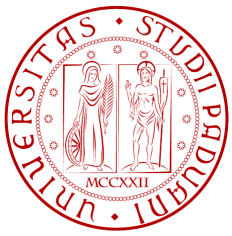
UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI FISICA E ASTRONOMIA "Galileo Galilei"



CORSO DI LAUREA MAGISTRALE IN FISICA

Study of variational symplectic algorithms for the numerical integration of guiding center equations of motion.

Relatore: Dr. Fabio Sattin
Correlatore: Prof. Luca Salasnich
Controrelatore: Prof. Marco Matone

Laureando: Filippo Zonta

A.A. 2013-2014

Contents

1 Introduction

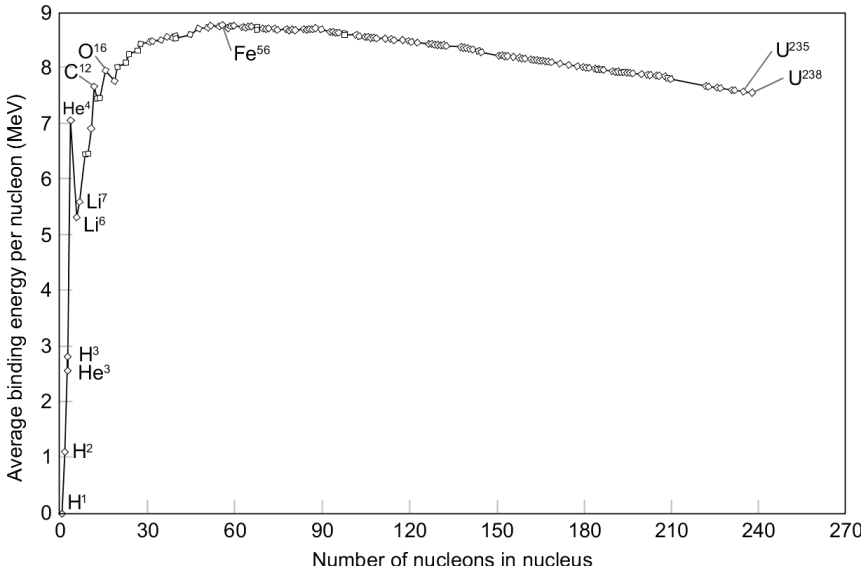
1.1 The need for Thermonuclear Fusion

The increasing demand for new energy production, expecially from rapidly industrialized countries, as well as the request of reducing the emissions and the environmental impact, has put the world in a difficult energy situation. A further complication is that, as well known, the present reserves of natural gas and oil will be exhausted in few decades.

Thermonuclear fusion exploits the energy surplus that is released when two light nuclei merge together into a heavier nucleus. Nuclear fusion has many attractive features in terms of safety, fuel reserves and minimal damage to the environment, hence it is a potential candidate to have a major role in the energy production of the future. In contrast with fission reactors, where it is needed to maintain a chain reaction in a large mass of fuel, fusion reactors must be continuosly fed with a small amount of fuel with a rate following the consumption needs, making catastrophic accidents like meltdown impossible. Fusion reactors do not produce greenhouse emissions like CO_2 nor other harmful chemicals. Attention must be paid to the emission of high energy neutrons by fusion reactions, which cause the activation of the exposed materials of the structure. However, the resulting nuclear wastes have the advantage of possessing only short-lived elements. With careful design and choice of materials, the level of radioactivity left by a fusion power plant after it has been closed for 100 years could be comparable to that of a coal fired power plant. Fusion has also disadvantages, the primary ones being related to the scientific and engineering challenges that are inherent in the fusion process. In particular, the requirement of confining a sufficient quantity of plasma for a sufficient long time at a sufficient high temperature has been the focus of the fusion research for the past 60 years.
Finally, a fusion reactor is inherently a complex facility and the ultimate question is whether a fusion reactor will be competitive with other power sources. Costings have been made for a number of reactor designs with a range of results: these typically lies in the range 1-3 times the cost of a fission reactor. However, there are many unknown variables, such as the future technological and economics developments and it is clearly difficult to predict the cost of fusion energy and its economic competitivity with respect to other sources 30-50 years in the future.

1.2 Relevant Fusion Reactions

Unlike fission reactions, where an isotope of Uranium (U^{235}) is bombarded with a slow neutron, thermonuclear fusion reactions involve two colliding light elements that fuse together, generating a heavier one and releasing energy in form of kinetic energies of the products.
The amount of energy released equals the difference of binding energy of the final and initial nuclei. This process is possible because the binding energy of light nuclei, as shown in fig. 1, is an increasing function with respect to the mass number, thus the released energy is positive, for nuclei up to Fe^{56}.



(a)

Figure 1: Binding energy per nucleon vs mass number

The reactions relevant for fusion research involve hydrogen isotopes Deuterium (D), Tritium (T), or Helium (He^3). A few are listed as follows:

$$^2\mathrm{D} + {}^2\mathrm{D} \rightarrow {}^3\mathrm{He} + {}^1\mathrm{n} + 3.27\ \mathrm{MeV} \tag{1.1a}$$

$$^2\mathrm{D} + {}^2\mathrm{D} \rightarrow \begin{cases} {}^3\mathrm{T} + {}^1\mathrm{H} + 4.03\ \mathrm{MeV} \\ {}^4\mathrm{He} + {}^1\mathrm{n} + 17.6\ \mathrm{MeV} \end{cases} \tag{1.1b}$$

$$^2\mathrm{D} + {}^2\mathrm{D} \rightarrow {}^4\mathrm{He} + {}^1\mathrm{H} + 18.3\ \mathrm{MeV} \tag{1.1c}$$

As a reference, the D-T reaction releases 17.6 MeV, or 3.52 MeV per nucleon, to be compared to 0.9 MeV per nucleon for a typical U^{235} fission reaction and to 0.9

eV of the combustion reaction of gasoline. 1 Kg of this fuel would release 10^8 kWh of energy, providing the requirements of a 1GW power plant for a day.

Deuterium is present naturally in ocean water, it can be easily extracted at a very low cost and it is virtually inexhaustible. Tritium is a radioactive isotope with a half life of about 12 years, hence it is not available naturally on Earth and must be obtained from a reaction involving Li^6. Geological estimates indicate that the reserves of Li^6 are of the order of 10^4 years at the present energy consumption rate.

Tritium may be bred from lithium using the neutron induced fission reactions:

$$Li^6 + n \rightarrow T + He^4 + 4.8 \text{ MeV} \tag{1.2a}$$

$$Li^7 + n \rightarrow T + He^4 + n - 2.5 \text{ MeV} \tag{1.2b}$$

The vacuum vessel of a D-T fusion reactor is surrounded by a blanket composed of a compound of lithium, which has the purpose of absorbing the neutrons produced by the fusion reaction, allowing both the trasmutation of lithium to tritium and the conversion of the neutrons energy to heat.

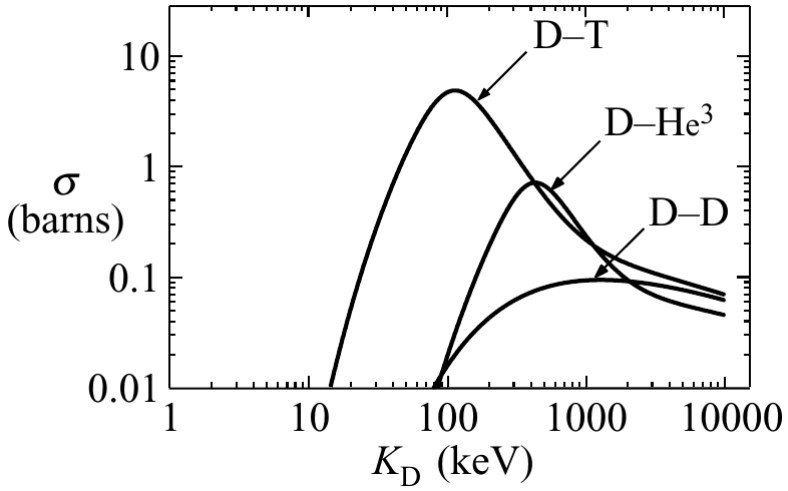


Figure 2: Experimentally measured cross sections for the D-T, D-He^3 , and D-D fusion reactions

Fig.2 shows the reaction rates for the three main reactions. It is clear the the D-T reaction is the most promising one, since it has the highest cross section except at impractical high energies.

The nuclear fusion reaction takes place due to the short-ranged nuclear force, the two nuclei must be very close to each other, making the inter particle Coulomb potential very repulsive and thus greatly reducing the likelihood of the reaction.

1.3 Confinement and Energy Balance of a Confined Plasma

The strategy behind all current fusion research is to confine the particles in a region of space for times long enough such that the particles can perform fusion reactions.

The thermonuclear power per unit volume produced in a D-T plasma is:

$$p_f = n_d n_t \langle \sigma v \rangle E_f \tag{1.3}$$

where n_d and n_t are the deuterium and tritium densities, $\langle \sigma v \rangle$ is the reactivity rate average of the cross section times the relative speed of the particles over a Maxwellian distribution and E_f is the energy released in a single fusion reaction, namely 17.6 MeV.

In an open system like a plasma there is a continuous loss of energy, through transport to the walls and radiation processes, which has to be balanced by heating in order for the process to be self-sustainable. For a thermalized plasma, the energy density per unit volume is $w = 3nT$ and its time evolution is an energy balance equation:

$$\frac{dw}{dt} = p_H + p_\alpha - p_L - p_R \tag{1.4}$$

The first term, p_H, includes any external supplied power. p_α is obtained from eq.(1.3), retaining only the energy of the alpha particles ($E_\alpha = 3.5$ MeV), since the neutrons leave the plasma without interaction. p_R is the energy lost by radiation, the main contribution of which is given by the Bremsstrahlung radiation that can be expressed as:

$$p_b = \alpha_b n^2 T^{1/2} \ W m^{-3} \tag{1.5}$$

where α_b is a constant equal to $5.35 \cdot 10^{-37} \ W m^3 keV^{-1/2}$ and the temperature is expressed in keV. Bremsstrahlung radiation is caused by Coulomb two-body collisions inside the plasma, thus it is an intrinsic and unvoidable term. The last term p_L, the power lost by transport phenomena, is the most difficult to quantify and it is empirically described through the energy confinement time τ_E:

$$P_L = \frac{W}{\tau_E} \tag{1.6}$$

where $W = wV$ is the total thermal energy of the plasma and τ_E represents the relaxation time of the plasma energy due to heat conduction. Experimentally, when

the fusion power term is negligible, it can be estimated from:

$$\tau_E = \frac{W}{P_H - P_R} \tag{1.7}$$

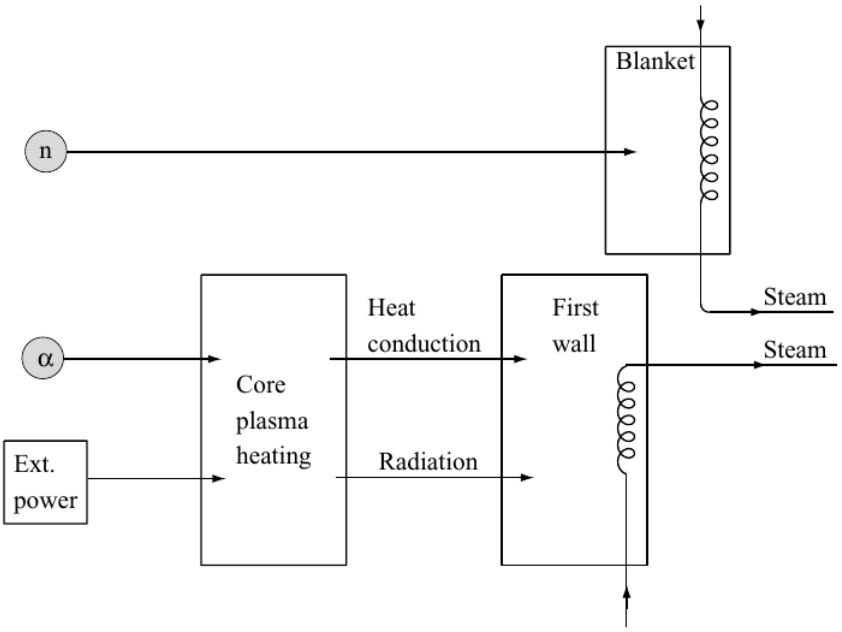


Figure 3: Schematic diagram of power flow inside a fusion reactor

The energy balance of the reactor is represented schematically in Fig.3. In a steady state power balance, assuming negligible heat conduction losses and no external heating, it is straightforward to derive a minimum value, sometimes referred as ideal ignition, of the temperature:

$$T \geq 4.4 \text{ keV} \tag{1.8}$$

Note that the corresponding value for D-D reactions is much higher, approximately 30keV. More realistically, in order to reach ignition the alpha power heating has to be sufficiently large to balance the combined Bremsstrahlung and thermal conduction losses, without the need of an external heating:

$$p_\alpha = p_L + p_b \tag{1.9}$$

which can be easily expressed as

$$n\tau_E \geq \frac{12T}{\langle \sigma v \rangle E_\alpha - 4\alpha_b T^{1/2}} \tag{1.10}$$

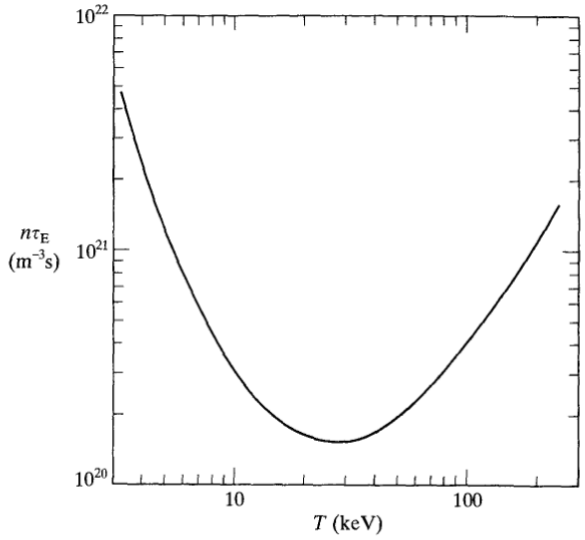


Figure 4: $n\tau_E$ condition necessary to reach ignition

At typical temperatures of about 10-20 keV, where the ignition curve is near its minimum (see Fig.4), the reactivity can be expressed using an approximated formula, giving a condition in the so called triple product:

$$n\tau_E T \geq 3 \cdot 10^{21} m^{-3} sMeV \tag{1.11}$$

As an example, with $T \simeq 10$ keV and $n \simeq 10^{20} m^{-3}$, the confinement time should be at least 3s. Altough ignition is a convenient condition, since the plasma would be self sustainable so that the external power source could be switched off, it is not necessary to obtain a positive energy balance. It is useful, to this regard, to consider the parameter Q, defined as the ratio of the thermonuclear power produced to the heating power required:

$$Q = \frac{P_f}{P_H} \simeq \frac{5P_\alpha}{P_H} \tag{1.12}$$

Thus, $Q = 1$ is the condition, called break-even, where the fusion power equals the external heating. At ignition, where P_H is reduced to 0, $Q \to \infty$.
It is possible to obtain a positive balance with a large Q without reaching ignition. However, the supplied power is a cost on the system that involves recycling of the power and a consequent loss of efficiency.

1.4 Paths to Ignited Plasmas

The confinement occours naturally in stars, where the gravitational force balances the kinetic pressures. However, gravity is a very weak force compared to the nuclear

one. The fusion power density in the core of the Sun is only 270 watts per cubic meter, many order of magnitude lower than that required for a commercial power plant.

Different approaches are needed in order for fusion to be viable on human-relevant scales, and two of them have been pursued.
One is inertial confinement: a small solid target made of deuterium and tritium is heated with laser beams leading to the evaporation of the external surface and the implosion of the internal part of the target. The increasing density, over 1000 times higher than that of normal liquid D-T, makes fusion reactions possible.

The second path is provided by magnetic confinement: a gaseous D-T mixture is heated up to temperatures relevant for the triggering of fusion reactions. At temperatures of interest, the gas is fully ionized, constitutes a neutral plasma made by ions and electrons. Material walls thus surround laboratory device but cannot be used directly to maintain confined plasmas, instead magnetic confinement is needed.
As will be discussed later, the average motion of any charged particle embedded into a magnetic field is dominantly aligned along the magnetic field line, whereas its orthogonal motion is relatively slow. Hence, a viable strategy is to tailor magnetic field configurations able to confine plasmas.
This is practically achieved by closing magnetic field lines onto themselves into the doughnut-shaped (toroidal) configuration typical of magnetic confinement devices, that will be described more extensively in the next section.

1.5 Particle Dynamics in a Magnetized Plasma and the Problem of Transport

Plasmas, being made of charged particles, can sustain self-produced flowing currents, that is, they can produce further magnetic fields in additions to those externally imposed. The resulting magnetic configuration can thus be different from the design one.
Indeed, some concepts of confinement devices, like the Reversed Field Pinches, rely on virtually stable magnetic configurations that are mostly self-produced by the plasma itself. There is in principle a premium in attaining such configurations, since they require the lesser effort (i.e., power comsumption) from the external circuitry. On the other hand, this is achieved at the expense of a magnetic topology which is now not known a priori. The basic element within this discussion is provided by magnetic *flux surfaces*. They will be explained in more detail in the next chapter.
For the purpose of this paragraph, it is important that magnetic field lines lie in

each of these surfaces. There is not a component of **B** perpendicular to the surface. Accordingly, as long as the identification between field line and particle trajectory holds, i.e. drifts and collisions aside, flux surfaces define also barriers to the the particle dynamics: as we will discuss later, they are KAM tori of Hamiltonian systems. The best configuration for confinement is one where all magnetic surfaces are nested and detached from material walls since it minimizes the flux of matter and energy away from the core plasma [1]

In actual scenarios, one encounters a mixture of (i) good flux surfaces, (ii) ergodic regions, i.e., regions where magnetic field lines wander erratically filling the whole available volume. Regions (i) and (ii) have wildly different transport properties, in particular stochastic zones (ii) act as short-circuits, effectively transferring mass and energy across distants regions via parallel streaming. (iii) Finally, a fraction of field lines may intercept the material walls (so called open field lines). These, too, act as efficient loss mechanisms.

It arises therefore the need of investigating what the dynamics of plasmas turns out to be within such scenarios. The full self-consistent problem, where the magnetic topology and the plasma evolution are computed together self-consistently represents ultimately the main question of equilibrium and heat and/or matter transport in confined plasmas, and is still an overwhelmingly complicated question.
In most devices one may reconstruct, at least approximately, the magnetic field starting from experimental measurements carried out using coils at the plasma edge, polarimetry and motional Stark effect in the core plasma.
On the basis of this knowledge plasma dynamics may be approximately inferred from the behaviour of a small number of test particles evolving within a given background magnetic field.
However, the trajectory of a charged particle can be approximated with the underlying magnetic field line only in a first approximation, under the hypothesis of homogeneous field.
The unavoidable existence of inhomogenities and curvature produces additional drift motions as we shall see in the next chapters.
Two-body Coulomb scattering must finally be accounted for, it causes impulsive deflections of the particles from the original trajectory. In a high-temperature plasma,

[1] The whole issue in an actual fusion reactor, is made much more complicated by the fact that conflicting demands are in order. From the one hand we wish to keep confined as much as possible the fuel, i.e. the hydrogen isotopes, as well as the energy needed to ignite it. On the other hand, it is necessary to extract the ashes of the combustion, i.e., the helium, alongside with the energy it carries along, since it the useful output of the fusion process. Also impurities, i.e., trace elements of atoms other than hydrogen, coming from the walls, need to be quickly removed, since they contribute to cool the plasma by radiation.

however, Coulomb collisions are relatively weak, the collision frequency scales like $(\text{temperature})^{-3/2}$, thus particles spend most of their time travelling along trajectories dictated by the background magnetic field alone, with two-body collisions intervening as sporadic perturbations: in a typical present-date laboratory plasma, an electron is allowed to make 10-100 toroidal revolutions along the major circumference between successive scattering events.

1.6 Outline of the Thesis

We are finally led to the purpose and content of this thesis. We note first of all that the intrinsic multiscale nature of plasmas force any investigation of the dynamics of charged particles to span an extremely wide range of time and length scales.
Concerning length scales, for instance, one goes from the electron Larmor radius (radius of the gyration along the magnetic field line), which in a typical laboratory plasma may be of order 10^{-4} m, to the total path travelled between, e.g., two Coulomb collisions $\approx 10^{2\div 3}$ m, with the further issue that particle trajectories may be highly irregular and possibly chaotic over some regions.
These complications call for special care in handling with the numerical integration of equations of motion.
The first fundamental simplification comes from the acknowledgment that the Larmor gyration of the particle around the local magnetic axis, which is extremely faster than any other time scale involved, can be averaged out, and the exact trajectory of the particle replaced by the motion of its gyration center, which moves over extremely longer time scales (guiding center theory).
The price to be paid is that the local inhomogenities sampled by the particles during its full evolution must now be retained as suitable averages in a new set of equations of motion for the guiding center. This topic will be addressed in Chapter 2.

It is well known that an accurate integration of the equations of motion greatly benefits from algorithms incorporating any conservation law that the particle is known to obey, since they prevent by construction the cumulative effect of accumulation of numerical errors over long integration times. Symplectic algorithms exploit the conservation properties built into those systems whose evolution is described by some Hamiltonian.
Luckily, it has been shown that the guiding center equations of motion do possess a Hamiltonian structure, thereby allowing for the development of symplectic integrators. Within this framework, time integration of a trajectory corresponds to a mapping from initial Hamiltonian phase space coordinates to final ones, it is a kind of canonical transformation, i.e., a transformation that preserves the form of Hamilton's equations.
On the other hand, it may be shown that the guiding center dynamics in generic

magnetic fields does not possess one single global canonical structure, but just a local one: the form of the mapping from initial to final coordinates depends from the point. This makes unpractical standard algorithms that instead rely on the postulate of the existence of a global canonical structure of the Hamiltonian.

One intriguing possibility recently suggested is to use variational symplectic integrators. These integrators are shown to possess in many cases good long time preservation properties, when applied to a noncanonical problem, thus in principle allow to override the previous shortcomings of canonical integrators. On the other hand a complete understanding of the theoretical and mathematical aspects involved with these integrators is yet not available. The aim of this thesis is to present an investigation of this topic.

The thesis is structured as follows. Chapters 2-3 are introductory. The main role of Chapter 2 is to provide the näive guiding center theory as encountered in standard textbooks. However, as corollary, some other basic concepts are introduced that will be useful later, in particular to scrutinize the results of Chaper 5.

Chapter 3 revisits the theory of the guiding center, this time within a Hamiltonian setting. This part, too, reviews standard material in textbooks.

Numerical integrators rely on a discretization of the equations of motion. Therefore, chapter 4 starts entering the main topic of the thesis by revisiting the symplectic integrators, with particular attention to canonical systems, this time with attention to the discrete aspects of the mechanics. Symplectic integrators are studied from a variational (Lagrangian) and from a Hamiltonian point of view. Finally, the correspondence between discrete integrators and continuous theory is established. This material comes from relatively new literature of the past 10 years. The basic elements of the theory of Hamiltonian systems, necessary for an understanding of the chapter, are separately summarized in Appendix A, for the convenience of the reader.

Chapter 5 is the core of the thesis, where we tackle the integration of non canonical Hamiltonian systems with symplectic algorithms. Some general properties and behaviour of this class of integrators are outlined in this chapter. The starting point is provided by recent published papers ([20] [21]) which have been further extended in an original way in this thesis.

Finally, chapter 6 deals with the original application of the theory developed in the previous chapter to the non canonical guiding center theory in terms of some

numerical excercises. Different discretizations approaches are proposed, and their stability properties assessed.

2 Charged particles motion in electromagnetic fields: the guiding-center picture

Consider a particle of mass m and charge q in an electromagnetic field. Its motion can be directly determined from the Lorentz force:

$$m\frac{d\mathbf{v}}{dt} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \tag{2.1a}$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{v} \tag{2.1b}$$

Eqs.(2.1) are a set of coupled ordinary differential equations and they constitute a quite complex problem since $\mathbf{B}(\mathbf{x}, t)$ and $\mathbf{E}(\mathbf{x}, t)$ are in general functions of space and time; however, the basic idea of magnetic confinement theory can be demonstrated by studying the particle motion in a uniform, time independent magnetic field. Assume the following fields configuration in a cartesian coordinate system: $\mathbf{E} = 0$, $\mathbf{B} = B\mathbf{e_z}$.
Taking the scalar product of (2.1) with $\mathbf{v}$ yields:

$$\frac{1}{2}m\mathbf{v} \cdot \mathbf{v} = const. \equiv W \tag{2.2}$$

Solving explicitely eq.(2.1) for the z-component leads to:

$$\frac{dv_z}{dt} = 0 \tag{2.3}$$

The full set of Newton's law reduces to:

$$\frac{dv_x}{dt} = \omega_c v_y \tag{2.4a}$$

$$\frac{dv_z}{dt} = -\omega_c v_x \tag{2.4b}$$

$$\frac{dv_z}{dt} = 0 \tag{2.4c}$$

where $\omega_c = qB/m$ is the gyro-frequency, sometimes referred as the **Larmor frequency**.

Eqs.(2.4) can be clearly separated into a uniform motion along the field lines of a guiding center

$$\mathbf{x}_{gc} = (x_0, y_0, z_0 + v_{\|}t) \tag{2.5}$$

and a circular motion, the gyromotion, around the field lines, with frequency ω_c and gyroradius $r_L = mv_\perp/qB$.
For a typical fusion plasma, the gyroradius is of the order of mm for ions, depending upon their energy, and is correspondingly scaled by the square root of the ratio of the masses for electrons at the same energy. Note that the direction of the gyromotion is opposite for opposite charges.

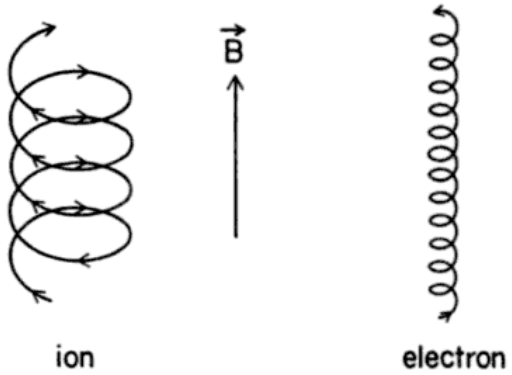


Figure 5: Trajectory of charged particles in a uniform magnetic field.

The combined parallel and perpendicular motion yields a helical motion around the magnetic field lines, as shown in fig.5.

2.1 Magnetic moment

The gyromotion of the charged particles produces a magnetic field that, from Ampere's law, is directed antiparallel to the **B** field (diamagnetism of the plasma). The magnetic moment associated is easily computed as:

$$\boldsymbol{\mu} = \left(-q\frac{\omega_c}{2\pi}\right)(\pi r_L^2)\hat{\mathbf{B}} = -\frac{W_\perp}{B}\mathbf{b} \tag{2.6}$$

where **b** is the unit vector parallel to the magnetic field. As a consequence, a plasma of particle density $2n$ possesses a diamagnetic behaviour, quantified by means of the ratio of the kinetic to the magnetic pressure, β:

$$\beta = \frac{nW_\perp}{B^2/2\mu_0} = -\frac{\delta B}{B} \tag{2.7}$$

2.2 Drift across a magnetic field

When the magnetic field is not uniform and constant, solution of eqs.(2.1) can differ substantially from the helical trajectory. The underlying assumption of guiding center theory is that the variations of B and E occur on a length scale longer than the gyroradius and on a time scale slower than the gyrofrequency. When this condition is satisfied, it is still possible to decouple the motion of the guiding center from the gyromotion. As it will be shown in the next sections, the main consequence of the variations of the field is the existence of drifts of the guiding center motion.

2.2.1 Curvature Drift

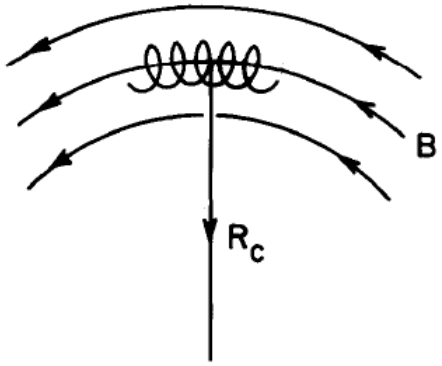


Figure 6: Charged Particle gyrating around a curved magnetic field

Let's consider a magnetic field constant in time and with a small curvature R_C, as sketched in fig.6. We can write the Lorentz force (eq. 2.1) and split the parallel and perpendicular components of the velocity with respect to the magnetic field:

$$\begin{aligned} q(\mathbf{v}\times\mathbf{B}) = m\frac{d\mathbf{v}}{dt} = m\frac{d\mathbf{v}_\perp}{dt} + m\frac{d\mathbf{v}_\parallel}{dt} = \\ = m\frac{d\mathbf{v}_\perp}{dt} + m\mathbf{b}\frac{dv_\parallel}{dt} + mv_\parallel\frac{d\mathbf{b}}{dt} \end{aligned} \tag{2.8}$$

The last term can be rewritten as:

$$mv_\parallel\frac{d\mathbf{b}}{dt} = mv_\parallel\mathbf{v}\cdot\nabla\mathbf{b} \tag{2.9}$$

If we average the last expression over one gyro orbit, the fast perpendicular oscillating motion cancels out and only the parallel component remains non vanishing:

$$\overline{mv_\parallel \mathbf{v}\cdot\nabla\mathbf{b}} = mv_\parallel^2 \mathbf{b}\cdot\nabla\mathbf{b} = -mv_\parallel^2 \frac{\mathbf{R}_C}{R_C^2} \tag{2.10}$$

Hence, returning to equation (2.8), the parallel component remains unaffected by the curvature, while an average centrifugal force $\mathbf{F_C}$ arises in the perpendicular direction:

$$m\frac{d\mathbf{v}_\perp}{dt} = \mathbf{F_C} + q\left(\mathbf{v}\times\mathbf{B}\right) \tag{2.11}$$

where $\mathbf{F_C}$ is:

$$\mathbf{F_C} = -mv_\parallel^2 \frac{\mathbf{R_C}}{R_C^2} \tag{2.12}$$

The perpendicular velocity can now be broken into two parts, $\mathbf{v}_\perp = \mathbf{v_c} + \mathbf{v_{gc}}$ such that the effect of $\mathbf{v_{gc}}$ is to cancel out $\mathbf{F_C}$

$$\mathbf{F_C} + q(\mathbf{v_{gc}}\times\mathbf{B}) = 0 \tag{2.13}$$

Hence, the motion perpendicular to the magnetic field comprises of a circular motion $\mathbf{v_c}$ and a drift of magnitude

$$\mathbf{v_{gc}} = \frac{1}{q}\frac{\mathbf{F_C}\times\mathbf{B}}{B^2} = -\frac{mv_\parallel^2}{qR_C^2}\frac{\mathbf{R_C}\times\mathbf{B}}{B^2} \tag{2.14}$$

2.2.2 Gradient Drift

Consider now the case in which the magnetic field is non uniform in space, but still along the same direction, $\mathbf{B} = B(\mathbf{x})\mathbf{u}_z$. Assuming small variations, we can approximate the magnetic field along a gyro orbit with:

$$\mathbf{B} \simeq \mathbf{B_0} + \tilde{\mathbf{x}}\cdot\nabla\mathbf{B} \tag{2.15}$$

where $\tilde{\mathbf{x}}$ represents the gyration motion around the gyrocenter. The equations of motion for the perpendicular component of the velocity now read:

$$m\frac{d\mathbf{v}_\perp}{dt} = q(\mathbf{v}_\perp\times\mathbf{B}_0) + q[\mathbf{v}_\perp\times(\tilde{\mathbf{x}}\cdot\nabla)\mathbf{B}] \tag{2.16}$$

The last term can be averaged over one gyro orbit:

$$\mathbf{F} \equiv \overline{q[\mathbf{v}_\perp\times(\tilde{\mathbf{x}}\cdot\nabla)\mathbf{B}]} = q\overline{v_y x}\frac{\partial}{\partial x}B - q\overline{v_x y}\frac{\partial}{\partial y}B \tag{2.17}$$

At first order, v_x and v_y are purely oscillatory (eq. 2.4) and the average in eq. (2.17) is an integral over a period of square of cosines and sines, hence:

$$\mathbf{F} = -\frac{qv_\perp^2}{2\omega_c}\nabla_\perp B = -\mu\nabla_\perp B \tag{2.18}$$

The force $\mathbf{F}$ produces a drift in the gyrocenter motion similar to that in the previous section (but now the component of the velocity perpendicular to B enters, rather than the parallel one)

$$\mathbf{v}_{\nabla_\perp} = \frac{\mu}{q}\frac{\mathbf{B}\times\nabla_\perp B}{B^2} \tag{2.19}$$

Magnetic mirror On the other hand, the presence of a parallel gradient $\mathbf{B} \parallel \nabla\mathbf{B}$, causes a phenomena known as **magnetic mirror**.
Using cylindrical coordinates, we can notice from the divergence condition on the magnetic field that a parallel gradient implies that the radial component of the field must be not null:

$$0 = \nabla\cdot\mathbf{B} = \frac{1}{r}\frac{\partial}{\partial r}(rB_r) + \frac{\partial B}{\partial z} \tag{2.20}$$

Assuming the gradient to be constant over a gyro orbit, equation (2.20) can be integrated to give:

$$B_r = -\frac{1}{r}\int_0^r r'\frac{\partial B_z}{\partial z}dr' = -\frac{1}{2}r\frac{\partial B_z}{\partial z} \tag{2.21}$$

The z-component of the Lorentz force is then computed from eq.(2.21) at the Larmor radius:

$$F_z = -qu_\perp B_r = -\frac{q}{2}u_\perp r_L\frac{\partial B_z}{\partial z} = -\mu\frac{\partial B_z}{\partial z} \tag{2.22}$$

Using (2.6) and (2.22), we can write the conservation of total kinetic energy, which still holds since the magnetic field is time independent, as:

$$0 = \frac{d}{dt}\left(\frac{1}{2}mv_\parallel^2 + \frac{1}{2}mv_\perp^2\right) = -\mu\frac{dB}{dt} + \frac{d(\mu B)}{dt} = B\frac{d\mu}{dt} \tag{2.23}$$

Therefore, both the total kinetic energy and the magnetic moment are constants of the motion.
From (2.6), it is clear that the fraction of kinetic energy due to orthogonal motion

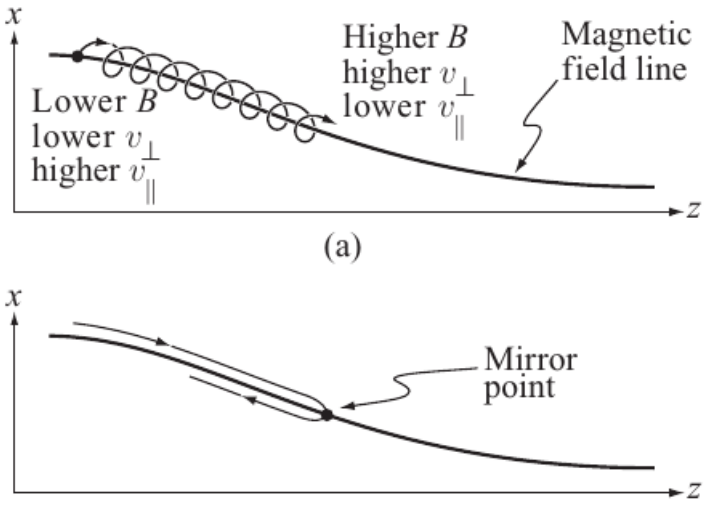


Figure 7: Inversion of parallel velocity due to mirror effect

has to increase when the particle moves into a region of higher **B**, leading eventually to the inversion of the parallel velocity when the magnetic field is sufficiently high (fig.7). More specifically, a particle is reflected when the ratio of the orthogonal to the parallel velocity exceeds a critical value:

$$\frac{v_\perp}{v_\parallel} > \left(\frac{B_{max}}{B} - 1\right)^{-\frac{1}{2}} \tag{2.24}$$

Actually, a simple linear device which exploits the mirror effect to confine the charged particles along the magnetic lines, could be designed. In fact, this was one of the earliest magnetic fusion configuration, known as the mirror machine. Two coil generates a magnetic field with a maximum under each coil and a minimum in the midway. However, the mirror machine in practical experiments does not work as expected. Particles not fulfilling Eq. (2.24) are lost from the machine, and Coulomb collisions, as well as the presence of both macroscopically and microscopically instabilities, contribute to continuosly replenish this region of the particles' velocity space at a rate that prevents the maximum theoretical energy gain factor Q from rasing above 1.

The approach that proved to be the most promising is to close the magnetic field lines in a toroidal configuration, in which a set of toroidal field coils encircling the plasma produce a toroidal field B_ϕ. The coordinate system used in this configuration is sketched in fig.8.
The curvature and nonuniformity of the magnetic field cause the particles to drift

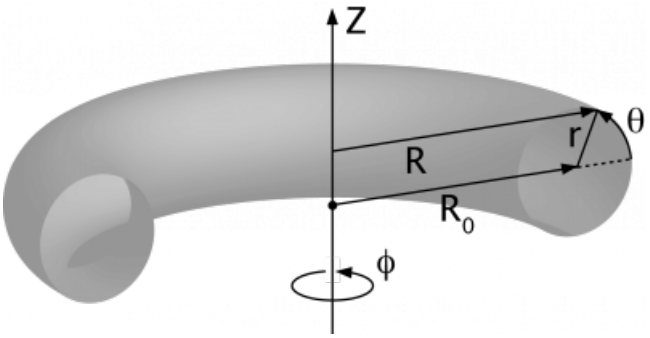


Figure 8: Toroidal (r, θ, ϕ) coordinate system

vertically with a velocity

$$\mathbf{v_d} = \frac{r_L v_\perp}{2} \frac{\mathbf{B} \times \nabla B}{B^2} + \frac{m v_\parallel}{qB} \frac{\mathbf{R_C} \times \mathbf{B}}{R_C^2 B} \tag{2.25}$$

which is in the opposite direction for ions and electrons. This charge separation results in an electric field that produces a drift directed radially outward:

$$\mathbf{v_d} = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \tag{2.26}$$

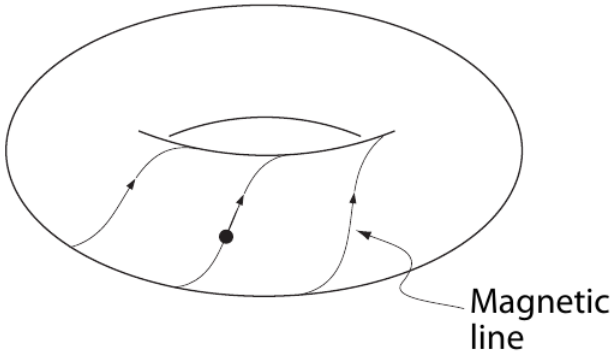


Figure 9: Charged particles gyrating around magnetic field lines in a toroidal device

In order to avoid particles hitting the wall, it is necessary to superimpose an additional magnetic field in the poloidal direction (θ), resulting in a helical magnetic field, as shown in fig.9. In this way, particles are carried outward in the upper poloidal orbit and inward otherwise, thus balancing the $\mathbf{E} \times \mathbf{B}$ drift.

2.3 Flux Surfaces

A surface defined by a function $f(\mathbf{x}) = const$ is said to be a **flux surface**, or a magnetic surface, if at any point the magnetic field lies within the surface. Mathematically, $\mathbf{B} \cdot \nabla f = 0$.
The existence of flux surfaces is important since it is a necessary condition for the stability and thus the confinement of the plasma. Flux surfaces provide a barrier to charged particle that is penetrable only by their relatively slow, perpendicular drift motion.

Let's introduce generic coordinates (r, θ, ϕ), such that θ and ϕ are periodic with period 2π.
The vector potential associated with the magnetic field can be expressed in a contravariant basis as:

$$\mathbf{A} = A_r \nabla r + A_\theta \nabla \theta + A_\phi \nabla \phi \tag{2.27}$$

Defining the scalar functions G, ψ and ψ_p such that $\partial_r G = A_r$, $\psi = A_\theta - \partial_\theta G$, $\psi_p = A_\phi - \partial_\phi G$, eq.(2.27) can be rewritten as:

$$\mathbf{A} = \nabla G + \psi \nabla \theta - \psi_p \nabla \phi \tag{2.28}$$

The magnetic field is then:

$$\mathbf{B} = \nabla \times \mathbf{A} = \nabla \psi \times \nabla \theta - \nabla \psi_p \times \nabla \phi \tag{2.29}$$

It is easy to verify that ψ and ψ_p are both flux functions if ψ_p depends only on ψ: $\psi_p = \psi_p(\psi)$.
In this case, the safety factor can be defined as:

$$\frac{1}{q(\psi)} = \frac{d\psi_p}{d\psi} \tag{2.30}$$

Eq. (2.29) is said to be a flux representation of the magnetic field. The coordinates (ψ, θ, ϕ) that permit this representation are called flux coordinates. The magnetic surfaces ($\psi = const.$) consist topologically of nested tori (fig. 10) and in general they differ from the toroidal coordinates, θ and ϕ not being the usual geometrically defined angular variables.
Only in the case of an axisymmetric magnetic field, the toroidal coordinates are good flux coordinates.

