HAL
open science

# Réductions hamiltoniennes en physique des plasmas autour de la gyrocinétique intrinsèque

Loïc de Guillebon

## To cite this version:

Loïc de Guillebon. Réductions hamiltoniennes en physique des plasmas autour de la gyrocinétique intrinsèque. Physique mathématique [math-ph]. Aix-Marseille Université, 2013. Français. NNT : . tel-00850139v3

**HAL Id: tel-00850139**

**https://theses.hal.science/tel-00850139v3**

Submitted on 3 Oct 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse de doctorat
Aix-Marseille Université
École doctorale : Physique et science de la matière
Mention : Physique théorique et mathématique

**Loïc de GUILLEBON de RESNES**

# Réductions hamiltoniennes en physique des plasmas autour de la gyrocinétique intrinsèque

**Hamiltonian reductions in plasma physics about intrinsic gyrokinetics**

Soutenue le 16 septembre 2013

| | | | |
|---|---|---|---|
| Jury : | Alain J. BRIZARD | Saint Michael's College - Vermont | |
| | Xavier GARBET | Commissariat à l'Energie Atomique - Cadarache | (président du jury) |
| | Robert G. LITTLEJOHN | University of California at Berkeley | |
| | Rafael de la LLAVE | Georgia Institute of Technology | (rapporteur) |
| | Philip J. MORRISON | University of Texas at Austin | |
| | Thierry PASSOT | Centre National de la Recherche Scientifique - Nice | |
| | Marco PETTINI | Aix-Marseille Université | |
| | Hong QIN | Princeton Plasma Physics Laboratory | (rapporteur) |
| | Michel VITTOT | Centre National de la Recherche Scientifique - Marseille | (directeur de thèse) |

## Abstract

Gyrokinetics is a key model for plasma micro-turbulence, commonly used for fusion plasmas or small-scale astrophysical turbulence, for instance. The model still suffers from several issues, which could imply to reconsider the equations. This thesis dissertation clarifies three of them. First, one of the coordinates caused questions, both from a physical and from a mathematical point of view; a suitable constrained coordinate is introduced, which removes the issues from the theory and explains the intrinsic structures underlying the questions. Second, the perturbative coordinate transformation for gyrokinetics was computed only at lowest orders; explicit induction relations are obtained to go arbitrary order in the expansion. Third, the introduction of the coupling between the plasma and the electromagnetic field was not completely satisfactory; using the Hamiltonian structure of the dynamics, it is implemented in a more appropriate way, with strong consequences on the gyrokinetic equations, especially about their Hamiltonian structure.

In order to address these three main points, several other results are obtained, for instance about the origin of the guiding-center adiabatic invariant, about a very efficient minimal guiding-center transformation, or about an intermediate Hamiltonian model between Vlasov-Maxwell and gyrokinetics, where the characteristics include both the slow guiding-center dynamics and the fast gyro-angle dynamics. In addition, various reduction methods are used, introduced or developed, e.g. a Lie-transform of the equations of motion, a litfing method to transfer particle reductions to the corresponding Hamiltonian field dynamics, or a truncation method related both to Dirac's theory of constraints and to a projection onto a Lie-subalgebra. Besides gyrokinetics, this is useful to clarify other Hamiltonian reductions in plasma physics, for instance for incompressible or electrostatic dynamics, for magnetohydrodynamics, or for fluid closures including moments of the Vlasov density of order two.



## Key words

- Non-linear dynamics
- Field dynamics
- Hamiltonian systems
- Dynamical reductions
- Perturbation theory
- Constrained systems
- Plasma physics
- Gyrokinetics
- Guiding-center theory
- Fluid dynamics



## Contacts

Centre de Physique Théorique
CNRS – UMR-7332
Aix Marseille Université
Campus de Luminy, Case 907
13288 Marseille cedex 9, France

## Summary

This thesis dissertation investigates Hamiltonian reductions in plasma physics, and especially it clarifies a few foundational questions about guiding-center theory and about gyrokinetics, which are respectively a reduced model for particle dynamics and for plasma dynamics in a strong magnetic field, and which constitute key ingredients to understand the dynamics of magnetic-fusion plasmas, and more generally small-scale plasma turbulence.

