$ at a point ϕ is the set $\mathsf{T}_\phi \mathcal{C}$ defined as

$$\mathsf{T}_\phi \mathcal{C}(\mathcal{Y}) = \left\{ V : \mathcal{U} \to \mathsf{T}\mathcal{Y} \mid \pi_{\mathcal{Y}, \mathsf{T}\mathcal{Y}} \circ V = \phi \right\}. \tag{2.184}$$

The elements V of $\mathsf{T}_\phi \mathcal{C}$ are called variations of the sections ϕ and have coordinate expressions

$$V : u \mapsto \left(\left(x^\mu(u), \varphi^a(u) \right), \left(V^\mu(u), V^a(u) \right) \right), \tag{2.185}$$

where the V^μ correspond to horizontal variations and the V^a to vertical variations. To each vector field V on $\mathsf{T}_\phi \mathcal{C}$ belongs a vector field $V_\mathcal{X}$ on $\mathcal{X}$, given by projection

$$V_\mathcal{X} \equiv \mathsf{T}\pi_{\mathcal{X}\mathcal{Y}} \circ V. \tag{2.186}$$

The projectors can be explicitly written as

$$\pi_{\mathcal{Y}, \mathsf{T}\mathcal{Y}} : \left((x^\mu, y^a), (V^\mu, V^a) \right) \mapsto (x^\mu, y^a), \tag{2.187}$$

$$\mathsf{T}\pi_{\mathcal{X}\mathcal{Y}} : \left((x^\mu, y^a), (V^\mu, V^a) \right) \mapsto (x^\mu, V^\mu), \tag{2.188}$$

such that $V_\mathcal{X}$ has the coordinate expression

$$V_\mathcal{X} : u \mapsto \left(x^\mu(u), V^\mu(u) \right). \tag{2.189}$$

The definition (2.184) of $\mathsf{T}_\phi \mathcal{C}$ can also be seen by considering the variation of a path in $\mathcal{C}(\mathcal{Y})$,

$$\phi^\epsilon : u \mapsto \left(x^\mu(\epsilon, u), \phi^a(\epsilon, u) \right). \tag{2.190}$$

The derivatives of this expression and its projection to $\mathcal{X}$ are

$$\left. \frac{d\phi^\epsilon}{d\epsilon} \right|_{\epsilon=0} = \left(V^\mu(u), V^a(u) \right), \tag{2.191}$$

which coincides with $V(u)$ in $\mathsf{T}_{\phi(u)} \mathcal{Y}$, and

$$\left. \frac{d\phi^\epsilon_\mathcal{X}}{d\epsilon} \right|_{\epsilon=0} = \frac{d}{d\epsilon} \left[\pi_{\mathcal{X}\mathcal{Y}} \circ \phi^\epsilon \right] \Big|_{\epsilon=0} = \mathsf{T}\pi_{\mathcal{X}\mathcal{Y}} V = V_\mathcal{X}. \tag{2.192}$$

To aid understanding some of these relations are depicted in the diagram below.

$$
\begin{array}{ccc}
& \mathsf{T}\mathcal{Y} \xrightarrow{\;\mathsf{T}\pi_{\mathcal{X}\mathcal{Y}}\;} \mathsf{T}\mathcal{X} \\
{\scriptstyle V}\nearrow & \downarrow{\scriptstyle \pi_{\mathcal{Y},\mathsf{T}\mathcal{Y}}} \quad \downarrow{\scriptstyle \pi_{\mathcal{X},\mathsf{T}\mathcal{X}}} \\
\mathcal{U} \xrightarrow{\;\phi\;} \mathcal{Y} \xrightarrow{\;\pi_{\mathcal{X}\mathcal{Y}}\;} \mathcal{X}
\end{array}
$$

Let us rephrase the action principle from the last section in this notation. The action functional

$$
\mathcal{A} : \mathcal{C}(\mathcal{Y}) \to \mathbb{R} \tag{2.193}
$$

is expressed as

$$
\mathcal{A}[\phi] = \int_{\mathcal{U}_{\mathcal{X}}} (j^1\varphi)^* \mathcal{L} = \int_{\mathcal{U}_{\mathcal{X}}} \mathcal{L}(j^1\varphi) = \int_{\mathcal{U}_{\mathcal{X}}} \mathcal{L}\big(j^1(\phi \circ \phi_{\mathcal{X}}^{-1})\big). \tag{2.194}
$$

As the action $\mathcal{A}[\phi]$ depends on ϕ only through φ, for any diffeomorphism $\gamma : \mathcal{U} \to \mathcal{U}$

$$
\mathcal{A}[\phi \circ \gamma] = \mathcal{A}[\phi]. \tag{2.195}
$$

As a consequence, the Euler-Lagrange equations only determine φ uniquely, not ϕ. However, as φ corresponds to the physical fields or trajectories, that is all we need. Hamilton's principle states that a section ϕ of $\mathcal{C}(\mathcal{Y})$ solves the Euler-Lagrange equations, iff the action is critical,

$$
\mathbf{d}\mathcal{A}[\phi] \cdot V = 0, \tag{2.196}
$$

for all variations $V \in \mathsf{T}_\phi \mathcal{C}(\mathcal{Y})$ which are zero on the boundary $\partial\mathcal{U}$ of $\mathcal{U}$. To derive the Cartan form, the last restriction has to be removed. The above expression corresponds to

$$
\mathbf{d}\mathcal{A}[\phi] \cdot V = \frac{d}{d\epsilon}\mathcal{A}[\phi^\epsilon]\bigg|_{\epsilon=0} = 0. \tag{2.197}
$$

A variation ϕ^ϵ of a section ϕ is induced by a transformation $\eta_{\mathcal{Y}}^\epsilon$ on the configuration space

$$
\eta_{\mathcal{Y}}^\epsilon : \mathcal{Y} \to \mathcal{Y} \qquad \text{with} \qquad \eta_{\mathcal{Y}}^0 = \mathrm{id} \tag{2.198}
$$

through

$$
\phi^\epsilon = \eta_{\mathcal{Y}}^\epsilon \circ \phi. \tag{2.199}
$$

We impose the condition that $\eta_{\mathcal{Y}}^\epsilon$ covers a diffeomorphism

$$
\eta_{\mathcal{X}}^\epsilon : \mathcal{X} \to \mathcal{X}. \tag{2.200}
$$

In coordinates

$$
\eta_{\mathcal{Y}}^\epsilon : \big(x, y\big) \mapsto \big(\eta_{\mathcal{X}}^\mu(x), \eta_{\mathcal{Y}}^a(x, y)\big). \tag{2.201}
$$

The diffeomorphism on the base space $\mathcal{X}$ is obtained through the projection

$$\eta_{\mathcal{X}}^{\epsilon} = \pi_{\mathcal{X}\mathcal{Y}} \circ \eta_{\mathcal{Y}}^{\epsilon}. \qquad (2.202)$$

The following diagram should help clarify these relations.

$$
\begin{array}{ccc}
\mathcal{Y} & \xrightarrow{\;\;\eta_{\mathcal{Y}}^{\epsilon}\;\;} & \mathcal{Y} \\
\varphi \uparrow & & \uparrow \varphi^{\epsilon} \\
\mathcal{U}_{\mathcal{X}} & \xrightarrow[\;\;\eta_{\mathcal{X}}^{\epsilon}\;\;]{} & \eta_{\mathcal{X}}^{\epsilon}(\mathcal{U}_{\mathcal{X}})
\end{array}
$$

We see now why it is necessary to introduce a parameter space $\mathcal{U}$. A physical section $\varphi = \phi \circ \phi_{\mathcal{X}}^{-1}$ is a section of $\pi_{\mathcal{U}_{\mathcal{X}},\mathcal{Y}}$. But the transformation $\eta^{\epsilon} \circ \phi$ induces a section $\varphi^{\epsilon} = \eta_{\mathcal{Y}}^{\epsilon} \circ (\phi \circ \phi_{\mathcal{X}}^{-1}) \circ (\eta_{\mathcal{X}}^{\epsilon})^{-1}$ of $\pi_{\eta_{\mathcal{X}}(\mathcal{U}_{\mathcal{X}}),\mathcal{Y}}$, i.e., the base space itself changes under the transformation. This becomes more evident by looking at the coordinate expressions

$$\phi : u \mapsto \left(x^{\mu}(u), \phi^{a}(u)\right), \qquad \phi^{\epsilon} : u \mapsto \left(x^{\mu}(u), \eta_{\mathcal{Y}}^{a}(x(u), \phi(u))\right), \qquad (2.203\text{a})$$

$$\varphi : x \mapsto \left(x^{\mu}, \phi^{a}(x)\right), \qquad \varphi^{\epsilon} : \tilde{x} \mapsto \left(\tilde{x}^{\mu}, \eta_{\mathcal{Y}}^{a}(x, \varphi(x))\right), \qquad (2.203\text{b})$$

where $\tilde{x} \in \eta_{\mathcal{X}}(\mathcal{U}_{\mathcal{X}})$ and $x = (\eta_{\mathcal{X}}^{\epsilon})^{-1}(\tilde{x}) \in \mathcal{U}_{\mathcal{X}}$. If we consider not variations of the physical sections $\varphi : \mathcal{U}_{\mathcal{X}} \to \mathcal{Y}$ but variations of the sections $\phi : \mathcal{U} \to \mathcal{Y}$, the point u in the base space $\mathcal{U}$ stays fixed.

Applying the transformation (2.198) to the action (2.194), the variation (2.197) becomes

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \frac{d}{d\epsilon}\mathcal{A}[\eta_{\mathcal{Y}}^{\epsilon} \circ \phi]\bigg|_{\epsilon=0} = \frac{d}{d\epsilon} \int_{\eta_{\mathcal{X}}^{\epsilon}(\mathcal{U}_{\mathcal{X}})} \mathcal{L}\left(j^1\varphi^{\epsilon}\right)\bigg|_{\epsilon=0} = \int_{\mathcal{U}_{\mathcal{X}}} \frac{d}{d\epsilon}\left(\eta_{\mathcal{X}}^{\epsilon}\right)^{*}\mathcal{L}\left(j^1\varphi^{\epsilon}\right)\bigg|_{\epsilon=0}. \qquad (2.204)$$

Application of the chain rule yields

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_{\mathcal{X}}} \frac{d}{d\epsilon}\left[\left(\eta_{\mathcal{X}}^{\epsilon}\right)^{*}\mathcal{L}\left(j^1\varphi^{0}\right)\right]\bigg|\bigg|_{\epsilon=0} + \int_{\mathcal{U}_{\mathcal{X}}} \frac{d}{d\epsilon}\left[\left(\eta_{\mathcal{X}}^{0}\right)^{*}\mathcal{L}\left(j^1\varphi^{\epsilon}\right)\right]\bigg|\bigg|_{\epsilon=0}. \qquad (2.205)$$

In the first integral, apply the dynamical definition of the Lie derivative as before in (2.155), and in the second integral realise that $\eta_{\mathcal{X}}^{0}$ is just the identity

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_{\mathcal{X}}} \pounds_{V_{\mathcal{X}}}\mathcal{L}\left(j^1\varphi\right) + \int_{\mathcal{U}_{\mathcal{X}}} \frac{d}{d\epsilon}\left[\mathcal{L}\left(j^1\varphi^{\epsilon}\right)\right]\bigg|\bigg|_{\epsilon=0}. \qquad (2.206)$$

Use Cartan's magic formula (2.81) in the first integral, and rewrite the second integral by making use of the identity

$$\frac{d}{d\epsilon}\left[\mathcal{L}\left(j^1\varphi^{\epsilon}\right)\right]\bigg|\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\left[L\left(j^1\varphi^{\epsilon}\right)\right]\bigg|\bigg|_{\epsilon=0} \omega = \left(\boldsymbol{\imath}_{j^1 V_{\varphi}}\mathbf{d}L\left(j^1\varphi\right)\right)\omega, \qquad (2.207)$$

such that

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_\mathcal{X}} \boldsymbol{\imath}_{V_\mathcal{X}} \mathbf{d}\mathcal{L}(j^1\varphi) + \int_{\mathcal{U}_\mathcal{X}} \mathbf{d}\big(\boldsymbol{\imath}_{V_\mathcal{X}} \mathcal{L}(j^1\varphi)\big) + \int_{\mathcal{U}_\mathcal{X}} \big(\boldsymbol{\imath}_{j^1 V_\varphi} \mathbf{d}L(j^1\varphi)\big)\omega. \tag{2.208}$$

The first integral vanishes as $\mathcal{L} = L\omega$ and therefore

$$\mathbf{d}\mathcal{L}(j^1\varphi) = \mathbf{d}L(j^1\varphi) \wedge \omega + L(j^1\varphi) \mathbf{d}\omega = L_\mu(j^1\varphi) \, dx^\mu \wedge \omega + L(j^1\varphi) \mathbf{d}\omega = 0, \tag{2.209}$$

but $\omega = dx^1 \wedge ... \wedge dx^n$ is a form of maximum order on the base space, such that $\mathbf{d}\omega = 0$ and $dx^\mu \wedge \omega = 0$ for all μ. By Stokes' theorem, the second integral can be transformed into a surface integral, with the result that

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\partial \mathcal{U}_\mathcal{X}} \boldsymbol{\imath}_{V_\mathcal{X}} \mathcal{L}(j^1\varphi) + \int_{\mathcal{U}_\mathcal{X}} \big(\boldsymbol{\imath}_{j^1 V_\varphi} \mathbf{d}L(j^1\varphi)\big)\omega. \tag{2.210}$$

Now we have to compute the vector field V_φ corresponding to the transformation of the physical section φ

$$V_\varphi = \frac{d}{d\epsilon}\varphi^\epsilon\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\Big[\eta_\mathcal{Y}^\epsilon \circ (\phi \circ \phi_\mathcal{X}^{-1}) \circ (\eta_\mathcal{X}^\epsilon)^{-1}\Big]\bigg|_{\epsilon=0}, \tag{2.211}$$

and its jet prolongation $j^1 V_\varphi$. Applying the chain rule, we get

$$V_\varphi = \frac{d}{d\epsilon}\Big[\eta_\mathcal{Y}^\epsilon \circ (\phi \circ \phi_\mathcal{X}^{-1}) \circ (\eta_\mathcal{X}^0)^{-1}\Big]\bigg|_{\epsilon=0} + \frac{d}{d\epsilon}\Big[\eta_\mathcal{Y}^0 \circ (\phi \circ \phi_\mathcal{X}^{-1}) \circ (\eta_\mathcal{X}^\epsilon)^{-1}\Big]\bigg|_{\epsilon=0}. \tag{2.212}$$

In the first term we use (2.191) and in the second term we use the fact that $d/d\epsilon(\eta_\mathcal{X}^\epsilon)^{-1}|_{\epsilon=0} = -V_\mathcal{X}$[19], such that

$$V_\varphi = V \circ \phi_\mathcal{X}^{-1} - \mathsf{T}(\phi \circ \phi_\mathcal{X}^{-1}) \circ V_\mathcal{X}. \tag{2.213}$$

The tangent lift of the vector field $V_\mathcal{X}$ is simply

$$\mathsf{T}(\phi \circ \phi_\mathcal{X}^{-1}) \circ V_\mathcal{X} = \Big((x^\nu, \varphi^a(x)), (V^\nu, \varphi_\mu^a V^\mu)\Big) \tag{2.214}$$

such that

$$V_\varphi = (0, \delta\varphi^a) = \big(0, V^a - \varphi_\mu^a V^\mu\big). \tag{2.215}$$

---

[19] This can be seen by the group property of the transformation

$$\eta_\mathcal{X}^{-\epsilon} = (\eta_\mathcal{X}^\epsilon)^{-1} \qquad \rightarrow \qquad -V_\mathcal{X} = \frac{d}{d\epsilon}\eta_\mathcal{X}^{-\epsilon}\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}(\eta_\mathcal{X}^\epsilon)^{-1}\bigg|_{\epsilon=0}$$

or by a simple calculation as follows

$$\eta_\mathcal{X}^\epsilon \circ (\eta_\mathcal{X}^\epsilon)^{-1} = \eta_\mathcal{X}\big(\epsilon, \eta_X^{-1}(\epsilon, x)\big) = \mathrm{id}$$

such that

$$\frac{d}{d\epsilon}\Big[\eta_\mathcal{X}^\epsilon \circ (\eta_\mathcal{X}^\epsilon)^{-1}\Big]\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\Big[\eta_\mathcal{X}^\epsilon \circ (\eta_\mathcal{X}^0)^{-1}\Big]\bigg|_{\epsilon=0} + \frac{d}{d\epsilon}\Big[\eta_\mathcal{X}^0 \circ (\eta_\mathcal{X}^\epsilon)^{-1}\Big]\bigg|_{\epsilon=0} = 0.$$

$\eta_\mathcal{X}^0 = \mathrm{id}$ and $(\eta_\mathcal{X}^0)^{-1} = \mathrm{id}^{-1} = \mathrm{id}$ as well, such that

$$\frac{d}{d\epsilon}\Big[\eta_\mathcal{X}^\epsilon \circ (\eta_\mathcal{X}^\epsilon)^{-1}\Big]\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\Big[\eta_\mathcal{X}^\epsilon\Big]\bigg|_{\epsilon=0} + \mathsf{T}\,\mathrm{id}\circ\frac{d}{d\epsilon}\Big[(\eta_\mathcal{X}^\epsilon)^{-1}\Big]\bigg|_{\epsilon=0} = V_\mathcal{X} + \frac{d}{d\epsilon}(\eta_\mathcal{X}^\epsilon)^{-1}\bigg|_{\epsilon=0} = 0,$$

where the tangent lift of the identity is the identity on the tangent space.

This is just the vertical component of the vector field V from (2.185)[20]. The jet prolongation of V_φ along $j^1\varphi$ is

$$j^1 V_\varphi = (0, \delta\varphi^a, \delta\varphi^a_\nu) = \left(0, V^a - \varphi^a_\mu V^\mu, \partial_\nu(V^a - \varphi^a_\mu V^\mu)\right). \tag{2.216}$$

With that we compute the action (2.210) as

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_\mathcal{X}} \left[\frac{\partial L}{\partial y^a}(j^1\varphi)\,\delta\varphi^a + \frac{\partial L}{\partial v^a_\mu}(j^1\varphi)\,\delta\varphi^a_\mu\right]\omega + \int_{\partial\mathcal{U}_\mathcal{X}} L(j^1\varphi)\,V^\mu\,\omega_\mu \tag{2.217}$$

where $\omega_\mu = \partial_\mu \lrcorner \omega$. Integrate by parts the second term of the first integral

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_\mathcal{X}} \left[\frac{\partial L}{\partial y^a}(j^1\varphi) - \frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial v^a_\mu}(j^1\varphi)\right)\right]\delta\varphi^a\,\omega$$

$$+ \int_{\partial\mathcal{U}_\mathcal{X}} \left[L(j^1\varphi)\,V^\mu + \frac{\partial L}{\partial v^a_\mu}(j^1\varphi)\,V^a - \frac{\partial L}{\partial v^a_\mu}(j^1\varphi)\,\varphi^a_\nu V^\nu\right]\omega_\mu. \tag{2.218}$$

To bring this expression into a coordinate-free form, consider a general vector field $W = (W^\nu, W^a, W^a_\nu)$. Its contraction with $dy^a \wedge \omega$ is given by

$$\boldsymbol{\imath}_W(dy^a \wedge \omega) = W^a\,\omega - (-1)^{p^\nu} W^\nu\,dy^a \wedge \omega_\nu \tag{2.219}$$

where p^ν is the number of permutations in the computation of $\omega_\nu = \partial_\nu \lrcorner \omega$ (the additional minus results from the permutation with dy^a). The pullback of this relation with $j^1\varphi$ is

$$(j^1\varphi)^*\boldsymbol{\imath}_W(dy^a \wedge \omega) = W^a(j^1\varphi)\,\omega - (-1)^{p^\nu} W^\nu\,\varphi^a_\nu\,dx^\nu \wedge \omega_\nu = (W^a(j^1\varphi) - W^\nu\,\varphi^a_\nu)\,\omega. \tag{2.220}$$

Applying this result to the variation of the action (2.218), we find

$$(j^1\varphi)^*\boldsymbol{\imath}_{j^1V}(dy^a \wedge \omega) = \delta\varphi^a\,\omega. \tag{2.221}$$

Further, consider the expression

$$\boldsymbol{\imath}_W(dy^a \wedge \omega_\nu) = W^a\omega_\nu - (-1)^{p^\mu} W^\mu\,dy^a \wedge \omega_{\nu\mu}. \tag{2.222}$$

With

$$dx^\sigma \wedge \omega_{\nu\mu} = \begin{cases} 0 & \sigma \neq \mu, \nu \\ \omega_\nu & \mu = \sigma \\ -\omega_\mu & \nu = \sigma \end{cases} \tag{2.223}$$

the pullback of (2.222) with $j^1\varphi$ is

$$(j^1\varphi)^*\boldsymbol{\imath}_W(dy^a \wedge \omega_\mu) = W^a\omega_\mu - (-1)^{p^\nu} W^\nu\,\varphi^a_\sigma\,dx^\sigma \wedge \omega_{\mu\nu}$$

$$= W^a\omega_\mu - W^\nu\varphi^a_\nu\omega_\mu + \varphi^a_\mu W^\nu\omega_\mu. \tag{2.224}$$

---

[20] Any vector $V \in \mathsf{T}_\phi\mathcal{C}$ can be decomposed into a horizontal and a vertical component $V = V^h + V^v$, where $V^h = \mathsf{T}\varphi \circ V_\mathcal{X}$ and $V^v = V - V^h$.

Applying this result to the variation of the action (2.218), we find

$$(j^1\varphi)^* \boldsymbol{\imath}_V (dy^a \wedge \omega_\mu) = V^a \omega_\mu - V^\sigma \varphi^a_\nu \, dx^\nu \wedge \omega_{\mu\sigma} = V^a \omega_\mu - V^\sigma \varphi^a_\nu \left(\omega_\mu \delta^\nu_\sigma - \omega_\sigma \delta^\nu_\mu \right)$$
$$= V^a \omega_\mu - V^\nu \varphi^a_\nu \omega_\mu + V^\nu \varphi^a_\mu \omega_\nu. \tag{2.225}$$

A final but simple computation shows

$$(j^1\varphi)^* \boldsymbol{\imath}_{j^1 V} \omega = V^\mu \omega_\mu. \tag{2.226}$$

Therefore, the variation of the action (2.218) can be written as

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_\chi} (j^1\varphi)^* \boldsymbol{\imath}_{j^1 V} \left[\frac{\partial L}{\partial y^a} - \frac{\partial}{\partial x^\mu} \frac{\partial L}{\partial v^a_\mu} \right] dy^a \wedge \omega$$
$$+ \int_{\partial \mathcal{U}_\chi} (j^1\varphi)^* \boldsymbol{\imath}_{j^1 V} \left[\frac{\partial L}{\partial v^a_\mu} \, dy^a \wedge \omega_\mu + \left(L - \frac{\partial L}{\partial v^a_\mu} \varphi^a_\mu \right) \omega \right]. \tag{2.227}$$

In the first integral we find the Euler-Lagrange equations as we derived them before by considering only vertical variations (2.169). The expression in square brackets in the second integral is the looked for Cartan form

$$\Theta_L = \frac{\partial L}{\partial v^a_\mu} \, dy^a \wedge \omega_\mu + \left(L - \frac{\partial L}{\partial v^a_\mu} \varphi^a_\mu \right) \omega \qquad \textbf{(Cartan Form)}. \tag{2.228}$$

In the case of particle mechanics, this is a one-form, otherwise it is a form of the order of the base manifold. The *symplectic form* is defined as the exterior derivative of the Cartan form

$$\Omega_L = -\mathbf{d}\Theta_L = dy^a \wedge \mathbf{d}\left(\frac{\partial L}{\partial v^a_\mu} \right) \wedge \omega_\mu - \mathbf{d}\left(L - \frac{\partial L}{\partial v^a_\mu} \varphi^a_\mu \right) \wedge \omega \quad \textbf{(Multisymplectic Form)}. \tag{2.229}$$

Both, the Cartan and the symplectic form, are defined on the first jet bundle, i.e., $\Theta_L \in \Omega_1$ and $\Omega_L \in \Omega^2$. Using these expressions, the variation of the action (2.227) can be written as

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_\chi} (j^1\varphi)^* (\boldsymbol{\imath}_{j^1 V} \Omega_L) + \int_{\partial \mathcal{U}_\chi} (j^1\varphi)^* (\boldsymbol{\imath}_{j^1 V} \Theta_L). \tag{2.230}$$

This will be the starting point to prove the preservation of the multisymplectic form along the Lagrangian flow in section (2.2.6). The expressions (2.228) and (2.229) are more general than the ones we derived previously, (2.171) and (2.172), in that they lift the restriction to a time-independent Lagrangian and describe field theories as well.

2.2.5. Preservation of the Symplectic Form

In this and the next section we want to prove the conservation of the symplectic and multisymplectic forms under Lagrangian flows. At first we do so on the tangent bundle, thereby restricting ourselves to the case of particle dynamics. The approach is then generalised to the framework of jet bundles, whereby we obtain a general proof that is valid for both particle and field systems.

Euler-Lagrange Map and Lagrangian Vector Fields

This section mostly aims at making the literature more easily accessible. It is not strictly necessary to understand the subsequent treatment.

Define the submanifold $\ddot{\mathcal{Q}}$ of $\mathsf{T}(\mathsf{T}\mathcal{Q})$ to be

$$\ddot{\mathcal{Q}} \equiv \left\{ w \in \mathsf{T}(\mathsf{T}\mathcal{Q}) \;\middle|\; \mathsf{T}\pi_{\mathcal{Q}}w = \pi_{\mathsf{T}\mathcal{Q}}w \right\} \subset \mathsf{T}(\mathsf{T}\mathcal{Q}). \tag{2.231}$$

This states that $\ddot{\mathcal{Q}}$ contains those elements of $\mathsf{T}(\mathsf{T}\mathcal{Q})$ for which the two projections $\pi_{\mathsf{T}\mathcal{Q}}$ and $\mathsf{T}\pi_{\mathcal{Q}}$ coincide.

$$
\begin{array}{ccc}
\mathsf{T}\mathcal{Q} & \xleftarrow[\mathsf{T}\pi_{\mathcal{Q}}]{\pi_{\mathsf{T}\mathcal{Q}}} & \mathsf{T}(\mathsf{T}\mathcal{Q}) \\
\downarrow{\scriptstyle \pi_{\mathcal{Q}}} & & \\
\mathcal{Q} & &
\end{array}
$$

To see what that means, write both expressions in coordinates

$$\pi_{\mathsf{T}\mathcal{Q}} : \big((q,v), (\dot{q},\dot{v}) \big) \mapsto (q,v), \tag{2.232a}$$
$$\mathsf{T}\pi_{\mathcal{Q}} : \big((q,v), (\dot{q},\dot{v}) \big) \mapsto (q,\dot{q}). \tag{2.232b}$$

Requiring that both projections are equivalent therefore means singling out those elements of $\mathsf{T}(\mathsf{T}\mathcal{Q})$ for which $v = \dot{q}$ and therefore also $\dot{v} = \ddot{q}$. These correspond to curves $q(t) \in \mathcal{Q}$ which are tangent lifted twice, first to $\mathsf{T}\mathcal{Q}$, then to $\mathsf{T}(\mathsf{T}\mathcal{Q})$. In other words, elements $w \in \ddot{\mathcal{Q}}$ are those elements of $\mathsf{T}(\mathsf{T}\mathcal{Q})$ that have the coordinate expression

$$w = \big((q,\dot{q}), (\dot{q},\ddot{q}) \big). \tag{2.233}$$

An alternative definition of the second order submanifold $\ddot{\mathcal{Q}}$ is therefore

$$\ddot{\mathcal{Q}} \equiv \left\{ \frac{d^2 q}{dt^2}(0) \in \mathsf{T}(\mathsf{T}\mathcal{Q}) \;\middle|\; q(t) \text{ a curve in } \mathcal{Q} \right\} \subset \mathsf{T}(\mathsf{T}\mathcal{Q}). \tag{2.234}$$

Given a Lagrangian L, there exists a map on $\ddot{\mathcal{Q}}$

$$D_{\mathrm{EL}}L : \ddot{\mathcal{Q}} \to \mathsf{T}^*\mathcal{Q} \tag{2.235}$$

referred to as the the Euler-Lagrange map. It defines a one-form in the dual space $\mathsf{T}^*\mathcal{Q}$ of $\mathsf{T}\mathcal{Q}$ with coordinate expression

$$D_{\mathrm{EL}}L = \frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial v}. \tag{2.236}$$

It is a function on $\ddot{\mathcal{Q}} \subset \mathsf{T}(\mathsf{T}\mathcal{Q})$ as

$$D_{\mathrm{EL}}L = \frac{\partial L}{\partial q}(q,\dot{q}) - \frac{\partial^2 L}{\partial v \, \partial q}(q,\dot{q}) \cdot \dot{q} - \frac{\partial^2 L}{\partial v \, \partial v}(q,\dot{q}) \cdot \ddot{q}. \tag{2.237}$$

With that, the variational principle (2.170) can be written

$$\mathbf{d}\mathcal{A}\big[q(t)\big]\cdot\delta q(t) = \int_{t_1}^{t_2} D_{\mathrm{EL}}L(\ddot{q})\cdot\delta q\, dt + \Big[\Theta_L(\dot{q})\cdot\delta\hat{q}\Big]_{t_1}^{t_2} \tag{2.238}$$

where $\ddot{q}$ refers to an element of $\ddot{\mathcal{Q}}$ and thus has coordinates $\big((q,\dot{q}),(\dot{q},\ddot{q})\big)$. Similarly, $\dot{q}$ refers to an element of $\mathsf{T}\mathcal{Q}$ with coordinates $(q,\dot{q})$. The variation $\delta\hat{q}$ is defined as

$$\delta\hat{q} \equiv \frac{d}{d\epsilon}\Big|_{\epsilon=0}\frac{d}{dt}\Big|_{t=0}q_\epsilon(t) \tag{2.239}$$

or in coordinates

$$\delta\hat{q}(t) = \Big((q(t),\dot{q}(t)),(\delta q(t),\delta\dot{q}(t))\Big). \tag{2.240}$$

This Euler-Lagrange map can now be used to define the *Lagrangian vector field*

$$X_L : \mathsf{T}\mathcal{Q} \to \ddot{\mathcal{Q}} \tag{2.241}$$

as a second order vector field on $\mathsf{T}\mathcal{Q}$ satisfying

$$D_{\mathrm{EL}}L \circ X_L = 0. \tag{2.242}$$

The flow of X_L is called the *Lagrangian flow*

$$F_L^t : \mathsf{T}\mathcal{Q} \to \mathsf{T}\mathcal{Q}. \tag{2.243}$$

By construction $q \in \mathcal{C}(\mathcal{Q})$ is a solution of the Euler-Lagrange equations iff $(q,\dot{q})$ is an integral curve of X_L. In the next section, the Lagrangian flow F_L will be defined without referring to the Euler-Lagrange map but by using coordinate expressions instead. The advantage is a somewhat easier treatment.

Lagrangian Flows and Preservation of the Symplectic Form

Denote the vector field on $\mathsf{T}\mathcal{Q}$ that solves the Euler-Lagrange equations by X_L. Its flow, referred to as the Lagrangian flow, is a map

$$F_L^s : \mathsf{T}\mathcal{Q} \to \mathsf{T}\mathcal{Q}, \tag{2.244}$$

taking initial values $(q_0,\dot{q}_0)$ to points of the corresponding phasespace trajectory at time s, that is

$$F_L^s : \big(q_0,\dot{q}_0\big) \mapsto \big(q(s),\dot{q}(s)\big), \tag{2.245}$$

such that

$$\frac{\partial L}{\partial q}\big((q(s),\dot{q}(s))\big) - \frac{d}{ds}\frac{\partial L}{\partial\dot{q}}\big((q(s),\dot{q}(s))\big) = 0. \tag{2.246}$$

The Lagrangian vector field is accordingly defined as

$$X_L(q_0,v_0) = \frac{d}{ds}F_L^s(q_0,v_0)\Big|_{s=0}. \tag{2.247}$$

Restrict the action $\mathcal{A}$ to the subspace $\mathcal{C}_L \subset \mathcal{C}(\mathcal{Q})$ of solutions of the Euler-Lagrange equations. Elements $q \in \mathcal{C}_L$ are integral curves of X_L, and therefore uniquely determined by the initial condition $v_q = \big(q(0), \dot{q}(0)\big) \in \mathsf{T}\mathcal{Q}$. Consequently, $\mathcal{C}_L$ may be identified with the space of initial conditions, i.e., $\mathcal{C}_L$ is isomorphic to $\mathsf{T}\mathcal{Q}$.

Associate to v_q the integral curve $s \mapsto F_L^s(v_q)$ with $s \in [0, t]$

$$q(s) = \pi_{\mathcal{Q}}\big(F_L^s(v_q)\big) \qquad \text{with} \qquad F_L^s(v_q) = \big(q(s), \dot{q}(s)\big). \tag{2.248}$$

The restricted action $\mathcal{A}_t$ corresponds to the value of $\mathcal{A}$ on that curve. It defines a map

$$\mathcal{A}_t : \mathsf{T}\mathcal{Q} \to \mathbb{R} \tag{2.249}$$

by

$$\mathcal{A}_t[v_q] = \mathcal{A}[q] \qquad \text{with} \qquad q \in \mathcal{C}_L \qquad \text{and} \qquad \big(q(0), \dot{q}(0)\big) = v_q. \tag{2.250}$$

or explicitly

$$\mathcal{A}_t[v_q] = \int_0^t L\big(q(s), \dot{q}(s)\big)\, ds = \int_0^t L\big(F_L^s(v_q)\big)\, ds \tag{2.251}$$

Calculating the variation of the restricted action, the first term in (2.170) vanishes, as $\mathcal{A}$ is restricted to solutions of the Euler-Lagrange equations

$$\mathbf{d}\mathcal{A}[q] \cdot \delta q = \left[\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \cdot \delta \hat{q}\right]_0^t \qquad \text{with} \qquad q \in \mathcal{C}_L. \tag{2.252}$$

As $\mathcal{A}_t$ is considered a real-valued function on $\mathsf{T}\mathcal{Q}$, this becomes

$$\mathbf{d}\mathcal{A}_t[v_q] \cdot w_{v_q} = \Theta_L\big(F_L^t(v_q)\big) \cdot \frac{d}{d\epsilon}F_L^t(v_q^\epsilon)\bigg|_{\epsilon=0} - \Theta_L(v_q) \cdot w_{v_q} \tag{2.253}$$

with v_q^ϵ an arbitrary curve in $\mathsf{T}\mathcal{Q}$, namely

$$v_q^\epsilon : \mathbb{R} \to \mathsf{T}\mathcal{Q} \qquad \text{such that} \qquad v_q^0 = v_q \qquad \text{and} \qquad w_{v_q} = \frac{d}{d\epsilon}v_q^\epsilon\bigg|_{\epsilon=0}. \tag{2.254}$$

Since w_{v_q} is arbitrary, (2.253) is equivalent to

$$\mathbf{d}\mathcal{A}_t = (F_L^t)^*\Theta_L - \Theta_L. \tag{2.255}$$

Taking the exterior derivative

$$0 = \mathbf{d}^2\mathcal{A}_t = \mathbf{d}(F_L^t)^*\Theta_L - \Theta_L = -(F_L^t)^*\Omega_L + \Omega_L. \tag{2.256}$$

leads to the conservation of the symplectic form Ω_L along the Lagrangian flow X_L

$$(F_L^t)^*\Omega_L = \Omega_L. \tag{2.257}$$

These results can be considered the Lagrangian equivalent of the well-known conservation of the symplectic form by Hamiltonian flows [82, 8, 49].

2.2.6. Preservation of the Multisymplectic Form

In this section we want to show that the multisymplectic form Ω_L from (2.229) is preserved under the Lagrangian flow, a generalisation of the results from the previous section. Therefore we recall equation (2.230) for the variation of the action, that is

$$\mathbf{d}\mathcal{A}[\phi] \cdot V = \int_{\mathcal{U}_x} (j^1\varphi)^* (\boldsymbol{\imath}_{j^1 V} \Omega_L) + \int_{\partial\mathcal{U}_x} (j^1\varphi)^* (\boldsymbol{\imath}_{j^1 V} \Theta_L). \tag{2.258}$$

The action $\mathcal{A}$ takes an extremum for $\phi \in \mathcal{C}(\mathcal{Y})$ if the first integral vanishes. The corresponding integrand vanishes not only for vector fields $j^1 V$, corresponding to vertical transformations, but for general vector fields W on $\mathsf{J}^1\mathcal{Y}$, that can be tangent to any $j^1\varphi$. As a consequence, ϕ is an extremum of the action, if the variation of the action (2.258) vanishes for all vectors $W \in \mathsf{T}(\mathsf{J}^1\mathcal{Y})$. Such ϕ are solutions of the Euler-Lagrange equations (2.169).

We define $\mathcal{C}_L$ to be the restriction of $\mathcal{C}(\mathcal{Y})$, defined in (2.175), to solutions of the Euler-Lagrange equations, i.e.,

$$\mathcal{C}_L = \big\{ \phi \in \mathcal{C}(\mathcal{Y}) \,\big|\, (j^1\varphi)^* [W \lrcorner \Omega_L] = 0 \quad \text{for all } W \in \mathsf{T}(\mathsf{J}^1\mathcal{Y}) \big\}, \tag{2.259}$$

such that φ^a is an element of $\mathcal{C}_L$ if

$$\frac{\partial L}{\partial y^a}(j^1\varphi) - \frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial v^a_\mu}(j^1\varphi) \right) = 0 \quad \text{in } \mathcal{U}_x. \tag{2.260}$$

A vector field $V \in \mathsf{T}\mathcal{C}_L$ is called a first variation. Its flow maps solutions ϕ of the Euler-Lagrange equations to other solutions of the Euler-Lagrange equations, such that sections $\phi \in \mathcal{C}_L$ are integral curves of V.

If we restrict the variation of the action to $\mathcal{C}_L$, the first integral in (2.258) becomes zero. Computing the exterior derivative of (2.258) and restricting it to two first variations $V, W \in \mathsf{T}\mathcal{C}_L$, gives

$$0 = \mathbf{d}^2\mathcal{A}[\phi] \cdot V \cdot W = \int_{\partial\mathcal{U}_x} (j^1\varphi)^* (V \lrcorner W \lrcorner \mathbf{d}\Theta_L). \tag{2.261}$$

This states that the multisymplectic form Ω_L is conserved

$$\int_{\partial\mathcal{U}_x} (j^1\varphi)^* (V \lrcorner W \lrcorner \Omega_L) = 0 \qquad \textbf{(Multisymplectic Form Formula)}. \tag{2.262}$$

A detailed proof of this expression is omitted but can be found in Marsden et al. [84].

2.2.7. Extended Lagrangians

In the variational treatment of field theoretic problems from plasma physics one faces the problem that most systems do not have a natural Lagrangian formulation. Similarly, even so most systems are Hamiltonian, they do not feature a canonical Hamiltonian formulation with respect to canonical conjugate variables, but only a so called noncanonical formulation. Therefore it is not possible to write a canonical Lagrangian for these systems. However, to apply the variational integrator formalism, a Lagrangian is indispensable. Salvation is brought by Ibragimov and his theory of integrating factors and adjoint equations [58, 59]. The basic idea is to extend the system by doubling the number of dependent

variables. This enables to write down a Lagrangian that is the product of the original equations and the added auxiliary variables, such that the variation with respect to those new variables results in the original equations.

In the following, we summarise the important definitions and statements of Ibragimov [58].

Integrating Factors

Integrating factors provide means to solve differential equations. Any first order differential equation of the form

$$a(x,y)\frac{dy}{dx} + b(x,y) = 0 \tag{2.263}$$

can also be written in differential form, i.e.,

$$a(x,y)\,dy + b(x,y)\,dx = 0. \tag{2.264}$$

This equations is said to be *exact* if its left hand side is the differential of some function $\alpha(x,y)$

$$a(x,y)\,dy + b(x,y)\,dx = d\alpha(x,y). \tag{2.265}$$

In general, (2.264) is not exact, but it can become exact upon multiplying by an appropriate function $\mu(x,y)$

$$\mu(x,y)\big(a(x,y)\,dy + b(x,y)\,dx\big) = d\alpha(x,y). \tag{2.266}$$

This function $\mu(x,y)$ is called an *integrating factor* for (2.264). As

$$\frac{\partial\alpha}{\partial x} = \mu b \qquad\qquad \text{and} \qquad\qquad \frac{\partial\alpha}{\partial y} = \mu a \tag{2.267}$$

the integrability condition for (2.267), $\alpha_{xy} = \alpha_{yx}$, yields an equation for determining the integrating factor

$$\frac{\partial(\mu a)}{\partial x} = \frac{\partial(\mu b)}{\partial y}. \tag{2.268}$$

In general, for an ordinary differential equation of order s,

$$a(x,y,y',y,...,y^{(s-1)})\,y^{(s)} + b(x,y,y',y,...,y^{(s-1)}) = 0, \tag{2.269}$$

a differential function $\mu(x,y,y',y,...,y^{(s-1)})$ is an integrating factor, if the multiplication by μ converts the left hand side of (2.269) into a total derivative of some function $\alpha(x,y,y',y,...,y^{(s-1)})$,

$$\mu a y^{(s)} + \mu b = D_x(\alpha). \tag{2.270}$$

Together, equations (2.269) and (2.270) imply $D_x(\alpha) = 0$, such that

$$\alpha(x,y,y',y,...,y^{(s-1)}) = const, \tag{2.271}$$

thereby reducing the order of the differential equation to solve. The integrating factor for (2.269) is determined by

$$\frac{\delta}{\delta y}\big(\mu a y^{(s)} + \mu b\big) = 0 \tag{2.272}$$

where $\delta/\delta y$ is the variational derivative.

Adjoint Equations

Consider a first order linear partial differential equation for a scalar field $u(x)$

$$\mathcal{D}[u] = a^\mu(x)\, u_\mu + b(x)\, u = f(x). \tag{2.273}$$

The first order linear differential operator $\mathcal{D}$, corresponding to this equation is

$$\mathcal{D}[u] = a^\mu(x)\, \partial_\mu + b(x). \tag{2.274}$$

The adjoint operator to $\mathcal{D}$ is a first-order linear differential operator $\mathcal{D}^*$ such that

$$v\,\mathcal{D}[u] - u\,\mathcal{D}^*[v] = \operatorname{div} C(x) \tag{2.275}$$

for all functions u and v and some vector field $C(x)$. The adjoint operator is uniquely determined

$$\mathcal{D}^*[v] = -\partial_\mu(a^\mu v) + bv. \tag{2.276}$$

It defines the adjoint equation to (2.273) by

$$\mathcal{D}^*[v] = -\partial_\mu(a^\mu v) + bv = 0. \tag{2.277}$$

If $\mathcal{D}[u] = \mathcal{D}^*[u]$ for any function u, the operator $\mathcal{D}$ is called self-adjoint.

All of these statements and definitions generalise straight forwardly to systems of m partial differential equations of arbitrary order s,

$$F_a(x, u, ..., u_{(s)}) = 0, \qquad\qquad a = 1, ..., m. \tag{2.278}$$

$F_a(x, u, ..., u_s)$ are differential functions with n independent variables x^μ and m dependent variables u^a. The system of adjoint equations to (2.278) is defined by

$$F_a^*(x, u, v, ..., u_{(s)}, v_{(s)}) = \frac{\delta(v^b F_b)}{\delta u^a} = 0, \qquad\qquad a = 1, ..., m, \tag{2.279}$$

where $v = v^a(x)$ are m new dependent variables, referred to as *auxiliary variables*. If the system obtained by substituting $v = u$ in the adjoint equations (2.279),

$$F_a^*(x, u, u, ..., u_{(s)}, u_{(s)}) = 0, \tag{2.280}$$

is identical with the original system (2.278), the system is called *self-adjoint*.

Extended Lagrangians

The extended system of differential equations, composed of the system of m partial differential equations of order s,

$$F_a(x, u, ..., u_{(s)}) = 0, \tag{2.281}$$

together with its adjoint equations,

$$F_a^*(x, u, v, ..., u_{(s)}, v_{(s)}) \equiv \frac{\delta(v^b F_b)}{\delta u^a} = 0, \tag{2.282}$$

has a Lagrangian given by

$$L = v^a F_a. \tag{2.283}$$

Obviously, the variation with respect to the auxiliary variables v^a, in this context also referred to as *Ibragimov multipliers*, yields the original equations (2.281),

$$\frac{\delta L}{\delta v^a} = F_a(x, u, ..., u_{(s)}). \tag{2.284}$$

Similarly, the variation with respect to the original variables u^a yields the adjoint equations (2.282),

$$\frac{\delta L}{\delta u^a} = F_a^*(x, u, v, ..., u_{(s)}, v_{(s)}). \tag{2.285}$$

The definition of the adjoint equations (2.279) suggested this result already.

Symmetries and Conserved Quantities

Here, we do not want to anticipate results from the next section, but a short comment seems appropriate. The original application Ibragimov had in mind for this method was the analysis of symmetries and conservation laws by applying Noether's theorem to systems without classical Lagrangian. However, an analysis of (2.283) will obviously lead to symmetries of the extended system of equations that are not necessarily symmetries of the original system. Therefore it is required to apply some appropriate restriction.

If the operator at hand is self-adjoint, the identification of the auxiliary variables v with the original variables u identifies the conserved flux of the extended system with the conserved flux of the original system at once. Most often, however, this is not the case. Ibragimov therefore defines the concept of quasi-self-adjointness [60], but for us a simpler idea suffices.

The symmetries of the extended system can be reduced to symmetries of the original system, if the auxiliary variables can be expressed with respect to the original variables and their derivatives, i.e.,

$$v^a = v^a(x, u, ..., u_{(s)}). \tag{2.286}$$

The adjoint equation (2.279) thereby becomes

$$F_a^*\big(x, u, v(x, u, ..., u_{(s)}), ..., u_{(s)}, v_{(s)}(x, u, ..., u_{(s)})\big) = 0 \tag{2.287}$$

and thus a function of x and u only. If it is possible to select the relation (2.286) such that (2.287) is automatically respected when u solves (2.281), then a conservation law for the extended system amounts to a physical conservation law.

Multisymplectic Form

The proof of preservation of the multisymplectic form along the Lagrangian flow from section 2.2.6 can be applied directly to extended Lagrangians. An open question is if the multisymplectic structure of the extended system can be restricted to the physical system, similar to the restriction of conservation laws of the extended system to conservation laws of the physical system.

It is however not clear if this is in general possible as the physical system is not necessarily Lagrangian and therefore might not even have a compatible multisymplectic structure. But in some cases, such as Hamiltonian systems like the Vlasov-Poisson system or ideal magnetohydrodynamics, the original system certainly has a multisymplectic structure, such that the development of a restriction method, that establishes a relation between the multisymplectic forms of the extended and the original system, appears worthwhile.

2.3. Noether Theorem

The Noether theorem [94, 63, 93] is one of the deepest and most influential insights of mathematical physics. It states that each continuous symmetry of a Lagrangian corresponds to a conservation law of the associated Euler-Lagrange equations and vice versa.

2.3.1. Point Transformations and One Parameter Groups

Consider an infinitesimal transformation on $\mathcal{Q}$ which is of the form

$$q(t) \to q^\epsilon(t) = \xi\big(q(t),\epsilon\big) = \xi^\epsilon\big(q(t)\big) \quad \text{with} \quad \xi^0 = \mathrm{id} \quad \text{such that} \quad q^0(t) = q(t). \qquad (2.288)$$

If the functional dependency of q^ϵ on the two parameters t and ϵ is of importance, we also write $q^\epsilon(t) = q(t,\epsilon)$. The transformation ξ maps each point $q \in \mathcal{Q}$ to a new point $q^\epsilon \in \mathcal{Q}$ and is therefore called a *point transformation*. It maps each trajectory $q(t) \in \mathcal{C}(\mathcal{Q})$ to a new trajectory $q^\epsilon(t) \in \mathcal{C}(\mathcal{Q})$. But what we are looking for are transformations of the Lagrangian and therefore transformations on $\mathsf{T}\mathcal{Q}$. Thus we need the tangent lift of ξ^ϵ which is defined as

$$\mathsf{T}\xi^\epsilon : (q^\epsilon) \mapsto (q^\epsilon, \dot{q}^\epsilon) \qquad \text{with} \qquad \dot{q}^\epsilon(t) = \frac{\partial}{\partial t} q(t,\epsilon). \qquad (2.289)$$

It maps each point $\dot{q} \in \mathsf{T}_q\mathcal{Q}$ to a new point $\dot{q}^\epsilon \in \mathsf{T}_{q^\epsilon}\mathcal{Q}$. Hence a transformation ξ^ϵ on $\mathcal{Q}$ induces a transformation $\mathsf{T}\xi^\epsilon$ on $\mathsf{T}\mathcal{Q}$.

We shall always assume that ξ^ϵ corresponds to a continuous family of transformations, such that the trajectories $q^\epsilon(t)$ are continuous in both, t and ϵ, and the transformation $\xi^\epsilon(q)$ constitutes a one-parameter group of curves.

Instead of specifying the transformation (2.288) directly, it can also be defined by its *generating vector field*

$$X(q,t) \equiv \frac{\partial q^\epsilon(t)}{\partial \epsilon}\bigg|_{\epsilon=0} \qquad \text{with} \qquad X = X^i \frac{\partial}{\partial q^i}, \qquad (2.290)$$

sometimes referred to as the infinitesimal symmetry direction.

2.3.2. Noether Theorem for Particle Systems

A Lagrangian has a symmetry if it is invariant under a point transformation $q^\epsilon(t) = \xi(q,\epsilon)$, that is

$$L\big(q^\epsilon(t), \dot{q}^\epsilon(t)\big) = L\big(q(t), \dot{q}(t)\big) \qquad\qquad \text{for all } \epsilon. \qquad (2.291)$$

This is equivalent to

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\big(q^\epsilon(t), \dot{q}^\epsilon(t)\big) = 0 \tag{2.292}$$

or explicitly

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\big(q^\epsilon, \dot{q}^\epsilon\big) = \frac{\partial L}{\partial q}\big(q, \dot{q}\big) \cdot X + \frac{\partial L}{\partial v}\big(q, \dot{q}\big) \cdot \dot{X} = 0. \tag{2.293}$$

If $q(t)$ solves the Euler-Lagrange equations (2.85), the first term on the right-hand side of (2.293) can be rewritten, such that the symmetry condition becomes

$$\left[\frac{d}{dt}\frac{\partial L}{\partial v}\big(q, \dot{q}\big)\right] \cdot X + \frac{\partial L}{\partial v}\big(q, \dot{q}\big) \cdot \dot{X} = 0. \tag{2.294}$$

This is a total time derivative and in fact a conservation law

$$\frac{d}{dt}\left[\frac{\partial L}{\partial v}\big(q, \dot{q}\big) \cdot X(q, t)\right] = 0. \tag{2.295}$$

It states that solutions q of the Euler-Lagrange equations preserve $\partial L/\partial v$ in direction X.

Example: Point Particle

Consider a simple transformation that amounts to a time-independent translation

$$q^\epsilon(t) = q(t) + \epsilon X, \qquad\qquad \dot{q}^\epsilon(t) = \dot{q}(t). \tag{2.296}$$

The corresponding transformed Lagrangian is

$$L\big(q^\epsilon(t), \dot{q}^\epsilon(t)\big) = \frac{m}{2}\big(\dot{q}(t)\big)^2 = L(q, \dot{q}) \tag{2.297}$$

which is obviously the same as the untransformed Lagrangian. The symmetry condition is therefore trivially fulfilled

$$\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0} L\big(q^\epsilon(t), \dot{q}^\epsilon(t)\big) = 0 \tag{2.298}$$

and the corresponding conservation law

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}}\big(q, \dot{q}\big) \cdot X\right] = \frac{d}{dt}\left[m\dot{q} \cdot X\right] = 0 \tag{2.299}$$

states that momentum is preserved in direction of X.

2.3.3. Noether Theorem for Field Theories

Now we want to generalise the Noether theorem for finite dimensional systems to infinite dimensional systems. Still we restrict to vertical transformations, i.e., transformations on the configuration space alone, as that will be sufficient for the following treatment.

The transformation of a (possibly vector valued) field y is described by a one-parameter group of transformations

$$\eta^\epsilon(x,y) = (x, y^\epsilon) \qquad \text{such that} \qquad \eta^0(x,y) = (x,y). \tag{2.300}$$

or in coordinates

$$\eta : (x,y) \mapsto \big(x^\nu, \eta^a(x,y,\epsilon)\big). \tag{2.301}$$

The infinitesimal generator of the transformation η is

$$X(x,y) = \left.\frac{d\eta}{d\epsilon}\right|_{\epsilon=0} = X^a \frac{\partial}{\partial y^a} \qquad \text{with components} \qquad X^a(x,y) = \left.\frac{d\eta^a}{d\epsilon}\right|_{\epsilon=0}, \tag{2.302}$$

and its first jet prolongation is given by

$$j^1 X(x,y,v) = X^a \frac{\partial}{\partial y^a} + \frac{\partial X^a}{\partial y^b} \frac{\partial y^b}{\partial x^\nu} \frac{\partial}{\partial v_\nu^a}. \tag{2.303}$$

The Lagrangian has a symmetry if it is invariant under this transformation

$$L\big(x^\nu, \eta^a(x,y), \eta_\nu^a(x,y,v)\big) = L\big(x^\nu, y^a, v_\nu^a\big) \qquad \text{for all } \epsilon. \tag{2.304}$$

This is equivalent to

$$j^1 X(L) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} L\big(x^\nu, \eta^a(x,y), \eta_\nu^a(x,y,v)\big) = 0 \tag{2.305}$$

or explicitly

$$j^1 X(L) = \frac{\partial L}{\partial y^a}(x,y,v) \cdot X^a + \frac{\partial L}{\partial v_\nu^a}(x,y,v) \cdot X_\nu^a = 0. \tag{2.306}$$

If $y(x)$ solves the Euler-Lagrange field equations (2.169), we can replace the first term on right-hand-side and obtain

$$j^1 X(L) = \left[\frac{\partial}{\partial x^\nu}\left(\frac{\partial L}{\partial v_\nu^a}\right)\right] \cdot X^a + \frac{\partial L}{\partial v_\nu^a} \cdot \left[\frac{\partial}{\partial x^\nu} X^a\right] = 0. \tag{2.307}$$

We immediately see that this is a divergence

$$j^1 X(L) = \frac{\partial}{\partial x^\nu}\left[\frac{\partial L}{\partial v_\nu^a} \cdot X^a\right] = 0. \tag{2.308}$$

The term in square brackets is called *Noether field*. Integration in the spatial dimensions yields a conservation law for solutions of the discrete Euler-Lagrange field equations,

$$\frac{d}{dt} \int \frac{\partial L}{\partial v_t^a} \cdot X^a \, dx = 0, \tag{2.309}$$

assuming that the boundary terms vanish.

2.3.4. Noether Theorem for Extended Lagrangians

Consider the generator of a transformation of the configuration bundle for a field u as in (2.302),

$$X(x,u) = X^u \frac{\partial}{\partial u}. \tag{2.310}$$

To be applied to the Lagrangian of an extended system of equations (2.283), this generator has to be extended to the auxiliary variable v by a certain function X^v as

$$Y(x,u,v) = X^u \frac{\partial}{\partial u} + X^v \frac{\partial}{\partial v}. \tag{2.311}$$

Everything else follows by applying the theory of the last section to the extended Lagrangian, considered as describing a theory of two fields (u,v). Only in the last step of integrating the Noether field, one has to find a relation between the auxiliary field v and the original field u, i.e., one has to find a functional expression for v in terms of u and its derivatives like in (2.286).

3. Variational Integrators

The seminal work in the development of a discrete equivalent of classical mechanics was presented by Veselov [132, 133]. His method, based on a discrete variational principle, leads to symplectic integration schemes that automatically preserve constants of motion. A comprehensive review of discrete mechanics can be found in Marsden and West [83]. This includes also a more thorough account on the historical development. The theory was extended to partial differential equations in form of first order Lagrangian theories by Marsden et al. [84] and soon also to second order Lagrangian theories by Kouranbaeva and Shkoller [65, 64]. Another extension was that of asynchronous variational integrators by Lew et al. [76], where each point in the spatial grid has its own timestep. This way, exact local energy conservation can be achieved in addition to an often observed speedup in runtime. In more recent developments, the variational integrator method was applied to Maxwell's equations by Stern et al. [123, 124] and fluid problems by Pavlov et al. [99, 98] and Gawlik et al. [43].

This chapter gives an overview of the theory of variational integrators for finite-dimensional as well as for infinite-dimensional systems. It follows mostly along the lines of Marsden and West [83], Marsden et al. [84], Kouranbaeva and Shkoller [65], and Kouranbaeva [64].

3.1. Discrete Particle Dynamics

The derivation of the discrete theory follows along the lines of the derivation of the continuous theory. The starting point is the discretisation of the action integral and the Lagrangian. There is some degree of freedom in the choice of the discrete quadrature rule as well as in the approximation of the generalised coordinates q and the generalised velocities $\dot{q}$. Everything else follows in a straight forward way, so that these choices are determining the respective form of the discrete Lagrangian as well as the resulting discrete equations of motion.

Time will be discretised uniformly, i.e., the timestep h is constant, q_k denotes the generalised coordinates at timepoint k, and $\dot{q}_k$ the generalised velocities at timepoint k. The discrete Lagrangian approximates the time integral of the continuous Lagrangian between two consecutive points in time, k and $k+1$

$$L_d(q_k, q_{k+1}) \approx \int_{t_k}^{t_{k+1}} L(q, \dot{q}) \, dt. \tag{3.1}$$

Its exact expression is determined by the quadrature rule used to approximate the integral. Here, we assume that the quadrature rule depends only on (q_k, q_{k+1}). The discrete action thus becomes merely a sum over the time index of discrete Lagrangians

$$\mathcal{A}_d = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}). \tag{3.2}$$
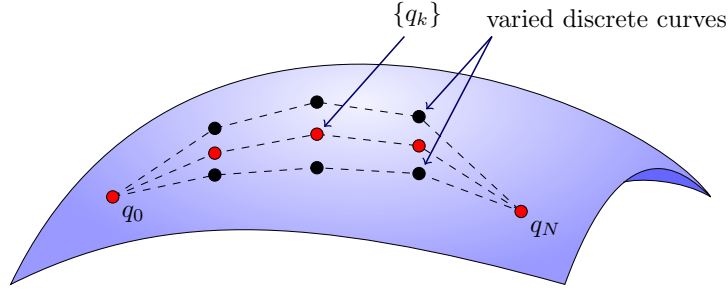
Fig. 3.1.: Variations of the discrete trajectory $\{q_k\}_{k=0}^N$.

The generalised velocities will usually be discretised by simple finite-difference expressions[1], i.e.

$$\dot{q} \approx \frac{q_{k+1} - q_k}{h} \qquad\qquad \text{for} \qquad\qquad t \in [t_k, t_{k+1}]. \qquad (3.3)$$

The quadrature (3.1) is most often realised by either the trapezoidal rule

$$L_d^{\text{tr}}(q_k, q_{k+1}) = \frac{h}{2} L\left(q_k, \frac{q_{k+1} - q_k}{h}\right) + \frac{h}{2} L\left(q_{k+1}, \frac{q_{k+1} - q_k}{h}\right) \qquad (3.4)$$

or the midpoint rule

$$L_d^{\text{mp}}(q_k, q_{k+1}) = h\, L\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}\right). \qquad (3.5)$$

The configuration manifold of the discrete theory is still $\mathcal{Q}$, but the discrete state space is $\mathcal{Q} \times \mathcal{Q}$ instead of $\mathsf{T}\mathcal{Q}$, such that the discrete Lagrangian L_d is a function

$$L_d : \mathcal{Q} \times \mathcal{Q} \to \mathbb{R}. \qquad (3.6)$$

3.1.1. Discrete Action Principle

The discrete trajectories $q_d = \{q_k\}_{k=0}^N$ are required to satisfy a discrete version of Hamilton's principle of least action

$$\delta\mathcal{A}_d[q_d] = \delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) = 0. \qquad (3.7)$$

The variation of the action is

$$\delta\mathcal{A}_d[q_d] = \sum_{k=0}^{N-1} \left[D_1 L_d(q_k, q_{k+1}) \cdot \delta q_k + D_2 L_d(q_k, q_{k+1}) \cdot \delta q_{k+1} \right] \qquad (3.8)$$

where D_i denotes the derivative with respect to to the ith argument. What follows corresponds to a discrete integration by parts, i.e., a reordering of the summation. The

---

[1] In the first term of the trapezoidal rule (3.4), this corresponds to a forward finite-difference, in the second term to a backward finite-difference, and in the midpoint rule (3.5) to a centred finite-difference.

$k = 0$ term is removed from the first part of the sum and the $k = N - 1$ term is removed from the second part

$$\delta\mathcal{A}_d[q_d] = D_1\, L_d(q_0, q_1) \cdot \delta q_0 + \sum_{k=1}^{N-1} D_1\, L_d(q_k, q_{k+1}) \cdot \delta q_k$$

$$+ \sum_{k=0}^{N-2} D_2\, L_d(q_k, q_{k+1}) \cdot \delta q_{k+1} + D_2\, L_d(q_{N-1}, q_N) \cdot \delta q_N. \tag{3.9}$$

As the variations at the endpoints, δq_0 and δq_N, are kept fixed, the corresponding terms vanish. At last, the summation range of the second sum is shifted upwards by one with the arguments of the discrete Lagrangian adapted correspondingly

$$\delta\mathcal{A}_d[q_d] = \sum_{k=1}^{N-1} \left[D_1\, L_d(q_k, q_{k+1}) + D_2\, L_d(q_{k-1}, q_k) \right] \cdot \delta q_k. \tag{3.10}$$

Hamilton's principle of least action requires the variation of the discrete action $\delta\mathcal{A}_d$ to vanish for any choice of δq_k. Consequently, the expression in the square brackets of (3.10) has to vanish. This defines the

Discrete Euler-Lagrange Equations

$$D_1\, L_d(q_k, q_{k+1}) + D_2\, L_d(q_{k-1}, q_k) = 0 \tag{3.11}$$

The discrete Euler-Lagrange equations (3.11) define an evolution map

$$F_{L_d} \,:\, \mathcal{Q} \times \mathcal{Q} \to \mathcal{Q} \times \mathcal{Q} \,:\, (q_{k-1}, q_k) \mapsto (q_k, q_{k+1}). \tag{3.12}$$

Starting from two configurations, $q_0 \approx q(t_0)$ and $q_1 \approx q(t_1 = t_0 + h)$, the successive solution of the discrete Euler-Lagrange equations (3.11) for q_2, q_3, etc., up to q_N, determines the discrete trajectory $\{q_k\}_{k=0}^N$. Quite often, however, it is more practical to prescribe an initial position and momentum instead of the configuration of the first two timesteps. We therefore define the discrete momentum at timestep k as[2]

$$p_k = -D_1 L_d(q_k, q_{k+1}) = D_2 L_d(q_{k-1}, q_k), \tag{3.13}$$

such that the variational integrator (3.11) can be rewritten in

Position-Momentum Form

$$p_k = -D_1 L_d(q_k, q_{k+1}) \tag{3.14a}$$
$$p_{k+1} = D_2 L_d(q_k, q_{k+1}) \tag{3.14b}$$

Given (q_k, p_k), the first equation can be solved for q_{k+1}. This is generally a nonlinearly implicit equation that has to be solved by some iterative technique like Newton's method. The second equation is an explicit function, so to obtain p_{k+1} we merely have to plug in q_k and q_{k+1}. The corresponding Hamiltonian evolution map is

$$\tilde{F}_{L_d} \,:\, \mathsf{T}^*\mathcal{Q} \to \mathsf{T}^*\mathcal{Q} \,:\, (q_k, p_k) \mapsto (q_{k+1}, p_{k+1}). \tag{3.15}$$

---

[2] The two expressions are equal by the discrete Euler-Lagrange equations (3.11).

Thus, starting with an initial position q_0 and an initial momentum p_0, the repeated solution of (3.14) gives the same discrete trajectory $\{q_k\}_{k=0}^{N}$ as (3.11). The position-momentum form, as a one-step method, is usually easier to implement than the discrete Euler-Lagrange equations (3.11). And for most problems, initial conditions are more naturally prescribed via the position and momentum of the particle at a given point in time, (q_0, p_0). If, however, only the position of the particle at two points in time, (q_0, q_1), is known, the Euler-Lagrange equations (3.11) are the more natural way of describing the dynamics.

This of course is just reflecting the difference in the Lagrangian and Hamiltonian point of view. For n degrees of freedom, the variational principle leads to n differential equations of second order. Hamilton's equations, on the other hand, are $2n$ differential equations of first order. Which form eventually is used largely depends on the problem at hand.

3.1.2. Discrete Tangent Space

The discrete path space is defined as

$$\mathcal{C}_d(\mathcal{Q}) = \left\{ c_d : \{t_k\}_{k=0}^{N} \to \mathcal{Q} \right\} \tag{3.24}$$

where

$$\{t_k\}_{k=0}^{N} = \{ t_k = kh \mid k = 0, ..., N \} \subset \mathbb{R} \tag{3.25}$$

is an increasing sequence of times and h is the discrete timestep. $\mathcal{C}_d(\mathcal{Q})$ contains all possible discrete trajectories c_d in $\mathcal{Q}$ and is isomorphic to $\mathcal{Q} \times ... \times \mathcal{Q}$ ($N+1$ copies)

$$\mathcal{C}_d(\mathcal{Q}) \cong \underset{N+1}{\bigtimes} \mathcal{Q}. \tag{3.26}$$

The subspace of $\mathcal{C}_d(\mathcal{Q})$ that contains all discrete trajectories with fixed endpoints q_0 and q_N is defined as

$$\mathcal{C}_d(q_0, q_N, \{t_k\}_{k=0}^{N}) = \left\{ c_d : \{t_k\}_{k=0}^{N} \to \mathcal{Q} \mid c(t_0) = q_0, c(t_N) = q_N \right\}. \tag{3.27}$$

The discrete action map $\mathcal{A}_d : \mathcal{C}_d(q_0, q_N, \{t_k\}_{k=0}^{N}) \to \mathbb{R}$ is defined as

$$\mathcal{A}_d[c_d] = \sum_{k=0}^{N-1} L_d\big(c(t_k), c(t_{k+1})\big) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) \qquad \text{where} \qquad q_k = c(t_k). \tag{3.28}$$

The tangent space $\mathsf{T}_{c_d}\mathcal{C}_d(q_0, q_N, \{t_k\}_{k=0}^{N})$ to $\mathcal{C}_d(q_0, q_N, \{t_k\}_{k=0}^{N})$ at c_d contains the variations of c_d. It is defined as the set of maps

$$v_{c_d} : \{t_k\}_{k=0}^{N} \to \mathsf{T}\mathcal{Q} \qquad \text{such that} \qquad \pi_{\mathcal{Q}} \circ v_{c_d} = c_d \qquad \text{and} \qquad v(t_0) = v(t_N) = 0, \tag{3.29}$$

where $\pi_{\mathcal{Q}}$ is the canonical projection $\pi_{\mathcal{Q}} : \mathsf{T}\mathcal{Q} \to \mathcal{Q}$ and local coordinates are given by

$$v_{c_d} = \{(q_k, v_k)\}_{k=0}^{N}. \tag{3.30}$$

v_{c_d} is called a discrete variation of the discrete path c_d and sometimes denoted $v_{c_d} = \delta c_d$. The variation of the discrete action can therefore be formulated as

$$\mathbf{d}\mathcal{A}_d[c_d] \cdot v_{c_d} = \sum_{k=0}^{N-1} \left[D_1 L(q_k, q_{k+1}) \cdot v_k + D_2 L(q_k, q_{k+1}) \cdot v_{k+1} \right]. \tag{3.31}$$

Example: Point Particle

Consider a particle with mass m, moving in some potential V. Its continuous Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2} m \dot{q}^2 - V(q) \tag{3.16}$$

Approximated by the trapezoidal rule, the discrete Lagrangian reads

$$L_d^{\text{tr}}(q_k, q_{k+1}) = h \left[\frac{m}{2} \left(\frac{q_{k+1} - q_k}{h} \right)^2 - \frac{V(q_k) + V(q_{k+1})}{2} \right]. \tag{3.17}$$

Applying the discrete Euler-Lagrange equations (3.11) to this expression results in discrete equations of motion

$$m \frac{q_{k+1} - 2 q_k + q_{k-1}}{h^2} = -\nabla V(q_k) \tag{3.18}$$

which clearly are a discrete version of Newton's second law

$$m \ddot{q} = -\nabla V = F. \tag{3.19}$$

For comparison, consider also the midpoint approximation

$$L_d^{\text{mp}}(q_k, q_{k+1}) = h \left[\frac{m}{2} \left(\frac{q_{k+1} - q_k}{h} \right)^2 - V \left(\frac{q_k + q_{k+1}}{2} \right) \right] \tag{3.20}$$

which leads to

$$m \frac{q_{k+1} - 2 q_k + q_{k-1}}{h^2} = -\frac{1}{2} \left[\nabla V \left(\frac{q_{k-1} + q_k}{2} \right) + \nabla V \left(\frac{q_k + q_{k+1}}{2} \right) \right] \tag{3.21}$$

and thus a different discretisation of (3.19). The position-momentum form (3.14) of the trapezoidal Lagrangian (3.17) can be written as

$$\frac{q_{k+1} - q_k}{h} = \frac{1}{m} \left[p_k - \frac{h}{2} \nabla V(q_k) \right] \tag{3.22a}$$

$$\frac{p_{k+1} - p_k}{h} = -\frac{1}{2} \left[\nabla V(q_k) + \nabla V(q_{k+1}) \right] \tag{3.22b}$$

and the one of the midpoint Lagrangian (3.20) reads

$$\frac{q_{k+1} - q_k}{h} = \frac{1}{m} \left[p_k - \frac{h}{2} \nabla V \left(\frac{q_k + q_{k+1}}{2} \right) \right] \tag{3.23a}$$

$$\frac{p_{k+1} - p_k}{h} = -\nabla V \left(\frac{q_k + q_{k+1}}{2} \right). \tag{3.23b}$$

This bears a close resemblance of Hamilton's equations of motion, where the additional term in the first equations can be interpreted as extrapolating the momentum p_k to $t_{k+1/2}$. As already noted, it is not always so easy to solve (3.14a) for q_{k+1}, but in general this is an implicit equation.

A reordering of the sum (discrete partial integration) leads to

$$
\mathbf{d}\mathcal{A}_d[c_d] \cdot v_{c_d} = \sum_{k=1}^{N-1} \big[D_1 L(q_k, q_{k+1}) + D_2 L(q_{k-1}, q_k) \big] \cdot v_k
$$
$$
+ D_1 L(q_0, q_1) \cdot v_0 + D_2 L(q_{N-1}, q_N) \cdot v_N. \tag{3.32}
$$

where the terms in the second line vanish as $v_0 = v(t_0) = 0$ as well as $v_N = v(t_N) = 0$ and thus

$$
\mathbf{d}\mathcal{A}_d[c_d] \cdot v_{c_d} = \sum_{k=1}^{N-1} \big[D_1 L(q_k, q_{k+1}) + D_2 L(q_{k-1}, q_k) \big] \cdot v_k. \tag{3.33}
$$

The arbitrariness of the v_k once more yields the discrete Euler-Lagrange equations.