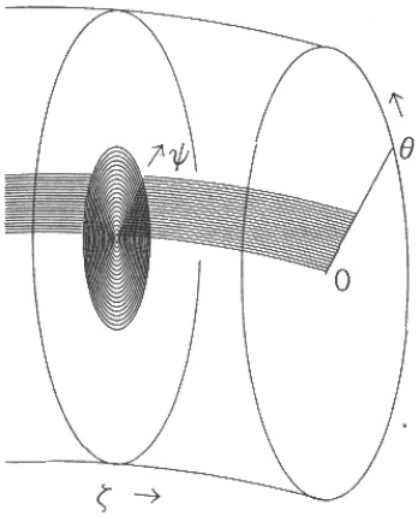


Figure 10: Toroidal and poloidal magnetic surfaces

The term flux function can be justified computing the toroidal and poloidal flux through a magnetic surface:

$$\frac{1}{2\pi}\int d^3x(\mathbf{B}\cdot\nabla\phi) = \frac{1}{2\pi}\int d\psi d\theta d\phi = 2\pi\psi \tag{2.31a}$$

$$\frac{1}{2\pi}\int d^3x(\mathbf{B}\cdot\nabla\theta) = \frac{1}{2\pi}\int d\psi_p d\theta d\phi = 2\pi\psi_p \tag{2.31b}$$

Therefore, ψ and ψ_p represent the magnetic flux per unit angle that passes through these surfaces.

The magnetic field lines on the θ-ϕ plane are defined by

$$\frac{d\theta}{d\phi} = \frac{\mathbf{B}\cdot\nabla\theta}{\mathbf{B}\cdot\nabla\phi} = \frac{1}{q(\psi)} \tag{2.32}$$

The safety factor $q(\psi)$ thus defines the field helicity in the θ, ϕ variables on the surface ψ.

2.4 Trapped Particles

As already stated in the previous sections, the magnetic field produced by the field coils is nonuniform. Using Ampere's law and taking the line integral of B_ϕ around any close loop inside the plasma, it is easy to verify that the magnetic field falls off as $1/R$ across the confinement region.

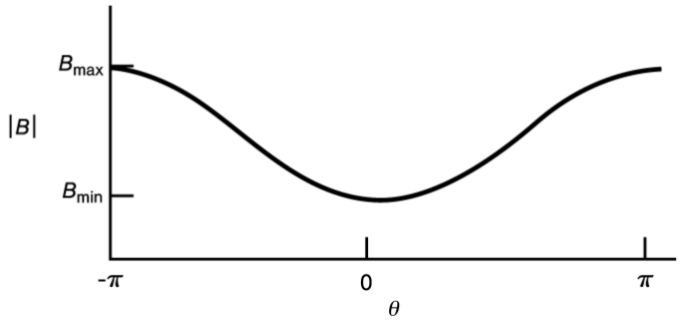


Figure 11: Toroidal magnetic field variation along a poloidal orbit

Hence, the magnetic field along a poloidal orbit of radius r has a maximum and a minimum value (see fig.11):

$$B_{max} = B_0 \frac{R_0}{R_0 + r} \tag{2.33a}$$

$$B_{min} = B_0 \frac{R_0}{R_0 - r} \tag{2.33b}$$

Following the discussion of section 2.2.2, a particle is trapped if its parallel velocity $v_\parallel$ at $\theta = 0$ satisfy the condition:

$$\frac{v_\parallel^2}{v^2} < 1 - \frac{B_{min}}{B_{max}} \simeq 2\frac{r}{R_0} \tag{2.34}$$

The trajectory of a trapped particle in the poloidal plane, shown in fig.12, is referred as the "banana" orbit. Note that the orbit shifts outward or inward with respect to the magnetic surface, depending on the sign of the initial parallel velocity.

Altough the fraction of trapped particles is usually small, it is the dominant contribution to the heat and transport losses inside the plasma. This can be explained by the fact that the parallel velocity of trapped particles is small and the orbital period is long, raising the probability to drift off the flux surface due to gradient and curvature forces.

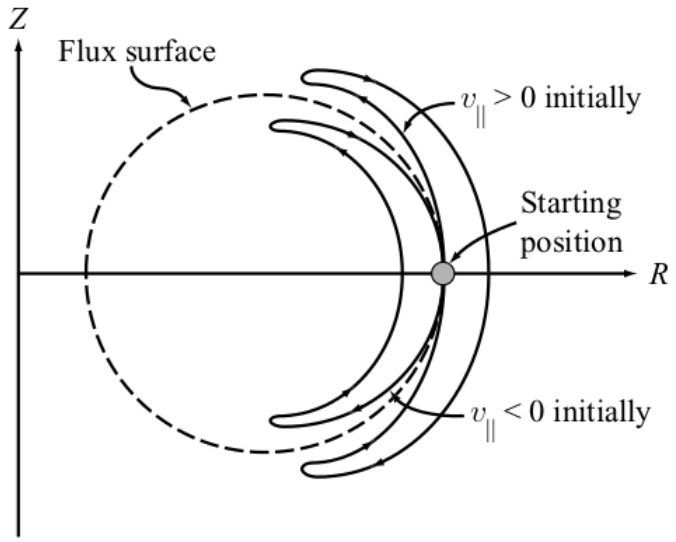


Figure 12: Trapped particle orbit projected onto the poloidal plane

3 Hamiltonian guiding center theory

3.1 Introduction

The original theory of the guiding center motion, as presented in the introductory chapter (chapter 2), has many disadvantages, as it doesn't clearly possess an energy conservation law, Noether theorem can't be applied and the motion of the particles is not consistent with the Liouville theorem.
In the 80's, flux coordinates ψ, ψ_p were used to obtain a canonical Hamiltonian theory of the Guiding Center, so that the equations of motion are described by the Hamilton's equations.
Although this theory has the advantage to be easily applicable to symplectic integrators, the main difficulty is that the Darboux-Lie theorem (appendix A.4.2) guarantees that canonical coordinates are well defined only locally: global canonical coordinates, defined in the whole region of integration, can be found only for particular choices of the magnetic field configuration.
This thesis will focus on the non canonical theory, following the work of Littlejohn [17], who derived coordinates valid in the whole space, starting from an extended variational principle.
For a full review of Hamiltonian theories of the guiding center, see Cary and Brizard[4].

In this chapter, we'll use the notion found in appendix A to build an Hamiltonian theory of the guiding center.
Of course, this theory possesses all the properties we need, such as the Liouville theorem, energy conservation and the Noether's theorem.
We'll start by showing that every Hamiltonian system descends from an extended variational principle in which we can define an extended Lagrangian, the so called **phase-space Lagrangian**.
We'll then present a non canonical guiding center theory using this formalism. In chapter 6 we'll apply the variational symplectic integrator to this Lagrangian.

3.2 Phase-space Lagrangian formalism

Given a regular Lagrangian system (eq. A.67), or equivalently a canonical Hamiltonian system, we already know that the equation of motions can be derived from a variational principle in the configuration space Q, namely the Hamilton's principle.

From the Darboux theorem (section A.4.2), every non canonical system can be brought in canonical coordinates, at least locally. However, this method arises difficulties in finding a suitable change of coordinates.
One can wonder if a generic Hamiltonian system can be derived directly from a

variational principle in its symplectic manifold. The answer is given in this section. A more deeper treatment can be found in literature, for example in [9], [19], or [4] for the phase-space Lagrangian formalism.

3.2.1 Phase-space variational principle

Let's start with a canonical Hamiltonian system with Hamiltonian H. The Hamilton's equations (eqs. A.102) are equivalent to the stationarity of the following integral:

$$\delta \int_{t_0}^{t_1} (\dot{q}(t) \cdot p(t) - H(q(t), p(t)))\, dt \tag{3.1}$$

The proof of this fact will be given in the following subsection.
Note that q, p are treated as independent variables, while in the Hamilton's principle only the varible q is meant to be varied arbitrarily, as long as the endpoints of the curve are kept fixed. For this reason, we call this variational principle extended.
We can wonder which are the conditions to impose at the endpoints to guarantee the correct dimensionality of the problem.
In this regard, Goldstein [9] pointed out that fixing the endpoints on the curve $p(t)$ is irrelevant, since the integral (3.1) is independent of the derivative $\dot{p}$ of p: Fixing or not the endpoints on $p(t)$ gives exactly the correct equations of motion.
For an Hamiltonian system on a symplectic manifold (M, Ω), the extended variational principle reads:

$$\delta \int_{t_0}^{t_1} (\Theta - H dt) = 0 \tag{3.2}$$

where Θ, known as the symplectic 1-form or the Cartan form, is a differential 1-form such that its exterior derivative is the symplectic form:

$$\Omega = -\mathbf{d}\Theta \tag{3.3}$$

3.2.2 Phase-space Lagrangians

The **phase-space** Lagrangian is defined to be the argument of the integral (3.1):

$$\mathcal{L}(q, \dot{q}, p, \dot{p}) = \dot{q} \cdot p - H(q, p) \tag{3.4}$$

The variational principle then reads:

$$\delta \int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t), p(t), \dot{p}(t)) dt = 0 \tag{3.5}$$

Note that although $\dot{p}$ is a variable, the Lagrangian does not depend on it and that all four variables are completely independent.
Since Euler-Lagrange equations are second-order equations, one can wonder if the problem remains physically the same, as we are doubling the space of variables and hence the degrees of freedom of the system.
As we'll see right now, Euler-Lagrange equations of a phase-space Lagrangian must be first-order equations.
For a canonical system (3.4), they are easily computed:

$$0 = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = \dot{p} + \frac{\partial H}{\partial q} \tag{3.6}$$

$$0 = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{p}} - \frac{\partial \mathcal{L}}{\partial p} = -\dot{p} + \frac{\partial H}{\partial p} \tag{3.7}$$

which are just the Hamilton's equations, as expected.
For a generic Hamiltonian system, denoting by $z \in M$ local coordinates of M, the phase-space Lagrangian is:

$$\mathcal{L}(z, \dot{z}) = \Theta_i(z)\dot{z}^i - H(z) \tag{3.8}$$

Note that both the Hamiltonian and the matrix of the 1-form Θ do not depend on $\dot{z}$ (by definition the 1-form is a map $\Theta : M \to T^*M$).
Thus, the Euler-Lagrange equations are:

$$0 = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{z}} - \frac{\partial \mathcal{L}}{\partial z} = \frac{\partial \Theta_i(z)}{\partial z^j}\dot{z}^j - \frac{\partial \Theta_j(z)}{\partial z^i}\dot{z}^j + \frac{\partial H}{\partial z^i} \tag{3.9}$$

The first term is just the matrix of the external derivative of Θ, namely the symplectic 2-form ($\Omega = -\mathbf{d}\Theta$). Hence:

$$\Omega_{ij}(z)\dot{z}^j = -\frac{\partial H}{\partial z^i} \tag{3.10}$$

where Ω_{ij} is:

$$\Omega_{ij} = \frac{\partial \Theta_i}{\partial z^j} - \frac{\partial \Theta_j}{\partial z^i} \tag{3.11}$$

Inverting the matrix Ω_{ij} we find:

$$\dot{z}^j = -(\Omega_{ij}(z))^{-1}\frac{\partial H}{\partial z^i} = B_{ij}(z)\frac{\partial H}{\partial z^i} \tag{3.12}$$

where $B(z)$ is the Poisson tensor. Hence the Euler-Lagrange equations of the phase-space Lagrangian (3.8) are just the equations of motion found in the classical Hamiltonian theory (eqs. A.88 and A.97).

Note that a phase-space Lagrangian is always singular, since

$$\frac{\partial \mathcal{L}}{\partial \dot{z} \partial \dot{z}} = 0 \tag{3.13}$$

At this point, the phase-space Lagrangian formalism constitutes a way to rewrite an hamiltonian system with a Lagrangian or a variational framework.

In chapter 5 and 6 we will build an integrator starting from these phase-space Lagrangians.
Such integrators, unlike continuous phase-space Lagrangians, have the unique feature of behaving like regular Lagrangian systems or, equivalently, like canonical hamiltonian systems on a doubled space T^*M, the cotangent bundle of M. Therefore, the degrees of freedom are twice the ones of the original system, and one of the critical points is how these integrators and their flows can be related to the continuous hamiltonian flow in M.

Constraints The theory of non canonical hamiltonian system can be treated from a different point of view. Singular Lagrangians are relevant in other areas of theoretical physics, specially in field theory. The quantization of these systems led Dirac and Bergmann to the development of the so called Dirac-Bergmann theory of constraints (see for example Bergmann [2] and Leon [6]).
The fact that for a singular Lagrangian the Legendre transform is not invertible with respect to the conjugate momenta p^z means that there exist non trivial relations (constraints) between the coordinates of the phase-space T^*M.
It can be proved that every phase-space Lagrangian, or equivalently every hamiltonian system, can be seen as a canonical Hamiltonian system in the extended phase space T^*M, subjected to primary constraints induced by the Legendre transform.
In coordinates, this means that the coordinates $z \in M$ and the conjugate momenta are related by the constraints:

$$\phi(z, p) = \Theta(z) - p = 0 \tag{3.14a}$$

$$p \equiv \frac{\partial \mathcal{L}}{\partial \dot{z}} \tag{3.14b}$$

This formulation is important for our purposes, since we will see that the conservation of these constraints remains valid in some cases in the discrete theory and this fact will be important for a better understanding of our integrators.

Review To summarize, our problem can be studied with four different formulations:

- A non canonical Hamiltonian system on the manifold M.
- A canonical Hamiltonian system defined on T^*Q, with $dim(T^*Q) = dim(M)$. From the Darboux-Lie theorem, this system is well defined only locally.
- A phase-space Lagrangian system on on the tangent bundle TM. It is necessarily singular and a regular second order vector field does not exist.
- A constrained canonical Hamiltonian system on T^*M. The constraints are defined by the Legendre transform of the phase-space Lagrangian.

3.3 Non canonical guiding center theory

The aim of a guiding center theory is to decouple the motion of the guiding center of the particle from its gyration motion.
For a constant magnetic field, this is done easily since the particle performs a simple circular motion, as we have seen at the beginning of chapter 2. For this case, we can choose the coordinates of the guiding center and the gyration angle (X, ζ) as:

$$\mathbf{x} = \mathbf{X} + \rho\hat{\mathbf{e}_\rho} \tag{3.15a}$$
$$\mathbf{v} = u\hat{\mathbf{u}} + w\hat{\mathbf{w}} \tag{3.15b}$$

where ρ is the displacement vector, u and w are the velocities parallel and perpendicular to the magnetic field and $\hat{\mathbf{e_1}}$ and $\hat{\mathbf{e_2}}$ are fixed coordinates where the gyration motion takes place.

With the aid of figure 13, the unit vectors $\hat{\mathbf{e}_\rho}$ and $\hat{\mathbf{e}_\mathbf{w}}$ are found to be:

$$\hat{\mathbf{e}_\rho} = \cos(\zeta)\hat{\mathbf{e_1}} - \sin(\zeta)\hat{\mathbf{e_2}} \tag{3.16a}$$
$$\hat{\mathbf{e}_\mathbf{w}} = -\sin(\zeta)\hat{\mathbf{e_1}} - \cos(\zeta)\hat{\mathbf{e_2}} \tag{3.16b}$$

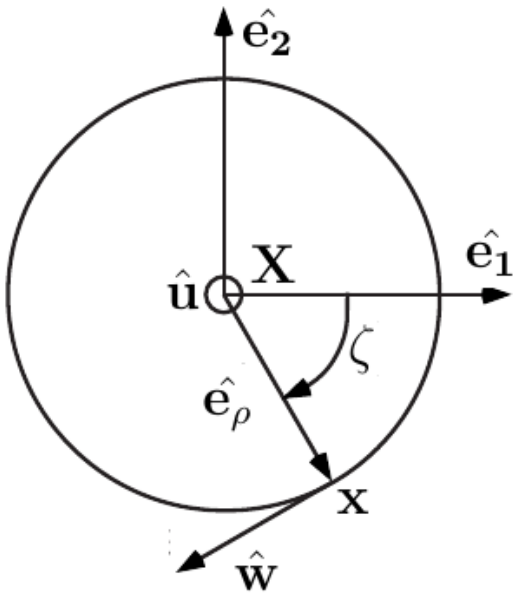


Figure 13: Guiding center coordinates. $\hat{\mathbf{e_1}}$ and $\hat{\mathbf{e_2}}$ represent the fixed frame. $\hat{\mathbf{e}_\rho}$ and $\hat{\mathbf{w}}$ represent the moving frame. $\hat{\mathbf{u}}$ is the unit vector parallel to the magnetic field.

A full set of guiding center coordinates is then $(\mathbf{X}, u, w, \zeta)$.
For a general electromagnetic field, the starting point is to maintain the same transformation (3.15) and find a Lagrangian whose gyration motion can be decoupled from the rest.

Consider a charged particle in an electromagnetic field. Its motion is described by the Lorentz force (2.1). Alternatively, we can write the Lagrangian of the system:

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = \frac{1}{2} m \dot{\mathbf{x}}^2 + \frac{e}{c} \dot{\mathbf{x}} \cdot A(\mathbf{x}, t) - e\Phi(\mathbf{x}, t) \tag{3.17}$$

where A and ϕ are respectively the vector and the scalar potential:

$$\begin{aligned} \nabla \times \mathbf{A} &= \mathbf{B} \\ -\nabla \phi &= \mathbf{E} \end{aligned} \tag{3.18}$$

Normalizing the constants such that $m = c = e = 1$ we can rewrite (3.17) as:

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = \frac{1}{2} \dot{\mathbf{x}}^2 + \dot{\mathbf{x}} \cdot A(\mathbf{x}, t) - \Phi(\mathbf{x}, t) \tag{3.19}$$

The canonical Hamiltonian is easily found with the Legendre transform:

$$\dot{\mathbf{x}} = \mathbf{p} - \mathbf{A} \tag{3.20a}$$

$$H(\mathbf{x}, p, t) = \frac{1}{2} \left(p - A(\mathbf{x}, t)\right)^2 + \Phi(\mathbf{x}, t) \tag{3.20b}$$

From section (3.2), we can write the phase-space Lagrangian simply as:

$$\mathcal{L}(\mathbf{x}, p, \dot{\mathbf{x}}, \dot{\mathbf{p}}) = \dot{\mathbf{x}} \cdot \mathbf{p} - H(\mathbf{x}, \mathbf{p}) \tag{3.21}$$

Since the phase-space Lagrangian formalism is obtained from a variational principle, we can choose arbitrary coordinates. Using the velocity $\mathbf{v} \equiv \dot{\mathbf{x}}$ instead of the momenta, equation (3.21) reads:

$$\mathcal{L}(\mathbf{x}, \mathbf{v}, \dot{\mathbf{x}}, \dot{\mathbf{v}}) = (\mathbf{v} + \mathbf{A}) \cdot \dot{\mathbf{x}} - \left(\Phi + \frac{m}{2} v^2\right) \tag{3.22}$$

The theory of Littlejohn relies on the small gyroradius and slowly varying fields approximation. For this reason, an ordering parameter ϵ is added to the Lagrangian:

$$\mathcal{L}(\mathbf{x}, \mathbf{v}, \dot{\mathbf{x}}, \dot{\mathbf{v}}) = \left(\mathbf{v} + \frac{1}{\epsilon} \mathbf{A}(\mathbf{x}, \epsilon t)\right) \cdot \dot{\mathbf{x}} - \left(\Phi(\mathbf{x}, \epsilon t) + \frac{m}{2} v^2\right) \tag{3.23}$$

The physical significance of the parameter ϵ is that for small values, the electromagnetic fields depend weakly on time and they dominate over the other kinetic energy terms.
A first order Lagrangian is then obtained by expressing every term of the Lagrangian (3.23) in gyroradius coordinates (3.15) with respect to ϵ and by retaining only the ϵ^{-1} and ϵ^0 terms.

Returning to the definition of the guiding center coordinates (3.15), we can observe that X and ρ are not uniquely determined for a general field configuration. Often, ρ is chosen to be:

$$\rho = \frac{w}{B}\hat{\mathbf{e}}_\rho \tag{3.24}$$

so that, remembering that the magnetic field has a $O(\epsilon^{-1})$ dependence, we can write:

$$\mathbf{x} = \mathbf{X} + \epsilon\frac{w}{B}\hat{\mathbf{e}}_\rho \tag{3.25}$$

Other choices of the displacement vector ρ are possible, and they lead to slightly different results (see Cary[4]).

By substituting the new coordinates in the Lagrangian (3.23), one can show that the first orders dependences on the gyration angle can be removed by a gauge transformation on the Lagrangian $\left(L \to L + \frac{df}{dt}\right)$, to give:

$$\boxed{\mathcal{L}(\mathbf{X}, u, \mu, \zeta, \dot{\mathbf{X}}, \dot{u}, \dot{\mu}, \dot{\zeta}, t) = \mathbf{A}^\dagger(\mathbf{X}, t)\cdot\dot{\mathbf{X}} + \mu\dot{\zeta} - \mu B(\mathbf{X}, t) - \frac{u^2}{2} - \Phi(\mathbf{X}, t)} \tag{3.26}$$

where μ is the magnetic moment: $\mu = \dfrac{w^2}{2B}$ and $A^\dagger$ is the modified vector potential:

$$\mathbf{A}^\dagger(\mathbf{X}, t) = \mathbf{A}(\mathbf{X}, t) + u\hat{\mathbf{b}}(\mathbf{X}, t) \tag{3.27}$$

We can recognize that the Lagrangian (3.26) is a phase space Lagrangian, or equivalently has an Hamiltonian structure, by splitting the first terms proportional to the first derivatives of the coordinates from the Hamiltonian terms:

$$\mathcal{L} = \mathbf{A}^\dagger\cdot\dot{\mathbf{X}} + \mu\dot{\zeta} - H(\mathbf{X}, u, \mu, \zeta, t) \tag{3.28a}$$

$$H(\mathbf{X}, u, \mu, \zeta, t) = \mu B + \frac{u^2}{2} + \Phi \tag{3.28b}$$

If we try to compute the Euler-Lagrange equations for the gyroangle ζ, we immediately notice that the magnetic moment is constant along the solution, so that we can safely remove the gyroangle dependency by imposing the constancy of the magnetic moment:

$$\boxed{\mathcal{L}(\mathbf{X}, u, \dot{\mathbf{X}}, \dot{u}, t) = \mathbf{A}^\dagger(\mathbf{X}, t) \cdot \dot{\mathbf{X}} - \mu B(\mathbf{X}, t) - \frac{u^2}{2} - \Phi(\mathbf{X}, t)} \tag{3.29}$$

This phase-space Lagrangian is the starting point of our variational symplectic integrator. In chapter 6 we will directly discretize this Lagrangian to find a suitable integrator. The symplectic one-form of the Lagrangian (3.29), which represents the conjugate momenta, is just:

$$\Theta_i = \begin{pmatrix} \mathbf{A}^\dagger \\ 0 \end{pmatrix} \tag{3.30}$$

The symplectic matrix of the Lagrangian (3.29) is easily found by deriving the symplectic one-form, as in eq. (3.12):

$$\Omega_{ij} = \begin{pmatrix} 0 & -B_z^\dagger & B_y^\dagger & \hat{b}_x \\ B_z^\dagger & 0 & -B_x^\dagger & \hat{b}_y \\ -B_y^\dagger & B_x^\dagger & 0 & \hat{b}_z \\ -\hat{b}_x & -\hat{b}_y & -\hat{b}_z & 0 \end{pmatrix} \tag{3.31}$$

where $\mathbf{B}^\dagger$ is the modified magnetic field:

$$\mathbf{B}^\dagger = \nabla \times \mathbf{A}^\dagger = \mathbf{B} + u \nabla \times \hat{\mathbf{b}} \tag{3.32}$$

The Hamiltonian equations of motion take the form:

$$\Omega_{ij} \dot{z}^j = -\frac{\partial H}{\partial z^i} \tag{3.33}$$

Solving further the equations of motion for u, we immediately find that, along a solution curve:

$$u = \hat{\mathbf{b}} \cdot \dot{\mathbf{X}} \tag{3.34}$$

Hence, the variable u is the velocity parallel to the magnetic field, as we expected. The full equations of motion read:

$$\dot{\mathbf{X}} = \frac{u\mathbf{B}^\dagger - \hat{\mathbf{b}} \times (\mathbf{E} - \mu\nabla B)}{\hat{\mathbf{b}} \cdot \mathbf{B}^\dagger} \tag{3.35a}$$

$$\dot{u} = \frac{\mathbf{B}^\dagger \cdot (\mathbf{E} - \mu\nabla B)}{\hat{\mathbf{b}} \cdot \mathbf{B}^\dagger} \tag{3.35b}$$

In equations (3.35), the fields have been assumed to be independent of time for simplicity. This assumptions will hold for the rest of this thesis.

4 Discrete Mechanics

4.1 Introduction

Symplectic Algorithms In the study of magnetized plasmas, it is often necessary to carry out simulations in a time scale much longer than that of the guiding center motion. In particular, the evaluation of the confinement of particles and their diffusion inside a fusion device, requires the integration of the particle trajectory over a very long time.

With standard integrators, such as the fourth-order Runge-Kutta, the error is guaranteed only to be small in each time steps, but it often accumulates coherently, resulting in large errors after long time. As a result, such numerical integrators has usually been limited to simulations over a limited time.

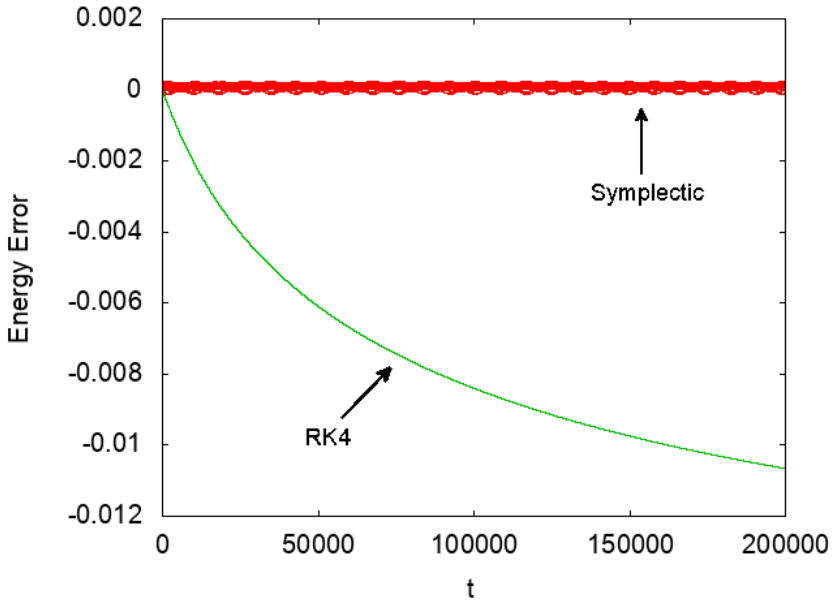


Figure 14: Comparison between the energy error of a fourth order Runge-Kutta and a symplectic method for a mechanical system

Symplectic integrators implicitly possess global conservation properties: momenta associated to symmetries are exactly preserved and the energy presents excellent long time stability features.
These properties make them ideal for simulating systems, either conservative or dissipative and forced, with very long time steps.

Fig.14 shows an example of the energy evolution for a mechanical system. The symplectic method, unlike Runge-Kutta, conserves exactly the energy within a small bounded error.

Symplectic integrators were first introduced by De Vogelaere in 1956 [7] and rediscovered in 1983 by Ruth [24]. Further developments were made in the following years by, for example, Forest and Ruth [8]. Longtime behaviour of symplectic methods for oscillatory and dissipative systems was studied by Hairer [10], Reich [22] and Benettin and Giorgilli [1].

A different approach using a variational view of discrete mechanics has been proposed in the '70s by Cadzow [3]. This method, which is the one used in this thesis, has been further developed recently by Marsden and West for dynamic systems with a well defined Lagrangian. [19]

Discrete Mechanics Since the aim of symplectic integrators is to discretize and to conserve the relevant quantities of continuous Lagrangian and hamiltonian mechanics, we'll follow a similar path of the continuous case (appendix A).
We'll see that the continuous and the discrete theories are strictly related. In particular, the flow of a symplectic integrator is determined by its generating function (called the discrete Lagrangian) and that the two theories are formally equivalent when the generating function is computed exactly.
In practical cases, this is not generally possible and we'll show that approximations of the generating function leads to integrators whose energy is close to the original one.

Symplectic integrators are presented from a Lagrangian and an Hamiltonian point of view. Subsequently, we'll show that the two theories are largely equivalent.

This chapter covers almost only canonical systems, for which a regular Lagrangian is well defined. The reason of this choice is that symplectic integrators for canonical systems are easy to build and they have been extensively studied in the last decades.

In chapter 5 and 6 we'll apply the concepts of this chapter to non canonical systems and in particular to the non canonical guiding center theory.

4.2 Discrete Lagrangian mechanics

Unlike standard integrators, where the equations of motion are derived and discretized, **variational symplectic integrators start from discretizing the ac-**

tion integral followed by extremizing it to obtain a set of discrete Euler-lagrange equations.

In the discrete case, rather than the tangent bundle $TQ \ni (q, \dot{q})$, a time step $h \in \mathbb{R}$ and the discrete state space $Q \times Q \ni (q_k, q_{k+1})$ are used.
Note that the two spaces are locally isomorphic through the transformation

$$(q_k, q_{k+1}) \to \left(q_k, \frac{q_{k+1} - q_k}{h} \right) \tag{4.1}$$

In the same way as in the continuous case (section A.3.1), we need a way to define a good discretization of second order vectors $v \in \ddot{Q}$. Given an element $w \in (Q \times Q) \times (Q \times Q)$, the discrete second order manifold is:

$$\ddot{Q}_d = \left\{ w = ((q_0, q_1), (q_1', q_2)) \in (Q \times Q) \times (Q \times Q) \middle| q_1 = q_1' \right\} \tag{4.2}$$

This is equivalent to the condition $\frac{dx}{dt} = v$ of the continuous case (A.49). To summarize, we have the following correspondences:

$$TQ \to Q \times Q \tag{4.3}$$
$$T(TQ) \to (Q \times Q) \times (Q \times Q) \tag{4.4}$$
$$\ddot{Q} \to \ddot{Q}_d \tag{4.5}$$

In the same way, we can give a discrete version of a curve $q(t)$ in the path space $\mathcal{C}(Q)$, that is, a collection of points $q_d = \{q_k\}_{k=0}^N$. The space of all these collections is called the discrete path space $\mathcal{C}_d(Q)$.
The next step is to discretize the Lagrangian and the action of the system:

$$L(q, \dot{q}) \to L_d(q_0, q_1, h) \tag{4.6a}$$
$$S(q_t) = \int_0^T L(q, \dot{q}) dt \to S_d(q_d) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}, h) \tag{4.6b}$$

where h is the **time step** which represents the time interval between two consecutive points q_k, q_{k+1}. In this work, we will limit to constant time steps, so that $hN = T$.
$L_d : Q \times Q \to \mathbb{R}$ and $S_d : \mathcal{C}_d \to \mathbb{R}$ are the **discrete Lagrangian** and the **discrete action**. The discrete Lagrangian has to be chosen so as to resemble as much as possible the continuous Lagrngian, at least with low time steps h.
In this spirit, we can define the **exact discrete Lagrangian** in this way:

$$L_d^E(q_0, q_1, h) = \int_0^h L(q(t), \dot{q}(t))dt \tag{4.7}$$

where $q(t)$ is a solution curve which passes through q_0 and q_1: $q(0) = q_0, q(h) = q_1$. Performing the sum in eq. (4.6), the discrete action turns out to be the continuous action written in coordinates $(q(0), q(T)$, rather then $(q(0), \dot{q}(0)$.
Note that computing the exact discrete Lagrangian for a solution $q(t)$ obviously requires the knowledge of the exact solution.
A priori, the exact discrete Lagrangian has no particular properties over other discretization choices. However, we will see in section 4.5 that this discretization generates the exact solution, and for this reason it can be use as a link between the continuous system and a particular discretization.

We can build the variations of the discrete action as the external derivative of the action, applied to a variation $\delta q_d \in T_{q_d}\mathcal{C}_d(Q)$:

$$\begin{aligned}\mathbf{d}S_d \cdot \delta q_d = &\sum_{k=1}^{N-1} \left[D_1 L_d(q_k, q_{k+1}, h) + D_2 L_d(q_{k-1}, q_k, h)\right] \cdot \delta q_k \\ &+ D_1 L_d(q_0, q_1, h) \cdot \delta q_0 + D_2 L_d(q_{N-1}, q_N, h) \cdot \delta q_N\end{aligned} \tag{4.8}$$

where D_1 and D_2 are the derivatives with respect to the first and second term. Then, again, we can identify the first and last terms respectively with the **discrete Euler-lagrange map** $D_{DEL} : \ddot{Q}_d \to T^*Q$ and the **discrete Lagrangian 1-forms** $\Theta_d^{\pm} : Q \times Q \to \mathbb{R}$:

$$D_{DEL} L_d((q_{k-1}, q_k), (q_k, q_{k+1})) = D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) \tag{4.9}$$

$$= \frac{\partial}{\partial q_k} \left(L_d(q_{k-1}, q_k) + L_d(q_k, q_{k+1})\right) \tag{4.10}$$

$$\Theta_d^+(q_k, q_{k+1}) = D_2 L_d(q_k, q_{k+1}) \tag{4.11a}$$

$$\Theta_d^-(q_k, q_{k+1}) = -D_1 L_d(q_k, q_{k+1}) \tag{4.11b}$$

If we fix the endpoints of the discrete curve q_d, such that $\delta q_0 = \delta q_1 = 0$, the boundary terms vanish and the summation term in eq.(4.8) forces the discrete Euler-lagrange map to vanish at each timestep, giving the **discrete Euler-lagrange equations**, which are the starting point for the variational symplectic algorithm:

$$\boxed{D_1 L_d(q_k, q_{k+1}, h) + D_2 L_d(q_{k-1}, q_k, h) = 0} \tag{4.12}$$

In coordinates $q = (q^1, .., q^i, .., q^n)$ equation 4.12 is rewritten as:

$$\frac{\partial}{\partial q_i(k)} \left[L_d \left(q(k), q(k+1), h \right) + L_d \left(q(k-1), q(k), h \right) \right] = 0 \tag{4.13}$$

These equations define a way to find, for each time step, the unknown variables q_{k+1} from q_k and q_{k-1}.
The fact that the continuous Euler-Lagrange equations are second-order in q is reflected in the discrete version by the requirement of the knowledge of two points in order to find the next one.
The discrete Lagrangian map $F_{L_d} : (Q \times Q) \to (Q \times Q)$ is the flow of the discrete Lagrangian vector field $X_{L_d} : Q \times Q \to (Q \times Q) \times (Q \times Q)$:

$$D_{DEL} \circ X_{L_d} = 0 \tag{4.14}$$

$$F_{L_d}(q_{k-1}, q_k) = (q_k, q_{k+1}) \tag{4.15}$$

Note that the initial conditions of the integrator are given by a point in $Q \times Q$ (q_0, q_1), rather than $q(0), \dot{q}(0)$.
The Lagrangian map can be evaluated at an arbitrary time step by composing it k times from the initial conditions:

$$(F_{L_d})^k(q_0, q_1) = (q_k, q_{k+1}) \tag{4.16}$$

One difference from the continuous mechanics is that there are two different Lagrangian 1-form and this will be reflected to the fact that there are two different Legendre transforms.
However, we can notice that removing the fixed endpoint condition, the external derivative of the action along a solution of the discrete Euler-lagrange equation takes the form:

$$\mathbf{d}L_d = \Theta^+_{L_d} - \Theta^-_{L_d} \tag{4.17}$$

Thus, taking another derivative of L_d, we find out that there is a unique 2-form $\Omega_d : (Q \times Q) \times (Q \times Q) \to \mathbb{R}$, the **discrete Lagrangian symplectic form**:

$$\mathbf{d}\Theta^+_{L_d} = \mathbf{d}\Theta^-_{L_d} \equiv \Omega_{L_d} \tag{4.18}$$

Following the same method of the continuous case, the discrete symplectic two-form Ω_{L_d} is shown to be conserved by the discrete map:

$$\mathbf{d}S_d = (F^{N-1})^*_{L_d}\Theta^+_{L_d} - \Theta^-_{L_d} \tag{4.19}$$

Hence, using eq.(4.18):

$$(F^{N-1})^*_{L_d}\Omega_{L_d} = \Omega_{L_d} \tag{4.20}$$

The discrete Lagrangian form has the following coordinate expression:

$$\boxed{\Omega_{L_d}(q_k, q_{k+1}) = \frac{\partial^2 L_d}{\partial q_k^i \partial q_{k+1}^j} dq_k^i \wedge dq_{k+1}^j} \tag{4.21}$$

From the discrete Euler-Lagrange equations (4.12), we can observe that the solutions are uniquely defined at every time step if the first term is invertible with respect to q_{k+1}.
Hence, the matrix of the discrete symplectic 2-form must be non singular to give rise to the discrete Lagrangian flow:

$$det\left(\frac{\partial^2 L_d}{\partial q_k \partial q_{k+1}}\right) \neq 0 \tag{4.22}$$

Although we haven't related yet a continuous system to its discretization, it is worth to point out that the conditions of non singularity of the symplectic forms are not equivalent a priori, so that a continuous system can be singular, while its discretization can be regular, and vice versa.

4.2.1 Discrete Noether Theorem

As in the continuous case (section A.3.4) , we can start by defining an action of a Lie Group $\Phi : G \times Q \to Q$ and its lift:

$$\Phi_g^{Q\times Q}(q_k, q_{k+1}) = (\Phi_g(q_k), \Phi_g(q_{k+1})) \tag{4.23}$$

The infinitesimal generator of a Lie algebra element $\mathfrak{g}$ is given by

$$\xi_{Q\times Q}(q_0, q_1) = (\xi_Q(q_0), \xi_Q(q_1)) \tag{4.24}$$

where ξ_Q is an infinitesimal generator on Q and it is the same as in the continuous mechanics:

$$\xi_Q(q) = \frac{d}{dg}\left(\Phi_g^Q(q)\right) \cdot \xi \tag{4.25}$$

If the action $\Phi^{Q\times Q}$ leaves the discrete Lagrangian invariant:

$$L_d \circ \Phi_g^{Q\times Q} = L_d \tag{4.26}$$

then the following discrete momentum map is conserved:

$$J_{L_d}(q_k, q_{k+1}) \cdot \xi = \Theta_{L_d}^{\pm} \cdot \xi_{Q\times Q}(q_k, q_{k+1}) \tag{4.27}$$

$$J_{L_d} \circ F_{L_d} = J_{L_d} \tag{4.28}$$

This theorem is of great importance and represents a striking difference from a standard integrator: if the discrete Lagrangian is symmetric under some trasformation, the quantities associated to Noether theorem are guaranteed to be exactly conserved. We can expect that the same argument doesn't hold for the energy. Fortunately, we will see that a symplectic integrator exhibits energy errors that remain globally bunded for indefinite times.

As in the continuous case, perhaps the simplest example of the Noether theorem is that of an ignorable variable q^m. In this case, the action Φ_g is just a translation along the axis:

$$\Phi_g(q_k, q_{k+1}) = (q_k + ge_x, q_{k+1} + ge_x) \tag{4.29}$$

From eq. (4.26), this condition is easily rewritten as

$$\frac{\partial}{\partial q_k^m} L_d(q_k, q_{k+1}) = -\frac{\partial}{\partial q_{k+1}^m} L_d(q_k, q_{k+1}) \tag{4.30}$$

then, for every time step k, l, the same quantities are conserved:

$$\frac{\partial}{\partial q_k^m} L_d(q_k, q_{k+1}) = \frac{\partial}{\partial q_l^m} L_d(q_l, q_{l+1}) \tag{4.31a}$$

$$\frac{\partial}{\partial q_{k+1}^m} L_d(q_k, q_{k+1}) = \frac{\partial}{\partial q_{l+1}^m} L_d(q_l, q_{l+1}) \tag{4.31b}$$

As we will see in the next subsection, this is just the conservation of the discrete conjugate momenta. The same conservation law will be derived immediately from eq. (4.30) using the so called momentum matching.

4.2.2 Discrete Legendre transforms and momentum matching

Given a discrete Lagrangian, the Legendre transforms $\mathbb{F}^{\pm} : Q \times Q \to T^*Q$ are defined by:

$$\boxed{\begin{aligned} \mathbb{F}^+_{L_d}(q_k, q_{k+1}) &\equiv (q_{k+1}, p^+_{k+1}) = (q_{k+1}, D_2 L_d(q_k, q_{k+1})) \\ \mathbb{F}^-_{L_d}(q_k, q_{k+1}) &\equiv (q_k, p^-_k) = (q_k, -D_1 L_d(q_k, q_{k+1})) \end{aligned}} \tag{4.32}$$

Hence, the discrete Euler-Lagrange equations (4.12) can be seen as an equivalence condition on the Legendre transforms of two consecutive time steps. This is called the **momentum matching**:

$$0 = D_1 L_d(q_k, q_{k+1}, h) + D_2 L_d(q_{k-1}, q_k, h) = -p^-_k + p^+_k \tag{4.33}$$

Thus, there is only one conjugate momentum p_k:

$$p_k = p_k^- = p_k^+ \tag{4.34}$$

Solving the discrete Euler-Lagrange equations is equivalent to find the momentum p_k with (4.32a) and to invert it with respect to q_{k+1} using eq. (4.32b):

$$\begin{aligned} p_{k+1} &= D_2 L_d(q_k, q_{k+1}) \\ p_{k+1} &= -D_1 L_d(q_{k+1}, q_{k+2}) \end{aligned} \tag{4.35}$$

Denoting by F_{H_d} the **discrete Hamiltonian map**, which brings (q_k, p_k) to (q_{k+1}, p_{k+1}), both the Lagrangian and Hamiltonian maps can be represented in terms of the Legendre transforms, as sketched in fig.(15):

$$F_{L_d} = (\mathbb{F}^- L_d)^{-1} \circ \mathbb{F}^+ L_d \tag{4.36a}$$

$$F_{H_d} = \mathbb{F}^+ L_d \circ (\mathbb{F}^- L_d)^{-1} \tag{4.36b}$$

Using equations (4.36), the symplecticity of the flow (4.20) can be rewritten as

$$F_{H_d}^* \Omega_H = \Omega_H \tag{4.37}$$

where Ω_H is

$$\Omega_H = (\mathbb{F}_-^{-1})^* \Omega_{L_d} \tag{4.38}$$

In coordinates, we can write Ω_H from (4.21) just by inverting q_{k+1} with the Legendre transform:

$$\boxed{\Omega_H = \sum_i q_i \wedge p_i} \tag{4.39}$$

Hence, although the space $Q \times Q$ is not directly related to TQ, we have found the important property that the integrator derived from the discrete Euler-Lagrange equations conserve the canonical symplectic form Ω_H in T^*Q.

