Aristotle University of Thessaloniki
Faculty of Sciences
Department of Physics

# Dynamics of Charged Particles in Electromagnetic Fields with Applications in Fusion Devices

Ph.D. Thesis
of
Nikos Kallinikos

Section of Astrophysics, Astronomy and Mechanics

SUPERVISOR: Assistant Professor E. Meletlidou

December 2016

*The reasonable man adapts himself to the world,*
*the unreasonable one persists in trying to adapt the world to himself.*
*Therefore, all progress depends on the unreasonable man.*

George Bernard Shaw

Siméon Poisson (1781-1840)

William Hamilton (1805-1865)

James Maxwell (1831-1879)

Sophus Lie (1842-1899)


Emmy Noether (1882-1935)

Élie Cartan (1869-1951)

The progress of Mathematical Physics much depended on the above "unreasonable" men and woman, whose names will be repeated now and then in the pages that follow.

## Abstract

In this thesis, we investigate the three-dimensional Newtonian motion of charged particles in electromagnetic fields both numerically and analytically.

On one hand, we explore the collective behavior of large numbers of particles in toroidal geometry and the presence of magnetic islands under the influence of wave-particle interaction. Focusing on applications in tokamak plasma, this study is based on simulations of the particles orbits including relativistic effects and collisions. For their need, an exact magnetic surface label for isolated perturbations is also derived in toroidal coordinates, and from the results we determine the current drive and wave absorption.

On the other hand, we carry out a symmetry group analysis for the equations of charged particle motion in the autonomous case. Resulting from the Lorentz force for arbitrary stationary electromagnetic fields that obey the homogeneous Maxwell's equations, we tackle the general autonomous system of three second-order ordinary differential equations that comes from a Lagrangian of three degrees of freedom with velocity-dependent potentials. Considering non-trivial cases of physical interest and, in particular, three-dimensional inhomogeneous, curved magnetic fields, we find the general group of Lie point symmetries when the system is nonlinear. A comparison between these symmetries and the symmetries of the magnetic field lines is also made. Subsequently, the group of augmented point equivalence transformations for this class of systems is found too, taking into account the homogeneous Maxwell's equations as auxiliary conditions. Under the equivalence group, we classify members of this class, that is, time-independent electromagnetic fields in terms of symmetry groups. Symmetries of Noether type are determined as well and the corresponding integrals of motion are constructed. Finally, based on the latter, aspects of complete integrability are discussed.

# Acknowledgements

If the path during my studies was governed by any dynamics at all, then the latter would certainly be far from integrability. However, many people kept me away from chaotic regions. I thank them deeply all. I can only hope I have not totally let them down. This scientific journey did not start at the beginning of this thesis (and hopefully will not stop at the end of it), and although not smooth it was probably continuous. Therefore, my gratitude cannot be limited only to those who supported me during the phd years, but unavoidably extends to other people who forged my research interests long before.

Starting chronologically, as a senior undergraduate student at the Department of Physics of Aristotle University of Thessaloniki I was lucky to have a diversity of skillful and encouraging teachers in the area of theoretical physics. Most of all and hopeless it may sound after many years since he passed away, I still feel the need to thank my supervisor of my undergraduate dissertation Professor Simos Ichtiaroglou, who showed me around the realm of Hamiltonian mechanics and symplectic geometry. His teaching skills and ability to impart sophisticated notions in the most simple way can only match his kindness and generosity. While his interest for my studies and his believing in me at a time when I had none kept me going so many times since then.

Next stop and a very important one for me was the Department of Mathematics at the University of Patras, where I had the privilege to work my master thesis under the supervision of Professor Dimitris Tsoubelis. He is the one who introduced me to the world of symmetries, among other methods in the theory of differential equations, and for his guidance, his support and the opportunities he opened up for me, which went on during my phd studies as well, I cannot thank him enough. To this day his knowledge and way of working are a constant source of inspiration to me. Together with Professors Vassilis Papageorgiou and Tassos Bountis, they all taught me many concepts of mathematical physics and more importantly showed me how to cope with the rest of them. Thank you all deeply.

Returning back to the Physics Department and my time there now as a phd student, I would like to thank my supervisor Assistant Professor Efi Meletlidou most of all for giving me the space to work on my own ideas and the freedom to take initiatives. I can only hope she enjoyed it as much as I did. Offering me loose working conditions and always valuing my opinion, she helped me grow as a scientist. I also want to thank her for proofreading the manuscript. I would like to express my sincere gratitude and appreciation to my advisor Professor Loukas Vlahos for welcoming me to the group of plasma physics from the beginning of my phd studies, and for giving me the opportunity to participate in a research programme of international level. His constant support and our collaboration proved several times the only motivation to complete this thesis. I am also very thankful to my advisor Professor Harry Varvoglis for his interest in my progress from the first to the last day, his valuable comments on parts of this thesis and for being there when needed.

I am more than grateful to Dr Heinz Isliker for the countless hours he spent in front of a computer trying to explain to me how things work inside one. Our endless discussions outside the observatory for a smoke or two have been a continuous source of valuable help and advice to me. Heinz, thank you for your patience, your persistence and a healthy collaboration all these years.

It gives me great pleasure to acknowledge the significance of my scientific interaction with Dr Stelios Dimas, who introduced me to the theory of equivalence transformations. Without his contribution, chapter 6 and [60] accordingly would not even begun to exist. Stelio, thank you for all the time and effort you put through from the other side of the world to get into my head over skype how classifying equations should be treated. May the force be with you.

A special thanks to Professor Fani Petalidou for so willingly and promptly tracking down and explaining Marle's article [81] on embeddings of presymplectic manifolds. I am only sorry we had not got the chance of more nonstop four-hour meetings.

Many thanks to my officemates Dr Jamal-Odysseas Maaita and Mariora Kavakioti for their friendly and easygoing spirit on a daily basis. Mariora, those long breaks and relaxed chats will surely be missed. Many wishes to both of you for a successful career.

Besides good teachers, I was also very lucky to have great fellow graduate students from Patras, whom also soon became even greater friends. Over the years and even when far away, they have helped me enhance my skills and knowledge in mathematics, computers, singing, reciting, crossing bridges, walking through the jungle, surviving in the woods, etc. Starting with Dr Sotiris Konstantinou-Rizos (aka sokor) who from wherever he is he would still manage

to solve whatever technical or other problem I am faced with whenever I ask (or not) for his assistance. The 2×Dr (for the time being) George Papamikos, who first showed me the benefits of working together, has always been my partner in crime, i.e. mathematical physics. Despite Nikitas Nikandros' taste for pure mathematics (or perhaps because of that), he and I have consistently shared plenty of laughs between tv series analyses, philosophical discussions on mathematics and poetry readings among other things. We have tried many times with Dr Stavros Anastassiou to collaborate, but repeatedly ended up for drinks and music performances (unfortunately by us); hope we get it right next time. On the other hand, with Dr Thodoris Kouloukas we completely skip the first part and go straight to the drinking part; our camping exhibitions as well as his ways through the wilderness are memorable. Congratulations to his wife Dr Helen Christodoulidi for putting up with him but also his friends including me. In this gang Dr Dimitris Nomikos is the missing link between having fun and painting the town red; too bad (or thank god) our studies in Patras did not overlap and we had to recover lost ground only with few times a year. Never leaving my mind and constantly figuring me out, Dr Ilia Roustemoglou is the little sister I never had. Finally, Dr Grigoris Protsonis has always been a special friend and the voice of reason in good and bad times; his personality (and house) has offered us all the calm and solicitude (if not childcare) when needed.

Other friends from Thessaloniki have been for so many years like family to me. Forever believing in me, their love was, is and will be a compass in my life that points me into pleasure and happiness at all times. Apologies to all for trading many gatherings for studying and working especially in the last few years. I am only looking forward to the next one.

It goes without saying that without the love and support of my parents Babis and Nana as well as my brother Christos and his lovely wife Betty my studies like most things would have neither been fulfilled nor probably even begun. Always supporting me in all my choices, even when they do not agree, and giving their best to see me happy, their help is more than I deserve. Moreover, my two beloved nieces Athina and Eftixia have played perhaps the most difficult part, that of clearing my head from unbreakable equations, impenetrable theorems and any frustration whatsoever every time I see them.

Last but certainly not least, the person who suffered and enjoyed probably even more than me the undertaking of this thesis is Natasha Theodorelou. The ways she filled my life and her unconditional love I cannot even begin to describe them. All I can say is that her nonstop patience, her nonnegiotable belief in me and her not-of-this-world kindness are more than I could ever wish for. And that nothing could make me happier than to return the favor.

Part I of this thesis was supported by (a) the National Programme for the Controlled Thermonuclear Fusion, Hellenic Republic, (b) the EJP Cofund Action SEP-210130335 EUROfusion. For the results of chapter 3, the author also gratefully acknowledges the access granted to the supercomputer Helios-IFERC in Japan. The sponsors do not bear any responsibility for the content of this work.

# Contents

# Introduction

Plasma is roughly speaking an ionised gas of charged particles (electrons and protons) and was first described by Crookes in 1879. The term itself was introduced later by Langmuir in 1928, borrowing the Greek word πλάσμα that basically means shapeless, of any form. In a sense of energy ordering, it has also been called the fourth state of matter, the other three being solid, liquid and gas. For under sufficient heating, as a solid transforms into a liquid and a liquid to a gas, a gas transforms into a plasma. Unlike the earth's surface, most of the visible universe is in the plasma state. Typical examples are the earth's ionosphere and magnetosphere, lightning, the polar aurorae, the solar corona and outer layers of other stars as well, the solar wind, the interplanetary space, the interstellar space, etc. Laboratory plasma on the other hand is of no less importance, serving as a vehicle for controlled thermonuclear fusion in the course of taming fusion energy. Plasma Physics is therefore very significant in understanding several phenomena either in Astrophysics or in Fusion research.

In the Theory of Plasma Physics there have been developed four successive models to (try to) describe the collective behavior of a very large number of charged particles. The first one encountered in any related textbook is the single particle model, which, given away by its name, simply focuses on the particles themselves and, in particular, the equations of motion that they obey, and determine their orbits. To quote Radu Balescu [4], one of the pioneers in the field, "The whole subject of Plasma Physics can be reduced, in last analysis, to the understanding of the motion of a set of charged particles in an electromagnetic field." Of course, as he immediately stresses afterwards, however simple this statement may be, the complexity of the problem is quite high and in more than one way. Still, this rather simplified picture is not without uncharted territories. What is even more challenging is that charged particle dynamics really dates back to a problem of Classical Mechanics, whenever relativistic effects can be neglected; quantum effects are almost always negligible in plasma physics, which is mostly regarded as a classical field theory.

In this approach, the motion of charged particles under the influence of electromagnetic fields is governed by Newton's second law, described via the Lorentz force and considered in three-dimensional space. For a particle of mass $m$ and charge $q$, the equations of motion are three (coupled) second-order ordinary differential equations, which in vector form read

$$m\ddot{\boldsymbol{x}} = q\left(\dot{\boldsymbol{x}} \times \boldsymbol{B} + \boldsymbol{E}\right), \tag{1}$$

where $\boldsymbol{x}$ is the position of the particle as a function of time $t$, $\boldsymbol{B}$ and $\boldsymbol{E}$ are the magnetic and electric fields, respectively, which may depend on both $\boldsymbol{x}$ and $t$ in general, while the dot stands for derivation with respect to time $t$. Of course, system (1) goes hand in hand with the well-known Maxwell's equations that describe the source and time evolution of the electromagnetic field. From the perspective of the single particle model, the electromagnetic field produced by the motion of the particle in turn is fairly negligible compared to the solutions $\boldsymbol{B}$ and $\boldsymbol{E}$ of Maxwell's equations. When studying, however, collectively large numbers of particles this can no longer be the case. (Ideally one would have to determine the total charge and current coming from the particles, insert them into Maxwell's equations, solve the latter all over again to find the modified electromagnetic fields, then replace the new solutions in the equations of motion, and so on.)

The motion of a particle under the action of the magnetic field alone along with its intriguing structure is characterized by many features that are quite unusual for Newtonian mechanics. In Griffiths' words [39], "it leads to some truly bizarre particle trajectories", which as Jackson [59] points out, "arise because of the peculiarities of the magnetic-force term." Unlike gravitational or electric fields that accelerate a particle along their field lines regardless the particle's velocity, the magnetic field introduces a force (see (1)) that differs in both ways, for it is *i*) perpendicular to the magnetic field lines, and more importantly *ii*) velocity-dependent. As a result, the particle dynamics and the dynamics of the magnetic field itself are in principle two very distinct systems, naturally raising the question of how these two systems relate. Moreover, in the typical case of divergence-free magnetic fields, the notion of the vector potential enters the picture with all its intrinsic geometrical aspects. Thus, due to the overall nature of the magnetic field, charged particle dynamics occupies a special place among Dynamical Systems.

Solutions to system (1) are only known in very simple situations (see [70, 71, 16]), which do not fall far from the simplest of all, i.e. when the electromagnetic field is homogeneous and stationary. Even then, however, the particle dynamics are not simple at all (see section 1.5), justifying Griffiths' comment. Neither are there many cases, for which a second constant of motion, besides the energy (i.e. the Hamiltonian) of the system, is known to exist, unless some

geometrical symmetry of the fields is apparent. For instance, the motion in a magnetic dipole, also known as the Störmer problem, involves a magnetic field having a rotational symmetry that is also admitted by the system and gives rise to the invariance of the canonical angular momentum (see Example 4.42). This problem has been extensively studied and proven to be non-integrable; a third constant of motion does not exist. Roughly speaking, most responsible for the highly complicated dynamics of charged particles with consequent integrability limitations is the magnetic term of system (1), which to a first approximation forces the particle to wind around the magnetic field. When we move on to more complicated systems, this motion is usually averaged out as a last resort. In other words, the rotation of the particle around the magnetic field lines is eliminated, providing the system with an adiabatic, instead of an exact, invariant related to the magnetic moment (see section 1.6). Still in these cases numerical integration is all the same often employed for further study.

The above scenery does not leave little room for questions, such as: Besides trivial cases, are there any other integrable cases? Which types of constants of motion arise, apart from energy conservation in the autonomous case? Are all the symmetries of the magnetic field lines admitted by the equations of motion? Or, what role do magnetic surfaces (invariant surfaces for the magnetic field) play on charged particle motion? Are there any other cases, besides homogeneous stationary magnetic fields with zero electric field, for which the magnetic moment is an integral of motion? Can we find analytical solutions for the guiding-center equations (the equations of motion resulting from the above mentioned averaging)? Only partial answers or answers in certain cases of interest to some of these questions have be given. For example, Lewis [72] has directly searched and found a type of quadratic integrals of motion along with the conditions on the (potentials of the) electromagnetic field. Balescu [5] describes the form of solutions for the guiding-center system in the case of an axisymmetric (meaning of rotational symmetry) magnetic field (in particular the standard magnetic field (2.19)). In the end, given an electromagnetic field one is interested to know what can he or she say about the dynamics of a charged particle and how should he or she treat it. Such a query requires exploring the structure of equations (1) more methodically.

Symmetry methods are one of the few systematic tools for studying nonlinear differential equations, either ordinary or partial, revealing, if any, their integrability properties. Note that one way or another the notion of symmetry was also present in the aforementioned cases of the magnetic dipole or the guiding center. In dynamical systems, continuous symmetries can be used to reduce the order of the system and in some cases even completely integrate it. Along

with Noether's theorem for variational problems, they may also yield conservation laws, that is constants of motion, which restrict solutions to an invariant manifold. Therefore, the direct search for symmetries in certain systems has received a great deal of attention over the past few decades. Another usage of symmetry analysis, which has been adopted more and more often, is to classify all possible symmetry groups admitted by a wide family of differential equations. Such results determine under which conditions a system of general form may possess one or more symmetries. To this end, equivalence transformations which represent maps between different members of a family are more proper to employ.

The role of symmetry and integrals of motion are of no less significance in applications. Axisymmetric magnetic fields are often used to describe the plasma equilibrium in toroidal configurations used in fusion research known as tokamaks. Starting with a system that admits one or more symmetries or constants of motion, one can investigate any instabilities arising in plasma as a perturbation and see how these symmetries or invariants break. Actually helical magnetic perturbations are quite often considered for studying such experiments, which are also related to a type of helical symmetry for the magnetic field lines. The existence of magnetic surfaces under these perturbations and their effect on the dynamics of charged particles could prove important. The interest of symmetry theory in Plasma Physics is in any case increasingly growing, most often for the Magnetohydrodynamic equations or the Grad-Shafranov equation [14, 22, 36, 105, 120, 23, 64].

## 0.1 What this thesis is about

In a word, equations (1) play the leading role in this thesis. In our study, we (mainly) focus on the autonomous case, magnetic effects being our primary concern. In fact, time-independent magnetic fields adopted throughout this thesis are often met in laboratory plasma covering a wide variety of phenomena. And with an exception in chapter 3, we also assume stationary electric fields. (Nevertheless, exploring time-dependent electromagnetic fields is within our future plans, for which the present work is a first and most likely necessary step.) The thesis' main target, in brief, and its structure accordingly are actually divided in two parts: *i*) an investigation for magnetic islands aiming towards applications, and *ii*) a more theoretical search for symmetries and first integrals of motion.

**Part I** explores the role of magnetic perturbations in toroidal geometry, closely related to problems in fusion experiments and especially tokamak devices.

For this purpose, we give first in **chapter 2** an analytical expression for the magnetic islands formed when an axisymmetric magnetic field is helically perturbed in terms of the poloidal and toroidal angles. Focusing on one dominant resonance mode, we exploit the well-known integrability of such Hamiltonian perturbations, describing the magnetic field lines directly in toroidal coordinates. In this way the previous toroidal magnetic surfaces of the unperturbed system are now replaced by helical ones around the torus. This magnetic surface quantity can then provide analytically any information required for the perturbed topology, such as position of saddles or centers of the islands, construction of Poincaré sections or the magnetic surfaces themselves, calculation of areas or volumes enclosed by magnetic surfaces, without the need of approximations or the risk of numerical errors.

Then, in **chapter 3**, we investigate the effect of magnetic islands on the wave-particle interaction in terms of our basis, the single-particle model. Most approaches on this topic are either linear or quasilinear approximations of the Vlasov equation. In order to explore the nonlinear aspects, we perform particle simulations, for the needs of which a numerical code was developed. The latter integrates the equations of motion, using either the Lorentz force law or the guiding-center equations. Relativistic effects for high energy particles as well as collisions are taken into account too, while the magnetic surface quantity of the previous chapter is employed for describing several features of the magnetic islands. The code also solves the plasma dispersion relation, providing the wavenumber and polarization of the wave. In addition, it has been parallelized over the individual particles either for multiprocessors with shared memory or computer systems with distributed memory. Thus, following large numbers of electrons, statistical results can be obtained either directly or after postprocessing. More specifically we calculate the current drive in the plasma, as well as the wave power absorption from the particles, along with other macroscopic quantities, such as mean kinetic energy, velocity distributions, etc.

**Part II** presents a symmetry analysis of the autonomous system (1) for arbitrary electromagnetic fields. Aiming for nontrivial cases of physical interest, we search for Lie as well as Noether point symmetries corresponding to first integrals of motion, besides the obvious time translations.

In **chapter 5**, we put the system under the symmetry test for arbitrary, inhomogeneous electromagnetic fields, focusing, in particular, on the nonlinear and genuinely three-dimensional case. Starting with unprescribed fields, we investigate how the symmetry condition, apart from the general form of the symmetry generator, places restrictions on the fields too. For consistency with any real physical problem, the homogeneous Maxwell's equations are also imposed. Thus,

the previous conditions are expressed and then solved in terms of the vector and scalar potentials, leading to five classes of solutions. In this way, we determine the general form of electromagnetic fields, for which systems of class (1) describing the motion of charged particles admit point symmetries, along with the general form of the symmetries themselves. The use of potentials is also more appropriate in order to identify symmetries of Noether type (and to compare them with Lie point ones), from which we construct first integrals of motion over and beyond the well-known Hamiltonian. We particularly derive cases, where the integrals are functionally independent of the Hamiltonian and in involution with it. In addition, a comparison in terms of symmetries is made between the equations of charged particle motion and the magnetic field lines. In the case of common symmetries, this in turn shows how magnetic surfaces and first integrals of motion are related.

In **chapter 6**, a symmetry group classification is presented either in terms of Lie or Noether point symmetries. The question for which electromagnetic fields does charged particle motion admits point symmetries now is reversed. Rather than expressing electromagnetic fields in terms of symmetries admitted, now symmetries are used to classify the electromagnetic fields, i.e. system (1). For this purpose, we find first the group of equivalence transformations of class (1), considering also the homogeneous Maxwell's equations as restrictions. Then we separate the system into equivalent subcases that are not connected through such transformations, and which admit more symmetries than time translations. Under the action of the equivalence group, each subcase is characterized by a representative electromagnetic potential as simple as possible. The classification is made using the Lie algebra of the equivalence group, separating it in one-, two- and three-dimensional equivalence subalgebras leaving aside the principal Lie algebra. In this way, we deduce accordingly when the system admits one-, two- and three-parameter extended symmetry groups, besides the well-known time translations, in the nonlinear and genuinely three-dimensional case. Based on this classification, aspects of integrability are also examined.

Particular conclusions and more technical remarks are discussed more suitably at the end of each chapter. An overall outline of the results of this thesis is given finally in **chapter 7** based on which plans for future work are also made.

Parts of this thesis, chapters 2 and 5, have been published in international scientific peer-reviewed journals, namely [62] and [63], respectively. Others, chapters 3 and 6, which correspond to [61] and [60] accordingly, will soon be submitted.

Part of the results presented here is a joint work with collaborators, as indicated in the jointly-authored of the aforementioned publications.

### Further remarks on the structure of the thesis

As probably guessed by now, **chapters 1 and 4** describe the theory used in the rest of parts I and II, respectively.

The first chapter provides the basic features of charged particle motion following a path from physical variables, to cylindrical velocities, and from there to guiding-center variables, given finally in relativistic form. In this short trip homogeneous electromagnetic fields are an intermmediate stop serving as a link passage to guiding-center theory. Our presentation is based on the Hamiltonian nature of the system and in particular the notion of Poisson brackets, which lately have received great attention in Gyrokinetics [20, 18]. Among these features one cannot overlook the relation of the particle's motion with the system of the magnetic field lines. Thus, the structure and geometrical aspects of the magnetic field are described as well and in particular its well-known Hamiltonian nature. An interpetration of the latter is explored in terms of presymplectic forms in section 1.3 based on subsections 1.2.1 and 1.2.2 and appendix B, which also explain the relation with the charged particle motion.

The fourth chapter is an introduction to symmetry theory of differential equations, focusing mostly on Lie point symmetries and their usage either in ordinary or partial differential equations. A few aspects of variational problems are also presented, within the context of classical mechanics, in order to arrive at the celebrated Noether's theorem providing a relation between symmetries (of a certain type) and first integrals of motion. Finally, the less-often exhibited theory of equivalence transformations is shortly described in an infinitesimal fashion much close to the one from symmetry methods.

**Appendices** have also been included to further support here and there several tools needed especially in the first chapter.

First, appendix A contains formulas from vector calculus that often come along with the structure of the magnetic field. Although elementary, they are quite necessary and were preferably displayed here, rather than constantly interrupting and referencing the reader to other sources with different notation as well.

The idea of appendix B is to serve as a vehicle for exploring the geometrical aspects of the magnetic field and the corresponding vector potential from the point of view of differential geometry. The ultimate goal is symplectic and in particular presymplectic forms used in section 1.3. For the needs of the latter and for connecting with the previous appendix, the particular case of the three-dimensional Euclidean space is also presented with usual vector operations expressed in the language of forms.

Finally, appendix C is a short description of (finite-dimensional) Hamiltonian systems without the need of canonical variables, often found unsuitable for either charged particle motion or magnetic field lines. Preference here is given to the notion of Poisson brackets, repeatedly employed in chapter 1, and its relation to the symplectic structure for consistency.

Let me close this section with some final remarks on the organization and writing style of this thesis, which hopefully would make reading it easier. One goal of primary interest was to have as much as possible a self-sufficient text. This probably turned into a necessity, when elements from different areas of Physics or Mathematics entered the picture. Avoiding as much as possible "leaving exercises to the reader", several calculations or proofs are also included. In an ideal scenario, these could help graduate students, who approach some of the thesis' aspects for the first time. In the same spirit, examples are given too for clarity as illustrations of the theory used. Overall, from my understanding, chapters 1 and 4 (as well as the appendices) are essential for connecting different parts of this thesis, and, in any case, I hope they will make an interesting reading.

That being said, the reader who is familiar with the theory of charged particle motion can skip chapter 1 and go directly to chapters 2 and 3, just as the one familiar with symmetry group theory can skip chapter 4 and go straightaway to chapters 5 and 6.

# Notation

For easy reference, below are listed the most frequently used symbols in this thesis. Borrowing elements from different areas, the notation adopted here tries to keep up with the dominant nomenclature of each scientific field and stay self-consistent at the same time. Even though exhausting almost entirely the Greek and English alphabets, unavoiadably there is some overlapping, which should be of no confuse by the context used each time. In any case, further notation comments are given at the beginning of each chapter. Some ground rules are: All vectors in $\mathbb{R}^3$ are denoted by bold symbols, e.g. $\boldsymbol{B}$, and their magnitude accordingly by $|\boldsymbol{B}|$. Einstein's summation convention has been adopted in chapters 1, 4, 5 and 6, as well as in the appendices, but not followed in chapters 2 and 3. And unless stated otherwise all indices in chapters 1, 5 and 6 take values from 1 to 3. The same goes for the appendix A, while in appendices B and C indices range from 1 to the dimension of the relevant (sub)manifold. In the thesis SI units have been used, but in most cases constants have been normalised to $m = q = 1$.

| | |
|---|---|
| $\boldsymbol{A}$ | vector potential of the magnetic field |
| $\boldsymbol{b}$ | unit vector of the magnetic field |
| $\boldsymbol{\beta}$ | binormal vector to the magnetic field |
| $B$ | magnetic 2-form, magnetic matrix |
| $\boldsymbol{B}$ | magnetic field |
| $c$ | speed of light; constant of integration |
| $\gamma$ | Lorentz factor |
| $D_x$ | total derivative with respect to $x$ |
| $\epsilon_{ijk}$ | Levi-Civita symbol |
| E | Euler operator |
| $\boldsymbol{E}$ | electric field |
| $\delta$ | auxiliary system for a class of differential equations |
| $\delta_{ij}$ | Kronecker's delta |

| | |
|---|---|
| $\boldsymbol{\delta}$ | curl of velocity |
| $\Delta$ | system of differential equations |
| $\zeta$ | gyrophase |
| $g$ | metric tensor |
| $\mathfrak{g}$ | Lie algebra of a Lie group |
| $G$ | Lie group |
| $H$ | Hamiltonian function |
| $\mathcal{H}$ | Hamiltonian function for the magnetic field |
| $h$ | helical magnetic flux; value of the Hamiltonian function |
| $\theta$ | poloidal angle |
| $\theta_{\text{in}}$ | intrinsic poloidal angle |
| $\vartheta$ | Liouville 1-form |
| $I$ | first integral of motion; identity matrix or operator |
| $J$ | Poisson matrix |
| J | Jacobian matrix |
| $\mathcal{J}$ | Poisson matrix for the magnetic field |
| $\boldsymbol{J}$ | current density |
| $k$ | wavenumber |
| $l$ | natural parameter of magnetic field lines |
| $L$ | Lagrangian function |
| $\mathcal{L}$ | Lagrangian function for the magnetic field |
| $\mathfrak{L}$ | functional of the Lagrangian function |
| $\mu$ | magnetic moment |
| $\boldsymbol{n}$ | principal normal vector to the magnetic field |
| $\boldsymbol{n}_1$ | perpendicular unit vector to the magnetic field |
| $\boldsymbol{n}_2$ | perpendicular unit vector both to the magnetic field and $\boldsymbol{n}_1$ |
| $\boldsymbol{p}$ | canonical momenta |
| $P$ | power |
| $q$ | safety factor |
| $r$ | minor radius in toroidal coordinates; radius in spherical coordinates |
| $R_0$ | major radius in toroidal coordinates |
| $\rho$ | gyroradius; radius in polar coordinates |
| $\boldsymbol{u}$ | proper velocity |

| | |
|---|---|
| $s$ | parameter of magnetic field lines |
| $t$ | time |
| $\tau$ | proper time |
| $\boldsymbol{v}$ | velocity |
| $\mathbf{v}$ | vector field |
| $\mathbf{v}^n$ | composition of $\mathbf{v}$ as an operator with itself $n$ times |
| $\mathbf{v}^{(n)}$ | prolongation of $\mathbf{v}$ of order $n$ |
| $\Phi$ | scalar potential of the electric field |
| $w$ | point of a Poisson manifold |
| w | winding number |
| $\phi$ | toroidal angle; polar angle |
| $\varphi$ | flow of a vector field; Lie group action |
| $x$ | point of a manifold; coordinate in the $x$-axis |
| $\boldsymbol{x}$ | point of $\mathbb{R}^3$ |
| $\psi$ | toroidal magnetic flux |
| $\omega$ | wavefrequency; differential form |
| $\omega_g$ | gyrofrequency |
| $\Omega$ | symplectic matrix |
| $\partial_x$ | partial derivative with respect to $x$ |
| $\parallel$ | parallel component to the magnetic field |
| $\perp$ | perpendicular component to the magnetic field |

# Part I

# Motion of Charged Particles : Theory and Applications

# Chapter 1

# Basic Principles of Charged Particle Motion

In this section we briefly go through key elements of the theory of charged particle motion that will be used later on. We start off with a short restatement of the problem and basic notation previously described in the Introduction.

In Classical Mechanics, a charged particle moving in an eletromagnetic field undergoes the Lorentz force and obeys Newton's second law. The space considered is the three-dimensional Euclidean space $\mathbb{R}^3$ and, unless stated otherwise, we assume that both the electric and magnetic field are stationary, i.e. time-independent. The equations of motion for the particle is expressed by a system of three second order, autonomous differential equations, which in vector form are

$$\ddot{\boldsymbol{x}} = \dot{\boldsymbol{x}} \times \boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{E}(\boldsymbol{x}), \tag{1.1}$$

where $\boldsymbol{B}$ and $\boldsymbol{E}$ are smooth vector functions of the position $\boldsymbol{x}$ alone, representing the magnetic and electric field, respectively. Using a Cartesian coordinate system, in which $\boldsymbol{B} = (B_1, B_2, B_3)$ and $\boldsymbol{E} = (E_1, E_2, E_3)$, the system (1.1) can also be written as

$$\ddot{x}_i = \epsilon_{ijk}\dot{x}_j B_k + E_i, \tag{1.2}$$

where $\epsilon_{ijk}$ is the Levi-Civita symbol. Einstein's summation convention has been adopted, assuming unless stated otherwise that all indices from now on and throughout the rest of this chapter take values from 1 to 3. By introducing the particle's velocity $\boldsymbol{v} = \dot{\boldsymbol{x}}$, equations (1.1) can be equivalently expressed as a first-order system,

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{v}, \\ \dot{\boldsymbol{v}} &= \boldsymbol{v} \times \boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{E}(\boldsymbol{x}). \end{aligned} \tag{1.3}$$

Coupled to the Lorentz force law is another set of physical laws that the electromagnetic field must obey, and which altogether constitute a system of first-order partial differential equations, known as Maxwell's equations. For stationary fields these are

$$\nabla \cdot \boldsymbol{B} = 0, \qquad \nabla \times \boldsymbol{E} = 0, \tag{1.4a}$$

$$\nabla \cdot \boldsymbol{E} = \sigma, \qquad \nabla \times \boldsymbol{B} = \boldsymbol{J}, \tag{1.4b}$$

where $\sigma$ and $\boldsymbol{J}$ are the charge and current densities. The last two equations express Gauss's and Ampére's laws, respectively, relating the fields with their sources. On the other hand, the first two through their solutions lead to the notion of the electromagnetic potential. This means that the magnetic field can be derived from a vector potential $\boldsymbol{A}(\boldsymbol{x})$ with Cartesian components $\boldsymbol{A} = (A_1, A_2, A_3)$, while the electric field from a scalar one $\Phi(\boldsymbol{x})$,

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}, \tag{1.5}$$

$$\boldsymbol{E} = -\nabla\Phi. \tag{1.6}$$

The minus sign in (1.6) physically means that an increasement of $\Phi$ for a positive charge results in an electric field $\boldsymbol{E}$ in the opposite direction.

**Remark 1.1.** Both the vector and the scalar potential are not real, measurable quantities and they can be defined arbitrarily, because of the invariance of the electromagnetic field under gauge transformations. In the stationary case considered here, the latter are separable in $\boldsymbol{A}$ and $\Phi$ and, as we can see from (1.6) and (1.5), split to the trivial transformation $\Phi \longrightarrow \Phi + c$ admitted by the electric field, where $c$ is an arbitrary constant, and more importantly to $\boldsymbol{A} \longrightarrow \boldsymbol{A} + \nabla g$ admitted by the magnetic field, where $g$ is an arbitrary function of $\boldsymbol{x}$.

## 1.1 Lagrangian and Hamiltonian formulation

Equations (1.4a) allow another viewpoint of charged particle motion, namely as an Euler-Lagrange system. For, in light of the potentials, equations (1.1) admit the Lagrangian function

$$L(\boldsymbol{x}, \dot{\boldsymbol{x}}) = \frac{1}{2}\dot{\boldsymbol{x}}^2 + \dot{\boldsymbol{x}} \cdot \boldsymbol{A}(\boldsymbol{x}) - \Phi(\boldsymbol{x}). \tag{1.7}$$

In other words, the motion of a charged particle can be described by a variational principle (namely the principle of least action), according to which the trajectory $\boldsymbol{x}(t)$ of the particle for the time interval $t_2 - t_1$ is an extremum of the functional

$$\int_{t_1}^{t_2} L(\boldsymbol{x}, \dot{\boldsymbol{x}})\, dt, \tag{1.8}$$

where the integrand $L$ is given by (1.7). The extrema $\boldsymbol{x}(t)$ of the action integral (1.8) satisfy the Euler-Lagrange equations (see section 4.7), which yield equations (1.1),

$$\begin{aligned} 0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} &= \frac{d}{dt}\left(\dot{x}_i + A_i\right) - \dot{x}_j \frac{\partial A_j}{\partial x_i} + \frac{\partial \Phi}{\partial x_i} = \ddot{x}_i + \left(\frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i}\right)\dot{x}_j + \frac{\partial \Phi}{\partial x_i} \\ &= \ddot{x}_i - \epsilon_{ijk}\dot{x}_j B_k - E_i\,. \end{aligned}$$

**Remark 1.2.** In terms of Analytical Mechanics, we easily see that the first term in (1.7) corresponds to the kinetic energy of the system. If we were to identify the remaining last two as the total "potential" of the system, then this would depend linearly on the velocities.

The typical route to the Hamiltonian description is to define the conjugate (to the variables $x_i$) momenta,

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = \dot{x}_i + A_i(\boldsymbol{x}), \tag{1.9}$$

which besides the velocity depend on the position of the particle, as well. Then, expressing $\dot{\boldsymbol{x}} = \boldsymbol{p} - \boldsymbol{A}(\boldsymbol{x})$, the Hamiltonian function can be constructed through the relation

$$H(\boldsymbol{x}, \boldsymbol{p}) = \boldsymbol{p}\cdot\dot{\boldsymbol{x}} - L\left(\boldsymbol{x}, \dot{\boldsymbol{x}}\right) = \frac{1}{2}\left(\boldsymbol{p} - \boldsymbol{A}(\boldsymbol{x})\right)^2 + \Phi(\boldsymbol{x}). \tag{1.10}$$

It is worth noting that $H$ is not separable in $\boldsymbol{x}$ and $\boldsymbol{p}$, in fact if we expand the first term we arrive again at a "potential" linear in the momenta, as in the Lagrangian case. Using (1.10), system (1.3) can be recovered by the well-known canonical form of Hamilton's equations,

$$\begin{aligned} \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \quad &\Rightarrow \quad \frac{dx_i}{dt} = p_i - A_i \quad \Rightarrow \quad \dot{x}_i = v_i\,, \\ \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} \quad &\Rightarrow \quad \frac{dp_i}{dt} = \left(p_j - A_j\right)\frac{\partial A_j}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \\ &\Rightarrow \quad \frac{dp_i}{dt} = v_j \frac{\partial A_j}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} - \frac{dA_i}{dt} + \frac{dA_i}{dt} \\ &\Rightarrow \quad \frac{dp_i}{dt} - \frac{dA_i}{dt} = v_j \frac{\partial A_j}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\frac{dx_j}{dt} \\ &\Rightarrow \quad \frac{d\left(p_i - A_i\right)}{dt} = \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\right) v_j - \frac{\partial \Phi}{\partial x_i} \\ &\Rightarrow \quad \dot{v}_i = \epsilon_{ijk} v_j B_k + E_i. \end{aligned}$$

However, the disadvantage of the canonical momentum not having a physical meaning turns our attention back to the velocity, raising the question of a Hamiltonian expression in terms of the latter. To this end, we must move towards the more general notion of a (finite-dimensional) Hamiltonian system, which is given in Definition C.1. In order not to get carried away by the

Hamiltonian aspects at this point, we have included all the necessary tools and conclusions in the appendix C, on which we would often rely at least in this section.

So, from the canonical case described there in general, first we recall the canonical Poisson matrix (C.12), given here by the $6 \times 6$ constant matrix

$$J(\boldsymbol{x}, \boldsymbol{p}) = \begin{pmatrix} O & I \\ -I & O \end{pmatrix}, \tag{1.11}$$

where $O$ is the $3 \times 3$ zero matrix and $I$ is the $3 \times 3$ unit matrix, as well as that the canonical variables satisfy the relations $\{x_i, x_j\} = \{p_i, p_j\} = 0$ and $\{x_i, p_j\} = \delta_{ij}$. Now, we are ready to move on to the physical variables $\widetilde{w} = (\boldsymbol{x}, \boldsymbol{v})$. To determine the Hamiltonian formulation (C.2) in terms of $\widetilde{w}$, we can rely on the canonical one in terms of the variables $w = (\boldsymbol{x}, \boldsymbol{p})$. We need just to perform the transformation $w \longrightarrow \widetilde{w}$, which even more conveniently involves only half of the variables and simply reduces to (1.9), i.e. $\boldsymbol{p} = \boldsymbol{v} + \boldsymbol{A}(\boldsymbol{x})$. Starting with the Hamiltonian, we easily get from (1.10)

$$H(\boldsymbol{x}, \boldsymbol{v}) = \frac{1}{2}\,\boldsymbol{v}^2 + \Phi(\boldsymbol{x}), \tag{1.12}$$

which straightaway can now be identified with the particle's energy. The absence of the magnetic field in (1.12) is the well-known property that it does no work. The only thing left to find is the Poisson matrix $J(\widetilde{w})$, which can be determined through the relations (C.9), $J_{ij}(\widetilde{w}) = \{\widetilde{w}_i, \widetilde{w}_j\}$, considering both $\widetilde{w}$ and the Poisson bracket in terms of $w$. We have

$$\begin{aligned}
&\{x_i, x_j\} = 0 \\
&\{x_i, v_j\} = \{x_i, p_j - A_j(\boldsymbol{x})\} = \{x_i, p_j\} - \{x_i, A_j(\boldsymbol{x})\} = \delta_{ij} - 0 = \delta_{ij} \\
&\{v_i, x_j\} = -\{x_j, v_i\} = -\delta_{ji} = -\delta_{ij} \\
&\{v_i, v_j\} = \{p_i - A_i(\boldsymbol{x}), p_j - A_j(\boldsymbol{x})\} = \{p_i, p_j\} - \{p_i, A_j(\boldsymbol{x})\} - \{A_i(\boldsymbol{x}), p_j\} + \{A_i(\boldsymbol{x}), A_j(\boldsymbol{x})\} \\
&\qquad = 0 + \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} + 0 = \epsilon_{ijk} B_k
\end{aligned}$$

where the last equation prevents $w \longrightarrow \widetilde{w}$ from being a canonical transformation. Therefore, the Poisson matrix in this case is a non-constant $6 \times 6$ matrix that depends on $\boldsymbol{x}$,

$$J(\boldsymbol{x}, \boldsymbol{v}) = \begin{pmatrix} O & I \\ -I & B(\boldsymbol{x}) \end{pmatrix}, \tag{1.13}$$

in which, at the expense of having the physical velocity instead of the canonical momenta, the

magnetic field appears through the matrix $B$ with elements $B_{ij} = \epsilon_{ijk} B_k$,

$$B(\boldsymbol{x}) = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}. \tag{1.14}$$

Thus, using (1.12) and (1.13) the equations of charged particle motion (1.3) can be expressed as a Hamiltonian system (C.2), or in more detail

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial v_i} \\ \frac{dv_i}{dt} &= -\frac{\partial H}{\partial x_i} + \epsilon_{ijk} B_k \frac{\partial H}{\partial v_j} \end{aligned} \tag{1.15}$$

According to (C.8) the Poisson bracket for this system is

$$\{F, G\} = \frac{\partial F}{\partial x_i}\frac{\partial G}{\partial v_i} - \frac{\partial F}{\partial v_i}\frac{\partial G}{\partial x_i} + \epsilon_{ijk} B_k \frac{\partial F}{\partial v_i}\frac{\partial G}{\partial v_j} \tag{1.16}$$

In conclusion, system (1.3) can be put in Hamiltonian form in two equivalent ways, either the usual canonical form of Hamilton's equations and the Hamiltonian function (1.10) or system (1.15) and the Hamiltonian (1.12). In the first one the magnetic field enters the Hamiltonian through the vector potential, leaving the Poisson matrix constant, while in the second one the magnetic field has left the Hamiltonian, entering the Poisson matrix instead. The canonical variables used in the first one although popular in Hamiltonian mechanics lack here physical meaning, for the conjugate momenta cannot be identified with the physical ones, which is also reflected in the corresponding Hamiltonian. On the other hand, the second description in terms of the physical velocities, restoring the related Hamiltonian to the energy of the particle, seems to be more suitable, despite its non-canonical character.

One final remark is that the variational principle (1.8) for (1.1), can be replaced by a more general but in fact equivalent one (see [1], p. 243-245) for (1.3). The latter is expressed by the Lagrangian function

$$L(\boldsymbol{x}, \boldsymbol{v}, \dot{\boldsymbol{x}}) = \boldsymbol{p} \cdot \dot{\boldsymbol{x}} - H(\boldsymbol{x}, \boldsymbol{v}) = (\boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{v}) \cdot \dot{\boldsymbol{x}} - \frac{1}{2}\boldsymbol{v}^2 - \Phi(\boldsymbol{x}) \tag{1.17}$$

as opposed to (1.7), and can be regarded as the inversion of the relation (1.10). Notice that in this treatment $\boldsymbol{x}$ and $\boldsymbol{v}$ are considered independent variables. The derivatives of the velocities are absent from (1.17), yielding easily the first set of equations $\partial L/\partial v_i = \dot{x}_i - v_i = 0$, and then the second one can be recovered same as before with the substitution $\dot{x}_i = v_i$ previously found. The above point of view of Lagrangian formulation would prove quite helpful in the forthcoming sections.

## 1.2 Magnetic field lines

Unlike the electric (or the gravitational) field, the magnetic field, as we have seen already, does not act directly on the particle, but through the Lorentz force. Thus, the dynamics of the magnetic field is in general quite different from the particle's motion and deserves special attention. So, in this section we consider another dynamical system consisting now of three first-order, autonomous ordinary differential equations

$$\frac{d\boldsymbol{x}}{ds} = \boldsymbol{B}(\boldsymbol{x}), \tag{1.18}$$

which describes the integral curves of the magnetic field (viewed as a vector field) commonly known as magnetic field lines. The independent variable here, denoted by $s$, is related to the line element of these curves. An alternative way to describe the tangency condition (1.18) of the integral curves $\boldsymbol{x}(s)$ to the magnetic field $\boldsymbol{B}(\boldsymbol{x})$ is to require that $\boldsymbol{x}'(s)$ and $\boldsymbol{B}(\boldsymbol{x})$ are parallel,

$$\boldsymbol{x}'(s) \times \boldsymbol{B}(\boldsymbol{x}) = 0. \tag{1.19}$$

Magnetic field lines are also equipped with a Lagrangian formulation, again due to the divergence free condition from Maxwell's equations (1.4a). Assuming the Lagrangian function

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{x}' \cdot \boldsymbol{A}(\boldsymbol{x}), \tag{1.20}$$

and the corresponding action integral, the Euler-Lagrange equations yield system (1.19),

$$0 = \frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{ds}\left(\frac{\partial \mathcal{L}}{\partial x_i'}\right) = x_j' \frac{\partial A_j}{\partial x_i} - \frac{dA_i}{ds} = \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\right) x_j' = \epsilon_{ijk} x_j' B_k \,.$$

If we try to construct a Hamiltonian formulation from the previously described Lagrangian, as in the case of the particle's motion, we quickly end up with a vanishing Hamiltonian (the conjugate momentum would now be the vector potential itself and so $\mathcal{H} = \boldsymbol{A} \cdot \boldsymbol{x}' - \mathcal{L} = 0$). Magnetic field lines do not have (globally) a Hamiltonian description in the sense of (C.2). They do, however, satisfy one important characteristic of Hamiltonian systems, namely Liouville's theorem. The latter states that the flow of a Hamiltonian system is incompressible or equivalently that Hamiltonian vector fields are divergence free, just like magnetic fields. Thus, system (1.18) falls into the wider category of Liouville dynamics, which refers to dynamical systems described by divergence free vector fields.

Nonetheless, from the above discussion it is evident that the divergence free condition is a necessary condition for Hamiltonian systems, a sign of a pre-Hamiltonian structure if you will. This point of view will be further explained in the next section based on a more theoretical

ground of Hamiltonian mechanics. For now, we will restrict ourselves to a few general cases, demonstrating, indeed, that the magnetic field lines can always be put in Hamiltonian form depending on choice of variables. To cover a wide variety of applications we will switch to curvilinear coordinates $\left(x^1, x^2, x^3\right)$ described in appendix A, which will be used many times as reference in the following two subsections. From formula (A.35), system (1.18) can be expressed in terms of the vector potential as

$$\begin{aligned} \frac{dx^1}{ds} &= \frac{1}{\sqrt{g}}\left(\frac{\partial A_3}{\partial x^2} - \frac{\partial A_2}{\partial x^3}\right) \\ \frac{dx^2}{ds} &= \frac{1}{\sqrt{g}}\left(\frac{\partial A_1}{\partial x^3} - \frac{\partial A_3}{\partial x^1}\right) \\ \frac{dx^3}{ds} &= \frac{1}{\sqrt{g}}\left(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2}\right) \end{aligned} \tag{1.21}$$

where $A_i$ are the covariant components of the potential $\boldsymbol{A}$ and $g$ is the determinant of the metric tensor defined by $\left(x^1, x^2, x^3\right)$, given in (A.21).

### 1.2.1 Magnetic surfaces

Firstly consider the case, where one of the components of the magnetic field vanishes, say $B^3 = 0$, which, as we can see from (A.28) and (A.31), can also be expressed as $\boldsymbol{B} \cdot \nabla x^3 = 0$. The latter is often referred to as a homogeneous *magnetic differential equation*. Either way this condition means that the field lines lie on the subset $x^3(x, y, z) = \text{const.}$ of $\mathbb{R}^3$, which is called a *magnetic surface*, and that

$$\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} = 0.$$

Then there exists a function $f(x^1, x^2, x^3)$ such that $A_1 = \partial f/\partial x^1$ and $A_2 = \partial f/\partial x^2$ and therefore the vector potential is $\boldsymbol{A} = (0, 0, \widetilde{A}_3) + \nabla f$, where $\widetilde{A}_3 = A_3 - \partial f/\partial x^3$. Under the gauge transformation, dicussed earlier in Remark 1.1, admitted by the magnetic field lines, we can equivalently consider $\boldsymbol{A} = (0, 0, \widetilde{A}_3)$. Thus, the magnetic field is

$$\boldsymbol{B} = \nabla \widetilde{A}_3 \times \nabla x^3 \tag{1.22}$$

using (A.34), and system (1.21) restricted to the surface $x^3 = \text{const.}$ can be expressed in this case as

$$\begin{aligned} \frac{dx^1}{ds} &= \frac{1}{\sqrt{g}} \frac{\partial \widetilde{A}_3}{\partial x^2} \\ \frac{dx^2}{ds} &= -\frac{1}{\sqrt{g}} \frac{\partial \widetilde{A}_3}{\partial x^1} \end{aligned} \tag{1.23}$$

The above equations are a Hamiltonian system in non-canonical form, where $\widetilde{A}_3$ plays the role of the Hamiltonian function, i.e. $\mathcal{H} = \widetilde{A}_3$, and

$$\mathcal{J}(x^1, x^2) = \frac{1}{\sqrt{g}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{1.24}$$

is the Poisson matrix. In fact, since $\mathcal{H}$ here is independent of $s$, it is a conserved quantity for system (1.23). Consequently, the latter is integrable and its solutions, meaning the magnetic field lines, lie on the intersection of the surfaces $x^3 = c_1$ and $\widetilde{A}_3 = c_2$, where $c_1$ and $c_2$ are constants.

**Remark 1.3.** In the simple case of Cartesian coordinates, where the above system describes a magnetic field in the $xy$-plane, the coordinates $x$ and $y$ are canonical.

### 1.2.2 General magnetic canonical representation

Now consider the more general case, where $B^3 \neq 0$. The procedure to a Hamiltonian description consists of three steps:

*i*) First, using again the gauge equivalence, we can still eliminate one of the three components of the vector potential, say $A_1$, by using the function $f = \int A_1 \, dx^1$. Therefore, $\boldsymbol{A} = (0, \widetilde{A}_2, \widetilde{A}_3) + \nabla f$, where $\widetilde{A}_2 = A_2 - \partial f/\partial x^2$ and $\widetilde{A}_3 = A_3 - \partial f/\partial x^3$, or equivalently $\boldsymbol{A} = (0, \widetilde{A}_2, \widetilde{A}_3)$. From (A.34), now the magnetic field is

$$\boldsymbol{B} = \nabla \widetilde{A}_2 \times \nabla x^2 + \nabla \widetilde{A}_3 \times \nabla x^3 \tag{1.25}$$

*ii*) Next we make the transformation $x^1 \longrightarrow \widetilde{x}^1(x^1, x^2, x^3)$, where $\widetilde{x}^1 = \widetilde{A}_2$. In fact, we already have the inverse transformation $\widetilde{x}^1 = \widetilde{A}_2(x^1, x^2, x^3)$ at hand, which is invertible, since the determinant of the related Jacobian matrix (see (A.2)) equals to $\partial \widetilde{x}^1/\partial x^1 = B^3 \neq 0$. Therefore, we can always replace $x^1$ by $\widetilde{x}^1$, that is, solve the equation $\widetilde{x}^1 = \widetilde{A}_2(x^1, x^2, x^3)$ in terms of $x^1$. Consequently, we may use instead $(\widetilde{x}^1, x^2, x^3)$ as coordinates, in terms of which the magnetic field lines (1.21) are expressed as

$$\begin{aligned} \frac{d\widetilde{x}^1}{ds} &= \frac{1}{\sqrt{\widetilde{g}}} \frac{\partial \widetilde{A}_3}{\partial x^2} \\ \frac{dx^2}{ds} &= -\frac{1}{\sqrt{\widetilde{g}}} \frac{\partial \widetilde{A}_3}{\partial \widetilde{x}^1} \\ \frac{dx^3}{ds} &= \frac{1}{\sqrt{\widetilde{g}}} \end{aligned} \tag{1.26}$$

where $\widetilde{g}$ is now the determinant of the metric defined by the coordinate system $\left(\widetilde{x}^1, x^2, x^3\right)$ and its relation to $g$ of the original system $\left(x^1, x^2, x^3\right)$ is $\sqrt{g} = \sqrt{\widetilde{g}}\,\partial\widetilde{x}^1/\partial x^1$.
*iii*) Finally, instead of $s$ we consider $x^3$ as the independent variable, in terms of which the magnetic field lines are parametrized, and hence (1.26) is transformed to

$$\begin{aligned} \frac{d\widetilde{x}^1}{dx^3} &= \frac{\partial \widetilde{A}_3}{\partial x^2} \\ \frac{dx^2}{dx^3} &= -\frac{\partial \widetilde{A}_3}{\partial \widetilde{x}^1} \end{aligned} \tag{1.27}$$

(1.27) is a Hamiltonian system in canonical variables $\widetilde{x}^1$ and $x^2$, the Hamiltonian function being $\mathcal{H} = \widetilde{A}_3$ again. Quite different from the previous case though, the Poisson matrix $\mathcal{J}$ given by (C.12) simply in two dimensions is now constant, while $\mathcal{H} = \mathcal{H}\left(\widetilde{x}^1, x^2, x^3\right)$ depends on "time" $x^3$. Thus, in general, the Hamiltonian is not conserved along the magnetic field lines nor is system (1.27) a priori integrable.

Returning to our previous discussion about the Hamiltonian formulation of the magnetic field, based on the results of 1.2.1 and 1.2.2, we conclude that :

**Corollary 1.4.** The field lines of any divergence-free magnetic field can be expressed as a Hamiltonian system.

In the next section, we will revisit this statement and explain it in more detail on the geometrical background of Hamiltonian systems shortly presented in appendices B and C.

### 1.2.3 The Frenet triad for the magnetic field lines

In this subsection, we introduce the local reference frame of an observer moving along the magnetic field lines $\boldsymbol{x}(s)$. The vector basis of this frame of reference consists of three vectors, namely the tangent, normal and binormal unit vectors given by the Frenet-Serret formulas with respect to the magnetic field. In terms of (1.18), we have of course already encountered the tangent vector to the curve $\boldsymbol{x}(s)$, being obviously $\boldsymbol{B}$. Divided by the field's strength, we can construct the unit tangent vector

$$\boldsymbol{b} = \frac{\boldsymbol{B}}{|\boldsymbol{B}|} = \frac{d\boldsymbol{x}}{dl}\,. \tag{1.28}$$

Here $l$, called the natural parameter, is a reparametrization of the magnetic field lines for which $l'(s) = |\boldsymbol{x}'(s)|$ holds. Continuing, the curvature vector is defined as the derivative of the unit tangent vector with respect to the natural parameter,

$$\boldsymbol{K} = \frac{d\boldsymbol{b}}{dl} = \frac{d^2\boldsymbol{x}}{dl^2}\,, \tag{1.29}$$

which is perpendicular to $\boldsymbol{b}$, for $2\,\boldsymbol{b}\cdot\boldsymbol{K} = 2\,\boldsymbol{b}\cdot\boldsymbol{b}' = (\boldsymbol{b}\cdot\boldsymbol{b})' = \left(|\boldsymbol{b}|^2\right)' = 0$. The normalisation of $\boldsymbol{K}$ in turn yields the (principal) normal

$$\boldsymbol{n} = \frac{\boldsymbol{K}}{|\boldsymbol{K}|} = \varkappa\boldsymbol{K}, \tag{1.30}$$

where $\varkappa = 1/|\boldsymbol{K}|$ is the curvature radius. Finally, in order to have a right-handed orthogonal coordinate system, the binormal vector is defined as

$$\boldsymbol{\beta} = \boldsymbol{b}\times\boldsymbol{n}\,. \tag{1.31}$$

The same argument used above shows that $d\boldsymbol{\beta}/dl$ is perpendicular to $\boldsymbol{\beta}$. In fact, it is parallel to $\boldsymbol{n}$, since $\boldsymbol{\beta}'\cdot\boldsymbol{b} = (\boldsymbol{\beta}\cdot\boldsymbol{b})' - \boldsymbol{\beta}\cdot\boldsymbol{b}' = -\,\boldsymbol{\beta}\cdot\boldsymbol{n} = 0$. Therefore, $d\boldsymbol{\beta}/dl = -\,\boldsymbol{n}/\sigma$, where the function $\sigma$ is the torsion radius.

Following [4], we make the following remarks, which will turn up useful when studying the guiding-center motion later on this chapter.

**Remark 1.5.** From the definition of the curvature vector (1.29), we can also obtain the following expression

$$\boldsymbol{K} = \frac{d\boldsymbol{b}}{dl} = \frac{\partial\boldsymbol{b}}{\partial x^i}\frac{dx^i}{dl} = b^i\frac{\partial\boldsymbol{b}}{\partial x^i} = (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{b} \tag{1.32}$$

**Remark 1.6.** Following the same steps as in the previous expression, we can obtain $d\boldsymbol{\beta}/dl = (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{\beta}$ and since $\sigma^{-1} = -\,\boldsymbol{n}\cdot d\boldsymbol{\beta}/dl$, we see that the torsion radius is

$$\sigma = -\,\frac{1}{\boldsymbol{n}\cdot[(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{\beta}]} \tag{1.33}$$

**Remark 1.7.** In the case of a straight, yet inhomogeneous magnetic field, we can see that $\boldsymbol{K} = 0$, which means that the curvature radius becomes infinite. Therefore, neither $\boldsymbol{n}$ nor $\boldsymbol{\beta}$ can be defined, they are arbitrary. A natural way to treat this case is to consider instead the only intrinsic direction left that is fixed by the field, which is the gradient of the field's strength. In order then to have an orthonormal basis, we define

$$\widetilde{\boldsymbol{n}} = \frac{\nabla|\boldsymbol{B}|}{|\nabla|\boldsymbol{B}||} - \left(\boldsymbol{b}\cdot\frac{\nabla|\boldsymbol{B}|}{|\nabla|\boldsymbol{B}||}\right)\boldsymbol{b} \tag{1.34}$$

$$\widetilde{\boldsymbol{\beta}} = \boldsymbol{b}\times\widetilde{\boldsymbol{n}} \tag{1.35}$$

**Remark 1.8.** If, however, the magnetic field is straight and parallel to the gradient of its strength, i.e. $\boldsymbol{K} = \widetilde{\boldsymbol{\beta}} = 0$, or homogeneous, then besides $\boldsymbol{b}$ itself, there is no other direction to characterize the magnetic field. We can therefore choose an appropriate coordinate system, where one of its axes coincides with $\boldsymbol{B}$.

## 1.3 Geometrical aspects

The structure of the magnetic field has an intriguing geometrical interpretation, giving insight to many of the features we met so far. The geometry of the magnetic field is described in this section in the language of differential and symplectic forms, briefly summarized in appendix B, which we will often refer to. We start off by the vector potential $\boldsymbol{A}$, viewed though as an 1-form in $\mathbb{R}^3$ in terms of curvilinear coordinates $\left(x^1, x^2, x^3\right)$

$$A = A_1 dx^1 + A_2 dx^2 + A_3 dx^3. \tag{1.36}$$

where the switch from $A_i$ to the components $A^i$ of the vector function $\boldsymbol{A}$ is described in (A.23). From the differential of $A$, given in (B.43) we can construct the following 2-form

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

Taking the Hodge star operator, we can define the magnetic field $\boldsymbol{B}$ according to (see (B.38)-(B.40) and the resulting equation (B.44) restated here)

$$\boldsymbol{B} = *B = \frac{1}{\sqrt{g}} \left\{ \left(\frac{\partial A_3}{\partial x^2} - \frac{\partial A_2}{\partial x^3}\right) \frac{\partial}{\partial x^1} + \left(\frac{\partial A_1}{\partial x^3} - \frac{\partial A_3}{\partial x^1}\right) \frac{\partial}{\partial x^2} + \left(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2}\right) \frac{\partial}{\partial x^3} \right\} \tag{1.37}$$

which precisely expresses the curl of the vector potential $\boldsymbol{A}$ in curvilinear coordinates according to (B.45).

As shown in appendix B and in particular (B.41), the divergence-free condition for the magnetic field $\boldsymbol{B}$ is equivalent to the closure property of the corresponding exact 2-form $B$, i.e. $\nabla \cdot \boldsymbol{B} = 0 \Leftrightarrow dB = 0$. Recalling Definition B.16, we immediately deduce that:

**Corollary 1.9.** Every divergence-free magnetic field $\boldsymbol{B}$ defined on some region of $\mathbb{R}^3$ corresponds to a presymplectic form $B$.

**Remark 1.10.** So far from the beginning of this chapter, we have taken for granted the existence of the vector potential as a solution to the divergence-free condition of the magnetic field, being one of Maxwell's equations (1.4a). Using (B.42) and (B.45), $\nabla \cdot \boldsymbol{B} = 0 \Rightarrow \boldsymbol{B} = \nabla \times \boldsymbol{A}$ translated to forms means that $dB = 0 \Rightarrow B = dA$. This assumption is justified when $\boldsymbol{B}$ is defined in (all of) $\mathbb{R}^3$, where every closed form is also exact (see Remark B.7). Therefore the above construction could be inverted, and start off more naturally with the real physical quantity, meaning the magnetic field: since $\boldsymbol{B}$ obeys the divergence-free condition it corresponds to a closed 2-form $B$, which in turn is coming from the differential of a 1-form $A$.

**Remark 1.11.** Regardless of the existence of the vector potential, we emphasize that, as already stated, Corollary 1.9 holds for any subspace of $\mathbb{R}^3$, simply-connected or not.

In light of the presymplectic form associated to every divergence-free magnetic field, the Hamiltonian nature of the magnetic field lines we saw in the previous section can now be explained. Let us see how the gap between Corollaries 1.4 and 1.9 can be filled.

### The symplectic structure of magnetic surfaces

First of all, in the case of magnetic surfaces §1.2.1 where $B^3 = 0$, the system of magnetic field lines reduces to the surface $x^3 = \text{const.}$, described by a two-dimensional system in terms of $x^1$ and $x^2$. And as we saw in Example B.13 and then again near the end of appendix C, a surface in $\mathbb{R}^3$ can always be equipped with the symplectic form $\omega = \sqrt{g}\, dx^2 \wedge dx^1$ giving rise to the Poisson matrix (C.14) of the system. The latter in none other than $\mathcal{J}$ in (1.24). Therefore, the problem reduces to a manifold that can always be regarded as symplectic, without involving the magnetic field whatsoever. Taking $A_3(x^1, x^2)$ as the Hamiltonian, we then arrive at the Hamiltonian system (1.23) for the magnetic field. Moreover, in this case the system is completely integrable due to the conservation of the Hamiltonian, $A_3 = \text{const.}$.

### The symplectic structure of the magnetic field in general

On the other hand, in the more general problem of §1.2.2, the magnetic field enters the symplectic structure. In this case, Proposition B.21 plays a crucial role.

To begin with, note that in terms of the equivalent potential and the variables introduced there the 1-form expressing the vector potential is $A = \widetilde{x}^1 dx^2 + \widetilde{A}_3 dx^3$ and the related presymplectic form is $B = d\widetilde{x}^1 \wedge dx^2 + (\partial \widetilde{A}_3/\partial \widetilde{x}^1)\, d\widetilde{x}^1 \wedge dx^3 + (\partial \widetilde{A}_3/\partial x^2)\, dx^2 \wedge dx^3$, both considered on the $\widetilde{x}^1 x^2 x^3$-space denoted by $M$. But since $B^3 \neq 0$, we can always (use the rectification lemma and) consider that the magnetic field lies on the $x^3$-axis. In other words, we can find new variables $(y^1, y^2, x^3)$ such that $\boldsymbol{B}(y^1) = \boldsymbol{B}(y^2) = 0$, while $\boldsymbol{B}(x^3) = 1$ is already set. In this coordinate system, $B = dy^1 \wedge dy^2$ is expressed in canonical form.

Following the steps of the proposition's proof, we construct then the dual $E^*$ of the characteristic line bundle $E$, where the latter is defined in (B.56) by the kernel of $B$. According to Remark B.19, sections of $E$ are vector fields parallel to the magnetic field and therefore multiples of $\partial_{x^3}$. Consequently sections of the dual bundle $E^*$ are 1-forms expressed in the dual basis $dx^3$. Thus, if $v$ is a 1-form on $M$, the generalized Liouville 1-form (B.58) on $E^*$ is $\vartheta_\lambda = v_3 dx^3$. Since the

presymplectic form $B$ of $M$ is of constant rank two, we can apply Proposition B.21 and define on $E^*$ the symplectic form (B.59), for which $\omega_\lambda = dv_3 \wedge dx^3$, leading to

$$\widetilde{\omega} = dv_3 \wedge dx^3 + \pi^*(B) \tag{1.38}$$

Let us see how the induced symplectic form can indeed explain the Hamiltonian structure of equations (1.27). First of all, to recover the dynamics of the magnetic field it is natural to consider the 1-form $A$ for the magnetic potential instead of any other $v$. In addition, we need of course to go back to the variables of §1.2.2. In other words, instead of $(y^1, y^2, x^3, v_3)$ we can use $(\widetilde{x}^1, x^2, x^3, \widetilde{a}_3)$ as local coordinates on $E^*$, where $\widetilde{a}_3$ is the value of the function $\widetilde{A}_3(\widetilde{x}^1, x^2, x^3)$ on $\mathbb{R}$. That is to say, take $\omega_\lambda = d\widetilde{a}_3 \wedge dx^3$. And then, when $B$ is dragged under $\pi^*$ from $M$ to $E^*$, $\widetilde{A}^3$ can no longer be treated as a function, meaning $\pi^*(B) = d\widetilde{x}^1 \wedge dx^2$ in these coordinates. Hence the above symplectic form (within a sign) is none other than

$$\widetilde{\omega} = dx^2 \wedge d\widetilde{x}^1 - d\widetilde{a}_3 \wedge dx^3 \tag{1.39}$$

Now, $\widetilde{\omega}$ in (1.39) can be viewed as the extended symplectic structure (C.18) on the extended four-dimensional phase space $E^*$ of the canonical symplectic form $\omega = dx^2 \wedge d\widetilde{x}^1$ on the two-dimensional $\widetilde{x}^1 x^2$-space. Just as $\widetilde{\mathcal{H}} = \widetilde{A}_3 - \widetilde{a}_3$ is the extended Hamiltonian (C.16) of the $x^3$-dependent Hamiltonian $\mathcal{H} = \widetilde{A}_3$ that describes the dynamics of the system on the hypersurface $\widetilde{\mathcal{H}} = 0$. In conclusion, $\widetilde{\omega}$ and $\widetilde{\mathcal{H}}$ are precisely the extended version of the Hamiltonian structure of the non-autonomous system (1.27) given in terms of the canonical variables $(\widetilde{x}^1, x^2)$ and the "time"-dependent Hamiltonian function $\widetilde{A}_3$.

Under the prism of presymplectic geometry, let us also revisit the widely used *Clebsch representation*, which states that for every divergence-free magnetic field there exist functions $a$ and $b$ such that $\boldsymbol{B} = \nabla a \times \nabla b$. It first appeared in [66] and in [38], where Grad and Rubin gave also a proof, which can be found more detailed in [28], section 5.2, as well. However, exploiting once more the correspondence between a divergence-free magnetic field $\boldsymbol{B}$ and a presymplectic form $B$ in terms of the 1-1 mapping (B.46) of the wedge to the cross product, the above statement expresses differently that $B = da \wedge db$. But this is just Darboux' theorem B.18 for presymplectic forms in the case of $\mathbb{R}^3$, providing canonical coordinates $a$ and $b$ for $B$, as stated in (B.55). On this more solid ground, a word of caution is in order. According to Darboux' theorem, the existence of $a$ and $b$ is in general only guaranteed locally. Although not emphasized, this is also mentioned in [38] unfortunately before the statements of the related theorems I and II of the appendix I, by considering "small regions of spaces". Only in certain cases, where magnetic surfaces exist, such a representation can be defined globally, as, for example, the magnetic field

(1.22) in 1.2.1. We should note that the magnetic surface requirement was also used in [66]. On the other hand, when dropped (or replaced by the existence of only locally defined magnetic surfaces one might say) as in [38], it leads to local considerations, i.e. $a$ and $b$ can be defined only locally, and they may be multivalued functions, as well, as opposed to the well behaved, single valued functions globally defined in the magnetic surface case.

### The symplectic structure of charged particle motion

Last but not least, Proposition B.20 gives a connection between the presymplectic form $B$ of the magnetic field lines and the symplectic form $\omega$ of the charged particle motion. For, by construction (see the related proof), the presymplectic Euclidean space $\mathbb{R}^3$ equipped with $B$ can be naturally embedded to the cotangent bundle $T^*\mathbb{R}^3 = \mathbb{R}^6$ inducing the symplectic form (B.57)

$$\omega = dv_i \wedge dx^i + B \tag{1.40}$$

(where the projection from $T^*\mathbb{R}^3$ to the base manifold $\mathbb{R}^3$ is here omitted). On the other hand, the Poisson matrix $J$ of the particle's motion given in (1.13) is nondegenerate, and in light of Proposition C.7 its inverse is the corresponding symplectic structure matrix for the Hamiltonian system (1.15) in $\mathbb{R}^6$. But the latter is no other than the matrix $\Omega = (\omega_{ij})$ of the above symplectic form $\omega$ (1.40), that is $\Omega = J^{-1}$. Consequently, the equations of the charged particle motion are described by the symplectic structure $\omega$ introduced by the presymplectic form $B$ of the magnetic field through (1.40).

## 1.4 Parallel and perpendicular motion

Returning to the problem of the particle's motion, another useful set of variables is a "cylindrical" coordinate system for the velocity. These are the parallel and perpendicular velocities with respect to the direction of the magnetic field, along with a third angle variable, widely known as the *gyrophase*. Although conceptually simple, these variables may lead to very complicated expressions, since the "cylinder" magnetic axis varies in space. The Hamiltonian formulation, however, provides a rather easy and consistent way, which we can rely on to describe the problem in terms of them. This set of velocity coordinates is actually the doorway to the guiding-center variables, introduced later on. The Hamiltonian way of presenting them has also the advantage of inducing the Hamiltonian version of the guiding-center motion.

We start off by considering the unit direction along the magnetic field $\boldsymbol{b}$, and we decompose the velocity vector $\boldsymbol{v}$ into two directions, parallel and perpendicular to $\boldsymbol{b}$, denoted by $\boldsymbol{v}_\parallel$ and $\boldsymbol{v}_\perp$,

respectively,

$$\boldsymbol{v} = \boldsymbol{v}_{\parallel} + \boldsymbol{v}_{\perp} = v_{\parallel}\boldsymbol{b} + v_{\perp}\boldsymbol{n}_1 \tag{1.41}$$

where $v_{\parallel} = \boldsymbol{v} \cdot \boldsymbol{b}$ and $v_{\perp} = (v^2 - v_{\parallel}^2)^{\frac{1}{2}}$. A third direction can be defined that is perpendicular to both $\boldsymbol{b}$ and $\boldsymbol{n}_1$ as $\boldsymbol{n}_2 = \boldsymbol{b} \times \boldsymbol{n}_1$, so that $(\boldsymbol{b}, \boldsymbol{n}_1, \boldsymbol{n}_2)$ is a right-handed orthogonal coordinate system. Finally, the gyrophase $\zeta$ can be defined as the angle between $\boldsymbol{n}_1$ the normal vector $\boldsymbol{n}$ of the magnetic field lines. The new unit vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are related to the the normal and binormal vectors $\boldsymbol{n}$ and $\boldsymbol{\beta}$, respectively, of the magnetic field lines as

$$\begin{aligned} \boldsymbol{n}_1(\boldsymbol{x}, \zeta) &= \cos\zeta\, \boldsymbol{n}(\boldsymbol{x}) + \sin\zeta\, \boldsymbol{\beta}(\boldsymbol{x}) \\ \boldsymbol{n}_2(\boldsymbol{x}, \zeta) &= -\sin\zeta\, \boldsymbol{n}(\boldsymbol{x}) + \cos\zeta\, \boldsymbol{\beta}(\boldsymbol{x}) \end{aligned} \tag{1.42}$$

Inserting $\boldsymbol{n}_1$ in (1.41) and dot multiplying with $\boldsymbol{n}$, we derive $v_{\perp}\cos\zeta = \boldsymbol{v} \cdot \boldsymbol{n}$, from which we can determine the angle $\zeta$.

In order to recover the equation of motion in terms of the previously used space coordinates and the above defined velocity coordinates $\left(\boldsymbol{x}, v_{\parallel}, v_{\perp}, \zeta\right)$ it suffices to find the Hamiltonian function and Poisson matrix in these new variables. The former is easily seen to be

$$H(\boldsymbol{x}, v_{\parallel}, v_{\perp}) = \frac{1}{2}\left(v_{\parallel}^2 + v_{\perp}^2\right) + \Phi(\boldsymbol{x}), \tag{1.43}$$

while for the latter we work as before, meaning calculate the Poisson brackets for the new variables using the bracket (1.16) in terms of the old ones $(\boldsymbol{x}, \boldsymbol{v})$. Only this time, since calculations are much more lengthy, we can also make use of the properties of the Poisson bracket to reduce them.

Obviously $\{x_i, x_j\} = 0$ once again, and note that for this bracket any two functions independent of the velocities in general satisfy $\{F(\boldsymbol{x}), G(\boldsymbol{x})\} = 0$. Next,

$$\left\{x_i, v_{\parallel}\right\} = \{x_i, v_j b_j(\boldsymbol{x})\} = \{x_i, v_j\}\, b_j + \{x_i, b_j(\boldsymbol{x})\}\, v_j = \delta_{ij} b_j = b_i\,.$$

Moving on, from Leibniz' rule (see Definition C.2), we have that $\left\{x_i, v_{\parallel}^2\right\} = 2v_{\parallel}\left\{x_i, v_{\parallel}\right\} = 2v_{\parallel}b_i$ and $\left\{x_i, v^2\right\} = \{x_i, v_j v_j\} = 2v_j\left\{x_i, v_j\right\} = 2v_j\delta_{ij} = 2v_i$, which help us arrive without complicated calculations at

$$\begin{aligned} \{x_i, v_{\perp}\} &= \frac{1}{2v_{\perp}}\left\{x_i, v_{\perp}^2\right\} = \frac{1}{2v_{\perp}}\left\{x_i, v^2 - v_{\parallel}^2\right\} = \frac{1}{2v_{\perp}}\left[\left\{x_i, v^2\right\} - \left\{x_i, v_{\parallel}^2\right\}\right] \\ &= \frac{1}{2v_{\perp}}\left(2v_i - 2v_{\parallel}b_i\right) = n_{1_i}. \end{aligned}$$

In a similar way, we work our way for the angle $\zeta$, for which, using property C.10, we note first $\{x_i, v_{\perp}\cos\zeta\} = \cos\zeta\,\{x_i, v_{\perp}\} + v_{\perp}\,\{x_i, \cos\zeta\} = \cos\zeta\,\{x_i, v_{\perp}\} - v_{\perp}\sin\zeta\,\{x_i, \zeta\}$. Therefore, with

the help of the previously found bracket as well,

$$\begin{aligned}\{x_i,\zeta\} &= \frac{1}{v_\perp\sin\zeta}\left[\cos\zeta\,\{x_i,v_\perp\}-\{x_i,\boldsymbol{v}\cdot\boldsymbol{n}\}\right] = \frac{1}{v_\perp\sin\zeta}\left[\cos\zeta\,n_{1_i}-\{x_i,v_j\}\,n_j-\{x_i,n_j\}\,v_j\right]\\ &= \frac{1}{v_\perp\sin\zeta}\left(\cos\zeta\,n_{1_i}-\delta_{ij}n_j\right) = \frac{1}{v_\perp\sin\zeta}\left(\cos^2\zeta\,n_i+\cos\zeta\sin\zeta\,\beta_i-n_i\right) = \frac{1}{v_\perp}\,n_{2_i}.\end{aligned}$$

The next three brackets between the cylindrical velocity coordinates require much more lengthy calculations in order to derive expressions in terms of these new variables. Nevertheless, a point that needs to be made clear may save us from going in circles: until the brackets are found the new variables $(v_\parallel, v_\perp, \zeta)$ are treated as functions of the old ones $(\boldsymbol{x},\boldsymbol{v})$; but from the moment we lose the brackets the old variables $\boldsymbol{v}$ are then considered as functions of the new ones $(\boldsymbol{x}, v_\parallel, v_\perp, \zeta)$, which of course are now independent of each other. To stress this change in treatment we denote the turning point by the letter $N$ above the equality sign. The final form of the expessions admits a nice representation if we also use the vector $\boldsymbol{\delta} = v_\parallel\nabla\times\boldsymbol{b}+v_\perp\nabla\times\boldsymbol{n}_1$, which is simply the curl of the velocity in the new coordinate system.

$$\begin{aligned}\{v_\parallel,v_\perp\} &= \frac{1}{2v_\perp}\{v_\parallel,v_\perp^2\} = \frac{1}{2v_\perp}\left\{v_\parallel,v^2-v_\parallel^2\right\} = \frac{1}{2v_\perp}\{v_\parallel,v^2\} = \frac{v_j}{v_\perp}\{v_\parallel,v_j\} = \frac{v_j}{v_\perp}\{v_ib_i,v_j\}\\ &= \frac{v_j}{v_\perp}\left[v_i\{b_i,v_j\}+b_i\{v_i,v_j\}\right] = \frac{v_j}{v_\perp}\left[v_i\frac{\partial b_i}{\partial x_j}+\epsilon_{ijk}b_iB_k\right] = \frac{1}{v_\perp}v_iv_j\frac{\partial b_i}{\partial x_j}\\ &\overset{N}{=} \frac{1}{v_\perp}v_j\left[\frac{\partial(v_ib_i)}{\partial x_j}-b_i\frac{\partial v_i}{\partial x_j}\right] = -\frac{1}{v_\perp}b_iv_j\frac{\partial v_i}{\partial x_j} = -\frac{1}{v_\perp}\boldsymbol{b}\cdot[(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}]\\ &= -\frac{1}{v_\perp}\boldsymbol{b}\cdot\left[\frac{1}{2}\nabla|\boldsymbol{v}|^2-\boldsymbol{v}\times(\nabla\times\boldsymbol{v})\right] = -\frac{1}{v_\perp}\boldsymbol{b}\cdot\left[\frac{1}{2}\nabla\left(v_\parallel^2+v_\perp^2\right)-\boldsymbol{v}\times(\nabla\times\boldsymbol{v})\right]\\ &= \frac{1}{v_\perp}(\boldsymbol{b}\times\boldsymbol{v})\cdot(\nabla\times\boldsymbol{v}) = (\boldsymbol{b}\times\boldsymbol{n}_1)\cdot\boldsymbol{\delta} = \boldsymbol{n}_2\cdot\boldsymbol{\delta}\end{aligned}$$

For the next bracket, we employ first $\{v_\parallel, v_\perp\cos\zeta\} = \cos\zeta\,\{v_\parallel,v_\perp\}-v_\perp\sin\zeta\,\{v_\parallel,\zeta\}$ just like before, and also recall from the previous calculation that $\{v_\parallel,v_j\} = v_i\partial b_i/\partial x_j$. Then, using the same vector identities (including (A.36) towards the end) and similar techniques,

$$\begin{aligned}\{v_\parallel,\zeta\} &= \frac{1}{v_\perp\sin\zeta}\left[\cos\zeta\,\{v_\parallel,v_\perp\}-\{v_\parallel,\boldsymbol{v}\cdot\boldsymbol{n}\}\right]\\ &= \frac{1}{v_\perp\sin\zeta}\left[\cos\zeta\,\boldsymbol{n}_2\cdot\boldsymbol{\delta}-\{v_\parallel,v_j\}\,n_j-\{v_\parallel,n_j\}\,v_j\right]\\ &= \frac{1}{v_\perp\sin\zeta}\left[\cos\zeta\,\boldsymbol{n}_2\cdot\boldsymbol{\delta}-v_in_j\frac{\partial b_i}{\partial x_j}-v_iv_j\{b_i,n_j\}-b_iv_j\{v_i,n_j\}\right]\\ &= \frac{1}{v_\perp\sin\zeta}\left[\cos\zeta\,\boldsymbol{n}_2\cdot\boldsymbol{\delta}-v_in_j\frac{\partial b_i}{\partial x_j}+b_iv_j\frac{\partial n_j}{\partial x_i}\right]\\ &\overset{N}{=} \frac{1}{v_\perp\sin\zeta}\left[\cos\zeta\,\boldsymbol{n}_2\cdot\boldsymbol{\delta}+b_in_j\frac{\partial v_i}{\partial x_j}+b_i\frac{\partial(v_jn_j)}{\partial x_i}-b_in_j\frac{\partial v_j}{\partial x_i}\right]\end{aligned}$$

$$= \frac{1}{v_\perp \sin\zeta}\left[\cos\zeta\, \boldsymbol{n}_2\cdot\boldsymbol{\delta} - b_i n_j\left(\frac{\partial v_j}{\partial x_i} - \frac{\partial v_i}{\partial x_j}\right)\right]$$
$$= \frac{1}{v_\perp \sin\zeta}\left[\cos\zeta\, \boldsymbol{n}_2\cdot\boldsymbol{\delta} - \boldsymbol{b}\cdot[\boldsymbol{n}\times(\nabla\times\boldsymbol{v})]\right] = \frac{1}{v_\perp \sin\zeta}\left[\cos\zeta\, \boldsymbol{n}_2\cdot\boldsymbol{\delta} - (\boldsymbol{b}\times\boldsymbol{n})\cdot(\nabla\times\boldsymbol{v})\right]$$
$$= \frac{1}{v_\perp \sin\zeta}\left(\cos\zeta\, \boldsymbol{n}_2\cdot\boldsymbol{\delta} - \boldsymbol{\beta}\cdot\boldsymbol{\delta}\right) = \frac{1}{v_\perp \sin\zeta}\left(\cos^2\zeta\,\boldsymbol{\beta} - \cos\zeta\sin\zeta\,\boldsymbol{n} - \boldsymbol{\beta}\right)\cdot\boldsymbol{\delta} = -\frac{1}{v_\perp}\,\boldsymbol{n}_1\cdot\boldsymbol{\delta}$$

For the last bracket, consider $\left\{v_\perp^2, \boldsymbol{v}\cdot\boldsymbol{n}\right\} = 2v_\perp\left\{v_\perp, v_\perp\cos\zeta\right\} = -2v_\perp^2\sin\zeta\left\{v_\perp,\zeta\right\}$, recalling from the previous one that $\left\{v_\parallel, \boldsymbol{v}\cdot\boldsymbol{n}\right\} = \boldsymbol{\beta}\cdot\boldsymbol{\delta}$, too. Therefore,

$$\begin{aligned}
\{v_\perp,\zeta\} &= -\frac{1}{2v_\perp^2\sin\zeta}\left\{v_\perp^2, \boldsymbol{v}\cdot\boldsymbol{n}\right\} = -\frac{1}{2v_\perp^2\sin\zeta}\left\{v^2 - v_\parallel^2, \boldsymbol{v}\cdot\boldsymbol{n}\right\} \\
&= \frac{1}{2v_\perp^2\sin\zeta}\left[\left\{v_\parallel^2, \boldsymbol{v}\cdot\boldsymbol{n}\right\} - \left\{v^2, \boldsymbol{v}\cdot\boldsymbol{n}\right\}\right] = \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\left\{v_\parallel, \boldsymbol{v}\cdot\boldsymbol{n}\right\} - v_i\left\{v_i, v_j n_j\right\}\right] \\
&= \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\boldsymbol{\beta}\cdot\boldsymbol{\delta} - v_i\left(n_j\left\{v_i, v_j\right\} - v_j\left\{v_i, n_j\right\}\right)\right] \\
&= \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\boldsymbol{\beta}\cdot\boldsymbol{\delta} - v_i\left(\epsilon_{ijk}n_j B_k - v_j\frac{\partial n_j}{\partial x_i}\right)\right] \\
&\overset{N}{=} \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\boldsymbol{\beta}\cdot\boldsymbol{\delta} + |\boldsymbol{B}|\,\boldsymbol{v}\cdot\boldsymbol{\beta} - v_i n_j\frac{\partial v_j}{\partial x_i}\right] \\
&= \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\boldsymbol{\beta}\cdot\boldsymbol{\delta} + v_\perp\sin\zeta\,|\boldsymbol{B}| - \boldsymbol{n}\cdot[(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}]\right] \\
&= \frac{|\boldsymbol{B}|}{v_\perp} + \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\boldsymbol{\beta}\cdot\boldsymbol{\delta} + \boldsymbol{n}\cdot[\boldsymbol{v}\times(\nabla\times\boldsymbol{v})]\right] \\
&= \frac{|\boldsymbol{B}|}{v_\perp} + \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\boldsymbol{\beta}\cdot\boldsymbol{\delta} + (\boldsymbol{n}\times\boldsymbol{v})\cdot(\nabla\times\boldsymbol{v})\right]\Big] \\
&= \frac{|\boldsymbol{B}|}{v_\perp} + \frac{1}{v_\perp^2\sin\zeta}\left[v_\parallel\boldsymbol{\beta}\cdot\boldsymbol{\delta} + \left(v_\parallel\,\boldsymbol{n}\times\boldsymbol{b} + v_\perp\,\boldsymbol{n}\times\boldsymbol{n}_1\right)\cdot\boldsymbol{\delta}\right]\Big] = \frac{1}{v_\perp}\left(\boldsymbol{b}\cdot\boldsymbol{\delta} + \omega_g\right)
\end{aligned}$$

where we set $\omega_g = |\boldsymbol{B}|$ meant only for the last bracket. For if we restore our equations to physical units, the second term in the last parenthesis would be the *Larmor frequency* or *gyrofrequency* $\omega_g = q|\boldsymbol{B}|/m$, while

$$\rho = \frac{v_\perp}{\omega_g} \tag{1.44}$$

is known as the *Larmor radius* or *gyroradius*. The nature of these quantities that justifies their names would be apparent in the next section.

Collecting all the above brackets, we can construct the Poisson matrix in terms of the variables $\left(\boldsymbol{x}, v_\parallel, v_\perp, \zeta\right)$

$$J\left(\boldsymbol{x}, v_\parallel, v_\perp, \zeta\right) = \begin{pmatrix} O & T \\ -T & D \end{pmatrix}, \tag{1.45}$$

where $T = T(\boldsymbol{x}, v_\perp) = (\boldsymbol{b}, \boldsymbol{n}_1, \boldsymbol{n}_2/v_\perp)$ is the matrix with columns $\boldsymbol{T}_i$ given by the cylindrical velocity directions, while $D = D\left(\boldsymbol{x}, v_\parallel, v_\perp, \zeta\right)$ is the matrix with elements $D_{ij} = \boldsymbol{T}_i\cdot(\boldsymbol{T}_j\times\boldsymbol{\delta}) +$

$\epsilon_{ijk}\delta_{1k}\rho^{-1}$. The last relation follows from the last three brackets when expressed as triple products, for example $\{v_\parallel, v_\perp\} = \boldsymbol{n}_2 \cdot \boldsymbol{\delta} = (\boldsymbol{b} \times \boldsymbol{n}_1) \cdot \boldsymbol{\delta} = \boldsymbol{b} \cdot (\boldsymbol{n}_1 \times \boldsymbol{\delta})$. The resulting equations of motion written as a Hamiltonian system (C.2) in terms of (1.43) and (1.45) are

$$\begin{aligned}
\frac{dx_i}{dt} &= b_i \frac{\partial H}{\partial v_\parallel} + n_{1_i} \frac{\partial H}{\partial v_\perp} \\
\frac{dv_\parallel}{dt} &= -\ b_i \frac{\partial H}{\partial x_i} + \boldsymbol{n}_2 \cdot \boldsymbol{\delta} \frac{\partial H}{\partial v_\perp} \\
\frac{dv_\perp}{dt} &= -\ n_{1_i} \frac{\partial H}{\partial x_i} - \boldsymbol{n}_2 \cdot \boldsymbol{\delta} \frac{\partial H}{\partial v_\parallel} \\
\frac{d\zeta}{dt} &= \left[ - n_{2_i} \frac{\partial H}{\partial x_i} + \boldsymbol{n}_1 \cdot \boldsymbol{\delta} \frac{\partial H}{\partial v_\parallel} - (\omega_g + \boldsymbol{b} \cdot \boldsymbol{\delta}) \frac{\partial H}{\partial v_\perp} \right] \frac{1}{v_\perp}
\end{aligned} \tag{1.46}$$

Comparing the above Poisson matrix to (1.13), we see that the identity matrix has been replaced by the matrix $T$ consisting of the velocity directions, while $B$ (1.14) has been replaced by $D$. As a consequence (1.45) now depends apart from the positions on the velocities as well, as opposed to (1.13). We note however that besides $J^{v_\parallel v_\perp} = \{v_\parallel, v_\perp\}$ and $J^{v_\parallel \zeta} = \{v_\parallel, \zeta\}$, the rest of the Poisson matrix as well as the Hamiltonian are independent of the gyrophase $\zeta$. Still, it cannot qualify as an *ignorable variable*, but system (1.46), although complicated, points the way towards the concept of the guiding center that averages out this angle. Before entering this theory, in the next section we study a simple case, which nonetheless has the advantages of a clear picture of several of the above concepts as well as an intuitive description of the guiding-center notion.

Closing this section, we substitute the Hamiltonian (1.43) in (1.46) and finally see that the equations of charged particle motion in terms of this rather mixed set of variables, consisting of Cartesian position coordinates $\boldsymbol{x}$ and cylindrical velocity coordinates $(v_\parallel, v_\perp, \zeta)$ with respect to the magnetic field, are of the form

$$\frac{d\boldsymbol{x}}{dt} = v_\parallel \boldsymbol{b} + v_\perp \boldsymbol{n}_1 \tag{1.47a}$$

$$\frac{dv_\parallel}{dt} = \boldsymbol{E} \cdot \boldsymbol{b} \ + v_\perp\, \boldsymbol{n}_2 \cdot \boldsymbol{\delta} \tag{1.47b}$$

$$\frac{dv_\perp}{dt} = \boldsymbol{E} \cdot \boldsymbol{n}_1 - v_\parallel\, \boldsymbol{n}_2 \cdot \boldsymbol{\delta} \tag{1.47c}$$

$$\frac{d\zeta}{dt} = \left(\boldsymbol{E} \cdot \boldsymbol{n}_2 + v_\parallel\, \boldsymbol{n}_1 \cdot \boldsymbol{\delta}\right) \frac{1}{v_\perp} - \omega_g - \boldsymbol{b} \cdot \boldsymbol{\delta} \tag{1.47d}$$

And last but not least, a Lagrangian formulation for this system is also at our disposal, inherited from (1.17) and given by the Lagrangian function

$$L\big(\boldsymbol{x}, v_\parallel, v_\perp, \zeta, \dot{\boldsymbol{x}}\big) = \big(\boldsymbol{A}(\boldsymbol{x}) + v_\parallel \boldsymbol{b}(\boldsymbol{x}) + v_\perp \boldsymbol{n}_1(\boldsymbol{x}, \zeta)\big) \cdot \dot{\boldsymbol{x}} - \frac{1}{2}\,(v_\parallel^2 + v_\perp^2) - \Phi(\boldsymbol{x}) \tag{1.48}$$

We notice the $\zeta$-dependence of $L$ through $\boldsymbol{n}_1$, and, as in (1.17), the independence of the derivatives of the velocities. Interestingly enough, the Euler-Lagrange equations that correspond to (1.48) express system (1.47) indirectly. Practicing for the guiding-center equations, let us see how. First of all,

$$\begin{aligned}
\frac{\partial L}{\partial v_\parallel} = 0 \Rightarrow \boldsymbol{b}\cdot\dot{\boldsymbol{x}} - v_\parallel = 0 \Rightarrow \dot{\boldsymbol{x}}\cdot\boldsymbol{b} = v_\parallel \\
\frac{\partial L}{\partial v_\perp} = 0 \Rightarrow \boldsymbol{n}_1\cdot\dot{\boldsymbol{x}} - v_\perp = 0 \Rightarrow \dot{\boldsymbol{x}}\cdot\boldsymbol{n}_1 = v_\perp \\
\frac{\partial L}{\partial \zeta} = 0 \Rightarrow v_\perp \frac{\partial \boldsymbol{n}_1}{\partial \zeta}\cdot\dot{\boldsymbol{x}} = 0 \Rightarrow \dot{\boldsymbol{x}}\cdot\boldsymbol{n}_2 = 0
\end{aligned}$$

describe the components of the vector $\dot{\boldsymbol{x}}$ in the three coordinate axes $\boldsymbol{b}$, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, which then have as an oblique result the expression (1.47a). The other set of equations,

$$\begin{aligned}
0 &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = \frac{d}{dt}\left(A_i + v_\parallel b_i + v_\perp n_{1_i}\right) - \left(\frac{\partial A_j}{\partial x_i} + v_\parallel \frac{\partial b_j}{\partial x_i} + v_\perp \frac{\partial n_{1_j}}{\partial x_i}\right)\dot{x}_j + \frac{\partial \Phi}{\partial x_i} \\
&= \dot{v}_\parallel b_i + \dot{v}_\perp n_{1_i} + v_\perp \frac{\partial n_{1_i}}{\partial \zeta}\dot{\zeta} + \left[\frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i} + v_\parallel\left(\frac{\partial b_i}{\partial x_j} - \frac{\partial b_j}{\partial x_i}\right) + v_\perp\left(\frac{\partial n_{1_i}}{\partial x_j} - \frac{\partial n_{1_j}}{\partial x_i}\right)\right]\dot{x}_j + \frac{\partial \Phi}{\partial x_i} \\
&= \dot{v}_\parallel b_i + \dot{v}_\perp n_{1_i} + v_\perp \dot{\zeta}\, n_{2_i} - \epsilon_{ijk}\dot{x}_j\left[B_k + v_\parallel\left(\nabla\times\boldsymbol{b}\right)_k + v_\perp\left(\nabla\times\boldsymbol{n}_1\right)_k\right] - E_i \\
&= \dot{v}_\parallel b_i + \dot{v}_\perp n_{1_i} + v_\perp \dot{\zeta}\, n_{2_i} - \epsilon_{ijk}\dot{x}_j\left(B_k + \delta_k\right) - E_i
\end{aligned}$$

corresponding now to the remaining variables $x_i$, yield altogether the vector relation

$$\dot{v}_\parallel \boldsymbol{b} + \dot{v}_\perp \boldsymbol{n}_1 + v_\perp \dot{\zeta}\,\boldsymbol{n}_2 = \dot{\boldsymbol{x}}\times\left(\boldsymbol{B}+\boldsymbol{\delta}\right) + \boldsymbol{E} \tag{1.49}$$

Then, dot multiplication of (1.49) with $\boldsymbol{b}$, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ and substitution of $\dot{\boldsymbol{x}}$ from (1.47a) give (1.47b)-(1.47d), respectively,

$$\begin{aligned}
\dot{v}_\parallel &= \left[\dot{\boldsymbol{x}}\times\left(\boldsymbol{B}+\boldsymbol{\delta}\right) + \boldsymbol{E}\right]\cdot\boldsymbol{b} = \boldsymbol{E}\cdot\boldsymbol{b} + \left[\boldsymbol{b}\times\left(v_\parallel\boldsymbol{b} + v_\perp\boldsymbol{n}_1\right)\right]\cdot\boldsymbol{\delta} = \boldsymbol{E}\cdot\boldsymbol{b} + v_\perp\boldsymbol{n}_2\cdot\boldsymbol{\delta} \\
\dot{v}_\perp &= \left[\dot{\boldsymbol{x}}\times\left(\boldsymbol{B}+\boldsymbol{\delta}\right) + \boldsymbol{E}\right]\cdot\boldsymbol{n}_1 = \boldsymbol{E}\cdot\boldsymbol{n}_1 + \left[\boldsymbol{n}_1\times\left(v_\parallel\boldsymbol{b} + v_\perp\boldsymbol{n}_1\right)\right]\cdot\left(\boldsymbol{B}+\boldsymbol{\delta}\right) = \boldsymbol{E}\cdot\boldsymbol{n}_1 - v_\parallel\boldsymbol{n}_2\cdot\boldsymbol{\delta} \\
v_\perp\dot{\zeta} &= \left[\dot{\boldsymbol{x}}\times\left(\boldsymbol{B}+\boldsymbol{\delta}\right) + \boldsymbol{E}\right]\cdot\boldsymbol{n}_2 = \boldsymbol{E}\cdot\boldsymbol{n}_2 + \left[\boldsymbol{n}_2\times\left(v_\parallel\boldsymbol{b} + v_\perp\boldsymbol{n}_1\right)\right]\cdot\left(\boldsymbol{B}+\boldsymbol{\delta}\right) \\
&= \boldsymbol{E}\cdot\boldsymbol{n}_2 + \left(v_\parallel\boldsymbol{n}_1 - v_\perp\boldsymbol{b}\right)\cdot\left(\boldsymbol{B}+\boldsymbol{\delta}\right) = \boldsymbol{E}\cdot\boldsymbol{n}_2 + \left(v_\parallel\boldsymbol{n}_1 - v_\perp\boldsymbol{b}\right)\cdot\boldsymbol{\delta} - v_\perp|\boldsymbol{B}|
\end{aligned}$$

recalling also that $\omega_g = |\boldsymbol{B}|$ in the last expression.