3.1.3. Discrete One- and Two-Form

As in the continuous case, the discrete one-form is obtained by computing the variation of the action for varying endpoints

$$
\mathbf{d}\mathcal{A}_d[q_d] \cdot \delta q_d = \sum_{k=0}^{N-1} \big[D_1 L_d(q_k, q_{k+1}) \cdot \delta q_k + D_2 L_d(q_k, q_{k+1}) \cdot \delta q_{k+1} \big]
$$
$$
= \sum_{k=1}^{N-1} \big[D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) \big] \cdot \delta q_k
$$
$$
+ D_1 L_d(q_0, q_1) \cdot \delta q_0 + D_2 L_d(q_{N-1}, q_N) \cdot \delta q_N. \tag{3.34}
$$

The two latter terms originate from the variation at the boundaries. They form the discrete counterpart of the Lagrangian one-form. However, there are two boundary terms that define two distinct one-forms on $\mathcal{Q} \times \mathcal{Q}$

$$
\begin{aligned}
\Theta_{L_d}^-(q_0, q_1) \cdot (\delta q_0, \delta q_1) &\equiv -D_1 L_d(q_0, q_1) \cdot \delta q_0, \\
\Theta_{L_d}^+(q_{N-1}, q_N) \cdot (\delta q_{N-1}, \delta q_N) &\equiv D_2 L_d(q_{N-1}, q_N) \cdot \delta q_N.
\end{aligned} \tag{3.35}
$$

In general, these one-forms are defined as

$$
\begin{aligned}
\Theta_{L_d}^-(q_k, q_{k+1}) &\equiv -D_1 L(q_k, q_{k+1}), \\
\Theta_{L_d}^+(q_k, q_{k+1}) &\equiv D_2 L(q_k, q_{k+1}).
\end{aligned} \tag{3.36}
$$

As $\mathbf{d}L_d = \Theta_{L_d}^+ - \Theta_{L_d}^-$ and $\mathbf{d}^2 L_d = 0$ one observes that

$$
\mathbf{d}\Theta_{L_d}^+ = \mathbf{d}\Theta_{L_d}^- \tag{3.37}
$$

such that the exterior derivative of both discrete one-forms defines the same *discrete Lagrangian two-form* or *discrete symplectic form*

$$
\Omega_{L_d} = \mathbf{d}\Theta_{L_d}^+ = \mathbf{d}\Theta_{L_d}^- = \frac{\partial^2 L_d}{\partial q_k \, \partial q_{k+1}}(q_k, q_{k+1}) \, dq_k \wedge dq_{k+1} \qquad \text{(no summation)}. \tag{3.38}
$$

3.1.4. Preservation of the Discrete Symplectic Form

Consider the exterior derivative of the discrete action (3.2). Upon insertion of the discrete Euler-Lagrange equations (3.11) it becomes

$$\mathbf{d}\mathcal{A}_d = D_1 L_d(q_0, q_1) \cdot \mathbf{d}q_0 + D_2 L_d(q_{N-1}, q_N) \cdot \mathbf{d}q_N = \Theta_{L_d}^+(q_{N-1}, q_N) - \Theta_{L_d}^-(q_0, q_1). \quad (3.39)$$

On the right hand side we find the just defined Lagrangian one-forms (3.36). Taking the exterior derivative of (3.39) gives

$$\Omega_d(q_0, q_1) = \Omega_d(q_{N-1}, q_N), \quad (3.40)$$

where q_{N-1} and q_N are connected with q_0 and q_1 through the discrete Euler-Lagrange equations (3.11). Therefore, (3.40) implies that the discrete symplectic structure Ω_d is preserved while the system advances from $t = 0$ to $t = Nh$ according to the discrete equations of motion (3.11). As the number of timesteps N is arbitrary, the discrete symplectic form Ω_d is preserved at all times of the simulation. Note that this does not automatically imply that the continuous symplectic structure Ω is preserved under the discrete map F_{L_d}.

3.1.5. Composition Methods

The composition of a one-step variational integrator $\tilde{F}_{L_d}$ with different step sizes is a simple method of obtaining higher order schemes. We assume that the initial scheme $\tilde{F}_{L_d}$ is symmetric, that is

$$\tilde{F}_{L_d}^h \circ \tilde{F}_{L_d}^{-h} = \mathrm{id}, \quad (3.41)$$

as this simplifies the construction. Alternatively, efficient methods can also be built by combining a non-symmetric method with its adjoint. The interested reader can find more information on these issues in Marsden and West [83] and Hairer et al. [52]. If a numerical method

$$\tilde{F}_{L_d} : \mathsf{T}^* \mathcal{Q} \times \mathbb{R} \to \mathsf{T}^* \mathcal{Q} \quad (3.42)$$

is symmetric, it can be used to compose higher order methods by splitting up each timestep into s substeps [52, 87, 83]

$$\hat{F}_{L_d}^h = \tilde{F}_{L_d}^{\gamma_s h} \circ \dots \circ \tilde{F}_{L_d}^{\gamma_i h} \circ \dots \circ \tilde{F}_{L_d}^{\gamma_1 h} \quad (3.43)$$

where the careful selection of the γ_i is crucial for the performance of the resulting scheme. In this section, we show that a variational integrator is self-adjoint and thereby symmetric if its discrete Lagrangian is self-adjoint, a condition that is easily checked. We present some fourth and sixth order composition methods that can be applied in most situations. It is worth mentioning that the composition can already be implemented at the level of the Lagrangian. We will outline this at the end of the section, and later on, when we come to the discrete Noether theorem, the consequences for the discrete conservation laws are described.

Adjoint of a Method and Adjoint Lagrangians

In this subsection, we show that a variational integrator is symmetric if its Lagrangian is self-adjoint. The adjoint $\tilde{F}^*_{L_d}$ of a method $\tilde{F}_{L_d}$ is defined as

$$\left(\tilde{F}^*_{L_d}(h)\right) \circ \left(\tilde{F}_{L_d}(-h)\right) = \mathrm{id}. \tag{3.44}$$

A method $\tilde{F}_{L_d}$ is self-adjoint if $\tilde{F}^*_{L_d} = \tilde{F}_{L_d}$, therefore a self-adjoint method is also symmetric. We would like to establish a condition of the discrete Lagrangian that tells us if the resulting method is self-adjoint or not. We therefore define the adjoint Lagrangian L^*_d of a discrete Lagrangian L_d as

$$L^*_d(q_k, q_{k+1}, h) \equiv -L_d(q_{k+1}, q_k, -h). \tag{3.45}$$

Hence the Lagrangian L_d is self-adjoint if

$$L_d(q_k, q_{k+1}, h) = -L_d(q_{k+1}, q_k, -h). \tag{3.46}$$

We want to show that if a discrete Lagrangian is self-adjoint so is the resulting method. We start by establishing that adjoint Lagrangians admit adjoint methods, i.e., if $\tilde{F}_{L_d}$ is the Hamiltonian map resulting from L_d and $\tilde{F}_{L^*_d}$ is the map resulting from L^*_d then $\tilde{F}^*_{L_d} = \tilde{F}_{L^*_d}$. In position momentum-form (3.14), the map $\tilde{F}_{L_d}$ is defined as

$$\tilde{F}_{L_d} : \begin{cases} p_k & = -D_1 L_d(q_k, q_{k+1}, h) \\ p_{k+1} & = D_2 L_d(q_k, q_{k+1}, h) \end{cases}. \tag{3.47}$$

Its adjoint method $\tilde{F}^*_{L_d}(h) = \left(\tilde{F}_{L_d}(-h)\right)^{-1}$ is the map

$$\tilde{F}^*_{L_d} : \begin{cases} p_k & = D_2 L_d(q_{k+1}, q_k, -h) \\ p_{k+1} & = -D_1 L_d(q_{k+1}, q_k, -h) \end{cases}. \tag{3.48}$$

And the map $\tilde{F}_{L^*_d}$ corresponding to the adjoint Lagrangian (3.45) is

$$\tilde{F}_{L^*_d} : \begin{cases} p_k & = -D_1 L^*_d(q_k, q_{k+1}, h) \\ p_{k+1} & = D_2 L^*_d(q_k, q_{k+1}, h) \end{cases}. \tag{3.49}$$

Computing the derivatives of the definition of the adjoint Lagrangian (3.45)

$$\begin{aligned} D_1 L^*_d(q_k, q_{k+1}, h) &= -D_2 L_d(q_{k+1}, q_k, -h) \\ -D_2 L^*_d(q_k, q_{k+1}, h) &= D_1 L_d(q_{k+1}, q_k, -h) \end{aligned} \tag{3.50}$$

establishes the equality of (3.48) and (3.49), i.e., if two Lagrangians are adjoint so are the resulting methods. Computing the derivatives of the definition of the self-adjoint Lagrangian (3.46)

$$D_1 L_d(q_k, q_{k+1}, h) = -D_2 L_d(q_{k+1}, q_k, -h) \tag{3.51}$$

$$-D_2 L_d(q_k, q_{k+1}, h) = D_1 L_d(q_{k+1}, q_k, -h) \tag{3.52}$$

establishes the equality of (3.47) and (3.48), i.e., if the Lagrangian is self-adjoint so is the resulting method. We have thereby obtained a condition for symmetry of a variational integrator that can easily be checked.

Fourth Order Composition Methods

If $\tilde{F}_{L_d}$ is a method of order r, a method $\hat{F}_{L_d}$ of order $r+2$ is obtained by the composition [52]

$$\hat{F}_{L_d}^h = \tilde{F}_{L_d}^{\gamma h} \circ \tilde{F}_{L_d}^{(1-2\gamma)h} \circ \tilde{F}_{L_d}^{\gamma h} \qquad \text{with} \qquad \gamma = \left(2 - 2^{1/(r+1)}\right)^{-1}. \qquad (3.53)$$

Hence, if $\tilde{F}_{L_d}$ is of second order, the resulting method $\hat{F}_{L_d}$ will be of fourth order. Note that symmetric methods are always of even order (for details see Marsden and West [83]). A method of the same order but with generally smaller errors is obtained by considering five steps

$$\hat{F}_{L_d}^h = \tilde{F}_{L_d}^{\gamma h} \circ \tilde{F}_{L_d}^{\gamma h} \circ \tilde{F}_{L_d}^{(1-4\gamma)h} \circ \tilde{F}_{L_d}^{\gamma h} \circ \tilde{F}_{L_d}^{\gamma h} \qquad \text{with} \qquad \gamma = \left(4 - 4^{1/(r+1)}\right)^{-1}. \qquad (3.54)$$

Multiple application of these compositions yields methods of orders higher than four.

Sixth Order Composition Methods

Higher order compositions can also be constructed directly (see Hairer et al. [52], section 3.2). A sixth order method with seven substeps is given by

$$\begin{aligned}
\gamma_1 = \gamma_7 &= +0.78451361047755726381949763, \\
\gamma_2 = \gamma_6 &= +0.23557321335935813368479318, \\
\gamma_3 = \gamma_5 &= -1.17767998417887100694641568, \\
\gamma_4 &= +1.31518632068391121888424973,
\end{aligned} \qquad (3.55)$$

but again smaller errors can be achieved by using nine steps

$$\begin{aligned}
\gamma_1 = \gamma_9 &= +0.39216144400731413927925056, \\
\gamma_2 = \gamma_8 &= +0.33259913678935943859974864, \\
\gamma_3 = \gamma_7 &= -0.70624617255763935980996482, \\
\gamma_4 = \gamma_6 &= +0.08221359629355080023149045, \\
\gamma_5 &= +0.79854399093482996339895035.
\end{aligned} \qquad (3.56)$$

The computational effort of these high order methods is quite large. Each step requires the solution of a nonlinear system of equations. Given the outstanding performance already second order variational integrators are able to deliver, the necessity for such high order methods is probably rarely found. Nevertheless, if extremely high accuracy is indispensable, these methods can be applied.

Composite Discrete Lagrangians

The composition schemes presented can all be derived as Euler-Lagrange equations from a composite discrete Lagrangian. There are several equivalent possibilities of constructing such a Lagrangian and the corresponding discrete Euler-Lagrange equations. We will present only one, for details on the alternatives see Marsden and West [83], section 2.5. The composite discrete Lagrangian of a method

$$\hat{F}_{L_d}^h = \tilde{F}_{L_d^s}^{\gamma_s h} \circ ... \circ \tilde{F}_{L_d^i}^{\gamma_i h} \circ ... \circ \tilde{F}_{L_d^1}^{\gamma_1 h}. \qquad (3.57)$$

with s substeps can be written as

$$\hat{L}_d(q_k^0, q_k^1, ..., q_k^s) = \sum_{i=1}^s L_d^i(q_k^{i-1}, q_k^i, \gamma_i h) \tag{3.58}$$

where we identify $q_k = q_k^0$ and $q_{k+1} = q_k^s$ such that $q_k^s = q_{k+1}^0$. The discrete action becomes

$$\mathcal{A}_d\big(\{q_k^0, q_k^1, ..., q_k^s\}_{k=0}^{N-1}\big) = \sum_{k=0}^{N-1} \hat{L}_d(q_k^0, q_k^1, ..., q_k^s) \tag{3.59}$$

and we obtain discrete Euler-Lagrange equations

$$D_2 L_d^s(q_{k-1}^{s-1}, q_{k-1}^s, \gamma_s h) + D_1 L_d^i(q_k^0, q_k^1, \gamma_1 h) = 0 \tag{3.60}$$

$$\vdots \tag{3.61}$$

$$D_2 L_d^i(q_k^{i-1}, q_k^i, \gamma_i h) + D_1 L_d^{i+1}(q_k^i, q_k^{i+1}, \gamma_{i+1} h) = 0 \tag{3.62}$$

$$\vdots \tag{3.63}$$

$$D_2 L_d^s(q_k^{s-1}, q_k^s, \gamma_s h) + D_1 L_d^1(q_{k+1}^0, q_{k+1}^1, \gamma_1 h) = 0. \tag{3.64}$$

The maps $\tilde{F}_{L_d^i}^{\gamma_i h}$ in the composition method (3.57) can therefore be written as

$$\tilde{F}_{L_d^i}^{\gamma_i h} : (q_k^{i-1}, p_k^{i-1}) \mapsto (q_k^i, p_k^i) \tag{3.65}$$

with

$$p_k^{i-1} = -D_1 L_d^i(q_k^{i-1}, q_k^i, \gamma_i h) \tag{3.66a}$$

$$p_k^i \;\; = \;\; D_2 L_d^i(q_k^{i-1}, q_k^i, \gamma_i h). \tag{3.66b}$$

The existence of a composite Lagrangian $\hat{L}_d$ corresponding to a composite method $\hat{F}_{L_d}$ is important in the analysis of conserved quantities. The discrete Noether theorem (next section) has to be applied to the composite Lagrangian to determine the quantities that are discretely conserved to the order of the composition method. It cannot be expected that the errors of the conserved quantities of the discrete Lagrangians L_d^i, which are used to build the composition scheme, scale with the order of the composition scheme.

3.1.6. Discrete Noether Theorem

The discrete Noether theorem, just as the continous Noether theorem, draws the connection between symmetries of a discrete Lagrangian and quantities that are conserved by the discrete Euler-Lagrange equations or, equivalently, the discrete Lagrangian flow. The continuous theory translates straight forwardly to the discrete case. Therefore, we repeat just the important steps, translated to the discrete setting.

Discrete Noether Theorem for Particle Systems

Consider a one parameter group of discrete curves $\{q_k^\epsilon\}_{k=0}^N$ such that $q_k^0(q_k) = q_k$. The discrete Lagrangian L_d has a symmetry if it is invariant under this transformation

$$L_d\big(q_k^\epsilon, q_{k+1}^\epsilon\big) = L_d\big(q_k, q_{k+1}\big) \qquad\qquad \text{for all } \epsilon \text{ and } k. \tag{3.67}$$

The direction of such a symmetry is

$$X_k = \left. \frac{\partial q_k^\epsilon}{\partial \epsilon} \right|_{\epsilon=0} \tag{3.68}$$

such that

$$\left. \frac{d}{d\epsilon} L_d\big(q_k^\epsilon, q_{k+1}^\epsilon\big) \right|_{\epsilon=0} = D_1 L_d\big(q_k, q_{k+1}\big) \cdot X_k + D_2 L_d\big(q_k, q_{k+1}\big) \cdot X_{k+1}. \tag{3.69}$$

If $\{q_k\}$ solves the discrete Euler-Lagrange equations

$$D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = 0 \tag{3.70}$$

we can replace the first term on the right hand side of (3.69) to get

$$0 = -D_2 L_d\big(q_{k-1}, q_k\big) \cdot X_k + D_2 L_d\big(q_k, q_{k+1}\big) \cdot X_{k+1}. \tag{3.71}$$

This amounts to a discrete conservation law of the form

$$D_2 L_d\big(q_{k-1}, q_k\big) \cdot X_k = D_2 L_d\big(q_k, q_{k+1}\big) \cdot X_{k+1} \quad \textbf{(Discrete Noether Theorem)}. \tag{3.72}$$

It states that solutions $\{q_k\}$ of the discrete Euler-Lagrange equations preserve the components of the momentum map $p_k = D_2 L_d\big(q_{k-1}, q_k\big)$ in direction X_k.

Example: Free Point Particle

Consider a transformation that amounts to an infinitesimal spatial translation

$$q_k^\epsilon = q_k + \epsilon X. \tag{3.73}$$

The discrete Lagrangian is invariant under this transformation

$$L_d(q^{1,\epsilon}, q^{2\epsilon}) = \frac{h}{2} \left(\frac{q^2 + \epsilon X - q^1 - \epsilon X}{h} \right)^2 = \frac{h}{2} \left(\frac{q^2 - q^1}{h} \right)^2 = L_d(q^1, q^2), \tag{3.74}$$

such that the symmetry condition is trivially fulfilled

$$\left. \frac{\partial}{\partial \epsilon} L_d\big(q_k^\epsilon, q_{k+1}^\epsilon\big) \right|_{\epsilon=0} = 0. \tag{3.75}$$

The discrete conservation law following from the symmetry of the Lagrangian under spatial translation

$$\left(\frac{q_k - q_{k-1}}{h} \right) \cdot X = \left(\frac{q_{k+1} - q_k}{h} \right) \cdot X \tag{3.76}$$

amounts to the preservation of the discrete momentum in direction of X.

Energy

In continuous particle dynamics, the conservation of energy follows from translational symmetry of the Lagrangian with respect to time. In discrete particle dynamics, with a fixed timestep h, it is not possible to consider infinitesimal translations with respect to time. In the setting we described, it is therefore not possible to prove conservation of the discrete energy by applying Noether's theorem. Indeed, most often we find that energy is not conserved exactly, but only approximately, in that the energy error is bounded by some threshold value. This behaviour is typical for symplectic methods (see e.g. Hairer et al. [52] and references therein).

Nevertheless, it is possible to achieve and prove exact energy conservation by making the timestep h a dynamical variable. It is thereby determined by the variational principle, such that energy is conserved exactly [62]. And in the Noether theorem, infinitesimal transformations of time can be considered as well.

However, we do not follow this path. Still we are interested in the energy conserving properties of our variational integrators. We therefore "read" the expression for the discrete energy from the Lagrangian. In the particle case, the Hamiltonian is an explicit part of the Lagrangian, such that its discrete counterpart follows directly from the discretisation of the Lagrangian.

Discrete Noether Theorem for Composite Lagrangians

As already pointed out in section 3.1.5, special care has to be taken in the case of composite discrete Lagrangians (3.58)

$$\hat{L}_d(q_k^0, q_k^1, ..., q_k^s) = \sum_{i=1}^{s} L_d^i(q_k^{i-1}, q_k^i, \gamma_i h). \tag{3.77}$$

$\hat{L}_d$ might have different discrete expressions of the conserved momenta than the L_d^i. And it might even have different conservational properties, i.e., not all conserved momenta of $\hat{L}_d$ might be conserved by the L_d^i or vice versa. Therefore the discrete Noether theorem (3.72) has to be applied to $\hat{L}_d$.

Similarly, the discrete expression for the energy is different for $\hat{L}_d$ and the L_d^i. To clarify this, let us look at an example. Consider the midpoint Lagrangian (3.5)

$$L_d^{\mathrm{mp}}(q_k, q_{k+1}) = h\, L\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}\right), \tag{3.78}$$

and the fourth order, three step composition method (3.53),

$$\hat{F}_{L_d}^h = \tilde{F}_{L_d}^{\gamma h} \circ \tilde{F}_{L_d}^{(1-2\gamma)h} \circ \tilde{F}_{L_d}^{\gamma h} \qquad \text{with} \qquad \gamma = \left(2 - 2^{1/(r+1)}\right)^{-1}. \tag{3.79}$$

The discrete Hamiltonian corresponds to

$$H_d = h\, H\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}\right), \tag{3.80}$$

such that the discrete Hamiltonian of the composite Lagrangian is

$$\hat{H}_d = h\left[\gamma\, H\left(\frac{q_k^0 + q_k^1}{2}, \frac{q_k^1 - q_k^0}{h}\right) + (1 - 2\gamma)\, H\left(\frac{q_k^1 + q_k^2}{2}, \frac{q_k^2 - q_k^1}{h}\right) + \gamma\, H\left(\frac{q_k^2 + q_k^3}{2}, \frac{q_k^3 - q_k^2}{h}\right)\right], \tag{3.81}$$

where $q_k^0 = q_k$ and $q_k^3 = q_{k+1}$. Only the error of this composite Hamiltonian will scale with the order r of the scheme. That might at first seem surprising, as symplectic methods are supposed to conserve a direct discretisation of the continuous Hamiltonian to at least order r, i.e.,

$$H(q_n, p_n) = H(q_0, p_0) + \mathcal{O}(h^r) \tag{3.82}$$

for exponentially long time intervals $kh \leq e^{h_0/2h}$ with some constant h_0. However, in the proof of this relation (see Hairer et al. [52] and references therein) it is assumed that the discrete flow map $\hat{F}_{L_d}$ preserves the continuous symplectic form Ω. But we have only proved that the discrete symplectic form Ω_d is preserved. We can therefore not assume that this result translates directly. What we can always assume (for conservative systems) is that there exists a discrete energy that is preserved to the the order of the scheme, and in the case of the composition schemes this is an expression analogous to (3.81).

3.2. Discrete Field Theory

The derivation of the discrete field theory is a straight forward generalisation of the derivation for particle dynamics. The only difference is that the basic physical quantity is not the Lagrangian, defined on a one-dimensional "grid" of time, but the Lagrangian density, defined over a multidimensional grid of spacetime or phasespacetime.

Again, the starting point is the discretisation of the action integral and the Lagrangian density. The discrete Lagrangian density approximates the (phase)spacetime integral of the continuous Lagrangian density over one cell of the (phase)spacetime grid, e.g. with one spatial dimension this is

$$\mathcal{L}_d\big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i,k+1}, \varphi_{i+1,k+1}\big) \approx \int\limits_{t_k}^{t_{k+1}} \int\limits_{x_i}^{x_{i+1}} \mathcal{L}\big(\varphi, \varphi_t, \varphi_x\big) \tag{3.83}$$

with the corresponding action being a sum over the whole grid

$$\mathcal{A}_d = \sum_{k=0}^{N_t-1} \sum_{i=0}^{N_x-1} \mathcal{L}_d\big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i,k+1}, \varphi_{i+1,k+1}\big). \tag{3.84}$$

To make manipulations more tractable, the discrete Lagrangian density is rewritten in a slightly more abstract way, namely in terms of cells rather than grid points. Let us consider a cell determined by its vertices $(\varphi^1, \varphi^2, \varphi^3, \varphi^4)$, like it is depicted in Fig. 3.2. For now, the horizontal axis shall be space, denoted by x, and the vertical axis shall be time, denoted by t.

Here, only a midpoint rule is considered, but the application of other quadrature rules is straight forward. The fields φ are thus approximated by

$$\varphi(t, x) \approx \frac{1}{4}\big(\varphi^1 + \varphi^2 + \varphi^3 + \varphi^4\big). \tag{3.85}$$

For the approximation of the derivatives, there are in principle two possibilities for each coordinate, e.g. an x derivative can be defined as $(\varphi^2 - \varphi^1)/h_x$ as well as $(\varphi^3 - \varphi^4)/h_x$. Again, a midpoint-like averaging over the two possibilities is employed, such that

$$\varphi_t(t, x) \approx \frac{1}{2}\bigg(\frac{\varphi^4 - \varphi^1}{h_t} + \frac{\varphi^3 - \varphi^2}{h_t}\bigg), \qquad \varphi_x(t, x) \approx \frac{1}{2}\bigg(\frac{\varphi^2 - \varphi^1}{h_x} + \frac{\varphi^3 - \varphi^4}{h_x}\bigg). \tag{3.86}$$

Applying this to (3.83), the resulting discrete Lagrangian density reads

$$\mathcal{L}_d(\varphi^1, \varphi^2, \varphi^3, \varphi^4) \approx h_t\, h_x\, \mathcal{L}\left(\frac{\varphi^1 + \varphi^2 + \varphi^3 + \varphi^4}{4}, \frac{\varphi^4 - \varphi^1}{2h_t} + \frac{\varphi^3 - \varphi^2}{2h_t}, \frac{\varphi^2 - \varphi^1}{2h_x} + \frac{\varphi^3 - \varphi^4}{2h_x}\right)$$
(3.87)

and the discrete action becomes

$$\mathcal{A}_d[\varphi_d] = \sum_{\substack{\text{grid}\\\text{boxes}}} \mathcal{L}_d(\varphi^1, \varphi^2, \varphi^3, \varphi^4)$$
(3.88)

where $\varphi_d = \{\{\varphi_{i,k}\}_{i=0}^{N_x-1}\}_{k=0}^{N_t-1}$ is the discrete field. The application of Hamilton's principle

$$\left.\frac{d}{d\epsilon} \mathcal{A}_d[\varphi_d^\epsilon]\right|_{\epsilon=0} = \left.\frac{d}{d\epsilon} \sum_{\substack{\text{grid}\\\text{boxes}}} \mathcal{L}_d(\varphi^{1,\epsilon}, \varphi^{2,\epsilon}, \varphi^{3,\epsilon}, \varphi^{4,\epsilon})\right|_{\epsilon=0}$$
(3.89)

leads to discrete Euler-Lagrange field equations (DELFEQs) just as it lead to Euler-Lagrange equations in the continuous case. With

$$\delta\mathcal{A}_d \equiv \left.\frac{d}{d\epsilon} \mathcal{A}_d[\varphi_d^\epsilon]\right|_{\epsilon=0} \qquad\text{and}\qquad \delta\varphi_{i,k} \equiv \left.\frac{d}{d\epsilon} \varphi_{i,k}^\epsilon\right|_{\epsilon=0}$$
(3.90)

the variation of the action can be written as

$$\delta\mathcal{A}_d = \sum_{\substack{\text{grid}\\\text{boxes}}} \frac{\partial \mathcal{L}_d}{\partial \varphi^a}(\varphi^1, \varphi^2, \varphi^3, \varphi^4) \cdot \delta\varphi^a \qquad (1 \leq a \leq 4).$$
(3.91)

As the variation of the action has to vanish for each $\delta\varphi_{i,k}$ on the spacetime grid, it is sufficient to consider the contributions of $\delta\mathcal{A}_d$ that are multiplied by the variation of φ at a given grid point (i, k)

$$\begin{aligned}
\delta\mathcal{A}_d = \quad &\dots + \frac{\partial \mathcal{L}_d}{\partial \varphi^1}\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) \cdot \delta\varphi_{i,k} + \dots \\
&\dots + \frac{\partial \mathcal{L}_d}{\partial \varphi^2}\Big(\varphi_{i-1,k}, \varphi_{i,k}, \varphi_{i,k+1}, \varphi_{i-1,k+1}\Big) \cdot \delta\varphi_{i,k} + \dots \\
&\dots + \frac{\partial \mathcal{L}_d}{\partial \varphi^3}\Big(\varphi_{i-1,k-1}, \varphi_{i,k-1}, \varphi_{i,k}, \varphi_{i-1,k}\Big) \cdot \delta\varphi_{i,k} + \dots \\
&\dots + \frac{\partial \mathcal{L}_d}{\partial \varphi^4}\Big(\varphi_{i,k-1}, \varphi_{i+1,k-1}, \varphi_{i+1,k}, \varphi_{i,k}\Big) \cdot \delta\varphi_{i,k} + \dots \quad = 0.
\end{aligned}$$
(3.92)
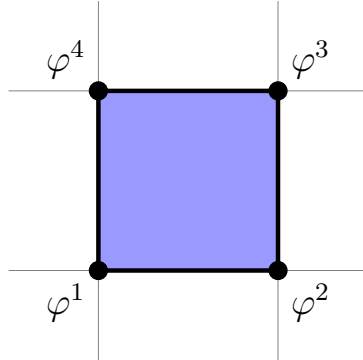
Fig. 3.2.: Basic element of a two-dimensional spacetime grid.
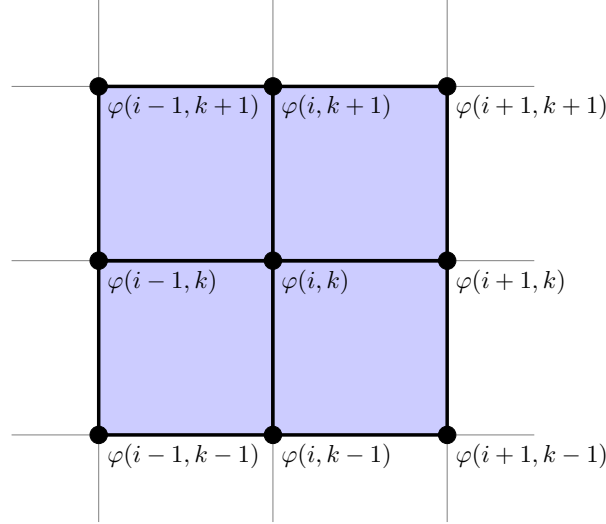
Fig. 3.3.: Contributions of a specific grid point (i, k) **to the variation of the discrete action.**

In total there are four such contributions, originating from the Lagrangian densities $\mathcal{L}_d$ of the four squares that touch the point (i, k) as is depicted in Fig. 3.3. The sum of the factors of $\delta\varphi_{ik}$ corresponds to the

Discrete Euler-Lagrange Field Equations

$$0 = \frac{\partial\mathcal{L}_d}{\partial\varphi^1}\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) + \frac{\partial\mathcal{L}_d}{\partial\varphi^2}\Big(\varphi_{i-1,k}, \varphi_{i,k}, \varphi_{i,k+1}, \varphi_{i-1,k+1}\Big)$$

$$+ \frac{\partial\mathcal{L}_d}{\partial\varphi^3}\Big(\varphi_{i-1,k-1}, \varphi_{i,k-1}, \varphi_{i,k}, \varphi_{i-1,k}\Big) + \frac{\partial\mathcal{L}_d}{\partial\varphi^4}\Big(\varphi_{i,k-1}, \varphi_{i+1,k-1}, \varphi_{i+1,k}, \varphi_{i,k}\Big). \tag{3.93}$$

Example: Wave Equation

The continuous Lagrangian density for the wave equations is

$$\mathcal{L}\Big(u_t(x,t), u_x(x,t)\Big) = \frac{1}{2}\left(\frac{\partial u(x,t)}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial u(x,t)}{\partial x}\right)^2. \tag{3.94}$$

A discretisation with the midpoint rule (3.86) leads to the discrete Lagrangian density

$$\mathcal{L}_d\Big(u^1, u^2, u^3, u^4\Big) = \frac{1}{2}\left(\frac{u^4 - u^1}{2h_t} + \frac{u^3 - u^2}{2h_t}\right)^2 - \frac{1}{2}\left(\frac{u^2 - u^1}{2h_x} + \frac{u^3 - u^4}{2h_x}\right)^2. \tag{3.95}$$

The four contributions to the discrete Euler-Lagrange field equations (3.93) are

$$\frac{\partial\mathcal{L}_d}{\partial u^1}\Big(u_{i,k}, u_{i+1,k}, u_{i+1,k+1}, u_{i,k+1}\Big) = -\frac{u_{i,k+1} - u_{i,k}}{4h_t^2} - \frac{u_{i+1,k+1} - u_{i+1,k}}{4h_t^2}$$

$$+ \frac{u_{i+1,k} - u_{i,k}}{4h_x^2} + \frac{u_{i+1,k+1} - u_{i,k+1}}{4h_x^2}, \tag{3.96a}$$

$$\frac{\partial \mathcal{L}_d}{\partial u^2}\Big(u_{i-1,k}, u_{i,k}, u_{i,k+1}, u_{i-1,k+1}\Big) = -\frac{u_{i-1,k+1} - u_{i-1,k}}{4h_t^2} - \frac{u_{i,k+1} - u_{i,k}}{4h_t^2}$$
$$- \frac{u_{i,k} - u_{i-1,k}}{4h_x^2} - \frac{u_{i,k+1} - u_{i-1,k+1}}{4h_x^2}, \tag{3.96b}$$

$$\frac{\partial \mathcal{L}_d}{\partial u^3}\Big(u_{i-1,k-1}, u_{i,k-1}, u_{i,k}, u_{i-1,k}\Big) = +\frac{u_{i-1,k} - u_{i-1,k-1}}{4h_t^2} + \frac{u_{i,k} - u_{i,k-1}}{4h_t^2}$$
$$- \frac{u_{i,k-1} - u_{i-1,k-1}}{4h_x^2} - \frac{u_{i,k} - u_{i-1,k}}{4h_x^2}, \tag{3.96c}$$

$$\frac{\partial \mathcal{L}_d}{\partial u^4}\Big(u_{i,k-1}, u_{i+1,k-1}, u_{i+1,k}, u_{i,k}\Big) = +\frac{u_{i,k} - u_{i,k-1}}{4h_t^2} + \frac{u_{i+1,k} - u_{i+1,k-1}}{4h_t^2}$$
$$+ \frac{u_{i+1,k-1} - u_{i,k-1}}{4h_x^2} + \frac{u_{i+1,k} - u_{i,k}}{4h_x^2}. \tag{3.96d}$$

Summing up all these terms, the discrete wave equation is obtained

$$\frac{u_{i-1,k+1} - 2\, u_{i-1,k} + u_{i-1,k-1}}{4h_t^2} + 2\,\frac{u_{i,k+1} - 2\, u_{i,k} + u_{i,k-1}}{4h_t^2} + \frac{u_{i+1,k+1} - 2\, u_{i+1,k} + u_{i+1,k-1}}{4h_t^2} =$$
$$= \frac{u_{i+1,k+1} - 2\, u_{i,k+1} + u_{i-1,k+1}}{4h_x^2} + 2\,\frac{u_{i+1,k} - 2\, u_{i,k} + u_{i-1,k}}{4h_x^2} + \frac{u_{i+1,k-1} - 2\, u_{i,k-1} + u_{i-1,k-1}}{4h_x^2}. \tag{3.97}$$

This clearly is a discrete version of the continuous wave equation

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial^2 u(x,t)}{\partial x^2},$$

with the following stencil

$$\frac{1}{4h_t^2}\begin{bmatrix} 1 & 2 & 1 \\ -2 & -4 & -2 \\ 1 & 2 & 1 \end{bmatrix} u = \frac{1}{4h_x^2}\begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 2 \\ 1 & -2 & 1 \end{bmatrix} u. \tag{3.98}$$

We observe that the derivative with respect to one direction is averaged in the other direction, i.e., the time derivative is averaged over three neighbouring points in space, and the spatial derivative is averaged over three neighbouring points in time. This averaging of derivatives is a common feature often found in variational integrators of field theories. It appears to be one of the decisive features that account for the superior performance of variational integrators.

3.2.1. Discrete Jet Space

The discrete phasespace, on which the Lagrangian density is defined, is the discrete first jet bundle $\mathsf{J}^1\mathcal{Y}$

$$\mathcal{L}_d : \mathsf{J}^1\mathcal{Y} \to \mathbb{R}. \tag{3.99}$$

To understand its structure, some considerations are in order. The discrete spacetime of dimension two is the grid of points

$$\mathcal{X} = \mathbb{Z} \times \mathbb{Z} = \left\{(i,k)\right\} \cong \left\{(ih_x, kh_t) \,\middle|\, i,k \in \mathbb{Z}\right\}. \tag{3.100}$$

It corresponds to a grid with elements $x_{i,j}$ in continuous spacetime. The discrete fibre bundle over $\mathcal{X}$ is

$$\mathcal{Y} = \mathcal{X} \times \mathcal{F}. \tag{3.101}$$

where $\mathcal{F}$ is a smooth manifold. Elements of $\mathcal{Y}$ over the point (i,k) are denoted by $y_{i,k}$ and the projection $\pi_{\mathcal{X}\mathcal{Y}}$ is given by

$$\pi_{\mathcal{X}\mathcal{Y}}(y_{i,k}) = (i,k). \tag{3.102}$$

A square $\square$ on $\mathcal{X}$ is an ordered quadruplet

$$\square = \left((i,k),(i+1,k),(i+1,k+1),(i,k+1)\right), \tag{3.103}$$

defining a grid cell, c.f. figure 3.2. The first component of $\square$, denoted $\square^1$, is the first vertex of the square, with equivalent definitions for the other three vertices

$$\square^1 = (i,k), \qquad \square^2 = (i+1,k), \qquad \square^3 = (i+1,k+1), \qquad \square^4 = (i,k+1). \tag{3.104}$$

A section φ of $\mathcal{Y}$ is a map

$$\varphi : \mathcal{U} \subseteq \mathcal{X} \to \mathcal{Y} \tag{3.105}$$

such that

$$\pi_{\mathcal{X}\mathcal{Y}} \circ \varphi = \mathrm{id}_{\mathcal{U}} \tag{3.106}$$

and the φ^a with $a \in \{1,2,3,4\}$ of figure 3.2 correspond to

$$\varphi^1 = \varphi(\square^1) = \varphi_{\square^1}, \quad \varphi^2 = \varphi(\square^2) = \varphi_{\square^2}, \quad \varphi^3 = \varphi(\square^3) = \varphi_{\square^3}, \quad \varphi^4 = \varphi(\square^4) = \varphi_{\square^4}. \tag{3.107}$$

The set of squares on $\mathcal{X}$ is denoted $\mathcal{X}^\square$. The first jet bundle of $\mathcal{Y}$ is given by

$$\begin{aligned} \mathsf{J}^1\mathcal{Y} &\equiv \left\{\left(\square,(y_{i,k}, y_{i+1,k}, y_{i+1,k+1}, y_{i,k+1})\right) \middle| \square \in \mathcal{X}^\square,\ (i,k) = \square^1,\ y_{i,k}, y_{i+1,k}, y_{i+1,k+1}, y_{i,k+1} \in \mathcal{F}\right\} \\ &\equiv \mathcal{X}^\square \times \mathcal{F}^4. \end{aligned} \tag{3.108}$$

The first jet prolongation of a section φ on $\mathcal{Y}$ is the map

$$j^1\varphi : \mathcal{X}^\square \mapsto \mathsf{J}^1\mathcal{Y} \tag{3.109}$$

defined by

$$j^1\varphi(\square) \equiv \left(\square, \varphi(\square^1), \varphi(\square^2), \varphi(\square^3), \varphi(\square^4)\right). \tag{3.110}$$

The first jet is defined to include first order derivatives, that at the discrete level are functions of $\square$ and $\varphi(\square^a)$. The restriction of a vector field V on $\mathcal{Y}$ to the fibre $\mathcal{Y}_{i,k}$ is denoted $V_{i,k}$, and similarly for vector fields on $\mathsf{J}^1\mathcal{Y}$. The first jet prolongation of a vector field V on $\mathcal{Y}$ is the vector field j^1V on $\mathsf{J}^1\mathcal{Y}$, defined by

$$j^1V(\square) \equiv \big(\square, V_{\square^1}(y_{\square^1}), V_{\square^2}(y_{\square^2}), V_{\square^3}(y_{\square^3}), V_{\square^4}(y_{\square^4})\big), \tag{3.111}$$

for any square $\square$.

A point $(i,k) \in \mathcal{X}$ is *touched* by a square, if it is a vertex of that square. A point $(i,k) \in \mathcal{U} \subseteq \mathcal{X}$ is an *interior point* of $\mathcal{U}$, if $\mathcal{U}$ contains all four squares of $\mathcal{X}$ that touch (i,k). The *interior* $\mathrm{int}\,\mathcal{U}$ of $\mathcal{U}$ is the collection of all interior points of $\mathcal{U}$. The closure $\mathrm{cl}\,\mathcal{U}$ of $\mathcal{U}$ is the union of all squares touching interior points of $\mathcal{U}$. A point (i,k) is a *boundary point* of $\mathcal{U}$ if it is a point in both, $\mathcal{U}$ and $\mathrm{cl}\,\mathcal{U}$, which is not an interior point. The boundary $\partial\mathcal{U}$ of $\mathcal{U}$ is the set of boundary points of $\mathcal{U}$, such that

$$\partial\mathcal{U} \equiv (\mathcal{U} \cap \mathrm{cl}\,\mathcal{U}) \setminus \mathrm{int}\,\mathcal{U}. \tag{3.112}$$

Discrete Action Principle

The Lagrangian density on a given square is a function

$$\mathcal{L}_{\square} : \mathcal{F}^4 \to \mathbb{R} \tag{3.113}$$

defined as

$$\mathcal{L}_{\square}\big(y^1, y^2, y^3, y^4\big) \equiv \mathcal{L}_d\big(\square, y^1, y^2, y^3, y^4\big). \tag{3.114}$$

Thus, the discrete Lagrangian density $\mathcal{L}_d$ can be regarded as the choice of a function $\mathcal{L}_{\square}$ on each square $\square$ of $\mathcal{X}$. The variables on the domain of $\mathcal{L}_{\square}$ are denoted y^1, y^2, y^3, y^4, independently of the actual $\square$.

If $\mathcal{C}(\mathcal{Y})$ is the set of sections of $\mathcal{Y}$ on a subset $\mathcal{U} \subseteq \mathcal{X}$, the discrete action $\mathcal{A}_d$ is a real-valued function on $\mathcal{C}(\mathcal{Y})$, defined by

$$\mathcal{A}_d[\varphi] = \sum_{\substack{\square \\ \square \subseteq \mathcal{U}}} \mathcal{L}_d \circ j^1\varphi(\square). \tag{3.115}$$

The variations φ^ϵ of a section $\varphi \in \mathcal{C}(\mathcal{Y})$ are described by a vertical map η^ϵ and its generating vector field $V \in \mathsf{T}_\varphi \mathcal{C}(\mathcal{Y})$. The map φ^ϵ corresponds to a one-parameter family of sections

$$\varphi^\epsilon(i,k) \equiv (\eta^\epsilon \circ \varphi)(i,k) \equiv \eta^\epsilon_{ik}\big(\varphi(i,k)\big), \tag{3.116}$$

where η^ϵ_{ik} is the flow of V_{ik} on $\mathcal{Y}_{ik}$. The action principle is to seek those sections φ for which

$$\mathbf{d}\mathcal{A}_d = \frac{d}{d\epsilon}\mathcal{A}_d\big[\varphi^\epsilon\big]\bigg|_{\epsilon=0} = 0 \tag{3.117}$$

for all vector fields V on $\mathcal{C}(\mathcal{Y})$. By focusing upon a fixed $(i,k) \in \mathrm{int}(\mathcal{U})$ and the same arguments as in the previous section, c.f. equations (3.91) and (3.92), we find the discrete Euler-Lagrange field equations for all $(i,k) \in \mathrm{int}(\mathcal{U})$

$$\sum_{\substack{a \\ (i,k)=\square^a}} \frac{\partial \mathcal{L}_{\square}}{\partial y^a}\big(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4}\big) = 0 \quad \textbf{(Discrete Euler-Lagrange Equations)}. \tag{3.118}$$

3.2.2. Discrete Cartan Form

Allowing for nonzero variations on the boundary $\partial \mathcal{U}$ will lead us to the discrete Cartan form. In that case, the vector field V does not necessarily vanish on $\partial \mathcal{U}$.

For each point (i, k) of the boundary $\partial \mathcal{U}$, find the squares in $\mathcal{U}$ that touch (i, k). There is at least one such square since $(i, k) \in \mathrm{cl}(\mathcal{U})$, but not four such squares since $(i, k) \notin \mathrm{int}(\mathcal{U})$. For each of the touching squares, (i, k) occurs at the a'th vertex for at most three values of $a = \{1, 2, 3, 4\}$, such that the expressions

$$\frac{\partial \mathcal{L}_d}{\partial y^1}\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) V_{ik}(\varphi_{ik}), \tag{3.119a}$$

$$\frac{\partial \mathcal{L}_d}{\partial y^2}\Big(\varphi_{i-1,k}, \varphi_{i,k}, \varphi_{i,k+1}, \varphi_{i-1,k+1}\Big) V_{ik}(\varphi_{ik}), \tag{3.119b}$$

$$\frac{\partial \mathcal{L}_d}{\partial y^3}\Big(\varphi_{i-1,k-1}, \varphi_{i,k-1}, \varphi_{i,k}, \varphi_{i-1,k}\Big) V_{ik}(\varphi_{ik}), \tag{3.119c}$$

$$\frac{\partial \mathcal{L}_d}{\partial y^4}\Big(\varphi_{i,k-1}, \varphi_{i+1,k-1}, \varphi_{i+1,k}, \varphi_{i,k}\Big) V_{ik}(\varphi_{ik}), \tag{3.119d}$$

give at most three contributions. The total contribution to $\mathbf{d}\mathcal{A}_d$ from the boundary is the sum of all such terms. In discrete particle mechanics, we found two one-forms. The above list suggests that there are four Cartan forms, which we define to be

$$\Theta^1_{L_d}\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) \cdot \Big(v_{y_{i,k}}, v_{y_{i+1,k}}, v_{y_{i+1,k+1}}, v_{y_{i,k+1}}\Big) =$$
$$= \frac{\partial \mathcal{L}_d}{\partial y^1}\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) \cdot \Big(v_{y_{ij}}, 0, 0, 0\Big), \quad \text{(3.120a)}$$

$$\Theta^2_{L_d}\Big(\varphi_{i-1,k}, \varphi_{i,k}, \varphi_{i,k+1}, \varphi_{i-1,k+1}\Big) \cdot \Big(v_{y_{i-1,k}}, v_{y_{i,k}}, v_{y_{i,k+1}}, v_{y_{i-1,k+1}}\Big) =$$
$$= \frac{\partial \mathcal{L}_d}{\partial y^2}\Big(\varphi_{i-1,k}, \varphi_{i,k}, \varphi_{i,k+1}, \varphi_{i-1,k+1}\Big) \cdot \Big(0, v_{y_{ij}}, 0, 0\Big), \quad \text{(3.120b)}$$

$$\Theta^3_{L_d}\Big(\varphi_{i-1,k-1}, \varphi_{i,k-1}, \varphi_{i,k}, \varphi_{i-1,k}\Big) \cdot \Big(v_{y_{i-1,k-1}}, v_{y_{i,k-1}}, v_{y_{i,k}}, v_{y_{i-1,k}}\Big) =$$
$$= \frac{\partial \mathcal{L}_d}{\partial y^3}\Big(\varphi_{i-1,k-1}, \varphi_{i,k-1}, \varphi_{i,k}, \varphi_{i-1,k}\Big) \cdot \Big(0, 0, v_{y_{ij}}, 0\Big), \quad \text{(3.120c)}$$

$$\Theta^4_{L_d}\Big(\varphi_{i,k-1}, \varphi_{i+1,k-1}, \varphi_{i+1,k}, \varphi_{i,k}\Big) \cdot \Big(v_{y_{i,k-1}}, v_{y_{i+1,k-1}}, v_{y_{i+1,k}}, v_{y_{i,k}}\Big) =$$
$$= \frac{\partial \mathcal{L}_d}{\partial y^4}\Big(\varphi_{i,k-1}, \varphi_{i+1,k-1}, \varphi_{i+1,k}, \varphi_{i,k}\Big) \cdot \Big(0, 0, 0, v_{y_{ij}}\Big). \quad \text{(3.120d)}$$

This quadruple $\big(\Theta^1_{L_d}, \Theta^2_{L_d}, \Theta^3_{L_d}, \Theta^4_{L_d}\big)$ is regarded as the discrete counterpart of the Cartan form Θ_L from (2.228). For a vector field V from $\mathsf{T}\mathcal{C}(\mathcal{Y})$, the expressions from the list become

$$\big[(j^1\varphi)^*(\boldsymbol{\imath}_{j^1V}\Theta^a_{L_d})\big](\square) =$$
$$= \Theta^a_{L_d}\big(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4}\big) \cdot \big(V_{\square^1}(y_{\square^1}), V_{\square^2}(y_{\square^2}), V_{\square^3}(y_{\square^3}), V_{\square^4}(y_{\square^4})\big). \quad \text{(3.121)}$$

We collect all the contributions from the boundary into one single object $\theta_{L_d}(\varphi) \cdot V$, where θ_{L_d} is the one-form on the space of sections $\mathcal{C}(\mathcal{Y})$, defined by

$$\theta_{L_d}(\varphi) \cdot V \equiv \sum_{\substack{\square \\ \square \cap \partial \mathcal{U} \neq \varnothing}} \bigg(\sum_{\substack{a \\ \square^a \in \partial \mathcal{U}}} \big[(j^1\varphi)^*(\boldsymbol{\imath}_{j^1V}\Theta^a_{L_d})\big](\square)\bigg). \tag{3.122}$$

Here we sum over all squares $\square$ that touch the boundary, and all vertices $\square^a$ of those squares that are elements of the boundary. With this, the variation of the discrete action (3.117) can be written as

$$\mathbf{d}\mathcal{A}_d \cdot V = \sum_{\substack{\square \\ \square \cap \operatorname{int}\mathcal{U} \neq \varnothing}} \left(\sum_{\substack{a \\ \square^a \in \operatorname{int}\mathcal{U}}} \left[(j^1\varphi)^*(\boldsymbol{\imath}_{j^1V}\Theta_{L_d}^a) \right](\square) \right)$$

$$+ \sum_{\substack{\square \\ \square \cap \partial\mathcal{U} \neq \varnothing}} \left(\sum_{\substack{a \\ \square^a \in \partial\mathcal{U}}} \left[(j^1\varphi)^*(\boldsymbol{\imath}_{j^1V}\Theta_{L_d}^a) \right](\square) \right) \qquad (3.123)$$

$$= \sum_{\substack{\square \\ \square \cap \operatorname{int}\mathcal{U} \neq \varnothing}} \mathbf{d}\mathcal{L}_\square \cdot V + \theta_{L_d}(\varphi) \cdot V = 0, \qquad (3.124)$$

which is similar to the continuous result (2.230).

3.2.3. Discrete Multisymplectic Form

The four Cartan forms (3.120) correspond to the exterior derivative of the discrete Lagrangian

$$\mathbf{d}\mathcal{L}_\square = \sum_{\substack{a \\ \square \subseteq \mathcal{U}}} \frac{\partial \mathcal{L}_\square}{\partial y^a}\left(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4}\right) dy_{\square^a} = \sum_{\substack{a \\ \square \subseteq \mathcal{U}}} \Theta_{L_d}^a\left(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4}\right), \qquad (3.125)$$

such that upon defining the discrete multisymplectic form as the exterior derivative of the discrete Cartan form,

$$\Omega_{L_d}^a = -\mathbf{d}\Theta_{L_d}^a, \qquad (3.126)$$

due to $\mathbf{d}^2\mathcal{L} = 0$ we get

$$\sum_{\substack{a \\ \square \subseteq \mathcal{U}}} \Omega_{L_d}^a\left(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4}\right) = \Omega_{L_d}^1(\square) + \Omega_{L_d}^2(\square) + \Omega_{L_d}^3(\square) + \Omega_{L_d}^4(\square) = 0. \qquad (3.127)$$

For a square $\square$ in $\mathcal{X}$, define the projection

$$\pi_\square : \mathcal{C}(\mathcal{Y}) \to \mathsf{J}^1\mathcal{Y} \qquad (3.128)$$

in analogy to the jet prolongation (3.109,3.110) by

$$\pi_\square(\varphi) \equiv \left(\square, \varphi(\square^1), \varphi(\square^2), \varphi(\square^3), \varphi(\square^4)\right), \qquad (3.129)$$

such that the forms $\pi_\square^*\Theta_L^a$ are computed as

$$(\pi_\square^*\Theta_{L_d}^a)(\varphi) \cdot V = \frac{\partial \mathcal{L}_d}{\partial y^a}\left(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4}\right) \cdot V(\square^a), \qquad (3.130)$$

and the one-form (3.122) becomes

$$\theta_{L_d} = \sum_{\substack{\square \\ \square \cap \partial\mathcal{U} \neq \varnothing}} \left(\sum_{\substack{a \\ \square^a \in \partial\mathcal{U}}} \pi_\square^*\Theta_{L_d}^a \right). \qquad (3.131)$$

Consider the subspace $\mathcal{C}_L \subset \mathcal{C}(\mathcal{Y})$ of the space of sections φ, that solve the discrete Euler-Lagrange equations (3.118). A first variation at a solution φ of the discrete Euler-Lagrange

field equations (3.118) corresponds to a vector field $V \in \mathsf{T}_\varphi \mathcal{C}_L$ such that the associated flow maps φ to other solutions of the discrete Euler-Lagrange field equations, i.e., sections $\varphi \in \mathcal{C}_L$ are integral curves of V. Restricting the action (3.123) to the subspace $\mathcal{C}_L$, the first sum in (3.123) becomes zero and only the one-form (3.122) is retained. Computing the exterior derivative of the variation of the action and restricting to two first variations $V, W \in \mathsf{T}_\varphi \mathcal{C}_L$, we obtain

$$0 = \mathbf{d}^2 \mathcal{A}_d \cdot V \cdot W = \mathbf{d}\theta_{L_d}(\phi)(V, W) = \sum_{\substack{\square \\ \square \cap \partial \mathcal{U} \neq \varnothing}} \left(\sum_{\square^a \in \partial \mathcal{U}}^a V \lrcorner W \lrcorner \pi_\square^* \Omega_{L_d}^a \right), \tag{3.132}$$

which is equivalent to

$$\sum_{\substack{\square \\ \square \cap \partial \mathcal{U} \neq \varnothing}} \left(\sum_{\square^a \in \partial \mathcal{U}}^a \left[(j^1\varphi)^* (V \lrcorner W \lrcorner \Omega_{L_d}^a) \right](\square) \right) = 0. \tag{3.133}$$

This is the discrete analogue to the multisymplectic form formula (2.262).

3.2.4. Discrete Noether Theorem

We restrict our treatment to the case of a scalar field theory, one spatial dimension, and vertical transformations, but at least the first two restrictions are easily lifted [84]. Consider a one-parameter group of

$$\varphi_{i,k}^\epsilon = \eta^\epsilon \circ \varphi_{i,k} \qquad \text{such that} \qquad \varphi_{i,k}^0 = \varphi_{i,k}. \tag{3.134}$$

The infinitesimal generator of the transformation $\varphi_{i,k}^\epsilon$ is

$$X_{i,k} = \frac{d}{d\epsilon} \varphi_{i,k}^\epsilon \Big|_{\epsilon=0} \qquad \text{or in abstract notation} \qquad X_{\square^a}(\varphi_{\square^a}) = \frac{d}{d\epsilon} \varphi_{\square^a}^\epsilon \Big|_{\epsilon=0}. \tag{3.135}$$

The discrete Lagrangian has a symmetry if it is invariant under this transformation

$$\mathcal{L}_d\left(\varphi_{\square^1}^\epsilon, \varphi_{\square^2}^\epsilon, \varphi_{\square^3}^\epsilon, \varphi_{\square^4}^\epsilon \right) = \mathcal{L}_d\left(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4} \right) \qquad \text{for all } \epsilon. \tag{3.136}$$
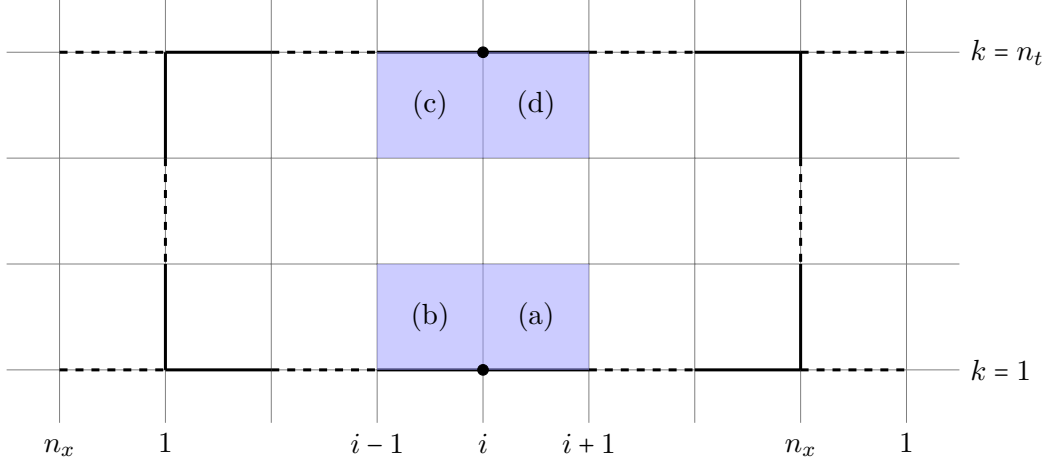
This is equivalent to

$$\mathbf{d}\mathcal{L}_\square \cdot X = \frac{d}{d\epsilon} \Big|_{\epsilon=0} \mathcal{L}_d\left(\varphi_{\square^1}^\epsilon, \varphi_{\square^2}^\epsilon, \varphi_{\square^3}^\epsilon, \varphi_{\square^4}^\epsilon \right) \tag{3.137}$$

$$= \sum_{\substack{a \\ \square \subseteq \mathcal{U}}} \left[\frac{\partial \mathcal{L}_d}{\partial y^a} \left(\varphi_{\square^1}, \varphi_{\square^2}, \varphi_{\square^3}, \varphi_{\square^4} \right) \cdot X_{\square^a}(\varphi_{\square^a}) \right] = 0, \tag{3.138}$$

or explicitly in grid coordinates,

$$\mathbf{d}\mathcal{L}_\square \cdot X = \frac{\partial \mathcal{L}_d}{\partial y^1} \left(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1} \right) \cdot X_{i,k}$$

$$+ \frac{\partial \mathcal{L}_d}{\partial y^2} \left(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1} \right) \cdot X_{i+1,k}$$

$$+ \frac{\partial \mathcal{L}_d}{\partial y^3} \left(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1} \right) \cdot X_{i+1,k+1}$$

$$+ \frac{\partial \mathcal{L}_d}{\partial y^4} \left(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1} \right) \cdot X_{i,k+1} = 0. \tag{3.139}$$

Fig. 3.4.: Boundary terms in a spatially periodic integration domain.

Since the φ are solutions of the discrete Euler-Lagrange field equations (3.118), the generating vector field X is a first variation, i.e., $X \in \mathsf{T}_\varphi \mathcal{C}_L$. This means that the sum in (3.124) vanishes, and we obtain

$$\theta_{L_d}(\varphi) \cdot X \equiv \sum_{\substack{\square \\ \square \cap \partial\mathcal{U} \neq \varnothing}} \left(\sum_{\substack{a \\ \square^a \in \partial\mathcal{U}}} \left[(j^1\varphi)^* (\boldsymbol{\imath}_{j^1 X} \Theta_{L_d}^a) \right](\square) \right) = 0, \tag{3.140}$$

or explicitly in grid coordinates,

$$0 = \sum_{\substack{i \\ (i,k) \in \partial\mathcal{U}}} \left[\frac{\partial \mathcal{L}_d}{\partial y^1}\Big(\varphi_{i,1}, \varphi_{i+1,1}, \varphi_{i+1,2}, \varphi_{i,2}\Big) \cdot X_{i,1} \right.$$

$$+ \frac{\partial \mathcal{L}_d}{\partial y^2}\Big(\varphi_{i-1,1}, \varphi_{i,1}, \varphi_{i,2}, \varphi_{i-1,2}\Big) \cdot X_{i,1}$$

$$+ \frac{\partial \mathcal{L}_d}{\partial y^3}\Big(\varphi_{i-1,n_t-1}, \varphi_{i,n_t-1}, \varphi_{i,n_t}, \varphi_{i-1,n_t}\Big) \cdot X_{i,n_t}$$

$$\left. + \frac{\partial \mathcal{L}_d}{\partial y^4}\Big(\varphi_{i,n_t-1}, \varphi_{i+1,n_t-1}, \varphi_{i+1,n_t}, \varphi_{i,n_t}\Big) \cdot X_{i,n_t} \right]. \tag{3.141}$$

The various contributions arise as depicted in figure 3.4. If we fix the spatial index i, there are four cells that touch the boundary, two at $(i,1)$ and two at (i,n_t), respectively. The first contribution arises from the cell marked (a). The point $(i,1)$ corresponds to $\square^1$, such that the derivative of the Lagrangian is computed with respect to $y^1 = y(\square^1)$. The other contributions follow in the same way. In principle, there are also contributions to (3.140) that arise from the spatial boundary at $i = 1$ and $i = n_x$. One either has to account for those or select boundary conditions, that automatically take care of these contributions. We shall always use periodic boundary conditions (as depicted), such that practically there is no boundary in the spatial dimension.

With the help of the discrete symmetry condition (3.139), we can replace the first two

lines in (3.141), such that

$$
\sum_{\substack{i \\ (i,k)\in\partial\mathcal{U}}} \left[\frac{\partial\mathcal{L}_d}{\partial y^3}\Big(\varphi_{i-1,n_t-1}, \varphi_{i,n_t-1}, \varphi_{i,n_t}, \varphi_{i-1,n_t}\Big)\cdot X_{i,n_t} + \frac{\partial\mathcal{L}_d}{\partial y^4}\Big(\varphi_{i,n_t-1}, \varphi_{i+1,n_t-1}, \varphi_{i+1,n_t}, \varphi_{i,n_t}\Big)\cdot X_{i,n_t} \right] =
$$

$$
= \sum_{\substack{i \\ (i,k)\in\partial\mathcal{U}}} \left[\frac{\partial\mathcal{L}_d}{\partial y^3}\Big(\varphi_{i-1,1}, \varphi_{i,1}, \varphi_{i,2}, \varphi_{i-1,2}\Big) \cdot X_{i,2} + \frac{\partial\mathcal{L}_d}{\partial y^4}\Big(\varphi_{i,1}, \varphi_{i+1,1}, \varphi_{i+1,2}, \varphi_{i,2}\Big) \cdot X_{i,2} \right].
$$

$$(3.142)$$

This is a conservation law, and as the number of timesteps n_t is arbitrary, can be written as

$$
\sum_{i=1}^{n_x-1} \left[\frac{\partial\mathcal{L}_d}{\partial y^3}\Big(\varphi_{i-1,k-1}, \varphi_{i,k-1}, \varphi_{i,k}, \varphi_{i-1,k}\Big) + \frac{\partial\mathcal{L}_d}{\partial y^4}\Big(\varphi_{i,k-1}, \varphi_{i+1,k-1}, \varphi_{i+1,k}, \varphi_{i,k}\Big) \right] \cdot X_{i,k} = const.
$$

$$(3.143)$$

The structure of this conservation law implies that if the continuous Lagrangian has a vertical symmetry, any consistent discretisation of the Lagrangian will lead to a discrete conservation law corresponding to that symmetry.

3.2.5. Discrete Momentum Maps

Geometrically, a conserved quantity is described by a momentum map

$$
J^a \equiv \boldsymbol{\iota}_{j^1 X}\Theta^a_{L_d} \qquad \text{with} \qquad X_{\square^a}(\varphi_{\square^a}) = \frac{d}{d\epsilon}\varphi^\epsilon_{\square^a}\bigg|_{\epsilon=0}, \tag{3.144}
$$

such that (3.140) takes the form

$$
\sum_{\substack{\square \\ \square\cap\partial\mathcal{U}\neq\varnothing}} \left(\sum_{\substack{a \\ \square^a\in\partial\mathcal{U}}} (j^1\varphi)^\star J^a(\square) \right) = 0, \tag{3.145}
$$

or explicitly in grid coordinates,

$$
\sum_{i=1}^{n_x-1} \bigg[J^1\Big(\varphi_{i,1}, \varphi_{i+1,1}, \varphi_{i+1,2}, \varphi_{i,2}\Big) + J^2\Big(\varphi_{i,1}, \varphi_{i+1,1}, \varphi_{i+1,2}, \varphi_{i,2}\Big)
$$

$$
+ J^3\Big(\varphi_{i,n_t-1}, \varphi_{i+1,n_t-1}, \varphi_{i+1,n_t}, \varphi_{i,n_t}\Big) + J^4\Big(\varphi_{i,n_t-1}, \varphi_{i+1,n_t-1}, \varphi_{i+1,n_t}, \varphi_{i,n_t}\Big) \bigg]. \tag{3.146}
$$

The invariance of the Lagrangian (3.136) and (3.125) implies

$$
J^1\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) + J^2\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big)
$$

$$
+ J^3\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) + J^4\Big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\Big) = 0, \tag{3.147}
$$

so that we can replace the last line of (3.146) to get

$$
\sum_{i=1}^{n_x-1} \bigg[J^1\Big(\varphi_{i,n_t-1}, \varphi_{i+1,n_t-1}, \varphi_{i+1,n_t}, \varphi_{i,n_t}\Big) + J^2\Big(\varphi_{i,n_t-1}, \varphi_{i+1,n_t-1}, \varphi_{i+1,n_t}, \varphi_{i,n_t}\Big) \bigg] =
$$

$$
= \sum_{i=1}^{n_x-1} \bigg[J^1\Big(\varphi_{i,1}, \varphi_{i+1,1}, \varphi_{i+1,2}, \varphi_{i,2}\Big) + J^2\Big(\varphi_{i,1}, \varphi_{i+1,1}, \varphi_{i+1,2}, \varphi_{i,2}\Big) \bigg]. \tag{3.148}
$$

As this is true for any n_t, we can infer the general expression

$$\sum_{i=1}^{n_x-1} \left[J^1\big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\big) + J^2\big(\varphi_{i,k}, \varphi_{i+1,k}, \varphi_{i+1,k+1}, \varphi_{i,k+1}\big) \right] = const. \quad (3.149)$$

This is equivalent to the conservation law (3.143).

3.3. Example: The Advection Equation

In this section, we apply the theory of the previous chapter to an interesting and instructive example, namely the advection equation. Its structure is similar to that of the Vlasov equation which we will study in chapter 5. The linear advection equation for a field $u(t,x)$ (in the following referred to as *density*) reads

$$\partial_t u + c\,\partial_x u = 0, \quad (3.150)$$

where the velocity, c, is a constant parameter. For initial conditions $u(t=0,x) = u_0(x)$, the analytic solution is

$$u(t,x) = u_0(x - ct). \quad (3.151)$$

The advection equation does not have a natural variational formulation, but we can apply the theory from section 2.2.7 to obtain an extended Lagrangian.

3.3.1. Extended Lagrangian

The extended Lagrangian for the advection equation is obtained by multiplying (3.150) with the auxiliary variable $v(t,x)$. The solution vector of the extended system is denoted $w = (u,v)$, such that the Lagrangian can be written as

$$\mathcal{L}(j^1 w) = v\big(u_t + cu_x\big)\,\omega, \quad (3.152)$$

where $\omega = dt \wedge dx$ is the volume form, so that the corresponding action is $\mathcal{A} = \int \mathcal{L}$. The variational derivative of the action with respect to the Ibragimov multiplier v retains the advection equation

$$\frac{\delta \mathcal{A}}{\delta v} = +u_t + cu_x = 0. \quad (3.153)$$

The variation with respect to the original variable u yields the adjoint equation

$$\frac{\delta \mathcal{A}}{\delta u} = -v_t - cv_x = 0. \quad (3.154)$$

It is immediately observed that the adjoint equation has the same solution as the original equation, such that if u is a solution of the advection equation, then $w = (u,u)$ solves the Euler-Lagrange equations of the extended Lagrangian (3.152). Here and in the remaining part of the section, w denotes the combined fields (u,v).

3.3.2. Variational Integrator

We discretise the Lagrangian (3.152) on a grid cell as depicted in figure 3.2 by

$$
\mathcal{L}_d(w^1, w^2, w^3, w^4) = h_t\, h_x\, \frac{1}{4}\Big(v^1 + v^2 + v^3 + v^4\Big)\times
$$

$$
\times \left[\frac{1}{2}\left(\frac{u^4 - u^1}{h_t} + \frac{u^3 - u^2}{h_t} \right) + \frac{c}{2}\left(\frac{u^2 - u^1}{h_x} + \frac{u^3 - u^4}{h_x} \right) \right] \quad (3.155)
$$

The discrete Euler-Lagrange field equations (3.93)

$$
0 = \frac{\partial \mathcal{L}_d}{\partial g^1}\Big(w_{i,k}, w_{i+1,k}, w_{i+1,k+1}, w_{i,k+1}\Big) + \frac{\partial \mathcal{L}_d}{\partial g^2}\Big(w_{i-1,k}, w_{i,k}, w_{i,k+1}, w_{i-1,k+1}\Big)
$$

$$
+ \frac{\partial \mathcal{L}_d}{\partial g^3}\Big(w_{i-1,k-1}, w_{i,k-1}, w_{i,k}, w_{i-1,k}\Big) + \frac{\partial \mathcal{L}_d}{\partial g^4}\Big(w_{i,k-1}, w_{i+1,k-1}, w_{i+1,k}, w_{i,k}\Big) \quad (3.156)
$$

are computed as

$$
0 = \frac{1}{4}\left[\frac{u_{i+1,k+1} - u_{i+1,k-1}}{2h_t} + 2\,\frac{u_{i,k+1} - u_{i,k-1}}{2h_t} + \frac{u_{i-1,k+1} - u_{i-1,k-1}}{2h_t} \right]
$$

$$
+ \frac{c}{4}\left[\frac{u_{i+1,k+1} - u_{i-1,k+1}}{2h_x} + 2\,\frac{u_{i+1,k} - u_{i-1,k}}{2h_x} + \frac{u_{i+1,k-1} - u_{i-1,k-1}}{2h_x} \right]. \quad (3.157)
$$

As in the continuous case, the discrete adjoint equation has the exact same form as the discrete advection equation. Again, we find the same kind of averaging we have already observed in the example of the wave equation (3.97).

3.3.3. Continuous Conservation Laws

We will shortly prove some conservation laws of the advection equation, namely the conservation of the average density, the L^1 and L^2 norms, the total linear momentum and kinetic energy. We consider vertical transformations with generators of the form

$$
X = X^u \frac{\partial}{\partial u} + X^v \frac{\partial}{\partial v} \qquad \text{with} \qquad X^a = X^a(u, v), \quad (3.158)
$$

as discussed in section 2.3. The jet prolongation (2.303) of such vector fields are

$$
j^1 X = X^u \frac{\partial}{\partial u} + X^v \frac{\partial}{\partial v} + \left[\frac{\partial X^u}{\partial u}\frac{\partial u}{\partial x^\mu} + \frac{\partial X^v}{\partial u}\frac{\partial u}{\partial x^\mu} \right]\frac{\partial}{\partial u_\mu} + \left[\frac{\partial X^u}{\partial v}\frac{\partial v}{\partial x^\mu} + \frac{\partial X^v}{\partial v}\frac{\partial v}{\partial x^\mu} \right]\frac{\partial}{\partial v_\mu},
$$
$$
(3.159)
$$

such that the invariance condition (2.306) becomes

$$
j^1 X(L) = X^v \Big(u_t + c\,u_x\Big) + v \Big(u_t + c\,u_x\Big)\left(\frac{\partial X^u}{\partial u} + \frac{\partial X^v}{\partial u} \right). \quad (3.160)
$$

Conservation laws (2.309) take the form

$$
\frac{d}{dt} \int \left[\frac{\partial L}{\partial u_t}\big(j^1 w\big) X^u + \frac{\partial L}{\partial v_t}\big(j^1 w\big) X^v \right] dx = \frac{d}{dt} \int v\, X^u\, dx. \quad (3.161)
$$

We can use these expressions directly to test for several possible symmetries.

Conservation of the Average Density

Consider a vertical transformation generated by $X = (X^u, X^v)$ with

$$X^u = 1, \qquad X^v = 0. \tag{3.162}$$

The Lagrangian is trivially invariant under this transformation

$$j^1 X(L) = 0. \tag{3.163}$$

The corresponding conservation law (3.161) is

$$\frac{d}{dt} \int v \, dx = 0. \tag{3.164}$$

If u is a solution of the advection equation, the pair $(u, v) = (u, u)$ solves the extended system. We can therefore reduce the conserved quantity in (3.164) by specialising it to $v = u$. This gives the conservation of the average density, namely

$$\frac{d}{dt} \int u \, dx = 0. \tag{3.165}$$

If, in addition, the advection equation (3.150) is provided with positive initial conditions, $u_0(x) \geq 0$, the positivity is preserved at later times, and thus (3.165) turns into a conservation law for the L^1 norm of u.

Conservation of the L^2 Norm

Consider a different vertical transformation generated by $X = (X^u, X^v)$ with

$$X^u = u, \qquad X^v = -v. \tag{3.166}$$

The Lagrangian density is invariant under this transformation as well

$$j^1 X(L) = -v\left(u_t + c \, u_x\right) + v\left(u_t + c \, u_x\right) = 0. \tag{3.167}$$

The corresponding conservation law (3.161) is

$$\frac{d}{dt} \int vu \, dx = 0. \tag{3.168}$$

Upon identifying v with u, this gives the conservation of the L^2 norm of u

$$\frac{d}{dt} \int u^2 \, dx = 0. \tag{3.169}$$

Conservation of Linear Momentum and Kinetic Energy

Lastly, consider the following vertical transformation generated by $X = (X^u, X^v)$ with

$$X^u = c, \qquad X^v = 0. \tag{3.170}$$

The Lagrangian density is trivially invariant also under this transformation

$$j^1 X(L) = 0. \tag{3.171}$$

The corresponding conservation law is

$$\frac{d}{dt} \int c\,v\,dx = 0. \tag{3.172}$$

Upon identifying v with u, this gives the conservation of linear momentum

$$\frac{d}{dt} \int c\,u\,dx = 0. \tag{3.173}$$

Conservation of kinetic energy follows exactly the same way by choosing $X^u = \frac{1}{2}c^2$, i.e.,

$$\frac{d}{dt} \int \tfrac{1}{2}c^2\,v\,dx = \frac{d}{dt} \int \tfrac{1}{2}c^2\,u\,dx = 0. \tag{3.174}$$

We have therefore proved conservation of the most important quantities related to the advection equation.