Recalling the last subsection, we can observe that the invariance along an axis (4.30) can be written in terms of the momenta as:

$$p_{k-1}^- = p_k^+ \tag{4.40}$$

so that, using eq. (4.34), we immediately find, for every time step:

$$p_{k-1} = p_k \tag{4.41a}$$

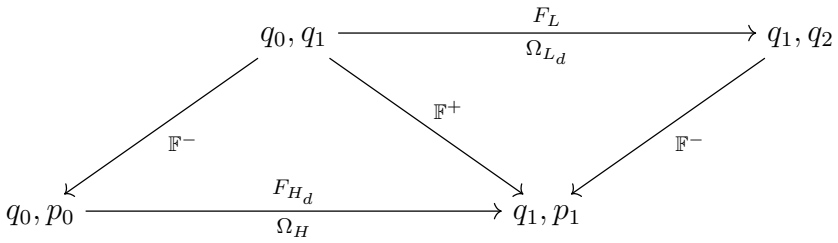


Figure 15: Discrete Lagrangian and Hamiltonian maps and their relations with the Legendre transforms

4.2.3 Examples of Discretization

In section 4.5 we will show that the exact discrete Lagrangian, defined in (4.7), is the best choice for a discretization. For now, we will take this as an assumption. Since the exact discrete Lagrangian is just a formal expression and can't be computed practically, the simplest method to find a discretized Lagrangian, or equivalently to approximate the integral in eq.(4.7), is accomplished by assuming a constant value of the Lagrangian over a time step. More specifically, the Lagrangian is evaluated at the following space and velocity points:

$$\dot{q}(t) \simeq \frac{q_{k+1} - q_k}{h} \tag{4.42a}$$

$$q(t) \simeq q_\alpha = (1-\alpha)q_k + \alpha q_{k+1}, \quad \alpha \in [0,1] \tag{4.42b}$$

The most common values of α are $\alpha = 0, 1, 1/2$:

$$L_d^0(q_k, q_{k+1}) = hL\left(q_k, \frac{q_{k+1} - q_k}{h}\right) \tag{4.43a}$$

$$L_d^1(q_k, q_{k+1}) = hL\left(q_{k+1}, \frac{q_{k+1} - q_k}{h}\right) \tag{4.43b}$$

$$L_d^{1/2}(q_k, q_{k+1}) = hL\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}\right) \tag{4.43c}$$

The latter discretization is called the **midpoint rule**, which is in general more precise than any other choice of α, as we will show in the following paragraph.

Order of the approximation In order to compute quantitatively the difference between the approximated and the exact Lagrangian, the latter can be expressed in a Taylor series in powers of h, expanding the continuous Lagrangian at $t = 0$ and computing the definite integral of each term in eq.(4.7):

$$L_d^E(q_0, q_1) = hL(q(0), \dot{q}(0)) + \frac{1}{2}h^2\left(\frac{\partial L}{\partial q}(q(0), \dot{q}(0)) \cdot \dot{q}(0) + \frac{\partial L}{\partial \dot{q}}(q(0), \dot{q}(0)) \cdot \ddot{q}(0)\right) + O(h^3) \tag{4.44}$$

Applying the same method to the approximated discrete Lagrangians (4.43), we find:

$$L_d^\alpha(q_0, q_1) = hL(q(0), \dot{q}(0)) + \frac{1}{2}h^2 \left(2\alpha \frac{\partial L}{\partial q}(q(0), \dot{q}(0)) \cdot \dot{q}(0) + \frac{\partial L}{\partial \dot{q}}(q(0), \dot{q}(0)) \cdot \ddot{q}(0) \right) + O(h^3) \tag{4.45}$$

Clearly, with every choice of α, the two Lagrangians are equal up to the first term, thus they are a first order approximation in h. Only the midpoint rule ($\alpha = 1/2$) is a second order approximation and hence generally preferable.
For particular Lagrangians, however, it is possible to obtain second and higher order approximations using methods different from the midpoint rule.

4.3 Discrete Hamiltonian mechanics

A different approach to symplectic algorithms is given by Hamiltonian integrators, whose aim is to build a solution of the system which conserves exactly the symplectic form, starting directly from the Hamiltonian and the hamilton's equations.
A common approach is by using generating functions of the system. The following treatment can be found in literature in standard books and articles (for example Hairer [11]).

Given a canonical Hamiltonian system, we already know from the continuous mechanics (section A.4.5) that the Hamiltonian $H : T^*Q \to \mathbb{R}$ and the generating function $S^1(q_0, p_1)$ is related through the Hamilton-Jacobi equation:

$$H\left(q_0 + \frac{\partial S^1}{\partial p_1}(q_0, p_1, h), p_1\right) - \frac{\partial S^1}{\partial h}(q_0, p_1, h) = 0 \tag{4.46}$$

where h is the time step.
If a generating function is known, the Hamiltonian flow can be built from (A.123) with the following equations:

$$p_1 - p_0 = -\frac{\partial}{\partial q_0} S^1(q_0, p_1, h) \tag{4.47a}$$

$$q_1 - q_0 = \frac{\partial}{\partial p_1} S^1(q_0, p_1, h) \tag{4.47b}$$

Again, this could be a good integrator, the solution of which being the exact solution, but of course the explicit knowledge of a generating function is not practically possible.
Fortunately, we can expand a generating function in a Taylor serie to obtain different generating functions. Specifically:

$$S^1(q_0, p_1, t) = hS_1^1(q_0, p_1) + h^2 S_2^1(q_0, p_1) + O(h^3) \tag{4.48}$$

The equations of motion (4.47) are then rewritten as:

$$p_1 - p_0 = -\frac{\partial}{\partial q_0}\left(hS_1^1(q_0, p_1) + h^2 S_2^1(q_0, p_1) + O(h^3)\right) \tag{4.49a}$$

$$q_1 - q_0 = \frac{\partial}{\partial p_1}\left(hS_1^1(q_0, p_1) + h^2 S_2^1(q_0, p_1) + O(h^3)\right) \tag{4.49b}$$

By comparing the terms of the expansion with the Hamilton-Jacobi equation, every term can be easily computed. For example, the first one and the second one

are:

$$S_1^1(q_0, p_1) = H(q_0, p_1) \tag{4.50a}$$

$$S_2^1(q_0, p_1) = \frac{1}{2}\left(\frac{\partial H}{\partial p}\frac{\partial H}{\partial q}\right)(q_0, p_1) \tag{4.50b}$$

Of course, truncating the serie won't give rise to a solution to the Hamilton-Jacobi equation for the Hamiltonian H, since the correct solution is the whole generating function S^1.
However, by virtue of the linearity of equations (4.47), we can notice that the truncation is a generating function itself, hence the corresponding flow must conserve the canonical symplectic form.

This constitutes a starting point for the construction of a symplectic integrator. The two simplest examples, the **symplectic Euler integrators**, are built with the generating functions S^1 and S^2, for which the same arguments hold, truncated to the first term:

$$\begin{cases} \dfrac{p_1 - p_0}{h} = -\partial_{q_0} H(q_0, p_1) \\ \dfrac{q_1 - q_0}{h} = \partial_{p_1} H(q_0, p_1) \end{cases} \tag{4.51}$$

$$\begin{cases} \dfrac{p_1 - p_0}{h} = -\partial_{q_0} H(q_1, p_0) \\ \dfrac{q_1 - q_0}{h} = \partial_{p_1} H(q_1, p_0) \end{cases} \tag{4.52}$$

Comparing (4.47) with (4.51), it is immediate to see that the flows of the symplectic Euler methods are first order in h, relative to the exact flow.
By using similar arguments, it is possible to prove that this slightly more sophisticated example, the **Hamiltonian midpoint rule**, is symplectic and its flow is second order:

$$\begin{cases} \dfrac{p_1 - p_0}{h} = -\partial_{\tilde{q}} H(\tilde{q}, \tilde{p}) \\ \dfrac{q_1 - q_0}{h} = \partial_{\tilde{p}} H(\tilde{q}, \tilde{p}) \end{cases} \tag{4.53}$$

where $\tilde{q}$ and $\tilde{p}$ are:

$$\tilde{q} = \frac{q_0 + q_1}{2}, \quad \tilde{p} = \frac{p_0 + p_1}{2} \tag{4.54}$$

4.4 Lagrangian and Hamiltonian integrators correspondence

Although Lagrangian and Hamiltonian integrators, as presented in these sections, are constructed in a different way, they are strictly related. We start by giving a direct relation for the simplest methods, the Euler and the midpoint. In a second stage, we will seek for a general relation.

Defining q_α and p_α by

$$q_\alpha = (1-\alpha)q_0 + \alpha q_1 \tag{4.55a}$$
$$p_\alpha = \alpha q_0 + (1-\alpha)q_1 \tag{4.55b}$$

Recall that simple first order Lagrangian methods can be constructed using the following discrete Lagrangian:

$$L_d^\alpha(q_0, q_1) = hL\left(q_\alpha, \frac{q_1 - q_0}{h}\right) \tag{4.56}$$

The conjugate momenta are easily computed with the discrete Legendre transforms (eq. 4.32):

$$p_0 = -\frac{\partial}{\partial q_0} L_d^\alpha(q_0, q_1) = -h(1-\alpha)\frac{\partial}{\partial q} L\left(q_\alpha, \frac{q_1 - q_0}{h}\right) + \frac{\partial}{\partial \dot{q}} L\left(q_\alpha, \frac{q_1 - q_0}{h}\right) \tag{4.57a}$$
$$p_1 = \frac{\partial}{\partial q_1} L_d^\alpha(q_0, q_1) = \quad h\alpha\frac{\partial}{\partial q} L\left(q_\alpha, \frac{q_1 - q_0}{h}\right) + \quad \frac{\partial}{\partial \dot{q}} L\left(q_\alpha, \frac{q_1 - q_0}{h}\right) \tag{4.57b}$$

Rearranging them, we find:

$$\frac{p_1 - p_0}{h} = \frac{\partial}{\partial q} L\left(q_\alpha, \frac{q_1 - q_0}{h}\right) \tag{4.58a}$$
$$p_\alpha = \frac{\partial}{\partial \dot{q}} L\left(q_\alpha, \frac{q_1 - q_0}{h}\right) \tag{4.58b}$$

The last equation is just the continuous Legendre transform, written in discretized variables.
If the Lagrangian is regular, denote by H the Hamiltonian of the correspondent canonical system. Then, by inverting (4.58b) and remembering the form of the hamilton's equations, we get:

$$\frac{q_1 - q_0}{h} = \frac{\partial H}{\partial p}(q_\alpha, p_\alpha) \tag{4.59a}$$
$$\frac{p_1 - p_0}{h} = -\frac{\partial H}{\partial q}(q_\alpha, p_\alpha) \tag{4.59b}$$

For $\alpha = 0, 1$ we reobtain the Hamiltonian symplectic Euler methods (4.51), while for $\alpha = \frac{1}{2}$ we get the midpoint rule (4.53).
It is worth noting again that this relation holds only for regular Lagrangian systems, such that the continuous Legendre transform equation (4.58b) is invertible (for a singular Lagrangian the canonical Hamiltonian is not even defined).

Generating functions Nonetheless, it is possible to find a link between Lagrangian and Hamiltonian algorithms for every regular integrator, even for the ones that derive from a singular Lagrangian. Lall and West [14] showed this in the context of optimal control theory. Here we give an explanation using generating functions.

Recalling the last sections, a variational integrator with discretete Lagrangian L_d conserves the canonical symplectic form, if written in coordinates (q_k, p_k):

$$F_{H_d}^* \Omega_H = \Omega_H \tag{4.60}$$

Also, the integrator is governed by the equations:

$$p_0 = -\frac{\partial}{\partial q_0} L_d(q_0, q_1) \tag{4.61a}$$

$$p_1 = \frac{\partial}{\partial q_1} L_d(q_0, q_1) \tag{4.61b}$$

Comparing this to equations (A.117), we can immediately recognize that the discrete Lagrangian is a generating function for the discrete flow.

Hence, given a generic symplectic integrator, we can find the generating functions $L_d(x_0, x_1)$, $S^1(x_0, p_1)$ or $S^2(x_1, p_0)$ and build the correspondent Lagrangian or Hamiltonian integrator, in form of the Euler symplectic method.

If a discrete Lagrangian L_d is already known, i.e. by discretizing the continuous action with some method, we can find the Hamiltonian integrator symply by changing coordinates as we did with equations (A.123).
Defining the **right** and **left Hamiltonian** $H^+(q_0, p_1) = \dfrac{S^1(q_0, p_1)}{h}$, $H^-(q_1, p_0) = \dfrac{S^2(q_1, p_0)}{h}$, equations (A.123) read:

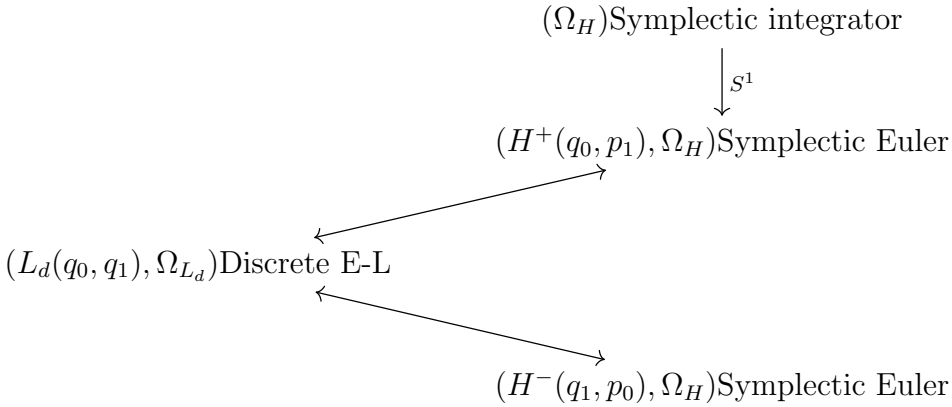


Figure 16: Relation between Lagrangian and Hamiltonian integrators

$$H^+(q_0,p_1) = \frac{p_1 \cdot (q_1(q_0,p_1) - q_0) - L_d(q_0, q_1(q_0,p_1))}{h} \tag{4.62}$$

$$H^-(q_1,p_0) = \frac{p_0 \cdot (q_1 - q_0(q_1,p_0)) - L_d(q_0(q_1,p_0), q_1)}{h} \tag{4.63}$$

where q_1 and q_0 are inverted with the discrete Legendre transforms $\mathbb{F}^{\pm}$.
Of course, this is just a formal relation: in practice writing H^+ or H^- explicitly with respect to the discrete Lagrangian is possible only when the Legendre tranforms can be easily inverted.
The equations for H^+ and H^- are respectively (4.51) and (4.52).
This scheme is skecthed in fig. (16)

One can immediately show that for a Lagrangian Euler method ($q_\alpha, \alpha = 0, 1$), the right and left Hamiltonian are just the continuous Hamiltonian. Hence a Lagrangian Euler method is just an Hamiltonian Euler method, as we have already checked in the last section.
For a Lagrangian midpoint rule, one obtains a different Hamiltonian H'. This is expected, since the midpoint rule is a second order integrator, hence its generating function includes the first (H) and other terms in the taylor expansion in (4.48). Fig. (17) illustrates this behaviour.

Example As a reference, let's consider a simple one dimensional system with an elastic potential. The Lagrangian and the Hamiltonian are:

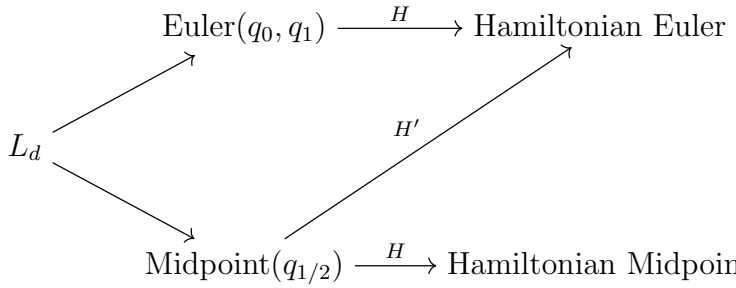


Figure 17: Relation between Lagrangian and Hamiltonian methods

$$L(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \tag{4.64a}$$

$$H(x, p) = \frac{p^2}{2m} + \frac{1}{2}kx^2 \tag{4.64b}$$

The discrete Lagrangian, for the midpoint rule, reads:

$$L_d^{1/2}(q_0, q_1) = hL\left(q_{1/2}, \frac{q_1 - q_0}{2}\right) = h\left(\frac{m(x_1 - x_0)^2}{2h^2} - \frac{1}{8}k(x_0 + x_1)^2\right) \tag{4.65}$$

The discrete momentum p_1 is:

$$p_1 = \frac{\partial}{\partial q_1}L_d(q_0, q_1) = h\left(\frac{m(x_1 - x_0)}{h^2} - \frac{1}{4}k(x_0 + x_1)\right) \tag{4.66}$$

Applying eq. (4.62), we get the following expression for H^+:

$$H^+(x_0, p_1) = -\frac{2\left(hkp_1x_0 + kmx_0^2 + p_1^2\right)}{h^2k - 4m} \tag{4.67}$$

Note that H^+ is just the continuous Hamiltonian $H(x_0, p_1)$ when h is set to 0. This agrees with our assumption that the generating function of the discrete flow is a truncation of the Taylor serie of the generating function of the exact continuous flow.

We can try to compute directly the first two terms of the generating function S^1:

$$S_0^1(x_0, p_1) + hS_1^1(x_0, p_1) = \frac{hkp_1x_0 + kmx_0^2 + p_1^2}{2m} \tag{4.68}$$

Hence, comparing the two expressions, the midpoint rule is a second order method, as we expected:

$$H^+(x_0, p_1) = S_0^1(x_0, p_1) + hS_1^1(x_0, p_1) + O(h^2) \tag{4.69}$$

4.5 Discrete-Continuous correspondence

In the last section we proved that the discrete Lagrangian is a generating function for the discrete flow. Recall that the exact discrete Lagrangian is defined by:

$$L_d^E(q_0, q_1, h) = \int_0^h L(q(t), \dot{q}(t))dt \tag{4.70}$$

Of course, this is just the action of the system computed in a time step $[0, h]$, which is, remembering when we defined the generating functions in continuous mechanics (section A.4.5), the Jacobi solution of the Hamilton-Jacobi equation.
Thus, it is a generating function for the continuous flow, too, and we conclude that the discrete Euler Lagrangian equations of an exact discrete Lagrangian give exactly the continuous solution.

More specifically, if q_k and $q(t)$ are respectively a solution for L_d^E and for L, and h is a fixed time step, there is an exact mapping between them, when the same initial conditions are used:

$$q_k = q(hk) \tag{4.71}$$
$$p_k = p(hk) \tag{4.72}$$

The main goal of this section is to study how the discrete and continuous flow are related when a discretization of the Lagrangian to some order is used. This can be done easily both in the Lagrangian and Hamiltonian side.

Given a Lagrangian L and its exact discretization L_d^E, let's say we have a discrete Lagrangian of order r:

$$L_d(q_0, q_1) = L_d^E(q_0, q_1) + O(h^{r+1}) \tag{4.73}$$

Fixing two points q_0, q_1 and denoting by $q(t)$ an exact solution which passes through them, we can relate the exact and the discrete Legendre transforms by substituting (4.70) into (4.73) and computing the discrete Legendre transform:

$$\begin{aligned} \mathbb{F}^- L_d(q_0, q_1) &= -\frac{\partial}{\partial q_0} L_d(q_0, q_1) = \\ &= -\int_0^h \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right) \frac{\partial q(t)}{\partial q_0} dt - \frac{\partial L}{\partial \dot{q}} \frac{\partial q(t)}{\partial q_0} \Big|_0^h + O(h^{r+1}) \end{aligned} \tag{4.74}$$

Since $q(t)$ is a solution, the first term vanishes, while the second term is null at $q(h)$:

$$\mathbb{F}^- L_d(q_0, q_1) = \frac{\partial L}{\partial \dot{q}}(q(0), \dot{q}(0)) + O(h^{r+1}) \tag{4.75}$$

The same result holds for the right Legendre transform:

$$\mathbb{F}^+ L_d(q_0, q_1) = \frac{\partial L}{\partial \dot{q}}(q(h), \dot{q}(h)) + O(h^{r+1}) \tag{4.76}$$

Remembering that the discrete Lagrangian and Hamiltonian flows are just a composition of the Legendre transforms, we can invert equations (4.75) and (4.76) to find:

$$F_{H_d} = F_H + O(h^{r+1}) \tag{4.77}$$

In particular, for the exact discrete Lagrangian, the discrete and the continuous flows are equal. This is just what we proved before.
Equation (4.77) tells us that the integrator is consistent with the original problem: when the time step h approaches 0, the discrete flow gets closer to the exact solution.

In the Hamiltonian formulation, the result is immediate.
If the generating function of the discrete flow agrees up to the r term with respect to the exact generating function:

$$S'^1(q_0, p_1) = hH^+(q_0, p_1) = \sum_{i=1}^{r} h^i S_i^1(q_0, p_1) + O(h^{r+1}) = S^1(q_0, p_1) + O(h^{r+1}) \tag{4.78}$$

then, from the equations for S^1, the Hamiltonian flow is of order r (equation 4.77).

4.5.1 Conservation of Energy

One of the main features of the symplectic integrators is their ability to bound the energy error for indefinitely long time. This behaviour is due to the fact the discrete flow is interpolated almost exactly by a slightly modified Hamiltonian system whose energy is close to the original one.
This result was proved in the 90's by Benettin and Giorgilli[1], Reich[22] and Hairer[10] by backward error analysis.

Assume we have a symplectic integrator, consistent to order r with an Hamiltonian system (M, Ω_M) with Hamiltonian H:

$$F_H^* \Omega_M = \Omega_M \tag{4.79a}$$
$$F_{H_d}^* \Omega_M = \Omega_M \tag{4.79b}$$
$$F_{H_d} = F_H + O(h^{r+1}) \tag{4.79c}$$

Following the discussion of section A.4, the discrete flow F_{H_d} must be locally Hamiltonian.
It can be proved (see Hairer [11]) that the Hamiltonian H' relative to the discrete flow, called the **modified Hamiltonian**, can be expressed as a serie with respect to h, where the first term is the Hamiltonian H:

$$H' = H + h^r H_{r+1} + h^{r+1} H_{r+2} + ... \tag{4.80}$$

Unfortunately, at high orders in h, the serie (4.80) has usually divergent terms. In practice, one has to truncate it to an order N:

$$H'_N = H + h^r H_{r+1} + ... + h^{N-1} H_N \tag{4.81}$$

Then, under weak assumptions on the flow, we have the following important result for the discrete flow z_k of the integrator (see references at the beginning of this section for the proof):

$$H'_N(z_{k+1}) = H'_N(z_k) + O(e^{\frac{-k}{2h}}) \tag{4.82a}$$
$$H(z_{k+1}) = H(z_k) + O(h^r) \tag{4.82b}$$

The first equation tells us that the Hamiltonian system with Hamiltonian H'_N interpolates almost exactly the discrete flow, while the second is the statement that the energy error of the discrete flow is bounded for all time steps.
Note that the last equation can be easily checked by studying the energy bounding when h is varied.

4.6 Review

In section 4.2 and 4.3, we have seen that given a canonical Hamiltonian system or a Lagrangian system, the construction of a symplectic integrator is easy: starting from a Lagrangian, one can discretize the action to find an integrator which conserves exactly a symplectic form Ω_{L_d} on $Q \times Q$.
Using the discrete Legendre transforms (eq. 4.32), the Hamiltonian flow is defined in a natural way and it is easily proved to conserve the canonical symplectic form on T^*Q.
The same can be done in the Hamiltonian side by truncating the generating function of the continuous flow in its Taylor expansion. The generating function thus found generates a canonical symplectic transformation.

We have seen in section 4.4 that the two methods are equivalent: given a generating function of an Hamiltonian integrator, the correspondent discrete Lagrangian is found by a change of coordinate. Figure 18 summarizes these facts.

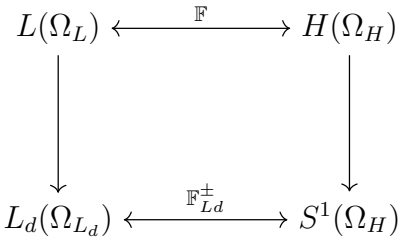


Figure 18: Continuous and discrete Lagrangian and hamiltonian systems and their relations. Ω_L and Ω_{L_d} are the continuous and the discrete Lagrangian symplectic forms. Ω_H is the canonical symplectic form. $\mathbb{F}$ and $\mathbb{F}_{\mathbb{L}}$ are the continuous and the discrete legendre transforms. The discrete Lagrangian L_d and the function S^1 are the generating functions of the Lagrangian and the hamiltonian integrators.

Finally, in the last section we proved that a symplectic integrator which conserves the same symplectic form of an Hamiltonian system and it is consistent with its flow is interpolated by a Hamiltonian system with energy close to the original one.
Note that this happens independently on the symplectic form considered, so it applies to canonical systems as well as non canonical ones.
Figure 19 summarizes this last statement.

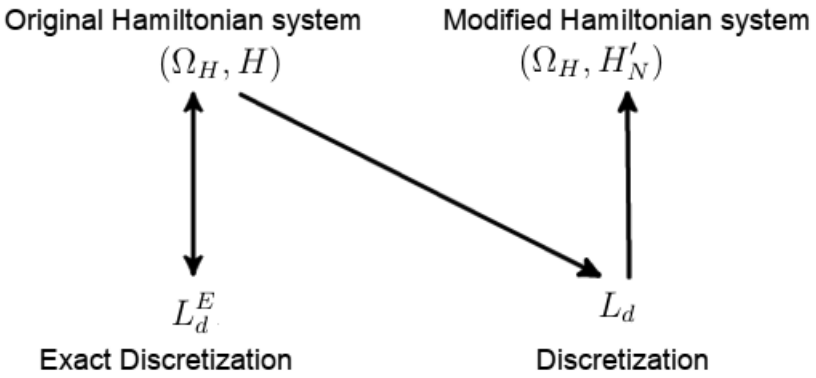


Figure 19: The original Hamiltonian system and its discretization along with the interpolating modified Hamiltonian system (H'_N))

5 Variational phase-space integrators

5.1 Introduction

In chapter 4 we saw that given an Hamiltonian system (Ω_M, M) with Hamiltonian H, the main requirement of a symplectic integrator is to conserve exactly the same symplectic form and to be consistent with the continuous system. (eqs. 4.79)
In this case, the integrator is interpolated almost exactly by a modified Hamiltonian system whose energy is close to the original one and bounded for exponentially long times.

We also showed that for canonical systems, the construction of such integrators is easy, at least for low order schemes, such as Euler symplectic or the mipoint rule. In fact, given a regular Lagrangian, or a canonical system, the discrete Euler-lagrange equations applied to a particular discretization give rise to a flow which conserves exactly the same canonical form.
Hence, if the discretization is consistent with the original problem, and this is the case for the midpoint rule or the Euler symplectics, as we proved in section 4.2.3, the energy will be correctly bounded.

However, for general non canonical systems, integrators that conserve exactly the non canonical symplectic structure are difficult to find and there are no general techniques (see Hairer [11] and Karasözen [13]).
Usually, one should make use of the Darboux-Lie theorem and resctrict to canonical coordinates, where a symplectic integrator can be easily applied. Canonical coordinates, though, are often difficult to find.

5.1.1 Continuous theory

The phase-space variational integrators are based on the idea that every Hamiltonian system can be obtained from a variational principle (see section 3.2.1).
In particular, the equations of motion are equivalent to the Euler-lagrange equations of the phase-space Lagrangian:

$$\mathcal{L} = \Theta(z) \cdot \dot{z} - H(z) \tag{5.1}$$

where Θ is the Hamiltonian one-form (eq. 3.30) and H is the Hamiltonian of the system.
The fact that this Lagrangian is necessarily singular means that in the continuous limit a regular flow $(z_0, \dot{z_0}) \to (z(t), \dot{z}(t))$ does not exist.

In practice, the flow of the system is $F : z_0 \to z(t)$ which arises from the non canonical hamiltonian system or from the first order Euler-Lagrange equations of the phase-space Lagrangian.

Alternatively, we can study the system (see section 3.2.2) as a constrained canonical system in the extended phase-space T^*M. The constraints which arises from the degeneracy of the Legendre transform are:

$$\phi(z,p) = \Theta(z) - p = 0 \tag{5.2a}$$

$$p \equiv \frac{\partial \mathcal{L}}{\partial \dot{z}} \tag{5.2b}$$

5.1.2 Discrete theory

As we will see shortly, variational integrators applied to a phase-space Lagrangian are in general regular.The fact that these integrators are obtained from the discrete Euler-lagrange equations guarantees that the discrete flow conserves exactly the Lagrangian symplectic form $\Omega_{\mathcal{L}_d}$ on $M \times M$ or equivalently the canonical symplectic form on T^*M.

This is very different from the continuous limit: in that case, the phase-space Lagrangian is singular and the flow conserves the (non canonical) symplectic form on the smaller space M. Therefore, the main problem is how the discrete flow can be related to the continuous non canonical Hamiltonian flow in M, since the main requirement of a symplectic integrator is to conserve exactly the same symplectic form of the original continuous problem.

Since the integrator is defined in $M \times M$, the degrees of freedom are twice the ones of the continuous case, so that particular attention must be paid on how the initial conditions are chosen.

Also, since the discrete flow conserves the canonical symplectic form on T^*M, one can wonder if the integrator behaves as a discrete version of the Dirac-Bergmann theory and in particular if it conserves the Dirac constraints (5.2).

These aspects will be analyzed in the following sections.

Figure 20 summarizes the spaces we will have to deal with.

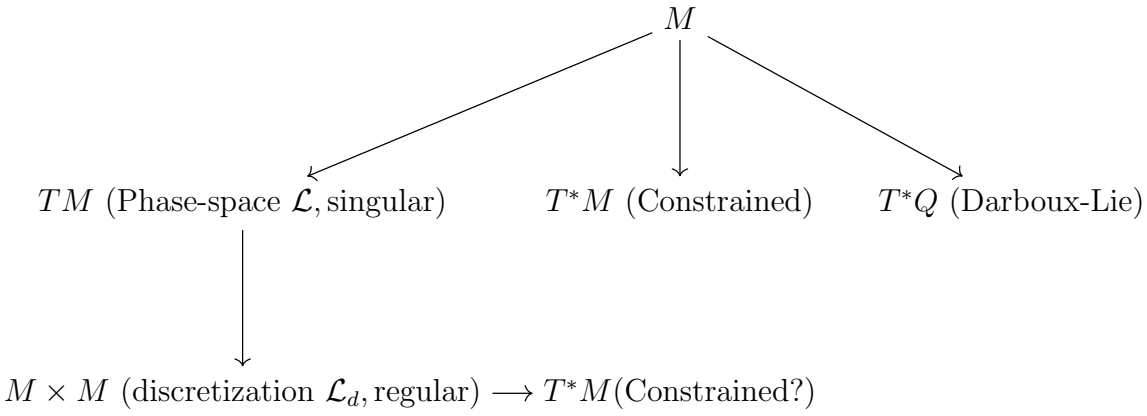


Figure 20: Hamiltonian, Lagrangian and phase-space Lagrangian systems and their relation

Summary Numerical integrations of non canonical systems using this variational approach were performed in a limited number of cases in the past and the theoretical aspects are still mainly unexplored.
Rowley and Marsden [23] studied a point vortex, while Qin and Guan [20] applied the variational integrator to the non canonical guiding center theory. The next chapter (chapter 6) is based mainly on the latter work.

In the following sections, we will study some general properties of the variational integrators applied to a phase-space Lagrangian.
We will show that integrators of Hamiltonian systems with a constant symplectic form conserve the Dirac constraints (eq. 5.2), thus their flow can be projected to the original submanifold M.
Also, for a general Hamiltonian system, the flow of the variational integrator splits in two distinct parts which are interpolated by two different modified canonical system in T^*M.

5.2 General Properties

5.2.1 Midpoint Rule

We limited our discussion to the midpoint discretization, as it is already a difficult task. Higher order methods, such as symplectic Runge-Kutta discretizations, constitute a good subject for future investigations.

Remember from section 4.2.3 that the midpoint is defined as:

$$\begin{aligned}\mathcal{L}_d(z_k, z_{k+1}) &= h\mathcal{L}\left(\tilde{z}_k, \frac{\Delta z_k}{h}\right) \\ &= \Theta(\tilde{z}_k)\cdot\Delta z_k - hH(\tilde{z}_k)\end{aligned} \tag{5.3}$$

where $\tilde{z}$ and Δz are:

$$\tilde{z_k} = \frac{z_k + z_{k+1}}{2} \tag{5.4a}$$

$$\Delta z_k = z_{k+1} - z_k \tag{5.4b}$$

Regularity For a regular Lagrangian system the Legendre tranform must be invertible:

$$det\left(\frac{\partial^2\mathcal{L}}{\partial\dot{z}\partial\dot{z}}\right) \neq 0 \tag{5.5}$$

The condition of regularity for the discretized midpoint rule is slightly different:

$$det\left(\frac{\partial^2\mathcal{L}_d}{\partial z_k\partial z_{k+1}}\right) \neq 0 \tag{5.6}$$

For a phase-space Lagrangian, this matrix is a sum of the symplectic matrix and other terms:

$$\frac{\partial^2\mathcal{L}_d}{\partial z_k\partial z_{k+1}} = \frac{\Omega(\tilde{z}_k)}{2} + M(z_k, z_{k+1}, h) \tag{5.7}$$

Since the symplectic matrix Ω is invertible by definition, the midpoint rule is in general regular.

Correspondent Hamiltonian method In section 4.4 we proved these two statements:
(i) Every regular symplectic integrator can be brought in Hamiltonian Euler form by using the generating function of its flow.
(ii) The Euler symplectic and the midpoint rule are equivalent to their correspondent Hamiltonian method if and only if the continuous Legendre transform is invertible.

For the latter argument, we conclude that the midpoint rule, applied to a phase-space Lagrangian, can't be treated as an Hamiltonian midpoint rule, as the regular Lagrangian case, rather one should find the correspondent Hamiltonian Euler method with equations (4.62).

5.2.2 Local correspondence between discrete and continuous flows

In chapter 4 we studied the order of the discrete Lagrangian and its Legendre transform by expanding the exact discrete Lagrangian in series of h and by comparing each term with a particular choice of the discrete Lagrangian (see eqs 4.44 and 4.45). If we try to apply the same arguments to a phase-space Lagrangian, we immediately notice that an exact discrete Lagrangian is not well defined, at least in the whole space $M \times M$, since a continuous solution required in the definition of L_d^E is completely determined by a single point, hence given two generic points z_0 and z_1, a solution curve passing through them doesn't exist in general.

One possible solution is to restrict our attention only to pairs of points consistent with a solution, so that a continuous solution curve which passes through them does exist. Since our purpose is to study an integrator whose flow is, hopefully, close to the continuous flow, we extend the definition of the exact discrete Lagrangian to points near a solution, in this way:

$$\mathcal{L}_d^E(z_0, \bar{z}_1, h) = \int_0^h L(\bar{z}(t), \dot{\bar{z}}(t))dt \tag{5.8}$$

where $\tilde{z_1}$ is a point close to a consistent point z_1 and $\bar{z}(t)$ is a curve close to a solution $z(t)$ passing through z_0 and z_1:

$$\bar{z}_1 = z_1 + O(h^N) \tag{5.9a}$$
$$\bar{z}(t) = z(t) + O(t^N) \tag{5.9b}$$

where the notation $o(h^N)$ indicates that the displacement from the exact solution is small with respect with the time scale.
This situation is sketched in fig.21. Note that when the two points are consistent, the curve used is just the continuous solution.

With reference to equations (4.44) and (4.45), we can notice that the form of the Taylor expansion is independent on the choice of the curve, so that L_d^α is still consistent with the exact discrete Lagrangian up to the first order generally and to second order for the midpoint rule:

$$\mathcal{L}_d^\alpha(z_0, z_1) = \mathcal{L}_d^E(z_0, z_1) + o(h^r) \tag{5.10}$$

for every points $(z_0, z_1) \in M \times M$.
The next question is how the discrete Legendre transforms are related to the continuous one.
Rewriting eq. (4.74) one can easily show that the term inside the integral gives rise

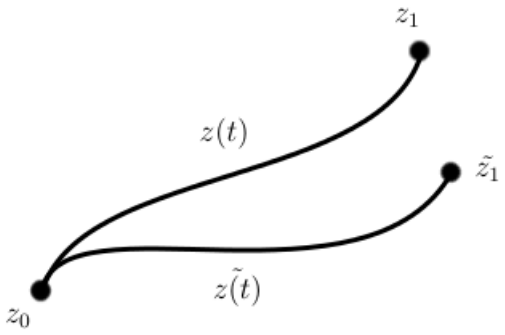


Figure 21: Curves used in the defintion of the exact discrete Lagrangian

to a term $o(h^N)$, so that:

$$\begin{aligned}\mathbb{F}^-\mathcal{L}_d^\alpha(z_0,z_1) &= -\frac{\partial}{\partial z_0}\left(\mathcal{L}_d^E(z_0,z_1)+O(h^r)\right) = \\ &= -\int_0^h\left(\frac{\partial\mathcal{L}}{\partial z}-\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{z}}\right)\frac{\partial z(t)}{\partial z_0}dt-\frac{\partial\mathcal{L}}{\partial\dot{z}}\frac{\partial z(t)}{\partial z_0}\Bigg|_0^h+O(h^N)+O(h^r)= \\ &= \frac{\partial\mathcal{L}}{\partial\dot{z}}(z_0)+O(h^N)+O(h^r)\end{aligned} \tag{5.11}$$

This leads to the following results:

- If two consistent points are chosen, either by imposing the continuous constraint $\phi(z_0,p_0)=0$ or by choosing two points (z_0,z_1) consistent with the continuous solution, the exact discrete Lagrangian, which is univocally defined in this case, gives rise to the continuous flow. In particular, the constraints $\phi(z_k,p_k)$ are conserved for every step.

- If two consistent points are chosen, an approximation of the exact discrete Lagrangian, such as $\mathcal{L}_d^\alpha$, gives rise to a flow $O(h^r)$ with respect to the continuous flow.
  If the initial points are selected with an $O(h^N)$ integrator, the error is dominated by the smaller order.

- If two generic points are chosen, the continuous and the discrete Legendre transforms are not related.

This means that if the initialization is (almost) consistent, the flow of the integrator is, at least locally, close to the continuous one.
It is worth to stress that, at this point, the behaviour of the relevant quantities such as energy error are not guaranteed to behave well globally in time.

These facts have an important consequence:

- When the time step is set to zero, the integrator **is not**, in general, **the identity transformation**.
  This can be trivially seen in the Lagrangian form just by setting as initial condition $z_0 \neq z_1$.
  In the Hamiltonian form, the integrator is the identity only when the constraints are imposed ($\phi = 0$).

5.2.3 Flow Splitting

Just as we did in chapter 4, we can try to expand the Hamiltonian flow in a serie of powers of h.
From the last section, our integrator is not in general a near the identity map, unless the constraints are imposed.
Therefore, we can write the Hamiltonian flow as:

$$z_1 = z_0 + g_z(\phi(z_0, p_0)) + O(h) \tag{5.12a}$$
$$p_1 = p_0 + g_p(\phi(z_0, p_0)) + O(h) \tag{5.12b}$$

where the condition on the constraints implies $g(0) = 0$.
This behaviour is quite problematic since an integrator should be consistent with an Hamiltonian system, and hence it should be a near the identity map, in order to be interpolated by a modified Hamiltonian system.
One possible solution arises if the constraints are conserved: in this case the function $g(z, p)$ is always 0 for every time step and it is possible to study the equations of motion in the constrained submanifold (i.e. M).
As we will see shortly, this happens only in few cases, namely when the symplectic form is constant.