## 1.5 Motion of a particle in simple electromagnetic fields

In this section we consider the case, where the electromagnetic field, besides stationary, is also homogeneous, meaning a constant vector in $\mathbb{R}^3$. Both for convenience and more insight, we study

first the particle's motion in the absence of the electric field. With this preliminary, conclusions can then be easily drawn for the general problem where both fields are present. See also Example 4.49 revealing how the next two subsections are connected.

### 1.5.1 $\boldsymbol{E} = \mathbf{0}$, Homogeneous $\boldsymbol{B}$

Consider first a homogeneous magnetic field when there is no electric field, that is $\boldsymbol{B}(\boldsymbol{x}) = \boldsymbol{B}$ is a constant vector and $\boldsymbol{E} = \mathbf{0}$. As already mentioned in Remark 1.8, a constant magnetic field, besides $\boldsymbol{b}(\boldsymbol{x}) = \boldsymbol{b}$ being a constant vector, means that the rest two vectors of the Frenet triad cannot be defined. Therefore, $\boldsymbol{n}$ and $\boldsymbol{\beta}$ in the previous formalism can be arbitrary (orthogonal) constant vectors in the perpendicular plane. In this case, we can always choose from the beginning the Cartesian frame of reference to be exactly $(\boldsymbol{b}, \boldsymbol{n}, \boldsymbol{\beta})$. Let us retain at this point this notation for easy reference to the previous as well as the next section, keeping in mind that $\boldsymbol{n}$ and $\boldsymbol{\beta}$ are not defined in the sense of §1.2.3 and have nothing to do with the magnetic field whatsoever.

Since $\boldsymbol{n}$ and $\boldsymbol{\beta}$ are constant in space, then $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are independent of $\boldsymbol{x}$, and, since $\boldsymbol{b}$ is, too, therefore $\boldsymbol{\delta} = \mathbf{0}$. If $\boldsymbol{x} = \big(x_{\|}, x_{\boldsymbol{n}}, x_{\boldsymbol{\beta}}\big)$ denote the three positions along the axes $\boldsymbol{b}$, $\boldsymbol{n}$ and $\boldsymbol{\beta}$, then the equations of motion (1.46) for this system are

$$\begin{aligned} \frac{dx_{\|}}{dt} &= v_{\|} & \frac{dv_{\|}}{dt} &= 0 \\ \frac{dx_{\boldsymbol{n}}}{dt} &= v_{\perp}\cos\zeta & \frac{dv_{\perp}}{dt} &= 0 \\ \frac{dx_{\boldsymbol{\beta}}}{dt} &= v_{\perp}\sin\zeta & \frac{d\zeta}{dt} &= -\,\omega_g \end{aligned} \tag{1.50}$$

where $\omega_g = |\boldsymbol{B}|$ is a constant. Note that $\zeta$ is still not ignorable. Also observe, that despite of the Frenet triad defining a fixed reference frame, $T \neq I$, and although $\boldsymbol{\delta} = \mathbf{0}$, $D \neq 0$ neither. In other words, the above set of variables are still noncanonical, as we can see from the corresponding Poisson matrix

$$J = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos\zeta & -\dfrac{\sin\zeta}{v_{\perp}} \\ 0 & 0 & 0 & 0 & \sin\zeta & \dfrac{\cos\zeta}{v_{\perp}} \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\cos\zeta & -\sin\zeta & 0 & 0 & \dfrac{\omega_g}{v_{\perp}} \\ 0 & \dfrac{\sin\zeta}{v_{\perp}} & -\dfrac{\cos\zeta}{v_{\perp}} & 0 & -\dfrac{\omega_g}{v_{\perp}} & 0 \end{pmatrix}, \tag{1.51}$$

Nevertheless, the above equations can easily be solved, starting from the last three. The fourth and fifth one tell us that the speeds in both directions are constants, $v_\parallel(t) = v_\parallel$ and $v_\perp(t) = v_\perp$, while the sixth one yields $\zeta(t) = \zeta_0 - \omega_g t$, where subscripts $_0$ stand for initial conditions. Then, integrating the first three equations we get

$$\begin{aligned} x_\parallel(t) &= x_{\parallel,0} + v_\parallel t \\ x_{\boldsymbol{n}}(t) &= x_{\boldsymbol{n},0} + \rho \sin\zeta_0 - \rho \sin\left(\zeta_0 - \omega_g t\right) \\ x_{\boldsymbol{\beta}}(t) &= x_{\boldsymbol{\beta},0} - \rho \cos\zeta_0 + \rho \cos\left(\zeta_0 - \omega_g t\right) \end{aligned}$$

where we recall $\rho = v_\perp/\omega_g$. Therefore, if $\boldsymbol{x}_0 = x_{\parallel,0}\boldsymbol{b} + x_{\boldsymbol{n},0}\boldsymbol{n} + x_{\boldsymbol{\beta},0}\boldsymbol{\beta}$ the particle's position in space for any $t$ is $\boldsymbol{x}(t) = \boldsymbol{x}_0 + v_\parallel t\boldsymbol{b} - \left[\rho\sin\left(\zeta_0 - \omega_g t\right) - \rho\sin\zeta_0\right]\boldsymbol{n} + \left[\rho\cos\left(\zeta_0 - \omega_g t\right) - \rho\cos\zeta_0\right]\boldsymbol{\beta}$ or, recalling (1.42), $\boldsymbol{x}(t) = \boldsymbol{x}_0 + v_\parallel t\boldsymbol{b} + \rho\left[\boldsymbol{n}_2(\zeta_0 - \omega_g t) - \boldsymbol{n}_2(\zeta_0)\right]$. Put more simply the particle's motion,

$$\boldsymbol{x}(t) = \widetilde{\boldsymbol{x}}_0 + v_\parallel t\boldsymbol{b} + \rho\boldsymbol{n}_2(\zeta_0 - \omega_g t)\,, \tag{1.52}$$

is a linear combination of the fixed vector $\widetilde{\boldsymbol{x}}_0 = \boldsymbol{x}_0 - \rho\boldsymbol{n}_2(\zeta_0)$, a uniform motion with velocity $v_\parallel$ along the constant direction $\boldsymbol{b}$ of the magnetic field and a constant motion along the uniformly varying direction $\boldsymbol{n}_2$, which means a uniform rotation of radius $\rho$ with angular frequency $-\omega_g$ around the point $\widetilde{\boldsymbol{x}}_0$ in the perpendicular plane. Thus the particle's trajectory, combining the parallel uniform translation and the perpendicular uniform rotation, results in a helix of radius $\rho$ and slope $v_\parallel/\rho$ wound around the magnetic field line that passes through the point $\widetilde{\boldsymbol{x}}_0$, with angular frequency $-\omega_g$. Note that the latter is positive for negatively charged particles.

Although solved and completely determined, let us investigate a bit further system (1.50), preparing the ground for the next section. For the moment we neglect the rotating part of the motion and focus only on the translation along the magnetic field. In other words, we don't follow the entire helical trajectory $\boldsymbol{x}(t)$ but only the motion of the centre $\widetilde{\boldsymbol{x}}(t)$ of the helix. This can be expressed by considering the vector

$$\widetilde{\boldsymbol{x}} = \boldsymbol{x} - \rho\boldsymbol{n}_2(\zeta) \tag{1.53}$$

analogously to the fixed vector $\boldsymbol{x}_0 - \rho\boldsymbol{n}_2(\zeta_0)$ which simply gives

$$\widetilde{\boldsymbol{x}}(t) = \widetilde{\boldsymbol{x}}_0 + v_\parallel t\boldsymbol{b} \tag{1.54}$$

If we want to restore the Hamiltonian formulation in terms of the new variables $\left(\widetilde{\boldsymbol{x}}, v_\parallel, v_\perp, \zeta\right)$, where $\widetilde{\boldsymbol{x}} = \left(x_\parallel, \widetilde{x}_{\boldsymbol{n}}, \widetilde{x}_{\boldsymbol{\beta}}\right)$ with $\widetilde{x}_{\boldsymbol{n}} = x_{\boldsymbol{n}} + \rho\sin\zeta$ and $\widetilde{x}_{\boldsymbol{\beta}} = x_{\boldsymbol{\beta}} - \rho\cos\zeta$, we can easily calculate the

Poisson brackets

$$
\begin{aligned}
\left\{x_\parallel, \widetilde{x}_{\boldsymbol{n}}\right\} &= \left\{x_\parallel, x_{\boldsymbol{n}}\right\} + \frac{1}{\omega_g}\left\{x_\parallel, v_\perp \sin\zeta\right\} = \frac{1}{\omega_g}\left[\sin\zeta\left\{x_\parallel, v_\perp\right\} + v_\perp \cos\zeta\left\{x_\parallel, \zeta\right\}\right] = 0 \\
\left\{x_\parallel, \widetilde{x}_{\boldsymbol{\beta}}\right\} &= \left\{x_\parallel, x_{\boldsymbol{n}}\right\} - \frac{1}{\omega_g}\left\{x_\parallel, v_\perp \cos\zeta\right\} = -\frac{1}{\omega_g}\left[\cos\zeta\left\{x_\parallel, v_\perp\right\} - v_\perp \sin\zeta\left\{x_\parallel, \zeta\right\}\right] = 0 \\
\{\widetilde{x}_{\boldsymbol{n}}, \widetilde{x}_{\boldsymbol{\beta}}\} &= \{x_{\boldsymbol{n}}, x_{\boldsymbol{\beta}}\} - \frac{1}{\omega_g}\{x_{\boldsymbol{n}}, v_\perp \cos\zeta\} + \frac{1}{\omega_g}\{v_\perp \sin\zeta, x_{\boldsymbol{\beta}}\} - \frac{1}{\omega_g^2}\{v_\perp \sin\zeta, v_\perp \cos\zeta\} \\
&= -\frac{1}{\omega_g}\cos\zeta\{x_{\boldsymbol{n}}, v_\perp\} + \frac{v_\perp}{\omega_g}\sin\zeta\{x_{\boldsymbol{n}}, \zeta\} + \frac{1}{\omega_g}\sin\zeta\{v_\perp, x_{\boldsymbol{\beta}}\} + \frac{v_\perp}{\omega_g}\cos\zeta\{\zeta, x_{\boldsymbol{\beta}}\} - \\
&\quad - \frac{1}{\omega_g^2}\sin\zeta\{v_\perp, v_\perp\cos\zeta\} - \frac{v_\perp}{\omega_g^2}\cos\zeta\{\zeta, v_\perp\cos\zeta\} \\
&= -\frac{1}{\omega_g} - \frac{1}{\omega_g} + \frac{v_\perp}{\omega_g^2}\sin^2\zeta\{v_\perp, \zeta\} - \frac{v_\perp}{\omega_g^2}\cos^2\zeta\{\zeta, v_\perp\} = -\frac{1}{\omega_g} \\
\left\{\widetilde{x}_{\boldsymbol{n}}, v_\parallel\right\} &= \left\{x_{\boldsymbol{n}}, v_\parallel\right\} + \frac{1}{\omega_g}\left\{v_\perp \sin\zeta, v_\parallel\right\} = \frac{1}{\omega_g}\left[\sin\zeta\left\{v_\perp, v_\parallel\right\} + v_\perp\cos\zeta\left\{\zeta, v_\parallel\right\}\right] = 0 \\
\left\{\widetilde{x}_{\boldsymbol{\beta}}, v_\parallel\right\} &= \left\{x_{\boldsymbol{n}}, v_\parallel\right\} - \frac{1}{\omega_g}\left\{v_\perp \cos\zeta, v_\parallel\right\} = -\frac{1}{\omega_g}\left[\cos\zeta\left\{v_\perp, v_\parallel\right\} - v_\perp\sin\zeta\left\{\zeta, v_\parallel\right\}\right] = 0 \\
\{\widetilde{x}_{\boldsymbol{n}}, v_\perp\} &= \{x_{\boldsymbol{n}}, v_\perp\} + \frac{1}{\omega_g}\{v_\perp \sin\zeta, v_\perp\} = \cos\zeta + \frac{v_\perp}{\omega_g}\cos\zeta\{\zeta, v_\perp\} = 0 \\
\{\widetilde{x}_{\boldsymbol{\beta}}, v_\perp\} &= \{x_{\boldsymbol{\beta}}, v_\perp\} - \frac{1}{\omega_g}\{v_\perp \cos\zeta, v_\perp\} = \sin\zeta + \frac{v_\perp}{\omega_g}\sin\zeta\{\zeta, v_\perp\} = 0 \\
\{\widetilde{x}_{\boldsymbol{n}}, \zeta\} &= \{x_{\boldsymbol{n}}, \zeta\} + \frac{1}{\omega_g}\{v_\perp \sin\zeta, \zeta\} = -\frac{\sin\zeta}{v_\perp} + \frac{1}{\omega_g}\sin\zeta\{v_\perp, \zeta\} = 0 \\
\{\widetilde{x}_{\boldsymbol{\beta}}, \zeta\} &= \{x_{\boldsymbol{\beta}}, \zeta\} - \frac{1}{\omega_g}\{v_\perp \cos\zeta, \zeta\} = \frac{\cos\zeta}{v_\perp} - \frac{1}{\omega_g}\cos\zeta\{v_\perp, \zeta\} = 0
\end{aligned}
$$

involving of course only the new variables $\widetilde{x}_{\boldsymbol{n}}$ and $\widetilde{x}_{\boldsymbol{\beta}}$, since the rest are already known. Therefore, in terms of the new coordinates the Poisson matrix is

$$
J = \begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & -\dfrac{1}{\omega_g} & 0 & 0 & 0 \\
0 & \dfrac{1}{\omega_g} & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \dfrac{\omega_g}{v_\perp} \\
0 & 0 & 0 & 0 & -\dfrac{\omega_g}{v_\perp} & 0
\end{pmatrix}, \tag{1.55}
$$

The Hamiltonian function of course remains the same as before, meaning the one for system

(1.50) given by equation (1.43) for $\Phi = 0$. Hence, system (1.50) is now expressed as

$$\begin{aligned} \frac{dx_\parallel}{dt} &= v_\parallel & \frac{dv_\parallel}{dt} &= 0 \\ \frac{d\widetilde{x}_{\boldsymbol{n}}}{dt} &= 0 & \frac{dv_\perp}{dt} &= 0 \\ \frac{d\widetilde{x}_{\boldsymbol{\beta}}}{dt} &= 0 & \frac{d\zeta}{dt} &= -\,\omega_g \end{aligned} \tag{1.56}$$

Note that the Poisson bracket for this system is very close to a canonical representation

$$\{F,G\} = \frac{\partial F}{\partial x_\parallel}\frac{\partial G}{\partial v_\parallel} - \frac{\partial F}{\partial v_\parallel}\frac{\partial G}{\partial x_\parallel} + \frac{1}{\omega_g}\left(\frac{\partial F}{\partial \widetilde{x}_{\boldsymbol{\beta}}}\frac{\partial G}{\partial \widetilde{x}_{\boldsymbol{n}}} - \frac{\partial F}{\partial \widetilde{x}_{\boldsymbol{n}}}\frac{\partial G}{\partial \widetilde{x}_{\boldsymbol{\beta}}}\right) + \frac{\omega_g}{v_\perp}\left(\frac{\partial F}{\partial v_\perp}\frac{\partial G}{\partial \zeta} - \frac{\partial F}{\partial \zeta}\frac{\partial G}{\partial v_\perp}\right) \tag{1.57}$$

The main advantage of the transformation (1.53) is the elimination of the variable $\zeta$ from the resulting equations (1.56) of the system. Even though this system has been already solved and integrals of motion are already apparent, for future purposes let us find the conserved physical quantity behind the ingnorable gyrophase $\zeta$ in this simple case. This discussion could be further elaborated in terms of symmetries and Noether's theorem presented later on. But for the time being it suffices to employ the typical characteristic of Hamiltonian systems that the conjugate variable $\mu$ to an ignonarable variable $\zeta$ is a first integral of motion. From (1.57), we see that the condition $\{\mu,\zeta\} = 1$ simply reduces to $\partial\mu/\partial v_\perp = v_\perp/\omega_g$ that yields $\mu = v_\perp^2/(2\omega_g)$ or restoring physical units

$$\mu = \frac{m v_\perp^2}{2|\boldsymbol{B}|}\,. \tag{1.58}$$

The quantity $\mu$ is known as the *magnetic moment*, and plays a crucial role in guiding-center theory. Here (1.58) is easily verified as a first integral, since both $v_\perp$ and $|\boldsymbol{B}|$ are constant.

In summary, we have transformed the original system to a much simpler one that *a*) admits $\zeta$ as an ignorable variable, i.e. a symmetry in the $\zeta$-direction $\boldsymbol{n}_2$. Moreover, in this new system *b*) the magnetic moment appeared as the conjugate variable to $\zeta$, yielding therefore an integral of motion. Of course, for system (1.50) this integral is functionally dependent on the already known first integral $v_\perp$. However, even with just the addition of a homogeneous electric field in the picture, presented in the next subsection, neither of the two velocities are conserved and no first integral is that obvious, besides the Hamiltonian. Therefore, it would be desirable to find (if they exist) such transformations in more general cases in order to discover (if they exist) integrals of motion. The two aspects *a*) and *b*) give us an illustrative example of forthcoming developments in this simple case and are also the key ingredients of the guiding-center transformation for more general electromagnetic fields.

### 1.5.2 Homogeneous $\boldsymbol{B}$, $\boldsymbol{E}$

We now assume that both the magnetic and the electric field are present and homogeneous, given by constant vectors $\boldsymbol{B}(\boldsymbol{x}) = \boldsymbol{B}$ and $\boldsymbol{E}(\boldsymbol{x}) = \boldsymbol{E}$ in $\mathbb{R}^3$. The fixed reference frame can still be given by the Frenet triad, but let us instead use a moving reference frame, which brings in perhaps the most easy treatment of this case. So, we consider a Galilean transformation

$$\bar{\boldsymbol{x}} = \boldsymbol{x} - \mathbf{v}t, \tag{1.59}$$

where $\mathbf{v}$ is the constant velocity of the moving frame. The reason why (1.59) is suitable becomes evident when we express the motion of a charged particle in the new variables $\ddot{\bar{\boldsymbol{x}}} = \ddot{\boldsymbol{x}} = \dot{\boldsymbol{x}} \times \boldsymbol{B} + \boldsymbol{E} = \dot{\bar{\boldsymbol{x}}} \times \boldsymbol{B} + \mathbf{v} \times \boldsymbol{B} + \boldsymbol{E}$. For if we choose

$$\mathbf{v} = \frac{1}{|\boldsymbol{B}|^2}\, \boldsymbol{E} \times \boldsymbol{B} \tag{1.60}$$

the last two terms at the right hand side of the equations cancel out, leaving only the parallel component of the electric field, $\mathbf{v} \times \boldsymbol{B} + \boldsymbol{E} = [(\boldsymbol{E} \cdot \boldsymbol{B})\boldsymbol{B} - (\boldsymbol{B} \cdot \boldsymbol{B})\boldsymbol{E}]/|\boldsymbol{B}|^2 + \boldsymbol{E} = \boldsymbol{E}_{\parallel}$.

Therefore, in the frame of reference that moves with the constant velocity (1.60) the system takes the form

$$\ddot{\bar{\boldsymbol{x}}} = \dot{\bar{\boldsymbol{x}}} \times \boldsymbol{B} + \boldsymbol{E}_{\parallel} \tag{1.61}$$

and can be decomposed in the parallel direction and the perpendicular plane with respect to the magnetic field,

$$\begin{aligned} \ddot{\bar{x}}_{\parallel} &= E_{\parallel} \\ \ddot{\bar{\boldsymbol{x}}}_{\perp} &= \dot{\bar{\boldsymbol{x}}}_{\perp} \times \boldsymbol{B} \end{aligned} \tag{1.62}$$

where $E_{\parallel}$ is a constant. The former can be easily integrated yielding the parabola (1.63), while the latter is the system studied in the previous subsection and whose solution was decribed in (1.52),

$$\bar{x}_{\parallel}(t) = \bar{x}_{\parallel,0} + \bar{v}_{\parallel,0}t + \frac{1}{2}E_{\parallel}t^2 \tag{1.63}$$

$$\bar{\boldsymbol{x}}_{\perp}(t) = \bar{\boldsymbol{x}}_{\perp,0} + \bar{\rho}\left[\bar{\boldsymbol{n}}_2\big(\bar{\zeta}_0 - \omega_g t\big) - \bar{\boldsymbol{n}}_2\big(\bar{\zeta}_0\big)\right] \tag{1.64}$$

whereas $\bar{v}_{\parallel}(t) = \bar{v}_{\parallel,0} + E_{\parallel}t$, $\bar{v}_{\perp}(t) = \bar{v}_{\perp}$ and $\bar{\zeta}(t) = \bar{\zeta}_0 - \omega_g t$. So, in the moving reference frame the particle undergoes the combined motion,

$$\bar{\boldsymbol{x}}(t) = \bar{\boldsymbol{x}}_0 + \left(\bar{v}_{\parallel,0}t + \frac{1}{2}E_{\parallel}t^2\right)\boldsymbol{b} + \bar{\rho}\left[\bar{\boldsymbol{n}}_2\big(\bar{\zeta}_0 - \omega_g t\big) - \bar{\boldsymbol{n}}_2\big(\bar{\zeta}_0\big)\right], \tag{1.65}$$

of a uniform acceleration $E_\parallel$ in the direction of the magnetic field and a uniform rotation of radius $\bar\rho = \bar v_\perp/|\boldsymbol{B}|$ around the point $\bar{\boldsymbol{x}}_0 - \bar\rho\bar{\boldsymbol{n}}_2(\bar\zeta_0)$ with angular frequency $-\omega_g$. An observer in this frame of reference will still see the particle's trajectory wrapped around the magnetic field line passing through $\bar{\boldsymbol{x}}_0 - \bar\rho\bar{\boldsymbol{n}}_2(\bar\zeta_0)$, only this time uniformly accelerated along that field line.

Although simple enough in the moving coordinate system, the solution needs to be expressed back in the original fixed coordinates. First of all, from (1.60), we see that $\mathrm{v}_\parallel = \mathbf{v}\cdot\boldsymbol{b} = 0$, i.e. the transformed reference frame moves perpendicularly to the magnetic field. Therefore, nothing changes in the parallel direction, meaning $\bar x_\parallel = x_\parallel$ and $\bar v_\parallel = v_\parallel$. On the other hand, all the transformation takes place in the perpendicular plane, that is, (1.59) with the choice (1.60) simply reduces to

$$\bar{\boldsymbol{x}}_\perp = \boldsymbol{x}_\perp - \mathbf{v}t \tag{1.66}$$

that also results in $\bar{\boldsymbol{v}}_\perp = \boldsymbol{v}_\perp - \mathbf{v}$. Thus, an observer in the moving frame of reference sees the particle's perpendicular velocity $\bar{\boldsymbol{v}}_\perp = \bar v_\perp\bar{\boldsymbol{n}}_1$ in a different direction than an observer in the fixed reference frame, for whom $\boldsymbol{v}_\perp = v_\perp\boldsymbol{n}_1$. Consequently, if $\zeta$ is as usual the angle between $\boldsymbol{n}_1$ and $\boldsymbol{n}$ measured by the second observer, then the first observer measures a different angle, say $\bar\zeta = \zeta - \chi$, between $\bar{\boldsymbol{n}}_1$ and (the fixed axis for both systems) $\boldsymbol{n}$. The relation between the two directions is given by

$$\begin{aligned}\bar{\boldsymbol{n}}_1(\bar\zeta) &= \boldsymbol{n}_1(\zeta-\chi)\,,\\ \bar{\boldsymbol{n}}_2(\bar\zeta) &= \boldsymbol{n}_2(\zeta-\chi)\end{aligned} \tag{1.67}$$

and then the second relation follows easily from $\bar{\boldsymbol{n}}_2 = \boldsymbol{b}\times\bar{\boldsymbol{n}}_1$. A little trigonometry shows that the difference between the two angles is $\chi = \arctan\left(-\bar{\boldsymbol{v}}_\perp\cdot\boldsymbol{n}_2/(\bar{\boldsymbol{v}}_\perp\cdot\boldsymbol{n}_1)\right)$ or

$$\chi = \arctan\left(\frac{\mathrm{v}_2}{v_\perp - \mathrm{v}_1}\right), \tag{1.68}$$

where $\mathrm{v}_1 = \mathbf{v}\cdot\boldsymbol{n}_1 = \boldsymbol{E}\cdot\boldsymbol{n}_2/|\boldsymbol{B}|$ and $\mathrm{v}_2 = \mathbf{v}\cdot\boldsymbol{n}_2 = -\boldsymbol{E}\cdot\boldsymbol{n}_1/|\boldsymbol{B}|$. Now, we have all we need to express the solution (1.65) in the original coordinate system, and so inverting (1.66) we get

$$\boldsymbol{x}(t) = \widetilde{\boldsymbol{x}}_0 + \left(v_{\parallel,0}t + \frac{1}{2}E_\parallel t^2\right)\boldsymbol{b} + \frac{|\boldsymbol{v}_{\perp,0}-\mathbf{v}|}{|\boldsymbol{B}|}\,\boldsymbol{n}_2(\zeta_0-\chi_0-\omega_g t) + \mathbf{v}t \tag{1.69}$$

for $\mathbf{v}$ (1.60), where $\widetilde{\boldsymbol{x}}_0 = \bar{\boldsymbol{x}}_0 - \bar\rho\bar{\boldsymbol{n}}_2(\bar\zeta_0) = \boldsymbol{x}_0 - |\boldsymbol{v}_{\perp,0}-\mathbf{v}|\,\boldsymbol{n}_2(\zeta_0-\chi_0)/|\boldsymbol{B}|$. More explicitly,

$$\begin{aligned} x_\parallel(t) &= x_{\parallel,0} + v_{\parallel,0}t + \frac{1}{2}E_\parallel t^2\\ x_{\boldsymbol{n}}(t) &= \widetilde{x}_{\boldsymbol{n},0} - \frac{|\boldsymbol{v}_{\perp,0}-\mathbf{v}|}{|\boldsymbol{B}|}\sin(\zeta_0-\chi_0-\omega_g t) + \mathrm{v}_{\boldsymbol{n}}t\\ x_{\boldsymbol{\beta}}(t) &= \widetilde{x}_{\boldsymbol{\beta},0} + \frac{|\boldsymbol{v}_{\perp,0}-\mathbf{v}|}{|\boldsymbol{B}|}\cos(\zeta_0-\chi_0-\omega_g t) + \mathrm{v}_{\boldsymbol{\beta}}t\end{aligned} \tag{1.70}$$

From the above expressions, we deduce that the particle's motion in the original reference frame is the superposition of a uniform acceleration $E_\parallel$ along the magnetic field, a uniform rotation of radius $|\boldsymbol{v}_{\perp,0}-\mathbf{v}|/|\boldsymbol{B}|$ around $\widetilde{\boldsymbol{x}}_0$ with angular frequency $-\omega_g$ in the perpendicular plane and a uniform motion with velocity $\mathbf{v}$ (1.60) perpendicular to both the magnetic and the electric field. Following the example of the previous subsection, if we extract the rotating part from the particle's motion by defining the vector $\widetilde{\boldsymbol{x}}=\bar{\boldsymbol{x}}-\bar{\rho}\bar{\boldsymbol{n}}_2(\bar{\zeta})=\boldsymbol{x}-|\boldsymbol{v}_{\perp,0}-\mathbf{v}|\,\boldsymbol{n}_2(\zeta-\chi-\omega_g t)/|\boldsymbol{B}|$, then the rest of the motion is described by

$$\widetilde{\boldsymbol{x}}(t)=\widetilde{\boldsymbol{x}}_0+\left(v_{\parallel,0}t+\frac{1}{2}E_\parallel t^2\right)\boldsymbol{b}+\frac{\boldsymbol{E}\times\boldsymbol{B}}{|\boldsymbol{B}|^2}\,t \tag{1.71}$$

Note that for $\boldsymbol{E}=\mathbf{0}$ we recover the solution (1.54). The differences now between the case studied here of a homogeneous electromagnetic field and the previous one of just a homogeneous magnetic field are better revealed, if we compare (1.54) and (1.71). From their comparison we identify the following two new features appearing due to the presence of the electric field that correspond to the last two terms of the above expression: *i*) the particle is uniformly accelerated along the magnetic field instead of moving with constant velocity, and more importantly *ii*) the center of the perpendicular rotation is no longer fixed but drifts at constant speed that is perpendicular to both fields.

The second characteristic breaks the helical symmetry of the particle's motion and results in a trajectory in the perpendicular plane known as a *trochoid* in general. The latter is expressed parametrically by the last two equations of (1.70), but instead of the directions $\boldsymbol{n}$ and $\boldsymbol{\beta}$, we may better use the directions of $\boldsymbol{w}$ and $\boldsymbol{E}_\perp$,

$$\begin{aligned} x_{\boldsymbol{w}}(t)&=\widetilde{x}_{\boldsymbol{w},0}-\frac{|\boldsymbol{v}_{\perp,0}-\mathbf{v}|}{|\boldsymbol{B}|}\sin\left(\zeta_0-\chi_0-\alpha-\omega_g t\right)+\mathrm{v}t\\ x_{\boldsymbol{E}_\perp}(t)&=\widetilde{x}_{\boldsymbol{E}_\perp,0}+\frac{|\boldsymbol{v}_{\perp,0}-\mathbf{v}|}{|\boldsymbol{B}|}\cos\left(\zeta_0-\chi_0-\alpha-\omega_g t\right)\end{aligned} \tag{1.72}$$

where $\alpha$ is the fixed angle between $\mathbf{v}$ and $\boldsymbol{n}$ and $\mathrm{v}=|\mathbf{v}|=|\boldsymbol{E}_\perp|/|\boldsymbol{B}|$. Now the interpretation of the trochoid becomes more clear; it describes the particle as a fixed point at a constant distance $|\boldsymbol{v}_{\perp,0}-\mathbf{v}|/|\boldsymbol{B}|$ from the center of a rolling circle of radius $\mathrm{v}/\omega_g$ that rolls (without sliding) along the direction of $\mathbf{v}$ in the perpendicular plane. For clarity and not to get lost in too many initial conditions, let us see a particular case.

**Example 1.12.** Take for example $\zeta_0-\chi_0=\alpha$, i.e. at $t=0$ the vector $\bar{\boldsymbol{n}}_1$ is parallel to $\mathbf{v}$, and assume that the particle starts with zero perpendicular initial velocity, $\boldsymbol{v}_{\perp,0}=\mathbf{0}$. Then (1.72)

reduce to

$$\begin{aligned} x_{\mathbf{v}}(t) &= x_{\mathbf{v},0} + \frac{\mathrm{v}}{\omega_g}\left(\omega_g t + \sin\left(\omega_g t\right)\right) \\ x_{\boldsymbol{E}_\perp}(t) &= x_{\boldsymbol{E}_\perp,0} + \frac{\mathrm{v}}{\omega_g}\left(1 + \cos\left(\omega_g t\right)\right) \end{aligned} \tag{1.73}$$

which describe a particular case of a trochoid, called *cycloid*. This is the case where the fixed point with respect to the rolling circle lies on the circle.

Other types of trochoids are encountered, depending on the initial velocity $\boldsymbol{v}_{\perp,0}$. For example, if $v_{\perp,0} < \mathrm{v}_{1,0}$ then the said fixed point lies inside the rolling circle and we have a contracted trochoid, while when $v_{\perp,0} > \mathrm{v}_{1,0}$ the fixed point lies outside the rolling circle and we get an extended trochoid.

The above behaviour of a moving rotation center that is entirely due to the presence of the electric field represented by the last term in (1.71) is commonly known as the $\boldsymbol{E}\times\boldsymbol{B}$ *drift* or the *electric drift*. Note that this drift motion is independent of the particle's charge, i.e. positive and negative charges drift in the same direction. Thus the electric drift cannot give rise to a current. We should also have in mind that in order to remain in the validity of the Newtonian description considered so far, we must demand $\mathrm{v} << c$ and therefore $|\boldsymbol{E}| << c|\boldsymbol{B}|$.

Closing this section, we conclude that, as already promised from the Introduction, the trajectories of charged particles even for simple electromagnetic fields are not simple at all. From the analysis presented here we understand that the main reason is that the magnetic field forces the charged particle in a rotating motion around the magnetic field lines, widely called *gyration*. When this rotation was absorbed in the coordinates $\widetilde{\boldsymbol{x}}$, the rest of the motion in either case was guided by the center of gyration. To avoid any confusion we comment that the Galilean transformation (1.59) adopted in the case of an additional homogeneous electric field $\boldsymbol{E} \neq \boldsymbol{0}$ has nothing to do whatsoever with the elimination of the gyrophase $\zeta$ giving rise to the magnetic moment $\mu$ as a first integral investigated in the $\boldsymbol{E} = \boldsymbol{0}$ case. Hoping to serve as an illustrative example in §1.5.1, this step with all its Hamiltonian aspects was intentionally not repeated in §1.5.2, left for more general electromagnetic fields in the next section, avoiding lengthy case by case calculations. For even just an additional homogeneous electric field, transformation (1.53) would not be enough to remove $\zeta$, and $\left(v_{\parallel}, v_{\perp}, \zeta\right)$ would need to be transformed too. This treatment and the limits of its validity for the general problem are the subject of the next section.

## 1.6 Guiding center motion

In retrospect, for the case of just a homogeneous magnetic field if we had not solved the original system (1.50), its solution could always be recovered even more easily from the solution of (1.56) and the inverse transformation (1.53). In fact, taking one step further, to a first approximation we could have completely disregarded the gyration of the particle, since most charged particles, such as electrons, protons, etc., have small masses, especially compared to their charge, and consequently very high gyrofrequencies and very small gyroradii. Therefore the particle's rotation around the magnetic field line is only a very small deviation from the motion along the magnetic field. What is more important is that in any case, besides this small rotation, the rest of the trajectory $\boldsymbol{x}(t)$ that governs the evolution of the particle in space is one-to-one followed by the trajectory $\widetilde{\boldsymbol{x}}(t)$ of an imaginary particle, known as the guiding center.

The above considerations and the concept of this imaginary particle can be approximately generalised for inhomogeneous electromagnetic fields, leading to a treatment that is widely used in charged particle motion. The basic idea behind the guiding-center approximation is that the particle's motion consists of a very fast rotation around the magnetic field and a motion almost along the magnetic field subject to certain drifts. The center of rotation that undergoes the second part of the motion is called the *guiding center*. As the reader might have guessed, the center of the helixes in 1.5.1 or the trochoids in 1.5.2 is the guiding center for these cases. The guiding-center picture is realistic, when the electromagnetic field is weakly inhomogeneous, and more precisely when the gyroradius $\rho$ (1.44) is much smaller than certain characteristic lengths of the magnetic field,

$$\rho \ll \lambda, \varkappa, \sigma \tag{1.74}$$

such as the scale length $\lambda$ and the curvature and torsion radii $\kappa$ and $\sigma$, respectively, of the magnetic field lines (see §1.2.3). Below we recall collectively the above quantities from previous sections, including the definition of the scale length,

$$\rho = \frac{v_\perp}{\omega_g}, \qquad \lambda = \frac{|\boldsymbol{B}|}{|\nabla|\boldsymbol{B}||}, \qquad \varkappa = \frac{1}{|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{b}|}, \qquad \sigma = -\frac{1}{\boldsymbol{n}\cdot[(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{\beta}]}$$

In this section we briefly go through the steps of the construction for the guiding-center averaging transformation, arriving at the guiding-center equations of motion in Hamiltonian form. We give a qualitative description, as full calculation details are quite lengthy and beyond the scope of this chapter. The path followed is based however on the Poisson bracket methods manifested in the previous sections, and similar in concept with former implementations. Although the results presented are originally due to Littlejohn [77], the construction below was

conceptually simplified by Weyssow and Balescu [114]. This route adopted here using the Poisson brackets has been lately very popular in statistical approaches, such as the Fokker-Planck equation, leading respectively to the so-called *Gyrokinetics*.

We start with the equations of motion (1.47) in terms of the parallel and perpendicular direction with respect to the magnetic field. As previously discussed, the gyrofrequency is quite high for most charged particles, when (1.74) is valid. In this case, $\omega_g$ is the dominant term in the equation (1.47d) that describes the evolution of the gyrophase. This very fast gyration can be realized by simply introducing a scaling parameter $\epsilon << 1$ and replacing the gyrofrequency with $\omega_g/\epsilon$. Then, focusing on the dependence of the terms involved and the special role of $\zeta$ among the rest of the variables $z = \left(\boldsymbol{x}, v_\parallel, v_\perp\right)$, equations (1.47) taken together are expressed as

$$\begin{aligned} \frac{dz}{dt} &= f(z,\zeta) \\ \frac{d\zeta}{dt} &= -\frac{1}{\epsilon}\,\omega_g(z) + f_6(z,\zeta) \end{aligned} \tag{1.75}$$

where $f = (f_1, f_2, f_3, f_4, f_5)$ is the right hand side of (1.47a)-(1.47c), and $f_6$ the rest of (1.47d).

The procedure continues motivated by earlier works of Kruskal [67] and Littlejohn [76], stating that there always exists a transformation

$$\widetilde{w} = w + \epsilon g(w) + O\left(\epsilon^2\right), \tag{1.76}$$

where $w = (z, \zeta)$, under which the system in the new variables $\widetilde{w} = (\widetilde{z}, \widetilde{\zeta})$, where $\widetilde{z} = \left(\widetilde{\boldsymbol{x}}, \widetilde{v}_\parallel, \widetilde{v}_\perp\right)$, will be independent of the transformed gyrophase $\widetilde{\zeta}$ at least up to first-order terms,

$$\begin{aligned} \frac{d\widetilde{z}}{dt} &= \widetilde{f}_0(\widetilde{z}) + \epsilon\widetilde{f}_1(\widetilde{z}) + O\left(\epsilon^2\right) \\ \frac{d\widetilde{\zeta}}{dt} &= -\frac{1}{\epsilon}\,\omega_g(\widetilde{z}) + \widetilde{f}_{60}\left(\widetilde{z}\right) + \epsilon\widetilde{f}_{61}\left(\widetilde{z}\right) + O\left(\epsilon^2\right) \end{aligned} \tag{1.77}$$

arguing that the elimination of $\widetilde{\zeta}$ could be carried over to higher order terms, as well[1]. Actually Littlejohn's line of work stays very close to the proof of Darboux' theorem for symplectic manifolds (see appendix B). Both he and Kruskal showed that (1.76) is not unique.

In their construction, Weyssow and Balescu exploiting the non-uniqueness of transformation (1.76) required the system (1.77), describing the motion of the guiding center, to be Hamiltonian, too. For this to be the case, they suggested the following two conditions at least up to first-order terms: *i*) the Poisson matrix in the new coordinates to be independent from the new gyrophase

1. The claim that this elimination can be extended to terms of all orders is only formally guaranteed; otherwise some sort of symmetry would be admitted by the equations, as we will see in chapter 4.

$\widetilde{\zeta}$, meaning $J(\widetilde{w}) = J_0(\widetilde{z}) + \epsilon J_1(\widetilde{z}) + O(\epsilon^2)$ and so $\{\widetilde{w}_i, \widetilde{w}_j\} = J_{ij,0}(\widetilde{z}) + \epsilon J_{ij,1}(\widetilde{z}) + O(\epsilon^2)$ or

$$\frac{\partial \{\widetilde{w}_i, \widetilde{w}_j\}}{\partial \zeta} = O(\epsilon^2)\,, \tag{1.78}$$

and still having enough freedom in the choice of the new variables *ii*) the form of the new Hamiltonian function will remain invariant, that is

$$H(\widetilde{\boldsymbol{x}}, \widetilde{v}_\parallel, \widetilde{v}_\perp) = \frac{1}{2}\,(\widetilde{v}_\parallel^2 + \widetilde{v}_\perp^2) + \Phi(\widetilde{\boldsymbol{x}}) + O(\epsilon^2)\,. \tag{1.79}$$

Observe that, given the non-uniqueness of $\widetilde{w}$, the partial differential equations (1.78) do not have to be solved in full generality, but special solutions suffice. Therefore, one can start off with arbitrary functions $g = (g_1, g_2, g_3, g_4, g_5, g_6)$ in (1.76), and implementing (1.78) and (1.79) solve for $g_i$ keeping solutions as simple as possible. Once $g$ and therefore $\widetilde{w}$ is found, then the construction of the Poisson matrix in terms of $\widetilde{w}$ can follow using the repeated technique in this chapter summarised in Remark C.5. And the resulting guiding-center equations (1.77) can be built as usual by formula (C.2). Thus, the guiding-center system naturally inherits the Poisson structure from the original one with a physically meaningful Hamiltonian function.

In this way the construction of the transformation (1.76) has been carried out through second-order terms. Below we see the guiding-center averaging transformation up to first-order terms.

$$\begin{aligned}
\widetilde{\boldsymbol{x}} &= \boldsymbol{x} - \epsilon \frac{v_\perp}{\omega_g}\,\boldsymbol{n}_2 \\
\widetilde{v}_\parallel &= v_\parallel + \epsilon v_\perp \frac{v_\perp}{\omega_g}\left\{\frac{1}{4}\,[2\boldsymbol{b}\cdot(\nabla\times\boldsymbol{b}) + \boldsymbol{n}_1\cdot(\nabla\times\boldsymbol{n}_1) - \boldsymbol{n}_2\cdot(\nabla\times\boldsymbol{n}_2)] + \frac{v_\parallel}{v_\perp}\,\boldsymbol{n}_1\cdot(\nabla\times\boldsymbol{b})\right\} \\
\widetilde{v}_\perp &= v_\perp - \epsilon v_\parallel \frac{v_\perp}{\omega_g}\left\{\frac{1}{4}\,[2\boldsymbol{b}\cdot(\nabla\times\boldsymbol{b}) + \boldsymbol{n}_1\cdot(\nabla\times\boldsymbol{n}_1) - \boldsymbol{n}_2\cdot(\nabla\times\boldsymbol{n}_2)] + \frac{v_\parallel}{v_\perp}\,\boldsymbol{n}_1\cdot(\nabla\times\boldsymbol{b}) + \frac{v_\parallel}{v_\perp}\,\boldsymbol{n}_2\cdot\boldsymbol{E}\right\} \\
\widetilde{\zeta} &= \zeta - \frac{\epsilon}{\omega_g}\left\{\frac{v_\parallel}{4}\,[\boldsymbol{n}_2\cdot(\nabla\times\boldsymbol{n}_1) + \boldsymbol{n}_1\cdot(\nabla\times\boldsymbol{n}_2)] + \frac{v_\parallel^2}{v_\perp}\,\boldsymbol{n}_2\cdot(\nabla\times\boldsymbol{b}) - v_\perp \boldsymbol{b}\cdot(\nabla\times\boldsymbol{n}_2) - \frac{1}{v_\perp}\,\boldsymbol{n}_1\cdot\boldsymbol{E} + \right. \\
&\qquad\qquad \left. + v_\perp \boldsymbol{n}_1\cdot\frac{\nabla|\boldsymbol{B}|}{|\boldsymbol{B}|}\right\}
\end{aligned}$$

Note that for a homogeneous magnetic field and zero electric field we recover transformation (1.53). As Balescu points out ([4], p. 60) in order to determine the Poisson brackets through order $\epsilon$ the new variables have to be calculated through order $\epsilon^2$, stressing the fact that the $\epsilon$-ordering of the Poisson matrix and the new variables is not the same. For details of the Poisson structure see [4]. Here we directly present the resulting equations under the above transformation which can be neatly expressed as

$$\frac{d\widetilde{\boldsymbol{x}}}{dt} = \widetilde{v}_\parallel \boldsymbol{b} + \frac{\epsilon}{\omega_g}\,\boldsymbol{b}\times\boldsymbol{N} \tag{1.80a}$$

$$\frac{d\widetilde{v}_\parallel}{dt} = -\widetilde{v}_\perp^2 \boldsymbol{b}\cdot\boldsymbol{G} + \boldsymbol{b}\cdot\boldsymbol{E} + \frac{\epsilon}{\omega_g}\widetilde{v}_\parallel \boldsymbol{K}\cdot(\boldsymbol{b}\times\boldsymbol{N}) \tag{1.80b}$$

$$\frac{d\widetilde{v}_\perp}{dt} = \widetilde{v}_\parallel\widetilde{v}_\perp \boldsymbol{b}\cdot\boldsymbol{G} \quad + \quad \frac{\epsilon}{\omega_g}\widetilde{v}_\perp \boldsymbol{G}\cdot(\boldsymbol{b}\times\boldsymbol{N}) \tag{1.80c}$$

$$\frac{d\widetilde{\zeta}}{dt} = -\frac{\omega_g}{\epsilon} - \widetilde{v}_\parallel \boldsymbol{b}\cdot(\nabla\boldsymbol{n}\cdot\boldsymbol{\beta}) + \frac{\widetilde{v}_\parallel}{2}\boldsymbol{b}\cdot(\nabla\times\boldsymbol{b}) \tag{1.80d}$$

where $\boldsymbol{N} = \widetilde{v}_\parallel^2\boldsymbol{K} + \widetilde{v}_\perp^2\boldsymbol{G} - \boldsymbol{E}$, recalling from (1.32) that $\boldsymbol{K} = (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{b}$ is the curvature vector, and defining $\boldsymbol{G} = \nabla|\boldsymbol{B}|/(2|\boldsymbol{B}|)$, while $\nabla\boldsymbol{n}\cdot\boldsymbol{\beta}$ in the last equation denotes the vector with components $\boldsymbol{\beta}\cdot\partial\boldsymbol{n}/\partial x_i$. The second term at the right hand side of (1.80a) includes the drift velocities of the guiding center. We easily identify the last term of this cross product $\boldsymbol{b}\times\boldsymbol{N}$ as the $\boldsymbol{E}\times\boldsymbol{B}$-drift we met earlier in §1.5.2 (see equation (1.60)). The other two terms that corresppond to $\boldsymbol{b}\times\boldsymbol{K}$ and $\boldsymbol{b}\times\boldsymbol{G}$ are known as the curvature and grad-$|\boldsymbol{B}|$ drifts, respectively.

Since none of the vectors $\boldsymbol{b}, \boldsymbol{n}, \boldsymbol{\beta}, \boldsymbol{G}, \boldsymbol{K}, \boldsymbol{E}$ depends on $\widetilde{\zeta}$, equations (1.80) are truly independent of the new gyrophase as intended from the beginning. Therefore, the first five equations can be solved independently from the last one, and then in principle their solution could be substituted in (1.80d) to solve it by quadrature. In fact the latter step can be omitted, if someone is not interested in the evolution of the gyrophase. Furthermore, given the Poisson structure of the system one can prove that, similarly to the homogeneous case we saw in the previous section, the ignorable variable $\widetilde{\zeta}$ for first-order terms corresponds to a conserved quantity $\widetilde{\mu}$ through first-order terms as well, i.e. an *adiabatic invariant*,

$$\frac{d\widetilde{\mu}}{dt} = 0 + O(\epsilon^2)$$

This is the canonical conjugate to the new gyrophase, meaning $\{\widetilde{\mu}, \widetilde{\zeta}\} = 1$, that can actually be defined as the magnetic moment in the new coordinates, and which for $\epsilon \longrightarrow 0$ coincides with the original magnetic moment,

$$\widetilde{\mu} = \frac{\widetilde{v}_\perp^2}{2\,|\boldsymbol{B}(\widetilde{\boldsymbol{x}})|} = \mu + O(\epsilon) \tag{1.81}$$

Therefore, we can replace $\widetilde{u}_\perp$ and use $\widetilde{\mu}$ as a variable instead, reducing the system by one more degree. This means that if the solution $\widetilde{\boldsymbol{x}}(t)$ is found from (1.80a), then the perpendicular velocity can always be recovered by inverting (1.81), that is $v_\perp = \sqrt{2\widetilde{\mu}|\boldsymbol{B}(\widetilde{\boldsymbol{x}})|}$, where $\widetilde{\mu} = \widetilde{\mu}_0$ is determined from the transformed magnetic moment's initial condition. Thus, in conclusion system (1.80) is only left with the first four equations (1.80a)-(1.80b). These are the so-called *guiding-center equations*, which describe approximately the evolution of the particle in space that is of main interest, leaving out the less significant, very fast gyromotion. If we define a

modified magnetic direction $\widetilde{\boldsymbol{b}} = \boldsymbol{b} + \epsilon\, \widetilde{v}_\parallel \boldsymbol{b} \times \boldsymbol{K}/\omega_g$ their form can also be given as

$$\frac{d\widetilde{\boldsymbol{x}}}{dt} = \widetilde{v}_\parallel \widetilde{\boldsymbol{b}} + \frac{\epsilon}{\omega_g}\, \boldsymbol{b} \times (\widetilde{\mu}\, \nabla |\boldsymbol{B}| - \boldsymbol{E}) \tag{1.82a}$$

$$\frac{d\widetilde{v}_\parallel}{dt} = -\, \widetilde{\boldsymbol{b}} \cdot (\widetilde{\mu}\, \nabla |\boldsymbol{B}| - \boldsymbol{E}) \tag{1.82b}$$

We emphasize that all vector fields either in (1.80) or the above equations (1.82) are evaluated at $\widetilde{\boldsymbol{x}}$.

Complementary to the Hamiltonian formulation a Lagrangian one for the guiding-center motion has also be given by Littlejohn in [78] starting from (1.48). However, the guiding-center variables adopted in this section, considered by Weyssow and Balescu [4, 114], are slightly different from Littlejohn's in that particular paper[2]. Therefore the variational principle considered here for the guiding-center system hardly deviates from the one presented in [78] and also followed by most authors, e.g. White [118]. That said, the Lagrangian function in terms of the variables $\widetilde{\boldsymbol{x}}$, $\widetilde{v}_\parallel$, $\widetilde{\mu}$ and $\widetilde{\zeta}$ is expressed as

$$L = \left[\frac{\boldsymbol{A}}{\epsilon} + \widetilde{v}_\parallel \boldsymbol{b} - \epsilon \widetilde{\mu} \left(\nabla \boldsymbol{n} \cdot \boldsymbol{\beta} - \frac{1}{2} \nabla \times \boldsymbol{b}\right)\right] \cdot \dot{\widetilde{\boldsymbol{x}}} - \epsilon \widetilde{\mu} \dot{\widetilde{\zeta}} - \frac{1}{2}\, \widetilde{v}_\parallel^2 - \widetilde{\mu}\, |\boldsymbol{B}| - \Phi + O\big(\epsilon^2\big) \tag{1.83}$$

up to first-order terms. In order to see how the corresponding Euler-Lagrange equations describe the guiding-center equations, we start off with

$$\frac{\partial L}{\partial \widetilde{v}_\parallel} = 0 \Rightarrow \boldsymbol{b} \cdot \dot{\widetilde{\boldsymbol{x}}} - \widetilde{v}_\parallel = 0 \Rightarrow \widetilde{v}_\parallel = \boldsymbol{b} \cdot \dot{\widetilde{\boldsymbol{x}}} \tag{1.84}$$

An advantage of the Lagrangian formulation is that, in light of (1.83), the independence of the new gyrophase now straightforwardly yields the conservation of the transformed magnetic moment up to first-order terms,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\widetilde{\zeta}}}\right) = 0 \Rightarrow \epsilon\, \frac{d\widetilde{\mu}}{dt} = 0$$

As expected the equation for the conjugate variable $\widetilde{\mu}$ describes the evolution of $\widetilde{\zeta}$ given in (1.80d) with the help however of (1.84),

$$\begin{aligned} \frac{\partial L}{\partial \widetilde{\mu}} = 0 \Rightarrow -\epsilon \left(\nabla \boldsymbol{n} \cdot \boldsymbol{\beta} - \frac{1}{2} \nabla \times \boldsymbol{b}\right) \cdot \dot{\widetilde{\boldsymbol{x}}} - \epsilon \dot{\widetilde{\zeta}} - |\boldsymbol{B}| = 0 \\ \Rightarrow \dot{\widetilde{\zeta}} = -\frac{\omega_g}{\epsilon} - \widetilde{v}_\parallel \boldsymbol{b} \cdot \left(\nabla \boldsymbol{n} \cdot \boldsymbol{\beta} - \frac{1}{2} \nabla \times \boldsymbol{b}\right) \end{aligned}$$

2. Nonetheless, inconsistencies or merely misprints between the two formulations can also be found between the two papers [77] and [78] of Littlejohn.

The same calculations that led from (1.48) to (1.49) show that the rest of the equations with respect to the variables $\widetilde{x}_i$ result in

$$\frac{\partial L}{\partial \widetilde{x}_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\widetilde{x}}_i}\right) = 0 \Rightarrow \epsilon \dot{\widetilde{v}}_{\parallel} \boldsymbol{b} = \dot{\widetilde{\boldsymbol{x}}} \times \left(\boldsymbol{B} + \epsilon \widetilde{v}_{\parallel} \nabla \times \boldsymbol{b}\right) + \epsilon \left(\boldsymbol{E} - \widetilde{\mu} \nabla |\boldsymbol{B}|\right) \tag{1.85}$$

keeping first-order terms, meaning the third term appearing inside the bracket in (1.83) contributing as an $\epsilon^2$-order term in (1.85) is neglected. Before continuing it would be advantageous to define the modification of the magnetic field typically used in the literature,

$$\widetilde{\boldsymbol{B}} = \boldsymbol{B} + \epsilon \widetilde{v}_{\parallel} \nabla \times \boldsymbol{b} \tag{1.86}$$

and $\widetilde{B}_{\parallel} = \widetilde{\boldsymbol{B}} \cdot \boldsymbol{b}$, in terms of which the modified magnetic direction makes more sense, as $\widetilde{\boldsymbol{b}} = \widetilde{\boldsymbol{B}}/\widetilde{B}_{\parallel}$. Note that these considerations are not just mathematically convenient, as the divergence-free condition $\nabla \cdot \widetilde{\boldsymbol{B}} = 0$ still preserved demonstrates. The latter gives rise to a modified vector potential

$$\widetilde{\boldsymbol{A}} = \boldsymbol{A} + \epsilon \widetilde{v}_{\parallel} \boldsymbol{b}, \tag{1.87}$$

too. Moving on, crossing (1.85) with $\boldsymbol{b}$ gives $(\boldsymbol{b} \cdot \widetilde{\boldsymbol{B}})\, \dot{\widetilde{\boldsymbol{x}}} - (\boldsymbol{b} \cdot \dot{\widetilde{\boldsymbol{x}}}) \widetilde{\boldsymbol{B}} = \epsilon\, \boldsymbol{b} \times (\widetilde{\mu} \nabla |\boldsymbol{B}| - \boldsymbol{E})$, which substituting (1.84) and solving for $\dot{\widetilde{\boldsymbol{x}}}$ yields (1.82a),

$$\dot{\widetilde{\boldsymbol{x}}} = \frac{1}{\widetilde{B}_{\parallel}} \left[\widetilde{v}_{\parallel} \widetilde{\boldsymbol{B}} + \epsilon\, \boldsymbol{b} \times (\widetilde{\mu} \nabla |\boldsymbol{B}| - \boldsymbol{E})\right] = \widetilde{v}_{\parallel} \widetilde{\boldsymbol{b}} + \frac{\epsilon}{|\boldsymbol{B}|}\, \boldsymbol{b} \times (\widetilde{\mu} \nabla |\boldsymbol{B}| - \boldsymbol{E})$$

recalling once again $\omega_g = |\boldsymbol{B}|$. Finally dotting (1.85) with $\widetilde{\boldsymbol{B}}$, we easily get (1.82b),

$$\epsilon \dot{\widetilde{v}}_{\parallel} \widetilde{\boldsymbol{B}} \cdot \boldsymbol{b} = \epsilon \widetilde{\boldsymbol{B}} \cdot (\boldsymbol{E} - \widetilde{\mu} \nabla |\boldsymbol{B}|) \Rightarrow \dot{\widetilde{v}}_{\parallel} = \widetilde{\boldsymbol{b}} \cdot (\boldsymbol{E} - \widetilde{\mu} \nabla |\boldsymbol{B}|)$$

Comparing with section 1.4 we see that, while the Lagrangian of the guiding center comes from the Lagrangian of the parallel and perpendicular motion, the corresponding Euler-Lagrange equations in the two cases require different manipulation to arrive at the equations of motion.

## 1.7 Relativistic motion

Electromagnetic fields often accelerate charged particles to very high energies, due to their small masses, like the electron's mass. When particles reach velocities close to the speed of light, relativistic effects must be taken into account. In order to do so, the equations of motion (1.3) have to be reformulated

$$\begin{aligned} \frac{d\boldsymbol{x}}{dt} &= \frac{\boldsymbol{u}}{\gamma(\boldsymbol{u})}, \\ \frac{d\boldsymbol{u}}{dt} &= \frac{1}{\gamma(\boldsymbol{u})}\, \boldsymbol{u} \times \boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{E}(\boldsymbol{x}). \end{aligned} \tag{1.88}$$

where $\boldsymbol{u}$ is the proper velocity. The latter is the spatial component of the four-velocity defined (as the tangent vector to a timelike world line) in the Minkwoski spacetime, and its relation to the velocity $\boldsymbol{v}$ is $\boldsymbol{u} = \gamma \boldsymbol{v}$, where $\gamma(v) = 1/\sqrt{1-(v/c)^2}$ is the Lorentz factor, $c$ being the speed of light. From the last two equations we can also express the Lorentz factor in terms of the proper velocity,

$$\gamma(u) = \sqrt{1 + \frac{u^2}{c^2}} \tag{1.89}$$

One way to pass from Newtonian Mechanics to the realm of Special Relativity is to suitably modify the Lagrangian or the Hamiltonian function for a given problem. Of course, this is not always possible, but the nature of the Lorentz force allows to do so, because, recalling Remark 1.2, it stems from a generalized potential, which is linear in the velocities. So, this route will also be followed here to recover the Hamiltonian formulation for the relativistic charged particle motion. In this context, the kinetic energy of the system $mv^2/2$ is replaced by $mc^2 = m_o \gamma c^2$, where $m_o$ is the rest mass. Therefore, from (1.12) and (1.89) the Hamiltonian function[3] in terms of the position and the proper velocity is

$$H(\boldsymbol{x}, \boldsymbol{u}) = \gamma(\boldsymbol{u})c^2 + \Phi(\boldsymbol{x}) = c\sqrt{c^2 + \boldsymbol{u}^2} + \Phi(\boldsymbol{x}). \tag{1.90}$$

We should stress that this cannot be considered as a change of variables from $\boldsymbol{v}$ to $\boldsymbol{u}$, meaning that the Poisson structure remains the same. Only the Hamiltonian changed based on physics arguments, regarded as the particle's energy. Thus, system (1.88) can be described as a Hamiltonian system with the same Poisson structure matrix (1.13), and the above Hamiltonian function,

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial u_i} \\ \frac{du_i}{dt} &= \epsilon_{ijk} B_k \frac{\partial H}{\partial u_j} - \frac{\partial H}{\partial x_i} \end{aligned} \tag{1.91}$$

This can be easily verified since, using again (1.89), the derivatives of $H$ with respect to the proper velocities are $\partial H/\partial u_i = c u_i/\sqrt{c^2 + u^2} = u_i/\gamma$.

The canonical formulation follows a similar path. The relativistic momenta are $\boldsymbol{p} = \boldsymbol{u} + \boldsymbol{A}$, in terms of which the previous Hamiltonian can be exressed as

$$H(\boldsymbol{x}, \boldsymbol{p}) = c\sqrt{c^2 + (\boldsymbol{p} - \boldsymbol{A}(\boldsymbol{x}))^2} + \Phi(\boldsymbol{x}). \tag{1.92}$$

3. setting $m_o = 1$ following the previously adopted normalisation

Using (1.92), the relativistic charged particle motion (1.88) can be described by the usual canonical form of Hamilton's equations,

$$\begin{aligned}
&\frac{dx_i}{dt}=\frac{\partial H}{\partial p_i} \quad \Rightarrow \quad \frac{dx_i}{dt}=c\frac{p_i-A_i}{\sqrt{c^2+(\boldsymbol{p}-\boldsymbol{A})^2}} \quad \Rightarrow \quad \frac{dx_i}{dt}=\frac{u_i}{\gamma}\,,\\
&\frac{dp_i}{dt}=-\frac{\partial H}{\partial x_i} \quad \Rightarrow \quad \frac{dp_i}{dt}=c\frac{p_j-A_j}{\sqrt{c^2+(\boldsymbol{p}-\boldsymbol{A})^2}}\frac{\partial A_j}{\partial x_i}-\frac{\partial \Phi}{\partial x_i}\\
&\qquad\Rightarrow \quad \frac{dp_i}{dt}=\frac{u_j}{\gamma}\frac{\partial A_j}{\partial x_i}-\frac{\partial \Phi}{\partial x_i}-\frac{dA_i}{dt}+\frac{dA_i}{dt}\\
&\qquad\Rightarrow \quad \frac{dp_i}{dt}-\frac{dA_i}{dt}=\frac{u_j}{\gamma}\frac{\partial A_j}{\partial x_i}-\frac{\partial \Phi}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\frac{dx_j}{dt}\\
&\qquad\Rightarrow \quad \frac{d\left(p_i-A_i\right)}{dt}=\frac{u_j}{\gamma}\left(\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\right)-\frac{\partial \Phi}{\partial x_i}\\
&\qquad\Rightarrow \quad \frac{du_i}{dt}=\frac{1}{\gamma}\,\epsilon_{ijk}u_jB_k+E_i.
\end{aligned}$$

Last but not least, as in the nonrelativistic case and in particular equation (1.17), a Lagrangian function can also be constructed for system (1.88) via the relation $L=\boldsymbol{p}\cdot\dot{\boldsymbol{x}}-H$ giving

$$L\left(\boldsymbol{x},\boldsymbol{u},\dot{\boldsymbol{x}}\right)=\left(\boldsymbol{A}(\boldsymbol{x})+\boldsymbol{u}\right)\cdot\dot{\boldsymbol{x}}-c\sqrt{c^2+\boldsymbol{u}^2}-\Phi(\boldsymbol{x}) \tag{1.93}$$

### 1.7.1 Magnetostatic fields

In the absence of an electric field, $\boldsymbol{E}=\boldsymbol{0}$, the conservation of the relativistic Hamiltonian (1.90) leads to a constant Lorentz factor $\gamma$. Let us consider the proper time $\tau$, that is the time measured by a clock travelling with the particle, whose relation with the time $t$ of the original inertial frame of reference is $d\tau=dt/\gamma$. Then one easily sees that in this case the relativistic equations of motion (1.88) in terms of $\tau$, $\boldsymbol{x}$ and $\boldsymbol{u}$ are the same as the nonrelativistic equations for a charged particle moving in the same magnetic field in terms of $t$, $\boldsymbol{x}$ and $\boldsymbol{v}$. Therefore, if the solution $(\boldsymbol{x}(t),\boldsymbol{v}(t))$ in the nonrelativistic limit is known, then the same problem (i.e. for the same magnetic field) in the relativistic case has the general solution $(\boldsymbol{x}(t/\gamma),\boldsymbol{v}(t/\gamma)/\gamma)$ for the position and velocity of the particle.

### 1.7.2 Guiding center motion

The guiding-center theory has also been extended for the relativistic motion, and in terms of a Hamiltonian formulation was first given by Grebogi and Littlejohn [41]. Simplifications to this theory via a covariant form were made later by Brizard and Chan first for magnetostatic fields [17] and then in the general case [104]. Essentially the route followed is identical with the

nonrelativistic case along with one practical rule, that is the replacement $v_i \longrightarrow u_i$ of ordinary velocities with proper ones in all guiding-center considerations. This means that, under the same guiding-center transformation only this time in terms of the proper velocity components $u_\parallel$ and $u_\perp$, system (1.88) yields the relativistic counterpart of the guiding-center equations (1.82) that read

$$\begin{aligned} \frac{d\widetilde{\boldsymbol{x}}}{dt} &= \frac{\widetilde{u}_\parallel}{\widetilde{\gamma}}\,\widetilde{\boldsymbol{b}} + \frac{\epsilon}{\omega_g}\,\boldsymbol{b}\times\left(\frac{\widetilde{\mu}}{\widetilde{\gamma}}\,\nabla|\boldsymbol{B}| - \boldsymbol{E}\right) \\ \frac{d\widetilde{u}_\parallel}{dt} &= -\,\widetilde{\boldsymbol{b}}\cdot\left(\frac{\widetilde{\mu}}{\widetilde{\gamma}}\,\nabla|\boldsymbol{B}| - \boldsymbol{E}\right) \end{aligned} \tag{1.94}$$

where $\widetilde{u}_\parallel$ is the parallel proper velocity of the guiding center, while now $\widetilde{\boldsymbol{b}} = \boldsymbol{b} + \epsilon\,\widetilde{u}_\parallel \boldsymbol{b}\times\boldsymbol{K}/\omega_g$, $\widetilde{\mu} = \widetilde{u}_\perp^2/(2\,|\boldsymbol{B}(\widetilde{x})|)$ and $\widetilde{\gamma} = \sqrt{1+\widetilde{u}_\parallel^2/c^2 + 2\,\widetilde{\mu}\,|\boldsymbol{B}|\,/c^2} = \gamma + O(\epsilon)$ is the modified Lorentz factor. Notice that the usual angular frequency $\omega_g = |\boldsymbol{B}|$ enters in these expressions, not to be confused with the relativistic one introduced in [41]. The Hamiltonian function in terms of $\left(\widetilde{\boldsymbol{x}}, \widetilde{u}_\parallel, \widetilde{u}_\perp\right)$ for comparison with (1.79) is

$$H\left(\widetilde{\boldsymbol{x}}, \widetilde{u}_\parallel, \widetilde{u}_\perp\right) = \widetilde{\gamma}\left(\widetilde{\boldsymbol{u}}\right)c^2 + \Phi(\widetilde{\boldsymbol{x}}) = c\sqrt{c^2 + \widetilde{u}_\parallel^2 + \widetilde{u}_\perp^2} + \Phi(\widetilde{\boldsymbol{x}}) + O\left(\epsilon^2\right). \tag{1.95}$$

while the Lagrangian in terms of the magnetic moment instead of the perpendicular velocity for comparison with (1.83) is accordingly

$$L = \left[\frac{\boldsymbol{A}}{\epsilon} + \widetilde{u}_\parallel\boldsymbol{b} - \epsilon\widetilde{\mu}\left(\nabla\boldsymbol{n}\cdot\boldsymbol{\beta} - \frac{1}{2}\nabla\times\boldsymbol{b}\right)\right]\cdot\dot{\widetilde{\boldsymbol{x}}} - \epsilon\widetilde{\mu}\dot{\widetilde{\zeta}} - c\sqrt{c^2 + \widetilde{u}_\parallel^2 + 2\,\widetilde{\mu}\,|\boldsymbol{B}|} - \Phi + O\left(\epsilon^2\right) \tag{1.96}$$

As with the exact relativistic equations the Poisson matrix, just like the rest quantities, remains formally the same under the substitution $v_i \longrightarrow u_i$.

## Final notes

The first developments of guiding-center theory were made by Northrop and others and relied purely on averaging methods, lacking though a Hamiltonian structure. Among other things, this was also a disadvantage from the statistical mechanics point of view of plasma physics, since most kinetic models start off from the celebrated Liouville's theorem, a key characteristic of a Hamiltonian single particle motion. Kruskal and in particular Littlejohn with his series of papers were the first to establish the guiding-center approach retaining the Hamiltonian formulation of the problem. Canonical representations were also given by White. Following earlier works of Northrop, Littlejohn derived the variational principle for the guiding-center equations, too.

The guiding-center variables used in section 1.6 are the primed overbarred variables used by Littlejohn in [77]. These have the advantage of the physically meaningful Hamiltonian (1.79) that does not have $\epsilon$-order terms, the disadvantage being, as we can see in [77], a non-block-diagonal Poisson matrix. The situation is reversed with the other set of variables, denoted by just overbars, adopted in that same article, in terms of which first-order terms appear now in the Hamiltonian function, but the Poisson matrix is block-diagonal. Although its simple form (see for example equation (17) in [41] for the corresponding Poisson bracket in the relativistic case; compare with (1.57) for the simple case of a homogeneous magnetostatic field), having a block-diagonal Poisson matrix is not a small restriction and better dropped. In the reduction procedure the canonical conjugacy between the new gyrophase and the magnetic moment suffices.

As opposed to the autonomous system only considered in this chapter, the guiding-center theory has been extended for the time-dependent case, as well. In [4], Balescu generalizes the guiding-center approximation to include electromagnetic fields that, besides a weak inhomogeneity, exhibit a slow variation in time. This can be easily achieved, using the standard Hamiltonian treatment of non-autonomous systems by considering the extended phase space, where the time $t$ and the Hamiltonian $H$ itself serve as two more independent (and in fact canonical) variables (see appendix C). A more general approach in charged particle dynamics includes the so-called *ponderomotive effects*. The latter appear in the case of high frequency electromagnetic waves, in which case the guiding-center notion is replaced by the *oscillation center theory*. The oscillation center traces the motion of a charged particle when both the fast gyrofrequency and the high frequency of the wave are averaged out. The Hamiltonian formulation of the oscillation center motion was studied first by Grebogi and Littlejohn [40, 41] and also later by Weyssow and Balescu [115, 116]. However, the notion of the oscillation center fails near the resonance of the wave-particle interaction, which is of main interest in chapter 3, and so was not presented here.

In more special cases, besides the magnetic moment, additional adiabatic invariants may also exist. Rosenbluth suggested the *longitudinal invariant* for charged particle motion between two *magnetic mirrors*, i.e. points of converging magnetic field lines where particles are reflected. Motivated by Van Allen radiation belts in the earth's magnetic field, Northrop and Teller defined a third adiabatic invariant expressing the *magnetic flux* through a guiding center's orbit for slowly time-varying electromagnetic fields. For a description of these adiabatic motions see for example [47, 70].

# Chapter 2

# Integrable Perturbed Magnetic Fields in Toroidal Geometry

One of the most important aspects of tokamak research is the appearance of neoclassical tearing modes (NTM), exhibiting magnetic islands, and their stabilization. Breaking the investigation of their control down to a single mode, a preliminary step required is the exact resolution of the perturbed topology of the magnetic field. To this end, an analytical form of the islands generated is requested, which can be provided by a magnetic surface quantity. The integrability of the magnetic field lines in the case of such simple perturbations allows to find one and, therefore, calculate the Poincare surface of section analytically. Thus, all the information related (position of o- and x-points, separatrix, island width, etc.) can be recovered in full precision, without the need of more or less approximate considerations or numerical analysis.

Given the Hamiltonian nature of the magnetic field, the plasma equilibrium, described by an axisymmetric integrable system, is better expressed in terms of action-angle variables. In this context, magnetohydrodynamic (MHD) instabilities can be represented by a perturbation term in the Hamiltonian function as an infinite series of modes. For NTM-like perturbations, which, in general, include only the modes with the same helicity, i.e. the same ratio, the system still remains integrable due to helical symmetry. Thus, a surface quantity can be derived from an effective Hamiltonian, which is related to the helical flux of the system. Consisting still though of infinite modes, a Taylor expansion in the neighborhood of the resonant surface must be made first, in order to obtain the island chain corresponding to a specific mode. As fully described in [47] and [118], the resemblance then with the pendulum Hamiltonian leads to direct conclusions. More about the formation of main and sattelite islands can be found in [86, 110].

In practice, however, when studying magnetic island effects on other phenomena, one particular mode is considered right from the start. For these kinds of situations, no series expansion is required and the problem reduces to simply analyzing the effective Hamiltonian. Nonetheless, the latter is not separable and may, in general, be quite complicated, far from typical mechanical systems. Thus, we present how to determine the perturbed topology and all its characteristics for these cases in a straightforward and simple manner. In doing so, we take on a geometric approach, the bottom-line being the projection of this new Hamiltonian in the desired poloidal cross section. The technique presented is not bound to the precise form of the Hamiltonian function, and could be applied in similar cases.

In addition - and although we use the large aspect ratio approximation for the equilibrium magnetic field - for the purpose of higher accuracy that is usually demanded, the actual toroidal geometry is adopted instead of the approximated cylindrical one generally used [34, 91, 93, 119]. All along, emphasis is also given on the Hamiltonian formulation, in terms of which the non-canonical form of the toroidal coordinates reflects the corrections one should consider from the cylindrical approximation to the toroidal one. We particularly focus on the construction of action-angle variables, which, apart from being more convenient, they are also more suitable to express the perturbation of the system properly. These aspects are pointed out, because of their significance in the structure of the magnetic field. For, in plasma terminology, canonical coordinates are Clebsch coordinates and vice-versa, while action-angle variables are flux coordinates, [28, 47]. The opposite, though, in not always true, which is actually the case, when the radial magnetic component of the background equilibrium is zero. The latter condition, met by the standard magnetic field that is quite often considered, relates the Hamiltonian description with more physical quantities. As will be shown, it also allows an easy manipulation of going back to the original toroidal coordinates. When carefully carried out, then the perturbed topology may be regarded directly in terms of them. As it turns out, toroidal coordinates, dominating among others, when it comes to simulating real tokamak experiments, can be utilized rather plainly with no approximations whatsoever. The outcome hopefully throws some light on these matters and provides a systematic, yet clear way of accessing all the island information needed.

## 2.1 Axisymmetric systems

The magnetic fields under consideration are the ones applied in toroidal configurations, such as tokamaks. Therefore, we introduce a (right-handed) toroidal coordinate system $\boldsymbol{x} = (r, \theta, \phi)$,

where $r$ is the minor radius of the torus, $\theta$ is the poloidal angle measured counterclockwise from the outer edge, and $\phi$ is the toroidal angle measured clockwise from the $y$-axis. The transformation from Cartesian coordinates is

$$\begin{aligned} x &= (R_0 + r\cos\theta)\sin\phi \\ y &= (R_0 + r\cos\theta)\cos\phi \\ z &= r\sin\theta \end{aligned}$$

where $R_0$ is the major radius of the torus and we usyally denote $R = R_0 + r\cos\theta$. As described in the appendix A the unit base is $(\widehat{\boldsymbol{e}}_r, \widehat{\boldsymbol{e}}_\theta, \widehat{\boldsymbol{e}}_\phi)$, while the covariant one $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\phi)$.

The background equilibrium of a tokamak plasma is very often approximated by an axisymmetric magnetic field $\boldsymbol{B}$, i.e. independent of $\phi$. Thus, as explained in chapter 5, we can equivalently start off with a vector potential $\boldsymbol{A}(r,\theta)$, being, too, independent of $\phi$, meaning (see equation (A.35))

$$\boldsymbol{B}(r,\theta) = \frac{1}{\sqrt{g}}\left[\frac{\partial A_\phi}{\partial\theta}\boldsymbol{e}_r - \frac{\partial A_\phi}{\partial r}\boldsymbol{e}_\theta + \left(\frac{\partial A_\theta}{\partial r} - \frac{\partial A_r}{\partial\theta}\right)\boldsymbol{e}_\phi\right] \tag{2.1}$$

where $g = (rR)^2$ is the determinant of the metric tensor defined by the toroidal coordinates, and $A_r, A_\theta, A_\phi$ are the covariant components of $\boldsymbol{A}$.

The dynamics of the magnetic field in three-dimensional Euclidean space is given by the set of equations (1.18) and (1.21) accordingly in terms of the vector potential. The Hamiltonian structure of the magnetic field lines has been investigated extensively [15, 21, 32, 58, 73, 87]. In the case of axisymmetry, system (1.21) can be easily casted into Hamiltonian form, in light of (2.1). We can go directly to the third step of §1.2.2 and adopt the usual technique of treating the ignorable coordinate $\phi$ as the new independent variable, while eliminating the previous one $s$, arriving at

$$\begin{aligned} \frac{d\theta}{d\phi} &= \frac{B^\theta}{B^\phi} \\ \frac{dr}{d\phi} &= \frac{B^r}{B^\phi} \end{aligned} \tag{2.2}$$

as long as $B^\phi \neq 0$. Then, considering $B^r, B^\theta$ and $B^\phi$ from (2.1) suggests choosing the Hamiltonian function as $H(r,\theta) = -A_\phi(r,\theta)$, while the symplectic structure as $\omega = \sqrt{g}B^\phi\, dr \wedge d\theta$. In other words, system (2.2) can be written as a two-dimensional Hamiltonian system in non-

canonical variables,

$$\begin{aligned} \frac{d\theta}{d\phi} &= \frac{1}{\sqrt{g}B^{\phi}}\frac{\partial H}{\partial r} \\ \frac{dr}{d\phi} &= -\frac{1}{\sqrt{g}B^{\phi}}\frac{\partial H}{\partial \theta} \end{aligned} \tag{2.3}$$

Due to the axisymmetry of $\boldsymbol{B}$, i.e. the independency of the toroidal angle, carried over to the Hamiltonian, this system is autonomous and therefore integrable. Thus, action-angle variables, $\psi$ and $\theta$, can be constructed, in terms of which the above system takes the form

$$\begin{aligned} \frac{d\theta_{\text{in}}}{d\phi} &= H'(\psi) \\ \frac{d\psi}{d\phi} &= 0 \end{aligned} \tag{2.4}$$

Its solutions, lying on surfaces that are topologically equivalent to the torus, are then simply $\psi = \text{const.}$ and $\theta_{\text{in}} = \text{w}(\psi)\phi + \theta_{\text{in},0}$, where $\text{w}(\psi) = H'(\psi)$ and $\theta_{\text{in},0}$ some constant. In fusion literature, $\theta_{\text{in}}$ is commonly known as the *intrinsic poloidal angle*, while the function $\text{w}(\psi)$ as the *winding number*. Its inverse, denoted by $q(\psi)$, is called the *safety factor* and has the meaning of the number of turns of the magnetic field lines around the toroidal angle $\phi$ per one turn along the intrinsic poloidal angle $\theta_{\text{in}}$. The triad $(\psi, \theta_{\text{in}}, \phi)$, in this case, qualifies as a flux coordinate system, in terms of which the magnetic field assumes the Clebsch representation.

## 2.2 Helical perturbations

The previous system, coming from a Hamiltonian function of the form $H_0(\psi)$, is a simplified picture of actual experiments. In real tokamak plasmas, MHD instabilities introduce small perturbations to the ideal equilibrium magnetic field. These can be modelled by considering a perturbed Hamiltonian function [2],

$$H(\psi, \theta_{\text{in}}, \phi) = H_0(\psi) + \epsilon H_1(\psi, \theta_{\text{in}}, \phi), \tag{2.5}$$

in terms of the action-angle variables $(\psi, \theta_{\text{in}})$ and the "time" $\phi$, where $\epsilon$ is the perturbation strength. We should note that these types of perturbations do not affect in any way the component $B^{\phi}$ of the magnetic field. Its Clebsch representation is also retained by the same variables, with respect to the new Hamiltonian, but these are no longer flux coordinates, since $H$ is no more independent of $\theta_{\text{in}}$ and $\phi$. In order for $H_1$ to be single-valued, it has to be a $2\pi$-periodic function of the angles $\theta_{\text{in}}$ and $\phi$. Thus, it can always be presented as a Fourier series:

$$H_1(\psi, \theta_{\text{in}}, \phi) = \sum_m \sum_n f_{mn}(\psi) \cos(m\theta_{\text{in}} - n\phi). \tag{2.6}$$

Tearing instabilities though involve helical perturbations ([47], section 7.3), i.e. terms with the same ratio $m/n$, reducing the above double sum to a single one. Moreover in order to investigate the effects of the perturbed magnetic topology, one is often bound to concentrate on simply one resonance mode $(m, n)$, i.e. one particular term of (2.6)

$$H_1(\psi, \theta_{\text{in}}, \phi) = f_{mn}(\psi)\cos(m\theta_{\text{in}} - n\phi) \tag{2.7}$$

These types of perturbations introduce systems, which are also integrable, owing now to the helicity of Hamiltonian $H(\psi, m\theta_{\text{in}} - n\phi)$, [47]. The integrability in this case can be shown in many ways, the easiest of which is probably by replacing $\theta_{\text{in}}$ with the new variable $\xi = m\theta_{\text{in}} - n\phi$, for which the equations for the magnetic field lines,

$$\begin{aligned} \frac{d\xi}{d\phi} &= m\frac{d\theta_{\text{in}}}{d\phi} - n = m\frac{\partial H}{\partial \psi} - n = \frac{\partial h}{\partial \psi} \\ \frac{d\psi}{d\phi} &= -\frac{\partial H}{\partial \theta_{\text{in}}} = -m\frac{\partial H}{\partial \xi} = -\frac{\partial h}{\partial \xi} \end{aligned}, \tag{2.8}$$

are casted again to Hamiltonian form [109], using $h(\psi, \xi) = mH(\psi, \xi) - n\psi$ as the new, *effective Hamiltonian*. The latter is independent of $\phi$, and therefore system (2.8) is integrable, meaning the magnetic field lines lie on the isosurfaces $h = \text{const.}$. Notice that $(\psi, \xi, \phi)$, besides being suitable for describing the symmetry of the system, they are Clebsch coordinates, as well, if $h$ is considered instead of $H$ in the representation of the magnetic field. Still, they, too, are not flux coordinates.

So, the previous invariant surfaces $\psi = \text{const.}$ of the unperturbed system are now replaced by $h(\psi, \xi) = \text{const.}$. A small perturbation, however, affects mostly the so-called resonant surfaces. These are the rational surfaces of the unperturbed system, for which $q(\psi_s) = m/n$. Their change in topology becomes more visible by the use of Poincare maps, which follow the (transversal) intersection of all solutions of the system with a certain plane or surface, called Poincare surface of section. The latter, being suitable for two-dimensional non-autonomous dynamical systems with bounded orbits, let us visualise the three-dimensional dynamics in order to study its structure.

## 2.3 Magnetic islands

As previously shown, the function $h$ characterizes the magnetic surfaces and, thus, can serve as a *flux surface label*. The latter can be useful in many ways, one of which is the analytical construction of Poincare plots.

In the case of magnetic field lines, where the role of "time" is played by the toroidal angle, Poincare surfaces are simply obtained by poloidal cross sections of the torus, that is $\phi = \text{const.}$. So, instead of loooking at the full three-dimensional space in terms of $\psi, \theta_{\text{in}}$ and $\phi$, we restrict ourselves to a Poincare surface of section in terms only of $\psi$ and $\theta_{\text{in}}$ chosen at a fixed $\phi$. For system (2.8) though, construction of a Poincare map numerically is no longer necessary, since we can directly obtain Poincare plots by exploiting the effective Hamiltonian.

Consider the restricted function $h_\phi(\psi, \theta_{\text{in}}) = h(\psi, \theta_{\text{in}}, \phi)$ for fixed $\phi$. Since the magnetic field lines lie on the surface $h\,(\psi, \xi) = \text{const.}$, their intersection with a poloidal cross section $\phi = \text{const.}$ is given by the equations $h_\phi(\psi, \theta_{\text{in}}) = \text{const.}$. Therefore, a contour plot of the function $h_\phi$ would simply yield the desired Poincare surface of section. The critical points of $h_\phi$ projected on the surface $(\psi, \theta_{\text{in}})$ correspond then to the equilibrium points of the Poincare map, representing periodic orbits of the system. The maxima or minima give rise to *centers*, simply called o-points, while the saddle points accordingly to *saddles*, simply called x-points.

In conclusion, when only one particular resonant $(m, n)$ is under investigation, meaning a perturbation of the form (2.7) is added to the unperturbed system $H_0$, magnetic surfaces can be characterized by a flux surface label,

$$h\,(\psi, \theta_{\text{in}}, \phi) = mH(\psi, \theta_{\text{in}}, \phi) - n\psi \tag{2.9}$$

and Poincare surfaces of section at any given $\phi = \phi_c$ can be constructed by the contour plot of the function $h_{\phi_c}$. Both o- and x-points can be determined from the condition $\nabla h_\phi = 0$, which, for $f_{mn}(\psi) \neq 0$ at least in a neighborhood around $\psi_s$, trivially results in

$$m\theta_{\text{in}} - n\phi = k\pi, \tag{2.10}$$

$$\mathrm{w}(\psi) + (-1)^k \epsilon f'_{mn}(\psi) = \frac{n}{m}, \tag{2.11}$$

where $k$ is any integer. From the first equation we can find the $2m$ angles $\theta_{\text{in},i}$ in the interval $[0, 2\pi)$ for different values of $k$, while from the second one the two actions $\psi_i$ depending on whether $k$ is even or odd. To determine which case corresponds to the o-points and which to the x-points we turn to the Hessian matrix of $h_\phi$, calculated at $(\psi_i, \theta_{\text{in},i})$. Thus, whenever (2.10)-(2.11) hold, the eigenvalues of this matrix are

$$\begin{aligned} \lambda_1 &= (-1)^{k+1} \epsilon m^3 f_{mn}(\psi_i) \\ \lambda_2 &= m\Big[\mathrm{w}'(\psi_i) + (-1)^k \epsilon f''_{mn}(\psi_i)\Big] \end{aligned} \tag{2.12}$$

When $\lambda_1 \lambda_2 > 0$ we have an o-point, while in the opposite case, $\lambda_1 \lambda_2 < 0$, we have an x-point. Finally, the equation for the *separatrix* is $h_\phi(\psi, \theta_{\text{in}}) = h_\phi(\psi_{\text{x}}, \theta_{\text{in,x}})$ for $(\psi_{\text{x}}, \theta_{\text{in,x}})$ any x-point.

In many cases, e.g. for a strictly monotonous profile of the safety factor and consequently of the winding number, as well, $\lambda_2$ is defined mostly by the first term, since $\epsilon$ is a small parameter. Therefore, the product $\lambda_1\lambda_2$ has the same sign as $(-1)^{k+1} f_{mn}(\psi_i)\mathrm{w}'(\psi_i)$ has. The latter, changing sign for successive values of $k$, explains the interchange of o- and x-points, causing the island chain formation we typically see. In light of (2.10), the number of islands is then equal to the poloidal mode number $m$. From equation (2.11), on the other hand, it is evident that o- and x-points are $\epsilon$-close to the previously formed resonant surface. Taylor expansion around $\psi_s$ suggests that the deviation from the resonant surface to a first order approximation is $\delta\psi = \psi_i - \psi_s = (-1)^{k+1}\epsilon f_{mn}(\psi_s)/\mathrm{w}'(\psi_s)$, recovering the results in [119]. Finally, it is also worth noting that the value $\theta_{\mathrm{in},0}$ of the intrinscic poloidal angle at $\phi = 0$, introduced from the very beginning, may serve through (2.10) as a rotation parameter of the islands, without affecting whatsoever the unperturbed system. The privilege of equations (2.10) and (2.11) being independent of each other does not allow the position $\psi_i$ of o- and x-points to be disturbed by the presence of $\theta_{\mathrm{in},0}$, too.

## 2.4 Tokamak fields

Before proceeding with a concrete detailed example, a special class of axisymmetric systems to begin with is considered, which is widely used in applications. These are unperturbed magnetic fields with vanishing radial component, $B^r = 0$. The latter reflects an equilibrium for large aspect ratio, meaning we do not take into account the poloidal current density, nor the Shafranov shift, though we retain the toroidal geometry, opposite to the cylindrical often used. A number of properties arising regarding the previous decription is the reason why such systems are investigated seperately as commonplaces of tokamaks.

First of all we have $dr/ds = 0$. Therefore, the invariant surfaces are defined simply by $r = \text{const.}$, meaning the magnetic field lines, in this case, lie on the surface of a torus of radius $r$. In terms of the Hamiltonian description (2.3)[1], this also implies that $H_0$ is a function of $r$ alone, representing the normalized *poloidal magnetic flux*,

$$H_0(r) = \int_0^r \sqrt{g} B^\theta \, dr = \frac{1}{2\pi} \int_0^{2\pi} \int_0^r \widehat{B}^\theta(r,\theta) R dr d\phi \tag{2.13}$$

1. Note that for $B^r = 0$ we could also adopt the alternative Hamiltonian description given in §1.2.1. Since however we want to study perturbations of such a field, we retain the previously described Hamiltonian considerations, following more likely the path shown in §1.2.2.

On the other hand, the construction of action-angle variables in the general case of an axisymmetric system would first require finding canonical coordinates for system (2.3). When $B^r = 0$, it may skip this step. Actually, the transformation from $(r, \theta)$ to $(\psi, \theta_{\text{in}})$ can be made via the relation $\sqrt{g}B^\phi\, dr \wedge d\theta = d\psi \wedge d\theta_{\text{in}}$. Since $r$ is an integral of motion, $\psi$ has to be a function of $r$ alone. Consequently, the previous condition can be simplified, yielding

$$\frac{d\psi}{dr}\frac{\partial\theta_{\text{in}}}{\partial\theta} = r\widehat{B}^\phi. \tag{2.14}$$

Thus, we can begin with a given function $\psi(r)$ and then construct $\theta_{\text{in}}(r, \theta)$ or vice versa. A typical choice, widely used in the literature, is expressing the action $\psi$ through the normalized *toroidal magnetic flux*,

$$\psi(r) = \frac{1}{2\pi}\int_0^{2\pi}\int_0^r \widehat{B}^\phi(r, \theta) r dr d\theta \tag{2.15}$$

while considering $\theta_{\text{in}}$ to be a $2\pi$-periodic function of $\theta$. In this case the effective Hamiltonian $h$ introduced earlier is the *helical flux*. It is also worth noticing that $\psi$ depending solely on $r$ (regardless the form of the function $\psi(r)$) indicates that the mixed variables $(r, \theta_{\text{in}}, \phi)$ are also flux (though, not Clebsch) coordinates for the unperturbed case.

Another consequence of the zero radial magnetic component is an alternate expression of the winding number directly in terms of $r$ instead of $\psi$. For when $B^r = 0$, then

$$\mathrm{w}(r) = \frac{d\theta_{\text{in}}}{d\phi} = \frac{\partial\theta_{\text{in}}}{\partial\theta}\frac{d\theta}{d\phi} = \frac{R\widehat{B}^\theta}{r\widehat{B}^\phi}\frac{\partial\theta_{\text{in}}}{\partial\theta} \tag{2.16}$$

Inverting the previous relation and taking into account that $\theta_{\text{in}}(r, \theta + 2\pi) = \theta_{\text{in}}(r, \theta) + 2\pi$ we recover the formula

$$q(r) = \frac{1}{2\pi}\int_0^{2\pi} \frac{r\widehat{B}^\phi}{R\widehat{B}^\theta}\, d\theta \tag{2.17}$$

found in [2], [113]. Equivalently, once $\psi$ is fixed, w and consequently $q$ can be calculated more simply using

$$\mathrm{w}(r) = \frac{dH_0}{d\psi} = \frac{1}{\psi'(r)}\frac{dH_0}{dr} = \frac{R\widehat{B}^\theta}{\psi'(r)} \tag{2.18}$$

At this point, we should comment that although action-angle variables have been used in Sections 3 and 4, this is by no means a restriction. But in the general case of an axisymmetric system the transformation to the actual toroidal coordinates $r$ and $\theta$ would be much complicated. In this case, however, the latter task would be a simple one. Actually, the unperturbed Hamiltonian and the safety factor have already been expressed naturally in terms of $r$. The determination of the o- and x-points for the perturbed system could also follow. Equation (2.11)

can be solved with respect to $r$ still independently from (2.10), and then replace its solutions $r_i$ in (2.10) to find $\theta_i$. From equation (2.16), we also deduce that when $\partial\theta_{\rm in}/\partial\theta \longrightarrow 1$, then the number of turns around $\phi$ per one turn along either $\theta_{\rm in}$ or $\theta$ is the same. So, in fact, all the above conclusions allow us to draw one more. Since $r$ is an integral of motion (for the unperurbed system) just like the action $\psi$, as long as the intrinsic poloidal angle behaves like the geometrical one, the island topology, realised in the artificial $\psi\theta_{\rm in}$-space, is carried over to the more realistic $r\theta$-plane. On this ground, $B^r = 0$ allows us to switch easily from the action-angle variables, appealing in theory, to the original toroidal coordinates, as desired in practice, and study the dynamics of the magnetic field lines therefrom.