3.3.4. Discrete Conservation Laws

The discrete generator $X_{i,j} = (X_{i,k}^u, X_{i,k}^v)$ of the transformation given by (3.162) is

$$X_{i,k}^u = 1, \qquad X_{i,k}^v = 0. \tag{3.175}$$

The discrete Lagrangian (3.155) is invariant under this transformation

$$
\begin{aligned}
\mathbf{d}\mathcal{L}_\square \cdot X =\ & \frac{\partial \mathcal{L}_d}{\partial u^1}\Big(w_{i,k}, w_{i+1,k}, w_{i+1,k+1}, w_{i,k+1}\Big) X_{i,k}^u \\
& + \frac{\partial \mathcal{L}_d}{\partial u^2}\Big(w_{i,k}, w_{i+1,k}, w_{i+1,k+1}, w_{i,k+1}\Big) X_{i+1,k}^u \\
& + \frac{\partial \mathcal{L}_d}{\partial u^3}\Big(w_{i,k}, w_{i+1,k}, w_{i+1,k+1}, w_{i,k+1}\Big) X_{i+1,k+1}^u \\
& + \frac{\partial \mathcal{L}_d}{\partial u^4}\Big(w_{i,k}, w_{i+1,k}, w_{i+1,k+1}, w_{i,k+1}\Big) X_{i,k+1}^u \\
=\ & \frac{h_t h_x}{8}\Big[v_{i,k} + v_{i+1,k} + v_{i+1,k+1} + v_{i,k+1}\Big]\Big[-\frac{1}{h_t} - \frac{c}{h_x}\Big] X_{i,k}^u \\
& + \frac{h_t h_x}{8}\Big[v_{i,k} + v_{i+1,k} + v_{i+1,k+1} + v_{i,k+1}\Big]\Big[-\frac{1}{h_t} + \frac{c}{h_x}\Big] X_{i+1,k}^u \\
& + \frac{h_t h_x}{8}\Big[v_{i,k} + v_{i+1,k} + v_{i+1,k+1} + v_{i,k+1}\Big]\Big[+\frac{1}{h_t} + \frac{c}{h_x}\Big] X_{i+1,k+1}^u \\
& + \frac{h_t h_x}{8}\Big[v_{i,k} + v_{i+1,k} + v_{i+1,k+1} + v_{i,k+1}\Big]\Big[+\frac{1}{h_t} - \frac{c}{h_x}\Big] X_{i,k+1}^u \\
=\ & 0.
\end{aligned}
\tag{3.176}
$$

The corresponding conservation law (3.143) is

$$
\begin{aligned}
\sum_{i=1}^{n_x-1} &\left[\frac{\partial \mathcal{L}_d}{\partial u^3}\Big(w_{i-1,k-1}, w_{i,k-1}, w_{i,k}, w_{i-1,k}\Big) + \frac{\partial \mathcal{L}_d}{\partial u^4}\Big(w_{i,k-1}, w_{i+1,k-1}, w_{i+1,k}, w_{i,k}\Big)\right] \cdot X_{i,k}^u \\
&= h_x \sum_{i=1}^{n_x-1} \frac{1}{4}\Big[v_{i,k} + v_{i+1,k} + v_{i+1,k+1} + v_{i,k+1}\Big] \\
&= h_x \sum_{i=1}^{n_x-1} \frac{1}{4}\Big[u_{i,k} + u_{i+1,k} + u_{i+1,k+1} + u_{i,k+1}\Big],
\end{aligned}
\tag{3.177}
$$

where the last equality arises from identifying v with u. The conservation of the discrete momentum and energy follows along the same lines.

4. Charged Particle Motion

This section addresses a reduced description of the motion of charged particles in a plasma, the so called guiding centre dynamics. It can be seen as a limit of gyrokinetic theory, the predominant model used in plasma physical particle-in-cell codes, that is valid when the magnetic field is very strong, such that the gyration orbit of the particle is very small, or when the electromagnetic field is almost spatially uniform, such that it varies only very little along the gyration orbit. It is also closely related to drift kinetic theory which is a reduced model of kinetic theory (see next chapter).

After a short summary of the Lagrangian formulation for this set of dynamical equations, a set of variational integrators based on different quadrature rules is derived. This integrators is adapted to the reduced dynamics in the poloidal plane similar to [103] as well as to the dynamics in full tokamak geometry. Several higher order methods are derived by composition methods. The advantages of the variational discretisations compared to widely used Runge-Kutta schemes are demonstrated.

4.1. Guiding Centre Dynamics

In a magnetic field, charged particles move along a helix. This motion can be decomposed into the gyration about a magnetic field line, and the motion of the centre of the gyration (guiding centre) along the field line. Mathematically, this amounts to a coordinate transformation from spatial coordinates x and the corresponding velocities $\dot{x}$ to guiding centre coordinates (X, Θ, u, μ), where X is the position of the guiding centre, Θ is the angle of the gyration (gyrophase), μ the magnetic moment, and $u = \dot{x} \cdot b$ is the velocity along the magnetic field lines (parallel velocity).

Littlejohn [77] was the first to devise a variational principle for the guiding centre motion and thereby find simple proofs for conservation of energy and angular momentum. His Lagrangian reads

$$L = A^* \cdot \dot{X} + \mu \dot{\Theta} - H \tag{4.1}$$

with the Hamiltonian

$$H = \frac{1}{2} u^2 + \mu B + \varphi \tag{4.2}$$

and the so called "modified vector potential" (first discovered by Morozov and Solov'ev [88])

$$A^* = A + u b. \tag{4.3}$$

Here, B is the magnetic field strength, b its unit vector, and φ is the electrostatic potential. Units are chosen such that $e = m = c = 1$, with charge e, particle mass m, and speed of light c.

The Lagrangian (4.1) is regarded as a function of the guiding centre variables (X, Θ, u, μ) and their time derivatives $(\dot{X}, \dot{\Theta}, \dot{u}, \dot{\mu})$. For the spatial components, the Euler-Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}^i}\right) - \frac{\partial L}{\partial X^i} = 0 \tag{4.4}$$

which explicitly amounts to

$$\frac{d}{dt}A_i^* - A_{j,i}^* \dot{X}^j - \mu B_{,i} - \varphi_{,i} = 0. \tag{4.5}$$

Computing the time derivative, this becomes

$$\left(A_{i,j}^* - A_{j,i}^*\right)\dot{X}^j - \mu B_{,i} - \varphi_{,i} = 0. \tag{4.6}$$

From the equation of the gyrophase

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\Theta}}\right) = 0 \qquad \text{or} \qquad \dot{\mu} = 0 \tag{4.7}$$

the conservation of the magnetic moment μ is obtained. The Euler-Lagrange equation of the parallel velocity is just

$$\frac{\partial L}{\partial u} = 0 \qquad \text{or} \qquad u = b \cdot \dot{X}, \tag{4.8}$$

i.e. the definition of the parallel velocity. And from the equation of μ

$$\frac{\partial L}{\partial \mu} = 0 \qquad \text{or} \qquad \dot{\Theta} = B \tag{4.9}$$

we obtain, upon restoration of physical units, the definition of the gyro frequency

$$\dot{\Theta} = \omega = \frac{eB}{m}. \tag{4.10}$$

If the variation of the background electromagnetic fields is small along the radius of the gyration, the particle's motion can be approximated by the motion of just the guiding centre, averaging over the gyrophase. The corresponding reduced Lagrangian is

$$L = A^* \cdot \dot{X} - H, \qquad\qquad H = \frac{1}{2}u^2 + \mu B - \varphi. \tag{4.11}$$

This is the starting point for the derivation of a set of variational integrators for the guiding centre motion of charged particles in a tokamak.

4.2. Variational Discretisation

At first, the general derivation of a variational integrator for guiding centre motion is reproduced similar to Qin et al. [103], where the trapezoidal rule is used to discretise the Lagrangian. In addition, we provide the derivation of a second integrator based on the midpoint rule that appears to be more stable and small timesteps and yields more accurate results. In the last section the construction of higher order schemes by composition of low order schemes is sketched.

To allow for a compact notation, we introduce the generalised coordinates $q^i = X^i$ with $i = \{1, 2, 3\}$ and $q^u = u$. Together, they are denoted q^ν with $\nu = \{1, 2, 3, u\}$ or just q. Correspondingly, the conjugate momenta are denoted p^ν or just p.

4.2.1. Trapezoidal Discretisation

Applying a trapezoidal discretisation to Littlejohn's guiding centre Lagrangian (4.1) gives

$$L_d(q_k, q_{k+1}) = \frac{h}{2} L\left(q_k, \frac{q_{k+1} - q_k}{h}\right) + \frac{h}{2} L\left(q_{k+1}, \frac{q_{k+1} - q_k}{h}\right) \tag{4.12}$$

$$= h\left[\frac{A_i^*(q_k) + A_i^*(q_{k+1})}{2} \frac{q_{k+1}^i - q_k^i}{h} - \frac{(q_k^u)^2 + (q_{k+1}^u)^2}{4}\right.$$
$$\left. - \nu \frac{B(q_k) + B(q_{k+1})}{2} - \frac{\phi(q_k) + \phi(q_{k+1})}{2}\right]. \tag{4.13}$$

This Lagrangian, however, results in a scheme with small stability region, as the expression resulting from the $(q^u)^2$ term is explicit in q^u. Qin et al. [103] replace this term with $q^k q^{k+1}$. We explore a different modification in this term to make the resulting expression implicit in q^u, that is

$$L_d(q_k, q_{k+1}) = h\left[\frac{A_i^*(q_k) + A_i^*(q_{k+1})}{2} \frac{q_{k+1}^i - q_k^i}{h} - \frac{1}{2}\left(\frac{q_k^u + q_{k+1}^u}{2}\right)^2\right.$$
$$\left. - \nu \frac{B(q_k) + B(q_{k+1})}{2} - \frac{\phi(q_k) + \phi(q_{k+1})}{2}\right]. \tag{4.14}$$

Questions about which discretisations of the Lagrangian can be regarded as "good", i.e., produce well working, stable schemes, and which discretisations work less well remain largely unanswered. Unfortunately, there exist no clear guidelines for the discretisation of the Lagrangian, but the preservation of symmetries in the course of discretisation certainly plays an important role.

Continuing with the derivation of the discrete Euler-Lagrange equations, these are defined as

$$\frac{\partial}{\partial q_k^j}\left[L_d(q_{k-1}, q_k) + L_d(q_k, q_{k+1})\right] =$$

$$= h\left[A_{i,j}^*(q_k) \frac{q_{k+1}^i - q_{k-1}^i}{2h} - \frac{A_j^*(q_{k+1}) - A_j^*(q_{k-1})}{2h} - \mu B_{,j}(q_k) - \varphi_{,j}(q_k)\right] = 0, \tag{4.15a}$$

$$\frac{\partial}{\partial q_k^u}\left[L_d(q_{k-1}, q_k) + L_d(q_k, q_{k+1})\right] =$$

$$= h\left[b_i(q_k) \frac{q_{k+1}^i - q_{k-1}^i}{2h} - \frac{q_{k-1}^u + 2\,q_k^u + q_{k+1}^u}{4}\right] = 0. \tag{4.15b}$$

This set of equations forms an implicit system for the solution of the guiding centre dynamics depending on data at three points in time, q_{k+1}, q_k and q_{k-1}. It constitutes a discrete map

$$(q_{k-1}, q_k) \mapsto (q_k, q_{k+1}). \tag{4.16}$$

Solving the system for q_{k+1} yields nonlinearly implicit iteration rules for integrating the discrete phasespace trajectory $\{q_k\}_{k=0}^N$ of the particle. A possible solution strategy is to use a Newton solver for the nonlinear iteration. If the initial guess, e.g., by the linearised scheme derived below, is sufficiently close to the solution, no more than two or three iterations are needed. Using an analytic solution in the Newton iteration and a fixed number

of iterations per timestep, such that an evaluation of the residual becomes unnecessary, the computational effort is about the same as for a standard fourth order Runge-Kutta method. The variational integrator is therefore computationally competitive to an explicit standard method while yielding superior results.

Linearisation

The linearisation of (4.15) allows for an easier comparison with the continuous equations of motion (4.6-4.9) and highlights the differences between the variational integrator and a direct discretisation. It is also possible to use this linearised scheme to compute an initial guess for a nonlinear solver applied to the above scheme.

Expand the $[A_j^*(q_{k+1}) - A_j^*(q_{k-1})]/2h$ term in (4.15a) into a Taylor series about q_k

$$\frac{A_j^*(q_{k+1}) - A_j^*(q_{k-1})}{2h} \approx A_{j,i}^*(q_k) \frac{q_{k+1}^i - q_{k-1}^i}{2h} + b_j(q_k) \frac{q_{k+1}^u - q_{k-1}^u}{2h}. \tag{4.17}$$

Upon insertion into (4.15a) we obtain a linearised set of equations

$$\left[A_{i,j}^*(q_k) - A_{j,i}^*(q_k) \right] \frac{q_{k+1}^i - q_{k-1}^i}{2h} - b_j(q_k) \frac{q_{k+1}^u - q_{k-1}^u}{2h} - \mu B_{,j}(q_k) - \varphi_{,j}(q_k) = 0, \tag{4.18a}$$

$$b_i(q_k) \frac{q_{k+1}^i - q_{k-1}^i}{2h} - \frac{q_{k-1}^u + 2\,q_k^u + q_{k+1}^u}{4} = 0. \tag{4.18b}$$

Except for the additional $\dot{u}b$ term, these equations resemble (4.6) and (4.8), while the original equation (4.15a) resembles (4.5).

Position Momentum Form

The position-momentum form (3.14) of the trapezoidal integrator (4.15) is given by

$$p_k^j = -\frac{\partial}{\partial q_k^j} L_d(q_k, q_{k+1}), \qquad\qquad p_{k+1}^j = \frac{\partial}{\partial q_{k+1}^j} L_d(q_k, q_{k+1}), \tag{4.19a}$$

$$p_k^u = -\frac{\partial}{\partial q_k^u} L_d(q_k, q_{k+1}), \qquad\qquad p_{k+1}^u = \frac{\partial}{\partial q_{k+1}^u} L_d(q_k, q_{k+1}). \tag{4.19b}$$

Explicitly computing these expressions gives

$$p_k^j = -\frac{1}{2} A_{i,j}^*(q_k) \left[q_{k+1}^i - q_k^i \right] + \frac{1}{2} \left[A_j^*(q_k) + A_j^*(q_{k+1}) \right] + \frac{h}{2} \mu B_{,j}(q_k) + \frac{h}{2} \phi_{,j}(q_k), \tag{4.20a}$$

$$p_k^u = -\frac{1}{2} b_i(q_k) \left[q_{k+1}^i - q_k^i \right] + \frac{h}{4} \left[q_k^u + q_{k+1}^u \right], \tag{4.20b}$$

$$p_{k+1}^j = +\frac{1}{2} A_{i,j}^*(q_{k+1}) \left[q_{k+1}^j - q_k^j \right] + \frac{1}{2} \left[A_j^*(q_k) + A_j^*(q_{k+1}) \right] - \frac{h}{2} \mu B_{,j}(q_{k+1}) - \frac{h}{2} \phi_{,j}(q_{k+1}), \tag{4.20c}$$

$$p_{k+1}^u = +\frac{1}{2} b_i(q_{k+1}) \left[q_{k+1}^i - q_k^i \right] - \frac{h}{4} \left[q_k^u - q_{k+1}^u \right]. \tag{4.20d}$$

This set of equations forms a nonlinearly implicit system for the solution of the guiding centre dynamics depending only on data at two points in time, q_{k+1} and q_k. The first two equations have to be solved for the q_{k+1}. Afterwards, the q_{k+1} are straight forwardly computed, as the other two equations are merely explicit functions. The discrete map corresponding to this formulation is

$$(q_k, p_k) \mapsto (q_{k+1}, p_{k+1}). \tag{4.21}$$

Solving this system, results in the same trajectory $\{q_k\}_{k=0}^N$ as solving the system (4.15).

Discrete Cartan Form

The discrete Cartan one-forms corresponding to the Lagrangian (4.14) are given by

$$\Theta_{L_d}^-(q_k, q_{k+1}) = -\frac{\partial}{\partial q_k^\nu} L_d(q_k, q_{k+1})\, dq_k^\nu, \qquad \Theta_{L_d}^+(q_k, q_{k+1}) = +\frac{\partial}{\partial q_{k+1}^\nu} L_d(q_k, q_{k+1})\, dq_{k+1}^\nu,$$

(4.22)

explicitly computed to be

$$\Theta_{L_d}^-(q_k, q_{k+1}) = -\left(\frac{1}{h}\, A_{i,j}^*(q_k) \left[q_{q_{k+1}}^i - q_k^i \right] - \frac{1}{h}\left[A_j^*(q_k) + A_j^*(q_{k+1}) \right] \right.$$
$$\left. - \frac{h}{2}\mu B_{,j}(q_k) - \frac{h}{2}\phi_{,j}(q_k) \right) dq_k^j$$
$$- \left(\frac{1}{2} b_j(q_k) \left[q_{k+1}^j - q_k^j \right] - \frac{h}{4}\left[q_k^u + q_{k+1}^u \right] \right) dq_k^u, \qquad (4.23)$$

$$\Theta_{L_d}^+(q_k, q_{k+1}) = +\left(\frac{1}{2}\, A_{i,j}^*(q_{k+1}) \left[q_{k+1}^i - q_k^i \right] + \frac{1}{2}\left[A_j^*(q_k) + A_j^*(q_{k+1}) \right] \right.$$
$$\left. - \frac{h}{2}\mu B_{,j}(q_{k+1}) - \frac{h}{2}\phi_{,j}(q_{k+1}) \right) dq_{k+1}^j$$
$$+ \left(\frac{1}{2} b_j(q_{k+1}) \left[q_{k+1}^j - q_k^j \right] - \frac{h}{4}\left[q_k^u + q_{k+1}^u \right] \right) dq_{k+1}^u. \qquad (4.24)$$

Together, these two forms determine the exterior derivative of the Lagrangian $L_d(q_k, q_{k+1})$

$$\mathbf{d}L_d(q_k, q_{k+1}) = \Theta_{L_d}^+(q_k, q_{k+1}) - \Theta_{L_d}^-(q_k, q_{k+1}). \qquad (4.25)$$

As $\mathbf{dd}L_d = 0$, the exterior derivative of both one-forms defines the same two-form

$$\Omega_d(q_k, q_{k+1}) = \mathbf{d}\Theta_{L_d}^+ = \mathbf{d}\Theta_{L_d}^-. \qquad (4.26)$$

This is the discrete symplectic two-form Ω_d. Its preservation along the Lagrangian flow is given by construction and has been shown in section 3.1.4.

4.2.2. Midpoint Discretisation

We now derive an alternative method based on a midpoint discretisation. The main advantage over the trapezoidal scheme from the last section is a higher accuracy of the method. We restrict ourselves to deriving the Euler-Lagrange equations and the position-momentum form of the equations of motion and do not repeat the derivation of the Cartan one-form.

Applying a midpoint discretisation to Littlejohn's guiding centre Lagrangian (4.1) gives

$$L_d(q_k, q_{k+1}) = h\, L\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h} \right) = h\, L\left(q_{k+1/2}, \frac{q_{k+1} - q_k}{h} \right) \qquad (4.27)$$

$$= h\left[A_i^*(q_{k+1/2}) \frac{q_{k+1}^i - q_k^i}{h} - \left(\frac{q_k^u + q_{k+1}^u}{2} \right)^2 - \mu B(q_{k+1/2}) - \phi(q_{k+1/2}) \right], \qquad (4.28)$$

where $q_{k+1/2} = \frac{1}{2}(q_k + q_{k+1})$. The discrete Euler-Lagrange equations of this midpoint discretisation are computed as

$$
\frac{\partial}{\partial q_k^j}\Big[L_d(q_{k-1}, q_k) + L_d(q_k, q_{k+1})\Big] =
$$

$$
= \frac{1}{2}\Big[A_{i,j}^*(q_{k+1/2})\Big]\Big[q_{k+1}^i - q_k^i\Big] + \frac{1}{2}\Big[A_{i,j}^*(q_{k-1/2})\Big]\Big[q_k^i - q_{k-1}^i\Big] - \Big[A_j^*(q_{k+1/2}) - A_j^*(q_{k-1/2})\Big]
$$

$$
- \frac{h}{2}\,\mu\Big[B_{,j}(q_{k-1/2}) + B_{,j}(q_{k+1/2})\Big] - \frac{h}{2}\Big[\varphi_{,j}(q_{k-1/2}) + \varphi_{,j}(q_{k+1/2})\Big] = 0,
$$

(4.29a)

$$
\frac{\partial}{\partial q_k^u}\Big[L_d(q_{k-1}, q_k) + L_d(q_k, q_{k+1})\Big] =
$$

$$
= \frac{1}{2}\Big[b_i(q_{k-1/2})\Big[q_k^i - q_{k-1}^i\Big] + b_i(q_{k+1/2})\Big[q_{k+1}^i - q_k^i\Big]\Big] - \frac{h}{4}\Big[q_{k-1}^u + 2\,q_k^u + q_{k+1}^u\Big] = 0.
$$

(4.29b)

The position-momentum form of the midpoint integrator is computed as

$$
p_k^i = -\frac{1}{2}\Big[A_{i,j}^*(q_{k+1/2})\Big]\Big[q_{k+1}^i - q_k^i\Big] + A_j^*(q_{k+1/2}) + \frac{h}{2}\,\mu B_{,j}(q_{k+1/2}) + \frac{h}{2}\,\varphi_{,j}(q_{k+1/2}) \quad (4.30a)
$$

$$
p_k^u = -\frac{1}{2}\,b_i(q_{k+1/2})\Big[q_{k+1}^i - q_k^i\Big] + \frac{h}{4}\Big[q_k^u + q_{k+1}^u\Big] \quad (4.30b)
$$

$$
p_{k+1}^i = +\frac{1}{2}\Big[A_{i,j}^*(q_{k+1/2})\Big]\Big[q_{k+1}^i - q_k^i\Big] + A_j^*(q_{k+1/2}) - \frac{h}{2}\,\mu B_{,j}(q_{k+1/2}) - \frac{h}{2}\,\varphi_{,j}(q_{k+1/2}) \quad (4.30c)
$$

$$
p_{k+1}^u = +\frac{1}{2}\,b_i(q_{k+1/2})\Big[q_{k+1}^i - q_k^i\Big] - \frac{h}{4}\Big[q_k^u + q_{k+1}^u\Big]. \quad (4.30d)
$$

The first two equations have to be solved for q_{k+1}. The solution of the second two equations for p_{k+1} is then straight forward.

4.2.3. Higher Order Schemes

If the discrete Lagrangian is self-adjoint, c.f. section 3.1.5,

$$
L_d^*(q_k, q_{k+1}, h) \equiv -L_d(q_{k+1}, q_k, -h) = L_d(q_{k+1}, q_k, h), \tag{4.31}
$$

the resulting variational integrator can be composed to yield higher order methods. Both, the trapezoidal Lagrangian (4.12)

$$
(L_d^{\mathrm{tr}})^*(q_k, q_{k+1}, h) = -(-h)\frac{1}{2}\Bigg[L\Big(q_{k+1}, \frac{q_k - q_{k+1}}{-h}\Big) - L\Big(q_k, \frac{q_k - q_{k+1}}{-h}\Big)\Bigg]
$$

$$
= \frac{h}{2}\Bigg[L\Big(q_k, \frac{q_{k+1} - q_k}{h}\Big) - L\Big(q_{k+1}, \frac{q_{k+1} - q_k}{h}\Big)\Bigg]
$$

$$
= L_d(q_k, q_{k+1}, h)
$$

and the midpoint Lagrangian (4.27)

$$
(L_d^{\mathrm{mp}})^*(q_k, q_{k+1}, h) = -(-h)\,L\Big(\frac{q_{k+1} + q_k}{2}, \frac{q_k - q_{k+1}}{-h}\Big)
$$

$$
= h\,L\Big(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}\Big)
$$

$$
= L_d^{\mathrm{mp}}(q_k, q_{k+1}, h)
$$

fulfil this property. The same is true for the modified trapezoidal Lagrangian (4.14), so that we can apply the composition rules from section 3.1.5 to the schemes just derived.

4.3. Particle Motion in the Poloidal Plane

In this section we want to apply the derivations from the previous section to the motion of a charged particle in axisymmetric tokamaks. The toroidal symmetry allows us to reduce the dynamics to the poloidal plane of a tokamak. At first, the derivation of Qin et al. [103] is reproduced, with the difference that cylindrical coordinates (R, Z) are used instead of toroidal coordinates (r, θ). The toroidal symmetry implies that the toroidal momentum p_φ is conserved and can be used to express the parallel velocity u as a function of (R, Z). Hence, upon prescribing the values of the toroidal momentum p_φ and the magnetic moment μ, only the coordinates of the poloidal plane (R, Z) are treated as dynamical variables.

In [103], the Lagrangian is discretised with the trapezoidal rule exclusively. We will also derive an integrator based on the midpoint rule as well as a composition scheme based thereon. Finally we will compare all three of these schemes with a standard explicit Runge-Kutta method. We restrict our treatment to the position-momentum form as that appears more natural with respect to the specification of initial conditions.

Magnetic Field and Vector Potential

For the magnetic field B and the vector potential A we will use analytic expressions following Qin et al. [103]. The vector potential is given as

$$A_R = \frac{B_0 R_0 Z}{2R}, \quad A_Z = -\ln\left(\frac{R}{R_0}\right)\frac{B_0 R_0}{2}, \quad A_\varphi = -\frac{B_0 r^2}{2qR}, \quad r = \sqrt{(R - R_0)^2 + Z^2}, \quad (4.32)$$

where subscripts R, Z and φ denote the radial, vertical and toroidal components, respectively. The magnetic field $B = \nabla \times A$ is

$$B_R = -\frac{B_0 Z}{qR}, \qquad B_Z = \frac{B_0(R - R_0)}{qR}, \qquad B_\varphi = -\frac{B_0 R_0}{R}, \qquad B = \frac{B_0 S}{qR},$$

with the normalised magnetic field being

$$b_R = -\frac{Z}{S}, \qquad b_Z = \frac{R - R_0}{S}, \qquad b_\varphi = -\frac{qR_0}{S}, \qquad S = \sqrt{r^2 + q^2 R_0^2}. \qquad (4.33)$$

Here, R_0 is the radial position of the magnetic axis, B_0 is the magnetic field at R_0, and q is the safety factor, regarded as constant. All the derivatives of the above expressions, which will be needed in the derived schemes, are listed in section 4.C.

Toroidal Momentum and Parallel Velocity

It will be practical to express the Lagrangian with respect to to the momenta (p_R, p_Z, p_φ), such that in cylinder coordinates (R, Z, φ) we have

$$L = p_R \dot{R} + p_Z \dot{Z} + p_\varphi \dot{\varphi} - H, \qquad\qquad H = \frac{1}{2}u^2 + \mu B - \phi, \qquad (4.34)$$

with

$$p_R = A_R^* = A_R + ub_R, \qquad p_Z = A_Z^* = A_Z + ub_Z, \qquad p_\varphi = RA_\varphi^* = R\left(A_\varphi + ub_\varphi\right). \qquad (4.35)$$

As p_φ is conserved ($\partial L/\partial \varphi = 0$),

$$p_\varphi = -R\left[\frac{B_0 r^2}{2qR} + u\frac{qR_0}{S}\right] = const, \qquad (4.36)$$

we can compute a functional expression for the parallel velocity u in which p_φ is a parameter

$$u = -\left[p_\varphi + \frac{B_0 r^2}{2q}\right]\frac{S}{qRR_0} = -\left[p_\varphi + \frac{B_0 r^2}{2q}\right]\frac{B}{B_0 R_0}. \qquad (4.37)$$

Reduced Lagrangian and Generalised Coordinates

Projecting the motion to the poloidal plane and assuming the absence of any electrostatic field, the Lagrangian reduces to

$$L = p_R\dot{R} + p_Z\dot{Z} - H = A_R^*\dot{R} + A_Z^*\dot{Z} - \frac{1}{2}u^2 - \mu B, \qquad (4.38)$$

where the components of the generalised magnetic potential read

$$A_R^* = \frac{B_0 R_0 Z}{2R} + \left[p_\varphi + \frac{B_0 r^2}{2q}\right]\frac{Z}{qRR_0}, \qquad A_Z^* = -\ln\left(\frac{R}{R_0}\right)\frac{B_0 R_0}{2} - \left[p_\varphi + \frac{B_0 r^2}{2q}\right]\frac{R - R_0}{qRR_0}, \qquad (4.39)$$

and the parallel velocity u is given by (4.37). We introduce generalised coordinates q_k^R and q_k^Z with discrete conjugate momenta p_k^R and p_k^Z

$$L = A_R^*\dot{q}^R + A_Z^*\dot{q}^Z - \frac{1}{2}u^2 - \mu B. \qquad (4.40)$$

This is the basis for the following discretisations.

4.3.1. Trapezoidal Discretisation

Applying a trapezoidal discretisation to the reduced guiding centre Lagrangian (4.40) gives

$$\begin{aligned}
L_d^{\mathrm{tr}}(q_k, q_{k+1}) &= \frac{h}{2}L\left(q_k, \frac{q_{k+1} - q_k}{h}\right) + \frac{h}{2}L\left(q_{k+1}, \frac{q_{k+1} - q_k}{h}\right) \\
&= h\left[\frac{A_R^*(q_k) + A_R^*(q_{k+1})}{2}\frac{q_{k+1}^R - q_k^R}{h} + \frac{A_Z^*(q_k) + A_Z^*(q_{k+1})}{2}\frac{q_{k+1}^Z - q_k^Z}{h}\right. \\
&\qquad \left. - \frac{1}{2}\left(\frac{u(q_k) + u(q_{k+1})}{2}\right)^2 - \mu\frac{B(q_k) + B(q_{k+1})}{2}\right]. \qquad (4.41)
\end{aligned}$$

The position-momentum form (3.14) of the trapezoidal integrator is computed as

$$p_k^R = \frac{1}{2}\Bigg[-A_{R,R}^*(q_k)\Big[q_{k+1}^R - q_k^R\Big] - A_{Z,R}^*(q_k)\Big[q_{k+1}^Z - q_k^Z\Big] + \Big[A_R^*(q_k) + A_R^*(q_{k+1})\Big]$$

$$+ \frac{h}{2}\, u_{,R}(q_k)\Big[u(q_k) + u(q_{k+1})\Big] + h\,\mu B_{,R}(q_k)\Bigg], \tag{4.42a}$$

$$p_k^Z = \frac{1}{2}\Bigg[-A_{R,Z}^*(q_k)\Big[q_{k+1}^R - q_k^R\Big] - A_{Z,Z}^*(q_k)\Big[q_{k+1}^Z - q_k^Z\Big] + \Big[A_Z^*(q_k) + A_Z^*(q_{k+1})\Big]$$

$$+ \frac{h}{2}\, u_{,Z}(q_k)\Big[u(q_k) + u(q_{k+1})\Big] + h\,\mu B_{,Z}(q_k)\Bigg], \tag{4.42b}$$

$$p_{k+1}^R = \frac{1}{2}\Bigg[+A_{R,R}^*(q_{k+1})\Big[q_{k+1}^R - q_k^R\Big] + A_{Z,R}^*(q_{k+1})\Big[q_{k+1}^Z - q_k^Z\Big] + \Big[A_R^*(q_k) + A_R^*(q_{k+1})\Big]$$

$$- \frac{h}{2}\, u_{,R}(q_{k+1})\Big[u(q_k) + u(q_{k+1})\Big] - h\,\mu B_{,R}(q_{k+1})\Bigg], \tag{4.42c}$$

$$p_{k+1}^Z = \frac{1}{2}\Bigg[+A_{R,Z}^*(q_{k+1})\Big[q_{k+1}^R - q_k^R\Big] + A_{Z,Z}^*(q_{k+1})\Big[q_{k+1}^Z - q_k^Z\Big] + \Big[A_Z^*(q_k) + A_Z^*(q_{k+1})\Big]$$

$$- \frac{h}{2}\, u_{,Z}(q_{k+1})\Big[u(q_k) + u(q_{k+1})\Big] - h\,\mu B_{,Z}(q_{k+1})\Bigg]. \tag{4.42d}$$

The first two equations constitute a nonlinear system determining q_{k+1}^R and q_{k+1}^Z and can be written as a function $F(q_k, p_k, q_{k+1}) = 0$. We solve it by Newton iteration with analytic Jacobian $\mathcal{J}$, which is determined by computing the variation of the two equations with respect to q_{k+1}^R and q_{k+1}^Z, such that

$$\mathcal{J}\,\delta q_{k+1}^{n+1} = \frac{1}{2}\begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}\begin{pmatrix} \delta q_{k+1}^R \\ \delta q_{k+1}^Z \end{pmatrix} = -\begin{pmatrix} F_R(q_k, p_k, q_{k+1}^n) \\ F_Z(q_k, p_k, q_{k+1}^n) \end{pmatrix}, \tag{4.43}$$

where n denotes the Newton step, such that

$$q_{k+1}^{n+1} = q_{k+1}^n + \delta q_{k+1}^{n+1}. \tag{4.44}$$

The components of the Jacobian can be found in section 4.B. After we solved for q_{k+1}^R and q_{k+1}^Z, it is straight forward to evaluate (4.42c) and (4.42d) as these are explicit functions for $p_{k+1}^R(q_k, p_k, q_{k+1})$ and $p_{k+1}^Z(q_k, p_k, q_{k+1})$.

4.3.2. Midpoint Discretisation

Applying a midpoint discretisation to the reduced guiding centre Lagrangian (4.40) gives

$$L_d^{\mathrm{mp}}(q_k, q_{k+1}) = h\, L\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}\right) = h\, L\left(q_{k+1/2}, \frac{q_{k+1} - q_k}{h}\right)$$

$$= h\left[A_R^*(q_{k+1/2})\,\frac{q_{k+1}^R - q_k^R}{h} + A_Z^*(q_{k+1/2})\,\frac{q_{k+1}^Z - q_k^Z}{h} - \frac{u^2(q_{k+1/2})}{2} - \mu B(q_{k+1/2})\right]. \tag{4.45}$$

The position-momentum form (3.14) of the midpoint integrator is computed as

$$p_k^R = \frac{1}{2}\Bigg[-A_{R,R}^\star(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] - A_{Z,R}^\star(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] + 2\,A_R^\star(q_{k+1/2})$$

$$+ h\,u(q_{k+1/2})\,u_{,R}(q_{k+1/2}) + h\,\mu B_{,R}(q_{k+1/2})\Bigg], \tag{4.46a}$$

$$p_k^Z = \frac{1}{2}\Bigg[-A_{R,Z}^\star(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] - A_{Z,Z}^\star(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] + 2\,A_Z^\star(q_{k+1/2})$$

$$+ h\,u(q_{k+1/2})\,u_{,Z}(q_{k+1/2}) + h\,\mu B_{,Z}(q_{k+1/2})\Bigg], \tag{4.46b}$$

$$p_{k+1}^R = \frac{1}{2}\Bigg[+A_{R,R}^\star(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] + A_{Z,R}^\star(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] + 2\,A_R^\star(q_{k+1/2})$$

$$- h\,u(q_{k+1/2})\,u_{,R}(q_{k+1/2}) - h\,\mu B_{,R}(q_{k+1/2})\Bigg], \tag{4.46c}$$

$$p_{k+1}^Z = \frac{1}{2}\Bigg[+A_{R,Z}^\star(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] + A_{Z,Z}^\star(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] + 2\,A_Z^\star(q_{k+1/2})$$

$$- h\,u(q_{k+1/2})\,u_{,Z}(q_{k+1/2}) - h\,\mu B_{,Z}(q_{k+1/2})\Bigg]. \tag{4.46d}$$

The solution strategy is the same as for the trapezoidal method.

4.3.3. Numerical Results

We want to compare the different variational integrators with each other, the different composition methods, and a standard fourth order Runge-Kutta discretisation, where our main focus lies on the energy error and the geometry of the particle orbit for long time integration. In the following, we consider the trapped particle example from Qin et al. [103], which is initialised by

$$R = R_0 + 0.05, \qquad Z = 0, \qquad \mu = 2.25 \times 10^{-6}, \qquad p_\varphi = -1.077 \times 10^{-3},$$

with

$$R_0 = 1, \qquad B_0 = 1, \qquad q = 2, \qquad \tau_b = 43107,$$

where τ_b is the bounce time, determining the timestep. We used the normalisation proposed by Qin et al. [103], where the parameters of the tokamak geometry are normalised by R_0 and B_0, to be able to compare with their results.

We compare the evolution of the particle orbit for the two variational integrators and the Runge-Kutta method at three different timesteps that correspond to 25, 50 and 100 timesteps per bounce period, respectively. In all three cases, we observe that the variational integrator follows the expected orbit accurately for long times, while the Runge-Kutta method exhibits substantial deviations (see figures 4.1 - 4.3). Due to the loss of energy in the Runge-Kutta simulations (see figures 4.4 - 4.6), the particle orbits shrinks until eventually it almost contracts to a point. On the contrary, the variational integrators exhibit an oscillating energy error, with a constant amplitude of the oscillation. The amplitude of the energy error scales according to the order of the scheme, which is

(a) Runge-Kutta   **(b)** Variational Midpoint
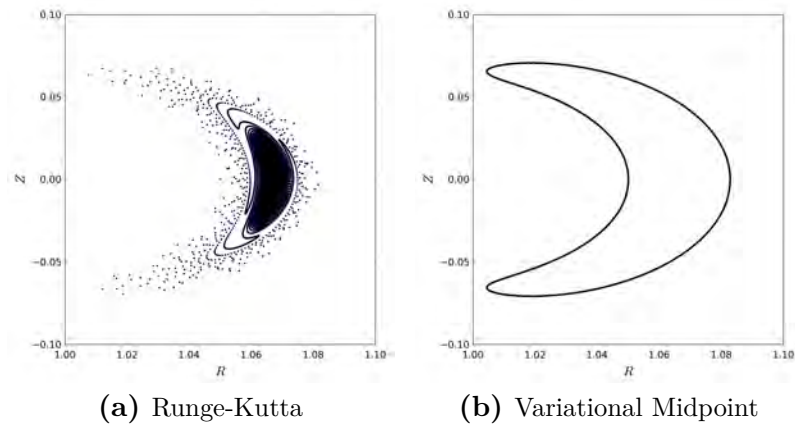
Fig. 4.1.: Trapped particle after 8.000 bounce periods with 25 steps per bounce period. The trapezoidal integrator is not stable for the timestep considered in this example.

(a) Runge-Kutta   **(b)** Variational Trapezoidal   **(c)** Variational Midpoint

Fig. 4.2.: Trapped particle after 200.000 bounce periods with 50 steps per bounce period.

(a) Runge-Kutta   **(b)** Variational Trapezoidal   **(c)** Variational Midpoint
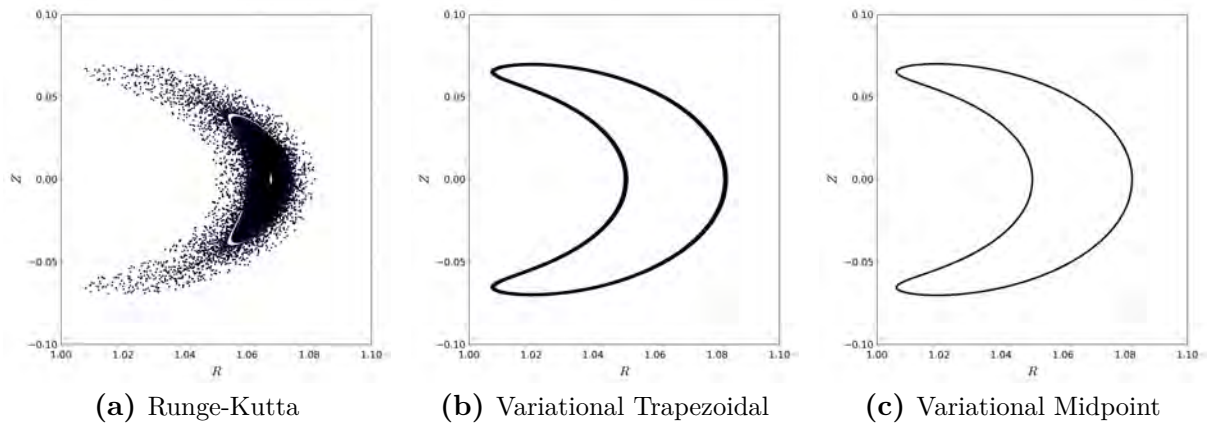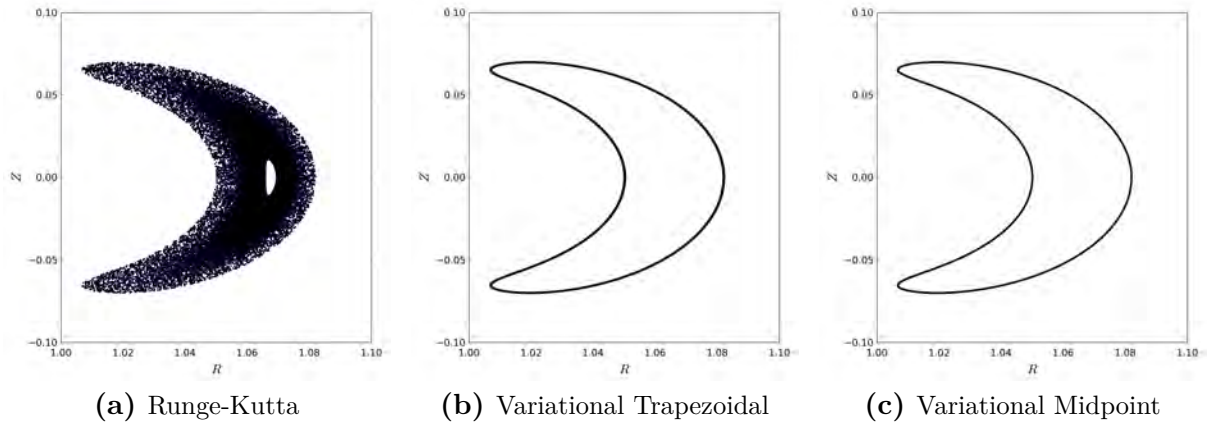
Fig. 4.3.: Trapped particle after two million bounce periods with 100 steps per bounce period.
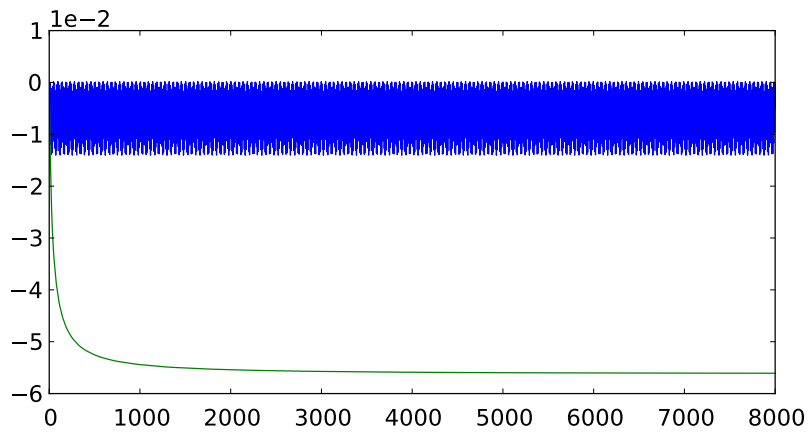
Fig. 4.4.: Energy error for trapped particle with 25 timesteps per bounce period (green: Runge-Kutta, blue: variational midpoint).
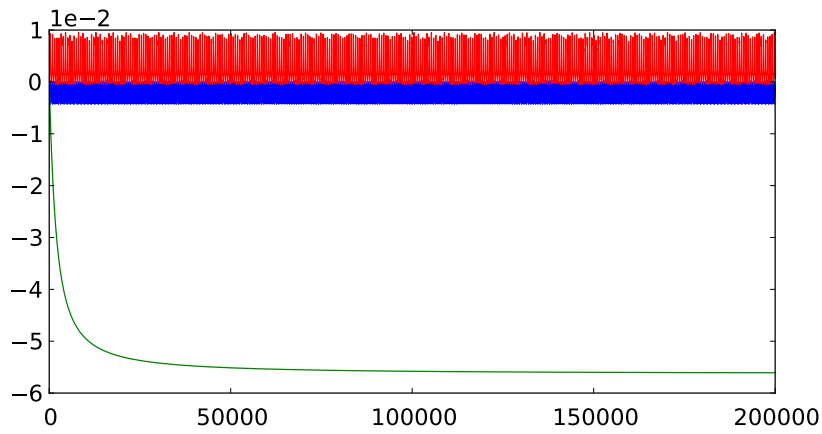
Fig. 4.5.: Energy error for trapped particle with 50 timesteps per bounce period (green: Runge-Kutta, blue: variational trapezoidal, red: variational midpoint).
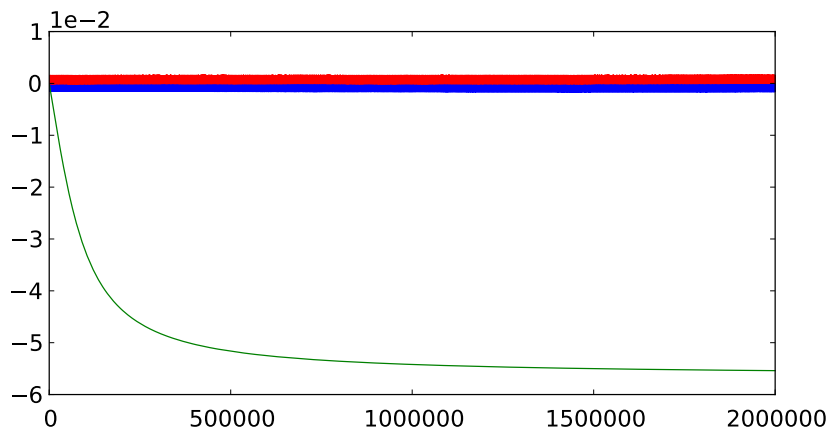
Fig. 4.6.: Energy error for trapped particle with 100 timesteps per bounce period (green: Runge-Kutta, blue: variational trapezoidal, red: variational midpoint).

second-order accurate, i.e., halving the timestep results in a reduction of the error by a factor of four.

The stability region of the trapezoidal integrator appears to be smaller than that of the midpoint integrator as simulations with 25 steps per bounce period are not possible. At 50 steps per bounce period, right from the beginning of the simulation, the trapezoidal integrator shows slight deviations of the orbit, but the global topology of the orbit is respected. These inaccuracies are recognisable as a slight smearing of the orbit and are not observed for the midpoint integrator (see figure 4.1).

Note that figures 4.3 and 4.6 are results from a simulation of two million characteristic times, corresponding to 200 million timesteps. These results agree well with those found by Qin et al. [103], but extend them in considering an alternative discretisation that appears to lead to more accurate results.

4.4. Particle Motion in the Tokamak

In this section, we want to discretise the particle motion in full tokamak geometry (R, Z, φ, u), such that also the toroidal coordinate φ and the parallel velocity u are treated dynamically, while μ is still regarded as a parameter. The resulting scheme is expected to be somewhat simpler than the previous one as we avoid the complicated functional expression for u. Although this requires the solution of a larger system of equations describing the same dynamics, in this formulation both energy and toroidal momentum are independently conserved quantities. This allows us to test the variational integrators when the dynamics is constrained by more than one constant of motion.

As before, we introduce generalised coordinates q_k^R, q_k^Z, q_k^φ, q_k^u with discrete conjugate momenta p_k^R, p_k^Z, p_k^φ, p_k^u. For simplicity, we neglect the electrostatic field, could, however, be added with only minor complications.

4.4.1. Trapezoidal Discretisation

The discrete trapezoidal Lagrangian is

$$
L_d^{\mathrm{tr}}(q_k, q_{k+1}) = h \left[\frac{A_R^*(q_k) + A_R^*(q_{k+1})}{2} \frac{q_{k+1}^R - q_k^R}{h} + \frac{A_Z^*(q_k) + A_Z^*(q_{k+1})}{2} \frac{q_{k+1}^Z - q_k^Z}{h} \right.
$$
$$
+ \frac{q_k^R A_\varphi^*(q_k) + q_{k+1}^R A_\varphi^*(q_{k+1})}{2} \frac{q_{k+1}^\varphi - q_k^\varphi}{h}
$$
$$
\left. - \frac{q_k^u q_{k+1}^u}{2} - \mu \frac{B(q_k) + B(q_{k+1})}{2} \right]. \tag{4.47}
$$

The position-momentum form (3.14) of the trapezoidal integrator is computed as

$$
p_k^R = \frac{1}{2} \left[- A_{R,R}^*(q_k) \left[q_{k+1}^R - q_k^R \right] - A_{Z,R}^*(q_k) \left[q_{k+1}^Z - q_k^Z \right] - q_k^R A_{\varphi,R}^*(q_k) \left[q_{k+1}^\varphi - q_k^\varphi \right] \right.
$$
$$
\left. + \left[A_R^*(q_k) + A_R^*(q_{k+1}) \right] - A_\varphi^*(q_k) \left[q_{k+1}^\varphi - q_k^\varphi \right] + h\mu B_{,R}(q_k) \right], \tag{4.48a}
$$

$$
p_k^Z = \frac{1}{2} \left[- A_{R,Z}^*(q_k) \left[q_{k+1}^R - q_k^R \right] - A_{Z,Z}^*(q_k) \left[q_{k+1}^Z - q_k^Z \right] - q_k^R A_{\varphi,Z}^*(q_k) \left[q_{k+1}^\varphi - q_k^\varphi \right] \right.
$$
$$
\left. + \left[A_Z^*(q_k) + A_Z^*(q_{k+1}) \right] + h\mu B_{,Z}(q_k) \right], \tag{4.48b}
$$

$$p_k^\varphi = \frac{1}{2}\left[+ \left[q_k^R\, A_\varphi^*(q_k) + q_{k+1}^R\, A_\varphi^*(q_{k+1}) \right] \right], \tag{4.48c}$$

$$p_k^u = \frac{1}{2}\left[+ h\, q_{k+1}^u - b_R(q_k)\left[q_{k+1}^R - q_k^R \right] - b_Z(q_k)\left[q_{k+1}^Z - q_k^Z \right] - q_k^R\, b_\varphi(q_k)\left[q_{k+1}^\varphi - q_k^\varphi \right] \right], \tag{4.48d}$$

$$p_{k+1}^R = \frac{1}{2}\left[+ A_{R,R}^*(q_{k+1})\left[q_{k+1}^R - q_k^R \right] + A_{Z,R}^*(q_{k+1})\left[q_{k+1}^Z - q_k^Z \right] + q_{k+1}^R\, A_{\varphi,R}^*(q_{k+1})\left[q_{k+1}^\varphi - q_k^\varphi \right] \right.$$
$$\left. + \left[A_R^*(q_k) + A_R^*(q_{k+1}) \right] + A_\varphi^*(q_{k+1})\left[q_{k+1}^\varphi - q_k^\varphi \right] - h\,\mu B_{,R}(q_{k+1}) \right], \tag{4.48e}$$

$$p_{k+1}^Z = \frac{1}{2}\left[+ A_{R,Z}^*(q_{k+1})\left[q_{k+1}^R - q_k^R \right] + A_{Z,Z}^*(q_{k+1})\left[q_{k+1}^Z - q_k^Z \right] + q_{k+1}^R\, A_{\varphi,Z}^*(q_{k+1})\left[q_{k+1}^\varphi - q_k^\varphi \right] \right.$$
$$\left. + \left[A_Z^*(q_k) + A_Z^*(q_{k+1}) \right] - h\,\mu B_{,Z}(q_{k+1}) \right], \tag{4.48f}$$

$$p_{k+1}^\varphi = \frac{1}{2}\left[+ \left[q_k^R\, A_\varphi^*(q_k) + q_{k+1}^R\, A_\varphi^*(q_{k+1}) \right] \right], \tag{4.48g}$$

$$p_{k+1}^u = \frac{1}{2}\left[- h\, q_k^u + b_R(q_{k+1})\left[q_{k+1}^R - q_k^R \right] + b_Z(q_{k+1})\left[q_{k+1}^Z - q_k^Z \right] + q_{k+1}^R\, b_\varphi(q_{k+1})\left[q_{k+1}^\varphi - q_k^\varphi \right] \right]. \tag{4.48h}$$

As before, we employ Newton's method to solve the system. The Jacobian is now a 4×4 matrix, given by

$$\mathcal{J}\,\delta q_{k+1}^{n+1} = \frac{1}{2}\begin{pmatrix} J_{11} & J_{12} & J_{13} & J_{14} \\ J_{21} & J_{22} & J_{23} & J_{24} \\ J_{31} & J_{32} & J_{33} & J_{34} \\ J_{41} & J_{42} & J_{43} & J_{44} \end{pmatrix}\begin{pmatrix} \delta q_{k+1}^R \\ \delta q_{k+1}^Z \\ \delta q_{k+1}^\varphi \\ \delta q_{k+1}^u \end{pmatrix} = -\begin{pmatrix} F_R(q_k, q_{k+1}^n) \\ F_Z(q_k, q_{k+1}^n) \\ F_\varphi(q_k, q_{k+1}^n) \\ F_u(q_k, q_{k+1}^n) \end{pmatrix}, \tag{4.49}$$

with components listed in section 4.B.

4.4.2. Midpoint Discretisation

The discrete midpoint Lagrangian is

$$L_d^{\mathrm{mp}}(q_k, q_{k+1}) = h\left[A_R^*(q_{k+1/2})\,\frac{q_{k+1}^R - q_k^R}{h} + A_Z^*(q_{k+1/2})\,\frac{q_{k+1}^Z - q_k^Z}{h} \right.$$
$$\left. + q_{k+1/2}^R\, A_\varphi^*(q_{k+1/2})\,\frac{q_{k+1}^\varphi - q_k^\varphi}{h} - \frac{1}{2}\left(\frac{q_k^u + q_{k+1}^u}{2} \right)^2 - \mu B(q_{k+1/2}) \right]. \tag{4.50}$$

The position-momentum form (3.14) of the midpoint integrator is computed as

$$p_k^R = \frac{1}{2}\left[- A_{R,R}^*(q_{k+1/2})\left[q_{k+1}^R - q_k^R \right] - A_{Z,R}^*(q_{k+1/2})\left[q_{k+1}^Z - q_k^Z \right] - q_{k+1/2}^R\, A_{\varphi,R}^*(q_{k+1/2})\left[q_{k+1}^\varphi - q_k^\varphi \right] \right.$$
$$\left. + 2\, A_R^*(q_{k+1/2}) - A_\varphi^*(q_{k+1/2})\left[q_{k+1}^\varphi - q_k^\varphi \right] + h\,\mu B_{,R}(q_{k+1/2}) \right], \tag{4.51a}$$

$$p_k^Z = \frac{1}{2}\left[- A_{R,Z}^*(q_{k+1/2})\left[q_{k+1}^R - q_k^R \right] - A_{Z,Z}^*(q_{k+1/2})\left[q_{k+1}^Z - q_k^Z \right] - q_{k+1/2}^R\, A_{\varphi,Z}^*(q_{k+1/2})\left[q_{k+1}^\varphi - q_k^\varphi \right] \right.$$
$$\left. + 2\, A_Z^*(q_{k+1/2}) + h\,\mu B_{,Z}(q_{k+1/2}) \right], \tag{4.51b}$$

$$p_k^\varphi = \frac{1}{2}\Bigg[+ 2\, q_{k+1/2}^R\, A_\varphi^*(q_{k+1/2}) \Bigg], \tag{4.51c}$$

$$p_k^u = \frac{1}{2}\Bigg[- b_R(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] - b_Z(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] - q_{k+1/2}^R\, b_\varphi(q_{k+1/2})\Big[q_{k+1}^\varphi - q_k^\varphi\Big]$$
$$+\, h\, q_{k+1/2}^u \Bigg], \tag{4.51d}$$

$$p_{k+1}^R = \frac{1}{2}\Bigg[+ A_{R,R}^*(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] + A_{Z,R}^*(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] + q_{k+1/2}^R\, A_{\varphi,R}^*(q_{k+1/2})\Big[q_{k+1}^\varphi - q_k^\varphi\Big]$$
$$+\, 2\, A_R^*(q_{k+1/2}) + A_\varphi^*(q_{k+1/2})\Big[q_{k+1}^\varphi - q_k^\varphi\Big] - h\,\mu B_{,R}(q_{k+1/2}) \Bigg], \tag{4.51e}$$

$$p_{k+1}^Z = \frac{1}{2}\Bigg[+ A_{R,Z}^*(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] + A_{Z,Z}^*(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] + q_{k+1/2}^R\, A_{\varphi,Z}^*(q_{k+1/2})\Big[q_{k+1}^\varphi - q_k^\varphi\Big]$$
$$+\, 2\, A_Z^*(q_{k+1/2}) - h\,\mu B_{,Z}(q_{k+1/2}) \Bigg], \tag{4.51f}$$

$$p_{k+1}^\varphi = \frac{1}{2}\Bigg[+ 2\, q_{k+1/2}^R\, A_\varphi^*(q_{k+1/2}) \Bigg], \tag{4.51g}$$

$$p_{k+1}^u = \frac{1}{2}\Bigg[+ b_R(q_{k+1/2})\Big[q_{k+1}^R - q_k^R\Big] + b_Z(q_{k+1/2})\Big[q_{k+1}^Z - q_k^Z\Big] + q_{k+1/2}^R\, b_\varphi(q_{k+1/2})\Big[q_{k+1}^\varphi - q_k^\varphi\Big]$$
$$-\, h\, q_{k+1/2}^u \Bigg]. \tag{4.51h}$$

4.4.3. Discrete Noether Theorem

In the four-dimensional treatment, the interesting question is that of conservation of the toroidal momentum p_φ. As we have discussed in the last section, p_φ is a conserved quantity of the continuous system and should therefore be exactly conserved by the variational integrator.

The corresponding transformation is

$$\varphi_k^\epsilon = \varphi_k + \epsilon X_k^\varphi. \tag{4.52}$$

Both, the trapezoidal (4.47) and the midpoint (4.50) Lagrangian are invariant under this transformation, as can easily be seen. The discrete conserved momenta are

$$\frac{\partial L_d^{\mathrm{tr}}}{\partial q_{k+1}^\varphi}(q_k, q_{k+1}) \cdot X_k^\varphi = \frac{1}{2}\Big(q_k^R\, A_\varphi^*(q_k) + q_{k+1}^R\, A_\varphi^*(q_{k+1})\Big) \cdot X_k^\varphi = p_k^\varphi \cdot X_k^\varphi \tag{4.53}$$

for the trapezoidal discretisation and

$$\frac{\partial L_d^{\mathrm{mp}}}{\partial q_{k+1}^\varphi}(q_k, q_{k+1}) \cdot X_k^\varphi = q_{k+1/2}^R\, A_\varphi^*(q_{k+1/2}) \cdot X_k^\varphi = p_k^\varphi \cdot X_k^\varphi \tag{4.54}$$

for the midpoint discretisation.

4.4.4. Numerical Results

We use the same initial conditions as in the previous section, that is

$$R = R_0 + 0.05, \qquad Z = 0, \qquad \varphi = 0, \qquad \mu = 2.25 \times 10^{-6},$$

with

$$R_0 = 1, \qquad B_0 = 1, \qquad q = 2, \qquad \tau_b = 43107.$$

The initial parallel velocity u is computed by evaluating equation (4.37) for the initial values of R, Z and p^φ. The initial momenta are given below, where $p^Z = A_Z^*(q_0)$, i.e.,

$$p^R = 0, \qquad p^Z = -2.438 \times 10^{-2}, \qquad p^\varphi = -1.077 \times 10^{-3}, \qquad p^u = 0. \tag{4.55}$$

As in the two-dimensional case, we compare the two variational integrators with the explicit fourth order Runge-Kutta method with different timesteps corresponding to 25, 50 and 100 steps per bounce period. The qualitative behaviour is the same as before. With the Runge-Kutta method, the particle orbits deviates severely from its expected shape while the variational integrators find the correct result (see figures 4.7 - 4.9). The trapezoidal integrator is not stable for 25 steps per bounce period and shows slight inaccuracies for 50 steps per bounce period. The midpoint integrator is stable also for 25 steps per bounce period and appears accurate already at this large timestep.

While the variational integrators exhibit an oscillating energy error with a bounded amplitude of the oscillation, the Runge-Kutta method dissipates energy monotonically (see figures 4.10, 4.12 and 4.10). For the variational integrators, the amplitude of the error oscillation scales with the order of the scheme which is second order accurate.

The most interesting question about the four-dimensional integrators is that of toroidal momentum conservation. As expected, the variational integrators exhibit only very small errors in the toroidal momentum, close to the machine accuracy (figures 4.11, 4.13 and 4.15) while the Runge-Kutta method dissipates the toroidal momentum as it appears monotonically. While the energy error of the Runge-Kutta method seems too approach a stationary value of order one percent or smaller, depending on the timestep, the toroidal momentum is dissipated almost completely during the course of the simulations, independently from the timestep.

The absolute value of the momentum error of the variational integrators seems to be mostly determined by the residual of the Newton iteration. If the residual is too large, errors tend to accumulate, leading to an almost monotonic growths of the momentum error during the simulation. But even in that case, the toroidal momentum error is of orders $\mathcal{O}(10^{-6})...\mathcal{O}(10^{-3})$ for the simulation times considered here, and thus much smaller than with the Runge-Kutta method.
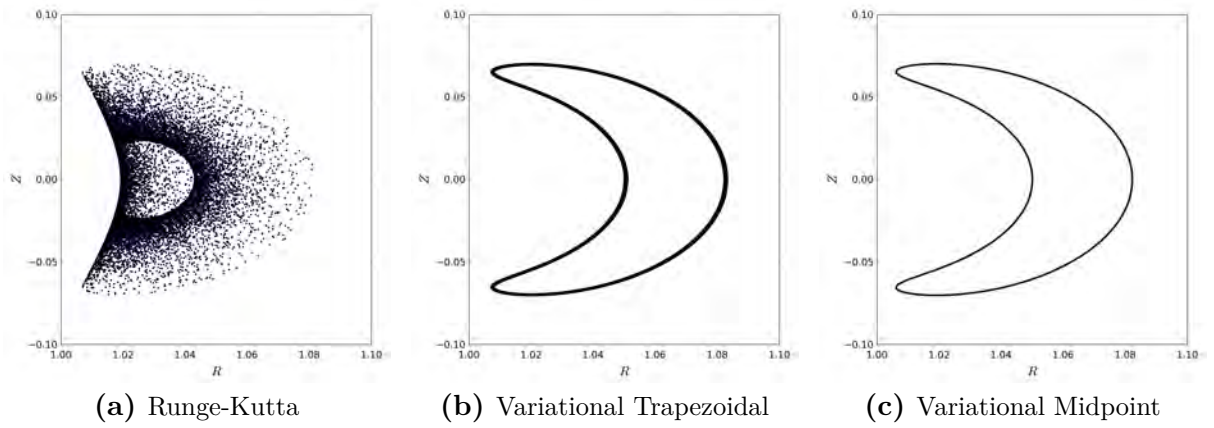
4.5. Variational PIC Scheme

The reason for deriving variational integrators for guiding centre dynamics is of course not to compute particle orbits in the poloidal plane but the aim of finding better integration techniques for large scale particle-in-cell codes as they are used in the simulation of plasma turbulence. In such codes, the electromagnetic fields are computed on a fixed grid, while the particles move in a mesh-free space. To be self-consistent, the electromagnetic fields have to be computed given the particle positions.

As the field dynamics can also be derived from an action principle, it is possible to combine the particle Lagrangian and the field Lagrangian, together with an interaction term, to get a Lagrangian description of the full system. This can then be used to derive a variational integrator for the complete system of particles and fields, thereby not only conserving the symplectic form of each particle, but the multisymplectic form of the full system as well.
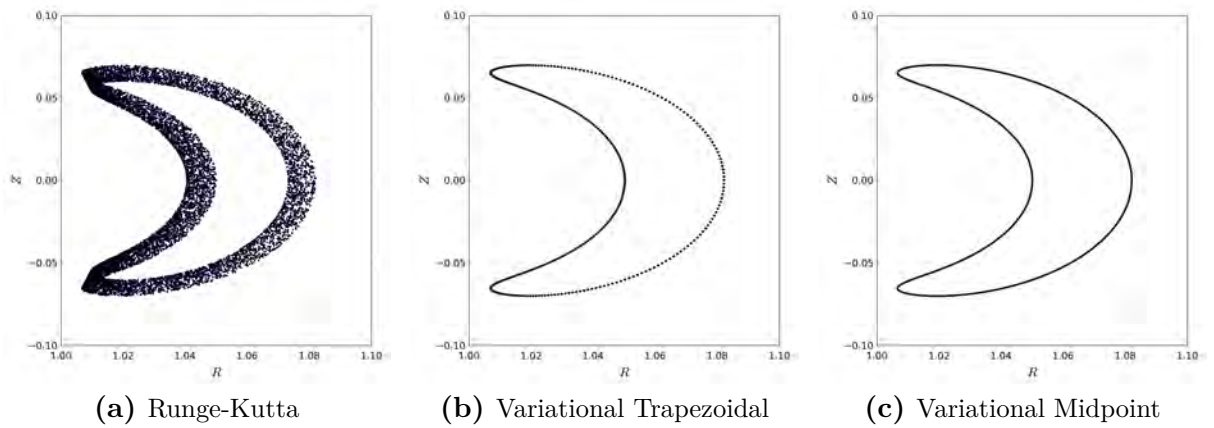
(a) Runge-Kutta (b) Variational Midpoint
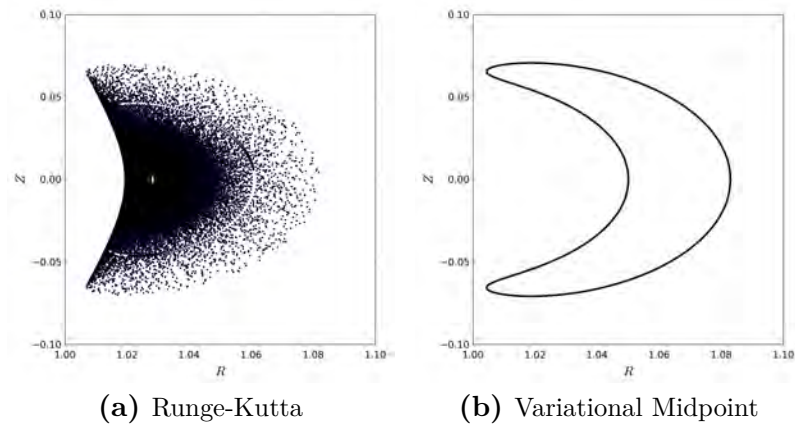
Fig. 4.7.: Trapped particle after 50.000 bounce periods with 25 steps per bounce period. The trapezoidal integrator is not stable for the timestep considered in this example.

(a) Runge-Kutta (b) Variational Trapezoidal (c) Variational Midpoint

Fig. 4.8.: Trapped particle after 200.000 bounce periods with 50 steps per bounce period.

(a) Runge-Kutta (b) Variational Trapezoidal (c) Variational Midpoint

Fig. 4.9.: Trapped particle after 250.000 bounce periods with 100 steps per bounce period.
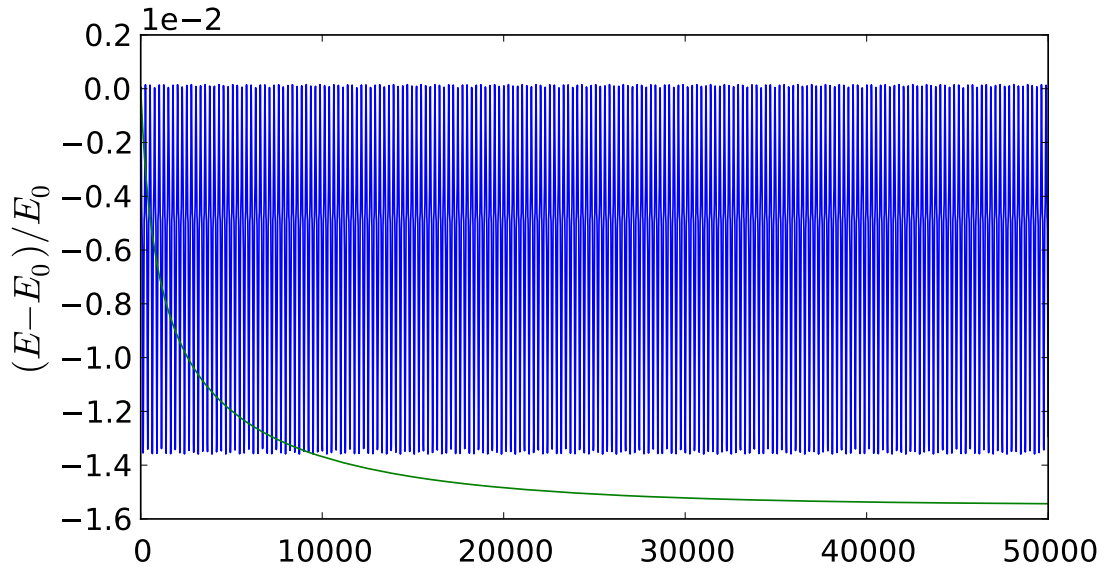
Fig. 4.10.: Energy error for trapped particle with 25 steps per bounce period (green: Runge-Kutta, blue: variational midpoint).
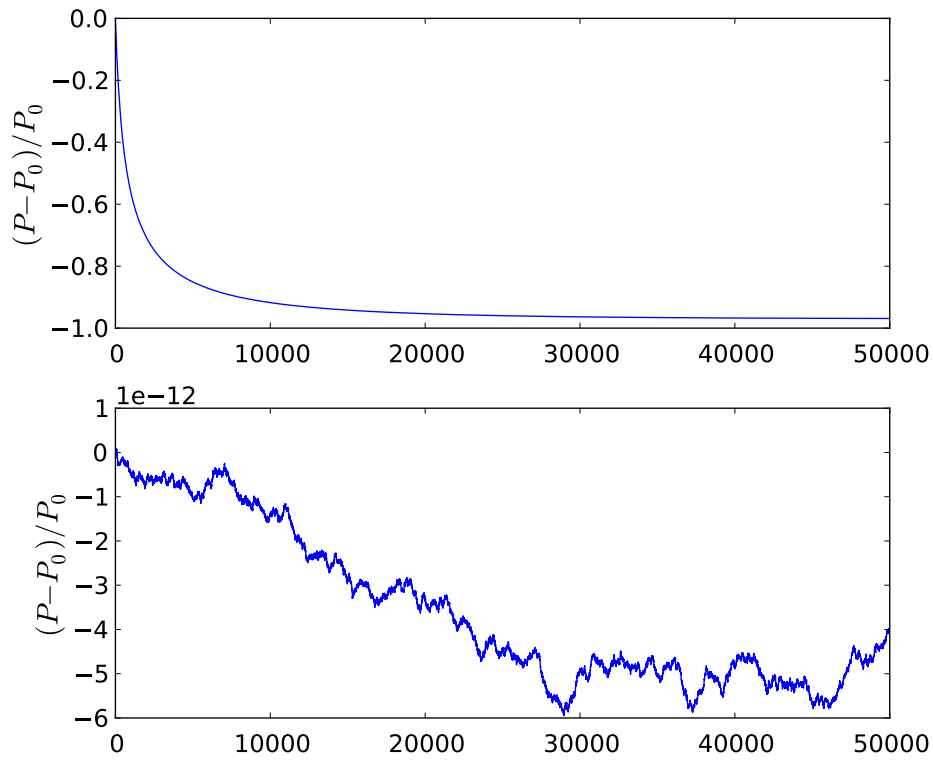
Fig. 4.11.: Total linear momentum for a trapped particle with 25 steps per bounce period (top: Runge-Kutta, bottom: variational midpoint).
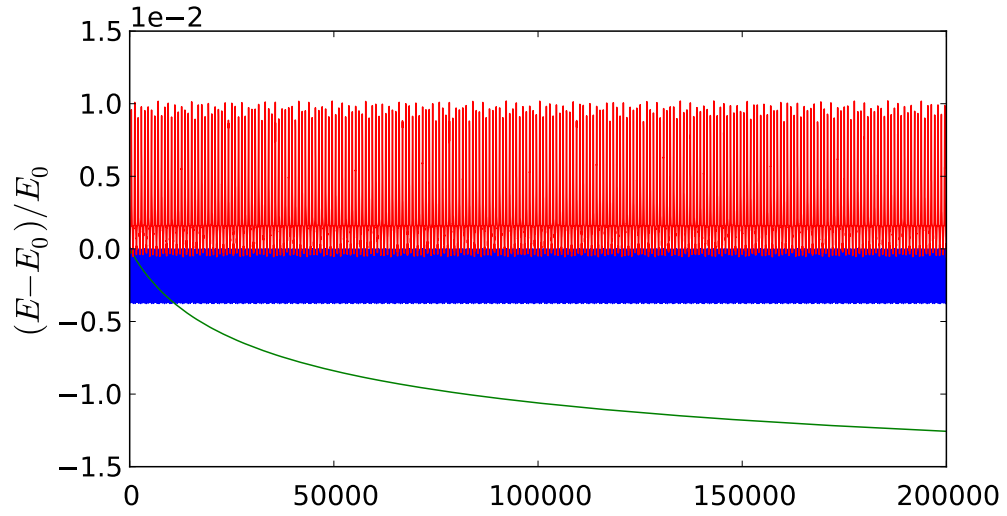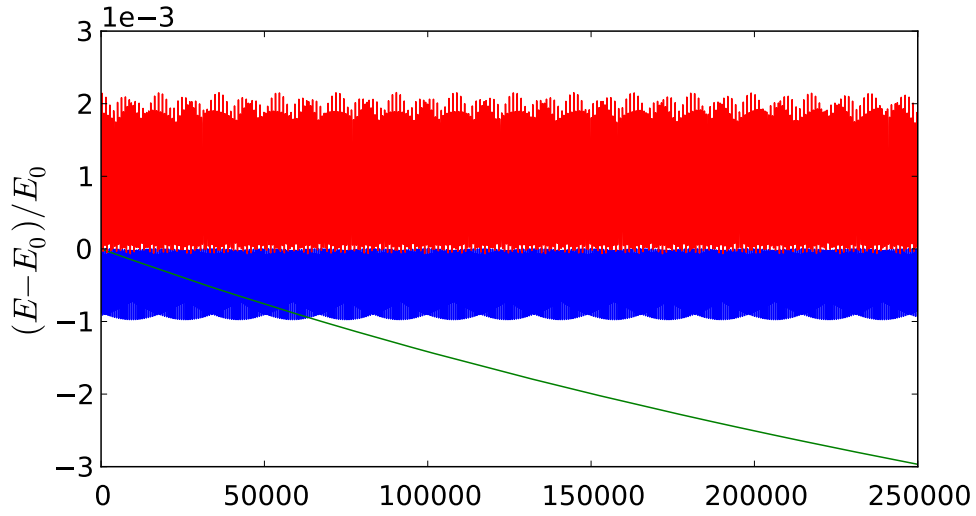
Fig. 4.12.: Energy error for trapped particle with 50 timesteps per bounce period (green: Runge-Kutta, blue: variational trapezoidal, red: variational midpoint).
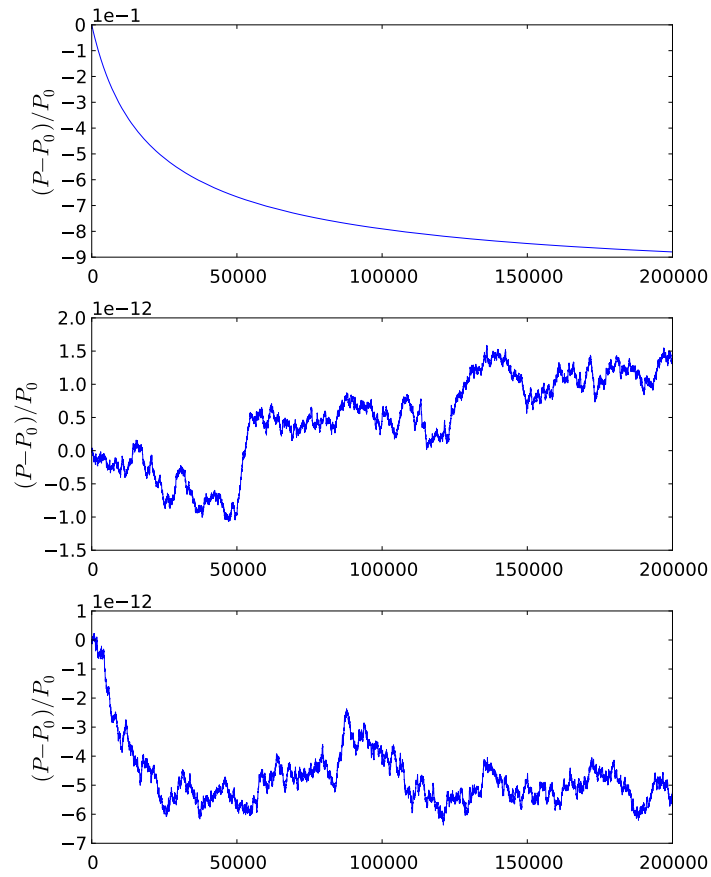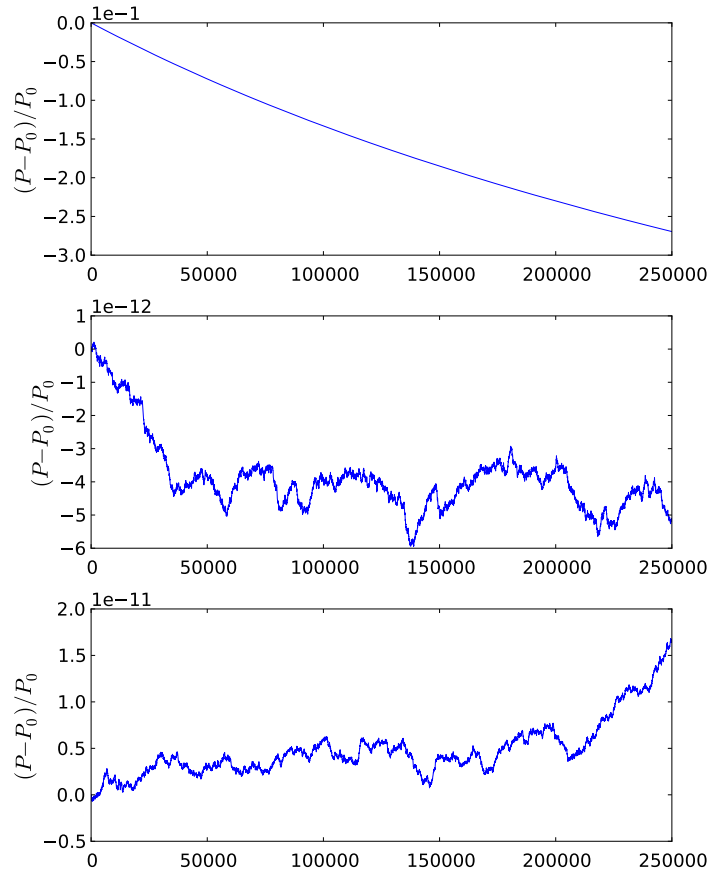

Fig. 4.13.: Total linear momentum for a trapped particle with 50 timesteps per bounce period (top: Runge-Kutta, middle: variational trapezoidal, bottom: variational midpoint).