About guiding-center theory, it introduces the higher-order reduction by obtaining explicit induction relations to go to arbitrary order in the perturbation expansion. The available freedom in the reduction is emphasized and explored, a very efficient minimal procedure is introduced, and a choice to obtain a maximal reduction is identified, whose lowest orders agree with previous works.

On another hand, the traditional questions or difficulties associated with the presence of a gyro-gauge in guiding-center theory are clarified, especially the gauge arbitrariness, the anholonomy and the non-global existence. A global, gauge-independent formulation is introduced, which removes the difficulties and emphasizes that the questions originated from intrinsic but regular properties of the system. It makes clearer what the true intrinsic entities of the theory are, and opens new possibilities interesting for guiding-center computations.

About gyrokinetics, this document investigates its Hamiltonian structure and studies how it is induced by the derivation, mainly a lifting procedure which transfers the guiding-center transformation from particle dynamics to the Vlasov-Maxwell Hamiltonian system, but also an application of Dirac's theory of constraints. This introduces the coupling between the plasma and the electromagnetic field in a more complete way than in previous works. Especially, it emphasizes the dependence of the gyro-center transformation on all the moments of the Vlasov density.

Last, about Hamiltonian reductions in plasma physics, Dirac's method is formulated as a projection of derivative operators for the Poisson bracket, and shown to be closely associated with a special, quarter-canonical structure in the bracket. A simplified procedure in this framework is identified, with a link towards an efficient reduction method based on a Lie-subalgebra. Various applications in plasmas are developed, both for kinetic and for fluid models, such as the Vlasov-Poisson system, magnetohydrodynamics, incompressible dynamics, and fluid closures.

In the process, several complementary aspects of the concerned subjects are considered, e.g. the mathematical origin of the lowest-order magnetic moment being an adiabatic invariant, a polynomial structure in guiding-center expansions, the role of Dirac's constraints to improve some Hamiltonian systems suffering from a conditional Jacobi identity, or a formulation of gyrokinetics as a Hamiltonian perturbation of particle dynamics. In addition, various reduction tools for dynamical systems are involved or developed, especially Hamiltonian and non-Hamiltonian perturbations, Lie-transforms, constrained systems, connections on fiber bundles, projection methods, or foundational aspects of a lifting mechanism relating reductions on a coordinate space to reductions on the corresponding field dynamics. Many possible applications or extensions of the work are emphasized.

# Contents


Abstract . . . 2
Key words . . . 2
Contact . . . 2
Summary . . . 3

**General introduction** **11**
Framework and subject . . . 11
Dynamical reductions in plasma physics . . . 11
Hamiltonian reductions . . . 11
The four main models of plasma dynamics . . . 12
Gyrokinetics . . . 14
The gyrokinetic reduction . . . 15
Electrostatic gyrokinetics . . . 15
Thesis work . . . 17
Questions addressed . . . 17
Complementary question . . . 17
Organization of the document . . . 18
Comments for the reader . . . 19

**I Intrinsic approach of the guiding-center reduction** **21**

**Introduction of the episode** **23**

**1 Intrinsic minimal guiding-center reduction** **27**
Introduction . . . 27
1.1 Lie-transforming the velocity vector field . . . 29
1.2 The reduction algorithm . . . 33
1.3 The result to second order . . . 37
1.4 On non-minimal guiding-center reductions . . . 41
Conclusion . . . 45

**2 Intrinsic full guiding-center reduction** **47**
Introduction . . . 47
2.1 Coordinates, method and requirements . . . 49
2.1.1 Choice for the gyro-angle coordinate . . . 50
2.1.2 Lie transforming the Lagrangian 1-form . . . 52
2.1.3 The hierarchy of requirements . . . 55
2.2 Derivation of the reduction . . . 57
2.2.1 Preliminary transformation and initial setting . . . 58
2.2.2 Immediate orders $-1$ and 0 . . . 61
2.2.3 Turning point: Order 1 . . . 63
2.2.4 Core of the second stage: order 2 . . . 68
2.2.5 End of the second stage: order 3 . . . 73
2.2.6 Third stage: algorithmic orders 4 and higher . . . 73