### 2.4.1 The standard magnetic field

A typical axisymmetric model for the background equilibrium that falls into the previous category is the so-called *standard magnetic field*, introduced by Balescu [5],

$$\boldsymbol{B}_0(r,\theta) = \frac{B_0}{R}\left(r\mathrm{w}_c(r)\,\widehat{\boldsymbol{e}}_\theta + R_0\,\widehat{\boldsymbol{e}}_\phi\right), \tag{2.19}$$

where $B_0$ is a constant, expressing the toroidal field on the magnetic axis, while $\mathrm{w}_c$ is the winding number in the approximation of cylindrical geometry, i.e. for $r/R_0 \to 0$. The Hamiltonian equations (2.3) for the standard magnetic field are

$$\frac{d\theta}{d\phi} = \frac{R_0 + r\cos\theta}{B_0 R_0\, r}\,\frac{dH_0}{dr}, \qquad \frac{dr}{d\phi} = 0, \tag{2.20}$$

where the Hamiltonian function can be deduced directly from the poloidal flux (2.13), yielding

$$H_0(r) = B_0 \int r\mathrm{w}_c(r)\,dr \tag{2.21}$$

Choosing $\psi$ as the toroidal flux (2.15) and then using (2.14), we end up with the following action-angle variables [2],

$$\psi = B_0 R_0 \left(R_0 - \sqrt{R_0^2 - r^2}\right) \tag{2.22}$$

$$\theta_{\rm in} = 2\arctan\left(\sqrt{\frac{R_0 - r}{R_0 + r}}\tan\frac{\theta}{2}\right) \tag{2.23}$$

Notice that in the large aspect ratio approximation, often used in applications, $\theta_{\rm in}$ is very similar to $\theta$, as requested. So, proceeding in terms of the toroidal coordinates, from equation (2.18) the actual winding number w with respect to the approximate one $\mathrm{w}_c$ is

$$\mathrm{w}(r) = \frac{\sqrt{R_0^2 - r^2}}{R_0}\,\mathrm{w}_c(r). \tag{2.24}$$

Finally, from (2.11), the radial position $r_i$ of o- and x-points for any mode $(m,n)$ we choose to perturb the system with can be found through

$$\mathrm{w}(r_i) + (-1)^k \epsilon f'_{mn}(r_i) \frac{\sqrt{R_0^2 - r_i^2}}{B_0 R_0 r_i} = \frac{n}{m} \tag{2.25}$$

and then the corresponding poloidal angle at any given cross section $\phi = \phi_p$ would be

$$\theta_i = 2 \arctan\left( \sqrt{\frac{R_0 + r_i}{R_0 - r_i}} \tan \frac{\theta_{\mathrm{in},i}}{2} \right) \tag{2.26}$$

where $\theta_{\mathrm{in},i} = (n\phi_p + k\pi)/m$ from (2.10).

Returning to an earlier comment, we see that the replacement $\theta_{\mathrm{in}} \longrightarrow \theta_{\mathrm{in}} - \theta_{\mathrm{in},0}$ enters only in (2.26), changing $\theta_i$ to some other angle $\theta_i + \widetilde{\theta}_0$, leaving, in general, $\theta$ unchanged. Thus, we can rotate the island topology, without rotating the whole system that would alter, for example, the typical $(1/R)$-dependence of the unperturbed toroidal field or the radial position of the o- and x-points. This is a common requirement in experiments such as NTM integration with electron cyclotron current drive, where electromagnetic waves are launched, targeting the o-points. This way of rotating the islands is independent of the specific Hamiltonian or action-angle variables.

### 2.4.2 An example: The (3,2) mode

The associated Hamiltonian system (2.3) with $H = H_0 + \epsilon H_1$ describes the field lines of the magnetic field $\boldsymbol{B} = \boldsymbol{B}_0 + \epsilon \boldsymbol{B}_1$. If the standard magnetic field (2.19) plays the role of the unperturbed system, corresponding to the unperturbed Hamiltonian $H_0$ (2.21) and giving rise accordingly to the action-angle variables (2.22)-(2.23), then any Hamiltonian perturbation $H_1$ of the form (2.7) in terms of the latter will in turn result in a magnetic perturbation

$$\boldsymbol{B}_1 = \frac{m\sqrt{R_0^2 - r^2}}{rR^2} f_{mn}(r) \sin\xi\, \widehat{\boldsymbol{e}}_r + \frac{1}{R}\left[ f'_{mn}(r)\cos\xi + \frac{mR_0 \sin\theta}{R\sqrt{R_0^2 - r^2}} f_{mn}(r) \sin\xi \right] \widehat{\boldsymbol{e}}_\theta \tag{2.27}$$

expressed back in the toroidal coordinate system, where $\xi = m\theta_{\mathrm{in}} - n\phi$. Notice that for small values of $r$ the last term of the poloidal component in (2.27) would be neglegible compared to the others.

Let us demonstrate the previous methods with an example of a $(m,n) = (3,2)$ resonance mode for the standard magnetic field with ITER-like parameters, $B_0 = 5.51\,\mathrm{T}$ and $R_0 = 6.2\,\mathrm{m}$. Starting from density and temperature profiles, we consider [85]

$$\mathrm{w}_c(r) = \frac{1}{4}\left(2 - \frac{r^2}{r_0^2}\right)\left(2 - 2\frac{r^2}{r_0^2} + \frac{r^4}{r_0^4}\right), \tag{2.28}$$

where $r_0 = 1.9\,\mathrm{m}$ is the maximum value of $r$, meaning the minor radius of the tokamak. Thus, from (2.21) the Hamiltonian of the unperturbed system is

$$H_0(r) = \frac{B_0 r^2}{2}\left(1 - \frac{3r^2}{4r_0^2} + \frac{r^4}{3r_0^4} - \frac{r^6}{16r_0^6}\right), \tag{2.29}$$

while the resonant surface is evaluated at $r_s = 96.5743\,\mathrm{cm}$. Following [91], we assume a perturbation amplitude of the form

$$f_{32}(r) = \frac{r}{3}\left(1 + \frac{r - r_s}{b}\right), \tag{2.30}$$

expressed directly in terms of $r$ instead of $\psi$, where $b = 12$ is a parameter that allows to describe asymmetry around the island. The effective hamiltonian, meaning the flux surface label (2.9), for this magnetic field in terms of $r$ and the helical variable $\xi = 3\theta_{\mathrm{in}} - 2\phi$ is accordingly

$$h(r, \xi) = 3H_0(r) - 2\psi(r) + 3\epsilon f_{32}(r)\cos\xi \tag{2.31}$$

Replacing $\theta_{\mathrm{in}}$ from (2.23) we can then arrive at an expression in the original toroidal coordinates. In the next figure, we see for $\epsilon = 0.005$ the level sets, that is, the contour plot of the restriction $h_\phi$ (2.32) of $h$ for $\phi = \pi/2$, i.e. in the poloidal cross section $y = 0$, revealing the island topology.

$$h_{\frac{\pi}{2}}(r, \theta) = 3H_0(r) - 2\psi(r) - 3\epsilon f_{32}(r)\cos(3\theta_{\mathrm{in}}(r, \theta)) \tag{2.32}$$

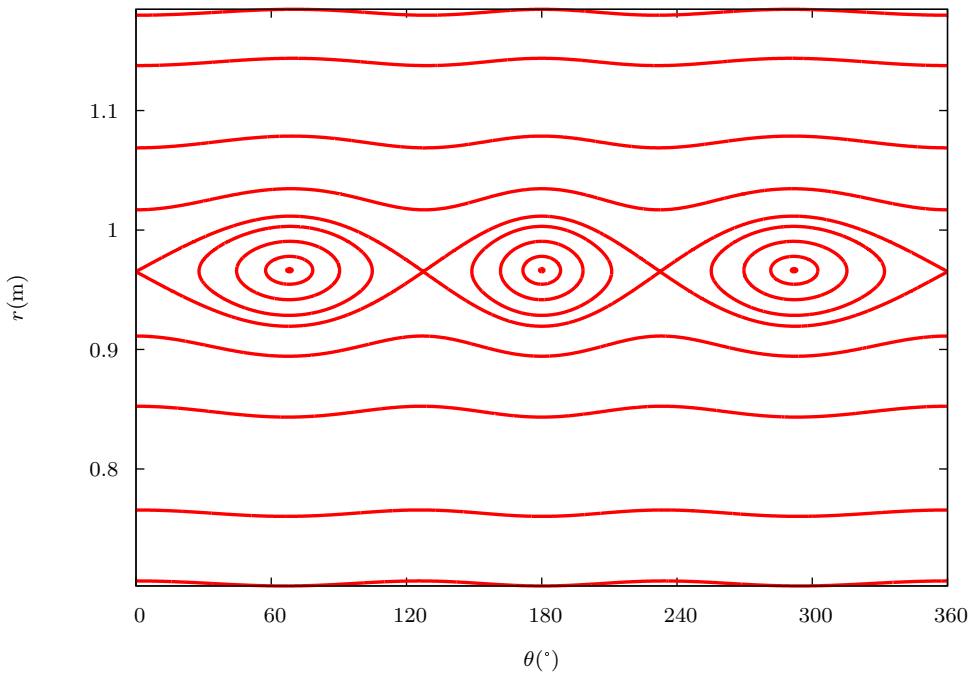


Figure 2.1: The island topology for the (3,2) resonance.

The contour plot of (2.32) back in the Cartesian coordinates $x$ and $z$ is shown in Figure 2.2. The same Poincare section was drawn using numerical integration of the full three-dimensional system $d\boldsymbol{x}/ds$ (1.18) with a 4th-order adaptive step-size Runge-Kutta scheme. Comparison of the two plots reveals actually no difference whatsoever, they look identical. Figure 2.3 shows the points from the numerical Poincare map of the magnetic field lines lying exactly on the contour lines of the analytical flux surface label $h_{\frac{\pi}{2}}$.

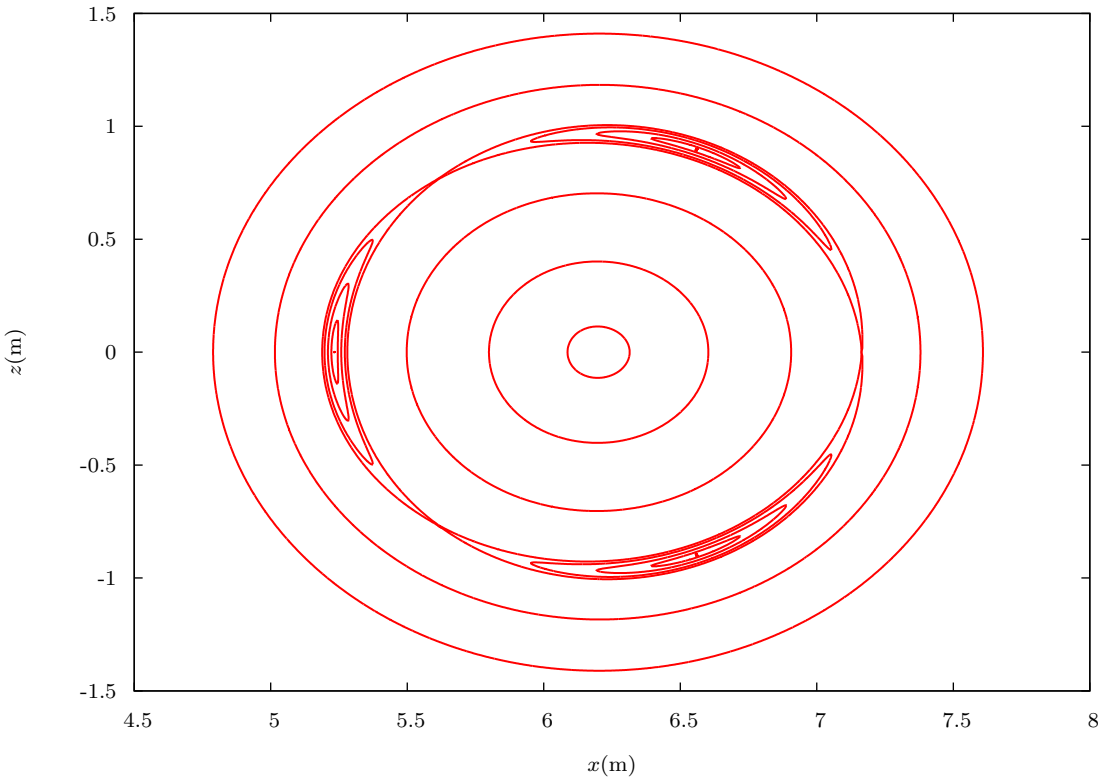


Figure 2.2: Contour plot of flux surface label.

For the calculation of the o- and x-points, we notice first that for $0 \leq r \leq r_0$ the perturbation $f_{32}$ is everywhere positive, the winding number w is decreasing monotonously and the toroidal flux $\psi$ is increasing instead. Thus, from the relations (2.12), assuming $\lambda_2 \approx 3\mathrm{w}'(r)/\psi'(r)$, we immediately deduce that when $k$ is odd, the critical point of the function $h_{\frac{\pi}{2}}$ would be a saddle point, resulting in an x-point, while when $k$ is even it would be a maximum, resulting in an o-point. So, relations (2.25) and (2.26) yield six (real) solutions altogether, three o-points at $r_\mathrm{o} = 96.6336\,\mathrm{cm}$, $\theta_\mathrm{o} = 68.0854°, 180°, 291.915$ and three x-points at $r_\mathrm{x} = 96.5149\,\mathrm{cm}$, $\theta_\mathrm{x} = 0°, 127.468°, 232.532°$, respectively. These values, also visualised in Figures 2.1 and 2.2 are

in absolute agreement with the ones from field line tracing. We see that o- and x-points located at $r_{\mathrm{o}}$ and $r_{\mathrm{x}}$ respectively are evenly distributed on different sides and at equal distance from the previous resonance surface located at $r_s$.

It is also worth noting that in the cylindrical approximation, often adopted in the large aspect ratio limit, we recover the usual considerations $\psi \approx B_0 r^2/2$, $\theta_{\mathrm{in}} \approx \theta$ (as previously requested), $\mathrm{w} \approx \mathrm{w}_c$ and accordingly $\boldsymbol{B}_1$ expresses a (3,2) mode in terms of $\theta$ and $\phi$ instead of $\theta_{\mathrm{in}}$ and $\phi$. The deviations, however, from the above treatment are not negligible, as, for example, the corresponding values $\theta_{\mathrm{in,o}} = 60°, 180°, 300°$ and $\theta_{\mathrm{in,x}} = 0°, 120°, 240°$ of the intrinsic poloidal angle indicate. As we can see in Figure 2.1, in a toroidal configuration these all lead to a slight difference between the central island and the other two, unlike the cylindrical case.

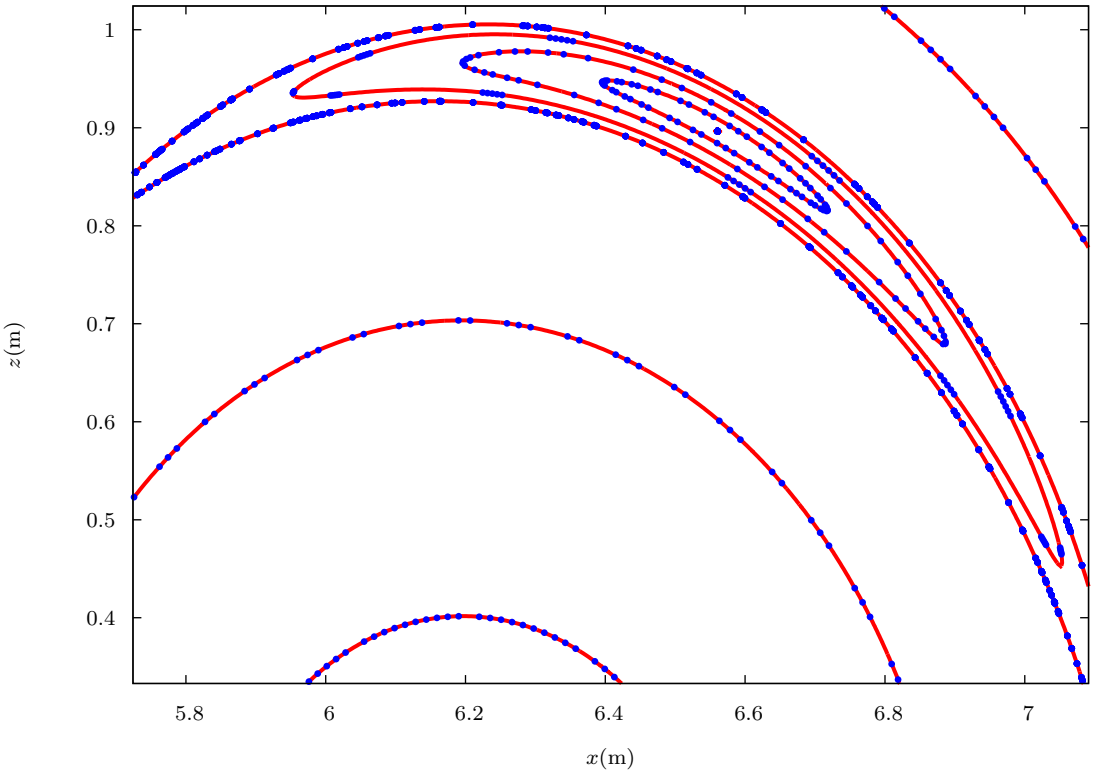


Figure 2.3: Comparison of the analytical flux surface label (red lines) and the numerical Poincare map (blue points), zoomed in the island region.

Finally, instead of using the restriction $h_\phi$ for Poincare sections, we can take by all means the original helical flux $h$ and from its level sets plot 3D magnetic surfaces, shown for example in Figure 2.4 back in Cartesian coordinates $x$, $y$ and $z$.

## 2.5 Discussion

In conclusion, a simple analytical way for determining the island topology of the magnetic field is at our disposal, when a single perturbation mode is introduced to the plasma equilibrium. Poincare sections of field lines can be constructed by contour plotting a flux surface label that is consistent with the magnetic field, and the positions of o- and x-points, as well as the separatrix, can be determined analytically. The method addressed follows the nontrivial Hamiltonian nature of the magnetic field lines in terms of their toroidal structure. The integrability of this kind of systems is widely known in the context of Hamiltonian mechanics, yet, at times, neglected in applications such as these. And though quite often employed in the cylindrical approximation, to the author's knowledge, it has not been fully utilized in the actual toroidal geometry of tokamaks. The technique described here, requiring no assumptions on the particular form of the integrable Hamiltonian, is quite general and could be applied elsewhere. The results presented in this chapter have also been published in [62].

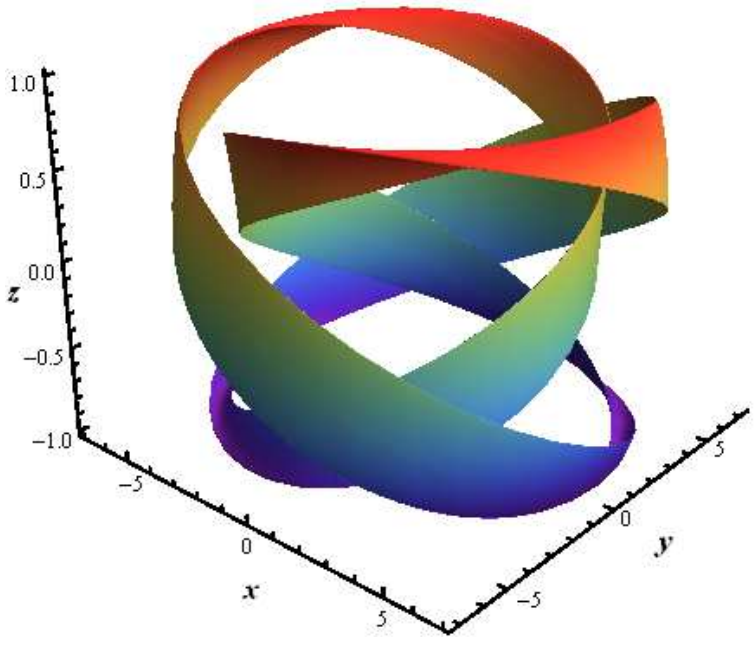


Figure 2.4: A magnetic surface for the (3,2) resonance.

# Chapter 3

# Magnetic Islands and Electromagnetic Waves in Plasma Physics

MHD instabilities in the form of neoclassical tearing modes are one of the most crucial factors in the operation of tokamak devices such as ITER. Introducing non-axisymmetric, helical magnetic perturbations on flux surfaces, which correspond to rational values $(m, n)$ of the safety factor, they result in the formation of magnetic islands that gradually lead to loss of confinement. Thus their control is an important issue that has received much attention, particularly focusing on the (2,1) and (3,2) modes, which are most likely to appear [68]. One of the most successful methods for the stabilization of these perturbations is to apply radio frequency waves, used for current drive in the plasma. When the wavefrequency is close to the cyclotron frequencies of the electrons, the method is known as Electron Cyclotron Resonance Heating (ECRH). In this mechanism, electromagnetic waves are launched into the plasma from an external source outside of the torus targeting the islands.

In most cases, ECRH in plasma is studied within either linear or quasilinear considerations using the Vlasov equation that describes the evolution of the particle distribution function. Our purpose is to investigate the combined effects of magnetic islands and electron cyclotron resonance in terms of the equations of charged particle motion and their full nonlinear character. So, in this chapter, we study numerically the relativistic motion of test particles in toroidal plasma configurations under the influence of a perturbed equilibrium magnetic field and an electromagnetic wave. We particularly focus on the wave-particle interaction in the presence of

the (3,2) island and an electron cyclotron beam, taking also into account collisional effects with background particles. Following large numbers of electrons, macroscopic results, such as velocity distribution functions, current drive, absorbed power, are obtained that are tested against results from the linear theory of wave-particle interaction, and we find reasonable qualitative agreement between the two approaches.

## 3.1 Magnetic perturbations

The model used for the magnetic field is the one described in the previous chapter and in particular §2.4.2. We briefly recall and repeat the basic features for easy reference. Starting with the coordinate system, we also adopt here the toroidal coordinates $(r, \theta, \phi)$ introduced at the beginning of section 2.1. As already mentioned, the latter define a right-handed orthogonal curvilinear system. In the literature, however, the toroidal angle is often measured counterclockwise from the $x$-axis, that is $\pi/2 - \phi$, leading though to a left-handed system, and consequently to several mix-ups, when handling quantities in toroidal geometry, like vector products, curls, etc., as in the case of the guiding-center approximation.

In the above defined coordinate system, the unperturbed magnetic field $\boldsymbol{B}_0$ expressing the background equilibrium is given by the standard magnetic field (2.19), where $\mathrm{w}_c$ is cylindrically modelled by (2.28). Characterized by a non-zero toroidal component, it allows magnetic field lines to be viewed as a Hamiltonian system (2.3) in variables $r$ and $\theta$ with $\phi$ playing the role of "time". Axisymmetry and zero radial component on the other hand result accordingly to a Hamiltonian function $H_0$ (2.29) that is independent of both $\phi$ and $\theta$. Therefore $\boldsymbol{B}_0$ gives rise to toroidal magnetic surfaces $r = \text{const}$.

The perturbed magnetic field $\boldsymbol{B}_1$ is coming from a Hamiltonian perturbation $H_1$ of the previous system properly defined in action-angle variables $\psi$ (2.22) and $\theta_{\mathrm{in}}$ (2.23). Focusing on a single perturbation mode $(m, n)$, we assume $H_1 = f_{mn}(\psi)\cos\xi$ which gives rise to $\boldsymbol{B}_1$ given in (2.27), where $\xi = m\theta_{\mathrm{in}} - n\phi$ is the helical variable. The overall magnetic field

$$\begin{aligned}\boldsymbol{B} = \boldsymbol{B}_0 + \epsilon\boldsymbol{B}_1 = \epsilon\,\frac{m\sqrt{R_0^2 - r^2}}{rR^2} f_{mn}(r)\sin\xi\,\widehat{\boldsymbol{e}}_r + \\ + \frac{1}{R}\left[B_0 r\mathrm{w}_c(r) + \epsilon\left(f'_{mn}(r)\cos\xi + \frac{mR_0\sin\theta}{R\sqrt{R_0^2 - r^2}} f_{mn}(r)\sin\xi\right)\right]\widehat{\boldsymbol{e}}_\theta + \frac{B_0 R_0}{R}\,\widehat{\boldsymbol{e}}_\phi\end{aligned} \tag{3.1}$$

projected in two dimensions results in an island chain, which replaces the resonant surface $r = r_s$ of the unperturbed system. The location of the island chain is defined by the relation $q(r_s) = m/n$, where $q$ is the safety factor given by the inverse of the winding number (2.24).

For the (3,2) mode, $f_{32}$ given in (2.30) and ITER-like parameters collected in Table 3.1, Figures 2.1 and 2.2 show the island formation in the poloidal cross section $y = 0$ distorting the previous straight lines or circles, respectively, $r = \text{const.}$ of the unperturbed system.

In this setup, the Hamiltonian formalism is not without a physical interpretation: the unperturbed Hamiltonian function $H_0$ is the poloidal flux, the action $\psi$ is the toroidal flux, while the effective Hamiltonian

$$h = mH - n\psi = mH_0 - n\psi + \epsilon m f_{mn} \cos\xi$$

is the helical flux, where $H = H_0 + \epsilon H_1$ is the perturbed Hamiltonian. Finally, the somewhat artificial angle $\theta_{\text{in}}$ behaves very much like the actual poloidal angle $\theta$, and in the large aspect ratio $r/R_0 \longrightarrow 0$ we have $\theta_{\text{in}} = \theta$. This is in fact the only price we pay, a rather small one, to have at our disposal an analytical, exact quantity to extract any information required regarding the magnetic field lines. For, as shown in section 2.2, the perturbed system is also integrable and the magnetic field $\boldsymbol{B}$ now lies on the helical surfaces $h = \text{const.}$ demonstrated in Figure 2.4. In other words, the function $h$ is an invariant of the magnetic field lines and can serve directly as a magnetic surface quantity, without the need of Taylor expansion, or other approximations that are often considered. Provided with a flux surface label $h$, we determine the periodic orbits of the field lines, as well as their stability, and consequently the location of o-, x-points and the separatrix all of them analytically and in full precision. The previously mentioned 2D Poincare sections 2.1 and 2.2 or the 3D one 2.4 have also been constructed analytically by contour plotting. In addition, later on in §3.4.2 the volumes of flux surfaces can too be calculated safely and accurately avoiding any numerical error likely to arise when approaching the o-point.

**Remark 3.1.** For practical purposes, after all calculations are made, the flux surface label $h$ is normalised in order to be consisent with other approaches often found the literature, i.e. the o-points correspond to the value $h = -1$, while the separatrix to $h = 1$. So, for example, the magnetic surface shown in Figure 2.4 corresponds to $h = -0.437$.

## 3.2 Electromagnetic wave

One of the most common ways of heating a plasma is the use of radio frequency waves in the form of a focused beam, exploiting ECRH. We assume a relatively simple yet realistic model of a Gaussian, monochromatic beam, launched from the low-field side outside of the torus, injected into the plasma and propagating in the direction

$$\boldsymbol{k} = -k\left(\cos\theta_l \sin\phi_l, \cos\theta_l \cos\phi_l, \sin\theta_l\right), \tag{3.2}$$

with respect to the Cartesian coordinate system, where $\theta_l$ and $\phi_l$ are the poloidal and toroidal launching angles, respectively. In consideration of having a wave beam that is confined through a Gaussian shape in the perpendicular to propagation direction, we assume that the electric field of the wave is multiplied by a factor $\exp(-d^2/2s^2)$, where $s$ is the width of the beam and $d$ the distance from a point $\boldsymbol{x}$ to the beam axis. If the latter passes through a fixed point $\boldsymbol{x}_\mathrm{o}$, then this distance is

$$d = |\boldsymbol{x} - \boldsymbol{x}_\mathrm{o}| \sin\zeta = \sqrt{|\boldsymbol{x} - \boldsymbol{x}_\mathrm{o}|^2 - \left[(\boldsymbol{x} - \boldsymbol{x}_\mathrm{o}) \cdot \hat{\boldsymbol{k}}\right]^2}, \tag{3.3}$$

where $\zeta$ is the angle between the propagation vector $\boldsymbol{k}$ and the vector $\boldsymbol{x} - \boldsymbol{x}_\mathrm{o}$, and $\hat{\boldsymbol{k}}$ the unit vector along $\boldsymbol{k}$.

Now, in order to determine the wavenumber $k$ and the polarization of the wave, we use the homogeneous-plasma wave equation, derived from Maxwell's equations, which after Fourier analysis in both time and space yields

$$\boldsymbol{k} \times (\boldsymbol{k} \times \boldsymbol{E}) + \frac{\omega^2}{c^2}\, \boldsymbol{\epsilon} \cdot \boldsymbol{E} = 0, \tag{3.4}$$

where $\omega$ is the wave frequency, $c$ the speed of light, $\boldsymbol{E}$ the electric field of the wave and $\boldsymbol{\epsilon}$ the linear dielectric tensor. The general form of the latter is given by a $3 \times 3$ complex matrix, whose elements, for hot, weakly-relativistic plasma and assuming that at the time of entry the background magnetic field $\boldsymbol{B}_0$ is parallel to the $y$-axis, are equal to [117]:

$$\begin{aligned}
\epsilon_{11} &= 1 - \beta_T \frac{\omega_p^2}{\omega^2} \sum_{l=-\infty}^{\infty} \frac{l^2}{\lambda} \Gamma_{|l|} \mathcal{F}_{|l|+3/2}, \\
\epsilon_{12} &= \beta_T \frac{\omega_p^2}{\omega\omega_g}\, n_\parallel n_\perp \sum_{l=-\infty}^{\infty} \frac{l}{\lambda} \Gamma_{|l|} \left(\mathcal{F}_{|l|+3/2} - \mathcal{F}_{|l|+5/2}\right), \\
\epsilon_{13} &= -i\beta_T \frac{\omega_p^2}{\omega^2} \sum_{l=-\infty}^{\infty} l \Gamma'_{|l|} \mathcal{F}_{|l|+3/2}, \\
\epsilon_{22} &= 1 - \beta_T \frac{\omega_p^2}{\omega^2} \sum_{l=-\infty}^{\infty} \Gamma_{|l|} \left[\beta_T n_\parallel^2 \left(\mathcal{F}_{|l|+7/2} - 2\mathcal{F}_{|l|+5/2} + \mathcal{F}_{|l|+3/2}\right) + \mathcal{F}_{|l|+5/2}\right], \\
\epsilon_{23} &= i\beta_T \frac{\omega_p^2}{\omega\omega_g}\, n_\parallel n_\perp \sum_{l=-\infty}^{\infty} \Gamma'_{|l|} \mathcal{F}_{|l|+5/2}, \\
\epsilon_{33} &= 1 - \beta_T \frac{\omega_p^2}{\omega^2} \sum_{l=-\infty}^{\infty} \left(\frac{l^2}{\lambda} \Gamma_{|l|} \mathcal{F}_{|l|+3/2} + 2\lambda \Gamma'_{|l|} \mathcal{F}_{|l|+5/2}\right),
\end{aligned} \tag{3.5}$$

while $\epsilon_{21} = \epsilon_{12}$, $\epsilon_{31} = \epsilon_{13}$ and $\epsilon_{32} = -\epsilon_{23}$. In the above, $\omega_g = q_e|\boldsymbol{B}|/m_e$ is the electron gyrofrequency, $\omega_p = q_e\sqrt{n_e/m_e\epsilon_o}$ is the plasma frequency, $\epsilon_o$ is the dielectric constant of free

space, $l$ is the cyclotron harmonic under consideration, $\boldsymbol{n} = c\boldsymbol{k}/\omega$ is introduced for simplicity, $\beta_T = m_e c^2/k_{\rm B}T_e$ is the inverse of the normalized thermal energy, $\Gamma_l(\lambda) = e^{-\lambda} I_l(\lambda)$ with $I_l$ the modified Bessel function of argument $\lambda = k_\perp^2 v_{\rm th}^2/\omega_g^2$ (the parameter $\lambda$ measures the ratio of the gyroradius over the wavelength), and $\mathcal{F}_q(\alpha, \zeta_l)$ are the Shkarofsky functions of arguments $\alpha = \beta_T n_\parallel^2/2$ and $\zeta_l = \beta_T\,(1 - l\omega_g/\omega)$.

Equation (3.4) is a linear, homogeneous system in terms of the Cartesian components $E_1$, $E_2$ and $E_3$ of the electric field. The matrix describing this system has elements $n_i n_j - n^2\delta_{ij} + \epsilon_{ij}$ (where $\delta_{ij}$ is Kronecker's delta) and in order to have non-trivial solutions for $\boldsymbol{E}$, its determinant must vanish, leading to the so-called *dispersion relation* that relates the wavenumber $k$ to the frequency $\omega$ of the wave,

$$D(\omega, \boldsymbol{k}) = \begin{vmatrix} \epsilon_{11} - n_2^2 - n_3^2 & \epsilon_{12} + n_1 n_2 & \epsilon_{13} + n_1 n_3 \\ \epsilon_{21} + n_1 n_2 & \epsilon_{22} - n_1^2 - n_3^2 & \epsilon_{23} + n_2 n_3 \\ \epsilon_{31} + n_1 n_3 & \epsilon_{32} + n_2 n_3 & \epsilon_{33} - n_1^2 - n_2^2 \end{vmatrix} = 0, \tag{3.6}$$

The left hand side is a six-degree polynomial in terms of $k$, having accordingly six roots. For hot plasma all these roots are complex, with the real part of the refraction index representing propagation and the imaginary part collisionless absorption. The complete solution of the (complex) dispersion relation is a much complicated numerical issue (see e.g. [103]). A typical simplification, which is followed also here, is to assume that the imaginary part of $k$ is an order smaller than its real part (weak-absorption ordering), and also to prescribe the parallel component $k_\parallel$ (following the experimental practice). In this way, only the Hermitian part of $\boldsymbol{\epsilon}$ is involved in the solution and an explicit expression for the real part may be achieved.

Having satisfied condition (3.6), the system of equations (3.4) then gives rise to the polarization of the wave, i.e. the relations between the different components of $\boldsymbol{E}$. Since the rank of the system is two, we need to use only two equations, and then the third one would be trivially satisfied. Thus, keeping for example $E_1$ arbitrary, we may solve for $E_2 = c_2 E_1/c_1$ and $E_3 = c_3 E_1/c_1$, where $c_i = (-1)^{3+i} D_{3i} D_{31}^*$ and $D_{ij}$ denotes the subdeterminant of $D$ by crossing out the $i$-th row and the $j$-th column. Since the components $E_i(\omega, \boldsymbol{k})$ are complex functions, they need to be transformed back to the real quantities $E_i(t, \boldsymbol{x})$. Taking the inverse Fourier transformation $\boldsymbol{E}(t, \boldsymbol{x}) = \mathrm{Re}\left(\boldsymbol{E}(\omega, \boldsymbol{k}) e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}\right)$, we arrive at

$$\begin{aligned} E_1(t, \boldsymbol{x}) &= E_0\, c_1 e^{-d^2/2s^2} \cos\left(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t\right) \\ E_2(t, \boldsymbol{x}) &= E_0\, c_2 e^{-d^2/2s^2} \cos\left(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t + \eta_2\right) \\ E_3(t, \boldsymbol{x}) &= E_0\, c_3 e^{-d^2/2s^2} \cos\left(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t + \eta_3\right) \end{aligned} \tag{3.7}$$

where the angles $\eta_2 = \arctan(\mathrm{Im}\, c_2/\mathrm{Re}\, c_2)$ and $\eta_3 = \arctan(\mathrm{Im}\, c_3/\mathrm{Re}\, c_3)$ express the phase differences of $E_2$ and $E_3$, respectively, with respect to $E_1$, and $E_0$ is a constant.

In the case of this Gaussian beam, given the wave power $P_0$ one can recover $E_0$ as follows: For plane waves, the Poynting vector is given as $\boldsymbol{S} = c\epsilon_o |\boldsymbol{E}|^2 \hat{\boldsymbol{k}}$, and so the instantaneous wave power for (3.7) is

$$P = \int_a \boldsymbol{S} \cdot d\boldsymbol{a} = c\epsilon_o \widetilde{E}_0 \int_0^{2\pi} \int_0^{\infty} e^{-\rho^2/s^2} d\rho d\phi = c\epsilon_o \pi s^2 \widetilde{E}_0^2 \tag{3.8}$$

where $a$ is the cross-sectional area of the beam, $(\rho, \phi)$ are polar coordinates in a beam cross section and $\widetilde{E}_0^2 = E_0^2 \left(c_1^2 \cos^2(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t) + c_2^2 \cos^2(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t + \eta_2) + c_3^2 \cos^2(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t + \eta_3)\right)$. Thus, the wave power, defined as the averaged power over the wave period $T$, is

$$\begin{aligned} P_0 = \langle P \rangle &= \frac{1}{T} \int_0^T P\, dt \\ &= \frac{c\epsilon_o \pi s^2 E_0^2}{T} \int_0^T \left[c_1^2 \cos^2(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t) + c_2^2 \cos^2(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t + \eta_2) + c_3^2 \cos^2(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t + \eta_3)\right] dt \\ &= \frac{1}{2} c\epsilon_o \pi s^2 E_0^2 \left(c_1^2 + c_2^2 + c_3^2\right) \end{aligned} \tag{3.9}$$

From equation (3.9) we can then determine $E_0$ in terms of $P_0$, arriving at

$$E_0 = \frac{1}{s} \sqrt{\frac{2P_0}{\pi c\epsilon_o \left(c_1^2 + c_2^2 + c_3^2\right)}} \tag{3.10}$$

Finally, taking into account the damping of the wave energy as the beam propagates in the plasma, we consider a profile for the wave power along the direction of the beam, as well,

$$P_0 = P_{\max} \left[b_1 \tanh(b_2(\sigma - b_3)) + b_4\right], \tag{3.11}$$

where $P_{\max}$ is the initial wave power, $\sigma$ measures the length along the beam and $b_i$ are constants. Equation (3.11) is based on the absorption modelled using ray tracing [56], for which a detailed description will be given later on.

## 3.3 Particle dynamics

For the test-particle simulations and the particle orbits, we integrate the relativistic equations of the charged particle motion. The latter include as force terms the perturbed magnetic field

(3.1) and the electric wave field (3.7), previously described. As a first approach, the magnetic field of the wave is not taken into account since it does not directly affect the energization (but it is accountable for other effects and should be considered later on). We follow large numbers of particles, starting from a specific flux-surface within the island and around the o-point, at which the wave beam passes by, with the particles initially having a Maxwellian velocity distribution with prescribed temperature $T_e$ (which is the local value of the electron plasma temperature in the island region).

### 3.3.1 Equations of motion

The code is given two options for the equations of motion that are being integrated, either the full Lorentz force or the guiding-center equations, both to be chosen in the relativistic or nonrelativistic limit. For the integration of both, the Lorentz force and the guiding-center equations, we use a 4th-order, adaptive step-size Runge-Kutta scheme, implemented in Fortran.

#### Lorentz force law

Considering the more general case when relativistic effects are not negligible, the exact equations of motion (1.88) in Cartesian coordinates, restoring physical units, are

$$\begin{aligned} \frac{d\boldsymbol{x}}{dt} &= \frac{\boldsymbol{u}}{\gamma}, \\ \frac{d\boldsymbol{u}}{dt} &= \frac{q_e}{m_e}\left(\frac{1}{\gamma}\,\boldsymbol{u}\times\boldsymbol{B}+\boldsymbol{E}\right). \end{aligned} \tag{3.12}$$

where $m_e$ and $q_e$ are the mass and charge of the electron, respectively. Given a time-step of the order of the particle gyration period, the numerical integration of the above equations seems to run very slow. On the other hand, for longer integration times, the conservation of energy is not well-behaved anyway, as it should in the collisionless case. The guiding-center approximation, bypassing the fine structure of the gyromotion, overcomes both of these setbacks.

#### Guiding-center approximation

As previously described in chapter 1, for weakly inhomogeneous fields, the previous six equations can be reduced, by eliminating the gyrophase at least through first-order terms. The Hamiltonian nature of the system relates the ignorable gyrophase with the adiabatic invariance of the conjucated magnetic momemt. Thus, the system is reduced by two, leaving four equations to describe the position of the guiding center and its parallel to the magnetic field velocity. In the

relativistic version (1.94), restoring again physical parameters, these are

$$\begin{aligned} \frac{d\widetilde{\boldsymbol{x}}}{dt} &= \frac{\widetilde{u}_{\parallel}}{\widetilde{\gamma}}\,\widetilde{\boldsymbol{b}} + \frac{1}{m_e\omega_g}\,\boldsymbol{b}\times\left(\frac{\widetilde{\mu}}{\widetilde{\gamma}}\,\nabla|\boldsymbol{B}| - q_e\boldsymbol{E}\right) \\ \frac{d\widetilde{u}_{\parallel}}{dt} &= -\frac{1}{m_e}\,\widetilde{\boldsymbol{b}}\cdot\left(\frac{\widetilde{\mu}}{\widetilde{\gamma}}\,\nabla|\boldsymbol{B}| - q_e\boldsymbol{E}\right) \end{aligned} \tag{3.13}$$

Simulation tests have shown a very good coincidence of the trajectories of the Lorentz force case with the guiding-center approximation (in absence of the wave). However, the conservation of energy and momentum (for the noncollisional case) are now much improved in the guiding-center case. In terms of the same integration step-size, the guiding-center code runs twice as fast as the Lorentz force one, with the order of the error being almost ten times smaller. Thus, given its high accuracy, much larger steps can be used with the guiding-center approach, reducing significantly the computational cost.

### 3.3.2 Collisions

In all the simulations, we take collisions into account, since the collision time is smaller than the total integration time. For simplicity, we consider the electrons to collide only with the background ions, and we apply the pitch angle scattering process as described in [46]. This implies that we solve an additional stochastic differential equation for the pitch angle $\alpha$ (defined as the angle between the particle's velocity and the magnetic field),

$$d\alpha = F_\alpha dt + D_\alpha W_t \tag{3.14}$$

with $D_\alpha = \sqrt{(\Gamma^{e/i}/2v^3)2dt}$, $F_\alpha = -(\Gamma^{e/i}/2v^3)(v_\parallel/v_\perp)$ and $W_t$ an independent and identically distributed Gaussian random variable with mean value zero and variance one. Here, $\Gamma^{a/b} = n_b q_a^2 q_b^2 \ln\Lambda^{a/b}/(4\pi\epsilon_0^2 m_a^2)$, where $q_a$, $m_a$ and $n_a$ stand for electric charge, mass and density, respectively, of species $a$ and $\ln\Lambda^{a/b}$ is the Coulomb logarithm (see e.g. [65]).

Equation (3.14) is solved with a finite time-step $dt$ (i.e. by using the Euler-Maruyama method), which is a small fraction of the thermal collision time $t_{\rm th}$ (given as $t_{\rm th}^{-1} = \Gamma^{e/i}/v_{\rm th}^3$, with $v_{\rm th} = \sqrt{k_{\rm B}T_e/m_e}$ the thermal velocity of the electrons). After every time-step $dt$, the integration of the equations of motion is interrupted, and the velocity components are pitch angle scattered, i.e. rotated with the angle $d\alpha$ given by (3.14), according to

$$\begin{aligned} v_{\parallel,i+1} &= v_{\parallel,i}\cos(d\alpha) - v_{\perp,i}\sin(d\alpha) \\ v_{\perp,i+1} &= v_{\parallel,i}\sin(d\alpha) + v_{\perp,i}\cos(d\alpha), \end{aligned} \tag{3.15}$$

after which the integration of the equations of motion is continued. For the parameters given in Table 3.1, the thermal collision time amounts to 0.2 msec.

### 3.3.3 Wave-particle interaction

The test particles interact strongly with the wave, when the following resonance condition holds :

$$\omega - k_{\parallel} v_{\parallel} - \frac{l\omega_g}{\gamma} = 0 \tag{3.16}$$

where $l$ is the cyclotron harmonic and $k_{\parallel}$ the component of the wave vector parallel to the magnetic field.

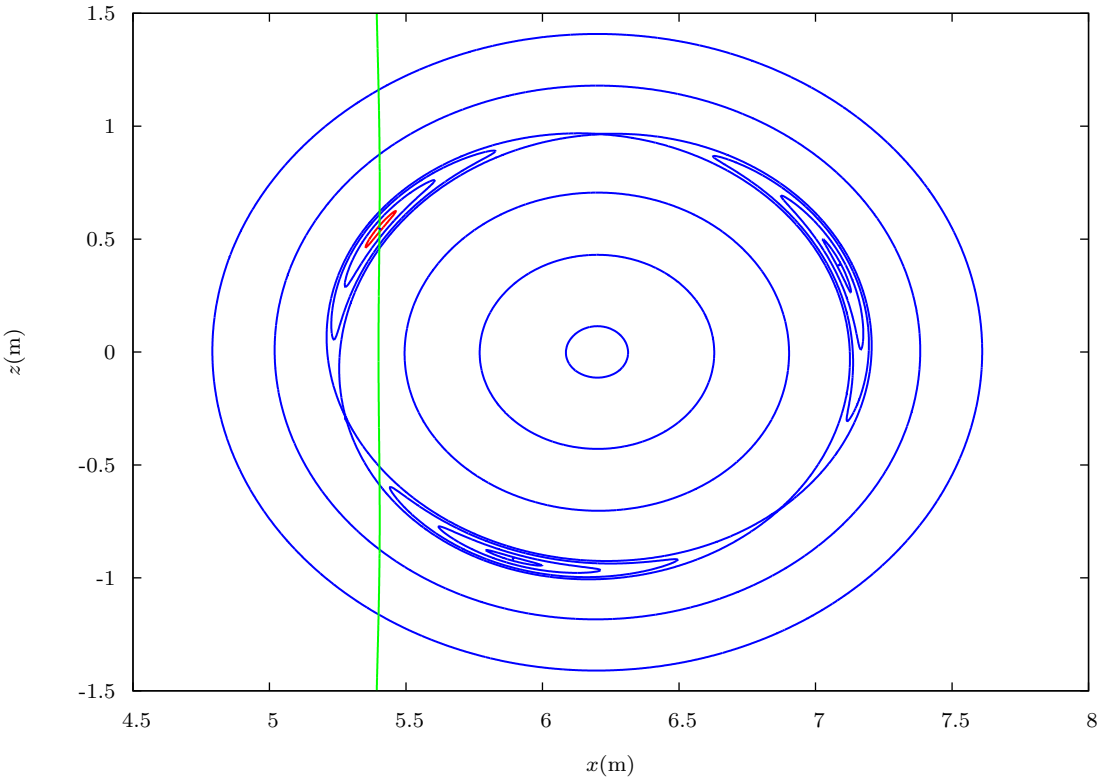


Figure 3.1: The island topology for the (3,2) mode suitably rotated for EC resonance. The red curve indicates the initial flux surface, while the green one the resonance region.

For the 1st harmonic, i.e. $l = 1$, using $v_{\parallel} = v_{\rm th}/\sqrt{2}$ and substituting in the gyrofrequency $\omega_g$ the magnitude of the magnetic field for $r = r_{\rm o}$ and $\phi = \pi/2$, we find the corresponding poloidal angle $\theta_{\rm res}$ that lies in the resonance regions. In order to investigate the interaction inside the islands, we rotate the magnetic field lines in the poloidal direction, as explained in chapter 2, such as, one of the o-points, say the first one, is placed at exactly $\theta_{\rm o} = \theta_{\rm res}$. For $\boldsymbol{x}_{\rm o}$ (see equation (3.3)) defined by the coordinates $(r_{\rm o}, \theta_{\rm res}, \pi/2)$, we also have the beam passing through that o-point. Setting the o-point in the resonance region, we follow orbits of electrons starting from

that o-point area, i.e. the initial positions are randomly picked from a specific flux surface, very close to the o-point, using the $h$-label introduced earlier. Figure 3.1 illustrates the situation in the poloidal cross section of the magnetic field lines at $y = 0$ for electrons with initial conditions on the flux surface $h_0 = -0.94$.

The guiding-center approximation is only valid when the high frequency electromagnetic wave vanishes, i.e. outside the beam. Thus, when a particle is approaching the wave beam region, the code switches from the guiding-center equations and a very large time step to the Lorentz force law and a much smaller time step, and vice versa. The very small Larmor radius of the electrons along with their high velocities allows one to do so, since the first order terms of the gyro-averaging transformation are six scales down from the zero order ones.

In our study, we use a reference case of ITER-like parameters that is summarized in Table 3.1. Our main results are deduced for the values listed, unless stated otherwise.

Before turning to statistical results, we focus on the dynamics of individual particles in their interaction with the electromagnetic wave. First, we consider very short integration times, of the order of about one collision time, in order to capture in more detail the wave-particle interaction. Every time the particles pass by the wave beam region, a "kick" in their energy appears, as can be seen in figure 3.2. Figure 3.3 shows the magnitude of the electric field that the same particle experiences, and of course the peaks are located at the exact same times at which the energy of the particle changes. The electric field zoomed into one of its peaks areas, i.e. a single beam crossing, is shown in figure 3.5, reflecting the Gaussian cross-section of the beam and oscillating due to the elliptical polarization of the wave (the oscillations actually are not resolved in the figure). Figure 3.4 presents how the energy changes during one single beam crossing, obviously oscillating in accordance with the electric field, having though a net energy gain on exit of the beam.

| | | | |
|---|---|---|---|
| $B_0$ | 5.51 T | $R_0$ | 6.2 m |
| $r_0$ | 1.9 m | $b$ | 12 |
| $\epsilon$ | 0.005 | $h_0$ | $-0.94$ |
| $\theta_l$ | 30° | $\phi_l$ | 80° |
| $\omega/2\pi$ | 170 GHz | $s$ | 3 cm |
| $P_{\max}$ | 10 MW | $T$ | 8 keV |
| $t$ | 1.5 msec | particles | 96000 |

Table 3.1: Typical parameters used in the test-particle simulations.

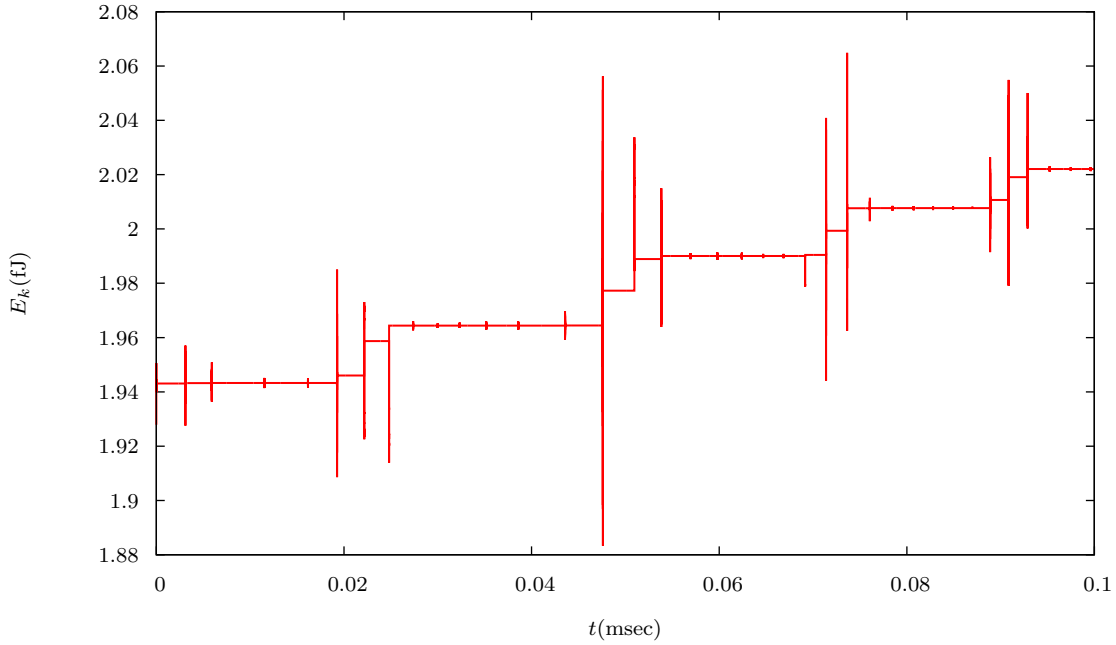


Figure 3.2: Kinetic energy of a typical particle as a function of time, for a short integration time.

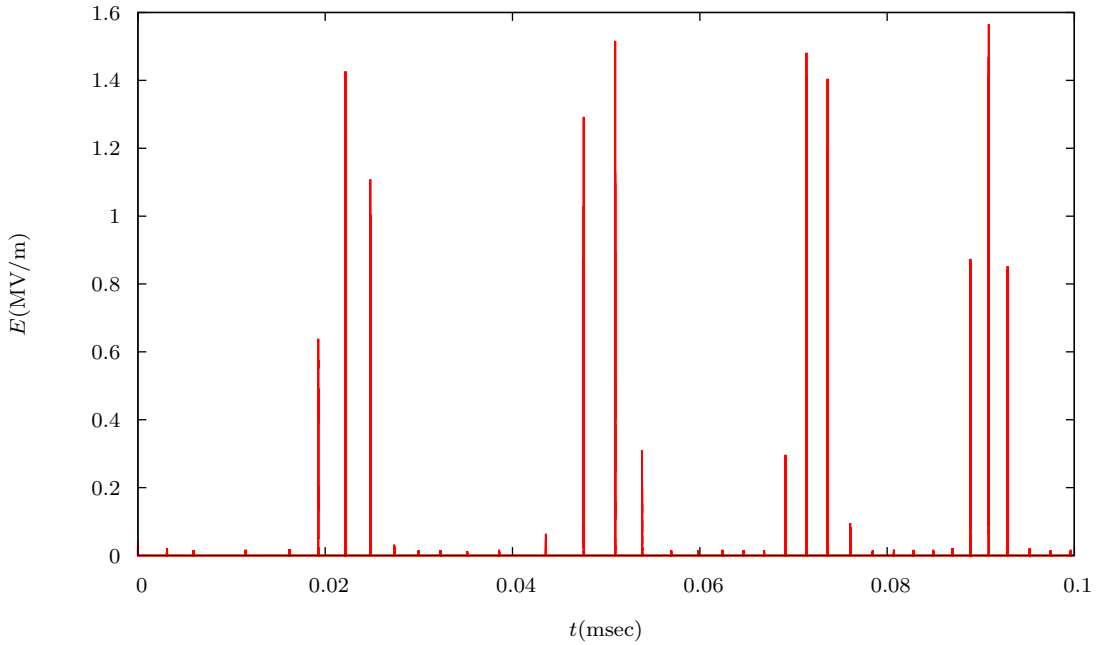


Figure 3.3: Electric wave field amplitude, as experienced by a typical particle, as a function of time, for a short integration time.

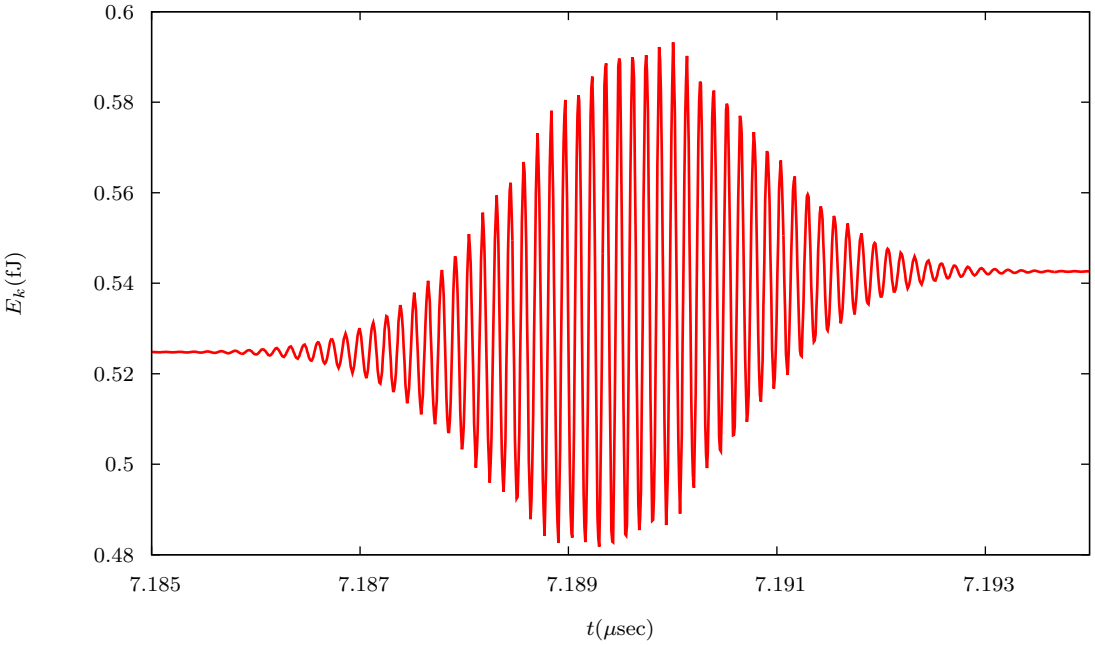


Figure 3.4: The kinetic energy oscillations during the wave-particle interaction.

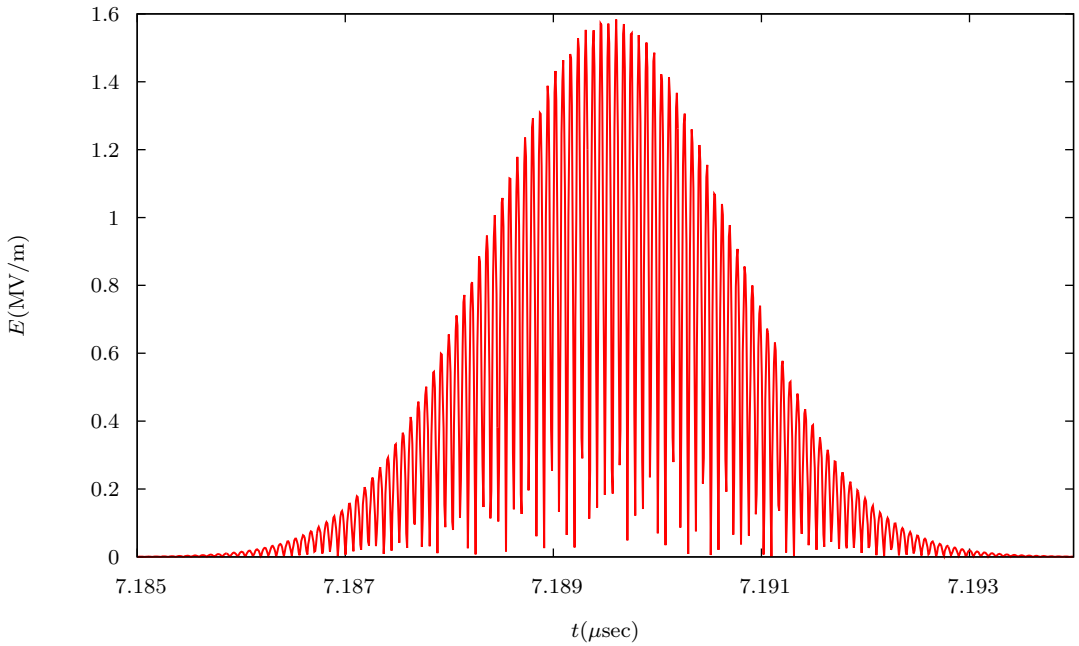


Figure 3.5: The Gaussian form of the electric wave field amplitude.

Turning now to integration times that are clearly larger than the collisions time, we show the time evolution of the kinetic energy of three different particles in Figure 3.6 (the integration

time used here corresponds to about 8 collision times). At various time-intervals many particles increase their energy almost systematically after several interactions with the electromagnetic wave. Collisions, however, can alter qualitatively the energy behavior, leading to a more complex picture. Overall, the kinetic energy of each particle seems to vary in a rather random-walk like manner. Only few particles keep on gaining energy almost constantly during extended phases, like the blue curve in Figure 3.6. Most of them undergo much bigger and more frequent changes like the other two curves. Some may even lose energy, as the red one clearly demonstrates.

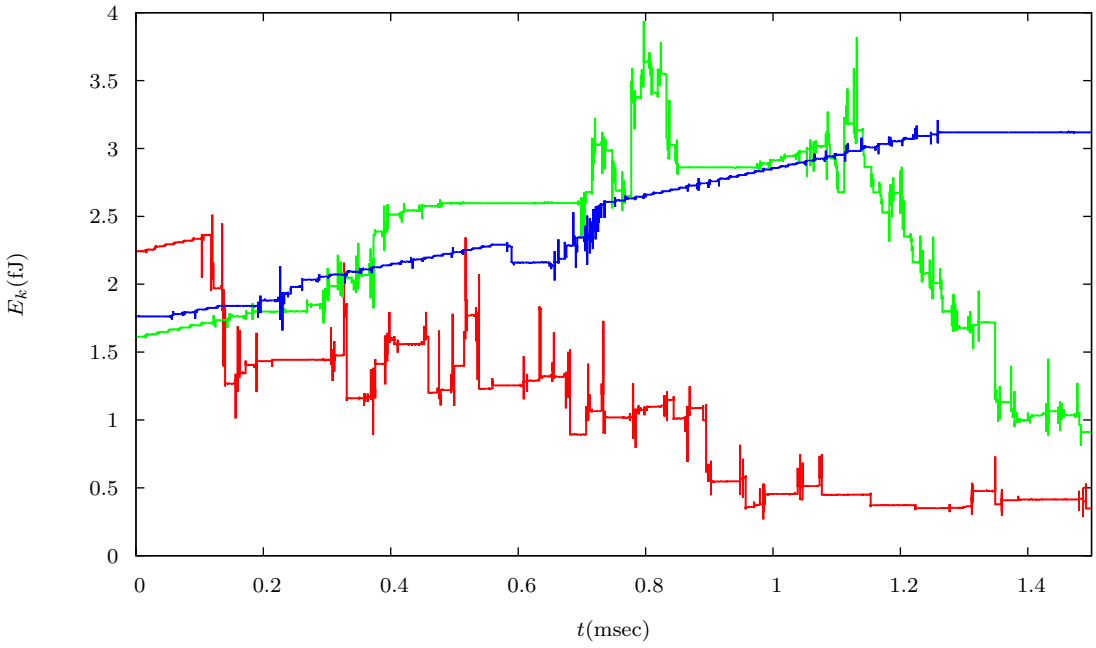


Figure 3.6: Time evolution of the kinetic energy for three different electrons, and for a relatively large final time.

### 3.3.4 Statistics

In order to test the single-particle model against other approaches to the wave-particle interaction, the particle simulations are oriented for obtaining macroscopic results. Reliable statistics require large numbers of particles to be tracked. For this purpose, the code has been parallelised over the individual particles, both with OpenMP (for multiprocessors with shared memory) and MPI (for computer systems with distributed memory). Since the particles move independently of each other, the simulations are highly scalable and profit a lot from massively parallel execution.

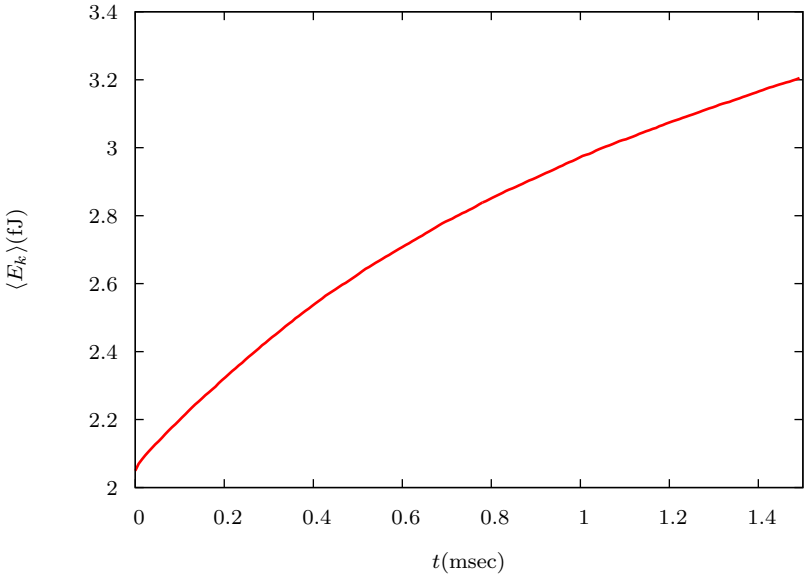


Figure 3.7: Mean kinetic energy of $10^5$ electrons as a function of time.

Figure 3.7 shows the mean kinetic energy of $10^5$ electrons as a function of time. The mean energy increases, so there is a clear gain of energy on the average. The figure also implies that no asymptotic state has been reached yet by the particles at the largest integration time considered here (longer integration times could not have been achieved, due to restrictions in computational power). In figure 3.8, the final velocity distribution as a function of the parallel and perpendicular velocity is shown. It exhibits a plateau-like depression of the distribution at intermediate velocities around the peak, as it is characteristic for wave-particle interaction and very reminiscent of Landau damping.

## 3.4 Current drive and wave absorption

### 3.4.1 Current drive

The driven current, meaning the plasma current induced by the wave in the direction parallel to the magnetic field, can be calculated by the velocities of the test-particles. At a fixed set of time-instances during the simulations, we keep track of the parallel and perpendicular velocities $v_\parallel$ and $v_\perp$ of each electron, as well as the flux surface $h$ it lies upon. For each time-instance, the particles are separated into different bins labeled by $h$ of sufficiently large width $\delta h$, starting from the separatrix $h_\mathrm{x}$ until we reach the o-point $h_\mathrm{o}$. The optimal choice for $\delta h$ should be a

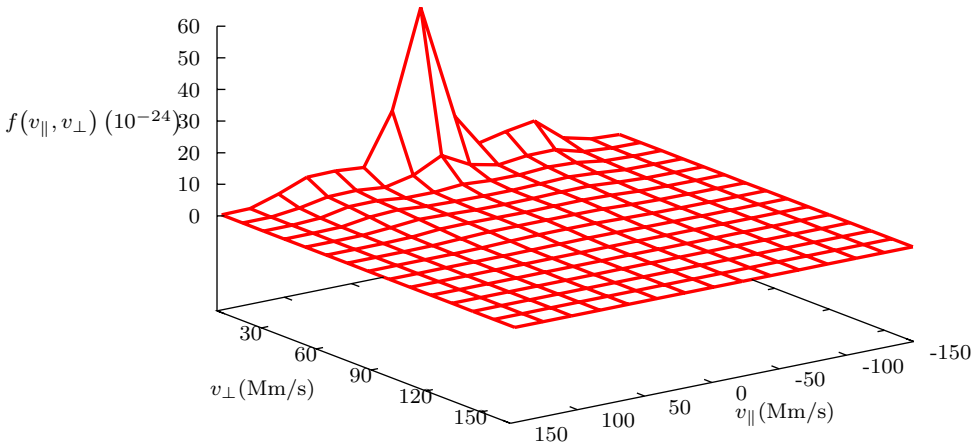


Figure 3.8: Final velocity distribution of $10^5$ electrons at $t = 1.5\,\mathrm{msec}$.

value large enough to include a statistically significant number of particles in every bin (2.5% of the particles being the minimum), but also small enough for good resolution in $h$. Calculating the average parallel velocity $\langle v_\parallel \rangle (h)$ over the particles lying between $h$ and $h + \delta h$, we can find the total current density

$$J_\parallel(h) = n_e q_e \langle v_\parallel \rangle (h) = \frac{n_e q_e}{N_h} \sum_{i=1}^{N_h} v_{\parallel,i}(h), \tag{3.17}$$

generated at the flux surface $h$, where $n_e$ is the electrons density and $N_h$ is the number of particles in the interval $[h, h + \delta h]$. In this way, the driven current density is obtained as a function of the flux surface label $h$.

In the next two figures, the current drive density for the typical case of Table 3.1 is presented in red. As we can see, the current density shows a negative peak around $-220\,\mathrm{MA/m^2}$ at $h = -0.97$ approximately. The deviation from the initial flux surface is actually quite small and consistent with the particle drifts towards the o-point. In Figure 3.9, a comparison is made between the reference case with initial conditions on $h_0 = -0.94$ and a case of a different initial flux surface, namely $h_0 = -0.7$, indicated in blue. The latter looks like a natural continuation of the red curve and, unlike the former, shows no peak at all, as expected since the particles are further away from the resonance region. In Figure 3.10, we see the reference case compared to

two cases with different toroidal launching angles $\phi_l = 100°$ and $\phi_l = 120°$ represented by the blue and green curves, respectively. The current density is much reduced at these new launching angles, and due to the noise-level, which is of the order of $\pm 10\,\mathrm{MA/m^2}$, it cannot be considered to be significantly different from zero.

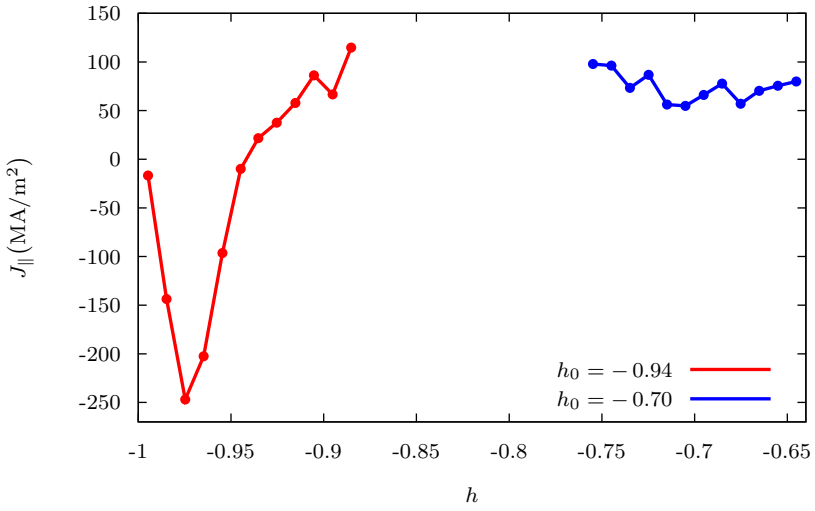


Figure 3.9: Current drive density as a function of $h$ at final time $t = 1.5\,\mathrm{msec}$ for two different initial flux surfaces.

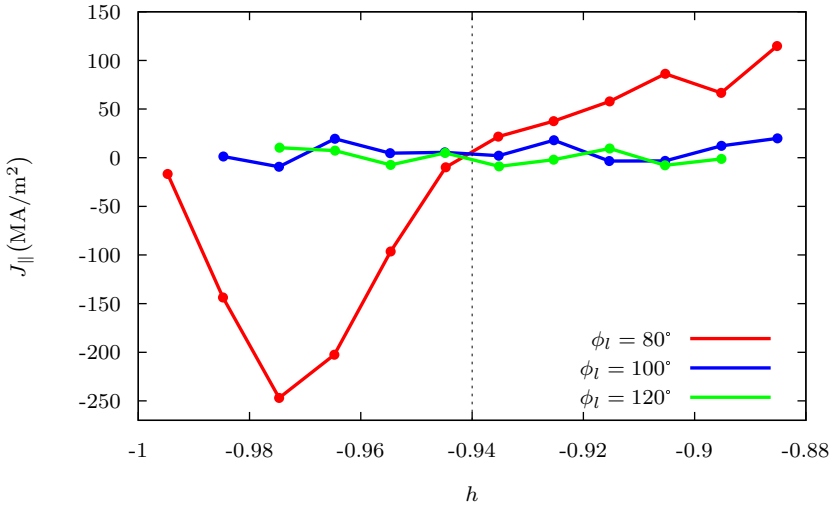


Figure 3.10: Current drive density as a function of $h$ at final time $t = 1.5\,\mathrm{msec}$ for three different toroidal launching angles. The dashed line corresponds to the initial conditions of the particles.

### 3.4.2 Wave absorption

The wave power density absorbed by the test-particles can be determined in a very similar manner, by computing the variation in the particles kinetic energy over time. First, the kinetic energy $E$ is determined from the velocities and then the difference $\delta E$ for each particle during the time interval $\delta t$ between two subsequent time-instances. Following the same binning in $h$ to the particles as before, the average power $\langle \delta E(h) \rangle / \delta t$ over the particles lying between $h$ and $h + \delta h$ yields the absorbed power density as a function of the flux surface label $h$,

$$P(h) = n_e \frac{\langle \delta E(h) \rangle}{\delta t} = \frac{n_e}{N_h \delta t} \sum_{i=1}^{N_h} \delta E_i(h) \tag{3.18}$$

The next two figures present the results for the absorbed power density corresponding to the current densities shown in Figures 3.9 and 3.10, respectively. As before, the red curve in both plots describes our reference case of Table 3.1, which shows a peak around $104\,\mathrm{MW/m^3}$ at $h = -0.97$ again, where the peak of the current drive lies. The same drift towards the o-point is also manifest here. As we can see in Figure 3.11 showing the comparison of initial conditions, the particles released at $h_0 = -0.7$ being too far from the resonance region, they basically carry zero power density. Figure 3.12 shows that the power density for the toloidal launching angles $\phi_l = 100°$ and $\phi_l = 120°$ is not significantly different from zero within the noise level, which here is of the order of $\pm 2\,\mathrm{MW/m^3}$.

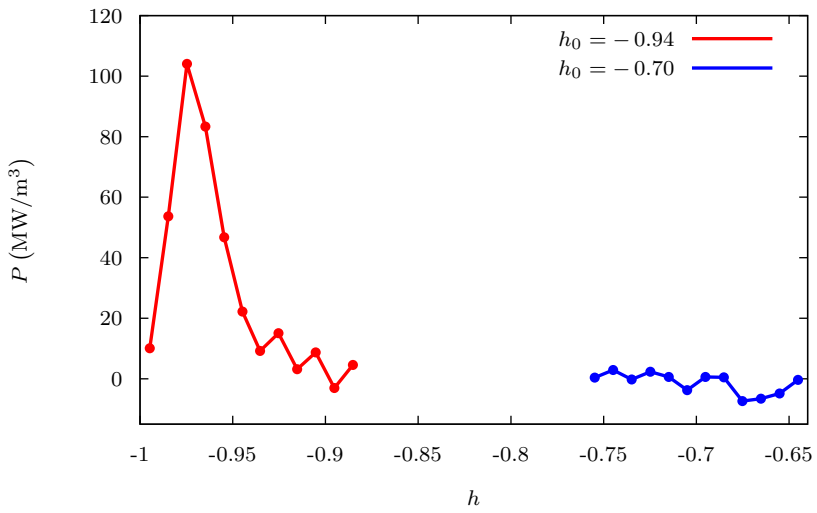


Figure 3.11: Absorbed power density as a function of $h$ at final time $t = 1.5\,\mathrm{msec}$ for two different initial flux surfaces.

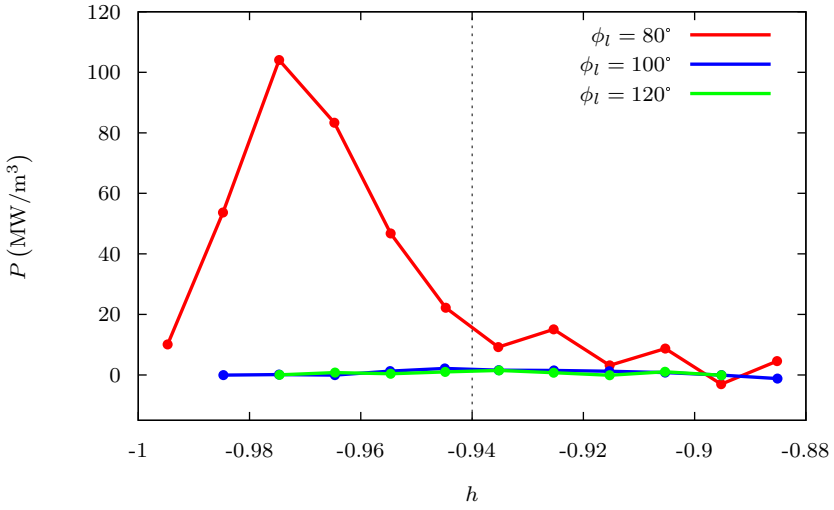


Figure 3.12: Absorbed power density as a function of $h$ at final time $t = 1.5\,\mathrm{msec}$ for different toroidal launching angles. The dashed line corresponds to the initial conditions of the particles.

## 3.5 Comparison with the linear theory

For comparison with the linear theory, we use the ray-tracing code described in [56]. In this code, the Hamiltonian equations of ray-propagation are based on the cold plasma dispersion relation, the wave absorption along the ray path is calculated in terms of the imaginary part of the wave vector, as determined from the weakly relativistic, hot plasma dispersion relation, and the current drive efficiency is calculated with the linear adjoint method, and it includes the effects of trapped particles, ion-electron collisions and the poloidal variation of the collision operator. Exactly the same magnetic background topology and perturbation as for the test-particles are used in the ray tracing code, and the calculation of the volumes enclosed by flux surfaces within the island is done by means of an analytical flux surface labeling that is quantitatively almost coincident with the one used for the test-particles.

In order to have a meaningful comparison, in the sense that the nonlinear and the linear approach explore the same resonance region, the island had been rotated such that its o-point lies close to the region where the ray is resonant, and the initial conditions of the test-particles have been chosen in this region. Also, for the ray to pass by the reference point of the beam in the particle treatment and for the beam's and the ray's wave-vector to be parallel, we applied back-tracing of the ray from the beam's reference point to the plasma edge, to determine the appropriate initial conditions for the actual ray tracing.

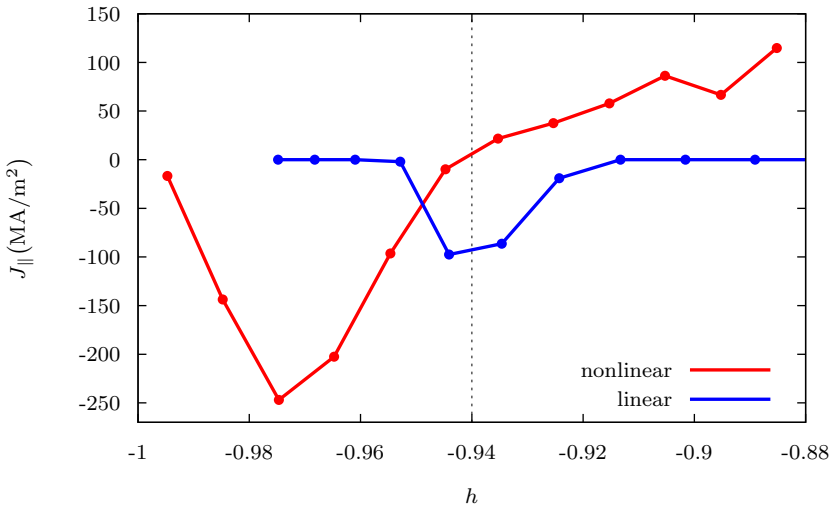


Figure 3.13: Comparison of current drive density between test particle simulations at final time $t = 1.5\,\mathrm{msec}$ (red curve) and ray-tracing (blue curve).

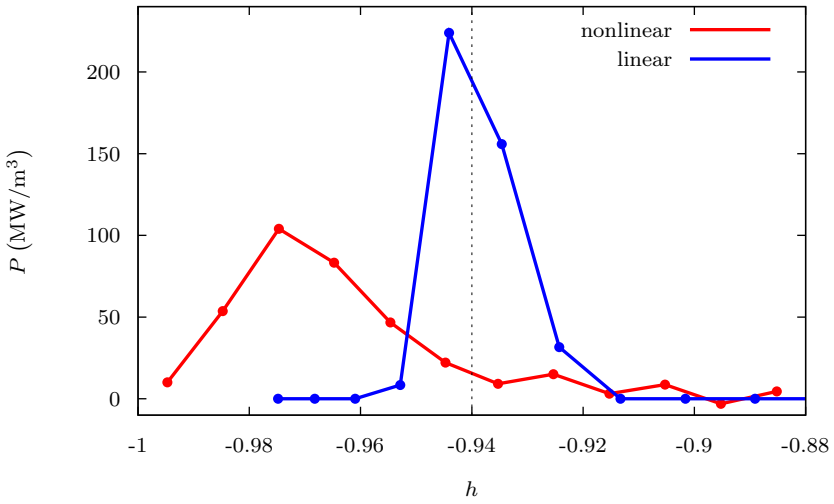


Figure 3.14: Comparison of absorbed power density between test particle simulations at final time $t = 1.5\,\mathrm{msec}$ (red curve) and ray-tracing (blue curve).

In Figures 3.13 and 3.14, we show the results from the particle simulations and the nonlinear character of the equations of motion compared with the ray tracing method and the linear theory, using the same initial wave power $P_{\max} = 10\,\mathrm{MW}$ and toroidal launching angle $\phi_l = 80°$. The current density from the linear code is roughly 2.5 times smaller than the one derived from

the nonlinear simulations, see Figure 3.13. In addition, and as described above, the peak in the nonlinear case is slightly offset towards the o-point from the linear one, which quite well coincides with the flux surface where the particles have initially been released. As mentioned, this difference can be explained by particle drifts, which the ray tracing code cannot take into account. On the contrary to the current density, the linearly determined absorbed power density is roughly 2 times larger than the one from the nonlinear simulations, as shown in figure 3.14, while the same nonlinear drift effect of the peak as for the current density is present.

Figures 3.15 and 3.16 are the counterparts of figures 3.10 and 3.12, showing the results from the ray-tracing code for different toroidal launching angles. Also here, current drive density and absorbed power density are much reduced for $\phi_l = 100°$ and $120°$. The current density is close to zero for both new launching angles, and thus the results are compatible with the nonlinearly derived ones, within the error-level of the latter. The same holds for the power density at $\phi_l = 120°$, which is close to zero both linearly and nonlinearly, only for $\phi_l = 100°$ there is a systematic difference, with a linear peak value of $50\,\mathrm{MW/m^3}$, which lies clearly above the noise level around 0 of the nonlinear simulations.

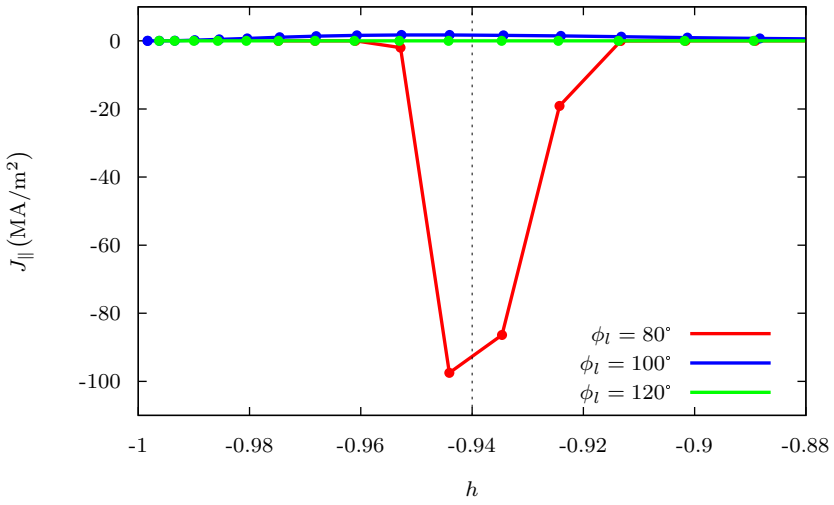


Figure 3.15: Current drive density as a function of $h$, for three different toroidal launching angles, calculated with the ray-tracing code.

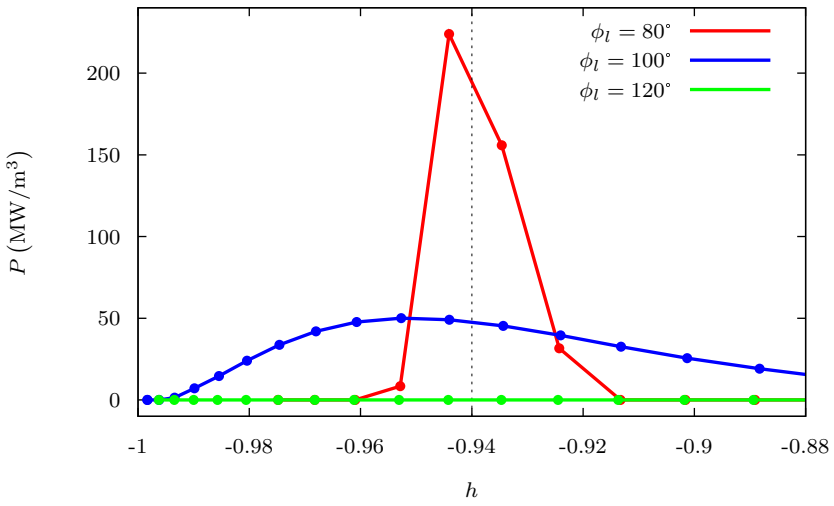


Figure 3.16: Absorbed power density as a function of $h$, for three different toroidal launching angles, calculated with the ray-tracing code.

## 3.6 Discussion

We have performed test particle simulations of the wave-particle interaction in ECRH within a magnetic island of an NTM perturbed equilibrium topology. Our code is simply based on direct numerical integration of the equations of motion for charged particles, suitably switching from the Lorentz force law to the guiding-center approximation and back. Both the perturbed magnetic field and the electromagnetic wave are analytically prescribed to the system. For the former, we also have at our disposal an analytical flux surface label (presented in the previous chapter) to determine in full precision any information required along the way for the magnetic islands. On the other hand the wave for ECRH is modeled by a narrow beam described (i.e. its wavenumber and polarization) in terms of the hot plasma theory. Electron-ion collisions are also implemented using a pitch angle scattering procedure outside the equations of motion. Initial positions for the particles can be randomly chosen on a given flux surface, while initial velocities are also random, following a Maxwellian distribution. Finally, parallel processing allows detecting large numbers of particles in order to investigate their collective behavior and arrive at statistical conclusions.

This work is still in progress. Although the code has rather reached a mature level and some first results have been produced, conclusions at this stage would be too hasty to draw. First of all, as previously commented in terms of Figure 3.7, we need to be able to extend our integration

times in order to reach some kind of asymptotic state. This is purely a matter of computational power, but very crucial for reliable validation and comparison against other codes or theories, which in their majority make predictions for stationary states. Once this obstacle is overcome, our tools will be able to target more focused problems, e.g. make more precise the effects coming from the magnetic islands or ECRH, either separetely or combined together. Other necessary improvements, which are now under development and will soon be completed, are the inclusion of the wave magnetic field and the electron-electron collisions. Both of them could alter significantly the energy of the particles : the latter directly affecting the amount of energy, while the former changing the way this energy is distributed over the parallel and perpendicular velocities. After completing the aforementioned tasks the results presented in this chapter will be submitted for publication [61].

In summary, we have developed what is hopefully a useful tool for studying numerically the charged particle motion. And even though suitably staged here for the magnetic island topology and ECRH, it could be employed elsewhere. Last but not least, although aiming towards results of more statistical nature, the code is, as one can imagine, equally suited for investigating individual orbits, i.e. exploring the motion of charged particles more like a dynamical system, and of course without having the aforementioned computational restrictions. For example, identifying banana orbits of trapped particles, or plotting Poincare sections that show the dynamics of the particles for longer times are all features included in the code and ready to be used for inspection.

# Part II

# Symmetry Analysis of Charged Particle Motion

# Chapter 4

# Symmetry Methods and Group Theory of Differential Equations

Simply consider the first-order ordinary differential equation

$$\frac{dx}{dt} = f(t, x) \tag{4.1}$$

where $t, x \in \mathbb{R}$ and $f$ a differentiable function. If $f$ does not depend on the variable $x$, then this equation can be easily solved by integration

$$x(t) = \int f(t)\, dt$$

Thus, in the general case, it is to our advantage to find new variables $\widetilde{t}$ and $\widetilde{x}$ which transform the above equation (4.1) into the simpler form

$$\frac{d\widetilde{x}}{d\widetilde{t}} = \widetilde{f}(\widetilde{t}) \tag{4.2}$$

For then from the solution $\widetilde{x}(\widetilde{t})$ of (4.2) we can return back to the original variables $t$ and $x$ and recover the solution $x(t)$ of (4.1).

**Example 4.1.** Take for instance the differential equation

$$\frac{dx}{dt} = tx$$

Note that the transformation $x \longrightarrow -x$ leaves the equation invariant. So without loss of generality we may assume $x > 0$ and replace the variable $x$ with $\widetilde{x} = \ln x$. Then the above equation transforms into

$$\frac{d\widetilde{x}}{dt} = t$$

which can be solved with direct integration, yielding

$$\widetilde{x}(t) = \frac{t^2}{2} + \widetilde{c}$$

and returning back to the original variable $x$, we arrive at the general solution

$$x(t) = c\, e^{\frac{t^2}{2}}$$

Of course the equation in this example can be solved using separation of variables and then integration. But all cases are not so simple. Let us see another example.

**Example 4.2.** In general, the Riccati equation

$$\frac{dx}{dt} = \frac{x+1}{t} + \frac{x^2}{t^3}$$

can be solved by finding a special solution $x_1(t)$. The transformation $x = x_1 + y^{-1}$ leads then to a linear first-order differential equation for $y(t)$, whose solution yields the general solution $x(t) = x_1(t) + 1/y(t)$.

Alternatively if we introduce new variables

$$\widetilde{t} = x/t$$

$$\widetilde{x} = -1/t$$

then we get the differential equation

$$\frac{d\widetilde{x}}{d\widetilde{t}} = \frac{1}{1+\widetilde{t}^{\,2}}$$

Integrating it, we find

$$\widetilde{x}\left(\widetilde{t}\right) = \arctan \widetilde{t} + c$$

and going back to the original varibales

$$x(t) = -\,t \tan\left(\frac{1}{t} + c\right)$$

As we can see we deduced the solution of the equation in a different way, which looking back one could say that looks simpler.

In the treatment of ordinary differential equations one often encounters various solving techniques, most of which are based somewhat on a trick which brings the equation into a much simpler, known form. These techniques, which seem unrelated to each other, apply only in certain equations, leaving the majority outside. Actually they are just a small part of a theory that can be applied to every differential equation of any order. At its core lies the concept of symmetry, and (for continuous symmetries only considered here) it is almost entirely based on the work of Sophus Lie. The transformations we used in the previous examples to integrate the differential equation in each case have not been found by accident, but through this theory.

## 4.1 Systems of differential equations

In order to introduce the notion of symmetry, first we need to describe differential equations on a more geometrical ground. Our goal is to keep our presentation as simple as possible and briefly go through the key concepts of symmetry theory. Therefore, in what follows we mainly consider Euclidean spaces, besides a small deviation in section 4.2, where there is no need to. A detailed and at the same time concrete account for the technicalities arising when working on manifolds is given in [89], section 3.5.

Although we are primarily concerned with ordinary differential equations (ODEs), symmetry methods do not distinguish ordinary from partial differential equations (PDEs). In our presentation therefore symmetry theory is introduced in terms of the more general case of PDEs, while demonstrated mostly with examples of ODEs. The multi independent variable case is all the same needed, when we move to the notion of equivalence transformations, as well. For avoiding confusion, we stress from the beginning the different notation between ODEs and PDEs, based on physical arguments: In ODEs, which usually decribe the time evolution of the spatial position for a system in general, as the problem studied in this thesis, the dependent variables are repsresented collectively by the space coordinates $x$ and the independent one by the time $t$, while in PDEs usually describing the evolution of a physical quantity both in space and time, dependent variables are denoted by $u$ and independent ones by $x$.

So, consider two Euclidean spaces $X = \mathbb{R}^q$ and $U = \mathbb{R}^p$ together with a smooth mapping $f : X \longrightarrow U$, whose coordinates at any points $x$ and $u$, respectively, express the independent variables

$$x = \left(x^1, x^2, \ldots, x^q\right)$$

of the problem at hand, and the dependent ones

$$u = f(x) = \left(u^1, u^2, \ldots, u^p\right)$$

Following the standard notation often employed in PDEs, where subscripts designate partial diferentiation, the $k$-th order partial derivative of the component $f^i(x)$ with respect to the variables $x^{j_1}, \ldots, x^{j_k}$ will be denoted as

$$u^i_J = \frac{\partial^k f^i}{\partial x^{j_1} \partial x^{j_2} \cdots \partial x^{j_k}}$$

Let $U_k$ represent the set of all these $k$-th order derivatives of all the components of $u$ for every choice of $J = (j_1, \ldots, j_k)$, whose dimension is $\dim U_k = p \binom{q+k-1}{k}$. Finally, $u^{(n)}$ stands for all

the derivatives of $u$ up to order $n$ and $U^{(n)} = U_1 \times \cdots \times U_n$ their set, which is isomorphic to the $p_n$-dimensional Euclidean space, where $p_n = \sum_{k=1}^{n} \dim U_k$ is the dimension of $U^{(n)}$.