Fig. 4.14.: Energy error for trapped particle with 100 timesteps per bounce period (green: Runge-Kutta, blue: variational trapezoidal, red: variational midpoint).

Fig. 4.15.: Total linear momentum for a trapped particle with 100 timesteps per bounce period (top: Runge-Kutta, middle: variational trapezoidal, bottom: variational midpoint).

A similar idea has recently been presented by Squire et al. [121]. In that work, however, the electromagnetic fields are represented by discrete differential forms, which is the geometrically correct approach, but not necessary if the electrostatic potential alone is considered. In the following we sketch the derivation of a variational PIC scheme based on the particle integrators from this chapter and a simple discretisation of the electrostatic potential. We do not compute the actual Euler-Lagrange equations as the aim of this section is merely to show the potential of this formulation and outline possible future directions of research.

4.5.1. Total Lagrangian and Euler-Lagrange Equations

If we restrict ourselves to the electrostatic case, the action reads

$$\mathcal{A} = \sum_p \int \dot{x}_p^2 \, dt + \frac{1}{2} \int \left(\frac{\partial \phi}{\partial x} \right)^2 dt \, dx - \int \rho(t, x) \, \phi(t, x) \, dt \, dx \tag{4.56}$$

where ρ is the charge density

$$\rho(t, x) = \sum_s q_s \int f_s(x, v, t) \, dv. \tag{4.57}$$

The distribution function of the species s is computed as the sum of the distribution functions of each particle of that species

$$f_s(t, x, v) = \sum_p f_p(t, x, v), \tag{4.58}$$

and the particle distribution function f_p is determined by

$$f_p(t, x, v) = N_p \, S_x \big(x - x_p(t) \big) \, S_v \big(v - v_p(t) \big), \tag{4.59}$$

where S_x and S_v are the shape functions of the particle in space and velocity. The simplest choice for S_v are just δ functions

$$S_v \big(v - v_p \big) = \delta \big(v^x - v_p^x \big) \, \delta \big(v^y - v_p^y \big) \, \delta \big(v^z - v_p^z \big). \tag{4.60}$$

Smoother shape functions are often used for S_x, e.g., given by B-splines of order l,

$$S_x \big(x - x_p \big) = \frac{1}{\Delta x_p \, \Delta y_p \, \Delta z_p} \, b_l \left(\frac{x - x_p}{\Delta x_p} \right) b_l \left(\frac{y - y_p}{\Delta y_p} \right) b_l \left(\frac{z - z_p}{\Delta z_p} \right), \tag{4.61}$$

where the b_l are recursively defined as

$$b_l(\xi) = \int_{-\infty}^{+\infty} d\xi' \, b_0(\xi - \xi') \, b_{l-1}(\xi') \tag{4.62a}$$

$$b_0(\xi) = \begin{cases} 1 & \text{if } |\xi| < \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{4.62b}$$

Using these expressions, the action (4.56) becomes

$$\mathcal{A} = \sum_p \int \dot{x}_p^2 \, dt + \frac{1}{2} \int \left(\frac{\partial \phi}{\partial x} \right)^2 dt \, dx - \sum_p q_p \int S_x \big(x - x_p(t) \big) \, \varphi(t, x) \, dt \, dx. \tag{4.63}$$

Computing the variation of the action with respect to x_p gives the equations of motion for the particle

$$\frac{\delta \mathcal{A}}{\delta x_p} = \ddot{x}_p + q_p \int \frac{\partial S_x}{\partial x} (x - x_p) \, \varphi(t, x) \, dx = 0, \tag{4.64}$$

and the variation with respect to ϕ gives the Poisson equation

$$\frac{\partial \mathcal{A}}{\partial \phi} = -\Delta \phi(t, x) - \sum_p q_p \, S_x(x - x_p). \tag{4.65}$$

This approach can now be used to obtain a fully variational discretisation of the system consisting of particles and fields.

4.5.2. Variational Integrator

The action (4.63) is discretised by a midpoint rule according to section (3.1) and (3.2) as

$$\mathcal{A}_d = \sum_k \left[\frac{h_t}{2} \left(\frac{x_p^{k+1} - x_p^k}{h_t} \right)^2 + \frac{h_t}{2} \sum_i \left(\frac{\phi^{i+1,k} - \phi^{i,k}}{2h_x} + \frac{\phi^{i+1,k+1} - \phi^{i,k+1}}{2h_x} \right) \tag{4.66}$$

$$- \frac{h_t}{16} \sum_i \left(\rho^{i,k} + \rho^{i+1,k} + \rho^{i+1,k+1} + \rho^{i,k+1} \right) \left(\phi^{i,k} + \phi^{i+1,k} + \phi^{i+1,k+1} + \phi^{i,k+1} \right) \right] \tag{4.67}$$

where the discrete expression of ρ, i.e., the quadrature rule to compute (4.57), depends on the order of the splines (4.62) that are used. For linear B-splines, a trapezoidal rule suffices, for quadratic B-splines, the Simpson rule should be used, and for cubic B-splines, Gauss' quadrature rule should be employed. The splines are integrated exactly by the corresponding quadrature rules. The integration domain $V^{i,j}$ for $\rho^{i,k}$ in a two-dimensional setting is selected as depicted below.
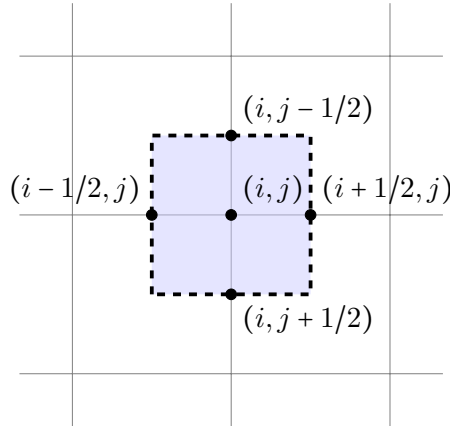
So in one spatial dimension, $V_i = [x^{i-1/2}, x^{i+1/2}]$, and the charge density becomes

$$\rho^{i,k} = \sum_p \frac{q_p}{V^i} \int_{V^i} S_x(x - x_p^k) \, dx. \tag{4.68}$$

Setting $\hat{x} = x - x_i$ gives

$$\rho^{i,k} = \sum_p \frac{q_p}{h_x} \int_{-h_x/2}^{+h_x/2} S_x(x_i - x_p^k + \hat{x}) \, d\hat{x}. \tag{4.69}$$

For some of the quadrature rules the integration domain has to be $(-1, +1)$, therefore we introduce another transformation $\tilde{x} = 2\hat{x}/h_x$ with $d\hat{x} = (h_x/2)\,d\tilde{x}$, so that

$$\rho^{i,k} = \sum_p \frac{q_p}{h_x} \int_{-1}^{+1} S_x(x_i - x_p^k + h_x\,\tilde{x}/2)\,d\tilde{x}. \tag{4.70}$$

We can now write the fully discrete expressions for the charge density, according to the different quadrature rules

$$\rho^{i,k}\big|_{\text{Trapezoidal}} = \frac{1}{2}\sum_p q_p\left[S_x\left(x_i - x_p^k - \frac{h_x}{2}\right) + S_x\left(x_i - x_p^k + \frac{h_x}{2}\right)\right], \tag{4.71a}$$

$$\rho^{i,k}\big|_{\text{Simpson}} = \frac{1}{6}\sum_p q_p\left[S_x\left(x_i - x_p^k - \frac{h_x}{2}\right) + 4\,S_x\left(x_i - x_p^k\right) + S_x\left(x_i - x_p^k + \frac{h_x}{2}\right)\right], \tag{4.71b}$$

$$\rho^{i,k}\big|_{\text{Gauss}} = \frac{1}{2}\sum_p q_p\left[S_x\left(x_i - x_p - \frac{h_x}{2\sqrt{3}}\right) + S_x\left(x_i - x_p + \frac{h_x}{2\sqrt{3}}\right)\right]. \tag{4.71c}$$

Now we have all the necessary ingredients for the derivation of a fully variational PIC scheme. The last step is to decide on the order of the B-splines and to actually compute the discrete Euler-Lagrange equations. This, however, is left for future research.

4.A. Calculation of Transit and Bounce Times

In the calculation of the bounce time, we follow Brizard [21]. We merely collect the necessary formulae. For details on the actual derivation we refer to Brizard's original work. We assume a circular tokamak with large aspect ratio, such that

$$\epsilon \equiv \frac{r}{R} \ll 1 \qquad\qquad \text{(inverse aspect ratio)}. \tag{4.72}$$

For a passing particle, the transit frequency is given by

$$\omega_t = \frac{\pi\,\omega_\parallel\,\sqrt{\kappa}}{\mathsf{K}(\kappa^{-1})} \tag{4.73}$$

while for a trapped particle, the bounce frequency is

$$\omega_b = \frac{\pi\,\omega_\parallel}{2\,\mathsf{K}(\kappa)}. \tag{4.74}$$

Here, $\omega_\parallel$ is the characteristic parallel frequency, defined as

$$\omega_\parallel = \frac{1}{qR}\sqrt{\epsilon\,\frac{\mu B_0}{m}}, \tag{4.75}$$

κ is the bounce-transit parameter, defined as

$$\kappa \equiv \frac{E - \mu B_0(1 - \epsilon)}{2\epsilon\,\mu B_0}, \tag{4.76}$$

and K is the complete elliptic integral of first kind. Furthermore, m is the mass of the particle, $E = \frac{1}{2}u^2 + \mu B$ its energy, μ the magnetic moment, and B the magnetic field. The transit and bounce times are accordingly computed as

$$\tau_t = \frac{2\pi}{\omega_t} = \frac{2\,\mathsf{K}(\kappa^{-1})}{\omega_\parallel\,\sqrt{\kappa}}, \qquad\qquad \tau_b = \frac{2\pi}{\omega_b} = \frac{4\,\mathsf{K}(\kappa)}{\omega_\parallel}. \tag{4.77}$$

For the trapped particle from the examples in this chapter, the bounce time is computed to be $\tau_b = 43107$.

4.B. Jacobians

2D Trapezoidal Method

The Jacobian is defined as

$$\mathcal{J}^{\text{2DTR}} = \frac{1}{2}\begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}, \tag{4.78}$$

with components

$$J_{11} = A^*_{R,R}(q_k) - A^*_{R,R}(q^n_{k+1}) - \tfrac{1}{2}\,h\,u_{,R}(q_k)\,u_{,R}(q^n_{k+1}), \tag{4.79a}$$

$$J_{12} = A^*_{Z,R}(q_k) - A^*_{R,Z}(q^n_{k+1}) - \tfrac{1}{2}\,h\,u_{,R}(q_k)\,u_{,Z}(q^n_{k+1}), \tag{4.79b}$$

$$J_{21} = A^*_{R,Z}(q_k) - A^*_{Z,R}(q^n_{k+1}) - \tfrac{1}{2}\,h\,u_{,Z}(q_k)\,u_{,R}(q^n_{k+1}), \tag{4.79c}$$

$$J_{22} = A^*_{Z,Z}(q_k) - A^*_{Z,Z}(q^n_{k+1}) - \tfrac{1}{2}\,h\,u_{,Z}(q_k)\,u_{,Z}(q^n_{k+1}). \tag{4.79d}$$

2D Midpoint Method

The Jacobian is defined as

$$\mathcal{J}^{\text{2DMP}} = \frac{1}{4}\begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}, \tag{4.80}$$

with components

$$J_{11} = 2\left[A^*_{R,R}(q^n_{k+1/2}) - A^*_{R,R}(q^n_{k+1/2})\right] + A^*_{R,RR}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right] + A^*_{Z,RR}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right]$$
$$- h\,u_{,R}(q_{k+1/2})\,u_{,R}(q_{k+1/2}) - h\,u(q_{k+1/2})\,u_{,RR}(q_{k+1/2}) - h\,\mu B_{,RR}(q_{k+1/2}), \tag{4.81a}$$

$$J_{12} = 2\left[A^*_{Z,R}(q^n_{k+1/2}) - A^*_{R,Z}(q^n_{k+1/2})\right]$$
$$+ A^*_{R,RZ}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right] + A^*_{Z,RZ}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right]$$
$$- h\,u_{,R}(q_{k+1/2})\,u_{,Z}(q_{k+1/2}) - h\,u(q_{k+1/2})\,u_{,RZ}(q_{k+1/2}) - h\,\mu B_{,RZ}(q_{k+1/2}), \tag{4.81b}$$

$$J_{21} = 2\left[A^*_{R,Z}(q_{k+1/2}) - A^*_{Z,R}(q_{k+1/2})\right]$$
$$+ A^*_{R,RZ}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right] + A^*_{Z,RZ}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right]$$
$$- h\,u_{,R}(q_{k+1/2})\,u_{,Z}(q_{k+1/2}) - h\,u(q_{k+1/2})\,u_{,RZ}(q_{k+1/2}) - h\,\mu B_{,RZ}(q_{k+1/2}), \tag{4.81c}$$

$$J_{22} = 2\left[A^*_{Z,Z}(q_{k+1/2}) - A^*_{Z,Z}(q_{k+1/2})\right] + A^*_{R,ZZ}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right] + A^*_{Z,ZZ}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right]$$
$$- h\,u_{,Z}(q_{k+1/2})\,u_{,Z}(q_{k+1/2}) - h\,u(q_{k+1/2})\,u_{,ZZ}(q_{k+1/2}) - h\,\mu B_{,ZZ}(q_{k+1/2}). \tag{4.81d}$$

4D Trapezoidal Method

The Jacobian is defined as

$$\mathcal{J}^{\text{4DTR}} = \frac{1}{2}\begin{pmatrix} J_{11} & J_{12} & J_{13} & J_{14} \\ J_{21} & J_{22} & J_{23} & J_{24} \\ J_{31} & J_{32} & J_{33} & J_{34} \\ J_{41} & J_{42} & J_{43} & J_{44} \end{pmatrix}, \tag{4.82}$$

with components

$$J_{11} = A^*_{R,R}(q_k) - A^*_{R,R}(q^n_{k+1}), \tag{4.83a}$$

$$J_{12} = A^*_{Z,R}(q_k) - A^*_{R,Z}(q^n_{k+1}), \tag{4.83b}$$

$$J_{13} = q^R_k A^*_{\varphi,R}(q_k) + A^*_\varphi(q_k), \tag{4.83c}$$

$$J_{14} = -b_R(q^n_{k+1}), \tag{4.83d}$$

$$J_{21} = A^*_{R,Z}(q_k) - A^*_{Z,R}(q^n_{k+1}), \tag{4.83e}$$

$$J_{22} = A^*_{Z,Z}(q_k) - A^*_{Z,Z}(q^n_{k+1}), \tag{4.83f}$$

$$J_{23} = q^R_k A^*_{\varphi,Z}(q_k), \tag{4.83g}$$

$$J_{24} = -b_Z(q^n_{k+1}), \tag{4.83h}$$

$$J_{31} = -q^R_{k+1} A^*_{\varphi,R}(q^n_{k+1}) - A^*_\varphi(q^n_{k+1}), \tag{4.83i}$$

$$J_{32} = -q^R_{k+1} A^*_{\varphi,Z}(q^n_{k+1}), \tag{4.83j}$$

$$J_{33} = 0, \tag{4.83k}$$

$$J_{34} = -q^R_{k+1} b_\varphi(q^n_{k+1}), \tag{4.83l}$$

$$J_{41} = b_R(q_k), \tag{4.83m}$$

$$J_{42} = b_Z(q_k), \tag{4.83n}$$

$$J_{43} = q^R_k b_\varphi(q_k), \tag{4.83o}$$

$$J_{44} = -\tfrac{1}{2} h. \tag{4.83p}$$

4D Midpoint Method

The Jacobian is defined as

$$\mathcal{J}^{\text{4DMP}} = \frac{1}{4} \begin{pmatrix} J_{11} & J_{12} & J_{13} & J_{14} \\ J_{21} & J_{22} & J_{23} & J_{24} \\ J_{31} & J_{32} & J_{33} & J_{34} \\ J_{41} & J_{42} & J_{43} & J_{44} \end{pmatrix}, \tag{4.84}$$

with components

$$
\begin{aligned}
J_{11} = {}& 2\left[A^*_{R,R}(q_{k+1/2}) - A^*_{R,R}(q_{k+1/2}) \right] + 2\, A^*_{\varphi,R}(q_{k+1/2}) \left[q^\varphi_{k+1} - q^\varphi_k \right] \\
& + A^*_{R,RR}(q_{k+1/2}) \left[q^R_{k+1} - q^R_k \right] + A^*_{Z,RR}(q_{k+1/2}) \left[q^Z_{k+1} - q^Z_k \right] \\
& + q^R_{k+1/2}\, A^*_{\varphi,RR}(q_{k+1/2}) \left[q^\varphi_{k+1} - q^\varphi_k \right] - h\,\mu B_{,RR}(q_{k+1/2}),
\end{aligned} \tag{4.85a}
$$

$$
\begin{aligned}
J_{12} = {}& 2\left[A^*_{Z,R}(q_{k+1/2}) - A^*_{R,Z}(q_{k+1/2}) \right] + A^*_{\varphi,Z}(q_{k+1/2}) \left[q^\varphi_{k+1} - q^\varphi_k \right] \\
& + A^*_{R,RZ}(q_{k+1/2}) \left[q^R_{k+1} - q^R_k \right] + A^*_{Z,RZ}(q_{k+1/2}) \left[q^Z_{k+1} - q^Z_k \right] \\
& + q^R_{k+1/2}\, A^*_{\varphi,RZ}(q_{k+1/2}) \left[q^\varphi_{k+1} - q^\varphi_k \right] - h\,\mu B_{,RZ}(q_{k+1/2}),
\end{aligned} \tag{4.85b}
$$

$$J_{13} = 2\, q^R_{k+1/2}\, A^*_{\varphi,R}(q_{k+1/2}) + 2\, A^*_\varphi(q_k), \tag{4.85c}$$

$$
\begin{aligned}
J_{14} = {}& b_{R,R}(q_{k+1/2}) \left[q^R_{k+1} - q^R_k \right] + b_{Z,R}(q_{k+1/2}) \left[q^Z_{k+1} - q^Z_k \right] + q^R_{k+1/2}\, b_{\varphi,R}(q_{k+1/2}) \left[q^\varphi_{k+1} - q^\varphi_k \right] \\
& + b_\varphi(q_{k+1/2}) \left[q^\varphi_{k+1} - q^\varphi_k \right] - 2\, b_R(q_{k+1/2}),
\end{aligned} \tag{4.85d}
$$

$$J_{21} = 2\left[A^*_{R,Z}(q_{k+1/2}) - A^*_{Z,R}(q_{k+1/2})\right] + A^*_{\varphi,Z}(q_{k+1/2})\left[q^\varphi_{k+1} - q^\varphi_k\right]$$
$$+ A^*_{R,RZ}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right] + A^*_{Z,RZ}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right]$$
$$+ q^R_{k+1/2} A^*_{\varphi,RZ}(q_{k+1/2})\left[q^\varphi_{k+1} - q^\varphi_k\right] - h\,\mu B_{,RZ}(q_{k+1/2}), \tag{4.85e}$$

$$J_{22} = 2\left[A^*_{Z,Z}(q_{k+1/2}) - A^*_{Z,Z}(q_{k+1/2})\right]$$
$$+ A^*_{R,ZZ}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right] + A^*_{Z,ZZ}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right]$$
$$+ q^R_{k+1/2} A^*_{\varphi,ZZ}(q_{k+1/2})\left[q^\varphi_{k+1} - q^\varphi_k\right] - h\,\mu B_{,ZZ}(q_{k+1/2}), \tag{4.85f}$$

$$J_{23} = 2\,q^R_{k+1/2}\, A^*_{\varphi,Z}(q_{k+1/2}), \tag{4.85g}$$

$$J_{24} = b_{R,Z}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right] + b_{Z,Z}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right] + q^R_{k+1/2}\, b_{\varphi,Z}(q_{k+1/2})\left[q^\varphi_{k+1} - q^\varphi_k\right]$$
$$- 2\,b_Z(q_{k+1/2}), \tag{4.85h}$$

$$J_{31} = -2\,q^R_{k+1/2}\, A^*_{\varphi,R}(q_{k+1/2}) - 2\,A^*_\varphi(q_{k+1/2}), \tag{4.85i}$$

$$J_{32} = -2\,q^R_{k+1/2}\, A^*_{\varphi,Z}(q_{k+1/2}), \tag{4.85j}$$

$$J_{33} = 0, \tag{4.85k}$$

$$J_{34} = -2\,q^R_{k+1/2}\, b_\varphi(q_{k+1/2}), \tag{4.85l}$$

$$J_{41} = 2\,b_R(q_{k+1/2}) + b_\varphi(q_{k+1/2})\left[q^\varphi_{k+1} - q^\varphi_k\right] + b_{R,R}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right]$$
$$+ b_{Z,R}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right] + q^R_{k+1/2}\, b_{\varphi,R}(q_{k+1/2})\left[q^\varphi_{k+1} - q^\varphi_k\right], \tag{4.85m}$$

$$J_{42} = 2\,b_Z(q_{k+1/2}) + b_{R,Z}(q_{k+1/2})\left[q^R_{k+1} - q^R_k\right]$$
$$+ b_{Z,Z}(q_{k+1/2})\left[q^Z_{k+1} - q^Z_k\right] + q^R_{k+1/2}\, b_{\varphi,Z}(q_{k+1/2})\left[q^\varphi_{k+1} - q^\varphi_k\right], \tag{4.85n}$$

$$J_{43} = 2\,q^R_{k+1/2}\, b_\varphi(q_{k+1/2}), \tag{4.85o}$$

$$J_{44} = -h. \tag{4.85p}$$

4.C. Derivatives

In this section, the reader can find an overview of the derivatives of the magnetic potential and the magnetic field that appear in the various integrators for guiding centre dynamics. Of course, all derivatives with respect to φ vanish as we assume axisymmetry.

Generalised Magnetic Potential in the Poloidal Plane

$$A^*_{R,R} = A_{R,R} + u_{,R}\, b_R + u\, b_{R,R} \qquad\qquad A^*_{Z,R} = A_{Z,R} + u_{,R}\, b_Z + u\, b_{Z,R} \tag{4.86a}$$
$$A^*_{R,Z} = A_{R,Z} + u_{,Z}\, b_R + u\, b_{R,Z} \qquad\qquad A^*_{Z,Z} = A_{Z,Z} + u_{,Z}\, b_Z + u\, b_{Z,Z} \tag{4.86b}$$

$$A^*_{R,RR} = A_{R,RR} + u_{,RR}\, b_R + u_{,R}\, b_{R,R} + u_{,R}\, b_{R,R} + u\, b_{R,RR} \tag{4.87a}$$

$$A^*_{R,RZ} = A_{R,RZ} + u_{,RZ}\, b_R + u_{,R}\, b_{R,Z} + u_{,Z}\, b_{R,R} + u\, b_{R,RZ} \tag{4.87b}$$

$$A^*_{R,ZZ} = A_{R,ZZ} + u_{,ZZ}\, b_R + u_{,Z}\, b_{R,Z} + u_{,Z}\, b_{R,Z} + u\, b_{R,ZZ} \tag{4.87c}$$

$$A^*_{Z,RR} = A_{Z,RR} + u_{,RR}\, b_Z + u_{,R}\, b_{Z,R} + u_{,R}\, b_{Z,R} + u\, b_{Z,RR} \tag{4.87d}$$

$$A^*_{Z,RZ} = A_{Z,RZ} + u_{,RZ}\, b_Z + u_{,R}\, b_{Z,Z} + u_{,Z}\, b_{Z,R} + u\, b_{Z,RZ} \tag{4.87e}$$

$$A^*_{Z,ZZ} = A_{Z,ZZ} + u_{,ZZ}\, b_Z + u_{,Z}\, b_{Z,Z} + u_{,Z}\, b_{Z,Z} + u\, b_{Z,ZZ} \tag{4.87f}$$

Generalised Magnetic Potential in Tokamak Geometry

$$A^*_{R,R} = A_{R,R} + u\, b_{R,R} \qquad A^*_{Z,R} = A_{Z,R} + u\, b_{Z,R} \qquad A^*_{\varphi,R} = A_{\varphi,R} + u\, b_{\varphi,R} \tag{4.88a}$$

$$A^*_{R,Z} = A_{R,Z} + u\, b_{R,Z} \qquad A^*_{Z,Z} = A_{Z,Z} + u\, b_{Z,Z} \qquad A^*_{\varphi,Z} = A_{\varphi,Z} + u\, b_{\varphi,Z} \tag{4.88b}$$

$$A^*_{R,u} = b_R \qquad\qquad A^*_{Z,u} = b_Z \qquad\qquad A^*_{\varphi,u} = b_\varphi \tag{4.88c}$$

$$A^*_{R,RR} = A_{R,RR} + u\, b_{R,RR} \quad A^*_{Z,RR} = A_{Z,RR} + u\, b_{Z,RR} \quad A^*_{\varphi,RR} = A_{\varphi,RR} + u\, b_{\varphi,RR} \tag{4.89a}$$

$$A^*_{R,RZ} = A_{R,RZ} + u\, b_{R,RZ} \quad A^*_{Z,RZ} = A_{Z,RZ} + u\, b_{Z,RZ} \quad A^*_{\varphi,RZ} = A_{\varphi,RZ} + u\, b_{\varphi,RZ} \tag{4.89b}$$

$$A^*_{R,ZZ} = A_{R,ZZ} + u\, b_{R,ZZ} \quad A^*_{Z,ZZ} = A_{Z,ZZ} + u\, b_{Z,ZZ} \quad A^*_{\varphi,ZZ} = A_{\varphi,ZZ} + u\, b_{\varphi,ZZ} \tag{4.89c}$$

$$A^*_{R,uR} = b_{R,R} \qquad\quad A^*_{Z,uR} = b_{Z,R} \qquad\quad A^*_{\varphi,uR} = b_{\varphi,R} \tag{4.89d}$$

$$A^*_{R,uZ} = b_{R,Z} \qquad\quad A^*_{Z,uZ} = b_{Z,Z} \qquad\quad A^*_{\varphi,uZ} = b_{\varphi,Z} \tag{4.89e}$$

$$A^*_{R,uu} = 0 \qquad\qquad A^*_{Z,uu} = 0 \qquad\qquad A^*_{\varphi,uu} = 0 \tag{4.89f}$$

Parallel Velocity in the Poloidal Plane

$$u_{,R} = -\left[p_\phi + \frac{r^2 B_0}{2q}\right]\frac{B_{,R}}{B_0 R_0} - \frac{B\,(R - R_0)}{qR_0} \tag{4.90a}$$

$$u_{,Z} = -\left[p_\phi + \frac{r^2 B_0}{2q}\right]\frac{B_{,Z}}{B_0 R_0} - \frac{BZ}{qR_0} \tag{4.90b}$$

$$u_{,RR} = -\left[p_\phi + \frac{r^2 B_0}{2q}\right]\frac{B_{,RR}}{B_0 R_0} - \frac{2\,(R - R_0)\,B_{,R} + B}{qR_0} \tag{4.91a}$$

$$u_{,RZ} = -\left[p_\phi + \frac{r^2 B_0}{2q}\right]\frac{B_{,RZ}}{B_0 R_0} - \frac{(R - R_0)\,B_{,Z} + Z B_{,R}}{qR_0} \tag{4.91b}$$

$$u_{,ZZ} = -\left[p_\phi + \frac{r^2 B_0}{2q}\right]\frac{B_{,ZZ}}{B_0 R_0} - \frac{2Z B_{,Z} + B}{qR_0} \tag{4.91c}$$

Magnetic Potential

$$A_{R,R} = -\frac{B_0 R_0 Z}{2R^2} \qquad A_{Z,R} = -\frac{B_0 R_0}{2R} \qquad A_{\varphi,R} = -\frac{A_\varphi}{R} - \frac{B_0\,(R - R_0)}{qR} \tag{4.92a}$$

$$A_{R,Z} = \frac{B_0 R_0}{2R} \qquad\quad A_{Z,Z} = 0 \qquad\qquad A_{\varphi,Z} = -\frac{B_0 Z}{qR} \tag{4.92b}$$

$$A_{R,RR} = \frac{B_0 R_0 Z}{R^3} \qquad\qquad A_{Z,RR} = \frac{B_0 R_0}{2R^2} \qquad\qquad A_{\varphi,RR} = -2\frac{A_{\varphi,R}}{R} - \frac{B_0}{qR} \qquad (4.93a)$$

$$A_{R,RZ} = -\frac{B_0 R_0}{2R^2} \qquad\qquad A_{Z,RZ} = 0 \qquad\qquad A_{\varphi,RZ} = \frac{B_0 Z}{qR^2} \qquad (4.93b)$$

$$A_{R,ZZ} = 0 \qquad\qquad A_{Z,ZZ} = 0 \qquad\qquad A_{\varphi,ZZ} = -\frac{B_0}{qR} \qquad (4.93c)$$

Magnetic Field

$$B_{,R} = B\left[\frac{R - R_0}{S^2} - \frac{1}{R}\right] \qquad\qquad\qquad S_{,R} = \frac{R - R_0}{S} \qquad (4.94a)$$

$$B_{,Z} = B\frac{Z}{S^2} \qquad\qquad\qquad\qquad S_{,Z} = \frac{Z}{S} \qquad (4.94b)$$

$$B_{,RR} = B_{,R}\left[\frac{R - R_0}{S^2} - \frac{1}{R}\right] + B\left[\frac{1}{R^2} + \frac{1}{S^2} - 2\frac{(R - R_0)^2}{S^4}\right] \qquad (4.95a)$$

$$B_{,RZ} = \frac{1}{S^2}\left[Z\,B_{,R} - 2\,(R - R_0)\,B_{,Z}\right] \qquad (4.95b)$$

$$B_{,ZZ} = \frac{B}{S^2}\left[1 - \frac{Z^2}{S^2}\right] \qquad (4.95c)$$

Normalised Magnetic Field

$$b_{R,R} = -\frac{1}{S}b_R b_Z \qquad\qquad b_{Z,R} = -\frac{1}{S}\left[b_Z^2 - 1\right] \qquad\qquad b_{\varphi,R} = -\frac{1}{S}b_Z b_\varphi \qquad (4.96a)$$

$$b_{R,Z} = \ \ \frac{1}{S}\left[b_R^2 - 1\right] \qquad\qquad b_{Z,Z} = \ \ \frac{1}{S}b_R b_Z \qquad\qquad b_{\varphi,Z} = \ \ \frac{1}{S}b_R b_\varphi \qquad (4.96b)$$

$$b_{R,RR} = \ \ \frac{b_R}{S^2}\left[3\,b_Z^2 - 1\right] \qquad b_{Z,RR} = \ \ \frac{b_Z}{S^2}\left[3\,b_Z^2 - 3\right] \qquad b_{\varphi,RR} = \ \ \frac{b_\varphi}{S^2}\left[3\,b_Z^2 - 1\right] \qquad (4.97a)$$

$$b_{R,RZ} = -\frac{b_Z}{S^2}\left[3\,b_R^2 - 1\right] \qquad b_{Z,RZ} = -\frac{b_R}{S^2}\left[3\,b_Z^2 - 1\right] \qquad b_{\varphi,RZ} = -\frac{3}{S^2}b_R b_Z b_\varphi \qquad (4.97b)$$

$$b_{R,ZZ} = \ \ \frac{b_R}{S^2}\left[3\,b_R^2 - 3\right] \qquad b_{Z,ZZ} = \ \ \frac{b_Z}{S^2}\left[3\,b_R^2 - 1\right] \qquad b_{\varphi,ZZ} = \ \ \frac{b_\varphi}{S^2}\left[3\,b_R^2 - 1\right] \qquad (4.97c)$$

5. Kinetic Theory

In the kinetic theory of plasma dynamics [68, 127], the system of charged particles constituting the plasma is described by a distribution function $f(t, x, v)$ that can be seen as a phasespace density. The integral of f over some phase space region Ω

$$\int_{\Omega} f(t, x, v)\, dx\, dv. \tag{5.1}$$

gives the number of particles in that region, such that its velocity integral yields the particle density

$$n(x) = \int_{-\infty}^{+\infty} f(t, x, v)\, dv, \tag{5.2}$$

at a given point in space, and the integral over full phase space gives the total number of particles N in the system

$$N = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(t, x, v)\, dx\, dv. \tag{5.3}$$

Conservation of phasespace volume along the particle trajectories implies that the distribution function evolves according to the Liouville equation

$$\frac{\partial f}{\partial t} + \dot{x} \cdot \frac{\partial f}{\partial x} + \dot{v} \cdot \frac{\partial f}{\partial v} = 0 \tag{5.4}$$

which is a linear advection equation in phasespace. When the particle motion can be described by a canonical Hamiltonian system, this equation can be expressed as

$$\frac{\partial f}{\partial t} + [f, h]_{xp} = 0 \tag{5.5}$$

where

$$[f, h]_{xp} = \frac{\partial f}{\partial x} \frac{\partial h}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial h}{\partial x}, \tag{5.6}$$

are the canonical Poisson brackets, h is the particle Hamiltonian for the system under consideration, and p is the canonical momentum conjugate to x.

5.1. The Vlasov-Poisson and Vlasov-Maxwell Systems

Replacing the acceleration with the Lorentz force

$$\dot{v} = \frac{1}{m} F_{\text{Lorentz}} = \frac{q}{m} \left(E + \frac{1}{c} v \times B \right) \tag{5.7}$$

we obtain the Vlasov equation (also referred to as the collisionless Boltzmann equation)

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial x} + \frac{q}{m} \left(E + \frac{1}{c} v \times B \right) \cdot \frac{\partial f}{\partial v} = 0. \tag{5.8}$$

The fields E and B are computed self-consistently with respect to the particle distribution given by f.

5.1.1. The Vlasov-Maxwell System

The Vlasov-Maxwell system consists of the Vlasov equation (5.8)

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial x} + \frac{q}{m} \left(E + \frac{1}{c} \, v \times B \right) \cdot \frac{\partial f}{\partial v} = 0 \tag{5.9}$$

and the dynamical Maxwell's equations

$$\nabla \times E = -\frac{1}{c} \frac{\partial B}{\partial t}. \qquad\qquad \nabla \times B = \frac{1}{c} \left(\frac{\partial E}{\partial t} + j \right). \tag{5.10}$$

The charge density ρ and the current density j are given as moments of the distribution function

$$\rho = q \int f(t, x, v) \, dv, \qquad\qquad j = q \int v \, f(t, x, v) \, dv, \tag{5.11}$$

and E and B satisfy the constraints

$$\nabla \cdot E = \rho, \qquad\qquad \nabla \cdot B = 0. \tag{5.12}$$

The particle Hamiltonian h has the form

$$h = \frac{1}{2m} \left(p - \frac{q}{c} \, A(x) \right)^2 + q\phi(x). \tag{5.13}$$

where the canonical momentum p is given by

$$p = mv + \frac{q}{c} \, A(x), \tag{5.14}$$

A is the magnetic vector potential and ϕ is the electrostatic potential, such that the electromagnetic fields are obtained by

$$E = -\nabla\phi - \frac{\partial A}{\partial t}, \qquad\qquad B = \nabla \times A. \tag{5.15}$$

5.1.2. The Vlasov-Poisson System

In the non-relativistic case, when $v \ll c$, the $v \times B$ in the Lorentz force is very small an can be neglected. If, moreover, B exhibits only little change over time, the electric and magnetic field decouple, and we can describe the system by the electrostatic Vlasov equation

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial x} + \frac{q}{m} \, E \cdot \frac{\partial f}{\partial v} = 0 \tag{5.16}$$

The electric field E can be replaced by the electrostatic potential $E = -\nabla\phi$, such that the Vlasov-equation becomes

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial x} - \frac{q}{m} \frac{\partial \phi}{\partial x} \cdot \frac{\partial f}{\partial v} = 0. \tag{5.17}$$

The electrostatic potential ϕ is determined through the Poisson equation

$$-\Delta\phi = \rho \tag{5.18}$$

where ρ is the charge density

$$\rho = qn = q \int f \, dv. \tag{5.19}$$

With the particle Hamiltonian h, consisting of the kinetic energy of the particles and their potential energy in the electrostatic field,

$$h = \frac{1}{2m} p^2 + q\phi, \tag{5.20}$$

the Vlasov equation (5.17) can be expressed with Poisson brackets as in (5.5). In the following treatment, it will however be more practical to express the Vlasov equation with respect to velocity phasespace variables (x, v) instead of (x, p) as in the canonical Poisson brackets (5.6). In that form, the Vlasov equation reads

$$\frac{\partial f}{\partial t} + [f, h]_{xv} = 0 \tag{5.21}$$

with particle Hamiltonian

$$h = \frac{m}{2} v^2 + q\phi \tag{5.22}$$

and noncanonical Poisson brackets

$$[\cdot, \cdot]_{x,p} = \frac{1}{m} [\cdot, \cdot]_{x,v}. \tag{5.23}$$

The additional mass factor will disappear in the normalisation procedure.

Dimensionless Equations

For the numerical treatment we normalise the Vlasov-Poisson equation to a dimensionless equation. Charges are normalised to e, where the electron charge is $q_e = -e$. Masses are normalised to the electron mass m_e. The speed of light c is set to one, and the average densities are also normalised to one,

$$\frac{1}{L_x} \int f(x, v) \, dv = 1, \tag{5.24}$$

with L_x the size of the system. If both, electrons and ions (assumed to have opposite charge $\pm e$), are treated dynamically, the mass ratio must be accounted for in the Hamiltonian of the ions

$$\frac{\partial f_e}{\partial t} + [f_e, h_e] = 0, \qquad\qquad h_e = \frac{1}{2} v^2 - \phi, \tag{5.25a}$$

$$\frac{\partial f_i}{\partial t} + [f_i, h_i] = 0, \qquad\qquad h_i = \frac{1}{2} \frac{m_i}{m_e} v^2 + \phi. \tag{5.25b}$$

In that case, the Poisson equation, of course, has contributions from both species

$$-\Delta\phi = \int (f_i - f_e) \, dv. \tag{5.26}$$

If the scale of the dynamics we are interested in is very fast the ions can be considered as fixed in space and only the electrons have to be treated dynamically. The ions therefore constitute a neutralising background that has to be taken into account in the Poisson equation in terms of the ions' average charge density (which is normalised to 1). Thereby we obtain the following set of equations

$$\frac{\partial f}{\partial t} + [f, h] = 0, \qquad h = \frac{1}{2} v^2 - \phi, \qquad -\Delta\phi = 1 - \int f \, dv. \tag{5.27}$$

We will restrict the following treatment to this case.

5.1.3. Conservation Properties

The Vlasov-Poisson (and Vlasov-Maxwell) system conserves a number of quantities that should in principle also be preserved in any numerical simulation. If this is not possible, their error can give a hint of the validity of a numerical simulation and should therefore always be monitored. These conserved quantities are[1]

- positivity and maximum principle (follows from characteristics theory)

$$0 \le f(x,v,t) \le \max_{(x,v)} f_0(x,v), \tag{5.28}$$

- total linear momentum P

$$\frac{d}{dt}\left(\int v\, f(x,v,t)\, dx\, dv \right) = 0, \tag{5.29}$$

- total energy E

$$\frac{d}{dt}\left(\int v^2\, f(x,v,t)\, dx\, dv + \frac{1}{2} \int \left(\frac{\partial \phi}{\partial x} \right)^2 dx \right) = 0, \tag{5.30}$$

- any functional of the form

$$\frac{d}{dt}\left(\int F(f)\, dx\, dv \right) = 0, \tag{5.31}$$

  with the especially important cases of integral norms L_p and entropy

- integral norms L_p

$$\frac{d}{dt}\left(\int \left(f(x,v,t) \right)^p dx\, dv \right) = 0 \qquad \text{for} \qquad 1 \le p \le \infty, \tag{5.32}$$

  where $p = 1$ corresponds to the total particle number N and $p = \infty$ is the maximum norm

- entropy S

$$\frac{d}{dt}\left(\int f \ln f\, dx\, dv \right) = 0. \tag{5.33}$$

In the derivation of the variational integrators, we concentrate on preserving the total particle number, the total linear momentum, and the total energy.

---

[1] Here, we merely list the conserved quantities, for proofs see e.g. the lecture notes of Sonnendrücker [120].

5.2. Review of Action Principles

As pointed out in the chapter on classical mechanics and field theory, action principles are very powerful tools for the description of physical theories. Not only do they allow us to derive equations of motions in a general and covariant way, but they also provide a machinery for finding conserved quantities for free, namely, the Noether theorem.

The very first action principle for the Vlasov-Maxwell system is due to Low and was published in 1958 [80]:

$$
L = \int \left[\frac{m}{2} \left(\frac{\partial x}{\partial t} \right)^2 - q\phi(x,t) + q\, \frac{\partial x}{\partial t} \cdot A(x,t) \right] f(x,v)\, dx\, dv + \frac{1}{2} \int \left(E^2 - B^2 \right) dx. \quad (5.34)
$$

Since then, a plethora of different action principles for the Vlasov-Poisson and Vlasov-Maxwell systems have been proposed. These action principles can be classified by the variables they use for the particles part; there are Lagrangian descriptions [80, 126], Eulerian descriptions [136, 70, 71, 39, 19, 20], and mixed Lagrangian-Eulerian (i.e., Hamilton-Jacobi) descriptions [100, 101, 90, 102]. For the electromagnetic fields, of course, in all action principles Eulerian variables are used.

While Lagrangian action principles are suitable for the derivation of numerical schemes for particle-in-cell (PIC) codes, the natural basis for a Vlasov code is a purely Eulerian action principle. These, however. have a severe difficulty, namely that the Vlasov system is inherently noncanonical - the distribution function f does not have a canonical conjugate field variable. As we will see below, there are several possibilities to circumvent this shortcoming. Unfortunately, none of these possibilities, while being quite elegant analytically, do lend themselves to a straight forward discretisation.

What all action principles have in common is the electrostatic Lagrangian $\mathcal{L}_\phi$ or the electromagnetic Lagrangian $\mathcal{L}_{\mathrm{EM}}$, depending on the Lagrangian describing the Vlasov-Poisson or the Vlasov-Maxwell system. These Lagrangian densities are given by

$$
\mathcal{L}_\phi = \frac{1}{2} \left(\frac{\partial \phi}{\partial x} \right)^2, \qquad\qquad \mathcal{L}_{\mathrm{EM}} = \frac{1}{2} \left(E^2 + B^2 \right). \quad (5.35)
$$

5.2.1. Parametrisation of the Distribution Function

An approach well known in fluid dynamics for a long time is the use of Clebsch variables, first applied to the Vlasov system by Ye and Morrison [136]. Here the field variable, in our case the distribution function f, is parametrised as the Poisson bracket of two Clebsch potentials α and β that constitute canonically conjugate field variables:

$$
f = [\alpha, \beta]. \quad (5.36)
$$

The corresponding Lagrangian density reads

$$
\mathcal{L}_f = \alpha \frac{\partial \beta}{\partial t} - [\alpha, \beta]\, h, \quad (5.37)
$$

with the Hamiltonian density

$$
\mathcal{H}_f = f\, h = [\alpha, \beta]\, h. \quad (5.38)
$$

and h the particle Hamiltonian. The variations with respect to β and α lead to Vlasov equations for α and β, respectively

$$\frac{\partial \alpha}{\partial t} + [\alpha, h] = 0, \qquad\qquad -\frac{\partial \beta}{\partial t} - [\beta, h] = 0. \tag{5.39}$$

It is easy to show, that if α and β obey the Vlasov equation, so does f. Just insert the parametrisation of f into the Vlasov equation,

$$\frac{\partial [\alpha, \beta]}{\partial t} + \Big[[\alpha, \beta], h \Big] = 0, \tag{5.40}$$

apply the Leibniz rule to the time derivative,

$$\frac{\partial [\alpha, \beta]}{\partial t} = \left[\frac{\partial \alpha}{\partial t}, \beta \right] + \left[\alpha, \frac{\partial \beta}{\partial t} \right], \tag{5.41}$$

and the Jacobi identity to the Poisson bracket,

$$\Big[[\alpha, \beta], h \Big] + \Big[[\beta, h], \alpha \Big] + \Big[[h, \alpha], \beta \Big] = 0, \tag{5.42}$$

to get

$$\left[\frac{\partial \beta}{\partial t} + [\alpha, h], \beta \right] - \left[\frac{\partial \beta}{\partial t} + [\beta, h], \alpha \right] = 0. \tag{5.43}$$

The outer Poisson brackets on the left hand side vanish, if α and β obey the Vlasov equation.

The problems with this parametrisation are the behaviour of the Clebsch potentials at the boundaries, especially for spatially periodic boundary conditions of f, their continuity over the domain, as well as the question of how to initialise α and β for a given distribution function f.

A related approach was taken by Flå [39, 40]. He parametrises f with respect to some reference distribution function f_0 from the same symplectic leaf, i.e., a distribution function with the same number of particles and the same energy. f_0 does not refer to the initial conditions, but is in general evolving along with f. The Poisson bracket of f_0 with some generator S then gives the deformation of f_0 towards the actual distribution function f:

$$f = f_0 + [S, f_0]. \tag{5.44}$$

Here, f_0 might be an equilibrium solution (e.g. a local Maxwellian distribution) while f describes the turbulent state of the system. In this description, however, there is no limitation on the difference $\Delta f = f - f_0$, besides that both, f and f_0, have to lie on the same symplectic leaf. So this is not comparable to a δf method.

The corresponding action is somewhat more complicated, as f and f_0 generate different electromagnetic fields, which has to be taken into account. The Lagrangian for the distribution function reads

$$\mathcal{L}_f = f_0 \frac{\partial S}{\partial t} - [S, f_0] \, h - e \, f_0 \, (\phi - \phi_0), \tag{5.45}$$

and the Lagrangian for the electrostatic potentials is

$$\mathcal{L}_\phi = \frac{1}{2} \left[\left(\frac{\partial \phi}{\partial x} \right)^2 - \left(\frac{\partial \phi_0}{\partial x} \right)^2 \right]. \tag{5.46}$$

So that the action reads

$$\mathcal{A}[f_0, f, \phi_0, \phi] = \int \left[f_0 \frac{\partial S}{\partial t} - [S, f_0]\, h - e\, f_0 \left(\phi - \phi_0 \right) \right] dt\, dx\, dp$$

$$+ \frac{1}{2} \int \left[\left(\frac{\partial \phi}{\partial x} \right)^2 - \left(\frac{\partial \phi_0}{\partial x} \right)^2 \right] dt\, dx. \tag{5.47}$$

The variation with respect to S yields the Vlasov equation for f_0,

$$\frac{\partial f_0}{\partial t} + [f_0, h] = 0. \tag{5.48}$$

The variation with respect to f_0 yields a Vlasov-like equation for S,

$$\frac{\partial S}{\partial t} + [S, h] = e \left(\phi - \phi_0 \right). \tag{5.49}$$

The variations with respect to ϕ and ϕ_0 yield the corresponding Poisson equations

$$\Delta \phi \;\; = -e \int \left(f_0 + [S, f_0] \right) dp = -e \int f\, dp \tag{5.50}$$

$$\Delta \phi_0 = -e \int f_0\, dp. \tag{5.51}$$

Flå's parametrisation is subject to similar problems as the Clebsch parametrisation. It is not obvious how to find the generating function S for given distribution functions f and f_0. Besides, it is much more suggestive to prescribe f_0 and S and obtain the initial f through the parametrisation as this procedure gives a physical meaning to S that would be lost when proceeding the other way around. Unfortunately, this strategy can not be followed, unless a set of generating functions S for some standard scenarios in plasma physical simulation is found.

5.2.2. Constrained Variations

Brizard [19, 20] suggested an action for the Vlasov-Maxwell system that is defined on an eight-dimensional extended phasespace, adding time and energy to position and momentum. He uses constraint variations $\delta f = [\delta S, f]_{\mathrm{cov}}$, where δS is the infinitesimal generator of the variation and $[\cdot, \cdot]_{\mathrm{cov}}$ denote Poisson brackets in the extended phasespace

$$[f, h]_{\mathrm{cov}} = [f, h] + \frac{\partial f}{\partial \epsilon} \frac{\partial h}{\partial t} - \frac{\partial f}{\partial t} \frac{\partial h}{\partial \epsilon}. \tag{5.52}$$

The action is written as

$$\mathcal{A}_f = \int f\, H_{\mathrm{cov}}\, dt\, d\epsilon\, dx\, dp, \tag{5.53}$$

with the covariant Hamiltonian $H_{\text{cov}}(t, \epsilon; x, p) = h(t, x, p) - \epsilon$, thus

$$\mathcal{A}_f = \int f\,(h - \epsilon)\,dt\,d\epsilon\,dx\,dp. \tag{5.54}$$

The variation of $\mathcal{A}$ is computed as follows

$$\delta \mathcal{A}_f = \int \delta f\,(h - \epsilon)\,dt\,d\epsilon\,dx\,dp \tag{5.55}$$

$$= \int [\delta S, f]_{\text{cov}}\,(h - \epsilon)\,dt\,d\epsilon\,dx\,dp \tag{5.56}$$

$$= \int \delta S\,[f, (h - \epsilon)]_{\text{cov}}\,dt\,d\epsilon\,dx\,dp. \tag{5.57}$$

Introducing a space-time-split in the Poisson bracket, we get

$$\delta \mathcal{A}_f = \int \delta S \left(\frac{\partial f}{\partial \epsilon} \frac{\partial (h - \epsilon)}{\partial t} - \frac{\partial f}{\partial t} \frac{\partial (h - \epsilon)}{\partial \epsilon} + [f, h] - [f, \epsilon] \right) dt\,d\epsilon\,dx\,dp. \tag{5.58}$$

As $\partial h / \partial \epsilon = \partial h / \partial t = 0$ and of course $\partial \epsilon / \partial t = 0$ as $(t, \epsilon; x, p)$ are independent variables, we get

$$\delta \mathcal{A}_f = \int \delta S \left(\frac{\partial f}{\partial t} + [f, h] \right) dt\,dx\,dp, \tag{5.59}$$

which holds for any variation δS and thus yields the Vlasov equation, i.e.,

$$\frac{\delta \mathcal{A}_f}{\delta S} = 0 \quad \Rightarrow \quad \frac{\partial f}{\partial t} + [f, h] = 0. \tag{5.60}$$

This action principle does not fit the formalism of chapter 3 very well. While it certainly is possible to use the variational integrator framework to compute a discrete variational derivative of the action (5.53), constrained to the form $\delta f = [\delta S, f]$, it is not easy to see what properties the resulting discrete equations will have. The application of the discrete Noether theorem from section 3.2.4 does not seem straight forward. And it is not obvious how to incorporate a symmetrisation of the Poisson brackets in a natural way. We will see in section 5.3 that this is crucial to retain some of the symmetries of the continuous system on the discrete level and obtain a robust numerical scheme (see also appendix B). Besides, the use of extended phasespace adds further complications, as the additional dimensions have to be removed by restriction of the dynamics to a hyperplane of constant energy after the application of the discrete action principle.

5.2.3. Euler-Poincaré Reduction

Another action principle based on constrained variations is the one by Cendra et al. [25], which is obtained from Low's action principle (5.34) by Euler-Poincaré reduction [55, 57, 82, 56]. Albeit this formulation is not directly applicable to the variational integrator approach as well, it deserves some attention as it constitutes the most natural geometric description of the Vlasov-Maxwell system that has been found so far. The formulation becomes even more interesting as it has recently been extended to gyrokinetics by Squire et al. [122]. Furthermore, work by Pavlov et al. [99] suggests that a descendent

of the variational integrator method as it is presented in this work can be applied to this formulation (for more details see section 7.2.2).

The basic idea is to reduce the system by using the invariance of the Lagrangian under particle relabelling

$$\psi(x_0, v_0) = \big(x(x_0, v_0), v(x_0, v_0)\big). \tag{5.61}$$

ψ is the particle evolution map. It maps particles with initial phasespace position (x_0, v_0) to their current phasespace position (x, v). The distribution function is therefore given as

$$f = f_0 \circ \psi^{-1}, \tag{5.62}$$

or explicitly

$$f\big(x(x_0, v_0, t), v(x_0, v_0, t)\big) = f_0(x_0, v_0), \tag{5.63}$$

i.e., f is just carried along the particle flow. The Lagrangian can be written

$$L_{f_0}(\psi, \dot{\psi}, \phi, \dot{\phi}, A, \dot{A}) = \int f_0(x_0, v_0) \left[\left(\frac{e}{c} A(x) + mv \right) \cdot \dot{x} - \frac{m}{2} v^2 - e\phi(x) \right] dx_0 \, dv_0$$
$$+ \frac{1}{8\pi} \int \left[\left(-\nabla\phi - \frac{\partial A}{\partial t} \right)^2 - \left(\nabla \times A \right)^2 \right] dx. \tag{5.64}$$

Invariance of the Lagrangian under the particle relabelling transformation ψ means

$$L_{f_0}(\psi, \dot{\psi}, \phi, \dot{\phi}, A, \dot{A}) = L_{f_0\psi^{-1}}(\psi\psi^{-1}, \dot{\psi}\psi^{-1}, \phi, \dot{\phi}, A, \dot{A}) \equiv l(u, \dot{\psi}, \phi, \dot{\phi}, A, \dot{A}), \tag{5.65}$$

where $u = (\dot{x}, \dot{v})$ is the phasespace velocity field

$$u(x, v) \equiv \dot{\psi} \circ \psi^{-1}(x, v), \tag{5.66}$$

such that

$$l(u, \dot{\psi}, \phi, \dot{\phi}, A, \dot{A}) = \int f(x, v) \left[\left(\frac{e}{c} A(x) + mv \right) \cdot u_x - \frac{m}{2} v^2 - e\phi(x) \right] dx \, dv$$
$$+ \frac{1}{8\pi} \int \left[\left(-\nabla\phi - \frac{\partial A}{\partial t} \right)^2 - \left(\nabla \times A \right)^2 \right] dx, \tag{5.67}$$

where x and v are now considered as coordinates rather than fields and u_x is the spatial component of the phasespace velocity u.

We now have to compute the variations of l with respect to u and f. Variations with respect to ϕ and A yield Maxwell's equation in the usual way. Variations of the particle evolution map ψ lead to variations in the phasespace velocity,

$$\delta u = \frac{\partial \eta}{\partial t} + [u, \eta], \tag{5.68}$$

which have the form of Lin constraints, well know in fluid dynamics [118]. Here, $[\cdot, \cdot]$ denotes not the Poisson but the Lie bracket, i.e.,

$$[u, \eta] = (u \cdot \nabla_z)\eta - (\eta \cdot \nabla_z)u, \tag{5.69}$$

where ∇_z denotes the nabla operator in phasespace. Variations of ψ also induce variations of the distribution function,

$$\delta f = -\nabla_z \cdot (f\eta). \tag{5.70}$$

According to (5.62), the evolution of f is is determined by the phasespace advection equation,

$$\frac{\partial f}{\partial t} + (u \cdot \nabla_z) f = 0, \tag{5.71}$$

which obviously is the Vlasov equation in conservation form. Computing the variational derivative of $\int l\, dt$ with respect to δu and δf leads to the Euler-Poincaré equations,

$$\frac{\partial}{\partial t} \frac{\delta l}{\delta u} + (u \cdot \nabla_z) \frac{\delta l}{\delta u} = f \nabla_z \frac{\delta l}{\delta f}. \tag{5.72}$$

With the reduced Lagrangian (5.67) this leads to

$$u_x = v, \qquad\qquad u_v = E + \frac{1}{c} v \times B, \tag{5.73}$$

such that (5.71) takes the expected form (5.9).

5.2.4. Lie Action Principles

A number of action principles have been derived based on Lie group methods [136, 70, 71, 39]. These are not only more complicated than the actions introduced above, they also suffer from similar problems, i.e., they employ constrained variations, auxiliary variables, generating functions, etc. We do not want to go into detail here, but just mention that all of these do not seem to be applicable for our purposes.

5.3. Variational Discretisation

We have seen in the previous section that, even though a variety of action principles for the Vlasov-Poisson and Vlasov-Maxwell systems exist, none of them appears directly applicable within the variational integrator framework. We therefore have to build an extended Lagrangian as described in section 2.2.7.

5.3.1. Extended Lagrangian

To write the action for the dimensionless Vlasov-Poisson system (5.27), we need two Ibragimov multipliers, $g(t,x,v)$ for the Vlasov equation and $\psi(t,x)$ for the Poisson equation,

$$\mathcal{A}[f,g,\phi,\psi] = \int g \left(\frac{\partial f}{\partial t} + [f,h] \right) dt\, dx\, dv + \int \psi \left(\Delta\phi + 1 - \int f\, dv \right) dt\, dx \tag{5.74}$$

Computing the variations results in the following equations of motion

$$\frac{\delta \mathcal{A}}{\delta g} = +\frac{\partial f}{\partial t} + [f,h] = 0, \qquad\qquad \frac{\delta \mathcal{A}}{\delta \psi} = \Delta\phi + \int f\, dv - 1 = 0 \tag{5.75a}$$

$$\frac{\delta \mathcal{A}}{\delta f} = -\frac{\partial g}{\partial t} - [g,h] - \psi = 0, \qquad\qquad \frac{\delta \mathcal{A}}{\delta \phi} = \Delta\psi + \int [g,f]\, dv = 0 \tag{5.75b}$$

A compatible solution of the auxiliary variables is given by $g = f$ and $\psi = 0$, such that the adjoint equation for the Vlasov equation becomes the Vlasov equation itself, and the Poisson bracket in the adjoint equation of the Poisson equation is identical zero as $[f, f] = 0$, thereby admitting a constant solution for ψ, where we choose $\psi = 0$ to obtain the first equality. The solution vector of the extended system of equations is thus $(f, f, \phi, 0)$.

5.3.2. Variational Integrator

The discretisation of the action (5.74) follows exactly along the lines of section 3.2, the only difference being that we have three dimensions now, time t, space x, and velocity v (see figure 5.1).
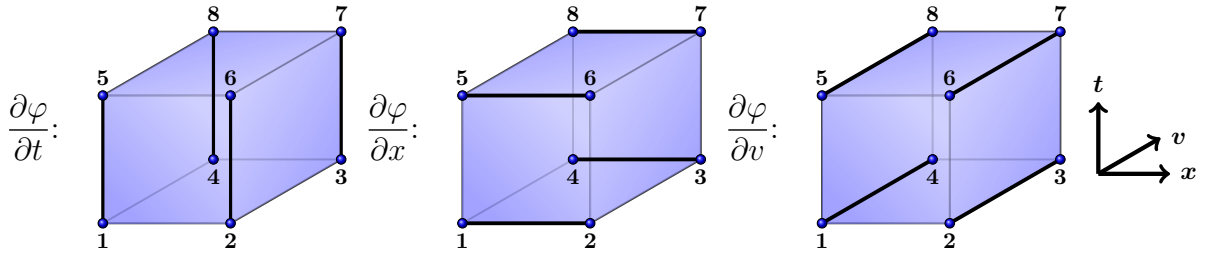
Fig. 5.1.: For a given phasespacetime grid cell, there are four possible ways of defining derivatives in the different coordinate directions (t, x, v), namely along the black lines. The actual discrete derivatives correspond to the averages over all four corresponding possibilities.

The approximations of the fields and derivatives are therefore

$$\varphi \quad \to \quad \varphi_{\bar{i},\bar{j},\bar{k}} \quad \equiv \frac{1}{8}\left(\varphi^1 + \varphi^2 + \varphi^3 + \varphi^4 + \varphi^5 + \varphi^6 + \varphi^7 + \varphi^8\right) \tag{5.76a}$$

$$\frac{\partial\varphi}{\partial t} \quad \to \quad \left(\frac{\partial\varphi}{\partial t}\right)_{\bar{i},\bar{j},k} \quad \equiv \frac{1}{4}\left(\frac{\varphi^5 - \varphi^1}{h_t} + \frac{\varphi^6 - \varphi^2}{h_t} + \frac{\varphi^7 - \varphi^3}{h_t} + \frac{\varphi^8 - \varphi^4}{h_t}\right) \tag{5.76b}$$

$$\frac{\partial\varphi}{\partial x} \quad \to \quad \left(\frac{\partial\varphi}{\partial x}\right)_{i,\bar{j},\bar{k}} \quad \equiv \frac{1}{4}\left(\frac{\varphi^2 - \varphi^1}{h_x} + \frac{\varphi^3 - \varphi^4}{h_x} + \frac{\varphi^6 - \varphi^5}{h_x} + \frac{\varphi^7 - \varphi^8}{h_x}\right) \tag{5.76c}$$

$$\frac{\partial\varphi}{\partial v} \quad \to \quad \left(\frac{\partial\varphi}{\partial v}\right)_{\bar{i},j,\bar{k}} \quad \equiv \frac{1}{4}\left(\frac{\varphi^4 - \varphi^1}{h_v} + \frac{\varphi^3 - \varphi^2}{h_v} + \frac{\varphi^8 - \varphi^5}{h_v} + \frac{\varphi^7 - \varphi^6}{h_v}\right). \tag{5.76d}$$

The bar over the indices indicates in which dimension averaging is applied as depicted in figure 5.1. Considering a given phasespacetime grid cell, the time derivative, for example, can be defined along each of the four highlighted edges of that cell. Our discrete time derivative is the average over all four of that possibilities, denoted by overbars $\bar{i}, \bar{j}$, but not $\bar{k}$ as that is the coordinate direction of the derivative. The fields themselves are averaged in all three dimensions, such that their indices have overbars $\bar{i}, \bar{j}, \bar{k}$.

In the Poisson equation, we do not have a v dimension, so we have to define a reduced field average and a reduced x derivative as follows

$$\varphi_{\bar{i},\bar{k}} \equiv \frac{1}{4}\left(\varphi^1 + \varphi^2 + \varphi^5 + \varphi^6\right), \qquad \left(\frac{\partial\varphi}{\partial x}\right)_{i,\bar{k}} \equiv \frac{1}{4}\left(\frac{\varphi^2 - \varphi^1}{h_x} + \frac{\varphi^6 - \varphi^5}{h_x}\right). \tag{5.77}$$

Some care has to be taken when discretising the Poisson bracket (see Salmon and Talley [110]). To retain the properties of the continuous bracket (antisymmetry and the Jacobi identity) at the discrete level, a symmetrisation has to be introduced in the Lagrangian. One has to realise that by partial integration the even permutations in the integrand are all identical (assuming boundary conditions such that the boundary terms of the partial integration vanish)

$$\int g\,[f,h]\,dx\,dv = \int f\,[h,g]\,dx\,dv = \int h\,[g,f]\,dx\,dv.$$

Hence, instead of one of the permutations a convex combination can be used just as well

$$\int g\,[f,h]\,dx\,dv = \int \Big(\alpha\,g\,[f,h] + \beta\,f\,[h,g] + \gamma\,h\,[g,f]\Big)\,dx\,dv \quad \text{with} \quad \alpha + \beta + \gamma = 1. \tag{5.78}$$

The symmetric case, i.e., the one that retains the properties of the bracket at the discrete level, corresponds to $\alpha = \beta = \gamma = 1/3$. We therefore write the action

$$\mathcal{A}[f,g,\phi,\psi] = \int \left[g\,\frac{\partial f}{\partial t} + \frac{1}{3}\Big(g\,[f,h] + f\,[h,g] + h\,[g,f]\Big) \right] dt\,dx\,dv$$

$$- \int \left[\frac{\partial \psi}{\partial x}\frac{\partial \phi}{\partial x} + \psi\left(\int f\,dv - 1\right) \right] dt\,dx, \quad (5.79)$$

where we also did a partial integration in the second integral to avoid second order derivatives. As the two integrals in the action have different integration domains, (t,x,v) for the Vlasov equations and (t,x) for the Poisson equation, we split the discrete Lagrangian into two parts

$$\mathcal{L}_d^V = g_{\bar{i},\bar{j},\bar{k}}\left(\frac{\partial f}{\partial t}\right)_{\bar{i},\bar{j},k} + \frac{1}{3}\Big(g_{\bar{i},\bar{j},\bar{k}}\,[f,h]_{i,j,k} + f_{\bar{i},\bar{j},\bar{k}}\,[h,g]_{i,j,k} + h_{\bar{i},\bar{j},\bar{k}}\,[g,f]_{i,j,k}\Big) \tag{5.80}$$

$$\mathcal{L}_d^P = -\left(\frac{\partial \psi}{\partial x}\right)_{i,\bar{k}}\left(\frac{\partial \phi}{\partial x}\right)_{i,\bar{k}} - \psi_{\bar{i},\bar{k}}\left(\sum_j f_{\bar{i},\bar{j},\bar{k}} - 1\right) \tag{5.81}$$

with the discrete Poisson bracket

$$[f,h]_{i,j,k} = \left(\frac{\partial f}{\partial x}\right)_{i,\bar{j},\bar{k}}\left(\frac{\partial h}{\partial v}\right)_{\bar{i},j,\bar{k}} - \left(\frac{\partial f}{\partial v}\right)_{\bar{i},j,\bar{k}}\left(\frac{\partial h}{\partial x}\right)_{i,\bar{j},\bar{k}}. \tag{5.82}$$

With these definitions the discrete action becomes

$$\mathcal{A}_d = h_t\,h_x\,h_v \sum_{i,j,k} \mathcal{L}_d^V + h_t\,h_x \sum_{i,k} \mathcal{L}_d^P \tag{5.83}$$

$$= h_t\,h_x\,h_v \sum_{i,j,k} \left[g_{\bar{i},\bar{j},\bar{k}}\left(\frac{\partial f}{\partial t}\right)_{\bar{i},\bar{j},k} + \frac{1}{3}\Big(g_{\bar{i},\bar{j},\bar{k}}\,[f,h]_{i,j,\bar{k}} + f_{\bar{i},\bar{j},\bar{k}}\,[h,g]_{i,j,\bar{k}} + h_{\bar{i},\bar{j},\bar{k}}\,[g,f]_{i,j,\bar{k}}\Big) \right]$$

$$- h_t\,h_x \sum_{i,k} \left[\left(\frac{\partial \psi}{\partial x}\right)_{i,\bar{k}}\left(\frac{\partial \phi}{\partial x}\right)_{i,\bar{k}} + \psi_{\bar{i},\bar{k}}\left(\sum_j f_{\bar{i},\bar{j},\bar{k}} - 1\right) \right]. \tag{5.84}$$

The discrete Euler-Lagrange field equations (3.93) are computed as

$$0 = \frac{\partial \mathcal{L}_d^V}{\partial g^1}\Big(y_{i,j,k}, y_{i+1,j,k}, y_{i+1,j+1,k}, y_{i,j+1,k}, y_{i,j,k+1}, y_{i+1,j,k+1}, y_{i+1,j+1,k+1}, y_{i,j+1,k+1}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^V}{\partial g^2}\Big(y_{i-1,j,k}, y_{i,j,k}, y_{i,j+1,k}, y_{i-1,j+1,k}, y_{i-1,j,k+1}, y_{i,j,k+1}, y_{i,j+1,k+1}, y_{i-1,j+1,k+1}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^V}{\partial g^3}\Big(y_{i-1,j-1,k}, y_{i,j-1,k}, y_{i,j,k}, y_{i-1,j,k}, y_{i-1,j-1,k+1}, y_{i,j-1,k+1}, y_{i,j,k+1}, y_{i-1,j,k+1}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^V}{\partial g^4}\Big(y_{i,j-1,k}, y_{i+1,j-1,k}, y_{i+1,j,k}, y_{i,j,k}, y_{i,j-1,k+1}, y_{i+1,j-1,k+1}, y_{i+1,j,k+1}, y_{i,j,k+1}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^V}{\partial g^5}\Big(y_{i,j,k-1}, y_{i+1,j,k-1}, y_{i+1,j+1,k-1}, y_{i,j+1,k}, y_{i,j,k}, y_{i+1,j,k}, y_{i+1,j+1,k}, y_{i,j+1,k}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^V}{\partial g^6}\Big(y_{i-1,j,k-1}, y_{i,j,k-1}, y_{i,j+1,k-1}, y_{i-1,j+1,k}, y_{i-1,j,k}, y_{i,j,k}, y_{i,j+1,k}, y_{i-1,j+1,k}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^V}{\partial g^7}\Big(y_{i-1,j-1,k-1}, y_{i,j-1,k-1}, y_{i,j,k-1}, y_{i-1,j,k-1}, y_{i-1,j-1,k}, y_{i,j-1,k}, y_{i,j,k}, y_{i-1,j,k}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^V}{\partial g^8}\Big(y_{i,j-1,k-1}, y_{i+1,j-1,k-1}, y_{i+1,j,k-1}, y_{i,j,k-1}, y_{i,j-1,k}, y_{i+1,j-1,k}, y_{i+1,j,k}, y_{i,j,k}\Big) \tag{5.85}$$

for the discrete Vlasov equation and

$$0 = \frac{\partial \mathcal{L}_d^P}{\partial \psi^1}\Big(y_{i,k}, y_{i+1,k}, y_{i+1,k+1}, y_{i,k+1}\Big) + \frac{\partial \mathcal{L}_d^P}{\partial \psi^2}\Big(y_{i-1,k}, y_{i,k}, y_{i,k+1}, y_{i-1,k+1}\Big)$$

$$+ \frac{\partial \mathcal{L}_d^P}{\partial \psi^3}\Big(y_{i-1,k-1}, y_{i,k-1}, y_{i,k}, y_{i-1,k}\Big) + \frac{\partial \mathcal{L}_d^P}{\partial \psi^4}\Big(y_{i,k-1}, y_{i+1,k-1}, y_{i+1,k}, y_{i,k}\Big) \tag{5.86}$$

for the discrete Poisson equation.