2.3 Comparison with previous works . . . 78
2.3.1 Agreeing results . . . 78
2.3.2 Polynomiality in the momentum coordinates . . . 82
2.3.3 A two-fold maximal reduction . . . 83
2.3.4 Gyro-gauge independence . . . 85
2.3.5 Maximal vs. minimal reduction . . . 86
Conclusion . . . 87
Appendix: Mechanism of the reduction . . . 88
Fundamental ingredients . . . 89
Resulting procedure in three stages . . . 92

**3 Intrinsic GC anholonomy and gyro-gauge arbitrariness 95**
Introduction . . . 95
3.1 A global gauge-independent coordinate for the gyro-angle . . . 96
3.2 Intrinsic counterpart of the gauge arbitrariness . . . 99
3.3 Intrinsic counterpart of the anholonomy . . . 102
3.4 Towards a scalar intrinsic gyro-angle . . . 106
3.5 An intrinsic formalism with no covariant derivative . . . 110
Conclusion . . . 111

**Conclusion of the episode 113**
Selected bibliography . . . 115

**II From particle guiding-center to Vlasov-Maxwell gyro-center 117**

**Introduction of the episode 119**

**4 Lifting the magnetic-moment to Vlasov-Maxwell 123**
Introduction . . . 123
4.1 Framework of the present lifting . . . 124
4.2 Lifting magnetic velocity coordinates . . . 125
4.2.1 Function chain rule . . . 126
4.2.2 Jacobian . . . 126
4.2.3 Functional chain rule . . . 127
4.3 Lifting spherical velocity coordinates $\mathbf{v} = V\hat{\mathbf{v}}$ . . . 128
4.4 Inclusion of a local dependence on $\mathbf{B}$ . . . 129
4.5 Inclusion of a nonlocal dependence on $\mathbf{B}$ . . . 131
4.6 Application to the magnetic moment . . . 132
Conclusion . . . 135

**5 Lifting the guiding-center to Vlasov-Maxwell 137**
Introduction . . . 137
5.1 Specificities of the Vlasov-Maxwell system . . . 138
5.2 Lifting for a general transformation . . . 139
5.3 Interpretation of the results . . . 142
5.4 Case of the guiding-center transformation . . . 144
5.4.1 Agreements with standard gyrokinetics . . . 144
5.4.2 Difference with standard gyrokinetics . . . 144
5.4.3 Consequences for particle dynamics . . . 145
5.4.4 Consequences for the gyro-center transformation . . . 146
5.4.5 Complementary comments . . . 148
Conclusion . . . 149
Appendix: Generalization to multiple coordinate spaces . . . 150

**6 Gyrokinetics as a Hamiltonian perturbation theory** **155**
Introduction . . . 155
6.1 Hamiltonian perturbation framework . . . 156
6.2 The magnetic-moment reduction at lowest order . . . 158
6.3 Canonical transformation at first order . . . 160
6.3.1 On the reduced Hamiltonian $H_1'$ . . . 160
6.3.2 Solving the first order . . . 162
6.4 Non-canonical approach . . . 163
Conclusion . . . 164

**Conclusion of the episode** **167**
Selected bibliography . . . 169

**III Hamiltonian reductions in plasma physics** **171**

**Introduction of the episode** **173**

**7 Role of projectors in Dirac reductions** **175**
Introduction . . . 175
7.1 Constrained and tainted brackets . . . 176
7.2 Formulation of the general method . . . 178
7.2.1 Projected functional derivatives . . . 178
7.2.2 Dirac brackets . . . 179
7.3 Example 1: magnetohydrodynamics . . . 181
7.3.1 Compressible magnetohydrodynamics . . . 181
7.3.2 Incompressible magnetohydrodynamics . . . 182
7.4 Example 2: Vlasov-Maxwell equations . . . 183
7.4.1 Vlasov-Maxwell modified bracket as a Dirac bracket . . . 183
7.4.2 From Vlasov-Maxwell to Vlasov-Poisson equations . . . 185
Conclusion . . . 186