Now, the first step to picture a differential equation is to glue together the independent and dependent variables by considering the space $M = X \times U$ with coordinates $(x, u)$. Then in order to include derivatives, the second one is an extension given by the space of independent and dependent variables now along with all the derivatives up to $n$-th order,

$$M^{(n)} = X \times U \times U^{(n)}$$

with coordinates $\left(x, u, u^{(n)}\right)$. The space $M^{(n)}$ is called the *jet space* of $n$-th order of the *base space* $M$, while smooth, real functions defined on $M^{(n)}$ are called *differential functions*. If $m = q+p$ is the dimension of $M$, then $m_n = q+p+p_n = q+p\binom{q+n}{n}$ denotes the dimension of $M^{(n)}$. Notice that $x$ and $u$ in terms of $M$ and, in general, $x$ and $u^{(n)}$ in $M^{(n)}$ are treated as coordinates on equal footing, independent of each other with no reference to $f$ anymore.

To recover the relation between $x$ and $u$, we may interprete a function $u = f(x)$ defined on a subset $S \subset X$ geometrically. And that is by identifying $f$ with its graph,

$$\Gamma_f = \{(x, u) : x \in S, u = f(x)\},$$

consisting of all points $(x, f(x))$ in $M$, for which $f$ is defined. The subset $\Gamma_f$ of $M$ can be extended to a subset $\Gamma_f^{(n)}$ of the $n$-jet space $M^{(n)}$ given by the graph of $f^{(n)}$,

$$\Gamma_f^{(n)} = \left\{(x, u, u^{(n)}) : x \in S, u = f(x), u^{(n)} = f^{(n)}(x)\right\}$$

Let $\Delta = (\Delta_1, \ldots, \Delta_l)$ be a vector function on $M^{(n)}$, where $\Delta_\nu$ for $\nu = 1, \ldots, l$ are differential functions. A *system of differential equations* of order $n$ is given by a set of equations

$$\Delta\left(x, u, u^{(n)}\right) = 0 \tag{4.3}$$

and can be viewed as the vanishing of the smooth map $\Delta : M^{(n)} \longrightarrow \mathbb{R}^l$. In this context, (4.3) defines a subset of the $n$-jet space $M^{(n)}$

$$D = \left\{(x, u, u^{(n)}) : \Delta\left(x, u, u^{(n)}\right) = 0\right\}$$

A smooth function $u = f(x)$ defined on the subset $S \subset X$ is called a *solution* of the system, if it satisfies equations (4.3) for any $x \in S$, meaning

$$x \in S \;\Rightarrow\; \Gamma_f^{(n)} \subset D \tag{4.4}$$

**Remark 4.3.** Several of the above geometrical notions can be relaxed. For example a given problem may not be defined over all of $M$, but only on an open subset thereof and $\Delta$ accordingly on an open subset of $M^{(n)}$. Further generalizations would be the case where $X$ and $U$ are in general differentiable manifolds, in which case $M^{(n)}$ is a jet bundle (see [95] for the geometry of jets).

So far, our discussion has been kept quite general. We should, however, impose some rather mild restrictions on $\Delta$ to exclude any degeneracies. We comment first that, in the above fashion, a system of differential equations can be viewed to some extent as a set of algebraic equations (in the sense of constraints) among the variables $z = \left(z^1, \dots, z^{m_n}\right)$, where $z = \left(x, u, u^{(n)}\right)$ is a point in $M^{(n)}$, meaning $z^\mu$ for $\mu = 1, \dots, m_n$ is any of the variables $x^j$, $u^i$ or $u^i_J$. Of course algebraic relations strictly among the independent variables do not qualify as differential equations and therefore are not allowed. In this algebraic viewpoint it is only natural to rule out any functional dependencies among the equations themselves. Therefore, for every system $\Delta(z) = 0$ we consider that the rank of the $l \times m_n$ Jacobian matrix

$$\mathrm{J}_\Delta = \left(\frac{\partial \Delta_\nu}{\partial z^\mu}\right)$$

equals $l$ when $\Delta = 0$, in which case the system is called of *maximal rank*. This is not much of a restriction and is often met in practice as the next example illustrates.

**Example 4.4.** Obviously there is no point considering the scalar PDE $(u_t - u_{xx})^2 = 0$ or the system $\{u_t - u_{xx} = 0, 3u_t - 3u_{xx} = 0\}$, both violating the maximal rank condition, rather than the heat equation $u_t - u_{xx} = 0$ itself which respects it.

In the particular case, where the maximal rank condition is satisfied for just $u^{(n)}$ and not all of $z$, i.e. the $l \times p_n$ submatrix $\left(\partial \Delta_\nu / \partial u^i_J\right)$ is of rank $l$, then by the implicit funtion theorem $l$ of the derivatives $u^i_J$ can be expressed in terms of the rest $m_n - l$ variables $z^\mu$, meaning that the system can be written in the familiar *solved form*.

On the other hand, a noticeable difference between differential and algebraic equations is that for the latter the set of solutions and the set defined by the equations themselves would be the same thing. On the contrary, for a system of differential equations (4.3) there is no guarantee in general that the collection of all the extended graphs $\Gamma^{(n)}_f$ for every solution $u = f(x)$ will cover all of $D$ defined by the system. There may be points in $D$, through which no (derivatives of) solutions pass. Examples of this notion are often found in systems of PDEs with nontrivial integrability conditions of order less or equal to $n$ that cannot be satisfied as an algebraic consequence of the equations of the system.

**Example 4.5.** For instance, the system $\{u_t = 0, u_x = t(u-1)\}$ defines a subset $D$ in the 1-jet space consisting of all the points of the form $(t, x, u, 0, t(u-1))$. However, the compatability condition $u_{tx} = u_{xt}$ results in the equation $0 = u - 1$ of order less than $n = 1$, i.e. an algebraic relation, that is new to the system. Therefore, the only solution of the system is $u(t, x) = 1$ and the extended set $\Gamma_1^{(1)}$ of solutions is given by the points $(t, x, 1, 0, 0)$. Obviously there are points in $D$, such as $(0, 0, 0, 0, 0)$, which do not correspond to solutions of the system.

A system $\Delta\left(x, u, u^{(n)}\right) = 0$ for which every point in $D$ corresponds to the graph $\Gamma_f^{(n)}$ of the derivative of a solution $u = f(x)$ is called *locally solvable*. Nonsingular ODEs and evolution equations in general or scalar PDEs are all locally solvable.

Technical the above features may sound, they will prove necessary for making precise the symmetry condition later on. So, in the following we assume that a system of differential equations is *nondegenerate*, that is both of maximal rank and locally solvable.

Let us also make a little notation index rule to avoid exhausting repetition. In this chapter, unless stated otherwise, the indices $i, j, k, \mu$ and $\nu$ take the following values: $i = 1, \dots, p$, $j = 1, \dots, q$ (with or without subscript), $k = 1, \dots, n$, $\nu = 1, \dots, l$ and $\mu = 1, \dots, m_n$. We also adopt Einstein's summation convention, according to which a repeated index appearing in a term means summation over all values of that index. In particular, summation with respect to the multi-index $J$ is meant for all $k$-tuples for all $k$.

## 4.2 Groups of transformations

In order to describe the way a system of differential equations transforms under a change of variables, we need to investigate the key notion of a vector field and its related consequences. Playing a central role in symmetry theory, the discussion below stays close to differential equations aspects, not so much following the more abstract line of differential geometry. Starting with the formal definition of a vector field, given in appendix B limited there just for the needs of differential forms, we move on the notion of the flow. Then, based on these concepts, we quickly generalize in the second subsection of this section, selectively borrowing elements from the theory of Lie groups.

### 4.2.1 The flow of a vector field

So, recalling Definition B.2, consider a *vector field* $\mathbf{v}$ on a $q$-dimensional manifold $X$ in general, which at every point $x \in X$ assigns a *tangent* vector $\mathbf{v}_x$ of the manifold. Along with this

geometrical view, preference here will be given to a more abstract one (for finite-dimensional manifolds), according to which a vector field $\mathbf{v}$ can be equivalently defined as a derivation, acting on the space $C^\infty(X)$ of smooth, real functions on the manifold $X$ (see (B.6)). In other words it maps a smooth function to another one, satisfying linearity with respect to addition and Leibniz' rule with respect to multiplication (see the properties of Definition (B.1)). Repeating (B.7), this operator in local coordinates is expressed as

$$\mathbf{v} = \xi^i(x)\frac{\partial}{\partial x^i} \tag{4.5}$$

where $\xi^i$ are smooth functions of $x$, stating that the partial derivatives $\partial/\partial x_i$ are a basis for the set of all vector fields on $X$.

A typical example of a vector field $\mathbf{v}$ is the velocity field of a fluid, which relates a velocity $\mathbf{v}_x$ of a particular direction and magnitude to every point $x$ of the fluid. In this picture, a fluid particle starting from some arbitrary point $x_0$, follows a trajectory $x = \varphi(t)$, along which its velocity at every point is equal to the field $\mathbf{v}_{\varphi(t)}$, that is

$$\frac{d\varphi}{dt} = \mathbf{v}_{\varphi(t)} \;\Rightarrow\; \frac{d\varphi^i}{dt} = \xi^i(\varphi(t))$$

The collection of all these trajectories, each one starting from a different point $x_0$, describes the flow of the fluid.

Returning to our general discussion on vector fields (4.5), but keeping the fluid visualization in mind, consider now a curve $\varphi(\epsilon)$ on $X$ parametrized by $\epsilon$ the latter lying in some open interval $\mathrm{I} \subset \mathbb{R}$. If at any point $x = \varphi(\epsilon)$ of a smooth curve, the vector field $\mathbf{v}$ is tangent to $\varphi$, meaning $\varphi'(\epsilon) = \mathbf{v}_{\varphi(\epsilon)}$, then $\varphi$ is called the *integral curve* of $\mathbf{v}$. In terms of local coordinates, the tangency condition reads

$$\frac{dx^i}{d\epsilon} = \xi^i(x) \tag{4.6}$$

and since $\xi^i$ are smooth functions, existeness and uniqueness for the solution $\varphi(\epsilon)$ of the system with given initial conditions $\varphi(0) = x_0$ (assuming 0 belongs to I) is guaranteed, at least in a neighborhood of $x_0$, by the well-known theorems for ODEs. Two curves $\varphi(\epsilon)$ for different initial conditions cannot intersect, for then from their intersection point two solutions of the system would set off violating the uniqueness property. Therefore, only one such curve $\varphi(\epsilon)$ passes through a fixed point $x$ of the manifold, which will be denoted by $\varphi_x(\epsilon)$. Letting $x$ vary, we consider the function $\varphi(\epsilon, x) = \varphi_x(\epsilon)$.

*Flow* then of a vector field $\mathbf{v}$ on a differentiable manifold $X$ is the set of solutions $\varphi(\epsilon, x)$ of system (4.6), each of which passes through a different point $x \in X$. In other words, it is

a map $\varphi : \mathcal{U} \longrightarrow X$, where $\mathcal{U}$ is an open subset of $\mathbb{R} \times X$ such that for each $x \in X$ the set $\mathrm{I}_x = \{\epsilon \in \mathbb{R} : (\epsilon, x) \in \mathcal{U}\}$ is an open interval containing 0, which satisfies the following properties

1. $\varphi(0, x) = x$, for all $x \in X$ (4.7a)

2. $\varphi(\epsilon, \varphi(\delta, x)) = \varphi(\epsilon + \delta, x)$, for all $\delta \in \mathrm{I}_x$ and $\epsilon \in \mathrm{I}_{\varphi(\delta,x)}$ such that $\epsilon + \delta \in \mathrm{I}_x$ (4.7b)

3. $\dfrac{\partial \varphi(\epsilon, x)}{\partial \epsilon} = \mathbf{v}_{\varphi(\epsilon,x)}$, for all $\epsilon \in \mathrm{I}_x$ (4.7c)

The first property expresses the initial conditions of the system, while the second one comes from the uniqueness of solutions. The third property is really system (4.6) and states that the vector field $\mathbf{v}$ for a fixed $x$ is tangent to the curve $\varphi(\epsilon, x)$ at each point. From (4.7a) and (4.7c) we can also find a way to recover the vector field from its flow,

$$\mathbf{v}_x = \left.\frac{\partial \varphi}{\partial \epsilon}\right|_{\epsilon=0}, \tag{4.8}$$

The map $\phi$ along with the properties (4.7a) and (4.7b) defines a local group action of the Lie group $\mathbb{R}$ on the manifold $X$, also known as a *one-parameter local group of transformations*. Since $\varphi$ is smooth, then from Taylor's theorem near the point $(0, x)$ it can be expressed as

$$\varphi(\epsilon, x) = \varphi(0, x) + \epsilon \left.\frac{\partial \varphi}{\partial \epsilon}\right|_{\epsilon=0} + \frac{\epsilon^2}{2} \left.\frac{\partial^2 \varphi}{\partial \epsilon^2}\right|_{\epsilon=0} + O(\epsilon^3)$$

The linear term in the above expansion is already given by (4.8), which recalling (B.5) can also be written as $\varphi_\epsilon(0, x) = \mathbf{v}(x)$. Moving on to higher derivatives we split (4.7c) in components, $\varphi^i_\epsilon = \xi^i(\varphi(\epsilon, x))$ and differentiate with respect to $\epsilon$, $\varphi^i_{\epsilon\epsilon} = \xi^i_{\varphi^j} \varphi^j_\epsilon$. Then set again $\epsilon = 0$ and using also (4.8) we arrive at $\varphi^i_{\epsilon\epsilon}(0, x) = \xi^i_{x^j} \xi^j$ or $\varphi_{\epsilon\epsilon}(0, x) = \mathbf{v}(\xi) = \mathbf{v}(\mathbf{v}(x)) \equiv \mathbf{v}^2(x)$. Therefore, substituting in the above Taylor expansion, we deduce that a point $x$ under the flow $\varphi$ of a vector field $\mathbf{v}$ transforms into another point $\widetilde{x}$ given as

$$\widetilde{x} = \varphi(\epsilon, x) = \left(I + \epsilon\, \mathbf{v} + \frac{\epsilon^2}{2}\, \mathbf{v}^2 + \cdots\right) x.$$

where $I$ is the identity operator. Successively differentiating (4.7c) and assuming convergence of the entire Taylor series, one can show that this transformation is described by the so-called *Lie series*

$$\widetilde{x} = \sum_{k=0}^{\infty} \frac{\epsilon^k}{k!}\, \mathbf{v}^k(x). \tag{4.9}$$

where $\mathbf{v}^k$ is defined as the operator given by the composition of $\mathbf{v}$ with itself $k$ times and $\mathbf{v}^0 = I$. The linear part of the above expression, $x + \epsilon\,\mathbf{v}(x)$, is called the *infinitesimal transformation* of $x$, while $\mathbf{v}$ the *infinitesimal generator*. Equation (4.9) is also written symbolically

$$\widetilde{x} = e^{\epsilon \mathbf{v}}(x) \tag{4.10}$$

and thus referred to as the *exponentiation* of $\mathbf{v}$. So, given a vector field one can find the transformation under its flow using the exponentiation formula (4.10), while in the opposite direction, one can recover the infinitesimal generator of a given transformation using (4.8).

**Example 4.6.** Consider the vector field $\mathbf{v} = \partial_x$ on $\mathbb{R}$, where the symbol $\partial_x$ stands for the partial derivative $\partial/\partial x$ in general. Obviously $\mathbf{v}(x) = 1$, while higher-order terms in the Lie series (4.9) will vanish, $\mathbf{v}^k(x) = \mathbf{v}^{k-1}(\mathbf{v}(x)) = \mathbf{v}^{k-1}(1) = 0$ for $k \geq 2$. Therefore, the transformation under the flow of this vector field is $\varphi(\epsilon, x) = \widetilde{x} = (I + \epsilon \mathbf{v})x = x + \epsilon$, namely a *translation*.

**Example 4.7.** Consider now the vector field $\mathbf{v} = x\partial_x$ on $\mathbb{R}$, for which $\mathbf{v}(x) = x$, and so is $\mathbf{v}^k(x) = \mathbf{v}^{k-1}(\mathbf{v}(x)) = \mathbf{v}^{k-1}(x) = \cdots = \mathbf{v}(x) = x$ for any $k \geq 1$. Therefore, from the Lie series (4.9), the transformation generated by this vector field is a *scaling*, also called a *dilatation*,

$$\varphi(\epsilon, x) = \widetilde{x} = \sum_{k=0}^{\infty} \frac{\epsilon^k}{k!}\, x = e^{\epsilon} x$$

**Example 4.8.** Finally, let $X = \mathbb{R}^2$ with coordinates $(x, y)$ and $\mathbf{v} = x\partial_y - y\partial_x$. We easily see that $\mathbf{v}(x) = -y$ and $\mathbf{v}(y) = x$, based on which we can find by induction $\mathbf{v}^{2k}(x) = (-1)^k x$, $\mathbf{v}^{2k+1}(x) = -(-1)^k y$, $\mathbf{v}^{2k}(y) = (-1)^k y$ and $\mathbf{v}^{2k+1}(y) = (-1)^k x$ for any $k \geq 0$. Therefore, from the Lie series (4.9) we have that

$$\begin{aligned}
\widetilde{x} = e^{\epsilon \mathbf{v}}(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k \epsilon^{2k}}{(2k)!}\, x - \sum_{k=0}^{\infty} \frac{(-1)^k \epsilon^{2k+1}}{(2k+1)!}\, y, \\
\widetilde{y} = e^{\epsilon \mathbf{v}}(y) &= \sum_{k=0}^{\infty} \frac{(-1)^k \epsilon^{2k}}{(2k)!}\, y + \sum_{k=0}^{\infty} \frac{(-1)^k \epsilon^{2k+1}}{(2k+1)!}\, x.
\end{aligned}$$

The above sums converge to the well-known trigonometric functions sine and cosine, and so the flow of this vector field is

$$\varphi(\epsilon, (x, y)) = (\widetilde{x}, \widetilde{y}) = (x \cos\epsilon - y \sin\epsilon, y \cos\epsilon + x \sin\epsilon)$$

Switching to polar coordinates $(\rho, \phi)$ via $x = \rho \cos\phi$ and $y = \rho \sin\phi$, this transformation is more suitably expressed as $(\widetilde{\rho}, \widetilde{\phi}) = (\rho, \phi + \epsilon)$ and now easily identified as a *rotation*.

Returning to our general discussion, the change of a function $f$ defined on the manifold $X$ under the flow of a vector field $\mathbf{v}$ can also be obtained in a similar manner. Starting from (B.6) we have $\mathbf{v}(f)(\widetilde{x}) = \mathbf{v}_{\widetilde{x}}(f)$, while from (4.7c) that $\partial\varphi^i/\partial\epsilon = \xi^i(\widetilde{x})$. Combined with just the chain rule $\partial f/\partial\epsilon = \partial f/\partial\widetilde{x}^i \cdot \partial\varphi^i/\partial\epsilon$ we arrive at $\partial f/\partial\epsilon = \xi^i(\widetilde{x})\partial f/\partial\widetilde{x}^i$ or

$$\frac{\partial f}{\partial \epsilon} = \mathbf{v}(f)(\widetilde{x}) \tag{4.11}$$

Then, just as before, we can Taylor expand $f$ near the point $(0, x)$ and determine each term from (4.11). Firstly setting $\epsilon = 0$ on the latter, we get

$$\left.\frac{\partial f}{\partial \epsilon}\right|_{\epsilon=0} = \mathbf{v}(f)(x)$$

Proceeding with the second derivative, we find

$$\frac{\partial^2 f}{\partial \epsilon^2} = \frac{\partial}{\partial \epsilon}\left(\frac{\partial f}{\partial \widetilde{x}^i}\frac{\partial \varphi^i}{\partial \epsilon}\right) = \frac{\partial^2 f}{\partial \widetilde{x}^i \partial \widetilde{x}^j}\frac{\partial \varphi^i}{\partial \epsilon}\frac{\partial \varphi^j}{\partial \epsilon} + \frac{\partial f}{\partial \widetilde{x}^i}\frac{\partial^2 \varphi^i}{\partial \epsilon^2}$$

which for $\epsilon = 0$ yields

$$\left.\frac{\partial^2 f}{\partial \epsilon^2}\right|_{\epsilon=0} = \frac{\partial^2 f}{\partial x^i \partial x^j}\xi^i(x)\xi^j(x) + \frac{\partial f}{\partial x^i}\frac{\partial \xi^i}{\partial x^j}\xi^j(x) = \xi^j(x)\frac{\partial}{\partial x^j}\left(\xi^i(x)\frac{\partial f}{\partial x^i}\right) = \mathbf{v}^2(f)(x)$$

Continuing in this fashion to higher derivatives and within the convergence assumption, the transformation of a function $f$ defined on the manifold $X$ under the flow of a vector field $\mathbf{v}$ is given by the Lie series

$$f(\widetilde{x}) = e^{\epsilon \mathbf{v}}(f(x)) = \sum_{k=0}^{\infty} \frac{\epsilon^k}{k!}\mathbf{v}^k(f)(x). \tag{4.12}$$

As we can see the action of a vector field $\mathbf{v}$ on a function $f$ gives the infinitesimal variation of the function under the flow of the vector field.

### 4.2.2 Lie groups and their Lie algebras

The previous notions of vector fields and flows can be generalized in terms of Lie algebras and Lie groups. In general, an *$r$-parameter Lie group* is an $r$-dimensional smooth manifold that also carries the structure of a group, such that the group operation and inversion are smooth maps from the manifold to itself. In particular, *local Lie groups* can always be identified with (meaning they are isomorphic to) the neighborhood near the identity of some (global) Lie group [89, 90]. On the other hand, Lie groups usually go together with another intriguing structure, that of a Lie algebra. Specifically, the *Lie algebra $\mathfrak{g}$ that corresponds to an $r$-parameter Lie group $G$* can be identified with (meaning is isomorphic to) the tangent space to the Lie group $G$ at the identity, i.e. $\mathfrak{g} = T_eG$, and therefore $\mathfrak{g}$ is $r$-dimensional, too. More generally,

**Definition 4.9.** A *Lie algebra* is a real vector space $\mathfrak{g}$ equipped with a bilinear operation $[\ ,\ ]:\mathfrak{g}\times\mathfrak{g}\longrightarrow\mathfrak{g}$ called the *Lie bracket* which for all $\mathbf{v},\mathbf{u},\mathbf{w}\in\mathfrak{g}$ satisfies the properties:

1. $[\mathbf{v},\mathbf{u}]=-[\mathbf{u},\mathbf{v}]$ (skew-symmetry)
2. $[[\mathbf{v},\mathbf{u}],\mathbf{w}]+[[\mathbf{w},\mathbf{v}],\mathbf{u}]+[[\mathbf{u},\mathbf{w}],\mathbf{v}]=0$ (Jacobi identity)

**Example 4.10.** As explained in appendix B, the space of all vector fields on a manifold $X$ denoted as $\mathcal{X}(X)$ is a vector space over $\mathbb{R}$ with $\partial/\partial x^i$ as a basis. Moreover $\mathcal{X}(X)$ is a Lie algebra, where the *Lie bracket of two vector fields* $\mathbf{v}$ and $\mathbf{u}$ is defined as the unique vector field

$$[\mathbf{v},\mathbf{u}]=\mathbf{vu}-\mathbf{uv} \tag{4.15}$$

For our purposes here (and for most applications as well) the above somewhat abstract discussion can be made more concrete, and realize Lie groups merely as groups of transformations, as they were actually first comprehended by Lie himself. As previously mentioned, the flow of a vector field is a one-parameter local group of transformations. A natural generalization to this notion is a local group of transformations that depends on (not just one, but) $r$ parameters $\epsilon_1,\ldots,\epsilon_r$, and which is accordingly defined by the action of (not $\mathbb{R}$, but) a local Lie group $G$ on the manifold $X$ with similar properties (4.7a)-(4.7b).

**Definition 4.11.** Let $X$ be a smooth manifold and $G$ an $r$-parameter local Lie group with $\cdot$ denoting the group operation. An $r$-parameter local group of transformations is the action of $G$ on $X$ defined as a smooth map $\varphi:\mathcal{U}\longrightarrow X$, where $\mathcal{U}$ is an open subset of $G\times X$ such that for each $x\in X$ the set $G_x=\{\epsilon\in G:(\epsilon,x)\in\mathcal{U}\}$ contains the identity element $e$ of $G$, and which satisfies the properties:

1. $\varphi(e,x)=x$, for all $x\in X$
2. $\varphi(\epsilon,\varphi(\delta,x))=\varphi(\epsilon\cdot\delta,x)$, for all $\delta\in G_x$ and $\epsilon\in G_{\varphi(\delta,x)}$ such that $\epsilon\cdot\delta\in G_x$

In fact, since under the action of $G$ to $X$ each element of $G$ is associated with a map from the manifold $X$ to itself, a local group of transformations of $X$ is practically identified with the local Lie group $G$.

For symmetry methods and their criteria the action of an $r$-parameter Lie group can actually be considered as a composition of $r$ one-parameter subgroups such as the flow of a vector field previously shown. Admittedly skipping important steps of Lie group theory along the way (see [112]), let us go right into the heart of it. And that is a generalization of relations (4.8)

and (4.10): Assuming connected Lie groups, the transformation $\widetilde{x} = \varphi(\epsilon, x)$ of a point $x$ of a manifold $X$ under the action of $G$ can be expressed similarly to (4.10) as

$$\widetilde{x} = e^{\epsilon_1 \mathbf{v}_1} e^{\epsilon_2 \mathbf{v}_2} \cdots e^{\epsilon_r \mathbf{v}_r}(x) \tag{4.16}$$

Here $\epsilon = (\epsilon_1, \ldots, \epsilon_r)$ are local coordinates in $G$ near the identity element $e$, giving rise to one-parameter subgroups $G_i$ of $G$ for $i = 1, \ldots, r$. In this context, the action $\varphi(\epsilon_i, x) = e^{\epsilon_i \mathbf{v}_i}(x)$ (no summation) of each subgroup $G_i$ on the manifold $X$ can be viewed as the flow of the vector field $\mathbf{v}_i$, meaning

$$\mathbf{v}_{i_x} = \left.\frac{\partial \varphi}{\partial \epsilon_i}\right|_{\epsilon_i = 0}, \tag{4.17}$$

Now, $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are $r$ linearly independent vector fields that form a basis of the Lie algebra $\mathfrak{g}$ that corresponds to the Lie group $G$. And as already mentioned and silently considered above there is a 1-1 correspondence between $\mathbf{v}_i$ and $G_i$. Moreover, for transformation groups the Lie algebra $\mathfrak{g}$ of a Lie group $G$ acting on a manifold $X$ can be mapped to the space $\mathcal{X}(X)$ of vector fields on $X$ while retaining its Lie bracket structure. This means that there is a homomorphism $\psi$ between $\mathfrak{g}$ and $\mathcal{X}(X)$, which gives rise to a Lie subalgebra $\psi(\mathfrak{g})$ of vector fields on $X$. The latter are defined precisely as in (4.17) namely the vector fields on $X$ whose flow coincides with the action of one-parameter subgroups of $G$. In practice $\mathfrak{g}$ is identified with its image $\psi(\mathfrak{g})$ and the homomorphism $\psi$ is omitted, so that (4.17) makes sense. Keeping this in mind, $\mathbf{v}_i$ and elements of $\mathfrak{g}$ in general are treated as vector fields on $X$, just like elements of $G$ are treated as transformations on $X$ as previously mentioned.

The reason is the following: As with any Lie algebra, the basis vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ of $\mathfrak{g}$ satisfy the fundamental relations

$$[\mathbf{v}_i, \mathbf{v}_j] = c_{ij}^k \mathbf{v}_k \tag{4.18}$$

for $i, j, k = 1, \ldots, r$, where $c_{ij}^k$ form a constant 3-tensor of mixed type and are known as the *structure constants* of $\mathfrak{g}$. Reflecting the Lie bracket properties, these are easily seen to satisfy

1. $c_{ij}^k = -c_{ji}^k$ (skew-symmetry)
2. $c_{ij}^k\, c_{kl}^m + c_{li}^k\, c_{kj}^m + c_{jl}^k\, c_{ki}^m = 0$ (Jacobi identity)

where all indices take values $1, \ldots, r$. Now, the justification for the above treatment is that up to isomorphisms the structure constants define the same Lie algebras and the same Lie groups accordingly (given the 1-1 correspondence between them explained below).

The vector fields of the Lie algebra $\mathfrak{g}$ are called the *infinitesimal generators* of the Lie group (action) $G$, and when they form a basis of $\mathfrak{g}$ exponentiation (4.16) can always yield the local

transformations of $G$. Equation (4.18) is the content of *Lie's second fundamental theorem*, while the two properties of the structure constants are given by *Lie's third fundamental theorem*. The *first fundamental theorem of Lie* in essence states that the flow of a vector field gives rise to a one-parameter local group of transformations, as we saw in the previous subsection.

Lie's truly brilliant idea in symmetry theory of differential equations that cannot be overestimated was the use of Lie algebras instead of the corresponding Lie groups. For the former carry almost all the information of the latter, but translated from the manifold to the much simpler linearization version of it offered by the tangent space. And while every Lie group gives rise to a Lie algebra, in some cases the converse is also true. For finite-dimensional Lie algebras, a direct consequence of Ado's theorem is a 1-1 correspondence between simply-connected Lie groups and Lie algebras (see [112] and references therein). The same holds between connected subgroups of a Lie group and subalgebras of the corresponding Lie algebras (see [89, 112]). These results allow us to switch easily from elements of $G$ to vector fields of $\mathfrak{g}$ and back.

An illustration of the above ideas that also characterizes the Lie bracket of two vector fields as the infinitesimal *commutator* is given by the next theorem; for proof see [89], p. 36.

**Theorem 4.12.** Let $X$ be a smooth manifold and $G$ a two-parameter connected Lie group acting on $X$. Then the flows generated by two vector fields $\mathbf{v}_1$ and $\mathbf{v}_2$ commute if and only if their Lie bracket vanishes everywhere, i.e. for all $x \in X$ and $\epsilon = (\epsilon_1, \epsilon_2) \in G$

$$e^{\epsilon_1 \mathbf{v}_1} e^{\epsilon_2 \mathbf{v}_2}(x) = e^{\epsilon_2 \mathbf{v}_2} e^{\epsilon_1 \mathbf{v}_1}(x) \quad \Leftrightarrow \quad [\mathbf{v}_1, \mathbf{v}_2] = 0$$

## 4.3 Transformations of differential equations

Consider now a vector field $\mathbf{v}$ on the space $M$ of the independent and dependent variables of a system of differential equations $\Delta\left(x, u, u^{(n)}\right) = 0$. Under the flow of $\mathbf{v}$ the transformation of the variables $x$ and $u$ to new variables

$$\begin{aligned} \widetilde{x} &= e^{\epsilon \mathbf{v}}(x) \\ \widetilde{u} &= e^{\epsilon \mathbf{v}}(u) \end{aligned} \tag{4.21}$$

naturally induces a transformation of the derivatives of the function $u = f(x)$ to the corresponding derivatives of the transformed function $\widetilde{u} = \widetilde{f}\,(\widetilde{x})$. In other words the one-parameter local group of transformations $G$ generated by the vector field $\mathbf{v}$, acting on $M$ through the flow

$$\varphi\left(\epsilon, (x, u)\right) = (\widetilde{x}, \widetilde{u})$$

is naturally extended to a one-parameter local group of transformations $G^{(n)}$ acting on the $n$-jet space $M^{(n)}$ by simply requesting that the transformation of the derivatives of $u$ with respect to $x$ is given by the derivatives of the transformed $\widetilde{u}$ with respect to the transformed $\widetilde{x}$, that is $\widetilde{u^{(n)}} = \widetilde{u}^{(n)}$. This action is defined by the map $\varphi^{(n)} : \mathcal{U}^{(n)} \longrightarrow M^{(n)}$, where $\mathcal{U}^{(n)}$ is an open subset of $\mathbb{R} \times M^{(n)}$, with

$$\varphi^{(n)}(\epsilon, (x, u, u^{(n)})) = (\widetilde{x}, \widetilde{u}, \widetilde{u}^{(n)})$$

and can be viewed as the flow generated by the vector field $\mathbf{v}^{(n)}$ on the space $M^{(n)}$ satisfying the relation

$$\mathbf{v}^{(n)}_{(x,u,u^{(n)})} = \left.\frac{\partial \varphi^{(n)}}{\partial \epsilon}\right|_{\epsilon=0} \tag{4.22}$$

Put it another way, knowledge of $\mathbf{v}^{(n)}$ is sufficient to describe the transformation of the system $\Delta\left(x, u, u^{(n)}\right) = 0$ under the change of variables (4.21), using again exponentiation

$$\begin{aligned} \widetilde{x} &= e^{\epsilon \mathbf{v}^{(n)}}(x) \\ \widetilde{u} &= e^{\epsilon \mathbf{v}^{(n)}}(u) \\ \widetilde{u}^{(n)} &= e^{\epsilon \mathbf{v}^{(n)}}(u^{(n)}) \end{aligned} \tag{4.23}$$

The field $\mathbf{v}^{(n)}$ defined by the above equation (4.22) is called the *$n$-th prolongation of the vector field* $\mathbf{v}$. And similar to (4.12), one can prove that the transformation of a function $f(x, u, u^{(n)})$ defined on the manifold $M^{(n)}$ under the flow $\varphi^{(n)}$ is given by the Lie series

$$f(\widetilde{x}, \widetilde{u}, \widetilde{u}^{(n)}) = \sum_{k=0}^{\infty} \frac{\epsilon^k}{k!} \left(\mathbf{v}^{(n)}\right)^k (f)(x, u, u^{(n)}). \tag{4.24}$$

where now the action of $\mathbf{v}^{(n)}$ on the function $f(x, u, u^{(n)})$ describes accordingly the infinitesimal variation of the function under the flow $\varphi^{(n)}$.

**Example 4.13.** The simplest case to begin with would be a constant vector field on $M = \mathbb{R}^{q+p}$ of the form

$$\mathbf{v} = \frac{\partial}{\partial x^1} + \cdots + \frac{\partial}{\partial x^q} + \frac{\partial}{\partial u^1} + \cdots + \frac{\partial}{\partial u^p}$$

Genaralizing Example 4.6, the flow $\varphi\left(\epsilon, (x, u)\right) = (\widetilde{x}, \widetilde{u})$ generated by this vector field is simply a translation in each component,

$$\left(\widetilde{x}^j, \widetilde{u}^i\right) = \left(x^j + \epsilon, u^i + \epsilon\right)$$

Therefore, the new derivatives remain unchanged,

$$\widetilde{u}^i_{\widetilde{x}^j} = \frac{\partial \widetilde{u}^i}{\partial \widetilde{x}^j} = \frac{\partial u^i}{\partial x^j} = u^i_{x^j}$$

and, of course, this is carried on to higher derivatives of any order, meaning $\widetilde{u}^{(n)} = u^{(n)}$, as well. Thus, the flow extended on $M^{(n)} = \mathbb{R}^{q+p+p_n}$ would be $\varphi^{(n)}\big(\epsilon, (x, u, u^{(n)})\big) = \big(\widetilde{x}, \widetilde{u}, u^{(n)}\big)$. If we differentiate the latter equation with respect to $\epsilon$ and set $\epsilon = 0$,

$$\left.\frac{\partial \varphi^{(n)}}{\partial \epsilon}\right|_{\epsilon=0} = (\underbrace{1, \dots, 1}_{q}, \underbrace{1, \dots, 1}_{p}, \underbrace{0, \dots, 0}_{p_n})$$

we arrive at the components of the vector field $\mathbf{v}^{(n)}$ that generates the flow $\varphi^{(n)}$. Therefore in this case $\mathbf{v}^{(n)} = \mathbf{v}$, that is the prolongation of this vector field is simply the vector field itself since no transformation occurs in the derivatives.

**Example 4.14.** Let us consider the case of one independent variable $t \in \mathbb{R}$ and one dependent $x \in \mathbb{R}$ and the vector field

$$\mathbf{v} = t\frac{\partial}{\partial t} - 2x\frac{\partial}{\partial x}$$

Following Example 4.7, the flow of this vector field according to (4.21) is the scaling

$$\varphi\left(\epsilon, (t, x)\right) = \big(\widetilde{t}, \widetilde{x}\big) = \big(te^{\epsilon}, xe^{-2\epsilon}\big)$$

From the chain rule we can find the derivative $\dot{\widetilde{x}}\left(\widetilde{t}\right)$ in terms of the derivative $\dot{x}(t)$

$$\dot{\widetilde{x}} = \frac{d\widetilde{x}}{d\widetilde{t}} = \frac{d\widetilde{x}}{dx}\frac{dx}{dt}\left(\frac{d\widetilde{t}}{dt}\right)^{-1} = e^{-2\epsilon}\frac{dx}{dt}e^{-\epsilon} = e^{-3\epsilon}\dot{x}$$

So, the transformation on $M = \mathbb{R}^2$ with coordinates $(t, x)$ extends to a transformation on $M^{(1)} = \mathbb{R}^3$ with coordinates $(t, x, \dot{x})$ given by

$$\varphi^{(1)}\left(\epsilon, (t, x, \dot{x})\right) = \big(te^{\epsilon}, xe^{-2\epsilon}, \dot{x}e^{-3\epsilon}\big)$$

Likewise, for the second derivative

$$\ddot{\widetilde{x}} = \frac{d\dot{\widetilde{x}}}{d\widetilde{t}} = \frac{d\dot{\widetilde{x}}}{d\dot{x}}\frac{d\dot{x}}{dt}\left(\frac{d\widetilde{t}}{dt}\right)^{-1} = e^{-3\epsilon}\frac{d\dot{x}}{dt}e^{-\epsilon} = e^{-4\epsilon}\ddot{x}$$

and so we get the transformation

$$\varphi^{(2)}\left(\epsilon, (t, x, \dot{x}, \ddot{x})\right) = \big(te^{\epsilon}, xe^{-2\epsilon}, \dot{x}e^{-3\epsilon}, \ddot{x}e^{-4\epsilon}\big)$$

Differentiating the latter equation with respect to $\epsilon$ and setting $\epsilon = 0$,

$$\left.\frac{\partial \varphi^{(2)}}{\partial \epsilon}\right|_{\epsilon=0} = (t, -2x, -3\dot{x}, -4\ddot{x})$$

yields the components of the vector field $\mathbf{v}^{(2)}$ that generates the flow $\varphi^{(2)}$ on $M^{(2)} = \mathbb{R}^4$ with coordinates $(t, x, \dot{x}, \ddot{x})$,

$$\mathbf{v}^{(2)} = t\frac{\partial}{\partial t} - 2x\frac{\partial}{\partial x} - 3\dot{x}\frac{\partial}{\partial \dot{x}} - 4\ddot{x}\frac{\partial}{\partial \ddot{x}}$$

The procedure followed in the previous example can be employed in general for any vector field and yields inductively the prolongation of any vector. The resulting algorithm is given by the next proposition.

**Proposition 4.15.** Let $\mathbf{v}$ be a vector field defined on the space $M$ of the independent and dependent variables of the form

$$\mathbf{v} = \xi^j(x,u)\frac{\partial}{\partial x^j} + \eta^i(x,u)\frac{\partial}{\partial u^i}$$

The $n$-th prolongation of $\mathbf{v}$ on the corresponding jet space $M^{(n)}$ of $n$-th order is given by the general formula

$$\mathbf{v}^{(n)} = \mathbf{v} + \eta_i^J(x,u,u^{(n)})\frac{\partial}{\partial u_J^i} \tag{4.25}$$

where

$$\eta_i^J(x,u,u^{(n)}) = \frac{d}{dx^J}\left(\eta^i - \xi^j\frac{\partial u^i}{\partial x^j}\right) + \xi^j\frac{\partial u_J^i}{\partial x^j}$$

and $d/dx^J = D_J = D_{j_1}D_{j_2}\cdots D_{j_k}$, $D_j$ being the *total derivative* with respect to $x^j$.

It is worth noting that the $n$-th prolongation $\mathbf{v}^{(n)}$ of the field $\mathbf{v}$ can be also viewed as the first prolongation of the field $\mathbf{v}^{(n-1)}$, just like the $n$-th derivative of a function is the derivative of the $(n-1)$-th derivative. That is why in practice the calculation of the components $\eta_i^J$ is made step by step, using the recursion relation

$$\eta_i^J = \frac{d\eta_i^{J/k}}{dx^{j_k}} - \frac{d\xi^j}{dx^{j_k}}\frac{\partial u_{J/k}^i}{\partial x^j} \tag{4.26}$$

where $J/k = (j_1,\ldots,j_{k-1})$. In the case of one independent variable $t$ and one dependent $x$, the above relation takes the form

$$\eta^{(k)} = \frac{d\eta^{(k-1)}}{dt} - \frac{d\xi}{dt}\frac{d^k x}{dt^k} \tag{4.27}$$

It is preferable at this point to give an illustrative example to avoid any confusions caused by the above notation.

**Example 4.16.** In order to be as thorough as possible, let us examine the case of two independent variables $t$ and $x$ and one dependent $u$. Consider the vector field

$$\mathbf{v} = \tau(t,x,u)\frac{\partial}{\partial t} + \xi(t,x,u)\frac{\partial}{\partial x} + \eta(t,x,u)\frac{\partial}{\partial u}$$

The first prolongation of this field is

$$\mathbf{v}^{(1)} = \mathbf{v} + \eta^t\frac{\partial}{\partial u_t} + \eta^x\frac{\partial}{\partial u_x}$$

where according to (4.26)

$$\eta^t = \frac{d\eta}{dt} - \frac{d\tau}{dt}u_t - \frac{d\xi}{dt}u_x = \eta_t + (\eta_u - \tau_t)\,u_t - \xi_t u_x - \tau_u u_t^2 - \xi_u u_x u_t$$
$$\eta^x = \frac{d\eta}{dx} - \frac{d\tau}{dx}u_t - \frac{d\xi}{dx}u_x = \eta_x + (\eta_u - \xi_x)\,u_x - \tau_x u_t - \xi_u u_x^2 - \tau_u u_x u_t$$

The second prolongation is

$$\mathbf{v}^{(2)} = \mathbf{v}^{(1)} + \eta^{tt}\frac{\partial}{\partial u_{tt}} + \eta^{xt}\frac{\partial}{\partial u_{xt}} + \eta^{xx}\frac{\partial}{\partial u_{xx}}$$

where for example

$$\begin{aligned}\eta^{xx} &= \frac{d\eta^x}{dx} - \frac{d\tau}{dx}u_{xt} - \frac{d\xi}{dx}u_{xx}\\ &= \eta_{xx} + (2\eta_{xu} - \xi_{xx})\,u_x - \tau_{xx}u_t + (\eta_{uu} - 2\xi_{xu})\,u_x^2 - 2\tau_{xu}u_x u_t - \xi_{uu}u_x^3 -\\ &\quad - \tau_{uu}u_x^2 u_t + (\eta_u - 2\xi_x)\,u_{xx} - 2\tau_x u_{xt} - 3\xi_u u_x u_{xx} - \tau_u u_t u_{xx} - 2\tau_u u_x u_{xt}\end{aligned}$$

We realize that as we move on to higher orders the calculation of the components of the prolonged vector fields becomes more and more complicated. For example the term $\eta^{xxx}$ of the third prolongation is

$$\begin{aligned}\eta^{xxx} = \frac{d\eta^{xx}}{dx} - \frac{d\xi}{dx}u_{xxx} - \frac{d\tau}{dx}u_{xxt} &= \eta_{xxx} + (3\eta_{xxu} - \xi_{xxx})\,u_x - \tau_{xxx}u_t + (3\eta_{xuu} - 3\xi_{xxu})\,u_x^2 -\\ &- 3\tau_{xxu}u_x u_t + (\eta_{uuu} - 3\xi_{xuu})\,u_x^3 - 3\tau_{xuu}u_x^2 u_t - \xi_{uuu}u_x^4 - \tau_{uuu}u_x^3 u_t + 3\,(\eta_{xu} - \xi_{xx})\,u_{xx} -\\ &- 3\tau_{xx}u_{xt} + 3\,(\eta_{uu} - 2\xi_{xu})\,u_x u_{xx} - 3\tau_{xu}u_t u_{xx} - 6\tau_{xu}u_x u_{xt} - 6\xi_{uu}u_x^2 u_{xx} - 3\tau_{uu}u_x u_t u_{xx} -\\ &- 3\tau_{uu}u_x^2 u_{xt} - 3\xi_u u_{xx}^2 - 3\tau_u u_{xx}u_{xt} + (\eta_u - 3\xi_x)\,u_{xxx} - 3\tau_x u_{xxt} - 4\xi_u u_x u_{xxx} -\\ &- 3\tau_u u_x u_{xxt} - \tau_u u_t u_{xxx}\end{aligned}$$

## 4.4 Symmetry groups

The transformations we described in the previous sections map a point $(x, u)$ to another $(\widetilde{x}, \widetilde{u})$ in the space of independent and dependent variables and have the structure of a Lie group. Therefore, they are widely known as *Lie point transformations*. A group of transformations will always mean a group of Lie point transformations in this thesis.

**Definition 4.17.** Let $\Delta\left(x, u, u^{(n)}\right) = 0$ be a system of differential equations. A local group $G$ of transformations $(x, u) \longrightarrow (\widetilde{x}, \widetilde{u})$ acting on the space $M$ of the independent and dependent variables with the property that if $u = f(x)$ is a solution of the system, then $\widetilde{u} = \widetilde{f}\,(\widetilde{x})$ is also a solution of the system is called a *symmetry group* for the system. Equivalently $G$ is a symmetry

group of the system if $\Delta\left(x,u,u^{(n)}\right)=0 \Rightarrow \Delta\left(\widetilde{x},\widetilde{u},\widetilde{u}^{(n)}\right)=0$, i.e. the equations of the system are invariant under the action of the group.

**Remark 4.18.** According to Definition 4.17, the existence of a symmetry group is independent of the choice of variables in which a system or its solution is expressed.

If one tries to employ the above definition to find a symmetry of a given system, he will be faced with a new system of differential equations much more difficult than the original one. For example, consider a first-order differential equation $\dot{x}=F(t,x)$ and a group of transformations $\left(\widetilde{t},\widetilde{x}\right)=\varphi\left(t,x\right)$ where $\varphi=\left(\varphi^1,\varphi^2\right)$. Then $\widetilde{x}\left(\widetilde{t}\right)$ is a solution if $\dot{\widetilde{x}}=F(\widetilde{t},\widetilde{x})$, which in terms of the original variables yields

$$\frac{\varphi^2_t+F(t,x)\varphi^2_x}{\varphi^1_t+F(t,x)\varphi^1_x}=F\big(\varphi^1(t,x),\varphi^2(t,x)\big) \tag{4.28}$$

Even if the equation we started with was linear, i.e. $F$ was linear, the latter partial differential equation would be quasilinear.

One of the key contributions of Lie's fundamental work was to linearize the symmetry condition of the above definition, expressing it in terms of the Lie algebra instead of the Lie group, meaning on the tangent space of $G$ rather than $G$ itself.

**Theorem 4.19.** Let $\Delta\left(x,u,u^{(n)}\right)=0$ be a nondegenerate system of differential equations. A vector field $\mathbf{v}$ is a generator of a symmetry group for the system if and only if $\mathbf{v}^{(n)}(\Delta)=0$ whenever $\Delta=0$.

*Proof.* Let us examine first the implications of Definition 4.17. Consider a symmetry transformation (4.21) generated by the vector field

$$\mathbf{v}=\xi^j(x,u)\frac{\partial}{\partial x^j}+\eta^i(x,u)\frac{\partial}{\partial u^i}.$$

Then if $u=f(x)$ is a solution of the system $\Delta_\nu(x,u,u^{(n)})=0$, so is $\widetilde{u}=\widetilde{f}\,(\widetilde{x})$. According to (4.4) this means that

$$\Delta_\nu(\widetilde{x},\widetilde{u},\widetilde{u}^{(n)})=0.$$

Differentiating with respect to $\epsilon$ and setting $\epsilon=0$, we get

$$\frac{\partial\Delta_\nu}{\partial x^j}\left.\frac{\partial\widetilde{x}^j}{\partial\epsilon}\right|_{\epsilon=0}+\frac{\partial\Delta_\nu}{\partial u^i}\left.\frac{\partial\widetilde{u}^i}{\partial\epsilon}\right|_{\epsilon=0}+\frac{\partial\Delta_\nu}{\partial u^i_J}\left.\frac{\partial\widetilde{u}^i_J}{\partial\epsilon}\right|_{\epsilon=0}=0.$$

From definition (4.22) and (4.25), the previous equation takes the form

$$\xi^j\frac{\partial\Delta_\nu}{\partial x^j}+\eta_i\frac{\partial\Delta_\nu}{\partial u^i}+\eta^J_i\frac{\partial\Delta_\nu}{\partial u^i_J}=0 \quad \text{or} \quad \mathbf{v}^{(n)}(\Delta_\nu)=0.$$

Therefore, the latter condition, assuming $\Delta_\nu = 0$ from the beginning, is necessary for $\mathbf{v}$ to be a symmetry generator of the system.

To prove that is also sufficient, let $\mathbf{v}$ be a vector field on $M$ for which $\mathbf{v}^{(n)}(\Delta_\nu) = 0$ for every solution $u = f(x)$ of the system, that is, whenever $\Delta_\nu = 0$. For convenience we collect all the variables together, $z = \left(z^1, \dots, z^{m_n}\right) = \left(x, u, u^{(n)}\right)$, as in section 4.1, where $z^\mu$ is any of the coordinates $x^j$, $u^i$ or $u^i_J$ of $M^{(n)}$. Since $\mathbf{v} \neq 0$, we can always choose local coordinates $\left(y^1, \dots, y^{m_n}\right)$ by a suitable transformation in $M$, in which the vector field $\mathbf{v}$ is expressed as

$$\mathbf{v} = \frac{\partial}{\partial y^1}$$

where $y^1$ is a function of the independent and dependent variables. Accordingly $\mathbf{v}^{(n)} = \mathbf{v}$ (see Example 4.13), and as the existence of a symmetry is indepenedent of the coordinate system used our assumption now reads $\mathbf{v}(\Delta_\nu) = 0$ when $\Delta_\nu = 0$. Due to the maximal rank condition of nondegenerate systems, this means that the equations of the system are genuinely independent of the variable $y^1$,

$$\frac{\partial \Delta_\nu}{\partial y^1} = 0$$

since the Jacobian matrix $(\partial \Delta_\nu / \partial y^\mu)$ is of rank $l$, when $\Delta_\nu = 0$. Therefore, the translation $\widetilde{y} = \left(y^1 + \epsilon, y^2, \dots, y^m\right)$ generated by the vector field $\mathbf{v}$ is admitted by the system. In other words, if $\Delta(y) = 0$ then $\Delta(\widetilde{y}) = 0$ too. And since the system is locally solvable, this implies that if $y$ is a solution of the system, then $\widetilde{y}$ is also a solution of the system. Therefore, $\mathbf{v}$ is a symmetry generator.

The above conclusions are not limited to one-parameter symmetry groups, for the same arguments hold for every infinitesimal generator $\mathbf{v}_i$ of an $r$-parameter group. □

Theorem 4.19 hands in a criterion that allows to find all the symmetries for a system of differential equations, meaning the most general symmetry group admitted by the system. Let us stress the fact that the condition $\mathbf{v}^{(n)}(\Delta) = 0$ needs to hold only on solutions of the system. Related to this, another important point is the following.

**Remark 4.20.** A solution of a system does not satisfy only its equations $\Delta_\nu = 0$, but also its differential consequences, that is, the equations $D_J \Delta_\nu = 0$ resulting from differentiation with respect to any of the independent variables $x^{j_1}, \dots, x^{j_k}$ for any $k$.

We should also emphasize that the nondegeneracy assumption was essential for the converse part of the infinitesimal symmetry criterion. The maximal rank condition and the local solvabil-

ity guarantee that neither more nor less, respectively, generators are found than the symmetries truly admitted. Their necessity is better illustrated by simple examples.

**Example 4.21.** For the equation $\Delta = (\ddot{x} - t)^2 = 0$, the Jacobian matrix $(\partial\Delta/\partial z^\mu)$, where $z = (t, x, \dot{x}, \ddot{x})$, has zero rank when $\Delta = 0$, and so the maximal rank condition does not hold. Take then for instance the vector field $\mathbf{v} = \partial_t$; it meets the symmetry criterion of Theorem 4.19, $\mathbf{v}^{(n)}(\Delta) = \mathbf{v}(\Delta) = 2(\ddot{x} - t) = 0$ when $\Delta = (\ddot{x} - t)^2 = 0$, but the translation $\widetilde{t} = t + \epsilon$ generated is obviously not a symmetry of the equation. The zero rank of the related Jacobian matrix in this case means that all the derivatives of $\Delta$ vanish on solutions of the system, and as a consequence one cannot conclude on which variables the equation does not depend on. In particular, $\partial\Delta/\partial t = 0$ on $\Delta = 0$ does not imply that $\Delta$ is independent of $t$ (this deduction was crucial in the proof of the previous theorem).

**Example 4.22.** In Example 4.5 we saw that the system $\{\Delta_1 = u_t = 0, \Delta_2 = u_x - t(u-1) = 0\}$, whose only solution is $u(t, x) = 1$ for any $t$ and $x$, is not locally solvable. Contrary to the previous example, now $\widetilde{t} = t + \epsilon$ is a symmetry for the system, for if $u(t, x)$ is a solution of the system, so is $u(\widetilde{t}, x)$. But now the generator $\mathbf{v} = \partial_t$ does not satisfy the infinitesimal symmetry criterion since $\mathbf{v}^{(n)}(\Delta_2) = \mathbf{v}(\Delta_2) = -(u-1)$ does not vanish when $\Delta_1 = \Delta_2 = 0$.

Let us now see how Theorem 4.19 works in practice to find the symmetries admitted by an ordinary differential equation.

**Example 4.23.** Consider the one-dimensional motion of a particle experiencing a nonlinear velocity-dependent force described by the second-order differential equation

$$\ddot{x} = \frac{\dot{x}^2}{x} - x^2$$

The second prolongation of the vector field $\mathbf{v} = t\partial_t - 2x\partial_x$, calculated earlier in the example 4.14, is $\mathbf{v}^{(2)} = t\partial_t - 2x\partial_x - 3\dot{x}\partial_{\dot{x}} - 4\ddot{x}\partial_{\ddot{x}}$, and

$$\mathbf{v}^{(2)}\left(\ddot{x} - \frac{\dot{x}^2}{x} + x^2\right) = -4\ddot{x} - 3\dot{x}\left(-\frac{2\dot{x}}{x}\right) - 2x\left(\frac{\dot{x}^2}{x^2} + 2x\right) = -4\left(\ddot{x} - \frac{\dot{x}^2}{u} + x^2\right).$$

Therefore,

$$\mathbf{v}^{(2)}\left(\ddot{x} - \frac{\dot{x}^2}{x} + x^2\right) = 0, \quad \text{when} \quad \ddot{x} = \frac{\dot{x}^2}{x} - x^2\,. \tag{4.29}$$

Consequently the vector field $\mathbf{v}$ generates a symmetry for the equation.

If, however, we want to find all the symmetries admitted, we have to work our way back, meaning request an arbitrary vector field

$$\mathbf{v} = \xi(t, x)\frac{\partial}{\partial t} + \eta(t, x)\frac{\partial}{\partial x}$$

to satisfy the condition of Theorem 4.19, that is, (4.29) in this case. From (4.25)

$$\mathbf{v}^{(2)} = \xi\frac{\partial}{\partial t} + \eta\frac{\partial}{\partial x} + \eta^t\frac{\partial}{\partial \dot{x}} + \eta^{tt}\frac{\partial}{\partial \ddot{x}}$$

while (4.27) gives

$$\eta^t = \frac{d\eta}{dt} - \frac{d\xi}{dt}\dot{x}$$
$$\eta^{tt} = \frac{d\eta^t}{dt} - \frac{d\xi}{dt}\ddot{x}$$

Restricting example 4.16 to the case of one independent and one dependent variables (by setting $\tau = 0$, $u_t = 0$ and relabelling the variables $(x, u)$ by $(t, x)$), we recover expressions for the above $\eta^t$ kai $\eta^{tt}$. Substituting $\ddot{x}$ with $\dot{x}^2/x - x^2$ from the equation itself, condition (4.29) takes the form

$$\left(-\xi_{xx} - \frac{\xi_x}{x}\right)\dot{x}^3 + \left(\frac{\eta}{x^2} - \frac{\eta_x}{x} + \eta_{xx} - 2\xi_{tx}\right)\dot{x}^2 + \left(-\frac{2\eta_t}{x} + 2\eta_{tx} - \xi_{tt} + 3\xi_x x^2\right)\dot{x} +$$
$$+ \left(2\eta x + \eta_{tt} - \eta_x x^2 + 2\xi_t x^2\right) = 0$$

The functions $\xi$ and $\eta$ depend only on the variables $t$ and $x$ and not on $\dot{x}$, and so the above relation, which holds for every $\dot{x}$, is equivalent to the system

$$\begin{aligned} -\xi_{xx} - \frac{\xi_x}{x} &= 0 \\ \frac{\eta}{x^2} - \frac{\eta_x}{x} + \eta_{xx} - 2\xi_{tx} &= 0 \\ -\frac{2\eta_t}{x} + 2\eta_{tx} - \xi_{tt} + 3\xi_x x^2 &= 0 \\ 2\eta x + \eta_{tt} - \eta_x x^2 + 2\xi_t x^2 &= 0 \end{aligned} \tag{4.30}$$

Starting from the top, we find

$$x\xi_{xx} + \xi_x = 0 \;\Rightarrow\; x\xi_x = f_1(t) \;\Rightarrow\; \boxed{\xi = f_1(t)\ln|x| + f_2(t)}$$

Substituting this expression in the second equation, we have

$$\eta_{xx} - \frac{\eta_x x - \eta}{x^2} - \frac{2f_1'(t)}{x} = 0 \quad\Rightarrow\quad \eta_x - \frac{\eta}{x} - 2f_1'(t)\ln|x| = f_3(t) \quad\Rightarrow$$
$$\frac{\eta_x x - \eta}{x^2} - 2f_1'(t)\frac{\ln|x|}{x} = \frac{f_3(t)}{x} \quad\Rightarrow\quad \frac{\eta}{x} - f_1'(t)\ln^2|x| = f_3(t)\ln|x| + f_4(t) \quad\Rightarrow$$

$$\boxed{\eta = f_1'(t)x\ln^2|x| + f_3(t)x\ln|x| + f_4(t)x}$$

Proceeding with the third one, we substitute $\xi$ and $\eta$ and end up with

$$3f_1''(t)\ln|x| + 3f_1(t)x + 2f_3'(t) - f_2''(t) = 0$$

The above equation is satisfied when the coefficients of the functions of $x$ are zero, so we get

$$\boxed{f_1(t) = 0} \quad \text{and} \quad f_2'(t) = 2f_3(t) + c_2,$$

where $c_2$ an arbitrary real constant. Finally, the fourth equation of (4.30) after substituting $\xi$ and $\eta$ takes the form

$$f_3(t)x^2 \ln|x| + f_3''(t)x \ln|x| + (f_4(t) + 3f_3(t) + 2c_2)\, x^2 + f_4''(t)x = 0$$

so, likewise, we get

$$\boxed{f_3(t) = 0}\,, \quad \boxed{f_4(t) = -2c_2} \quad \text{and so} \quad \boxed{f_2(t) = c_2 t + c_1}$$

where $c_1$ is also a constant. Summing up, the general solution of system (4.30) is

$$\xi(t, x) = c_2 t + c_1$$
$$\eta(t, x) = -2c_2 x$$

In other words, every vector field of the form $\mathbf{v} = (c_2 t + c_1)\,\partial_t - 2c_2 x \partial_x$ generates a two-parameter symmetry group for the equation. So, the algebra of the symmetry group is spanned by the vector fields

$$\mathbf{v}_1 = \partial_t$$
$$\mathbf{v}_2 = t\partial_t - 2x\partial_x$$

and the corresponding transformations, meaning the one-parameter groups generated, are

$$G_1 : (\widetilde{t}, \widetilde{x}) \longrightarrow (t + \epsilon_1, x)$$
$$G_2 : (\widetilde{t}, \widetilde{x}) \longrightarrow \left(e^{\epsilon_2} t, e^{-2\epsilon_2} x\right)$$

Therefore, according to Definition 4.17, if $x = f(t)$ is a solution of the equation, so are

$$x_1 = f\,(t - \epsilon_1)$$
$$x_2 = e^{-2\epsilon_2} f\left(e^{-\epsilon_2} t\right)$$

The previous example demonstrates the way to calculate the symmetries admitted by a differential equation or in general a system of differential equations. The whole procedure consists actually of the following steps: *i*) Apply condition $\mathbf{v}^{(n)}(\Delta) = 0$ of Theorem 4.19, *ii*) substitute

the restrictions $\Delta = 0$ and write the symmetry condition as a polynomial in the derivatives, *iii*) equate the coefficients of the derivative monomials to zero, *iv*) solve the resulting system of equations to find the most general form of the vector field $\mathbf{v}$, and finally *v*) construct the symmetry group for the system. The equations derived in the third step of the above algorithm are called the *determining equations* of the symmetry group for a given system. These are always linear, homogeneous partial differential equations, and in most cases constitute an overdetermined system, meaning there are more equations than unknowns.

**Remark 4.24.** An exception to the latter rule is in particular a differential equation of the simplest form: a first-order ODE say $\dot{x} = F(t, x)$. For in this case if $\mathbf{v} = \xi(t, x)\partial_t + \eta(t, x)\partial_x$, the restriction $\dot{x} = F$ to solutions of the system in the second step of the algorithm leaves the one and only determining equation $\eta^t - \xi F_t - \eta F_x = 0$ (see Example 4.23 for $\eta^t$) with no derivatives. Therefore, it does not split to more equations, and so we are left with only one determining equation with two unknowns, i.e. the two components of the generator,

$$\frac{\partial \eta}{\partial t} + \frac{\partial \eta}{\partial x}F - \frac{\partial \xi}{\partial t}F - \frac{\partial \xi}{\partial x}F^2 - \xi F_t - \eta F_x = 0 \tag{4.31}$$

(compare with (4.28)), and which cannot be solved without making any ansatzes. For example, once $f$ is given, one may prescribe one of the components $\xi$ or $\eta$ of the vector field to find the other. In this way, a particular solution of the above equation can be obtained.

**Remark 4.25.** Another kind of differential equations that resist to symmetry analysis are actually linear ones in the sense that the linearity advantage of the determining equations over the original system is lost. The solution of the determining equations that yields the components of the symmetry generator requires in this case the solution of the original system. For example consider the general $n$-th order linear, inhomogeneous ordinary differential equation

$$f_n(t)x^{(n)} + \cdots + f_1(t)\dot{x} + f_0(t)x - g(t) = 0$$

and the $n$ linearly independent solutions $x_k(t)$ of the corresponding homogeneous equation. Due to the linearity, if $x(t)$ is a solution of the above equation, then so is $x(t) + \epsilon x_k(t)$. Therefore, $\widetilde{x} = x + \epsilon x_k(t)$ is a symmetry with generator $\mathbf{v}_k = x_k(t)\partial_x$ having the solution $x_k(t)$ of the original equation as its coefficient.

This also shows that *an n-th order linear ODE admits at least an n-parameter symmetry group*. For the same reason a linear PDE always has an infinite-dimensional symmetry algebra generated by the solution of the equation given now by an arbitrary function.

On the other hand, the symmetry properties of linear equations can help us identify linearizable equations, that is, equations that can be transformed to linear ones in suitable variables. For instance, an $n$-th order nonlinear ODE that has $n$ or more symmetries may be a linear equation in disguise.

Getting back to our discussion on the calculation of symmetries, the fourth stage of the algorithm presented that involves finding the solution of the determining equations is obviously the most difficult one. It is common practice to solve the equations one at a time starting from the one that corresponds to the highest derivative and working our way down. Since this is an overdetermined system, compatability conditions can also be of use. In many cases this technique proves quite efficient. Let us see another example of calculating the symmetry group this time for a partial differential equation.

**Example 4.26.** Encountered in many physical phenomena, describing the way waves (usually water waves, but also sound waves and even plasma waves) propagate, the celebrated Korteweg-de Vries (KdV) equation

$$u_t = u_{xxx} + uu_x$$

has also many interesting mathematical aspects, as well, leading the way in the field of integrable systems. To find the symmetry group admitted, we consider the vector field

$$\mathbf{v} = \tau(t,x,u)\frac{\partial}{\partial t} + \xi(t,x,u)\frac{\partial}{\partial x} + \eta(t,x,u)\frac{\partial}{\partial u}$$

and following the algorithm previously described, first we require

$$\mathbf{v}^{(3)}(u_t - u_{xxx} - uu_x) = 0 \qquad \text{when} \qquad u_t = u_{xxx} + uu_x$$

which takes the form

$$\eta^t - \eta^{xxx} - \eta u_x - \eta^x u = 0$$

where $\eta^x, \eta^t, \eta^{xxx}$ have been calculated in Example 4.16. From the equation itself we substitute $u_t$ as well as her derivatives $u_{xt} = u_{xxxx} + u_x^2 + uu_{xx}$ and $u_{xxt} = u_{xxxxx} + 3u_x u_{xx} + uu_{xxx}$. The (reduced) system of the determining equations (after substitutions among the original equations themselves) we end up with is

$$\tau_u = 0$$

$$\tau_x = 0$$

$$\xi_u = 0$$

$$\begin{aligned} 3\xi_x - \tau_t &= 0 \\ \eta_{uu} &= 0 \\ \eta_{xu} - \xi_{xx} &= 0 \\ \eta + \xi_t + 2\left(u\xi_x + \eta_{xxu}\right) &= 0 \\ -\eta_t + u\eta_x + \eta_{xxx} &= 0 \end{aligned}$$

From the first two equations of the above system, we have

$$\tau = 3f_1(t)$$

while form the third and fourth

$$\xi = f_1'(t)x + f_2(t)$$

The latter combined with the fifth and sixth yields

$$\eta = \big(f_1'(t) + f_3(t)\big)u + f_4(t,x)$$

Substituting $\xi$ and $\eta$ in the seventh one, we get $\left(3f_1'(t) + f_3(t)\right)u + f_1''(t)x + f_4(t,x) + f_2'(t) = 0$ resulting in

$$\begin{aligned} f_3(t) &= -3f_1'(t) \\ f_4(t,x) &= -f_1''(t)x - f_2'(t) \end{aligned}$$

Taking these into account, the last of the determining equations yields $f_1''(t)u + f_1'''(t)x + f_2''(t) = 0$ from which we deduce $f_1''(t) = f_2''(t) = 0$ and therefore

$$\begin{aligned} f_1(t) &= c_4t + c_1 \\ f_2(t) &= c_3t + c_2 \end{aligned}$$

where $c_1, c_2, c_3, c_4$ are arbitrary real constants. Putting it all together and replacing $3c_1 \to c_1$, we arrive at

$$\begin{aligned} \tau(t,x,u) &= 3c_4t + c_1 \\ \xi(t,x,u) &= c_4x + c_3t + c_2 \\ \eta(t,x,u) &= -2c_4u - c_3 \end{aligned}$$

Thus, the KdV equation admits a 4-parameter symmetry group generated by the vector fields

$$\mathbf{v}_1 = \partial_t$$

$$\mathbf{v}_2 = \partial_x$$
$$\mathbf{v}_3 = t\partial_x - \partial_u$$
$$\mathbf{v}_4 = 3t\partial_t + x\partial_x - 2u\partial_u$$

The corresponding symmetry transformations are

$$G_1 : \left(\widetilde{t}, \widetilde{x}, \widetilde{u}\right) \longrightarrow (t + \epsilon_1, x, u)$$
$$G_2 : \left(\widetilde{t}, \widetilde{x}, \widetilde{u}\right) \longrightarrow (t, x + \epsilon_2, u)$$
$$G_3 : \left(\widetilde{t}, \widetilde{x}, \widetilde{u}\right) \longrightarrow (t, x + \epsilon_3 t, u - \epsilon_3)$$
$$G_4 : \left(\widetilde{t}, \widetilde{x}, \widetilde{u}\right) \longrightarrow \left(e^{3\epsilon_4} t, e^{\epsilon_4} x, e^{-2\epsilon_4} u\right)$$

Therefore, if $u = f(t, x)$ is a solution for the KdV, so are

$$u_1 = f\left(t - \epsilon_1, x\right)$$
$$u_2 = f\left(t, x - \epsilon_2\right)$$
$$u_3 = f\left(t, x - \epsilon_3 t\right) - \epsilon_3$$
$$u_4 = e^{-2\epsilon_4} f\left(e^{-3\epsilon_4} t, e^{-\epsilon_4} x\right)$$

If we interpret $t$ and $x$ as time and space coordinates, respectively, then the subgroups $G_1$ and $G_2$ represent the symmetry of KdV under time and space translations, accordingly. If we regard $u$ as velocity, the symmetry subgroup $G_3$ is a Galilean boost for an observer moving with velocity $\epsilon_3$ with respect to an inertial frame of reference. Finally, $G_4$ expresses the symmetry of the equation under scaling transformations.

By definition, a symmetry group for a system maps a solution of the system to another solution. So, symmetries can be used, as in the previous examples, to construct new solutions $\widetilde{u} = \widetilde{f}\left(\widetilde{x}\right)$ from already known ones $u = f(x)$. Apparently, this method for finding solutions has the disadvantage that a solution must first be known. Moreover, if the symmetries found are rather usual transformations, say for example translations or scalings, then the new solutions are qualitatively not different than the ones we began with.

However, there is another way to manipulate the symmetry information, which although distinguishes between ordinary and partial differential equations, it is essentially the same. This method relies on the concept of invariance, and for ODEs leads to integration by quadratures, while for PDEs constructing invariant solutions.

## 4.5 Integration of ODEs

For treating the case of ODEs, first we recall that every $n$-th order ordinary differential equation of the form $x^{(n)} = f\left(t, x, \dot{x}, \dots, x^{(n-1)}\right)$ can be transformed to a system of $n$ first-order ordinary differential equations. On the other hand, the converse is not true; not every system of first-order ODEs can be transformed to a single ODE. Therefore, we consider the more general class of (non-autonomous) systems

$$\frac{dx}{dt} = F(t, x) \tag{4.32}$$

where $t$ is the independent variable and $x = \left(x^1, \dots, x^p\right)$ are the dependent ones, while $F$ is a smooth vector function. Note that this is a nondegenerate system.

**Theorem 4.27.** The existence of a one-parameter symmetry group for a system of first-order ordinary differential equations reduces the order of the system by one.

*Proof.* Let $\mathbf{v}$ be the generator of an one-parameter symmetry group for system (4.32). In every point $(t, x)$, where $\mathbf{v} \neq 0$, we can change to new coordinates $(s, y)$, choosing $s$ as the new independent variable and $y = \left(y^1, \dots, y^p\right)$ the new dependent ones, such that

$$\mathbf{v} = \frac{\partial}{\partial y^1}$$

The system in these coordinates takes the form $y'(s) = \widetilde{F}(y, s)$, while $\mathbf{v}^{(1)} = \mathbf{v}$. Therefore, the symmetry condition is expressed as

$$\mathbf{v}^{(1)}\left(\frac{dy^i}{ds} - \widetilde{F}_i(y, s)\right) = 0 \quad \Rightarrow \quad \frac{\partial \widetilde{F}_i}{\partial y^1} = 0$$

meaning the functions $\widetilde{F}_i$ do not depend on the variable $y^1$. Thus, the system we are faced with has an ignorable variable,

$$\begin{aligned}
\frac{dy^1}{ds} &= \widetilde{F}_1\left(s, y^2, \dots, y^p\right) \\
\frac{dy^2}{ds} &= \widetilde{F}_2\left(s, y^2, \dots, y^p\right) \\
&\vdots \\
\frac{dy^p}{ds} &= \widetilde{F}_p\left(s, y^2, \dots, y^p\right)
\end{aligned}$$

The last $(p-1)$ equations can be solved independently from the first one, and if their solution $\left(y^2(s), \dots, y^p(s)\right)$ is found, then the first one can be solved by simply integrating

$$y^1(s) = \int \widetilde{F}_1\left(s, y^2(s), \dots, y^p(s)\right) ds$$

□

**Remark 4.28.** The variables $\left(s, y^2, \ldots, y^p\right)$ are actually the characteristics of the first-order linear PDE $\mathbf{v}(f) = 0$, while $y^1$ can be any particular solution of $\mathbf{v}(f) = 1$.

**Remark 4.29.** Recall that for first-order ODEs the corresponding determining equation is underdetermined as explained in Remark 4.24. This drawback is often paraphrased as *a first-order ODE has infinite symmetries*, considering an underdetermined system has infinite solutions. It could be argued that solving (4.31), although linear, is a much more difficult task than solving the original equation. In light of the above theorem, however, the disadvantage of not being able to solve the determining equation in complete generality is of less significance since it suffices to find just one particular solution, that is, just one symmetry to integrate a first-order ODE.

In fact, underdetermined systems of determining equations is a characteristic that generalizes to every system of first-order ODEs; the number of the determining equations is always less by one from the number of the unknown components of the symmetry generator. For $n$-th order systems $n \geq 2$, however, a typical way out to this problem is to transform them to an equivalent $n$-th order equation, when this is possible. Thus, we can perform the symmetry analysis for that equation and then go back to the system. Careful handling is needed though regarding the transformation of the symmetries: the symmetry generators of the system are the ($n$-1)-th prolongations of the symmetry generators of the related equation.

**Example 4.30.** We have seen earlier that the differential equation $\ddot{x} = \dot{x}^2/x - x^2$ has a one-parameter symmetry group generated by the vector field $\mathbf{v}_2 = t\partial_t - 2x\partial_x$. Note that the transformation $\widetilde{t} = -t$ leaves the equation also invariant; this is an example of a *discrete symmetry*. Thus, there is no loss of generality in assuming $t > 0$. The equation can be equivalently transformed to the system

$$\dot{x} = v$$
$$\dot{v} = \frac{v^2}{x} - x^2$$

Notice that the expression of $\mathbf{v}_2$ in the new variables $(t, x, v = \dot{x})$ actually requires its first prolongation,

$$\mathbf{v}_2 = t\partial_t - 2x\partial_x - 3v\partial_v$$

We seek new variables $(s, y, u)$ in which the symmetry generator takes the form $\mathbf{v}_2 = \partial_y$. This leads to $\mathbf{v}_2(s) = \mathbf{v}_2(u) = 0$ and $\mathbf{v}_2(y) = 1$, which can be satisfied by choosing $s = xt^2$, $y = \ln t$ and $u = vt^3$. Using these variables the above system is transformed to

$$\frac{dy}{ds} = \frac{1}{2s + u}$$

$$\frac{du}{ds} = \frac{u^2 - s^3 + 3su}{s\,(2s + u)}$$

As we can see, the second equation can be solved independently from the first one. And if we substitute the solution $u(s)$ in the first one, then $y(s)$ can be simply found by integration. Therefore, the order of the system has indeed been reduced, since it suffices to solve only one first-order equation to find the general solution of the system.

From the above discussion, it is naturally to assume that an $r$-parameter symmetry group can reduce the order of a system by $r$. But this is not always the case. Recall that for the differential equation of the previous example, we had found in Example 4.23 a two-parameter symmetry group. However, when we reduced the system in the previous example, we end up with another one, in which the information of the second symmetry had been lost.

The notion we need to avoid such cases is that of a *solvable* symmetry group. This guarantees that an $r$-parameter group of transformations splits into $r$ one-parameter transformation subgroups each of which is a normal subgroup of the other, and therefore acts independently with respect to the others. Avoiding a more formal definition, we straightforwardly proceed with the following criterion.

**Proposition 4.31.** An $r$-parameter Lie group is solvable if and only if its Lie algebra has a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ such that

$$[\mathbf{v}_i, \mathbf{v}_j] = \sum_{k=1}^{j-1} c_{ij}^k \mathbf{v}_k \qquad \text{for} \qquad i < j \tag{4.33}$$

Obviously every *abelian* Lie algebra, for which the Lie bracket of any two basis vectors is always zero, is solvable. This was the case of Theorem 4.12 stating that the actions of the corresponding one-parameter groups commute. We should note that every two-dimensional Lie algebra is also solvable. For solvable symmetry groups the following theorem holds; for proof see [89], p. 151 for the similar (though sub-)case of an $n$-th order ODE.

**Theorem 4.32.** Consider a system of $p$ first-order ordinary differential equations that admits an $r$-parameter solvable Lie group of symmetries, acting regularly. Then the order of the system can be reduced by $r$, resulting in a reduced system consisting of $(p - r)$ first-order equations. In particular, if $r = p$, then the general solution of the system can be found by quadratures alone.

The implementation of the reduction procedure, just like the theorem's proof, follows the specific order in which (4.33) is expressed. In other words, we start off the reduction of the system using $\mathbf{v}_1$, then $\mathbf{v}_2$, and we continue in this order until $\mathbf{v}_r$. Thus, following this order,

which may not be unique, the remaining symmetries in every step are still admitted by the reduced system.

**Example 4.33.** Let us revisit the system from Example (4.30) assuming now $x > 0$. We first calculate the Lie bracket of the two vector fields $\mathbf{v}_1 = \partial_t$ and $\mathbf{v}_2 = t\partial_t - 2x\partial_x - 3v\partial_v$ that generate the two-parameter symmetry group for the system,

$$[\mathbf{v}_1, \mathbf{v}_2] = \partial_t \left(t\partial_t - 2x\partial_x - 3v\partial_v\right) - \left(t\partial_t - 2x\partial_x - 3v\partial_v\right)\partial_t = \partial_t = \mathbf{v}_1$$

Therefore, we begin the reduction with $\mathbf{v}_1$, admitted by every autonomous system. This vector field is already in the desired form, so we retain the same variables, considering though $t = t(x)$ and $v = v(x)$ as functions of $x$. In this way, the system is expressed

$$\frac{dt}{dx} = \frac{1}{v}$$
$$\frac{dv}{dx} = \frac{v}{x} - \frac{x^2}{v}$$

whose order is smaller from the original by one. Continuing with the vector field $\mathbf{v}_2$, whose first prolongation with respect to these variables is $\mathbf{v}_2^{(1)} = -2x\partial_x + t\partial_t - 3v\partial_v + 3t'\partial_{t'} - v'\partial_{v'}$, we verify that it is still a symmetry for the system,

$$\mathbf{v}_2^{(1)}\left(t' - \frac{1}{v}\right) = 0 \qquad \text{when} \qquad t' - \frac{1}{v} = 0$$
$$\mathbf{v}_2^{(1)}\left(v' - \frac{v}{x} + \frac{x^2}{v}\right) = 0 \qquad \text{when} \qquad v' - \frac{v}{x} + \frac{x^2}{v} = 0$$

Making a new change of variables from $(x, v)$ to $(y, u)$, where $y = x^3/v^2$ and $u = -(\ln x)/2$, such that $\mathbf{v}_2(y) = 0$ and $\mathbf{v}_2(u) = 1$, the system finally takes the form

$$\frac{dt}{dy} = \frac{e^u}{\sqrt{y}\,(2y+1)}$$
$$\frac{du}{dy} = -\frac{1}{2y\,(2y+1)}$$

and can be solved by quadratures. So, using the solvable symmetry group in the prescribed order of the generators, we were able to reduce the system by two.

Now, we can find the solution of the system simply by integration. From the second equation we find

$$u = \ln\left(\frac{2y+1}{c_1^2 y}\right)^{\frac{1}{2}}$$

where $c_1$ is a real constant. Then, we substitute $u$ in the first one and integrate

$$t = \frac{1}{c_1}\ln\frac{\sqrt{2y+1}-1}{\sqrt{2y+1}+1} + c_2 \quad \Rightarrow \quad y = \frac{2e^{c_1(t-c_2)}}{\left(1 - e^{c_1(t-c_2)}\right)^2}$$

assuming $c_1 > 0$. Taking the inverse transformation, we have

$$x = e^{-2u} = c_1^2 \frac{y}{2y+1} = \frac{2c_1^2 e^{c_1(t-c_2)}}{\left(1+e^{c_1(t-c_2)}\right)^2}$$

and if we make the substitution $c_1 \to 2c_1$, we find that the solution of the system (meaning the original equation) is

$$x(t) = 2c_1^2 \operatorname{sech}^2(c_1(t-c_2))$$

Since sech is an even function so is $\operatorname{sech}^2$, and therefore $c_1$ can be arbitrary.

## 4.6 Invariant solutions of PDEs

In the treatment of partial differential equations, we are often forced to resort into some sort of ansatz, not being able to find the general solution. This is not as disappointing as it sounds, for these special types of solutions can give us information for the asymptotic or even the dominant behavior of the general solution, while in addition they could have great value from the physics point of view. A typical example is the search for travelling waves in wave equations, as well as the method commonly known as separation of variables. The symmetry analysis may allows us to determine from the beginning the form of some families of solutions admitted by the equation at hand.

Following the same route with the previous section, we consider now a nondegenerate system of partial differential equations

$$\Delta\left(x, u, u^{(n)}\right) = 0$$

and let $\mathbf{v}$ be the infinitesimal generator of a one-parameter symmetry group for the system. At each point $(x, u)$ where $\mathbf{v} \neq 0$ we can change to new variables $(y, w)$, such that

$$\mathbf{v} = \frac{\partial}{\partial y^1}$$

and consequently $\mathbf{v}^{(n)} = \mathbf{v}$. The symmetry condition for the transformed system

$$\widetilde{\Delta}\left(y, w, w^{(n)}\right) = 0$$

is now expressed as

$$\mathbf{v}^{(n)}(\widetilde{\Delta}_\nu) = 0 \quad \Rightarrow \quad \frac{\partial \widetilde{\Delta}_\nu}{\partial y^1} = 0$$

Since the variable $y^1$ does not directly appear in this system, we may assume that $w$ do not depend on it. Thus, we can rest assured that this ansatz will not prove wrong for the system in question.

In order to find the transformation of $\Delta$, we need to express the variables $u$ in terms of the new ones $y$ and $w$. In this expression there may be left some of the old independent variables $x$, which, hoewever, appear only as parameters and should not concern us. So, we have $u = u(x, y, w)$, and differentiation of this relation according to the chain rule

$$u^i_{x^j} = \frac{\partial u^i}{\partial x^j} + \frac{\partial u^i}{\partial y^m}\frac{\partial y^m}{\partial x^j} + \frac{\partial u^i}{\partial w^n}\frac{\partial w^n}{\partial y^m}\frac{\partial y^m}{\partial x^j}$$

results in a similar expression of the form $u_x = u_x(x, y, w, w_y)$, taking into account that $w_{y^1} = 0$. Proceeding with the higher derivatives, we substitute all the expressions found in the original system $\Delta$ and thus arrive at the form of the new one $\widetilde{\Delta}$.

**Example 4.34.** We have seen earlier that one of the symmetries of the KdV equation

$$u_t = u_{xxx} + uu_x$$

is generated by the vector field

$$\mathbf{v}_3 = t\partial_x - \partial_u$$

We choose new variables $(s, y, w) = (x/t, t, x + tu)$, such that $\mathbf{v}_3(s) = 1$, $\mathbf{v}_3(y) = \mathbf{v}_3(w) = 0$, where $w$ is considered as a function only of $y$. Then, solve $w = x + tu$ for $u$,

$$u = \frac{w - x}{y}$$

and calculate the derivatives

$$u_t = \frac{yw_y + x - w}{y^2}, \qquad u_x = -\frac{1}{y}, \qquad u_{xxx} = 0$$

Substitution in KdV, yields the equation

$$\frac{dw}{dy} = 0$$

and from its solution $w = c$, where $c$ is a real constant, we finally have

$$u(t, x) = \frac{c - x}{t}$$

We note, from the above example, too, that the variable $y^1$ that we eliminate does not enter the picture at all. Indeed, to make the transformation to the system $\widetilde{\Delta}$, it was sufficient to know all the rest $\left(y^2, \ldots, y^p, w^1, \ldots, w^q\right)$, for which the infinitesimal generator $\mathbf{v}$ of the symmetry vanishes,

$$\mathbf{v}(y^2) = \cdots = \mathbf{v}(y^p) = \mathbf{v}(w^1) = \cdots = \mathbf{v}(w^q) = 0$$

and then by the chain rule, to express the derivatives of the old variables $u$ in terms of the now ones $w$. These variables are called *invariants* of the symmetry group, since as we can see from equation (4.12) they remain constant under the group transformations.

If a second symmetry for the system exists, then to make a second ansatz, meaning to eliminate a second variable, the symmetry must be conserved in the transformed system. In other words, just as in the reduction of order for ODEs, an $r$-parameter symmetry group must be solvable in order to eliminate $r$ independent variables of the system of PDEs. Based on this requirement, using $p-1$ symmetries, where $p$ is the number of independent variables, we can arrive at a system of ordinary differential equations. And if $r \geq p$, then from the symmetries left we can continue as we have shown in the previous section by reducing the order of this system.

Instead of finding though the invariants $\left(y^2, \ldots, y^p, w^1, \ldots, w^q\right)$ of the first symmetry $\mathbf{v}_1$, then determine a second one $\mathbf{v}_2$ that satisfies the relation $[\mathbf{v}_1, \mathbf{v}_2] = c_{12}^1 \mathbf{v}_1$, and then find its invariants, $\left(\widetilde{y}^3, \ldots, \widetilde{y}^p, \widetilde{w}^1, \ldots, \widetilde{w}^q\right)$, and so on, we can find directly the invariants of the full symmetry group generated by the vector fields $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$. That is, instead of using the symmetries one by one, we can employ them all in one step, by finding directly functions $\psi(x, u)$ such that

$$c_1 \mathbf{v}_1(\psi) + \cdots + c_r \mathbf{v}_r(\psi) = 0 \tag{4.34}$$

After determining the invariants $\psi$ of the entire symmtery group from the above relation, we suitably separate them into independent and dependent variables for the new system, and then move on to the transformation of the derivatives just like before. Of course, if the dimension of the symmetry algebra found is greater than the number of the independent variables, as in the case of the KdV, then we can choose any subalgebra.

**Example 4.35.** Let us for example take the linear combination of the first two symmetry generators of the KdV,

$$\mathbf{v}_{12} = \partial_t + c\partial_x$$

giving rise to the transformation subgroup $G_{12}$

$$\left(\widetilde{x}, \widetilde{t}, \widetilde{u}\right) \longrightarrow (x + c\epsilon, t + \epsilon, u)$$

where $c$ is a real constant. Invariants of this group are $(y, w) = (x - ct, u)$. Note that by choosing $x - ct$ as the independent variable with $t$ playing the role of time and $x$ that of space, the function $u(x - ct)$ describes a disturbance propagating in the $x$-direction with velocity $c$, known as *travelling wave*. Substituting $u = w$, $u_t = -cw_y$, $u_x = w_y$ and $u_{xxx} = w_{yyy}$ in KdV,

we end up with the ordinary differential equation

$$cw_y + w_{yyy} + ww_y = 0$$

which can be directly integrated,

$$cw + w_{yy} + \frac{w^2}{2} + k = 0.$$

Multiplying with $w_y$ and integrating again yield

$$\frac{cw^2}{2} + \frac{w_y^2}{2} + \frac{w^3}{6} + kw + l = 0$$

where $k$ and $l$ are arbitrary constants. If we are looking for solutions that vanish sufficiently rapidly at infinity, then from the last two relations, we conclude $k = l = 0$. In this case, the above equation has the general solution

$$w = -3c\,\mathrm{sech}^2\left(\frac{\sqrt{c}}{2}\,y + b\right)$$

assuming that the velocity $c$ is positive, while $b$ is an arbitrary constant. Returning back to the original variables, we finally have

$$u(x,t) = -3c\,\mathrm{sech}^2\left(\frac{\sqrt{c}}{2}\,(x - c\,t) + b\right)$$

This type of solutions describe a special class of waves with remarkable properties that appear in nonlinear PDEs. These waves are called *solitons* and are a key characteristic of integrable systems.

## 4.7 Noether's theorem in Classical Mechanics

Many problems in physics describing basic theories from Classical Mechanics to Electromagnetism or even the General Theory of Relativity are coming from a variational principle. It is therefore of particular interest to study the role of symmetry in variational problems. The investigation of symmetries from this point of view was made by Emmy Noether, resulting in her famous theorem according to which there is a one-to-one correspondence between a certain subclass of symmetries of such systems and the conservation laws that they obey. It is worth mentioning that her contribution to symmetry theory does not stop here (see final notes at the end of this chapter).

In this section we will restrict our discussion strictly for ODEs and in particular second-order systems appearing in mechanics, where as usual the independent variable is denoted by $t$ and

the dependent ones collectively by $x = \left(x^1, \ldots, x^p\right)$ (the route for more general systems of either ordinary or partial differential equations follows a very similar path, but concrete results, e.g. the derivation of conservation laws, depend much on the order of the system). From calculus of variations, we borrow the notion of a functional $\mathfrak{L}$, which in this case is given by the definite integral of a differential function of the form $L\left(t, x, \dot{x}\right)$ over the time interval $t_2 - t_1$,

$$\mathfrak{L}\left[x\right] = \int_{t_1}^{t_2} L\left(t, x, \dot{x}\right) dt\,. \tag{4.35}$$

A *variational problem* consists then of determining the smooth functions $x(t)$, for which the integral $\mathfrak{L}$ attains maximum or minimum values. Just as the extrema of an ordinary function are found among its critical points, so too the extrema $x(t)$ of the functional $\mathfrak{L}$ are found among the critical "points" of $\mathfrak{L}$. The latter are the functions where the *variational derivative* $\delta\mathfrak{L}$ with components $\delta\mathfrak{L}/\delta x^i$, defined as

$$\int_{t_1}^{t_2} \frac{\delta\mathfrak{L}}{\delta x^i}\, y^i dt = \left.\frac{d\mathfrak{L}\left[x + \epsilon y\right]}{d\epsilon}\right|_{\epsilon=0} \tag{4.36}$$

where $y(t)$ is another smooth function defined on $t_2 - t_1$, vanishes. Interchanging the order of differentiation and integration, and then using the divergence theorem to integrate by parts,

$$\begin{aligned}\left.\frac{d\mathfrak{L}\left[x + \epsilon y\right]}{d\epsilon}\right|_{\epsilon=0} &= \int_{t_1}^{t_2} \left.\frac{dL\left(t, x + \epsilon y, \dot{x} + \epsilon\dot{y}\right)}{d\epsilon}\right|_{\epsilon=0} dt = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x^i}\, y^i + \frac{\partial L}{\partial \dot{x}^i}\, \dot{y}^i\right) dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\right] y^i\, dt\end{aligned}$$

we see that the *variational derivative of* $\mathfrak{L}$ *with respect to* $x^i$ in this case is given by the relation

$$\frac{\delta\mathfrak{L}}{\delta x^i} = \frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)$$

Therefore, the extrema $x(t)$ of the functional $\mathfrak{L}\left[x\right]$ (4.35) satisfy the so-called *Euler-Lagrange equations*

$$\mathrm{E}_i(L) = \frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = 0 \tag{4.37}$$

where $\mathrm{E}_i$ is known as the $i$-th *Euler operator*, while the function $L$ that characterizes the variational problem is called the *Langrangian* of the system. Clearly, for a first-order Lagrangian with one independent variable the Euler-Lagrange equations $\Delta \equiv \mathrm{E}(L) = 0$ are systems of second-order ODEs, i.e. $\Delta = \Delta(t, x, \dot{x}, \ddot{x})$. Newtonian systems with conservative forces are perhaps

one of the most popular cases of Euler-Lagrange systems. The system studied in this thesis coming from the Lorentz force is another more general example described in section 1.1.

If the Lagrangian function $L$ is the total derivative of another function $f$, meaning $L = df/dt$ (note that $f$ must be independent of $\dot{x}$ in order $L$ to be independent of $\ddot{x}$), then the corresponding functional $\mathfrak{L}[x] = f(t_2, x(t_2)) - f(t_1, x(t_1))$ depends only on the behavior of $f$ at the endpoints of the interval $t_2 - t_1$. Therefore, if $x(t)$ is an extremum of $\mathfrak{L}$, then any other function simply satisfying the same boundary conditions with $f$ will be an extremum, too. It is not difficult to see that the Euler-Lagrange equations in this case vanish identically for all $t$ and $x$,

$$\begin{aligned} \mathrm{E}_i\left(\frac{df}{dt}\right) &= \frac{\partial}{\partial x^i}\left(\frac{df}{dt}\right) - \frac{d}{dt}\left[\frac{\partial}{\partial \dot{x}^i}\left(\frac{df}{dt}\right)\right] = \frac{\partial}{\partial x^i}\left(\frac{df}{dt}\right) - \frac{d}{dt}\left[\frac{\partial}{\partial \dot{x}^i}\left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^j}\,\dot{x}^j\right)\right] \\ &= \frac{\partial}{\partial x^i}\left(\frac{df}{dt}\right) - \frac{d}{dt}\left(\frac{\partial f}{\partial x^i}\right) = 0 \end{aligned}$$

since the two operators $d/dt$ and $\partial/\partial x^i$ commute. Obviously this type of functions are of no interest, and in order to exclude them we will consider two different functionals $\mathfrak{L}_1[x]$ and $\mathfrak{L}_2[x]$ defined on the same interval $t_2 - t_1$ equivalent if and only if $L_1 = L_2 + df/dt$ for any differential function $f$.

Considering that every solution of the Euler-Lagrange equations (4.37) is not necessarily an extremal of the functional (4.35), it is natural to focus on the symmetries that leave the functional invariant and not the equations themselves.

**Definition 4.36.** Let $\mathfrak{L}[x]$ be a variational problem for the Lagrangian function $L(t, x, \dot{x})$. A local group $G$ of transformations $(t, x) \longrightarrow (\widetilde{t}, \widetilde{x})$ acting on the space $M$ of the independent and dependent variables is called a *variational* or *Noether symmetry group* of the system if $\mathfrak{L}[\widetilde{x}] = \mathfrak{L}[x]$, i.e. the functional of the system is invariant under the action of the group,

$$\int_{\widetilde{t}_1}^{\widetilde{t}_2} L\left(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}}\right) d\widetilde{t} = \int_{t_1}^{t_2} L(t, x, \dot{x})\, dt \tag{4.38}$$

As with Lie symmetries, the existence of a Noether symmetry group is independent of the choice of variables. The next theorem describes the analogous infinitesimal criterion for Noether point symmetries.