These discrete variations yield the following discrete Vlasov-Poisson system

$$0 = \frac{\bar{f}_{k+1} - \bar{f}_{k-1}}{2h_t} + \frac{1}{8}\Big(A(f_{k+1}, h_{k+1}) + A(f_{k+1}, h_k) + A(f_k, h_{k+1})$$

$$+ 2A(f_k, h_k) + A(f_k, h_{k-1}) + A(f_{k-1}, h_k) + A(f_{k-1}, h_{k-1})\Big) \tag{5.87}$$

$$0 = \Delta_{fd}\phi_{ik+1} - \frac{1}{4}\left(\sum_{j=0}^{n_v} f_{i-1jk+1} + 2\sum_{j=0}^{n_v} f_{ijk+1} + \sum_{j=0}^{n_v} f_{i+1jk+1}\right) + 1$$

$$+ 2\Delta_{fd}\phi_{ik} - \frac{1}{2}\left(\sum_{j=0}^{n_v} f_{i-1jk} + 2\sum_{j=0}^{n_v} f_{ijk} + \sum_{j=0}^{n_v} f_{i+1jk}\right) + 2$$

$$+ \Delta_{fd}\phi_{ik-1} - \frac{1}{4}\left(\sum_{j=0}^{n_v} f_{i-1jk-1} + 2\sum_{j=0}^{n_v} f_{ijk-1} + \sum_{j=0}^{n_v} f_{i+1jk-1}\right) + 1. \tag{5.88}$$

The time derivative of the distribution function f is an average of centred-finite-differences over 9 grid points in phasespace, weighted as depicted below.
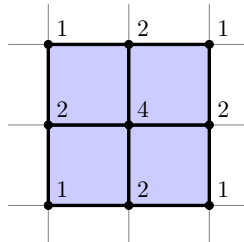
The Poisson bracket $A(\cdot, \cdot)$ is discretised by the well known Arakawa scheme [4] (see also appendix B). It is noteworthy that the Arakawa discretisation arises naturally from the variational principle. This is of course a consequence of the applied symmetrisation, but that symmetrisation is imperative to retain the symmetries of the continuous Lagrangian on the discrete level.

The discrete Laplace operator Δ_{fd} is just the standard finite difference stencil $[+1 \ -2 \ +1]/h^2$. The Poisson equation is, however, averaged over three points in time, and the charge density is averaged over three points in space.

Simplifications

Overall, we obtain a well working, nonlinearly implicit scheme. It allows, however, for some simplifications. The first issue is the representation of the time derivative by a second order discretisation. This requires the prescription of initial conditions at two consecutive points in time, which is unnatural as the Vlasov equation requires only one initial condition. Rewriting the discrete Vlasov equation (5.87) as

$$
0 = \frac{\bar{f}_{k+1} - \bar{f}_k}{2h_t} + \frac{1}{8}\Big(A(f_{k+1}, h_{k+1}) + A(f_{k+1}, h_k) + A(f_k, h_{k+1}) + A(f_k, h_k) \Big)
$$
$$
+ \frac{\bar{f}_k - \bar{f}_{k-1}}{2h_t} + \frac{1}{8}\Big(A(f_k, h_k) + A(f_k, h_{k-1}) + A(f_{k-1}, h_k) + A(f_{k-1}, h_{k-1}) \Big) \tag{5.89}
$$

the solution to this issue becomes apparent. Consider the reduced equation

$$
0 = \frac{\bar{f}_{k+1} - \bar{f}_k}{h_t} + \frac{1}{4}\Big(A(f_{k+1}, h_{k+1}) + A(f_{k+1}, h_k) + A(f_k, h_{k+1}) + A(f_k, h_k) \Big). \tag{5.90}
$$

If we use this equation to determine f_1 for given initial conditions f_0, and use this as initial conditions for (5.89), the solution of (5.89) will always also be a solution of (5.90) and vice versa. We can therefore solve the reduced system (5.90) instead of (5.89), retaining all conservation properties, but replacing the time derivative of the distribution function with a discrete derivative of first order.

By a similar argument we can remove the time average in the discrete Poisson equation (5.88). If we prescribe an initial distribution function at $k = 0$ and only use the third line of (5.88) to compute the corresponding potential, all three parts of the discrete Poisson equation have to hold separately at all times. We can thus replace the discrete Poisson equation by

$$
0 = \Delta_{fd}\phi_{ik+1} - \frac{1}{4}\left(\sum_{j=0}^{n_v} f_{i-1jk+1} + 2\sum_{j=0}^{n_v} f_{ijk+1} + \sum_{j=0}^{n_v} f_{i+1jk+1} \right) + 1. \tag{5.91}
$$

Interestingly, this reduction can already be implemented on the level of the discrete Lagrangian by modifying some of the averaging, i.e., by removing the time average from g

in the Vlasov equation and all time averages in the Poisson equation

$$
\mathcal{L}_d^V = g_{\bar{i},\bar{j},k} \left(\frac{\partial f}{\partial t} \right)_{\bar{i},\bar{j},k} + \frac{1}{3} g_{\bar{i},\bar{j},k} \left[\left(\frac{\partial f}{\partial x} \right)_{i,\bar{j},\bar{k}} \left(\frac{\partial h}{\partial v} \right)_{\bar{i},j,\bar{k}} - \left(\frac{\partial f}{\partial v} \right)_{\bar{i},j,\bar{k}} \left(\frac{\partial h}{\partial x} \right)_{i,\bar{j},\bar{k}} \right]
$$

$$
+ \frac{1}{3} f_{\bar{i},\bar{j},\bar{k}} \left[\left(\frac{\partial h}{\partial x} \right)_{i,\bar{j},\bar{k}} \left(\frac{\partial g}{\partial v} \right)_{\bar{i},j,k} - \left(\frac{\partial h}{\partial v} \right)_{\bar{i},j,\bar{k}} \left(\frac{\partial g}{\partial x} \right)_{i,\bar{j},k} \right]
$$

$$
+ \frac{1}{3} h_{\bar{i},\bar{j},\bar{k}} \left[\left(\frac{\partial g}{\partial x} \right)_{i,\bar{j},k} \left(\frac{\partial f}{\partial v} \right)_{\bar{i},j,\bar{k}} - \left(\frac{\partial g}{\partial v} \right)_{\bar{i},j,k} \left(\frac{\partial f}{\partial x} \right)_{i,\bar{j},\bar{k}} \right] \tag{5.92}
$$

$$
\mathcal{L}_d^P = - \left(\frac{\partial \psi}{\partial x} \right)_{i,k} \left(\frac{\partial \phi}{\partial x} \right)_{i,k} - \psi_{\bar{i},k} \left(\sum_j f_{\bar{i},\bar{j},k} - 1 \right). \tag{5.93}
$$

This is an important point as it allows us to study and compare the discrete symmetries and conservation laws of both the original and the reduced scheme. Without having done the analysis (which should be addressed in future work), we observe numerically that both schemes preserve the total particle number, the total linear momentum, and the total energy of the system exactly, i.e., up to machine accuracy. Furthermore, integral norms of the distribution function f are preserved, and as the scheme is symplectic by construction, the phasespace volume is preserved as well. Solely, positivity of the distribution function and the maximum principle are not preserved automatically.

5.3.3. Linearised Lagrangian

Another simplification that can be introduced on the level of the Lagrangian is a linearisation with respect to time. The above scheme has extraordinary conservation properties, but in some situations the solution of a nonlinear system of equations might be too demanding in terms of computational time and might thus not be an option. Moreover, to get quick convergence of the nonlinear iteration (e.g. in a Newton solver) a good predictor or initial guess is necessary. We therefore derive a linearised scheme that can be used in both of these cases and examine which of the properties of the fully nonlinear scheme are retained. Again, we apply a different averaging strategy, defining separate averages, space and velocity derivatives for timepoints k and $k+1$

$$
\varphi_{\bar{i},\bar{j},k} \equiv \frac{1}{4} \left(\varphi^1 + \varphi^2 + \varphi^3 + \varphi^4 \right), \qquad \varphi_{\bar{i},\bar{j},k+1} \equiv \frac{1}{4} \left(\varphi^5 + \varphi^6 + \varphi^7 + \varphi^8 \right) \tag{5.94a}
$$

$$
\left(\frac{\partial \varphi}{\partial x} \right)_{i,\bar{j},k} \equiv \frac{1}{2} \left[\frac{\varphi^2 - \varphi^1}{h_x} + \frac{\varphi^3 - \varphi^4}{h_x} \right], \qquad \left(\frac{\partial \varphi}{\partial x} \right)_{i,\bar{j},k+1} \equiv \frac{1}{2} \left[\frac{\varphi^6 - \varphi^5}{h_x} + \frac{\varphi^7 - \varphi^8}{h_x} \right] \tag{5.94b}
$$

$$
\left(\frac{\partial \varphi}{\partial v} \right)_{\bar{i},j,k} \equiv \frac{1}{2} \left[\frac{\varphi^4 - \varphi^1}{h_v} + \frac{\varphi^3 - \varphi^2}{h_v} \right], \qquad \left(\frac{\partial \varphi}{\partial v} \right)_{\bar{i},j,k+1} \equiv \frac{1}{2} \left[\frac{\varphi^8 - \varphi^5}{h_v} + \frac{\varphi^7 - \varphi^6}{h_v} \right] \tag{5.94c}
$$

and replace the symmetrised Poisson bracket by

$$\overline{\left(g\left[f,h\right]+f\left[h,g\right]+h\left[g,f\right]\right)}_{i,j,k} =$$

$$= \frac{1}{12}\bigg(g_{\bar{i},\bar{j},k}\left(\left[f_k,h_{k+1}\right]_{ij}+\left[f_{k+1},h_k\right]_{ij}\right)+g_{\bar{i},\bar{j},k+1}\left(\left[f_k,h_{k+1}\right]_{ij}+\left[f_{k+1},h_k\right]_{ij}\right)$$

$$+ f_{\bar{i},\bar{j},k}\left(\left[h_{k+1},g_k\right]_{ij}+\left[h_{k+1},g_{k+1}\right]_{ij}\right)+f_{\bar{i},\bar{j},k+1}\left(\left[h_k,g_k\right]_{ij}+\left[h_k,g_{k+1}\right]_{ij}\right)$$

$$+ h_{\bar{i},\bar{j},k}\left(\left[g_k,f_{k+1}\right]_{ij}+\left[g_{k+1},f_{k+1}\right]_{ij}\right)+h_{\bar{i},\bar{j},k+1}\left(\left[g_k,f_k\right]_{ij}+\left[g_{k+1},f_k\right]_{ij}\right)\bigg)$$

$$(5.95)$$

with

$$\left[f_k,h_{k+1}\right]_{i,j} = \left(\frac{\partial f}{\partial x}\right)_{i,\bar{j},k}\left(\frac{\partial h}{\partial v}\right)_{\bar{i},j,k+1}-\left(\frac{\partial f}{\partial v}\right)_{\bar{i},j,k}\left(\frac{\partial h}{\partial x}\right)_{i,\bar{j},k+1} \tag{5.96}$$

such that in combinations of f and h, both fields are always taken at different times. In the time derivative and in the Poisson equation we apply the simplifications from the previous section, thereby obtaining the linearised discrete Lagrangians

$$\mathcal{L}_d^V = g_{\bar{i},\bar{j},k}\left(\frac{\partial f}{\partial t}\right)_{\bar{i},\bar{j},k}+\overline{\left(g\left[f,h\right]+f\left[h,g\right]+h\left[g,f\right]\right)}_{i,j,k} \tag{5.97}$$

$$\mathcal{L}_d^P = -\left(\frac{\partial\psi}{\partial x}\right)_{ik}\left(\frac{\partial\phi}{\partial x}\right)_{ik}-\psi_{\bar{i}k}\left(\sum_j f_{\bar{i}jk}-1\right). \tag{5.98}$$

The resulting scheme is

$$0 = \frac{\bar{f}_{k+1}-\bar{f}_k}{h_t}+\frac{1}{2}\Big(A(f_{k+1},h_k)+A(f_k,h_{k+1})\Big) \tag{5.99}$$

$$0 = \Delta_{fd}\phi_{ik+1}-\frac{1}{4}\left(\sum_{j=0}^{n_v}f_{i-1jk+1}+2\sum_{j=0}^{n_v}f_{ijk+1}+\sum_{j=0}^{n_v}f_{i+1jk+1}\right)+1.$$

As we will see in the numerical examples, this scheme still preserves the total particle number, linear momentum, and integral norms of the distribution function, but it does not preserve the energy exactly. Instead the usual energy behaviour of symplectic methods is observed, i.e., the energy error oscillates about zero with a bounded amplitude of the oscillation.

The loss of exact energy conservation is almost certainly explained by destruction of some symmetry (namely the one responsible for energy conservation) in the discrete Lagrangian in the course of the linearisation procedure. Again, a detailed analysis of the discrete symmetries and discrete conservation laws should clarify this point.

5.4. Velocity Space Collision Operator

A well known problem with low order finite difference schemes like ours is the development of oscillations when phasespace filaments of the order of the grid size develop. In other discretisation techniques, e.g. finite elements or semi-Lagrangian methods, interpolation procedures are employed which, as a side effect, damp these oscillations. An alternative is to add a velocity space collision operator.

5.4.1. Continuous Collision Operator

We start by considering the Lenard-Bernstein operator [74] which conserves the total particle number but not momentum and energy

$$C_{\mathrm{LB}}[f] = \nu \frac{\partial}{\partial v}\left[\frac{\partial f}{\partial v} + vf\right].$$
(5.100)

We try to fix this by adding correction terms that restore these conservation properties and obtain the same collision operator as Filbet and Sonnendrücker [38]. The general expression of the collision operator with collision frequency ν is

$$C[f] = \nu \frac{\partial}{\partial v}\left[\frac{\partial f}{\partial v} + A(v, f)\, f\right]$$
(5.101)

where the correction term is of the form

$$A(v, f) = \sum_{n=1}^{K} A_n(f)\, v^{n-1}.$$
(5.102)

If the operator shall preserve the total particle number, linear momentum, and energy (i.e., the zeroth, first and second moment of the distribution function), the velocity integral of the collision operator, multiplied with $\{1, v, v^2\}$ has to vanish. In general, to preserve the first K moments, $C[f]$ has to fulfil

$$\int_{-v_{\max}}^{+v_{\max}} v^k\, C[f]\, dv = \nu \int_{-v_{\max}}^{+v_{\max}} v^k \frac{\partial}{\partial v}\left[\frac{\partial f}{\partial v} + A(v, f)\, f\right] dv = 0 \quad \text{for} \quad k = 0, ..., K.$$
(5.103)

The integration domain should be the whole real line, $(-\infty, +\infty)$, but in the discrete case, it suffices if the velocity domain is large enough to ensure that f and its derivatives vanish or are at least very small at the boundaries. A partial integration with respect to v (neglecting the collision frequency ν) gives

$$\left[v^k \left(\frac{\partial f}{\partial v} + A(v, f)\, f\right)\right]\Bigg|_{-v_{\max}}^{+v_{\max}} - k \int_{-v_{\max}}^{+v_{\max}} v^{k-1}\left[\frac{\partial f}{\partial v} + A(v, f)\, f\right] dv$$
(5.104)

which means that f and $\partial_v f$ have to be (close to) zero at $v \pm v_{\max}$ such that

$$\left[v^k \left(\frac{\partial f}{\partial v} + A(v, f)\, f\right)\right]\Bigg|_{-v_{\max}}^{+v_{\max}} \approx 0.$$
(5.105)

Assuming that this is fulfilled, conservation of the moments of f requires the following expression to vanish

$$\int_{-v_{\max}}^{+v_{\max}} v^{k-1}\left[\frac{\partial f}{\partial v} + A(v, f)\, f\right] dv.$$
(5.106)

Partial integration of the first term gives

$$\left[v^{k-1} f\right]\Big|_{-v_{\max}}^{+v_{\max}} - \int_{-v_{\max}}^{+v_{\max}}\left[(k-1)\, v^{k-2} - A(v, f)\, v^{k-1}\right] f\, dv$$
(5.107)

where the surface term vanishes (approximately) for $f \approx 0$ at $v \pm v_{\max}$. Plugging (5.102) into the integral and writing M_i for the ith moment of f, we get the set of conditions

$$(k-1)\, M_{k-2}(f) = \sum_{n=1}^{K} A_n(f)\, M_{k+n-2}(f) \qquad \text{for} \qquad k = 1, ..., K. \tag{5.108}$$

We want to preserve all moments up to $K = 2$, so we compute

$$k = 1 : \quad 0 = A_1 M_0 + A_2 M_1,$$
$$k = 2 : \quad M_0 = A_1 M_1 + A_2 M_2.$$

With the definition of the moments

$$M_0 = n = \int f\, dv, \qquad M_1 = nu = \int v f\, dv, \qquad M_2 = n\varepsilon = \int v^2 f\, dv \tag{5.109}$$

we get the system

$$0 = A_1 n + A_2 nu,$$
$$n = A_1 nu + A_2 n\varepsilon.$$

the solution of which determines the correction factors to be

$$A_1 = \frac{u}{u^2 - \varepsilon}, \qquad A_2 = -\frac{1}{u^2 - \varepsilon}. \tag{5.110}$$

where the factors A_i depend on the distribution function f through the momenta n, u and ϵ. The full expression of our operator is thus

$$C[f] = \nu \frac{\partial}{\partial v}\left[\frac{\partial}{\partial v} + \frac{v - u}{\varepsilon - u^2}\right] f. \tag{5.111}$$

The denominator represents the temperature $T = \varepsilon - u^2$ of the plasma, such that the correction factor $(v - u)/T$ corresponds to the thermal spread of the particles about the average velocity.

Comparison With Other Operators

The operator (5.111) can be shown to be related to the one presented by Oppenheim [97], Ong and Yu [96], as well as Clemmow and Dougherty [32], by multiplying the right hand side of (5.111) with the denominator of the second term

$$C'[f] = \nu \frac{\partial}{\partial v}\left[(\varepsilon - u^2)\frac{\partial}{\partial v} + (v - u)\right] f. \tag{5.112}$$

Upon insertion of the definition of ε we get

$$C'[f] = \nu \frac{\partial}{\partial v}\left[\left(\frac{1}{n}\int v^2 f\, dv - u^2\right)\frac{\partial}{\partial v} + (v - u)\right] f$$
$$= \nu \frac{\partial}{\partial v}\left[\left(\frac{1}{n}\int (v - u)^2 f\, dv\right)\frac{\partial}{\partial v} + (v - u)\right] f \tag{5.113}$$

where the last equality holds as

$$\frac{1}{n}\int (v-u)^2 f\, dv = \frac{1}{n}\int \left(v^2 - 2uv + u^2\right) f\, dv = \frac{1}{n}\left(\int v^2 f\, dv - 2nu^2 + nu^2\right) = \frac{1}{n}\int v^2 f\, dv - u^2.$$

Equation (5.113) is the expression presented in the above references [97, 96, 32]. It has similar properties as our operator, i.e., it preserves the total particle number, the total linear momentum, and the total energy. It also relaxes towards a Maxwellian, but the coefficient of the diffusion term modifies its behaviour, such that the strength of the diffusion scales with the thermal energy $\epsilon - u^2$.

5.4.2. Discrete Collision Operator

To obtain the discrete collision operator we repeat the derivation at the discrete level in the same spirit we derived the discrete action principle, i.e., by mimicking the continuous derivation. We discretise the derivatives in the collision operator (5.101) by

$$C_d[j] = \nu \left[\frac{f(j-1) - 2f(j) + f(j+1)}{h_v^2} + A_{1,d}\frac{f(j+1) - f(j-1)}{2h_v} \right.$$
$$\left. + A_{2,d}\frac{v(j+1)\,f(j+1) - v(j-1)\,f(j-1)}{2h_v} \right]. \tag{5.114}$$

where we dropped the spatial and time indices as the collision operator is always computed at a single point in spacetime (i,k). The coefficients $A_{1,d}$ and $A_{2,d}$ are choosen to enforce the discrete conservation properties

$$0 = \sum_j \Big(v(j) + v(j+1)\Big)\Big(C_d(j) + C_d(j+1)\Big), \tag{5.115a}$$

$$0 = \sum_j \Big(v^2(j) + v^2(j+1)\Big)\Big(C_d(j) + C_d(j+1)\Big). \tag{5.115b}$$

Discrete partial integration, i.e., reordering of the sums, then leads to the following expressions for the correction factors $A_{1,d}$ and $A_{2,d}$

$$A_{1,d} = -\frac{u_d}{\varepsilon_d - u_d^2}, \qquad A_{2,d} = \frac{1}{\varepsilon_d - u_d^2} \tag{5.116}$$

which, no surprises, are exactly the same as their continuous counter parts (5.110). The important result is that we automatically obtain the correct energy and momentum preserving discretisation of the moments

$$n_d = h_v \sum_j f(j), \qquad u_d = \frac{h_v}{n_d}\sum_j v(j)f(j), \qquad \varepsilon_d = \frac{h_v}{n_d}\sum_j v^2(j)f(j). \tag{5.117}$$

The complete discretised collision operator is (replacing the subscript d with grid coordinates)

$$C_{i,j,k} = \nu \left[\frac{f(i,j-1,k) - 2f(i,j,k) + f(i,j+1,k)}{h_v^2} \right.$$
$$\left. + \frac{[v(j+1) - u(i,k)]\,f(i,j+1,k) - [v(j-1) - u(i,k)]\,f(i,j-1,k)}{2h_v\,[\varepsilon(i,k) - u^2(i,k)]} \right]. \tag{5.118}$$

We add the discrete operator to the simplified nonlinear Vlasov equation (5.90) by employing a spacetime averaging approach mimicking the result of the discrete variational principle

$$\frac{\bar{f}_{k+1} - \bar{f}_k}{h_t} + \frac{1}{4}\Big(A(f_{k+1}, h_{k+1}) + A(f_{k+1}, h_k) + A(f_k, h_{k+1}) + A(f_k, h_k) \Big) = C_{\bar{i}, j, \bar{k}}[f] \quad (5.119)$$

where

$$C_{\bar{i}, j, \bar{k}} = \frac{1}{8}\Big[C_{i-1, j, k} + 2\, C_{i, j, k} + C_{i+1, j, k} + C_{i-1, j, k+1} + 2\, C_{i, j, k+1} + C_{i+1, j, k+1} \Big].$$

This averaging does not have an effect on the conservation properties as the discrete collision operator is designed to conserve the total particle number, the linear momentum and energy locally, i.e., for each spacetime grid point (i, k) separately.

5.5. Numerical Examples

In this section we consider several numerical examples that can be considered standard benchmark cases [53, 134, 15, 38, 5, 91, 28].
If not noted otherwise, the simulation domain is $[0, 2\pi/k] \times [-v_{\max}, +v_{\max}]$. The resolution is always $n_x = 201$, $n_v = 401$, and except for the simulations with the linear integrator, the timestep is $h_t = 0.1$ in units of the inverse plasma frequency. For the linear integrator the timestep is $h_t = 0.01$.
Most of the examples are initialised as a perturbation of a Maxwellian distribution, given by

$$f_M(x, v) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\tfrac{1}{2} v^2 \right\}. \tag{5.120}$$

The temperature is set to one such that the thermal velocity is also one. The density is normalised to one. The initial potential is determined by the initial distribution function via the Poisson equation.

5.5.1. Simulation Code

The variational integrator for the Vlasov-Poisson system, equations (5.90) and (5.91), constitutes a nonlinearly implicit system of equations. The nonlinearity is solved by Newton's method where in each Newton step a direct linear solver based on LU decomposition and a GMRES correction is employed. The initial guess for the Newton solver is either computed by the linear variational integrator from section 5.3.3, or, in linear or weakly nonlinear examples, the last timestep is used. Depending on the problem, the Newton solver usually needs 1-3 iterations to converge with a residual smaller 10^{-11}. In most cases, the LU decomposition of the Jacobian needs only be carried out once per timestep.
The implementation of efficient solvers is a topic left for of future research, but preliminary results suggest that for a sufficiently good initial guess, the LU decomposition can be replaced by an iterative method (GMRES), where only a few iterations are needed to solve the linear system.
The code is implemented in Python and Cython using PETSc [14, 13] to solve the nonlinear system and take care of the parallel communication and MUMPS [129] for the LU decomposition.

5.5.2. Diagnostics

We have not yet carried out a detailed analysis of the discrete conservation laws of the Vlasov-Poisson system. We therefore assume a discrete representation of the conservation properties according to the discretisation of the Lagrangian, i.e., a midpoint representation.

The total particle number is computed as

$$N_k = \frac{1}{4} \sum_{i=1}^{n_x-1} \sum_{j=1}^{n_v-1} \left(f_{i,j,k} + f_{i+1,j,k} + f_{i+1,j+1,k} + f_{i,j+1,k} \right) h_x h_v, \tag{5.121}$$

where $i = n_x$ corresponds to $i = 1$ as we use periodic boundary conditions in space. Furthermore, the velocity domain should always be chosen large enough, such that $f_{i,1} = f_{i,n_v} = 0$, the above expression is really just a sum of f over the whole phasespace grid (i, j),

$$N_k = \frac{1}{4} \sum_{i=1}^{n_x} \sum_{j=1}^{n_v} f_{i,j,k} h_x h_v. \tag{5.122}$$

Similar to (5.121), the L^2 norm is computed as

$$L_k^2 = \sum_{i=1}^{n_x-1} \sum_{j=1}^{n_v-1} \left[\frac{1}{4} \left(f_{i,j,k} + f_{i+1,j,k} + f_{i+1,j+1,k} + f_{i,j+1,k} \right) \right]^2 h_x h_v, \tag{5.123}$$

momentum is computed as

$$P_k = \frac{1}{8} \sum_{i=1}^{n_x-1} \sum_{j=1}^{n_v-1} \left(f_{i,j,k} + f_{i+1,j,k} + f_{i+1,j+1,k} + f_{i,j+1,k} \right) \left(v_j + v_{j+1} \right) h_x h_v, \tag{5.124}$$

energy is computed as

$$E_k = \frac{1}{16} \sum_{i=1}^{n_x-1} \sum_{j=1}^{n_v-1} \left(f_{i,j,k} + f_{i+1,j,k} + f_{i+1,j+1,k} + f_{i,j+1,k} \right) \times$$
$$\times \left(h_{i,j,k} + h_{i+1,j,k} + h_{i+1,j+1,k} + h_{i,j+1,k} \right) h_x h_v, \tag{5.125}$$

and entropy is computed as

$$S_k = \frac{1}{4} \sum_{i=1}^{n_x-1} \sum_{j=1}^{n_v-1} \left(f_{i,j,k} + f_{i+1,j,k} + f_{i+1,j+1,k} + f_{i,j+1,k} \right) \times$$
$$\times \log \left(\frac{1}{4} \left(f_{i,j,k} + f_{i+1,j,k} + f_{i+1,j+1,k} + f_{i,j+1,k} \right) \right) h_x h_v. \tag{5.126}$$

It is expected that a rigorous calculation of the discrete conservation laws from the discrete Noether theorem (section 3.2.4) may improve on the quality of the discrete conservation laws. Even with the foregoing heuristic diagnostics, we obtain very satisfying results so that a more precise analysis is left for future work.

5.5.3. Landau Damping

Landau damping is probably the most popular benchmark for the Vlasov equation, first because it is a purely kinetic effect involving phase mixing, and second because there are analytical results available to compare with (at least in the linear case). The initial distribution function is given by

$$f(x,v) = f_M \left(1 + A \cos(kx)\right), \tag{5.127}$$

where f_M is a Maxwellian distribution (5.120). With $k = 0.5$, the spatial simulation domain is $[0, 4\pi]$, and $v_{\max} = 10$. The resolution is $n_x = 201$, $n_v = 401$, and the timestep is $h_t = 0.1$ in units of the inverse plasma frequency. The spatial step width $h_x = 2\pi/n_x k$ depends on the chosen wave number k.

Linear Landau Damping

At first, we consider the linear case, which can be compared with theoretical results, that is an initial perturbation with $A = 0.01$, $k = 0.5$, and without collisions ($\nu = 0$). Figure 5.2) shows the time traces of the errors of the total particle number, the total energy, and linear momentum, while the evolution of the electrostatic energy is displayed in figure 5.3. Using only the marked maxima, the damping rate is computed to be $\gamma = -0.152$, which is very close to the theoretical value of $\gamma = -0.153$. Using only the first ten maxima, we obtain the predicted value $\gamma = -0.153$. The total particle number and the total linear momentum are preserved optimally (see figure 5.2) and exhibit the expected oscillatory behaviour about a constant value. The error in the total energy is very small but seems to grow during the simulation. Indeed, the error of the particle number oscillates about zero for $t < 40$, and then jumps to 10^{-13}. The energy error appears to grow monotonically, although it remains very small. This is attributed to the formation of structures on a scale length shorter than the grid step size, namely, subgrid modes which are discussed below.

We can therefore conclude that, without additional ad hoc devices (such as hyperdiffusion), the integrator shows remarkable conservation properties and accuracy, as far as the grid is sufficient to resolve the phase-mixing structure of the distribution function. For long-time integration, special care of subgrid modes should be taken.

Subgrid Modes and Collision Operator

At about $t = 40$, subgrid modes start to develop. Consequently, large gradients in the distribution function appear, which in turn lead to an unphysically large electrostatic potential. Therefore, the total energy error increases, as can be seen in figure 5.2, and the damping rate becomes spurious (figure 5.3). To remove these subgrid modes, we employ the collision operator described in section 5.4. It dissipates the L^2 norm but retains the conservation of total particle number, total energy, and total linear momentum.

At a collision frequency of $\nu = 10^{-4}$, the error of the conserved quantities is almost optimal (figure 5.4), and the electrostatic field is damped up to the machine accuracy (figure 5.5). When measuring the absorption coefficient γ from the first timesteps, a good agreement with the theoretical value is observed. Long-time measurements of γ result, however, in values too small compared with the theoretical value. For the marked maxima in figure 5.5, we obtain $\gamma = -0.144$. This behaviour is explained by subgrid modes which are not damped completely by the collision operator and are therefore still active.

To obtain the approximately correct value of $\gamma = -0.152$, we have to increase the collision frequency to $\nu = 4 \times 10^{-4}$ (figure 5.6). Even so, there is almost no visible difference in the time traces of the energy error for $\nu = 1 \times 10^{-4}$ and $\nu = 4 \times 10^{-4}$, figures 5.4 and 5.6, respectively, the difference is obvious in the damping of the electrostatic potential.

These results suggests that the linear case, for which an analytical solution is known, can be used to tune the collision frequency for a given step width h_v in velocity space. We will see in the following nonlinear examples, that in all cases a collision frequency of $\nu = 4 \times 10^{-4}$ is necessary to obtain accurate conservation of particle number, energy and momentum on long timescales.

Nonlinear Landau Damping

In the case of nonlinear Landau damping, $A = 0.5$ and $k = 0.5$, the previous observations manifest more clearly. The effects of the subgrid modes are much more pronounced as the nonlinear character of the dynamics tends to develop smaller scale structures in phasespace. In particular, the phase mixing that comes along with Landau damping quickly develops very small phasespace structures that cannot be resolved.

With a collision frequency of $\nu = 10^{-4}$, the conservation of energy and the total particle number is severely violated. Only the error in the linear momentum is very small (see figure 5.8). In contrast to the previous example, here a larger contribution to the error seems to come from the kinetic part. The relatively large error in the total particle number is directly reflected in the error of the kinetic energy. With a collision frequency of $\nu = 4 \times 10^{-4}$, instead, conservation of the total particle number, energy as well as the total linear momentum is optimal, see figure 5.10.

The initial damping rate is hardly changed by the collisions. For $\nu = 0$ we find $\gamma_1 = -0.2854$ (no figure), for $\nu = 10^{-4}$ we find $\gamma_1 = -0.2856$ (figure 5.9) and for $\nu = 4 \times 10^{-4}$ we find $\gamma_1 = -0.2864$ (figure 5.11). All numbers are equal to two digits and compare well with the existing literature, e.g., Cheng and Knorr computed $\gamma_1 = -0.281$, Nakamura and Yabe computed $\gamma_1 = -0.280$, and Heath et al. computed $\gamma_1 = -0.287$.

The effect of the collisions on the second phase, where the electrostatic potential is growing again, is more pronounced. Without collisions, we obtain $\gamma_2 = 0.0860$, with $\nu = 10^{-4}$ we obtain $\gamma_2 = 0.0830$ and for $\nu = 4 \times 10^{-4}$ the we find the growth rate to be $\gamma_2 = 0.0746$ and hence considerably reduced. The results of Cheng and Knorr, who computed $\gamma_2 = 0.084$, and Nakamura and Yabe, who computed $\gamma_2 = 0.0845$, are closer to our results with less or no collisions. The result of Heath et al., $\gamma_2 = 0.0746$, on the other side, matches ours exactly (for a summary see table 5.1).

Not surprisingly, the collisions damp the electrostatic field, and more so for larger collision frequencies ν. The important question is whether they are just removing unphysical contributions to the field energy that originate from subgrid modes, or whether they damp the field too much. As already described, the relatively large error in the energy and particle number for $\nu = 10^{-4}$ suggests that subgrid modes are not sufficiently damped. Therefore the electrostatic potential is likely to be affected as well, such that part of the electrostatic energy is due to subgrid modes and therefore unphysical. Consequently, the electrostatic energy is likely to be overestimated in that case. On the other hand, it cannot be anticipated that the smaller growth and consecutively stronger damping for $\nu = 4 \times 10^{-4}$ is closer to the real situation as we are counteracting a numerical effect with an effective collision operator, which neither allows us to draw conclusions for the collisionless case, nor represents the physical collision process. Furthermore, the electrostatic energy is much larger in the nonlinear case than in the linear case. Therefore the error in the total

energy can not be attributed to the kinetic or the potential part without ambiguity. The correlation between the errors in the total particle number and the total energy suggests, however, that the energy error arises mainly from the kinetic energy. A definite conclusion is not possible, but simulations with higher resolution and higher order integrators should indicate which effects are physical and which are numerical.

In figure 5.12, the time evolution of the distribution function is plotted. The phase mixing is nicely visible as is the action of the collision operator. At about $t = 30$ the phasespace structures start to become too small to be resolvable and get therefore damped by the collisions. At about $t = 50$, the fine scale structures have disappeared almost completely, but a large scale oscillation is still visible. At $t = 200$ this oscillation has been further damped, such that it does not appear in the plot anymore.

Linear Integrator

For both, linear and nonlinear Landau damping, we also did simulations with the linear integrator from section 5.3.3. The simulation parameters are the same, except for the timestep, which was chosen as $h_t = 0.01$. We did only simulations without collisions, as in the linear scheme, the collision operator can only be treated explicitly (as it is inherently nonlinear) and therefore even smaller timesteps would be necessary. For both, linear and nonlinear Landau damping, we find the same behaviour with the linear method as we do with the fully nonlinear method.

In the case of linear Landau damping, the total particle number and the total linear momentum are well preserved (see figure 5.13). The conservation of the total energy is good, but the error is larger than with the nonlinear integrator. In fact, we observe a behaviour of the energy error that is typical for multisymplectic integrators, namely, the energy is not preserved exactly, but its error is bounded, often oscillating, where the amplitude of the oscillation depends on the timestep. This is the reason why we choose a smaller timestep for the simulations with the linear integrator, i.e., to still get good energy conservation. In the case of linear Landau damping, the amplitude of the oscillation is $\mathcal{O}(10^{-9})$. However, after the initial perturbation is damped, it becomes much smaller again.

In the case of nonlinear Landau damping, initially, the total particle number and total linear momentum are well preserved and the energy error shows a similar behaviour as in the linear case, albeit with a larger amplitude of the error which is $\mathcal{O}(10^{-5})$ (see figure 5.14). Eventually the energy error grows larger due to subgrid modes, which will also spoil the momentum and particle number conservation when running for longer times.

We see that the linear integrator poses a viable alternative if solving a nonlinear system is not an option. That the linear integrator is working rather well is probably attributed to the fact that the Vlasov-Poisson system consists of two linear equations. The nonlinearity arises only through the coupling of the two equations.

L^2 Norm and Entropy

Before we move to the next example, a comment on the evolution of entropy and the L^2 norm is in order. The variational integrator for the Vlasov-Poisson system (5.90, 5.91) preserves the L^2 norm of the distribution function exactly (see figures 5.15 and 5.17, top). Through the application of the collision operator, the L^2 norm is dissipated (figures 5.15 and 5.17, middle and bottom).

Integrator	ν	γ_1	γ_2
Linear VI	0	-0.285	$+0.087$
Nonlinear VI	0	-0.285	$+0.086$
Nonlinear VI	1×10^{-4}	-0.286	$+0.083$
Nonlinear VI	4×10^{-4}	-0.286	$+0.075$
Cheng and Knorr [28]	-	-0.281	$+0.084$
Nakamura and Yabe [91]	-	-0.280	$+0.085$
Heath et al. [53]	-	-0.287	$+0.075$

Tab. 5.1.: Damping and growth rates in nonlinear Landau damping simulations with variational integrators and comparison with previous works.

Entropy is not conserved by our variational integrator, but without collisions it can be used as a diagnostic for the appearance of subgrid modes. In the case of linear Landau damping, where the effect of the subgrid modes onto the distribution function is rather small, the entropy grows slowly but steadily as can be seen in the top of figure 5.16. In the case of nonlinear Landau damping, where the effect of the subgrid modes onto the distribution is more severe, the entropy is growing slowly at first, just as in the linear case, but starting from about $t = 15$ it is growing much more rapidly (top of figure 5.18). This sudden growth indicates the appearance of subgrid modes, long-time before they are visible in the energy diagnostics. In the simulations with collisions (bottom of figures 5.16 and 5.18) the entropy is increasing more smoothly, not showing such sudden jumps.

5.5.4. Twostream Instability

The distribution function is initialised as

$$f(x,v) = v^2 f_M(x,v) \big(1 + A\cos(kx)\big), \tag{5.128}$$

with amplitude $A = 0.05$ and wave number $k = 0.5$. The simulation parameters are the same as before. The spatial domain is $[0, 4\pi]$, $v_{\max} = 10$, $n_x = 201$, $n_v = 401$, $h_t = 0.1$. The collision frequency is taken to be either $\nu = 0$ or $\nu = 4 \times 10^{-4}$, following the tuning by linear Landau damping, described in the previous section. The distribution function describes two particle beams, propagating in opposite direction and having a small perturbation imposed on them.

In simulations without collisions ($\nu = 0$), we find a good conservation of the total particle number and the total linear momentum, but conservation of the total energy is violated (figure 5.19). Employing the collision operator, optimal energy conservation can be restored (figure 5.20). The collision frequency necessary to retain the correct energy throughout the whole simulation is $\nu = 4 \times 10^{-4}$, as it was estimated in the linear Landau damping simulations.

The distribution function, figure 5.21, exhibits the correct qualitative behaviour. After an initial growth of the instability, particles become trapped and a hole in phasespace forms. Between $t = 100$ and $t = 200$, the distribution function takes an almost steady state.

5.5.5. Jeans Instability

Finally, we are considering a test case from gravitational dynamics, the Jeans instability [17, 29]. The only difference compared with plasma dynamics is that the gravitational

field is always attractive. This results in a change of sign in the Poisson equation. The distribution function is initialised as

$$f = f_M \left(1 + A \cos(kx) \right), \tag{5.129}$$

with $A = 0.01$ and $k = 0.8$. The spatial domain is $[0, 2.5\,\pi]$, $v_{\max} = 10$, $n_x = 201$, $n_v = 401$, $h_t = 0.1$. For $k < 1$, the distribution function is unstable and collapses towards the centre of the simulation domain. For $k > 1$ initial perturbations are damped.

The conservation properties are very similar as in the case of the twostream instability. Without collisions, the total particle number and the total linear momentum are well conserved but not the total energy (figure 5.22). Adding collisions, with the same frequency of $\nu = 4 \times 10^{-4}$ as before (h_v is still the same), energy conservation is retained (figure 5.23). The qualitative behaviour of the distribution function (figure 5.24) meets the expectations. As we choose $k < 1$, the initial perturbation is unstable and develops a swirl about the centre of the simulation domain, corresponding to a gravitational collapse. Between $t = 50$ and $t = 100$ the distribution functions reaches a steady state and barely changes until $t = 200$.
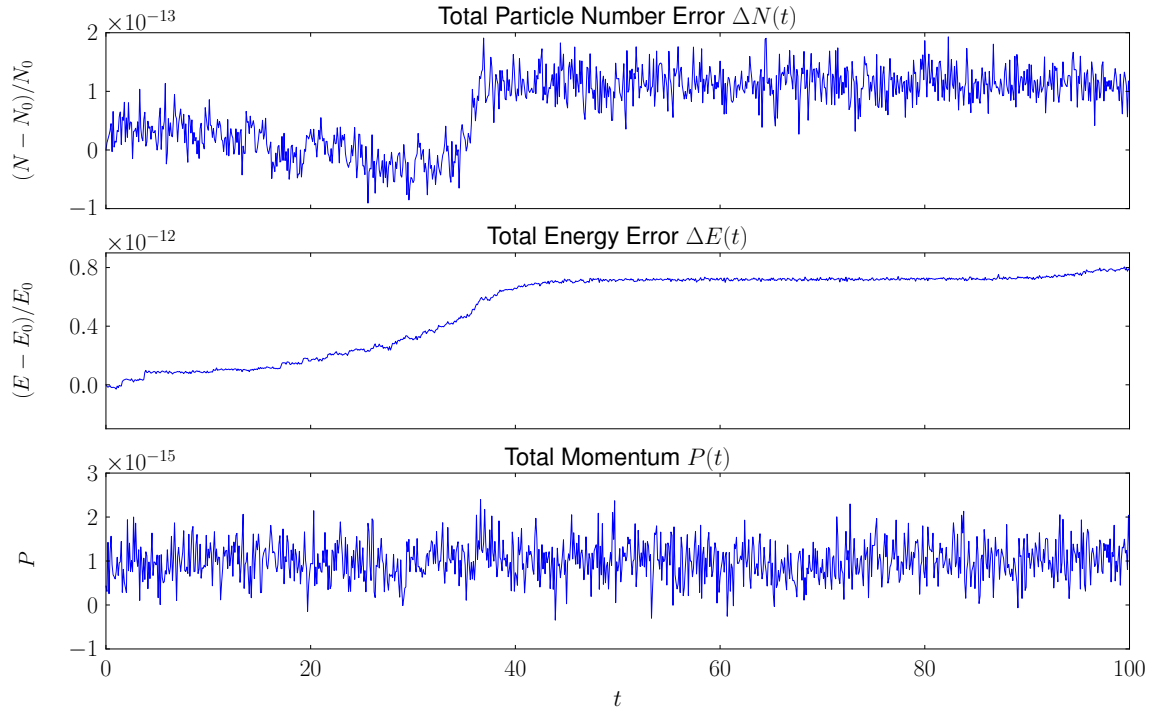
Fig. 5.2.: Linear Landau damping without collisions. The total particle number and the total linear momentum are well preserved. Conservation of the total energy is slightly violated due to subgrid mode effects.
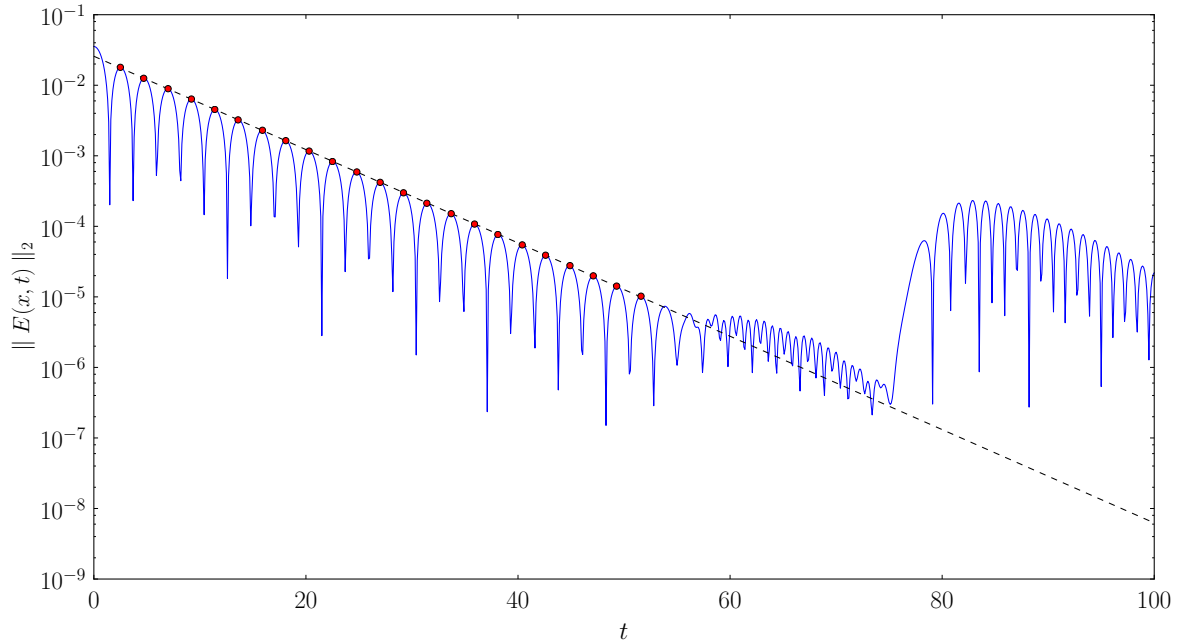
Fig. 5.3.: Linear Landau damping without collisions. Evolution of the electrostatic energy. At about $t = 40$, subgrid modes start to develop, eventually spoiling the damping of the initial perturbation.
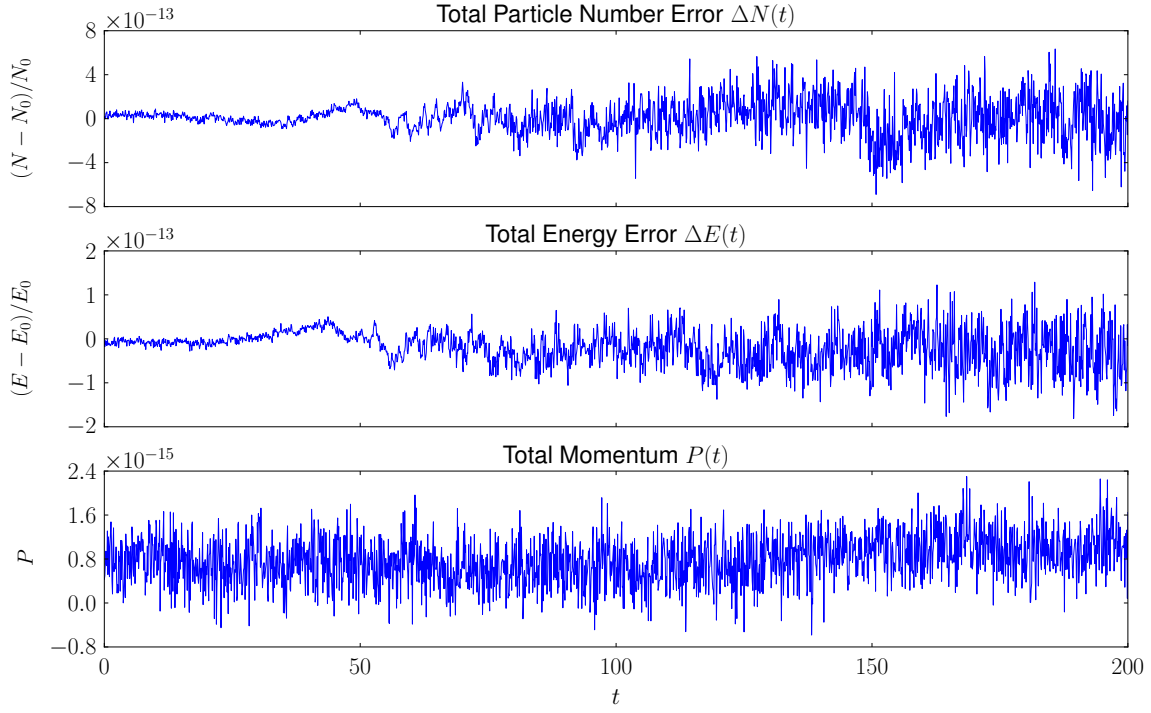
Fig. 5.4.: Linear Landau damping with collision frequency $\nu = 10^{-4}$. With collisions, the total particle number, energy, and linear momentum are all preserved up to machine precision.
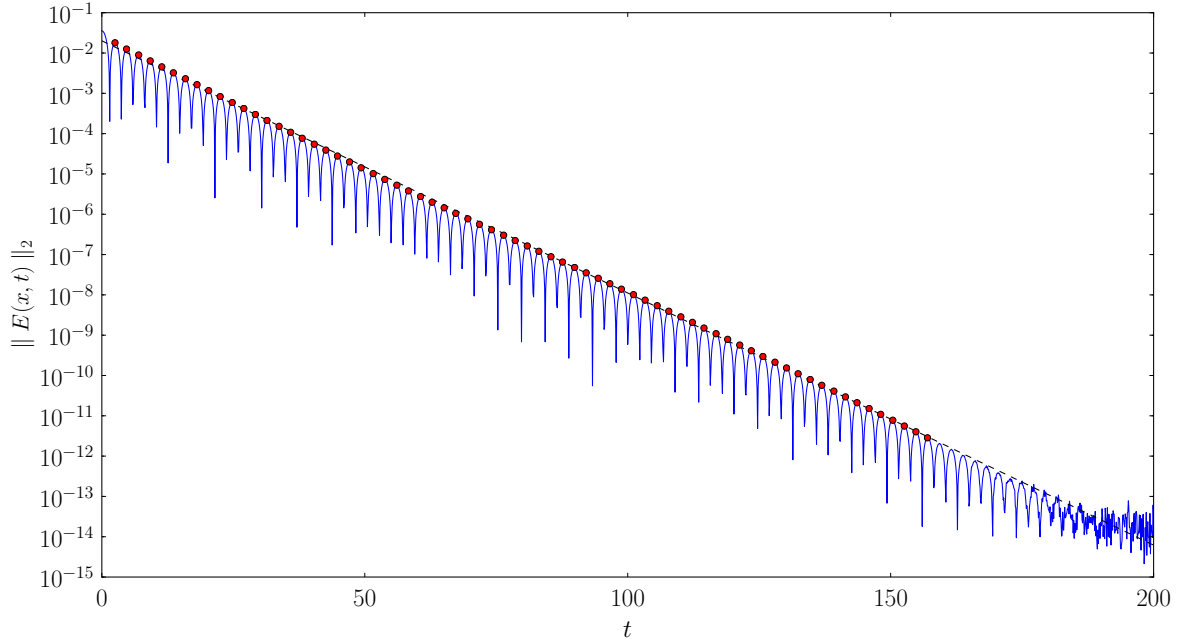
Fig. 5.5.: Linear Landau damping with collision frequency $\nu = 10^{-4}$. Evolution of the electrostatic energy. The qualitative behaviour of the linear Landau damping is restored by the collisions, but the damping rate, $\gamma = -0.144$, is to low.
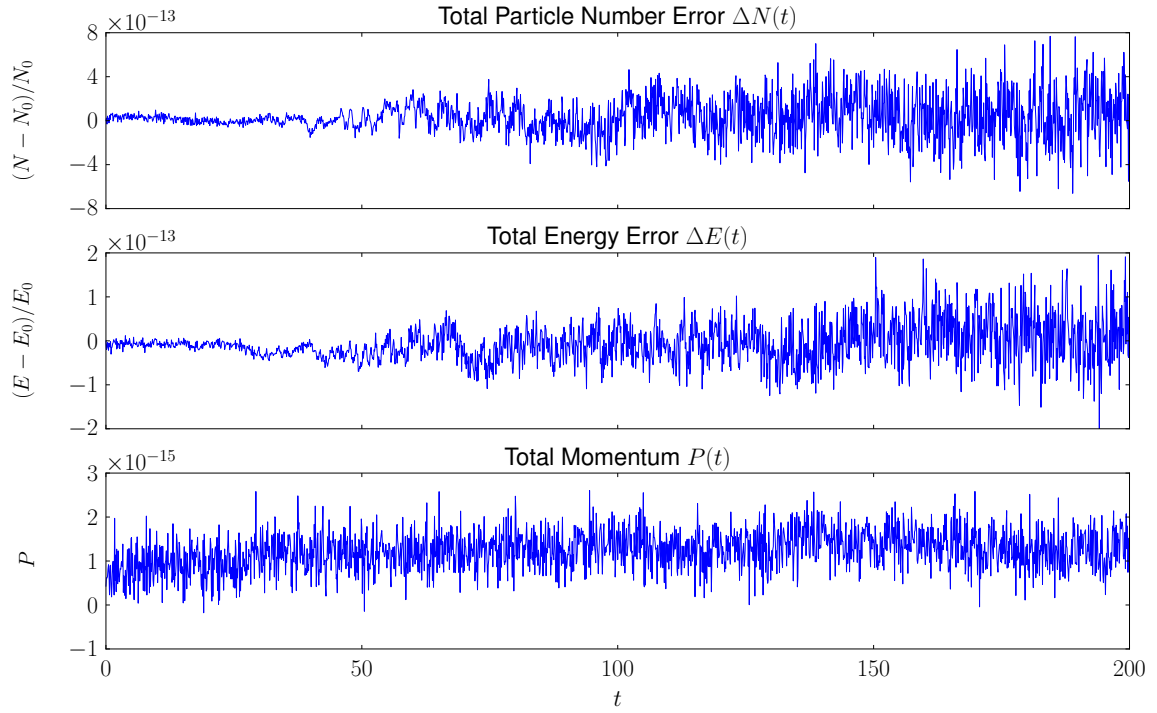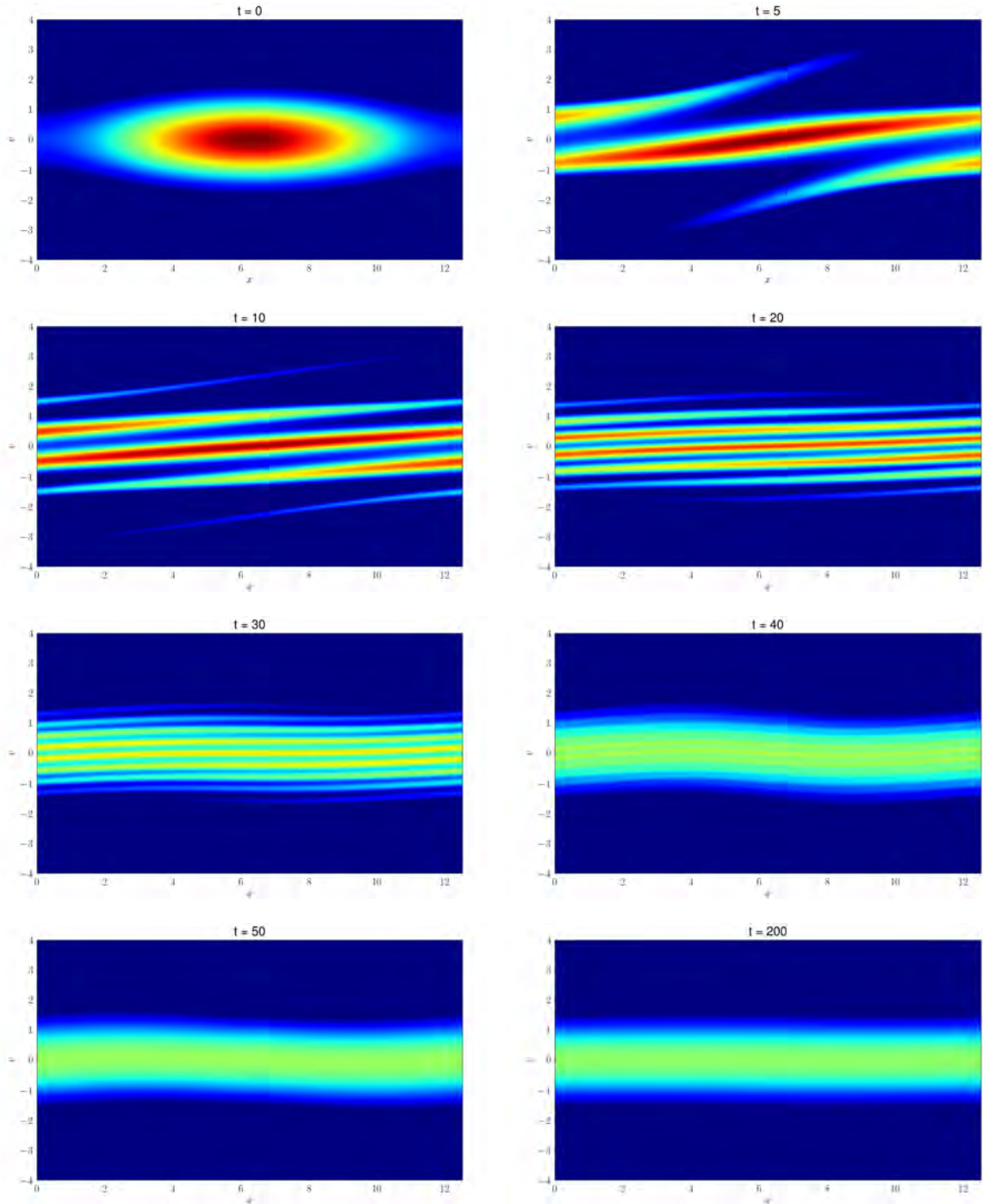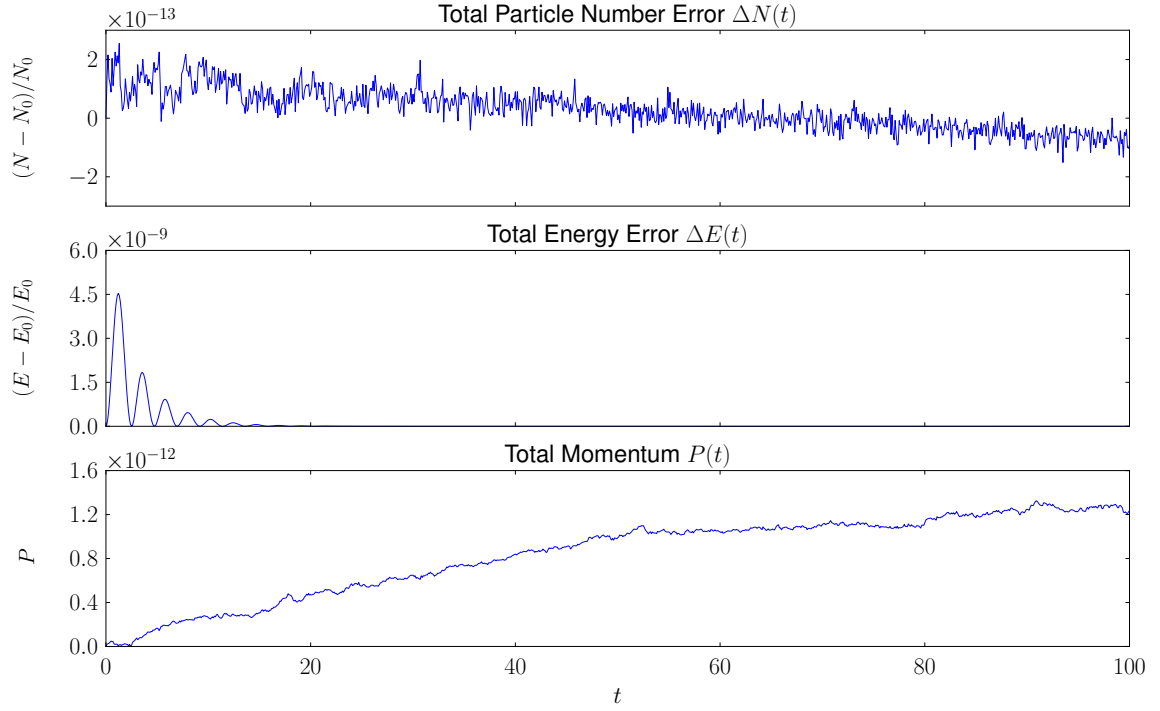
Fig. 5.6.: Linear Landau damping with collision frequency $\nu = 4 \times 10^{-4}$. With collisions, the total particle number, energy, and linear momentum are all preserved up to machine precision. Hardly any differences to the case with $\nu = 10^{-4}$ are visible.
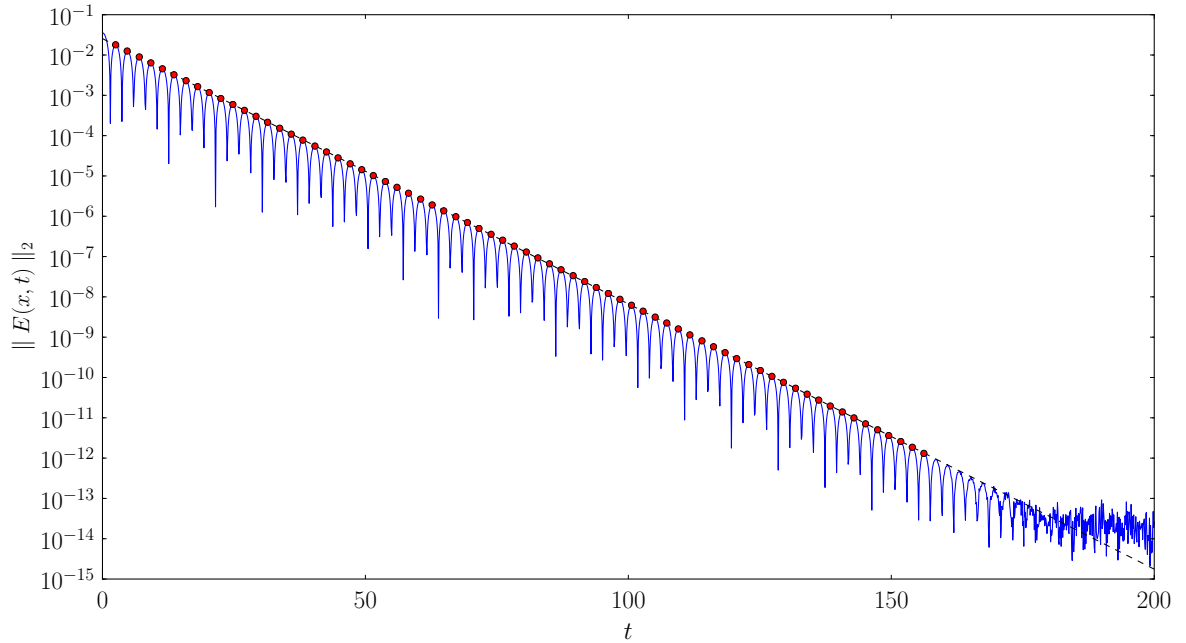
Fig. 5.7.: Linear Landau damping with collision frequency $\nu = 4 \times 10^{-4}$. Evolution of the electrostatic energy. Both, the correct qualitative behaviour and the correct damping rate, $\gamma = -0.152$, are obtained for $\nu = 4 \times 10^{-4}$.
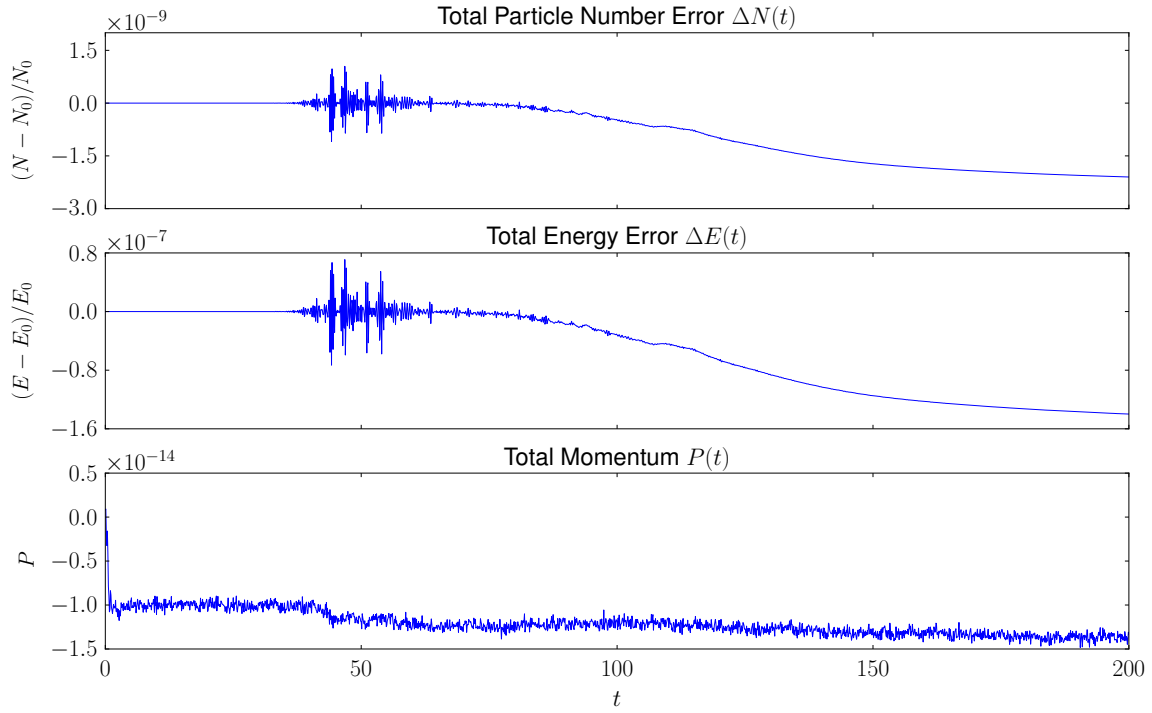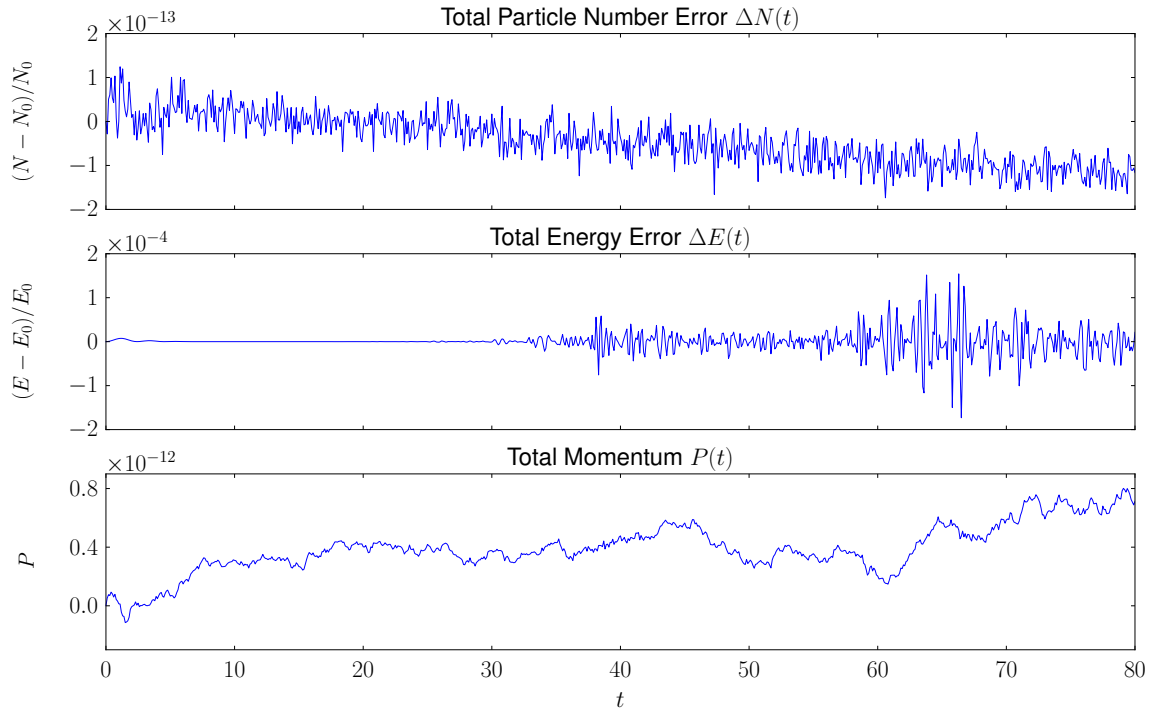
Fig. 5.8.: Nonlinear Landau damping with collision frequency $\nu = 10^{-4}$. Conservation of the total particle number and energy is violated due to subgrid mode effects. Total linear momentum is preserved exactly.
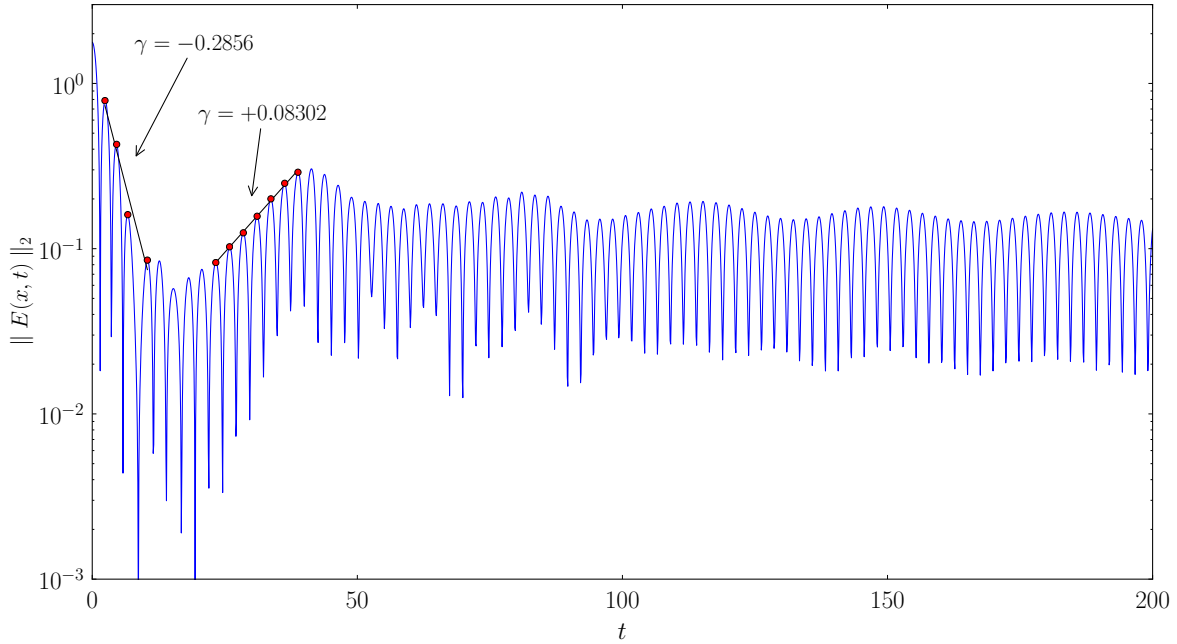
Fig. 5.9.: Nonlinear Landau damping with collision frequency $\nu = 10^{-4}$. Evolution of the electrostatic energy.
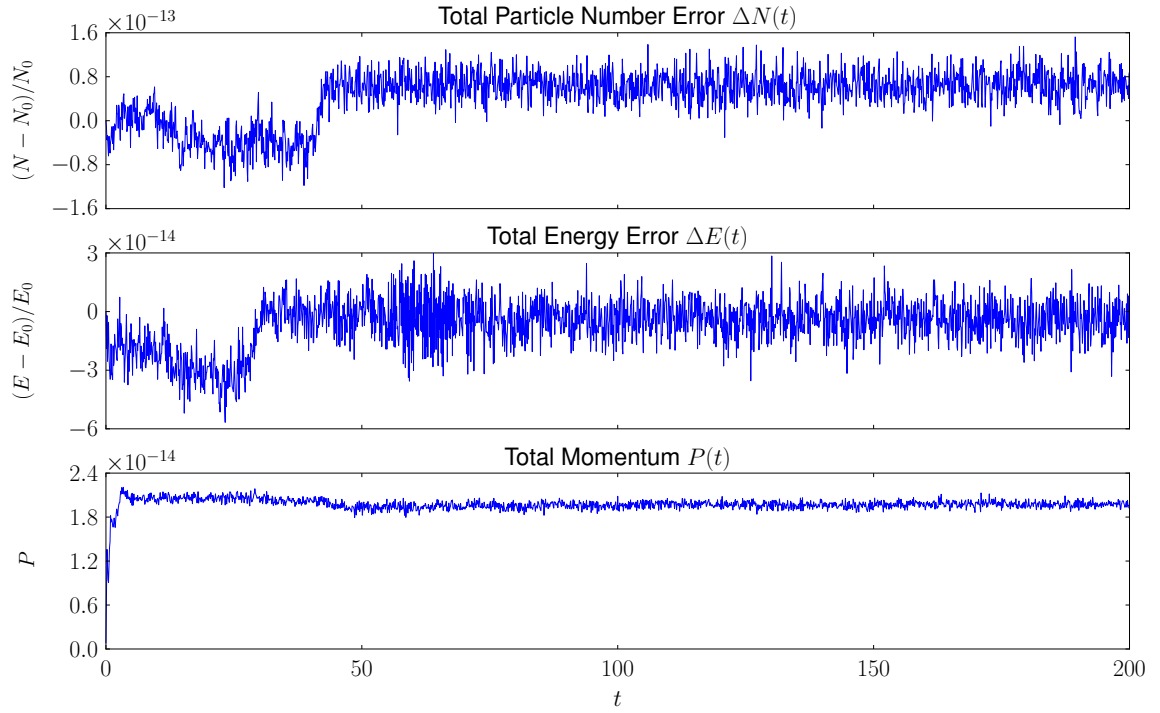
Fig. 5.10.: Nonlinear Landau damping with collision frequency $\nu = 4 \times 10^{-4}$. Collisions retain the conservation of the total particle number and energy in addition to exact preservation of the linear momentum.
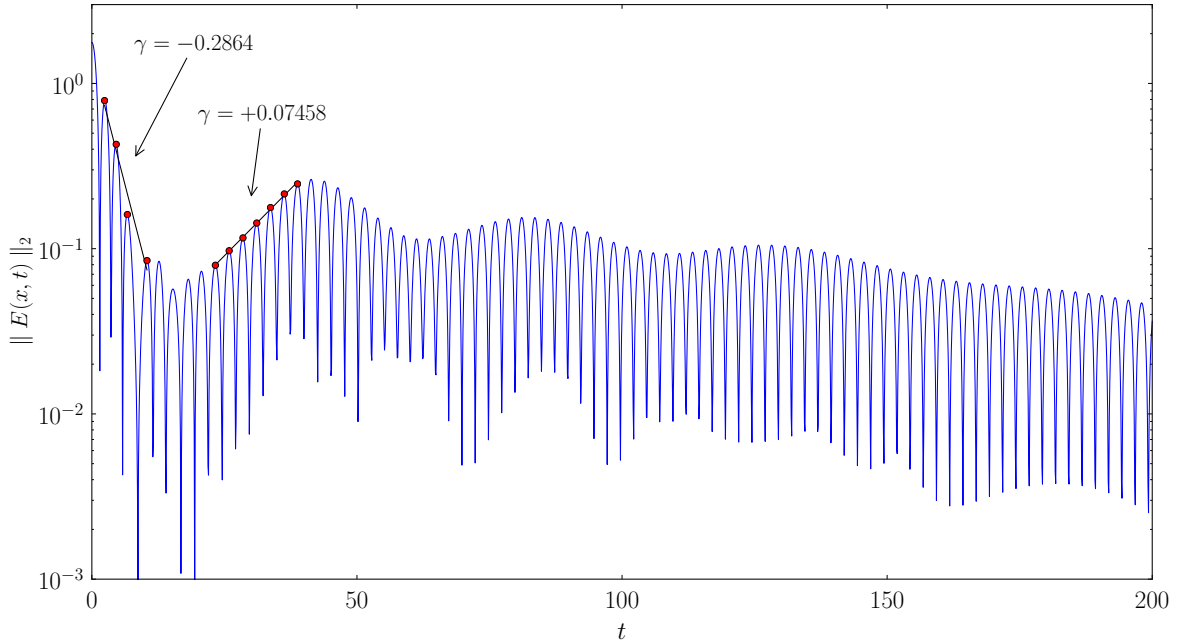
Fig. 5.11.: Nonlinear Landau damping with collision frequency $\nu = 4 \times 10^{-4}$. Evolution of the electrostatic energy.