**8 Dirac projectors and quarter-canonical truncations** **187**
Introduction . . . 187
8.1 Introductory example . . . 188
8.2 The quarter-canonical structure in Dirac reductions . . . 190
8.2.1 Condition for a Dirac truncation . . . 190
8.2.2 The quarter-canonical structure . . . 191
8.2.3 Application to bracket extensions . . . 193
8.2.4 Application to fluid and kinetic incompressibility . . . 193
8.2.5 Incidence of non-quarter-canonical coordinates . . . 195
8.3 Consequences for Dirac procedure and truncation methods . . . 195
8.4 Application to the gyrokinetic reduction . . . 196
8.4.1 Framework and goals of the reduction . . . 196
8.4.2 Introductory reduction of particle dynamics . . . 198
8.4.3 Dirac reduction of plasma dynamics . . . 198
8.4.4 Complementary comments . . . 202
8.4.5 The reduction in the presence of the coupling . . . 203
Conclusion . . . 204

**9 On Hamiltonian fluid reductions from Vlasov-Maxwell** **207**
Introduction . . . 207
9.1 Principles of the reduction . . . 208
9.2 Reduction to the ideal-fluid model . . . 209
9.3 Fluid reduction at orders higher than 1 . . . 211

9.4 Fluid brackets expressed with pressure-like moments . . . 213
Conclusion . . . 213
Appendix: Verifying the Jacobi identity for truncated brackets . . . 214

**Conclusion of the episode** **217**
Selected bibliography . . . 219

**General conclusion** **221**
Positioning of the matter . . . 221
Main points of the work . . . 221
For the future . . . 222

**IV Appendix chapters** **225**

**10 Intrinsic magnetic-moment reduction** **227**
Introduction . . . 227
10.1 Building the normal form . . . 228
10.1.1 Equation for the normal form . . . 228
10.1.2 The adiabatic invariant . . . 232
10.1.3 The normal form to second order . . . 234
10.1.4 The normal form to higher order . . . 237
10.2 The reduced Poisson bracket . . . 238
10.2.1 Change of variables for the symplectic 2-form . . . 239
10.2.2 Change of variables for the Poisson bracket . . . 240
Conclusion . . . 242
App. A: Gyro-rotations and related operators . . . 243
The operators $\mathcal{D}$, $\mathcal{G}$, $\mathtt{avg}$ and $\mathtt{osc}$ . . . 243
Computational properties of $\mathcal{D}$, $\mathcal{G}$, $\mathtt{avg}$ and $\mathtt{osc}$ . . . 245
App. B: Explicit expression for the operator $\mathtt{avg}\left(\overline{\mathtt{v}}\nabla\alpha_1\right)$ . . . 246
App. C: Computation of the right-hand side of Equation (10.31) . . . 248

**11 Lifting coordinates changes for Hamiltonian fluids** **251**
Introduction . . . 251
11.1 Framework, operators and meta-operators . . . 252
11.2 Transformed Hamiltonian functional and Poisson bracket . . . 253
11.3 Functional chain-rule formula . . . 254
11.3.1 Difficulty for a bare chain rule . . . 254
11.3.2 Derivation of a chain-rule formula . . . 256
11.3.3 Interpretation: implicit function and moving point . . . 258
11.4 Transformed functional derivatives . . . 259
11.4.1 Transformed functional derivatives: unexpected form . . . 260
11.4.2 Intermediate lemma . . . 261
11.4.3 Transformed functional derivatives in the good form . . . 263
11.5 Transformed dynamics . . . 263
Conclusion . . . 264

**12 Dirac truncations and non-quarter-canonical coordinates** **265**
Introduction . . . 265
12.1 Incidence of the quarter-canonical structure . . . 266
12.1.1 Fundamental case of existence . . . 266
12.1.2 Reasons for the occurrence . . . 267
12.2 On the identification of suited coordinates . . . 269
12.2.1 Case of quarter-canonical coordinates . . . 269
12.2.2 Case of semi-canonical coordinates . . . 270

12.2.3 Towards weakened requirements for the quarter-canonical structure . . . . . 272
12.3 Investigation of an explicit generic bracket . . . . . 273
12.3.1 Simplified projector . . . . . 273
12.3.2 Relation with the quarter-canonical structure . . . . . 274
12.3.3 A projection by changing coordinates twice . . . . . 275
12.3.4 Orthogonal projector . . . . . 275
12.4 Coordinate-independent formulation . . . . . 275
12.4.1 Simplified projector . . . . . 276
12.4.2 Example of application . . . . . 276
12.4.3 Case of quarter-canonical coordinates . . . . . 277
12.4.4 When some coefficients are operators . . . . . 277
12.4.5 A few consequences . . . . . 277
Conclusion . . . . . 278