**Theorem 4.37.** Let $\mathfrak{L}[x]$ be a variational problem described by the Lagrangian function $L(t, x, \dot{x})$. A vector field $\mathbf{v} = \xi(t, x)\partial_t + \eta^i(t, x)\partial_{x^i}$ is a generator of a variational symmetry group for the system if and only if there exists a function $f(t, x)$, such that for all $t$ and $x$

$$\mathbf{v}^{(1)}(L) + L\,\frac{d\xi}{dt} = \frac{df}{dt} \tag{4.39}$$

*Proof.* Let $\mathbf{v}$ be a variational symmetry generator for the functional $\mathfrak{L}[x]$, and $\mathbf{v}^{(1)} = \mathbf{v} + \eta_i^t \partial_{\dot{x}^i}$ its first prolongation. Then Definition 4.36 implies that $\mathfrak{L}[x] = \mathfrak{L}[\widetilde{x}]$, and consequently the equality (4.38). The latter, if we parametrize the integral on the left hand side in terms of the original variable $t$, can be written as

$$\int_{t_1}^{t_2} L\left(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}}\right) \frac{d\widetilde{t}}{dt}\, dt = \int_{t_1}^{t_2} L\left(t, x, \dot{x}\right) dt$$

where $(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}}) = e^{\epsilon \mathbf{v}}(t, x, \dot{x})$ are given by the flow of $\mathbf{v}$. Differentiation with respect to $\epsilon$ on both sides,

$$\int_{t_1}^{t_2} \left[ \left( \frac{\partial L}{\partial \widetilde{t}} \frac{\partial \widetilde{t}}{\partial \epsilon} + \frac{\partial L}{\partial \widetilde{x}^i} \frac{\partial \widetilde{x}^i}{\partial \epsilon} + \frac{\partial L}{\partial \dot{\widetilde{x}}} \frac{\partial \dot{\widetilde{x}}}{\partial \epsilon} \right) \frac{d\widetilde{t}}{dt} + L \left( \frac{d\xi}{dt} + O(\epsilon) \right) \right] dt = 0$$

and then setting $\epsilon = 0$, we have

$$\int_{t_1}^{t_2} \left( \frac{\partial L}{\partial t}\, \xi + \frac{\partial L}{\partial x^i}\, \eta^i + \frac{\partial L}{\partial \dot{x}^i}\, \eta_i^t + L \frac{d\xi}{dt} \right) dt = 0 \qquad \text{or} \qquad \int_{t_1}^{t_2} \left( \mathbf{v}^{(1)}(L) + L\, \frac{d\xi}{dt} \right) dt = 0$$

In accordance with our equivalence relation among functionals, the last equation is satisfied when the integrand is equal to the total derivative of an arbitrary function $f(t, x)$, resulting in (4.39).

For the converse, let us first comment on the commutation relation between the total derivative $D_t = d/dt$ of a function and the prolongation $\mathbf{v}^{(n)}$ of a vector field $\mathbf{v}$. In the simplest case of $n = 1$, where $\mathbf{v}^{(1)} = \mathbf{v} + \eta_i^t \partial_{\dot{x}^i}$ and $\eta_i^t = D_t \eta^i - \dot{x}^i D_t \xi$, for every function of the form $g(t, x)$ we have

$$\begin{aligned}
\frac{d}{dt}\left(\mathbf{v}(g)\right) = \frac{d}{dt}\left(\xi g_t + \eta^i g_{x^i}\right) &= \frac{d\xi}{dt}\, g_t + \xi \frac{dg_t}{dt} + \frac{d\eta^i}{dt}\, g_{x^i} + \eta^i \frac{dg_{x^i}}{dt} \\
&= \xi \frac{\partial}{\partial t}\left(\frac{dg}{dt}\right) + \eta^i \frac{\partial}{\partial x^i}\left(\frac{dg}{dt}\right) + \left(\frac{d\eta^i}{dt} - \frac{d\xi}{dt}\, \dot{x}^i\right) g_{x^i} + \frac{d\xi}{dt}\, g_t + \frac{d\xi}{dt}\, \dot{x}^i g_{x^i} \\
&= \xi \frac{\partial}{\partial t}\left(\frac{dg}{dt}\right) + \eta^i \frac{\partial}{\partial x^i}\left(\frac{dg}{dt}\right) + \eta_i^t \frac{\partial}{\partial \dot{x}^i}\left(\frac{dg}{dt}\right) + \frac{d\xi}{dt}\left(g_t + g_{x^i} \dot{x}^i\right) \\
&= \mathbf{v}^{(1)}\left(\frac{dg}{dt}\right) + \frac{d\xi}{dt} \frac{dg}{dt}
\end{aligned}$$

or

$$D_t \mathbf{v} = \mathbf{v}^{(1)} D_t + D_t(\xi) D_t \tag{4.40}$$

Keeping in mind that both sides act on the same type of functions, we might as well write $\left[D_t, \mathbf{v}^{(1)}\right] = D_t(\xi) D_t$. In fact, this relation can be generalized for the prolongation of any order,

meaning $\left[D_t, \mathbf{v}^{(n)}\right] = D_t(\xi)D_t$ in the same sense, that is both sides are considered acting on functions of the form $g(t, x, \dot{x}, \ldots, x^{(n-1)})$. For $\xi = 0$, the flow of the vector field transforms only the dependent variables and the two operators commute. This kind of fields are known as evolutionary vector fields, and are particularly useful for several generalizations; see [89], sections 5.1 and 5.3 for more on this topic.

Now, suppose (4.39) holds. According to the Lie series (4.24) the expansion of the function $L(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}})$ for instance up to second-order terms is

$$L(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}}) = L\left(t, x, \dot{x}\right) + \epsilon \mathbf{v}^{(1)}(L)\left(t, x, \dot{x}\right) + \frac{\epsilon^2}{2}\mathbf{v}^{(1)}\Big(\mathbf{v}^{(1)}(L)\Big)\left(t, x, \dot{x}\right) + O\big(\epsilon^3\big)$$

while, since $\widetilde{t} = t + \epsilon\mathbf{v}(t) + \frac{\epsilon^2}{2}\mathbf{v}(\mathbf{v}(t)) + O\big(\epsilon^3\big) = t + \epsilon\xi + \frac{\epsilon^2}{2}\mathbf{v}(\xi) + O\big(\epsilon^3\big)$, in light of (4.40) we also have

$$\frac{d\widetilde{t}}{dt} = 1 + \epsilon\frac{d\xi}{dt} + \frac{\epsilon^2}{2}\frac{d}{dt}\left(\mathbf{v}(\xi)\right) + O\big(\epsilon^3\big) = 1 + \epsilon\frac{d\xi}{dt} + \frac{\epsilon^2}{2}\left[\mathbf{v}^{(1)}\left(\frac{d\xi}{dt}\right) + \left(\frac{d\xi}{dt}\right)^2\right] + O\big(\epsilon^3\big)$$

Therefore,

$$\begin{aligned}
L(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}})\frac{d\widetilde{t}}{dt} &= L + \epsilon\left[\mathbf{v}^{(1)}(L) + L\frac{d\xi}{dt}\right] + \\
&\quad + \frac{\epsilon^2}{2}\left\{\mathbf{v}^{(1)}(\mathbf{v}^{(1)}(L)) + 2\mathbf{v}^{(1)}(L)\frac{d\xi}{dt} + L\left[\mathbf{v}^{(1)}\left(\frac{d\xi}{dt}\right) + \left(\frac{d\xi}{dt}\right)^2\right]\right\} + O\big(\epsilon^3\big) \\
&= L + \epsilon\left[\mathbf{v}^{(1)}(L) + L\frac{d\xi}{dt}\right] + \\
&\quad + \frac{\epsilon^2}{2}\left\{\mathbf{v}^{(1)}\left(\mathbf{v}^{(1)}(L) + L\frac{d\xi}{dt}\right) + \frac{d\xi}{dt}\left[\mathbf{v}^{(1)}(L) + L\frac{d\xi}{dt}\right]\right\} + O\big(\epsilon^3\big)
\end{aligned}$$

where the last line follows from $\mathbf{v}^{(1)}(L)D_t\xi + L\mathbf{v}^{(1)}(D_t\xi) = \mathbf{v}^{(1)}(LD_t\xi)$, since the prolongation of a vector field $\mathbf{v}$ of any order satisfies the Leibniz' rule, too. Hence, according to our assumption, and using (4.40) once more we arrive at

$$L(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}})\frac{d\widetilde{t}}{dt} = L + \epsilon\frac{df}{dt} + \frac{\epsilon^2}{2}\left[\mathbf{v}^{(1)}\left(\frac{df}{dt}\right) + \frac{d\xi}{dt}\frac{df}{dt}\right] + O\big(\epsilon^3\big) = L + \epsilon\frac{df}{dt} + \frac{\epsilon^2}{2}\frac{d\left(\mathbf{v}(f)\right)}{dt} + O\big(\epsilon^3\big)$$

Continuing in this fashion, one can show that the terms of any order are given by the total derivative of some function. Thus, $L(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}})D_t\widetilde{t} = L(t, x, \dot{x}) + D_tF$ and so

$$\int_{\widetilde{t}_1}^{\widetilde{t}_2} L\left(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}}\right) d\widetilde{t} = \int_{t_1}^{t_2} L\left(\widetilde{t}, \widetilde{x}, \dot{\widetilde{x}}\right)\frac{d\widetilde{t}}{dt}\, dt = \int_{t_1}^{t_2} L\left(t, x, \dot{x}\right) dt$$

i.e. the vector field $\mathbf{v}$ is a variational symmetry generator for the functional $\mathfrak{L}[x]$. □

Note that while the infinitesimal criterion of Theorem 4.19 for the usual Lie symmetry groups has to hold only on solutions $x(t)$ of the system, the corresponding one (4.39) of Theorem 4.37 for Noether symmetry groups is an identity for all $t$ and $x$. Considering variational problems, a natural question posed is the relation between these two concepts, between these two groups. Leaving the functional $\mathfrak{L}[x]$ invariant and consequently the Euler-Lagrange equations, a Noether symmetry $(t,x) \longrightarrow (\widetilde{t},\widetilde{x})$ transforms extremals $x(t)$ of $\mathfrak{L}[x]$ to extremals $\widetilde{x}(\widetilde{t})$ of the same functional $\mathfrak{L}[\widetilde{x}] = \mathfrak{L}[x]$, and consequently solutions of the Euler-Lagrange equations to solutions of the same equations, accordingly. Therefore, by definition the next proposition holds; a computational proof is also given here in terms of the previously mentioned infinitesimal criterions.

**Proposition 4.38.** Let $\mathfrak{L}[x]$ be a variational problem described by the Lagrangian function $L(t,x,\dot{x})$. If $G$ is a variational symmetry group of the functional $\mathfrak{L}[x]$, then $G$ is a symmetry group of the corresponding Euler-Lagrange equations $\mathrm{E}(L) = 0$.

*Proof.* Let us first note that the Lie bracket[1] of the first prolongation $\mathbf{v}^{(1)} = \mathbf{v} + \eta_i^t \partial_{\dot{x}^i}$ of a vector field $\mathbf{v} = \xi\partial_t + \eta^i\partial_{x^i}$, where $\eta_i^t = D_t\eta^i - \dot{x}^i D_t\xi$, and the partial derivative $\partial/\partial x^i$ is

$$\left[\mathbf{v}^{(1)}, \frac{\partial}{\partial x^i}\right] = \left[\xi\frac{\partial}{\partial t} + \eta^j\frac{\partial}{\partial x^j} + \eta_j^t\frac{\partial}{\partial \dot{x}^j}, \frac{\partial}{\partial x^i}\right] = -\frac{\partial\xi}{\partial x^i}\frac{\partial}{\partial t} - \frac{\partial\eta^j}{\partial x^i}\frac{\partial}{\partial x^j} - \frac{\partial\eta_j^t}{\partial x^i}\frac{\partial}{\partial \dot{x}^j}$$

and therefore for any given function $L(t,x,\dot{x})$ the two operators can be interchanged as follows

$$\mathbf{v}^{(1)}\left(\frac{\partial L}{\partial x^i}\right) = \frac{\partial}{\partial x^i}\left(\mathbf{v}^{(1)}(L)\right) - \frac{\partial\xi}{\partial x^i}\frac{\partial L}{\partial t} - \frac{\partial\eta^j}{\partial x^i}\frac{\partial L}{\partial x^j} - \left(\frac{d\eta_{x^i}^j}{dt} - \frac{d\xi_{x^i}}{dt}\dot{x}^j\right)\frac{\partial L}{\partial \dot{x}^j} \tag{4.41}$$

Also the Lie bracket of $\mathbf{v}^{(1)}$ and the partial derivative $\partial/\partial\dot{x}^i$ is

$$\left[\mathbf{v}^{(1)}, \frac{\partial}{\partial \dot{x}^i}\right] = \left[\xi\frac{\partial}{\partial t} + \eta^j\frac{\partial}{\partial x^j} + \eta_j^t\frac{\partial}{\partial \dot{x}^j}, \frac{\partial}{\partial \dot{x}^i}\right] = -\frac{\partial\eta_j^t}{\partial \dot{x}^i}\frac{\partial}{\partial \dot{x}^j}$$

and hence

$$\begin{aligned}\mathbf{v}^{(1)}\left(\frac{\partial L}{\partial \dot{x}^i}\right) &= \frac{\partial}{\partial \dot{x}^i}\left(\mathbf{v}^{(1)}(L)\right) - \frac{\partial\eta_j^t}{\partial \dot{x}^i}\frac{\partial L}{\partial \dot{x}^j} = \frac{\partial}{\partial \dot{x}^i}\left(\mathbf{v}^{(1)}(L)\right) - \frac{\partial}{\partial \dot{x}^i}\left(\frac{d\eta^j}{dt} - \frac{d\xi}{dt}\dot{x}^j\right)\frac{\partial L}{\partial \dot{x}^j} \\ &= \frac{\partial}{\partial \dot{x}^i}\left(\mathbf{v}^{(1)}(L)\right) - \frac{\partial}{\partial \dot{x}^i}\left[\frac{\partial\eta^j}{\partial t} + \frac{\partial\eta^j}{\partial x^k}\dot{x}^k - \left(\frac{\partial\xi}{\partial t} + \frac{\partial\xi}{\partial x^k}\dot{x}^k\right)\dot{x}^j\right]\frac{\partial L}{\partial \dot{x}^j}\end{aligned}$$

1. See [89], p. 301 and in particular equation (5.23) therein for the definition of the Lie bracket of prolonged vector fields, and in general the so-called *generalized* vector fields whose components depend on the derivatives of the dependent variables, too. In the two Lie brackets considered here no complications arise and the commutator follows naturally. On the other hand, the case of the previous formula (4.40) would require all the technicalities of a formal definition and that is why it was rather presented in a more straightforward manner.

$$= \frac{\partial}{\partial \dot{x}^i}\left(\mathbf{v}^{(1)}(L)\right) - \left(\frac{\partial \eta^j}{\partial x^i} - \frac{\partial \xi}{\partial x^i}\,\dot{x}^j\right)\frac{\partial L}{\partial \dot{x}^j} + \frac{d\xi}{dt}\frac{\partial L}{\partial \dot{x}^i} \tag{4.42}$$

accordingly. Equations (4.41) and (4.42) are useful here to simplify the infinitesimal symmetry criterion to be investigated.

Now, let $\mathbf{v}$ be a generator of the variational symmetry group $G$. As previously mentioned, equation (4.40) generalizes to higher orders, e.g. $D_t\mathbf{v}^{(1)} = \mathbf{v}^{(2)}D_t + D_t(\xi)D_t$. Therefore,

$$\begin{aligned}\mathbf{v}^{(2)}(\mathrm{E}_i(L)) &= \mathbf{v}^{(1)}\left(\frac{\partial L}{\partial x^i}\right) - \mathbf{v}^{(2)}\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\right]\\ &= \mathbf{v}^{(1)}\left(\frac{\partial L}{\partial x^i}\right) - \frac{d}{dt}\left[\mathbf{v}^{(1)}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\right] + \frac{d\xi}{dt}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\end{aligned}$$

Then, using (4.41) and (4.42) we have:

$$\begin{aligned}\mathbf{v}^{(2)}(\mathrm{E}_i(L)) &= \frac{\partial}{\partial x^i}\left(\mathbf{v}^{(1)}(L)\right) - \frac{\partial \xi}{\partial x^i}\frac{\partial L}{\partial t} - \frac{\partial \eta^j}{\partial x^i}\frac{\partial L}{\partial x^j} - \left(\frac{d\eta^j_{x^i}}{dt} - \frac{d\xi_{x^i}}{dt}\,\dot{x}^j\right)\frac{\partial L}{\partial \dot{x}^j} -\\ &\quad - \frac{d}{dt}\left[\frac{\partial}{\partial \dot{x}^i}\left(\mathbf{v}^{(1)}(L)\right) - \left(\frac{\partial \eta^j}{\partial x^i} - \frac{\partial \xi}{\partial x^i}\,\dot{x}^j\right)\frac{\partial L}{\partial \dot{x}^j} + \frac{d\xi}{dt}\frac{\partial L}{\partial \dot{x}^i}\right] + \frac{d\xi}{dt}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\\ &= \mathrm{E}_i\left(\mathbf{v}^{(1)}(L)\right) - \frac{\partial \xi}{\partial x^i}\frac{\partial L}{\partial t} - \frac{\partial \eta^j}{\partial x^i}\frac{\partial L}{\partial x^j} + \frac{\partial \eta^j}{\partial x^i}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right) - \frac{\partial \xi}{\partial x^i}\,\ddot{x}^j\frac{\partial L}{\partial \dot{x}^j} -\\ &\quad - \frac{\partial \xi}{\partial x^i}\,\dot{x}^j\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right) - \frac{d^2\xi}{dt^2}\frac{\partial L}{\partial \dot{x}^i}\\ &= \mathrm{E}_i\left(\mathbf{v}^{(1)}(L) + L\frac{d\xi}{dt}\right) - \mathrm{E}_i\left(L\frac{d\xi}{dt}\right) - \frac{\partial \xi}{\partial x^i}\frac{\partial L}{\partial t} - \frac{\partial \eta^j}{\partial x^i}\,\mathrm{E}_i(L) - \frac{\partial \xi}{\partial x^i}\,\ddot{x}^j\frac{\partial L}{\partial \dot{x}^j} -\\ &\quad - \frac{\partial \xi}{\partial x^i}\,\dot{x}^j\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right) - \frac{d^2\xi}{dt^2}\frac{\partial L}{\partial \dot{x}^i}\end{aligned}$$

Now, recalling the infinitesimal criterion (4.39) for variational symmetries and that the Euler operator on total derivatives vanishes, the first term on the right-hand side of the above equation equals to zero. So,

$$\begin{aligned}\mathbf{v}^{(2)}(\mathrm{E}_i(L)) &= -\frac{d\xi}{dt}\frac{\partial L}{\partial x^i} - L\frac{d\xi_{x^i}}{dt} + \frac{d}{dt}\left[L\frac{\partial}{\partial \dot{x}^i}\left(\frac{d\xi}{dt}\right) + \frac{d\xi}{dt}\frac{\partial L}{\partial \dot{x}^i}\right] - \frac{\partial \xi}{\partial x^i}\frac{\partial L}{\partial t} - \frac{\partial \eta^j}{\partial x^i}\,\mathrm{E}_i(L) -\\ &\quad - \frac{\partial \xi}{\partial x^i}\,\ddot{x}^j\frac{\partial L}{\partial \dot{x}^j} - \frac{\partial \xi}{\partial x^i}\,\dot{x}^j\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right) - \frac{d^2\xi}{dt^2}\frac{\partial L}{\partial \dot{x}^i}\\ &= -\left(\frac{d\xi}{dt} + \frac{\partial \eta^j}{\partial x^i}\right)\mathrm{E}_i(L) + \frac{\partial \xi}{\partial x^i}\frac{dL}{dt} - \frac{\partial \xi}{\partial x^i}\frac{\partial L}{\partial t} - \frac{\partial \xi}{\partial x^i}\,\ddot{x}^j\frac{\partial L}{\partial \dot{x}^j} - \frac{\partial \xi}{\partial x^i}\,\dot{x}^j\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right)\\ &= -\left(\frac{d\xi}{dt} + \frac{\partial \eta^j}{\partial x^i}\right)\mathrm{E}_i(L) + \frac{\partial \xi}{\partial x^i}\left[\frac{\partial L}{\partial x^j}\,\dot{x}^j + \frac{\partial L}{\partial \dot{x}^j}\,\ddot{x}^j - \ddot{x}^j\frac{\partial L}{\partial \dot{x}^j} - \dot{x}^j\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right)\right]\\ &= -\left(\frac{d\xi}{dt} + \frac{\partial \eta^j}{\partial x^i}\right)\mathrm{E}_i(L) + \frac{\partial \xi}{\partial x^i}\,\dot{x}^j\,\mathrm{E}_j(L)\end{aligned}$$

from which we deduce that $\mathbf{v}^{(2)}(\mathrm{E}(L)) = 0$ when $\mathrm{E}(L) = 0$, hence $\mathbf{v}$ is also a symmetry generator of the Euler-Lagrange equations. □

The converse of this proposition is not true, since every solution of the Euler-Lagrange equations is not an extremal of the functional. Thus, the Noether symmetry group is a subgroup of the Lie symmetry group for Euler-Lagrange systems. The significance of variational symmetries starts with their connection with conservation laws established by the celebrated Noether's theorem, which is stated here in terms of classical mechanics.

**Theorem 4.39 (Noether).** Let $\mathfrak{L}[x]$ be a variational problem described by the Lagrangian function $L(t,x,\dot{x})$. Each variational symmetry generator $\mathbf{v}=\xi(t,x)\partial_t+\eta^i(t,x)\partial_{x^i}$ for the functional $\mathfrak{L}[x]$ corresponds to a first integral of motion $I(t,x,\dot{x})$ for the Euler-Lagrange equations $\mathrm{E}(L)=0$, given by the relation

$$I=\xi L+\left(\eta^i-\xi\dot{x}^i\right)\frac{\partial L}{\partial\dot{x}^i}-f \tag{4.43}$$

*Proof.* Let $\mathbf{v}$ be the generator of a variational symmetry, meaning $\mathbf{v}^{(1)}(L)+L\,D_t\xi=D_tf$, where $\mathbf{v}^{(1)}=\mathbf{v}+\eta^t_i\partial_{\dot{x}^i}$ and $\eta^t_i=D_t\eta^i-\dot{x}^iD_t\xi$. Straightforward calculations,

$$\begin{aligned}
0&=\mathbf{v}^{(1)}(L)+L\frac{d\xi}{dt}-\frac{df}{dt}=\xi\frac{\partial L}{\partial t}+\eta^i\frac{\partial L}{\partial x^i}+\left(\frac{d\eta^i}{dt}-\frac{d\xi}{dt}\,\dot{x}^i\right)\frac{\partial L}{\partial\dot{x}^i}+L\frac{d\xi}{dt}-\frac{df}{dt}\\
&=\frac{d(\xi L)}{dt}-\xi\frac{dL}{dt}+\xi\frac{\partial L}{\partial t}+\eta^i\frac{\partial L}{\partial x^i}+\frac{d\eta^i}{dt}\frac{\partial L}{\partial\dot{x}^i}-\frac{d\xi}{dt}\,\dot{x}^i\frac{\partial L}{\partial\dot{x}^i}-\frac{df}{dt}\\
&=\frac{d}{dt}\left(\xi L+\eta^i\frac{\partial L}{\partial\dot{x}^i}-\xi\dot{x}^i\frac{\partial L}{\partial\dot{x}^i}-f\right)-\eta^i\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^i}\right)+\xi\frac{d}{dt}\left(\dot{x}^i\frac{\partial L}{\partial\dot{x}^i}\right)-\xi\frac{\partial L}{\partial x^i}\,\dot{x}^i-\xi\frac{\partial L}{\partial\dot{x}^i}\ddot{x}^i+\eta^i\frac{\partial L}{\partial x^i}\\
&=\frac{dI}{dt}+\left(\eta^i-\xi\dot{x}^i\right)\left[\frac{\partial L}{\partial x^i}-\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^i}\right)\right],
\end{aligned}$$

show that this is this case if and only if the function $I$ is a constant of motion for the solutions of the Euler-Lagrange equations. □

The juxtaposition of Lie symmetries with Noether ones is also better revealed when it comes to reduction of order for ODEs. We need first though the following proposition (first presented by Lutzky) that as Sarlet and Cantrijn [94] point out is often neglected.

**Proposition 4.40.** The first integral of motion $I$ (4.43) corresponding to a variational symmetry generated by the vector field $\mathbf{v}$ is an invariant of its first prolongation, $\mathbf{v}^{(1)}(I)=0$.

*Proof.* Let $\mathbf{v}=\xi(t,x)\partial_t+\eta^i(t,x)\partial_{x^i}$ be the variational symmetry generator and $\mathbf{v}^{(1)}=\mathbf{v}+\eta^t_i\partial_{\dot{x}^i}$ its first prolongation, where $\eta^t_i=D_t\eta^i-\dot{x}^iD_t\xi$. Then, using firstly the Leibniz' rule and dropping the prolongation where not needed we have

$$\mathbf{v}^{(1)}(I)=L\mathbf{v}(\xi)+\xi\mathbf{v}^{(1)}(L)+\frac{\partial L}{\partial\dot{x}^i}\,\mathbf{v}^{(1)}\left(\eta^i-\xi\dot{x}^i\right)+\left(\eta^i-\xi\dot{x}^i\right)\mathbf{v}^{(1)}\left(\frac{\partial L}{\partial\dot{x}^i}\right)-\mathbf{v}(f)$$

From (4.39) and (4.42) first, and then separating terms with respect to $L$ and its derivatives, followed by rather straightforward calculations, we get

$$\begin{aligned}
\mathbf{v}^{(1)}(I) = & L\left(\xi\frac{\partial\xi}{\partial t}+\eta^i\frac{\partial\xi}{\partial x^i}\right)+\xi\left(\frac{df}{dt}-L\frac{d\xi}{dt}\right)+\left[\xi\frac{\partial\eta^i}{\partial t}+\eta^j\frac{\partial\eta^i}{\partial x^j}-\left(\xi\frac{\partial\xi}{\partial t}+\eta^j\frac{\partial\xi}{\partial x^j}\right)\dot{x}^i-\right.\\
& \left.-\xi\left(\frac{d\eta^i}{dt}-\frac{d\xi}{dt}\,\dot{x}^i\right)\right]\frac{\partial L}{\partial\dot{x}^i}+\left(\eta^i-\xi\dot{x}^i\right)\left[\frac{\partial}{\partial\dot{x}^i}\left(\frac{df}{dt}-L\frac{d\xi}{dt}\right)-\left(\frac{\partial\eta^j}{\partial x^i}-\frac{\partial\xi}{\partial x^i}\,\dot{x}^j\right)\frac{\partial L}{\partial\dot{x}^j}+\right.\\
& \left.+\frac{d\xi}{dt}\frac{\partial L}{\partial\dot{x}^i}\right]-\xi\frac{\partial f}{\partial t}-\eta^i\frac{\partial f}{\partial x^i}\\
= & -\left[\xi\frac{\partial f}{\partial t}+\eta^i\frac{\partial f}{\partial x^i}-\xi\frac{df}{dt}-\left(\eta^i-\xi\dot{x}^i\right)\frac{\partial}{\partial\dot{x}^i}\left(\frac{df}{dt}\right)\right]+\\
& +\left[\xi\frac{\partial\xi}{\partial t}+\eta^i\frac{\partial\xi}{\partial x^i}-\xi\frac{d\xi}{dt}-\left(\eta^i-\xi\dot{x}^i\right)\frac{\partial}{\partial\dot{x}^i}\left(\frac{d\xi}{dt}\right)\right]L+\\
& +\left[-\xi\dot{x}^j\frac{\partial\eta^i}{\partial x^j}+\eta^j\frac{\partial\eta^i}{\partial x^j}+\xi\dot{x}^i\dot{x}^j\frac{\partial\xi}{\partial x^j}+\eta^j\frac{\partial\xi}{\partial x^j}\,\dot{x}^i-\left(\eta^j-\xi\dot{x}^j\right)\left(\frac{\partial\eta^i}{\partial x^j}-\frac{\partial\xi}{\partial x^j}\,\dot{x}^i\right)\right]\frac{\partial L}{\partial\dot{x}^j}\\
= & -\left[\eta^i\frac{\partial f}{\partial x^i}-\xi\dot{x}^i\frac{\partial f}{\partial x^i}-\left(\eta^i-\xi\dot{x}^i\right)\frac{\partial f}{\partial x^i}\right]+\left[\eta^i\frac{\partial\xi}{\partial x^i}-\xi\dot{x}^i\frac{\partial\xi}{\partial x^i}-\left(\eta^i-\xi\dot{x}^i\right)\frac{\partial\xi}{\partial x^i}\right]L=0
\end{aligned}$$

□

Compared now to Theorem 4.27, the next theorem unravels the power of Noether symmetries stemming from the variational structure.

**Theorem 4.41.** The existence of a one-parameter variational symmetry group for an Euler-Lagrange system of ordinary differential equations with Lagrangian $L(t,x,\dot{x})$ reduces the order of the system by two.

*Proof.* Let $\mathbf{v}$ be the generator of a one-parameter variational symmetry group for the Euler-Lagrange system $\Delta\left(t,x,\dot{x},\ddot{x}\right)=0$, $\Delta=\mathrm{E}(L)$ of order $2p$. Following our usual technique, in every point $(t,x)$, where $\mathbf{v}\neq 0$, we can change to new coordinates $(s,y)$, choosing $s$ as the new independent variable and $y=\left(y^1,\ldots,y^p\right)$ the new dependent ones, such that $\mathbf{v}=\partial_{y^1}$. The system in these coordinates takes the form $\widetilde{\Delta}\left(s,y,y',y''\right)=0$, $\widetilde{\Delta}=\mathrm{E}(\widetilde{L})$, where $\widetilde{L}(s,y,y')$ is the transformed Lagrangian, while $\mathbf{v}^{(2)}=\mathbf{v}^{(1)}=\mathbf{v}$. By Proposition 4.38, $\mathbf{v}$ is also a generator of a one-parameter symmetry group for the system, too, and the symmetry condition reads

$$\frac{\partial\widetilde{\Delta}}{\partial y^1}=0$$

meaning the equations of the system do not depend on the variable $y^1$. On the other hand, the first equation of the system, using the variational symmetry condition, now expressed as $\partial\widetilde{L}/\partial y^1=df/ds$, easily yields the first integral $I=\partial\widetilde{L}/\partial y^{1'}-f$, which is none other than

(4.43). Furthermore Proposition 4.40 guarantees that $\mathbf{v}(I) = 0$ and hence $I(s, y, y')$ is also independent of $y^1$. Therefore, we can solve $I(s, y, y') = \text{const.}$ for $y^{1'}$ and replace the solution

$$y^{1'} = F_1(s, y^2, \dots, y^p, y^{2'}, \dots, y^{p'})$$

in the remaining $p-1$ equations. Consequently the latter will be independent of either $y^1$ or $y^{1'}$ forming a system of order $2p-2$,

$$\widetilde{\Delta}_i(s, y^2, \dots, y^p, y^{2'}, \dots, y^{p'}) = 0, \qquad i = 2, \dots, p$$

which can be solved independently of $y^1$, and once their solution $\big(y^2(s), \dots, y^p(s)\big)$ is found, then $y^1(s)$ can be determined by simply integrating

$$y^1(s) = \int F_1(s, y^2(s), \dots, y^p(s), y^{2'}(s), \dots, y^{p'}(s))ds$$

□

One can also prove that the reduced system is an Euler-Lagrange system subject to a constraint (see [89], p. 258 for this treatment and [24], p. 216 for variational problems with conditions). Then, one naturally asks what happens in the case of a second variational symmetry. The answer is that unless an $r$-parameter variational symmetry group is abelian, there is no guarantee that the order of the system can be reduced by $2r$. This is the counterpart of the well-known *Arnold-Liouville's theorem of complete integrability* for finite Hamiltonian systems, according to which if a $2p$-dimensional Hamiltonian system admits $p$ integrals of motion in involution, then its solutions can be found by quadratures alone. The analogy between these two statements is provided by the one-to-one correspondence between the Lie algebra of Hamiltonian symmetries (these are variational symmetries for Hamiltonian systems for which a Hamiltonian version of Noether's theorem holds; see [89]) and the Poisson algebra of integrals of motion, having the same structure constants.

Much more can be said about the reduction of either Euler-Lagrange or Hamiltonian systems with all their geometrical artillery. This, however, lies outside the purpose of this thesis and thus we refer to the vast literature. Here we close this discussion by an illustrative example from charged particle motion too, the motion in a *magnetic dipole*, also known as the *Störmer problem*.

**Example 4.42.** Consider a charged particle moving in zero electric field and a magnetic field of the form $\boldsymbol{B} = \left(3xz, 3yz, 2z^2 - x^2 - y^2\right)/r^5$, where $r = \sqrt{x^2 + y^2 + z^2}$ is only introduced for

brevity. The equations of motion (1.1) read

$$\begin{aligned}\ddot{x} &= \frac{\left(2z^2-x^2-y^2\right)\dot{y}-3yz\dot{z}}{\left(x^2+y^2+z^2\right)^{\frac{5}{2}}}\\ \ddot{y} &= -\frac{\left(2z^2-x^2-y^2\right)\dot{x}-3xz\dot{z}}{\left(x^2+y^2+z^2\right)^{\frac{5}{2}}}\\ \ddot{z} &= -\frac{3z\left(x\dot{y}-y\dot{x}\right)}{\left(x^2+y^2+z^2\right)^{\frac{5}{2}}}\end{aligned} \tag{4.44}$$

The magnetic field is coming from the vector potential $\boldsymbol{A} = (-y, x, 0)/r^3$, and according to relation (1.7), the Lagrangian function of this problem is

$$L(x,y,z,\dot{x},\dot{y},\dot{z}) = \frac{1}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right) + \frac{x\dot{y}-y\dot{x}}{\left(x^2+y^2+z^2\right)^{\frac{3}{2}}} \tag{4.45}$$

As we can see, the vector field $\mathbf{v} = -y\partial_x + x\partial_y$, whose first prolongation is $\mathbf{v}^{(1)} = \mathbf{v} - \dot{y}\partial_{\dot{x}} + \dot{x}\partial_{\dot{y}}$, generates a variational symmetry,

$$\mathbf{v}^{(1)}(L) = -\dot{y}\left(\dot{x}-\frac{y}{r^3}\right)+\dot{x}\left(\dot{y}+\frac{x}{r^3}\right) - y\,\frac{\dot{y}r^3-(\dot{x}y-y\dot{x})3rx}{r^6} - x\,\frac{\dot{x}r^3+(\dot{x}y-y\dot{x})3ry}{r^6} = 0$$

satisfying criterion (4.39) for $f = 0$. Therefore, it corresponds to the first integral of motion (4.43) $I = -y\partial L/\partial\dot{x} + x\partial L/\partial\dot{y}$ or after substitution of the Langrangian,

$$I = x\dot{y} - y\dot{x} + \frac{x^2+y^2}{\left(x^2+y^2+z^2\right)^{\frac{3}{2}}} \tag{4.46}$$

Let us now follow the steps of the proof of Theorem 4.41, and change first to polar coordinates $(\rho,\phi,z)$ in which $\mathbf{v} = \partial_\phi$ and $I = \rho^2(\dot{\phi}+r^{-3})$. The equations of motion (4.44) in the new variables are $\ddot{\rho} = \rho\dot{\phi}\left[\dot{\phi}+(2z^2-\rho^2)r^{-5}\right]$ and $\ddot{z} = -3zr^{-5}\rho^2\dot{\phi}$, while the third one simply reduces to the invariance of $I$. By solving $I = c$, where $c$ is a constant, for $\dot{\phi}$, the latter yields

$$\dot{\phi} = \frac{c}{\rho^2} - \frac{1}{\left(\rho^2+z^2\right)^{\frac{3}{2}}} \tag{4.47}$$

As promised, all equations are independent of $\phi$, and taking one step further, we can use the above relation (4.47) to eliminate $\dot{\phi}$ as well from the first two,

$$\begin{aligned}\ddot{\rho} &= \frac{c^2}{\rho^3} - \frac{3c\rho}{\left(\rho^2+z^2\right)^{\frac{5}{2}}} + \rho\,\frac{2\rho^2-z^2}{\left(\rho^2+z^2\right)^4}\\ \ddot{z} &= -\frac{3cz}{\left(\rho^2+z^2\right)^{\frac{5}{2}}} + \frac{3z\rho^2}{\left(\rho^2+z^2\right)^4}\end{aligned} \tag{4.48}$$

Therefore, the original system of sixth order has been reduced by two, since if we solve the fourth-order one (4.48) for $\rho(t)$ and $z(t)$, then $\phi(t)$ can be found by integrating (4.47). Last but not least the reduced system (4.48) can be recovered from the Lagrangian

$$L(\rho, z, \dot{\rho}, \dot{z}) = \frac{1}{2}\left(\dot{\rho}^2 + \dot{z}^2\right) + \frac{1}{2}\left[\frac{c^2}{\rho^2} - \frac{\rho^2}{\left(\rho^2 + z^2\right)^3}\right] \tag{4.49}$$

The latter is coming from the original Lagrangian function expressed in terms of the new coordinates. Due to the symmetry $\partial_\phi$, it does not depend on $\phi$, while the dependence in $\dot{\phi}$ was also removed using (4.47). Note however that $c$ in (4.49) is not to be treated as a constant, but as the function $I(\rho, z; \dot{\phi})$, which only after extracting the Euler-Lagrange equations can be replaced by $I = c$.

## 4.8 Equivalent differential equations

In many circumstances, arbitrary elements, such as functions or parameters, appear in a differential equation, in which case we speak of a family of differential equations, and it is important to know for which of these elements, i.e. for which members of the family, if and which symmetries appear. Repeating Ovsiannikov [90], the correspondence between differential equations and symmetry groups is not one-to-one : two differential equations may have the same symmetry group, while there is only one symmetry group admitted by a differential equation. Therefore, the answer to the above question is preferably given by a classification (of members of a family) of differential equations in terms of their symmetry groups rather than the other way round.

As previously implied, this can help us identify different members of a class of differential equations admitting the same symmetry groups. As a result, we may be able to bring complicated-looking equations to a much simpler form (for instance it could be a linear one if they have the "right" type of symmetry, as was pointed out in Remark 4.25), and map solutions of one differential equation to another. For example it has been proven [79] that a second-order ODE has 0, 1, 2, 3 or 8 symmetries, and that it is linear or linearizable by a point transformation if and only if it has 8 symmetries, or that for $n \geq 3$ a linear or linearizable $n$-th order ODE has $n+1$, $n+2$ or $n+4$ symmetries [80]. All the same symmetry classification of differential equations is not just of theoretical interest, since many problems in Mathematical Physics provide us with a variety of experimental data that allows some arbitrariness of parameters or functions involved, from which we can choose the ones with desirable analytical properties.

This line of work was also initiated by Lie himself, but more extensively investigated by Cartan and Tresse formulating the so-called equivalence problem geometrically. Ovsiannikov [90]

further exploited the group classification problem in a context very similar to Lie's framework for symmetry transformations, by defining transformations that leave invariant an entire class of differential equations. Then later on Ibragimov and coworkers first presented the analogous infinitesimal version with numerous applications [51, 53, 54]. A comprehensive treatment of this theory carefully inspecting and working out in detail many of its aspects was given by Lisle [75]. For a first encounter we refer to a manuscript of Ibragimov included in [53], p. 67 and illustrative examples therein, and for further reading Lisle's thesis for formal and complete definitions, constructions and proofs of theorems, written in an articulate way, too. Here we content ourselves in a brief description of the basic concepts, key points for direct application, generalized almost directly from symmetry theory.

In order to investigate relations between differential equations, first we must make precise and define properly the notion of a class of differential equations. We return to our general notation of section 4.1, meaning $x = \left(x^1, \dots, x^q\right)$ and $u = f(x) = \left(u^1, \dots, u^p\right)$ are the independent and dependent variables, respectively. Recall the corresponding spaces $X$ and $U$, their Cartesian product $M = X \times U$ and the $n$-jet space $M^{(n)}$ of $x$, $u$ and the derivatives $u^{(n)}$ of $u$ with respect to $x$ up to $n$-th order. Now, in addition, we also consider the space $A = \mathbb{R}^r$ of arbitrary functions $a = g(x, u) = \left(a^1, \dots, a^r\right)$ of $x$ and $u$, whose coordinates are given by $a^h, h = 1, \dots, r$, and accordingly define $N = M \times A$ and the $s$-jet space $N^{(s)}$ of $(x, u)$, $a$ and the derivatives of $a$ with respect to either $x$ or $u$ up to $s$-th order denoted as $a^{(s)}$. Note that for the extension $M^{(n)}$ $u$ are treated as the dependent variables, but for the extension $N^{(s)}$ as independent ones. Let $\Delta = (\Delta_1, \dots, \Delta_l)$ be differential functions on $M^{(n)} \times N^{(s)}$ and $\delta = (\delta_1, \dots, \delta_\lambda)$ differential functions on $N^{(s)}$. In general, a *class $\mathcal{C}$ of differential equations* is given by two decoupled systems of equations

$$\Delta\left(x, u, u^{(n)}; a, a^{(s)}\right) = 0 \tag{4.50a}$$

$$\delta\left(x, u, a, a^{(s)}\right) = 0 \tag{4.50b}$$

where the first one is the *primary system* of interest and the second one is known as the *auxiliary system*. The order $s$ of the highest derivatives of $a$ appearing in the two sets of equations need not to be the same, but for simplicity we assume that $s$ is the maximum between the two sets. In practice, the auxiliary system expresses conditions that the arbitrary functions $a$ entering system (4.50a) of primary concern must obey. These could express any restrictions involved in a physical problem. In the absence of physical arguments, they can simply describe the functional dependence of $a$ in terms of $x$ and $u$.

**Example 4.43.** The Fokker-Planck equation for time-independent friction and diffusion coefficients $b$ and $a$, respectively,

$$u_t = -\left(b(x)u\right)_x + \left(a(x)u\right)_{xx}$$

viewed as the primary system of a class of differential equations with arbitrary functions $a$ and $b$, can be enlraged by the auxiliary equations

$$a_t = a_u = b_t = b_u = 0$$

We emphasize that the two sets forming a class of differential equations are decoupled (even when $u$ are present in the auxiliary system). In fact, one has to solve first the auxiliary system and then insert its solution into the primary one. For every solution $a = g(x,u)$ of the auxiliary system, the primary system corresponds to a *member* $\mathcal{C}(g)$ *of the class*,

$$\Delta\left(x, u, u^{(n)}, g(x,u), g^{(s)}(x,u)\right) = 0 \tag{4.51}$$

A function $u = f(x)$ is a solution of $\mathcal{C}(g)$ if $g$ satisfies (4.50b) and $f$ satisfies (4.51). This means that a solution of a system of differential equations in some class needs two functions to be specified: $g$ to determine which system and $f$ to determine which solution.

Consider now all previous notions of vector fields, flows, local group of transformations, prolongations on the space $M \times A$. Starting with point transformations $(x,u,a) \longrightarrow (\widetilde{x},\widetilde{u},\widetilde{a})$, we particularly focus on *augmented transformations*,

$$\begin{aligned} \widetilde{x} &= \widetilde{x}\,(x,u) \\ \widetilde{u} &= \widetilde{u}\,(x,u) \\ \widetilde{a} &= \widetilde{a}\,(x,u,a) \end{aligned} \tag{4.52}$$

that is augmented by point transformations $(x,u) \longrightarrow (\widetilde{x},\widetilde{u})$ on $M$. A vector field $\mathbf{V}$ defined on $M \times A$ generating an augmented transformation is expressed as

$$\mathbf{V} = \xi^j(x,u)\frac{\partial}{\partial x^j} + \eta^i(x,u)\frac{\partial}{\partial u^i} + \mu^h(x,u,a)\frac{\partial}{\partial a^h}$$

and its flow describes a one-parameter group of augmented transformations, which in turn is expressed via exponentiation of $\mathbf{V}$

$$\begin{aligned} \widetilde{x} &= e^{\epsilon\mathbf{V}}(x) \\ \widetilde{u} &= e^{\epsilon\mathbf{V}}(u) \\ \widetilde{a} &= e^{\epsilon\mathbf{V}}(a) \end{aligned} \tag{4.53}$$

Using its projection to $M$, meaning the vector field $\mathbf{v} = \xi^j(x,u)\partial_{x^j} + \eta^i(x,u)\partial_{u^i}$, we can also write

$$\mathbf{V} = \mathbf{v} + \mu^h(x,u,a)\frac{\partial}{\partial a^h}$$

Next we want to see how this transformation must be extended in order to describe transformations of a class $\mathcal{C}$ of differential equations (4.50). Remember that two quite different space extensions were needed to define $\mathcal{C}$, the jet spaces $M^{(n)}$ and $N^{(s)}$. We begin with the auxiliary system for which the $s$-th prolongation of $\mathbf{V}$ is required regarding $x$ and $u$ as independent variables and $a$ as the dependent ones, $\mathbf{V}^{(s)} = \mathbf{V} + \mu^Z_h(x,u,a,a^{(s)})\partial_{a^h_Z}$. In order then to apply to the primary system we need the usual $n$-th prolonagtion of $\mathbf{v}$ regarding $x$ as the independent variables and $u$ as the dependent ones, $\mathbf{v} + \eta^J_i(x,u,u^{(n)})\partial_{u^i_J}$. Putting the two together, the $(n,s)$-th prolongation of $\mathbf{V}$ on the corresponding space $M^{(n)} \times N^{(s)}$ is given by the general formula

$$\mathbf{V}^{(n,s)} = \mathbf{V} + \eta^J_i(x,u,u^{(n)})\frac{\partial}{\partial u^i_J} + \mu^Z_h(x,u,a,a^{(s)})\frac{\partial}{\partial a^h_Z} \tag{4.54}$$

where $a^h_Z$ is the $\kappa$-th derivative of $a^h$ with respect to $\kappa$ of the variables $(x,u)$ denoted by the multi-index $Z = \left(x^{j_1},\dots,x^{j_{\kappa_1}},u^{i_1},\dots,u^{i_{\kappa_2}}\right)$ for $\kappa_1+\kappa_2 = \kappa$, and $\mu^Z_h$ is given by the same formula as $\eta^J_i$ only carefully keeping track the different roles of independent and dependet variables,

$$\mu^Z_h(x,u,a,a^{(s)}) = D_Z\left(a^h - \xi^j\frac{\partial a^h}{\partial x^j} - \eta^i\frac{\partial a^h}{\partial u^i}\right) + \xi^j\frac{\partial a^h_Z}{\partial x^j} + \eta^i\frac{\partial a^h_Z}{\partial u^i}$$

Here $D_Z = D_{j_1}D_{j_2}\cdots D_{j_{\kappa_1}}D_{i_1}D_{i_2}\cdots D_{i_{\kappa_2}}$, where $D_j$ is the total derivative with respect to $x^j$ and $D_i$ the total derivative with respect to $u^i$, forgetting at this point the dependence of $u$ in terms of $x$. Note that (4.54) can also be written as

$$\mathbf{V}^{(n,s)} = \mathbf{v}^{(n)} + \mu^h(x,u,a)\frac{\partial}{\partial a^h} + \mu^Z_h(x,u,a,a^{(s)})\frac{\partial}{\partial a^h_Z} \tag{4.55}$$

Recursion relations similar to (4.26)-(4.27) can also be given, but we rather see an example instead to get the hang of it, from which it will also be clear that $\mu^Z_h$ for augmented transformations are not so difficult to calculate due to the $a$-independence of $\xi$ and $\eta$.

**Example 4.44.** Consider for example that $(t,x) \in \mathbb{R}^2$ are the independent and $u \in \mathbb{R}$ the dependent variables for the primary system, while $a \in \mathbb{R}$ is the arbitrary function. Let

$$\mathbf{V} = \tau(t,x,u)\frac{\partial}{\partial t} + \xi(t,x,u)\frac{\partial}{\partial x} + \eta(t,x,u)\frac{\partial}{\partial u} + \mu(t,x,u,a)\frac{\partial}{\partial a}$$

From formula (4.55)

$$\mathbf{V}^{(1,1)} = \mathbf{v}^{(1)} + \mu\frac{\partial}{\partial a} + \mu^t\frac{\partial}{\partial a_t} + \mu^x\frac{\partial}{\partial a_x} + \mu^u\frac{\partial}{\partial a_u}$$

where $\mathbf{v} = \tau\partial_t + \xi\partial_x + \eta\partial_u$ and its first prolongation $\mathbf{v}^{(1)}$ is given in Example 4.16, while

$$\begin{aligned}
\mu^t &= \frac{d\mu}{dt} - \frac{d\tau}{dt}\,a_t - \frac{d\xi}{dt}\,a_x - \frac{d\eta}{dt}\,a_u = \mu_t + (\mu_a - \tau_t)\,a_t - \xi_t a_x - \eta_t a_u \\
\mu^x &= \frac{d\mu}{dx} - \frac{d\tau}{dx}\,a_t - \frac{d\xi}{dx}\,a_x - \frac{d\eta}{dx}\,a_u = \mu_x + (\mu_a - \xi_x)\,a_x - \tau_x a_t - \eta_x a_u \\
\mu^u &= \frac{d\mu}{du} - \frac{d\tau}{du}\,a_t - \frac{d\xi}{du}\,a_x - \frac{d\eta}{du}\,a_u = \mu_u + (\mu_a - \eta_u)\,a_u - \tau_u a_t - \xi_u a_x
\end{aligned}$$

**Definition 4.45.** Let $\mathcal{C}$ be a class of differential equations consisting of the primary system $\Delta\left(x, u, u^{(n)}; a, a^{(s)}\right) = 0$ and the auxiliary one $\delta\left(x, u, a, a^{(s)}\right) = 0$. A local group $G$ of augmented transformations $(x, u, a) \longrightarrow (\widetilde{x}, \widetilde{u}, \widetilde{a})$ acting on the space of the independent, dependent variables and the arbitrary functions for the primary system with the property that if $u = f(x)$ is a solution of the system $\mathcal{C}(g)$, then $\widetilde{u} = \widetilde{f}\,(\widetilde{x})$ is a solution of the system $\mathcal{C}(\widetilde{g})$ of the same class is called an *equivalence group* for the class. Equivalently $G$ is an equivalence group of $\mathcal{C}$ if it is a symmetry group for the auxiliary system and $\Delta\left(x, u, u^{(n)}; a, a^{(s)}\right) = 0 \Rightarrow \Delta\left(\widetilde{x}, \widetilde{u}, \widetilde{u}^{(n)}; \widetilde{a}, \widetilde{a}^{(s)}\right) = 0$, i.e. the equations of the class are invariant under the action of the group.

**Example 4.46.** Consider again the class for the Fokker-Planck equation of Example 4.43 and in particular for $b = a'$, which results to

$$\begin{aligned}
u_t &= (a u_x)_x \\
a_t &= a_u = 0
\end{aligned} \tag{4.56}$$

Obviously solutions of the auxiliary system are $a = g(x)$. For any parameter $\epsilon$, we can easily see that $(\widetilde{t}, \widetilde{a}) = (\epsilon t, a/\epsilon)$ is an equivalence transformation: if $a = g(x)$ is a solution of $a_t = a_u = 0$ then so is $\widetilde{a} = g(x)/\epsilon$, and $u_t = (a u_x)_x$ divided by $\epsilon$ results in $u_{\widetilde{t}} = (\widetilde{a} u_x)_x$.

An intriguing example given in [75] shows that *equivalence transformations for Hamilton's equations in canonical form (C.1) are actually canonical transformations.*

According to the previous definition, an equivalence transformation maps a member of a class of differential equations to another one in the same class. As with the symmetry definition, this one, too, is in most cases not practical to employ. Therefore, a corresponding infinitesimal criterion is of use given by the next theorem. Like Theorem 4.19, the converse of the next one holds only under certain nondegeneracy conditions that ensure the complete group of equivalence transformations (in the above sense of point augmented transformations) is found. Again, these are the maximal rank and the local solvability conditions, now referred to both the auxiliary

system and the class of differential equations as a whole; see [75], section 3.2 and subsection 3.3.2 for more details and proof of the next theorem, too, which more or less follows a scheme very close to that of Theorem 4.19.

**Theorem 4.47.** Let $\mathcal{C}$ be a nondegenerate class of differential equations consisting of the primary system $\Delta\left(x,u,u^{(n)};a,a^{(s)}\right)=0$ and the auxiliary one $\delta\left(x,u,a,a^{(s)}\right)=0$. A vector field $\mathbf{V}$ is a generator of an equivalence group for the class if and only if $\mathbf{V}^{(n,s)}(\delta)=\mathbf{V}^{(n,s)}(\Delta)=0$ whenever $\delta=\Delta=0$.

**Example 4.48.** The generator of the equivalence transformation $(\widetilde{t},\widetilde{a})=(\epsilon t,a/\epsilon)$ for the class of differential equations (4.56) we saw in the previous example is $\mathbf{V}=t\partial_t-a\partial_a$ and its $(2,1)$-th prolongation (leaving out the components of $\partial/\partial u_{tt}$ and $\partial/\partial u_{tx}$ which are not needed) is

$$\mathbf{V}^{(2,1)}=t\frac{\partial}{\partial t}-a\frac{\partial}{\partial a}-u_t\frac{\partial}{\partial u_t}-2a_t\frac{\partial}{\partial a_t}-a_x\frac{\partial}{\partial a_x}-a_u\frac{\partial}{\partial a_u}$$

as recovered from Examples 4.16 and 4.44. As we can see, starting from the auxiliary system and then continuing with the primary one, $\mathbf{V}$ satisfies the infinitesimal equivalence criterion,

$$\mathbf{V}^{(2,1)}(a_t)=-2a_t=0,\qquad \mathbf{V}^{(2,1)}(a_u)=-2a_u=0\qquad \text{when}\qquad a_t=a_u=0$$
$$\mathbf{V}^{(2,1)}(u_t-a_xu_x-au_{xx})=-u_t+a_xu_x+au_{xx}=0\qquad \text{when}\qquad a_t=a_u=u_t-a_xu_x-au_{xx}=0$$

**Example 4.49.** Let us go back to section 1.5 and see how the two systems of charged particle motion of §1.5.1 and §1.5.2 are related. Consider a homogeneous electromagnetic field and let $(x_{\boldsymbol{n}},x_{\boldsymbol{\beta}},x_{\parallel})$ be denoted here simply as $(x,y,z)$. Therefore, $\boldsymbol{B}=\left(0,0,B^3\right)$ and $\boldsymbol{E}=\left(E^1,E^2,E^3\right)$ are constant vectors and the equations of motion in the form (1.2) are

$$\begin{aligned}\ddot{x}&=\dot{y}B^3+E^1\\ \ddot{y}&=-\dot{x}B^3+E^2\\ \ddot{z}&=E^3\end{aligned}\tag{4.57}$$

Having the four components of the electromagnetic field as arbitrary constants, the above equations play the role of the primary system and can be enlarged by the auxiliary equations

$$B^3_t=B^3_{x^j}=E^i_t=E^i_{x^j}=0\tag{4.58}$$

Consider then augmented infinitesimal transformations generated by a vector field given as a linear combination

$$\begin{aligned}\mathbf{V}=&\left(c_4x-c_1t\right)\frac{\partial}{\partial x}+\left(c_4y-c_2t\right)\frac{\partial}{\partial y}+\left(c_4z-c_3t^2\right)\frac{\partial}{\partial z}+\\&+\left(c_4E^1+c_2B^3\right)\frac{\partial}{\partial E^1}+\left(c_4E^2-c_1B^3\right)\frac{\partial}{\partial E^2}+\left(c_4E^3-2c_3\right)\frac{\partial}{\partial E^3}\end{aligned}$$

where $c_i$, $i = 1, \dots, 4$ are arbitrary constants. One can easily show that the $(2,1)$-th prolongation of $\mathbf{V}$ leaves invariant the auxiliary system (4.58) (since its components in the $(\boldsymbol{B}, \boldsymbol{E})$-space are really constants on solutions of the system). Therefore we focus on the usual second prolongation of the projection $\mathbf{v} = (c_4 x - c_1 t)\,\partial_x + (c_4 y - c_2 t)\,\partial_y + \left(c_4 z - c_3 t^2\right)\partial_z$,

$$\mathbf{v}^{(2)} = \mathbf{v} + (c_4\dot{x} - c_1)\frac{\partial}{\partial \dot{x}} + (c_4\dot{y} - c_2)\frac{\partial}{\partial \dot{y}} + (c_4\dot{z} - 2c_3 t)\frac{\partial}{\partial \dot{z}} + c_4\ddot{x}\frac{\partial}{\partial \ddot{x}} + c_4\ddot{y}\frac{\partial}{\partial \ddot{y}} + (c_4\ddot{z} - 2c_3)\frac{\partial}{\partial \ddot{z}}$$

that is only required in applying $\mathbf{V}^{(2,1)}$ on the primary system (4.57), and we easily see that the infinitesimal equivalence criterion of Theorem 4.47 is satisfied, i.e.

$$\begin{aligned}
\mathbf{V}^{(2,1)}\left(\ddot{x} - \dot{y}B^3 - E^1\right) &= c_4\ddot{x} - (c_4\dot{y} - c_2)\,B^3 - c_4E^1 - c_2B^3 = c_4\left(\ddot{x} - \dot{y}B^3 - E^1\right) = 0\\
\mathbf{V}^{(2,1)}\left(\ddot{y} + \dot{x}B^3 - E^2\right) &= c_4\ddot{y} + (c_4\dot{x} - c_1)\,B^3 - c_4E^2 + c_1B^3 = c_4\left(\ddot{y} + \dot{x}B^3 - E^2\right) = 0\\
\mathbf{V}^{(2,1)}\left(\ddot{z} - E^3\right) &= c_4\ddot{z} - 2c_3 - c_4E^3 + 2c_3 = c_4\left(\ddot{z} - E^3\right) = 0
\end{aligned}$$

on solutions of the system.

Therefore $\mathbf{V}$ generates a 4-parameter group of equivalence transformations that can be expressed altogether as

$$\begin{aligned}
\widetilde{t} &= t & \widetilde{B}^3 &= B^3\\
\widetilde{x} &= \epsilon_4\,(x - \epsilon_1 t) & \widetilde{E}^1 &= \epsilon_4\left(E^1 + \epsilon_2 B^3\right)\\
\widetilde{y} &= \epsilon_4\,(y - \epsilon_2 t) & \widetilde{E}^2 &= \epsilon_4\left(E^2 - \epsilon_1 B^3\right)\\
\widetilde{z} &= \epsilon_4\left(z - \epsilon_3 t^2\right) & \widetilde{E}^3 &= \epsilon_4\left(E^3 - 2\epsilon_3\right)
\end{aligned} \tag{4.59}$$

where $\epsilon_i$, $i = 1, \dots, 4$ are arbitrary constants with $\epsilon_4 \neq 0$. Now let us see how knowledge of the equivalence group can be exploited in this case: If we choose $\epsilon_1 = E^2/B^3$, $\epsilon_2 = -E^1/B^3$, $\epsilon_3 = E^3/2$ and $\epsilon_4 = 1$ then solutions $(x(t), y(t), z(t))$ of the original system (4.57) are mapped to the solutions $(\widetilde{x}(t), \widetilde{y}(t), \widetilde{z}(t))$ of another member of this class, the equivalent system

$$\begin{aligned}
\ddot{\widetilde{x}} &= \dot{\widetilde{y}}B^3\\
\ddot{\widetilde{y}} &= -\dot{\widetilde{x}}B^3\\
\ddot{\widetilde{z}} &= 0
\end{aligned} \tag{4.60}$$

where the magnetic field is the same, but the electric field $\widetilde{E}^i = 0$ has vanished! Thus we can solve the much simpler system (4.60), and then the solution to the original one (4.57) can be retrieved by the inverse transformation $(x, y, z) = \left(\widetilde{x} + \epsilon_1 t, \widetilde{y} + \epsilon_2 t, \widetilde{z} + \epsilon_3 t^2\right)$. This is actually what (1.70) describes; $\epsilon_1$ and $\epsilon_2$ are the components $\mathrm{v}_{\boldsymbol{n}}$ and $\mathrm{v}_{\boldsymbol{\beta}}$, respectively, of the perpendicular velocity (1.60) and $\epsilon_3 = E_{\parallel}/2$ accordingly. And the transformation in the $xy$-plane is nothing more than the Galilean boost (1.59) in the perpendicular plane.

Of course if we want to find all the equivalence transformations of a class of differential equations, we have to enforce the condition of the above theorem and work our way back. The program followed is the analogue of the procedure for finding symmetries: *i*) Apply the two conditions $\mathbf{V}^{(n,s)}(\delta) = 0$ and $\mathbf{V}^{(n,s)}(\Delta) = 0$ of Theorem 4.47, *ii*) starting with the auxiliary system, substitute the restrictions $\delta = 0$, then move on to the primary system, substitute the restrictions $\delta = \Delta = 0$ and write the two equivalence conditions as a polynomial in the derivatives of $u$, $a$ and the derivatives of $a$, *iii*) equate the coefficients of the related monomials to zero, *iv*) solve the resulting system of equations to find the most general form of the vector field $\mathbf{V}$, and finally *v*) construct the equivalence group for the class. The equations derived in the third step of the above algorithm determining the Lie algebra of the equivalence group are also linear, homogeneous and in most cases overdetermined. Interestingly enough, even for classes of first-order ODEs these may be overdetermined too, as shown by example in [75], p. 57.

From Definition 4.45, one realizes that among the equivalence transformations of a certain class there may be some which map solutions of a member of the class to solutions of the same member, and hence by definition are symmetries for that system. In other words, the equivalence group for a class contains as subgroups symmetry groups for different systems of the class. The next proposition tells us when.

**Proposition 4.50.** Let $\mathbf{V}$ be a generator of a one-parameter equivalence group for a class $\mathcal{C}$ of differential equations. Then the projection $\mathbf{v}$ is a symmetry generator for the primary system $\mathcal{C}(g)$ if and only if $a = g(x,u)$ is a group invariant solution for the auxiliary system, meaning $\mathbf{V}(a - g(x,u)) = 0$ whenever $a = g(x,u)$ is a solution of the auxiliary system.

*Proof.* If $a = g(x,u)$ is an invariant solution for the auxiliary system, then the group of transformations generated by $\mathbf{V}$ maps the solution $a$ to itself $\widetilde{a} = a$, and consequently the member $\mathcal{C}(g)$ of the class to itself. Hence according to Definition 4.45 the group of transformations generated by $\mathbf{v}$ maps a solution of the system $\mathcal{C}(g)$ to a solution of the same system $\mathcal{C}(g)$, and thus it is a symmetry group. On the other hand if $\mathbf{v}$ is a symmetry group for $\mathcal{C}(g)$, then the system is invariant $\mathcal{C}(\widetilde{g}) = \mathcal{C}(g)$ under the action of the group generated by $\mathbf{V}$ and so $\widetilde{g}(\widetilde{x},\widetilde{u}) = g(x,u)$ or $\widetilde{a} = a$, that is $a = g(x,u)$ is a group invariant solution. □

The intersection of all the symmetry algebras of each system $\mathcal{C}(g)$ is called the *principal Lie algebra*. It consists of all the vector fields $\mathbf{v}$ that generate symmetries admitted by all systems $\mathcal{C}(g)$ for any $g$, and which can be found by the next corollary of the above proposition.

**Corollary 4.51.** The projection $\mathbf{v}$ of an equivalence generator $\mathbf{V}$ is a symmetry generator of the primary system with arbitrary functions $a$ if and only if $\mathbf{V}(a - g(x,u)) = 0$ is satisfied for any solution $g(x,u)$ of the auxiliary system.

Now, let us return back to our discussion at the beginning of this section and see how the use of equivalence transformations can lead to a symmetry classification. For the latter becomes necessary, since one cannot list all the invariant solutions of the auxiliary system, i.e. all the members of the class admitting symmetry groups. To this end, two systems or, in general, two subclasses of a class of differential equations will be considered distinct if they are not related through an equivalence transformation; otherwise their symmetry group is essentially the same. Such a distinction can be expressed back in terms of Lie subgroups of the equivalence group and accordingly to Lie subalgebras of the equivalence algebra: Two equivalence subalgebras will be considered *similar* if they are isomorphic under an equivalence transformation. Following [90], these isomorphisms are called *inner automorphisms* of the equivalence algebra; see sections 14.3, 14.4 and 14.7 therein for more on these matters.

From the above arguments it becomes evident that we need to determine the inner automorphisms of the generators $\mathbf{V}_i$, $i = 1, \dots, r$ of an $r$-parameter equivalence group $G$. Recalling from (4.18) the Lie algebra structure, we see that any element $\mathbf{W} = c_i \mathbf{V}_i$ of the equivalence algebra $\mathfrak{g}$ under the action of the Lie bracket is mapped to an element $[\mathbf{V}_j, \mathbf{W}] = [\mathbf{V}_j, c_k \mathbf{V}_k] = c^i_{jk} c_k \mathbf{V}_i$ of $\mathfrak{g}$, where $c^k_{ij}$ are the structure constants. The linear mapping $c_i \mathbf{V}_i \longmapsto c^i_{jk} c_k \mathbf{V}_i$ denoted as $\operatorname{ad} \mathbf{V}_j$ and defined in general for any Lie algebra as

$$\operatorname{ad} \mathbf{V}\,(\mathbf{W}) = [\mathbf{V}, \mathbf{W}]$$

is called the *adjoint action* or *representation*, given in infinitesimal form. The inner automorphisms of the Lie algebra to itself are given by the related adjoint transformations,

$$\begin{aligned} \operatorname{Ad}\left(e^{\epsilon \mathbf{V}}\right)(\mathbf{W}) &= e^{\epsilon \operatorname{ad} \mathbf{V}}(\mathbf{W}) \\ &= \mathbf{W} + \epsilon\,[\mathbf{V}, \mathbf{W}] + \frac{\epsilon^2}{2}\,[\mathbf{V}, [\mathbf{V}, \mathbf{W}]] - \cdots \end{aligned} \tag{4.61}$$

which form the so-called *adjoint group* of the corresponding Lie group. For simplicity, note that every element $\mathbf{V} = c_i \mathbf{V}_i$ of an $r$-dimensional Lie algebra can be represented by the vector $c = (c_1, \dots, c_r)$. Thus, instead of the above mapping one can consider the adjoint action on the constants $c_i$ generated by the vector fields [51],

$$\mathbf{C}_j = c^i_{jk} c_k \frac{\partial}{\partial c_i} \tag{4.62}$$

and determine the adjoint group of the equivalence group from their exponentiation. It is also worth mentioning that the *center* of a Lie algebra, commuting with every other of its elements, is the kernel of the adjoint map ad [90].

In conclusion, the whole problem of symmetry classification of a class of differential equations is reduced to a classification of its equivalence algebra. Under the above considerations, the equivalence algebra is decomposed into subalgebras which are not connected by an inner automorphism, meaning classification is made under the adjoint action. The goal is to arrive at an *optimal system* of the least possible subalgebras, in which no two subalgebras are similar and every other subalgebra is similar to one of the subalgebras in the optimal system. In this classification the center of the equivalence algebra cannot be changed in any way, and thus directly included in any subalgebra. Unfortunately, besides perhaps the use of invariants of the adjoint group in some cases, no systematic method is available for constructing the optimal system. See [90], section 14.8 for the case of low-dimensional subalgebras.

Nonetheless, there could be symmetries lying outside the equivalence group. Therefore, the symmetry classification described above is only *partial* or *preliminary*, as it is often called. In many cases though that is all one can do, not to mention the severe decrease in the amount of labor to do so. On the other hand, the equivalence information can be handled in different ways to arrive at a complete symmetry classification. For example, one can employ equivalence transformations to classify the determining equations deduced from the symmetry condition before solving them.

## Final notes

The notion of symmetry in differential equations has been generalized in lots of ways extending further its applicability. A first example are *generalized symmetries* also known as *Lie-Bäcklund transformations* that besides the independent and dependent variables they also include transformations of the derivatives. True generalized symmetries first considered by Noether herself can only be defined as transformations on the infinite-dimensional function space of solutions (see [89], p. 297) (for ordinary differential equations many authors also use the term dynamical or internal symmetries). Lie and Bäcklund on the other hand probably considered genuine geometrical transformations on a jet space of finite order; when $n = 1$ these are called *contact symmetries*. For partial differential equations, however, a transformation on an $n$-jet space can always be recovered by the $n$-th prolongation of a transformation on the base space.

Other generalizations are *nonlocal symmetries* (a special case of which are *potential symmetries* [10]) that describe symmetries which depend on functions that are connected with the dependent variables through differential relations, or *conditional symmetries* that are symmetries of the original system of interest, which also leave invariant an attached system of differential equations. One may also consider symmetries for problems with initial-boundary conditions, which are admitted, apart from the system, by the initial-boundary conditions and the initial-boundary surface. *Approximate symmetries* for perturbed systems have also been introduced, which are symmetries that depend on the perturbation parameter, leaving invariant the system through first order of perturbation. Several if not most parts of the machinery of the point symmetries considered in this chapter, such as group structure, generators, prolongations, infinitesimal criterion, reduction of order, Noether's theorem, are sustained in many cases at least under suitable modifications. For example, approximate Noether symmetries correspond to adiabatic invariants. On the other hand some features may be lost, as for instance generalized symmetries for second-order ODEs suffer the same problem (see Remark 4.24) that point symmetries do for first-order ones, or conditional symmetries no longer enjoy the linearity of the corresponding determining equations and are hard to find.

Among the advantages of symmetry methods is that they are amenable to computer programming. The five-step algorithm for finding symmetry (or equivalence) groups can be executed using computer algebra programs, which can automate all the tedious calculations involved. Recall from Example 4.16 how cumbersome can be the construction of the required prolongation for high-order equations. Nevertheless, the algorithmic nature of prolonging a vector field allows a computer implementation. Moreover and more importantly, the linear system of the determining equations may turn out so large that pencil and paper are almost impractical. To this end, more advanced methods have also been developed based on the works of Riquier and Janet, using the so-called Janet bases that bring linear homogeneous systems of PDEs into a simple form and which were lately revived by Reid and Schwarz [97]; see also [7], chapter 10. In conclusion, the main goal is to have a program that given a differential equation as input it will produce the symmetry group as output. In this way the margin of human error, not to mention time and effort, is less. Examples in a user-friendly environment such as *Mathematica* are "MathLie" by Baumann [7] and more recently "Sym" by Dimas [29, 30].

The handling of the determining equations is actually again related to the symmetry classification problem, for which Reid's method exploiting all the compatability conditions can extract the size and structure of the symmetry algebra without even solving them. On the other hand

lies essentially the work of Cartan who employed the equivalence group to break down the class of differential equations into canonical forms. Lisle in his thesis [75] explores the combination of Cartan-Tresse geometrical methods with Reid's analytic algorithm, suggesting his approach that includes both of them as limits in opposite cases.

# Chapter 5

# Symmetries of Charged Particle Motion

In this chapter, we investigate the symmetry properties of the three-dimensional Newtonian motion of charged particles given by the Lorentz force law (1.1). Assuming a Cartesian frame of reference, we consider arbitrary, time-independent electromagnetic fields $\boldsymbol{B} = \left(B^1, B^2, B^3\right)$ and $\boldsymbol{E} = \left(E^1, E^2, E^3\right)$ as smooth functions of the position $\boldsymbol{x}$ alone. The system, consisting of three second-order, autonomous ordinary differential equations, is repeated here from the component form (1.2)

$$\ddot{x}^i = \epsilon_{ijk}\dot{x}^j B^k(\boldsymbol{x}) + E^i(\boldsymbol{x}), \tag{5.1}$$

where all indices throughout this chapter and the next one take values from 1 to 3, unless stated otherwise. We also recall the well-known Lagrangian formulation of (5.1), previously described in detail in chapter 1.

For $\boldsymbol{B} = \boldsymbol{0}$, the above system reduces to a classical problem in mechanics, for which a complete symmetry group classification can be found in [25, 107] in two dimensions and [26, 27, 108] in three. A symmetry analysis for the two-dimensional case of system (5.1), where the magnetic field has a constant direction in space, has also been made [43, 44], even for time-dependent electromagnetic fields. Therefore this case, along with trivial ones of less physical interest, where either one of the fields is homogeneous, will not be considered here. Our intention is to cover a variety of applications, in which the electromagnetic field may be quite complicated, such as those occuring in plasma physics and fusion devices. A series of simpler problems for particular choices of the functions $\boldsymbol{B}$ and $\boldsymbol{E}$ is presented in [92].

## 5.1 Lie point symmetries

We consider infinitesimal point transformations generated by the vector field

$$\mathbf{v} = \xi(t, \boldsymbol{x})\frac{\partial}{\partial t} + \eta^i(t, \boldsymbol{x})\frac{\partial}{\partial x^i}. \tag{5.2}$$

According to Theorem 4.19 the symmetry condition reads

$$\mathbf{v}^{(2)}(\ddot{x}^i - \epsilon_{ijk}\dot{x}^j B^k(\boldsymbol{x}) - E^i(\boldsymbol{x})) = 0, \tag{5.3}$$

whenever equations (5.1) hold, where $\mathbf{v}^{(2)}$ is the second prolongation of $\mathbf{v}$ given by Proposition 4.15. Applying formula (4.25) we have

$$\mathbf{v}^{(2)} = \mathbf{v} + \eta^t_i(t, \boldsymbol{x}, \dot{\boldsymbol{x}})\frac{\partial}{\partial \dot{x}^i} + \eta^{tt}_i(t, \boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}})\frac{\partial}{\partial \ddot{x}^i} \tag{5.4}$$

and from the recursion relation (4.26),

$$\begin{aligned}
\eta^t_i &= D_t\eta^i - \dot{x}^i D_t\xi = \eta^i_t + \eta^i_{x^j}\dot{x}^j - \xi_t\dot{x}^i - \xi_{x^j}\dot{x}^i\dot{x}^j \\
\eta^{tt}_i &= D_t\eta^t_i - \ddot{x}^i D_t\xi = \eta^i_{tt} + 2\eta^i_{tx^j}\dot{x}^j + \eta^i_{x^jx^k}\dot{x}^j\dot{x}^k + \eta^i_{x^j}\ddot{x}^j - \xi_{tt}\dot{x}^i - 2\xi_{tx^j}\dot{x}^i\dot{x}^j - 2\xi_t\ddot{x}^i - \\
&\quad - \xi_{x^jx^k}\dot{x}^i\dot{x}^j\dot{x}^k - 2\xi_{x^j}\dot{x}^j\ddot{x}^i - \xi_{x^j}\dot{x}^i\ddot{x}^j
\end{aligned}$$

After substituting the second derivatives, equations (5.3) take the form of polynomials of third degree in terms of the first derivatives $\dot{x}^i$,

$$\begin{aligned}
0 &= \eta^{tt}_i - \epsilon_{ijk}\eta^t_j B^k - \epsilon_{ijk}\dot{x}^j\eta^l B^k_{x^l} - \eta^l E^i_{x^l} \\
&= \eta^i_{tt} + 2\eta^i_{tx^j}\dot{x}^j + \eta^i_{x^jx^k}\dot{x}^j\dot{x}^k + \eta^i_{x^l}\epsilon_{ljk}\dot{x}^j B^k + \eta^i_{x^j}E^j - \xi_{tt}\dot{x}^i - 2\xi_{tx^j}\dot{x}^i\dot{x}^j - 2\xi_t\epsilon_{ijk}\dot{x}^j B^k - \\
&\quad - 2\xi_t E^i - \xi_{x^jx^k}\dot{x}^i\dot{x}^j\dot{x}^k - 2\xi_{x^j}\dot{x}^j\epsilon_{ikl}\dot{x}^k B^l - 2\xi_{x^j}\dot{x}^j E^i - \xi_{x^j}\dot{x}^i\epsilon_{jkm}\dot{x}^k B^m - \xi_{x^j}\dot{x}^i E^j - \\
&\quad - \epsilon_{ijk}\eta^j_t B^k - \epsilon_{ilk}\eta^l_{x^j}\dot{x}^j B^k + \epsilon_{ijk}\xi_t\dot{x}^j B^k + \epsilon_{ijl}\xi_{x^k}\dot{x}^j\dot{x}^k B^l - \epsilon_{ijk}\dot{x}^j\eta^l B^k_{x^l} - \eta^l E^i_{x^l}
\end{aligned} \tag{5.5}$$

Being identities for all $t$, $x^i$ and $\dot{x}^i$, they finally break up into a larger set of partial differential equations, the so-called determining equations,

$$\xi_{x^jx^k} = 0 \tag{5.6a}$$

$$2\eta^i_{x^jx^k} - \epsilon_{ijl}\xi_{x^k}B^l + \epsilon_{ilk}\xi_{x^j}B^l - \delta_{ij}\left(2\xi_{tx^k} - \epsilon_{lmk}\xi_{x^l}B^m\right) - \delta_{ik}\left(2\xi_{tx^j} + \epsilon_{ljm}\xi_{x^l}B^m\right) = 0 \tag{5.6b}$$

$$2\eta^i_{tx^j} + \epsilon_{ljk}\eta^i_{x_l}B^k - \epsilon_{ijk}\xi_t B^k - 2\xi_{x^j}E^i - \epsilon_{ilk}\eta^l_{x_j}B^k - \epsilon_{ijk}\eta^l B^k_{x_l} - \delta_{ij}(\xi_{tt} + \xi_{x^l}E^l) = 0 \tag{5.6c}$$

$$\eta^i_{tt} + \eta^i_{x^j}E^j - 2\xi_t E^i - \epsilon_{ijk}\eta^j_t B^k - \eta^j E^i_{x^j} = 0 \tag{5.6d}$$

Dropping Kronecker's deltas and Levi-Civita symbols, the above system in more detail is given by the equations listed below. Equations (5.6a), (5.6b) and (5.6d) are repeated exactly the same. Equations (5.6c) on the other hand are slightly manipulated for convenience: equations (5.6c1) are (5.6c) for $i = j$, just as (5.6c2) are the half of (5.6c) for $i \neq j$. Finally, (5.6c3) are replacing the rest of (5.6c) by adding the $ij$- and $ji$-equations in pairs for $i \neq j$.

$$\begin{aligned}
\xi_{xx} &= 0 \\
\xi_{yy} &= 0 \\
\xi_{zz} &= 0 \\
\xi_{xy} &= 0 \\
\xi_{xz} &= 0 \\
\xi_{yz} &= 0
\end{aligned} \tag{5.6a}$$

$$\begin{aligned}
\eta^1_{yy} - \xi_y B^3 &= 0 \\
\eta^1_{zz} + \xi_z B^2 &= 0 \\
\eta^2_{zz} - \xi_z B^1 &= 0 \\
\eta^2_{xx} + \xi_x B^3 &= 0 \\
\eta^3_{xx} - \xi_x B^2 &= 0 \\
\eta^3_{yy} + \xi_y B^1 &= 0 \\
2\eta^1_{yz} - \xi_z B^3 + \xi_y B^2 &= 0 \\
2\eta^2_{xz} - \xi_x B^1 + \xi_z B^3 &= 0 \\
2\eta^3_{xy} - \xi_y B^2 + \xi_x B^1 &= 0 \\
\eta^1_{xx} - 2\xi_{tx} - \xi_z B^2 + \xi_y B^3 &= 0 \\
\eta^2_{yy} - 2\xi_{ty} - \xi_x B^3 + \xi_z B^1 &= 0 \\
\eta^3_{zz} - 2\xi_{tz} - \xi_y B^1 + \xi_x B^2 &= 0 \\
2\eta^1_{xy} - 2\xi_{ty} - 2\xi_x B^3 + \xi_z B^1 &= 0 \\
2\eta^1_{xz} - 2\xi_{tz} + 2\xi_x B^2 - \xi_y B^1 &= 0 \\
2\eta^2_{yz} - 2\xi_{tz} - 2\xi_y B^1 + \xi_x B^2 &= 0 \\
2\eta^2_{xy} - 2\xi_{tx} + 2\xi_y B^3 - \xi_z B^2 &= 0 \\
2\eta^3_{xz} - 2\xi_{tx} - 2\xi_z B^2 + \xi_y B^3 &= 0 \\
2\eta^3_{yz} - 2\xi_{ty} + 2\xi_z B^1 + \xi_x B^3 &= 0
\end{aligned} \tag{5.6b}$$

$$
\begin{aligned}
2\eta^1_{tx} - \xi_{tt} - \left(\eta^1_y + \eta^2_x\right) B^3 + \left(\eta^1_z + \eta^3_x\right) B^2 - 2\xi_x E^1 - \xi_{x^i} E^i &= 0 \\
2\eta^2_{ty} - \xi_{tt} - \left(\eta^2_z + \eta^3_y\right) B^1 + \left(\eta^2_x + \eta^1_y\right) B^3 - 2\xi_y E^2 - \xi_{x^i} E^i &= 0 \\
2\eta^1_{tz} - \xi_{tt} - \left(\eta^3_x + \eta^1_z\right) B^2 + \left(\eta^3_y + \eta^2_z\right) B^1 - 2\xi_z E^3 - \xi_{x^i} E^i &= 0
\end{aligned} \tag{5.6c1}
$$

$$
\begin{aligned}
2\eta^1_{ty} + \left(\eta^1_x - \eta^2_y\right) B^3 - \eta^1_z B^1 + \eta^3_y B^2 - \xi_t B^3 - 2\xi_y E^1 - \eta^i B^3_{x^i} &= 0 \\
2\eta^1_{tz} + \left(\eta^3_z - \eta^1_x\right) B^2 + \eta^1_y B^1 - \eta^2_z B^3 - \xi_t B^2 - 2\xi_z E^1 - \eta^i B^2_{x^i} &= 0 \\
2\eta^2_{tz} + \left(\eta^2_y - \eta^3_z\right) B^1 - \eta^2_x B^2 + \eta^1_z B^3 - \xi_t B^1 - 2\xi_z E^2 - \eta^i B^1_{x^i} &= 0
\end{aligned} \tag{5.6c2}
$$

$$
\begin{aligned}
2\left(\eta^1_{ty} + \eta^2_{tx}\right) - \left(\eta^1_z + \eta^3_x\right) B^1 + \left(\eta^2_z + \eta^3_y\right) B^2 + 2\left(\eta^1_x - \eta^2_y\right) B^3 - 2\xi_y E^1 - 2\xi_x E^2 &= 0 \\
2\left(\eta^1_{tz} + \eta^3_{tx}\right) + \left(\eta^1_y + \eta^2_x\right) B^1 - \left(\eta^2_z + \eta^3_y\right) B^3 + 2\left(\eta^3_z - \eta^1_x\right) B^2 - 2\xi_z E^1 - 2\xi_x E^3 &= 0 \\
2\left(\eta^2_{tz} + \eta^3_{ty}\right) - \left(\eta^1_y + \eta^2_x\right) B^2 + \left(\eta^1_z + \eta^3_x\right) B^3 + 2\left(\eta^2_y - \eta^3_z\right) B^1 - 2\xi_z E^2 - 2\xi_y E^3 &= 0
\end{aligned} \tag{5.6c3}
$$

$$
\begin{aligned}
\eta^1_{tt} + \eta^1_{x^i} E^i - 2\xi_t E^1 + \eta^3_t B^2 - \eta^2_t B^3 - \eta^i E^1_{x^i} &= 0 \\
\eta^2_{tt} + \eta^2_{x^i} E^i - 2\xi_t E^2 + \eta^3_t B^1 - \eta^1_t B^3 - \eta^i E^2_{x^i} &= 0 \\
\eta^3_{tt} + \eta^3_{x^i} E^i - 2\xi_t E^3 + \eta^2_t B^1 - \eta^1_t B^2 - \eta^i E^3_{x^i} &= 0
\end{aligned} \tag{5.6d}
$$

Starting from the top, the first set of equations shows that $\xi_{x^i}$ are functions only of $t$. Moving on to the second one, first observe that mixed second order partial derivatives $\eta^i_{x^j x^k}$ and $\eta^i_{x^k x^j}$ are already considered equal. Imposing further integrability conditions on (5.6b), meaning $\eta^i_{x^j x^k x^l} = \eta^i_{x^l x^j x^k}$, we find, after some investigation, that $\xi_{x^i} = 0$ unless $\boldsymbol{B}$ is constant, which in turn means that $\boldsymbol{E}$ is linear from (5.6c), and so consequently system (5.1) is altogether linear and not of much physical interest. Thus, for nonlinear systems $\xi_{x^i} = 0$ and hence $\eta^i_{x^j x^k} = 0$, i.e.

$$\xi = \xi(t), \tag{5.7}$$

$$\eta^i = f_{ij}(t) x^j + f_i(t) \tag{5.8}$$

where $\xi, f_{ij}, f_i$ are arbitrary functions of $t$. After substitution of the above expressions, equations (5.6c1) take the form

$$
\begin{aligned}
2\dot{f}_{11} - \ddot{\xi} + \left(f_{13} + f_{31}\right) B^2 - \left(f_{12} + f_{21}\right) B^3 &= 0, \\
2\dot{f}_{22} - \ddot{\xi} + \left(f_{12} + f_{21}\right) B^3 - \left(f_{23} + f_{32}\right) B^1 &= 0, \\
2\dot{f}_{33} - \ddot{\xi} + \left(f_{23} + f_{32}\right) B^1 - \left(f_{13} + f_{31}\right) B^2 &= 0,
\end{aligned} \tag{5.9}
$$

while (5.6c3), which is the other set of equations (5.6c) not containing any derivatives of $B^i$, become

$$
\begin{aligned}
2\left(\dot{f}_{12} + \dot{f}_{21}\right) - \left(f_{13} + f_{31}\right) B^1 + \left(f_{23} + f_{32}\right) B^2 + 2\left(f_{11} - f_{22}\right) B^3 &= 0, \\
2\left(\dot{f}_{13} + \dot{f}_{31}\right) + \left(f_{12} + f_{21}\right) B^1 - \left(f_{23} + f_{32}\right) B^3 + 2\left(f_{33} - f_{11}\right) B^2 &= 0, \\
2\left(\dot{f}_{23} + \dot{f}_{32}\right) - \left(f_{12} + f_{21}\right) B^2 + \left(f_{13} + f_{31}\right) B^3 + 2\left(f_{22} - f_{33}\right) B^1 &= 0.
\end{aligned} \tag{5.10}
$$

Before proceeding with the rest of (5.6c), direct inspection of systems (5.9) and (5.10) shows that unless

$$f_{ij} = -f_{ji}, \quad i \neq j \tag{5.11}$$

hold, then the magnetic field is at best of constant direction, a case excluded from our study as it has already been investigated in [44]. Inserting this skew-symmetry back to (5.9), we find

$$2f_{11} = 2f_{22} = 2f_{33} = \dot{\xi} + k. \tag{5.12}$$

where $k$ is a constant.

Then, the rest of the third set of the determining equations, system (5.6c2), place inevitably restrictions upon $\boldsymbol{B}$,

$$(f_{jk}x^k + f_j)\frac{\partial B^i}{\partial x^j} = (k - f_{jj})\, B^i + f_{ij}B^j + \epsilon_{ijk}\dot{f}_{jk}, \tag{5.13}$$

The latter show that, for a time-independent magnetic field, the functions $f_{ij}, f_i$ must be constants. This can be seen, if we differentiate the last three equations with respect to $x_i$, and then treat $f_{ij}, f_i$ and $c$ as the unknowns. Hence, we arrive at an algebraic system of nine equations in eight unknowns (considering (5.11)-(5.12)), which is also linear and homogeneous. In the generic case the rank of this system is seven, and therefore we can solve for $f_{ij}$ and $f_i$ in terms of $k$, which is an absolute constant, and the derivatives of the magnetic field. Since $\boldsymbol{B}$ does not depend explicitly on time, all these solutions have to be constants. And, from (5.14), as a result $\xi$ is a linear function of $t$. Therefore,

$$\begin{aligned} f_1(t) &= c_1 & f_{11}(t) &= c_7 \\ f_2(t) &= c_2 & f_{12}(t) &= -c_4 \\ f_3(t) &= c_3 & f_{13}(t) &= c_5 \\ \xi(t) &= c_8 t + c_0 & f_{23}(t) &= -c_6 \end{aligned} \tag{5.14}$$

where $c_i$, $i = 0, \ldots, 8$ are all constants, and the remaining of the determining equations, systems (5.13) and (5.6d)

$$\begin{aligned} (\boldsymbol{\eta}\cdot\nabla)\boldsymbol{B} &= Q_1\boldsymbol{B}, \\ (\boldsymbol{\eta}\cdot\nabla)\boldsymbol{E} &= Q_2\boldsymbol{E} \end{aligned} \tag{5.15}$$

where $\boldsymbol{\eta}$ is the vector with entries $\eta^i$ and $Q_1 = -c_8 I + Q$, $Q_2 = (c_7 - 2c_8)\, I + Q$ are $3\times 3$ square matrices, $I$ being the unit matrix and $Q$ the skew-symmetric matrix

$$Q = \begin{pmatrix} 0 & -c_4 & c_5 \\ c_4 & 0 & -c_6 \\ -c_5 & c_6 & 0 \end{pmatrix}. \tag{5.16}$$

In conclusion, when system (5.1) is nonlinear and the magnetic field is not straight, then (summarizing relations (5.7), (5.8), (5.11), (5.12) and (5.14)) the most general form of the symmetry generator is

$$\begin{aligned} \xi &= c_8 t + c_0, \\ \eta^1 &= c_7 x - c_4 y + c_5 z + c_1, \\ \eta^2 &= c_7 y + c_4 x - c_6 z + c_2, \\ \eta^3 &= c_7 z - c_5 x + c_6 y + c_3. \end{aligned} \tag{5.17}$$

when the electromagnetic field satisfies (5.15). We should note, however, that at this stage only $c_0$ is completely arbitrary, while all other constants are related through (5.15). Thus, the symmetry algebra is spanned by $\mathbf{v}_0 = \partial_t$ and a symmetry generator of the general form

$$\mathbf{v} = \sum_{i=1}^{8} c_i \mathbf{v}_i \tag{5.18}$$

which is actually a sum of translations, rotations and dilatations

$$\begin{aligned} \mathbf{v}_1 &= \partial_x \\ \mathbf{v}_2 &= \partial_y \\ \mathbf{v}_3 &= \partial_z \\ \mathbf{v}_4 &= x\partial_y - y\partial_x \\ \mathbf{v}_5 &= z\partial_x - x\partial_z \\ \mathbf{v}_6 &= y\partial_z - z\partial_y \\ \mathbf{v}_7 &= x\partial_x + y\partial_y + z\partial_z \\ \mathbf{v}_8 &= t\partial_t \end{aligned} \tag{5.19}$$

While the former, expected for every autonomous system, appears whatever the (stationary) fields $\boldsymbol{B}$ and $\boldsymbol{E}$ may be, the latter is admitted only when $\boldsymbol{B}$ and $\boldsymbol{E}$ satisfy equations (5.15). The dimension of the symmetry algebra, in general, will be determined from the number of independent constants $c_i$, $i = 1, \ldots, 8$ inserted in (5.18).

So, leaving the anticipated $\mathbf{v}_0$ aside, we focus on the existence of the general symmetry $\mathbf{v}$ (5.18). The next task to complete the symmetry analysis is to find the form of the electromagnetic field, respecting the symmetry condition, i.e. solve equations (5.15).

## 5.2 The form of the electromagnetic field

The solutions to (5.15) describe the general form the vector functions $\boldsymbol{B}$ and $\boldsymbol{E}$ must have in order system (5.1) to admit symmetry (5.18). In physics, however, the electromagnetic field obeys Maxwell's equations (1.4a)-(1.4b). And while Gauss's and Ampére's laws (1.4b) simply determine the source of the electric and magnetic field respectively, i.e. the charge and current densities, the first two (1.4a) represent conditions that $\boldsymbol{B}$ and $\boldsymbol{E}$ have to satisfy, and are inextricably connected with every real problem (5.1). Instead of imposing these conditions, we recall their solution in terms of the vector potential $\boldsymbol{A} = (A_1, A_2, A_3)$ and the scalar one $\Phi$.

Replacing the solutions (1.5)-(1.6) in (5.15), provides us then with the restrictions that now $\boldsymbol{A}$ and $\Phi$ must satisfy in order to respect the Lie point symmetry condition. The equations we end up with, after integration, can be cast into the following form

$$(\boldsymbol{\eta} \cdot \nabla)\boldsymbol{A} = Q_3 \boldsymbol{A} + \nabla f, \tag{5.20}$$

$$(\boldsymbol{\eta} \cdot \nabla)\Phi = 2\,(c_7 - c_8)\,\Phi + c_9, \tag{5.21}$$

where $Q_3 = (c_7 - c_8)\,I + Q$, $f$ is an arbitrary function of $\boldsymbol{x}$ and $c_9$ some constant. Furthermore, the gauge invariance $\boldsymbol{A} \longrightarrow \boldsymbol{A} + \nabla g$ of the vector potential (see Remark 1.1) can, in fact, guarantee the existence of an equivalent $\boldsymbol{A}$, such that

$$(\boldsymbol{\eta} \cdot \nabla)\boldsymbol{A} = Q_3 \boldsymbol{A} \tag{5.22}$$

for $g$ satisfying the relation

$$\boldsymbol{\eta} \cdot \nabla g + (c_8 - 2c_7)\, g = f, \tag{5.23}$$

Thus, in case of true electromagnetic fields, that certainly have to comply with Maxwell's equations (1.4a), the solutions to (5.22)-(5.21) describe through (1.5) and (1.6) the ones for which system (5.1) has the symmetry (5.18).