Fig. 5.12.: Nonlinear Landau damping with collision frequency $\nu = 4 \times 10^{-4}$. Contours of the distribution function in phasespace. Contours are linear and constant.

Fig. 5.13.: **Linear Landau damping with linear integrator and without collisions.** The total particle number and the total linear momentum are well preserved. Conservation of the total energy is good, but the error is larger than with the nonlinear integrator.
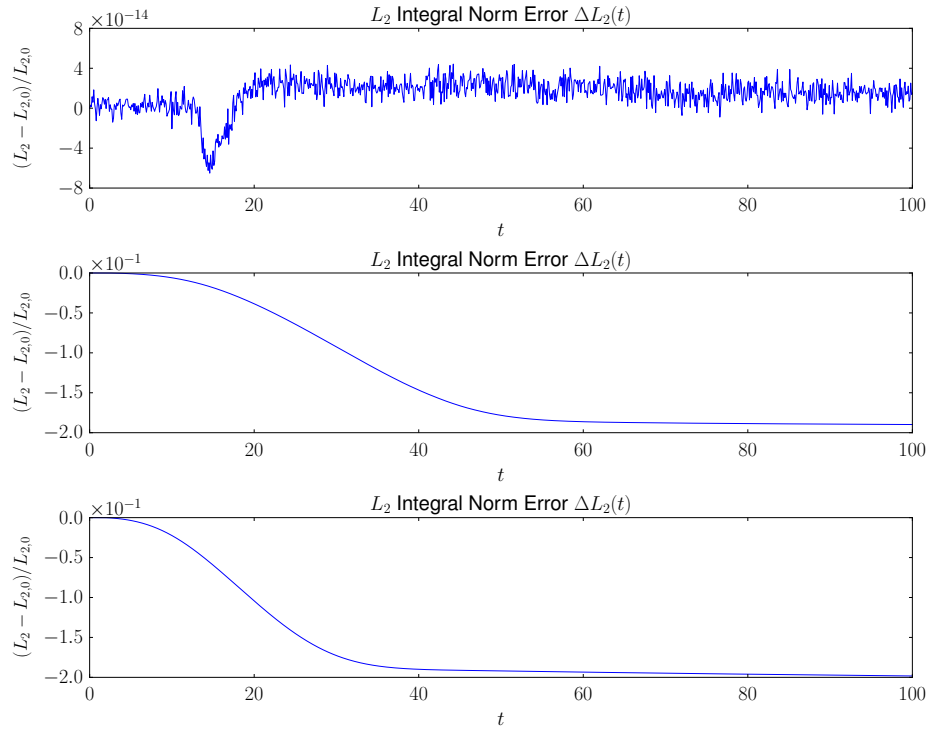

Fig. 5.14.: **Nonlinear Landau damping with linear integrator and without collisions.** Total particle number and total linear momentum are well preserved, but the energy error is soon dominated by subgrid mode effects.
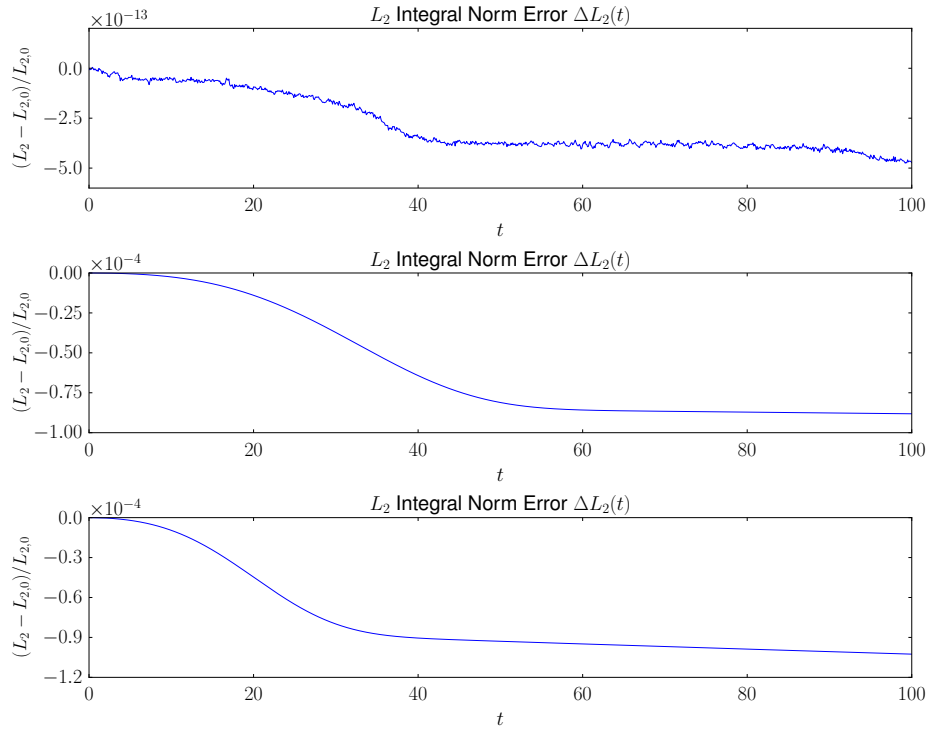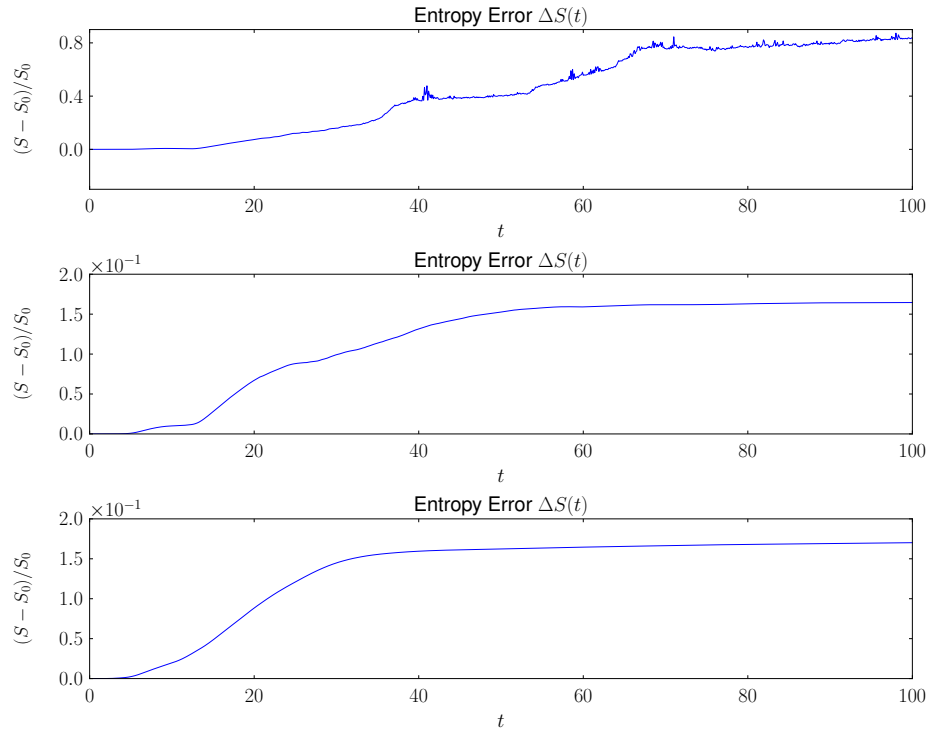
Fig. 5.15.: Linear Landau damping. Evolution of the L^2 norm.
Top: $\nu = 0$**, Middle:** $\nu = 10^{-4}$**, Bottom:** $\nu = 4 \times 10^{-4}$**.**
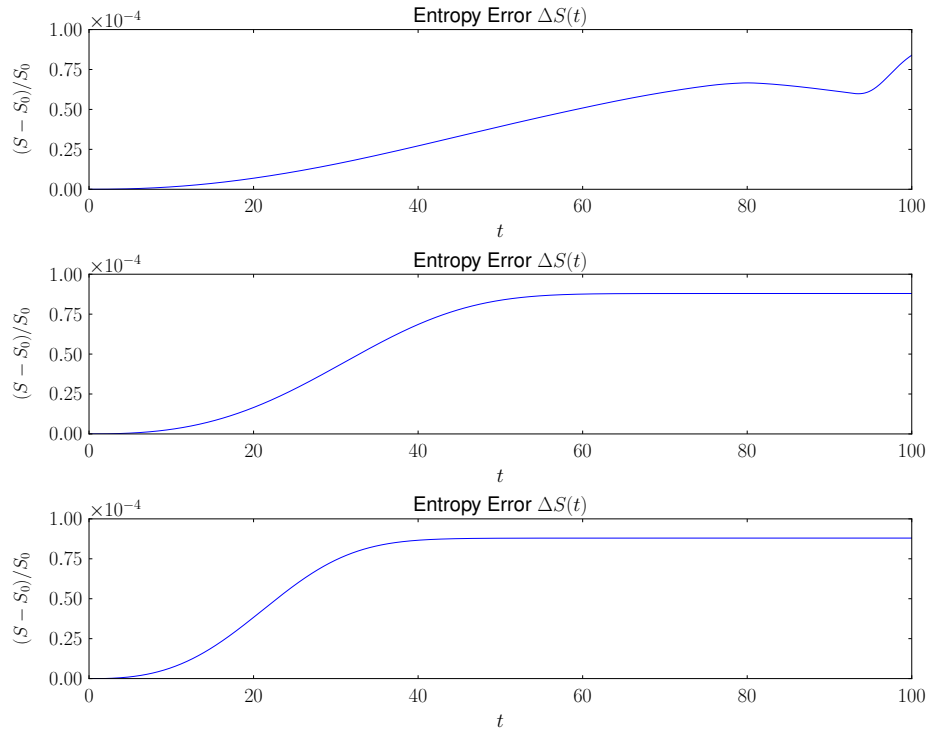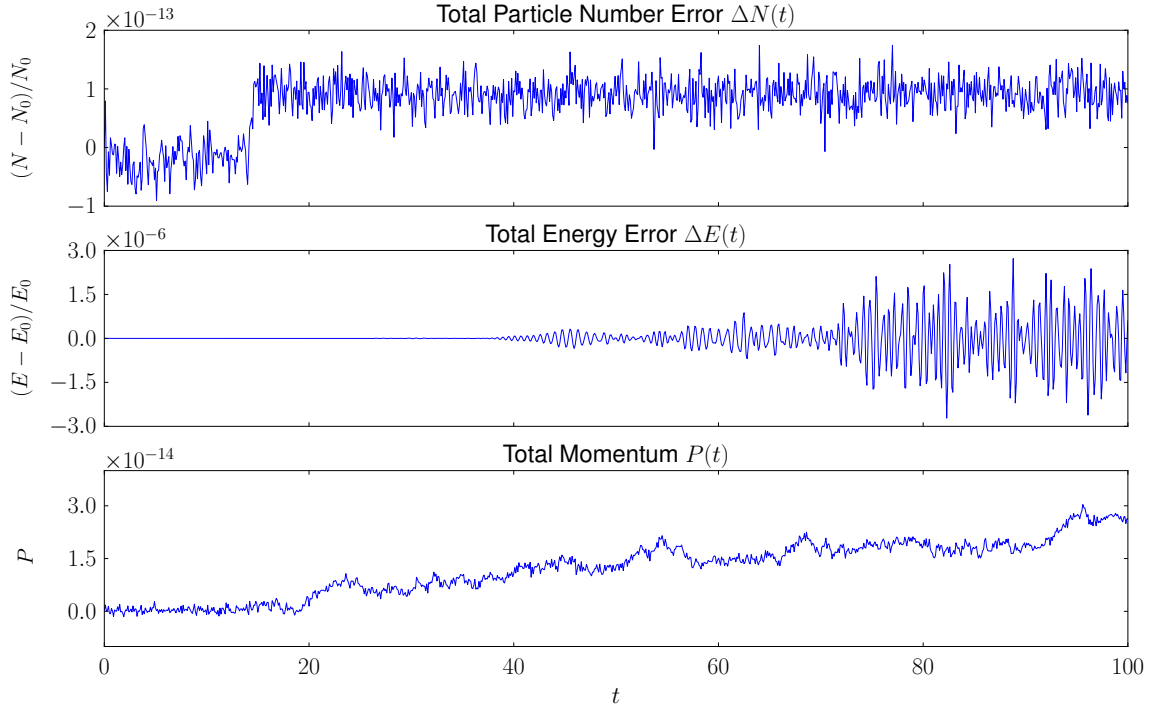

Fig. 5.16.: Linear Landau damping. Evolution of the entropy S.
**Top:** $\nu = 0$**, Middle:** $\nu = 10^{-4}$**, Bottom:** $\nu = 4 \times 10^{-4}$**.**

Fig. 5.17.: Nonlinear Landau damping. Evolution of the L^2 norm.
Top: $\nu = 0$, **Middle:** $\nu = 10^{-4}$, **Bottom:** $\nu = 4 \times 10^{-4}$.

Fig. 5.18.: Nonlinear Landau damping. Evolution of the entropy S.
Top: $\nu = 0$, **Middle:** $\nu = 10^{-4}$, **Bottom:** $\nu = 4 \times 10^{-4}$.

Fig. 5.19.: Twostream instability without collisions. Total particle number and linear momentum are well preserved but energy conservation is violated due to subgrid modes.
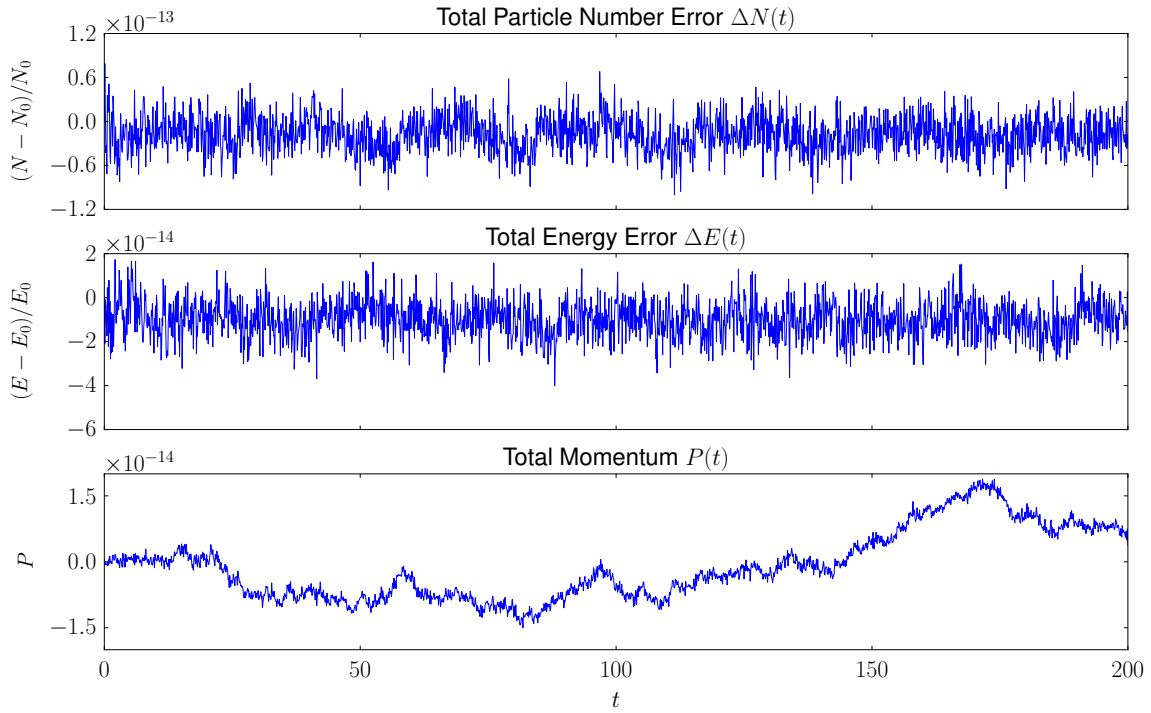
Fig. 5.20.: Twostream instability with collision frequency $\nu = 4 \times 10^{-4}$. Collisions retain exact energy conservation in addition to exact preservation of the total particle number and linear momentum.
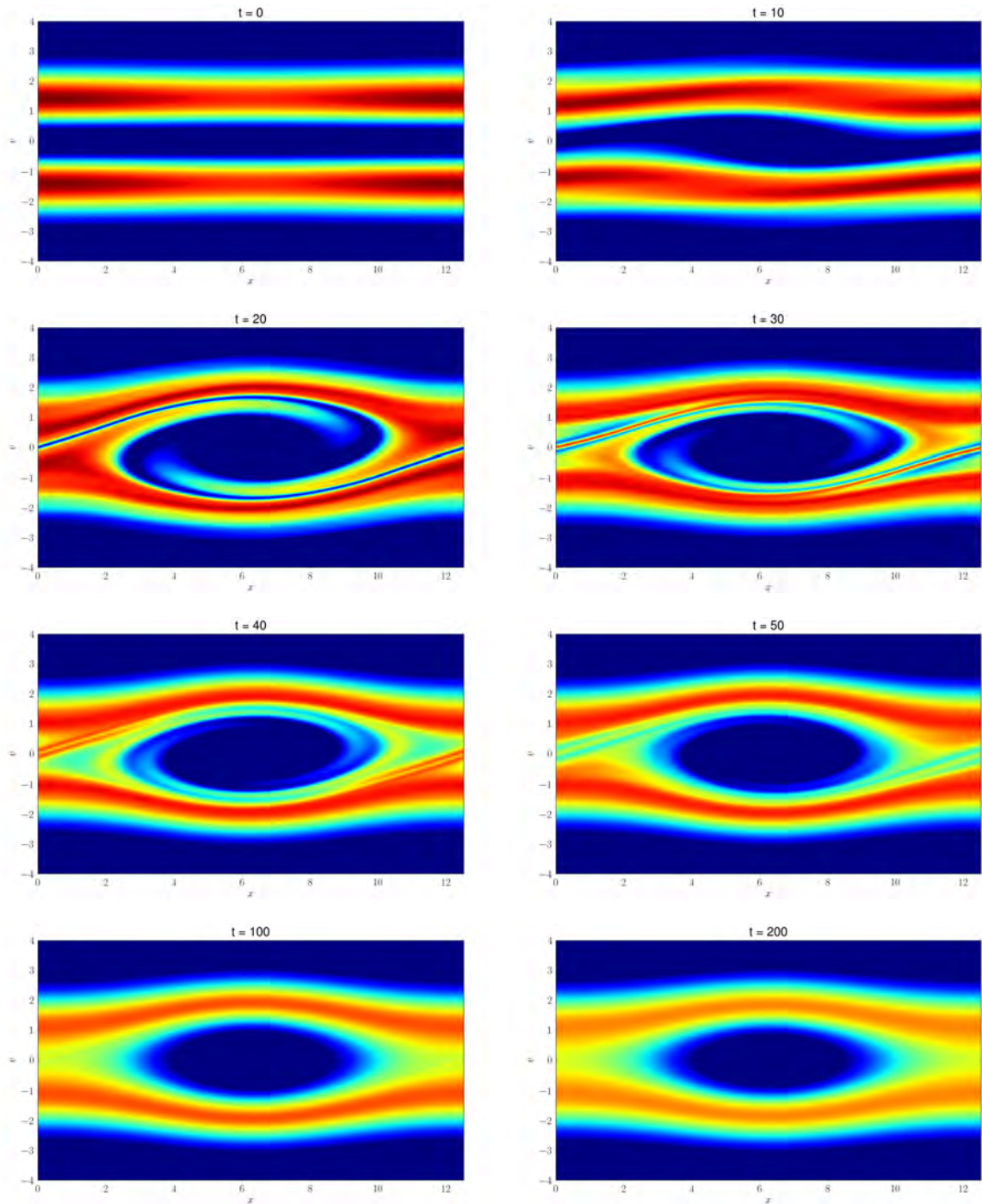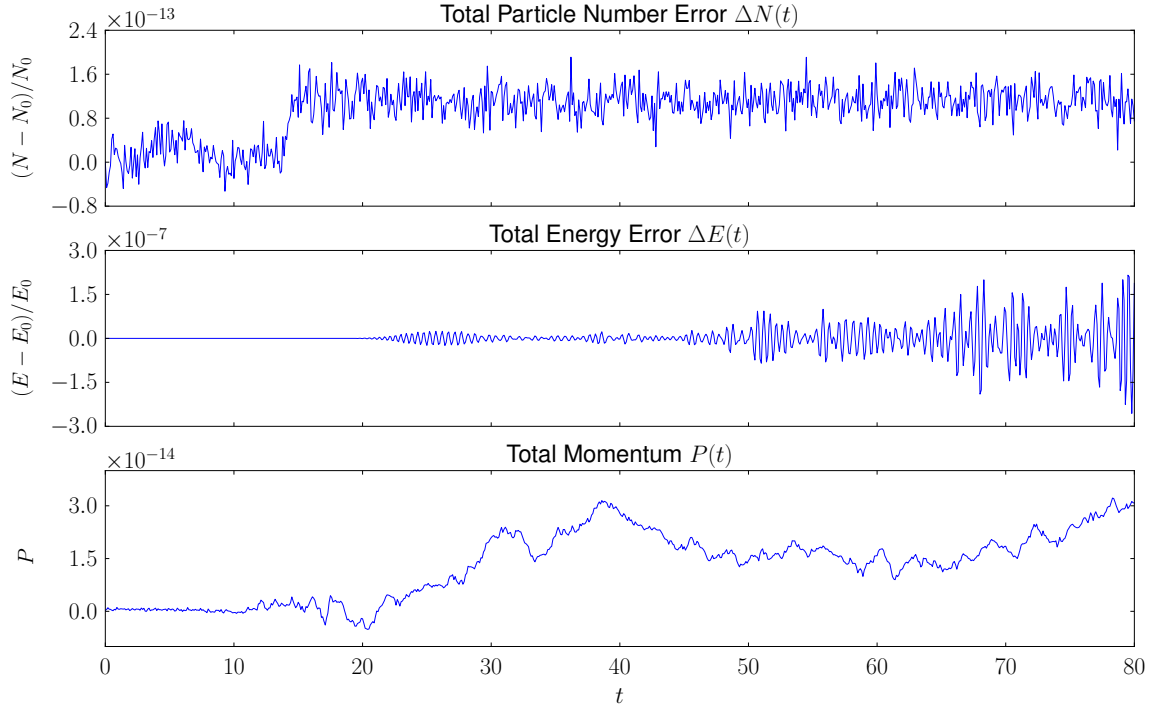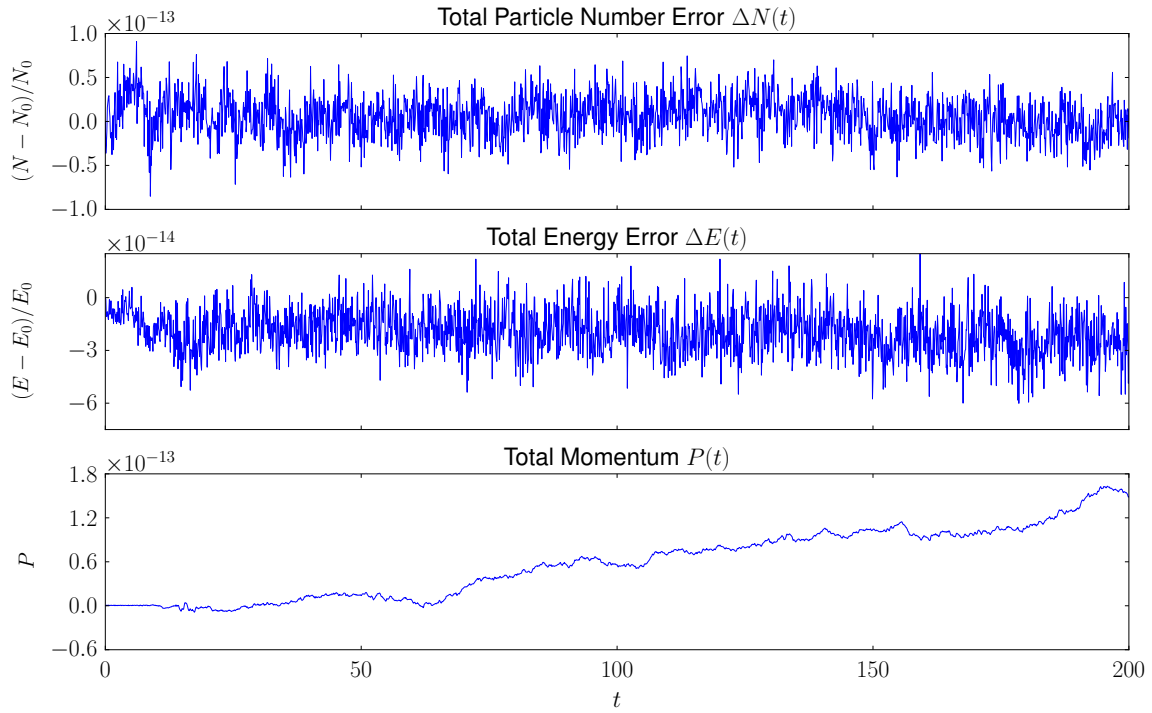
Fig. 5.21.: Twostream instability with collision frequency $\nu = 4 \times 10^{-4}$. Contours of the distribution function in phasespace. Contours are linear and constant.

Fig. 5.22.: Jeans instability without collisions. Total particle number and linear momentum are well preserved, but energy conservation is violated due to subgrid modes.

Fig. 5.23.: Jeans instability with collision frequency $\nu = 4 \times 10^{-4}$. Collisions retain exact energy conservation in addition to exact preservation of the total particle number and linear momentum.
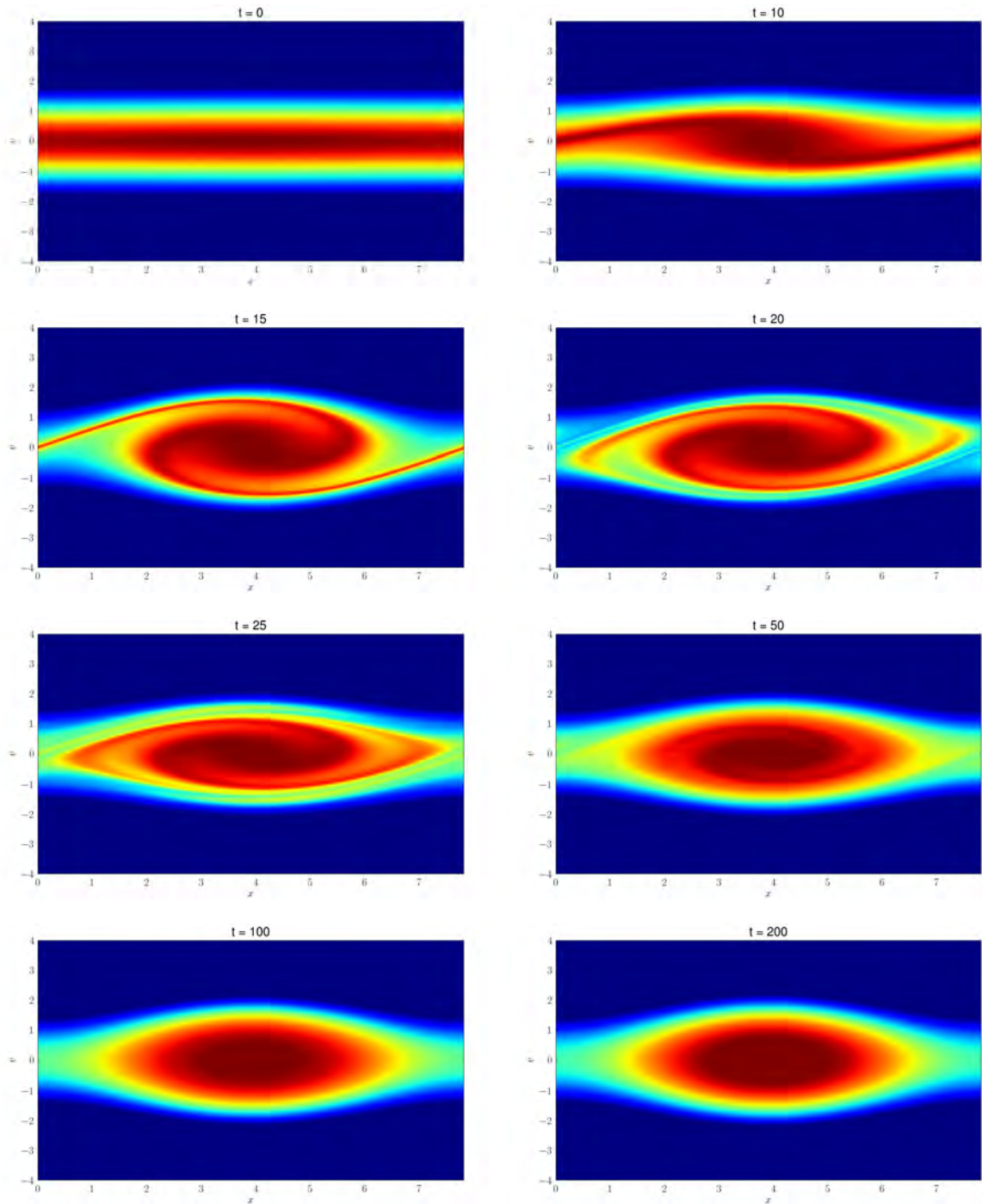
Fig. 5.24.: Jeans instability with collision frequency $\nu = 4 \times 10^{-4}$. Contours of the distribution function in phasespace. Contours are linear and constant.

6. Magnetohydrodynamics

Magnetohydrodynamics (MHD) describes the dynamics of electrically conducting fluids like plasmas or liquid metals. It is one of the most widely applied theories in laboratory as well as astrophysical plasmas physics [115, 48, 18, 34, 41], used to describe macroscopic phenomena like equilibrium states in tokamaks or stellarators, large scale turbulence, or dynamos that generate magnetic fields of stars and planets. The structure of the equations is very similar to hydrodynamics, albeit in MHD the fluid equations are coupled with Maxwell's equations, thereby allowing for an even richer variety of phenomena.

6.1. Incompressible ideal MHD

The equations of magnetohydrodynamics result from the combination of the Navier-Stokes equation for an incompressible fluid,

$$\frac{\partial V}{\partial t} + (V \cdot \nabla)V = -\nabla p + \mu \nabla^2 V + F, \tag{6.1}$$

with Maxwell's equations of electrodynamics. As usual for an incompressible flow, the mass densities are taken constant. V is the fluid velocity, satisfying $\nabla \cdot V = 0$, p is the pressure, F is a force term, and μ is the viscosity of the fluid. The right-hand side is normalised to the density ρ. The fluid carries an electric current density J and is immersed into a magnetic field B, which produces the force $F = J \times B$. Throughout this chapter, we use natural units for electromagnetic quantities, i.e., $\mu_0 = \epsilon_0 = c = 1$.

To obtain an evolution equation for the magnetic field, we combine Faraday's law

$$\frac{\partial B}{\partial t} = -\nabla \times E \tag{6.2}$$

with Ohm's law for a resistive plasma

$$E + V \times B = \eta J, \tag{6.3}$$

η being the plasma resistivity and E the electric field, and Ampere's law

$$J = \nabla \times B \tag{6.4}$$

with the displacement current neglected, to get

$$\frac{\partial B}{\partial t} = \nabla \times (V \times B) - \eta \nabla \times \nabla \times B. \tag{6.5}$$

With the identities

$$\nabla \times \nabla \times B = \nabla (\nabla \cdot B) - \nabla^2 B, \tag{6.6}$$

$$\nabla \times (A \times B) = A (\nabla \cdot B) - B (\nabla \cdot A) + (B \cdot \nabla) A - (A \cdot \nabla) B, \tag{6.7}$$

the induction equation (6.5) becomes

$$\frac{\partial B}{\partial t} = V(\nabla \cdot B) - B(\nabla \cdot V) + (B \cdot \nabla)V - (V \cdot \nabla)B - \eta \nabla(\nabla \cdot B) + \eta \nabla^2 B. \tag{6.8}$$

Both V and B are divergence-free, so the induction equation simplifies to

$$\frac{\partial B}{\partial t} + (V \cdot \nabla)B = (B \cdot \nabla)V + \eta \nabla^2 B. \tag{6.9}$$

The force term F in the Navier-Stokes equation is rewritten as

$$F = J \times B = (\nabla \times B) \times B = (B \cdot \nabla)B - \frac{1}{2}\nabla(B \cdot B). \tag{6.10}$$

Again, we used Ampere's law and the identity

$$B \times (\nabla \times B) = \tfrac{1}{2}\nabla(B \cdot B) - (B \cdot \nabla)B. \tag{6.11}$$

Summing up, we obtain the following system of equations

Incompressible MHD Equations

$$\frac{\partial V}{\partial t} + (V \cdot \nabla)V = (B \cdot \nabla)B + \mu \nabla^2 V - \nabla P, \qquad\qquad \nabla \cdot V = 0, \tag{6.12}$$

$$\frac{\partial B}{\partial t} + (V \cdot \nabla)B = (B \cdot \nabla)V + \eta \nabla^2 B, \qquad\qquad \nabla \cdot B = 0, \tag{6.13}$$

with the generalised pressure P being the sum of the kinetic gas pressure and the magnetic pressure

$$P \equiv p + \frac{1}{2}B^2. \tag{6.14}$$

The first equation (6.12) is called the *momentum equation*, the second equation (6.13) the *induction equation*. Both V and B are divergence-free, V as we are considering an incompressible fluid, and B as there are no magnetic monopoles. But while $\nabla \cdot B = 0$ is implied by the induction equation (provided that the initial magnetic field $B(t = 0)$ is divergence-free), $\nabla \cdot V = 0$ is a dynamical constraint, for which the pressure P is a Lagrange multiplier.

The left-hand sides of (6.12,6.13) represent the advective derivatives of the velocity field V and the magnetic field B. The force term $(B \cdot \nabla)B$ is the directional derivative of B in direction of B. It describes the magnetic tension force, a restoring force that straightens magnetic field lines. This force is perpendicular to B and inversely proportional to the radius of the field line curvature so that the fluid is accelerated towards the local centre of the curvature. As the magnetic field is advected by the fluid, the field lines are dragged with the fluid and thus straightened.

The term $\mu \nabla^2 V$ describes diffusion of the fluid due to viscosity and results from the divergence of the anisotropic part of the stress tensor. As we are dealing with an incompressible fluid, this term only amounts to shear stress.

The pressure gradient ∇p arises from the isotropic part of the stress tensor which describes normal forces. The effect of this term is that fluid flows from regions of high pressure to regions of low pressure. The magnetic pressure has the same effect, fluid flows from regions

of high magnetic pressure to regions of low magnetic pressure. Due to the advection of the magnetic field with the fluid velocity this then leads to the field lines being pushed apart (imagine a bundle of field lines driven apart).

When $\eta = 0$, (6.13) states that the magnetic field is advected with the fluid flow, which implies the conservation of the magnetic flux through a surface moving with the fluid [9]. In a resistive plasma, $\eta \nabla^2 B$ describes diffusive effects, for which the magnetic field are not just dragged along with the field, but are free to change their topology.

In ideal MHD, viscosity and resistivity are neglected, thus $\mu = \eta = 0$. As an effect, the topology of the magnetic field lines is conserved. They are not allowed to open up and reconnect. A property that we would like to maintain on the discrete level. Two important conserved quantities of ideal MHD in two dimensions [9] are the total energy

$$E = \frac{1}{2} \int \left[\|V\|^2 + \|B\|^2 \right] dx\,dy,$$
(6.15)

and cross helicity

$$H = \int V \cdot B \, dx\,dy.$$
(6.16)

Conservation of both quantities are desirable in numerical simulations.

6.1.1. Lie Derivative Formulation

To elucidate the link with work related to ours [43], we rewrite the MHD equations with Lie derivatives, thereby also emphasising the advective character of the equations a bit further. Write the ideal MHD equations in component form, use covariant components in the momentum equation, and add and subtract $V^j \partial_i V_j$ and $B^j \partial_i B_j$

$$\partial_t V_i + V^j \partial_j V_i + V^j \partial_i V_j = B^j \partial_j B_i + B^j \partial_i B_j + V^j \partial_i V_j - B^j \partial_i B_j - \partial_i P,$$
(6.17a)

$$\partial_t B^i + V^j \partial_j B^i - B^j \partial_j V^i = 0,$$
(6.17b)

where V_i and V^i are co- and contravariant components, respectively, and analogously for B_i and B^i. The second and third term on the left-hand side of the momentum equation are the Lie derivative of a 1-form, $V^\flat = V_i \, dx^i$, along its corresponding vector field V. The first two terms on the right-hand side are the Lie derivative of another 1-form, $B^\flat = B_i \, dx^i$, also along its corresponding vector field B. In the induction equation, the second and third term are the Lie derivative of a vector field, B, along V (recall section 2.1.7). The reformulated ideal MHD equations read

Lie Derivative Formulation of Ideal Magnetohydrodynamics

$$\frac{\partial V^\flat}{\partial t} + \pounds_V V^\flat = \pounds_B B^\flat - \mathbf{d}\left[P + \frac{1}{2} \|B\|^2 - \frac{1}{2} \|V\|^2 \right],$$
(6.18)

$$\frac{\partial B}{\partial t} + \pounds_V B = 0.$$
(6.19)

Now the interpretation of the time evolution of the fields becomes even more apparent. The evolution of the fluid velocity is determined by three different mechanisms. The Lie derivative of the velocity along itself describes the change of velocity along the fluid flow. The Lie derivative of the magnetic field along itself describes how the magnetic field changes along field lines. The resulting force, as already discussed, pushes the velocity field

towards the local centre of the magnetic field-line curvature, and thus, as the magnetic field is advected by the fluid flow, balances variations in the magnetic field. The gradient of the pressure states that fluid flows from regions of high (kinetic and magnetic) pressure to regions of low (kinetic and magnetic) pressure. The Lie derivative in the induction equation just states that the magnetic field is advected along the fluid flow.

But what is the meaning of the additional terms? Essentially they just remove physics that we added with the Lie derivative but that was not present in the original equations. The Lie derivative describes all actions that happen along the given vector field, namely a (rigid) translation, a (rigid) rotation, and a deformation, but the original term, e.g. $(u \cdot \nabla)u$, describes only a translation. Let us try to better understand this by having a look at the external derivative of the velocity term (in component form)

$$\frac{1}{2}\,\partial_i \|V\|^2 = V^j\,\partial_i V_j = \frac{1}{2}\left(V^j\,\partial_i V_j + V^j\,\partial_j V_i\right) + \frac{1}{2}\left(V^j\,\partial_i V_j - V^j\,\partial_j V_i\right) = \mathrm{T}_{ij}V^j + \mathrm{S}_{ij}V^j.$$

(6.20)

Here we split $\partial_i V_j$ in a symmetric part T_{ij} and an antisymmetric part S_{ij}. The symmetric tensor T describes the rate of stretching of the vector V along the direction of the eigenvectors or T. As we are discussing incompressible fluids only, the divergence of V vanishes and the trace of T is zero. The fluid element gets deformed but its volume stays constant. Taking the dot-product of a vector with the antisymmetric tensor S describes the rate of rotation of the vector V with angular velocity vector $\frac{1}{2}\xi$, with $\xi = \omega = \nabla \times V$ being the vorticity. Analogously, $\frac{1}{2}\,\partial_i \|B\|^2$ splits into two contributions, describing the rate of stretching and rotation of B, where now $\xi = J = \nabla \times B$.

6.1.2. Potential Formulation in Two Dimensions

Another formulation of magnetohydrodynamics, especially popular in reconnection studies, is the so called potential formulation. Here, the dynamics is not described in terms of the velocity field and the magnetic field, but in terms of their potentials, the streaming function ψ and the magnetic vector potential A.

Potential Formulation of Ideal Magnetohydrodynamics

$$\frac{\partial(\Delta \psi)}{\partial t} + [\psi, \Delta\psi] = [A, \Delta A],$$

(6.21)

$$\frac{\partial A}{\partial t} + [A, \psi] = 0.$$

(6.22)

Here, $[\cdot, \cdot]$ are Poisson brackets with respect to the spatial variables (x, y). The velocity and the magnetic field are computed as

$$V = \nabla \times \psi, \quad B = \nabla \times A, \quad \text{where} \quad \psi = (0, 0, \psi) \quad \text{and} \quad A = (0, 0, A), \quad (6.23)$$

such that in this formulation, the constraints $\nabla \cdot V$ and $\nabla \cdot B$ are automatically fulfilled. With the help of vorticity and current density

$$\omega = \nabla \times V, \quad J = \nabla \times B, \quad \text{where} \quad \omega = (0, 0, \omega) \quad \text{and} \quad J = (0, 0, J), \quad (6.24)$$

we can rewrite the above equations as

$$\frac{\partial \omega}{\partial t} + [\omega, \psi] = [J, A], \qquad\qquad -\Delta \psi = \omega, \qquad\qquad (6.25)$$

$$\frac{\partial J}{\partial t} + [\omega, A] = [J, \psi] + 2[\nabla \psi, \nabla A], \qquad -\Delta A = J, \qquad\qquad (6.26)$$

thereby reducing the highest order of derivatives that appear in the equations from three to two. An important point, since this simplifies the derivation of variational integrators. Looking at this formulation, one might get the expression that we can directly apply the discretisation from the last chapter to this formulation. After all, we already discretised the time derivative, Poisson brackets, and the Laplace operator. Unfortunately, it is not that simple, as in the above potential formulation, additional derivatives appear within the Poisson brackets. To account for those properly, one has to do so at the level of the discrete action. Consequently, one has to repeat the whole derivation, work that is left for future research.

6.2. Variational Discretisation

We will base the derivation of the variational integrator for ideal MHD on the equations (6.12) and (6.13). Their respective components are

$$\frac{\partial V_x}{\partial t} + V_x \partial_x V_x + V_y \partial_y V_x - B_x \partial_x B_x - B_y \partial_y B_x + \partial_x P = 0, \qquad (6.27a)$$

$$\frac{\partial V_y}{\partial t} + V_x \partial_x V_y + V_y \partial_y V_y - B_x \partial_x B_y - B_y \partial_y B_y + \partial_y P = 0, \qquad (6.27b)$$

$$\frac{\partial B_x}{\partial t} + V_x \partial_x B_x + V_y \partial_y B_x - B_x \partial_x V_x - B_y \partial_y V_x = 0, \qquad (6.27c)$$

$$\frac{\partial B_y}{\partial t} + V_x \partial_x B_y + V_y \partial_y B_y - B_x \partial_x V_y - B_y \partial_y V_y = 0. \qquad (6.27d)$$

As we will see in the next section, we have to use a staggered grid approach for the discretisation of the MHD equations. To be able to discretise (6.27) on a single grid cell as depicted in figure 6.1, we have to transform these equations as follows. At first, consider the momentum equations. Like in section (6.1.1), add and subtract $V_y \partial_x V_y$ in the x component, add and subtract $V_x \partial_y V_x$ in the y component, and do the same for the corresponding B terms, such that

$$\partial_t V_x + V_y \left(\partial_y V_x - \partial_x V_y \right) - B_y \left(\partial_y B_x - \partial_x B_y \right) + \partial_x \left[P + \frac{1}{2} \|V\|^2 - \frac{1}{2} \|B\|^2 \right] = 0, \qquad (6.28a)$$

$$\partial_t V_y + V_x \left(\partial_x V_y - \partial_y V_x \right) - B_x \left(\partial_x B_y - \partial_y B_x \right) + \partial_y \left[P + \frac{1}{2} \|V\|^2 - \frac{1}{2} \|B\|^2 \right] = 0. \qquad (6.28b)$$

In the round brackets, we find the rotation operator, which is the z component of the curl, i.e.,

$$(\nabla \times V)_z \equiv \partial_x V_y - \partial_y V_x, \qquad (6.29)$$

and in the square brackets, we find a modified pressure term

$$\tilde{P} = P + \frac{1}{2} \|V\|^2 - \frac{1}{2} \|B\|^2 = p + \frac{1}{2} \|V\|^2, \qquad (6.30)$$

that contains the kinetic instead of the magnetic energy in addition to the gas pressure p. Next, consider the induction equation. By the divergence free constraint of V and B make the replacements $\partial_x V_x = -\partial_y V_y$ in equation (6.28a), $\partial_y V_y = -\partial_x V_x$ in equation (6.28b), and corresponding terms for B, such that

$$\partial_t B_x + V_y\,\partial_y B_x - V_x\,\partial_y B_y + B_x\,\partial_y V_y - B_y\,\partial_y V_x = 0, \tag{6.31a}$$

$$\partial_t B_y + V_x\,\partial_x B_y - V_y\,\partial_x B_x + B_y\,\partial_x V_x - B_x\,\partial_x V_y = 0, \tag{6.31b}$$

which can be condensed into

$$\partial_t B_x + \partial_y\big(V_y B_x - V_x B_y\big) = 0, \tag{6.32a}$$

$$\partial_t B_y + \partial_x\big(V_x B_y - V_y B_x\big) = 0, \tag{6.32b}$$

where the term in brackets is of course the z component of the curl of $V \times B$. To make the following derivations more tractable, we define two operators with components

$$\psi_x(V) \equiv -V_y\big(\partial_x V_y - \partial_y V_x\big), \qquad\qquad \phi_x(V, B) \equiv -\partial_y\big(V_x B_y - V_y B_x\big), \tag{6.33a}$$

$$\psi_y(V) \equiv +V_x\big(\partial_x V_y - \partial_y V_x\big), \qquad\qquad \phi_y(V, B) \equiv +\partial_x\big(V_x B_y - V_y B_x\big), \tag{6.33b}$$

which is the same notation used by Gawlik et al. [43]. With that, the incompressible, ideal MHD equations read

$$\partial_t V + \psi(V) - \psi(B) + \nabla \tilde{P} = 0, \tag{6.34a}$$

$$\partial_t B + \phi(V, B) = 0. \tag{6.34b}$$

The corresponding extended Lagrangian, is readily written upon introducing three auxiliary variables α, β, γ, where α and β are vector fields and γ is a scalar field. The extended Lagrangian for (6.34) reads

$$\mathcal{L}(V, B, \tilde{P}, \alpha, \beta, \gamma) = \alpha \cdot \Big[\partial_t V + \psi(V) - \psi(B) + \nabla \tilde{P}\Big] + \beta \cdot \Big[\partial_t B + \phi(V, B)\Big] + \gamma\Big[\nabla \cdot V\Big]. \tag{6.35}$$

This will be the basis for the derivation of the variational integrator. It is worth mentioning that $\nabla \cdot V = 0$ is a dynamical equation determining the pressure, c.f. comments after equation (6.14). The Ibragimov multipliers α, β and γ correspond to the physical variables V, B and $\tilde{P}$, respectively.

6.2.1. Staggered Grid

Straight forward centred finite difference discretisations of the Navier-Stokes equation, where the components of the velocity vector and the pressure are located at the same grid points, are known to be prone to instabilities (see e.g. Langtangen et al. [69] or McDonough [86]). The pressure often becomes highly oscillatory as a symmetric difference operator, e.g., with stencil $\begin{bmatrix} -1 & 0 & +1 \end{bmatrix}$, annihilates pressures which oscillate between +1 and −1 between neighbouring grid points. This is often referred to as *checkerboarding*.
An efficacious remedy for this problem is the introduction of a staggered grid, where the pressure is located at the centre of a grid cell and the velocity components at the vertices, like it is depicted in figure 6.1a. The location of the physical quantities comes natural when viewed as differential forms. The pressure is a zero-form and is therefore collocated

at the centre of a cell. The velocity (and in two dimensions also the magnetic field) is a one-form and is therefore collocated at the edges of a cell.

This can also be seen by considering the discrete divergence-free constraint of the velocity field

$$\frac{V_x(i, j + \frac{1}{2}, k) - V_x(i - 1, j + \frac{1}{2}, k)}{h_x} + \frac{V_y(i + \frac{1}{2}, j, k) - V_y(i + \frac{1}{2}, j - 1, k)}{h_y} = 0, \tag{6.36}$$

which is defined such that the logical location of the divergence coincides with the location of the pressure. The function of the pressure in incompressible fluid dynamics can be described as taking care of the divergence of the velocity field. By this discretisation, only one pressure point takes care of the divergence of the neighbouring velocity points. And as the divergence is computed by a simple forward finite difference, without symmetric stencil, checkerboarding will not be an issue.

(a) Divergence Constraint　　　　　　　**(b)** Momentum Equation
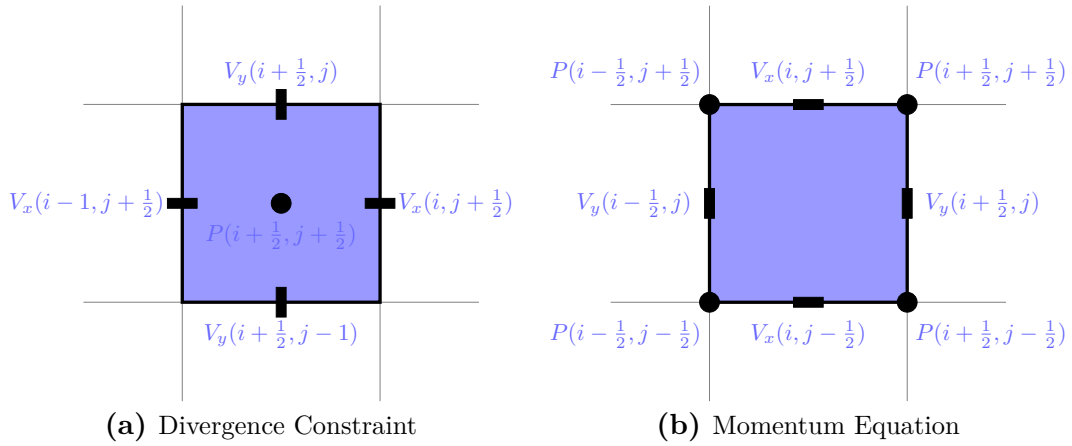
Fig. 6.1.: Staggered Grid. Left: Natural positions for the pressure and the velocity components. Right: Dual grid.

On this grid, the time derivatives are defined pointwise (without spatial averaging)

$$\frac{\partial V_x}{\partial t} \quad \to \quad \left(\frac{\partial V_x}{\partial t}\right)_{i,j+\frac{1}{2},k+\frac{1}{2}} \equiv \frac{V_x(i, j + \frac{1}{2}, k + 1) - V_x(i, j + \frac{1}{2}, k)}{h_t}, \tag{6.37a}$$

$$\frac{\partial V_y}{\partial t} \quad \to \quad \left(\frac{\partial V_y}{\partial t}\right)_{i+\frac{1}{2},j,k+\frac{1}{2}} \equiv \frac{V_y(i + \frac{1}{2}, j, k + 1) - V_y(i + \frac{1}{2}, j, k)}{h_t}. \tag{6.37b}$$

For the spatial derivatives of the vectors, we use a midpoint averaging with respect to time, i.e.,

$$\frac{\partial V_x}{\partial x} \quad \to \quad \left(\frac{\partial V_x}{\partial x}\right)_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} \equiv \frac{1}{2}\left[\frac{V_x(i + 1, j + \frac{1}{2}, k) - V_x(i, j + \frac{1}{2}, k)}{h_x}\right.$$
$$\left. + \frac{V_x(i + 1, j + \frac{1}{2}, k + 1) - V_x(i, j + \frac{1}{2}, k + 1)}{h_x}\right], \tag{6.38a}$$

$$\frac{\partial V_x}{\partial y} \quad \to \quad \left(\frac{\partial V_x}{\partial y}\right)_{i,j,k+\frac{1}{2}} \equiv \frac{1}{2}\left[\frac{V_x(i, j + \frac{1}{2}, k) - V_x(i, j - \frac{1}{2}, k)}{h_y}\right.$$
$$\left. + \frac{V_x(i, j + \frac{1}{2}, k + 1) - V_x(i, j - \frac{1}{2}, k + 1)}{h_y}\right], \tag{6.38b}$$

$$\frac{\partial V_y}{\partial x} \quad\rightarrow\quad \left(\frac{\partial V_y}{\partial x}\right)_{i,j,k+\frac{1}{2}} \quad\equiv\quad \frac{1}{2}\left[\frac{V_y(i+\frac{1}{2},j,k)-V_y(i-\frac{1}{2},j,k)}{h_x}\right.$$
$$\left.+\frac{V_y(i+\frac{1}{2},j,k+1)-V_y(i-\frac{1}{2},j,k+1)}{h_x}\right],$$

$$\frac{\partial V_y}{\partial y} \quad\rightarrow\quad \left(\frac{\partial V_y}{\partial y}\right)_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} \quad\equiv\quad \frac{1}{2}\left[\frac{V_y(i+\frac{1}{2},j+1,k)-V_y(i+\frac{1}{2},j,k)}{h_y}\right.$$
$$\left.+\frac{V_y(i+\frac{1}{2},j+1,k+1)-V_y(i+\frac{1}{2},j,k+1)}{h_y}\right],$$

Note, that the x derivative of V_x and the y derivative of V_y are defined on the grid in figure 6.1a, while the y derivative of V_x and the x derivative of V_y are defined on the dual grid in figure 6.1b. The indices of the derivatives denote the logical collocation of the derivative, which is always the cell centre.

Derivatives of the pressure can only be defined on the dual grid, figure 6.1b. They are naturally defined on the edges of the cells. The staggering approach is applied to P also with respect to time, i.e., the pressure nodes are $(i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2})$. Taking all of this into account, we define

$$\frac{\partial P}{\partial x} \quad\rightarrow\quad \left(\frac{\partial P}{\partial x}\right)_{i,j+\frac{1}{2},k+\frac{1}{2}} \equiv \frac{P(i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2})-P(i-\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2})}{h_x}, \tag{6.40a}$$

$$\frac{\partial P}{\partial y} \quad\rightarrow\quad \left(\frac{\partial P}{\partial y}\right)_{i+\frac{1}{2},j,k+\frac{1}{2}} \equiv \frac{P(i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2})-P(i+\frac{1}{2},j-\frac{1}{2},k+\frac{1}{2})}{h_y}. \tag{6.40b}$$

Averages are needed only on the dual grid, so we are defining them only there and only for the vector fields. For V, the averaging is applied with respect to both, space and time,

$$\langle V_x\rangle_{i,j,k+\frac{1}{2}} \equiv \frac{V_x(i,j-\frac{1}{2},k)+V_x(i,j+\frac{1}{2},k)}{4} + \frac{V_x(i,j-\frac{1}{2},k+1)+V_x(i,j+\frac{1}{2},k+1)}{4},$$
$$\tag{6.41a}$$

$$\langle V_y\rangle_{i,j,k+\frac{1}{2}} \equiv \frac{V_y(i-\frac{1}{2},j,k)+V_y(i+\frac{1}{2},j,k)}{4} + \frac{V_y(i-\frac{1}{2},j,k+1)+V_y(i+\frac{1}{2},j,k+1)}{4},$$
$$\tag{6.41b}$$

but as α will be collocated at $k+\frac{1}{2}$ (see comment in the next section), its averages do not feature a time average,

$$\langle \alpha_x\rangle_{i,j,k+\frac{1}{2}} \equiv \frac{\alpha_x(i,j-\frac{1}{2},k+\frac{1}{2})+\alpha_x(i,j+\frac{1}{2},k+\frac{1}{2})}{2}, \tag{6.42a}$$

$$\langle \alpha_y\rangle_{i,j,k+\frac{1}{2}} \equiv \frac{\alpha_y(i-\frac{1}{2},j,k+\frac{1}{2})+\alpha_y(i+\frac{1}{2},j,k+\frac{1}{2})}{2}. \tag{6.42b}$$

6.2.2. Navier-Stokes Equation

We start the derivation of the variational integrator by restricting our attention to the incompressible Navier-Stokes equation

$$\partial_t V + \psi(V) + \nabla P = 0, \qquad\qquad \nabla \cdot V = 0, \tag{6.43}$$

neglecting the force term through the magnetic field. The generalisation to magnetohydrodynamics is straight forward, as the magnetic field appears with the same advection term $\psi(B)$ as the velocity field, and the analytical expression of the pressure P in terms of v and B^2 does not play any role. The action integral of the extended Lagrangian (6.35), reduced to this subsystem, is

$$\mathcal{A} = \int \left[... + \alpha \cdot \left[\partial_t V + \psi(V) + \nabla P \right] + \gamma \left[\nabla \cdot V \right] + ... \right] dt\, dx\, dy. \tag{6.44}$$

To be able to discretise all of the derivatives in the first term of the Lagrangian, we have to switch to the dual grid, as depicted in figure 6.1b. The time derivatives are approximated using the trapezoidal rule,

$$\alpha_x \left(\partial_t V_x \right) \quad \rightarrow \quad \alpha_x(i, j - \tfrac{1}{2}, k + \tfrac{1}{2}) \left(\frac{\partial V_x}{\partial t} \right)_{i, j - \frac{1}{2}, k + \frac{1}{2}} + \alpha_x(i, j + \tfrac{1}{2}, k + \tfrac{1}{2}) \left(\frac{\partial V_x}{\partial t} \right)_{i, j + \frac{1}{2}, k + \frac{1}{2}}, \tag{6.45a}$$

$$\alpha_y \left(\partial_t V_y \right) \quad \rightarrow \quad \alpha_y(i - \tfrac{1}{2}, j, k + \tfrac{1}{2}) \left(\frac{\partial V_y}{\partial t} \right)_{i - \frac{1}{2}, j, k + \frac{1}{2}} + \alpha_y(i + \tfrac{1}{2}, j, k + \tfrac{1}{2}) \left(\frac{\partial V_y}{\partial t} \right)_{i + \frac{1}{2}, j, k + \frac{1}{2}}. \tag{6.45b}$$

Here, we are exploiting the same ideas, explained in section 5.3, but here that is implemented directly into the Lagrangian rather than in the final integrator. We omit the time averaging of α, which implies that α is collocated at $k + \tfrac{1}{2}$, just as the time derivative. We use a trapezoidal approximation to avoid spatial averaging of the time derivatives in the resulting scheme, as that might again lead to grid oscillations (checkerboarding), this time in the velocity field. We apply the same approximation to the pressure gradient term, for the same reason, i.e., to avoid oscillations, and as the structure of the terms is identical (e.g., $\partial_t V_x$ and the x component of the pressure gradient $\partial_x P$ are both objects collocated at the same logical position),

$$\alpha_x \left(\partial_x P \right) \quad \rightarrow \quad \alpha_x(i, j - \tfrac{1}{2}, k + \tfrac{1}{2}) \left(\frac{\partial P}{\partial x} \right)_{i, j - \frac{1}{2}, k + \frac{1}{2}} + \alpha_x(i, j + \tfrac{1}{2}, k + \tfrac{1}{2}) \left(\frac{\partial P}{\partial x} \right)_{i, j + \frac{1}{2}, k + \frac{1}{2}}, \tag{6.46}$$

$$\alpha_y \left(\partial_y P \right) \quad \rightarrow \quad \alpha_y(i - \tfrac{1}{2}, j, k + \tfrac{1}{2}) \left(\frac{\partial P}{\partial y} \right)_{i - \frac{1}{2}, j, k + \frac{1}{2}} + \alpha_y(i + \tfrac{1}{2}, j, k + \tfrac{1}{2}) \left(\frac{\partial P}{\partial y} \right)_{i + \frac{1}{2}, j, k + \frac{1}{2}}. \tag{6.47}$$

As previously mentioned, the pressure is collocated at $k + \tfrac{1}{2}$, such that no time average of P is needed. The ψ operator (6.33) is discretised by a midpoint approximation, both with respect to space and time, i.e.,

$$\alpha_x \psi_x(V) \quad \rightarrow \quad - \langle \alpha_x \rangle_{i,j,k+\frac{1}{2}} \langle V_x \rangle_{i,j,k+\frac{1}{2}} \left[\left(\frac{\partial V_y}{\partial x} \right)_{i,j,k+\frac{1}{2}} - \left(\frac{\partial V_x}{\partial y} \right)_{i,j,k+\frac{1}{2}} \right], \tag{6.48}$$

$$\alpha_y \psi_y(V) \quad \rightarrow \quad + \langle \alpha_y \rangle_{i,j,k+\frac{1}{2}} \langle V_y \rangle_{i,j,k+\frac{1}{2}} \left[\left(\frac{\partial V_y}{\partial x} \right)_{i,j,k+\frac{1}{2}} - \left(\frac{\partial V_x}{\partial y} \right)_{i,j,k+\frac{1}{2}} \right]. \tag{6.49}$$

As, e.g., α_x and V_y are objects collocated at different logical positions, they cannot be multiplied directly, but products can only be computed of their averages, which are collocated at the cell centres. Therefore, we have to use a midpoint approximation in this term. The rotation is logically located at the centre of the cell, therefore posing no problems.

The discretisation of the second term in (6.44) is implemented on the grid in figure 6.1a. Recognising that γ is a scalar field and thus collocated at the same position as the pressure, the discretisation follows directly from (6.36), i.e.,

$$\gamma\left(\nabla\cdot V\right) \quad\rightarrow\quad \gamma(i+\tfrac{1}{2},j+\tfrac{1}{2},k)\left[\left(\frac{\partial V_x}{\partial y}\right)_{i,j-\frac{1}{2},\bar{k}}+\left(\frac{\partial V_x}{\partial y}\right)_{i,\bar{j},\bar{k}}\right]. \tag{6.50}$$

Summing up all contributions gives the discrete Lagrangian.

6.2.3. Induction Equation

Now we consider those terms of the action that will yield the induction equation.

$$\mathcal{A} = \int\left[\dots+\beta\cdot\left[\partial_t B+\phi(V,B)\right]+\dots\right]dt\,dx\,dy. \tag{6.51}$$

To find a discretisation of the operator $\phi(V,B)$ on a single grid cell, we have to first do a partial integration, such that

$$\mathcal{A} = \int\left[\dots+\beta_x\left[\partial_t B_x-\partial_y(V_x B_y-V_y B_x)\right]+\beta_y\left[\partial_t B_y+\partial_x(V_x B_y-V_y B_x)\right]+\dots\right]dt\,dx\,dy$$

$$= \int\left[\dots+\beta_x\,\partial_t B_x+(\partial_y\beta_x)(V_x B_y-V_y B_x)\right.$$

$$\left.+\beta_y\,\partial_t B_y-(\partial_x\beta_y)(V_x B_y-V_y B_x)+\dots\right]dt\,dx\,dy. \tag{6.52}$$

The discretisation of the time derivative is the same as in the case of the momentum equation, i.e., using a trapezoidal rule,

$$\beta_x\left(\partial_t B_x\right) \rightarrow \beta_x(i,j-\tfrac{1}{2},k+\tfrac{1}{2})\left(\frac{\partial B_x}{\partial t}\right)_{i,j-\frac{1}{2},k+\frac{1}{2}}+\beta_x(i,j+\tfrac{1}{2},k+\tfrac{1}{2})\left(\frac{\partial B_x}{\partial t}\right)_{i,j+\frac{1}{2},k+\frac{1}{2}}, \tag{6.53}$$

$$\beta_y\left(\partial_t B_y\right) \rightarrow \beta_y(i-\tfrac{1}{2},j,k+\tfrac{1}{2})\left(\frac{\partial B_y}{\partial t}\right)_{i-\frac{1}{2},j,k+\frac{1}{2}}+\beta_y(i+\tfrac{1}{2},j,k+\tfrac{1}{2})\left(\frac{\partial B_y}{\partial t}\right)_{i+\frac{1}{2},j,k+\frac{1}{2}}. \tag{6.54}$$

The factors of the operator ϕ are collocated at different positions of the grid, e.g., in the first equation, $\partial_y\beta_x$ is collocated at (i,j), V_x and B_x are collocated at $(i,j+\tfrac{1}{2})$, and V_y and B_y are collocated at $(i+\tfrac{1}{2},j)$. Therefore, we have to use a midpoint rule, i.e.,

$$(\partial_y\beta_x)(V_x B_y-V_y B_x) \rightarrow \left(\frac{\partial\beta_x}{\partial y}\right)_{i,j,k+\frac{1}{2}}\left[\langle V_x\rangle_{i,j,k+\frac{1}{2}}\langle B_y\rangle_{i,j,k+\frac{1}{2}}-\langle V_y\rangle_{i,j,k+\frac{1}{2}}\langle B_x\rangle_{i,j,k+\frac{1}{2}}\right], \tag{6.55}$$

$$(\partial_x\beta_y)(V_x B_y-V_y B_x) \rightarrow \left(\frac{\partial\beta_y}{\partial x}\right)_{i,j,k+\frac{1}{2}}\left[\langle V_x\rangle_{i,j,k+\frac{1}{2}}\langle B_y\rangle_{i,j,k+\frac{1}{2}}-\langle V_y\rangle_{i,j,k+\frac{1}{2}}\langle B_x\rangle_{i,j,k+\frac{1}{2}}\right]. \tag{6.56}$$

With that, we have all the ingredients for a complete discretisation of the action integral corresponding to (6.35).

6.2.4. Variational Integrator

Computing the variation of the discrete action with respect to α, β and γ, we obtain the discrete ideal MHD equations

$$\frac{V_x^{i,j+1/2,k+1} - V_x^{i,j+1/2,k}}{h_t} + \psi_x^{i,j+1/2}\left(\frac{V^k + V^{k+1}}{2}\right) - \psi_x^{i,j+1/2}\left(\frac{B^k + B^{k+1}}{2}\right)$$
$$+ \frac{P^{i+1/2,j+1/2,k+1/2} - P^{i-1/2,j+1/2,k+1/2}}{h_x} = 0, \tag{6.57}$$

$$\frac{V_y^{i+1/2,j,k+1} - V_y^{i+1/2,j,k}}{h_t} + \psi_y^{i+1/2,j}\left(\frac{V^k + V^{k+1}}{2}\right) - \psi_y^{i+1/2,j}\left(\frac{B^k + B^{k+1}}{2}\right)$$
$$+ \frac{P^{i+1/2,j+1/2,k+1/2} - P^{i+1/2,j-1/2,k+1/2}}{h_y} = 0, \tag{6.58}$$

$$\frac{B_x^{i,j+1/2,k+1} - B_x^{i,j+1/2,k}}{h_t} + \phi_x^{i,j+1/2}\left(\frac{V^k + V^{k+1}}{2}, \frac{B^k + B^{k+1}}{2}\right) = 0, \tag{6.59}$$

$$\frac{B_y^{i+1/2,j,k+1} - B_y^{i+1/2,j,k}}{h_t} + \phi_y^{i+1/2,j}\left(\frac{V^k + V^{k+1}}{2}, \frac{B^k + B^{k+1}}{2}\right) = 0, \tag{6.60}$$

$$\left[\frac{V_x^{i+1,j+1/2,k} - V_x^{i,j+1/2,k}}{h_x} + \frac{V_x^{i+1,j+1/2,k+1} - V_x^{i,j+1/2,k+1}}{h_x}\right.$$
$$\left. + \frac{V_y^{i+1/2,j+1,k} - V_y^{i+1/2,j,k}}{h_y} + \frac{V_y^{i+1/2,j+1,k+1} - V_y^{i+1/2,j,k+1}}{h_y}\right] = 0. \tag{6.61}$$

with the discrete operators defined by

$$\psi_x^{i,j+1/2}(V^{k+1/2}) = -\langle V_x\rangle_{i,j,k+\frac{1}{2}}\left[\left(\frac{\partial V_y}{\partial x}\right)_{i,j,k+\frac{1}{2}} - \left(\frac{\partial V_x}{\partial y}\right)_{i,j,k+\frac{1}{2}}\right]$$
$$-\langle V_x\rangle_{i,j+1,k+\frac{1}{2}}\left[\left(\frac{\partial V_y}{\partial x}\right)_{i,j+1,k+\frac{1}{2}} - \left(\frac{\partial V_x}{\partial y}\right)_{i,j+1,k+\frac{1}{2}}\right], \tag{6.62}$$

$$\psi_y^{i+1/2,j}(V^{k+1/2}) = +\langle V_y\rangle_{i,j,k+\frac{1}{2}}\left[\left(\frac{\partial V_y}{\partial x}\right)_{i,j,k+\frac{1}{2}} - \left(\frac{\partial V_x}{\partial y}\right)_{i,j,k+\frac{1}{2}}\right]$$
$$+\langle V_y\rangle_{i+1,j,k+\frac{1}{2}}\left[\left(\frac{\partial V_y}{\partial x}\right)_{i+1,j,k+\frac{1}{2}} - \left(\frac{\partial V_x}{\partial y}\right)_{i+1,j,k+\frac{1}{2}}\right], \tag{6.63}$$

and

$$\phi_x^{i,j+1/2}(V^{k+1/2}, B^{k+1/2}) = +\left[\langle V_x \rangle_{i,j+1,k+\frac{1}{2}} \langle B_y \rangle_{i,j+1,k+\frac{1}{2}} - \langle V_y \rangle_{i,j+1,k+\frac{1}{2}} \langle B_x \rangle_{i,j+1,k+\frac{1}{2}} \right]$$
$$- \left[\langle V_x \rangle_{i,j,k+\frac{1}{2}} \langle B_y \rangle_{i,j,k+\frac{1}{2}} - \langle V_y \rangle_{i,j,k+\frac{1}{2}} \langle B_x \rangle_{i,j,k+\frac{1}{2}} \right],$$
(6.64)

$$\phi_y^{i+1/2,j}(V^{k+1/2}, B^{k+1/2}) = -\left[\langle V_x \rangle_{i+1,j,k+\frac{1}{2}} \langle B_y \rangle_{i+1,j,k+\frac{1}{2}} - \langle V_y \rangle_{i+1,j,k+\frac{1}{2}} \langle B_x \rangle_{i+1,j,k+\frac{1}{2}} \right]$$
$$+ \left[\langle V_x \rangle_{i,j,k+\frac{1}{2}} \langle B_y \rangle_{i,j,k+\frac{1}{2}} - \langle V_y \rangle_{i,j,k+\frac{1}{2}} \langle B_x \rangle_{i,j,k+\frac{1}{2}} \right].$$
(6.65)

This discretisation of the operators ψ and ϕ is the very same as those found by Gawlik et al. [43] and Liu and Wang [78]. However, Gawlik et al. resolve the nonlinearity in a different way, such that in their scheme, cross helicity is preserved exactly, but the energy error shows a monotonic growth. In our scheme, energy is preserved exactly as is cross helicity (up to machine accuracy). They follow a different but related path in their derivation, based on Euler-Poincaré reduction. The crucial difference is, that in their method, only the velocity field is treated variational, and the magnetic field is treated as a quantity advected with the velocity field, whereas in our method, the velocity field and the magnetic field are treated on equal footing, fully variationally. The scheme of Liu and Wang uses an explicit Runge-Kutta method for time integration, so that conservation laws are broken in long time simulations.