Let's start by writing the discrete Euler-Lagrange equations in position-momentum form for the midpoint rule, as we did in eqs.(4.57) with the discrete Legendre transforms:

$$p_0 = -\frac{h}{2}\frac{\partial \mathcal{L}}{\partial z}\left(\tilde{z_0}, \frac{\Delta z_0}{h}\right) + \frac{\partial \mathcal{L}}{\partial \dot{z}}\left(\tilde{z_0}, \frac{\Delta z_0}{h}\right) \tag{5.13a}$$

$$p_1 = \frac{h}{2}\frac{\partial \mathcal{L}}{\partial z}\left(\tilde{z_0}, \frac{\Delta z_0}{h}\right) + \frac{\partial \mathcal{L}}{\partial \dot{z}}\left(\tilde{z_0}, \frac{\Delta z_0}{h}\right) \tag{5.13b}$$

Bringing the time step to 0 and writing explicitly all terms of the phase-space Lagrangian (5.1), equation (5.13) reads:

$$p_0 = \Theta(\tilde{z}_0) - \Theta^T_{(1)}(\tilde{z}_0)\frac{\Delta z_0}{2} \tag{5.14a}$$

$$p_1 = \Theta(\tilde{z}_0) + \Theta^T_{(1)}(\tilde{z}_0)\frac{\Delta z_0}{2} \tag{5.14b}$$

where Θ is the Hamiltonian one-form (eq. 5.1) and $\Theta_{(k)}$ is its k-derivative. Since z_0 and p_0 are variables and can be chosen arbitrarily, (z_1, p_1) is in general different from (z_0, p_0): it is straightforward to show that eq. (5.14) is the identity transformation only if the constraints (eq. 5.2) at time step 0 are imposed:

$$p_0 = \Theta(z_0) \tag{5.15}$$

There is, however, a different phenomena which we call **flow splitting**, that occurs for every system discretized with the midpoint rule, regardless of which symplectic form is chosen.

Writing equations (5.13) for two consecutive time steps, and substituting the explicit expression for a general phase-space Lagrangian, we find:

$$p_1 = \Theta(\tilde{z}_1) + \frac{h}{2}\nabla H(\tilde{z}_1) - \Theta^T_{(1)}(\tilde{z}_1)\frac{\Delta z_1}{2} \tag{5.16a}$$

$$= \Theta(\tilde{z}_0) - \frac{h}{2}\nabla H(\tilde{z}_0) + \Theta^T_{(1)}(\tilde{z}_0)\frac{\Delta z_0}{2} \tag{5.16b}$$

Performing a Taylor expansion around the point z_1, we find the following expressions:

$$p_1 = \Theta_1 + \frac{h}{2}\nabla H_1 + \sum_{k=1}^{\infty}\left[\frac{\Theta_{1(k)}}{k!} - \frac{\Theta^T_{1(k)}}{(k-1)!} + \frac{h\nabla H_{1(k)}}{2k!}\right]\overbrace{\left(\frac{\Delta z_1}{2},\ldots,\frac{\Delta z_1}{2}\right)}^{\text{k-times}} \tag{5.17a}$$

$$= \Theta_1 - \frac{h}{2}\nabla H_1 + \sum_{k=1}^{\infty}\left[\frac{\Theta_{1(k)}}{k!} - \frac{\Theta^T_{1(k)}}{(k-1)!} + \frac{(-1)^k h\nabla H_{1(k)}}{2k!}\right]\left(\frac{-\Delta z_0}{2},\ldots,\frac{-\Delta z_0}{2}\right) \tag{5.17b}$$

At first sight, this may seem complicated. However, we can notice that the two expressions are sum of terms with equal or opposite sign, so that we can write:

$$p_1 - \Theta_1 = f(z_1, \Delta z_1, h) = f(z_1, -\Delta z_0, -h) \tag{5.18}$$

Hence, at $h = 0$, the only way to satisfy the latter equation is to make the other arguments equal:

$$\Delta z_0 = -\Delta z_1 \tag{5.19a}$$

As a consequence, we get:

$$z_2 - z_0 = \Delta z_1 + \Delta z_0 = 0 \tag{5.20a}$$

$$\tilde{z_0} = \tilde{z_1} \tag{5.20b}$$

$$p_2 - p_0 = \Delta p_0 + \Delta p_1 = \Theta^T_{(1)}(\tilde{z_1})\Delta z_1 + \Theta^T_{(1)}(\tilde{z_0})\Delta z_0 = 0 \tag{5.20c}$$

Hence, the Hamiltonian flow is splitted in two parts, both of which have a continuous limit regardless of which initial points are chosen:

$$z_2 = z_0 + O(h) \tag{5.21a}$$

$$p_2 = p_0 + O(h) \tag{5.21b}$$

This may indicate that in a general case, in particular when the initial constraints are not conserved, one should study the two flows separately and find an interpolating Hamiltonian. A basic simulation showing the splitting at $h = 0$ is reported in figure 22.

5.3 Constant symplectic form

Before discussing the variational integrator for the guiding center theory, we start by studying a simpler case, that of an Hamiltonian system with a constant symplectic form, or equivalently, with a linear one-form.

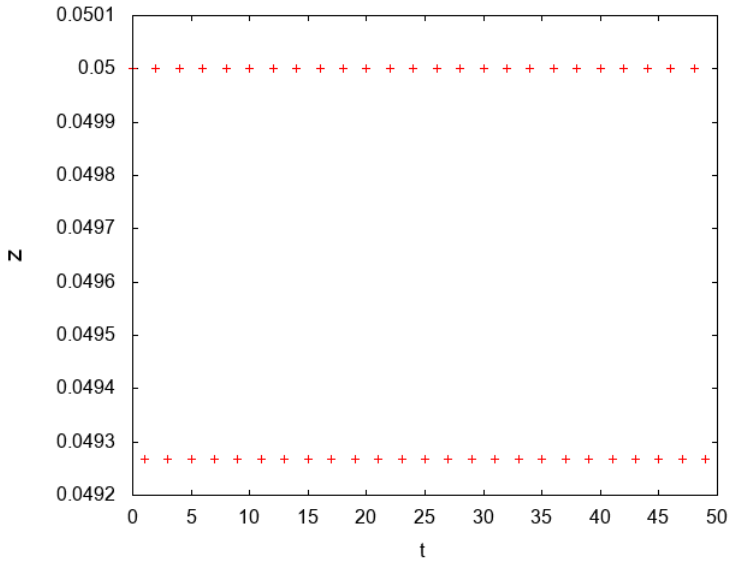


Figure 22: Flow splitting for a general initialization at h=0

The reason of this choice is not only due to simplicity: as we will see, in this case the symplectic integrator is guaranteed to behave correctly in every situation. Furthermore, we can choose examples in which the discrete flow is explicitly invertible and this will simplify the whole analysis.

Returning to the expansion made in eqs (5.17), we can notice that the condition on the constancy of the symplectic form implies that only the first two terms of $\Theta(z)$ are not null, so that we can write:

$$p_1 - \Theta_1 = \frac{h}{2}\nabla H(\tilde{z}_1) + \Omega\left(\frac{\Delta z_1}{2}\right) \tag{5.22a}$$

$$p_1 - \Theta_1 = -\frac{h}{2}\nabla H(\tilde{z}_0) - \Omega\left(\frac{\Delta z_0}{2}\right) \tag{5.22b}$$

Rewriting eq.(5.22a) for the timestep 0, we get:

$$p_0 - \Theta_0 = \frac{h}{2}\nabla H(\tilde{z}_0) + \Omega\left(\frac{\Delta z_0}{2}\right) = -(p_1 - \Theta_1) \tag{5.23}$$

Hence, the quantity $\phi_k = p_k - \Theta_k$ is inverted at each step and it is conserved by both parts of the splitted flow:

$$p_{k+1} - \Theta_{k+1} = -(p_k - \Theta_k) = p_{k-1} - \Theta_{k-1} \tag{5.24a}$$
$$\phi_{k+1} = \phi_{k-1} \tag{5.24b}$$

We will see that this invariance of ϕ is caused by a translational Noether symmetry along the z axes.
It was proved by Jalnapurkar [12] that the projection of the flow of a variational integrator induced by a conserved Noether current conserves the projection of the canonical symplectic form and can be written as a solution of a set of reduced discrete Euler-Lagrange equations.
If we initialize the integrator imposing the constraints ($\phi(z_0, p_0) = 0$) then $\phi_k = 0$ for every step. In this case, the integrator is not splitted and we can study both the whole flow or each part separately.
Inverting the discrete Legendre transforms (eqs.5.22) for $\phi = 0$, we get:

$$\Delta z_0 = -2\Omega^{-1} \left(\sum_{k=0}^{\infty} \frac{h}{2} \frac{\nabla H_{0(k)}}{k!} (\Delta z_0/2, \ldots, \Delta z_0/2) \right) \tag{5.25a}$$
$$= -h\Omega^{-1} (\nabla H_0) + o(h^2) \tag{5.25b}$$

Hence, the projection of the integrator is symplectic and it is consistent with the original continuous problem:

$$\dot{z} = -\Omega^{-1} \nabla H \tag{5.26}$$

When a generic initialization is performed, the flow is splitted and it is possible to define a good projection for each part.
In this case, however, additional terms dependent on the value of ϕ arise for each term in the time step expansion of the flow.
It is possible that these terms have a small influence on the interpolating Hamiltonian, at least when the value of $\phi(z, p)$ is small.

5.3.1 Canonical case

If we restrict to the canonical case, we find an additional property: the projection of a midpoint phase-space integrator, built by imposing the constraints as initialization, is just a midpoint rule applied to the standard Lagrangian.

Let's start by rewriting the phase-space Lagrangian of a canonical system with Lagrangian L and Hamiltonian H:

$$\begin{aligned} \mathcal{L}(z, \dot{z}) = \mathcal{L}(x, p, \dot{x}, \dot{p}) &= \dot{x}p - H(x, p) \\ &= \dot{x}p - \dot{x}(x, p)p + L(x, \dot{x}(x, p)) \end{aligned} \tag{5.27}$$

In this case the Hamiltonian one-form is just:

$$\Theta(x,p) = \begin{pmatrix} p \\ 0 \end{pmatrix} \tag{5.28}$$

Following eq.(5.2), the continuous constraints are:

$$\phi(x,p,p^x,p^p) = \begin{pmatrix} p^x \\ p^p \end{pmatrix} - \Theta(x,p) = \begin{pmatrix} p^x - p \\ p^p \end{pmatrix} \tag{5.29a}$$

where $p^z = (p^x, p^p)$ are the momenta conjugated to x and p:

$$\begin{pmatrix} p^x \\ p^p \end{pmatrix} = \begin{pmatrix} \dfrac{\partial \mathcal{L}}{\partial \dot{x}} \\ \dfrac{\partial \mathcal{L}}{\partial \dot{p}} \end{pmatrix} \tag{5.30}$$

Discretization Since we are dealing with only the midpoint discretization, let's simplify the notation by saying that given a function f, $[f]_k^D$ is its discretization with the midpoint rule, i.e.:

$$[f(x,\dot{x})]_k^D = f\left(\tilde{x}_k, \frac{\Delta x_k}{h}\right) = f\left(\frac{x_k + x_{k+1}}{2}, \frac{x_{k+1} - x_k}{h}\right) \tag{5.31}$$

As sketched in figure 23 and 24, the flows of the standard and the phase-space integrators are then found by matching the discrete Legendre transforms (see section 4.2.2):

$$\begin{aligned} p_0^z &= \left[-\frac{h}{2}\frac{\partial \mathcal{L}}{\partial z} + \frac{\partial \mathcal{L}}{\partial \dot{z}}\right]_D^0 \\ p_1^z &= \left[\frac{h}{2}\frac{\partial \mathcal{L}}{\partial z} + \frac{\partial \mathcal{L}}{\partial \dot{z}}\right]_D^0 \end{aligned} \tag{5.32}$$

for the phase-space integrator, and

$$\begin{aligned} p_0 &= \left[-\frac{h}{2}\frac{\partial L}{\partial x} + \frac{\partial L}{\partial \dot{x}}\right]_D^0 \\ p_1 &= \left[\frac{h}{2}\frac{\partial L}{\partial x} + \frac{\partial L}{\partial \dot{x}}\right]_D^0 \end{aligned} \tag{5.33}$$

for the standard integrator. Also, since L is a regular Lagrangian, the midpoint rule applied to L is regulated by the equations (4.58) and (4.59). In particular:

$$[p]_D = \left[\frac{\partial L}{\partial \dot{x}}\right]_D \tag{5.34a}$$

$$[\dot{x}]_D = \left[\frac{\partial H}{\partial p}\right]_D = [\dot{x}(x,p)]_D \tag{5.34b}$$

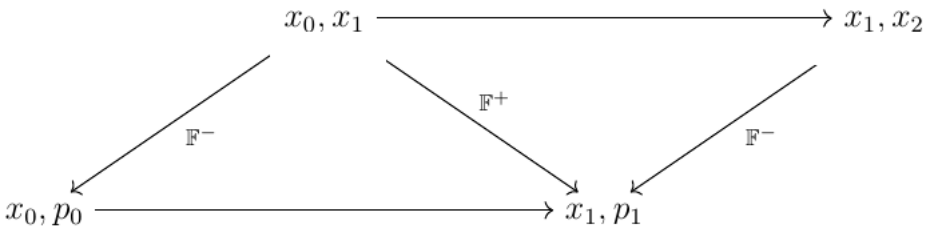


Figure 23: Legendre transforms and flows of the standard integrator

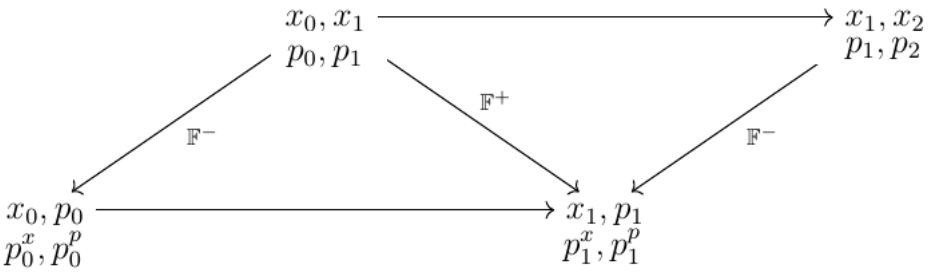


Figure 24: Legendre transforms and flows of the phase-space integrator

Now, let's fix an initial point (x_0, p_0) and let (x_1, p_1) be the following step found with the standard midpoint rule.
Computing the momenta p_0^x and p_0^p with the left Legendre transform of the phase-space integrator (eq. 5.32) and using eqs. (5.27) and (5.34), we find:

$$p_0^x = \left[-\frac{h}{2}\frac{\partial \mathcal{L}}{\partial x} + \frac{\partial \mathcal{L}}{\partial \dot{x}} \right]_D \tag{5.35a}$$

$$= \left[p - \frac{h}{2}\frac{\partial L}{\partial x} - \frac{h}{2}\frac{\partial \dot{x}}{\partial x}\left(\frac{\partial L}{\partial \dot{x}} - p\right) \right]_D \tag{5.35b}$$

$$= \left[\frac{\partial L}{\partial \dot{x}} - \frac{h}{2}\frac{\partial L}{\partial x} \right]_D = p_0 \tag{5.35c}$$

$$p_0^p = \left[-\frac{h}{2}\frac{\partial \mathcal{L}}{\partial p} \right]_D \tag{5.36a}$$

$$= \left[-\frac{h}{2}\left(\dot{x} - \dot{x}(x,p) + \frac{\partial \dot{x}}{\partial p}\left(\frac{\partial L}{\partial \dot{x}} - p\right)\right) \right]_D = 0 \tag{5.36b}$$

The constraints (5.29) are therefore respected.
The contrary is obviously true: if the initial points of the phase-space integrator are chosen by imposing the constraints $\phi = 0$, the points (x_1, p_1) are just the ones found by evolving the standard integrator.
Studying the right Legendre transform, we find the identical result, so that the projection of the phase-space integrator is just the standard one.

5.3.2 Example: One-dimensional spring

Let's consider a one dimensional problem with an elastic potential. The standard and the phase-space Lagrangians are:

$$L(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \tag{5.37a}$$

$$\mathcal{L}(x, p, \dot{x}, \dot{p}) = \dot{x}p - \frac{p^2}{2m} - \frac{1}{2}kx^2 \tag{5.37b}$$

The discretizion of the phase-space Lagrangian with the midpoint rule gives:

$$\mathcal{L}_d(x_0, p_0, x_1, p_1) = \frac{4m(p_0+p_1)(x_1-x_0) - h\left(km(x_0+x_1)^2 + (p_0+p_1)^2\right)}{8m} \tag{5.38}$$

Hamitonian Flow The discrete Euler-Lagrange equations can be solved explicitly. The Hamiltonian flow, written in powers of h is:

$$x_1 = -2p_0^p + x_0 + \frac{p_0^x}{m}h + o(h^2) \tag{5.39a}$$
$$p_1 = -p_0 + 2p_0^x + (kp_0^p - kx_0)h + o(h^2) \tag{5.39b}$$
$$p_1^x = p_0^x + (kp_0^p - kx_0)h + o(h^2) \tag{5.39c}$$
$$p_1^p = -p_0^p \tag{5.39d}$$

If we impose the constraints at the step 0, we immediately recognize that the integrator has a continuous limit, it is consistent with the continuous Hamilton equations and the constraints are conserved ($p_1^p = 0, p_1^x = p_1$):

$$x_1 = x_0 + \frac{p_0}{m}h + o(h^2) \tag{5.40a}$$
$$p_1 = p_0 - kx_0h + o(h^2) \tag{5.40b}$$
$$p_1^x = p_0 - kx_0h + o(h^2) \tag{5.40c}$$
$$p_1^p = 0 \tag{5.40d}$$

Splitted Flow Lagrangian One interesting aspect arises if we try to write the Lagrangian of a single part of the splitted flow, which is the composition of two consecutive steps.
One can easily show that this is just the sum of two consecutive discrete Lagrangians with the middle point (z_1) explicited with respect to the other two points:

$$L_d^s(z_0, z_2) = L_d(z_0, z_1(z_0, z_2)) + L_d(z_1(z_0, z_2), z_2) \tag{5.41}$$

For this problem we find the following Lagrangian:

$$\begin{aligned}\mathcal{L}_d(x_0, p_0, x_2, p_2) = -\frac{1}{16hkm}\big[&h^2k\left((p_0 - p_2)^2 + km(x_0 - x_2)^2\right) - \\ &- 4m\left((p_0 - p_2)^2 + km(x_0 - x_2)^2\right) + 8hkm(p_0 - p_2)(x_0 + x_2)\big]\end{aligned} \tag{5.42}$$

One can again show, by computing the Legendre transforms, that the constraints ϕ are conserved.
Looking at the Lagrangian we have just found, we can see that it is invariant under translations along the p axis in $M \times M$. This is the reason of the conservation of p_p. The conservation of $p_x - p$ follows from the fact that translations along the x axis leave the Lagrangian invariant up to terms proportional to $(p_2 - p_0)$ which cancel out in the discrete action sum (eq. 4.6).
Equivalently, the discrete Euler-lagrange equations for these terms are always null.

More generally, it can be easily proved that two Lagrangians that differ by a total

time derivative of linear or quadratic forms give rise to discrete Lagrangians with the same discrete Euler lagrange equations:

$$\begin{aligned} \tilde{L}(q,\dot{q}) &= L(q,\dot{q}) + \frac{d}{dt}f(q) \\ D_{DEL}\tilde{L}_d(q_0,q_1) &= D_{DEL}L_d(q_0,q_1) \end{aligned} \tag{5.43}$$

where $f(q)$ is a linear or a quadratic form:

$$f(q) = \sum_{i=1}^{n} a_i q_i + \sum_{i,j=1}^{n} b_{ij} q_i q_j \tag{5.44}$$

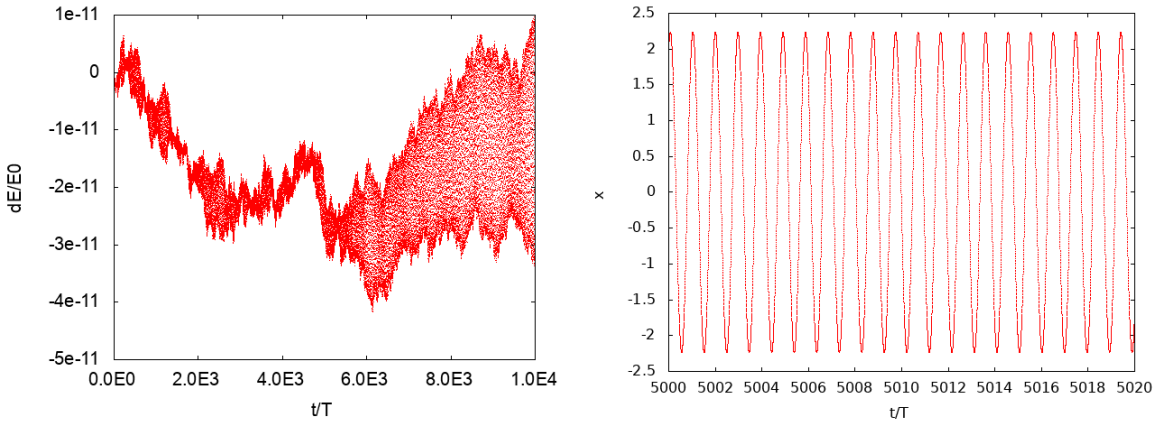
Of course, this implies that equivalent continuous Lagrangians can give rise to different discrete integrators.

Numerical Results For our test, we chose $k = m = 1$ and $h = 0.1$ which is about $h \simeq \frac{T}{65}$, where T is the period of the orbit.
A first simulation was performed imposing the constraints as initial conditions. The results as sketched in fig. 25.
The error on the energy was checked by comparing the continuous Hamiltonian at a step to its initial value at step 0:

$$\frac{dE_k}{E_0} = \frac{H(z_k) - H(z_0)}{H(z_0)} \tag{5.45}$$

The small oscillations seen in the energy error and in the conservation of constraints are due to numerical errors, so they're all exactly conserved.
Of course, while the conservation of constraints are just what we expected, the fact that the energy is exactly conserved is just a fortunate coincidence of this problem and won't happen in general.

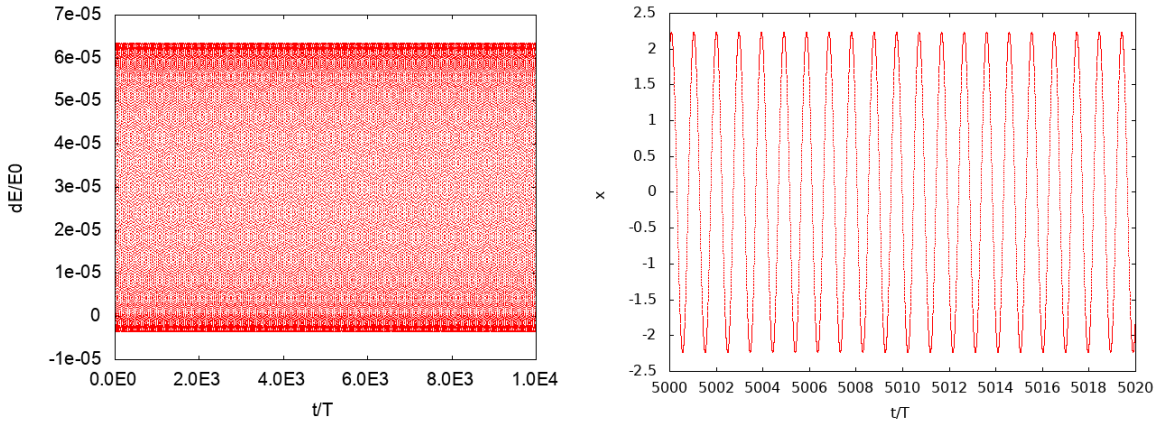
A different simulation is sketched in fig. 26. This time, we chose the points z_0 and z_1 to be consistent with the continuous solution. We call this a "Lagrangian" initialization.
In practice, z_1 was computed using a different, non symplectic integrator, with a small time step so that z_1 is almost exactly consistent with the solution.
We can notice that the energy is not exactly conserved as before, but the long time behaviour is still bounded.
As we expected from the theory, the integrator is splitted in two parts, each of which conserves the quantity ϕ.



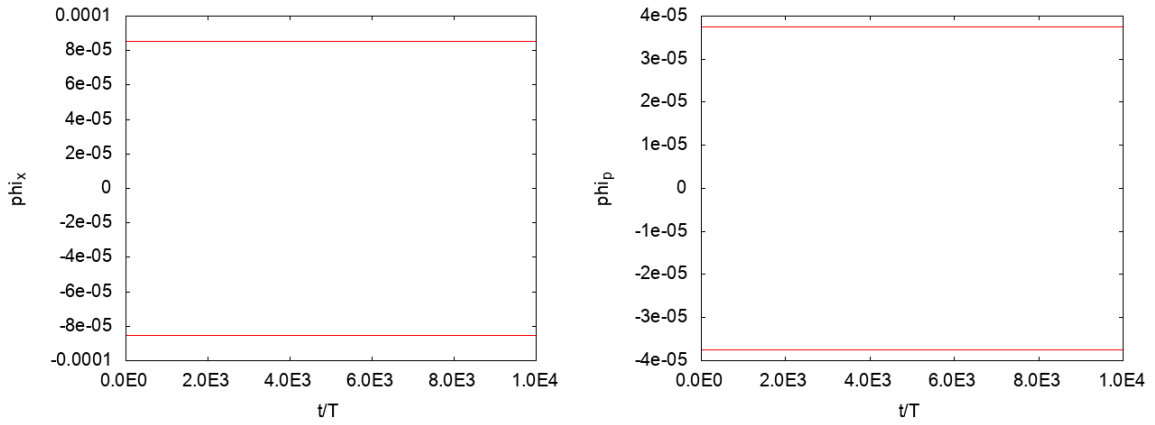
(a) (b)

(c) (d)

Figure 25: 1D spring integrator with constrained initialization. (a) is the energy error, (b) is the evolution of x. The conservation of constraints are plotted in (c) and (d)



(a) (b)



(c) (d)

Figure 26: 1D spring integrator with Lagrangian initialization.

6 Guiding Center Integrators: Numerical Results

6.1 Introduction

In the last chapter, we studied the behaviour of a variational symplectic integrator when applied to a phase-space Lagrangian.
In particular, we proved that its flow is in general splitted in two parts each of which admits a continuous limit.
Particular attention must be paid to the conservation of the continuous constraints. In this regard, we saw that for a Hamiltonian system with a constant symplectic form, such as canonical systems, the constraints are conserved and a good projection can be defined, such that it is symplectic and consistent with the original continuous problem.
The same simple arguments do not hold for a non constant symplectic form: the Taylor expansion made in eqs.(5.17) doesn't stop at the first terms and additional terms with the same sign arise. Hence, equation (5.22) does not hold and the constraints $\phi_k = \Theta_k - p_k$ are not conserved and it is not possible to define a trivial projection to the submanifold M.
However, we will see that the integrator does still possess good long-term conservation properties. The reason of this behaviour is still unknown.
It is possible that non trivial quantities dependent on the time step h are conserved by each part of the splitted flow.

In this chapter, we will study the variational symplectic integrator applied to the non canonical guiding center theory.
We will follow the guidelines given by Qin and coworkers in a series of paper ([20] and [21]), where different versions of the integrator were studied, in particular an explicit linearization of the flow.
We want to anticipate here that all the integrators are very sensible to the initial conditions, and it is not clear at the moment which is the best choice of initialization, given a particular choice of the algorithm.
The complete understanding of the theory underlying these types of integrators for non constant symplectic forms is of great importance, since it can lead to better choice of initial conditions and possibly to new versions of the algorithm.

Let's start by recalling some basic results of the non canonical guiding center theory of chapter 3.
The guiding center Hamiltonian one-form and symplectic form are given by:

$$\Theta = \begin{pmatrix} \mathbf{A}^\dagger \\ 0 \end{pmatrix}$$

$$\Omega = \begin{pmatrix} 0 & -B_z^\dagger & B_y^\dagger & \hat{b}_x \\ B_z^\dagger & 0 & -B_x^\dagger & \hat{b}_y \\ -B_y^\dagger & B_x^\dagger & 0 & \hat{b}_z \\ -\hat{b}_x & -\hat{b}_y & -\hat{b}_z & 0 \end{pmatrix}$$

where $A^\dagger$ and $B^\dagger$ are respectively a modified vector potential and a modified magnetic field:

$$\mathbf{A}^\dagger(\mathbf{X}, t) = \mathbf{A}(\mathbf{X}, t) + u\hat{\mathbf{b}}(\mathbf{X}, t) \tag{6.1a}$$

$$\mathbf{B}^\dagger = \nabla \times \mathbf{A}^\dagger = \mathbf{B} + u\nabla \times \hat{\mathbf{b}} \tag{6.1b}$$

The resulting phase-space Lagrangian is given by:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, u, \dot{\mathbf{x}}, \dot{u}) &= \Theta(\mathbf{x}, u) \cdot (\dot{\mathbf{x}}, \dot{u}) - H(\mathbf{x}, u) \\ &= A^\dagger \cdot \dot{\mathbf{x}} - H(\mathbf{x}, u) \end{aligned} \tag{6.2}$$

where $H(\mathbf{x}, u)$ is the non canonical Hamiltonian:

$$H(\mathbf{x}, u) = \frac{u^2}{2} + \mu B \tag{6.3}$$

The equations of motion read:

$$\begin{pmatrix} \dot{\mathbf{x}} \\ \dot{u} \end{pmatrix} = -\Omega(\mathbf{x}, u)^{-1} \nabla H(\mathbf{x}, u) \tag{6.4}$$

or equivalently:

$$\dot{\mathbf{X}} = \frac{u\mathbf{B}^\dagger - \hat{\mathbf{b}} \times (\mathbf{E} - \mu\nabla B)}{\hat{\mathbf{b}} \cdot \mathbf{B}^\dagger} \tag{6.5a}$$

$$\dot{u} = \frac{\mathbf{B}^\dagger \cdot (\mathbf{E} - \mu\nabla B)}{\hat{\mathbf{b}} \cdot \mathbf{B}^\dagger} \tag{6.5b}$$

With equations (6.5) we can easily build a standard, i.e. non symplectic, integrator, such as Euler or Runge-Kutta.
We can use these integrators as a direct comparison with the symplectic ones. Also, since they share the same degrees of freedom of the continuous system, they can be used to build an initialization starting from a point z_0.
Finally, they are useful as a first guess for the solution of the implicit equations of the variational integrator.

6.2 A few reference magnetic field configurations

In this thesis, the variational integrators were tested with three different magnetic field configurations: a uniform magnetic field, a tokamak configuration and a force free magnetic field.

6.2.1 Reference Case A: Two Dimensional Uniform Magnetic Field

As a first example, we can choose a magnetic field directed along the z axis and dependent weakly on the other two variables:

$$\mathbf{A}(\mathbf{x}) = -\frac{0.05}{3}y^3\hat{\mathbf{e}}_x + \left(\frac{0.05}{12}x^3 + x\right)\hat{\mathbf{e}}_y \tag{6.6a}$$

$$\mathbf{B}(\mathbf{x}) = 1 + 0.05\left(\frac{x^2}{4} + y^2\right)\hat{\mathbf{e}}_z \tag{6.6b}$$

$$\tag{6.6c}$$

Since there is no dependence on the z axis, the third component of the modified vector field (see eq.6.2) is a conserved quantity:

$$\frac{\partial \mathcal{L}}{\partial \dot{z}} = A_z^\dagger = A_z + u\hat{b}_z = u = const. \tag{6.7}$$

The variable z appears in the Lagrangian only with a term $k\dot{z}$ and it can be safely dropped and we get the following reduced Lagrangian:

$$\mathcal{L}(x, y, \dot{x}, \dot{y}) = \begin{pmatrix} A_x \\ A_y \end{pmatrix} \cdot \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} - \mu B(x, y) \tag{6.8}$$

Hence, the magnitude of the magnetic field along an orbit is conserved and the dynamics in the x-y plane is a closed orbit with equation:

$$\frac{x^2}{4} + y^2 = const \tag{6.9}$$

From the theory presented in chapter 4, we expect that a symplectic integrator should conserve exactly the quantities associated with Noether symmetries, in this case the velocity u, while the magnitude of the magnetic field, which represents the energy, should possess good long time bounding properties.

Parameters We chose the time step such that the integrator completes one orbit of period T in about 50 steps: $h \simeq \frac{T}{50}$, which turns out to be $h \simeq 1 \times 10^6$.
The initial position of the particle is chosen to be $\mathbf{x}_0 = (0.05, 0, 0)$ with an initial velocity $u_0 = 3.9 \times 10^{-4}$.

6.2.2 Reference Case B: Tokamak magnetic field

As a more complicated example, we can choose a typical magnetic field configuration used in tokamak fusion reactor. Denoting by z the toroidal coordinate and by x, y the coordinates of the poloidal plane, the vector potential and the magnetic field are:

$$A = -B_0 R_0 ln\left(\frac{R_0+x}{R_0}\right)\hat{\mathbf{y}} + \frac{B_0}{2q(R_0+x)}\left[2R_0(R_0+x)ln\left(\frac{R_0+x}{R_0}\right) - 2R_0x - 2x^2 - y^2\right]\hat{\mathbf{z}} \tag{6.10a}$$

$$B = -\frac{B_0 y}{q(R_0+x)}\hat{\mathbf{x}} + \frac{B_0}{q}\frac{2R_0x + 2x^2 - y^2}{2(R_0+x)^2}\hat{\mathbf{y}} - \frac{B_0R_0}{R_0+x}\hat{\mathbf{z}} \tag{6.10b}$$

In this case, the magnetic field is toroidally symmetric and its field lines are almost circular concentric in the poloidal plane, as illustrated in fig. 27.

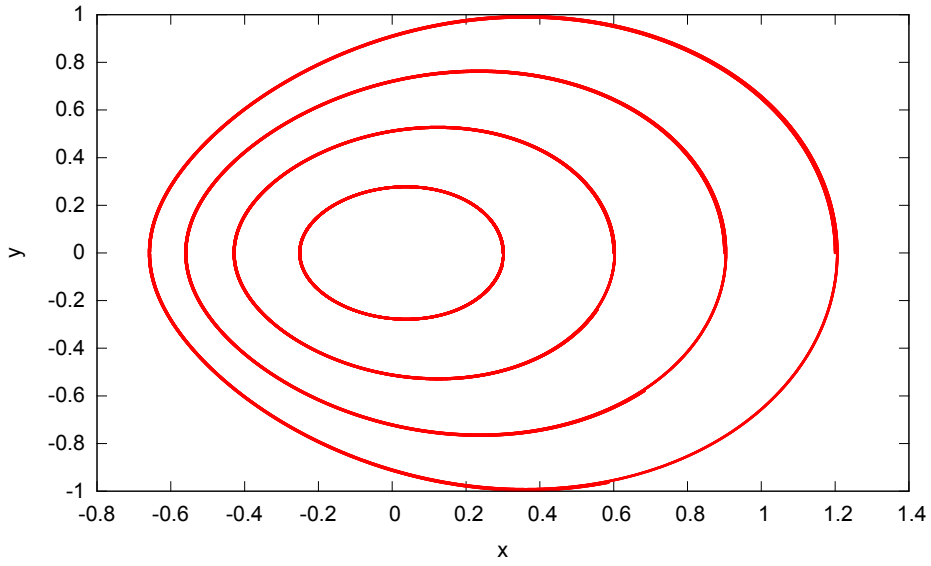


Figure 27: Magnetic field lines for the tokamak B configuration (section 6.2.2)

As before, the fact that the fields are toroidally symmetric implies that the relative conjugate momentum, which is the z-th component of the modified vector potential, is a conserved quantity:

$$A_z^\dagger(\mathbf{x}, u) = A_z(\mathbf{x}) + u\hat{b}(\mathbf{x}) \tag{6.11}$$

The dynamics in the poloidal plane is a closed orbit which can be elliptic or a banana orbit if the energy of the particle is sufficiently low so that the parallel gradient drift is enough to bounce back the particle (see section 2.4).

Parameters The magnetic field and the minor radius B_0, R_0 are chosen to be normalized to unity with a safety factor $q = 2$ and a magnetic moment of $\mu = 2.25 \times 10^{-6}$.
Using the same initial conditions of the configuration A, the particle performs banana orbits in the poloidal plane.
The time step used is $h = 800$ which corresponds to $h \simeq \frac{T}{50}$. Summing up:

h	800
R_0	1
B_0	1
q	2
μ	2.25×10^{-6}
$\mathbf{x}_0$	$(0.05, 0, 0)$
u_0	3.9×10^{-4}

6.2.3 Reference Case C: force-free field

The equilibrium state of a plasma is described by the equilibrium of the magnetic and pressure forces:

$$\mathbf{J} \times \mathbf{B} = \nabla p \tag{6.12}$$

Also, the condition of null divergence of the magnetic field and the Ampere's law read:

$$0 = \nabla \cdot \mathbf{B} \tag{6.13a}$$
$$\mathbf{J} = \nabla \times \mathbf{B} \tag{6.13b}$$

When the kinetic pressure of a plasma is negligible, its magnetic field assumes the following form at equilibrium:

$$0 = \mathbf{J} \times \mathbf{B} = (\nabla \times \mathbf{B}) \times \mathbf{B} \tag{6.14}$$

hence, the curl of the magnetic field must be parallel to the field:

$$\nabla \times \mathbf{B} = a\mathbf{B} \tag{6.15}$$

where a is a parameter that can depend in principle on the space. These types of magnetic field are called force-free and they are first-order approximations of the ones found in astrophysics or in the RFP fusion reactor, where the condition on the kinetc pressure is almost satisfied.
In our case we assume the parameter a to be constant. The magnetic field in this case is said to be linear force-free.

The solution of equation (6.15) was first given by Namikawa [26] and Chandrasekhar [5].
Using the fact that the magnetic field has null divergence, we can express eq.(6.14) as:

$$\nabla^2 \mathbf{B} + a^2 \mathbf{B} = 0 \tag{6.16}$$

which is the vector Helmholtz equation. It was proved by Chandrasekhar and Namikawa that given a solution ψ of the scalar Helmholtz equation:

$$\nabla^2 \psi + a^2 \psi = 0 \tag{6.17}$$

then a solution in cylindrical coordinates (r, θ, z) of the force-free equation (6.15) can be found with:

$$\mathbf{B} = \frac{1}{\mu} \nabla \times (\nabla \times \hat{\mathbf{z}} \psi) + \nabla \times \hat{\mathbf{z}} \psi \tag{6.18}$$

where μ is the magnetic moment. The corresponding vector field is then:

$$\mathbf{A} = \frac{1}{a} \nabla \times \hat{\mathbf{z}} \psi + \hat{\mathbf{z}} \psi \tag{6.19}$$

Since we are assuming periodicity along the toroidal and polodial axes, we can express a generic solution in terms of the poloidal and toroidal numbers and it turns out that each term is just the Bessel function:

$$\psi(r, \theta, z) = \sum_{m,n} J_m \left(r a_n \right) e^{im\theta + inz} \tag{6.20a}$$

$$a_n = \sqrt{a^2 - n^2} \tag{6.20b}$$

Reference Case C1: Axial Symmetric field Using equation (6.18) we can find the first term $(n = 0, m = 0)$ in expansion (6.20):

$$\mathbf{B}_{00}(r, \theta, z) = B_0 J_1(ar) \hat{\mathbf{e}}_\theta + B_0 J_0(ar) \hat{\mathbf{e}}_z \tag{6.21}$$

the vector potential is:

$$\mathbf{A}_{00}(r, \theta, z) = \frac{B_0}{a} J_1(ar) \hat{\mathbf{e}}_\theta + \frac{B_0}{a} J_0(ar) \hat{\mathbf{e}}_z \tag{6.22}$$

As illustrated in figure 28, the toroidal magnetic field is inverted at a specific radius point, that is $ar \simeq 2.4$. This is a typical behaviour found in the RFP experiments.

Similarly to the tokamak configuration, when the energy of the particle is sufficiently low the dynamics in the poloidal plane is a closed banana orbit.