Consequently, we can either treat $\boldsymbol{B}$ and $\boldsymbol{E}$ only as functions entering the system, or we can view them as part of a bigger physical problem, that also includes Maxwell's equations (1.4a). In what follows we focus on that second case and give the solutions $\boldsymbol{A}$ and $\Phi$ of (5.21)-(5.22). For reasons that will be apparent, the potentials that the electromagnetic field comes from, besides carrying more information, are more convenient to use in this situation. The form of $\boldsymbol{B}$ and $\boldsymbol{E}$ can then be found through (1.5)-(1.6). If, however, one wishes to determine them without taking into account (1.4a), the task would be similar. The two systems in (5.15) can be solved independently of each other, and obviously by the same means as (5.22) can. Either one of them

is a system of three coupled first order linear partial differential equations, and they all have almost the exact same form, since the matrices $Q_i$ are all of the same structure differing only by a minor factor. Thus, we present in detail a way of solving (5.22), whose explicit form is given below along with (5.21), and then either set of (5.15) can be treated accordingly.

$$(c_7x - c_4y + c_5z + c_1)\frac{\partial A_1}{\partial x} + (c_7y + c_4x - c_6z + c_2)\frac{\partial A_1}{\partial y} + (c_7z - c_5x + c_6y + c_3)\frac{\partial A_1}{\partial z} = (c_7 - c_8)A_1 - c_4A_2 + c_5A_3$$

$$(c_7x - c_4y + c_5z + c_1)\frac{\partial A_2}{\partial x} + (c_7y + c_4x - c_6z + c_2)\frac{\partial A_2}{\partial y} + (c_7z - c_5x + c_6y + c_3)\frac{\partial A_2}{\partial z} = (c_7 - c_8)A_2 + c_4A_1 - c_6A_3$$

$$(c_7x - c_4y + c_5z + c_1)\frac{\partial A_3}{\partial x} + (c_7y + c_4x - c_6z + c_2)\frac{\partial A_3}{\partial y} + (c_7z - c_5x + c_6y + c_3)\frac{\partial A_3}{\partial z} = (c_7 - c_8)A_3 - c_5A_1 + c_6A_2$$

$$(c_7x - c_4y + c_5z + c_1)\frac{\partial \Phi}{\partial x} + (c_7y + c_4x - c_6z + c_2)\frac{\partial \Phi}{\partial y} + (c_7z - c_5x + c_6y + c_3)\frac{\partial \Phi}{\partial z} = 2\,(c_7 - c_8)\,\Phi + c_9$$

First of all we want to uncouple the system. The form of the equations allows one to do so, when $Q_3$ is diagonalizable. The above matrix however is not, since it has two complex eigenvalues, $\lambda_{1,2} = c_7 - c_8 \pm ic$, and only one real, $\lambda_3 = c_7 - c_8$, where $c = \left(c_4^2 + c_5^2 + c_6^2\right)^{\frac{1}{2}}$. Nevertheless, we can still separate one equation from the other two, by setting $\boldsymbol{A} = P\bar{\boldsymbol{A}}$, where $P = (u\ v\ e)$ is the matrix of the eigenvectors $u \pm iv, e$ of $Q_3$ corresponding to the eigenvalues $\lambda_{1,2}, \lambda_3$, respectively. Then, left multiplication of equation (5.22) with the inverse of $P$, leads to

$$(\boldsymbol{\eta} \cdot \nabla)\bar{\boldsymbol{A}} = \bar{Q}_3\bar{\boldsymbol{A}}, \tag{5.24}$$

where $\bar{Q}_3 = P^{-1}Q_3P = (c_7 - c_8)\,I + \bar{Q}$ and $\bar{Q} = P^{-1}QP$ is the normal form of the skew-symmetric matrix $Q$,

$$\bar{Q} = \begin{pmatrix} 0 & -c & 0 \\ c & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.25}$$

Thus, the third of equations (5.24) for $\bar{A}_3$ is detached from the others. To uncouple the rest two we need a nonlinear transformation. Their form, as it can be seen from the matrix $\bar{Q}$, naturally implies taking polar coordinates in the $\bar{A}_1\bar{A}_2$-plane. In conclusion, the following transformation in the dependent variables, i.e. the components of the vector potential,

$$\begin{aligned} \widetilde{A}_1 &= \sqrt{\bar{A}_1^2 + \bar{A}_2^2} \\ \widetilde{A}_2 &= \arctan\left(\frac{\bar{A}_2}{\bar{A}_1}\right) \qquad \text{and} \qquad \bar{\boldsymbol{A}} = P^{-1}\boldsymbol{A}, \\ \widetilde{A}_3 &= \bar{A}_3 \end{aligned} \tag{5.26}$$

where $P$ is defined through the eigenvectors of $Q$, uncouples system (5.22). Indeed, the latter now takes the form

$$\begin{aligned}
&(c_7x - c_4y + c_5z + c_1)\frac{\partial \widetilde{A}_1}{\partial x} + (c_7y + c_4x - c_6z + c_2)\frac{\partial \widetilde{A}_1}{\partial y} + (c_7z - c_5x + c_6y + c_3)\frac{\partial \widetilde{A}_1}{\partial z} = (c_7 - c_8)\,\widetilde{A}_1\\
&(c_7x - c_4y + c_5z + c_1)\frac{\partial \widetilde{A}_2}{\partial x} + (c_7y + c_4x - c_6z + c_2)\frac{\partial \widetilde{A}_2}{\partial y} + (c_7z - c_5x + c_6y + c_3)\frac{\partial \widetilde{A}_2}{\partial z} = c\\
&(c_7x - c_4y + c_5z + c_1)\frac{\partial \widetilde{A}_3}{\partial x} + (c_7y + c_4x - c_6z + c_2)\frac{\partial \widetilde{A}_3}{\partial y} + (c_7z - c_5x + c_6y + c_3)\frac{\partial \widetilde{A}_3}{\partial z} = (c_7 - c_8)\,\widetilde{A}_3
\end{aligned}$$

Now, each equation can be solved independently by the method of characteristics. Even more conveniently, all three of them are essentially (the first and third exactly) the same.

This means that three out of four characteristic equations for the above equations are common in each case, forming a linear dynamical system,

$$\frac{d\boldsymbol{x}}{ds} = C\boldsymbol{x} + \boldsymbol{c}, \tag{5.27}$$

where the $3 \times 3$ square matrix $C = c_7 I + Q$ and the column vector $\boldsymbol{c} = (c_1, c_2, c_3)$. Since the matrices $Q_3$ and $C$ share the same eigenvectors, the homogeneous counterpart of the above equations can be easily resolved, very similarly to (5.22). Only now, due to $\boldsymbol{c}$, we need to make a slight adjustment. More specifically, consider new independent variables,

$$\begin{aligned}
&\widetilde{x} = \sqrt{\bar{x}^2 + \bar{y}^2}\\
&\widetilde{y} = \arctan\left(\frac{\bar{y}}{\bar{x}}\right) \qquad \text{and} \qquad \bar{\boldsymbol{x}} = P^{-1}\left(\boldsymbol{x} - \boldsymbol{k}\right),\\
&\widetilde{z} = \bar{z}
\end{aligned} \tag{5.28}$$

where $P$ is the same as before and $\boldsymbol{k} = (k_1, k_2, k_3)$ is a constant vector soon to be defined case by case. Then, transformations (5.26) and (5.28) bring system (5.22) into a much simpler form,

$$\begin{aligned}
&c_7\widetilde{x}\,\frac{\partial \widetilde{A}_1}{\partial \widetilde{x}} + c\,\frac{\partial \widetilde{A}_1}{\partial \widetilde{y}} + (c_7\widetilde{z} + \widetilde{c}_3)\,\frac{\partial \widetilde{A}_1}{\partial \widetilde{z}} = (c_7 - c_8)\,\widetilde{A}_1,\\
&c_7\widetilde{x}\,\frac{\partial \widetilde{A}_2}{\partial \widetilde{x}} + c\,\frac{\partial \widetilde{A}_2}{\partial \widetilde{y}} + (c_7\widetilde{z} + \widetilde{c}_3)\,\frac{\partial \widetilde{A}_2}{\partial \widetilde{z}} = c,\\
&c_7\widetilde{x}\,\frac{\partial \widetilde{A}_3}{\partial \widetilde{x}} + c\,\frac{\partial \widetilde{A}_3}{\partial \widetilde{y}} + (c_7\widetilde{z} + \widetilde{c}_3)\,\frac{\partial \widetilde{A}_3}{\partial \widetilde{z}} = (c_7 - c_8)\,\widetilde{A}_3.
\end{aligned} \tag{5.29}$$

where $\widetilde{c}_3$ is also a constant depending on each case. The solution of this system can now be found easily. Notice that $\widetilde{x}, \widetilde{y}, \widetilde{z}$ describe cylindrical coordinates in the $\bar{x}\bar{y}\bar{z}$-space, which in turn is a linear transformation of the original space.

Summing up, transformation (5.26) uncouples the equations of the system, the same way (5.28) yields its characteristics. Under these two changes of variables we arrive at system (5.29) and then solve it. After the solution is found, the inverse transformations of (5.26), (5.28) give the solution of the original system (5.22). Though this treatment is general, we distinguish four characteristic cases, depending on the form of the matrix $P$, the vector $\boldsymbol{k}$ and the constant $\widetilde{c}_3$, as well as a last and more trivial one, where no transformation is necessary at all.

### 5.2.1 The case $c_7 \neq 0$ and $c_5$ or $c_6 \neq 0$

In this general case, owing to the skew-symmetry of $Q$ the matrix $P$ consisting of the eigenvectors of $Q_3$ can be orthogonal, $P^{-1} = P^T$. If we let $\epsilon_4 = \arctan(c_5/c_6)$ and $\epsilon_5 = \arctan\left(\sqrt{c_5^2 + c_6^2}/c_4\right)$, then $P = R_3(\epsilon_4)R_2(\epsilon_5)$, where $R_3$ and $R_2$ are the rotation matrices around the $z$- and $y$-axes, respectively. More explicitly

$$P = \frac{1}{c\left(c_5^2 + c_6^2\right)^{\frac{1}{2}}} \begin{pmatrix} c_4 c_6 & -c c_5 & c_6\left(c_5^2 + c_6^2\right)^{\frac{1}{2}} \\ c_4 c_5 & c c_6 & c_5\left(c_5^2 + c_6^2\right)^{\frac{1}{2}} \\ -\left(c_5^2 + c_6^2\right) & 0 & c_4\left(c_5^2 + c_6^2\right)^{\frac{1}{2}} \end{pmatrix}, \tag{5.30}$$

while $\boldsymbol{k} = -C^{-1}\boldsymbol{c}$, making (5.27) homogeneous in terms of $\bar{x}, \bar{y}, \bar{z}$. So, the new independent variables, according to (5.28), are

$$\begin{aligned} \widetilde{x} &= \left[(x-k_1)^2 + (y-k_2)^2 + (z-k_3)^2 - \widetilde{z}^2\right]^{\frac{1}{2}} \\ \widetilde{y} &= \arctan\left(\frac{c\left[c_6\left(y-k_2\right) - c_5\left(x-k_1\right)\right]}{c_4 c_6\left(x-k_1\right) + c_4 c_5\left(y-k_2\right) - \left(c_5^2 + c_6^2\right)\left(z-k_3\right)}\right) \\ \widetilde{z} &= \frac{1}{c}\left[c_6\left(x-k_1\right) + c_5\left(y-k_2\right) + c_4\left(z-k_3\right)\right] \end{aligned} \tag{5.31}$$

Together with (5.26), they transform system (5.22) into (5.29) for $\widetilde{c}_3 = 0$. Its solution is now easily found to be $\widetilde{A}_{1,3} = \widetilde{z}^{1-\frac{c_8}{c_7}}\,\widetilde{F}_{1,3}(u_1, u_2)$ and $\widetilde{A}_2 = \widetilde{y} + \widetilde{F}_2(u_1, u_2)$, where $\widetilde{F}_1$, $\widetilde{F}_2$ and $\widetilde{F}_3$ are arbitrary functions of the characteristics

$$\begin{aligned} u_1 &= \frac{\widetilde{z}}{\widetilde{x}}, \\ u_2 &= c\ln\widetilde{z} - c_7\widetilde{y}. \end{aligned} \tag{5.32}$$

Taking the inverse transformation of (5.26), the previous solution can be expressed back in the original components of the vector potential, arriving at

$$\boldsymbol{A}\left(\widetilde{x}, \widetilde{y}, \widetilde{z}\right) = \widetilde{z}^{1-\frac{c_8}{c_7}}\, P R_3(\widetilde{y})\, \boldsymbol{F}\left(u_1, u_2\right), \tag{5.33}$$

where $\boldsymbol{F} = (F_1, F_2, F_3)$ is an arbitrary vector function of $u_1$ and $u_2$, coming from the previous $\widetilde{\boldsymbol{F}} = (\widetilde{F}_1, \widetilde{F}_2, \widetilde{F}_3)$. In terms of the original coordinates, solution (5.33) is further reduced, though some of the transformed variables are still kept for simplicity,

$$\begin{aligned} A_1 &= \widetilde{z}^{-\frac{c_8}{c_7}} \left[ (c(x-k_1) - c_6\widetilde{z})\, F_1 + (c_5(z-k_3) - c_4(y-k_2))\, F_2 + c_6\widetilde{z}\, F_3 \right] \\ A_2 &= \widetilde{z}^{-\frac{c_8}{c_7}} \left[ (c(y-k_2) - c_5\widetilde{z})\, F_1 + (c_4(x-k_1) - c_6(z-k_3))\, F_2 + c_5\widetilde{z}\, F_3 \right] \\ A_3 &= \widetilde{z}^{-\frac{c_8}{c_7}} \left[ (c(z-k_3) - c_4\widetilde{z})\, F_1 + (c_6(y-k_2) - c_5(x-k_1))\, F_2 + c_4\widetilde{z}\, F_3 \right] \end{aligned} \tag{5.34}$$

Equation (5.21) on the other hand is very similar to either one of system (5.29), and can be solved likewise. So, the scalar potential is

$$\Phi = \begin{cases} \widetilde{z}^{2\left(1-\frac{c_8}{c_7}\right)} G, & c_8 \neq c_7 \\ \dfrac{c_9}{c}\,\widetilde{y} + G, & c_8 = c_7 \end{cases} \tag{5.35}$$

up to some additive constant, where $G$ is too an arbitrary of $u_1$ and $u_2$.

Since $k_i$'s are quite complicated expressions of $c_i$'s, $i = 1, 2, 3$, it is preferable in this case for practical purposes to consider the latter in terms of the former, and thus, to conclude that the potentials of the form (5.34)-(5.35) yield symmetry (5.18) for $c_i = -C_{ij}k_j$, $i = 1, 2, 3$.

### 5.2.2 The case $c_7 = 0$ and $c_5$ or $c_6 \neq 0$

Here we may use again the matrix $P$ given in (5.30), but, since $C$ is no longer invertible, we define $\boldsymbol{k} = c^{-2}C\boldsymbol{c}$ differently, $k_1 = (c_5c_3 - c_4c_2)/c^2$, $k_2 = (c_4c_1 - c_6c_3)/c^2$ and $k_3 = (c_6c_2 - c_5c_1)/c^2$. The new coordinates resemble the ones of the previous case,

$$\begin{aligned} \widetilde{x} &= \left[ (x-k_1)^2 + (y-k_2)^2 + (z-k_3)^2 - \widetilde{z}^2 \right]^{\frac{1}{2}} \\ \widetilde{y} &= \arctan\left( \frac{c\left[c_6\,(y-k_2) - c_5\,(x-k_1)\right]}{c_4c_6\,(x-k_1) + c_4c_5\,(y-k_2) - \left(c_5^2 + c_6^2\right)(z-k_3)} \right) \\ \widetilde{z} &= \frac{1}{c}\left(c_6x + c_5y + c_4z\right), \end{aligned} \tag{5.36}$$

but now $\widetilde{c}_3 = (c_6c_1 + c_5c_2 + c_4c_3)/c$. Thus, $\widetilde{A}_{1,3} = e^{-\frac{c_8}{c}\widetilde{y}}\,\widetilde{F}_{1,3}(u_1, u_2)$ and $\widetilde{A}_2 = \widetilde{y} + \widetilde{F}_2\,(u_1, u_2)$, where the arbitrary functions $\widetilde{F}_1, \widetilde{F}_2, \widetilde{F}_3$ now depend on the characteristics

$$\begin{aligned} u_1 &= \widetilde{x}, \\ u_2 &= c\widetilde{z} - \widetilde{c}_3\widetilde{y}, \end{aligned} \tag{5.37}$$

that actually being the essential difference with the first case. Thus, the vector potential in the original components is

$$\boldsymbol{A}\,(\widetilde{x}, \widetilde{y}, \widetilde{z}) = e^{-\frac{c_8}{c}\widetilde{y}} P R_3(\widetilde{y})\, \boldsymbol{F}\,(u_1, u_2)\,, \tag{5.38}$$

or more simply

$$\begin{aligned} A_1 &= e^{-\frac{c_8}{c}\widetilde{y}}\left[(c(x-k_1)-c_6\widetilde{z})\,F_1+(c_5(z-k_3)-c_4(y-k_2))\,F_2+c_6\,F_3\right] \\ A_2 &= e^{-\frac{c_8}{c}\widetilde{y}}\left[(c(y-k_2)-c_5\widetilde{z})\,F_1+(c_4(x-k_1)-c_6(z-k_3))\,F_2+c_5\,F_3\right] \\ A_3 &= e^{-\frac{c_8}{c}\widetilde{y}}\left[(c(z-k_3)-c_4\widetilde{z})\,F_1+(c_6(y-k_2)-c_5(x-k_1))\,F_2+c_4\,F_3\right] \end{aligned} \tag{5.39}$$

and the scalar one,

$$\Phi=\begin{cases} e^{-\frac{2c_8}{c}\widetilde{y}}\,G, & c_8\neq 0 \\ \dfrac{c_9}{c}\,\widetilde{y}+G, & c_8=0 \end{cases} \tag{5.40}$$

where $\boldsymbol{F}$ and $G$ are now arbitrary functions of $u_1$ and $u_2$ (5.37).

### 5.2.3 The case $c_7 \neq 0$ and $c_5 = c_6 = 0$

Now the third equation of system (5.22) is already isolated and so it can be solved independently, yielding $A_3$ directly. This means that simply $P = I$, which is actually the only difference with the first case. All the results obtained there can be reproduced here by simply making this substitution. In particular, setting again $\boldsymbol{k} = -C^{-1}\boldsymbol{c}$ and $\widetilde{c}_3 = 0$, the form of system (5.29) in the new variables

$$\begin{aligned} \widetilde{x} &= \sqrt{(x-k_1)^2+(y-k_2)^2} \\ \widetilde{y} &= \arctan\left(\frac{y-k_2}{x-k_1}\right) \\ \widetilde{z} &= z-k_3 \end{aligned} \tag{5.41}$$

is exactly as in the first case and consequently its solution remains the same. Therefrom, we end up with the following expression for the vector potential

$$\boldsymbol{A}\left(\widetilde{x},\widetilde{y},\widetilde{z}\right)=\widetilde{z}^{\,1-\frac{c_8}{c_7}}R_3(\widetilde{y})\,\boldsymbol{F}\left(u_1,u_2\right), \tag{5.42}$$

where the arbitrary function $\boldsymbol{F}$ depends again on $u_1$ and $u_2$ following the description (5.32), but for the variables $\widetilde{x},\widetilde{y},\widetilde{z}$ defined above, noting that $c = c_4$ in this case, i.e.

$$\begin{aligned} u_1 &= \frac{z-k_3}{\sqrt{(x-k_1)^2+(y-k_2)^2}}, \\ u_2 &= c_4\ln(z-k_3)-c_7\arctan\left(\frac{y-k_2}{x-k_1}\right), \end{aligned} \tag{5.43}$$

where $k_1 = -\left(c_7 c_1 + c_4 c_2\right)/\left(c_7^2 + c_4^2\right)$, $k_2 = \left(c_4 c_1 - c_7 c_2\right)/\left(c_7^2 + c_4^2\right)$ and $k_3 = -c_3/c_7$. Thus, in terms of the original coordinates the above solution is written as

$$\begin{aligned} A_1 &= (z-k_3)^{-\frac{c_8}{c_7}} \left[(x-k_1) F_1 - (y-k_2) F_2\right] \\ A_2 &= (z-k_3)^{-\frac{c_8}{c_7}} \left[(y-k_2) F_1 + (x-k_1) F_2\right] \\ A_3 &= (z-k_3)^{1-\frac{c_8}{c_7}} F_3 \end{aligned} \tag{5.44}$$

The scalar potential accordingly is

$$\Phi = \begin{cases} (z-k_3)^{2\left(1-\frac{c_8}{c_7}\right)} G, & c_8 \neq c_7 \\ \dfrac{c_9}{c_7} \ln(z-k_3) + G, & c_8 = c_7 \end{cases}, \tag{5.45}$$

where $G$ is an arbitrary function of $u_1$ and $u_2$ given in (5.43), as well.

When $c_4 = 0$, then the characteristic $u_2$ is just $\widetilde{y}$. Thus, solution (5.42) simply reduces to $\boldsymbol{A} = (z-k_3)^{1-\frac{c_8}{c_7}} \boldsymbol{F}(u_1, \widetilde{y})$. In other words there is no need for transformation (5.26) or (5.28) at all, because the system (5.22) in this case is completely uncoupled.

### 5.2.4 The case $c_5 = c_6 = c_7 = 0,\ c_4 \neq 0$

This is actually a combination of the last two cases. The new variables are now defined for $P = I$, as in the previous case, but $\boldsymbol{k} = c^{-2} C \boldsymbol{c}$ as in the second case, meaning

$$\begin{aligned} \widetilde{x} &= \sqrt{(x-k_1)^2 + (y-k_2)^2} \\ \widetilde{y} &= \arctan\left(\frac{y-k_2}{x-k_1}\right) \\ \widetilde{z} &= z \end{aligned} \tag{5.46}$$

where $k_1 = -c_2/c_4, k_2 = c_1/c_4$ and $k_3 = 0$. The transformed system (5.29) has the same form with the one of the second case for $c = c_4$ and $\widetilde{c}_3 = c_3$. Thus, the vector potential is

$$\boldsymbol{A}(\widetilde{x}, \widetilde{y}, \widetilde{z}) = e^{-\frac{c_8}{c_4}\widetilde{y}} R_3(\widetilde{y})\, \boldsymbol{F}(u_1, u_2), \tag{5.47}$$

or again

$$\begin{aligned} A_1 &= e^{-\frac{c_8}{c_4}\widetilde{y}} \left[(x-k_1) F_1 - (y-k_2) F_2\right] \\ A_2 &= e^{-\frac{c_8}{c_4}\widetilde{y}} \left[(y-k_2) F_1 + (x-k_1) F_2\right] \\ A_3 &= e^{-\frac{c_8}{c_4}\widetilde{y}} F_3 \end{aligned} \tag{5.48}$$

and the scalar one,

$$\Phi = \begin{cases} e^{-\frac{c_8}{c_4}\widetilde{y}} G, & c_8 \neq 0 \\ \dfrac{c_9}{c_4}\,\widetilde{y} + G, & c_8 = 0 \end{cases}, \tag{5.49}$$

where $\boldsymbol{F}$ and $G$ are arbitrary functions of the characteristics, which can be defined again through (5.37), but for the variables introduced in this case, meaning

$$\begin{aligned} u_1 &= \sqrt{(x-k_1)^2 + (y-k_2)^2}, \\ u_2 &= c_4 z - c_3 \arctan\left(\frac{y-k_2}{x-k_1}\right). \end{aligned} \tag{5.50}$$

### 5.2.5 The case $c_4 = c_5 = c_6 = c_7 = 0,\ c_1 \neq 0$

Now no transformation is needed at all. Equations (5.22) are uncoupled, and each one of them can be solved independently. The solution for the vector potential is

$$\boldsymbol{A} = e^{-\frac{c_8}{c_1}x} \boldsymbol{F}\left(c_2 x - c_1 y, c_3 x - c_1 z\right), \tag{5.51}$$

while for the scalar

$$\Phi = \begin{cases} e^{-2\frac{c_8}{c_1}x} G\left(c_2 x - c_1 y, c_3 x - c_1 z\right), & c_8 \neq 0 \\ \dfrac{c_9}{c_1}\, x + G\left(c_2 x - c_1 y, c_3 x - c_1 z\right), & c_8 = 0 \end{cases} \tag{5.52}$$

for some arbitrary vector function $\boldsymbol{F}$ and a scalar one $G$.

**Remark 5.1.** In all of the above cases, the constants $c_i$, $i = 1, \ldots, 8$ that appear in the symmetry $\mathbf{v}$ (5.18), are also present in the expressions for $\boldsymbol{A}$ and $\Phi$. In this way they define both the electromagnetic field and at the same time the symmetry admitted by system (5.1). In other words, the cases presented in this section represent families of potentials, which, apart from time translations, are compatible with Lie point symmetries of the form (5.18). So, $\mathbf{v}$ generates at least one symmetry, and along with $\mathbf{v}_0$ they span $n$-dimensional symmetry algebras, $n \geq 2$ for each case respectively. Further restrictions on the arbitrariness of $\boldsymbol{A}$ and $\Phi$ can specify exactly $n$, but these cases will be considered in the next chapter.

## 5.3 Noether point symmetries

In retrospect, the introduction of the potentials of the electromagnetic field leads immediately to a different viewpoint regarding equations (5.1) as an Euler-Lagrange system, stemming from the function (1.7) repeated here

$$L\left(\boldsymbol{x}, \dot{\boldsymbol{x}}\right) = \frac{1}{2}\,\dot{\boldsymbol{x}}^2 + \dot{\boldsymbol{x}} \cdot \boldsymbol{A}(\boldsymbol{x}) - \Phi(\boldsymbol{x}). \tag{5.53}$$

Subsequently, the question of Noether symmetries, providing us with first integrals of motion, naturally arises. Unlike the case of Lie point ones, now Maxwell's equations (1.4a) are obviously implied from the beginning.

Considering that Noether point symmetries are a subset of the Lie point ones, already found, we don't have to employ a new query from the beginning. We may check instead if and when the general form (5.17) of the symmetry generator is of variational type. Applying Theorem 4.37, the vector field (5.2) defines a Noether point symmetry if and only if it satisfies condition (4.39), or equivalently if a function $F(t, \boldsymbol{x})$ exists, such that

$$\xi \frac{\partial L}{\partial t} + \eta^i \frac{\partial L}{\partial x^i} + \left( \frac{d\eta^i}{dt} - \frac{d\xi}{dt} \dot{x}^i \right) \frac{\partial L}{\partial \dot{x}^i} + L \frac{d\xi}{dt} = \frac{dF}{dt}. \tag{5.54}$$

Substituting the time-independent Lagrangian given in (5.53) and the components of the generator from (5.17), results in a second degree polynomial in terms of the velocity components,

$$\eta^i \dot{x}^j \frac{\partial A_j}{\partial x^i} - \eta^i \frac{\partial \Phi}{\partial x^i} + \eta^i_{x^j} \dot{x}^i \dot{x}^j + \eta^i_{x^j} \dot{x}^j A_i - \frac{1}{2} \xi_t \dot{x}^i \dot{x}^i - \xi_t \Phi = F_t + F_{x^i} \dot{x}^i$$

This equation must hold identically for all $t, x^i, \dot{x}^i$ and so every coefficient of this polynomial must vanish, yielding the following five equations

$$2c_7 - c_8 = 0, \tag{5.55a}$$

$$(\boldsymbol{\eta} \cdot \nabla) \boldsymbol{A} + (c_7 I - Q)\, \boldsymbol{A} = \nabla F, \tag{5.55b}$$

$$-(\boldsymbol{\eta} \cdot \nabla) \Phi - c_8 \Phi = F_t. \tag{5.55c}$$

The first equation clearly rules the scalings $\mathbf{v}_7$ and $\mathbf{v}_8$ out as Noether symmetry candidates, allowing only $\mathbf{v}_7 + 2\mathbf{v}_8$. On the other hand, the integrability conditions of (5.55b) and (5.55c), $F_{x^i x^j} = F_{x^j x^i}$ and $F_{t x^i} = F_{x^i t}$, which guarantee the existence of $F$, expressed in terms of $\boldsymbol{B}$ and $\boldsymbol{E}$, lead back to (5.15) again for $c_8 = 2c_7$. Furthermore, from (5.55b) and (5.55c), we can easily deduce that

$$F(t, \boldsymbol{x}) = -c_9 t + f(\boldsymbol{x})$$

where $c_9$ is a constant and $f$ arbitrary. Thus, we recover the restrictions (5.20) and (5.21), for $c_8 = 2c_7$, which, going the other way round, were derived, when (5.15) were integrated.

**Remark 5.2.** It is worth noting that the spatial part $f$ of the so-called *gauge term* $F$ for the Noether symmetry condition was able to be absorbed by the gauge equivalence $\boldsymbol{A} \longrightarrow \boldsymbol{A} + \nabla g$ of the vector potential through (5.23).

So, the Noether symmetry condition in essence involves just (5.55a) and no other constraint. Of course the very expression of the Lagrange function, requires Maxwell's equations (1.4a) from the start. For a real physical problem these would be implemented in the Lie symmetry case, as well. On this ground, the conditions placed upon the electromagnetic field, in terms of $\boldsymbol{A}$ and $\Phi$, are the previous ones for $c_8 = 2c_7$. Once again the gauge invariance of the vector potential can be used as in section 5.2, where the restriction (5.23) for $c_8 = 2c_7$ now reduces to $\boldsymbol{\eta} \cdot \nabla g = f$. In conclusion, equations (5.1) admit Noether point symmetries of the general form

$$\begin{aligned} \xi &= 2c_7 t + c_0, \\ \eta^1 &= c_7 x - c_4 y + c_5 z + c_1, \\ \eta^2 &= c_7 y + c_4 x - c_6 z + c_2, \\ \eta^3 &= c_7 z - c_5 x + c_6 y + c_3 \end{aligned} \tag{5.56}$$

for electromagnetic fields coming from the potentials described earlier in the previous section, by setting $c_8 = 2c_7$. We emphasize again that only $c_0$ is completely arbitrary, corresponding to the symmetry $\mathbf{v}_0$ of time translations. The rest of the constants define a general Noether symmetry,

$$\mathbf{v} = \sum_{i=1}^{6} c_i \mathbf{v}_i + c_7 \left( \mathbf{v}_7 + 2\mathbf{v}_8 \right), \tag{5.57}$$

which the system has for a particular form of the electromagnetic field, specified again by them. While the invariant that corresponds to the first symmetry $\mathbf{v}_0$ is the well known Hamiltonian of the system, other integrals of motion may arise from $\mathbf{v}$.

## 5.4 Integrals of motion

According to Theorem 4.39, the integral of motion, which corresponds to a symmetry of (5.53), is given by relation (4.43). Inserting the Lagrangian for the system in the latter yields

$$I = \boldsymbol{\eta} \cdot (\dot{\boldsymbol{x}} + \boldsymbol{A}) - \xi \left( \frac{1}{2} \dot{\boldsymbol{x}}^2 + \Phi \right) - F. \tag{5.58}$$

Equations (5.55b) and (5.55c) for $\boldsymbol{\eta} = 0$ and $\xi = 1$ trivially result in a constant function $F$. Thus, from the symmetry $\mathbf{v}_0$, we recover through (5.58) the well-known Hamiltonian function (1.12) of the system, expressing the particle's energy. Before finding the integral that corresponds to the symmetry (5.57), we comment that another constant of motion would be of real value, if it is functionally independent of the Hamiltonian $H$ and in involution with it with respect to the corresponding Poisson bracket. For this to be the case we require $F$ to be time-independent,

that is $c_9 = 0$, and set $c_8 = 0$ that also leads through (5.55a) to $c_7 = 0$. Then, using the gauge transformation $\boldsymbol{A} \longrightarrow \boldsymbol{A} + \nabla g$ of section 5.2, $F = f$ will not enter at all in (5.58), since equation (5.23) for $c_8 = 2c_7$ yields $\boldsymbol{\eta} \cdot \nabla g = F$. Thus, we restrict our attention to linear integrals,

$$I\left(\boldsymbol{x}, \dot{\boldsymbol{x}}\right) = \boldsymbol{\eta} \cdot \left(\dot{\boldsymbol{x}} + \boldsymbol{A}\right), \tag{5.59}$$

either in terms of the velocities or the canonical momentums $p_i = L_{\dot{x}^i} = \dot{x}^i + A_i$. This is actually a generalisation of the relevant two-dimensional results in [31] and [48], and not included in Lewis's search for quadratic invariants [72]. It is worth noticing that in this way the scalar potential defines the form of the Hamiltonian, while the vector potential the form of the second integral.

We distinguish the following three cases, which correspond to the second, fourth and fifth one described in section 5.2 for $c_8 = c_9 = 0$.

### 5.4.1 The case $c_7 = 0$ and $c_5$ or $c_6 \neq 0$

The vector and scalar potential in this case are given in (5.39) and (5.40), respectively, for $c_8 = c_9 = 0$. The corresponding integral is

$$I = \left(-c_4 y + c_5 z + c_1\right)\dot{x} + \left(c_4 x - c_6 z + c_2\right)\dot{y} + \left(-c_5 x + c_6 y + c_3\right)\dot{z} + c^2 \widetilde{x}^{\,2} F_2 + c\,\widetilde{c}_3\, F_3 \tag{5.60}$$

defined only by two of the arbitary functions of the vector potential, $F_{2,3} = F_{2,3}\left(\widetilde{x}, \widetilde{c}_3 \widetilde{y} - c\widetilde{z}\right)$, where $\widetilde{x}, \widetilde{y}, \widetilde{z}$ are given in (5.36).

### 5.4.2 The case $c_5 = c_6 = c_7 = 0,\ c_4 \neq 0$

The vector and scalar potential in this case are given in (5.48) and (5.49), respectively, for $c_8 = c_9 = 0$. The corresponding integral is

$$I = \left(-c_4 y + c_1\right)\dot{x} + \left(c_4 x + c_2\right)\dot{y} + c_3 \dot{z} + c_4 \widetilde{x}^{\,2} F_2 + c_3 F_3 \tag{5.61}$$

defined only by two of the arbitary functions of the vector potential, $F_{2,3} = F_{2,3}\left(\widetilde{x}, c_3 \widetilde{y} - c_4 \widetilde{z}\right)$, where $\widetilde{x}, \widetilde{y}, \widetilde{z}$ are given in (5.46).

### 5.4.3 The case $c_4 = c_5 = c_6 = c_7 = 0,\ c_1 \neq 0$

The vector and scalar potential in this case are given in (5.51) and (5.52), respectively, for $c_8 = c_9 = 0$. The corresponding integral is

$$I = c_1 \dot{x} + c_2 \dot{y} + c_3 \dot{z} + c_1 F_1 + c_2 F_2 + c_3 F_3 \tag{5.62}$$

where $F_{1,2,3} = F_{1,2,3}\,(c_2x - c_1y, c_3x - c_1z)$ are the arbitrary functions entering the vector potential.

**Remark 5.3.** In all of the above cases the scalar potential reduces to an arbitrary function of the related characteristics. For each case there also exists a suitable coordinate system, where the vector potential is an arbitrary function of the corresponding characteristics, as well.

There is another justification for focusing on these types of integrals, which can be apparent when investigating the dynamics of the magnetic field itself. The latter offer a better insight into the magnetic fields found in section 5.2.

## 5.5 Magnetic field lines

In this section we compare, in terms of symmetries, system (5.1), which describes the particle's orbits, with the system of equations (1.18)

$$\frac{dx^i}{ds} = B^i(\boldsymbol{x}), \tag{5.63}$$

which describes the magnetic field lines. Recall that the independent variable $s$ in the above equations is related to the line element of these curves. From the physical point of view, since $t$ and $s$ carry very different meanings, such a comparison can only be made on the ground of spatial symmetries independent of them, i.e. generated by a vector field of the form $\mathbf{v} = \eta^i(\boldsymbol{x})\partial_{x^i}$.

For system (5.1) these can be recovered from section 5.1 for $\xi = 0$, meaning $c_0 = c_8 = 0$. So, in this case, the coefficients $\eta^i$ result from the beginning to the form (5.17), while the conditions (5.15) for the magnetic field are

$$(\boldsymbol{\eta}\cdot\nabla)\boldsymbol{B} = Q\boldsymbol{B}\,. \tag{5.64}$$

On the other hand, by prolonging $\mathbf{v}$ up now to the first derivatives $dx^i/ds$, the symmetry condition for system (5.63), $\mathbf{v}^{(1)}\big(dx^i/ds - B^i\big) = 0$ on its solutions, leads simply to

$$(\boldsymbol{\eta}\cdot\nabla)\boldsymbol{B} = (\boldsymbol{B}\cdot\nabla)\,\boldsymbol{\eta}, \tag{5.65}$$

without predefining the form of the symmetries. In order to compare these two equations, note that $Q\boldsymbol{B} = -c_7\boldsymbol{B} + (\boldsymbol{B}\cdot\nabla)\,\boldsymbol{\eta}$ for $\boldsymbol{\eta}$ given in (5.17).

From the above we conclude that, for an inhomogeneous, curved magnetic field, which satisfies equations (5.64), the symmetries of (5.1) cannot meet the requirements (5.65) for system (5.63) when $c_7 \neq 0$. On the other hand, for $c_7 = 0$, where conditions (5.64) and (5.65) are identical,

system (5.63) may have more symmetries than the linear ones of system (5.1). Thus, in general, *for inhomogeneous, curved magnetic fields not every symmetry of the charged particle motion is a symmetry of the magnetic field lines and vice versa. If, however, we limit our choices to symmetries of the form*

$$\mathbf{v} = \sum_{i=1}^{6} c_i \mathbf{v}_i, \tag{5.66}$$

*then these are admitted by both systems for the same magnetic field* $\boldsymbol{B}$.

Actually, we have already encountered $\mathbf{v}$: it is a Noether symmetry of (5.1), that corresponds to the integral (5.59) described in the previous section. However, recall that (5.63) is an Euler-Lagrange system, too, coming from the Lagrangian function (1.20) $\mathcal{L}(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{x}' \cdot \boldsymbol{A}(\boldsymbol{x})$. The Noether symmetry condition for these equations,

$$\eta^i \frac{\partial \mathcal{L}}{\partial x^i} + \frac{d\eta^i}{ds} \frac{\partial \mathcal{L}}{\partial x^{i'}} = \frac{df}{ds}, \tag{5.67}$$

is also satisfied by $\mathbf{v}$, giving $f_s = 0$ and the previous constraints on the vector potential, i.e. equations (5.20) for $c_7 = c_8 = 0$. Using the same gauge transformation to determine $\boldsymbol{A}$, the corresponding integral, $\eta^i L_{x^{i'}} - f$, for system (1.18) becomes

$$\mathcal{I} = \boldsymbol{\eta} \cdot \boldsymbol{A} \tag{5.68}$$

This is a projection of the integral (5.59) on the original configuration space $\mathbb{R}^3$, and therefore, according to the form of the symmetry and the vector potential, it, too, is seperated into three cases, which are the ones of the previous section without the velocities. In conclusion,

**Corollary 5.4.** Whenever the motion of the charged particle is confined in the hypersurface $\boldsymbol{\eta} \cdot (\dot{\boldsymbol{x}} + \boldsymbol{A}) = \text{const.}$, the magnetic field lines lie on the surface $\boldsymbol{\eta} \cdot \boldsymbol{A} = \text{const.}$, where $\boldsymbol{\eta} = Q\boldsymbol{x} + \boldsymbol{c}$. In this case both systems enjoy a symmetry of the form (5.66).

**Remark 5.5.** The system of the magnetic field lines is integrable, when it admits a symmetry $\mathbf{v}$ (5.66), as shown, for example, in [35]. Its Hamiltonian formulation in this case has also been studied in [84, 106]. Further results on the reduction of divergence-free vector fields, like the magnetic field, with divergence-free symmetries, like the above, can be found in [45] for the three-dimensional case and in [49, 121] for the $n$-dimensional one.

**Remark 5.6.** Note that the symmetries (5.66) consist only of translations and rotations, which are the Killing vector fields of $\mathbb{R}^3$ that leave invariant its metric: For example, if we switch to a coordinate system, in which the metric tensor is independent of one of the coordinates, say $\widetilde{x}$, then the vector field $\partial_{\widetilde{x}}$ is a Killing vector field.

## 5.6 Discussion

We have found five classes of stationary electromagnetic fields in terms of the potentials, which yield Lie point symmetries, besides time translations, for the three-dimensional autonomous non-relativistic charged particle motion. The analysis showed that when the system is nonlinear and in particular in the case of inhomogeneous and curved magnetic fields the only possible symmetries are linear, consisting of translations, rotations and dilatations. Rotations and translations have proven to be of Noether type, as well, while time and space dilatations can only survive the Noether symmetry condition as a specific linear combination. In any case no further restrictions on the potentials of the electromagnetic field were required. The corresponding invariants have been constructed, in particular focusing on three cases, where the integrals are functionally independent of the Hamiltonian and in involution with it. Thus, a total reduction of four can be made, and then further investigate the system in only two variables.

Another aspect of these three cases is that the same symmetry is also admitted by the system of the magnetic field lines, yielding an integral of motion for the latter, too. Time-independent magnetic fields lying on a surface, with some geometrical (usually axial or helical) symmetry are very often used to describe the equilibrium state of plasma configurations in the context of ideal Magnetohydrodynamics. The results obtained in this work may be useful for the determination of such symmetric magnetic surfaces, by finding new solutions of the Grad-Shafranov equation, for example. In addition, the comparison between the two integrals, (5.59) and (5.68), could also relate the behaviour, and possibly the confinement, of the charged particle to the dynamics of the magnetic field. This relation could be further analysed, after reduction of the particle's trajectories and integration of the magnetic field lines.

Finally, in order to see where do these results stand let us revisit axisymmetric systems and helical perturbations encountered in chapter 2. An axisymmetric magnetic field would be independent of the angle $\phi = \arctan(y/x)$ and therefore admitting the symmetry $\partial_\phi = x\partial_y - y\partial_x$. This is just a simple subcase of the ones found in section 5.2 and in particular 5.2.4, where all $c_i$ are zero except $c_4$. It corresponds accordingly just to $\mathbf{v}_4$ of (5.66), which is of Noether type for both the magnetic field lines and the charged particle motion, and so results in first integrals for both of them, as well. Then consider a helical perturbation first in the cylindrical geometry that is often used in applications as an approximation to the toroidal one. A helical cylindrical magnetic field of helicity $(m, n)$ admits the symmetry $n\mathbf{v}_3 + m\mathbf{v}_4$, which still belongs to the case 5.2.4 for $c_3 = n$ and $c_4 = m$ and all else $c_i$ equal to zero. Hence it is again admitted by both systems and also of the Noether kind for both of them too. Now let us come to helical

perturbations in the actual toroidal geometry, where most of these features are lost, mainly because (the determinant of) the metric defined by the toroidal coordinates depends on the poloidal angle $\theta$[1]. For instance, in terms of the Hamiltonian system (2.3) for the magnetic field lines observe that since the metric enters the symplectic structure of the system, even when the Hamiltonian function is independent of $\theta$, this does not mean that $\theta$ is an ignorable variable. Another implication is that even when the vector potential $\boldsymbol{A}$ (in covariant components) is independent of $\theta$, the magnetic field $\boldsymbol{B}$ (in contravariant components) is not. Thus we need to be specific about which components we are referring to, when we helically perturb a magnetic field in a toroidal configuration. The usual case for toroidal magnetic fields is considered in the unit base that is feasible for applications. In terms of this base, a "helical" toroidal magnetic field in the sense that its physical components depend on the poloidal and toroidal angles only through expressions of the form $m\theta - n\phi$ does not have helical symmetry; the system of the magnetic field lines involves contravariant components and will still depend on $\theta$. So, "helical" perturbations in the unit base do not actually reflect helical symmetry $n\partial_\theta + m\mathbf{v}_4$. This is the very reason why the intrinsic poloidal angle $\theta_{\text{in}}$ was introduced in the first place in chapter 2, which behaves very similarly to $\theta$. And as shown in section 2.2 helical perturbations in terms of $\theta_{\text{in}}$ and $\phi$ instead of $\theta$ and $\phi$ do result in symmetry $n\partial_{\theta_{\text{in}}} + m\mathbf{v}_4$ corresponding to the invariance of the effective Hamiltonian $h$. Moreover in contrast with the cylindrical case, neither one of these two symmetries is admitted by the charged particle motion judging by the nonlinear form of either $\partial_\theta$ or worse $\partial_{\theta_{\text{in}}}$ as opposed to the linear ones (5.18).

From the above, we conclude that the symmetry analysis presented in this chapter does not lie far from applications. Although the symmetries found were only linear, they seem to capture a non-negligible part of the considerations used in real problems. In any case, it can be used as a base to draw first conclusions. Furthermore, for the unperturbed systems we see that compared to axisymmetric systems and the single symmetry $\mathbf{v}_4$ there is room for more general magnetic fields among the cases of 5.2.2 and 5.2.4 and the more general symmetry $\mathbf{v}$ (5.66) admitted again by both the magnetic field lines and the charged particle motion as a Noether symmetry. And under the aforementioned or similar perturbations symmetries may still survive with cylindrical models standing a better chance than toroidal ones.

The results presented in this chapter have also been published in [63].

1. In light of Remark 5.6, this already rules out $\partial_\theta$ from (5.66); on the contrary note that $\phi$ either used as polar angle in cylindrical coordinates or toroidal angle in toroidal ones does not enter the corresponding metric.

# Chapter 6

# Group Classification of Charged Particle Motion

In the previous chapter and particularly section 5.2, we have found the form of electromagnetic fields allowing at least one additional symmetry besides time translations for the autonomous charged particle motion. Their description, however, is specified by the symmetry as if the latter was given. But instead it is preferable to have the fields determining the symmetries in order to tell whether a particular system, i.e. for a given field, admits any extra symmetries or not. On the other hand, from the cases presented one cannot detect whether a given electromagnetic field belongs to more than one subcase and therefore resulting in two or more extra symmetries. A complete answer to both of these questions would require a vast list of every single case involved. This obstacle can be overcome with the aid of equivalence transformations, which can be used to identify systems that are mapped to each other and therefore can be grouped together. So, in this chapter we find first the equivalence group for the class of the three-dimensional nonrelativistic motion of charged particles in time-independent electromagnetic fields. And then using the equivalence transformations we give a symmetry group classification for this class. Finally, we arrive at some first results regarding the integrability of the system.

## 6.1 Equivalence transformations in terms of fields

We start by rewriting system (5.1) as a class of differential equations enlarged by the auxiliary equations expressing the time-independence of the fields,

$$\begin{aligned} B^i_t &= 0 \\ E^i_t &= 0 \end{aligned} \tag{6.1a}$$

$$\ddot{x}^i - \epsilon_{ijk}\dot{x}^j B^k - E^i = 0 \tag{6.1b}$$

and we consider augmented equivalence transformations for the class (6.1) generated by the vector field

$$\mathbf{V} = \xi(t,\boldsymbol{x})\frac{\partial}{\partial t} + \eta^i(t,\boldsymbol{x})\frac{\partial}{\partial x^i} + \mu^i(t,\boldsymbol{x},\boldsymbol{B},\boldsymbol{E})\frac{\partial}{\partial B^i} + \nu^i(t,\boldsymbol{x},\boldsymbol{B},\boldsymbol{E})\frac{\partial}{\partial E^i}. \tag{6.2}$$

Following formula (4.55), we prolong $\mathbf{V}$ up to second-order derivatives in terms of $x^i$ and first-order ones in terms of $B^i$ and $E^i$,

$$\mathbf{V}^{(2,1)} = \mathbf{v}^{(2)} + \mu^i\frac{\partial}{\partial B^i} + \nu^i\frac{\partial}{\partial E^i} + \mu^t_i\frac{\partial}{\partial B^i_t} + \mu^{x^j}_i\frac{\partial}{\partial B^i_{x^j}} + \nu^t_i\frac{\partial}{\partial E^i_t} + \nu^{x^j}_i\frac{\partial}{\partial E^i_{x^j}} \tag{6.3}$$

(actually spatial derivatives for the fields are not required here, but we include them anyway for future reference), where $\mathbf{v} = \xi\partial_t + \eta^i\partial_{x^i}$ is as always the projection of $\mathbf{V}$ and so $\mathbf{v}^{(2)}$ is again given by equation (5.4), while

$$\begin{aligned}
\mu^t_i &= D_t\mu^i - B^i_t D_t\xi - B^i_{x^j} D_t\eta^j\\
\nu^t_i &= D_t\nu^i - E^i_t D_t\xi - E^i_{x^j} D_t\eta^j\\
\mu^{x^j}_i &= D_{x^j}\mu^i - B^i_t D_{x^j}\xi - B^i_{x^k} D_{x^j}\eta^k\\
\nu^{x^j}_i &= D_{x^j}\nu^i - E^i_t D_{x^j}\xi - E^i_{x^k} D_{x^j}\eta^k
\end{aligned}$$

Applying then Theorem 4.47, the vector field $\mathbf{V}$ is a generator of equivalence transformations if and only if

$$\begin{aligned}
\mathbf{V}^{(2,1)}(B^i_t) &= 0\\
\mathbf{V}^{(2,1)}(E^i_t) &= 0\\
\mathbf{V}^{(2,1)}(\ddot{x}^i - \epsilon_{ijk}\dot{x}^j B^k - E^i) &= 0
\end{aligned} \tag{6.4}$$

whenever (6.1) hold. After substitution of the latter, the above conditions (6.4) for the auxiliary and the primary system,

$$\begin{aligned}
0 &= \mu^i_t - \eta^j_t B^i_{x^j}\\
0 &= \nu^i_t - \eta^j_t E^i_{x^j}\\
0 &= \eta^i_{tt} + 2\eta^i_{tx^j}\dot{x}^j + \eta^i_{x^jx^k}\dot{x}^j\dot{x}^k + \eta^i_{x^t}\epsilon_{ljk}\dot{x}^j B^k + \eta^i_{x^j}E^j - \xi_{tt}\dot{x}^i - 2\xi_{tx^j}\dot{x}^i\dot{x}^j - 2\xi_t\epsilon_{ijk}\dot{x}^j B^k -\\
&\quad - 2\xi_t E^i - \xi_{x^jx^k}\dot{x}^i\dot{x}^j\dot{x}^k - 2\xi_{x^j}\dot{x}^j\epsilon_{ikl}\dot{x}^k B^l - 2\xi_{x^j}\dot{x}^j E^i - \xi_{x^j}\dot{x}^i\epsilon_{jkm}\dot{x}^k B^m - \xi_{x^j}\dot{x}^i E^j -\\
&\quad - \epsilon_{ijk}\eta^j_t B^k - \epsilon_{ilk}\eta^l_{x^j}\dot{x}^j B^k + \epsilon_{ijk}\xi_t\dot{x}^j B^k + \epsilon_{ijl}\xi_{x^k}\dot{x}^j\dot{x}^k B^l - \epsilon_{ijk}\dot{x}^j\mu^k - \nu^i
\end{aligned} \tag{6.5}$$

are now viewed as polynomials in terms of the derivatives of both $x^i$ and $B^i$, $E^i$ as well. Therefore, as identities for all $t$, $x^i$, $B^i$, $E^i$, $\dot{x}^i$, $B^i_{x^j}$ and $E^i_{x^j}$ they break up into a larger set of partial

differential equations again. These are the equations that determine the equivalence algebra, coming now from the coefficients of the monomials in the first derivatives of $\boldsymbol{x}$, $\boldsymbol{B}$ and $\boldsymbol{E}$,

$$\begin{aligned}
\eta^i_t &= 0\\
\mu^i_t &= 0\\
\nu^i_t &= 0\\
\xi_{x^jx^k} &= 0\\
\eta^i_{x^jx^k} &= 0\\
\xi_{tx^i} &= 0\\
\xi_{x^i} &= 0\\
\xi_{tt} &= 0\\
\eta^i_{tx^j} &= 0\\
\eta^i_{x^j} + \eta^j_{x^i} &= 0 \quad \text{for} \quad i \neq j\\
\epsilon_{ljk}\eta^i_{x_l}B^k - \epsilon_{ijk}\xi_t B^k - \epsilon_{ilk}\eta^l_{x_j}B^k - \epsilon_{ijk}\mu^k &= 0 \quad \text{for} \quad i \neq j\\
\eta^i_{tt} + \eta^i_{x^j}E^j - 2\xi_t E^i - \epsilon_{ijk}\eta^j_t B^k - \nu^i &= 0
\end{aligned} \tag{6.6}$$

It is worth mentioning that at the beginning the equivalence condition is not entirely different from the symmetry condition; note the resemblance between (5.5) and the third of equations (6.5). Thus, for example, the terms of (5.6b) appear here also exactly the same way. But while $B^i$ and $E^i$ are treated there as functions of $t$ and $\boldsymbol{x}$, here at this stage they are considered as variables independent of them. This is why equations (5.6b) now split even further to the fifth, sixth and seventh of equations (6.6) listed above. Note how the different treatment of the arbitrary functions $B^i$ and $E^i$ simplified things, compared to the symmetry condition. Continuing, system (6.6) is easily reduced to

$$\begin{aligned}
\xi_{tt} &= 0\\
\xi_{x^i} &= 0\\
\eta^i_t &= 0\\
\eta^i_{x^jx^k} &= 0\\
\eta^i_{x^j} + \eta^j_{x^i} &= 0 \quad \text{for } i \neq j\\
\eta^i_{x^j} - \eta^l_{x^k} &= 0 \quad \text{for } i = j,\ l = k\\
-\big(\eta^1_x + \xi_t\big)B^i + \eta^i_{x^j}B^j - \mu^i &= 0\\
-2\xi_t E^i + \eta^i_{x^j}E^j - \nu^i &= 0
\end{aligned} \tag{6.7}$$

Again, one cannot help noticing how conditions (5.7), (5.8), (5.11), (5.12) and $\dot{f}_{ij} = 0$ for inhomogeneous, curved magnetic fields yielded there after detailed investigation have now emerged most of them directly and in any case much more simply. From (6.7) we deduce the form of the equivalence generator

$$\begin{aligned}
\xi &= c_8 t + c_0 \\
\eta^1 &= c_7 x - c_4 y + c_5 z + c_1 \\
\eta^2 &= c_7 y + c_4 x - c_6 z + c_2 \\
\eta^3 &= c_7 z - c_5 x + c_6 y + c_3 \\
\mu^1 &= -c_8 B^1 - c_4 B^2 + c_5 B^3 \\
\mu^2 &= -c_8 B^2 + c_4 B^1 - c_6 B^3 \\
\mu^3 &= -c_8 B^3 - c_5 B^1 + c_6 B^2 \\
\nu^1 &= (c_7 - 2c_8)\, E^1 - c_4 E^2 + c_5 E^3 \\
\nu^2 &= (c_7 - 2c_8)\, E^2 + c_4 E^1 - c_6 E^3 \\
\nu^3 &= (c_7 - 2c_8)\, E^3 - c_5 E^1 + c_6 E^2
\end{aligned} \tag{6.8}$$

where $c_i, i = 0, \ldots, 8$ are arbitrary constants. Hence the group of equivalence transformations for the class (6.1) is generated by a 9-dimensional Lie algebra spanned by the vector fields

$$\begin{aligned}
\mathbf{V}_0 &= \partial_t \\
\mathbf{V}_1 &= \partial_x \\
\mathbf{V}_2 &= \partial_y \\
\mathbf{V}_3 &= \partial_z \\
\mathbf{V}_4 &= x\partial_y - y\partial_x + B^1\partial_{B^2} - B^2\partial_{B^1} + E^1\partial_{E^2} - E^2\partial_{E^1} \\
\mathbf{V}_5 &= z\partial_x - x\partial_z + B^3\partial_{B^1} - B^1\partial_{B^3} + E^3\partial_{E^1} - E^1\partial_{E^3} \\
\mathbf{V}_6 &= y\partial_z - z\partial_y + B^2\partial_{B^3} - B^3\partial_{B^2} + E^2\partial_{E^3} - E^3\partial_{E^2} \\
\mathbf{V}_7 &= x\partial_x + y\partial_y + z\partial_z + E^1\partial_{E^1} + E^2\partial_{E^2} + E^3\partial_{E^3} \\
\mathbf{V}_8 &= t\partial_t - B^1\partial_{B^1} - B^2\partial_{B^2} - B^3\partial_{B^3} - 2E^1\partial_{E^1} - 2E^2\partial_{E^2} - 2E^3\partial_{E^3}
\end{aligned} \tag{6.9}$$

As we can see the projections of (6.9) are none other than $\mathbf{v}_i$, $i = 0, \ldots, 8$ encountered in section 5.1. From Corollary 4.51, we see that $\mathbf{V}(B^i - B^i(\boldsymbol{x})) = \mathbf{V}(E^i - E^i(\boldsymbol{x})) = 0$ for any solution $(\boldsymbol{B}(\boldsymbol{x}), \boldsymbol{E}(\boldsymbol{x}))$ of the auxiliary system (6.1a) if and only if $\eta^i = \mu^i = \nu^i = 0$. Thus, the principal Lie algebra generating symmetries admitted by the primary system (6.1b) for any $\boldsymbol{B}$ and $\boldsymbol{E}$ is just the well expected $\mathbf{v}_0 = \partial_t$. Note that for the rest equivalence algebra $\mathbf{V}_8$ is the center.

## 6.2 The homogeneous Maxwell's equations as auxiliary conditions

Now let us focus again to electromagnetic fields respecting Maxwell's equations (1.4a), as well. The latter can also be inserted as auxiliary equations, further enlarging the auxiliary system and thus restricting the original class (6.1) to

$$\begin{aligned} B^i_t &= 0 \\ E^i_t &= 0 \end{aligned} \tag{6.10a}$$

$$\begin{aligned} B^i_{x^i} &= 0 \\ \epsilon_{ijk} E^k_{x^j} &= 0 \end{aligned} \tag{6.10b}$$

$$\ddot{x}^i - \epsilon_{ijk} \dot{x}^j B^k - E^i = 0 \tag{6.10c}$$

Here, the equivalence condition for this class apart from (6.4) now also includes

$$\begin{aligned} \mathbf{V}^{(2,1)}(B^i_{x^i}) &= 0 \\ \mathbf{V}^{(2,1)}(\epsilon_{ijk} E^k_{x^j}) &= 0 \end{aligned} \tag{6.11}$$

whenever (6.10a)-(6.10b) hold. Substituting at first (6.10a), relations (6.11) add to system (6.5) accordingly the set of equations

$$\begin{aligned} \mu^i_{x^i} + \mu^i_{B^j} B^j_{x^i} + \mu^i_{E^j} E^j_{x^i} - \eta^j_{x^i} B^i_{x^j} &= 0 \\ \epsilon_{ijk} (\nu^i_{x^j} + \nu^i_{B^l} B^l_{x^j} + \nu^i_{E^l} E^l_{x^j} - \eta^l_{x^j} E^i_{x^l}) &= 0 \end{aligned}$$

and followed by replacements of spatial derivatives using (6.10b), they yield in turn additional equations, supplementary to the original ones (6.7) and interestingly enough without altering them,

$$\begin{aligned} \mu^i_{x^i} &= 0 \\ \mu^i_{B^j} - \eta^i_{x^j} &= 0 \quad \text{for} \quad i \neq j \\ \mu^i_{B^j} - \eta^i_{x^j} - \mu^l_{B^k} + \eta^l_{x^k} &= 0 \quad \text{for} \quad i = j,\ l = k \\ \mu^i_{E^j} - \mu^j_{E^i} &= 0 \quad \text{for} \quad i \neq j \\ \mu^i_{E^j} &= 0 \quad \text{for} \quad i = j \\ \nu^i_{x^j} - \nu^j_{x^i} &= 0 \quad \text{for} \quad i \neq j \\ \nu^i_{B^j} &= 0 \\ \nu^i_{E^j} + \eta^j_{x^i} &= 0 \quad \text{for} \quad i \neq j \\ \nu^i_{E^j} + \eta^j_{x^i} - \nu^l_{E^k} - \eta^k_{x^l} &= 0 \quad \text{for} \quad i = j,\ l = k \end{aligned} \tag{6.12}$$

It is easily verified that the latter satisfy the solution (6.8) of the former without any further restrictions whatsoever. In other words, we conclude that:

*The class of differential equations (6.1) admits the same equivalence group with the subclass (6.10).*

In retrospect, recalling Definition 4.45, note that $\mathbf{V}_i$ for $i = 0, \dots, 6$ are indeed well-known symmetry generators of Maxwell's equations altogether in vacuum (meaning system (1.4) for $\sigma = \boldsymbol{J} = 0$) that belong to the Lie algebra of the Lorentz group (see for example [55]). However, that system only admits the linear combination $\mathbf{V}_7 + \mathbf{V}_8$, contrary to (1.4a) which as we can see admits both $\mathbf{V}_7$ and $\mathbf{V}_8$ separately.

On the other hand, the above conclusion is rather unusual in the following sense: additional constraints may either lead to new equivalence transformations, known as conditional, or they can further restrict the equivalence generator resulting in an equivalence subgroup. From this point of view, it is worthy of noting that the equivalence group stayed the same.

## 6.3 Equivalence transformations in terms of potentials

Since Maxwell's equations (1.4a) did not change the scenery, it is preferable to express the equivalence problem for charged particle motion in terms of the potentials of the electromagnetic field. The reason is that although a classification can be given in terms of the fields, solutions of the auxiliary part (6.10b) can only be given in terms of the potentials. Therefore, we express the class (6.10) equally as

$$\begin{aligned} A^i_t &= 0 \\ \Phi_t &= 0 \end{aligned} \tag{6.13a}$$

$$\ddot{x}^i - \dot{x}^j(A^j_{x^i} - A^i_{x^j}) + \Phi_{x^i} = 0 \tag{6.13b}$$

for which we consider augmented equivalence transformations generated by the vector field

$$\mathbf{V} = \xi(t,\boldsymbol{x})\frac{\partial}{\partial t} + \eta^i(t,\boldsymbol{x})\frac{\partial}{\partial x^i} + \mu^i(t,\boldsymbol{x},\boldsymbol{A},\Phi)\frac{\partial}{\partial A^i} + \nu(t,\boldsymbol{x},\boldsymbol{A},\Phi)\frac{\partial}{\partial \Phi}. \tag{6.14}$$

As previously shown, we prolong $\mathbf{V}$ again up to second-order derivatives for $x^i$ and first-order temporal ones for $A^i$ and $\Phi$, and apply the equivalence condition

$$\begin{aligned} \mathbf{V}^{(2,1)}(A^i_t) &= 0 \\ \mathbf{V}^{(2,1)}(\Phi_t) &= 0 \\ \mathbf{V}^{(2,1)}(\ddot{x}^i - \dot{x}^j(A^j_{x^i} - A^i_{x^j}) + \Phi_{x^i}) &= 0 \end{aligned} \tag{6.15}$$

whenever equations (6.13) hold. After substitutions the equations coming from (6.15) that determine the equivalence algebra are reduced to the system

$$\begin{aligned}
\xi_{tt} &= 0\\
\xi_{x^i} &= 0\\
\eta^i_t &= 0\\
\eta^i_{x^jx^k} &= 0\\
\eta^i_{x^j} + \eta^j_{x^i} &= 0 \quad \text{for } i \neq j\\
\eta^i_{x^j} - \eta^l_{x^k} &= 0 \quad \text{for } i = j,\ l = k\\
\mu^i_t &= 0\\
\mu^i_\Phi &= 0\\
\mu^i_{x^j} - \mu^j_{x^i} &= 0 \quad \text{for } i \neq j\\
\mu^i_{A^j} + \eta^j_{x^i} &= 0 \quad \text{for } i \neq j\\
\mu^i_{A^j} - \eta^i_{x^j} + \xi_t &= 0 \quad \text{for } i = j\\
\nu_t &= 0\\
\nu_{x^i} &= 0\\
\nu_{A^i} &= 0\\
\nu_\Phi - 2\left(\eta^i_{x^j} - \xi_t\right) &= 0 \quad \text{for } i = j
\end{aligned} \tag{6.16}$$

Their solution yields the following form for the equivalence generator for the class of differential equations (6.13)

$$\begin{aligned}
\xi &= c_8 t + c_0\\
\eta^1 &= c_7 x - c_4 y + c_5 z + c_1\\
\eta^2 &= c_7 y + c_4 x - c_6 z + c_2\\
\eta^3 &= c_7 z - c_5 x + c_6 y + c_3\\
\mu^1 &= (c_7 - c_8) A^1 - c_4 A^2 + c_5 A^3 + f_x\\
\mu^2 &= (c_7 - c_8) A^2 + c_4 A^1 - c_6 A^3 + f_y\\
\mu^3 &= (c_7 - c_8) A^3 - c_5 A^1 + c_6 A^2 + f_z\\
\nu &= 2(c_7 - c_8)\Phi + c_9
\end{aligned} \tag{6.17}$$

where $c_i, i = 0, \dots, 9$ are arbitrary constants and $f$ is an arbitrary function of $\boldsymbol{x}$. Thus the Lie algebra of equivalence transformations consists of the 10-dimensional subalgebra spanned by the

vector fields

$$\begin{aligned}
\mathbf{V}_0 &= \partial_t \\
\mathbf{V}_1 &= \partial_x \\
\mathbf{V}_2 &= \partial_y \\
\mathbf{V}_3 &= \partial_z \\
\mathbf{V}_4 &= x\partial_y - y\partial_x + A^1\partial_{A^2} - A^2\partial_{A^1} \\
\mathbf{V}_5 &= z\partial_x - x\partial_z + A^3\partial_{A^1} - A^1\partial_{A^3} \\
\mathbf{V}_6 &= y\partial_z - z\partial_y + A^2\partial_{A^3} - A^3\partial_{A^2} \\
\mathbf{V}_7 &= x\partial_x + y\partial_y + z\partial_z + A^1\partial_{A^1} + A^2\partial_{A^2} + A^3\partial_{A^3} + 2\Phi\partial_\Phi \\
\mathbf{V}_8 &= t\partial_t - A^1\partial_{A^1} - A^2\partial_{A^2} - A^3\partial_{A^3} - 2\Phi\partial_\Phi \\
\mathbf{V}_9 &= \partial_\Phi
\end{aligned} \tag{6.18}$$

and the infinite-dimensional one $\mathbf{V}_f = \nabla f \cdot \partial_{\boldsymbol{A}}$. We see again that the symmetry generators included in the equivalence algebra are the same $\mathbf{v}_i$, $i = 0, \ldots, 8$ of section 5.1, while the principal Lie algebra consists only of $\mathbf{v}_0$. Accordingly, the equivalence group consists of the transformations

$$\begin{aligned}
\widetilde{t} &= \epsilon_8 t + \epsilon_0 \\
\widetilde{\boldsymbol{x}} &= \epsilon_7 R_1(\epsilon_6) R_2(\epsilon_5) R_3(\epsilon_4)\boldsymbol{x} + \boldsymbol{\epsilon} \\
\widetilde{\boldsymbol{A}} &= \epsilon_7 \epsilon_8^{-1} R_1(\epsilon_6) R_2(\epsilon_5) R_3(\epsilon_4)\boldsymbol{A} + \nabla g \\
\widetilde{\Phi} &= \epsilon_7^2 \epsilon_8^{-2}\Phi + \epsilon_9
\end{aligned} \tag{6.19}$$

where $\epsilon_i$, $i = 0, \ldots, 9$, are arbitrary constants for $\epsilon_{7,8} \neq 0$, $g$ is an arbitrary function of $\boldsymbol{x}$, $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3)$, and $R_i$ are the rotation matrices around $x^i$, respectively. Besides the well-known gauge equivalence, one also recognizes the linear part of the transformations (5.26) and (5.28) used in the previous chapter as a subgroup of (6.19) (see (5.30) for $P^{-1}$). Note also that the reflections ($\widetilde{t} = -t$, $\widetilde{\boldsymbol{x}} = -\boldsymbol{x}$) and ($\widetilde{t} = -t$, $\widetilde{\boldsymbol{A}} = -\boldsymbol{A}$) yield discrete equivalence transformations.

## 6.4 Classifying equations

Having found the equivalence transformations, we proceed with the symmetry analysis of the class (6.13) for particular forms of the electromagnetic field. Consider the linear combination

$$\overline{\mathbf{V}} = \sum_{i=1}^{9} c_i \mathbf{V}_i, \tag{6.20}$$

of the equivalence generators (6.18) leaving aside the principal generator $\mathbf{V}_0$. As previously noted, the projection $\mathbf{v}$ of $\mathbf{V} = \overline{\mathbf{V}} + \mathbf{V}_f$ to the $(t, \boldsymbol{x})$-space is the linear combination of the symmetries $\mathbf{v}_i$ (5.19),

$$\mathbf{v} = \sum_{i=1}^{8} c_i \mathbf{v}_i \tag{6.21}$$

Applying Proposition 4.50, the vector field $\mathbf{v}$ is a symmetry generator for the primary system (6.13b) for specific values

$$\begin{aligned} A^i &= A^i(\boldsymbol{x}) \\ \Phi &= \Phi(\boldsymbol{x}) \end{aligned} \tag{6.22}$$

if and only if (6.22) are invariant solutions for the auxiliary system (6.13a), that is $\mathbf{V}$ itself is admitted by system (6.22),

$$\begin{aligned} \mathbf{V}\left(A^i - A^i(\boldsymbol{x})\right) &= 0 \\ \mathbf{V}\left(\Phi - \Phi(\boldsymbol{x})\right) &= 0 \end{aligned} \tag{6.23}$$

Noting that, in general, $\mu^i$ and $\nu$ from (6.17) are just the right-hand sides of (5.20)-(5.21), the above conditions (6.23) lead back again to restrictions (5.20)-(5.21) encountered in the previous chapter, also repeated here,

$$\begin{aligned} (\boldsymbol{\eta} \cdot \nabla)\boldsymbol{A} &= \ (c_7 - c_8)\boldsymbol{A} + Q\boldsymbol{A} + \nabla f, \\ (\boldsymbol{\eta} \cdot \nabla)\Phi &= 2(c_7 - c_8)\Phi \ + c_9 \end{aligned} \tag{6.24}$$

where $\boldsymbol{\eta}$ denotes again the vector with entries $\eta^i$ listed here in (6.17) and $Q$ is the matrix (5.16). Of course, the general solution to the above system for the electromagnetic potential expressed in terms of the symmetry generator has been found in the previous chapter and classified in five characteristic cases in section 5.2. There $\boldsymbol{A}$ and $\Phi$ were described as general functions of the constants $c_i$, $i = 1, \ldots, 9$ that determine the general form of the symmetry generator (6.21). In other words, the vector and scalar potentials were treated as the unknowns as if the symmetries admitted were given. Now this treatment is reversed: preferably we want to know given an electromagnetic field if and which symmetries are admitted. In a manner of speaking we want to solve (6.24) for the symmetry generator in terms of the potentials, and ultimately classify these solutions in terms of symmetries. The classification severely reduces the amount of labor to a few cases that correspond to systems that cannot be mapped to one another through an equivalence transformation (6.19).

One way to begin with would be to classify the solutions of (6.24) already found in section 5.2 under (6.19), i.e. separate them into classes of electromagnetic fields that cannot be mapped to one another through (6.19). Their highly-complicated form leaves quite a space for human error though. An alternative would be to make the classification first, before solving them. However, employing (6.19) directly to (6.24) and separating it to disjoint subcases still involves a great amount of work. Instead we can take the equivalence algebra and decompose it into disjoint subalgebras, as described towards the end of section 4.8. Starting with one-dimensional subalgebras, we can continue to higher-dimensional ones that will provide us with cases of potentials corresponding to more than one additional symmetry, which could be very difficult to identify from the solutions of section 5.2.

## 6.5 Classification of equivalence subalgebras

In order to put into action the classification scheme outlined at the end of section 4.8, we start off with the finite-dimensional subalgebra and find first the Lie brackets of the equivalence generators $\mathbf{V}_i$ given by the next table.

| $[\ ,\ ]$ | $\mathbf{V}_1$ | $\mathbf{V}_2$ | $\mathbf{V}_3$ | $\mathbf{V}_4$ | $\mathbf{V}_5$ | $\mathbf{V}_6$ | $\mathbf{V}_7$ | $\mathbf{V}_8$ | $\mathbf{V}_9$ |
|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{V}_1$ | 0 | 0 | 0 | $\mathbf{V}_2$ | $-\mathbf{V}_3$ | 0 | $\mathbf{V}_1$ | 0 | 0 |
| $\mathbf{V}_2$ | 0 | 0 | 0 | $-\mathbf{V}_1$ | 0 | $\mathbf{V}_3$ | $\mathbf{V}_2$ | 0 | 0 |
| $\mathbf{V}_3$ | 0 | 0 | 0 | 0 | $\mathbf{V}_1$ | $-\mathbf{V}_2$ | $\mathbf{V}_3$ | 0 | 0 |
| $\mathbf{V}_4$ | $-\mathbf{V}_2$ | $\mathbf{V}_1$ | 0 | 0 | $\mathbf{V}_6$ | $-\mathbf{V}_5$ | 0 | 0 | 0 |
| $\mathbf{V}_5$ | $\mathbf{V}_3$ | 0 | $-\mathbf{V}_1$ | $-\mathbf{V}_6$ | 0 | $\mathbf{V}_4$ | 0 | 0 | 0 |
| $\mathbf{V}_6$ | 0 | $-\mathbf{V}_3$ | $\mathbf{V}_2$ | $\mathbf{V}_5$ | $-\mathbf{V}_4$ | 0 | 0 | 0 | 0 |
| $\mathbf{V}_7$ | $-\mathbf{V}_1$ | $-\mathbf{V}_2$ | $-\mathbf{V}_3$ | 0 | 0 | 0 | 0 | 0 | $-2\mathbf{V}_9$ |
| $\mathbf{V}_8$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $2\mathbf{V}_9$ |
| $\mathbf{V}_9$ | 0 | 0 | 0 | 0 | 0 | 0 | $2\mathbf{V}_9$ | $-2\mathbf{V}_9$ | 0 |

Table 6.1: Lie brackets between the equivalence generators.

From the above matrix we can construct the generators (4.62) of the adjoint action on the parameters $c_i$, $i = 1, \dots, 9$, which can be put into form

$$\begin{aligned}
\mathbf{C}_1 &= c_7\partial_{c_1} + c_4\partial_{c_2} - c_5\partial_{c_3}\\
\mathbf{C}_2 &= -c_4\partial_{c_1} + c_7\partial_{c_2} + c_6\partial_{c_3}\\
\mathbf{C}_3 &= c_5\partial_{c_1} - c_6\partial_{c_2} + c_7\partial_{c_3}\\
\mathbf{C}_4 &= c_2\partial_{c_1} - c_1\partial_{c_2} - c_6\partial_{c_5} + c_5\partial_{c_6}\\
\mathbf{C}_5 &= -c_3\partial_{c_1} + c_1\partial_{c_3} + c_6\partial_{c_4} - c_4\partial_{c_6}\\
\mathbf{C}_6 &= c_3\partial_{c_2} - c_2\partial_{c_3} - c_5\partial_{c_4} + c_4\partial_{c_5}\\
\mathbf{C}_7 &= -c_1\partial_{c_1} - c_2\partial_{c_2} - c_3\partial_{c_3}\\
\mathbf{C}_8 &= c_9\partial_{c_9}\\
\mathbf{C}_9 &= (c_7 - c_8)\partial_{c_9}
\end{aligned} \tag{6.25}$$

The corresponding one-parameter groups of (6.25), which express the inner automorphisms of the 9-dimensional equivalence algebra, are listed below:

1. $\widetilde{c}_1 = c_1 + \epsilon_1 c_7, \qquad \widetilde{c}_2 = c_2 + \epsilon_1 c_4, \qquad \widetilde{c}_3 = c_3 - \epsilon_1 c_5$ (6.26a)
2. $\widetilde{c}_1 = c_1 - \epsilon_2 c_4, \qquad \widetilde{c}_2 = c_2 + \epsilon_2 c_7, \qquad \widetilde{c}_3 = c_3 + \epsilon_2 c_6$ (6.26b)
3. $\widetilde{c}_1 = c_1 + \epsilon_3 c_5, \qquad \widetilde{c}_2 = c_2 - \epsilon_3 c_6, \qquad \widetilde{c}_3 = c_3 + \epsilon_3 c_7$ (6.26c)
4. $\widetilde{c}_1 = c_1\cos\epsilon_4 + c_2\sin\epsilon_4, \qquad \widetilde{c}_5 = c_5\cos\epsilon_4 - c_6\sin\epsilon_4,$
   $\widetilde{c}_2 = c_2\cos\epsilon_4 - c_1\sin\epsilon_4, \qquad \widetilde{c}_6 = c_6\cos\epsilon_4 + c_5\sin\epsilon_4$ (6.26d)
5. $\widetilde{c}_1 = c_1\cos\epsilon_5 - c_3\sin\epsilon_5, \qquad \widetilde{c}_4 = c_4\cos\epsilon_5 + c_6\sin\epsilon_5,$
   $\widetilde{c}_3 = c_3\cos\epsilon_5 + c_1\sin\epsilon_5, \qquad \widetilde{c}_6 = c_6\cos\epsilon_5 - c_4\sin\epsilon_5$ (6.26e)
6. $\widetilde{c}_2 = c_2\cos\epsilon_6 + c_3\sin\epsilon_6, \qquad \widetilde{c}_4 = c_4\cos\epsilon_6 - c_5\sin\epsilon_6,$
   $\widetilde{c}_3 = c_3\cos\epsilon_6 - c_2\sin\epsilon_6, \qquad \widetilde{c}_5 = c_5\cos\epsilon_6 + c_4\sin\epsilon_6$ (6.26f)
7. $\widetilde{c}_1 = \epsilon_7 c_1, \qquad \widetilde{c}_2 = \epsilon_7 c_2, \qquad \widetilde{c}_3 = \epsilon_7 c_3$ (6.26g)
8. $\widetilde{c}_9 = \epsilon_8 c_9$ (6.26h)
9. $\widetilde{c}_9 = (c_7 - c_8)\epsilon_9 + c_9$ (6.26i)

where $\epsilon_i$, $i = 1, \dots, 9$ are arbitrary constants with $\epsilon_{7,8} \neq 0$. From (6.26), we easily identify $c_7$ and $c_8$ as invariants of the adjoint group action, while a third one exists as well, namely $c = \left(c_4^2 + c_5^2 + c_6^2\right)^{\frac{1}{2}}$. Note also that when $c_7 = c_8$ the adjoint algebra changes, since $\mathbf{C}_9$ is zero, meaning $\mathbf{V}_9$ cannot be removed.

Now, let us come to the infinite-dimensional algebra $\mathbf{V}_f$. In this case, we cannot follow relation (4.62), but must resort to the more general formula (4.61). First of all, note that in light of (6.16) the commutator of any $\mathbf{V}_g = \nabla g \cdot \partial_{\boldsymbol{A}}$ with any equivalence generator,

$$\begin{aligned}
[\mathbf{V}, \mathbf{V}_g] = [\overline{\mathbf{V}}, \mathbf{V}_g] &= \left(\eta^i \frac{\partial^2 g}{\partial x^i \partial x^j} - \frac{\partial g}{\partial x^i}\frac{\partial \mu^j}{\partial A^i}\right)\frac{\partial}{\partial A^j} \\
&= \left[\frac{\partial}{\partial x^j}\left(\eta^i \frac{\partial g}{\partial x^i}\right) - \left(\frac{\partial \eta^i}{\partial x^j} + \frac{\partial \mu^j}{\partial A^i}\right)\frac{\partial g}{\partial x^i}\right]\frac{\partial}{\partial A^j} \\
&= \left[\frac{\partial}{\partial x^j}\left(\eta^i \frac{\partial g}{\partial x^i}\right) - (2c_7 - c_8)\delta_{ij}\frac{\partial g}{\partial x^i}\right]\frac{\partial}{\partial A^j} = \frac{\partial}{\partial x^j}\left[\eta^i \frac{\partial g}{\partial x^i} + (c_8 - 2c_7)\, g\right]\frac{\partial}{\partial A^j} \\
&= \nabla\left(\boldsymbol{\eta} \cdot \nabla g + (c_8 - 2c_7)\, g\right) \cdot \partial_{\boldsymbol{A}}
\end{aligned} \tag{6.27}$$

gives another vector field of the same form, i.e. in the infinite-dimensional algebra. Following (4.61) we have $\operatorname{Ad}\mathbf{V}_g(\mathbf{V}) = \overline{\mathbf{V}} + \mathbf{V}_f - \nabla\left(\boldsymbol{\eta} \cdot \nabla g + (c_8 - 2c_7)\, g\right) \cdot \partial_{\boldsymbol{A}}$, meaning under the adjoint action of $\mathbf{V}_g$ the generator $\overline{\mathbf{V}} + \mathbf{V}_f$ is mapped to $\overline{\mathbf{V}} + \mathbf{V}_{\widetilde{f}}$, where

$$\widetilde{f} = f - \boldsymbol{\eta} \cdot \nabla g - (c_8 - 2c_7)\, g \tag{6.28}$$

The latter relation is an inner automorphism just like the rest nine this time for the infinite-dimensional equivalence algebra.

Now, using (6.26) and (6.28) we can decompose the equivalence algebra generated by $\mathbf{V}_i$, $i = 1, \ldots, 9$ and $\mathbf{V}_f$ starting with one-dimensional subalgebras.

### 6.5.1 One-dimensional equivalence subalgebras

In this section, we give the optimal system of one-dimensional equivalence subalgebras. In other words, we consider the equivalence generator $\mathbf{V}$ and employing the above transformations, we try to reduce as many parameters $c_i$, $i = 1, \ldots, 9$ and also $f$ as much as possible. In this way we arrive at a generator $\mathbf{Y}$ that cannot be reduced furthermore.

Starting with the part $\overline{\mathbf{V}}$ that belongs to the finite-dimensional equivalence algebra, this procedure breaks down to subcases depending on the elimination of $c_i$, $i = 1, \ldots, 9$. Considering first the invariant $c$ we distinguish two large classes of subalgebras to begin with. Then, using the other two invariants, $c_7$ and $c_8$, we work our way down, where each of the previous classes is separated into two subcases, each of which splits further to two more.

*i*) $c \neq 0$. In this case, at least one of $c_4$, $c_5$ or $c_6$ is nonzero. Without loss of generality let $c_4 \neq 0$ and for simplicity assume $c_4 = 1$. Then using (6.26e) for $\epsilon_5 = \arctan c_6$ we can eliminate $c_6$, and likewise using (6.26f) next for $\epsilon_6 = -\arctan c_5$ we eliminate $c_5$. Moving

on, from (6.26a) and (6.26b) applied successively for $\epsilon_1 = -(c_7c_1 + c_2)/(c_7^2 + 1)$ and $\epsilon_2 = (c_1 - c_7c_2)/(c_7^2 + 1)$ we cancel $c_1$ and $c_2$ too. This case now is separated into two subcases depending on $c_7$, each of which splits into two more depending on $c_7 - c_8$: If $a$) $c_7 \neq 0$, then using (6.26c) for $\epsilon_3 = -c_3/c_7$ we can eliminate $c_3$. Furthermore, if $a1$) $c_7 \neq c_8$ we can also eliminate $c_9$ using (6.26i) for $\epsilon_9 = (c_8 - c_7)/c_9$. Thus the equivalence operator can be expressed as $\mathbf{Y}_1 = \mathbf{V}_4 + k_1\mathbf{V}_7 + k_2\mathbf{V}_8$, $k_2 \neq k_1 \neq 0$. On the other hand, $a2$) $c_7 = c_8$ results in the equivalence operator $\mathbf{Y}_2 = \mathbf{V}_4 + k\,(\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9)$, $k \neq 0$. On the contrary, if $b$) $c_7 = 0$ and $b1$) $c_8 \neq 0$, then $c_9$ can be removed using (6.26i) for $\epsilon_9 = c_8/c_9$ and we arrive at the equivalence operator $\mathbf{Y}_3 = \mathbf{V}_4 + k_1\mathbf{V}_3 + k_2\mathbf{V}_8$, $k_2 \neq 0$. If $b2$) $c_8 = 0$ as well then the equivalence operator is simply $\mathbf{Y}_4 = \mathbf{V}_4 + k\mathbf{V}_3 + \lambda\mathbf{V}_9$.

*ii*) $c = 0$. Now in this case $c_4 = c_5 = c_6 = 0$. Note that if $c_7 = 0$, then at least one of $c_1$, $c_2$ or $c_3$ must be nonzero, for otherwise (6.24) yield $\boldsymbol{A} = \Phi = 0$ for $c_8 \neq 0$, and if $c_8 = 0$ then we have no symmetry. Accordingly we consider the following subcases: If $a$) $c_7 \neq 0$, say $c_7 = 1$, then using (6.26a)-(6.26c) for $\epsilon_i = -c_i$ we can eliminate $c_i$ for $i = 1, 2, 3$. Furthermore, if $a1$) $c_7 \neq c_8$ we can also eliminate $c_9$ using (6.26i) for $\epsilon_9 = (c_8 - c_7)/c_9$ and thus the equivalence operator is $\mathbf{Y}_5 = \mathbf{V}_7 + k\mathbf{V}_8$, $k \neq 1$. On the other hand, for $a2$) $c_7 = c_8$ we arrive at the equivalence operator $\mathbf{Y}_6 = \mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9$. In the opposite case, if $b$) $c_7 = 0$, then as explained above let $c_3 \neq 0$, say $c_3 = 1$. Next, using (6.26e) for $\epsilon_5 = \arctan c_1$ we can eliminate $c_1$, and from (6.26f) for $\epsilon_6 = -\arctan c_2$ we eliminate $c_2$, as well. If $b1$) $c_8 \neq 0$, then same as before $c_9$ can be removed using (6.26i) for $\epsilon_9 = c_8/c_9$. Thus, the equivalence operator is $\mathbf{Y}_7 = \mathbf{V}_3 + k\mathbf{V}_8$, $k \neq 0$. If $b2$) $c_8 = 0$ too, then the equivalence operator is $\mathbf{Y}_8 = \mathbf{V}_3 + \lambda\mathbf{V}_9$.

For each of the above cases, the infinite-dimensional algebra $\mathbf{V}_f$, if included as in $\mathbf{Y} + \mathbf{V}_f$, then it can always be removed using (6.28) and a suitable function $g$ such that $\widetilde{f} = 0$. The latter is nothing more than (5.23), which was used in the previous chapter to eliminate $f$ from (5.20) and arrive at (5.22). Same usage here, this is always possible since (5.23) is a first-order linear partial differential equation with $\boldsymbol{\eta}$ also linear and independent of the dependent variable (see (6.17)). Of course, $\mathbf{V}_f$ themselves, which are the only remaining vector fields outside the 9-dimensional algebra, are not projected to any symmetries. Consequently the above classification of one-dimensional subalgebras under the adjoint group based on the finite-dimensional algebra represented by $\overline{\mathbf{V}}$ is essentially the same for the infinite-dimensional one $\overline{\mathbf{V}} + \mathbf{V}_f$.

In conclusion, the optimal system of one-dimensional equivalence subalgebras comprises of the operators $\mathbf{Y}_i$, $i = 1, \dots, 8$ collected in the next table.

| | 1-dimensional subalgebras | |
|---|---|---|
| 1 | $\mathbf{V}_4 + k_1\mathbf{V}_7 + k_2\mathbf{V}_8,$ | $k_2 \neq k_1 \neq 0$ |
| 2 | $\mathbf{V}_4 + k\,(\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9),$ | $k \neq 0$ |
| 3 | $\mathbf{V}_4 + k_1\mathbf{V}_3 + k_2\mathbf{V}_8,$ | $k_2 \neq 0$ |
| 4 | $\mathbf{V}_4 + k\mathbf{V}_3 + \lambda\mathbf{V}_9$ | |
| 5 | $\mathbf{V}_7 + k\mathbf{V}_8,$ | $k \neq 1$ |
| 6 | $\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9$ | |
| 7 | $\mathbf{V}_3 + k\mathbf{V}_8,$ | $k \neq 0$ |
| 8 | $\mathbf{V}_3 + \lambda\mathbf{V}_9$ | |

Table 6.2: Optimal system of one-dimensional equivalence subalgebras.

### 6.5.2 Two-dimensional equivalence subalgebras

For the construction of the optimal system of two-dimensional equivalence subalgebras, we follow the method outlined in [90], section 14.8 and also in [53]. Basically first one considers that the two-dimensional subalgebras are spanned by one of the operators $\mathbf{Y}$ of the optimal system of one-dimensional subalgebras and the general generator $\mathbf{V}$. Then their Lie bracket must lie again in the subalgebra, meaning for arbitrary constants $\alpha$ and $\beta$ the following relation must hold,

$$[\mathbf{Y}, \mathbf{V}] = \alpha\mathbf{Y} + \beta\mathbf{V} \tag{6.29}$$

Same as before, we also start here with the part $\overline{\mathbf{V}}$ of $\mathbf{V}$ in the finite-dimensional equivalence subalgebra. In this case, using the commutator table 6.1, the above equation becomes an algebraic system in terms of $\alpha$, $\beta$ and the coefficients $c_i$, $i = 1, \ldots, 9$ of $\overline{\mathbf{V}}$, which can be solved accordingly. After each solution is found, we can use again transformations (6.26) to reduce further $\overline{\mathbf{V}}$ without changing $\mathbf{Y}$. Let $\overline{\mathbf{Z}}$ stand for the reduced operator $\overline{\mathbf{V}}$ and $\{\mathbf{Y}, \overline{\mathbf{Z}}\}$ the resulting two-dimensional algebra spanned by $\mathbf{Y}$ and $\overline{\mathbf{Z}}$.

In this process, many cases may be repeatedly appear. Below we give an outline, avoiding tedious details (especially for repeated (sub)cases), which although may not seem so obvious are rather trivial. Following the order of the one-dimensional subalgebras obtained in the previous subsection, we select one by one the operators presented in Table 6.2 and put them to the above test for $\mathbf{Y}$.

*i*) Consider $\mathbf{Y} = \mathbf{V}_4 + k_1\mathbf{V}_7 + k_2\mathbf{V}_8$, $k_2 \neq k_1 \neq 0$. Condition (6.29) yields $c_1 = c_2 = c_5 = c_6 = 0$, and $\alpha = 0$. If $\beta = 0$ then $c_3 = c_9 = 0$ and we recover subcases of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_4$ soon considered. Therefore we take $\beta \neq 0$, which implies $c_7 = c_8 = 0$ and $\beta = -k_1$ in turn. If *a*) $c_9 = 0$, then the second generator is $\overline{\mathbf{Z}} = \mathbf{V}_3$ and therefore we obtain the two-dimensional equivalence subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_1 = \{\mathbf{V}_4 + k_1\mathbf{V}_7 + k_2\mathbf{V}_8, \mathbf{V}_3\}$, $k_2 \neq k_1 \neq 0$. If *b*) $c_9 \neq 0$ meaning $\overline{\mathbf{Z}} = \mathbf{V}_3 + \lambda\mathbf{V}_9$, then $k_1 = 2k_2$ and thus we obtain the two-dimensional equivalence subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_2 = \{\mathbf{V}_4 + k\,(2\mathbf{V}_7 + \mathbf{V}_8)\,, \mathbf{V}_3 + \lambda\mathbf{V}_9\}$, $k \neq 0$.

*ii*) Consider $\mathbf{Y} = \mathbf{V}_4 + k\,(\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9)$, $k \neq 0$. Only slightly different from the previous case, now condition (6.29) results in $c_1 = c_2 = c_5 = c_6 = 0$ and $\alpha = 0$ again. If $\beta = 0$ then $c_3 = \lambda(c_7 - c_8) = 0$, and for $\lambda = 0$ we recover a subcase of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_4$ and for $c_7 = c_8$ case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_6$ given below. For $\beta \neq 0$ then $c_7 = c_8 = c_9 = 0$ and consequently $\beta = -k$. Hence $\overline{\mathbf{Z}} = \mathbf{V}_3$ and we obtain the two-dimensional subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_3 = \{\mathbf{V}_4 + k\,(\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9)\,, \mathbf{V}_3\}$, $k \neq 0$.

*iii*) Consider $\mathbf{Y} = \mathbf{V}_4 + \bar{k}_1\mathbf{V}_3 + \bar{k}_2\mathbf{V}_8$, $\bar{k}_2 \neq 0$. Condition (6.29) gives $c_1 = c_2 = c_5 = c_6 = c_9 = 0$, $\alpha = \beta = 0$ and $\bar{k}_1 c_7 = 0$. Let $\bar{k}_1 = 0$: If *a*) $c_7 \neq 0$, say $c_7 = 1$, then using transformation (6.26c) for $\epsilon_3 = -\,c_3$ we remove $c_3$. Therefore $\overline{\mathbf{Z}} = \mathbf{V}_7 + k_2\mathbf{V}_8$ and the resulting two-dimensional subalgebra is $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_4 = \{\mathbf{V}_4 + k_1\mathbf{V}_8, \mathbf{V}_7 + k_2\mathbf{V}_8\}$, $k_1 \neq 0$. If *b*) $c_7 = 0$, then $\overline{\mathbf{Z}} = \mathbf{V}_3 + k_1\mathbf{V}_8$ and thus the two-dimensional subalgebra obtained is $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_5 = \{\mathbf{V}_4 + k_2\mathbf{V}_8, \mathbf{V}_3 + k_1\mathbf{V}_8\}$, $k_2 \neq 0$. For $\bar{k}_1 \neq 0$ we end up with subcases of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_5$.