Note the absence of any spatial averaging of the time derivatives and the pressure gradient. This is on purpose, as we wanted to prevent the emergence of grid-scale oscillations in the fields by the introduction of the staggered grid.

6.3. Numerical Examples

In this section, we consider four quite different examples of ideal magnetohydrodynamics problems taken from the previous literature [33, 42, 43]: Alfvénic waves, the passive advection of a magnetic loop, the development of current sheaths in an Orszag-Tang vortex, and the development of magnetic islands along a current sheath.

6.3.1. Diagnostics

Energy

The total energy of the system is the sum of kinetic energy and magnetic energy, which are computed by

$$E_{\text{kin},d}^k = h_x\, h_y\, \frac{1}{2} \sum_{i,j} V_x^2(i, j + 1/2, k) + h_x\, h_y\, \frac{1}{2} \sum_{i,j} V_y^2(i + 1/2, j, k)$$
(6.66)

$$E_{\text{mag},d}^k = h_x\, h_y\, \frac{1}{2} \sum_{i,j} B_x^2(i, j + 1/2, k) + h_x\, h_y\, \frac{1}{2} \sum_{i,j} B_y^2(i + 1/2, j, k).$$
(6.67)

As there is no dissipation term in the ideal MHD equations, the total energy is always preserved.

Cross Helicity

The cross helicity is the integral of the scalar product of the velocity and magnetic fields,

$$
\begin{aligned}
H_d^k = h_x\,h_y\,\frac{1}{2}\sum_{i,j} V_x(i,j+1/2,k)\,B_x(i,j+1/2,k) \\
+ h_x\,h_y\,\frac{1}{2}\sum_{i,j} V_y(i+1/2,j,k)\,B_y(i+1/2,j,k).
\end{aligned} \tag{6.68}
$$

In ideal MHD, the parallel components of the velocity and magnetic fields do not interact, such that the integral of their product over the spatial domain stays constant.

Vector Potential

In two dimensions, the magnetic field is given by

$$
B_x = \partial_y A \qquad \text{and} \qquad B_y = -\partial_x A. \tag{6.69}
$$

The magnetic potential is collocated at the cell centres of figure 6.1b. Therefore these equations discretise as

$$
B_x(i,j+\tfrac{1}{2}) = \frac{A(i,j+1)-A(i,j)}{h_y} \quad \text{and} \quad B_y(i,j+\tfrac{1}{2}) = -\frac{A(i+1,j)-A(i,j)}{h_x}. \tag{6.70a}
$$

The potential field can be reconstructed by fixing $A(1,1)$ and looping over the whole grid, computing

$$
A(i,j+1) = A(i,j) + h_y\,B_x(i,j+\tfrac{1}{2}) \quad \text{and} \quad A(i+1,j) = A(i,j) - h_x\,B_x(i,j+\tfrac{1}{2}), \tag{6.71}
$$

using the first equation to compute columns and the second to jump between rows, or the other way around. To which value $A(1,1)$ is fixed is not important as A is determined only up to a constant. The contour lines of the magnetic potential A correspond to field lines of the magnetic field B. Hence, A is an important diagnostic.

Current Density

The current is given by the curl of the magnetic field, or in two dimensions by the z component of the curl. The discrete version of that is
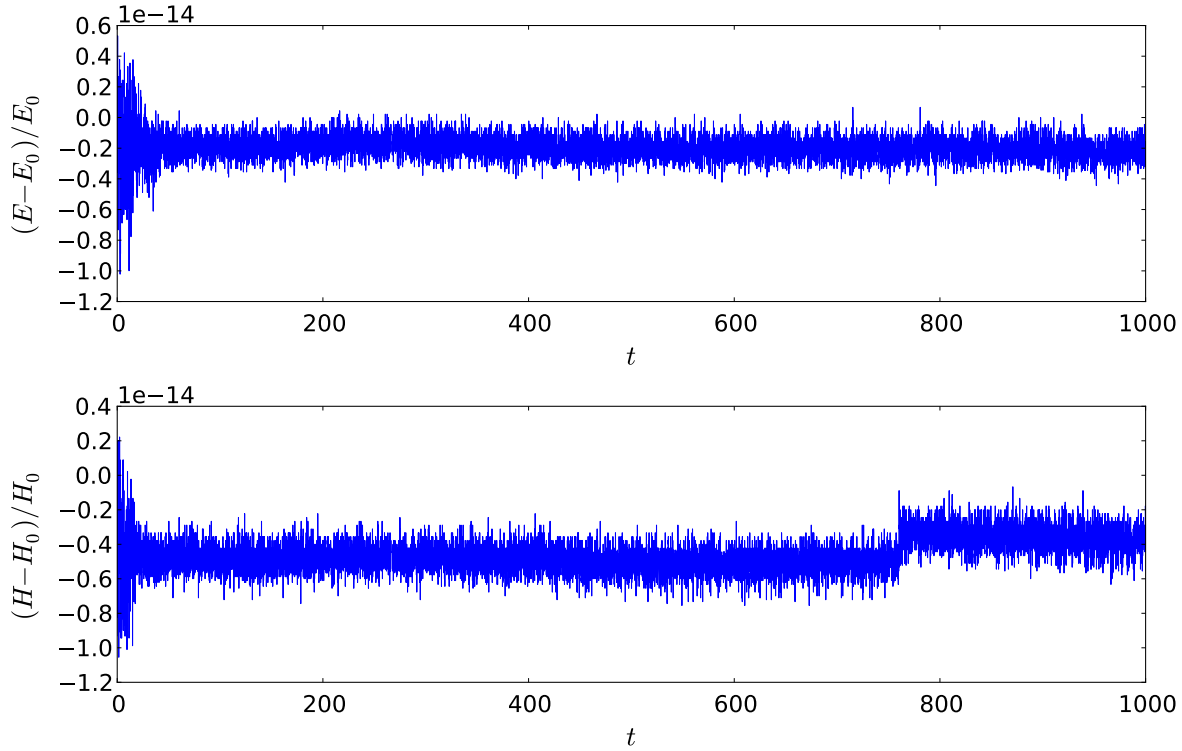
$$
J(i,j,k) = \frac{B_y(i+1/2,j,k)-B_y(i-1/2,j,k)}{h_x} - \frac{B_x(i,j+1/2,k)-B_x(i,j-1/2,k)}{h_y}. \tag{6.72}
$$

As the vector potential, the current is collocated at cell centres.

6.3.2. Alfvén Waves

In the first example, we consider a travelling Alfvén wave, initialised by

$$
V_x = 0, \qquad V_y = V_0 \sin(\pi x), \qquad B_x = B_0, \qquad B_y = B_0 \sin(\pi x), \qquad P = 0.1,
$$

Fig. 6.2.: Conservation of energy and cross helicity for a travelling Alfvén wave.

with $V_0 = 1$ and $B_0 = 1$. The simulation domain is $[0, 2] \times [0, 2]$ with periodic boundaries and a resolution of $n_x \times n_y = 30 \times 30$. The timestep is $h_t = 0.1$ in units of the Alfvén time (i.e., the Alfvén velocity is one).

Although this example is rather simple, the results of our variational integrator are already remarkable. Figure 6.2 shows the time traces of the errors in the total energy and the cross helicity. For most of the simulation, the amplitude of the oscillations is $\mathcal{O}(10^{-15})$, i.e., machine precision. The sudden jump in the error of the cross helicity at about $t = 800$ might be explained by the residual of the Newton iteration being slightly larger for some timesteps than it is during the rest of the simulation. We want to stress that during the runtime of 1000 characteristic times there is no change in the energy within the machine accuracy. It appears as if the Alfvén wave would continue travelling practically forever. It also worth mentioning, that this is a fully nonlinear wave, i.e., the amplitudes of the perturbations of the magnetic field as well as the velocity field are $\mathcal{O}(1)$.

6.3.3. Loop Advection

We now consider a case with very small magnetic field, such that the momentum and induction equations are almost decoupled and the magnetic field is passively advected by the fluid. The initial conditions are

$$V_x = V_0 \cos(\theta), \qquad V_y = V_0 \sin(\theta), \qquad B_x = \ \partial_y A, \qquad B_y = -\partial_x A, \qquad P = 1.0,$$

with

$$A = A_0 \left(R - \sqrt{x^2 + y^2} \right), \qquad V_0 = \sqrt{5}, \qquad \theta = \tan^{-1}(0.5), \qquad A_0 = 10^{-3}, \qquad R = 0.3.$$

The spatial domain is $[-1, +1] \times [-0.5, +0, 5]$ with periodic boundaries. We consider two resolutions, 128×64 and 256×128, and a timestep $h_t = 0.01$.

The problem is setup such that the loop should return to its initial position after integer times $t = 1, 2, 3, \ldots$. In figure 6.5 it can be seen that this is initially the case, but after some time, the loop gets deformed, such that its centre is slightly displaced from its initial position at integer times. Note, that in figure 6.5 only contours in the interval $[1.2 \times 10^{-4}, 3.0 \times 10^{-4}]$ are plotted. Otherwise, the plot would be too polluted since the velocity and magnetic field are of course not entirely decoupled but back react onto each other. A possible reason for this behaviour is that our scheme is only of second order and therefore not extremely accurate. Higher order variational integrators might achieve better accuracy, but this is a topic left for future research. Nevertheless, the energy and cross helicity are preserved optimally throughout the whole simulation (figures 6.3 and 6.4). Interestingly, the errors are somewhat larger, when the resolution is increased. The amplitude of the energy error, for example, is about 5×10^{-16} on a grid of 128×64 points, but 3×10^{-15} on a grid of 256×128 points. First of all, one must not forget how close to machine precision these values are, so slight deviations for similar but not identical simulations are no surprise. However, this issue can be explained more specifically by the larger number of degrees of freedom in the case with higher resolution that results in a stronger error accumulation.

6.3.4. Orszag-Tang Vortex

Next we consider the evolution of current sheets in an Orszag-Tang vortex, where we use the same initial conditions as Cordoba and Marliani [33], namely

$$V_x = \partial_y \psi, \qquad V_y = -\partial_x \psi, \qquad B_x = \partial_y A, \qquad B_y = -\partial_x A, \qquad P = 0.1.$$

with

$$\psi = 2\sin(y) - 2\cos(x), \qquad A = \cos(2y) - 2\cos(x).$$

The spatial domain is $[0, 2\pi] \times [0, 2\pi]$ with periodic boundaries. We consider two resolutions, 64×64 and 128×128, and a timestep $h_t = 0.01$.

The Orszag Tang vortex constitutes a turbulent setting that leads to the growths of current sheaths. These are areas of large current density due to a change of sign in the magnetic field. In figure 6.8, the current density computed by (6.72) is plotted. The current sheaths are located in those parts of the plot where the colour changes from blue to red within a small region. Starting from about $t = 60$, the simulation is under resolved and subgrid modes start to appear. The situation is similar using 128×128 grid points. Note that in the original work, Cordoba and Marliani [33] use an adaptive mesh refinement approach with an initial resolution of 1024×1024 points. The important observation is, that even with low resolution energy and cross helicity are preserved optimally (see figures 6.6 and 6.7). During the simulation, only a slight growth in the errors is observed, probably due to the subgrid modes, but the errors stay $\mathcal{O}(10^{-15})$ for the whole time. As before, we observe that the energy error is slightly larger in the case of higher resolution. As before, this is probably due to the higher number of degrees of freedom, and therefore increased error accumulation.

6.3.5. Current Sheath

In the following, we consider as initial conditions for the magnetic field three different current sheath models that appear in reconnection studies: a sharp jump of the magnetic

field

$$B_y^{\text{sharp}} = \begin{cases} +B_0 & x < x_1 \\ -B_0 & x_1 \le x \le x_2 \\ +B_0 & x > x_2 \end{cases} \qquad \text{with} \qquad x_1 = 0.5, \qquad x_2 = 1.5,$$

a tanh profile

$$B_y^{\text{tanh}} = B_0 \tanh(\pi x),$$

and a cosh profile

$$B_y^{\text{cosh}} = \frac{B_0}{\cosh(\pi x)},$$

with $B_x = 0$ and $B_0 = 1$ in all three cases. The initial conditions for the fluid quantities are

$$V_x = V_0 \sin(\pi y), \qquad\qquad V_y = 0, \qquad\qquad P = 0.1$$

for all three cases. The spatial domains are $[0, 2] \times [0, 2]$ for the sharp jump, $[-4, +4] \times [0, 4]$ for the tanh case and $[-1, +1] \times [-1, +1]$ for the cosh case. For the sharp case we consider different resolutions, namely 32×32 and 64×64 grid points, to analyse the qualitative and conservative behaviour. To compare the field line evolution for the different models, we use a common resolution of $n_x \times n_y = 30 \times 30$. In all three cases we use periodic boundaries and a timestep of $h_t = 0.1$.

In all of the considered cases, the conservation of energy and cross helicity is optimal (see figures 6.9 - 6.12). The jumps in the energy error of the cosh case (figure 6.11) are probably due to the Newton solver converging to a slightly larger residual than during the rest of the simulation. Nevertheless, the energy error is extraordinary small. Thus here we want to focus on the conservation of field line topology.

In figure 6.13, the field line evolution for the sharp jump is plotted. Initially all field lines are parallel. Due to the perturbation in the velocity field, the magnetic field lines get bend, but for most of the time, they do not break up and reconnect. At $t = 1$ and $t = 5$ we see what appears to be reconnection events, but as these islands disappear very soon after they form, this is likely to be attributed to artefacts of the plotting function (like an inaccurate interpolation or integration of the contour lines). At about $t = 20$, however, real islands form an consecutively grow as can be seen at $t = 25$. At this point, the solution can not be regarded physical anymore.

We have to stress here, that this set of initial conditions is quite challenging for most numerical schemes due to the discontinuity and other methods break down much earlier. To investigate the preservation of the magnetic field line topology on longer time scales, we consider therefore the two less severe examples of current sheaths, defined by cosh and tanh profiles, that are frequently used in reconnection studies [50, 51, 128]. In both, the cosh and the tanh case, the magnetic field changes sign not suddenly but more smoothly. Under this condition, we can run the simulation much longer. Both, figures 6.14 and 6.15, show the field line evolution for the case of the cosh sheet and the tanh sheet, respectively, up to $t = 100$. We observe, that for a smooth magnetic field, magnetic field lines are only bend but do not reconnect, as is expected from the theory but rarely observed in numerical simulations, especially on the time scales we are considering here. In our description of

magnetohydrodynamics, reconnection can only occur if the resistive term, proportional to $\nabla^2 B$, in (6.13) is present. But in ideal MHD, $\eta = 0$, such that the topology of the magnetic fields lines is preserved. Most numerical schemes, however, do feature a certain amount of numerical resistivity, leading to unphysical reconnection. In the variational integrator, such spurious resistivity appears to be completely absent, at least in the case of a continuous magnetic field.
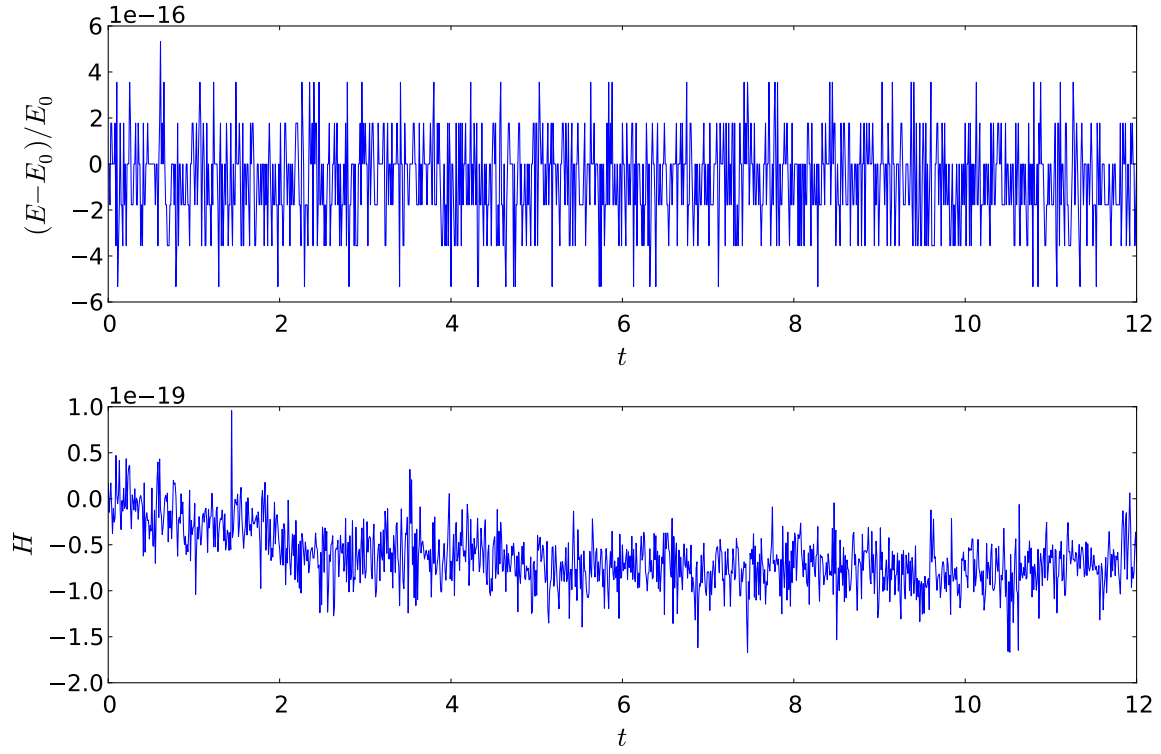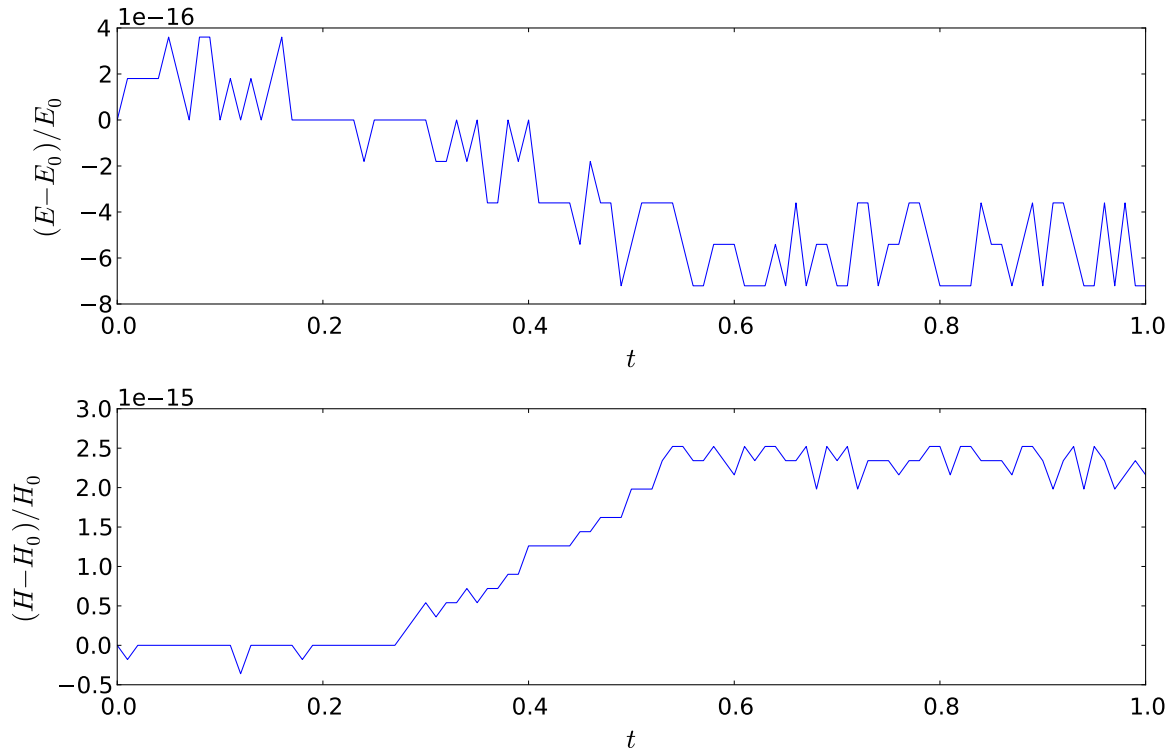
Fig. 6.3.: Loop advection, 128×64 grid points. Conservation of energy and cross helicity.
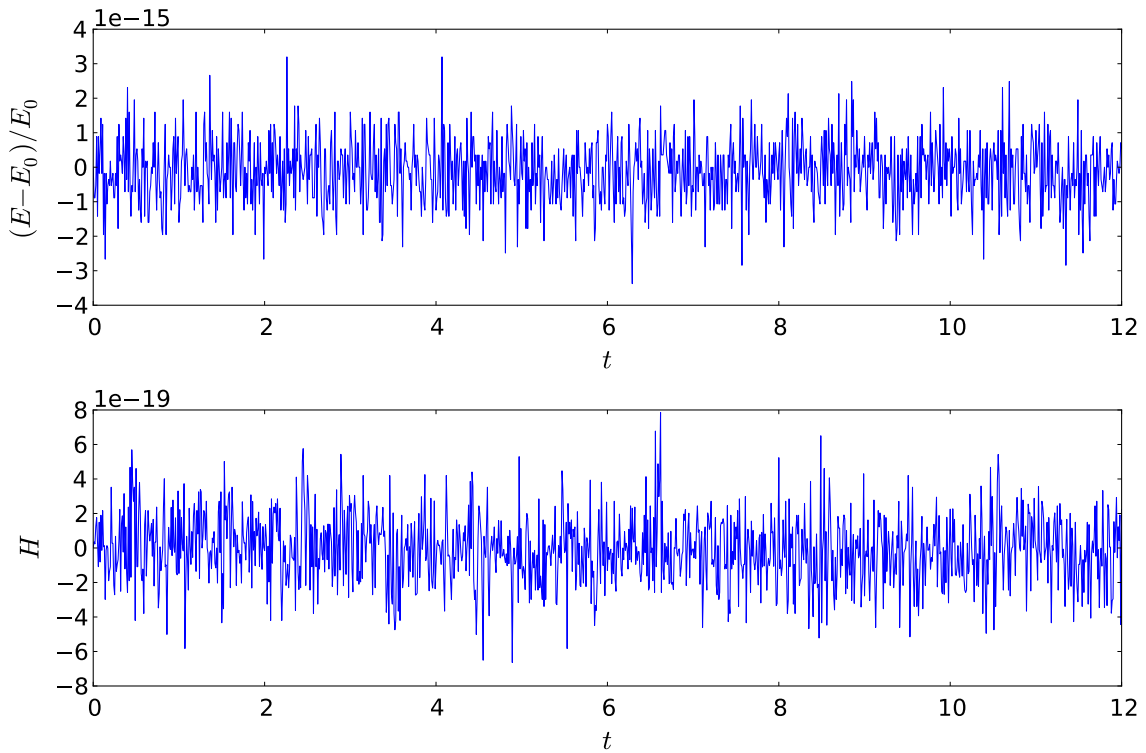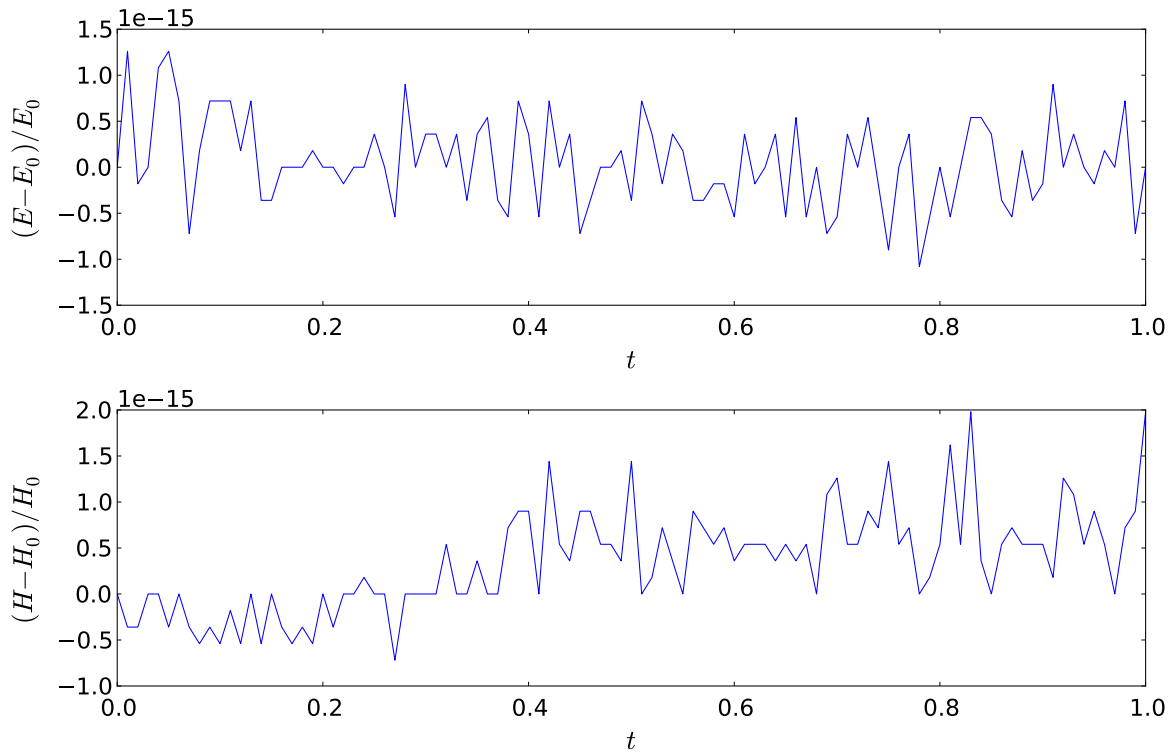

Fig. 6.4.: Loop advection, 256×128 grid points. Conservation of energy and cross helicity.
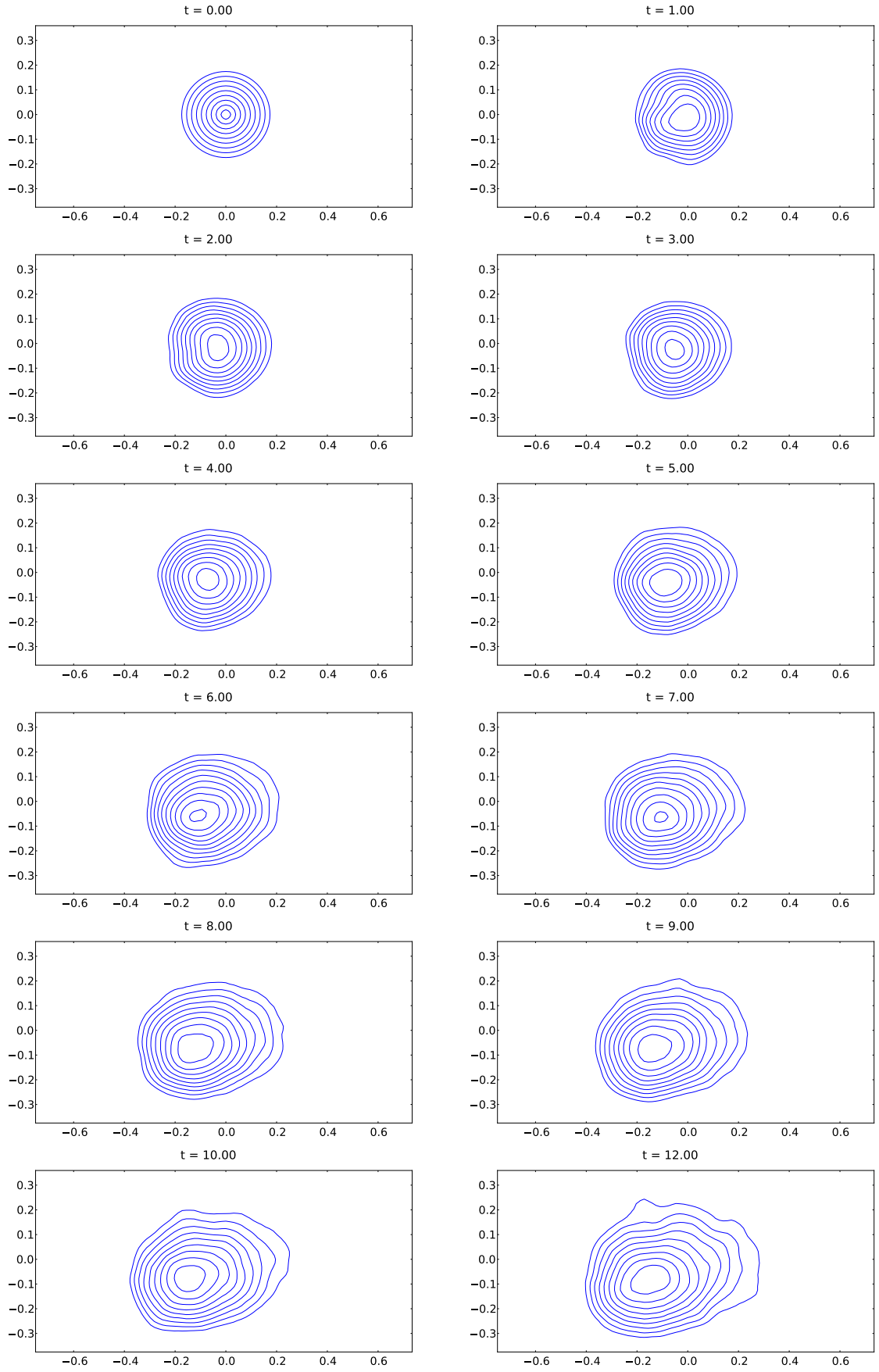
Fig. 6.5.: Loop advection, 128×64 grid points. Evolution of the magnetic loop.

Fig. 6.6.: Orszag Tang Vortex, 64×64 **grid points. Conservation of energy and cross helicity.**

Fig. 6.7.: Orszag Tang Vortex, 128×128 **grid points. Conservation of energy and cross helicity.**
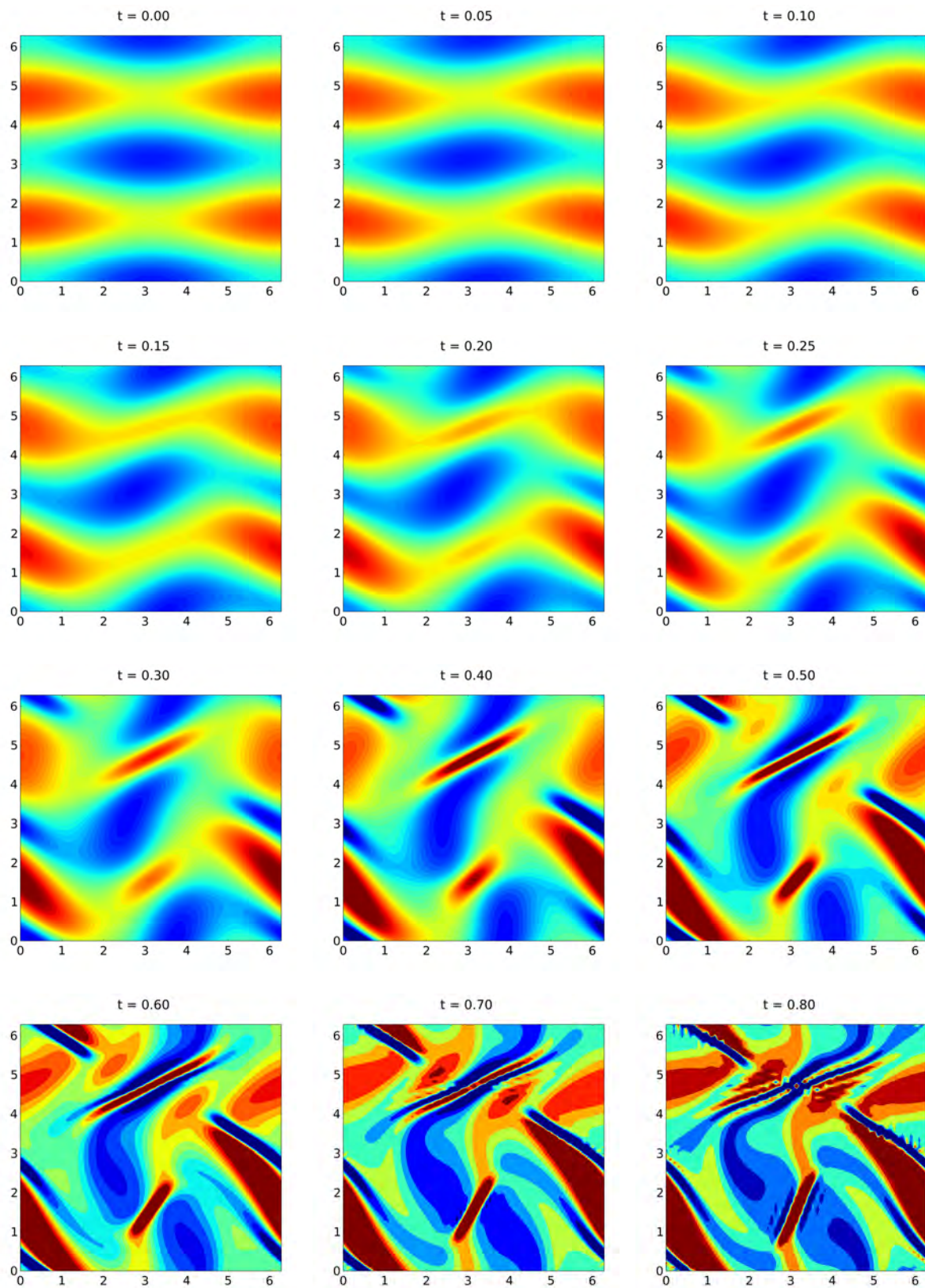
Fig. 6.8.: Orszag Tang Vortex, 64×64 **grid points. Current density. Fixed colour scale.**

Fig. 6.9.: Current sheath, 32×32 grid points. Conservation of energy and cross helicity.

Fig. 6.10.: Current sheath, 64×64 grid points. Conservation of energy and cross helicity.

Fig. 6.11.: Cosh current sheath, 30×30 grid points. Conservation of energy and cross helicity.
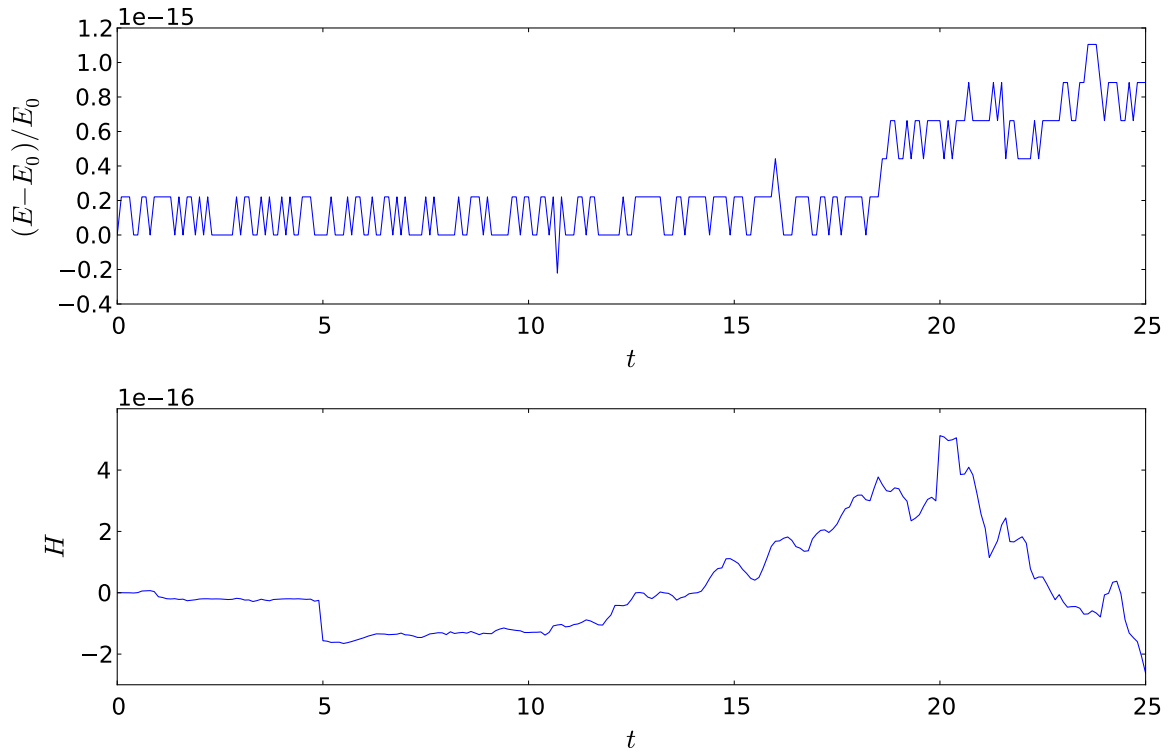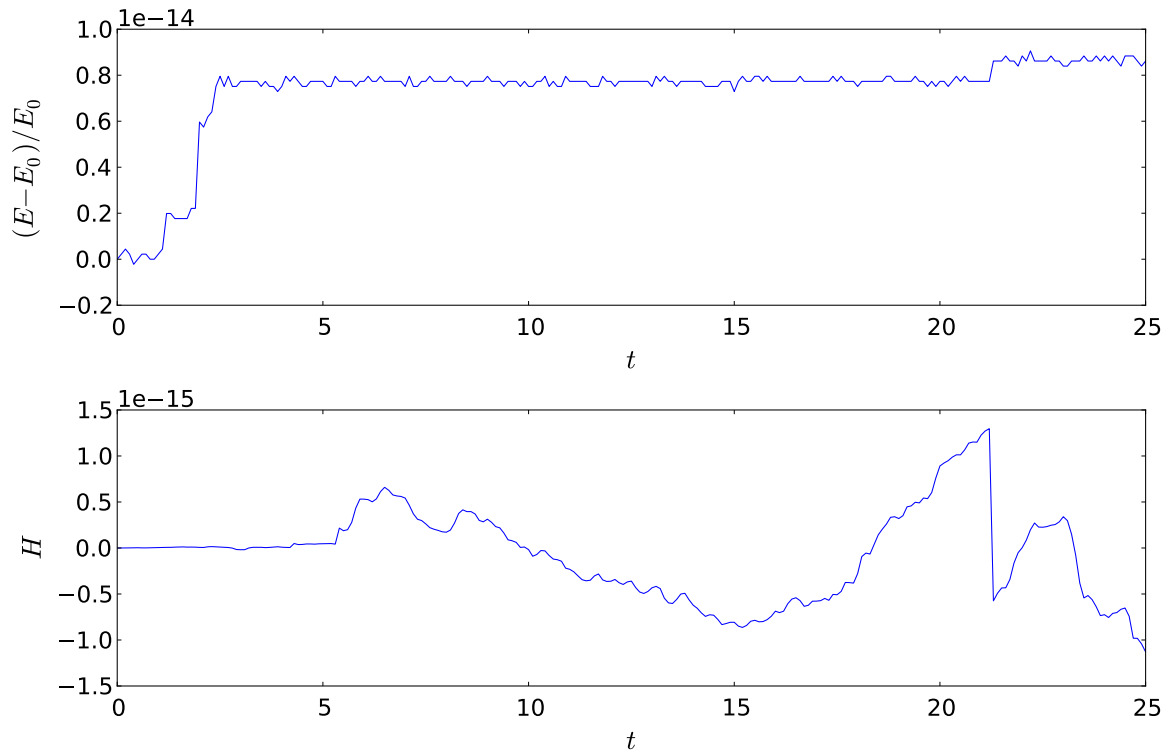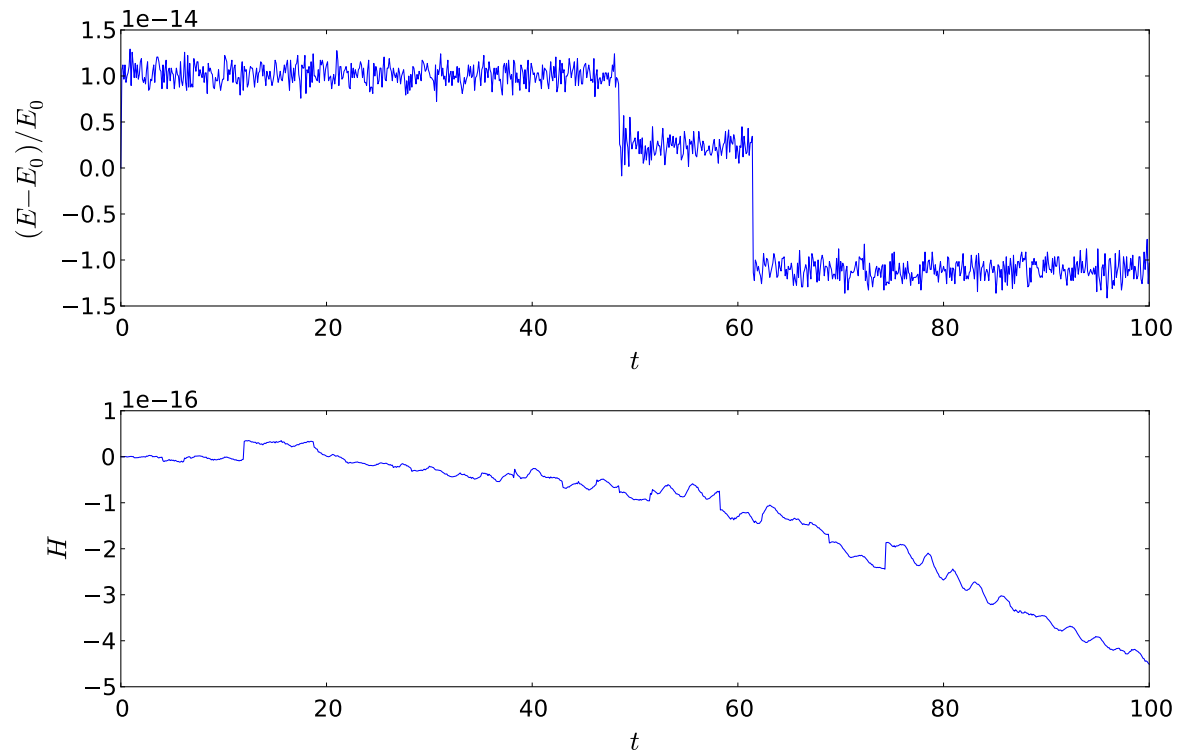

Fig. 6.12.: Tanh current sheath, 30×30 grid points. Conservation of energy and cross helicity.
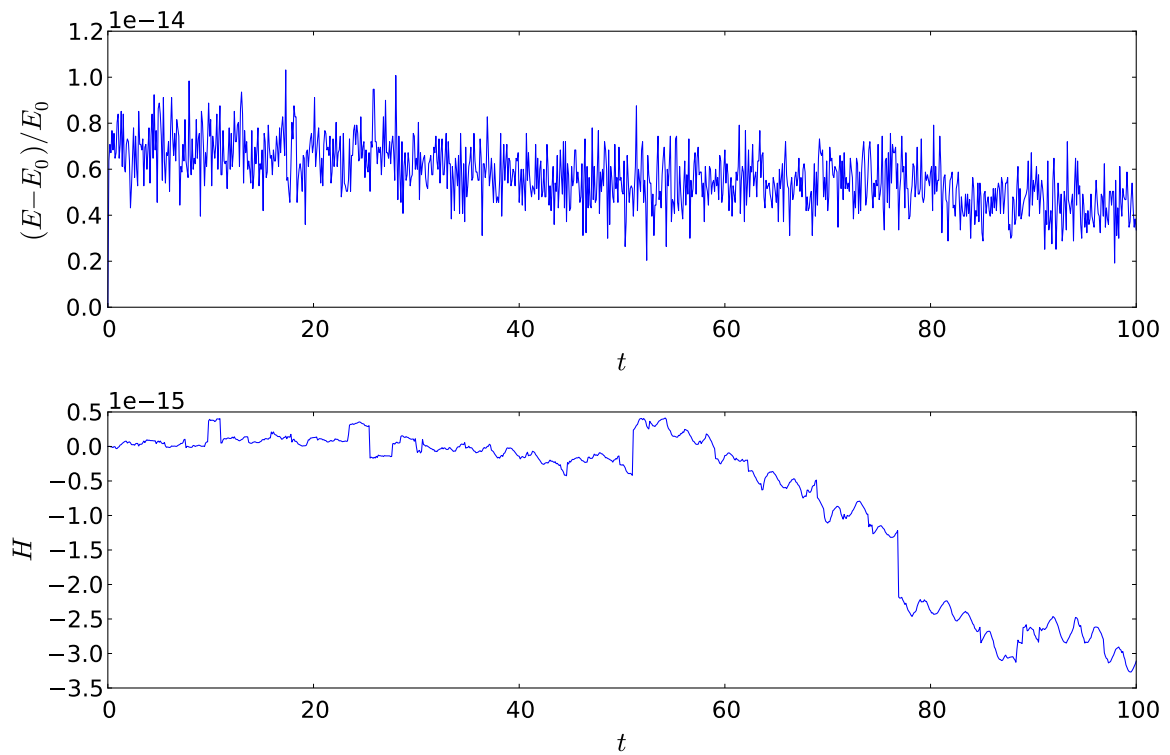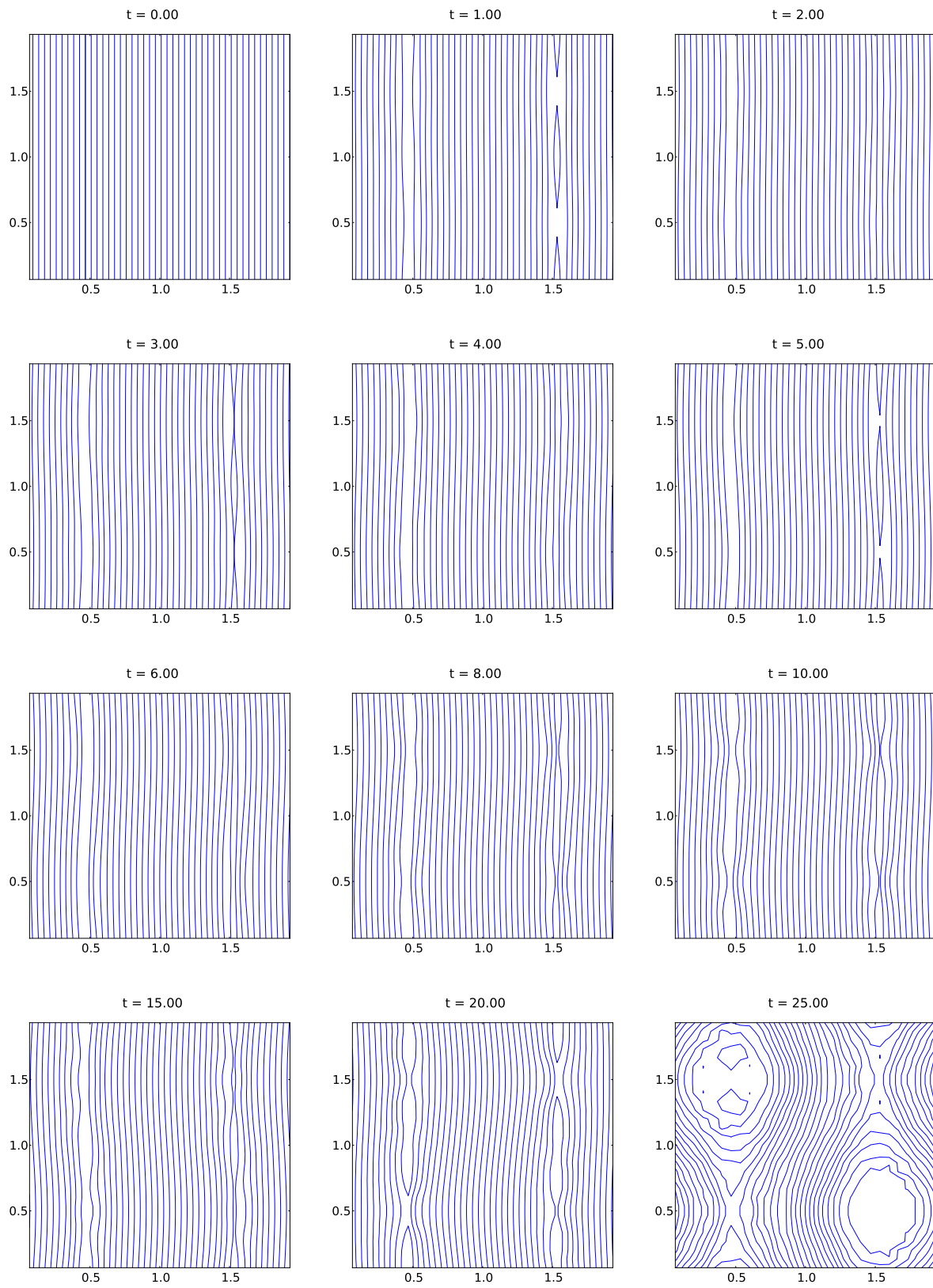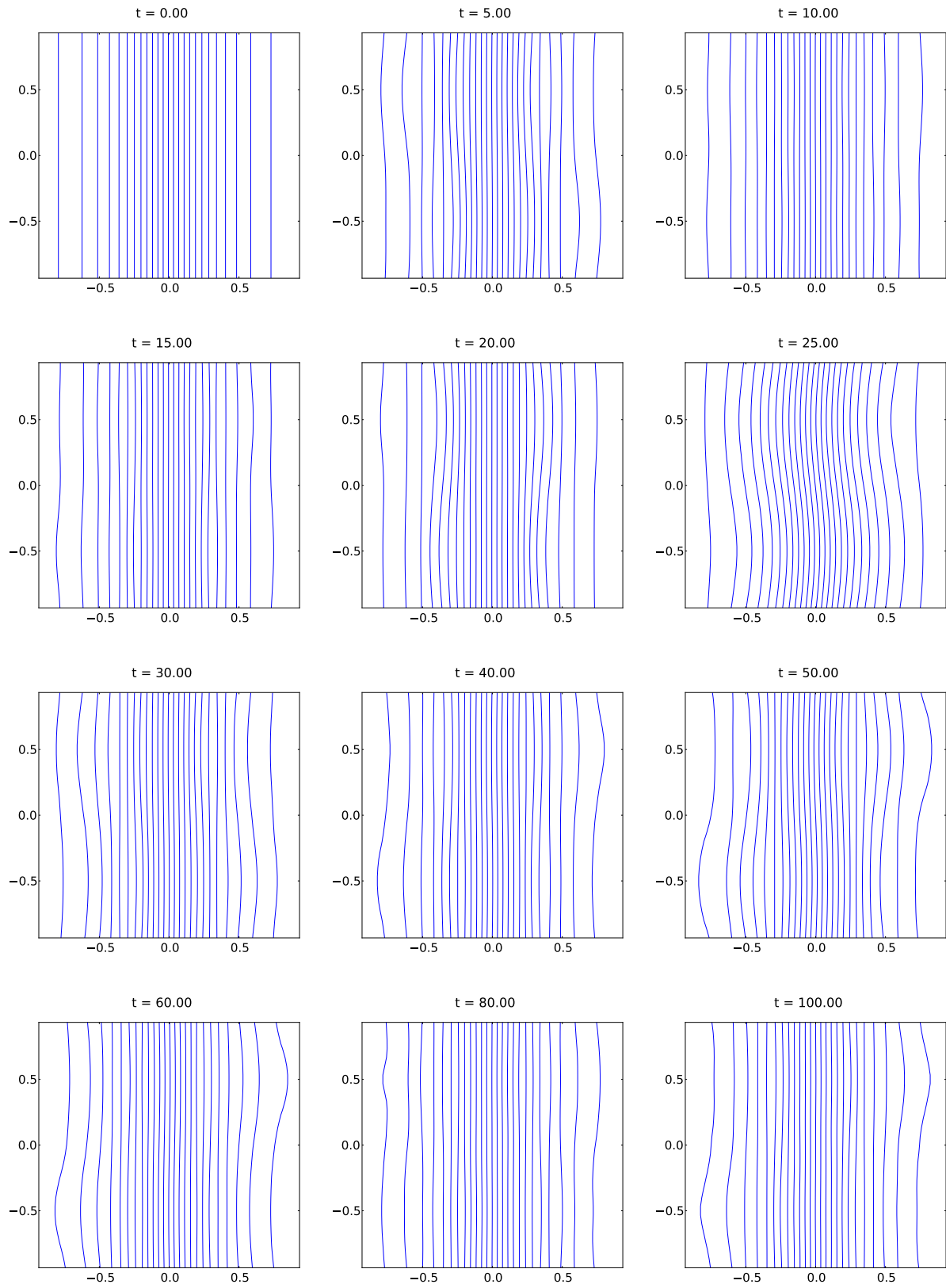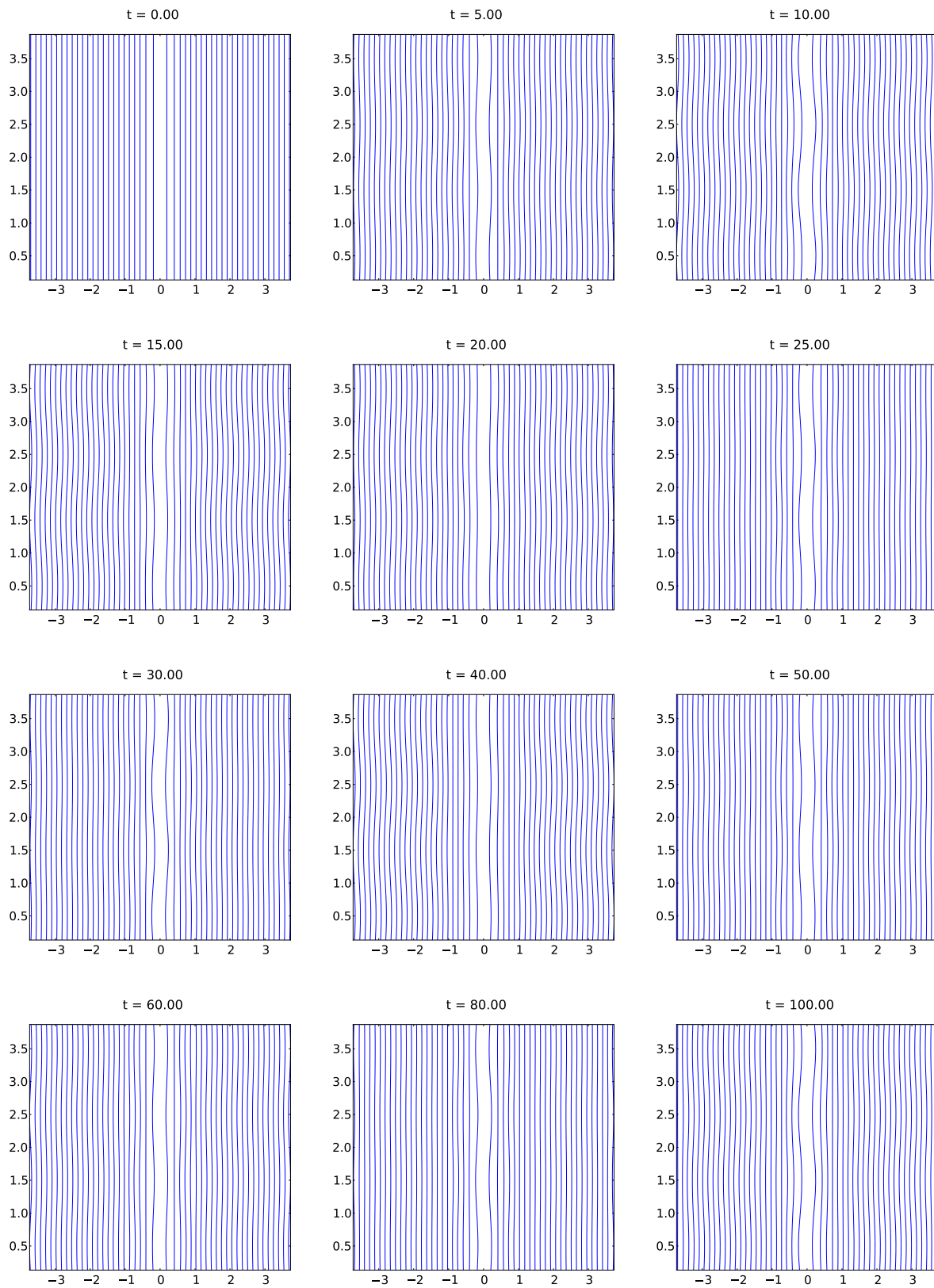
Fig. 6.13.: Sharp current sheath, 30×30 **grid points. Magnetic field lines.**

Fig. 6.14.: Cosh current sheath, 30×30 **grid points. Magnetic field lines.**

Fig. 6.15.: Tanh current sheath, 30×30 grid points. Magnetic field lines.

7. Summary and Outlook

7.1. Results

In this work, it was successfully demonstrated that the application of variational integrators to models from plasma physics has great potential in the development of novel numerical algorithms that map the real world to the discrete more realistically than most standard methods. In the following our results will be summarised.

7.1.1. Theory

Most systems from plasma physics, especially infinite dimensional ones (field theories), do not admit a natural Lagrangian formulation to which the variational integrator method is directly applicable. To be able to apply variational integrators to plasma physics problems it was therefore necessary to extend the theory. A combination of Ibragimov's theory of extended Lagrangians with the theory of variational integrators made it possible to use this method for systems not envisaged in the original theory. This includes several systems, like advection type problems, that are typical for plasma physics. The impact of this result reaches far beyond the realm of the examples considered in this work, as the class of problems that is admissible to the variational integrator theory is now greatly enlarged.

An important feature of Ibragimov's theory is the recovery of Noether's theorem, which allows us to analyse the conservation properties of a system on the continuous as well as on the discrete level, thereby allowing us to find the exact discrete expressions of the quantities that are conserved by the variational integrators.

7.1.2. Particle Dynamics

In chapter four, previous work on the development of variational integrators for guiding centre dynamics in two dimensions was extended. A new integrator based on a midpoint discretisation was developed and compared with the previous one based on a trapezoidal discretisation. The properties of both integrators are then evaluated against the standard explicit fourth order Runge-Kutta method. The midpoint integrator appears to be more accurate and more stable at smaller timesteps than the trapezoidal integrator. But both are by far superior to the Runge-Kutta method, for which the particle is found to severely deviate from its expected orbit. The variational integrators exhibit a much better long term stability as the global error of the energy is limited and does not grow in the course of a simulation, i.e., there is no numerical dissipation. The particle orbit is preserved for millions of characteristic times and hundreds of millions of timesteps. Indeed, the energy is not constant but found to oscillate about a fixed value. The important point is that the amplitude of the oscillation stays constant during the whole simulation.

These integrators were then extended to four dimensions where an additional conserved quantity, the toroidal momentum, is present. The performance of the integrators with

respect to energy conservation and accuracy of the particle orbit were similar to those of the two dimensional case. The error in the toroidal momentum is determined by the residual of the nonlinear solver.

All of the derived integrators are nonlinearly implicit such that iterative methods have to be used to solve the corresponding system of equations. In the two dimensional case, an analytic solution of the linear system comprising each Newton step can be used with a fixed number of iterations, namely three plus an explicit initial guess, such that the computational effort is about the same as with the fourth order Runge-Kutta method. In the four dimensional case, a matrix solver is used to solve the system of equations, which is computationally more demanding than the corresponding Runge-Kutta integrator, but most probably more efficient ways of solving this system could be found. That, however, is an issue that was not part of this work and is left for future research.

7.1.3. Kinetic Theory

In chapter five, several variational integrators for the Vlasov-Poisson system in one dimension (actually, one spatial plus one velocity dimension) were derived. This system can be regarded as the stepping stone to more complicated models like gyrokinetics, which is the backbone of nowadays' small scale turbulence simulations in plasma physics.

A fully nonlinear integrator was constructed, that has outstanding conservational properties, preserving the total particle number, the total energy, linear momentum and the L^2 norm up to machine precision. The only limit to the performance of this integrator is the grid resolution. If not sufficient, subgrid modes may develop, leading to large gradients in the distribution function, therefore altering the solution unphysically and spoiling the conservation properties. This is, however, a well known phenomenon with low-order finite difference methods like ours. To counteract these effects, a collision operator was constructed, that dissipates the L^2 norm and damps the subgrid modes, thereby retaining the conservation of the total particle number, energy and momentum. It was shown that with the collision operator the variational integrator allows for long time simulations of different standard test cases from kinetic theory, linear and nonlinear ones, for example Landau damping and the twostream instability.

Furthermore, a linearly implicit integrator was constructed, which is computationally less demanding but retains most of the conservation properties, even though it does not reach the accuracy in the errors of the conserved quantities as the nonlinear integrator. It is important to note that the linearisation is implemented at the level of the Lagrangian. It is not the discrete Euler-Lagrange equations that are linearised. This is a crucial point, as it guarantees that the properties of variational integrators, conservation of discrete momenta and the multisymplectic form, are kept intact. What is lost is exact energy preservation. Instead, an oscillating energy error is found, as it is typical for symplectic and multisymplectic methods. The amplitude of the oscillation depends largely on the strength of the nonlinearity as well as on the timestep. In most cases, however, the linearised system still yields very good results. The energy error is still bounded and there is still no numerical dissipation. This integrator can be used to compute an initial guess for an iterative solver of the nonlinear integrator but also standalone, if the solution of a nonlinearly implicit scheme is computationally too expensive.

Albeit implementing efficient solution techniques for the derived schemes was not an objective of this work, first results suggest that efficient iterative methods (GMRES) can be used effectively if a sufficiently accurate initial guess is available. If the linear integrator

could be solved efficiently with the GMRES method, the computational effort could be in the vicinity of explicit methods.

7.1.4. Fluid Dynamics

In chapter six, a variational discretisation of a fluid plasma model, namely ideal incompressible magnetohydrodynamics, was developed and analysed. The ideal model shares many features with more sophisticated and physically comprehensive descriptions used in simulations of large scale turbulence or magnetic reconnection, to name just two examples. The scheme is similar to the one for the Vlasov-Poisson system in that it is fully nonlinear and exhibits excellent conservation properties. Both, the total energy and the cross helicity are preserved up to machine precision over long simulation times. Nonlinear Alfvénic waves, a core phenomenon of magnetohydrodynamics, were found to keep travelling for thousand characteristic times without loosing energy. In turbulent settings, where small scale structures can develop, problems with subgrid modes were observed if the resolution was not sufficient, similar to the Vlasov-Poisson integrator. As in realistic simulations usually some form of dissipation is present in the model, this is not expected to be a problem impairing the applicability of the derived scheme. In these cases, the use of a conservative scheme can ensure that the dissipation mechanism is consistent with the physical dissipation and not numerically generated.

The analysis of several reconnection models has shown the preservation of magnetic field line topology by the variational integrator for very long times, provided that the magnetic field is continuous. For a discontinuous magnetic field, the topology is still preserved for times longer than with most standard methods, but at some point the solution was found to become spurious, probably due to error accumulation. This is, however, not discouraging as the variational integrator was not constructed with discontinuities in mind for which usually specially designed methods are employed.

7.1.5. Semi-Discretisations

In the appendix we describe two different approaches of using variational integrators or closely related methods to obtain semi-discretisations for the Vlasov equation and similar systems.

In appendix A, the spatial dimensions are transformed into Fourier space and only time or time and the velocity dimensions are discretised variationally. In appendix B, discretisation methods for different kinds of brackets are presented, where only the phasespace dimensions are discretised but not time. A new discretisation for Nambu three brackets is derived, that has potential use in simulations of the gyrokinetic Vlasov equation in axisymmetric tokamaks. Furthermore, it is shown that discretisation strategies for Lie-Poisson brackets are applicable also to problems from plasma physics. Interestingly, these bracket discretisations are quite similar to variational integrators, e.g., for the Vlasov-Poisson system.

7.2. Future Work

This work has shown that the application of variational integrators in plasma physics is both viable and attractive. Thereby it has lead the way to a number of possible directions for future research, some of which are sketched below.

7.2.1. Theory

The theory of variational integrators is still young and therefore not fully developed. So far, there are no clear guidelines on which discretisations of the Lagrangian lead to "good" (e.g., stable) numerical schemes. A possible solution to this might be the combination of the present methods with Arnold's theory of finite element exterior calculus (more on this below).

In general, the variational integrator method does not lead to schemes that are (multi-) symplectic, momentum and energy preserving at the same time, without implementing one or another form of timestep adaption, either global timestep adaption or asynchronous variational integrators [62, 76]. However, the methods we derived for the Vlasov-Poisson system as well as for ideal magnetohydrodynamics turned out to be exactly energy preserving (to machine precision), momentum preserving (also to machine precision), and they are multisymplectic by design. It would be interesting and important to understand such properties and find general criteria for which discretisations of the Lagrangian lead to such optimal integrators.

Finite Element Discrete Exterior Calculus

Recently, a discrete theory of exterior calculus based on finite elements has been developed by Arnold et al. [6, 7]. In contrast to the theory of variational integrators, this theory of finite element exterior calculus is embedded in an abstract Hilbert space framework that makes the analysis of stability and convergence of the derived discretisations much more systematic. A connection between these two theories might allow for a rigorous numerical analysis of variational integrators and therefore help to find answers to some of the questions raised in the previous paragraph.

Analysis of Discrete Conservation Laws and Preservation of the Multisymplectic Form

To judge the performance of the variational integrators derived in chapters five and six, only heuristic diagnostics were used. Instead, a detailed analysis of the discrete conservation laws using the discrete Noether theorem of section 3.2.4 should be carried out to find exact expressions of the conserved quantities.

It was shown how the solution space of a system described by an extended Lagrangian (section 2.2.7) can be restricted to the solution space of the original system to recover the conserved quantities. An open problem is to find out if the multisymplectic structure of the extended system endows the physical system with a compatible multisymplectic structure and if so how the structure of the extended system can be restricted to obtain the one of the physical system. The variational integrators preserve a discrete counterpart of the multisymplectic structure by construction, but so far no statement about the conservation of the multisymplectic structure of the original system is possible.

7.2.2. Vlasov-Poisson and Vlasov-Maxwell

Some obvious extensions of the variational integrator for the Vlasov-Poisson system from chapter five include higher dimensional domains, the Vlasov-Maxwell system, and higher order schemes. Not many surprises are expected in deriving variational integrators for the Vlasov-Poisson system in higher dimensions, especially extending the integrator of

the Vlasov equation should prove to be straight forward. Treating the electrodynamic equations of the Vlasov-Maxwell system properly might however necessitate the use of a staggered grid approach [137] and therefore suggests a treatment based of discrete exterior calculus [123, 124].

The derivation of higher order schemes as well appears to be a straight forward generalisation of the results presented in this work. The most interesting question in this respect is if Arakawa's fourth order discretisation of the Poisson brackets can be derived by a variational method similarly to his second order discretisation.

Nambu Bracket Discretisation

In appendix B, it was shown that the Lie-Poisson bracket formulation of the Vlasov equation can be used to obtain semi-discretisations of the phasespace part of the Vlasov equation by first transforming them to Nambu brackets. A Lie-Poisson bracket also exists for the Vlasov-Maxwell system [82]

$$
\{F, G\} = \int f \left[\frac{\delta F}{\delta f}, \frac{\delta G}{\delta f} \right] dx\, dv + \int \left(\frac{\delta F}{\delta E} \cdot \left(\nabla \times \frac{\delta G}{\delta B} \right) - \frac{\delta G}{\delta E} \cdot \left(\nabla \times \frac{\delta F}{\delta B} \right) \right) dx\, dv
$$
$$
+ \int \left(\frac{\delta F}{\delta E} \cdot \frac{\delta f}{\delta v} \frac{\delta G}{\delta f} - \frac{\delta G}{\delta E} \cdot \frac{\delta f}{\delta v} \frac{\delta F}{\delta f} \right) dx\, dv + \int f\, B \cdot \left(\frac{\partial}{\partial v} \frac{\delta F}{\delta f} \times \frac{\partial}{\partial v} \frac{\delta G}{\delta f} \right) dx\, dv, \quad (7.1)
$$

where f is the distribution function, E and B the electric and magnetic fields, respectively, and H is the Hamiltonian energy functional, given by

$$
H(f, E, B) = \frac{1}{2} \int \|v\|^2\, f(t, x, v)\, dx\, dv + \frac{1}{2} \int \left(\|E(t, x)\|^2 + \|B(t, x)\|^2 \right) dx, \quad (7.2)
$$

such that the evolution of all functionals $F(f, E, B)$ is determined by

$$
\dot{F} = \{F, H\}. \tag{7.3}
$$

Preliminary results suggest that this Lie-Poisson bracket has a Nambu bracket formulation as well which could be used to derive semi-discretisations for the Vlasov-Maxwell system, where in contrast to the Vlasov-Poisson system not only the distribution function but also the electric and magnetic fields obey dynamical equations.

Dirac Bracket Discretisation

Recently, there has been some effort to derive Dirac brackets for several models of plasma physics [27, 26], including the Vlasov-Poisson and Vlasov-Maxwell systems. These brackets are constructed to respect constraints of the dynamics automatically, e.g. $\Delta \phi - \int f\, dv = 0$ for the Vlasov-Poisson system or $\nabla \cdot V = 0$ for incompressible fluids. The Dirac brackets for the Vlasov-Poisson system,

$$
\{F, G\}_*(f, E) = \int f \left[\frac{\delta F}{\delta f} - \Delta^{-1} \nabla \cdot \frac{\delta F}{\delta E}, \frac{\delta G}{\delta f} - \Delta^{-1} \nabla \cdot \frac{\delta G}{\delta E} \right] dx\, dv, \tag{7.4}
$$

can be transformed into Lie-Poisson brackets,

$$
\{F, G\}(f, \Phi) = \int f \left[\frac{\delta F}{\delta f} - \frac{\delta F}{\delta \Phi}, \frac{\delta G}{\delta f} - \frac{\delta G}{\delta \Phi} \right] dx\, dv, \tag{7.5}
$$

where

$$\Phi = \Delta\phi = -\nabla \cdot E. \tag{7.6}$$

Again it might be possible to find a relation with Nambu brackets and use that relation to obtain semi-discretisations. An interesting peculiarity of this formulation is that due to the Poisson equation being included as a constraint two dynamical equations, the usual one for the distribution function f and another one for the potential vorticity Φ, have to be solved. This might turn out to be an advantage for parallel implementations of the resulting methods.

Euler-Poincaré Action Principle

The Euler-Poincaré action principle reviewed in section 5.2.3 might pose the starting point for the derivation of variational integrators in a similar way as was done by Pavlov et al. [99, 98] and Gawlik et al. [43] for incompressible fluids. As already noted by Squire et al. [122] this might prove nontrivial as, e.g., a discretisation of the group of symplectomorphisms, which describes the dynamics of the Vlasov-Poisson and Vlasov-Maxwell systems, has to be found (for a more thorough discussion see [122]).