**13 Quarter-canonical effects on Dirac and subalgebra methods 279**
Introduction . . . . . 279
13.1 On the invertibility of the matrix of constraints . . . . . 280
13.1.1 Quarter-canonical matrices of constraints are invertible . . . . . 280
13.1.2 Examples: guiding-center and Euler-Poisson . . . . . 281
13.1.3 Consequence on the choice of secondary constraints . . . . . 282
13.1.4 The Euler-Poisson by constraining the electric field . . . . . 282
13.2 On the procedure for secondary constraints . . . . . 282
13.2.1 Natural $H$-constraints, but structural $J$-constraints . . . . . 283
13.2.2 Frequent equivalence . . . . . 283
13.2.3 Application to fluid incompressibility . . . . . 284
13.3 A simplified subalgebra reduction . . . . . 285
13.3.1 Bracket truncations and subalgebras . . . . . 286
13.3.2 Counterexample but proper justification . . . . . 286
13.3.3 Examples . . . . . 287
13.4 Flexibility and q-can. effects in subalgebra reductions . . . . . 288
13.4.1 Example with a single constraint . . . . . 288
13.4.2 Presence of multiple subalgebras . . . . . 288
13.4.3 Application to bracket extensions . . . . . 290
13.4.4 Quarter-canonical effects on subalgebras . . . . . 290
Conclusion . . . . . 292

**14 Hamiltonian derivation of magnetohydrodynamics 293**
Introduction . . . . . 293
14.1 Field coordinates for the total plasma . . . . . 294
14.2 Reduction to a single-fluid Poisson bracket . . . . . 297
14.3 Reduction of the Hamiltonian functional . . . . . 298
14.4 Extension to multi-fluid, relativistic, and anisotropic cases . . . . . 301
Conclusion . . . . . 302

**Appendix A: List of papers 303**

**Appendix B: Complementary paper 305**

**Bibliography 309**

**Index 317**

**List of main symbols 319**
Notational conventions . . . . . 319
List of main symbols . . . . . 320

# General introduction

## Framework and subject

### Dynamical reductions in plasma physics

Plasmas are quite elementary media. Roughly speaking, they just correspond to a gas of charged particles, ideally neither too dense nor too rarefied [7, 36, 54, 120]. But they offer a rich domain in physics, both with their many applications and with the wide range of phenomena they involve.

Fluorescent lamps, plasma display panels, flares, laser-matter interactions, particle colliders, nuclear fusion, polar aurora, solar winds, ionosphere, the Sun, stars, the interplanetary, interstellar, and intergalactic media are all concerned with plasma physics. Larmor gyration, bounce average, magnetic mirror, Debye screening, ambipolar diffusion, Alfven waves, kink instability, two-stream instability, Landau damping, dynamo effect, Hall effect, magnetic reconnection, zonal flows are all classical phenomena occurring in plasmas.

In fact, plasma physics remains intrinsically a complex domain, because a plasma is characterized by a kinetic individual behaviour with strong local electromagnetic fields, coexisting with an electroneutral collective behaviour at larger scale. It is intrinsically a multiscale medium, with several ranges, such as the ones delimited by the Debye length, the Larmor radius, the plasma frequency or the Alfven time, without mentioning the skin depth, the collision frequency, the trapping rate, the thermal and sound scales, etc. and each of these scales exists typically twice, once for ions and once for electrons.

As a result, when studying a specific question about plasmas, it is not possible and not desirable to take into account all the existing phenomena, and one has to focus on the ones that have a significant contribution to the physics of interest. This is done by building a suitable model, which discards some of the aspects of plasma physics, considered as negligible for the purpose, and retains or focuses on other aspects, considered as essential or significant. Thus, plasma physics involves quite *a lot of models*, each one corresponding to a specific description, or perspective, of the whole physics, and having its range of validity and accuracy.