*iv*) Consider $\mathbf{Y} = \mathbf{V}_4 + k\mathbf{V}_3 + \lambda\mathbf{V}_9$. Condition (6.29) yields $c_1 = c_2 = c_5 = c_6 = 0$, $\alpha = \beta = 0$, $kc_7 = 0$ and $\lambda(c_7 - c_8) = 0$. For $k = 0$ and $c_7 = c_8$ we have the next two cases: If *a*) $c_7 \neq 0$, say $c_7 = 1$, then using transformation (6.26c) for $\epsilon_3 = -\,c_3$ we eliminate $c_3$. Thus the second generator is $\overline{\mathbf{Z}} = \mathbf{V}_7 + \mathbf{V}_8 + \lambda_2\mathbf{V}_9$ and we obtain the two-dimensional equivalence subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_6 = \{\mathbf{V}_4 + \lambda_1\mathbf{V}_9, \mathbf{V}_7 + \mathbf{V}_8 + \lambda_2\mathbf{V}_9\}$. If *b*) $c_7 = 0$, then $\overline{\mathbf{Z}} = \mathbf{V}_3 + \lambda_1\mathbf{V}_9$ and the two-dimensional equivalence subalgebra obtained is $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_7 = \{\mathbf{V}_4 + \lambda_2\mathbf{V}_9, \mathbf{V}_3 + \lambda_1\mathbf{V}_9\}$. If $k = \lambda = 0$, then for $c_7 \neq 0$ we arrive at a subcase of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_4$ and for $c_7 = 0$ at a subcase of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_7$. For $c_7 = c_8$ we also recover $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_7$, and for $c_7 = \lambda = 0$ a subcase of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_5$.

*v*) Consider $\mathbf{Y} = \mathbf{V}_7 + k\mathbf{V}_8$, $k \neq 1$. Now, condition (6.29) yields at first $\alpha = 0$ and next we separate cases depending on $\beta$. If *a*) $\beta = 0$, then $c_1 = c_2 = c_3 = c_9 = 0$, and therefore one of $c_4$, $c_5$ or $c_6$ must be nonzero. Let $c_4 \neq 0$, say $c_4 = 1$. Then, as previously shown, $c_6$ and $c_5$ can be cancelled using (6.26e) for $\epsilon_5 = \arctan c_6$ and (6.26f) for $\epsilon_6 = -\arctan c_5$, respectively. Thus $\overline{\mathbf{Z}} = \mathbf{V}_4 + k_1\mathbf{V}_8$ and the two-dimensional subalgebra obtained is $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_4$ again but now for $k_2 \neq 1$. If *b*) $\beta \neq 0$, then $c_4 = c_5 = c_6 = c_8 = 0$, which in turn also means $\beta = -1$.

Consequently one of $c_1$, $c_2$ or $c_3$ must be nonzero. Let $c_3 \neq 0$, say $c_3 = 1$. Thus $c_1$ and $c_2$ can be cancelled using (6.26e) for $\epsilon_5 = \arctan c_1$ and (6.26f) for $\epsilon_6 = -\arctan c_2$, respectively. Then if $b1$) $c_9 = 0$ we have $\overline{\mathbf{Z}} = \mathbf{V}_3$ and we obtain the two-dimensional equivalence subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_8 = \{\mathbf{V}_7 + k\mathbf{V}_8, \mathbf{V}_3\}$, $k \neq 1$. If $b2$) $c_9 \neq 0$ then $\overline{\mathbf{Z}} = \mathbf{V}_3 + \lambda\mathbf{V}_9$ and the resulting two-dimensional equivalence subalgebra is $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_9 = \{2\mathbf{V}_7 + \mathbf{V}_8, \mathbf{V}_3 + \lambda\mathbf{V}_9\}$.

*vi*) Consider $\mathbf{Y} = \mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9$. Condition (6.29) gives $\alpha = 0$ again. If $\beta = 0$, then $c_1 = c_2 = c_3 = \lambda c_8 = 0$ and for $\lambda = 0$ we recover a subcase of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_4$ and for $c_8 = 0$ of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_6$. Therefore we consider $\beta \neq 0$, which implies $c_4 = c_5 = c_6 = c_8 = 0$, and $c_9 = 0$, $\beta = -1$ accordingly. Once again one of $c_1$, $c_2$ or $c_3$ must be nonzero. Let $c_3 \neq 0$, say $c_3 = 1$. Thus $c_1$ and $c_2$ can be cancelled using (6.26e) for $\epsilon_5 = \arctan c_1$ and (6.26f) for $\epsilon_6 = -\arctan c_2$, respectively. Then we obtain the equivalence generator $\overline{\mathbf{Z}} = \mathbf{V}_3$ and accordingly the two-dimensional subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_{10} = \{\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9, \mathbf{V}_3\}$.

*vii*) Consider $\mathbf{Y} = \mathbf{V}_3 + k\mathbf{V}_8$, $k \neq 0$. Condition (6.29) yields $\alpha = \beta = 0$ and consequently $c_5 = c_6 = c_7 = c_9 = 0$. If *a*) $c_4 \neq 0$, say $c_4 = 1$, then using transformation (6.26a) for $\epsilon_1 = -c_2$ we eliminate $c_2$, and likewise from (6.26b) for $\epsilon_2 = c_1$ we eliminate $c_1$ too. Thus $\overline{\mathbf{Z}} = \mathbf{V}_4 + k_2\mathbf{V}_8$ and we obtain again the two-dimensional subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_5$ but now for $k_1 \neq 0$. If *b*) $c_4 = 0$, then one of $c_1$ or $c_2$ must be nonzero. Let $c_2 \neq 0$, say $c_2 = 1$. Thus we can remove $c_1$ using (6.26d) for $\epsilon_4 = -\arctan c_1$. Hence $\overline{\mathbf{Z}} = \mathbf{V}_2 + k_1\mathbf{V}_8$ and we arrive at the two-dimensional subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_{11} = \{\mathbf{V}_2 + k_1\mathbf{V}_8, \mathbf{V}_3 + k_2\mathbf{V}_8\}$, $k_2 \neq 0$.

*viii*) Finally consider $\mathbf{Y} = \mathbf{V}_3 + \lambda\mathbf{V}_9$. Condition (6.29) gives $\beta = 0$, and $\alpha = c_7$, $\lambda(c_7 - 2c_8) = 0$ in turn. Here many previous (sub)cases are repeated: If $\lambda = 0$ and $c_4 \neq 0$, then if $c_8 \neq c_7 \neq 0$ case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_1$ reappears or if $c_8 \neq c_7 = 0$ a subcase of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_5$, while if $c_7 = c_8 \neq 0$ we find again case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_3$. If $\lambda = c_4 = 0$, then if $c_8 \neq c_7 \neq 0$ we recover case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_8$ or if $c_8 \neq c_7 = 0$ a subcase of $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_{11}$, while if $c_7 = c_8 \neq 0$ case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_{10}$ reappears. On the other hand if $c_7 = 2c_8 \neq 0$ then if $c_4 \neq 0$ we have case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_2$ again, while if $c_4 = 0$ case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_9$. Finally if $c_7 = c_8 = 0$ then if $c_4 \neq 0$ we end up with case $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_7$ again, while for $c_4 = c_7 = c_8 = 0$ we obtain the generator $\overline{\mathbf{Z}} = \mathbf{V}_2 + \lambda_1\mathbf{V}_9$ and therefore the two-dimensional equivalence subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}\}_{12} = \{\mathbf{V}_2 + \lambda_1\mathbf{V}_9, \mathbf{V}_3 + \lambda_2\mathbf{V}_9\}$.

Now, let us also take into account the infinite-dimensional algebra $\mathbf{V}_f$. Just like before, any two-dimensional equivalence subalgebras lying completely outside the 9-dimensional algebra will not project to two-dimensional symmetry algebras. Therefore inclusion of $\mathbf{V}_f$ is only made in the already existing two-dimensional subalgebras found above. In other words, we consider

subalgebras spanned by $\mathbf{Y}$ again and now the generator $\overline{\mathbf{Z}}+\mathbf{V}_f$ instead of $\overline{\mathbf{Z}}$. Then, in addition to the previous algebraic relations, condition (6.29) yields $[\mathbf{Y},\mathbf{V}_f]=\beta\mathbf{V}_f$. Using (6.27), the latter gives a differential equation for $f$, i.e.

$$\boldsymbol{\eta}\cdot\nabla f+(c_8-2c_7)\,f=\beta f+\bar{\lambda} \tag{6.30}$$

where the constants $c_i$, $i=1,\dots,9$ correspond to the operator $\mathbf{Y}$, while $\bar{\lambda}$ is an arbitrary constant. Note that $\beta$ is already determined for each case.

The question rises whether for every function $f$ satisfying (6.30) there exists a function $g$ which under the inner automorphism (6.28) leaves $\mathbf{Y}$ unaltered, that is

$$\boldsymbol{\eta}\cdot\nabla g+(c_8-2c_7)\,g=d_1 \tag{6.31}$$

where $d_1$ is a constant, and at the same time eliminates $\mathbf{V}_f$ from $\overline{\mathbf{Z}}+\mathbf{V}_f$, meaning

$$\bar{\boldsymbol{\eta}}\cdot\nabla g+(\bar{c}_8-2\bar{c}_7)\,g=d_2+f \tag{6.32}$$

where $d_2$ is a constant, while the constants $\bar{c}_i$, $i=1,\dots,9$ now correspond to the operator $\overline{\mathbf{Z}}$ and in an obvious way the entries $\bar{\eta}^i$ of the vector $\bar{\boldsymbol{\eta}}$ are $\eta^i$ (6.17) for $\bar{c}_j$. Rephrasing the question, is system (6.31)-(6.32) compatible for any solution $f$ of (6.30)? The answer is in the affirmative for the subalgebras $\{\mathbf{Y},\overline{\mathbf{Z}}\}_i$, $i=1,2,4$ for $(k_1,k_2)\neq(0,2),5,6,8,9,11$. Thus, these algebras are representatives in the classification under the adjoint group even if the infinite-dimensional equivalence algebra is used.

For the rest two-dimensional subalgebras $\{\mathbf{Y},\overline{\mathbf{Z}}\}_i$, $i=3,7,10,12,13$, where $\{\mathbf{Y},\overline{\mathbf{Z}}\}_{13}$ stands from now on for $\{\mathbf{Y},\overline{\mathbf{Z}}\}_4$ for $k_1=0,k_2=2$, the answer is negative and therefore these algebras need to be replaced by $\{\mathbf{Y},\overline{\mathbf{Z}}+\mathbf{V}_f\}_i$ in the classification scheme. Still, we find that in all of these cases there always exists a function $g$ satisfying (6.31) and (6.32) for the general solution $f$ of the homogeneous counterpart of equation (6.30), i.e. for $\bar{\lambda}=0$. Hence, owing to the linearity of $\mathbf{V}_f$ with respect to $f$, the generator $\overline{\mathbf{Z}}+\mathbf{V}_f$ can always be reduced under (6.28) to $\mathbf{Z}=\overline{\mathbf{Z}}+\mathbf{V}_{\widetilde{f}}$ without affecting $\mathbf{Y}$, where $\widetilde{f}$ is any particular solution of the inhomogeneous equation (6.30). Thus, we have $\widetilde{f}=\bar{\lambda}\phi$ for $i=3,7,13$, $\widetilde{f}=\bar{\lambda}\ln z$ for $i=10$ and $\widetilde{f}=\bar{\lambda}z$ for $i=12$.

In conclusion, the optimal system of two-dimensional equivalence subalgebras comprises of the algebras $\{\mathbf{Y},\overline{\mathbf{Z}}\}_i$, $i=1,2,4,5,6,8,9,11$, $\{\mathbf{Y},\overline{\mathbf{Z}}+\bar{\lambda}\mathbf{V}_\phi\}_i$, $i=3,7,13$, $\{\mathbf{Y},\overline{\mathbf{Z}}+\bar{\lambda}\mathbf{V}_{\ln z}\}_{10}$ and $\{\mathbf{Y},\overline{\mathbf{Z}}+\bar{\lambda}\mathbf{V}_z\}_{12}$, where $\mathbf{V}_\phi=\rho^{-2}(x\partial_{A^2}-y\partial_{A^1})$, $\mathbf{V}_{\ln z}=z^{-1}\partial_{A^3}$ and $\mathbf{V}_z=\partial_{A^3}$, while $\rho=\sqrt{x^2+y^2}$ and $\phi=\arctan(y/x)$ are polar coordinates in the $xy$-plane. Not preferably in the same order these are collected in the next table.

| | 2-dimensional subalgebras | | |
|---|---|---|---|
| 1 | $\mathbf{V}_3$, | $\mathbf{V}_4 + k_1\mathbf{V}_7 + k_2\mathbf{V}_8$, | $2k_2, k_2 \neq k_1 \neq 0$ |
| 2 | $\mathbf{V}_3 + \lambda\mathbf{V}_9$, | $\mathbf{V}_4 + k(2\mathbf{V}_7 + \mathbf{V}_8)$, | $k \neq 0$ |
| 3 | $\mathbf{V}_3 + \lambda_1\mathbf{V}_\phi$, | $\mathbf{V}_4 + k\,(\mathbf{V}_7 + \mathbf{V}_8 + \lambda_2\mathbf{V}_9)$, | $k \neq 0$ |
| 4 | $\mathbf{V}_3 + k_1\mathbf{V}_8$, | $\mathbf{V}_4 + k_2\mathbf{V}_8$, | $k_1 \neq 0$ or $k_2 \neq 0$ |
| 5 | $\mathbf{V}_3 + \lambda_1\mathbf{V}_9 + \lambda_3\mathbf{V}_\phi$, | $\mathbf{V}_4 + \lambda_2\mathbf{V}_9$ | |
| 6 | $\mathbf{V}_4 + k_1\mathbf{V}_8$, | $\mathbf{V}_7 + k_2\mathbf{V}_8$, | $k_1 \neq 0$ or $k_2 \neq 1, 2$ |
| 7 | $\mathbf{V}_4 + \lambda_1\mathbf{V}_9$, | $\mathbf{V}_7 + \mathbf{V}_8 + \lambda_2\mathbf{V}_9$ | |
| 8 | $\mathbf{V}_4$, | $\mathbf{V}_7 + 2\mathbf{V}_8 + \lambda\mathbf{V}_\phi$ | |
| 9 | $\mathbf{V}_3$, | $\mathbf{V}_7 + k\mathbf{V}_8$, | $k \neq 1/2, 1$ |
| 10 | $\mathbf{V}_3 - \lambda_1\mathbf{V}_{\ln z}$, | $\mathbf{V}_7 + \mathbf{V}_8 + \lambda_2\mathbf{V}_9$ | |
| 11 | $\mathbf{V}_3 + \lambda\mathbf{V}_9$, | $2\mathbf{V}_7 + \mathbf{V}_8$ | |
| 12 | $\mathbf{V}_2 + k_1\mathbf{V}_8$, | $\mathbf{V}_3 + k_2\mathbf{V}_8$, | $k_2 \neq 0$ |
| 13 | $\mathbf{V}_2 + \lambda_1\mathbf{V}_9 + \lambda_3\mathbf{V}_z$, | $\mathbf{V}_3 + \lambda_2\mathbf{V}_9$ | |

Table 6.3: Optimal system of two-dimensional equivalence subalgebras.

### 6.5.3 Three-dimensional equivalence subalgebras

Three-dimensional equivalence subalgebras could follow accordingly from two-dimensional ones. In other words, we consider three-dimensional algebras spanned by $\mathbf{Y}$, $\mathbf{Z}$ and the general generator $\mathbf{V}$, containing a two-dimensional subalgebra $\{\mathbf{Y}, \mathbf{Z}\}$ of the optimal system in Table 6.3. Necessary and sufficient conditions (see [90], p. 189) for these cases are

$$\begin{aligned} [\mathbf{Y}, \mathbf{V}] &= \alpha_1\mathbf{Y} + \beta_1\mathbf{Z} + \gamma_1\mathbf{V} \\ [\mathbf{Z}, \mathbf{V}] &= \alpha_2\mathbf{Y} + \beta_2\mathbf{Z} + \gamma_2\mathbf{V} \end{aligned} \tag{6.33}$$

As usual, we deal first with generators from the finite-dimensional algebra, meaning we take $\overline{\mathbf{Z}}$ and $\overline{\mathbf{V}}$ instead of $\mathbf{Z}$ and $\mathbf{V}$, respectively. Now, equations (6.33) result in an algebraic system for the unknowns $\alpha_1$, $\alpha_2$, $\beta_1$, $\beta_2$, $\gamma_1$, $\gamma_2$ and $c_i$, $i = 1, \ldots, 9$. Solving this system in each case, we can then use the inner automorphisms (6.26) to reduce $\overline{\mathbf{V}}$, without changing either $\mathbf{Y}$ or $\overline{\mathbf{Z}}$. Let $\overline{\mathbf{W}}$ denote the reduced $\overline{\mathbf{V}}$, and $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}$ the algebra obtained spanned by $\mathbf{Y}$, $\overline{\mathbf{Z}}$ and $\overline{\mathbf{W}}$.

Following the order of Table 6.3, below we give a synopsis of the procedure described above case by case, avoiding too many details.

*i* − *v*) From the first five cases of Table 6.3, we recover one of the next three-dimensional subalgebras.

*vi*) Consider $\{\mathbf{Y},\overline{\mathbf{Z}}\} = \{\mathbf{V}_4 + k_1\mathbf{V}_8, \mathbf{V}_7 + k_2\mathbf{V}_8\}$, $k_1 \neq 0$ or $k_2 \neq 1, 2$. Conditions (6.33) yield first $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$ and $c_1 = c_2 = c_5 = c_6 = 0$. $\gamma_1 = 0$ must also hold which implies $k_1 c_9 = 0$, as well as $\gamma_2 \neq 0$ which yields in turn $\gamma_2 = -1$, $c_8 = 0$ and $(2k_2 - 1)c_9 = 0$. For *a*) $c_9 = 0$ we find $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_1 = \{\mathbf{V}_4 + k_1\mathbf{V}_8, \mathbf{V}_7 + k_2\mathbf{V}_8, \mathbf{V}_3\}$, $k_1 \neq 0$ or $k_2 \neq 1, 2$. For *b*) $k_1 = 0$, $k_2 = 1/2$ we obtain $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_2 = \{\mathbf{V}_4, 2\mathbf{V}_7 + \mathbf{V}_8, \mathbf{V}_3 + \lambda\mathbf{V}_9\}$.

*vii*) Consider $\{\mathbf{Y},\overline{\mathbf{Z}}\} = \{\mathbf{V}_4 + \lambda_1\mathbf{V}_9, \mathbf{V}_7 + \mathbf{V}_8 + \lambda_2\mathbf{V}_9\}$. Conditions (6.33) yield first $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$ and $c_1 = c_2 = c_5 = c_6 = 0$. $\gamma_1 = 0$ again must also hold which now implies $\lambda_1 c_8 = 0$, as well as $\gamma_2 \neq 0$ which yields in turn $\gamma_2 = -1$, $c_8 = 0$ and $c_9 = 0$ accordingly. Thus we obtain $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_3 = \{\mathbf{V}_4 + \lambda_1\mathbf{V}_9, \mathbf{V}_7 + \mathbf{V}_8 + \lambda_2\mathbf{V}_9, \mathbf{V}_3\}$.

*viii*) Consider $\{\mathbf{Y},\overline{\mathbf{Z}}\} = \{\mathbf{V}_4, \mathbf{V}_7 + 2\mathbf{V}_8\}$. Similar to *vib*) we obtain the three-dimensional subalgebra $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_4 = \{\mathbf{V}_4, \mathbf{V}_7 + 2\mathbf{V}_8, \mathbf{V}_3\}$.

*ix*) Consider $\{\mathbf{Y},\overline{\mathbf{Z}}\} = \{\mathbf{V}_3, \mathbf{V}_7 + k\mathbf{V}_8\}$, $k \neq 1, 1/2$. Conditions (6.33) give $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$, and again $\gamma_1 = 0$, which results in $c_5 = c_6 = 0$. For $\gamma_2 = 0$ we recover $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_1$. On the other hand for $\gamma_2 \neq 0$ we have $c_4 = c_8 = 0$, $\gamma_2 = -1$ and consequently $c_9 = 0$. Then one of $c_1$ or $c_2$ must be nonzero. Let $c_2 \neq 0$, say $c_2 = 1$. Using (6.26d) for $\epsilon_4 = -\arctan c_1$ we can cancel $c_1$. Thus we arrive at the three-dimensional subalgebra $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_5 = \{\mathbf{V}_3, \mathbf{V}_7 + k\mathbf{V}_8, \mathbf{V}_2\}$, $k \neq 1, 1/2$.

*x*) Consider $\{\mathbf{Y},\overline{\mathbf{Z}}\} = \{\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9, \mathbf{V}_3\}$. Conditions (6.33) give $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$ and also $\gamma_2 = 0$, which implies $c_5 = c_6 = 0$. For $\gamma_1 = 0$ we recover again $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_3$ and a subcase of $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_1$. For $\gamma_1 \neq 0$ we have $c_4 = c_8 = 0$, $\gamma_1 = -1$ and consequently $c_9 = 0$. Same as before one of $c_1$ or $c_2$ must then be nonzero; we assume $c_2 \neq 0$, say $c_2 = 1$. Likewise, using (6.26d) for $\epsilon_4 = -\arctan c_1$ we remove $c_1$ and arrive at the three-dimensional subalgebra $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_6 = \{\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9, \mathbf{V}_3, \mathbf{V}_2\}$.

*xi*) Consider $\{\mathbf{Y},\overline{\mathbf{Z}}\} = \{\mathbf{V}_3 + \lambda\mathbf{V}_9, 2\mathbf{V}_7 + \mathbf{V}_8\}$. Conditions (6.33) give $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$ and also $\gamma_1 = 0$, which implies $c_5 = c_6 = 0$ and $\lambda c_8 = 0$. Same as before, for $\gamma_2 = 0$ we recover again $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_2$ and a subcase of $\{\mathbf{Y},\overline{\mathbf{Z}},\overline{\mathbf{W}}\}_1$. For $\gamma_2 \neq 0$ we have $c_4 = c_8 = 0$ and

$\gamma_2 = -2$. Once again one of $c_1$ or $c_2$ must then be nonzero. Assuming $c_2 \neq 0$, say $c_2 = 1$, then using (6.26d) for $\epsilon_4 = -\arctan c_1$ we can remove $c_1$. Hence we obtain the three-dimensional subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_7 = \{\mathbf{V}_3 + \lambda_1\mathbf{V}_9, 2\mathbf{V}_7 + \mathbf{V}_8, \mathbf{V}_2 + \lambda_2\mathbf{V}_9\}$.

*xii*) Consider $\{\mathbf{Y}, \overline{\mathbf{Z}}\} = \{\mathbf{V}_2 + k_2\mathbf{V}_8, \mathbf{V}_3 + k_3\mathbf{V}_8\}$, $k_3 \neq 0$. Conditions (6.33) give $\gamma_1 = \gamma_2 = 0$, which yield $c_4 = c_5 = c_9 = 0$. As a result, we have $c_6 = c_7 = 0$ and $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$. Therefore we obtain the subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_8 = \{\mathbf{V}_2 + k_2\mathbf{V}_8, \mathbf{V}_3 + k_3\mathbf{V}_8, \mathbf{V}_1 + k_1\mathbf{V}_8\}$.

*xiii*) Consider $\{\mathbf{Y}, \overline{\mathbf{Z}}\} = \{\mathbf{V}_3 + \lambda_3\mathbf{V}_9, \mathbf{V}_2 + \lambda_2\mathbf{V}_9\}$. Conditions (6.33) give $\gamma_1 = \gamma_2 = 0$, which now yield $c_4 = c_5 = 0$. Then we have $c_6 = -\alpha_1 = \beta_2$, $c_7 = \alpha_2 = \beta_1$, $c_6(\lambda_2^2 + \lambda_3^2) = 0$ and $(c_7 - 2c_8)(\lambda_2^2 + \lambda_3^2) = 0$. If *a*) $c_6 = c_7 - 2c_8 = 0$ and *a*1) $c_7 \neq 0$, then we recover the subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_7$, while if *a*2) $c_7 = 0$, then we obtain the three-dimensional subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_9 = \{\mathbf{V}_3 + \lambda_3\mathbf{V}_9, \mathbf{V}_2 + \lambda_2\mathbf{V}_9, \mathbf{V}_1 + \lambda_1\mathbf{V}_9\}$. On the other hand, if *b*) $\lambda_2 = \lambda_3 = 0$, we can first of all, for convenience, use transformation (6.26e) for $\epsilon_5 = \pi/2$ and replace $c_6$ with $c_4$ and $c_1$ with $c_3$. The latter also means that instead of the original subalgebra $\{\mathbf{V}_3, \mathbf{V}_2\}$ we begun with, we now have $\{\mathbf{V}_1, \mathbf{V}_2\}$. We consider $c_4 \neq 0$, say $c_4 = 1$, for otherwise we recover previous subcases. If *b*1) $c_7 \neq 0$, we can eliminate $c_3$ using (6.26c) for $\epsilon_3 = -c_3/c_7$. Furthermore if *b*1*i*) $c_8 \neq c_7$ then we can also eliminate $c_9$ using (6.26i) for $\epsilon_9 = (c_8 - c_7)/c_9$. Thus we arrive at the equivalence subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_{10} = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_4 + k_1\mathbf{V}_7 + k_2\mathbf{V}_8\}$, $k_2 \neq k_1 \neq 0$. If *b*1*ii*) $c_7 = c_8$ then we obtain $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_{11} = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_4 + k\,(\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9)\}$, $k \neq 0$. Accordingly if *b*2) $c_7 = 0$ then if *b*2*i*) $c_8 \neq 0$ then we can eliminate $c_9$ using (6.26i) for $\epsilon_9 = c_8/c_9$ and hence obtain the equivalence subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_{12} = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_4 + k_1\mathbf{V}_3 + k_2\mathbf{V}_8\}$, $k_2 \neq 0$. While if *b*2*ii*) $c_8 = 0$ then we arrive at $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_{13} = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_4 + k\mathbf{V}_3 + \lambda\mathbf{V}_9\}$.

So far we have considered three-dimensional algebras that originate from two-dimensional ones. The only exception to this rule (see [90], p .190) is $\mathfrak{so}(3)$ generated by the rotations $\mathbf{V}_4$, $\mathbf{V}_5$ and $\mathbf{V}_6$, which does not contain any two-dimensional subalgebras. Therefore to the above three-dimensional equivalence subalgebras we must also add $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_{14} = \{\mathbf{V}_4, \mathbf{V}_5, \mathbf{V}_6\}$.

Same as before, we finally include the infinite-dimensional one $\mathbf{V}_f$. As already explained the latter itself does not correspond to any symmetries. Therefore instead of the previous subalgebras $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}$ found, we only need to consider three-dimensional equivalence subalgebras spanned by $\mathbf{Y}$, $\mathbf{Z} = \overline{\mathbf{Z}} + \mathbf{V}_{f_1}$ and $\overline{\mathbf{W}} + \mathbf{V}_{f_2}$. Apart from $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_{14}$, the functions $f_1$ for each case have already been determined in the previous subsection and presented in Table 6.3. Now, conditions (6.33) yield additionally $[\mathbf{Y}, \mathbf{V}_{f_2}] = \beta_1\mathbf{V}_{f_1} + \gamma_1\mathbf{V}_{f_2}$ and $[\mathbf{Z}, \mathbf{V}_{f_2}] + [\mathbf{V}_{f_1}, \mathbf{W}] = \beta_2\mathbf{V}_{f_1} + \gamma_2\mathbf{V}_{f_2}$.

In light of (6.27), the latter relations are expressed as

$$\begin{aligned} \boldsymbol{\eta}\cdot\nabla f_2+(c_8-2c_7)\,f_2&=\beta_1 f_1+\gamma_1 f_2+\bar\lambda_1\\ \bar{\boldsymbol{\eta}}\cdot\nabla f_2+(\bar c_8-2\bar c_7)\,f_2-\bar{\bar{\boldsymbol{\eta}}}\cdot\nabla f_1-(\bar{\bar c}_8-2\bar{\bar c}_7)\,f_1&=\beta_2 f_1+\gamma_2 f_2+\bar\lambda_2 \end{aligned} \tag{6.34}$$

where $\bar\lambda_1$ and $\bar\lambda_2$ are arbitrary constants, while $\beta_1$, $\beta_2$, $\gamma_1$ and $\gamma_2$ have already been found for each case. Following the notation of the previous subsection, the constants $c_i$, $i=1,\dots,9$ correspond to the operator $\mathbf{Y}$, $\bar c_i$, $i=1,\dots,9$ correspond to the operator $\overline{\mathbf{Z}}$ and $\bar{\bar c}_i$, $i=1,\dots,9$ to $\overline{\overline{\mathbf{W}}}$, and in an obvious way the entries $\bar\eta^i$ of the vector $\bar{\boldsymbol{\eta}}$ are $\eta^i$ (6.17) for $\bar c_j$ and the entries $\bar{\bar\eta}^i$ of $\bar{\bar{\boldsymbol{\eta}}}$ are $\eta^i$ (6.17) for $\bar{\bar c}_j$.

Again we want to know given $f_1$ whether for every function $f_2$ satisfying (6.34) there exists a function $g$, which under the inner automorphism leaves $\mathbf{Y}$ and $\mathbf{Z}$ unchanged, meaning

$$\begin{aligned} \boldsymbol{\eta}\cdot\nabla g+(c_8-2c_7)\,g&=d_1\\ \bar{\boldsymbol{\eta}}\cdot\nabla g+(\bar c_8-2\bar c_7)\,g&=d_2 \end{aligned} \tag{6.35}$$

where $d_1$, $d_2$ are constants, and at the same time eliminates $\mathbf{V}_{f_2}$ from $\overline{\mathbf{W}}+\mathbf{V}_{f_2}$, that is

$$\bar{\bar{\boldsymbol{\eta}}}\cdot\nabla g+(\bar{\bar c}_8-2\bar{\bar c}_7)\,g=d_3+f_2 \tag{6.36}$$

where $d_3$ is also a constant. In other words, is system (6.35)-(6.36) compatible for any solution $f_2$ of (6.34)? We find that the answer is in the affirmative for the subalgebras $\{\mathbf{Y},\mathbf{Z},\overline{\mathbf{W}}\}_i$, $i=1,2,4,5$ for $k\neq 0,7,8,10$ for $k_2\neq 0,11,12$. Thus, these algebras are still representatives in the classification under the adjoint group when the infinite-dimensional equivalence algebra is also included. We also need to stress that the compatability conditions of (6.34) require $f_1=0$ in the last three cases, that is for $i=10$ for $k_2\neq 0,11,12$.

For the three-dimensional subalgebras $\{\mathbf{Y},\mathbf{Z},\overline{\mathbf{W}}\}_i$, $i=3,6,9,13,15,16$, where $\{\mathbf{Y},\mathbf{Z},\overline{\mathbf{W}}\}_{15}$ denotes $\{\mathbf{Y},\mathbf{Z},\overline{\mathbf{W}}\}_5$ for $k=0$ and $\{\mathbf{Y},\mathbf{Z},\overline{\mathbf{W}}\}_{16}$ stands for $\{\mathbf{Y},\mathbf{Z},\overline{\mathbf{W}}\}_{10}$ for $k_2=0$, the answer is negative and therefore these algebras must be substituted by $\{\mathbf{Y},\mathbf{Z},\overline{\mathbf{W}}+\mathbf{V}_{f_2}\}_i$ in the classification scheme. However, we find that in all of these cases there always exists a function $g$ satisfying (6.35) and (6.36) for the general solution $f_2$ of the homogeneous counterpart of equations (6.34), i.e. for $\bar\lambda_1=\bar\lambda_2=f_1=0$. Hence, owing to the linearity of $\mathbf{V}_f$ with respect to $f$, the generator $\overline{\mathbf{W}}+\mathbf{V}_{f_2}$ can always be reduced under (6.28) to $\mathbf{W}=\overline{\mathbf{W}}+\mathbf{V}_{\widetilde{f}_2}$ without altering $\mathbf{Y}$ or $\mathbf{Z}$, where $\widetilde{f}_2$ is any particular solution of the inhomogeneous system (6.34). Thus, we find $\widetilde{f}_2=\bar\lambda_1\phi-\bar\lambda_2\ln z$ for $i=3$, $\widetilde{f}_2=\bar\lambda_1\ln y$ for $i=6$, $\widetilde{f}_2=\bar\lambda_1 z+\bar\lambda_2 y$ for $i=9$, $\widetilde{f}_2=\bar\lambda_1 z$ for $i=15$ and $\widetilde{f}_2=\lambda_1(x^2-y^2)/2$ for $i=13,16$. Note that in the latter cases the constant $\lambda_1$ from the function $f_1$ reappears.

Finally, the treatment of the subalgebra $\{\mathbf{Y}, \overline{\mathbf{Z}}, \overline{\mathbf{W}}\}_{14}$ not only changes but becomes much more difficult. First of all, the commutator relations among the three generators require now, besides (6.34), a third relation, namely $[\mathbf{Y}, \mathbf{V}_{f_1}] = \mathbf{V}_{f_2}$. Moreover we want to reduce $\overline{\mathbf{Z}}+\mathbf{V}_{f_1}$ and $\overline{\mathbf{W}} + \mathbf{V}_{f_2}$ simultaneously to $\mathbf{Z} = \overline{\mathbf{Z}} + \mathbf{V}_{\widetilde{f}_1}$ and $\mathbf{W} = \overline{\mathbf{W}} + \mathbf{V}_{\widetilde{f}_2}$, respectively, without predefining $f_1$. After quite lengthy calculations we find that $f_1$ and $f_2$ can be replaced by $\widetilde{f}_1 = \lambda yr/\rho^2$ and $\widetilde{f}_2 = \lambda xr/\rho^2$.

In conclusion, the optimal system of three-dimensional equivalence subalgebras comprises of the algebras listed in the next table (again not in the same order as they were previously derived).

| | 3-dimensional subalgebras | | | |
|---|---|---|---|---|
| 1 | $\mathbf{V}_3$, | $\mathbf{V}_4 + k_1\mathbf{V}_8$, | $\mathbf{V}_7 + k_2\mathbf{V}_8$, | $k_1 \neq 0$ or $k_2 \neq 1/2, 1, 2$ |
| 2 | $\mathbf{V}_3 + \lambda\mathbf{V}_9$, | $\mathbf{V}_4$, | $2\mathbf{V}_7 + \mathbf{V}_8$ | |
| 3 | $\mathbf{V}_3 + \lambda_1\mathbf{V}_\phi - \lambda_2\mathbf{V}_{\ln z}$, | $\mathbf{V}_4 + \lambda_3\mathbf{V}_9$, | $\mathbf{V}_7 + \mathbf{V}_8 + \lambda_4\mathbf{V}_9$ | |
| 4 | $\mathbf{V}_3$, | $\mathbf{V}_4$, | $\mathbf{V}_7 + 2\mathbf{V}_8 + \lambda\mathbf{V}_\phi$ | |
| 5 | $\mathbf{V}_4$, | $\mathbf{V}_5 + \lambda\mathbf{V}_{yr/\rho^2}$, | $\mathbf{V}_6 + \lambda\mathbf{V}_{xr/\rho^2}$ | |
| 6 | $\mathbf{V}_1$, | $\mathbf{V}_2$, | $\mathbf{V}_4 + k_1\mathbf{V}_7 + k_2\mathbf{V}_8$, | $k_1k_2 \neq 0, k_1 \neq k_2$ |
| 7 | $\mathbf{V}_1$, | $\mathbf{V}_2 + \lambda\mathbf{V}_x$, | $\mathbf{V}_4 + k\mathbf{V}_7 + \lambda\mathbf{V}_{(x^2-y^2)/2}$, | $k \neq 0$ |
| 8 | $\mathbf{V}_1$, | $\mathbf{V}_2$, | $\mathbf{V}_4 + k\,(\mathbf{V}_7 + \mathbf{V}_8 + \lambda\mathbf{V}_9)$, | $k \neq 0$ |
| 9 | $\mathbf{V}_1$, | $\mathbf{V}_2$, | $\mathbf{V}_4 + k_1\mathbf{V}_3 + k_2\mathbf{V}_8$, | $k_1k_2 \neq 0$ |
| 10 | $\mathbf{V}_1$, | $\mathbf{V}_2 + \lambda_1\mathbf{V}_x$, | $\mathbf{V}_4 + k\mathbf{V}_3 + \lambda_2\mathbf{V}_9 + \lambda_1\mathbf{V}_{(x^2-y^2)/2}$, | $k \neq 0$ |
| 11 | $\mathbf{V}_2$, | $\mathbf{V}_3$, | $\mathbf{V}_7 + k\mathbf{V}_8$, | $k \neq 0, 1/2, 1$ |
| 12 | $\mathbf{V}_2 + \lambda\mathbf{V}_z$, | $\mathbf{V}_3$, | $\mathbf{V}_7$ | |
| 13 | $\mathbf{V}_2 + \lambda_1\mathbf{V}_9$, | $\mathbf{V}_3 + \lambda_2\mathbf{V}_9$, | $2\mathbf{V}_7 + \mathbf{V}_8$ | |
| 14 | $\mathbf{V}_2 - \lambda_2\mathbf{V}_{\ln y}$, | $\mathbf{V}_3 - \lambda_3\mathbf{V}_{\ln z}$, | $\mathbf{V}_7 + \mathbf{V}_8 + \lambda_1\mathbf{V}_9$ | |
| 15 | $\mathbf{V}_1 + k_1\mathbf{V}_8$, | $\mathbf{V}_2 + k_2\mathbf{V}_8$, | $\mathbf{V}_3 + k_3\mathbf{V}_8$, | $k_3 \neq 0$ |
| 16 | $\mathbf{V}_1 + \lambda_1\mathbf{V}_9 + \lambda_4\mathbf{V}_y + \lambda_5\mathbf{V}_z$, | $\mathbf{V}_2 + \lambda_2\mathbf{V}_9 + \lambda_6\mathbf{V}_z$, | $\mathbf{V}_3 + \lambda_3\mathbf{V}_9$ | |

Table 6.4: Optimal system of three-dimensional equivalence subalgebras.

## 6.6 Symmetry Classification

Once the classification of the equivalence algebra is made, now we can return to section 6.4 and, in particular, the classifying equations (6.24). As explained there, the projection of the equivalence subalgebras spanned by any $\mathbf{V}$ of the form (6.20) yield symmetry subalgebras spanned by $\mathbf{v}$ (6.21) if and only if the potentials $\boldsymbol{A}$ and $\Phi$ satisfy the corresponding subsystem of (6.24). The symmetry algebras obtained are extensions of the principal Lie algebra consisting only of the generator $\mathbf{v}_0$ of time translations.

### 6.6.1 Systems with two Lie point symmetries

Consider the symmetry generators $\mathbf{X}$ coming from the optimal system of one-dimensional subalgebras of Table 6.2, which yield a second symmetry for system (6.13b) besides time translations. Solutions of equations (6.24) in this case can be recovered from section 5.2 of the previous chapter. The results are shown in the next table, where $F_i = F_i(u_1, u_2)$ and $G = G(u_1, u_2)$ are arbitrary functions, while $\boldsymbol{F} = (F_1, F_2, F_3)$ is an arbitrary vector function of the same form. For brevity, we also use cylindrical coordinates $(\rho, \phi, z)$ in many cases. Finally $k$'s and $\lambda$'s are constants, the difference being that the former also define the symmetry.

| | Symmetry generator | Electromagnetic potential | |
|---|---|---|---|
| 1 | $\mathbf{X} = \mathbf{v}_4 + k_1\mathbf{v}_7 + k_2\mathbf{v}_8$<br>$k_2 \neq k_1 \neq 0$ | $A_1 = z^{-\frac{k_2}{k_1}}(xF_1 - yF_2)$<br>$A_2 = z^{-\frac{k_2}{k_1}}(yF_1 + xF_2)$<br>$A_3 = z^{1-\frac{k_2}{k_1}}F_3$<br>$\Phi = z^{2\left(1-\frac{k_2}{k_1}\right)}G$ | $u_1 = \rho/z$<br>$u_2 = \ln z - k_1\phi$ |
| 2 | $\mathbf{X} = \mathbf{v}_4 + k(\mathbf{v}_7 + \mathbf{v}_8)$<br>$k \neq 0$ | $A_1 = z^{-1}(xF_1 - yF_2)$<br>$A_2 = z^{-1}(yF_1 + xF_2)$<br>$A_3 = F_3$<br>$\Phi = \lambda \ln z + G$ | $u_1 = \rho/z$<br>$u_2 = \ln z - k\phi$ |
| 3 | $\mathbf{X} = \mathbf{v}_4 + k_1\mathbf{v}_3 + k_2\mathbf{v}_8$<br>$k_2 \neq 0$ | $A_1 = e^{-k_2\phi}(xF_1 - yF_2)$<br>$A_2 = e^{-k_2\phi}(yF_1 + xF_2)$<br>$A_3 = e^{-k_2\phi}F_3$<br>$\Phi = e^{-2k_2\phi}G$ | $u_1 = \rho$<br>$u_2 = z - k_1\phi$ |

| | | | |
|---|---|---|---|
| 4 | $\mathbf{X} = \mathbf{v}_4 + k\mathbf{v}_3$ | $A_1 = xF_1 - yF_2$<br>$A_2 = yF_1 + xF_2$<br>$A_3 = F_3$<br>$\Phi = \lambda\phi + G$ | $u_1 = \rho$<br>$u_2 = z - k\phi$ |
| 5 | $\mathbf{X} = \mathbf{v}_7 + k\mathbf{v}_8, \quad k \neq 1$ | $\boldsymbol{A} = x^{1-k}\boldsymbol{F}$<br>$\Phi = x^{2(1-k)}G$ | $u_1 = y/x$<br>$u_2 = z/y$ |
| 6 | $\mathbf{X} = \mathbf{v}_7 + \mathbf{v}_8$ | $\boldsymbol{A} = \boldsymbol{F}$<br>$\Phi = \lambda \ln z + G$ | $u_1 = y/x$<br>$u_2 = z/y$ |
| 7 | $\mathbf{X} = \mathbf{v}_3 + k\mathbf{v}_8, \quad k \neq 0$ | $\boldsymbol{A} = e^{-kz}\boldsymbol{F}$<br>$\Phi = e^{-2kz}G$ | $u_1 = x$<br>$u_2 = y$ |
| 8 | $\mathbf{X} = \mathbf{v}_3$ | $\boldsymbol{A} = \boldsymbol{F}$<br>$\Phi = \lambda z + G$ | $u_1 = x$<br>$u_2 = y$ |

Table 6.5: Vector and scalar potentials of the electromagnetic field for two-parameter symmetry groups generated by $\mathbf{v}_0$ and $\mathbf{X}$.

### 6.6.2 Systems with three Lie point symmetries

Now, we consider the symmetry algebras spanned by $\mathbf{X}_1$ and $\mathbf{X}_2$ coming from the optimal system of two-dimensional subalgebras of Table 6.3, which give two more symmetries for system (6.13b) besides time translations. In this case, taking one of the equivalence generators we arrive at a solution of equations (6.24), i.e. one of the potentials listed in Table 6.5. Then each of these solutions is replaced back to the subsystem of (6.24) corresponding to the second generator. Solutions to the latter equations are rather easily found usually by eliminating one of the variables $u_1$ or $u_2$ of the potentials. The results are shown in the next table, using the same notation with the previous table only now $F_i = F_i(u)$ and $G = G(u)$.

| | Symmetry generators | Electromagnetic potential | |
|---|---|---|---|
| 1 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4 + k_1\mathbf{v}_7 + k_2\mathbf{v}_8$<br>$2k_2, k_2 \neq k_1 \neq 0$ | $A_1 = e^{-k_2\phi}(xF_1 - yF_2)$<br>$A_2 = e^{-k_2\phi}(yF_1 + xF_2)$<br>$A_3 = e^{(k_1-k_2)\phi}F_3$<br>$\Phi = e^{2(k_1-k_2)\phi}G$ | $u = \ln\rho - k_1\phi$ |

| | | | |
|---|---|---|---|
| 2 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4 + k\,(2\mathbf{v}_7 + \mathbf{v}_8)$<br>$k \neq 0$ | $A_1 = e^{-k\phi}\,(xF_1 - yF_2)$<br>$A_2 = e^{-k\phi}\,(yF_1 + xF_2)$<br>$A_3 = e^{k\phi}F_3$<br>$\Phi = \lambda z + e^{2k\phi}G$ | $u = \ln\rho - 2k\phi$ |
| 3 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4 + k\,(\mathbf{v}_7 + \mathbf{v}_8)$<br>$k \neq 0$ | $A_1 = \rho^{-1}\left[xF_1 - y\left(F_2 + \lambda_1 z\rho^{-1}\right)\right]$<br>$A_2 = \rho^{-1}\left[yF_1 + x\left(F_2 + \lambda_1 z\rho^{-1}\right)\right]$<br>$A_3 = F_3$<br>$\Phi = \lambda_2 \ln\rho + G$ | $u = \ln\rho - k\phi$ |
| 4 | $\mathbf{X}_1 = \mathbf{v}_3 + k_1\mathbf{v}_8$<br>$\mathbf{X}_2 = \mathbf{v}_4 + k_2\mathbf{v}_8$<br>$k_1 \neq 0$ or $k_2 \neq 0$ | $A_1 = e^{-(k_1 z + k_2\phi)}\,(xF_1 - yF_2)$<br>$A_2 = e^{-(k_1 z + k_2\phi)}\,(yF_1 + xF_2)$<br>$A_3 = e^{-(k_1 z + k_2\phi)}F_3$<br>$\Phi = e^{-2(k_1 z + k_2\phi)}G$ | $u = \rho$ |
| 5 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4$ | $A_1 = xF_1 - yF_2 - \lambda_3 yz\rho^{-2}$<br>$A_2 = yF_1 + xF_2 + \lambda_3 xz\rho^{-2}$<br>$A_3 = F_3$<br>$\Phi = \lambda_1 z + \lambda_2\phi + G$ | $u = \rho$ |
| 6 | $\mathbf{X}_1 = \mathbf{v}_4 + k_1\mathbf{v}_8$<br>$\mathbf{X}_2 = \mathbf{v}_7 + k_2\mathbf{v}_8$<br>$k_1 \neq 0$ or $k_2 \neq 1, 2$ | $A_1 = z^{-k_2}e^{-k_1\phi}\,(xF_1 - yF_2)$<br>$A_2 = z^{-k_2}e^{-k_1\phi}\,(yF_1 + xF_2)$<br>$A_3 = z^{1-k_2}e^{-k_1\phi}F_3$<br>$\Phi = z^{2(1-k_2)}e^{-2k_1\phi}G$ | $u = \rho/z$ |
| 7 | $\mathbf{X}_1 = \mathbf{v}_4$<br>$\mathbf{X}_2 = \mathbf{v}_7 + \mathbf{v}_8$ | $A_1 = z^{-1}\,(xF_1 - yF_2)$<br>$A_2 = z^{-1}\,(yF_1 + xF_2)$<br>$A_3 = F_3$<br>$\Phi = \lambda_1\phi + \lambda_2 \ln z + G$ | $u = \rho/z$ |
| 8 | $\mathbf{X}_1 = \mathbf{v}_4$<br>$\mathbf{X}_2 = \mathbf{v}_7 + 2\mathbf{v}_8$ | $A_1 = \rho^{-2}\left[xF_1 - y\left(F_2 + \lambda\ln\rho\right)\right]$<br>$A_2 = \rho^{-2}\left[yF_1 + x\left(F_2 + \lambda\ln\rho\right)\right]$<br>$A_3 = z^{-1}F_3$<br>$\Phi = z^{-2}G$ | $u = \rho/z$ |
| 9 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_7 + k\mathbf{v}_8$<br>$k \neq 1, 1/2$ | $\boldsymbol{A} = x^{1-k}\boldsymbol{F}$<br>$\Phi = x^{2(1-k)}G$ | $u = y/x$ |

| | | | |
|---|---|---|---|
| 10 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_7 + \mathbf{v}_8$ | $A_1 = F_1$<br>$A_2 = F_2$<br>$A_3 = \lambda_1 \ln y + F_3$<br>$\Phi = \lambda_2 \ln y + G$ | $u = y/x$ |
| 11 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = 2\mathbf{v}_7 + \mathbf{v}_8$ | $\boldsymbol{A} = \sqrt{x}\boldsymbol{F}$<br>$\Phi = \lambda z + xG$ | $u = y/x$ |
| 12 | $\mathbf{X}_1 = \mathbf{v}_2 + k_1\mathbf{v}_8$<br>$\mathbf{X}_2 = \mathbf{v}_3 + k_2\mathbf{v}_8, \quad k_2 \neq 0$ | $\boldsymbol{A} = e^{-(k_1 y + k_2 z)}\boldsymbol{F}$<br>$\Phi = e^{-2(k_1 y + k_2 z)}G$ | $u = x$ |
| 13 | $\mathbf{X}_1 = \mathbf{v}_2$<br>$\mathbf{X}_2 = \mathbf{v}_3$ | $A_1 = 0$<br>$A_2 = F_2$<br>$A_3 = \lambda_3 y + F_3$<br>$\Phi = \lambda_1 y + \lambda_2 z + G$ | $u = x$ |

Table 6.6: Vector and scalar potentials of the electromagnetic field for three-parameter symmetry groups generated by $\mathbf{v}_0$, $\mathbf{X}_1$ and $\mathbf{X}_2$.

**Remark 6.1.** Note that the sixth case of the above table for $k_1 = 0$, $k_2 = 3$, $F_1 = F_3 = 0$ and $F_2(u) = (u^2 + 1)^{-3/2}$ recovers the vector potential of the magnetic dipole described earlier in Example 4.42. Interestingly enough this means that, besides the additional Noether symmetry generated by $\mathbf{X}_1 = \mathbf{v}_4$ corresponding to the integral $I$ (4.46), the system for the magnetic dipole admits another symmetry generated by $\mathbf{X}_2 = \mathbf{v}_7 + 3\mathbf{v}_8$. Violating the Noether condition (5.55a), the latter is not a Noether symmetry and neither preserves $I$ nor $H$. (Related to this, see [13] for an extended notion of integrability demonstrated too by an example of charged particle motion on the surface of a torus.) Nonetheless $\mathbf{X}_2$ is still admitted by the reduced system using either $\mathbf{X}_1$ (as we can see from (4.48)) or $\mathbf{v}_0$. Worthy of mention then are the commutation relations $[\mathbf{X}_1, \mathbf{X}_2] = [\mathbf{v}_0, \mathbf{X}_1] = 0$ and $[\mathbf{v}_0, \mathbf{X}_2] = 3\mathbf{v}_0$.

### 6.6.3 Systems with four Lie point symmetries

Finally, consider the symmetry algebras spanned by $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$ that originate from the optimal system of three-dimensional subalgebras of Table 6.4, which give three more symmetries for system (6.13b) besides time translations. Now each of the solutions of the previous subsection is replaced back to the subsystem of (6.24) for the third generator. Solutions to the latter equations yield the results shown in the next table, where $a_i$, $i = 1, \ldots, 4$ are arbitrary constants.

| | Symmetry generators | Electromagnetic potential |
|---|---|---|
| 1 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4 + k_1\mathbf{v}_8$<br>$\mathbf{X}_3 = \mathbf{v}_7 + k_2\mathbf{v}_8$<br>$k_1 \neq 0$ or $k_2 \neq 1/2, 1, 2$ | $A_1 = e^{-k_1\phi}\rho^{-k_2}\left(a_1 x - a_2 y\right)$<br>$A_2 = e^{-k_1\phi}\rho^{-k_2}\left(a_1 y + a_2 x\right)$<br>$A_3 = a_3 e^{-k_1\phi}\rho^{1-k_2}$<br>$\Phi = a_4 e^{-2k_1\phi}\rho^{2(1-k_2)}$ |
| 2 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4$<br>$\mathbf{X}_3 = 2\mathbf{v}_7 + \mathbf{v}_8$ | $A_1 = \sqrt{\rho}^{-1}\left(a_1 x - a_2 y\right)$<br>$A_2 = \sqrt{\rho}^{-1}\left(a_1 y + a_2 x\right)$<br>$A_3 = a_3\sqrt{\rho}$<br>$\Phi = \lambda z + a_4\rho$ |
| 3 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4$<br>$\mathbf{X}_3 = \mathbf{v}_7 + \mathbf{v}_8$ | $A_1 = \rho^{-1}\left[a_1 x - y\left(a_2 + \lambda_1 z\rho^{-1}\right)\right]$<br>$A_2 = \rho^{-1}\left[a_1 y + x\left(a_2 + \lambda_1 z\rho^{-1}\right)\right]$<br>$A_3 = \lambda_2 \ln\rho$<br>$\Phi = \lambda_3\phi + \lambda_4 \ln\rho$ |
| 4 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4$<br>$\mathbf{X}_3 = \mathbf{v}_7 + 2\mathbf{v}_8$ | $A_1 = \rho^{-2}\left[a_1 x - y\left(a_2 + \lambda \ln\rho\right)\right]$<br>$A_2 = \rho^{-2}\left[a_1 y + x\left(a_2 + \lambda \ln\rho\right)\right]$<br>$A_3 = a_3\rho^{-1}$<br>$\Phi = a_4\rho^{-2}$ |
| 5 | $\mathbf{X}_1 = \mathbf{v}_4$<br>$\mathbf{X}_2 = \mathbf{v}_5$<br>$\mathbf{X}_3 = \mathbf{v}_6$ | $A_1 = \quad \lambda y z r^{-1}\rho^{-2}$<br>$A_2 = -\lambda x z r^{-1}\rho^{-2}$<br>$A_3 = 0$<br>$\Phi = G(r)$ |
| 6 | $\mathbf{X}_1 = \mathbf{v}_1$<br>$\mathbf{X}_2 = \mathbf{v}_2$<br>$\mathbf{X}_3 = \mathbf{v}_4 + k_1\mathbf{v}_7 + k_2\mathbf{v}_8$<br>$k_1 k_2 \neq 0, k_1 \neq k_2$ | $A_1 = a_1 z^{1-\frac{k_2}{k_1}}\cos\left(\ln(z/k_1) + a_2\right)$<br>$A_2 = a_1 z^{1-\frac{k_2}{k_1}}\sin\left(\ln(z/k_1) + a_2\right)$<br>$A_3 = 0$<br>$\Phi = a_4 z^{2\left(1-\frac{k_2}{k_1}\right)}$ |
| 7 | $\mathbf{X}_1 = \mathbf{v}_1$<br>$\mathbf{X}_2 = \mathbf{v}_2$<br>$\mathbf{X}_3 = \mathbf{v}_4 + k\mathbf{v}_7$<br>$k \neq 0$ | $A_1 = a_1 z\cos\left(\ln(z/k) + a_2\right) + \lambda y$<br>$A_2 = a_1 z\sin\left(\ln(z/k) + a_2\right)$<br>$A_3 = 0$<br>$\Phi = a_4 z^2$ |

| | | |
|---|---|---|
| 8 | $\mathbf{X}_1 = \mathbf{v}_1$<br>$\mathbf{X}_2 = \mathbf{v}_2$<br>$\mathbf{X}_3 = \mathbf{v}_4 + k\,(\mathbf{v}_7 + \mathbf{v}_8)$<br>$k \neq 0$ | $A_1 = a_1 \cos\left(\ln(z/k) + a_2\right)$<br>$A_2 = a_1 \sin\left(\ln(z/k) + a_2\right)$<br>$A_3 = 0$<br>$\Phi = \lambda \ln z$ |
| 9 | $\mathbf{X}_1 = \mathbf{v}_1$<br>$\mathbf{X}_2 = \mathbf{v}_2$<br>$\mathbf{X}_3 = \mathbf{v}_4 + k_1\mathbf{v}_3 + k_2\mathbf{v}_8$<br>$k_1 k_2 \neq 0$ | $A_1 = a_1 e^{-\frac{k_2}{k_1} z} \cos\left(z/k_1 + a_2\right)$<br>$A_2 = a_1 e^{-\frac{k_2}{k_1} z} \sin\left(z/k_1 + a_2\right)$<br>$A_3 = 0$<br>$\Phi = a_4 e^{-2\frac{k_2}{k_1} z}$ |
| 10 | $\mathbf{X}_1 = \mathbf{v}_1$<br>$\mathbf{X}_2 = \mathbf{v}_2$<br>$\mathbf{X}_3 = \mathbf{v}_4 + k\mathbf{v}_3$<br>$k \neq 0$ | $A_1 = a_1 \cos\left(z/k + a_2\right) + \lambda_1 y$<br>$A_2 = a_1 \sin\left(z/k + a_2\right)$<br>$A_3 = 0$<br>$\Phi = \lambda_2 z/k$ |
| 11 | $\mathbf{X}_1 = \mathbf{v}_2$<br>$\mathbf{X}_2 = \mathbf{v}_3$<br>$\mathbf{X}_3 = \mathbf{v}_7 + k\mathbf{v}_8$<br>$k \neq 0, 1/2, 1$ | $A_1 = 0$<br>$A_2 = a_2 x^{1-k}$<br>$A_3 = a_3 x^{1-k}$<br>$\Phi = a_4 x^{2(1-k)}$ |
| 12 | $\mathbf{X}_1 = \mathbf{v}_2$<br>$\mathbf{X}_2 = \mathbf{v}_3$<br>$\mathbf{X}_3 = \mathbf{v}_7$ | $A_1 = 0$<br>$A_2 = a_2 x$<br>$A_3 = a_3 x + \lambda y$<br>$\Phi = a_4 x^2$ |
| 13 | $\mathbf{X}_1 = \mathbf{v}_2$<br>$\mathbf{X}_2 = \mathbf{v}_3$<br>$\mathbf{X}_3 = 2\mathbf{v}_7 + \mathbf{v}_8$ | $A_1 = 0$<br>$A_2 = a_2\sqrt{x}$<br>$A_3 = a_3\sqrt{x}$<br>$\Phi = a_4 x + \lambda_1 y + \lambda_2 z$ |
| 14 | $\mathbf{X}_1 = \mathbf{v}_2$<br>$\mathbf{X}_2 = \mathbf{v}_3$<br>$\mathbf{X}_3 = \mathbf{v}_7 + \mathbf{v}_8$ | $A_1 = 0$<br>$A_2 = \lambda_2 \ln x$<br>$A_3 = \lambda_3 \ln x$<br>$\Phi = \lambda_1 \ln x$ |

| | | |
|---|---|---|
| 15 | $\mathbf{X}_1 = \mathbf{v}_1 + k_1\mathbf{v}_8$<br>$\mathbf{X}_2 = \mathbf{v}_2 + k_2\mathbf{v}_8$<br>$\mathbf{X}_3 = \mathbf{v}_3 + k_3\mathbf{v}_8$<br>$k_3 \neq 0$ | $A_1 = a_1 e^{-(k_1x+k_2y+k_3z)}$<br>$A_2 = a_2 e^{-(k_1x+k_2y+k_3z)}$<br>$A_3 = a_3 e^{-(k_1x+k_2y+k_3z)}$<br>$\Phi = a_4 e^{-2(k_1x+k_2y+k_3z)}$ |
| 16 | $\mathbf{X}_1 = \mathbf{v}_1$<br>$\mathbf{X}_2 = \mathbf{v}_2$<br>$\mathbf{X}_3 = \mathbf{v}_3$ | $A_1 = 0$<br>$A_2 = \lambda_4 x$<br>$A_3 = \lambda_5 x + \lambda_6 y$<br>$\Phi = \lambda_1 x + \lambda_2 y + \lambda_3 z$ |

Table 6.7: Vector and scalar potentials of the electromagnetic field for four-parameter symmetry groups generated by $\mathbf{v}_0$, $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$.

## 6.7 Classification in terms of Noether symmetries

In light of section 5.3 of the previous chapter, a Noether symmetry classification can also be given and rather easily too. Since, as explained there, the Noether symmetry condition required only the extra constraint $c_8 = 2c_7$. From the latter and the Tables 6.5, 6.6 we can thus classify system (6.13b) in terms of Noether symmetries. Following the same notation, the results are collected in the next two tables.

| | Noether symmetry | Electromagnetic potential | |
|---|---|---|---|
| 1 | $\mathbf{X} = \mathbf{v}_4 + k\,(\mathbf{v}_7 + 2\mathbf{v}_8)$<br>$k \neq 0$ | $A_1 = z^{-2}\,(xF_1 - yF_2)$, $A_3 = z^{-1}F_3$<br>$A_2 = z^{-2}\,(yF_1 + xF_2)$, $\Phi = z^{-2}G$ | $u_1 = \rho/z$<br>$u_2 = \ln z - k\phi$ |
| 2 | $\mathbf{X} = \mathbf{v}_4 + k\mathbf{v}_3$ | $A_1 = xF_1 - yF_2$, $A_3 = F_3$<br>$A_2 = yF_1 + xF_2$, $\Phi = \lambda\phi + G$ | $u_1 = \rho$<br>$u_2 = z - k\phi$ |
| 3 | $\mathbf{X} = \mathbf{v}_7 + 2\mathbf{v}_8$ | $\boldsymbol{A} = x^{-1}\boldsymbol{F}$<br>$\Phi = x^{-2}G$ | $u_1 = y/x$<br>$u_2 = z/y$ |
| 4 | $\mathbf{X} = \mathbf{v}_3$ | $\boldsymbol{A} = \boldsymbol{F}$<br>$\Phi = \lambda z + G$ | $u_1 = x$<br>$u_2 = y$ |

Table 6.8: Vector and scalar potentials of the electromagnetic field for two-parameter Noether symmetry groups generated by $\mathbf{v}_0$ and $\mathbf{X}$.

From formula (5.58) we can find the invariants that correspond to the Noether symmetries in each represenative case. Thus, we can draw conclusions regarding the existence of additional constants of motion.

**Corollary 6.2.** For inhomogeneous and curved magnetic fields, the autonomous system (6.13b) of charged particle motion admits a first integral of motion $I$ that corresponds to a Noether point symmetry $\mathbf{X}$, which is functionally independent of the Hamiltonian function $H$ and in involution with it, in two representative cases under the equivalence transformations (6.19) :

1. case 2 of Table 6.8 for $\lambda = 0$, where $I = x\dot{y} - y\dot{x} + k\dot{z} + (x^2 + y^2)F_2 + kF_3$,
2. case 4 of Table 6.8 for $\lambda = 0$, where $I = \dot{z} + F_3$.

| | Noether symmetries | Electromagnetic potential | |
|---|---|---|---|
| 1 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4 + k\,(\mathbf{v}_7 + 2\mathbf{v}_8)$<br>$k \neq 0$ | $A_1 = e^{-2k\phi}\,(xF_1 - yF_2)$<br>$A_2 = e^{-2k\phi}\,(yF_1 + xF_2)$<br>$A_3 = e^{-k\phi}F_3$<br>$\Phi = e^{-2k\phi}G$ | $u = \ln\rho - k\phi$ |
| 2 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_4$ | $A_1 = xF_1 - yF_2 - \lambda_3 yz\rho^{-2}$<br>$A_2 = yF_1 + xF_2 + \lambda_3 xz\rho^{-2}$<br>$A_3 = F_3$<br>$\Phi = \lambda_1 z + \lambda_2\phi + G$ | $u = \rho$ |
| 3 | $\mathbf{X}_1 = \mathbf{v}_4$<br>$\mathbf{X}_2 = \mathbf{v}_7 + 2\mathbf{v}_8$ | $A_1 = \rho^{-2}\,[xF_1 - y\,(F_2 + \lambda\ln\rho)]$<br>$A_2 = \rho^{-2}\,[yF_1 + x\,(F_2 + \lambda\ln\rho)]$<br>$A_3 = z^{-1}F_3$<br>$\Phi = z^{-2}G$ | $u = \rho/z$ |
| 4 | $\mathbf{X}_1 = \mathbf{v}_3$<br>$\mathbf{X}_2 = \mathbf{v}_7 + 2\mathbf{v}_8$ | $\boldsymbol{A} = x^{-1}\boldsymbol{F}$<br>$\Phi = x^{-2}G$ | $u = y/x$ |
| 5 | $\mathbf{X}_1 = \mathbf{v}_2$<br>$\mathbf{X}_2 = \mathbf{v}_3$ | $A_1 = 0$<br>$A_2 = F_2$<br>$A_3 = \lambda_3 y + F_3$<br>$\Phi = \lambda_1 y + \lambda_2 z + G$ | $u = x$ |

Table 6.9: Vector and scalar potentials of the electromagnetic field for three-parameter Noether symmetry groups generated by $\mathbf{v}_0$, $\mathbf{X}_1$ and $\mathbf{X}_2$.

Subsequently, we can start investigating aspects of complete integrability in terms of Noether point symmetries. In other words, we study the construction of two additional first integrals of motion $I_1$ and $I_2$, which are functionally independent of the Hamiltonian $H$ and all three $H$, $I_1$ and $I_2$ are pairwise in involution, based on Noether point symmetries. First of all, the cases where $I_1$ and $I_2$ correspond directly to point symmetries $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively, would lie among the potentials of Table 6.9 for three-dimensional Noether symmetry algebras. Higher-dimensional classification results, as previously pointed out, would retrieve subcases of Table 6.9 except for case 5 of Table 6.7, which does not contain any three-dimensional subalgebra. In the latter case, (where the magnetic field represents the hypothetical magnetic monopole, $\boldsymbol{B} = \lambda \boldsymbol{e}_r/r^2$) all three symmetry generators $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$ are of Noether type. But the corresponding integrals $\bar{I}_1 = x\dot{y} - y\dot{x} - \lambda z/r$, $\bar{I}_2 = z\dot{x} - x\dot{z} - \lambda y/r$ and $\bar{I}_3 = y\dot{z} - z\dot{y} - \lambda x/r$, following the structure of the symmetry algebra, are not in involution, $\{\bar{I}_1, \bar{I}_2\} = -\bar{I}_3$, $\{\bar{I}_2, \bar{I}_3\} = -\bar{I}_1$ and $\{\bar{I}_1, \bar{I}_3\} = \bar{I}_2$. However, similar to the classical central-force problem, taking $I_1 = \bar{I}_1^2 + \bar{I}_2^2 + \bar{I}_3^2$ and any $\bar{I}_i$ as $I_2$, we have $\{I_1, I_2\} = \{\bar{I}_j\bar{I}_j, \bar{I}_i\} = 2\bar{I}_j\{\bar{I}_j, \bar{I}_i\} = 2\epsilon_{ijk}\bar{I}_j\bar{I}_k = 0$. Note that in this case $I_1$ is quadratic in the velocities and corresponds to a contact symmetry generated by $2\bar{I}_i\mathbf{X}_i$. By construction though either $I_1$ as a function of the integrals $\bar{I}_i$ that correspond to $\mathbf{X}_i$, or $2\bar{I}_i\mathbf{X}_i$ as a linear combination of $\mathbf{X}_i$ with coefficients the corresponding integrals is really coming from the rotations generated by $\mathbf{X}_i$, i.e. point symmetries. Under these considerations:

**Corollary 6.3.** For inhomogeneous and curved magnetic fields, the autonomous system (6.13b) of charged particle motion is completely integrable via Noether point symmetries in three representative cases under the equivalence transformations (6.19):

1. case 2 of Table 6.9 for $\lambda_i = 0$, where $I_1 = \dot{z} + F_3$ and $I_2 = x\dot{y} - y\dot{x} + (x^2 + y^2)F_2$,
2. case 5 of Table 6.9 for $\lambda_i = 0$, where $I_1 = \dot{y} + F_2$ and $I_2 = \dot{z} + F_3$,
3. case 5 of Table 6.7, where $I_1 = (x\dot{y} - y\dot{x})^2 + (z\dot{x} - x\dot{z})^2 + (y\dot{z} - z\dot{y})^2$ and $I_2 = \bar{I}_i$ for any $\bar{I}_1 = x\dot{y} - y\dot{x} - \lambda z/r$, $\bar{I}_2 = z\dot{x} - x\dot{z} - \lambda y/r$, $\bar{I}_3 = y\dot{z} - z\dot{y} - \lambda x/r$.

## 6.8 Discussion

We have found and classified one-, two- and three-parameter symmetry group extensions of time translations admitted by the three-dimensional autonomous non-relativistic charged particle motion. The classification was made under the action of the equivalence group, which also proved to preserve the homogeneous Maxwell's equations with no restrictions at all. Therefore each

symmetry class is described in terms of the vector and scalar potentials of the electromagnetic field in a representative form as simple as possible under equivalence transformations. In other words, the members that belong to each case can be found by transforming the typical potentials presented in the previous tables using (6.19). The corresponding Noether symmetry classification has also led to some first conclusions about the existence of first integrals of motion besides the well-known Hamiltonian.

As previously mentioned at the end of section 4.8, this type of classification is considered preliminary in the sense that there could be more symmetries lying outside the equivalence group. However, from the inspection of the determining equations resulting from the symmetry condition in chapter 5, we concluded that, when the system is nonlinear and particularly for inhomogeneous and curved magnetic fields, the only symmetries admitted are the ones that belong to the equivalence group. Therefore, in this case Tables 6.5-6.7 can be considered as a full classification in terms of point symmetries and up to three-parameter symmetry groups. Of course, more symmetries can be expected when the potentials listed in the previous tables result in magnetic fields of constant direction or even linear equations of motion, as for example in the last case of Table 6.7.

# Chapter 7

# Conclusions and future work

In this chapter we take an overview of the outcome of this thesis, triggering potential extensions of this work. More detailed and technical discussions on conclusions of individual results were preferably given at the end of each of the chapters 2, 3, 5 and 6.

In this thesis we have approached the subject of charged particle motion from two different angles. On one side, we investigated the problem from the point of view of applications and in particular plasma fusion research. A current issue of interest especially in toroidal devices is the appearance of magnetic islands and their effects as well as their control. For their investigation, we presented first (chapter 2) an analytical description of an isolated island, coming from a helical magnetic perturbation, which is often used for modelling. Then (chapter 3), we developed a numerical code for large numbers of particles in order to study the wave-particle interaction in electron cyclotron resonance inside and near the center of a magnetic island. In this treatment, relativistic effects, collisions and the magnetic surface label of the previous chapter have been included. From the simulations, we have determined the driven current and absorbed power densities as well as other macroscopic quantities such as velocity distributions, etc.

On the other hand, we have set out a more theoretical study concerning the structure of the equations of motion for charged particles. On this quest, we have employed symmetry group methods, which have been widely used the last few decades. Considering the autonomous system in general, we have found (chapter 5) the general form of electromagnetic fields admitting more symmetries than time translations along with the general form of the symmetries themselves. Focusing on Lie point, i.e. geometrical, symmetries we have looked into Noether point symmetries as well and deduced, in particular, integrals of motion that are in involution with the Hamiltonian. Comparison in terms of symmetries with the system of the magnetic field lines

is also made. Finally (chapter 6), recruiting the infinitesimal methods of equivalence transformations, which lately have spread out, we have made a symmetry classification for the system. From the investigation of the symmetry condition in the previous chapter this partial classification turns out to be complete (in terms of point symmetries) when the system is nonlinear and genuinely considered in three dimensions. Restricting to Noether symmetries, we also give a classification in terms of first integrals, from which integrable cases have been found.

Parts I and II, summarized in the previous two paragraphs, respectively, may sound disconnected at this point. And of course not all questions posed have been answered and not fully either. However, the work presented in this thesis does not stop merely here, but the tools developed may serve as a basis for future plans. Let us see how.

First of all, the symmetry analysis started here from the autonomous system could be continued for the general case of time-dependent electromagnetic fields. Although it seems unlikely to carry out a general investigation of the symmetry condition like in chapter 5, the method of equivalence transformations applied in chapter 6 could pay off. The resulting classification would surface with less difficulties, having ruled out the wide class of the time-independent fields found here. This investigation may yield more interesting cases that are more close to applications. For instance, monochromatic plane or other polarised electromagnetic waves, like the ones used in chapter 3, are likely candidates. In this context it is also worth noting the relation between the charged-particle motion and the Vlasov equation and, in particular, that the first integrals of the former are solutions of the latter.

Another extension would be to consider generalised symmetries, i.e. symmetry transformations more general than the geometrical explored in this work that also include derivatives of the dependent variables. This idea comes naturally given that, in light of (5.58), the first integrals from Noether point symmetries which are functionally independent of the Hamiltonian are necessarily linear in the velocities. Of course for this treatment one must begin with an ansatz for the order of derivatives involved, and in some cases, as in this one, the dependence of the symmetry generators on these derivatives. However, since the system is of second order and linear in the velocities, the most simple choice of first-order derivatives appearing linearly in the generators seems already promising.

The usage of approximate symmetries might also open a new perspective to study the guiding-center equations as a perturbation of the magnetic field lines that feature the particle's motion. In this viewpoint, one could inspect which symmetries of the magnetic field are stable, meaning carried over as approximate symmetries to the guiding-center motion, and what is their relation

with the corresponding symmetries of the original system. The adiabatic invariance of the magnetic moment in the first place, which resulted in a reduction by two, originates from an approximate symmetry with respect to the gyrophase. More symmetries could further reduce the system, which may stand a better chance of finding solutions. This is actually the case for axisymmetric systems, whose orbits could be classified in free and trapped (banana) ones, given at least a qualitative if not full description. Helical magnetic fields exhibiting magnetic islands perhaps could follow.

Last but not least, reduction of order, which is perhaps the ultimate goal of symmetry analysis, and already previously suggested, could also be employed in continuation of either of the above investigations, including the present one, for particular cases of interest. Even if not completely integrable, a reduced system could prove important not only for further analytical investigation, but numerical treatments as well. This is where the two parts of this thesis could meet. For under a symmetry reduction, numerical simulations could be much improved with less computational time and higher accuracy at the same time. In this way, the barrier of the very demanding Lorentz force we encountered when following large numbers of particles could be overcome.

# Appendix A

# Curvilinear Coordinates

In this short description we are primarily concerned about the way vectors and a few vector operations are expressed in curvilinear coordinates. Let $(x, y, z)$ be Cartesian coordinates in the Euclidean space $\mathbb{R}^3$ and $\left(x^1, x^2, x^3\right)$ a system of curvilinear coordinates. The former are related to the latter by a set of relations

$$x = x\left(x^1, x^2, x^3\right), \qquad y = y\left(x^1, x^2, x^3\right), \qquad z = z\left(x^1, x^2, x^3\right), \tag{A.1}$$

which are invertible, given that the determinant of the *Jacobian matrix* for the above transformation,

$$\mathrm{J} = \frac{\partial\left(x, y, z\right)}{\partial\left(x^1, x^2, x^3\right)} = \begin{pmatrix} \dfrac{\partial x}{\partial x^1} & \dfrac{\partial x}{\partial x^2} & \dfrac{\partial x}{\partial x^3} \\ \dfrac{\partial y}{\partial x^1} & \dfrac{\partial y}{\partial x^2} & \dfrac{\partial y}{\partial x^3} \\ \dfrac{\partial z}{\partial x^1} & \dfrac{\partial z}{\partial x^2} & \dfrac{\partial z}{\partial x^3} \end{pmatrix} \tag{A.2}$$

is not zero, and thus leading to

$$x^1 = x^1(x, y, z), \qquad x^2 = x^2(x, y, z), \qquad x^3 = x^3(x, y, z). \tag{A.3}$$

In terms of the Cartesian coordinates, a point in $\mathbb{R}^3$ can be denoted by $(x, y, z)$ or its position vector

$$\boldsymbol{r} = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}, \tag{A.4}$$

where $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ are the unit vectors along the $x$-, $y$- and $z$-axes, respectively. The same point in the curvilinear system is now denoted by $\left(x^1, x^2, x^3\right)$, while if $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ represent a vector basis, then the position vector will be expressed as

$$\boldsymbol{r} = x^1\boldsymbol{e}_1 + x^2\boldsymbol{e}_2 + x^3\boldsymbol{e}_3. \tag{A.5}$$

Similarly to the $x$-, $y$- and $z$-axes, three *coordinate curves* are naturally induced in the curvilinear system, namely

$$\begin{aligned} \boldsymbol{r}_1(x^1) &= \{x^2 = c_2, x^3 = c_3\} \\ \boldsymbol{r}_2(x^2) &= \{x^1 = c_1, x^3 = c_3\} \\ \boldsymbol{r}_3(x^3) &= \{x^1 = c_1, x^2 = c_2\} \end{aligned} \tag{A.6}$$

where $c_i$ are constants. Since these are not straight lines, as opposed to the axes of the Cartesian frame of refence, a vector basis can only be defined in terms of the tangent vectors of these curves. Indeed, using the expression (A.5) for (A.6), we easily see that $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are the tangent vectors to $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$, respectively,

$$\frac{d\boldsymbol{r}_1}{dx^1} = \boldsymbol{e}_1, \qquad \frac{d\boldsymbol{r}_2}{dx^2} = \boldsymbol{e}_2, \qquad \frac{d\boldsymbol{r}_3}{dx^3} = \boldsymbol{e}_3.$$

On the other hand, if we differentiate both (A.4) and (A.5) with respect to $x^i$ and equate the two expressions, we have

$$\boldsymbol{e}_i = \frac{\partial \boldsymbol{r}}{\partial x^i} = \frac{\partial x}{\partial x^i}\,\boldsymbol{i} + \frac{\partial y}{\partial x^i}\,\boldsymbol{j} + \frac{\partial z}{\partial x^i}\,\boldsymbol{k} \tag{A.7}$$

or in matrix form

$$\begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \boldsymbol{e}_3 \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x}{\partial x^1} & \dfrac{\partial y}{\partial x^1} & \dfrac{\partial z}{\partial x^1} \\ \dfrac{\partial x}{\partial x^2} & \dfrac{\partial y}{\partial x^2} & \dfrac{\partial z}{\partial x^2} \\ \dfrac{\partial x}{\partial x^3} & \dfrac{\partial y}{\partial x^3} & \dfrac{\partial z}{\partial x^3} \end{pmatrix} \begin{pmatrix} \boldsymbol{i} \\ \boldsymbol{j} \\ \boldsymbol{k} \end{pmatrix} \tag{A.8}$$

From (A.8), we clearly deduce that J being invertible guarantees that $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are linearly independent.

Now, consider a vector $\boldsymbol{A} = A_x\boldsymbol{i} + A_y\boldsymbol{j} + A_z\boldsymbol{k}$. In the curvilinear coordinates this would be expressed with respect to the tangent basis as

$$\boldsymbol{A} = A^i\boldsymbol{e}_i \tag{A.9}$$

If we substitute (A.7) in the above relation, equate the two expressions and solve in terms of $A^i$, we have

$$\begin{pmatrix} A^1 \\ A^2 \\ A^3 \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x^1}{\partial x} & \dfrac{\partial x^1}{\partial y} & \dfrac{\partial x^1}{\partial z} \\ \dfrac{\partial x^2}{\partial x} & \dfrac{\partial x^2}{\partial y} & \dfrac{\partial x^2}{\partial z} \\ \dfrac{\partial x^3}{\partial x} & \dfrac{\partial x^3}{\partial y} & \dfrac{\partial x^3}{\partial z} \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \tag{A.10}$$

i.e. the components $A^i$ are related to the Cartesian components of $\boldsymbol{A}$ exactly the opposite way that $\boldsymbol{e}_i$ are related to the Cartesian basis vectors. Considering the inverse transformation (A.3), we identify the matrix in (A.10) as $\mathrm{J}^{-1}$.

Since the tangent vectors $\boldsymbol{e}_i$ do not constitute a unit basis, we also consider the normalised vectors

$$\widehat{\boldsymbol{e}}_i = \frac{\boldsymbol{e}_i}{|\boldsymbol{e}_i|}\,, \qquad \text{(no summation)} \tag{A.11}$$

which, unlike $\boldsymbol{e}_i$, they depend on $x^j$, because of the denominators. These unit vectors are often refered to as the natural or physical basis, in terms of which a vector field is expressed as

$$\boldsymbol{A} = \widehat{A}^i\, \widehat{\boldsymbol{e}}_i, \tag{A.12}$$

where from (A.9) and (A.11) we easily find $\widehat{A}^i = A^i\,|\boldsymbol{e}_i|$ (no summation).

Let us comment that behind the identification of a point $\boldsymbol{x}$ to its position vector lies the isomorphism of the tangent space $T_{\boldsymbol{x}}\mathbb{R}^3$ of $\mathbb{R}^3$ at every point $\boldsymbol{x}$ with $\mathbb{R}^3$ itself. There is, however, another vector space that also coincides with $\mathbb{R}^3$, namely the dual space of the tangent space, consisting of all the linear, real maps of $T_{\boldsymbol{x}}\mathbb{R}^3$. Without going in to further details, the dual basis $\boldsymbol{e}^i$ is defined by the relations

$$\boldsymbol{e}^i(\boldsymbol{e}_j) = \boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i_j\,. \tag{A.13}$$

Since for every vector $\boldsymbol{B} = B^i\boldsymbol{e}_i$ we have $\boldsymbol{e}^i \cdot \boldsymbol{B} = \boldsymbol{e}^i \cdot \left(B^j\boldsymbol{e}_j\right) = B^j\boldsymbol{e}^i \cdot \boldsymbol{e}_j = B^j\delta^i_j = B^i$, and $\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{A}\cdot\left(B^i\boldsymbol{e}_i\right) = (\boldsymbol{A}\cdot\boldsymbol{e}_i)\,B^i = (\boldsymbol{A}\cdot\boldsymbol{e}_i)\,\boldsymbol{e}^i\cdot\boldsymbol{B}$ for any two vectors $\boldsymbol{A} = A^i\boldsymbol{e}_i$ and $\boldsymbol{B}$, then

$$\boldsymbol{A} = (\boldsymbol{A}\cdot\boldsymbol{e}_i)\,\boldsymbol{e}^i, \tag{A.14}$$

meaning every vector can be written as a linear combination of $\boldsymbol{e}^i$. Therefore, $\boldsymbol{e}^i$ is a basis, in terms of which the components of a vector $\boldsymbol{A}$ are $A_i = \boldsymbol{A}\cdot\boldsymbol{e}_i$. Now, if we let $\boldsymbol{A} = \nabla x^i = \partial x^i/\partial \boldsymbol{r}$ in (A.14) then, recalling (A.7), we easily arrive at $\nabla x^i \cdot \boldsymbol{e}_j = \delta^i_j$ and thus deduce that $\boldsymbol{e}^i$ are simply the gradients of the functions $x^i(x,y,z)$ in the Cartesian coordinates,

$$\boldsymbol{e}^i = \frac{\partial x^i}{\partial \boldsymbol{r}} = \frac{\partial x^i}{\partial x}\,\boldsymbol{i} + \frac{\partial x^i}{\partial y}\,\boldsymbol{j} + \frac{\partial x^i}{\partial z}\,\boldsymbol{k} \tag{A.15}$$

or in matrix form

$$\begin{pmatrix} \boldsymbol{e}^1 \\ \boldsymbol{e}^2 \\ \boldsymbol{e}^3 \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x^1}{\partial x} & \dfrac{\partial x^1}{\partial y} & \dfrac{\partial x^1}{\partial z} \\ \dfrac{\partial x^2}{\partial x} & \dfrac{\partial x^2}{\partial y} & \dfrac{\partial x^2}{\partial z} \\ \dfrac{\partial x^3}{\partial x} & \dfrac{\partial x^3}{\partial y} & \dfrac{\partial x^3}{\partial z} \end{pmatrix} \begin{pmatrix} \boldsymbol{i} \\ \boldsymbol{j} \\ \boldsymbol{k} \end{pmatrix} \tag{A.16}$$

Complementary to this algebraic viewpoint, a geometrical interpretation is associated with the *coordinate surfaces*, which can be considered as the counterparts of the $yz$-, $xz$- and $xy$-planes in the Cartesian coordinates, and are defined as

$$\begin{aligned} \boldsymbol{r}_{23}(x^2, x^3) &= \{x^1 = c_1\} \\ \boldsymbol{r}_{13}(x^1, x^3) &= \{x^2 = c_2\} \\ \boldsymbol{r}_{12}(x^1, x^2) &= \{x^3 = c_3\} \end{aligned} \tag{A.17}$$

The vectors defined in (A.15) can now be viewed as the normal vectors to these surfaces giving rise to a vector basis, as the tangent vectors to the coordinate curves did.

So, a vector $\boldsymbol{A}$ in $\mathbb{R}^3$ can also be written as

$$\boldsymbol{A} = A_i \boldsymbol{e}^i \tag{A.18}$$

where $A_i = \boldsymbol{A} \cdot \boldsymbol{e}_i$. Using Cartesian components in the last expression, we get

$$\begin{pmatrix} A_1 \\ \\ A_2 \\ \\ A_3 \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x}{\partial x^1} & \dfrac{\partial y}{\partial x^1} & \dfrac{\partial z}{\partial x^1} \\ \dfrac{\partial x}{\partial x^2} & \dfrac{\partial y}{\partial x^2} & \dfrac{\partial z}{\partial x^2} \\ \dfrac{\partial x}{\partial x^3} & \dfrac{\partial y}{\partial x^3} & \dfrac{\partial z}{\partial x^3} \end{pmatrix} \begin{pmatrix} A_x \\ \\ A_y \\ \\ A_z \end{pmatrix} \tag{A.19}$$

We stress here the difference between the components $A^i$ and $A_i$ of a vector $\boldsymbol{A}$ in the curvilinear coordinate system. In light of (A.10) and (A.19) we see that they are related to the Cartesian components of $\boldsymbol{A}$ in exactly the opposite way. This difference is further emphasized by calling the upperscript components $A^i$ *contravariant* and the lowerscript ones $A_i$ *covariant*. In fact, this terminology rule will be generalized, meaning that everything superscripted will be called contravariant and everything subscripted will be called covariant. The justification of this logic stems from the way geometrical objects transform. For example, since $\boldsymbol{e}^i$ share the same transformation matrix with $A^i$ they will be refer to as the contravariant basis, while $\boldsymbol{e}_i$ sharing the same transformation matrix with $A_i$ will be called the covariant basis (see also [28], p. 19).

Now if we want to directly transform from contravariant components $A^i$ to covariant ones $A_i$, we may combine (A.19) and the inversed (A.10) or simply recall $A_i = \boldsymbol{A} \cdot \boldsymbol{e}_i$ to deduce

$$A_i = g_{ij} A^j, \tag{A.20}$$

arriving at the notion of the metric tensor, expressed by the matrix $G = \mathrm{J}^T \mathrm{J}$ with elements

$$g_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j = \frac{\partial x}{\partial x^i}\frac{\partial x}{\partial x^j} + \frac{\partial y}{\partial x^i}\frac{\partial y}{\partial x^j} + \frac{\partial z}{\partial x^i}\frac{\partial z}{\partial x^j}, \tag{A.21}$$

which is clearly symmetric. Since $g = |G| = |\mathrm{J}|^2$ and J is invertible, therefore $g \neq 0$ (and in fact positive), i.e. $G$ is also invertible. Consistent with the above notation, the elements (A.21) of $G$ are the covariants components of the metric, while the elements of $G^{-1}$ are the contravariant given by

$$g^{ij} = \frac{\partial x^i}{\partial x}\frac{\partial x^j}{\partial x} + \frac{\partial x^i}{\partial y}\frac{\partial x^j}{\partial y} + \frac{\partial x^i}{\partial z}\frac{\partial x^j}{\partial z}\,. \tag{A.22}$$

These can be recovered in a similar way as the covariant had by use of (A.10) and the inversed (A.19), yielding the inverse relation

$$A^i = g^{ij} A_j. \tag{A.23}$$

In order to proceed with the basic vector operations, we first note that $\left|\mathrm{J}^T\right| = |\mathrm{J}|$ could actually be written as the mixed product of $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, just as the determinant $\left|\mathrm{J}^{-1}\right| = |\mathrm{J}|^{-1}$ can be written as the mixed product of $\boldsymbol{e}^1$, $\boldsymbol{e}^2$ and $\boldsymbol{e}^3$, given (A.7) and (A.15). In terms of the determinant of the metric tensor previously discussed, we have the following

$$\boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_3) = \sqrt{g} \tag{A.24}$$

$$\boldsymbol{e}^1 \cdot (\boldsymbol{e}^2 \times \boldsymbol{e}^3) = \frac{1}{\sqrt{g}} \tag{A.25}$$

Next we observe that $\boldsymbol{e}_1 \cdot \boldsymbol{e}^2 = \boldsymbol{e}_1 \cdot \boldsymbol{e}^3 = 0$ means that $\boldsymbol{e}_1 = c\,\boldsymbol{e}^2 \times \boldsymbol{e}^3$, where $c$ is a constant. By dot-multiplication of the last equation with the vector $\boldsymbol{e}^1$ we find $c = \sqrt{g}$, given (A.25). That is

$$\boldsymbol{e}_1 = \sqrt{g}\,\boldsymbol{e}^2 \times \boldsymbol{e}^3 \tag{A.26}$$

$$\boldsymbol{e}^1 = \frac{1}{\sqrt{g}}\,\boldsymbol{e}_2 \times \boldsymbol{e}_3 \tag{A.27}$$

where the second equation can be derived in the same way. Of course (A.24)-(A.27) can be completed with similar relations obtained by cyclic permutation.

Now we are ready to derive general formulas for the vector operations of interest, in which, as we will see, contravariant or covariant components are more preferable than physical ones. The *dot or scalar product* of two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is

$$\boldsymbol{A} \cdot \boldsymbol{B} = A^i \boldsymbol{e}_i \cdot B^j \boldsymbol{e}_j = A^i B^j \boldsymbol{e}_i \cdot \boldsymbol{e}_j = A^i B^j g_{ij} = A^i B_i = A_i B^i \tag{A.28}$$

The *cross or vector product* of two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is

$$\boldsymbol{A} \times \boldsymbol{B} = A^j \boldsymbol{e}_j \times B^k \boldsymbol{e}_k = A^j B^k \boldsymbol{e}_j \times \boldsymbol{e}_k = \sqrt{g}\,\epsilon_{ijk} A^j B^k \boldsymbol{e}^i \tag{A.29}$$

$$\text{or } \boldsymbol{A} \times \boldsymbol{B} = A_j \boldsymbol{e}^j \times B_k \boldsymbol{e}^k = A_j B_k \boldsymbol{e}^j \times \boldsymbol{e}^k = \frac{1}{\sqrt{g}}\,\epsilon^{ijk} A_j B_k \boldsymbol{e}_i \tag{A.30}$$

where $\epsilon_{ijk} = \epsilon^{ijk}$ is the Levi-Civita symbol that does not transform like tensors, but follows the rules $\epsilon^{ijk} = g g^{il} g^{jm} g^{kn} \epsilon_{lmn}$ and $\epsilon_{ijk} = g^{-1} g_{il} g_{jm} g_{kn} \epsilon^{lmn}$.

The *gradient of a scalar field* $\Phi\left(x^1, x^2, x^3\right)$ is naturally induced in the contravariant basis

$$\nabla\Phi = \frac{\partial\Phi}{\partial x}\,\boldsymbol{i} + \frac{\partial\Phi}{\partial y}\,\boldsymbol{j} + \frac{\partial\Phi}{\partial z}\,\boldsymbol{k} = \frac{\partial\Phi}{\partial x^i}\frac{\partial x^i}{\partial x}\,\boldsymbol{i} + \frac{\partial\Phi}{\partial x^i}\frac{\partial x^i}{\partial y}\,\boldsymbol{j} + \frac{\partial\Phi}{\partial x^i}\frac{\partial x^i}{\partial z}\,\boldsymbol{k} = \frac{\partial\Phi}{\partial x^i}\,\boldsymbol{e}^i \tag{A.31}$$

Perhaps the easiest derivation way of the divergence formula for a vector field $\boldsymbol{A}\left(x^1, x^2, x^3\right)$ is the following, using basic vector properties. As previously mentioned in the definition of the contravariant basis or easily verified by the above formula, $\boldsymbol{e}^i = \nabla x^i$ hold. So,

$$\begin{aligned}\nabla\cdot\left(\frac{\boldsymbol{e}_i}{\sqrt{g}}\right) &= \nabla\cdot\left(\epsilon_{ijk}\,\boldsymbol{e}^j\times\boldsymbol{e}^k\right) = \epsilon_{ijk}\nabla\cdot\left(\nabla x^j\times\nabla x^k\right) = \\ &= \epsilon_{ijk}\nabla\cdot\left[\nabla\times\left(x^j\nabla x^k\right) - x^j\,\nabla\times\nabla x^k\right] = \epsilon_{ijk}\nabla\cdot\left[\nabla\times\left(x^j\nabla x^k\right)\right] = 0,\end{aligned}$$

where the next to last equality follows from the property that the curl of a gradient is always zero, and the last one from the property that the divergence of a curl is always zero. Therefore, the *divergence of a vector field* $\boldsymbol{A}\left(x^1, x^2, x^3\right)$ is

$$\nabla\cdot\boldsymbol{A} = \nabla\cdot\left(A^i\boldsymbol{e}_i\right) = \nabla\cdot\left(\sqrt{g}A^i\frac{\boldsymbol{e}_i}{\sqrt{g}}\right) = \sqrt{g}A^i\,\nabla\cdot\left(\frac{\boldsymbol{e}_i}{\sqrt{g}}\right) + \frac{\boldsymbol{e}_i}{\sqrt{g}}\cdot\nabla\left(\sqrt{g}A^i\right) = \frac{\boldsymbol{e}_i}{\sqrt{g}}\cdot\nabla\left(\sqrt{g}A^i\right)$$

and using the gradient formula (A.31) we have

$$\nabla\cdot\boldsymbol{A} = \frac{1}{\sqrt{g}}\frac{\partial\left(\sqrt{g}A^i\right)}{\partial x^i} \tag{A.32}$$

Last but certainly not least, as it is of primer concern for the magnetic field, the *curl of a vector field* $\boldsymbol{A}\left(x^1, x^2, x^3\right)$ is

$$\nabla\times\boldsymbol{A} = \nabla\times\left(A_k\boldsymbol{e}^k\right) = A_k\nabla\times\boldsymbol{e}^k + \nabla A_k\times\boldsymbol{e}^k = A_k\nabla\times\nabla x^k + \nabla A_k\times\boldsymbol{e}^k = \nabla A_k\times\boldsymbol{e}^k. \tag{A.33}$$

We pause here for a moment our calculation in order to derive first a widely known expression in the literature for magnetic fields,

$$\nabla\times\boldsymbol{A} = \nabla A_k\times\nabla x^k. \tag{A.34}$$

Continuing from (A.33), we find

$$\nabla\times\boldsymbol{A} = \frac{\partial A_k}{\partial x^j}\,\boldsymbol{e}^j\times\boldsymbol{e}^k = \frac{1}{\sqrt{g}}\,\epsilon^{ijk}\frac{\partial A_k}{\partial x^j}\,\boldsymbol{e}_i \tag{A.35}$$

We finally calculate the cross product of a vector field $\boldsymbol{v}\left(x^1, x^2, x^3\right)$ with the curl of another vector field $\boldsymbol{A}\left(x^1, x^2, x^3\right)$ obviously because of its appearance in the Lorentz force. According

to (A.29), (A.35) and the identity $\epsilon_{ijk}\epsilon^{lmk} = \delta_i^l\delta_j^m - \delta_i^m\delta_j^l$ we have

$$\begin{aligned}\boldsymbol{v}\times(\nabla\times\boldsymbol{A}) &= \sqrt{g}\,\epsilon_{ijk}v^j\,(\nabla\times\boldsymbol{A})^k\,\boldsymbol{e}^i = \sqrt{g}\,\epsilon_{ijk}v^j\frac{1}{\sqrt{g}}\,\epsilon^{klm}\frac{\partial A_m}{\partial x^l}\,\boldsymbol{e}^i = v^j\frac{\partial A_m}{\partial x^l}\,\epsilon_{ijk}\epsilon^{lmk}\boldsymbol{e}^i \\ &= v^j\frac{\partial A_m}{\partial x^l}\left(\delta_i^l\delta_j^m - \delta_i^m\delta_j^l\right)\boldsymbol{e}^i\end{aligned}$$

or

$$\boldsymbol{v}\times(\nabla\times\boldsymbol{A}) = v^j\left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}\right)\boldsymbol{e}^i \tag{A.36}$$

**Remark A.1.** So we see that while the contravariant basis comes naturally in the gradient of a function, the covariant one is more appropriate for the curl of a vector. On the other hand, while contravariant components of a vector are suitable for its divergence, covariant ones, on the contrary, are more convenient for its curl. Disappointingly enough, the physical components in any case, while necessary for real experiments, lack easy mathematical manipulation.