Despite the possible difficulties it seems worthwhile to pursue this path as the Euler-Poincaré reduced system appears to be the most natural, most geometric formulation of the family of Vlasov systems known to date.

Gyrokinetics

From the point of view of applying variational integrators in large scale plasma physics codes, the kinetic model appears to be less attractive than the gyrokinetic model. Gyrokinetics is a version of kinetic theory reduced to five phasespace dimensions instead of six, thereby lowering the computational burden.

An extension of the variational integrators derived in this work to gyrokinetics is thus an important topic of future work, overlapping with the extension to higher dimensions and the Vlasov-Maxwell system.

7.2.3. Ideal and Reduced MHD

The most obvious extensions of the variational integrator for magnetohydrodynamics from chapter six are similar to those of the Vlasov-Poisson case: moving to three dimensional domains, higher order discretisations, and more comprehensive models. Some others are shortly explained below.

Potential Formulation

In studies of magnetic reconnection, the potential formulation presented in section (6.1.2) is very popular. To derive variational integrators for this formulation, an extended version of the theory for second or third order field theories has to be applied [65, 64]. Apart from this there appear to be no obvious obstacles.

It would be quite interesting to compare the performance of variational integrators for the two different approaches, i.e., the description in terms of the potentials A and ψ and the description in terms of the fields B and V.

Nambu and Dirac Brackets

Several flavours of magnetohydrodynamics can also be described by Lie-Poisson brackets, e.g., for the ideal case in potential formulation that is

$$\{F,G\}(A,\omega) = \int A \left(\left[\frac{\delta F}{\delta A}, \frac{\delta G}{\delta \omega} \right] - \left[\frac{\delta F}{\delta \omega}, \frac{\delta G}{\delta A} \right] \right) dx\,dy + \int \omega \left[\frac{\delta F}{\delta \omega}, \frac{\delta G}{\delta \omega} \right] dx\,dy, \qquad (7.7)$$

with the Hamiltonian energy functional

$$H(A,\omega) = \frac{1}{2} \int \left(\psi\omega - Aj \right) dx\,dy. \qquad (7.8)$$

Here, A is the magnetic vector potential, $j = -\Delta A$ the current density, $\omega = -\Delta\psi$ is the vorticity, and ψ the streaming function. With the Casimir invariant

$$C = \int A\omega\,dx\,dy, \qquad (7.9)$$

this can be transformed into a Nambu bracket

$$\{F,G,C\}(A,\omega) = \int \frac{\delta C}{\delta \omega} \left(\left[\frac{\delta F}{\delta A}, \frac{\delta G}{\delta \omega} \right] - \left[\frac{\delta F}{\delta \omega}, \frac{\delta G}{\delta A} \right] \right) dx\,dy + \int \frac{\delta C}{\delta A} \left[\frac{\delta F}{\delta \omega}, \frac{\delta G}{\delta \omega} \right] dx\,dy. \qquad (7.10)$$

A discretisation approach as described in appendix B.3.3 should be straight forwardly applicable to this expression.

As for the family of Vlasov systems, there has been recent research on Dirac bracket formulations for magnetohydrodynamics as well [27, 26], which might open new possibilities for deriving discretisations if a relation between Dirac brackets and Nambu brackets could be drawn as outlined above.

A. Mixed Spectral-Variational Schemes

In this appendix, we sketch how to derive mixed spectral-variational methods for the Vlasov-Poisson system. We do the derivation for a system very similar to Vlasov-Poisson, but somewhat simpler, namely the vorticity equation in two spatial dimensions. The generalisation to the Vlasov-Poisson equation, e.g., in two spatial and two velocity dimensions is mostly straight forward.

A.1. The Vorticity Equation in 2D

The vorticity equation describes the evolution of the vorticity of a fluid element in an incompressible ideal fluid

$$\frac{\partial \omega}{\partial t} + \frac{\partial \omega}{\partial x} \frac{\partial \psi}{\partial y} - \frac{\partial \omega}{\partial y} \frac{\partial \psi}{\partial x} = 0, \tag{A.1}$$

where ω is the vorticity of the fluid and ψ is the streaming function, determined by

$$-\Delta \psi = \omega. \tag{A.2}$$

In two dimensions, the vorticity equations takes a particularly interesting form which has a structure similar to the one of the Vlasov-Poisson system. The analogy is not exact, as the Poisson equation (A.2) is two-dimensional.

In Fourier representation ω and ψ take the form

$$\omega(r,t) = \sum_k \exp\{-i\,k \cdot r\}\,\hat{\omega}(k,t), \tag{A.3a}$$

$$\psi(r,t) = \sum_k \exp\{-i\,k \cdot r\}\,\hat{\psi}(k,t), \tag{A.3b}$$

where $k \in \mathbb{Z}^2$ and $r = (x,y) \in \mathbb{I}^2$. With the shorthand notation ω_k for $\hat{\omega}(k,t)$ and ψ_k for $\hat{\psi}(k,t)$, the vorticity and Poisson equations become

$$\frac{\partial \omega_k}{\partial t} = \hat{z} \cdot \sum_{p+q=k} (p \times q)\,\omega_q\,\psi_p, \tag{A.4}$$

$$-k^2\,\psi_k = \omega_k. \tag{A.5}$$

The reality condition

$$\omega(-k,t) = \omega^*(k,t) \tag{A.6}$$

determines the Fourier modes in the lower half plane and will play an important role in the analysis of conserved quantities.

A.2. Conservation Laws

We will concentrate on a particular invariant of the spectral vorticity equation, the L^2 norm of the vorticity

$$G = \sum_k |\omega_k|^2 = \int_{\mathbb{I}^2} \omega^2 \, dx \, dy. \tag{A.7}$$

In physical space, $dG/dt = 0$ is a consequence of the anti-symmetry of the Poisson brackets in (A.1). In Fourier space, the conservation of G follows by multiplying the spectral vorticity equation with ω_{-k}

$$\omega_{-k} \frac{\partial \omega_k}{\partial t} = \hat{z} \cdot \sum_{p+q=k} p \times q \, \omega_{-k} \, \omega_q \, \psi_p \tag{A.8}$$

and adding of the result to the same equation with the sign of k flipped

$$\omega_k \frac{\partial \omega_{-k}}{\partial t} = \hat{z} \cdot \sum_{p+q=-k} p \times q \, \omega_k \, \omega_q \, \psi_p \tag{A.9}$$

to get

$$\omega_{-k} \frac{\partial \omega_k}{\partial t} + \omega_k \frac{\partial \omega_{-k}}{\partial t} = \hat{z} \cdot \left[\sum_{p+q=k} p \times q \, \omega_{-k} \, \omega_q \, \psi_p + \sum_{p+q=-k} p \times q \, \omega_k \, \omega_q \, \psi_p \right]. \tag{A.10}$$

With the reality condition, the left-hand side becomes

$$\omega_{-k} \frac{\partial \omega_k}{\partial t} + \omega_k \frac{\partial \omega_{-k}}{\partial t} = \frac{\partial}{\partial t} \left(\omega_k^* \omega_k \right) = \frac{\partial}{\partial t} |\omega_k|^2 . \tag{A.11}$$

We sum the full equation over k and rewrite the sums on the right-hand side

$$\frac{\partial}{\partial t} \sum_k |\omega_k|^2 = \hat{z} \cdot \sum_k \sum_q \left[(k-q) \times q \, \omega_{-k} \, \omega_q \, \psi_{k-q} + (-k-q) \times q \, \omega_k \, \omega_q \, \psi_{-k-q} \right]$$

$$= \hat{z} \cdot \sum_k \sum_q k \times q \left[\omega_{-k} \, \omega_q \, \psi_{k-q} - \omega_k \, \omega_q \, \psi_{-k-q} \right]. \tag{A.12}$$

On the right-hand side, exchange k and q (we can do this as both, the sum of k and the sum of q, run over the whole wave number space) and add the result to the original equation

$$\frac{\partial}{\partial t} \sum_k |\omega_k|^2 = \frac{1}{2} \hat{z} \cdot \sum_k \sum_q \left(k \times q \left[\omega_{-k} \, \omega_q \, \psi_{k-q} - \omega_k \, \omega_q \, \psi_{-k-q} \right] + q \times k \left[\omega_{-q} \, \omega_k \, \psi_{q-k} - \omega_q \, \omega_k \, \psi_{-q-k} \right] \right)$$

$$= \frac{1}{2} \hat{z} \cdot \sum_k \sum_q k \times q \left[\omega_{-k} \, \omega_q \, \psi_{k-q} - \omega_k \, \omega_q \, \psi_{-k-q} + \omega_q \, \omega_k \, \psi_{-q-k} - \omega_{-q} \, \omega_k \, \psi_{q-k} \right]$$

$$= \frac{1}{2} \hat{z} \cdot \sum_k \sum_q k \times q \left[\omega_{-k} \, \omega_q \, \psi_{k-q} - \omega_{-q} \, \omega_k \, \psi_{q-k} \right]. \tag{A.13}$$

Now change the sign of both k and q, and add the result to the original equation

$$\frac{\partial}{\partial t} \sum_k |\omega_k|^2 = \frac{1}{4} \hat{z} \cdot \sum_k \sum_q k \times q \left[\omega_{-k} \, \omega_q \, \psi_{k-q} - \omega_q \, \omega_{-k} \, \psi_{-q+k} + \omega_k \, \omega_{-q} \, \psi_{-k+q} - \omega_{-q} \, \omega_k \, \psi_{q-k} \right]$$

$$= 0. \tag{A.14}$$

Thus we have proved that G is conserved.

A.3. Extended Lagrangian

The extended Lagrangian one-form for the spectral vorticity equation is

$$\mathcal{L}\big(\omega_k, \xi_k\big) = \xi_k \left[\frac{\partial \omega_k}{\partial t} - \hat{z} \cdot \sum_{p+q=k} p \times q \, \omega_q \, \psi_p \right] dt. \tag{A.15}$$

The Poisson equation (A.5) is not explicitly time dependent and therefore not included. The variational principle applied to the action

$$\mathcal{A} = \int \mathcal{L}\big(\omega_k, \xi_k\big) \tag{A.16}$$

yields two equations

$$\frac{\partial \omega_k}{\partial t} = \hat{z} \cdot \sum_{p+q=k} p \times q \, \omega_q \, \psi_p, \tag{A.17a}$$

$$\frac{\partial \xi_k}{\partial t} = \hat{z} \cdot \sum_{p+q=k} p \times q \, \xi_q \, \psi_p, \tag{A.17b}$$

that are identical. So we can assume that ξ has the same solution as ω, which is important for the analysis of conserved quantities with Noether's theorem.

A.4. Variational Integrator

In the discrete system we apply the Fourier-Galerkin truncation where the sum is taken only over a finite subset of all possible wave numbers k. We will only discretise time with the help of a discrete variational principle.

The discrete extended Lagrangian density is

$$\mathcal{L}_d\big(\omega_k^1, \omega_k^2; \xi_k^1, \xi_k^2\big) = \frac{1}{2}\big(\xi_k^1 + \xi_k^2\big)\left[\frac{\omega_k^2 - \omega_k^1}{h_t} - \frac{1}{4}\,\hat{z} \cdot \sum_{p+q=k} p \times q \left(\omega_q^1 + \omega_q^2\right)\big(\psi_p^1 + \psi_p^2\big) \right] h_t. \tag{A.18}$$

With that one finds the discrete Euler-Lagrange equations

$$\frac{\partial \mathcal{L}_d}{\partial \xi^1}\big(\omega^j, \omega^{j+1}\big) + \frac{\partial \mathcal{L}_d}{\partial \xi^2}\big(\omega^{j-1}, \omega^j\big) = 0 \tag{A.19}$$

to be

$$\frac{\omega_k^{j+1} - \omega_k^j}{2h_t} + \frac{\omega_k^j - \omega_k^{j-1}}{2h_t} = \frac{1}{8}\,\hat{z} \cdot \sum_{p+q=k} p \times q \left[\left(\omega_q^{j-1} + \omega_q^j\right)\big(\psi_p^{j-1} + \psi_p^j\big) + \left(\omega_q^j + \omega_q^{j+1}\right)\big(\psi_p^j + \psi_p^{j+1}\big) \right]. \tag{A.20}$$

This corresponds to the sum of two equations (c.f. the discussion in section 5.3.2),

$$\frac{\omega_k^{j+1} - \omega_k^j}{2h_t} = \frac{1}{8}\,\hat{z} \cdot \sum_{p+q=k} p \times q \left(\omega_q^j + \omega_q^{j+1}\right)\big(\psi_p^j + \psi_p^{j+1}\big), \tag{A.21a}$$

$$\frac{\omega_k^j - \omega_k^{j-1}}{2h_t} = \frac{1}{8}\,\hat{z} \cdot \sum_{p+q=k} p \times q \left(\omega_q^{j-1} + \omega_q^j\right)\big(\psi_p^{j-1} + \psi_p^j\big). \tag{A.21b}$$

If ω_k is a solution of the first equation, it is also a solution of the second equation, as well as of the original equation (for a detailed discussion see section 5.3). So the discrete equation we use to advance the spectral vorticity in time is

$$\frac{\omega_k^{j+1} - \omega_k^j}{h_t} = \frac{1}{4} \, \hat{z} \cdot \sum_{p+q=k} p \times q \left(\omega_q^j + \omega_q^{j+1} \right) \left(\psi_p^j + \psi_p^{j+1} \right). \tag{A.22}$$

The proof of conservation of the discrete L^2 norm follows exactly the same path as in the continuous case and is therefore omitted.

A.5. The Vlasov-Poisson System

The Vlasov-Poisson system (see also chapter 5)

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial x} - \frac{\partial \phi}{\partial x} \cdot \frac{\partial f}{\partial v} = 0 \tag{A.23}$$

$$\Delta \phi = - \int f \, dv. \tag{A.24}$$

describes the dynamics of a charged particle system, characterised by the particle distribution function f, in an electrostatic potential ϕ. In Fourier representation, f and ϕ read

$$f(t, x, v) = \sum_k \exp\{-i\,k \cdot x\} \, \hat{f}(t, k, v), \tag{A.25}$$

$$\phi(t, x) = \sum_k \exp\{-i\,k \cdot x\} \, \hat{\phi}(t, k), \tag{A.26}$$

and the Vlasov and Poisson equations become

$$\frac{\partial \hat{f}}{\partial t} - i\,v \cdot k\,\hat{f} + i\,\hat{\phi}\,k \cdot \frac{\partial \hat{f}}{\partial v} = 0, \tag{A.27}$$

$$k^2 \, \hat{\phi} = - \int \hat{f} \, dv. \tag{A.28}$$

The variational integrator is derived in the same way as for the vorticity equation, only that now, in addition to time, also the velocity dimensions are considered in the derivation. However, the analysis of conservation laws is supposedly more complicated than in the previous case.

B. Discretisation of Brackets

In this appendix we want to discuss the discretisation of Poisson brackets, Lie-Poisson brackets and both finite and infinite dimensional Nambu brackets [92]. This treatment is based on ideas of Salmon and Talley. In [110], they describe a general way of discretising Poisson brackets $[\cdot, \cdot]$ by a method based on a discrete functional derivative that is very similar to the variational integrator formalism. In [109], Salmon generalises these ideas to infinite dimensional Nambu brackets (Nambu field brackets), which are related to Lie-Poisson brackets as they often appear in the Hamiltonian description of plasma physics models.

B.1. Canonical Poisson Brackets

The starting point for Salmon and Talley [110] is the rephrasing of the equation at hand, e.g.,

$$\mathcal{D} f(t, x, p) = 0, \tag{B.1}$$

where $\mathcal{D}$ is any operator, in a weak formulation, that is

$$\int g(x, p) \, \mathcal{D} f(t, x, p) \, dx \, dp = 0 \qquad \text{for any test function } g(x, p). \tag{B.2}$$

If the solutions of (B.2) are regular and (B.2) vanishes for any $g(x, y)$ it is equivalent to the original equation (B.1). Consider as an example the Vlasov equation from chapter 5,

$$\partial_t f + [f, h] = 0, \tag{B.3}$$

for which the corresponding weak formulation reads

$$\int g(x, p) \left(\partial_t f + [f, h] \right) dx \, dp = 0. \tag{B.4}$$

Here, $f(t, x, p)$ is the distribution function and h is the particle Hamiltonian. This is almost identical to the extended Lagrangian (5.74) from chapter 5. What is missing is the integral over time. So from this formulation, a semi-discretisation of the phasespace part of the equation is obtained. Such a semi-discretisation has also been considered by León et al. [75] in the framework of variational integrators.

In the following we want to concentrate on the brackets, ignoring the time derivative. Therefore, we consider the energy functional

$$\mathcal{E} = \int g\,[f, h] \, dx \, dp, \tag{B.5}$$

such that the Poisson bracket $[f, h]$ is retained by a functional derivative with respect to g, that is

$$\frac{\delta \mathcal{E}}{\delta g} = [f, h]. \tag{B.6}$$

It is now important to realise, integrating by parts with appropriate boundary conditions, that the following expressions are identical

$$\int g\,[f,h]\,dx\,dp = \int f\,[h,g]\,dx\,dp = \int h\,[g,f]\,dx\,dp. \tag{B.7}$$

The energy functional (B.5) can thus be written as a convex combination of these three expressions

$$\mathcal{E} = \int \Big[\alpha\,g\,[f,h] + \beta\,f\,[h,g] + \gamma\,h\,[g,f]\Big]\,dx\,dp, \tag{B.8}$$

where $\alpha + \beta + \gamma = 1$. This observation is important for the discretisation of the brackets. To retain the antisymmetry properties of the brackets on the discrete level Salmon and Talley observed that equal factors, $\alpha = \beta = \gamma = 1/3$, have to be used, such that

$$\mathcal{E} = \frac{1}{3}\int \Big[g\,[f,h] + f\,[h,g] + h\,[g,f]\Big]\,dx\,dp. \tag{B.9}$$

B.1.1. Discrete Poisson Brackets on a Rectangular Mesh

The next step in the derivation is the discretisation of this integral. Salmon and Talley discretise the derivatives along the diagonals of a grid cell as depicted below.
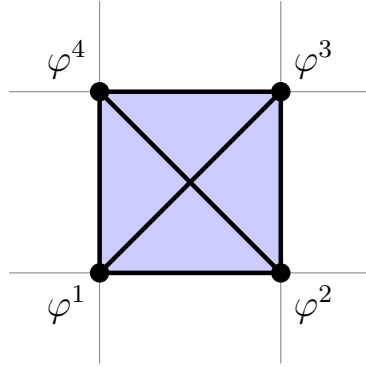
Where the discrete derivatives are defined as

$$\varphi_x^{\boxtimes}(x,p) \approx \frac{\varphi^2 - \varphi^4}{h_x}, \qquad\qquad \varphi_p^{\boxtimes}(x,p) \approx \frac{\varphi^3 - \varphi^1}{h_p}. \tag{B.10}$$

This definition of the derivatives appears quite unnatural and ambiguous as both the x and the p derivative could be defined along each of the two diagonals. The actual choice seems mostly motivated by the desired result, to obtain Arakawa's discretisation of the brackets (details follow below). However, we will see that this specific discretisation is also obtained by using the discrete derivatives from chapter 3 instead. I.e., we define the derivatives along the edges of the grid cell, that is

$$\varphi_x^{\square}(x,p) \approx \frac{1}{2}\left(\frac{\varphi^2 - \varphi^1}{h_x} + \frac{\varphi^3 - \varphi^4}{h_x}\right), \qquad \varphi_p^{\square}(x,p) \approx \frac{1}{2}\left(\frac{\varphi^4 - \varphi^1}{h_p} + \frac{\varphi^3 - \varphi^2}{h_p}\right). \tag{B.11}$$

As in chapter 3, the fields are approximated by averages over all four vertices of the grid cell,

$$\varphi^{\square}(x,p) \approx \frac{1}{4}\left(\varphi^1 + \varphi^2 + \varphi^3 + \varphi^4\right). \tag{B.12}$$

Upon defining

$$S^{\square}(g,f,h) \equiv \frac{1}{4}\left(g^1 + g^2 + g^3 + g^4\right)\left[\frac{1}{2}\left(\frac{f^2 - f^1}{h_x} + \frac{f^3 - f^4}{h_x}\right)\frac{1}{2}\left(\frac{h^4 - h^1}{h_p} + \frac{h^3 - h^2}{h_p}\right)\right.$$
$$\left. - \frac{1}{2}\left(\frac{f^4 - f^1}{h_p} + \frac{f^3 - f^2}{h_p}\right)\frac{1}{2}\left(\frac{h^2 - h^1}{h_x} + \frac{h^3 - h^4}{h_x}\right)\right], \quad \text{(B.13)}$$

the discrete equivalent of the energy functional (B.9) becomes

$$\mathcal{E}_d = \frac{1}{3}\sum_{\square}\left(S^{\square}(g,f,h) + S^{\square}(f,h,g) + S^{\square}(h,g,f)\right). \tag{B.14}$$

The discrete Poisson brackets at a grid point (i,j) are obtained by computing a discrete functional derivative, the same way as we computed the discrete variation, that is

$$[f,h]_{i,j} = \frac{\partial \mathcal{E}_d}{\partial g_{i,j}}. \tag{B.15}$$

Only four terms of the sum (B.15) include $g_{i,j}$, such that those four terms define the discrete Poisson bracket

$$[f,h]_{i,j} = \frac{\partial \mathcal{E}_d}{\partial g^1}\left(\varphi_{i,j}, \varphi_{i+1,j}, \varphi_{i+1,j+1}, \varphi_{i,j+1}\right) + \frac{\partial \mathcal{E}_d}{\partial g^2}\left(\varphi_{i-1,j}, \varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i-1,j+1}\right)$$
$$+ \frac{\partial \mathcal{E}_d}{\partial g^3}\left(\varphi_{i-1,j-1}, \varphi_{i,j-1}, \varphi_{i,j}, \varphi_{i-1,j}\right) + \frac{\partial \mathcal{E}_d}{\partial g^4}\left(\varphi_{i,j-1}, \varphi_{i+1,j-1}, \varphi_{i+1,j}, \varphi_{i,j}\right). \tag{B.16}$$

The result of this computation is the well-known Arakawa scheme [4].

B.1.2. Arakawa's Discretisation

In his original work [4], Arakawa considers different discretisations of the Poisson brackets with the aim of preserving the total number of particles

$$\int [f,h]\,dx\,dp = 0 \qquad \rightarrow \qquad \int f(t)\,dx\,dp = \int f(0)\,dx\,dp, \tag{B.17}$$

the total energy

$$\int [f,h]\,h\,dx\,dp = 0 \qquad \rightarrow \qquad \int f(t)\,h(t)\,dx\,dp = \int f(0)\,h(0)\,dx\,dp, \tag{B.18}$$

and the L^2 norm of the distribution function

$$\int [f,h]\,f\,dx\,dp = 0 \qquad \rightarrow \qquad \int f^2(t)\,dx\,dp = \int f^2(0)\,dx\,dp. \tag{B.19}$$

He defines four different discretisations of $J(f, h) = [f, h]$, that is

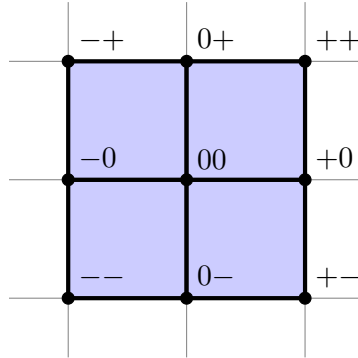$$J^{++} = \frac{1}{4 h_x h_p} \left(\left(f_{+0} - f_{-0} \right) \left(h_{0+} - h_{0-} \right) - \left(f_{0+} - f_{0-} \right) \left(h_{+0} - h_{-0} \right) \right), \tag{B.20a}$$

$$J^{+\times} = \frac{1}{4 h_x h_p} \left[f_{+0} \left(h_{+-} - h_{++} \right) - f_{-0} \left(h_{--} - h_{-+} \right) - f_{0+} \left(h_{-+} - h_{++} \right) + f_{0-} \left(h_{--} - h_{+-} \right) \right]$$

$$= \frac{1}{4 h_x h_p} \left[h_{++} \left(f_{0+} - f_{+0} \right) - h_{--} \left(f_{-0} - f_{0-} \right) - h_{-+} \left(f_{0+} - f_{-0} \right) + h_{+-} \left(f_{+0} - f_{0-} \right) \right], \tag{B.20b}$$

$$J^{\times+} = \frac{1}{4 h_x h_p} \left[f_{++} \left(h_{+0} - h_{0+} \right) - f_{--} \left(h_{0-} - h_{-0} \right) - f_{-+} \left(h_{-0} - h_{0+} \right) + f_{+-} \left(h_{0-} - h_{+0} \right) \right]$$

$$= \frac{1}{4 h_x h_p} \left[h_{+0} \left(f_{++} - f_{+-} \right) - h_{-0} \left(f_{-+} - f_{--} \right) - h_{0+} \left(f_{++} - f_{-+} \right) + h_{0-} \left(f_{+-} - f_{--} \right) \right], \tag{B.20c}$$

$$J^{\times\times} = \frac{1}{8 h_x h_p} \left(\left(f_{++} - f_{--} \right) \left(h_{-+} - h_{+-} \right) - \left(f_{-+} - f_{+-} \right) \left(h_{++} - h_{--} \right) \right), \tag{B.20d}$$

where the subscript 00 refers to the grid point where the brackets are defined, +0 the grid point on the right, −0 the grid point on the left, and so on, as depicted below.


Arakawa then explores different combinations of those expressions to find that all three of the above conservation properties (B.17) - (B.19) are only respected for

$$J = \frac{1}{3} \left(J^{++} + J^{+\times} + J^{\times+} \right), \tag{B.21}$$

which is exactly the expression we obtain from (B.16). The above expressions correspond to a second order discretisation of the Poisson bracket. Arakawa also provides fourth order expressions. It would be very interesting to see, if these can be derived by a similar approach. This, however, is a problem left for future research.

B.1.3. Discrete Poisson Brackets on a Triangular Mesh

A similar discretisation of the brackets can be performed on a mesh of triangles, leading to the scheme of Sadourny et al. [106]. The only complication comes with the fact that we have two kinds of triangles, namely those pointing upward and those pointing downward, and we have, of course, to consider the contribution of both.
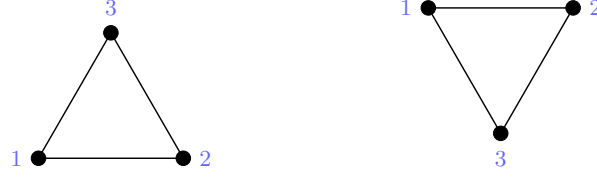
Therefore we need to define cell averages and discrete derivatives for each type of triangle separately. The grid points of the triangles are given by

$$\triangle = \big((i,j),(i+1,j),(i,j+1)\big) \qquad \triangledown = \big((i,j),(i+1,j),(i+1,j-1)\big), \tag{B.22}$$

such that the vertices are denoted by

$$\triangle^1 = (i,j), \qquad \triangle^2 = (i+1,j), \qquad \triangle^3 = (i,j+1), \tag{B.23a}$$
$$\triangledown^1 = (i,j), \qquad \triangledown^2 = (i+1,j), \qquad \triangledown^3 = (i+1,j-1). \tag{B.23b}$$

Field averages are the same on both kind of triangles but the derivatives are different, that is

$$\varphi^\triangle = \frac{1}{3}\left(\varphi^{\triangle^1} + \varphi^{\triangle^2} + \varphi^{\triangle^3}\right), \qquad\qquad \varphi^\triangledown = \frac{1}{3}\left(\varphi^{\triangledown^1} + \varphi^{\triangledown^2} + \varphi^{\triangledown^3}\right), \tag{B.24a}$$

$$\varphi_x^\triangle = \frac{\varphi^{\triangle^2} - \varphi^{\triangle^1}}{h_x}, \qquad\qquad \varphi_x^\triangledown = \frac{\varphi^{\triangledown^2} - \varphi^{\triangledown^1}}{h_x}, \tag{B.24b}$$

$$\varphi_p^\triangle = \frac{1}{2}\left(\frac{\varphi^{\triangle^3} - \varphi^{\triangle^1}}{h_p} + \frac{\varphi^{\triangle^3} - \varphi^{\triangle^2}}{h_p}\right), \qquad \varphi_p^\triangledown = \frac{1}{2}\left(\frac{\varphi^{\triangledown^1} - \varphi^{\triangledown^2}}{h_p} + \frac{\varphi^{\triangledown^1} - \varphi^{\triangledown^3}}{h_p}\right). \tag{B.24c}$$

Upon defining

$$S^\triangle(g,f,h) \equiv \frac{1}{3}\left(\varphi^{\triangle^1} + \varphi^{\triangle^2} + \varphi^{\triangle^3}\right)\left[\frac{1}{2}\left(\frac{f^{\triangle^2} - f^{\triangle^1}}{h_x}\right)\left(\frac{h^{\triangle^3} - h^{\triangle^1}}{h_p} + \frac{h^{\triangle^3} - h^{\triangle^2}}{h_p}\right)\right.$$
$$\left. - \frac{1}{2}\left(\frac{f^{\triangle^3} - f^{\triangle^1}}{h_p} + \frac{f^{\triangle^3} - f^{\triangle^2}}{h_p}\right)\left(\frac{h^{\triangle^2} - h^{\triangle^1}}{h_x}\right)\right], \tag{B.25a}$$

$$S^\triangledown(g,f,h) \equiv \frac{1}{3}\left(\varphi^{\triangledown^1} + \varphi^{\triangledown^2} + \varphi^{\triangledown^3}\right)\left[\frac{1}{2}\left(\frac{f^{\triangledown^2} - f^{\triangledown^1}}{h_x}\right)\left(\frac{h^{\triangledown^1} - h^{\triangledown^2}}{h_p} + \frac{h^{\triangledown^1} - h^{\triangledown^3}}{h_p}\right)\right.$$
$$\left. - \frac{1}{2}\left(\frac{f^{\triangledown^1} - f^{\triangledown^2}}{h_p} + \frac{f^{\triangledown^1} - f^{\triangledown^3}}{h_p}\right)\left(\frac{h^{\triangledown^2} - h^{\triangledown^1}}{h_x}\right)\right], \tag{B.25b}$$

the discrete energy functional (B.9) becomes

$$\mathcal{E}_d = \sum_\triangle \frac{1}{3}\Big(S^\triangle(g,f,h) + S^\triangle(f,h,g) + S^\triangle(h,g,f)\Big)$$
$$+ \sum_\triangledown \frac{1}{3}\Big(S^\triangledown(g,f,h) + S^\triangledown(f,h,g) + S^\triangledown(h,g,f)\Big). \tag{B.26}$$

The discrete Poisson brackets at a grid point (i,j) are obtained by computing the discrete functional derivative as in (B.15), that is

$$[f,h]_{i,j} = \frac{\partial \mathcal{E}_d}{\partial g_{i,j}}. \tag{B.27}$$

Six terms of the sum (B.26) include $g_{i,j}$, such that those terms define the discrete Poisson bracket

$$[f,h]_{i,j} = \frac{\partial \mathcal{E}_d}{\partial g^{\triangle 1}}\Big(\varphi_{i,j}, \varphi_{i+1,j}, \varphi_{i,j+1}\Big) + \frac{\partial \mathcal{E}_d}{\partial g^{\triangle 2}}\Big(\varphi_{i-1,j}, \varphi_{i,j}, \varphi_{i-1,j+1}\Big)$$
$$+ \frac{\partial \mathcal{E}_d}{\partial g^{\triangle 3}}\Big(\varphi_{i,j-1}, \varphi_{i+1,j-1}, \varphi_{i,j}\Big) + \frac{\partial \mathcal{E}_d}{\partial g^{\triangledown 1}}\Big(\varphi_{i,j}, \varphi_{i,j+1}, \varphi_{i-1,j+1}\Big)$$
$$+ \frac{\partial \mathcal{E}_d}{\partial g^{\triangledown 2}}\Big(\varphi_{i,j-1}, \varphi_{i,j}, \varphi_{i-1,j}\Big) + \frac{\partial \mathcal{E}_d}{\partial g^{\triangledown 3}}\Big(\varphi_{i+1,j-1}, \varphi_{i+1,j}, \varphi_{i,j}\Big).$$
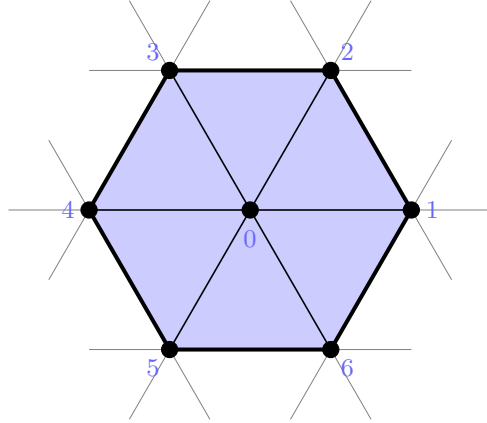
(B.28)

The result of this computation is the scheme of Sadourny et al. [106],

$$[f,h]_d = \frac{1}{A_{\bigcirc}} \sum_{a=1}^{6} \frac{1}{2} f_a \Big(h_{a+1} - h_{a-1}\Big)$$

(B.29)

where the area of the hexagon, $A_{\bigcirc}$, is

$$A_{\bigcirc} = 6\,A_{\triangle} = 3\,h_x\,h_p \cos 30°,$$

(B.30)

a denotes the vertices of the hexagon as depicted below, $a = 6 + 1$ is replaced with $a = 1$, and $a = 1 - 1$ is replaced with $a = 6$.

B.2. Nambu Three Brackets

The ideas of the previous section can be extended to the three brackets defined by Nambu [92]

$$[f,g,h] \equiv \epsilon^{abc} f_{,a}\, g_{,b}\, h_{,c} = \frac{\partial f}{\partial x}\big[g,h\big]_{yz} + \frac{\partial f}{\partial y}\big[g,h\big]_{zx} + \frac{\partial f}{\partial z}\big[g,h\big]_{xy}, \quad (a,b,c) \in \{x,y,z\},$$

(B.31)

where $[\cdot,\cdot]_{xy}$ denote Poisson brackets with respect to x and y, etc.. The energy functional is completely analogous to the previous one, that is

$$\mathcal{E} = \int k\,[f,g,h]\,dx\,dy\,dz$$

(B.32)

$$= \frac{1}{4}\int \Big(k\,[f,g,h] + f\,[k,h,g] + g\,[k,f,h] + h\,[k,g,f]\Big)dx\,dy\,dz,$$

(B.33)

where we applied the same symmetrisation ideas as before.

B.2.1. Discrete Nambu Brackets

The discrete derivatives are defined similar as before, only that now we have to consider a three dimensional grid, such that the field averages and discrete derivatives become

$$\varphi^\square = \frac{1}{8} \left(\varphi^1 + \varphi^2 + \varphi^3 + \varphi^4 + \varphi^5 + \varphi^6 + \varphi^7 + \varphi^8 \right) \tag{B.34a}$$

$$\varphi_x^\square = \frac{1}{4} \left(\frac{\varphi^2 - \varphi^1}{h_x} + \frac{\varphi^3 - \varphi^4}{h_x} + \frac{\varphi^6 - \varphi^5}{h_x} + \frac{\varphi^7 - \varphi^8}{h_x} \right), \tag{B.34b}$$

$$\varphi_y^\square = \frac{1}{4} \left(\frac{\varphi^4 - \varphi^1}{h_y} + \frac{\varphi^3 - \varphi^2}{h_y} + \frac{\varphi^8 - \varphi^5}{h_y} + \frac{\varphi^7 - \varphi^6}{h_y} \right), \tag{B.34c}$$

$$\varphi_z^\square = \frac{1}{4} \left(\frac{\varphi^5 - \varphi^1}{h_z} + \frac{\varphi^6 - \varphi^2}{h_z} + \frac{\varphi^7 - \varphi^3}{h_z} + \frac{\varphi^8 - \varphi^4}{h_z} \right). \tag{B.34d}$$

The discrete action follows exactly along the lines of the previous derivations. The exact form of the discrete brackets is quite complex and therefore not explicitly repeated here.

B.2.2. Application to Gyrokinetics

An application of this formulation and our discretisation is the gyrokinetic Vlasov equation on an extruded triangular mesh as depicted below [117],

$$\frac{\partial f}{\partial t} + \frac{1}{\sqrt{g}\, B_\parallel^*} \left[h, f, A_\varphi^* \right]_{xyp_z} = 0, \tag{B.35}$$

where (x, y) are coordinates in the poloidal plane of an axisymmetric tokamak, φ is the toroidal coordinate and p_z is the parallel momentum, g is the metric, f the distribution function and h the particle Hamiltonian, $A^* = A + \frac{c}{e} p_z b$ is the generalised vector potential and $B_\parallel^* = b \cdot (\nabla \times A^*)$ the parallel magnetic field strength.
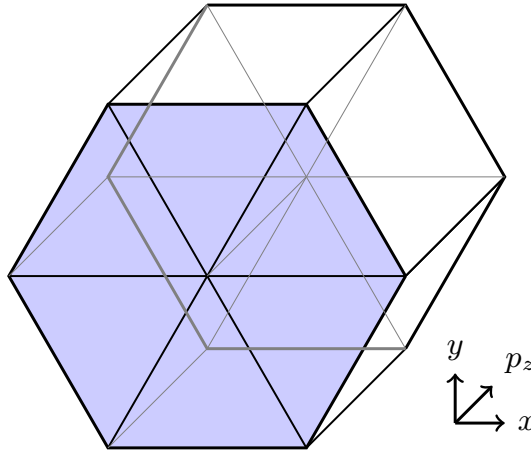
Here, we have to combine the ideas of the previous sections. The spatial derivatives are defined in the poloidal plane over triangles and averaged in the momentum coordinate. Again we have to consider upward pointing and downward pointing triangles separately. The momentum derivative is defined in the extruded direction and averaged over the vertices of each triangle. Field averages are defined over all vertices of a three-dimensional grid cell. The rest of the derivation follows straight forwardly.

B.3. Lie-Poisson and Nambu Field Brackets

The natural Hamiltonian formulation of the Vlasov equation is via Lie-Poisson brackets. As we discussed in chapter five, the Vlasov-Poisson system does not have a canonical Hamiltonian formulation, nevertheless it is a Hamiltonian system. In this section we want to explore a discretisation approach for such Lie-Poisson brackets.

After sketching the noncanonical Hamiltonian description of systems like Vlasov-Poisson, we review Salmon's approach [109] of discretising Lie-Poisson brackets. He does not discretise the brackets directly, but first finds equivalent infinite dimensional Nambu brackets (Nambu field brackets) and then discretises those by similar ideas as they were presented in the previous section.

B.3.1. Noncanonical Hamiltonian Field Theory

The dynamics of Hamiltonian systems is usually expressed with the help of canonically conjugate variables (q, p) and Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p} \qquad \text{and} \qquad \dot{p} = -\frac{\partial H}{\partial q}. \tag{B.36}$$

The evolution of any functional $F(q, p)$ can be obtained with Poisson brackets, i.e.,

$$\dot{F}(q, p) = [F, H]. \tag{B.37}$$

A large class of Hamiltonian systems, especially infinite-dimensional ones and especially systems from plasma physics (e.g., Vlasov-Poisson, reduced and ideal MHD, incompressible Fluid dynamics), do not fit into the form of (B.36). They can, however, be described by a generalisation of (B.37). The dynamics of a functional $F(\xi)$ of state variables $\xi(t, x)$ (e.g., distribution function, vorticity, density, temperature) of a Hamiltonian system is determined by

$$\dot{F}(\xi) = \{F, H\}, \tag{B.38}$$

where $H(\xi)$ is the Hamiltonian functional and $\{\cdot, \cdot\}$ are generalised Poisson brackets, that means they are antisymmetric and fulfil the Leibniz rule and the Jacobi identity.

B.3.2. Lie-Poisson and Nambu Brackets in the Vlasov Equation

The Vlasov equation can be express in terms of Lie-Poisson brackets [82] as

$$\dot{F} = \{F, H\} \equiv \int f \left[\frac{\delta F}{\delta f}, \frac{\delta H}{\delta f} \right] dx \, dp, \tag{B.39}$$

where F is any functional of f and H is the total energy functional

$$H = \int \frac{|p|^2}{2m} f(x, p) \, dx \, dp + \frac{1}{2} \int |\nabla \phi(x)|^2 \, dx. \tag{B.40}$$

Following Salmon's considerations for the vorticity equation [109], the Lie-Poisson bracket of the Vlasov equation can be expressed as a Nambu field bracket (see also [16]). Therefore we just have to replace the single f in (B.39) with the functional derivative of the L^2 norm,

$$Z = \frac{1}{2} \int f^2 \, dx \, dp. \tag{B.41}$$

Hence, the Lie-Poisson bracket in (B.39) becomes a Nambu three bracket,

$$\dot{F} = \{F, H, Z\} \equiv \int \frac{\delta Z}{\delta f} \left[\frac{\delta F}{\delta f}, \frac{\delta H}{\delta f} \right] dx \, dp. \tag{B.42}$$

This bracket is antisymmetric in its three parameters, a property that is important in the discretisation procedure.

### B.3.3. Discretisation of Nambu Field Brackets

The functionals Z and H are approximated by a simple quadrature rule as

$$Z = \frac{1}{2} \sum_{i,j} f_{i,j}^2, \qquad H = \sum_{i,j} f_{i,j} \, h_{i,j} = \sum_{i,j} f_{i,j} \left(p_j^2/m + q\phi_{i,j} \right). \tag{B.43}$$

As in the previous sections, the key observation to a successful discretisation is the equivalence of the following expressions (integrating by parts with appropriate boundary conditions)

$$\int \frac{\delta Z}{\delta f} \left[\frac{\delta F}{\delta f}, \frac{\delta H}{\delta f} \right] dx \, dp = \int \frac{\delta F}{\delta f} \left[\frac{\delta H}{\delta f}, \frac{\delta Z}{\delta f} \right] dx \, dp = \int \frac{\delta H}{\delta f} \left[\frac{\delta Z}{\delta f}, \frac{\delta F}{\delta f} \right] dx \, dp. \tag{B.44}$$

The functional derivatives are defined on each vertex of a grid cell as we know it from the previous chapters and we define averages and derivatives just as before, c.f. equation (B.34).
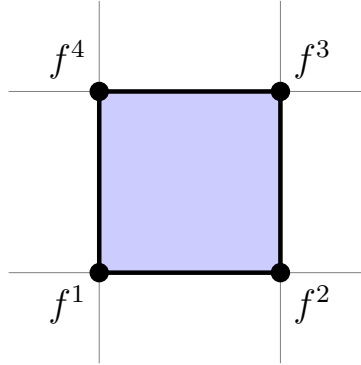
The functional derivative of the L^2 norm is therefore discretised as

$$\frac{\delta Z}{\delta f} \approx \frac{1}{4} \left(\frac{\partial Z}{\partial f^1} + \frac{\partial Z}{\partial f^2} + \frac{\partial Z}{\partial f^3} + \frac{\partial Z}{\partial f^4} \right), \tag{B.45}$$

and the Poisson brackets in B.42 are discretised by

$$\left[\frac{\delta F}{\delta f}, \frac{\delta H}{\delta f} \right]_d = \frac{1}{4h_x h_p} \left(\left(\frac{\partial F}{\partial f^2} - \frac{\partial F}{\partial f^1} + \frac{\partial F}{\partial f^3} - \frac{\partial F}{\partial f^4} \right) \left(\frac{\partial H}{\partial f^4} - \frac{\partial H}{\partial f^1} + \frac{\partial H}{\partial f^3} - \frac{\partial H}{\partial f^2} \right) \right.$$
$$\left. - \left(\frac{\partial F}{\partial f^4} - \frac{\partial F}{\partial f^1} + \frac{\partial F}{\partial f^3} - \frac{\partial F}{\partial f^2} \right) \left(\frac{\partial H}{\partial f^2} - \frac{\partial H}{\partial f^1} + \frac{\partial H}{\partial f^3} - \frac{\partial H}{\partial f^4} \right) \right), \tag{B.46}$$

such that the discrete Nambu field bracket can be written as

$$\{F, H, Z\}_{\widetilde{d}} = \sum_{\text{grid boxes}} \frac{1}{4} \left(\frac{\partial Z}{\partial f^1} + \frac{\partial Z}{\partial f^2} + \frac{\partial Z}{\partial f^3} + \frac{\partial Z}{\partial f^4} \right) \left[\frac{\delta F}{\delta f}, \frac{\delta H}{\delta f} \right]_d. \tag{B.47}$$

To retain the antisymmetry property of the continuous Nambu bracket on the the discrete level, this expression has to be symmetrised, taking into account all even as well as all odd permutations

$$\{F, H, Z\}_d = \frac{1}{6} \Big(\{F, H, Z\}_{\widetilde{d}} + \{H, Z, F\}_{\widetilde{d}} + \{Z, F, H\}_{\widetilde{d}}$$
$$- \{F, Z, H\}_{\widetilde{d}} - \{H, F, Z\}_{\widetilde{d}} - \{Z, H, F\}_{\widetilde{d}} \Big). \tag{B.48}$$

With the discretisations we described above, considering only the even permutations suffices, but in general this is not the case. The semi-discrete analogue of the Vlasov-Poisson equation is then

$$\frac{\partial f_{ij}}{\partial t} = \{f_{ij}, H, Z\}_d. \tag{B.49}$$

With our discretisation of the derivatives, this will again lead to the Arakawa discretisation. It is an interesting observation that derivations on the Lagrangian side (variational integrators) and on the Hamiltonian side (discrete Nambu brackets) lead to similar discretisations of the equations of motion. This is due to the fact that both the extended Lagrangian formulation and the Lie-Poisson brackets (B.39) are constructed on the basis of the particle brackets $[\cdot, \cdot]$. This appears to be a consequence of the use of extended Lagrangians.

It will be interesting to further develop this approach as Lie-Poisson brackets exists for a wide range of system from plasma physics, like the Vlasov-Maxwell system and different flavours of magnetohydrodynamics.

Bibliography

[1] Ralph Abraham and Jerrold E. Marsden. *Foundations of Mechanics.* American Mathematical Society, 1978. URL `http://authors.library.caltech.edu/25029/`.

[2] Ralph Abraham, Jerrold E. Marsden, and Tudor S. Ratiu. *Manifolds, Tensor Analysis and Application.* Springer, 1988.

[3] V. Aldaya and J. A. De Azcarraga, 1980, Geometric formulation of classical mechanics and field theory, *La Rivista del Nuovo Cimento*, **3**(10) 1–66. doi: 10.1007/BF02906204.

[4] Akio Arakawa, 1966, Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. part i, *Journal of Computational Physics*, **1** 119–143. doi: 10.1016/0021-9991(66)90015-5.

[5] T. D. Arber and R. G. L. Vann, 2002, A Critical Comparison of Eulerian-Grid-Based Vlasov Solvers, *Journal of Computational Physics*, **180** 339–357. doi: 10.1006/jcph.2002.7098.

[6] Douglas N. Arnold, Richard S. Falk, and Ragnar Winther, 2006, Finite element exterior calculus, homological techniques, and applications, *Acta Numerica*, **15** 1–155. doi: 10.1017/S0962492906210018. URL `http://journals.cambridge.org/article_S0962492906210018`.

[7] Douglas N. Arnold, Richard S. Falk, and Ragnar Winther, 2010, Finite element exterior calculus: From hodge theory to numerical stability, *Bulletin of the American Mathematical Society*, **47** 281–354. doi: http://dx.doi.org/10.1090/S0273-0979-10-01278-4. arXiv:0906.4325.

[8] Vladimir I. Arnold. *Mathematical Methods of Classical Mechanics.* Springer, 1989.

[9] Vladimir I. Arnold and Boris A. Khesin. *Topological Methods in Hydrodynamics.* Springer, 1998.

[10] Blanca de Dios Ayuso and Soheil Hajian. High order and energy preserving discontinuous galerkin methods for the vlasov-poisson system. arXiv:1209.4025, 2012.

[11] Blanca De Dios Ayuso, José A. Carrillo, and Chi-Wang Shu. Discontinuous Galerkin Methods for the One-Dimensional Vlasov-Poisson System. Technical Report 2009-41, Scientific Computing Group, Brown University, Providence, RI, USA, 2009.

[12] Blanca De Dios Ayuso, José A. Carrillo, and Chi-Wang Shu, 2012, Discontinuous galerkin methods for the multi-dimensional vlasov–poisson problem, *Mathematical Models and Methods in Applied Sciences*, **22** 1250042. doi: 10.1142/S021820251250042X.

[13] Satish Balay, Jed Brown, , Kris Buschelman, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Barry F. Smith, and Hong Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 3.3, Argonne National Laboratory, 2012.

[14] Satish Balay, Jed Brown, Kris Buschelman, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Barry F. Smith, and Hong Zhang. PETSc Web page, 2012. http://www.mcs.anl.gov/petsc.

[15] Nicolas Besse and Eric Sonnendrücker, 2003, Semi-lagrangian schemes for the vlasov equation on an unstructured mesh of phase space, *Journal of Computational Physics*, **191** 341–376. doi: 10.1016/S0021-9991(03)00318-8.

[16] Iwo Bialynicki-Birula and Philip J. Morrison, 1991, Quantum mechanics as a generalization of Nambu dynamics to the Weyl-Wigner formalism, *Physics Letters A*, **158** 453–457. doi: http://dx.doi.org/10.1016/0375-9601(91)90458-K.

[17] James J. Binney and Scott Tremaine. *Galactic Dynamics.* Princeton University Press, 2008.

[18] Dieter Biskamp. *Magnetohydrodynamic Turbulence.* Cambridge University Press, 2003.

[19] Alain J. Brizard, 2000, New variational principle for the vlasov-maxwell equations, *Physical Review Letters*, **84** 5768–5771. doi: 10.1103/PhysRevLett.84.5768.

[20] Alain J. Brizard, 2000, Variational principle for nonlinear gyrokinetic vlasov–maxwell equations, *Physics of Plasmas*, **7** 4816. doi: 10.1063/1.1322063.

[21] Alain J. Brizard, 2011, Compact Formulas for Guiding-Center Orbits in Axisymmetric Tokamak Geometry, *Physics of Plasmas*, **18** 022508. doi: 10.1063/1.3554696. arXiv:1011.2401.

[22] Chris J. Budd and Matthew D. Piggott. Geometric integration and its applications. In *in Handbook of numerical analysis*, Pages 35–139. North-Holland, 2000.

[23] William L. Burke. *Applied Differential Geometry.* Cambridge University Press, 1985.

[24] Sean Carroll. *Spacetime and Geometry: An Introduction to General Relativity.* Addison-Wesley, 2003.

[25] Hernán Cendra, Darryl D. Holm, Mark J. W. Hoyle, and Jerrold E. Marsden, 1998, The Maxwell–Vlasov Equations in Euler–Poincaré Form, *Journal of Mathematical Physics*, **39** 3138 – 3157. doi: 10.1063/1.532244. arXiv:chao-dyn/9801016.

[26] Cristel Chandre, Loïc De Guillebon, Aurore Back, Emanuele Tassi, and Philip J. Morrison. On the use of projectors for hamiltonian systems and their relationship with dirac brackets. arXiv:1205.2347, 2012.

[27] Cristel Chandre, Philip J. Morrison, and Emanuele Tassi, 2012, On the hamiltonian formulation of incompressible ideal fluids and magnetohydrodynamics via diracê¼s theory of constraints, *Physics Letters A*, **376**(5) 737–743. doi: 10.1016/j.physleta.2011.12.015. arXiv:1110.6891.

[28] C. Z. Cheng and Georg Knorr, 1976, The integration of the vlasov equation in config-
uration space, *Journal of Computational Physics*, **22** 330–351. doi: 10.1016/0021-
9991(76)90053-X.

[29] Yingda Cheng and Irene M Gamba, 2012, Numerical study of one-dimensional
Vlasov-Poisson equations for infinite homogeneous stellar systems, *Communi-
cations in Nonlinear Science and Numerical Simulation*, **17** 2052–2061. doi:
10.1016/j.cnsns.2011.10.004.

[30] Yingda Cheng, Irene M. Gamba, and Philip J. Morrison, 2013, Study of conservation
and recurrence of runge–kutta discontinuous galerkin schemes for vlasov–poisson
systems, *Journal of Scientific Computing*, 1–31. doi: 10.1007/s10915-012-9680-x.
URL `http://dx.doi.org/10.1007/s10915-012-9680-x`.

[31] Snorre H. Christiansen, Hans Z. Munthe-Kaas, and Brynjulf Owren, 2011,
Topics in structure-preserving discretization, *Acta Numerica*, **20** 1–119. doi:
10.1017/S096249291100002X. URL `http://journals.cambridge.org/article_`
`S096249291100002X`.

[32] Phillip C. Clemmow and John P. Dougherty. *Electrodynamics of Particles and
Plasmas*. Addison-Wesley, 1969.

[33] Diego Cordoba and Christiane Marliani, 2000, Evolution of current sheets
and regularity of ideal incompressible magnetic fluids in 2d, *Communications
on Pure and Applied Mathematics*, **53** 512–524. doi: 10.1002/(SICI)1097-
0312(200004)53:4<512::AID-CPA4>3.0.CO;2-R.

[34] Peter A. Davidson. *An Introduction to Magnetohydrodynamics*. Cambridge Univer-
sity Press, 2001.

[35] Arturo Echeverria-Enríquez, Miguel C. Muñoz-Lecanda, and Narciso Román-Roy,
1996, Geometry of lagrangian first-order classical field theories, *Fortschritte der
Physik*, **44** 235–280. doi: 10.1002/prop.2190440304. arXiv:dg-ga/9505004.

[36] Arturo Echeverria-Enríquez, Miguel C. Muñoz-Lecanda, and Narciso Román-Roy,
2000, Geometry of multisymplectic hamiltonian first-order field theories, *Journal of
Mathematical Physics*, **41** 7402. doi: 10.1063/1.1308075. arXiv:math-ph/0004005.

[37] Marcelo Epstein. *The Geometrical Language of Continuum Mechanics*. Cambridge
University Press, 2010.

[38] Francis Filbet and Eric Sonnendrücker, 2003, Comparison of eulerian vlasov
solvers, *Computer Physics Communications*, **150** 247–266. doi: 10.1016/S0010-
4655(02)00694-X.

[39] Tor Flå, 1994, Action principle and the hamiltonian formulation for the maxwell–
vlasov equations on a symplectic leaf, *Physics of Plasmas*, **1** 2409. doi:
10.1063/1.870569.

[40] Tor Flå and Michael Kraus. Private communication. 2011.

[41] Jeffrey P. Freidberg. *Ideal Magnetohydrodynamics*. Plenum Press, 1987.

[42] Thomas A. Gardiner and James M. Stone, 2005, An unsplit godunov method for ideal mhd via constrained transport, *Journal of Computational Physics*, **205** 509–539. doi: 10.1016/j.jcp.2004.11.016. arXiv:astro-ph/0501557.

[43] Evan S. Gawlik, Patrick Mullen, Dmitry Pavlov, Jerrold E. Marsden, and Mathieu Desbrun, 2011, Geometric, variational discretization of continuum theories, *Physica D: Nonlinear Phenomena*, **240** 1724–1760. doi: 10.1016/j.physd.2011.07.011. arXiv:1010.4851.

[44] Israel M. Gelfand and Sergei V. Fomin. *Calculus of Variations*. Prentice-Hall, 1963.

[45] Giovanni Giachetta, Luigi Mangiarotti, and Gennadi A. Sardanashvili. *New Lagrangian and Hamiltonian Methods in Field Theory*. World Scientific, 1997. URL `http://gnsardan.appfarm.ru/lp_mo.html`.

[46] Giovanni Giachetta, Luigi Mangiarotti, and Gennadi A. Sardanashvili. *Advanced Mechanics. Mathematical Introduction*. 2009. arXiv:0911.0411.

[47] Giovanni Giachetta, Luigi Mangiarotti, and Gennadi A. Sardanashvili. *Advanced Classical Field Theory*. World Scientific, 2010. URL `http://gnsardan.appfarm.ru/lp_mo.html`.

[48] Johan Peter Goedbloed and Stefaan Poedts. *Principles of Magnetohydrodynamics*. Cambridge University Press, 2004.

[49] Mark J. Gotay, James Isenberg, and Jerrold E. Marsden. Momentum maps and classical fields, 1998.

[50] Daniela Grasso, Francesco Califano, Francesco Pegoraro, and Francesco Porcelli, 2001, Phase mixing and island saturation in hamiltonian reconnection, *Phys. Rev. Lett.*, **86** 5051–5054. doi: 10.1103/PhysRevLett.86.5051.

[51] Daniela Grasso, Emanuele Tassi, and François L. Waelbroeck, 2010, Nonlinear gyrofluid simulations of collisionless reconnection, *Physics of Plasmas*, **17** 082312. doi: 10.1063/1.3475440.

[52] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration*. Springer, 2006.

[53] R. E. Heath, Irene M. Gamba, Philip J. Morrison, and C. Michler, 2012, A discontinuous galerkin method for the vlasov–poisson system, *Journal of Computational Physics*, **231** 1140–1174. doi: 10.1016/j.jcp.2011.09.020.

[54] M. P. Hobson, G. P. Efstathiou, and A. N. Lasenby. *General Relativity: An Introduction for Physicists*. Cambridge University Press, 2006.

[55] Darryl D. Holm. *Geometric Mechanics*. 2011. URL `http://www.ma.ic.ac.uk/~dholm/classnotes/`.

[56] Darryl D. Holm, Jerrold E. Marsden, and Tudor S. Ratiu, 1998, The Euler–Poincaré Equations and Semidirect Products with Applications to Continuum Theories, *Advances in Mathematics*, **137** 1 – 81. doi: 10.1006/aima.1998.1721. arXiv:chaodyn/9801015.

[57] Darryl D. Holm, Tanya Schmah, and Cristina Stoica. *Geometric Mechanics and Symmetry.* Oxford University Press, 2009.

[58] Nail H. Ibragimov, 2006, Integrating factors, adjoint equations and lagrangians, *Journal of Mathematical Analysis and Applications*, **318** 742—757. doi: 10.1016/j.jmaa.2005.11.012.

[59] Nail H. Ibragimov, 2007, A new conservation theorem, *Journal of Mathematical Analysis and Applications*, **333** 311—328. doi: 10.1016/j.jmaa.2006.10.078.

[60] Nail H. Ibragimov, 2007, Quasi-self-adjoint differential equations, *Archives of ALGA*, **4** 55—60.

[61] Jorge V. Jose and Eugene J. Saletan. *Classical Dynamics: A Contemporary Approach.* Cambridge University Press, 1998.

[62] C. Kane, Jerrold E. Marsden, and Michael Ortiz, 1999, Symplectic-energy-momentum preserving variational integrators, *Journal of Mathematical Physics*, **40** 3353–3371. doi: 10.1063/1.532892.

[63] Yvette Kosmann-Schwarzbach. *The Noether Theorems: Invariance and Conservation Laws in the Twentieth Century.* Springer, 2010.

[64] Shinar O. Kouranbaeva. *Geometry and Analysis of the Camassa-Holm Equation: Variational Approach to Second-Order Multisymplectic Field Theory.* PhD thesis, University of California, Santa Cruz, 1999.

[65] Shinar O. Kouranbaeva and Steve Shkoller, 2000, A Variational Approach to Second-Order Multisymplectic Field Theory, *Journal of Geometry and Physics*, **25** 333–366. doi: 10.1016/S0393-0440(00)00012-7. arXiv:math/9909100.

[66] Iosif S. Krasil'shchik and Aleksandr M. Vinogradov. *Symmetries and Conservation Laws for Differential Equations of Mathematical Physics.* American Mathematical Society, 1999.

[67] Demeter Krupka, 1973, Some geometric aspects of variational problems in fibred manifolds, *Folia Fac. Sci. Nat. Univ. Purk. Brunensis, Physica*, **14**. arXiv:math-ph/0110005.

[68] Lev D. Landau, Evgeny M. Lifshitz, and Lev P. Pitaevskii. *Course of Theoretical Physics, Volume X, Physical Kinetics.* Elsevier, 1981.

[69] Hans Petter Langtangen, Kent-Andre Mardal, and Ragnar Winther, 2002, Numerical methods for incompressible viscous flow, *Advances in Water Resources*, **25** 1125–1146. doi: 10.1016/S0309-1708(02)00052-0.

[70] Jonas Larsson, 1992, An action principle for the vlasov equation and associated lie perturbation equations. part 1. the vlasov—poisson system, *Journal of Plasma Physics*, **48** 13–35. doi: 10.1017/S0022377800016342.

[71] Jonas Larsson, 1993, An action principle for the vlasov equation and associated lie perturbation equations. part 2. the vlasov–maxwell system, *Journal of Plasma Physics*, **49** 255–270. doi: 10.1017/S0022377800016974.

[72] John M. Lee. *Introduction to Smooth Manifolds.* Springer, 2012.

[73] Benedict Leimkuhler and Sebastian Reich. *Simulating Hamiltonian Dynamics.* Cambridge University Press, 2004.

[74] Andrew Lenard and Ira B. Bernstein, 1958, Plasma oscillations with diffusion in velocity space, *Physical Review*, **112** 1456–1459. doi: 10.1103/PhysRev.112.1456.

[75] Manuel León, Juan C. Marrero, and David Martín Diego, 2008, Some applications of semi-discrete variational integrators to classical field theories, *Qualitative Theory of Dynamical Systems*, **7** 195–212. doi: 10.1007/s12346-008-0011-4. URL `http://dx.doi.org/10.1007/s12346-008-0011-4`. arXiv:math-ph/0611073.

[76] Adrian Lew, Jerrold E. Marsden, Michael Ortiz, and Matthew West, 2003, Asynchronous Variational Integrators, *Archive for Rational Mechanics and Analysis*, **167** 85–146. doi: 10.1007/s00205-002-0212-y.

[77] Robert G. Littlejohn, 1983, Variational principles of guiding center motion, *Journal of Plasma Physics*, **29**(1) 111–125. doi: 10.1017/S002237780000060X.

[78] Jian-Guo Liu and Wei-Cheng Wang, 2001, An Energy-Preserving MAC–Yee Scheme for the Incompressible MHD Equation, *Journal of Computational Physics*, **174** 12–37. doi: 10.1006/jcph.2001.6772.

[79] Stephen T. Lovett. *Differential Geometry of Manifolds.* Taylor & Francis, 2010.

[80] Francis E. Low, 1958, A Lagrangian Formulation of the Boltzmann-Vlasov Equation for Plasmas, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **248** 282–287. doi: 10.1098/rspa.1958.0244.

[81] Jerrold E. Marsden. *Lectures on Mechanics.* 2009. URL `http://resolver.caltech.edu/CaltechAUTHORS:20110103-125952366`.

[82] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to Mechanics and Symmetry.* Springer, 2002.

[83] Jerrold E. Marsden and Matthew West, 2001, Discrete mechanics and variational integrators, *Acta Numerica*, **10** 357–514. doi: 10.1017/S096249290100006X.

[84] Jerrold E. Marsden, George W. Patrick, and Steve Shkoller, 1998, Multisymplectic Geometry, Variational Integrators, and Nonlinear PDEs, *Communications in Mathematical Physics*, **199** 351 – 395. doi: 10.1007/s002200050505. arXiv:math/9807080.

[85] Jerrold E. Marsden, Sergey Pekarsky, Steve Shkoller, and Matthew West, 2001, Variational Methods, Multisymplectic Geometry and Continuum Mechanics, *Journal of Geometry and Physics*, **38** 253–284. doi: 10.1016/S0393-0440(00)00066-8. arXiv:math/0005034.

[86] James M. McDonough. *Lectures in Computational Fluid Dynamics of Incompressible Flow.* University of Kentucky, 2007.

[87] Robert McLachlan, 1995, On the numerical integration of ordinary differential equations by symmetric composition methods, *SIAM Journal on Scientific Computing*, **16** 151–168.

[88] Aleksandr I. Morozov and Leonid S. Solov'ev. *Motion of Charged Particles in Electromagnetic Fields*, Volume 2 of *Reviews of Plasma Physics*, Chapter 4.3, Page 229. 1966.

[89] Philip J. Morrison, 1998, Hamiltonian Description of the Ideal Fluid, *Reviews of Modern Physics*, **70**(2) 467–521. doi: 10.1103/RevModPhys.70.467.

[90] Philip J. Morrison and Dieter Pfirsch, 1989, Free energy expressions for vlasov-maxwell equilibria, *Physical Review*, **40A** 3898–3910. doi: 10.1103/PhysRevA.40.3898.

[91] Takashi Nakamura and Takashi Yabe, 1999, Cubic interpolated propagation scheme for solving the hyper-dimensional vlasov—poisson equation in phase space, *Journal of Computational Physics*, **120** 122–154. doi: 10.1016/S0010-4655(99)00247-7.

[92] Yoichiro Nambu, 1973, Generalized Hamiltonian Dynamics, *Physical Review D*, **7** 2405–2412.

[93] Dwight E. Neuenschwander. *Emmy Noether's Wonderful Theorem.* Johns Hopkins University Press, 2010.

[94] Emmy Noether, 1918, Invariante variationsprobleme, *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen*, 235–257. arXiv:physics/0503066.

[95] Peter J. Olver. *Equivalence, Invariants and Symmetry.* Cambridge University Press, 1995.

[96] R. S. B. Ong and M. Y. Yu, 1969, The effect of weak collisions on ion-acoustic wave instabilities in a current-carrying plasma, *Journal of Plasma Physics*, **3** 425–433. doi: 10.1017/S0022377800004505.

[97] Alan Oppenheim, 1965, Wave Motion in a Plasma Based on a Fokker-Planck Equation, *Physics of Fluids*, **8** 900–911. doi: 10.1063/1.1761334.

[98] Dmitry Pavlov. *Structure-Preserving Discretization of Incompressible Fluids.* PhD thesis, California Institute of Technology, 2009. URL `http://resolver.caltech.edu/CaltechETD:etd-05222009-125630`.

[99] Dmitry Pavlov, Patrick Mullen, Yiying Tong, Eva Kanso, Jerrold E. Marsden, and Mathieu Desbrun, 2011, Structure-preserving discretization of incompressible fluids, *Physica D: Nonlinear Phenomena*, **240** 443–458. doi: 10.1016/j.physd.2010.10.012. arXiv:0912.3989.

[100] Dieter Pfirsch, 1984, New variational formulation of maxwell-vlasov and guiding center theories local charge and energy conservation laws, *Zeitschrift für Naturforschung A*, **39a** 1–8.

[101] Dieter Pfirsch and Philip J. Morrison, 1985, Local conservation laws for the maxwell-vlasov and collisionless guiding-center theories, *Physical Review A*, **32** 1714–1721. doi: 10.1103/PhysRevA.32.1714.

[102] Dieter Pfirsch and Philip J. Morrison, 1991, The energy-momentum tensor for the linearized maxwell-vlasov and kinetic guiding center theories, *Physics of Fluids B*, **3** 271–283. doi: 10.1063/1.859735.

[103] Hong Qin, Xiaoyin Guan, and William M. Tang, 2009, Variational symplectic algorithm for guiding center dynamics and its application in tokamak geometry, *Physics of Plasmas*, **16** 042510. doi: 10.1063/1.3099055.

[104] Tudor S. Ratiu. *A Crash Course in Geometric Mechanics.* 2005. URL `http://cel.archives-ouvertes.fr/cel-00391890`.

[105] Lewis Ryder. *Introduction to General Relativity.* Cambridge University Press, 2009.

[106] Robert Sadourny, Akio Arakawa, and YALE MINTZ, 1968, Integration Of The Nondivergent Barotropic Vorticity Equation With An Icosahedral-Hexagonal Grid For The Sphere, *Monthly Weather Review*, **96** 351–356. doi: doi:10.1175/1520-0493(1968)096<0351:IOTNBV>2.0.CO;2.

[107] Eugene J. Saletan and Alan H. Cromer. *Theoretical Mechanics.* John Wiley & Sons, 1971.

[108] Rick Salmon, 1988, Hamiltonian Fluid Mechanics, *Annual Review of Fluid Mechanics*, **20** 225—256. doi: 10.1146/annurev.fl.20.010188.001301.

[109] Rick Salmon, 2005, A general method for conserving quantities related to potential vorticity in numerical models, *Nonlinearity*, **18** R1. URL `http://stacks.iop.org/0951-7715/18/i=5/a=R01`.

[110] Rick Salmon and Lynne D. Talley, 1989, Generalizations of arakawa's jacobian, *Journal of Computational Physics*, **83** 247–259. doi: 10.1016/0021-9991(89)90118-6.

[111] Jesús María Sanz-Serna and Manuel P. Calvo. *Numerical Hamiltonian Problems.* Chapman and Hall, 1994.

[112] Sardanashvily. *Fibre Bundles, Jet Manifolds and Lagrangian Theory. Lectures for Theoreticians.* 2009. arXiv:0908.1886.

[113] David J. Saunders. *The Geometry of Jet Bundles.* Cambridge University Press, 1989.

[114] David J. Saunders. *Jet manifolds and natural bundles, Handbook of Global Analysis*, Pages 1035—1068. Elsevier, 2008.

[115] Dalton D. Schnack. *Lectures in Magnetohydrodynamics.* Springer, 2009.

[116] Bernard F. Schutz. *Geometrical Methods of Mathematical Physics.* Cambridge University Press, 1980.

[117] B.D. Scott, A. Kendl, and T. Ribeiro, 2010, Nonlinear dynamics in the tokamak edge, *Contributions to Plasma Physics*, **50** 228–241. doi: 10.1002/ctpp.201010039. URL `http://dx.doi.org/10.1002/ctpp.201010039`.

[118] R. L. Seliger and G. B. Whitham, 1968, Variational Principles in Continuum Mechanics, *Proceedings of the Royal Society A*, **305** 1–25. doi: 10.1098/rspa.1968.0103.

[119] Reyer Sjamaar. *Manifolds and Differential Forms.* 2006. URL `http://www.math.cornell.edu/~sjamaar/papers/manifold.pdf`.

[120] Eric Sonnendrücker. Numerical methods for the vlasov equation. Lecture Notes, 2013.

[121] Jonathan Squire, Hong Qin, and William M. Tang, 2012, Geometric integration of the Vlasov-Maxwell system with a variational particle-in-cell scheme, *Physics of Plasmas*, **19** 084501. doi: 10.1063/1.4742985.

[122] Jonathan Squire, Hong Qin, William M. Tang, and Christel Chandre, 2013, The Hamiltonian Structure and Euler-Poincaré Formulation of the Vlasov-Maxwell and Gyrokinetic Systems, *Physics of Plasmas*, **20** 022501. doi: 10.1063/1.4791664. arXiv:1301.6066.

[123] Ari Stern, Yiying Tong, Mathieu Desbrun, and Jerrold E. Marsden. Geometric computational electrodynamics with variational integrators and discrete differential forms. arXiv:0707.4470, 2009.

[124] Ari Stern, Yiying Tong, Mathieu Desbrun, and Jerrold E. Marsden. *Geometric discretization of Lagrangian mechanics and field theories.* PhD thesis, California Institute of Technology, 2009. URL `http://resolver.caltech.edu/CaltechETD:etd-12312008-173851`.

[125] E. C. G. Sudarshan and N. Mukunda. *Classical Dynamics: A Modern Perspective.* Wiley, 1974.

[126] Hideo Sugama, 2000, Gyrokinetic field theory, *Physics of Plasmas*, **7** 466. doi: 10.1063/1.873832.

[127] Donald Gary Swanson. *Plasma Kinetic Theory.* CRC Press, 2011.

[128] Emanuele Tassi, Philip J. Morrison, Daniela Grasso, and Francesco Pegoraro, 2010, Hamiltonian four-field model for magnetic reconnection: nonlinear dynamics and extension to three dimensions with externally applied fields, *Nuclear Fusion*, **50** 034007.

[129] MUMPS Team. MUMPS Web page, 2012. http://graal.ens-lyon.fr/MUMPS/.

[130] Gerardo F. Torres del Castillo. *Differentiable Manifolds: A Theoretical Physics Approach.* Birkhäuser, 2011.

[131] Loring W. Tu. *An Introduction to Manifolds.* Springer, 2011.

[132] Alexander P. Veselov, 1988, Integrable discrete-time systems and difference operators, *Functional Analysis and Its Applications*, **22** 83–93. doi: 10.1007/BF01077598.

[133] Alexander P. Veselov, 1991, Integrable lagrangian correspondences and the factorization of matrix polynomials, *Functional Analysis and Its Applications*, **25** 112–122. doi: 10.1007/BF01079590.

[134] Tomo-Hiko Watanabe and Hideo Sugama, 2005, Vlasov and drift kinetic simulation methods based on the symplectic integrator, *Transport Theory and Statistical Physics*, **34** 287–309. doi: 10.1080/00411450500274550.

[135] Matthew West. *Variational Integrators.* PhD thesis, California Institute of Technology, 2004. URL `http://resolver.caltech.edu/CaltechETD:etd-06072004-161416`.

[136] Huanchun Ye and Philip J. Morrison, 1992, Action principles for the vlasov equation, *Physics of Fluids B*, **4** 771–776. doi: 10.1063/1.860231.

[137] Kane Yee, 1966, Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media, *IEEE Transactions on Antennas and Propagation*, **14** 302–307. doi: 10.1109/TAP.1966.1138693.