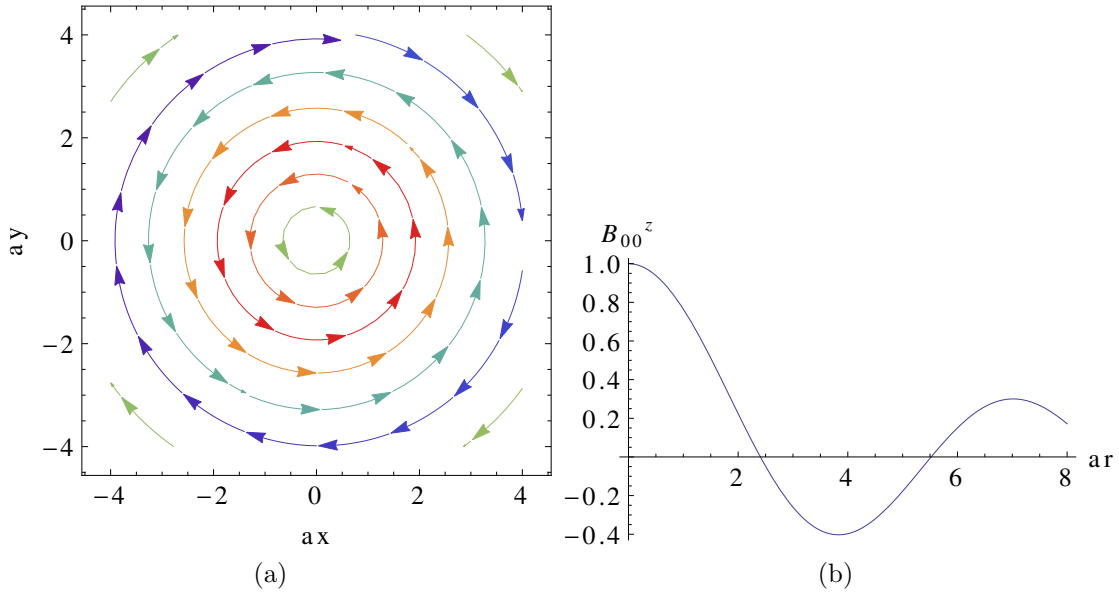


Figure 28: Field lines in the poloidal plane (a) and in the (z, r) plane (b) of the magnetic field configuration C1

Reference Case C2: Asymmetric field We tried a slightly more sophisticated configuration by adding small perturbations with numbers $(n = 1, m = 1)$ and $(n = 1, m = 2)$:

$$\mathbf{B}(r,\theta,z) = \mathbf{B}_{00}(r,\theta,z) + \alpha\mathbf{B}_{11}(r,\theta,z) + \beta\mathbf{B}_{12}(r,\theta,z) \tag{6.23}$$

where $\mathbf{B}_{11}$ and $\mathbf{B}_{12}$ are found to be:

$$\begin{aligned}\mathbf{B}_{11}(r,\theta,z) = &\left[\frac{a_1}{a}J_0(a_1 r) + \frac{a-1}{ar}J_1(a_1 r)sin(\theta+z)\right]\hat{\mathbf{e}}_r \\ &+ \left[-a_{11}J_0(a_1 r) + \frac{a-1}{ar}J_1(a_1 r)cos(\theta+z)\right]\hat{\mathbf{e}}_\theta \\ &+ \left[\frac{a^2-1}{a}J_1(a_1 r)cos(\theta+z)\right]\hat{\mathbf{e}}_z\end{aligned} \tag{6.24}$$

$$\begin{aligned}\mathbf{B}_{12}(r,\theta,z) = &\left[\frac{2a_2}{a}J_0(a_2 r) + \frac{a-2}{ar}J_1(a_2 r)sin(\theta+2z)\right]\hat{\mathbf{e}}_r \\ &+ \left[-a_{12}J_0(a_2 r) + \frac{a-2}{ar}J_1(a_2 r)\right]cos(\theta+2z)\hat{\mathbf{e}}_\theta \\ &+ \left[\frac{a_2}{a}J_1(a_2 r)cos(\theta+2z)\right]\hat{\mathbf{e}}_z\end{aligned} \tag{6.25}$$

Note that the condition on the coefficients (6.20b) requires that the constant a must be greater than 2 for the perturbations to exist.
The effect of the perturbations on the field lines is plotted in fig. (29).
Note that we don't expect anymore closed orbits in the poloidal plane since the magnetic field is not perfectly symmetric along the toroidal axis.

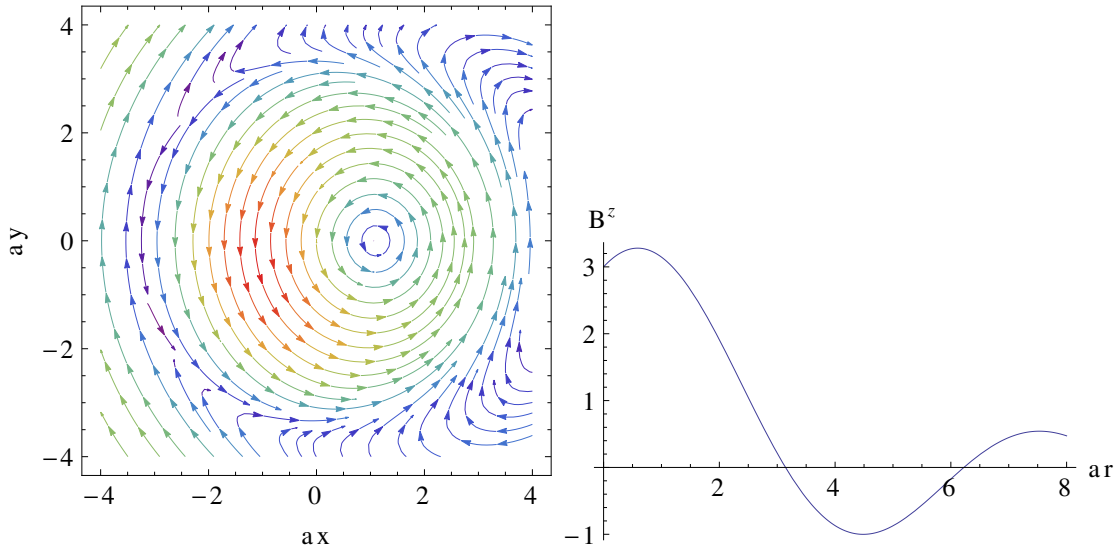


Figure 29: Field lines in the poloidal plane (a) and in the (z, r) plane (b) of the magnetic field configuration C2

Parameters We chose the parameter a to be $a = 3$, which leads to a reasonable RFP configuration. The radius r is assumed to be normalized to the minor radius of the reactor.
The velocity of the particle is selected to produce elliptic orbits for the symmetric version and banana orbits for the perturbed version.
The time step h corresponds to $h \simeq \frac{T}{50}$, where T is the period of the orbit.

h (C1)	240
h (C2)	600
a	3
$\mathbf{x}_0$	$(0.05, 0, 0)$
u_0 (C1)	3.9×10^{-4}
u_0 (C2)	3.9×10^{-5}
α	0.01
β	0.01

6.3 Initialization

The choice of the initial points plays an important role, since different initializations, as we already saw in the last chapter, can lead to very different flows.
In the following, we will use two different initializations.

Denoting by $\mathbf{z} = (\mathbf{x}, u)$ the coordinates of the phase-space, recall from section 4.2.2 that for a variational integrator we can study two different flows: the Lagrangian flow $(\mathbf{z}_0, \mathbf{z}_1) \to (\mathbf{z}_1, \mathbf{z}_2)$ and the hamiltonian flow $(\mathbf{z}_0, \mathbf{p}_0) \to (\mathbf{z}_1, \mathbf{p}_1)$.
Hence, this two flows induce two natural initializations: we call them Hamiltonian and Lagrangian initialization:

- **Hamiltonian Initialization**: given an initial point $\mathbf{z}_0$, $\mathbf{p}_0$ is selected by imposing the constraints:

$$\phi_0 = \mathbf{p}_0 - \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{z}}} = \mathbf{p}_0 - \Theta_0 = \mathbf{p}_0 - \begin{pmatrix} A^\dagger(\mathbf{x}_0) \\ 0 \end{pmatrix} \equiv 0 \tag{6.26}$$

- **Lagrangian Initialization**: given an initial point $\mathbf{z}_0$, $\mathbf{z}_1$ is selected by evolving the continuous system to $\mathbf{z}_1$.
  In practice, an additional integrator, such as RK4, must be used. Of course, the point $\mathbf{z}_1$ can be find with arbitrary precision, for example by choosing small timesteps of the auxiliary integrator.

We recall that the initial points $\mathbf{z}_1$ and $\mathbf{p}_0$ are related by the left Legendre transform (see paragraph. 4.2.2).
It makes sense, for the reasons we saw in the last chapter, that the Hamiltonian initialization might be the most correct one, since it guarantees that the flow is close to the continuous one for at least one time step and for the fact that it produces the correct result for the case of constant symplectic forms.
However, we will see that this is not always the case.

6.4 Numerical Schemes and Results

In the following, the continuous phase-space Lagrangian (eq. 6.2) is discretized with different methods: the midpoint rule, which produces generally the best results, a slightly modified version along with an explicit linearization, both suggested by Qin [20] and a midpoint discretization applied to a reduced Lagrangian.
Finally, new explicit linearizations are proposed and we will show some interesting benefits over the old algorithms.

For now on, we will refer to z^i as the coordinates of the phase-space

$$\mathbf{z} = (\mathbf{x}, u) \tag{6.27}$$

and to $\tilde{z}$ and Δz as the midpoint discretization of $\mathbf{z}$ and $\dot{\mathbf{z}}$:

$$\begin{aligned} \tilde{\mathbf{z}}_k &= \frac{\mathbf{z}_k + \mathbf{z}_{k+1}}{2} \\ \Delta \mathbf{z}_k &= \mathbf{z}_{k+1} - \mathbf{z}_k \end{aligned} \tag{6.28}$$

Also, by $f_{i,j}$ we denote the derivative of the i-th component of the function f with respect to the j-th component:

$$f_{i,j}(\mathbf{z}) = \frac{\partial f_i(z)}{\partial \mathbf{z}_j} \tag{6.29}$$

Energy error Every integrator was tested by checking the conservation of the energy: denoting by H_0 and by H_k the hamiltonian evaluated at the time step 0 and k, the energy error is defined as:

$$\frac{dE}{E_0} = \frac{H_k - H_0}{H_0} \tag{6.30}$$

6.4.1 Implicit Scheme 1: Midpoint Rule

The discretization of the guiding center Lagrangian (6.2) with the midpoint rule (see eq.4.43c) leads to the following expression:

$$\mathcal{L}_d(\mathbf{z}_0, \mathbf{z}_1) = A^\dagger(\tilde{\mathbf{z}}_0) \cdot \Delta \mathbf{x}_0 - h\mu B(\tilde{\mathbf{x}}_0) - h\frac{\tilde{u}_0^2}{2} \tag{6.31}$$

Recall from eq. (4.57) that the left and right discrete Legendre transforms for the midpoint rule are:

$$\begin{aligned}\mathbf{p}_k &= \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{z}}}\left(\tilde{\mathbf{z}_k}, \frac{\Delta \mathbf{z}_k}{2}\right) - \frac{h}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{z}}\left(\tilde{\mathbf{z}_k}, \frac{\Delta \mathbf{z}_k}{2}\right) \\ \mathbf{p}_{k+1} &= \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{z}}}\left(\tilde{\mathbf{z}_k}, \frac{\Delta \mathbf{z}_k}{2}\right) + \frac{h}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{z}}\left(\tilde{\mathbf{z}_k}, \frac{\Delta \mathbf{z}_k}{2}\right)\end{aligned} \tag{6.32}$$

For the guiding center we get the following expressions:

$$\mathbf{p}_0^{\mathbf{x}} = -\frac{1}{2}\left(A^\dagger_{i,j}(\tilde{\mathbf{x}}_0)\right)\cdot(\Delta \mathbf{x}_0) + A^\dagger_j(\tilde{\mathbf{x}}_0) + \frac{h}{2}\mu B_{,j}(\tilde{\mathbf{x}}_0) \tag{6.33a}$$

$$p_0^u = -\frac{1}{2}\hat{b}(\tilde{\mathbf{x}_0})\cdot(\Delta \mathbf{x}_0) + \frac{h}{2}\tilde{u}_0 \tag{6.33b}$$

$$\mathbf{p}_1^{\mathbf{x}} = \frac{1}{2}\left(A^\dagger_{i,j}(\tilde{\mathbf{x}}_0)\right)\cdot(\Delta \mathbf{x}_0) + A^\dagger_j(\tilde{\mathbf{x}}_0) - \frac{h}{2}\mu B_{,j}(\tilde{\mathbf{x}}_0) \tag{6.33c}$$

$$p_1^u = \frac{1}{2}\hat{b}(\tilde{\mathbf{x}_0})\cdot(\Delta \mathbf{x}_0) - \frac{h}{2}\tilde{u}_0 \tag{6.33d}$$

Given an initial points $\mathbf{x}_0, u_0, p_0^{\mathbf{x}}, p_0^p$, the first two equations of (6.33) are inverted to find the points $\mathbf{x}_1, u_1$ and the momenta $p_1^{\mathbf{x}}, p_1^p$ are found with the last two equations. Of course, this is a way to compute the Hamiltonian flow.

It is equivalent to start with two points $\mathbf{x}_0, u_0, \mathbf{x}_1, u_1$ and find $\mathbf{x}_2, u_2$ by matching the momenta of eqs (6.33):

$$\frac{1}{2}\left(A^\dagger_{i,j}(\tilde{\mathbf{x}}_0)\right)\cdot(\Delta \mathbf{x}_0) + \frac{1}{2}\left(A^\dagger_{i,j}(\tilde{\mathbf{x}}_1)\right)\cdot(\Delta \mathbf{x}_1) + A^\dagger_j(\tilde{\mathbf{x}}_0) - A^\dagger_j(\tilde{\mathbf{x}}_1) - \frac{h}{2}\mu\left[B_{,j}(\tilde{\mathbf{x}}_0) + B_{,j}(\tilde{\mathbf{x}}_1)\right] = 0 \tag{6.34a}$$

$$\frac{1}{2}\hat{b}(\tilde{\mathbf{x}_0})\cdot(\Delta \mathbf{x}_0) + \frac{1}{2}\hat{b}(\tilde{\mathbf{x}_1})\cdot(\Delta \mathbf{x}_1) - \frac{h}{2}(u_2 + 2u_1 + u_0) = 0 \tag{6.34b}$$

In either way, the resolution of the flow requires the inversion of an implicit equation.
This is done by using an auxiliary integrator, in our case RK4 applied to the continuous equations of motion (3.35), and by converging to the implicit equations (6.33) with sufficiently high number of Newton iterations. Usually, three or four iterations are enough to guarantee a good accuracy.

We expect for the field configurations A,B and C1, for which there is a translational symmetry along the z axis, that the corresponding discrete momentum is conserved by the integrator:

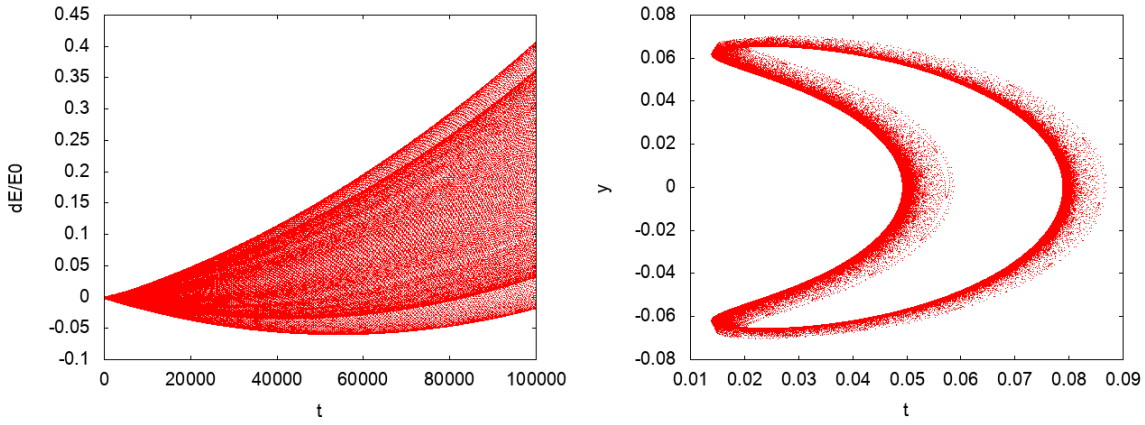
$$p_k^z = \frac{\partial \mathcal{L}}{\partial \dot{z}}\left(\tilde{\mathbf{z}_k}, \frac{\Delta \mathbf{z}_k}{2}\right) = A^\dagger_z(\tilde{\mathbf{x}}_k) = const. \tag{6.35}$$

Also, since the Lagrangian is independent from $\dot{u}$, we expect that the conjugate momentum of u is inverted at each step. In fact, from the definition of the discrete Legendre transforms (eqs. 6.32), we get:

$$p_k^u = -\frac{h}{2}\frac{\partial \mathcal{L}}{\partial u}\left(\tilde{\mathbf{z}_k}, \frac{\Delta \mathbf{z}_k}{2}\right) = -p_{k+1}^u \tag{6.36}$$

Of course, this is just a particular case of the previous chapter, where we considered constant symplectic forms, or equivalently linear Cartan one-forms Θ.
In particular, if we restrict to the Hamiltonian initialization, for which the constraints are imposed at the first step (see paragraph 6.3), we expect that p^u should be exactly conserved.

Initialization In section (6.3), we said that different initializations can lead to different results.
At a first stage, the midpoint discretization was tested with a Lagrangian initialization, so that given a point z_0, the point z_1 was found with a RK4 integrator applied to the equations of motion (6.5) with a small time step in order to guarantee the best accuracy.
Figure 30 illustrates the results for a tokamak configuration B.
The energy error oscillates with increasing amplitude and it becomes very inaccurate even after few steps.
This bad behaviour is similar for the other field configurations and is shared by all the implicit schemes we tested. This agrees with the discussions of the previous chapter, where we proved that, at least with simple cases, the Hamiltonian initialization was the most correct one. Hence, in the following we'll focus only on the Hamiltonian initialization.



(a) (b)

Figure 30: Energy error (a) and particle orbit in the poloidal plane (b) for the midpoint rule with Lagrangian initialization applied to field configuration B.

Numerical Results Figure 31 shows the numerical results for the 2D field configuration A and the tokamak field B.
In both cases the momenta p_u and p_z are exactly conserved as expected. The energy is well bounded for long times, in particular for the configuration A the energy is exactly conserved.

For the tokamak field B we lowered the time step to $h = 10$ and plotted in figure 33 the conservation of the energy and the x component of the constraints:

$$\phi_x = p^x - A^{\dagger x} \tag{6.37}$$

We immediately notice that the constraint is not exactly conserved. As a consequence, the flow is splitted in two parts. The same behaviour happens for all the integrators we used in the rest of the thesis.

Unfortunately, the behaviour of the midpoint rule with the force-free fields C1 is slightly worse.
As shown in figure 34, the flow of the integrator is splitted in two parts with distinct energy.
While the momenta p^z and p^u are exactly conserved as expected, the energy of each part has a noticeable drift.

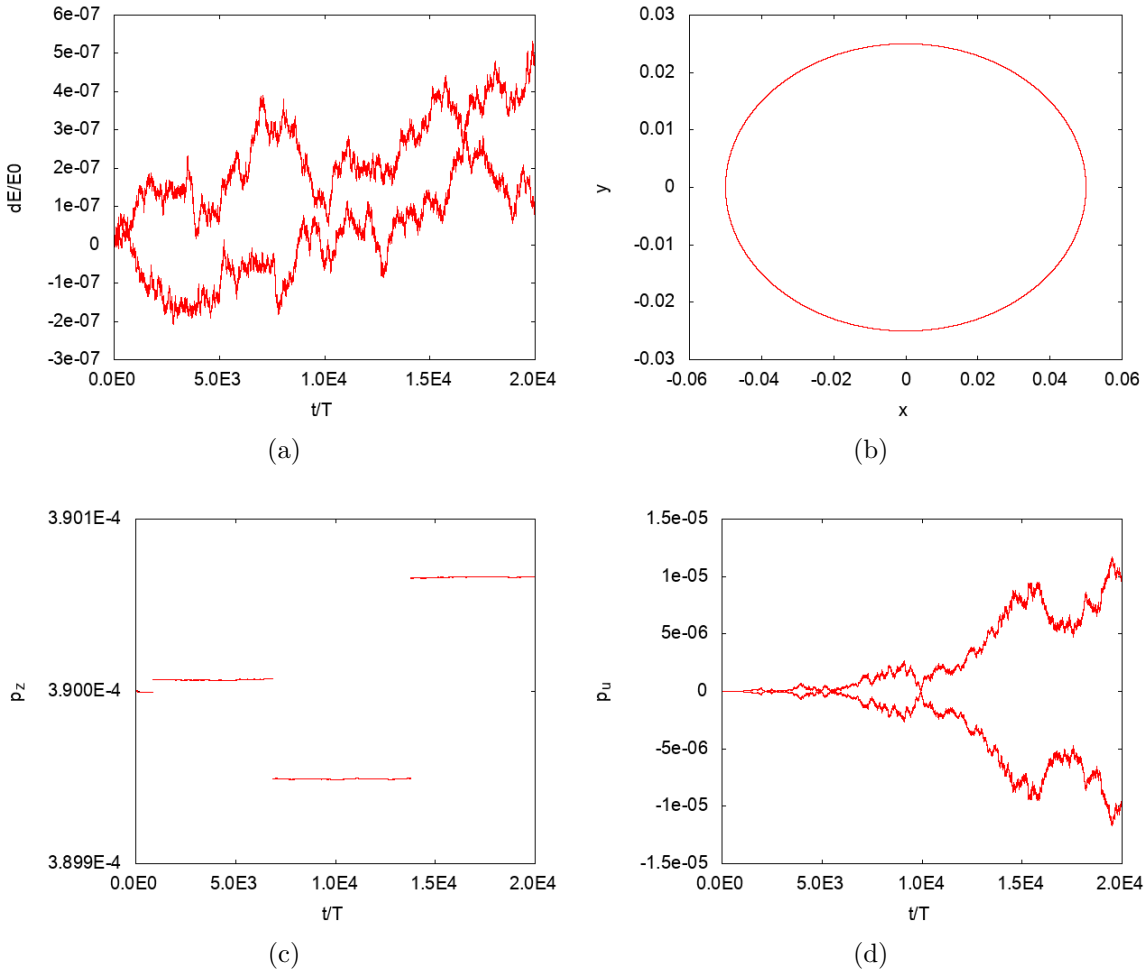
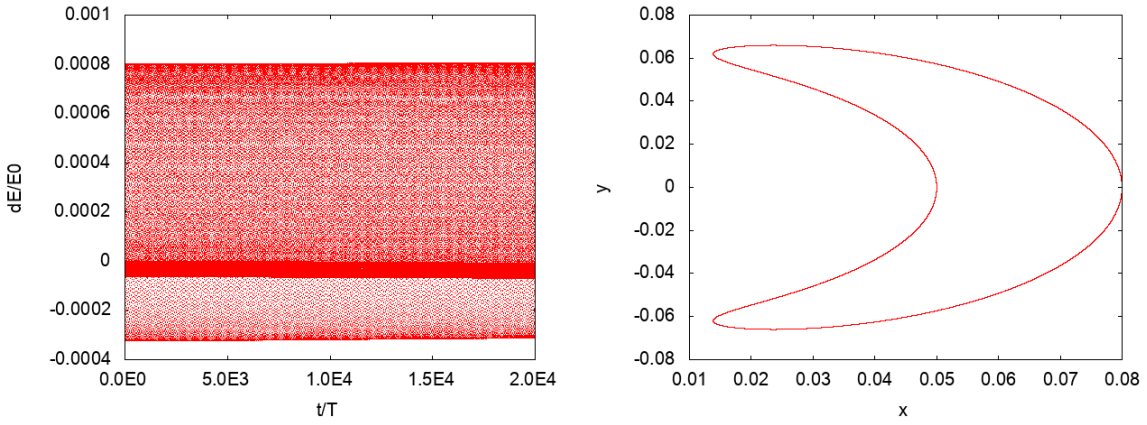


Figure 31: Midpoint rule applied to field configuration A. (a) is the energy error, (b) is the particle trajectory in the (x, y) plane, (c) and (d) are the z and u components of the discrete momentum (eq. 6.35)

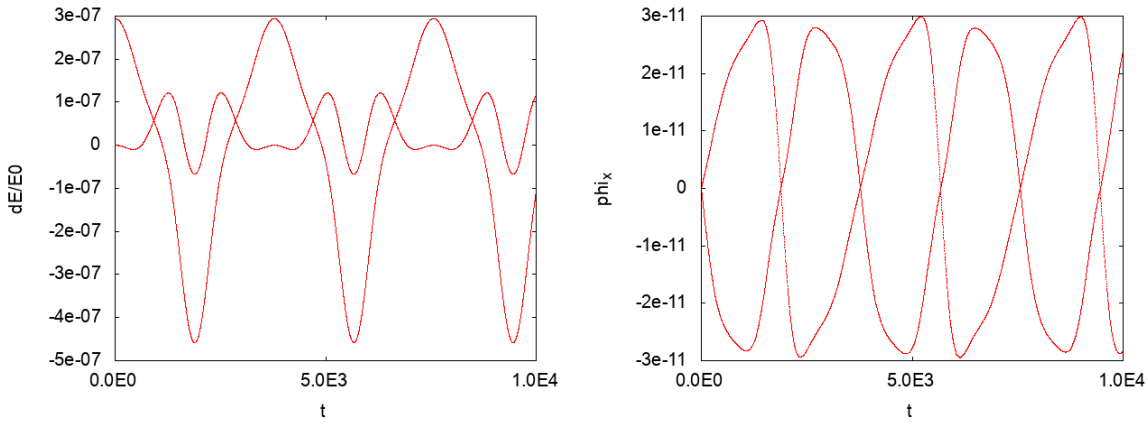
For the configuration C2, the particle performs small banana orbits in the poloidal plane and a bigger elliptic closed orbit. The results are plotted in figure 35. A zoom of the trajectory with a low time step is reported in diagram (c) to highlight the banana orbits.
In this case, the energy is bounded correctly for long times and the momentum p^u is correctly conserved, while the momentum p^z is only bounded. This is expected, since the perturbations (eqs. 6.24) break the toroidal symmetry.



(a) (b)

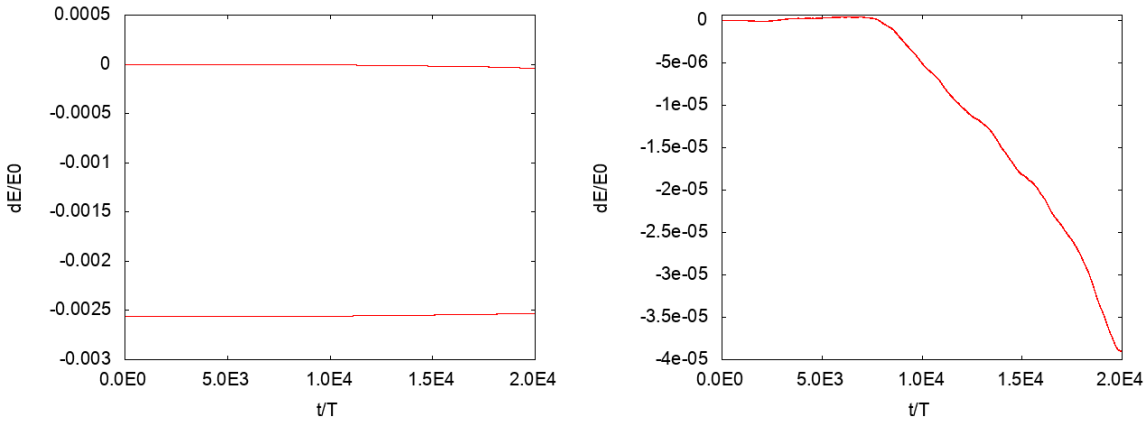
(c) (d)

Figure 32: Midpoint rule applied to field configuration B. (a) is the energy error, (b) is the particle trajectory in the (x, y) plane, (c) and (d) are the z and u components of the discrete momentum (eq. 6.35)



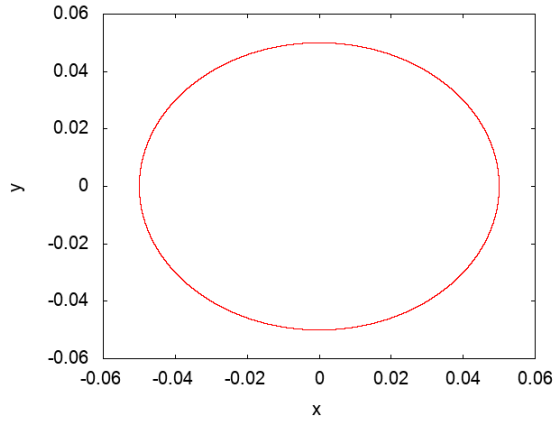
(a) (b)

Figure 33: Energy error (a) and x component of the constraints (eq. 6.37) (b) for the field configuration B with low time step ($h = 10$)

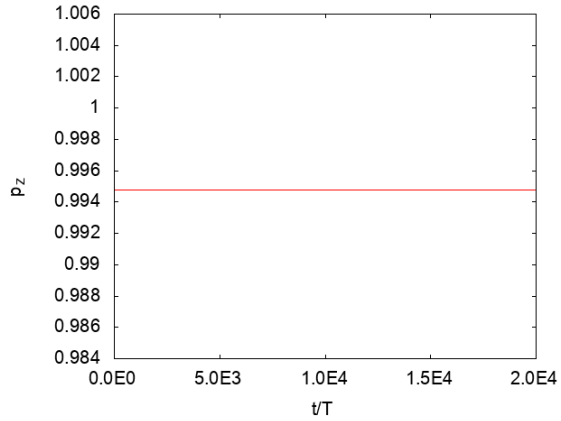


(a)

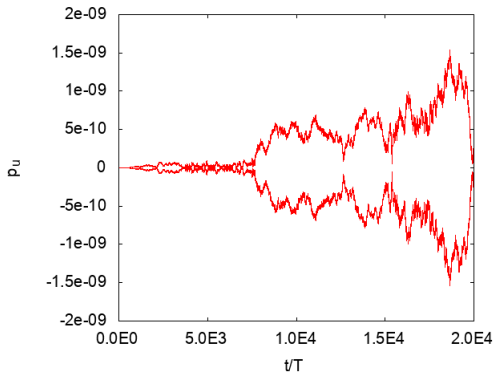
(b)



(c)



(d)



(e)

Figure 34: Midpoint rule applied to field configuration C1. (a) and (b) are the energy error respectively for the whole splitted flow and for a single part, (c) is the particle trajectory in the (x, y) plane, (d) and (e) are the z and u components of the discrete momentum (eq. 6.35)

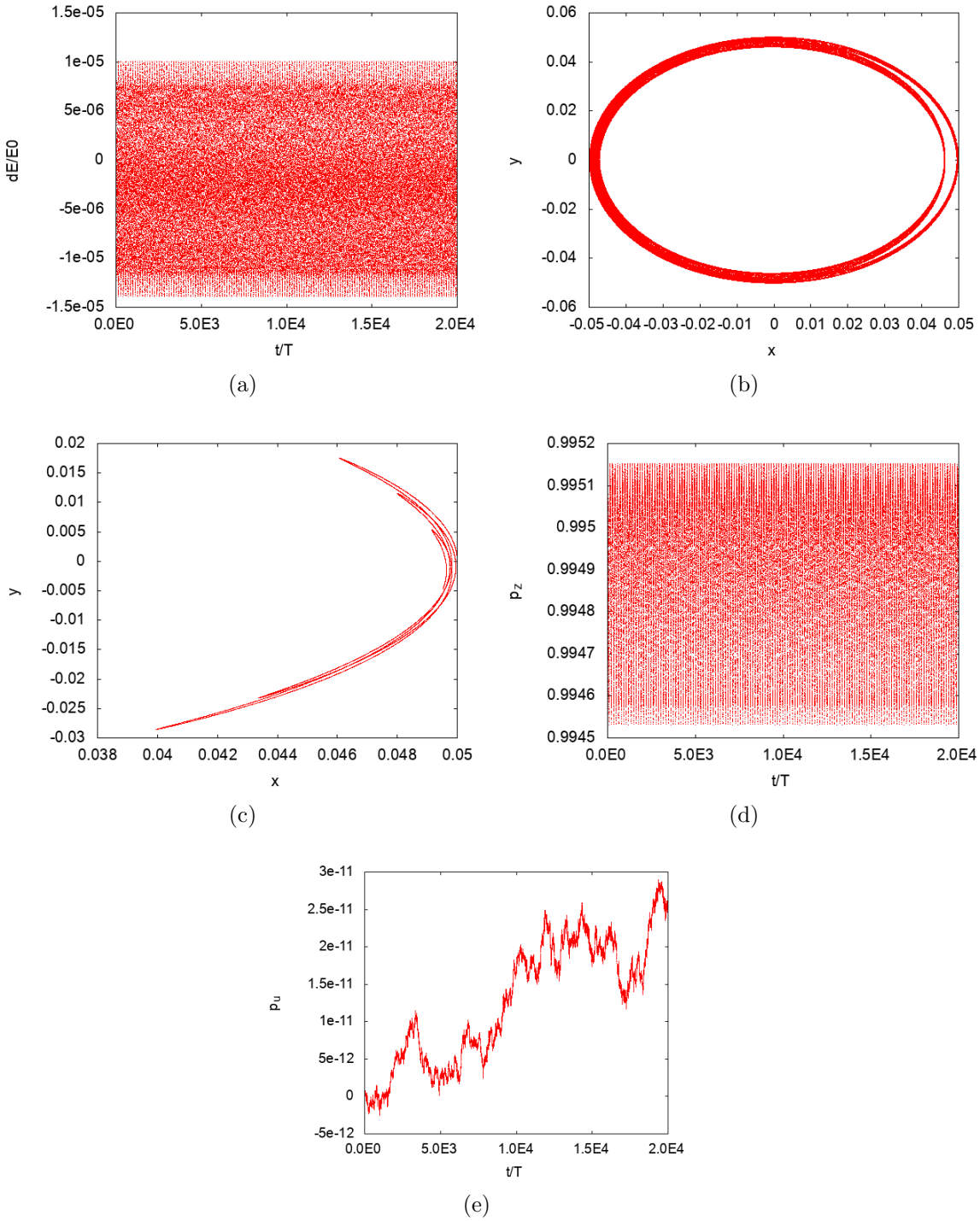


Figure 35: Midpoint rule applied to field configuration C2. (a) is the energy error, (b) is the particle trajectory in the (x, y) plane, (c) is a zoom showing the banana orbits, (d) and (e) are the z and u components of the discrete momentum (eq. 6.35)

6.4.2 Implicit Scheme 2: Three dimensional Lagrangian

We have already seen in chapter 2 that the continuous equations of motion can be solved trivially for u to give:

$$u = \hat{\mathbf{b}} \cdot \dot{\mathbf{x}} \tag{6.38}$$

It is quite easy to show that the Euler-Lagrange equations for the guiding center phase-space Lagrangian with the velocity u explicited with respect to the other variables are equivalent to the original Euler-Lagrange equations.
Thus, it is reasonable to wonder how a variational integrator would behave if applied to such Lagrangian:

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \mathbf{A}(\mathbf{x}) \cdot \dot{\mathbf{x}} + \frac{1}{2} \left(\mathbf{b}(\mathbf{x}) \cdot \dot{\mathbf{x}}\right)^2 - \mu B(\mathbf{x}) \tag{6.39}$$

The midpoint discretization gives:

$$\mathcal{L}_d(\mathbf{x}_0, \mathbf{x}_1) = \mathbf{A}(\tilde{\mathbf{x}}_0) \cdot \Delta\mathbf{x}_0 + \frac{1}{2h} \left(\mathbf{b}(\tilde{\mathbf{x}}_0) \cdot \Delta\mathbf{x}_0\right)^2 - h\mu B(\tilde{\mathbf{x}}_0) \tag{6.40}$$

The discrete integrator is then defined by matching of the discrete momenta:

$$\begin{aligned} \mathbf{p}_k &= \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} \left(\tilde{\mathbf{x}_k}, \frac{\Delta\mathbf{x}_k}{2}\right) - \frac{h}{2} \frac{\partial \mathcal{L}}{\partial \mathbf{x}} \left(\tilde{\mathbf{x}_k}, \frac{\Delta\mathbf{x}_k}{2}\right) \\ \mathbf{p}_{k+1} &= \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} \left(\tilde{\mathbf{x}_k}, \frac{\Delta\mathbf{x}_k}{2}\right) + \frac{h}{2} \frac{\partial \mathcal{L}}{\partial \mathbf{x}} \left(\tilde{\mathbf{x}_k}, \frac{\Delta\mathbf{x}_k}{2}\right) \end{aligned} \tag{6.41}$$

It is important to stress that this new Lagrangian is still degenerate (the degrees of freedom of the guiding center are two, hence a regular Lagrangian must be defined in a two dimensional space), and it doesn't correspond to a phase-space or an Hamiltonian formalism.
This is reflected to the fact that the Legendre transform is singular and new primary constraints arise. In fact, the Legendre transform reads:

$$\mathbf{p} \equiv \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} = (\hat{\mathbf{b}} \cdot \dot{\mathbf{x}})\hat{\mathbf{b}} + \mathbf{A} \tag{6.42}$$

Hence, it is straightforward to show that there are two constraints:

$$\phi = \begin{pmatrix} A^x + \dfrac{b^x}{b^z}(p^z - A^z) - p^x \\ A^y + \dfrac{b^y}{b^z}(p^z - A^z) - p^y \end{pmatrix} \tag{6.43}$$

Also, from the Hamiltonian function of the original 4D system, we know that the following function H is conserved:

$$H = \frac{u^2}{2} + \mu B = \frac{(\hat{\mathbf{b}} \cdot \dot{\mathbf{x}})^2}{2} + \mu B = \frac{\|\mathbf{p} - \mathbf{A}\|^2}{2} + \mu B \tag{6.44}$$

This function is conserved along the orbits of the particle. However, it is important to stress that we can't just apply hamilton's equations to this function, as one could think, since the function H is identical in form to an Hamiltonian of a charged particle in an electromagnetic field.
In fact, the Lagrangian 6.39 does not possess an hamiltonian structure: it is neither a regular Lagrangian nor a phase-space Lagrangian. The only correct equations are the Euler-Lagrange equations for this Lagrangian, which are not the hamilton's equations for H.
These equations are again a set of implicit equations and have to be solved with a first guess integrator.
The procedure used is the following: we can map a point $(\mathbf{x}, \mathbf{p})$ to a 4D point $z = (\mathbf{x}, u)$ with the following equation:

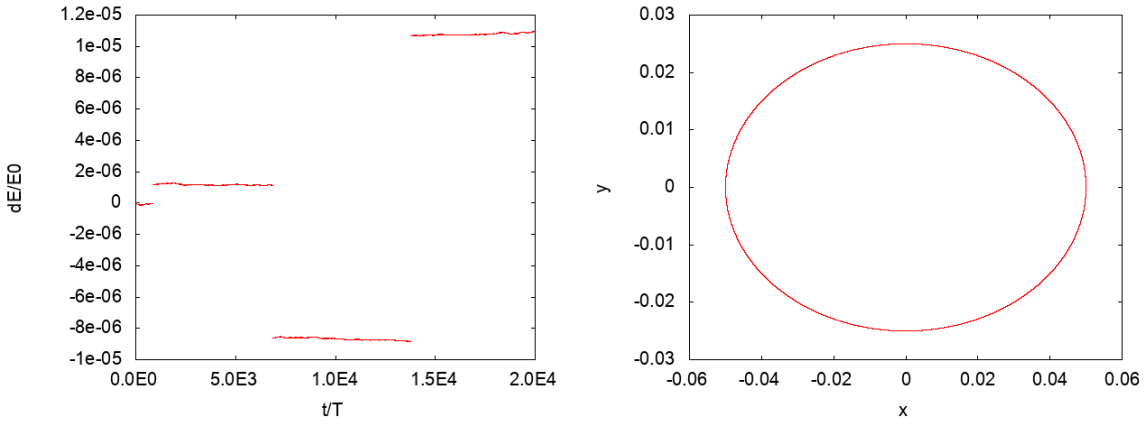
$$u = \|\mathbf{p} - \mathbf{A}\| \tag{6.45}$$

Hence, we can find the velocity u_0 from a point $(\mathbf{x}_0, \mathbf{p}_0)$ of this 3D integrator, RK4 is applied to find the following point x_1, u_1 and finally Newton's method allows to converge to the implicit equations.
Alternatively, we can find the velocity u_0 from $\mathbf{x}_0$ and the 4D momenta p_0^z by imposing the 4D constraints (eq. 6.26). Finally, the point z_1, p_1^z is found with one of the linearizations explained afterwards and Newton iterations are used for converging the implicit equations.
We chose the latter method, since it gives the best results and it is the most performant.

Since all variables are velocity dependent, we expect that only the conjugate momenta relative to Noether symmetries are conserved.
Hence, field configurations A,B and C1 which are toroidally symmetric should conserve the z component of the discrete momentum:

$$p_k^z = \frac{\partial \mathcal{L}}{\partial \dot{z}} \left(\tilde{\mathbf{x}}_k, \frac{\Delta \mathbf{x}_k}{h} \right) = \frac{1}{h} (\hat{\tilde{\mathbf{b}}}_k \cdot \Delta \mathbf{x}_k) \tilde{b}_k^z + \tilde{A}_k^z = const. \tag{6.46}$$

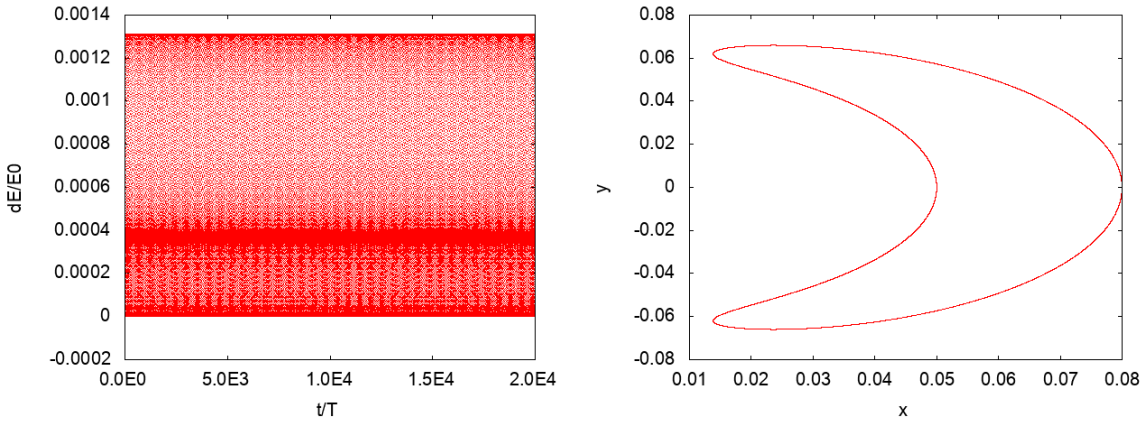
Numerical Results Figure 36, 37, 38 and 39 show the numerical results respectively for the field configurations A,B, C1 and C2.
The toroidal conjugate momentum p^z is correctly conserved for the fields A,B and C1. Also, the energy is bounded for long times for all the cases in consideration.