**Remark A.2.** In the case of Cartesian coordinates, the covariant, contravariant and physical bases all coincide with $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$, while the metric is of course the unit matrix. Thus, there is no need for distinguishing vector components with superscripts or subscripts, and so only subscripts are used for Cartesian components in this thesis.

# Appendix B

# Differential Forms

In continuation of the previous appendix, the geometry of the magnetic field lines can be better understood and further investigated when we move on to the notion of a differential form. Borrowing the least possible notions from Differential Geometry, in order not to get carried away, we merely touch few concepts on this wide subject for the purpose of just introducing the so-called symplectic forms. The latter will prove quite helpful in particular for exhibiting the intrinsic Hamiltonian nature of the magnetic field. Therefore in this short descriptive chapter we continue our previous discussion from appendix A and bring in little of the artillery of differentiable manifolds, passing from $\mathbb{R}^3$ to $\mathbb{R}^m$ (actually $m = 2, 3, 6$ will concern us in this thesis).

## B.1 Vector Fields

First, consider a vector in the Euclidean space $\mathbb{R}^m$ or more generally in an $m$-dimensional differentiable manifold $M$ in terms of the covariant basis at a particular point $x = \left(x^1, \ldots, x^m\right)$, denoted by $\mathbf{v}_x = v_x^i \boldsymbol{e}_i$. For any such vector, we can define a map, denoted again by $\mathbf{v}_x$, which for any smooth function $f(x)$ assigns the directional derivative of $f$ along the vector $\mathbf{v}_x$,

$$\mathbf{v}_x(f) = v_x^i \left.\frac{\partial f}{\partial x^i}\right|_x \tag{B.1}$$

In fact, this relation gives rise to an isomorphism between tangent vectors and derivations for any finite-dimensional differentiable manifold. Let $C^\infty(M)$ denote the space of all smooth, real functions defined on $M$.

**Definition B.1.** A *derivation* is a differential operator $\mathbf{v}_x : C^\infty(U) \longrightarrow \mathbb{R}$ defined on the space

of smooth functions on a neighborhood $U$ of $M$, which for every $f, g \in C^\infty(U)$ and $a, b \in \mathbb{R}$ satisfies the following properties:

1. $\mathbf{v}_x(af + bg) = a\mathbf{v}_x(f) + b\mathbf{v}_x(g)$ (linearity)
2. $\mathbf{v}_x(fg) = f(x)\mathbf{v}_x(g) + g(x)\mathbf{v}_x(f)$ (Leibniz' rule)

Therefore, in any finite-dimensional smooth manifold $M$, a tangent vector at a point $x$ can also be expressed as a derivation defined in some neighborhood $U$ of $x$. For example, the covariant basis vectors of the tangent space $T_xM$ if we consider equation (B.1) act as the partial derivatives at the point $x$,

$$\boldsymbol{e}_i = \frac{\partial}{\partial x^i}\bigg|_x \tag{B.4}$$

for any function $f$ on $U$. Note that in terms of this expression, relation (A.7) is just the chain rule.

If, however, we are not interested at a particular point $x$, meaning if, for example, we consider a vector function varying from point to point that needs to be defined globally, then we arrive at the notion of a vector field.

**Definition B.2.** A *vector field* $\mathbf{v}$ on a smooth manifold $M$ is a smooth map which at every point $x$ assigns a tangent vector $\mathbf{v}_x$ of $T_xM$,

$$\mathbf{v}(x) = \mathbf{v}_x \tag{B.5}$$

Equivalently it can be defined as a derivation $\mathbf{v} : C^\infty(M) \longrightarrow C^\infty(M)$ determined by the map $\mathbf{v}(f) : M \longrightarrow \mathbb{R}$ with

$$\mathbf{v}(f)(x) = \mathbf{v}_x(f) \tag{B.6}$$

The reason that the mapping (B.5) takes place in the tangent space is that a manifold is not in general equipped with the structure of a vector space, as in the case for example of the Euclidean space $\mathbb{R}^m$, while the tangent space is. According to the above definition, we have again an isomorphism between vector fields and derivations on $M$, just like tangent vectors and derivations at a point $x$ of $M$. It is easy to see that the space $X(M)$ of all vector fields defined on $M$ has the structure of a vector space over $\mathbb{R}$, and it can be shown that the partial derivatives $\partial/\partial x^i$ form a basis in $X(M)$ in agreement with intuitively removing the index $x$ from (B.4). Thus, every vector field can be expressed as

$$\mathbf{v} = v^i(x)\frac{\partial}{\partial x^i} \tag{B.7}$$

where $v^i(x) = v^i_x = \mathbf{v}_x\big(x^i\big)$ are smooth functions. Staying consistent with the notation adopted here, the passage from tangent vectors (B.1) to vector fields (B.7) is reflected on simply dropping the subscripts $x$ denoting the local nature.

The above considerations for contravariant components and covariant basis vectors can be carried over to covariant components and contravariant basis vectors.

## B.2 1-Forms

Recall first from the previous section that at any point $x$ every linear form of the tangent space, called *covector*, is an element of the dual space, called the cotangent space $T^*_x M$, whose vector basis $\boldsymbol{e}^i$ was defined by the relations $\boldsymbol{e}^i(\boldsymbol{e}_j) = \delta^i_j$ that we saw earlier in (A.13) described by the dot product for the case of $\mathbb{R}^3$. As with the covariant basis, the contravariant one can also be interpreted by a differential operator.

**Definition B.3.** Let $M$ be a differentiable manifold and $f$ a smooth function defined in some neighborhood $U$ of a point $x$. *Differential* of $f$ is a linear map $(df)_x : T_x M \longrightarrow \mathbb{R}$, i.e. a covector, defined by the relation

$$(df)_x(\mathbf{v}_x) = \mathbf{v}_x(f) \tag{B.8}$$

From the above definition and (B.1), we see that if $f(x) = x^i$ and $\mathbf{v}_x = \boldsymbol{e}_j$, then the differentials of the local coordinates satisfy the definition of the dual basis,

$$\big(dx^i\big)_x(\boldsymbol{e}_j) = \delta^i_j$$

In contrast to (B.4), this means that the contravariant basis for the cotangent space is

$$\boldsymbol{e}^i = \big(dx^i\big)_x \tag{B.9}$$

Contrary to the notation used in the previous appendix, where the tangent and cotangent spaces were all isomorphic to $\mathbb{R}^3$, here covectors to be distinguished from tangent vectors will be denoted by $\omega_x$, where the index $x$ is used again to indicate the local nature. In terms of the contravariant basis (B.9) every covector can be expressed as

$$\omega_x = \omega_{i_x}\big(dx^i\big)_x \tag{B.10}$$

and its action on vectors in general is given by the relation

$$\omega_x(\mathbf{v}_x) = \omega_{i_x} v^i_x \tag{B.11}$$

Returning to the differential of a function, for any tangent vector $\mathbf{v}_x = v_x^i e_i$, we also have

$$(df)_x(\mathbf{v}_x) = (df)_x\big(v_x^i \boldsymbol{e}_i\big) = v_x^i (df)_x(\boldsymbol{e}_i) = \delta_j^i v_x^j \left.\frac{\partial f}{\partial x^i}\right|_x = \left.\frac{\partial f}{\partial x^i}\right|_x v_x^j \left(dx^i\right)_x(\boldsymbol{e}_j) =$$
$$= \left.\frac{\partial f}{\partial x^i}\right|_x \left(dx^i\right)_x\big(v_x^j \boldsymbol{e}_j\big) = \left.\frac{\partial f}{\partial x^i}\right|_x \left(dx^i\right)_x(\mathbf{v}_x)$$

Therefore, the differential of a function $f$ is indeed expressed as a linear combination of the basis vectors $\left(dx^i\right)_x$,

$$(df)_x = \left.\frac{\partial f}{\partial x^i}\right|_x \left(dx^i\right)_x \tag{B.12}$$

in terms of which its (covariant) components are given by the components of the gradient of $f$ at the point $x$.

Similarly to the generalization from vectors to vector fields, again when we are not referring to a fixed point on the manifold, covectors generalize to the notion of a field, called a (differential) 1-form.

**Definition B.4.** A *1-form* on a differentiable manifold $M$ is a smooth map $\omega$ which at every point $x$ assigns a cotangent vector $\omega_x$ of $T_x^*M$,

$$\omega(x) = \omega_x \tag{B.13}$$

Equivalently it can be defined as the smooth map $\omega : X(M) \longrightarrow C^\infty(M)$ determined from the relation

$$\omega(\mathbf{v})(x) = \omega_x(\mathbf{v}_x) \tag{B.14}$$

It is not difficult to see that the space of 1-forms is a vector space, and just like (B.7), dropping the index $x$, every 1-form can be written as

$$\omega = \omega_i(x) dx^i \tag{B.15}$$

where $\omega_i(x) = \omega_{i_x} = \omega_x(\boldsymbol{e}_i)$ are smooth functions. So, in local coordinates, the action (B.14) of the 1-form (B.15) on the vector field (B.7) can be written as $\omega(\mathbf{v})(x) = \omega_i(x) v^i(x)$ or

$$\omega(\mathbf{v}) = \omega_i v^i \tag{B.16}$$

Besides the two naturally defined operations of the vector space structure, we introduce another two giving rise to higher order forms and present their counterparts with the aid of a third one for the particular case of $\mathbb{R}^3$.

## B.3 Form Operations

### Wedge product

Starting constructively in terms of 1-forms, the *wedge product* $\wedge$ is the unique map of any two 1-forms $\omega^1$ and $\omega^2$ that satisfies the relation[1]

$$\omega^1 \wedge \omega^2 (\mathbf{v}_1, \mathbf{v}_2) = \det\big(\omega^i(\mathbf{v}_j)\big) = \det\Big(\omega^i_k v^k_j\Big) . \tag{B.17}$$

Noting the similarity to (B.14), which defines a 1-form $\omega$ as a map from a vector field $\mathbf{v}$ to a smooth function, the result of the wedge product $\omega^1 \wedge \omega^2$ is actually a 2-form mapping two vector fields $\mathbf{v}_1$ and $\mathbf{v}_2$ to a smooth function, i.e. a smooth map $X(M) \times X(M) \longrightarrow C^\infty(M)$, defined by (B.17) in this case. Considering a generalization of (B.17) for the case of $k$ vector fields $\mathbf{v}_1, \ldots, \mathbf{v}_k$, taking the wedge product of $k$ 1-forms $\omega^1, \ldots, \omega^k$, we can construct a $k$-form $\omega^1 \wedge \cdots \wedge \omega^k$. The wedge product can be further generalized for any $k_1$-form $\omega^1$ and $k_2$-form $\omega^2$, yielding a $(k_1{+}k_2)$-form, and satisfies the following fundamental properties:

1. $\omega^1 \wedge \big(a\omega^2 + b\omega^3\big) = a\,\omega^1 \wedge \omega^2 + b\,\omega^1 \wedge \omega^3$,
   $\big(a\omega^1 + b\omega^2\big) \wedge \omega^3 = a\,\omega^1 \wedge \omega^3 + b\,\omega^2 \wedge \omega^3$ (bilinearity)
2. $\omega^1 \wedge \omega^2 = (-1)^{k_1 k_2} \omega^2 \wedge \omega^1$ (anticommutativity)
3. $\omega^1 \wedge \big(\omega^2 \wedge \omega^3\big) = \big(\omega^1 \wedge \omega^2\big) \wedge \omega^3$ (associativity)

for any $a, b \in \mathbb{R}$. For 1-forms the above properties follow easily from (B.17) and the usual properties of determinants.

We pause here for a moment to comment that, in general, $k$-forms on a manifold $M$ are covariant alternating $k$-tensor fields, and by a covariant $k$-tensor at a point $x$ we mean a multilinear (actually $k$-linear) form of the Cartesian product of $T_xM$ with itself $k$ times, just as covectors are linear forms of $T_xM$, meaning covariant 1-tensors. The wedge product plays a fundamental role in tensor analysis, since it provides a basis for the space of alternating tensors. Closing this brief aside, for our puproses here, we only need to keep in mind that $dx^1 \wedge \cdots \wedge dx^k$ is a basis for $k$-forms and every $k$-form $\omega$ can be expressed as

$$\omega = \omega_{i_1 i_2 \cdots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k} \tag{B.21}$$

where $I = (i_1, i_2, \ldots, i_k)$ is any strictly increasing $k$-tuple, meaning $1 \leq i_1 < i_2 < \cdots < i_k \leq m$. Given the above local expression of a $k$-form and the definition of the wedge product, we can

1. Notice that superscripts on $\omega$ indicate different forms, while subscripts stand for their components.

also derive the action, for example, of a 2-form $\omega = \omega_{ij} dx^i \wedge dx^j$ on two vector fields $\mathbf{v} = v^i \partial_{x^i}$ and $\mathbf{w} = w^i \partial_{x^i}$. For every $i, j$ we have

$$\begin{aligned}\omega(\mathbf{v}, \mathbf{w}) &= \frac{1}{2}\,\omega_{ij}\, dx^i \wedge dx^j(\mathbf{v}, \mathbf{w}) = \frac{1}{2}\,\omega_{ij}\big[dx^i(\mathbf{v})dx^j(\mathbf{w}) - dx^i(\mathbf{w})dx^j(\mathbf{v})\big] \\ &= \frac{1}{2}\,\omega_{ij}\big(v^i w^j - w^i v^j\big) = \frac{1}{2}\, v^i w^j\,(\omega_{ij} - \omega_{ji}) = \omega_{ij} v^i w^j \end{aligned} \tag{B.22}$$

### Exterior derivative

Another basic operation in the calculus of differentiable manifolds is the *differential or exterior derivative*. Given away by its name, we have already encountered this operation in terms of the covector $(df)_x$. First we define the 1-form analogue of the latter, removing the local nature of (B.8) and (B.12).

**Definition B.5.** Let $M$ be a differentiable manifold and $f$ a smooth function defined on $M$. Differential of $f$ is the 1-form $df : X(M) \longrightarrow C^\infty(M)$ defined by the relation

$$df(\mathbf{v}) = \mathbf{v}(f) \tag{B.23}$$

From the above definition, the components of $df$ are easily found to be the partial derivatives of $f$, meaning

$$df = \frac{\partial f}{\partial x^i}\, dx^i \tag{B.24}$$

Now let $\omega = \omega_i(x) dx^i$ be a 1-form on a smooth manifold $M$. The differential of $\omega$ is a 2-form defined as

$$d\omega = d\omega_i \wedge dx^i \tag{B.25}$$

where $d\omega_i$ is the differential of the component functions $\omega_i$ in the sense of Definition B.5 and when further analyzed according to (B.24),

$$\begin{aligned} d\omega &= \frac{\partial \omega_i}{\partial x^j}\, dx^j \wedge dx^i = \frac{1}{2}\left(\frac{\partial \omega_i}{\partial x^j}\, dx^j \wedge dx^i + \frac{\partial \omega_i}{\partial x^j}\, dx^j \wedge dx^i\right) \\ &= \frac{1}{2}\left(\frac{\partial \omega_j}{\partial x^i}\, dx^i \wedge dx^j - \frac{\partial \omega_i}{\partial x^j}\, dx^i \wedge dx^j\right) = \frac{1}{2}\left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j}\right) dx^i \wedge dx^j \end{aligned}$$

for all $i, j$. Note, however, that in the last double summation the interchange of $i$ and $j$ yield the same term, meaning we can write

$$d\omega = \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j}\right) dx^i \wedge dx^j, \tag{B.26}$$

where $i < j$ as usual, following the common practice for higher order forms.

Going through the case of a 1-form with the least possible indices, one can perceive the generalization to higher order forms. Differential of a $k$-form $\omega = \omega_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ is a $(k{+}1)$-form that can be defined as

$$d\omega = d\omega_{i_1 i_2 \cdots i_k} \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k} \tag{B.27}$$

As the previous case for 1-forms showed, the implications of the above expression, when one tries to write $d\omega$ using a basis in terms of $dx^i$, require certain caution regarding ordered indices. In the proof of Proposition B.14 the differential of a 2-form is required, revealing how the complexity of calculations increases as we move to higher order forms.

**Remark B.6.** If $\omega = \omega_{1\cdots n} dx^1 \wedge \cdots \wedge dx^m$ is an $m$-form, then $d\omega = 0$, since $dx^i \wedge dx^j = 0$ for any $i = j$ due to the anticommutativity of the wedge product. Actually, the same argument indicates that every $(m{+}1)$-form on an $m$-dimensional manifold vanishes, silently implied by the restriction $i_k \leq m$ on the local expression of $k$-forms in (B.21).

Although the above definition of the differential is given in local coordinates, the operator $d$ taking $k$-forms to $(k{+}1)$-forms is actually uniquely determined by the following properties:

1. $d(a\omega^1 + b\omega^2) = a\, d\omega^1 + b\, d\omega^2$ (linearity)
2. $d(d\omega) = 0$ (closure)
3. $d\big(\omega^1 \wedge \omega^2\big) = d\omega^1 \wedge \omega^2 + (-1)^{k_1} \omega^1 \wedge d\omega^2$ (antiderivation)

for any forms $\omega, \omega^1, \omega^2$ and $a, b \in \mathbb{R}$, where $k_1$ is the order of $\omega^1$, along with Definition B.5 for the differential of a function, regarded as a 0-form. Linearity is straightforward, while antiderivation is a direct consequence of Leibniz' rule, that is, considering $\omega^1 = f dx^{i_1} \wedge \cdots \wedge dx^{i_{k_1}}$ and $\omega^2 = g\, dx^{j_1} \wedge \cdots \wedge dx^{j_{k_2}}$, then

$$\begin{aligned}
d\left(\omega^1 \wedge \omega^2\right) &= d\left(\left(f dx^{i_1} \wedge \cdots \wedge dx^{i_{k_1}}\right) \wedge \left(g\, dx^{j_1} \wedge \cdots \wedge dx^{j_{k_2}}\right)\right) \\
&= d\left(fg\, dx^{i_1} \wedge \cdots \wedge dx^{i_{k_1}} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_{k_2}}\right) \\
&= d(fg) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k_1}} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_{k_2}} \\
&= (g df + f dg) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k_1}} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_{k_2}} \\
&= \left(df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k_1}}\right) \wedge \left(g\, dx^{j_1} \wedge \cdots \wedge dx^{j_{k_2}}\right) + \\
&\qquad + (-1)^{k_1} \left(f dx^{i_1} \wedge \cdots \wedge dx^{i_{k_1}}\right) \wedge \left(dg \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_{k_2}}\right) \\
&= d\omega^1 \wedge \omega^2 + (-1)^{k_1} \omega^1 \wedge d\omega^2
\end{aligned}$$

To prove the second property, i.e. that $d\omega$ is *closed*, it suffices to prove it for a 0-form, meaning that $df$ is closed,

$$\begin{aligned} d(df) &= d\left(\frac{\partial f}{\partial x^j}\,dx^j\right) = \frac{\partial^2 f}{\partial x^i \partial x^j}\,dx^i \wedge dx^j \quad \text{(for all } i,j\text{)} \\ &= \left(\frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial^2 f}{\partial x^j \partial x^i}\right) dx^i \wedge dx^j \qquad \text{(for } i<j\text{)} \end{aligned}$$

which simply reduces to mixed second-order partial derivatives being equal. Using $d(df) = 0$ and the third property, we can easily derive $d(d\omega) = 0$ for any $k$-form.

**Remark B.7.** Every $k$-form, say $B$, that comes from the differential of a ($k$-1)-form, say $A$, meaning $B = dA$, is called *exact*, and the above second property verifies that every exact form is closed. The converse of this statement is not always true, that is, every closed form is not necessarily exact. The question of equivalent classes of closed forms is closely related to the topological characteristics of the manifold considered, described by the so-called *de Rham cohomology* theory. However, the celebrated *Poincare's lemma* assures us that in simply connected manifolds, like $\mathbb{R}^3$ which we are mostly concerned with, every closed form is exact.

Before moving to the third and final operation, we briefly go through two necessary notions to define it, which otherwise require whole chapters for their description.

**Remark B.8.** A counterpart notion to $k$-forms, are *$k$-vector fields*, which are contravariant alternating $k$-tensor fields, contravariant $k$-tensors at a point $x$ being $k$-linear forms of the Cartesian product of $T_x^*M$ with itself $k$ times. Obviously $k = 1$ corresponds to the familiar vector fields we encountered. The construction of $k$-vector fields follows a similar path to that of $k$-forms through related operations, such as the wedge product, the exterior derivative, etc. Just for comparison with $k$-forms, a $k$-vector field $V$ for example is expressed as

$$V = V^{i_1 i_2 \cdots i_k} \frac{\partial}{\partial x^{i_1}} \wedge \frac{\partial}{\partial x^{i_2}} \wedge \cdots \wedge \frac{\partial}{\partial x^{i_k}} \tag{B.31}$$

where again $1 \le i_1 < i_2 < \cdots < i_k \le m$. Limiting this remark here and mainly to the above equation, we refer to classical textbooks of Differential Geometry for further study.

**Remark B.9.** An $n$-dimensional manifold $M$ is said to be *orientable* if the Jacobian matrix defined for any change of coordinate bases of $T_xM$ has the same sign at each point $x \in M$. It turns out that a necessary and sufficient condition for orientation is the existence of a nowhere vanishing $n$-form $\omega = f dx^1 \wedge \cdots \wedge dx^m$.

### Hodge star operator

So, finally, a useful operation that clarifies the geometrical description of the magnetic field in $\mathbb{R}^3$ and its various expressions in different contexts is the Hodge star operator. Keeping in mind the previous remarks, in an orientable manifold equipped with a nowhere vanishing form $\omega = f dx^1 \wedge \cdots \wedge dx^m$, the *Hodge star operation* $*$ takes a $k$-form $B = B_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ to an ($m$-$k$)-vector field $*B$ with components

$$(*B)^{i_{k+1} \cdots i_m} = \frac{1}{f} \epsilon^{i_1 \cdots i_k i_{k+1} \cdots i_m} B_{i_1 \cdots i_k} \tag{B.32}$$

where $\epsilon^{i_1 \cdots i_m}$ is the generalized Levi-Civita symbol, which is not a tensor (and for which sub- or superscripts don't make any difference and are only retained just to be consistent with our summation rule), and let us stress again that $i_1 < i_2 < \cdots < i_k$. Following [96], the above definition is given in local coordinates, without having the notion of Riemannian metric entering the picture. In every manifold however there is an isomorphism between $k$-forms and ($m$-$k$)-forms (as well as between $k$-vector fields and ($m$-$k$)-vector fields). For Riemannian manifolds this isomorphism can be naturally defined in terms of the Hodge star operator. And since every smooth manifold admits a Riemannian metric, in most textbooks, the $*$ operator is defined in a coordinate-free fashion, taking a $k$-form to an ($m$-$k$)-form (or a $k$-vector field to an ($m$-$k$)-vector field). This is by no means an ambiguity, since in Riemannian manifolds the isomorphism between 1-vector fields and 1-forms allows to use the common practice of what is widely known as *lowering or raising indices* and transform a $k$-vector field to a $k$-form and vice versa, leading equivalently to the above definition adopted here. Note that for the same reason, one could consider the Hodge star operator taking a $k$-vector field to an ($m$-$k$)-form. All the same, since our target is basically the three-dimensional Euclidean space, definition (B.32) suffices, avoiding to introduce the concept of the Riemannian metric in general. Instead, the special case of the metric defined in the previous appendix in terms of curvilinear coordinates (see (A.21), (A.20) and (A.23)) serves our purposes for $\mathbb{R}^3$ just as well, as we will see right after the next subsection.

### Push-forward and Pull-back

For the needs of the next section and in particular the cotangent bundle construction, we close this one by also describing how vectors and covectors are mapped between two manifolds.

The differential of a smooth map $g : M \longrightarrow N$ between two manifolds $M$ and $N$, in general, is defined as the map $(dg)_x : T_x M \longrightarrow T_{g(x)} N$ where $(dg)_x(\mathbf{v}_x) = \mathbf{v}_{g(x)}$ and $\mathbf{v}_{g(x)}(f) = \mathbf{v}_x(f \circ g)$ for any $x \in M$ and any smooth real function $f$ defined on $N$. On the other hand, the *codifferential* of

$g$, defined as the dual map to $(dg)_x$, is the map $(dg)^*_x : T^*_{g(x)}N \longrightarrow T^*_x M$ where $(dg)^*_x(\omega_{g(x)}) = \omega_x$. The relation between the two expresses that either way we arrive at the same number,

$$\omega_x(\mathbf{v}_x) = \omega_{g(x)}(\mathbf{v}_{g(x)}) \tag{B.33}$$

If we extend these pointwise maps for any point $x$, we arrive at the notion of the push-forward and the pull-back maps. Starting with the codifferential, considering 1-forms instead of covectors, the *pull-back* $g^*$ of a 1-form $\omega$ on $N$ by $g$ is the 1-form $g^*(\omega)$ on $M$ defined as $(g^*(\omega))_x = \omega_x$, where $\omega_x$ is defined as above. Although no extra requirements are needed for the definition of $g^*$, on the contrary, if $g$ is not a diffeomorphism[2] then $\mathbf{v}_{g(x)}$ may not be assigned for every point of $N$ nor may be unique. Thus, if $g$ is a diffeomorphism, the *push-forward* $g_*$ of a vector field $\mathbf{v}$ on $M$ by $g$ is a vector field $g_*(\mathbf{v})$ on $N$ defined as $(g_*(\mathbf{v}))_y = \mathbf{v}_y$, for any point $y$ of $N$, where $\mathbf{v}_y = (dg)_{g^{-1}(y)}(\mathbf{v}_{g^{-1}(y)})$. If $y = g(x)$, equation (B.33) can then be expressed as

$$g^*(\omega)(\mathbf{v})(x) = \omega(g_*(\mathbf{v}))(y) \tag{B.34}$$

On this ground, these maps can be defined for any tensor field, for example the pull-back of a 2-form $\omega$ on $N$ can be constructed in terms of the push-forward of vector fields $\mathbf{v}$ and $\mathbf{w}$ on $M$, as in

$$g^*(\omega)(\mathbf{v}, \mathbf{w})(x) = \omega(g_*(\mathbf{v}), g_*(\mathbf{w}))(y) \tag{B.35}$$

Some properties needed in the following are $(f \circ g)_* = f_* \circ g_*$, $(f \circ g)^* = g^* \circ f^*$ and commutativity with the exterior derivative, $d(g^*\omega) = g^*(d\omega)$ for any $k$-form $\omega$.

Last but not least, it is crucial at this point to retrieve many of the above notions and summarize them for the case of $\mathbb{R}^3$, the ultimate goal being an easy passage from the usual vector calculus to differential forms and back, where needed.

### The Euclidean space $\mathbb{R}^3$

Consider a right-handed curvilinear coordinate system $\left(x^1, x^2, x^3\right)$, giving $\mathbb{R}^3$ a (positive) orientation related to a nonvanishing volume form

$$\omega = \sqrt{g}\, dx^1 \wedge dx^2 \wedge dx^3 \tag{B.36}$$

Here $g$ is the determinant of the metric matrix defined in (A.21), which as we saw in (A.24) yields the volume of the parallelepiped spanned by the covariant basis vectors. Recall from

2. that is, besides smooth, onto (surjective) and 1-1 (injective) with a smooth inverse

the previous appendix that we can switch easily from vector fields to 1-forms (see (A.20) and (A.23)), following the notation of this appendix though, 1-forms won't be denoted by bold letters as opposed to the corresponding vector fields.

We start off by comparing (A.15) and (B.9) or (A.31) and (B.24), clearly indicating that the gradient of a scalar function $f$ is described by the differential of the 0-form $f$, meaning

$$\nabla f = df \tag{B.37}$$

Interpreting the above equation backwards, every closed 1-form is the gradient of a function, and every 1-form in general can be written as a linear combination of gradients.

Moving on to higher order forms, we can see that in terms of (B.36) every 2-form in $\mathbb{R}^3$

$$B = B_{ij}\, dx^i \wedge dx^j \tag{B.38}$$

under the Hodge star operation can be mapped to a vector field $\boldsymbol{B}$ with components given by (B.32)

$$B^k = \frac{1}{\sqrt{g}}\, \epsilon^{ijk} B_{ij} \qquad (\text{for } i < j) \tag{B.39}$$

(taking $i < j$ implies no summation), that is

$$\boldsymbol{B} = *B = \frac{1}{\sqrt{g}}\, \epsilon^{ijk} B_{ij} \frac{\partial}{\partial x^k} \qquad (\text{for } i < j) \tag{B.40}$$

Now according to (B.51) given later on, the differential of the 2-form (B.38) is

$$dB = \left( \frac{\partial B_{23}}{\partial x^1} + \frac{\partial B_{31}}{\partial x^2} + \frac{\partial B_{12}}{\partial x^3} \right) dx^1 \wedge dx^2 \wedge dx^3$$

and in terms of (B.39) vanishes whenever

$$\frac{\partial B_{23}}{\partial x^1} + \frac{\partial B_{31}}{\partial x^2} + \frac{\partial B_{12}}{\partial x^3} = 0 \Leftrightarrow \frac{\partial \left(\sqrt{g} B^1\right)}{\partial x^1} + \frac{\partial \left(\sqrt{g} B^2\right)}{\partial x^2} + \frac{\partial \left(\sqrt{g} B^3\right)}{\partial x^3} = 0$$

Recalling then (A.32), we conclude that the closure of $B$ is nothing more than the divergence-free condition for $\boldsymbol{B}$,

$$dB = 0 \Leftrightarrow \nabla \cdot \boldsymbol{B} = 0 \tag{B.41}$$

Actually we can take one more time the Hodge star operator to deduce that the divergence of a vector function $\boldsymbol{B}$ is the Hodge star operation of the differential of the corresponding 2-form $B$,

$$\nabla \cdot \boldsymbol{B} = *\, dB \tag{B.42}$$

Since we have seen where the differential of 0-forms and 2-forms leads to, it is only natural to consider now a 1-form

$$A = A_i dx^i$$

and take the differential of $A$ given in (B.26),

$$B = dA = \left( \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} \right) dx^i \wedge dx^j \qquad (\text{for } i < j) \tag{B.43}$$

Then the Hodge star operator maps $B$, again through relations (B.39), to the vector field

$$\boldsymbol{B} = *B = \frac{1}{\sqrt{g}} \epsilon^{ijk} \frac{\partial A_k}{\partial x^j} \frac{\partial}{\partial x^i} \tag{B.44}$$

Compared to (A.35), we easily identify the curl of a vector function $\boldsymbol{A}$ with the Hodge star operation of the differential of the corresponding 1-form $A$,

$$\nabla \times \boldsymbol{A} = * \, dA \tag{B.45}$$

Let us close the this subsection by recovering one more useful correspondence in $\mathbb{R}^3$ between the cross and the wedge product. First,

$$\begin{aligned} df \wedge dg &= \left( \frac{\partial f}{\partial x^i} dx^i \right) \wedge \left( \frac{\partial g}{\partial x^j} dx^j \right) = \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j} dx^i \wedge dx^j \\ &= \frac{1}{2} \left( \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j} dx^i \wedge dx^j + \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j} dx^i \wedge dx^j \right) \\ &= \frac{1}{2} \left( \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j} dx^i \wedge dx^j + \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^i} dx^j \wedge dx^i \right) = \frac{1}{2} \left( \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j} - \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^i} \right) dx^i \wedge dx^j \end{aligned}$$

for all $i, j$. Then, returning to the usual form notation for $i < j$, and taking the Hodge star, based once more on (B.39), results in

$$df \wedge dg = \left( \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j} - \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^i} \right) dx^i \wedge dx^j \quad \Rightarrow \quad *(df \wedge dg) = \frac{1}{\sqrt{g}} \epsilon^{ijk} \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^k} \frac{\partial}{\partial x^i}$$

So, in light of (A.30), we see that the cross product of the gradients of two functions $f$ and $g$ is the Hodge star of the wedge product of the differentials of $f$ and $g$,

$$\nabla f \times \nabla g = *(df \wedge dg) \tag{B.46}$$

**Remark B.10.** We finally comment that since the Hodge star operator is a 1-1 map, consequently in all three relations (B.42), (B.45) and (B.46) we can switch straightaway from usual vector operations to differential forms, e.g. from $\nabla f \times \nabla g$ to $df \wedge dg$.

## B.4 Symplectic and Presymplectic Forms

Based on the language of forms developed, we can now introduce the notion and some first implications of a symplectic structure. Being a differential form, its description is given here as a natural continuation of the previous sections. Its connection, however, with Hamiltonian systems is left for the next appendix studied through Poisson structures, avoiding further tools of Differential Geometry. We also define the less-known presymplectic forms, which fit perfectly in the description of the magnetic field, differing from symplectic ones by the nondegeneracy condition.

A 2-form $\omega$ is called *nondegenerate* if $\omega(\mathbf{v}, \mathbf{w}) = 0$ for any vector field $\mathbf{w}$ implies $\mathbf{v} = 0$. This means that the kernel of the map $\omega^\flat$ defined as the 1-form $\omega^\flat(\mathbf{v}) = \omega(\ , \mathbf{v})$, meaning $\omega^\flat(\mathbf{v}) = \omega_{ij} v^j dx^i$,

$$\ker \omega^\flat = \{\mathbf{v} : \omega(\mathbf{v}, \mathbf{w}) = 0, \forall\, \mathbf{w}\} = \left\{\mathbf{v} : \omega_{ij} v^j = 0\right\}, \tag{B.47}$$

(see also (B.22)), consisting of all the vector fields $\mathbf{v}$ that annihilate $\omega$ for any vector field $\mathbf{w}$, is empty.

**Definition B.11.** A closed, nondegenerate 2-form $\omega$ on a manifold $M$ is called a *symplectic form*, and $M$ equipped with $\omega$ is called a symplectic manifold.

**Proposition B.12.** Every symplectic manifold is even-dimensional.

*Proof.* Let $M$ be an $m$-dimensional symplectic manifold with $\omega = \omega_{ij}(x) dx^i \wedge dx^j$ being the symplectic form in local coordinates. Since $\omega$ is a 2-form, meaning an alternating 2-tensor field, $\omega_{ij} = -\omega_{ji}$, and so the matrix $\Omega = (\omega_{ij})$ is skew-symmetric, i.e. $\Omega = -\Omega^T$. Therefore, $|\Omega| = \left|-\Omega^T\right| = (-1)^m \left|\Omega^T\right| = (-1)^m |\Omega|$. On the other hand, since $\omega$ is nondegenerate, we have that $\omega(\mathbf{v}, \partial_{x^j}) = 0 \Rightarrow \mathbf{v} = 0$, or from (B.22) that $\Omega\, \mathbf{v} = 0 \Rightarrow \mathbf{v} = 0$, meaning $\Omega$ is invertible, $|\Omega| \neq 0$. Thus, $|\Omega| = (-1)^m |\Omega| \Rightarrow 1 = (-1)^m \Rightarrow m$ is even. □

**Example B.13.** Recalling Remarks B.6 and B.9, we can show that every two-dimensional orientable manifold is a symplectic manifold. Since in this case the closure condition is trivially satisfied for any 2-form, while the nondegeneracy is guaranteed by the orientation. For, in local coordinates, the latter means that there exists a 2-form $\omega = f dx^1 \wedge dx^2$, where $f\left(x^1, x^2\right) \neq 0$, which is obviously nondegenerate and can therefore serve as a symplectic form. In particular, if the manifold is a surface in $\mathbb{R}^3$ equipped with the area form induced by the Euclidean metric, then we can always choose $\omega = \sqrt{g}\, dx^1 \wedge dx^2$, where $g$ is the determinant of the surface's metric.

**Proposition B.14.** Let $M$ be a $2n$-dimensional differentiable manifold. A $2n \times 2n$ matrix $\Omega(x)$ of functions of $x \in M$ defines a symplectic form on $M$, $\omega = \omega_{ij}(x) dx^i \wedge dx^j$, if and only if

1. $|\Omega| \neq 0$ (nondegeneracy)
2. $\omega_{ij} = -\omega_{ji}$ (skew-symmetry)
3. $\dfrac{\partial \omega_{ij}}{\partial x^k} + \dfrac{\partial \omega_{jk}}{\partial x^i} + \dfrac{\partial \omega_{ki}}{\partial x^j} = 0$ (Jacobi identity)

The matrix $\Omega$ satisfying the above three properties will be called the *symplectic structure matrix*.

*Proof.* We have already seen that the first two properties express directly the nondegeneracy and skew-symmetry of the symplectic form. Finally, the third one comes from the closure, for all $i, j$ and $k$ we have

$$\begin{aligned} d\omega &= d\left(\frac{1}{2}\,\omega_{ij} dx^i \wedge dx^j\right) = \frac{1}{2}\frac{\partial \omega_{ij}}{\partial x^k}\, dx^k \wedge dx^i \wedge dx^j \\ &= \frac{1}{6}\left(\frac{\partial \omega_{ij}}{\partial x^k} + \frac{\partial \omega_{ij}}{\partial x^k} + \frac{\partial \omega_{ij}}{\partial x^k}\right) dx^i \wedge dx^j \wedge dx^k \\ &= \frac{1}{6}\left(\frac{\partial \omega_{ij}}{\partial x^k}\, dx^i \wedge dx^j \wedge dx^k + \frac{\partial \omega_{jk}}{\partial x^i}\, dx^j \wedge dx^k \wedge dx^i + \frac{\partial \omega_{ki}}{\partial x^j}\, dx^k \wedge dx^i \wedge dx^j\right) \\ &= \frac{1}{6}\left(\frac{\partial \omega_{ij}}{\partial x^k} + \frac{\partial \omega_{jk}}{\partial x^i} + \frac{\partial \omega_{ki}}{\partial x^j}\right) dx^i \wedge dx^j \wedge dx^k \end{aligned}$$

or, since there are six permutations of the three indices $i, j, k$, if we return to the typical ordered triple $i < j < k$ we get

$$d\omega = \left(\frac{\partial \omega_{ij}}{\partial x^k} + \frac{\partial \omega_{jk}}{\partial x^i} + \frac{\partial \omega_{ki}}{\partial x^j}\right) dx^i \wedge dx^j \wedge dx^k \tag{B.51}$$

Therefore $d\omega = 0$ if and only if the latter 3-form vanishes, i.e. all of its components are zero. □

One of the most characteristic examples of symplectic manifolds that often appears in many problems of Classical Mechanics is the cotangent bundle of a smooth manifold.

**Example B.15. The cotangent bundle.** Let $M$ be an $m$-dimensional smooth manifold. The disjoint union of the cotangent spaces $T^*_x M$ for every point $x$ of $M$ is called the *cotangent bundle* $T^*M$. A point on the manifold $T^*M$ is denoted by $(x, v_x)$, where $x$ is a point of $M$ and $v_x$ a covector of $M$ at $x$, and local coordinates are $\left(x^1, \ldots, x^m, v_{1_x}, \ldots, v_{m_x}\right)$. Note that the cotangent bundle is $(2m)$-dimensional. This construction comes also equipped with the natural projection $\pi : T^*M \longrightarrow M$ defined as $\pi(x, v_x) = x$ for every $x \in M$ and $v_x \in T^*_x M$. The cotangent

bundle is naturally equipped with a symplectic structure, coming from the so called *Liouville or tautological 1-form* $\vartheta$ defined on $T^*M$ as

$$\vartheta(\boldsymbol{X})(w) = v(\pi_*(\boldsymbol{X}))(x) \tag{B.52}$$

for every vector field $\boldsymbol{X}$ on $T^*M$, where $w = (x, v_x)$ and $v(x) = v_x$. The above relation is described in terms of the push-forward of the natural projection which assigns a vector field $\boldsymbol{X}$ on $T^*M$ to a vector field on $M$. However, recalling (B.34), we can take instead the pull-back of $\pi$ which drags every 1-form $v = v_i\, dx^i$ of $M$ to a 1-form in $T^*M$. In other words, from (B.52) we have $\vartheta(\boldsymbol{X})(w) = \pi^*(v)(\boldsymbol{X})(w)$ and therefore $\vartheta = \pi^*(v)$ or in local coordinates

$$\vartheta = v_i\, dx^i, \tag{B.53}$$

Taking then the differential of the Liouville 1-form $\vartheta$, the resulting 2-form is obviously closed and nondegenerate,

$$\omega = d\vartheta = dv_i \wedge dx^i \tag{B.54}$$

and therefore symplectic. In conclusion, the cotangent bundle of every manifold carries naturally a canonical (soon to be defined) symplectic structure expressed directly in local coordinates.

Alhough the more popular symplectic forms were defined here first, presymplectic ones actually precede them, relaxing the nondegeneracy restriction.

**Definition B.16.** A closed 2-form $B$ on a manifold $M$ is called a *presymplectic form*, and $M$ equipped with $B$ is called a presymplectic manifold.

**Remark B.17.** Note that without the nondegeneracy condition presymplectic manifolds need not to be even-dimensional. However, using the same argument as in the proof of Proposition B.12, we conclude that the rank of a presymplectic form (which may not be constant) is even.

A natural question raised, when one sees the many applications of symplectic manifolds especially through Hamiltonian mechanics is what do they have in common? More precisely, is there a way of classifying them? Locally, the answer came from Darboux in 1882 proving that every symplectic manifold locally looks like $\mathbb{R}^{2n}$ (which will be visited as a whole in the next appendix), while an extension for presymplectic manifolds also followed. The theorem bearing his name is stated below for the generalized version of presymplectic forms, treating symplectic ones as a special case. Proofs of this theorem for the original symplectic case can be found for example in [3], p. 175, or [69], p. 571, both due to Moser and Weinstein. A less elegant, but less

demanding, too, proof is outlined in [69], p. 594 (and also fully given in the previous edition of [69], still available online). For a proof of the presymplectic version, which is based on the symplectic case, see [3], p. 371 or [74], p. 140.

**Theorem B.18** (**Darboux**). Let $M$ be a $(2n+k)$-dimensional presymplectic manifold equipped with a presymplectic form $B$ of constant rank $2n$. In the neighborhood of any point $x$ of $M$ there exist local coordinates $(x^1, \dots, x^{2n}, z^1, \dots, z^k)$, called *canonical*, such that

$$B = dx^{n+i} \wedge dx^i \,. \tag{B.55}$$

Moreover, every presymplectic manifold locally splits in a submanifold $S$ and a *leaf* $N$ that intersect transversally at $x$ (when $k = 0$, $B$ is symplectic, $N$ are zero-dimensional and we recover the original Darboux' theorem for symplectic manifolds). The double union of all the tangent spaces to $N$ for every $x$ and every leaf, is spanned by the vector fields that belong to the kernel of the map $B^\flat$, which repeating (B.47) is

$$\ker B^\flat = \{\mathbf{v} : B(\mathbf{v}, \mathbf{w}) = 0, \forall\, \mathbf{w}\} = \left\{\mathbf{v} : B_{ij} v^j = 0\right\}. \tag{B.56}$$

The space $E = \ker B^\flat$ has the structure of a vector bundle of dimension $2n+2k$ and is known as the *characteristic bundle* of $B$. When the rank of $B$ is constant, it is in fact a subbundle of the tangent bundle $TM$, the latter being the disjoint union of the tangent spaces $T_xM$ for every point $x$ of $M$. For the description of vector bundles in general see [69] and for (B.56) in particular we refer to [3].

**Remark B.19.** In a three-dimensional manifold $M$ every 2-form $B$ has either zero rank (being trivial) or two. If it is of rank two everywhere then its kernel $E_x$ at any point $x$ (i.e. the fibers of $E$) is one-dimensional, in which case we speak of a characteristic line bundle $E$. Moreover, if $M$ is orientable then mapping $B$ under the Hodge star operation to a vector field $\boldsymbol{B} = *B$ via (B.39), the condition $B_{ij} v^j = 0$ can be equivalently expressed as $\epsilon_{ijk} v^j B^k = 0$. Thus, the vector fields that belong to $E$ can be determined in this case as the ones parallel to $\boldsymbol{B}$.

Two symplectic structures of interest in this thesis related to a presymplectic form are given by the next two propositions. First of all, observe that if $M$ is a symplectic manifold equipped with a symplectic form $\omega$ and $B$ is any presymplectic form defined on $M$, then $\widetilde{\omega} = \omega + B$ is also symplectic. Let us look into this for the special case of cotangent bundles where the following result holds [74].

**Proposition B.20.** Let $M$ be a presymplectic manifold equipped with a presymplectic form $B$. Then there is a symplectic structure $\widetilde{\omega}$ on the cotangent bundle $T^*M$ and an embedding $g : M \longrightarrow T^*M$, such that $g^*(\widetilde{\omega}) = B$.

*Proof.* Let $\omega$ be the canonical symplectic form (B.54) on $T^*M$ and $\pi$ the natural projection from $T^*M$ to $M$, both as constructed above, and consider $B = B_{ij}(x)dx^i \wedge dx^j$ in local coordinates. Define the 2-form

$$\widetilde{\omega} = \omega + \pi^*(B) \tag{B.57}$$

Then obviously $\widetilde{\omega} = dv_i \wedge dx^i + B_{ij}\, dx^i \wedge dx^j$ is closed and nondegenerate and hence symplectic. Let $g$ be the embedding defined as $g(x) = (x, 0_x)$, where $0_x$ is the zero covector on $M$, and $\mathbf{w} = w^i\partial_{x^i}$ and $\mathbf{r} = r^i\partial_{x^i}$ two vector fields on $M$. From (B.35) then we can easily see that $g^*(\omega)\,(\mathbf{w}, \mathbf{r}) = \omega\,(g_*(\mathbf{w}), g_*(\mathbf{r})) = dv_i \wedge dx^i\,\left(w^j\partial_{x^j}, r^k\partial_{x^k}\right) = r^i w^j dv_i(\partial_{x^j}) = 0$, and consequently we have $g^*(\widetilde{\omega}) = g^*(\pi^*(B)) = (\pi \circ g)^*(B) = B$. □

A similar construction, given by the next proposition, yields a lower-dimensional symplectic manifold than the cotangent bundle. Here we follow (a simplified version of) the proof given by Marle [81]; another one was given by Gotay [37]. For this case, canonical transformations for presymplectic systems were also developed by Carinena et al. in [19].

**Proposition B.21.** Let $M$ be a $(2n+k)$-dimensional presymplectic manifold equipped with a presymplectic form $B$ of constant rank $2n$. Then there is a symplectic structure $\widetilde{\omega}$ on the $(2n+2k)$-dimensional dual bundle $E^*$ of the characteristic bundle $E = \ker B^\flat$ and an embedding $g : M \longrightarrow E^*$, such that $g^*(\widetilde{\omega}) = B$.

*Proof.* A preliminary needed first is the notion of the *generalized Liouville 1-form* $\vartheta_\lambda$ associated to a projection $\lambda : TM \longrightarrow E$. The 1-form $\vartheta_\lambda$ is defined on $E^*$ in terms of the Liouville 1-form $\vartheta$ by the relation

$$\vartheta_\lambda = ({}^t\lambda)^*(\vartheta) \tag{B.58}$$

where ${}^t\lambda : E^* \longrightarrow T^*M$ is the transpose map of $\lambda$. Let $\widetilde{\pi} : T^*M \longrightarrow M$ and $\pi : E^* \longrightarrow M$ denote natural projections. For our needs it suffices to restrict to the simple case, where $\lambda$ is also the natural projection. Note then that $\widetilde{\pi} \circ {}^t\lambda = \pi$. Recalling from Example B.15 that $\vartheta = \widetilde{\pi}^*(v)$, we have $\vartheta_\lambda = ({}^t\lambda)^*\widetilde{\pi}^*(v) = (\widetilde{\pi} \circ {}^t\lambda)^*(v) = \pi^*(v)$, where $v$ is a 1-form on $M$.

Now, according to Darboux' theorem, at any point of $M$ we can find canonical coordinates $(x^1, \ldots, x^{2n}, z^1, \ldots, z^k)$ in terms of which $B = dx^{n+i} \wedge dx^i$. Local coordinates in $E^*$ are in

this case $(x^1, \ldots, x^{2n}, z^1, \ldots, z^k, v_{2n+1}, \ldots, v_{2n+k})$ and thus the generalized Liouville 1-form is $\vartheta_\lambda = v_{2n+j}\, dz^j$. Define then the 2-form

$$\widetilde{\omega} = \omega_\lambda + \pi^*(B) \tag{B.59}$$

where $\omega_\lambda = d\vartheta_\lambda$, i.e. $\widetilde{\omega} = dv_{2n+j} \wedge dz^j + dx^{n+i} \wedge dx^i$. Obviously $\widetilde{\omega}$ is closed and nondegenerate, therefore symplectic. Let $g$ be the embedding defined as the map of $M$ to the zero-section of $E^*$. Similarly as in the previous proof, $g^*(\omega_\lambda) = 0$ and hence $g^*(\widetilde{\omega}) = B$. □

Presymplectic forms are also found in the literature as *magnetic terms*, revealing their close connection with the structure of the magnetic field. In this context, they often arise in the cotangent bundle reduction process [83].

For further study, see [99] for Souriau's fundamental ideas and contributions on the connection of presymplectic manifolds and mechanics among other things.

# Appendix C

# Hamiltonian Systems of ODEs

The equivalence of Newtonian and Hamiltonian Mechanics is reflected if not based on the alternative way of writing Newton's second law as a first-order system

$$\begin{aligned} \frac{dx^i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial x^i} \end{aligned} \tag{C.1}$$

widely known as Hamilton's equations. The coordinates $\left(x^i, p_i\right)$ are called canonical variables and $H = H(x^i, p_i)$ is called the Hamiltonian function. However, other problems, in fact whole branches of Physics are characterized by a Hamiltonian structure that may not be feasible in the above fashion. The charged particle motion as well as magnetic field lines are full of Hamiltonian aspects that cannot always be expressed through formalism (C.1). In this section, we give a brief review of a generalized notion of Hamiltonian systems, strictly for ordinary differential equations as the equations of interest in this thesis. Actually, our main concern here is to escape from the demanding reliance on canonical coordinates and the related canonical transformations, too. In order to do so, we focus on the key concept of Poisson structures, rather than symplectic ones, avoiding almost entirely the sophisticated language of differential forms that is often adopted in most textbooks of Classical Mechanics. An exception to this rule is [89], which we refer to for further reading, as well as [57] where symplectic matrices are employed rather than forms.

Let us start by introducing first some geometrical considerations. Instead of separating the variables of a problem in "positions" and "momenta" as in (C.1), we group them all together denoted as $w = (w^1, \ldots, w^m)$. For finite-dimensional problems considered here, $w$ may be regarded as a point in $\mathbb{R}^m$, or, in general, in an $m$-dimensional differentiable manifold $M$. The dependent variables then of a first-order system of $m$ ordinary differential equations (often called

dynamical system) are the local coordinates on $M$, while the independent variable $t$ lies in some interval of $\mathbb{R}$.

**Definition C.1.** A system of first-order ordinary differential equations is called *Hamiltonian* if it can be expressed as

$$\frac{dw}{dt} = J(w)\nabla H(w), \tag{C.2}$$

where $w = (w^1, \ldots, w^m)$ are the dependent variables, $t$ the independent one, $H$ is a smooth, real function of $w$, $\nabla$ the gradient with respect to $w$ ($dw/dt$ and $\nabla H$ written as column matrices), and $J$ an $m \times m$ matrix of functions of $w$ that satisfies the following two properties:

1. $J^{ij} = -J^{ji}$ (skew-symmetry)
2. $J^{il}\dfrac{\partial J^{jk}}{\partial w^l} + J^{kl}\dfrac{\partial J^{ij}}{\partial w^l} + J^{jl}\dfrac{\partial J^{ki}}{\partial w^l} = 0$ (Jacobi identity)

The function $H$ is called then the *Hamiltonian function* of the system and $J$ the *Poisson structure matrix*.

The main structure behind equations (C.2) dwells in the matrix $J$ and its properties, giving rise to a Poisson bracket. In order to understand what lies beneath the geometry of Hamiltonian systems, we need first to define the Poisson structure.

**Definition C.2.** Let $M$ be a differentiable manifold and $C^\infty(M)$ the space of smooth, real functions on $M$. A bilinear map $\{\ ,\ \} : C^\infty(M) \times C^\infty(M) \longrightarrow C^\infty(M)$ that satisfies the following properties:

1. $\{F, G\} = -\{G, F\}$ (skew-symmetry)
2. $\{\{F, G\}, H\} + \{\{H, F\}, G\} + \{\{G, H\}, F\} = 0$ (Jacobi identity)
3. $\{F\,G, H\} = \{F, H\}\,G + F\,\{G, H\}$ (Leibniz' rule)

for every $F, G, H \in C^\infty(M)$ and $a, b \in \mathbb{R}$ is called *Poisson bracket*, while $M$ equipped with a Poisson bracket is called a Poisson manifold.

In the above form, Poisson structures were first introduced by Lichnerowicz, but their theory can be traced back to Lie himself. Although this definition may sound abstract to an application-oriented reader, the following proposition gives us a concrete expression of a Poisson bracket in local coordinates.

**Proposition C.3.** Let $M$ be an $m$-dimensional differentiable manifold. An $m \times m$ matrix $J(w)$ of functions of $w \in M$ defines a Poisson bracket on $M$ according to

$$\{F, G\} = (\nabla F)^T J \nabla G = J^{ij}(w) \frac{\partial F}{\partial w^i} \frac{\partial G}{\partial w^j} \tag{C.8}$$

for every pair of smooth functions $F$ and $G$ of $w$ if and only if it is skew-symmetric and satisfies the Jacobi identity.

*Proof.* Bilinearity and Leibniz' rule are automatically satisfied for the bracket (C.8), due to the gradient operator. The equivalence of skew-symmetry between the matrix $J$ and the bracket is also apparent, since $\{G, F\} = (\nabla G)^T J \nabla F = \left((\nabla F)^T J^T \nabla G\right)^T = (\nabla F)^T J^T \nabla G$ and therefore $\{F, G\} = -\{G, F\}$ if and only if $J = -J^T$. Finally, for the equivalence of the Jacobi identity we start off for example from the first term

$$\begin{aligned}\{\{F, G\}, H\} &= J^{lk} \frac{\partial}{\partial w^l} \left( J^{ij} \frac{\partial F}{\partial w^i} \frac{\partial G}{\partial w^j} \right) \frac{\partial H}{\partial w^k} \\ &= J^{lk} \frac{\partial J^{ij}}{\partial w^l} \frac{\partial F}{\partial w^i} \frac{\partial G}{\partial w^j} \frac{\partial H}{\partial w^k} + J^{lk} J^{ij} \frac{\partial^2 F}{\partial w^l \partial w^i} \frac{\partial G}{\partial w^j} \frac{\partial H}{\partial w^k} + J^{lk} J^{ij} \frac{\partial F}{\partial w^i} \frac{\partial^2 G}{\partial x^l \partial w^j} \frac{\partial H}{\partial w^k}.\end{aligned}$$

Likewise we also obtain the relations

$$\begin{aligned}\{\{H, F\}, G\} &= J^{lj} \frac{\partial J^{ki}}{\partial w^l} \frac{\partial H}{\partial w^k} \frac{\partial F}{\partial w^i} \frac{\partial G}{\partial w^j} + J^{lj} J^{ki} \frac{\partial^2 H}{\partial w^l \partial w^k} \frac{\partial F}{\partial w^i} \frac{\partial G}{\partial w^j} + J^{lj} J^{ki} \frac{\partial H}{\partial w^k} \frac{\partial^2 F}{\partial w^l \partial w^i} \frac{\partial G}{\partial w^j}, \\ \{\{G, H\}, F\} &= J^{li} \frac{\partial J^{jk}}{\partial w^l} \frac{\partial G}{\partial w^j} \frac{\partial H}{\partial w^k} \frac{\partial F}{\partial w^i} + J^{li} J^{jk} \frac{\partial^2 G}{\partial w^l \partial w^j} \frac{\partial H}{\partial w^k} \frac{\partial F}{\partial w^i} + J^{li} J^{jk} \frac{\partial G}{\partial w^j} \frac{\partial^2 H}{\partial w^l \partial w^k} \frac{\partial F}{\partial w^i},\end{aligned}$$

and we suitably interchange the indices, as for example

$$J^{lj} J^{ki} \frac{\partial H}{\partial w^k} \frac{\partial^2 F}{\partial w^l \partial w^i} \frac{\partial G}{\partial w^j} = J^{ij} J^{kl} \frac{\partial H}{\partial w^k} \frac{\partial^2 F}{\partial w^i \partial w^l} \frac{\partial G}{\partial w^j}.$$

Thus, if we add the above three equations by parts, using the skew-symmetry of $J$ the terms with second-order derivatives cancel each other, and so

$$\{\{F, G\}, H\} + \{\{H, F\}, G\} + \{\{G, H\}, F\} = \left( J^{li} \frac{\partial J^{jk}}{\partial w^l} + J^{lk} \frac{\partial J^{ij}}{\partial w^l} + J^{lj} \frac{\partial J^{ki}}{\partial w^l} \right) \frac{\partial F}{\partial w^i} \frac{\partial G}{\partial w^j} \frac{\partial H}{\partial w^k}$$

As we can see the Jacobi identity of the Poisson bracket (C.8) is satisfied if and only if the Jacobi identity of the matrix $J$ is. □

Consequently, the above proposition bridges Definitions C.1 and C.2, stating that the existence of the matrix $J$, playing a crucial role for a Hamiltonian system, and the Poisson bracket are equivalent.

**Corollary C.4.** The evolution of a Hamiltonian system (C.2) takes place in a Poisson manifold with Poisson bracket (C.8).

From (C.8) we also deduce that the elements of the structure matrix $J$ can be defined from the Poisson bracket as

$$J^{ij}(w) = \left\{w^i, w^j\right\}. \tag{C.9}$$

In terms of the Poisson bracket, the equations of the Hamiltonian system can also be written as

$$\frac{dw}{dt} = \{w, H\}$$

**Remark C.5.** We should comment here that (C.9) is of great importance for the presentation of the charged particle motion in chapter 1. For when we make a change of variables in a Hamiltonian system and want to recover the Poisson matrix in the new variables, we don't have to blindly search from the beginning. We only need to determine the brackets of the new variables using the Poisson bracket in terms of the old ones. So, this methodology is often employed in order to express the Poisson matrix and consequently the equations of motion for charged particles in different (velocity) coordinate systems, i.e. conjucate momenta, physical velocities, parallel and perpendicular velocities, etc. In the Hamiltonian literature the question arises of whether a transformation of variables retains the Poisson structure, in which case is called a Poisson map (or, in the case of a symplectic structure discussed right below, a canonical transformation), or not. For our purpose, however, a negative answer to this question, although undesirable, does not prevent the outcome. In fact, none of the transformations in chapter 1 are canonical, and the terminology "pseudo-canonical" to describe them in [4] won't be necessary nor adopted here.

Finally, another useful relation, often employed in chapter 1 in the spirit decribed in the previous remark, concerns the Poisson bracket of two functions $F$ and $H \circ G = H(G)$ on $M$, where one of them is functionally dependent of a third function. Then,

$$\{F, H \circ G\} = J^{ij}\frac{\partial F}{\partial w^i}\frac{\partial (H \circ G)}{\partial w^j} = J^{ij}\frac{\partial F}{\partial w^i}H'(G)\frac{\partial G}{\partial w^j} = H'(G)\{F, G\} \tag{C.10}$$

Now, that we have established the local expression of the Poisson bracket, we may ask how it is related to the notion of a symplectic form traditionally used to describe Hamiltonian systems. If we compare the properties of a symplectic matrix given in Proposition B.14, with the properties of a Poisson matrix stated in Definition C.1, we note two things: *i*) both matrices satisfy the Jacobi identity, which, however, is expressed quite differently for each one: In terms of the (covariant) elements of the symplectic matrix it results in a system of linear PDEs, but for the (contravariant) ones of the Poisson matrix leads to a system of non-linear PDEs, and *ii*) the symplectic matrix is additionally invertible, while the Poisson matrix is not necessarily. Both

these points are better clarified in light of the following proposition. But first, we note that the symplectic matrix being inverse was crucial for proving Proposition B.12, which in fact leads us to the next remark.

**Remark C.6.** The dimension of a Poisson manifold and therefore of a related Hamiltonian system (C.2) is not necessarily an even number, as opposed to that of a symplectic one. Still, based on the same argument used in the proof of Proposition B.12, the rank of a Poisson matrix (which may not be constant) is even.

**Proposition C.7.** A Poisson manifold is symplectic if and only if the Poisson structure matrix $J$ is invertible. In this case $\Omega = J^{-1}$, where $\Omega$ is the symplectic structure matrix.

*Proof.* On the grounds of Propositions B.14 and C.3, we need only to compare the matrices $J$ and $\Omega$. The equivalence of skew-symmetry can easily be seen: $J = -J^T \Leftrightarrow J^{-1} = \left(-J^T\right)^{-1} \Leftrightarrow \Omega = -\left(J^{-1}\right)^T = -\Omega^T$. For the Jacobi identity, first we recall that

$$J\Omega = I \Rightarrow J^{lm}\omega_{mj} = \delta^l_j \Rightarrow \frac{\partial J^{lm}}{\partial w^k}\,\omega_{mj} + J^{lm}\frac{\partial \omega_{mj}}{\partial w^k} = 0 \Rightarrow \omega_{il}\frac{\partial J^{lm}}{\partial w^k}\,\omega_{mj} + \omega_{il}J^{lm}\frac{\partial \omega_{mj}}{\partial w^k} = 0$$
$$\Rightarrow \omega_{il}\frac{\partial J^{lm}}{\partial w^k}\,\omega_{mj} + \frac{\partial \omega_{ij}}{\partial w^k} = 0 \Rightarrow \Omega\,\frac{\partial J}{\partial w^k}\,\Omega + \frac{\partial \Omega}{\partial w^k} = 0 \Rightarrow \frac{\partial \Omega}{\partial w^k} = -\Omega\,\frac{\partial J}{\partial w^k}\,\Omega$$

Therefore starting with the Jacobi identity in terms of $\Omega$, we substitute the above relation on the left-hand side and use the skew-symmetry of $\Omega$,

$$\begin{aligned}\frac{\partial \omega_{ij}}{\partial w^k} + \frac{\partial \omega_{jk}}{\partial w^i} + \frac{\partial \omega_{ki}}{\partial w^j} &= -\,\omega_{il}\frac{\partial J^{lm}}{\partial w^k}\,\omega_{mj} - \omega_{jl}\frac{\partial J^{lm}}{\partial w^i}\,\omega_{mk} - \omega_{kl}\frac{\partial J^{lm}}{\partial w^j}\,\omega_{mi} \\ &= \omega_{il}\omega_{jm}\frac{\partial J^{lm}}{\partial w^k} + \omega_{jl}\omega_{km}\frac{\partial J^{lm}}{\partial w^i} + \omega_{kl}\omega_{im}\frac{\partial J^{lm}}{\partial w^j}\,.\end{aligned}$$

Multiplying successively by $J^{ai}$, $J^{bj}$ and $J^{ck}$ (and note that since $J$ is invertible we can always go backwards), we have

$$\begin{aligned}&\frac{\partial \omega_{ij}}{\partial w^k} + \frac{\partial \omega_{jk}}{\partial w^i} + \frac{\partial \omega_{ki}}{\partial w^j} = 0 \Leftrightarrow \delta^a_l\delta^b_m J^{ck}\frac{\partial J^{lm}}{\partial w^k} + \delta^b_l\delta^c_m J^{ai}\frac{\partial J^{lm}}{\partial w^i} + \delta^a_m\delta^c_l J^{bj}\frac{\partial J^{lm}}{\partial w^j} = 0 \Leftrightarrow \\ &J^{ck}\frac{\partial J^{ab}}{\partial w^k} + J^{ai}\frac{\partial J^{bc}}{\partial w^i} + J^{bj}\frac{\partial J^{ca}}{\partial w^j} = 0 \Leftrightarrow J^{al}\frac{\partial J^{bc}}{\partial w^l} + J^{cl}\frac{\partial J^{ab}}{\partial w^l} + J^{bl}\frac{\partial J^{ca}}{\partial w^l} = 0,\end{aligned}$$

i.e. the two properties are equivalent. Thus, if a Poisson manifold is symplectic, then the Poisson structure matrix is related to the symplectic one through $J = \Omega^{-1}$ and therefore is invertible, while if $J$ is invertible then its inverse $\Omega = J^{-1}$ defines a symplectic structure. □

**Remark C.8.** More geometrical insight to Poisson manifolds shows that the Poisson bracket actually gives rise to a 2-vector field (see (B.31)), expressed in local coordinates as

$$\Pi = J^{ij}\frac{\partial}{\partial w^i} \wedge \frac{\partial}{\partial w^j} \tag{C.11}$$

as opposed to the 2-form $\omega$ related to the symplectic structure. And as with $\omega^\flat$ now the map $\Pi^\#$ is introduced, defined as the vector field $\Pi^\#(a) = \Pi(\ ,a)$, meaning $\Pi^\#(a) = J^{ij}a_j\partial_{w^i}$, for every 1-form $a$. This geometrical picture of the Poisson bracket is often considered in many textbooks of Classical Mechanics, such as [74, 83].

**The canonical case.** Let us now revisit the case of canonical variables to recover standard expressions. For system (C.1), the space considered is $\mathbb{R}^{2n}$, where $w = \left(x^i, p_i\right)$, and the Poisson matrix is the $2n \times 2n$ constant matrix

$$J = \begin{pmatrix} O & I \\ -I & O \end{pmatrix}, \tag{C.12}$$

where $O$ is the $n \times n$ zero matrix and $I$ is the $n \times n$ identity matrix. Note that $J$ in (C.12) trivially satisfies the Jacobi identity, since it is constant, and that its determinant is $|J| = 1 \neq 0$. Therefore, a canonical symplectic form is defined as $\omega = dp_i \wedge dx^i$, while the canonical Poisson bracket is

$$\{F, G\} = \frac{\partial F}{\partial x^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial x^i} \tag{C.13}$$

for any two differentiable functions $F$ and $G$ of $w$. Finally, the following relations also hold, $\left\{x^i, x^j\right\} = \{p_i, p_j\} = 0$ and $\left\{x^i, p_j\right\} = \delta^i_j$. As we saw in the previous appendix, the canonical form is naturally induced to the cotangent bundle of every differentiable manifold, $\mathbb{R}^{2n}$ being just the case for $\mathbb{R}^n$, from the exterior derivative of the Liouville 1-form $p_i dx^i$.

In spite of the generalization of the canonical case, necessary to include a whole variety of Hamiltonian systems, its significance is by no means diminished. The reason is that locally every Poisson structure can be expressed as in (C.13). This is the analogue of Darboux' theorem for Poisson manifolds generalized by Weinstein with his famous splitting theorem, stated below for consistency, without proof though, as it is beyond the scope of this thesis. For the constant rank case presented here, a nice proof is given in [89] being the Poisson counterpart of the one outlined in [69], p. 594 for the symplectic case, that we mentioned earlier in the previous appendix.

**Theorem C.9.** Let $M$ be a $(2n+k)$-dimensional Poisson manifold equipped with a Poisson bracket of constant rank $2n$. In the neighborhood of any point $w$ of $M$ there exist local canonical coordinates $(w^1, \ldots, w^{2n}, z^1, \ldots, z^k)$, such that

$$\{F, G\} = \frac{\partial F}{\partial w^i}\frac{\partial G}{\partial w^{n+i}} - \frac{\partial F}{\partial w^{n+i}}\frac{\partial G}{\partial w^i}.$$

Furthermore, Weinstein's theorem describes how every Poisson manifold locally splits in a trivial Poisson submanifold $N$ at $w$ and a *symplectic leaf* $S$ passing through $w$ that intersect transversally at this point. The double union of all the tangent spaces to $S$ for every $w$ and every symplectic leaf, is spanned by the vector fields that belong to the image of the map $\Pi^{\#}$ (see Remark C.8),

$$\operatorname{Im}\Pi^{\#} = \left\{\mathbf{v} = \Pi^{\#}(a) : a \in T^*M\right\}$$

The importance of this space lies on the fact that it is in fact isomorphic to the span of all the *Hamiltonian vector fields* on $M$,

$$\operatorname{Im}\Pi^{\#} = \left\{\mathbf{v} = \Pi^{\#}(dF) : F \in C^\infty(M)\right\} = \left\{\mathbf{v} : v^i = J^{ij}F_{w^j}, F \in C^\infty(M)\right\}$$

where put more simply $\Pi^{\#}(dF) = \{\ ,F\}$.

**The 2-dimensional case.** Let $M$ be a two-dimensional manifold. One could argue that every skew-symmetric matrix defined on $M$ is a Poisson matrix, showing that the Jacobi identity is trivially satisfied. Just as that every 2-form on $M$ is a presymplectic form, recalling that the corresponding closure condition is trivial, too. However, in both cases, since the rank of either the Poisson or the presymplectic structure, respectively, is even, besides the trivial case of zero rank, we are only left with the case of maximal rank 2, resulting though to a symplectic structure.

Therefore, in continuation of Example B.13 we consider here the case of a two-dimensional manifold in $\mathbb{R}^3$ equipped with a metric, whose determinant is $g$, induced by the Euclidean metric. As previously shown, if $(w^1, w^2)$ are local coordinates, we can always define on $M$ the symplectic form $\omega = \sqrt{g}\, dw^2 \wedge dw^1$, which gives rise to the Poisson matrix

$$J = \frac{1}{\sqrt{g}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{C.14}$$

**The non-autonomous case.** As we can see, Definition C.1 refers to autonomous systems only. However, it can still be valid, if the Hamiltonian function depends also on $t$, with no discriminations whatsoever. A proper treatment though that retains all the geometrical background is to extend the $m$-dimensional non-autonomous problem to an $(m+2)$-dimensional autonomous Hamiltonian system.

The way to do so is to consider the time $t$ and the Hamiltonian $H$ as two extra dependent variables canonically conjugate to each other. More specifically, as described in [102], the value

$h$ of the Hamiltonian function is employed to be distinguished from the function $H(w,t)$ itself. In this way, we enlarge the original $m$-dimensional space $M$ by $w^{m+1} = t$ and $w^{m+2} = -h$ to an $(m+2)$-dimensional space denoted as $\widetilde{M}$, and system (C.2) by the trivial equations

$$\begin{aligned} \frac{dw^{m+1}}{dt} &= 1 \\ \frac{dw^{m+2}}{dt} &= -\frac{\partial H}{\partial t} \end{aligned} \tag{C.15}$$

to a Hamiltonian system in $\widetilde{M}$, respectively. The extended Hamiltonian on $\widetilde{M}$ can then be defined as the function

$$\widetilde{H}(w^1, \ldots, w^{m+2}) = H(w^1, \ldots, w^{m+1}) + w^{m+2} \tag{C.16}$$

The new Hamiltonian is now autonomous and the motion of the system is realized on the hypersurface $\widetilde{H} = 0$. Note that in terms of $\widetilde{H}$, the two additional equations (C.15) are in canonical form. Thus, if $J$ is the Poisson matrix of the non-autonomous system,

$$\widetilde{J} = \begin{pmatrix} J & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

is the extended Poisson matrix defined on $\widetilde{M}$ and correspondingly the extended Poisson bracket for any two smooth functions $F$ and $G$ on $\widetilde{M}$ is

$$\widetilde{\{F,G\}} = \{F,G\} - \frac{\partial F}{\partial t}\frac{\partial G}{\partial h} + \frac{\partial F}{\partial h}\frac{\partial G}{\partial t}\,. \tag{C.17}$$

Likewise, in case of a symplectic structure $\omega$ the following extended symplectic form can be considered

$$\widetilde{\omega} = \omega - dh \wedge dt \tag{C.18}$$

See [74], p. 329 for the symplectic case in the cotangent bundle. It is also worth noting that for an autonomous system much of the above considerations can be used in the opposite direction, meaning to reduce the problem to a non-autonomous system. See, for example, [3], p. 391 the so-called Hamiltonian flow box theorem, where the treatment of time and (time-independent) Hamiltonian function as canonical variables is employed not for enlarging the system but ultimately instead for reducing it.

# Bibliography


[1] **Arnold V. I.** (1978) *Mathematical Methods of Classical Mechanics*, Graduate Texts in Mathematics, Springer-Verlag.

[2] **Abdullaev S.** (2006) *Construction of Mappings for Hamiltonian Systems and Their Applications*, Springer.

[3] **Abraham R., Marsden J. E.** (1987) *Foundations of Mechanics*, 2nd edition, Addison-Wesley Publishing Company, Inc..

[4] **Balescu R.** (1988) *Transport Processes in Plasmas* vol. I : *Classical Transport*, North-Holland.

[5] **Balescu R.** (1988) *Transport Processes in Plasmas* vol. II : *Neoclassical Transport*, North-Holland.

[6] **Barut A. O.** (1980) *Electrodynamics and Classical Theory of Fields and Particles*, Dover Publications Inc..

[7] **Baumann G.** (2000) *Symmetry Analysis of Differential Equations with Mathematica*, Springer-Verlag.

[8] **Bellan P. M.** (2006) *Fundamentals of Plasma Physics*, Cambridge Univeristy Press.

[9] **Βλάχος Λ.** (2000) *Φυσική Πλάσματος, η τέταρτη Κατάσταση της Ύλης*, Εκδόσεις Τζιόλα, Θεσσαλονίκη.

[10] **Bluman G. W., Kumei S.** (1989) *Symmetries and Differential Equations*, Applied Mathematical Sciences, Springer-Verlag.

[11] **Bluman G. W., Anco S. C.** (2002) *Symmetry and Integration Methods for Differential Equations*, Applied Mathematical Sciences, Springer-Verlag.

[12] **Bluman G. W., Cheviakov A. F., Anco S. C.** (2010) *Applications of Symmetry Methods to Partial Differential Equations*, Applied Mathematical Sciences, Springer.

[13] **Bogoyavlenskij O. I.** (1998) Extended integrability and bi-Hamiltonian systems, *Commun. Math. Phys.* **196**, 19-51.

[14] **Bogoyavlenskij O. I.** (2001) Infinite symmetries of the ideal MHD equilibrium equations, *Phys. Lett.* A **291**, 256-264.

[15] **Boozer A. H.** (1983) Evaluation of the structure of ergodic fields, *Phys. Fluids* **26**, 1288-1291.

[16] **Bouquet S., Bourdier A.** (1998) Notion of integrability for time-dependent Hamiltonian systems, *Phys. Rev.* E **57**, 1273-1283.

[17] **Brizard A. J., Chan A.** (1999) Nonlinear relativistic gyrokinetic Vlasov-Maxwell equations, *Phys. Plasmas* **6**, 4548-4558.

[18] **Burby J. W.** (2015) *Chasing Hamiltonian Structure in Gyrokinetic Theory*, Ph.D. Thesis, Princeton University.

[19] **Cariñena J. F., Gomis J., Ibort L. A., Román N.** (1985) Canonical transformations theory for presymplectic systems, *J. Math. Phys.* **26**, 1961-1969.

[20] **Cary J. R., Brizard A. J.** (2009) Hamiltonian theory of guiding center motion, *Rev. Mod. Phys.* **81**, 693-738.

[21] **Cary J. R., Littlejohn R. G.** (1983) Noncanonical Hamiltonian mechanics and its application to magnetic field line flow, *Ann. Phys.* **151**, 1-34.

[22] **Cheviakov A. F.** (2004) *Symmetries and Exact Solutions of Plasma Equilibrium Equations*, Ph.D. Thesis, Queen's University, Ontario, Canada.

[23] **Cimpoiasu R.** (2014) Generalized conditional symmetries and related solutions of the Grad-Shafranov equation, *Phys. Plasmas* **21**, 042118.

[24] **Courant R., Hilbert D.** (1989) *Methods of Mathematical Physics*, Vol. I, Wiley Classics Lirary, John Wiley & Sons.

[25] **Damianou P. A., Sophocleous C.** (1999) Symmetries of Hamiltonian systems with two degrees of freedom, *J. Math. Phys.* **40**, 210-235.

[26] **Damianou P. A., Sophocleous C.** (2000) Symmetry Group Classification of Three-Dimensional Hamiltonian Systems, *Appl. Math. Lett.* **13**, 63-70.

[27] **Damianou P. A., Sophocleous C.** (2004) Classification of Noether Symmetries for Lagrangians with Three Degrees of Freedom, *Nonlinear Dynamics* **36**, 3-18.

[28] **D'haeseleer W. D., Hitchon W. N. G., Callen J. D., Shohet J. L.** (1990) *Flux Coordinates and Magnetic Field Structure*, Springer-Verlag.

[29] **Dimas S., Tsoubelis D.** (2005) SYM : A new symmetry-finding package for Mathematica, *Proceedings of the 10th International Conference in Modern Group Analysis* (editors : Ibragimov N.H., Sophocleous C., Damianou P.A.), Nicosia, Univeristy of Cyprus, 64-70.

[30] **Dimas S., Tsoubelis D.** (2006) A new Mathematica-based program for solving overdetermined systems of PDEs, *Proceedings of the 8th International Mathematica Symposium*, Avignon, France.

[31] **Dorizzi B., Grammaticos B., Ramani A., Winternitz P.** (1985) Integrable Hamiltonian systems with velocity-dependent potentials, *J. Math. Phys.* **26**, 3070-3079.

[32] **Elsässer K.** (1986) Magnetic field line flow as a Hamiltonian problem, *Plasma Phys. Control. Fusion* **28**, 1743-1752.

[33] **Elskens Y., Escande D.** (2003) *Microscopic Dynamics of Plasmas and Chaos*, Series in Plasma Physics, IOP Publishing.

[34] **Fitzpatrick R.** (1995) Helical temperature perturbations associated with tearing modes in tokamak plasmas, *Phys. Plasmas* **2**, 825-838.

[35] **Gascon F. G.** (1997) Divergence-free vector fields and integration via quadratures, *Phys. Lett.* A **225**, 269-273.

[36] **Gascon F. G., Peralta-Salas D.** (2005) Some properties of the magnetic fields generated by symmetric configurations of wires, *Physica* D **206**, 109-120.

[37] **Gotay M.** (1982) On coisotropic imbeddings of presymplectic manifolds, *Proc. Amer. Math. Soc.* **84**, 111-114.

[38] **Grad H., Rubin H.** (1958) Hydromagnetic equilibria and force-free fields, *Proc. 2nd Int. Conf. on Peaceful uses of Atomic Energy* **31**, 190-197.

[39] **Griffiths D. J.** (1999) *Introduction to Electrodynamics*, 2nd edition, Prentice-Hall, Inc..

[40] **Grebogi C., Kaufman A., Littlejohn R. G.** (1979) Hamiltonian theory of ponderomotive effects of an electromagnetic wave in a nonuniform magnetic field, *Phys. Rev. Lett.* **43**, 1668-1671.

[41] **Grebogi C., Littlejohn R. G.** (1984) Relativistic ponderomotive Hamiltonian, *Phys. Fluids* **27**, 1996-2004.

[42] **Guillemin V., Sternberg S.** (1990) *Symplectic Techniques in Physics*, Cambridge University Press.

[43] **Haas F., Goedert J.** (1999) Noether symmetries for two-dimensional charged-particle motion, *J. Phys. A: Math. Gen.* **32**, 6837-6852.

[44] **Haas F., Goedert J.** (2000) Lie symmetries for two-dimensional charged-particle motion, *J. Phys. A: Math. Gen.* **33**, 4661-4677.

[45] **Haller G., Mezić I.** (1998) Reduction of three-dimensional, volume-preserving flows with symmetry, *Nonlinearity* **11**, 319-339.

[46] **Hamamatsu K., Takizuka T., Hayashi N., Ozeki T.** (2007) Numerical simulation of electron cyclotron current drive in magnetic islands of neo-classical tearing mode, *Plasma Phys. Control. Fusion* **49**, 1955-1969.

[47] **Hazeltine R. D., Meiss J. D.** (2003) *Plasma Confinement*, 2nd edition, Dover Publications Inc..

[48] **Hietarinta J.** (1987) Direct methods for the search of the second invariant, *Phys. Rep.* **147**, 87-154.

[49] **Huang D., Zhao X., Liu Z.** (1998) Divergence-free vector field and reduction, *Phys. Lett.* A **244**, 377-382.

[50] **Hydon P. E.** (2000) *Symmetry Methods for Differential Equations : A Beginner's Guide*, Cambridge University Press.

[51] **Ibragimov N. H.** (1995) *Lie Group Analysis of Differential Equations* vol. II : *Applications in Engineering and Physical Sciences*, CRC Press, Inc..

[52] **Ibragimov N. H.** (1999) *Elementary Lie Group Analysis and Ordinary Differential Equations*, John Wiley & Sons.

[53] **Ibragimov N. H.** (2006) *Selected Works* vol. II, ALGA Publications, Blekinge Institute of Technology, Karlskrona, Sweden.

[54] **Ibragimov N. H.** (2009) *Selected Works* vol. IV, ALGA Publications, Blekinge Institute of Technology, Karlskrona, Sweden.

[55] **Ibragimov N. H.** (2008) Symmetries, Lagrangian and Conservation Laws for the Maxwell Equations, *Acta Appl. Math.* **105**, 157-188.

[56] **Isliker H., Chatziantonaki I., Tsironis C. and Vlahos L.** (2012) Electron-cyclotron wave propagation, absorption and current drive in the presence of neoclassical tearing modes, *Plasma Phys. Control. Fusion* **54**, 095005.

[57] **Ιχτιάρογλου Σ.** (2003) *Εισαγωγή στην Μηχανική Hamilton*, 2η έκδοση, Υπηρεσία Δημοσιευμάτων Α.Π.Θ.

[58] **Janaki M. S., Ghosh G.** (1987) Hamiltonian formulation of magnetic field line equations, *J. Phys. A: Math. Gen.* **20**, 3679-3685.

[59] **Jackson J. D.** (1998) *Classical Electrodynamics*, 3rd edition, John Wiley & Sons, Inc..

[60] **Kallinikos N.** (2016) Group classification of charged particle motion in stationary electromagnetic fields, *(to be submitted)*.

[61] **Kallinikos N., Isliker H., Tsironis C., Vlahos L.** (2016) Test-particle simulations of wave-particle interaction in the presence of magnetic islands, *(in preparation)*.

[62] **Kallinikos N., Isliker H., Vlahos L., Meletlidou E.** (2014) Integrable perturbed magnetic fields in toroidal geometry: An exact analytical flux surface label for large aspect ratio, *Phys. Plasmas* **21**, 064504.

[63] **Kallinikos N., Meletlidou E.** (2013) Symmetries of charged particle motion under time-independent electromagnetic fields, *J. Phys. A: Math. Theor.* **46**, 305202.

[64] **Kaltsas D., Throumoulopoulos G.** (2016) Exact solutions of the Grad-Shafranov equation via similarity reduction and applications to magnetically confined plasmas, *Phys. Lett.* A **380**, 3373.

[65] **Karney F. F.** (1986) Fokker-Planck and quasilinear codes, *Comput. Phys. Rep.* **4**, 183-244.

[66] **Kruskal M. D., Kulsrud R. M.** (1958) Equilibrium of a magnetically confined plasma in a toroid, *Phys. Fluids* **1**, 265-274.

[67] **Kruskal M. D.** (1962) Asymptotic theory of Hamiltonian and other systems with all solutions nearly periodic, *J. Math. Phys.* **20**, 806.

[68] **La Haye R. J.** (2006) Neoclassical tearing modes and their control, *Phys. Plasmas* **13**, 055501.

[69] **Lee J. M.** (2000) *Introduction to Smooth Manifolds*, 2nd edition, Graduate Texts in Mathematics, Springer.

[70] **Lehnert B.** (1964) *Dynamics of Charged Particles*, North-Holland Publishing Company.

[71] **Lewis H. R.** (1968) Motion of a time-dependent harmonic oscillator, and of a charged particle in a class of time-dependent, axially symmetric electromagnetic fields, *Phys. Rev.* **172**, 1313-1315.

[72] **Lewis H. R.** (1984) Exact invariants quadratic in the momentum for a particle in a three-dimensional electromagnetic field, *J. Math. Phys.* **25**, 1139-1143.

[73] **Lewis H. R.** (1990) Representation of magnetic fields with toroidal topology in terms of field-line invariants, *Phys. Fluids* **2**, 2551-2562.

[74] **Libermann P., Marle C.-M.** (1987) *Symplectic Geometry and Analytical Mechanics*, Mathematics and its Applications, D. Reidel Publishing Company.

[75] **Lisle I.** (1992) *Equivalence Transformations for Classes of Differential Equations*, Ph.D. Thesis, University of British Columbia, Vancouver, Canada.

[76] **Littlejohn R. G.** (1979) A guiding center Hamiltonian : A new approach, *J. Math. Phys.* **20**, 2445-2458.

[77] **Littlejohn R. G.** (1981) Hamiltonian formulation of guiding center motion, *Phys. Fluids* **24**, 1730-1749.

[78] **Littlejohn R. G.** (1983) Variational principles of guiding center motion, *J. Plasma Phys.* **29**, 111-125.

[79] **Mahomed F. M., Leach P. G. L.** (1989) Lie algebras associated with scalar second-order ordinary differential equations, *J. Math. Phys.* **30**, 2770-2777.

[80] **Mahomed F. M., Leach P. G. L.** (1990) Symmetry Lie algebras of $n$th order ordinary differential equations, *J. Math. Anal. Appl.* **151**, 80-107.

[81] **Marle C.-M.** (1983) Sous-variétés de rang constant d'une variété symplectique, *Astérisque* **107-108**, 69-86.

[82] **Marsden J. E.** (1982) A group theoretic approach to the equations of plasma physics, *Canad. Math. Bull.* **25**, 129-142.

[83] **Marsden J. E., Ratiu T. S.** (1999) *Introduction to Mechanics and Symmetry*, 2nd edition, Texts in Applied Mathematics, Springer.

[84] **Mezić I., Wiggins S.** (1994) On the integrability and perturbation of three-dimensional fluid flows with symmetry, *J. Nonlinear Sci.* **4**, 157-194.

[85] **Misguich J. H., Reuss J. D., Constantinescu D., Steinbrecher D., Vlad M., Spineanu F., Weyssow B., Balescu R.** (2003) Noble internal transport barriers and radial subdiffusion of toroidal magnetic lines, *Ann. Phys. Fr.* **28**, 1-101.

[86] **Monteiro L. H. A., Okano V., Kucinski M. Y., Caldas I. L.** (1994) Magnetic structure of toroidal helical fields in tokamaks, *Phys. Lett.* A **193**, 89-93.

[87] **Morrison P. J.** (2000) Magnetic field lines, Hamiltonian dynamics, and nontwist systems, *Phys. Plasmas* **7**, 2279-2289.

[88] **Nakahara M.** (2003) *Geometry, Topology and Physics*, 2nd edition, Graduate Student Series in Physics, Institute of Physics.

[89] **Olver P. J.** (1993) *Applications of Lie Groups to Differential Equations*, 2nd edition, Graduate Texts in Mathematics, Springer.

[90] **Ovsiannikov L. V.** (1982) *Group Analysis of Differential Equations*, Academic Press, Inc..

[91] **Ren C., Callen J. D., Gianakon T. A., Hegna C. C., Chang Z., Fredrickson E. D., McGuire K. M., Taylor G., Zarnstorff M. C.** (1998) Measuring $\Delta'$ from electron temperature fluctuations in the Tokamak fusion test reactor, *Phys. Plasmas* **5**, 450-454.

[92] **Ritter O. M.** (2000) Symmetries and invariants for some cases involving charged particles and general electromagnetic fields: a brief review, *Braz. J. Phys.* **30**, 438-454.

[93] **Rutherford P. H.** (1973) Nonlinear growth of the tearing mode, *Phys. Fluids* **16**, 1903-1908.

[94] **Sarlet W., Cantrijn F.** (1981) Generalizations of Noether's theorem in classical mechanics, *SIAM Rev.* **23**, 467-494.

[95] **Saunders D. J.** (1989) *The Geometry of Jet Bundles*, London Mathematical Society Lecture Note Series, Cambridge University Press.

[96] **Schutz B.** (1980) *Geometrical Methods of Mathematical Physics*, Cambridge University Press.

[97] **Schwarz F.** (2008) *Algorithmic Lie Theory for Solving Ordinary Differential Equations*, Pure and Applied Mathematics, Chapman & Hall/CRC.

[98] **Sophocleous C., Moyo S., Leach P. G. L., Damianou P. A.** (2000) Noether symmetries and integrals in one, two and three dimensions, *Department of Mathematics and Statistics, University of Cyprus*, TR/16/2000.

[99] **Souriau J.-M.** (1997) *Structure of Dynamical Systems : A Symplectic View of Physics*, Progress in Mathematics, Birkhäuser.

[100] **Stephani H.** (1989) *Differential Equations : Their Solution Using Symmetries*, Cambridge University Press.

[101] **Stix T. H.** (1992) *Waves In Plasmas*, AIP Press, Springer-Verlag.

[102] **Struckmeier J.** (2005) Hamiltonian dynamics on the symplectic extended phase space for autonomous and non-autonomous systems, *J. Phys. A: Math. Gen.* **38**, 1257-1278.

[103] **Swanson W.** (2003) *Plasma Waves*, 2nd edition, Series in Plasma Physics and Fluid Dynamics, IOP Publishing.

[104] **Tao X., Chan A., Brizard A. J.** (2007) Hamiltonian theory of adiabatic motion of relativistic charged particles, *Phys. Plasmas* **14**, 092107.

[105] **Tassi E., Pegoraro F., Cicogna G.** (2008) Solutions and symmetries of force-free magnetic fields, *Phys. Plasmas* **15**, 092113.

[106] **Torres del Castillo G. F.** (1995) Debye potentials and Hamiltonian for magnetic field line flow, *J. Math. Phys.* **36**, 3413-3418.

[107] **Tsamparlis M., Paliathanasis A.** (2011) Two-dimensional dynamical systems which admit Lie and Noether symmetries, *J. Phys. A: Math. Theor.* **44**, 175202.

[108] **Tsamparlis M., Paliathanasis A., Karpathopoulos L.** (2012) Autonomous three-dimensional Newtonian systems which admit Lie and Noether point symmetries, *J. Phys. A: Math. Theor.* **45**, 275201.

[109] **Turner L.** (1985) Hamiltonian representation for helically symmetric magnetic fields, *J. Math. Phys.* **26**, 991-993.

[110] **Viana R. L.** (2000) Chaotic magnetic field lines in a tokamak with resonant helical windings, *Chaos, Solitons and Fractals* **11**, 765-778.

[111] **Vilasi G.** (2001) *Hamiltonian Dynamics*, World Scientific.

[112] **Warner F. W.** (1983) *Foundations of Differentiable Manifolds and Lie Groups*, 2nd edition, Graduate Texts in Mathematics, Springer-Verlag.

[113] **Wesson J. A.** (2004) *Tokamaks*, 3rd edition, International Series of Monographs on Physics, Oxford University Press.

[114] **Weyssow B., Balescu R.** (1986) Hamiltonian theory of guiding center motion revisited, *J. Plasma Phys.* **35**, 449-471.

[115] **Weyssow B., Balescu R.** (1987) A manifestly gauge-invariant Hamiltonian theory of the oscillation center dynamics, *J. Plasma Phys.* **37**, 467-486.

[116] **Weyssow B., Balescu R.** (1988) Hamiltonian theory of the generalized oscillation center transformation, *J. Plasma Phys.* **39**, 81-102.

[117] **Westerhof E.** (1998) Hot plasma dielectric tensor, *Trans. Fus. Tech.* **33**, 139-144.

[118] **White R. B.** (2001) *The Theory of Toroidally Confined Plasmas*, 2nd edition, Imperial College Press.

[119] **White R. B., Monticello D. A., Rosenbluth M. N., Waddell B. V.** (1977) Saturation of the tearing mode, *Phys. Fluids* **20**, 800-805.

[120] **White R. B., Hazeltine R. D.** (2009) Symmetry analysis of the Grad-Shafranov equation, *Phys. Plasmas* **16**, 123101.

[121] **Zheng Y., Huang D., Liu Z.** (2003) Reduction of volume-preserving flows on an $n$-dimensional manifold, *Acta Math. Appl. Sin.* (English Series) **19**, 129-134.