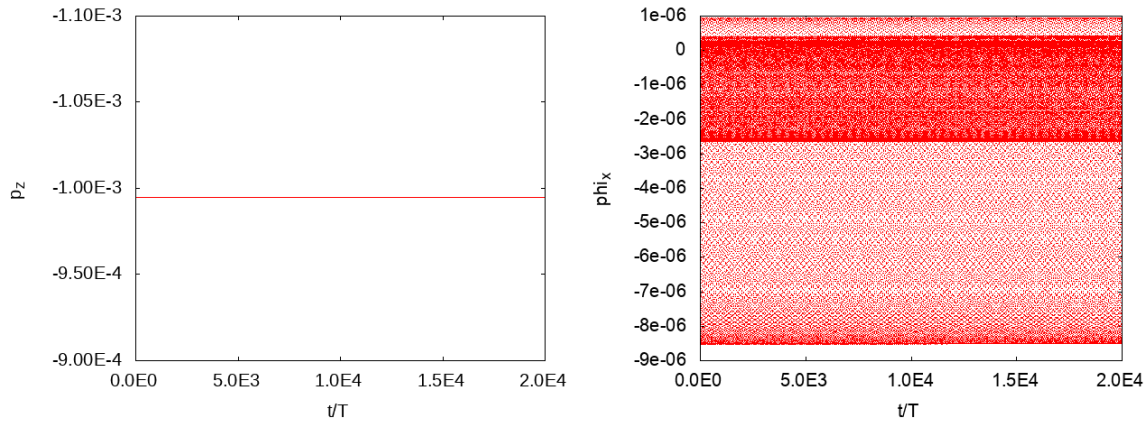
(a) (b)

(c) (d)

Figure 36: 3D Midpoint rule applied to field configuration A. (a) is the energy error, (b) is the particle trajectory in the (x, y) plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the first component of the constraints (eq. 6.43)



(a) (b)



(c) (d)

Figure 37: 3D Midpoint rule applied to field configuration B. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the first component of the constraints (eq. 6.43)

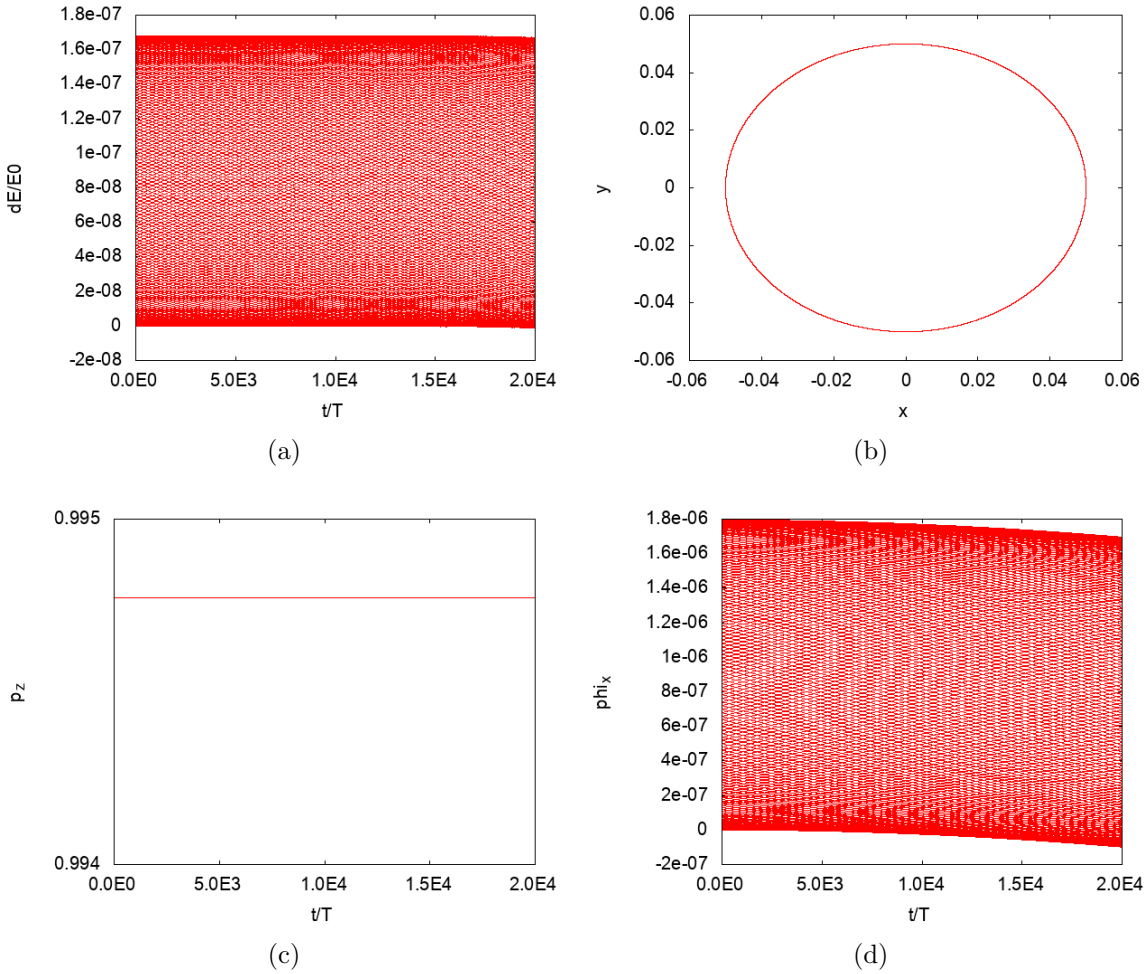


Figure 38: 3D Midpoint rule applied to field configuration C1. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the first component of the constraints (eq. 6.43)

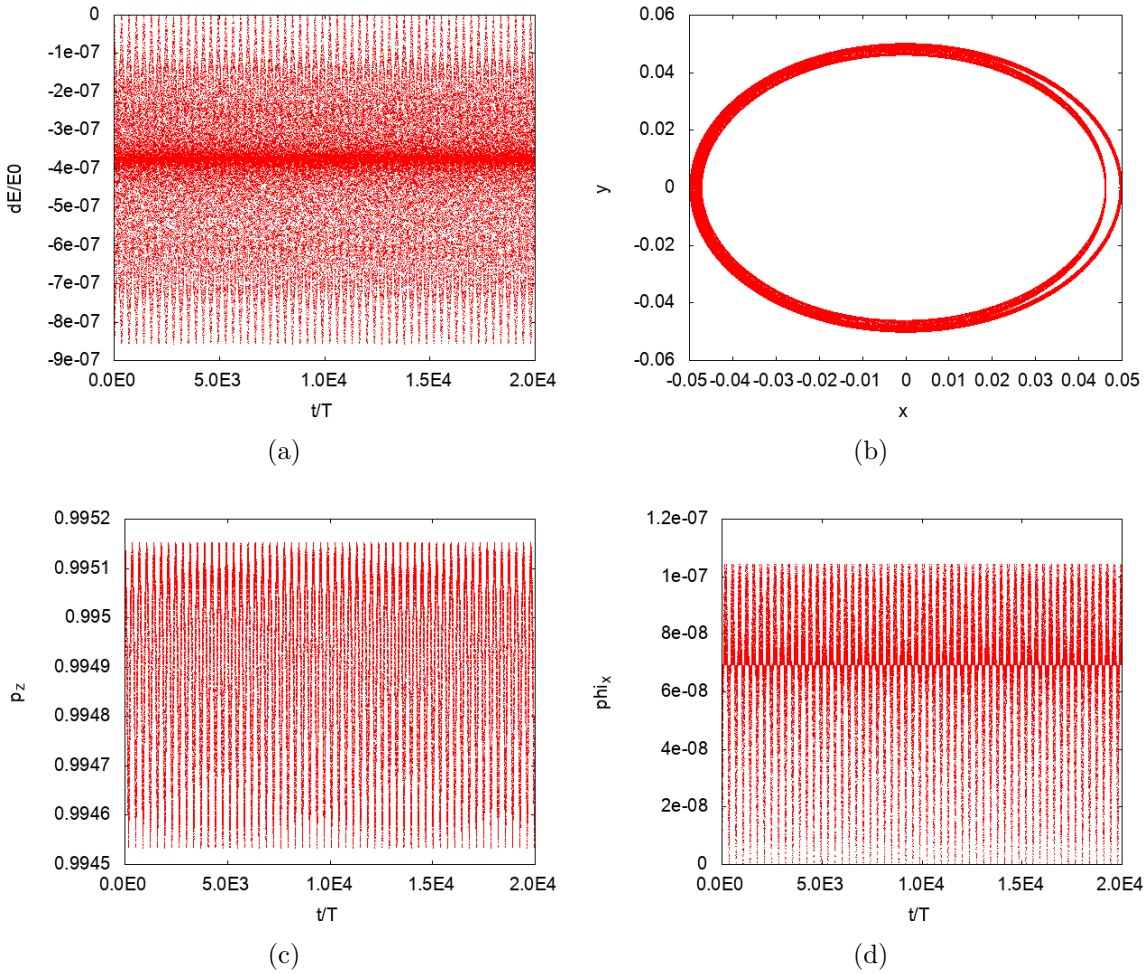


Figure 39: 3D Midpoint rule applied to field configuration C2b. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the first component of the constraints (eq. 6.43)

6.4.3 Semiexplicit Scheme

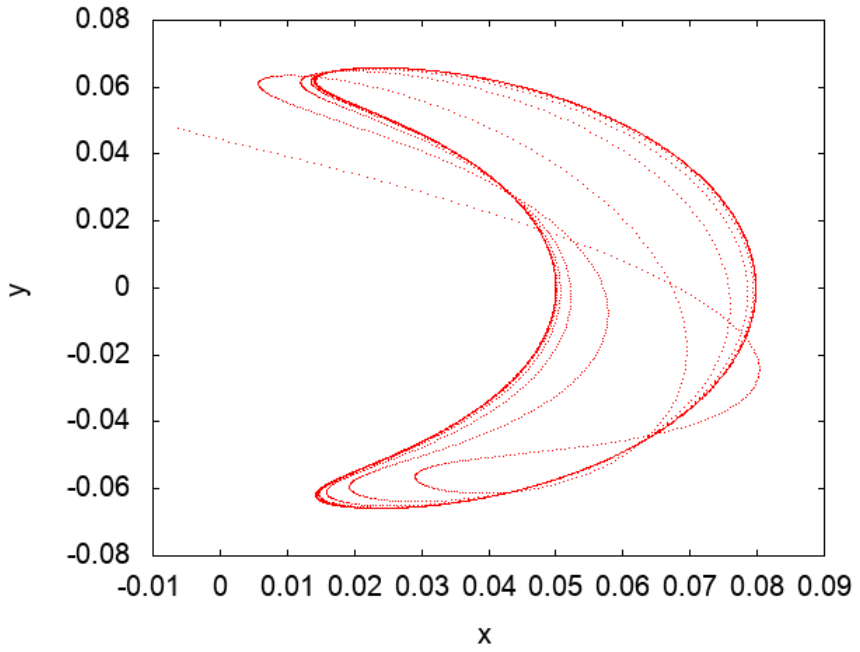
In the original papers of Qin ([20] and [21]) a slightly different discrete Lagrangian was used, apparently for stability reasons and for other reasons that will be clear after the linearization done in the following paragraphs:

$$\mathcal{L}_d(\mathbf{z}_0, \mathbf{z}_1) = \frac{A_1^\dagger + A_0^\dagger}{2} \cdot \Delta x_0 - h\frac{u_0 u_1}{2} - h\mu B_0 \tag{6.47}$$

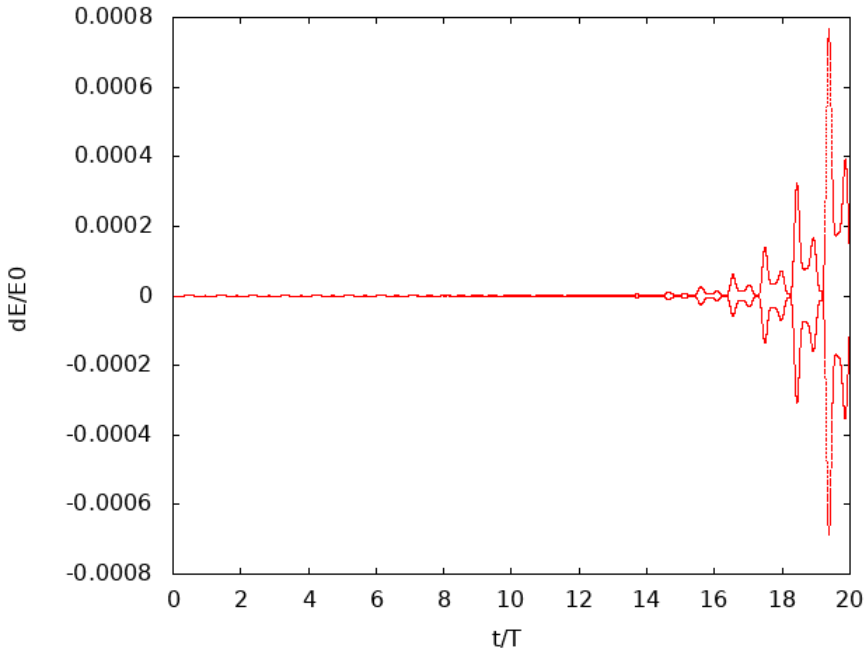
The main differences are the modified vector potential and the magnetic field evaluated exactly at each step rather than at the midpoint and the velocity term in the Hamiltonian.
We can notice that this has no longer a midpoint discretization form. Of course, the flow of the variational integrator applied to this modified Lagrangian is still symplectic and the discrepancy from the flow of the original midpoint Lagrangian is small.

Instabilities This integrator has shown to produce instabilites in all the magnetic field configuration we used. This behaviour is illustrated in figure 40, where a tokamak configuration was used. The instabilities are characterized by oscillations which are initially small and increase exponentially in times, breaking the integrator in a small amount of time and hence making it unusable in practice. Furthermore, they're present in some of the explicit schemes we will study below.
Squire and Qin [25] suggested that these instabilities could be due to an incorrect behaviour of the conserved symplectic forms when the time step is brought to 0, and that they could be mitigated by searching for particular gauge transformations of the Lagrangian.

In the following paragraphs, we will suggest new explicit linearization of the symplectic integrator which are apparently free from instabilities.



(a)



(b)

Figure 40: Instabilities produced by the linearized integrator with field configuration B. (a) is the particle trajectory in the poloidal plane (x, y), (b) is the energy error

6.4.4 Explicit Scheme 1

The purpose of this section is to find an explicit integrator which remains close to the solution. Of course, such an algorithm would be much faster than an implicit one. The standard technique is the linearization of the implicit scheme.

Let's start with equations (5.16):

$$\mathbf{p}_1 = \Theta(\tilde{\mathbf{z}}_1) + \frac{h}{2}\nabla H(\tilde{\mathbf{z}}_1) - \Theta^T_{(1)}(\tilde{\mathbf{z}}_1)\frac{\Delta \mathbf{z}_1}{2} \tag{6.48a}$$

$$= \Theta(\tilde{\mathbf{z}}_0) - \frac{h}{2}\nabla H(\tilde{\mathbf{z}}_0) + \Theta^T_{(1)}(\tilde{\mathbf{z}}_0)\frac{\Delta \mathbf{z}_0}{2} \tag{6.48b}$$

The idea of the linearization is to expand these equations around z_1, as we did in eqs (5.17) and to retain only the first two terms of the symplectic forms and the first term of the Hamiltonian:

$$\mathbf{p}_1 = \Theta_1 + \frac{h}{2}\nabla H_1 + \Omega_1\frac{\Delta \mathbf{z}_1}{2} \tag{6.49a}$$

$$= \Theta_1 - \frac{h}{2}\nabla H_1 - \Omega_1\frac{\Delta \mathbf{z}_0}{2} \tag{6.49b}$$

Hence, starting from two points $\mathbf{z}_0, \mathbf{p}_0$, we can find $\mathbf{z}_1$ and $\mathbf{p}_1$ with the following explicit equations:

$$\mathbf{z}_1 = \mathbf{z}_0 + 2\Omega_0^{-1}\left(\mathbf{p}_0 - \Theta_0 - \frac{h}{2}\begin{pmatrix} \mu\nabla B_0 \\ u_0 \end{pmatrix}\right) \tag{6.50}$$

$$\mathbf{p}_1 = \Theta_1 - \Omega_1\frac{\Delta \mathbf{z}_0}{2} - \frac{h}{2}\begin{pmatrix} \mu\nabla B_0 \\ u_0 \end{pmatrix} \tag{6.51}$$

Alternatively, subtracting eqs. (6.49), we can find the equations for $\mathbf{z}_2$ in Lagrangian form:

$$\frac{\Omega_1}{2}(\mathbf{z}_2 - \mathbf{z}_0) + h\begin{pmatrix} \mu\nabla B_1 \\ u_1 \end{pmatrix} = 0 \tag{6.52}$$

Following the same discussion done in the previous sections, we can notice that the integrator will still be splitted in two parts.
Again, the constraints will not be conserved as the symplectic form is not constant.

A legitimate question is what does happen if we choose to retain also the term linear in Δz in eq. (6.48). We will see in section 6.4.7 that this will have beneficial effects on the integrator.

Out of all of the linearizations we tried, the Lagrangian initialization performed much better than the Hamiltonian one, in contrast with the behaviour of the implicit integrators.
The Hamiltonian initialization proved to possess frequently drifts in the energy error and to be always imprecise. For this reason, we will use the Lagrangian initialization for all the linearized integrators.

Numerical Results Figure 41 shows the linearization applied to the 2D field A. The energy is bounded for long times and the momenta p_z and p_u are exactly conserved. The conservation of momenta in this case is just a fortunate coincidence and it is a consequence of the simplicity of the magnetic field considered.
All the other field configurations are unstable, similarly to the semiexplicit scheme.

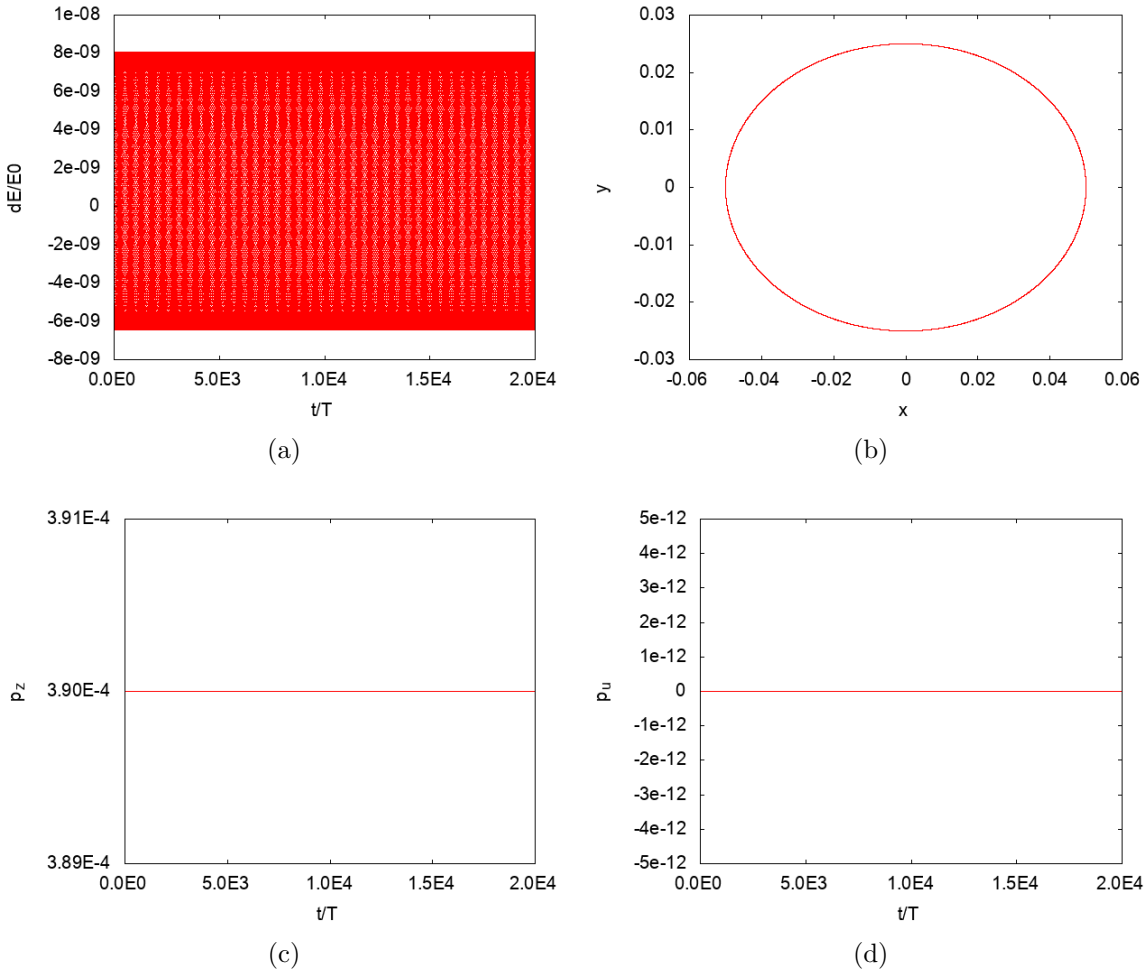


Figure 41: 3D Midpoint linearization applied to field configuration A. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

6.4.5 Explicit Scheme 2: Qin's Version

Starting from the semiexplicit Lagrangian (eq. 6.47), one can easily show that the linearization of all terms give rise to the same equations of the linearized midpoint Lagrangian with the exception of the quadratic velocity terms in the Hamiltonian. This modified velocity has proven to give better stability results than the original linearization.

The new discrete equations in position-momentum form take the following form:

$$\mathbf{p}_1 = \Theta_1 + \frac{h}{2}\nabla H_1 + \Omega_1^- \frac{\Delta \mathbf{z}_1}{2} \tag{6.53a}$$

$$= \Theta_1 - \frac{h}{2}\nabla H_1 - \Omega_1^+ \frac{\Delta \mathbf{z}_0}{2} \tag{6.53b}$$

where $\Omega^{\pm}$ are the following modified symplectic forms:

$$\Omega^{\pm} = \begin{pmatrix} 0 & -B_z^\dagger & B_y^\dagger & \hat{b}_x \\ B_z^\dagger & 0 & -B_x^\dagger & \hat{b}_y \\ -B_y^\dagger & B_x^\dagger & 0 & \hat{b}_z \\ -\hat{b}_x & -\hat{b}_y & -\hat{b}_z & \mp h \end{pmatrix} \tag{6.54}$$

The Lagrangian form reads:

$$\frac{\Omega_1^-}{2}(\mathbf{z}_2 - \mathbf{z}_1) + \frac{\Omega_1^+}{2}(\mathbf{z}_1 - \mathbf{z}_0) + h \begin{pmatrix} \mu \nabla B_1 \\ u_1 \end{pmatrix} = 0 \tag{6.55}$$

Writing explicitly the symplectic terms we get:

$$\begin{cases} \frac{1}{2}\left[A_1^{\dagger i,j} - A_1^{\dagger j,i}\right](x_2^i - x_0^i) - \frac{b_1^j}{2}\left[u_2 - u_0\right] = h\mu B_1^{,j} \\ \frac{1}{2} b_1^i (x_2^i - x_0^i) = \frac{h}{2}(u_2 + u_0) \end{cases} \tag{6.56}$$

Eventually, we can decouple the spatial part of the symplectic form from the velocity by substituting the velocity terms in the first three equations:

$$\begin{cases} \frac{1}{2}\left[A_1^{\dagger i,j} - A_1^{\dagger j,i}\right](x_2^i - x_0^i) - \frac{b_0^j}{2}\left[2u_0 - \frac{b_1^i}{h}(x_2^i - x_0^i)\right] = h\mu B_1^{,j} \\ \frac{1}{2} b_1^i (x_2^i - x_0^i) = \frac{h}{2}(u_2 + u_0) \end{cases} \tag{6.57}$$

so that the system can be solved by inverting a 3×3 matrix and by evaluating the velocity equation.

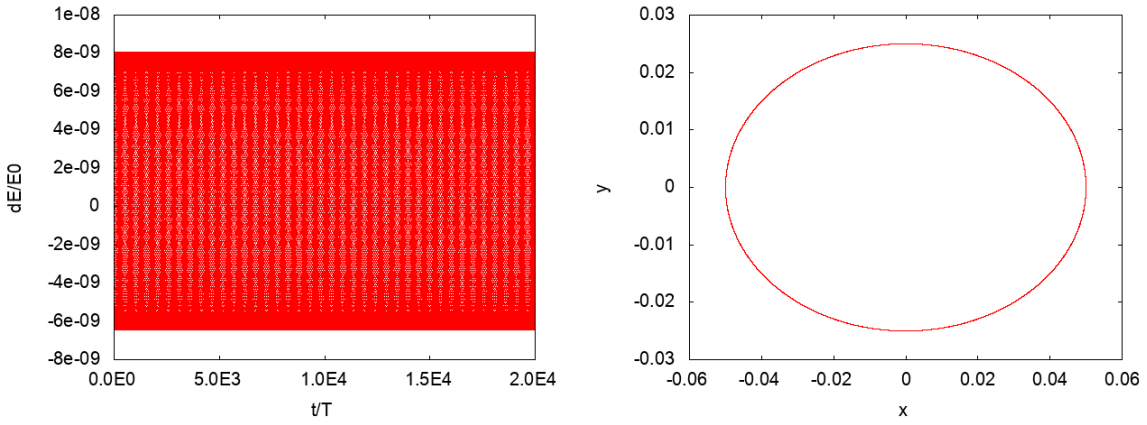
Numerical Results Similarly to the original linearization (section 6.4.4), the force-free configurations C1 and C2 are always unstable.
However, the region of instability of the tokamak field B is slightly smaller for this version and the instabilities proved to be very sensible to the initial conditions used.

Figure 42 and 43 show the results for the 2D field A and the tokamak field B with a Lagrangian initialization applied to the initial point

$$\mathbf{z}_0 = (\mathbf{x}_0, u_0) = (0.08461, 0.00228, 4.51147, -0.00057) \tag{6.58}$$

which proved to produce a stable flow.

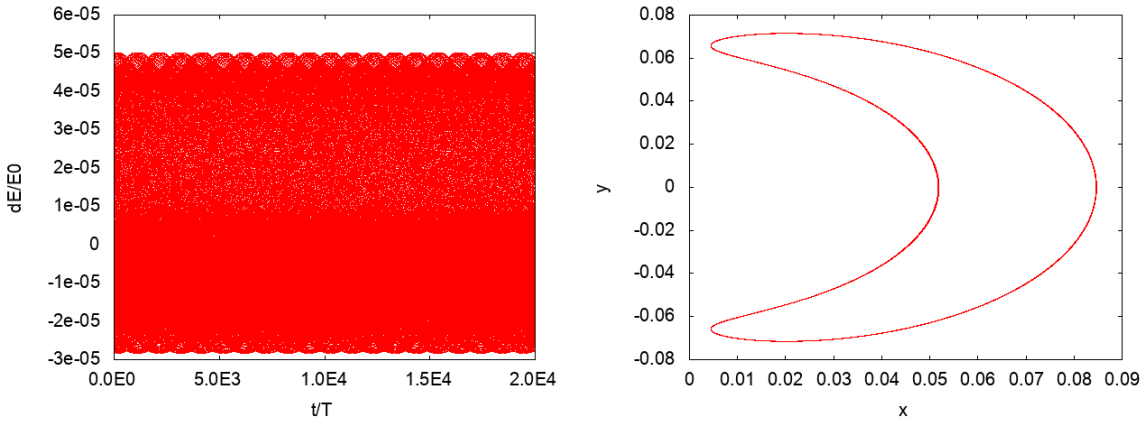
The energy is bounded for both cases. Note that the z component of the momentum is no longer conserved, since the discrete Noether theorem is guaranteed to work only with the implicit symplectic integrator.



(a) (b)

(c) (d)

Figure 42: Qin linearization applied to field configuration A. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum



(a) (b)

(c) (d)

Figure 43: Qin linearization applied to field configuration B. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

6.4.6 Explicit Scheme 3

In the same spirit of section 6.4.2, we tried a different version of the linearization made by Qin.
The term $u_2 - u_0$ that appears in equations (6.57) is just the discretization of terms proportional to the derivative of u in the Taylor expansion of the one-form Θ.
This term was substituted by:

$$\frac{u_2 - u_0}{2h} \simeq \hat{b} \cdot \ddot{x} \simeq \frac{b_1^i (x_2^i - 2x_1^i + x_0^i)}{h^2} \tag{6.59}$$

This approximation seems quite rude; however, we will see that this integrator has some interesting properties.
Writing explicitly the discrete Euler-Lagrange equations, we get:

$$\begin{cases} \frac{1}{2}\left[A_1^{\dagger i,j} - A_1^{\dagger j,i}\right](x_2^i - x_0^i) - \frac{b_1^j b_1^i}{h}\left[x_2^i - 2x_1^i + x_0^i\right] = h\mu B_1^{\dagger j} \\ \frac{1}{2} b_1^i (x_2^i - x_0^i) = \frac{h}{2}(u_2 + u_0) \end{cases} \tag{6.60}$$

or equivalently in position-momentum form:

$$\mathbf{p}_1 = \Theta_1 + \frac{h}{2}\nabla H_1 + \Omega_1^- \frac{\Delta \mathbf{z}_1}{2} \tag{6.61a}$$

$$= \Theta_1 - \frac{h}{2}\nabla H_1 - \Omega_1^+ \frac{\Delta \mathbf{z}_0}{2} \tag{6.61b}$$

$$\Omega^{\pm} = \begin{pmatrix} 0 & -B_z^\dagger & B_y^\dagger & 0 \\ B_z^\dagger & 0 & -B_x^\dagger & 0 \\ -B_y^\dagger & B_x^\dagger & 0 & 0 \\ -\hat{b}_x & -\hat{b}_y & -\hat{b}_z & \mp h \end{pmatrix} \mp \frac{2 b^i b^j}{h} \tag{6.62}$$

Numerical Results Figure 44, 45, 46 and 47 show the numerical results for the configurations A,B,C1 and C2.
Unlike the previous cases, the integrator is stable for every configuration we tested. The energy is bounded for long times and the momenta p^z and p^u are conserved only in the 2D field configuration and bounded in the other cases.

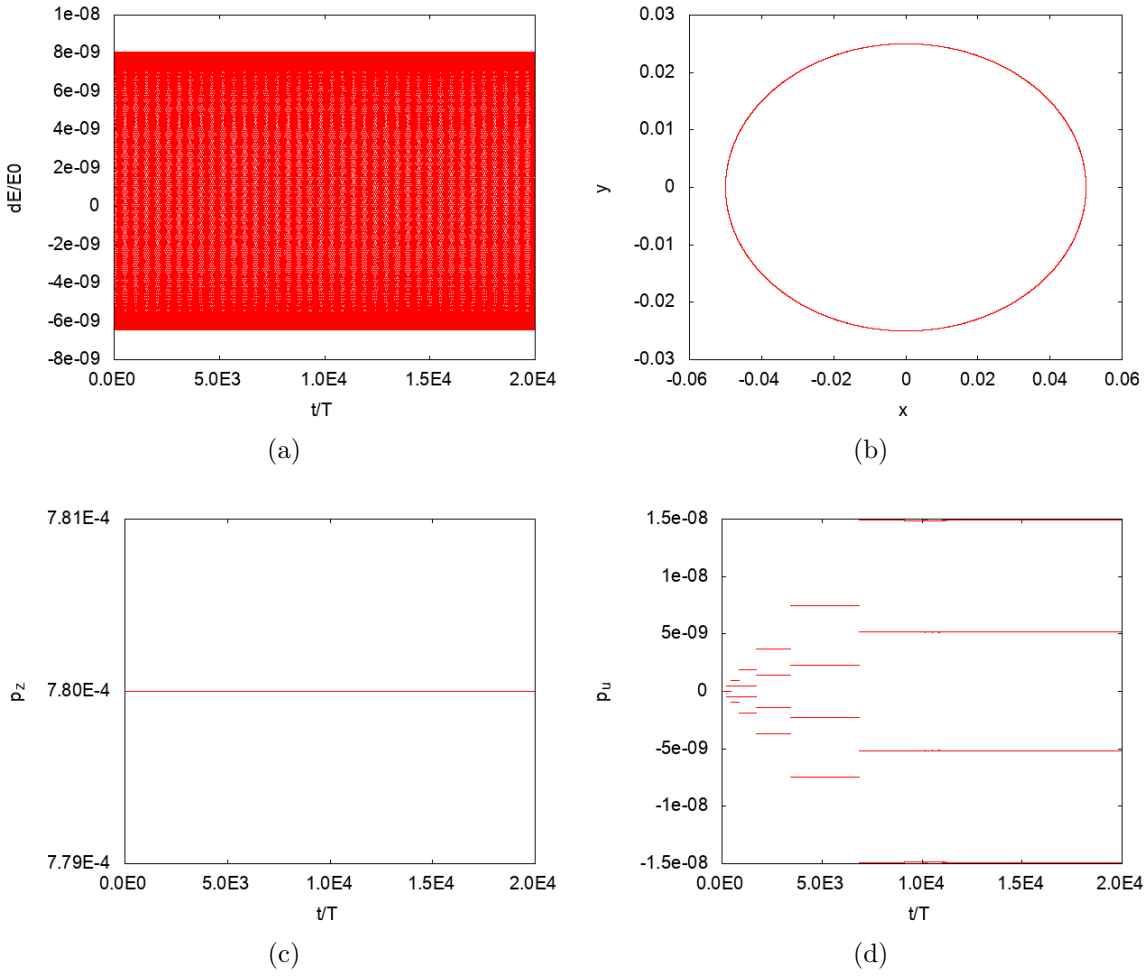


Figure 44: Modified Qin linearization applied to field configuration B. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

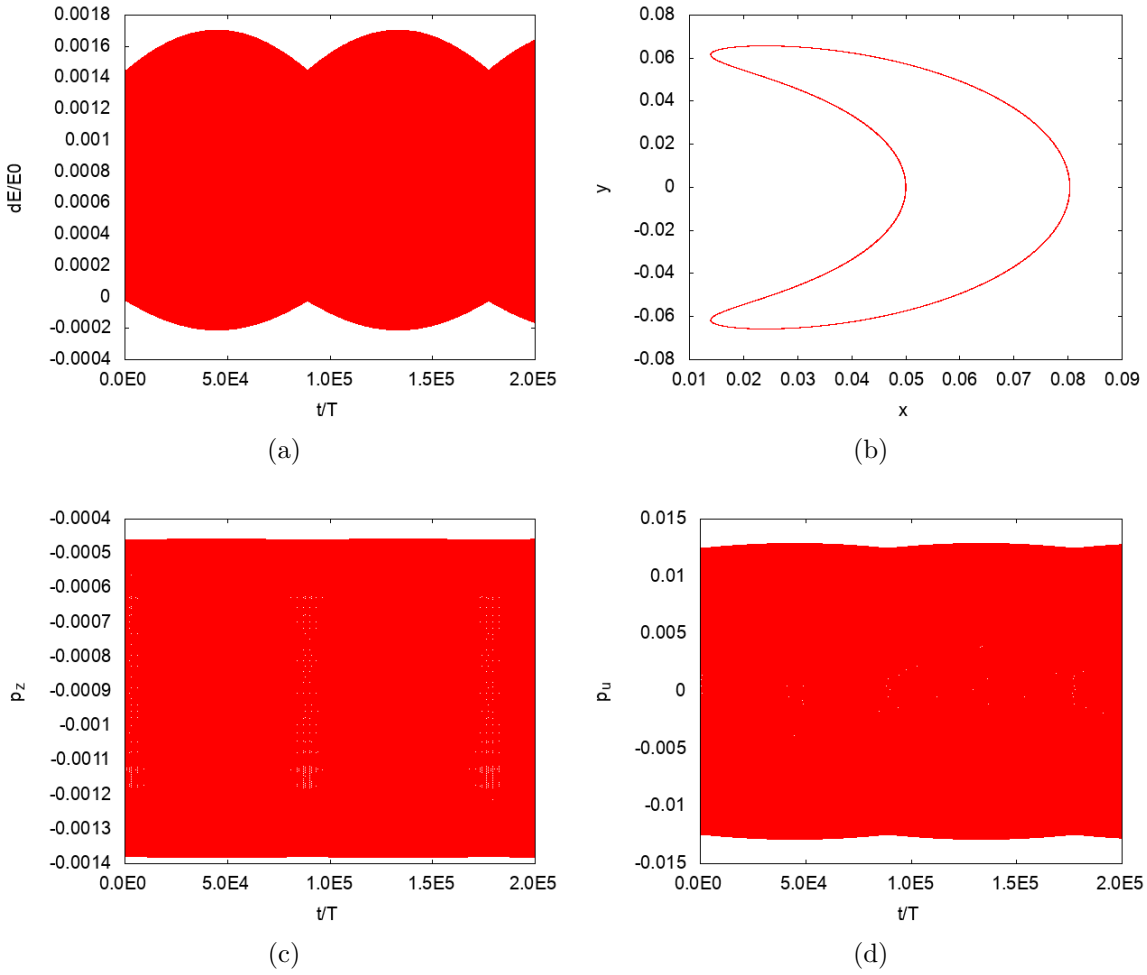


Figure 45: Modified Qin linearization applied to field configuration B. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

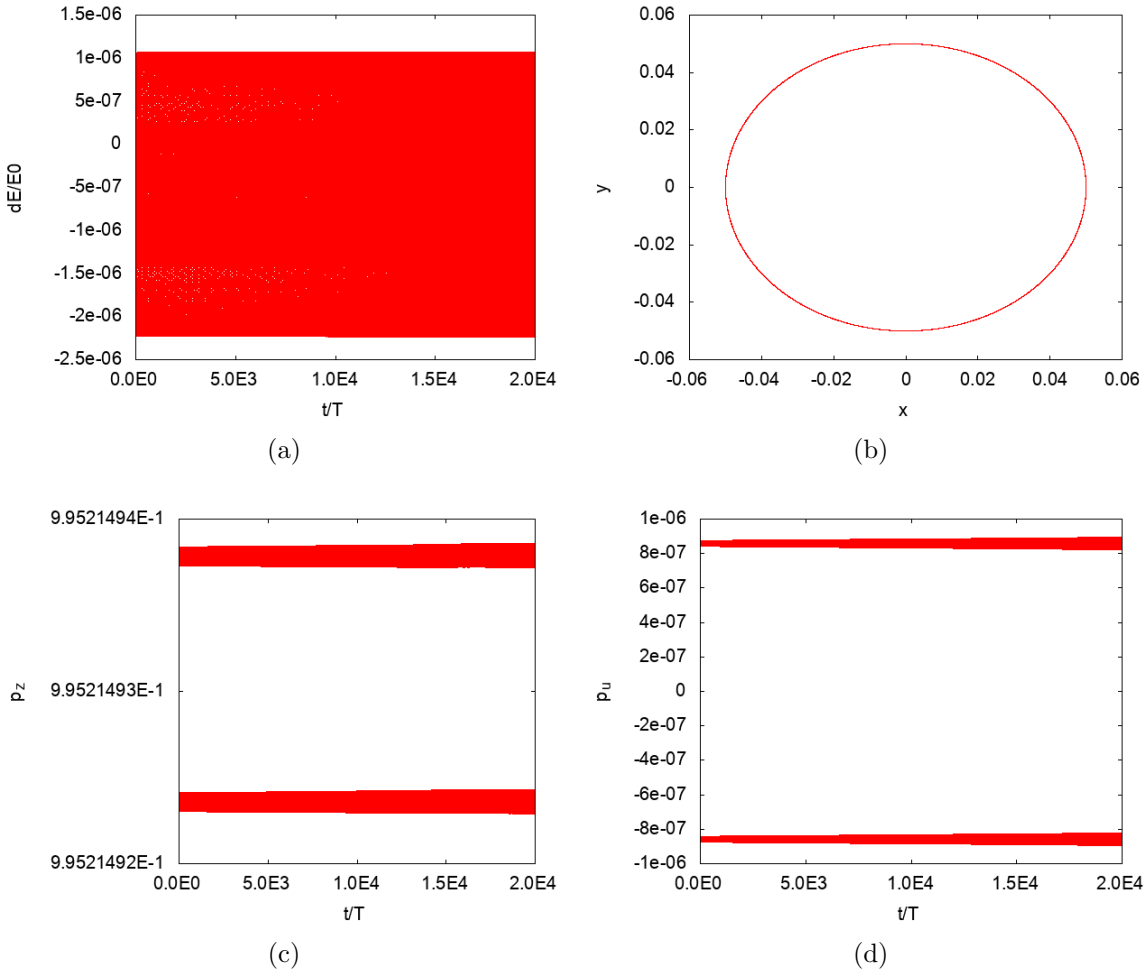
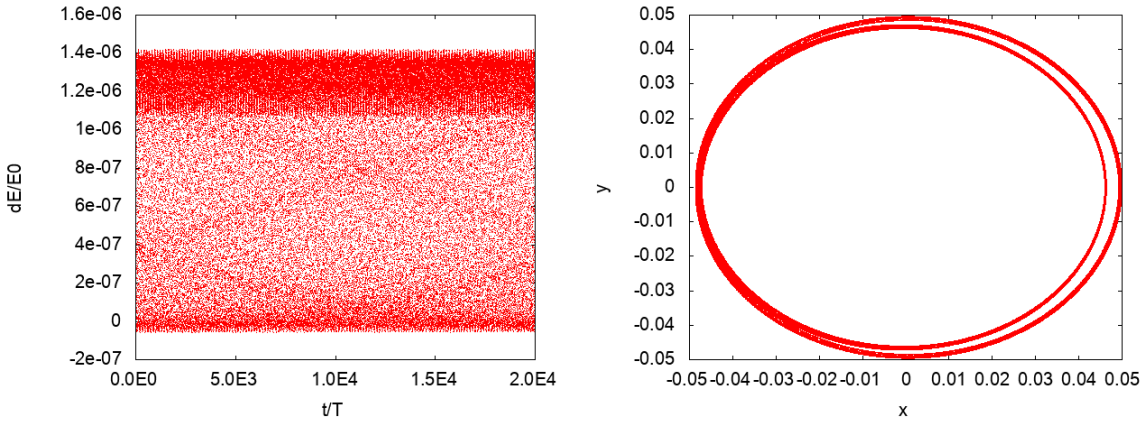


Figure 46: Modified Qin linearization applied to field configuration C1. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum



(a) (b)

(c) (d)

Figure 47: Modified Qin linearization applied to field configuration C2b. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

6.4.7 Explicit Scheme 4: First order Hamiltonian

When we performed the truncation of the Taylor serie in equation 6.49, we chose to exclude the first order term of the Hamiltonian.
We can wonder how the linearization behave if we truncate the series to the first order, so that we can write:

$$\mathbf{p}_1 = \Theta_1 + \frac{h}{2}\nabla H_1 + \Omega_1 \frac{\Delta \mathbf{z}_1}{2} + \frac{h}{2}\nabla H_{1(1)} \frac{\Delta \mathbf{z}_1}{2} \tag{6.63a}$$

$$= \Theta_1 - \frac{h}{2}\nabla H_1 - \Omega_1 \frac{\Delta \mathbf{z}_0}{2} + \frac{h}{2}\nabla H_{1(1)} \frac{\Delta \mathbf{z}_0}{2} \tag{6.63b}$$

where $\nabla H_{1(1)}$ is the Hessian of the Hamiltonian. Matching the momenta we get:

$$\frac{\Omega_1}{2}(\mathbf{z}_2 - \mathbf{z}_0) + h \begin{pmatrix} \mu \nabla B_1 \\ u_1 \end{pmatrix} + \frac{h}{4} \begin{pmatrix} \mu B_{,i,j}(\mathbf{x}_1) & 0 \\ 0 & 1 \end{pmatrix} (\mathbf{z}_2 - 2\mathbf{z}_1 + \mathbf{z}_0) = 0 \tag{6.64}$$

As we will see shortly, the effect of these new terms is the removing of the instabilities typical of the original linearizations.

Dropping the magnetic terms We tried a sightly modified version by removing the second derivative of the magnetic field, so that eq. (6.64) is rewritten as:

$$\frac{\Omega_1}{2}(\mathbf{z}_2 - \mathbf{z}_0) + h \begin{pmatrix} \mu \nabla B_1 \\ u_1 \end{pmatrix} + \frac{h}{4} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} (\mathbf{z}_2 - 2\mathbf{z}_1 + \mathbf{z}_0) = 0 \tag{6.65}$$

The integrator thus found is almost unchanged. Computationally, this leads to a more effective algorithm, since the second derivative of the magnetic field could be expensive to compute, specially for the force-free configuration.

Numerical Results Figure 48, 49, 50 and 51 show the numerical results for the field A,B,C1 and C2.
The energy is bounded for all the configurations, the momenta p^z and p^u are only conserved for the 2D field and bounded for the other cases.
In all the tests we performed, no instabilities were observed.

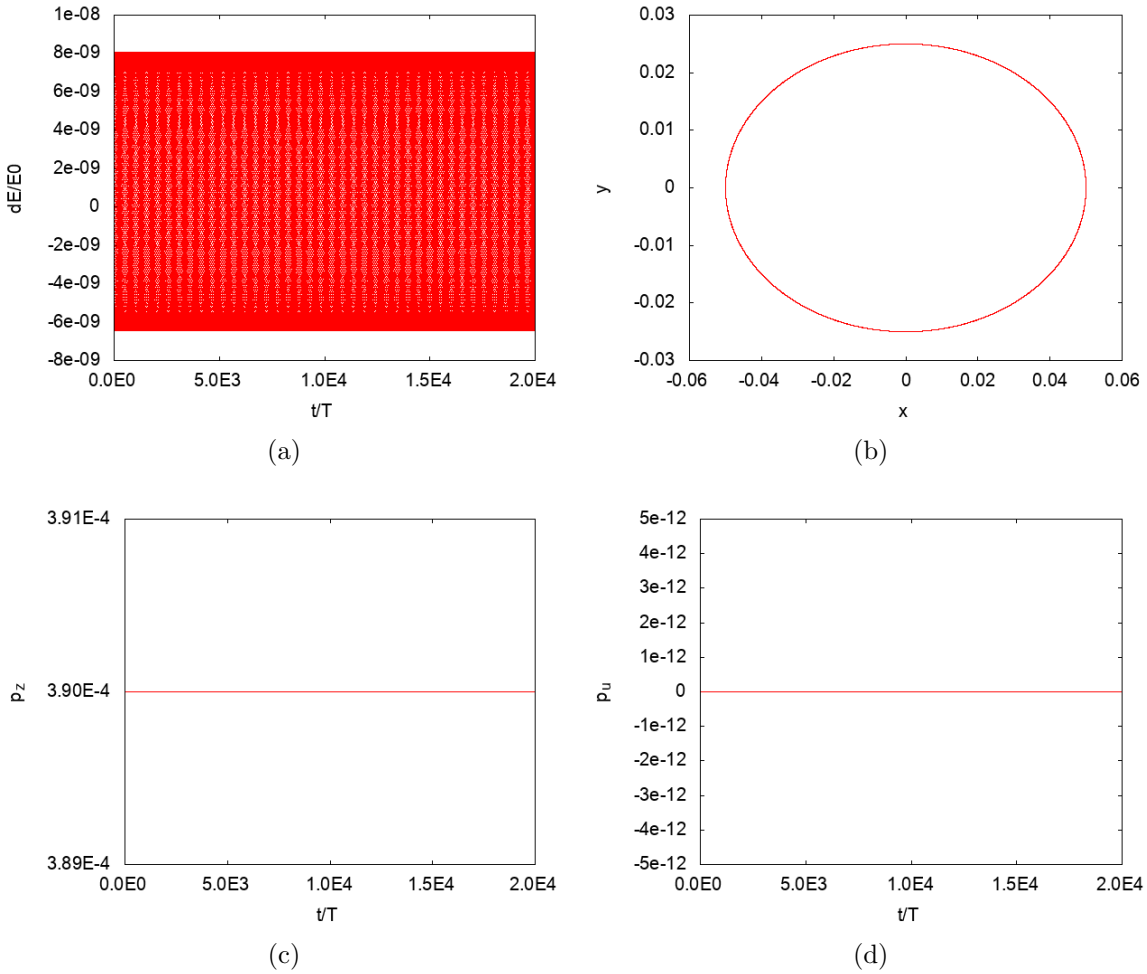


Figure 48: Linearization with first order Hamiltonian applied to field configuration A. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

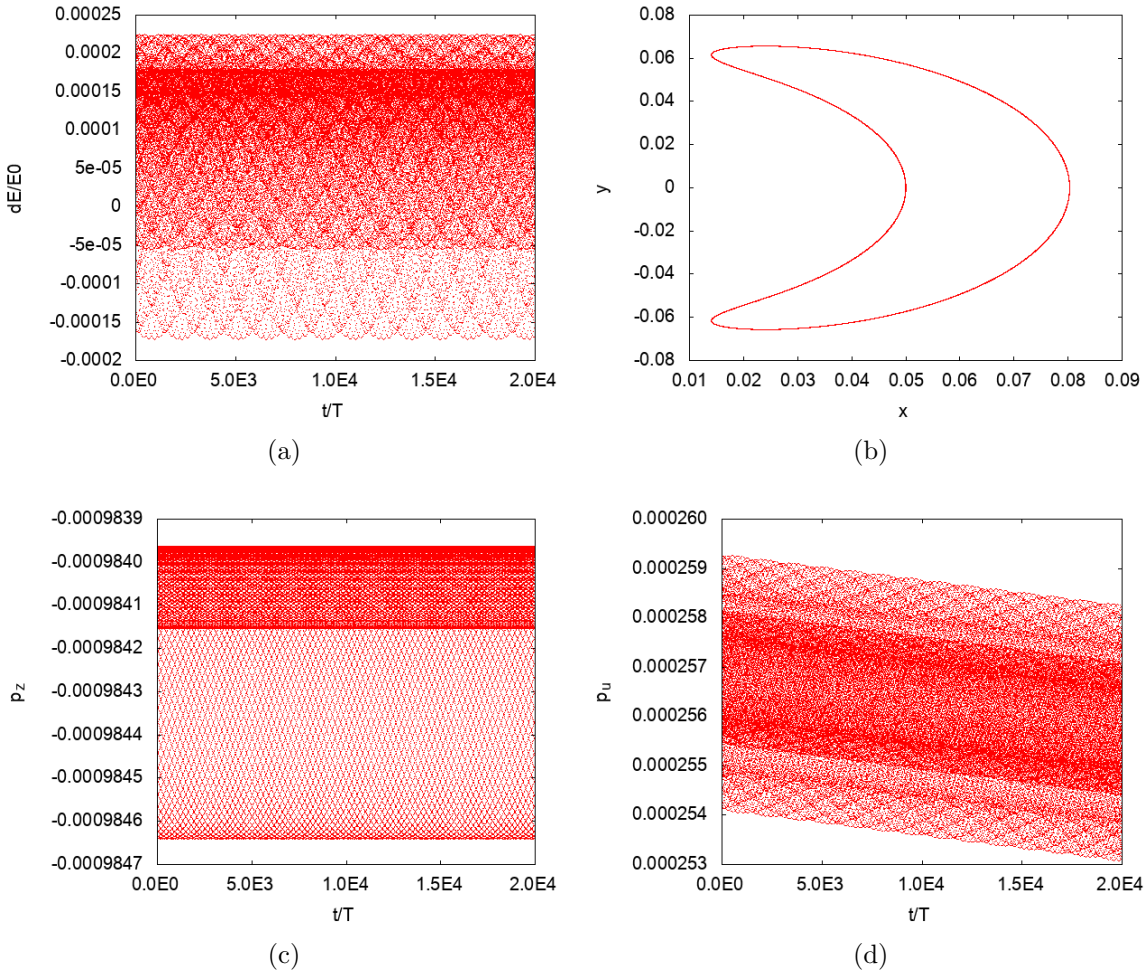


Figure 49: Linearization with first order Hamiltonian applied to field configuration B. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

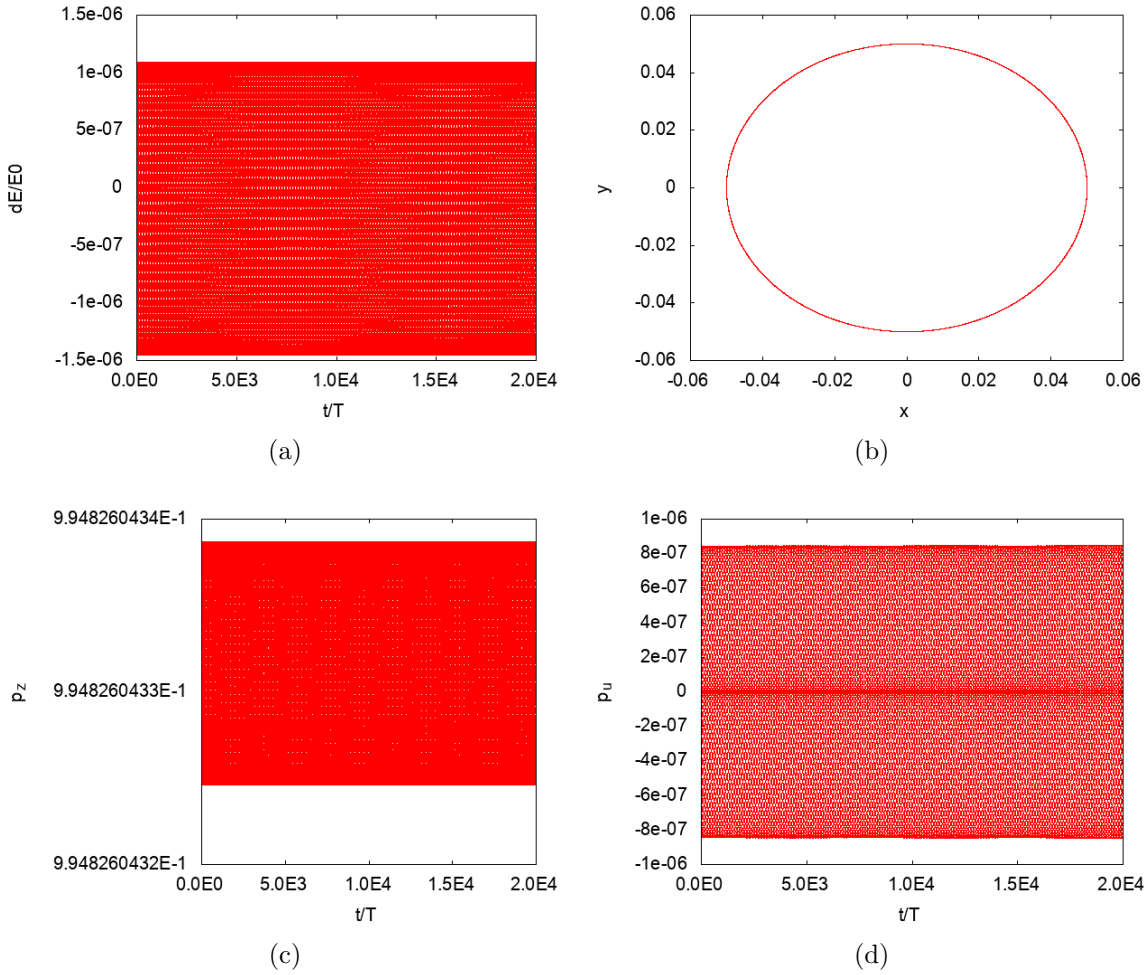


Figure 50: Linearization with first order Hamiltonian applied to field configuration C1. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

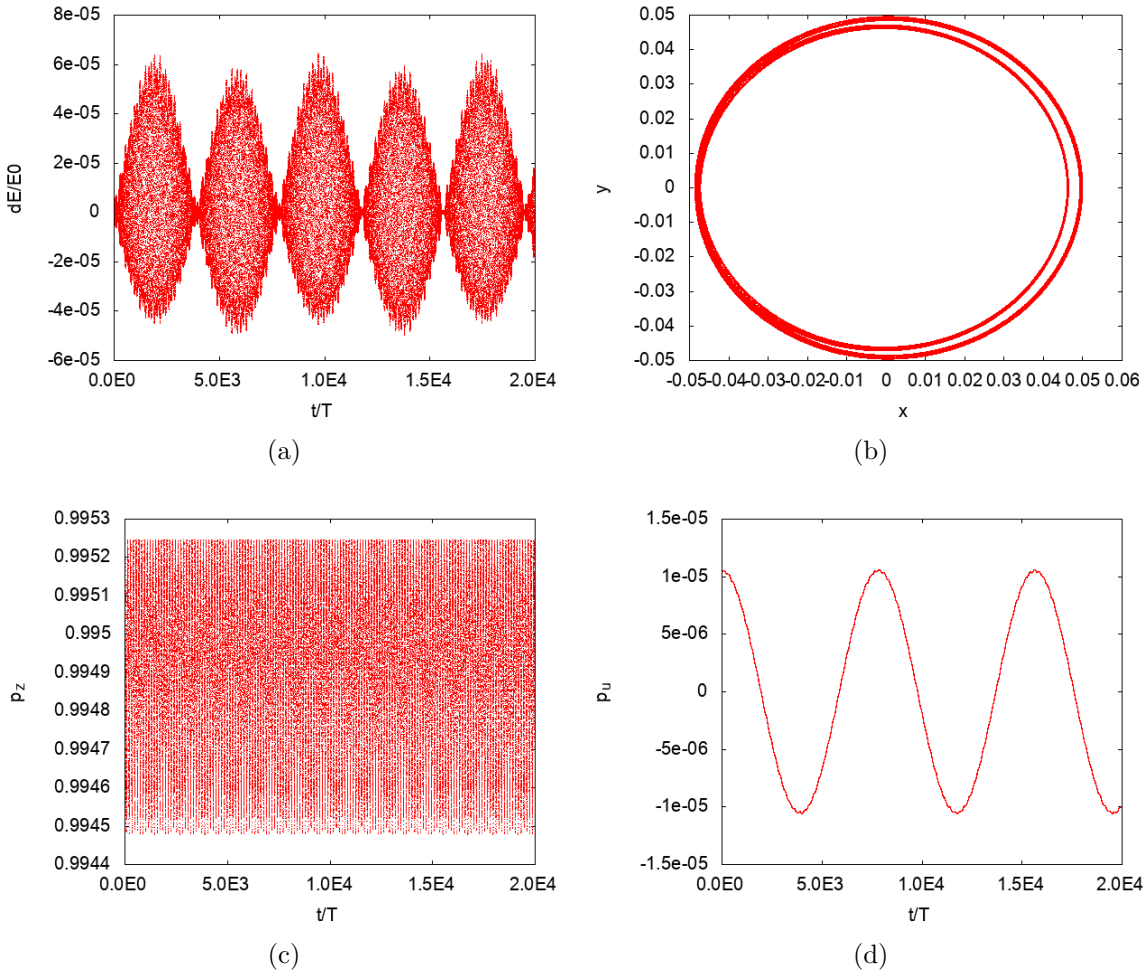


Figure 51: Linearization with first order Hamiltonian applied to field configuration C2. (a) is the energy error, (b) is the particle trajectory in the (x,y)plane, (c) is the z component of the discrete momentum (eq. 6.46), (d) is the u component of the discrete momentum

6.5 Review and comparison with RK4

Figure 53 shows the energy error for the field configuration B plotted as a function of time for the fourth order Runge-Kutta, the implicit scheme 1 and the explicit scheme 4.
In the case of RK4, the energy quickly decays, showing the inefficiency of this integrator for high time steps and high integration times.
All the symplectic integrators bounds the energy for long times, with the exception of the midpoint rule applied to the force-free configuration C1, which presents a small energy drift. The reason of this behaviour is unknown.
Note that the linearizations studied in this thesis has the same, and sometimes even better, accuracy than the non linearized midpoint rule.

The quantities associated to Noether symmetries are exactly conserved only with the non linearized integrators, both the 4D (section 6.4.1) and the 3D version (section 6.4.2). In the explicit versions, the Noether quantities are not exactly conserved, but still bounded.

It is worth noticing that all the implicit integrators require an Hamiltonian initialization, since the Lagrangian initialization produces very high drifts in the energy error.
This behaviour is the opposite for the linearized versions, which require a Lagrangian initialization to work correctly.
The complete comprehension of the theoretical aspects of this subject is very important and could lead to a better choice of the initialization.

The following table summarizes all these properties.

Integrator	Section	Energy Bounding	Noether	Initialization
RK4		No	No	-
Implicit 1	6.4.1	**Yes** (for A,B,C2)	**Yes**	Hamiltonian
Implicit 2	6.4.2	**Yes**	**Yes**	Hamiltonian
Semiexplicit	6.4.3	No†	**Yes**	Hamiltonian
Explicit 1	6.4.4	No†	No	Lagrangian
Explicit 2	6.4.5	No†	No	Lagrangian
Explicit 3	6.4.6	**Yes**	No	Lagrangian
Explicit 4	6.4.7	**Yes**	No	Lagrangian

Table 1: Comparison between the numerical integrators used in this chapter for the error bounding (see section 4.5.1), the conservation of Noether quantities (section 4.2.1), and the initialization used (section 6.3).
† Altough the energy seems to be bounded, exponential oscillations which grows very rapidly in a small amount of time are present (see section 6.4.3)

6.5.1 Stability Analysis

The explicit scheme 1 (section 6.4.4) and the Qin's version (section 6.4.5) are usually unstable and thus unusable in many cases.
However, we have found two linearizations, the explicit scheme 3 (section 6.4.6) and 4 (6.4.7) which are apparently free from instabilities, at least with the cases we studied.

One way to show the stability properties of these integrators is by studying the **Lyapunov exponents**. Usually, the Lyapunov exponents are used to study the rate of separation of infinitesimally close trajectories of a dynamical system.

Specifically, if the starting points of two trajectories are separated by a small quantity δx_0, the separation of the trajectories evolves approximately as:

$$|\delta x(t)| \simeq e^{\lambda t}|\delta x_0| \tag{6.66}$$

where λ is the Lyapunov exponent. For a discrete multidimensional system, it can be shown [15] that the whole spectrum of Lyapunov exponents can be found with:

$$\lambda_i(\mathbf{z}_0) = lim_{k\to\infty}\frac{1}{2k}\ln|\mu_i\left(\Phi_k^T\Phi_k\right)| \tag{6.67}$$

where Φ_k is the product of the jacobians of the discrete flow from the initial point to the time step k:

$$\Phi_k = F'(\mathbf{z}_k)\cdots F'(\mathbf{z}_0) \tag{6.68}$$

and μ_i denotes the i-th eigenvalue.

The maximum value of the Lyapunov exponents, called the **maximal Lyapunov exponent** (MLE), is useful to study the stability or the chaotic behaviour of a system. In particular, if the MLE is negative or zero, the system is stable, and it is conservative if the MLE is zero. A positive MLE is an indication of a chaotic system, where two initial nearby points will diverge to an arbitrary separation, no matter how close they are.

Numerical results The Lyapunov exponents of the stable versions of our integrator have been analyzed with the above method with different initial conditions, and all of them have proved to be Lyapunov stable. Figure 52 shows the MLE for the tokamak configuration B, tested with the explicit scheme 4.

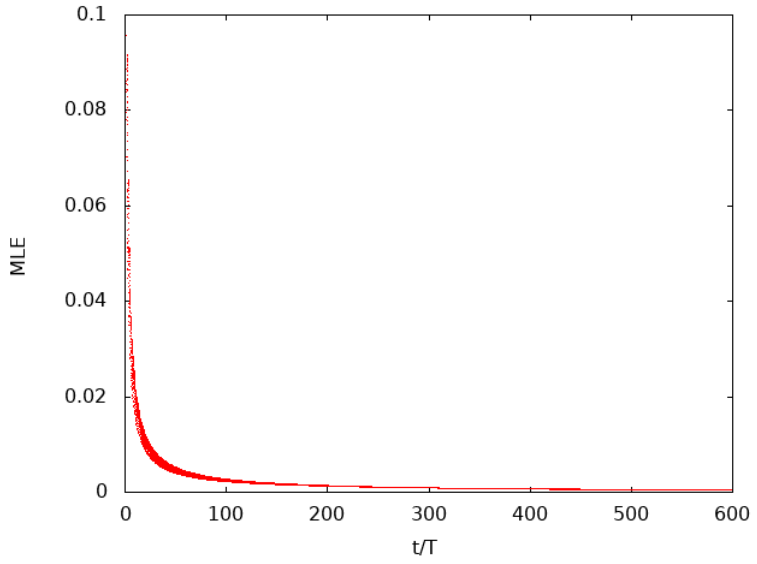


Figure 52: Maximal Lyapunov Exponent for the Explicit Scheme 4 (Section 6.4.7) and magnetic configuration B.

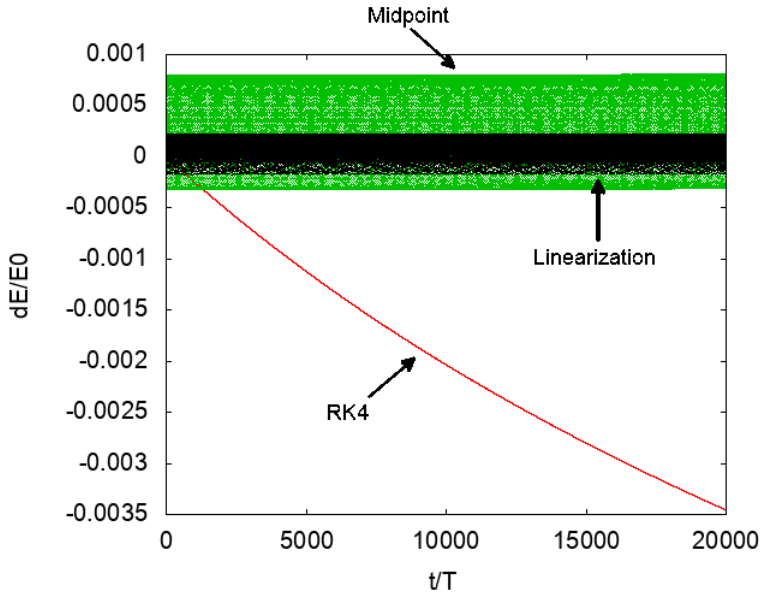
6.5.2 Computational Costs

Since the implicit integrators require generally two to four Newton iterations for the convergence plus an explicit first guess and one iteration requires one evaluation of the magnetic fields, we expect that these integrators should be slightly more costly, but still comparable, to RK4, which requires four field evalutions.
In particular, the 3D version of the midpoint rule should be slightly faster than the 4D version since we have one less degree of freedom.
The explicit integrators are the most efficient ones. By comparing eqs. (6.52) and

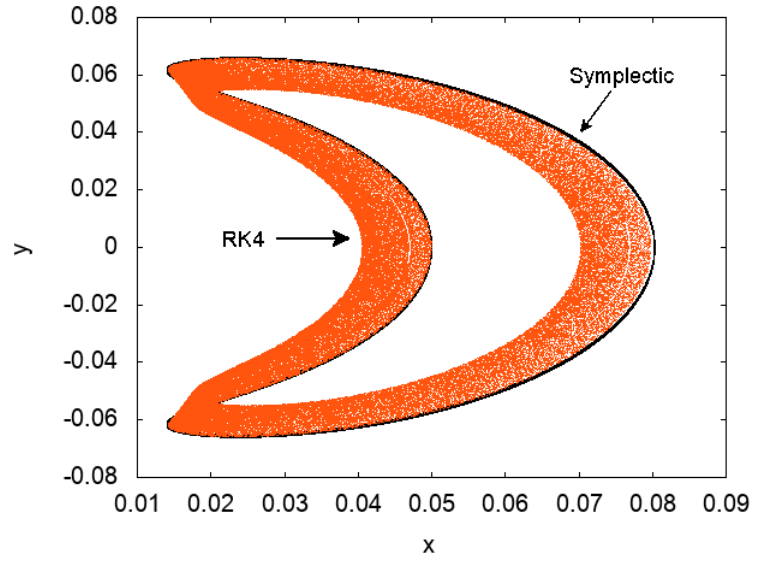
(6.4), we recognize at once that an explicit integrator is equivalent to an Euler method. A typical simulation of a tokamak like magnetic field for 2×10^4 orbits was performed by the integrator in about 10 to 20 seconds in an old notebook.

6.5.3 Source Code

All the numerical simulations in this thesis are performed with a code written in C language. The source code can be found in internet at `https://github.com/m3l7/Symplectic-Guiding-Center`.



(a)



(b)

Figure 53: Comparison between RK4 and the variational symplectic integrators for the tokamak field configuration B. (a) is the energy error, (b) is the particle trajectory in the poloidal (x,y) plane.

7 Summary and Outlook

This Thesis dealt with the problem of numerical integration of Hamiltonian systems, in particular for the motion of the guiding center of charged particles in electromagnetic fields, for which the different time scales involved require simulations with high time steps and long integration times.

Symplectic integrators are well known since some decades and their ability to conserve the geometry of an Hamiltonian system for indefinitely long times is a crucial feature that fits well for our purposes. Under this regard, this work explored a new technique, based on variational integrators, for the integration of non canonical Hamiltonian systems.

Every Hamiltonian system can be described by a variational principle, or equivalently by an extended, or phase-space, Lagrangian. The first order Euler-Lagrange equations which arises from this Lagrangian are exactly the Hamilton's equations. The basic idea underlying this thesis is to apply a variational symplectic integrator to these phase-space Lagrangians.

The standard theory of variational integrators can't be applied in this context: a variational integrator behaves always as a regular Lagrangian system rather than a phase-space one. As a consequence, the degrees of freedom of a discretized phase-space Lagrangian are twice the ones of the original continuous system and the discrete flow is proven to conserve an extended canonical structure, different from the original non canonical one.
For this reason, the role of choosing the initial additional variables is important: different initializations can lead to very different discrete flows.

Some properties of these integrators have been shown: altough the discrete trajectory is not a near the identity map, it is splitted in two separate smooth parts. Also, for simple cases, namely when the symplectic structure is constant in space, it was shown that the integrator is constrained, it is not splitted and its projection is a symplectic integrator of the original Hamiltonian system.

The variational symplectic integrator was then applied to the non canonical guiding center theory. The benefit of this theory over the canonical one is that simple global coordinates are guaranteed to exist, without the need of patches.
Different versions of the variational integrator, both implicit and linearized, were proposed and all of them were tested with three configurations: a uniform magnetic field, a tokamak configuration and a force-free magnetic field.

The experimental results are quite satisfactory for many integrators that were tested. In particular, the implicit versions, for a specific choice of initial points, which we called the Hamiltonian initialization, bounds the energy for indefinitely long times and conserve exactly the quantities associated to Noether symmetries.
A reduced three dimensional implicit version was proposed. It is based on the idea that the velocity can be solved analitically separately from the other variables.

On the other hand, the linearizations need to be initialized in a different way, called the Lagrangian initialization. Altough they are very fast, they suffer from numerical instabilities which make them hard to use in a practical scenario. However, two new linearizations, which are apparently free from these instabilities, at least with the magnetic field configurations we studied, were proposed.

In conclusion, we proposed four candidates of possible integrators: the implicit scheme 1 (section 6.4.1) which bounds the energy and conserves the Noether symmetries, the explicit scheme 4 (6.4.7) which bounds both the energy and the Noether symmetries and it is as fast as an Euler method and two reduced integrators, the implicit scheme 2 (6.4.2) and the explicit scheme 3 (6.4.6) which can work only in the guiding center context since they exploit the reuction of the system by solving independently the velocity with respect to the other spatial variables.

There are some open issues left. First, the theoretical results found for the constant symplectic case should be extended to the most general situation: it would be interesting to find conserved quantities, at least with a suitable initialization, for which the discrete flow can be projected in the original continous manifold. Even though this won't be possible, it should be possible to understand which is the best initialization for a given problem and a given discretization.
Also, this work dealt only with the midpoint rule discretization: an analysis with higher order discretization schemes, such as Runge-Kutta symplectic integrators, is needed.

Last, it would be interesting to obtain a better understanding of the stable linearizations proposed, both from a numerical point of view, since these versions were tested in a quite limited number of magnetic field configurations, and from a theoretical persective.

Appendices

A Continuous Mechanics

A.1 Introduction

In this appendix, some basic concepts of classical mechanics, both from Lagrangian and Hamiltonian point of view, are studied. Particular attention is paid to the geometric formulation of these two theories, since our algorithms are based on the direct discretizations of the geometrical objects underlying the theories of hamiltonian and Lagrangian mechanics.

Lagrangian mechanics was introduced in the late '700 by Euler and Lagrange. One of the main reasons that led to the formulation of this theory was to seek for equations of motion invariant under coordinate transformations, so that arbitrary coordinates could be introduced without loosing the formal results of the theory. These equations are now known as the Euler-Lagrange equations.
In the early '800 Hamilton extended the theory by noting that the Euler-Lagrange equations were equivalent to a variational principle, which is now called the Hamilton's principle.

The Lagrangian theory constitutes the main building block of this thesis, in particular the integrators we deal with in chapters 5 and 6 are based on the discretization of the variational principle applied to the guiding center theory.

However, the Lagrangian setting, as it is generally known, is not sufficient to describe our problem, and we need a generalization of it, which is given by the Hamiltonian mechanics or by the the theory of phase-space Lagrangians.

This section is organized as follows: first, the geometrical mathematical foundations of Lagrangian and Hamitonian mechanics are given.
Then, the Lagrangian and the Hamiltonian theories along with their variational principles are presented from a geometrical point of view.

The notions of this section can be easily found in standard mechanics textbooks, for example [18] and [16].

A.2 Geometric Foundations of Mechanics

In this section some basic concepts of differential geometry are given, in particular the theory of manifold and differential forms, which will be useful for the comprehension of the following sections, when a geometric formulation of Lagrangian and Hamiltonian systems will be developed.
The aim of this chapter is to focus the reader on the geometric structures that characterize a dynamical system. These structures constitute the starting point of the geometric, or symplectic, integrators, since their exact conservation is one of the main reason of their exceptional good behaviour, in particular of the long term energy conservation.
For reasons of simplicity, we will restrict to time independent, finite dimensional systems.

A.2.1 Manifolds

Intuitively, a **manifold** of dimension n is a set M such that the neighbourhood of every point looks like a subset of $\mathbb{R}^n$. More precisely, for every point $m \in M$, there is an open neighbourhood, U_m, homeomorphic to $\mathbb{R}^n$, namely with a continuous invertible map $\phi_m : M \to \mathbb{R}^n$.
Note that the map ϕ_m defines local coordinates on M and we can denote them with $(x^1, ..., x^n) = \phi_m(x)$. The set (U_m, ϕ_m) is said to be a chart on M. For simple manifolds, like the euclidean space, one chart is sufficient to get a global parametrization, but of course this is not true for more complex manifolds, for example the sphere S^2.
Our aim is to find a way to define functions and derivatives of functions globally on a manifold.
Two charts $(U_1, \phi_1), (U_2, \phi_2)$ are **compatible** if their transition function, defined on the intersection $U_1 \cap U_2$, is C^∞:

$$g_{12} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2) \tag{A.1a}$$
$$g_{12} = \phi_2 \circ \phi_1^{-1} \tag{A.1b}$$

A manifold M is said to be **smooth**, or **differentiable** if it is covered by a collection of compatible charts (U_i, ϕ_i): $M = \bigcup_i U_i$.
Hence, given a smooth manifold, there is a way to assign coordinates to every point and we can change the coordinate system whenever the intersection of two charts is non empty simply by using the transition functions g_{ij}. We can then define what is the meaning of differentiable.

A function on a smooth manifold $f : M \to \mathbb{R}$ is **differentiable**, or smooth, if the function $f \circ \phi^{-1}$ is C^∞.
Similarly, a function between two manifold $f : M \to N$ is differentiable if given

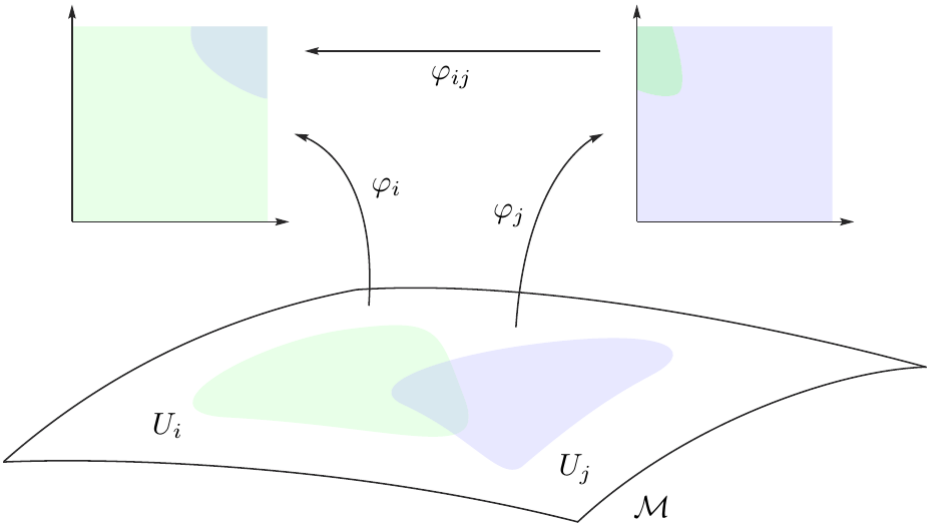


Figure 54: Charts on a manifold and their transition function

two charts on M and N, the function $\phi \circ f \circ \psi^{-1}$ is C^∞. Furthermore, f is a **diffeomorphism** if it is invertible and both f and its inverse are differentiable.

A.2.2 Tangent vectors and Tangent space

In the theory of dynamical systems, we will have to deal with vectors and curves and we will have to perform derivatives. Hence, we need a way to define rigorously these object in a manifold.

Let's consider two curves on a manifold M:

$$\gamma_1, \gamma_2 : \mathbb{R} \to M \tag{A.2}$$

such that they map to the same element of the manifold at $t = 0$. We say that the the two curves are equivalent if their velocity at $t = 0$ is the same:

$$\gamma_1(0) = \gamma_2(0) \tag{A.3a}$$

$$\frac{d}{dt}(\phi \circ \gamma_1)(0) = \frac{d}{dt}(\phi \circ \gamma_2)(0) \tag{A.3b}$$

The set of all curves passing on a point $m \in M$, with the equivalence relation (A.3b), is a vector space of dimension $dim(M)$, and it's called the **tangent space** $T_m M$.

Hence, a vector $v \in T_m M$ can be imagined as the set of curves which pass through m and has the same direction in m. It can be then represented by coordinates in $\mathbb{R}^n$ using the components of the derivative:

$$v^i = \left.\frac{dx^i}{dt}\right|_{t=0} = \left.\frac{d}{dt}(\phi \circ \gamma)^i\right|_{t=0} \tag{A.4}$$

The **tangent bundle** TM is the union of the tangent spaces of the manifold M:

$$TM = \bigcup_m T_m M \tag{A.5}$$

An element of the tangent bundle, which has dimension $2n$, is a point of M together with a vector in its tangent space. Hence, we can find local coordinates as:

$$(x^1, ..., x^n, v^1, ..., v^n) \in TM \tag{A.6}$$

A.2.3 Cotangent space

There is an alternative definition of the tangent space that reveals a connection between tangent vectors and functions defined on a manifold.
Let $f : M \to \mathbb{R}$ be a differentiable scalar function on a manifold. The partial derivatives $\frac{\partial}{\partial x^i}$ are linear operators defined as:

$$\left(\frac{\partial}{\partial x^i}\right) f = \frac{\partial(f \circ \phi^{-1})}{\partial x^i} \tag{A.7}$$

It can be proved that the following set

$$(e_1, ..., e_n) = \left(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}\right) \tag{A.8}$$

is a basis of a vector space, the tangent bundle, such that every element $v \in T_m M$ can be uniquely written as:

$$v = \sum_i v^i \frac{\partial}{\partial x^i} \tag{A.9}$$

If a vector of the tangent space is applied to a function, we obtain the **directional derivative**:

$$v(f) = \sum_i \frac{\partial(f \circ \phi^{-1})}{\partial x^i} v^i \tag{A.10}$$

Hence, the relation (A.9) is intuitively equivalent to the fact that specifying the directional derivatives completely determines a vector.

The **differential** of a function $f : M \to \mathbb{R}$ is a linear map $df : T_mM \to \mathbb{R}$. Hence, it is an element of the dual of the tangent space, the **cotangent space**: $df \in T_m^*M$. It assigns a vector to the directional derivative of the function:

$$df(v) \equiv v(f) = \sum_i \frac{\partial(f \circ \phi^{-1})}{\partial x^i} v^i \tag{A.11}$$

Identifying the function that gives the i-th coordinate of an element of a manifold as x^i, we get immediately the relation:

$$dx^i(\frac{\partial}{\partial x^j}) = \frac{\partial}{\partial x^j} x^i = \delta_{ij} \tag{A.12}$$

This means that the dual base of $\left(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}\right)$ is $(dx^1, ..., dx^n)$, which is a base for the cotangent space T_m^*M. Thus, every differential, or covector, can be written as:

$$df = \frac{\partial f}{\partial x^i} dx^i \tag{A.13}$$

As before, we can take the disjoint union of the cotangent spaces to obtain the cotangent bundle:

$$T^*M = \bigcup_m T_m^*M \tag{A.14}$$

In a (co)tangent bundle, there are two functions that are implicitly defined: the **natural projection** $\tau_M : TM \to M$, which gives the point m associated to an element of the (co)tangent bundle, and its inverse, called the **fiber**, which assigns to a point its tangent space.

A.2.4 Lifts and vector fields

We can generalize the notion of the differential to functions between manifolds: given a smooth function $f : M \to N$, the derivative is the function $T_mf : T_mM \to T_{f(m)}N$:

$$T_mf(v) = \frac{d}{dt} f(\gamma(t))\Big|_{t=0} \tag{A.15}$$

where $\gamma(t)$ is a curve that represents the vector v. It is immediate to verify that this definition reduces to the differential when f is a scalar function.
As a particular case, we can take as function a curve on a manifold, seen as a function that assign a point in M to another point in M.

Precisely, the **tangent lift** of a curve assigns a curve in the tangent space to a curve in the manifold, in this way:

$$T\gamma(t) = \left(\gamma(t), \frac{d}{dt}\gamma(t)\right) \tag{A.16}$$

We can then define what is a **vector field**, that is, a function $X : M \to TM$ which gives a vector in the tangent space for every point in the manifold. An integral curve of a vector field is a curve γ such that its tangent lift is the vector field:

$$T\gamma(t) = \left(\gamma(t), \frac{d}{dt}\gamma(t)\right) = (\gamma(t), X(\gamma(t))) \tag{A.17}$$

Of course, for a given vector field, there are infinite integral curves that satisfy (A.17). We can get a particular curve by specifying the initial condition $\gamma(0) = m_0$. The **flow** $\phi_t(m) : \mathbb{R} \times M \to M$ of a vector field is the collection of integral curves that pass through m at $t = 0$. It is differentiable with respect to both m and t and it is a one-parameter group of transformations on M:

$$\phi_{t+t'} = \phi_t \circ \phi_{t'} \tag{A.18}$$

Pull back and Push forward Given a vector field on a manifold, there is a natural way to induce a vector field on a different manifold, when a diffeomorphism between the two manifolds is available. This method can be extended to other mathematical objects, as we will see later for differential forms.
Precisely, let M, N be two manifolds, $X : N \to TN$ a vector field and $\varphi : M \to N$ a diffeomorphism. The **pull back** of X is a vector field on M defined as:

$$(\varphi^* X)(m) = (T\varphi^{-1} \circ X \circ \varphi)(m) \tag{A.19}$$

In other words, given a point $m \in M$, get the correspondent point in N through φ, apply the vector field X and return to the manifold M with the lift of the inverse of the diffeomorphism.

The **push forward** is just the inverse of the pullback:

$$\varphi_* = (\varphi^*)^{-1} \tag{A.20}$$

A.2.5 Differential Forms

The theory of differential forms allows to implicitely define vector calculus operations, like integrals, curl and div, in a manifold, indepentently from coordinates. We have already studied a differential form, the differential defined in (A.11). This notion can be generalized as follows:

A **differential k-form** on a manifold M selects at each point $m \in M$ a skew symmetric k-multilinear function $T_mM \times ... \times T_mM \to \mathbb{R}$, that is,

$$\alpha_m(av_1 + bv_1', v_2, ..., v_k) = a\alpha_m(v_1, v_2, ..., v_k) + b\alpha_m(v_1', v_2, ..., v_k) \tag{A.21}$$

with $v_1, v_1', ..., v_k \in T_mM$, and such that whenever two arguments are interchanged, α changes its sign.
Note that the differential of a function df is a differential one-form. The contrary is not necessarily true.

Product of forms Given a k-form and a j-form, we can multiply them with the **wedge product**, which gives a $(k+j)$-form:

$$(\alpha \wedge \beta)(v_1, ..., v_{k+j}) = \frac{(k+j)!}{k!j!} A(\alpha(v_1, ..., v_k)\beta(v_{k+1}, ..., v_{k+j})) \tag{A.22}$$

where A is the alternating operator:

$$A(\alpha(v_1, ..., v_k)) = \frac{1}{k!} \sum_{\sigma} sgn(\sigma)\alpha(v_{\sigma(1)}, ..., v_{\sigma(k)}) \tag{A.23}$$

Remembering the notation used for the tangent and cotangent base, a k-form can be representend in the following way:

$$\alpha_m = \alpha_{i_i,...,i_k}(m)dx^{i_1} \wedge ... \wedge dx^{i_k} \tag{A.24}$$

where a sum over the index i_j is performed (from 1 to $dim(M)$), with the condition $i_1 < ... < i_k$.
Hence, one-forms and two-forms are written respectively as:

$$\Theta_m = \Theta_i(m)dx^i \tag{A.25a}$$

$$\omega_m = \left(\omega_{ij}(m)dx^i \wedge dx^j\right)_{(i<j)} = \frac{1}{2}\omega_{ij}(m)dx^i \wedge dx^j \tag{A.25b}$$

Note that ω is represented by an antisymmetric matrix:

$$\omega_{ij}(m) = -\omega_{ji}(m) \tag{A.26}$$

Interior Product Given a k-form and a vector $v \in T_m M$, we can build their contraction as a $(k-1)$-form in this way:

$$\mathbf{i}_v \alpha(v_2, ..., v_k) = \alpha(v, v_2, ..., v_k) \tag{A.27}$$

The same definition can be extended straightforwardly to a vector field $X : M \to TM$:

$$\mathbf{i}_X \alpha(v_2, ..., v_k) = \alpha(X(m), v_2, ..., v_k) \tag{A.28}$$

Exterior Derivative Given a k-form α written in coordinates representation (A.24), its **exterior derivative** $d\alpha$ is a $(k+1)$-form given by:

$$d\alpha = \frac{\partial \alpha_{i_1,...,i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge ... \wedge dx^{i_k} \tag{A.29}$$

One can immediately show that the exterior derivative is a linear operator and that $d^2\alpha = 0$ for every k-form α. Keeping in mind this fact, the following definitions are given:
a k-form α is **closed** if $d\alpha = 0$ and **exact** if there is a $(k-1)$-form ω such that $d\omega = \alpha$. Obviously, an exact form is always closed, but not every closed form is exact.
However, it is possible to prove that every closed form α is locally exact, so that for every point $m \in M$, there is an open neighbourhood in which $\alpha = d\omega$ for some $(k-1)$-form ω. This is just the statement of the **Poincaré lemma**.

Remembering the general representation of the differential:

$$df = \frac{\partial f}{\partial x_i} dx^i \tag{A.30}$$

we can notice that it is the exterior derivative of the 0-form f, hence it is an exact (and closed) one-form.

Pull back and Push forward The idea of pull-back and push-forward used previously for vector fields can be applied easily for differential forms:
If α is a k-form on a manifold N, its pull-back under the diffeomorphism $\varphi : M \to N$ is a k-form on M:

$$(\varphi^* \alpha)_m(v_1, ..., v_k) = \alpha_{\varphi(m)}(T_m\varphi(v_1), ..., T_m\varphi(v_k)) \tag{A.31}$$

For 0-forms $f : M \to \mathbb{R}$, the pull-back is computed as:

$$(\varphi^* f)(m) = (f \circ \varphi)(m) \tag{A.32}$$

An interesting property of pull-backs is that they commute with the exterior derivative:

$$\varphi^*(d\alpha) = d(\varphi^*\alpha) \tag{A.33}$$

As before, the push-forward is just the inverse of the pull-back:

$$\varphi_* = (\varphi^{-1})^* \tag{A.34}$$

A.2.6 The Lie Derivative

We already know how to perform derivatives of vector fields and differential forms in a manifold, namely with the exterior derivative. There is another, slightly different and perhaps more direct, way to accomplish the same task.
The basic idea is to study how a differential form or a vector field changes along the flow of a vector field.

Given a k-form α and a vector field X on a manifold M, the **Lie derivative** of α along X is a k-form defined as:

$$\mathcal{L}_X\alpha = \frac{d}{dt}\varphi_t^*\alpha\bigg|_{t=0} = \lim_{t\to 0}\frac{1}{t}\left((\varphi_t^*\alpha) - \alpha\right) \tag{A.35}$$

where φ_t^* is the pull-back under the flow of X, seen as a smooth function from M to itself.
The following important relations hold for the Lie derivative:

$$\frac{d}{dt}\varphi_t^*\alpha = \varphi_t^*\mathcal{L}_X\alpha \tag{A.36}$$

$$\mathcal{L}_X\alpha = d\mathbf{i}_X\alpha + \mathbf{i}_X d\alpha \tag{A.37}$$

In particular, the last one (A.37), which relates Lie derivatives and external derivatives, is known as **Cartan's magic formula**.
Thus, we can observe, from (A.37) or directly from (A.35), that the Lie and the external derivatives agree when a scalar function (0-form) is used and both return the directional derivatives along the vector field:

$$X(f) \equiv \mathcal{L}_X f = df(X) \tag{A.38}$$

Given two vector field X and Y, the **Poisson-Lie Bracket** is the unique vector field $[X, Y]$ such that, for every smooth scalar function f, the following equivalence holds:

$$[X, Y](f) = X(Y(f)) - Y(X(f)) \tag{A.39}$$

Intuitively, the Poisson-Lie Bracket is the directional derivative of Y in the direction generated by X.

A.2.7 Volume Forms

A manifold of dimension n is called **orientable** if there exists a differential n-form μ which is not null for every point in the manifold. In this case μ is said to be a **volume form**. A diffeomorfism $\varphi : M \to M$ is called **volume preserving** if the volume form is conserved under the pullback of φ:

$$\varphi^* \mu = \mu \tag{A.40}$$

More generally, the pullback of a k-form is a k-form, so we can write the pullback of ω as a combination of a function of φ and ω:

$$\varphi^* \mu = J(\varphi)\omega \tag{A.41}$$

Hence, $J(\varphi)$, referred as the **Jacobian** of φ, must be 1 for a volume preserving transformation.

A.3 Lagrangian Mechanics

A.3.1 Introduction

We start by giving some basic notions on the spaces we will deal with: recall that the tangent lift of a curve $\gamma : \mathbb{R} \to Q$ is:

$$T\gamma(t) = (\gamma(t), \dot{\gamma}(t)) = \left(\gamma(t), \frac{d}{dt}\gamma(t)\right) \tag{A.42}$$

Denote by $\mathcal{C}([t_0, t_1], Q)$ the spaces of C^2 curves (the **path space**) from the real interval $[t_0, t_1]$ to the manifold Q and by $\mathcal{C}(q_0, q_1, [t_0, t_1], Q)$ the path space of curves with fixed endpoints:

$$\mathcal{C}(q_0, q_1, [t_0, t_1], Q) = \left\{ \gamma \in \mathcal{C}([t_0, t_1], Q) \middle| \gamma(t_0) = q_0, \gamma(t_1) = q_1 \right\} \tag{A.43}$$

The path space is itself a smooth manifold, and we can build a tangent vector to a curve in the same way as we did in (A.4).
If $\gamma_\epsilon : \mathbb{R} \to \mathcal{C}$ is a curve in $\mathcal{C}$, such that $\gamma_0(t) = \gamma(t)$, then a tangent vector $w \in T_\gamma\mathcal{C}$ to a curve γ is:

$$w = \frac{d}{d\epsilon}\gamma_\epsilon \bigg|_{\epsilon=0} \tag{A.44}$$

Thus, for every t, $w(t)$ is a vector in the tangent space of Q: $w(t) \in T_{\gamma(t)}Q$, so that w is just a curve in TQ:

$$T_\gamma\mathcal{C}([t_0, t_1], Q) = \mathcal{C}([t_0, t_1], TQ) \tag{A.45}$$

Intuitively, if we change slightly a curve in the path space, w represents the variations of the curve. For this reason, we can write $\delta\gamma$ in place of w.
One can immediately show that if we keep the endpoints fixed, as in (A.43), the variations at the endpoints are null:

$$T_\gamma\mathcal{C}(q_0, q_1, [t_0, t_1], Q) = \left\{ w \in \mathcal{C}([t_0, t_1], TQ) \middle| w(t_0) = w(t_1) = 0 \right\} \tag{A.46}$$

We already know that the equations of motion for a mechanical system are second order differential equations. For this reason we'll have to deal with vectors in the tangent of the tangent space $T(TQ)$. Given an element (x, v) in TQ, we can build a vector in its tangent space by deriving a curve that passes through (x, v):

$$(x, v), \left(\frac{d}{dt}x, \frac{d}{dt}v\right) \in T(TQ) \tag{A.47}$$

Hence, for a generic element in $T(TQ)$ it is not guaranteed that $\frac{d}{dt}x = v$. The correct manifold is a submanifold of $T(TQ)$, the so called **second order manifold** $\ddot{Q}$:

$$\ddot{Q} = \left\{ w \in T(TQ) \middle| T\tau_Q(w) = \tau_{TQ}(w) \right\} \tag{A.48}$$

The function $T\tau_Q(w)$ takes an element $(x, v) \in TQ$, projects it to the configuration manifold Q and returns its tangent lift:

$$T\tau_Q(w) = \left(x, \frac{d}{dt}x \right) \equiv \tau_{TQ}(w) = (x, v) \tag{A.49}$$

This it just the requirement we were seeking for.
In the same way, a **second order vector field** $X : TQ \to T(TQ)$, satisfies:

$$T\tau_{TQ} \circ X = id \tag{A.50}$$

A.3.2 The variational principle and the Euler-Lagrange equations

We are now ready to study the relevant quantities of Lagrangian mechanics.
Let Q be a smooth manifold (the configuration manifold). The **Lagrangian** $L : TQ \to \mathbb{R}$ is a scalar function on the tangent bundle:

$$L(q, v) \tag{A.51}$$

The **action** $S : \mathcal{C}([t_0, t_1], Q) \to \mathbb{R}$ is defined as:

$$S(q_t) = \int_{t_0}^{t_1} L\left(q(t), \dot{q}(t)\right) dt \tag{A.52}$$

where $q(t)$ is a curve in the path space and $\dot{q}(t)$ is its tangent lift. Since the action is a scalar function on a smooth manifold, we can take its differential (the external derivative) $dS : \mathcal{C} \to T^*\mathcal{C}$, which is a directional derivative along a variation $\delta q_t \in T_{q_t}\mathcal{C}$, to give:

$$\begin{aligned} dS(q_t)\delta q_t &= \frac{d}{d\epsilon} S(q_{t,\epsilon}) \bigg|_{\epsilon=0} = \\ &= \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i} \right) \right] \cdot \delta q^i dt + \left[\frac{\partial L}{\partial \dot{q}^i} \cdot \delta q^i \right]_{t_0}^{t_1} \end{aligned} \tag{A.53}$$

where $q_{t,\epsilon}$ is a curve that represents the variation δq_t:

$$\frac{d}{d\epsilon} q_{t,\epsilon} = \delta q_t \tag{A.54}$$

and, at every time t, δq^i is the i-th component of the vector $\delta q(t) \in TQ$:

$$\delta q^i \equiv dx^i(\delta q(t)) \tag{A.55}$$

Intuitively, given a curve on the manifold, we change the curve by small variations δq retaining the endpoint fixed, and we observe how the action (eq. A.52) varies. Figure 55 illustrates the curve and its variations.

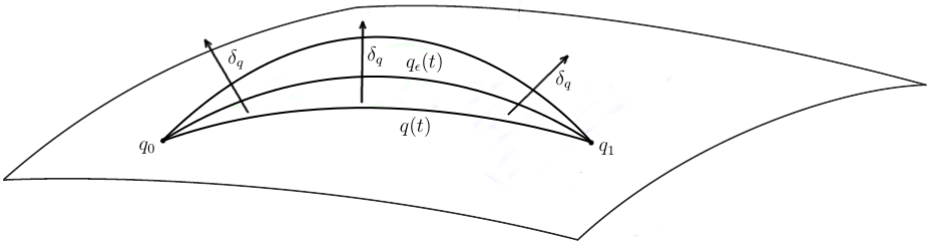


Figure 55: A curve $q(t)$ and its variation with fixed endpoints

The first and second term of equation (A.53) are referred respectively as the **Euler-Lagrange map** $D_{EL} : \ddot{Q} \to T^*Q$ and the **Lagrangian one-form** $\Theta_L : TQ \to T^*(TQ)$:

$$D_{EL} = \frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \tag{A.56}$$

$$\Theta_L = \frac{\partial L}{\partial \dot{q}^i} dq^i \tag{A.57}$$

The arguments of these maps, $\ddot{q} \in \ddot{Q}$ and $\dot{q} \in TQ$ are evaluated by lifting the curve q_t at every t.

We can now state the **Hamilton's variational principle** from a geometric point of view: let S be the action of the system restricted to the path space with endpoints fixed (A.43). $q_t \in \mathcal{C}(q_0, q_1, [t_0, t_1], Q)$ is a solution if the exterior derivative of

S vanishes for every $\delta q_t \in T_{q_t}\mathcal{C}(q_0, q_1, [t_0, t_1], Q)$:

$$dS(q_t)\delta q_t = 0 \tag{A.58}$$

From (A.53), the Lagrangian one-form is a boundary part that vanishes when variations to the action with fixed endpoints are imposed. The Hamilton's principle is then equivalent to the vanishing of the Euler-Lagrange map, for every time $t \in [t_0, t_1]$:

$$D_{EL}(\ddot{q}) = 0 \tag{A.59}$$

Equivalently, we can define the **Lagrangian vector field** as the second order vector field $X : TQ \to T(TQ)$ which satisfies the Euler-Lagrange equations:

$$D_{EL} \circ X = 0 \tag{A.60}$$

Hence, its flow $F_L : TQ \times \mathbb{R} \to TQ$ gives the state of a system following the euler-lagrange equations after a time T:

$$F_L^T(q_0, \dot{q}_0) = (q(T), \dot{q}(T)) \tag{A.61}$$

If q_t is a solution and we remove the condition on the endpoints, the first term of equation (A.53) vanishes, while the contribution from the boundary term is still not null.
It is easy to show that the Lagrangian one-form computed at t_1 is just a pull-back under the Lagrangian vector field, so that:

$$dS(q_t) = F_t^*\Theta_L(q_0, \dot{q}_0) - \Theta_L(q_0, \dot{q}_0) \tag{A.62}$$

Hence, taking another derivative of S and remembering that $d^2 = 0$ for a differential form:

$$(F_L^T)^* d\Theta_L = d\Theta_L \tag{A.63}$$

The conserved two-form $\Omega_L = -d\Theta_L$ is referred to as the **Lagrangian symplectic form**. We can write it in coordinates to give:

$$\boxed{\Omega_L(q, \dot{q}) = \frac{\partial^2 L}{\partial q^i \partial \dot{q}^i} dq^i \wedge dq^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^i} d\dot{q}^i \wedge dq^j} \tag{A.64}$$

It will be clear that for a symplectic integrator there is a discrete analogous of the Lagrangian symplectic form and the same conservation law applies as well.

More generally, we talk of a **symplectic form** $\Omega_q : T_qQ \times T_qQ \to \mathbb{R}$ on a manifold Q when it is closed:

$$\mathbf{d}\Omega_q = 0 \tag{A.65}$$

and non degenerate, such that the map $\Omega_q^\flat : T_qQ \to T_q^*Q$ is an isomorphism:

$$\Omega_q^\flat(v) = \Omega_q(v, \cdot) \tag{A.66}$$

The degeneracy condition is equivalent to the matrix associated to the 2-form, as in (A.26), being singular.
We can write the Lagrangian symplectic two-form in matrix form:

$$\begin{pmatrix} A\left(\dfrac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\right) & \dfrac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \\ -\dfrac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} & 0 \end{pmatrix}$$

where A is the alternating operator, defined in (A.23). It follows that Ω_L is non degenerate only if:

$$\det\left(\frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}}\right) \neq 0 \tag{A.67}$$

As was already stated in the first section, only in this case the Lagrangian vector field X_L is guaranteed to be uniquely determined and hence to give a solution to the system.
When condition A.67 holds, the Lagrangian L is said to be **regular**, otherwise it is said to be **singular** or **degenerate**.
Therefore, for singular Lagrangians, a second order vector field does not exist generally. However, for special singular Lagrangians (see chapter 3.2), the so called **phase-space Lagrangians**, one can find a first order vector field that represents the solution of the Euler-Lagrange equations.

Mechanical systems For classical mechanics problems, the Lagrangian usually takes the form of the kinetic energy minus the potential energy:

$$L(q, \dot{q}) = \frac{1}{2} m ||\dot{q}||^2 - V(q) \tag{A.68}$$

One can immediately show that in this case the Euler-Lagrange equations are just the Newton's second law:

$$\frac{d}{dt}(m\dot{q}) = -\frac{\partial V}{\partial q} \tag{A.69}$$

A.3.3 Legendre transform

Given a Lagrangian $L : TQ \to \mathbb{R}$, the **Legendre transform** is the function $\mathbb{F}L : TQ \to T^*Q$, such that, for two vectors $v, w \in T_qQ$:

$$\mathbb{F}L(v) \cdot w = \frac{d}{d\epsilon} L(v + \epsilon w)\bigg|_{\epsilon=0} \tag{A.70}$$

Note that T_qQ is a vector space, so it makes sense to sum two vectors and to multiply them by a scalar. The point $q \in Q$ remains unchanged.
Written in a local chart, so that $v = (q^i, \dot{q}^i) \in TQ$, the Legendre transform takes the form:

$$\mathbb{F}L(q^i, \dot{q}^i) = \left(q^i, \frac{\partial L}{\partial \dot{q}^i}\right) = (q^i, p_i) \tag{A.71}$$

where $p_i \in T_q^*Q$ are known as the **conjugate momenta**.
The Legendre transform is useful to relate the Lagrangian and Hamiltonian formalism. We can use it in the Lagrangian theory to relate the symplectic form directly to the Lagrangian vector field in what follows:
if v is a vector in the tangent bundle, $v \in TQ$, the **energy function** is the map $E : TQ \to \mathbb{R}$:

$$E(v) = \mathbb{F}L(v) \cdot v - L(v) \tag{A.72}$$

Then, it can be proved that, for all $v \in TQ$ and $w \in T_v(TQ)$, the Lagrangian vector field (such that $D_{EL} \circ X_L = 0$), the symplectic form and the energy function are related by:

$$\boxed{\mathbf{i}_{X_L}\Omega_L(v) \cdot w = \mathbf{d}E(v) \cdot w} \tag{A.73}$$

Often, equation (A.73) is used as a geometric definition of Lagrangian systems, without passing through the variational framework and the Euler-Lagrange equations: given a regular Lagrangian L, there is a unique second-order vector field X_L, which satisfies eq. (A.73).
From (A.73), it is easy to show that the energy function is conserved along a solution, yielding the **conservation of the energy**:

$$F_L^* E = E \tag{A.74}$$

A.3.4 Symmetries and Noether's theorem

The study of simmetries has been a main research topic in theoretical physics since the last century. One of the most important results, the Noether theorem constitutes the basic tools for modern physics theories. Noether theorem, derived by

Emmy Noether in 1915, states that for every symmetry of a Lagrangian system, there is an associated conservation law.

Consider a Lie group, that is, a manifold M with a group structure and such that the group multiplication:

$$\mu : G \times G \to G \tag{A.75}$$

is C^∞. The left translation map $L_g : G \to G$ is defined as:

$$L_g(h) = gh \tag{A.76}$$

The Lie algebra of G, $\mathfrak{g}$, is the vector space T_eG, together with a Lie Bracket structure (see section A.2.6), that is, given two elements $\xi, \eta \in T_eG$:

$$[\xi, \eta] = [X_\xi, X_\eta](e) \tag{A.77}$$

where X_ξ is a vector field defined as:

$$X_\xi(g) = T_eL_g(\xi) \tag{A.78}$$

Now, consider a Lie Group G which acts on Q by an action $\Phi : G \times Q \to Q$ and on TQ by its tangent lift, defined as:

$$\Phi_g^{TG}(q, \dot{q}) = \left(\Phi_g(q), \frac{\partial \Phi_g}{\partial q}(q)\dot{q}\right) \tag{A.79}$$

The infinitesimal generator $\xi_{TQ} : TQ \to T(TQ)$ on TQ corresponding to a Lie algebra element $\xi \in \mathfrak{g}$ is

$$\xi_{TQ}(q, \dot{q}) = \frac{d}{dg}\left(\Phi_g^{TQ}(q, \dot{q})\right) \cdot \xi \tag{A.80}$$

If a Lagrangian system is invariant under the lift of the action Φ:

$$L \circ \Phi_g^{TQ} = L \tag{A.81}$$

then the momentum map $J_L : TQ \to \mathfrak{g}^*$, defined as:

$$J_L(q, \dot{q}) \cdot \xi = \Theta_L(q, \dot{q}) \cdot \xi_{TQ}(q, \dot{q}) \tag{A.82}$$

is conserved by the Lagrangian flow:

$$J_L \circ F_L^T = J_L \tag{A.83}$$

Example Perhaps, the simplest example of the Noether theorem is the case of a Lagrangian that has no dependence on a variable q^i. The Lagrangian is then invariant under the lift of the action ϕ_α with correspondent generator ξ:

$$\phi_\alpha(q) = q + \alpha e_i \tag{A.84a}$$
$$\xi(q) = e_i \tag{A.84b}$$

which is a translation along the axis e_i. The quantity $p_i(q, \dot{q})$, referred as the conjugate momentum of q^i, is conserved:

$$p_i(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}^i} \tag{A.85}$$

For mechanical systems (A.68), $p_i = m\dot{q}^i$ is just the linear momentum of the particle.

A.4 Hamiltonian mechanics

The basic idea of Hamiltonian mechanics is to study a mechanical system using the coordinates q and their conjugate momenta p, rather than the velocities $\dot{q}$. This can lead to advantages in some situations.
The Hamiltonian formulation admits a wider range of coordinate transformations, leading for example to the study of non canonical systems. This point can be treated in some extent in the Lagrangian side by extending the space of coordinates, giving rise to the so called phase-space Lagrangian formalism. (see chapter 3)
Second, Hamiltonian systems possess a direct and simple geometric interpretation, thus studying some of their properties is easier in this context.

Recall from section A.2 that a symplectic form Ω is a 2-form which is closed ($\mathbf{d}\Omega = 0$) and non degenerate, such that $\Omega^\flat$ is an isomorhpism.
A smooth manifold M, together with a symplectic form (M, Ω) is called a **symplectic manifold**.

Given a symplectic manifold (M, Ω_H), we talk of a **Hamiltonian system** if there is a vector field $X_H : M \to TM$ and a function $H : M \to \mathbb{R}$, such that:

$$\boxed{\mathbf{i}_{X_H}\Omega_H = \mathbf{d}H} \tag{A.86}$$

X_H and H are referred respectively as the **Hamiltonian vector field** and the **Hamiltonian**.
Note that equation (A.86) is similar to equation (A.73). We will see that the two are strictly related for canonical systems. Just as in that case, the Hamiltonian function H represents the energy of the system.
By definition, the evolution of the system satisfies:

$$\dot{z} = X_H(z) \tag{A.87}$$

Writing equation (A.86) in coordinates, we find:

$$\dot{z}^i = X_H(z)^i = -(\Omega_H^{-1})^{ij}\frac{\partial H}{\partial z^j} \tag{A.88}$$

where Ω_H^{ij} is the matrix of the 2-form Ω_H.
From Cartan's formula (A.37) and (A.35), and from the fact that both Ω_H and $\mathbf{d}H$ are closed forms, we get:

$$\frac{d}{dt}\varphi_t^*\Omega_H = \varphi_t^*\mathcal{L}_{X_H}\Omega_H = \varphi_t^*\left(\mathbf{i}_{X_H}\mathbf{d}\Omega_H + \mathbf{d}\mathbf{i}_{X_H}\Omega_H\right) = 0 \tag{A.89}$$

where φ_t is the flow of the Hamiltonian vector field. Hence, we have proved that the Hamiltonian symplectic form is conserved along the flow:

$$\varphi_t^* \Omega_H = \Omega_H \tag{A.90}$$

Similarly, if X is a vector field whose flow is symplectic (eq. A.90), we can rewrite eq. (A.89):

$$0 = \frac{d}{dt} \varphi_t^* \Omega_H = \mathbf{d} \mathbf{i}_X \Omega_H \tag{A.91}$$

Therefore, from Poincarè lemma (section A.2.5), there exists locally a function H which satisfies

$$\mathbf{i}_X \Omega_H = \mathbf{d} H \tag{A.92}$$

In conclusion, the flow of an Hamiltonian vector field is symplectic, and a symplectic map is (locally) Hamiltonian.

A.4.1 Poisson brackets and tensor

Given two functions $F, G : M \to \mathbb{R}$, the **Poisson brackets** and the **Poisson tensor**, which associates a point $z \in M$ to a contravariant skew-symmetric 2-tensor $B : T_z^* M \times T_z^* M \to \mathbb{R}$, are defined by:

$$\{F, G\}(z) = B_z(\mathbf{d}F_z, \mathbf{d}G_z) = \Omega_z(X_F(z), X_G(z)) \tag{A.93}$$

where X_F and X_G are the Hamiltonian vector fields associated to F and G. The evolution of a function is related to the Poisson brackets by:

$$\frac{d}{dt}(F \circ \varphi_t) = \{F, H\} \circ \varphi_t \tag{A.94}$$

so that, using the Hamiltonian, we get the **conservation of energy** by skew-simmetry:

$$\frac{d}{dt}(H \circ \varphi_t) = \{H, H\} \circ \varphi_t = 0 \tag{A.95}$$

As we did for differential forms, we can express the Poisson tensor in coordinates, using a basis of the tangent space and an antisymmetric matrix:

$$B_z = B^{ij}(z) \frac{\partial}{\partial z^i} \frac{\partial}{\partial z^j} \tag{A.96}$$

Hence, the evolution of the system, in terms of the Poisson tensor, has the following coordinate representation:

$$\dot{z}^i = \{z^i, H\} = B_z(\mathbf{d}z^i, \mathbf{d}H) = B^{ij}(z)\frac{\partial H}{\partial z^j} \tag{A.97}$$

Comparing this with equation (A.88), we see that the matrices of the Poisson tensor and the Hamiltonian symplectic form are related by:

$$B^{ij} = -(\Omega^{-1})_{ij} \tag{A.98}$$

A.4.2 Canonical systems

Let the manifold M of the previous section be the cotangent bundle T^*Q of some smooth manifold Q. Note that the dimension of $M = T^*Q$ is $2\dim(Q)$. Denoting by (q^i, p_i) local coordinates on T^*Q, the following (locally) constant 2-form $\Omega_H : T(T^*Q) \times T(T^*Q) \to \mathbb{R}$ is obviously a symplectic form:

$$\Omega_H = \sum_i dq^i \wedge dp_i \tag{A.99}$$

Hence, the symplectic form Ω_H, together with a function $H : T^*Q \to \mathbb{R}$ is an hamiltonian system, and it is called an **hamiltonian canonical system**.
The matrix associated with (A.99) is usually denoted by $\mathbb{J}$:

$$\mathbb{J} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$$

The canonical symplectic form is exact, since it is the external derivative of the following canonical one-form $\Theta_H : T(T^*Q) \to \mathbb{R}$:

$$\Theta_H = \sum_i p_i dq^i \tag{A.100}$$

$$\Omega_H = -\mathbf{d}\Theta_H \tag{A.101}$$

If H is the Hamiltonian of a canonical system, we can use formula (A.88) to obtain the equation of motion in coordinates:

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \tag{A.102a}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q^i} \tag{A.102b}$$

which are known as the **Hamilton's equations**

One can wonder if it is possible to choose canonical coordinates for every manifold. The answer is given by the **Darboux-Lie theorem**, which states that for every symplectic manifold (M, Ω), there are local coordinates at every point, such that $m = (q^i, p_i) \in M$, and $\Omega = \Omega_H$.
This is equivalent to say that a symplectic manifold looks like T^*Q for some smooth manifold Q in the neihbourhood of every point. Note that this also means that the dimension of a symplectic manifold must always be even.

A.4.3 Liouville Theorem

One of the fundamental characteristics of Hamiltonian systems is that there is a volume form, called the Liouville measure, that is conserved by the Hamiltonian flow:

$$\omega = \Omega \wedge ... \wedge \Omega \tag{A.103}$$

$$\varphi_t^* \omega = \omega \tag{A.104}$$

where the wedge products in (A.103) are performed n times. When canonical coordinates are available, the Liouville measure takes the form:

$$\omega = dq^i \wedge ... \wedge dq^n \wedge dp_i \wedge ... \wedge dp_n \tag{A.105}$$

This is of great importance, for example, in statistical mechanics, so that the Liouville theorem is just the statement that the distribution function in the phase-space T^*Q is constant along a solution.

A.4.4 Correspondence between Hamiltonian and Lagrangian mechanics

We have already seen that a regular Lagrangian system, for which the regularity condition (A.67) holds and a second order vector field X_L exists, can be written as:

$$\mathbf{i}_{X_L} \Omega_L(v) = \mathbf{d}E(v) \tag{A.106}$$

Hence, every regular Lagrangian system is also Hamiltonian, defined on the symplectic manifold (TQ, Ω_L).
Recall that the Legendre transform maps the tangent bundle to the cotangent bundle: $\mathbb{F}L : TQ \to T^*Q$ and that we can write it in coordinates as:

$$\mathbb{F}L(q^i, \dot{q}^i) = \left(q^i, \frac{\partial L}{\partial \dot{q}^i} \right) = (q^i, p_i) \tag{A.107}$$

Using the legendre transform, we can find canonical coordinates given a Lagrangian system. The symplectic forms and the vector fields are related just by pull-back

under the Legendre transform:

$$\Theta_L = (\mathbb{F}L^*)\Theta_H \tag{A.108}$$
$$\Omega_L = (\mathbb{F}L^*)\Omega_H \tag{A.109}$$
$$X_L = (\mathbb{F}L^*)X_H \tag{A.110}$$
$$F_L^t = (\mathbb{F}L)^{-1} \circ F_H^t \circ (\mathbb{F}L) \tag{A.111}$$

so that we can compute the Lagrangian flow by going to the cotangent bundle, computing the Hamiltonian flow and returning back to the tangent bundle. Finally, we can obtain the Hamiltonian from the Lagrangian with:

$$H(q,p) = \mathbb{F}L(q,v) \cdot v - L(q,v) \tag{A.112}$$

Writing eq. (A.112) in coordinates, we find:

$$H(q,p) = \sum_j p_j \dot{q}^j(q,p) - L(q, \dot{q}(q,p)) \tag{A.113}$$

Note that with singular Lagrangians, $\dot{q}$ can't be inverted with respect to (q,p), thus in those cases the last equation doesn't make sense.

A.4.5 Generating functions

Given a canonical Hamiltonian system, we can relate it to the Lagrangian formalism with a different approach.
Fixing a time t, the Hamiltonian flow $F_H^t : T^*Q \to T^*Q$ is a symplectic transformation, since it conserves the canonical symplectic form, hence we can write:

$$(F_t^t)^*\Omega_H - \Omega_H = \mathbf{d}(\Theta_H - (F_t^t)^*\Theta_H) = 0 \tag{A.114}$$

Therefore, by Poincarè lemma, there exists locally a funtion S, the **generating function** of the flow F_H^t, such that:

$$\mathbf{d}S = \Theta_H - (F_t^t)^*\Theta_H \tag{A.115}$$

Denoting by (q_0, p_0) and by (q_1, p_1) the coordinates respectively on T^*Q and on T^*Q, transformed by F_H^t (i.e. $(q_1, p_1) = F_H^t(q_0, p_0)$), we can write equation (A.115) in coordinates (q_0, q_1):

$$p_0 dq_0 - p_1 dq_1 = \frac{\partial S}{\partial q_0} dq_0 + \frac{\partial S}{\partial q_1} dq_1 \tag{A.116}$$

thus:

$$p_0 = \frac{\partial S}{\partial q_0} \tag{A.117}$$

$$p_1 = -\frac{\partial S}{\partial q_1} \tag{A.118}$$

This means that if we know a generating function $S(q_0, q_1)$ of the Hamiltonian flow, we can find (p_0, p_1) from (q_0, q_1) by equations (A.117). We can generalize this notions as follows:

The graph of F_H^t is a subset of $T^*Q \times T^*Q$. We can extend the symplectic 1-form to a 1-form on $T^*Q \times T^*Q$:

$$\hat{\Theta} = \tau_1^* \Theta_H - \tau_2^* \Theta_H \tag{A.119}$$

where τ is the projection from $T^*Q \times T^*Q$ to the first and second cotangent bundle respectively. Then, if i is the inclusion map, from the graph of F_H^t to $T^*Q \times T^*Q$, the analogous of equation (A.114) is:

$$\mathbf{d}(i^* \hat{\Theta}) = 0 \tag{A.120}$$

Hence, there is a function $S : (T^*Q \times T^*Q) \to \mathbb{R}$, such that:

$$i^* \hat{\Theta} = d\mathbf{S} \tag{A.121}$$

Note that, being S defined on $T^*Q \times T^*Q$, we can choose any local coordinates. The choice $S(q_0, q_1)$ gives equations (A.117). Generally, choosing coordinates (q_0, p_1) and (q_1, p_0), a slightly different but of course equivalent definition of the generating functions, S^1 and S^2, is given: (see [11] for further reference)

$$q_1 \cdot dp_1 + p_0 \cdot dq_0 = \mathbf{d}(p_1 \cdot q_0 + S^1(q_0, p_1)) \tag{A.122a}$$

$$p_1 \cdot dq_1 + q_0 \cdot dp_0 = \mathbf{d}(p_0 \cdot q_1 - S^2(q_1, p_0)) \tag{A.122b}$$

Comparing equations (A.122) and (A.116), we can relate the three generating functions as follows:

$$S^1 = q_1 \cdot p_1 - p_1 \cdot q_0 - S = p_1 \cdot (q_1 - q_0) - S \tag{A.123a}$$

$$S^2 = p_0 \cdot q_1 - p_0 \cdot q_0 - S = p_0 \cdot (q_1 - q_0) - S \tag{A.123b}$$

Note that knowing explicitly a generating function from another one is not feasible in practice, since the knowledge of the Hamiltonian flow is needed.

The generating functions and the Hamiltonian of the system are related through the **Hamilton-Jacobi equation**:

$$H\left(q_0, \frac{\partial S}{\partial q_0}(q_0, q_1, t)\right) + \frac{\partial S}{\partial t}(q_0, q_1, t) = 0 \tag{A.124}$$

A similar equation holds for S^1:

$$H\left(q_0 + \frac{\partial S^1}{\partial p_1}(q_0, p_1, t), p_1\right) - \frac{\partial S^1}{\partial t}(q_0, p_1, t) = 0 \tag{A.125}$$

Knowing a generating function from the Hamiltonian, and vice versa, is difficult for the same reason as before. Jacobi found out that the action of the system, expressed in coordinates (q_0, q_1), is a solution, the **Jacobi's solution**, to the Hamilton-Jacobi equation:

$$S(q_0, q_1, t) = \int_{q_0}^{q_1} L(q(\tau), \dot{q}(\tau))d\tau \tag{A.126}$$

where $q(t)$ is a solution to the Lagrangian system which passes through q_0 and q_1: $q(0) = q_0, q(t) = q_1$.
Eq.(A.126) is very important for our purposes, since this thesis deals with numerical integrators whose flow are generated by an approximation of the Jacobi's solution.

References

[1] Giancarlo Benettin and Antonio Giorgilli. On the hamiltonian interpolation of near-to-the identity symplectic mappings with application to symplectic integration algorithms. *Journal of Statistical Physics*, 74(5-6):1117–1143, 1994.

[2] Peter G. Bergmann. Hamilton-jacobi theory with mixed constraints. *Transactions of the New York Academy of Sciences*, 33(1 Series II):108–115, 1971.

[3] James A. Cadzow. Discrete calculus of variations. *International Journal of Control*, 11(3):393–407, 1970.

[4] J. R. Cary and A. J. Brizard. Hamiltonian theory of guiding-center motion. *Reviews of Modern Physics*, 81(3):693–738, 2009.

[5] S. Chandrasekhar. On force-free magnetic fields. *Proceedings of the National Academy of Sciences of the United States of America*, 42(1):1–5, 1956.

[6] M. de León, J. C. Marrero, D. M. de Diego, and M. Vaquero. On the Hamilton-Jacobi theory for singular lagrangian systems. *Journal of Mathematical Physics*, 54(3):032902, 2013.

[7] R. de Vogelaere. Methods of integration which preserve the contact transformation property of the hamiltonian equations. *Report No. 4, Dept. Math., Univ. of Notre Dame, Notre Dame*, 1956.

[8] E. Forest and R. D. Ruth. Fourth-order symplectic integration. *Phys. D*, 43(1):105–117, 1990.

[9] H. Goldstein, C.P. Poole, and J.L. Safko. *Classical Mechanics, 3e.* Addison-Wesley Longman, Incorporated, 2002.

[10] Ernst Hairer. Backward analysis of numerical integrators and symplectic methods. In *Stiff and Differential-Algebraic Problems*. Springer-Verlag, 1994.

[11] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31. Springer, 2003.

[12] Sameer M Jalnapurkar, Melvin Leok, Jerrold E Marsden, and Matthew West. Discrete routh reduction. *Journal of Physics A: Mathematical and General*, 39(19):5521, 2006.

[13] B. Karasözen. Poisson integrators. *Mathematical and Computer Modelling*, 40(1112):1225–1244, 2004.

[14] S Lall and M West. Discrete variational hamiltonian mechanics. *Journal of Physics A: Mathematical and General*, 39(19):5509, 2006.

[15] Changpin Li and Guanrong Chen. Estimating the lyapunov exponents of discrete systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 14(2):343–346, 2004.

[16] Paulette Libermann and Charles-Michel Marle. *Symplectic geometry and analytical mechanics*, volume 35. Springer, 1987.

[17] Robert G. Littlejohn. Variational principles of guiding centre motion. *Journal of Plasma Physics*, 29(1):111–125, 1983.

[18] Jerrold E Marsden and Tudor S Ratiu. *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*, volume 17. Springer, 1999.

[19] Jerrold E Marsden and Matthew West. Discrete mechanics and variational integrators. *Acta Numerica*, 10:357–514, 2001.

[20] Hong Qin and Xiaoyin Guan. Variational symplectic integrator for long-time simulations of the guiding-center motion of charged particles in general magnetic fields. *Physical review letters*, 100(3):035006, 2008.

[21] Hong Qin, Xiaoyin Guan, and William M Tang. Variational symplectic algorithm for guiding center dynamics and its application in tokamak geometry. *Physics of Plasmas*, 16(4):042510, 2009.

[22] Sebastian Reich. Backward error analysis for numerical integrators. *SIAM Journal on Numerical Analysis*, 36(5):1549–1570, 1996.

[23] Clarence W Rowley and Jerrold E Marsden. Variational integrators for degenerate lagrangians, with application to point vortices. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 2, pages 1521–1527. IEEE, 2002.

[24] R. D. Ruth. A Canonical Integration Technique. *IEEE Transactions on Nuclear Science*, 30(4):2669, 1983.

[25] J. Squire, H. Qin, and W. M. Tang. Gauge properties of the guiding center variational symplectic integrator. *Physics of Plasmas*, 19(5):052501, 2012.

[26] Namikawa T. Magneto-hydrodynamic oscillations of a conducting liquid mass rotating in a uniform magnetic field. *Journal of Geomagnetism and Geoelectricity*, 7(4):97, 1955.