The *links between models* play an important role in plasma physics, as is emphasized both by textbooks and by our whole bibliography. It may sometimes seem to be purely theoretical or mathematical, because it often focuses on algebraic or formal structures and possibly on axiomatic arguments. However, it is a way to better understand the physical content and limits of existing models, to verify their relevance, their accuracy and the confidence we can have in them, and possibly to improve them or to develop new models.

This explains the general framework of the thesis work, which is concerned with dynamical reductions in plasma physics, i.e. with the derivation of models and with the study of relations between models.

### Hamiltonian reductions

A remarkable phenomenon in plasmas is the omnipresence of *Hamiltonian models*, as appears in [104, 159] for instance, among many others[1]. Indeed, physical models are in general dissipative,

[1] A definition of Hamiltonian systems is given at Eq. (2.10). An example in finite dimension can be found in Eqs. (2.1), (2.11), and (2.12). An example for field dynamics is provided by Eqs. (4.2)-(4.4). For an introduction about

i.e. they do not conserve the physical energy, but the non-conservation of energy can be related to few coefficients, related to physical-dissipation phenomena. When these coefficients are set to zero, most of the plasma models are not only energy-conserving but also Hamiltonian. This feature is highly non-trivial, because a Hamiltonian structure is a huge requirement for a system, much stronger than just the existence of a conserved quantity, the energy.

So, the Hamiltonian character seems to play a role in the physical relevance of models in plasma physics [108], even for dissipative models, where the backbone remains a Hamiltonian structure and only a few dissipative corrections depart from a pure Hamiltonian system.

From a practical point of view also, this fact is important because the Hamiltonian structure has deep mathematical foundations and somehow constitutes a rigid structure which by itself brings about many interesting features, for instance integrability and conservation properties. It implies the presence of tools and methods that are not available in a general dynamical system. In addition, because of that, a special care was carried on Hamiltonian systems and some tools that are available in other dynamical systems have been more developed in the Hamiltonian context. So, Hamiltonian systems come along with a very attractive artillery [107, 108, 159], for instance in perturbation theory [27, 86, 88, 92], in control theory [30, 38, 39, 76, 151, 153], in the study of equilibria and their stability [70, 116, 140, 148, 149], in conservation laws and symmetries [20, 27, 43, 139], etc. This may be of great value both in plasma theory and in numerical simulations.

With regard to our general purpose, the relations between plasma models, the Hamiltonian structure offers powerful reduction tools. Not only this may be interesting to better understand the existing models or open efficient ways to derive new models but also it may be decisive in verifying the relevance of existing models, since physically important models are expected to have a *Hamiltonian backbone* (i.e. to be Hamiltonian when physical dissipation coefficients are set to zero).

Indeed, most models are initially derived by working on the equations of motion, which does not guarantee to keep them Hamiltonian, whereas the parent model is Hamiltonian (when dissipative terms are negligible), and the same is expected for the reduced model if no fake dissipation is introduced in the reduction process [108]. Usually, the preservation of the Hamiltonian structure is considered a posteriori, but it may be quite difficult to conclude about it, and it may occur that the model is not Hamiltonian and has to be corrected in order to become so. An alternative way is to work directly on the Hamiltonian structure and to use Hamiltonian reduction methods, which automatically preserve the Hamiltonian structure, so that if the relevant constraints are imposed, the resulting equations can be expected to be physically relevant. For an example, see [136, 138] for instance, or the sequence [29, 86, 122].

This explains the more restricted positioning of the thesis work, which is interested in *Hamiltonian reductions in plasma physics*. Notice that it will not be considered as a limitation, and non-Hamiltonian reductions will be worked on when useful; for instance, in Part 1 of the manuscript, it will be the case of the first and third chapters. All the same, when mentioning a plasma model, we will most often implicitly have in mind its conservative, Hamiltonian version.

### The four main models of plasma dynamics

Among the various plasma models, the thesis is primarily interested in *gyrokinetics* [17, 29, 80]. In order to better position it, let us remind that plasma theory can be considered as roughly based on four main dynamical systems [7, 36, 54, 120], each of them being Hamiltonian, either intrinsically or when physical dissipation is negligible.

In the first of them, the plasma is considered as a set of *particles*, where each particle corresponds to a six-dimensional position-momentum phase space, and particles are coupled through the electromagnetic field, which evolves through Maxwell's equations. For a Hamiltonian formulation of this dynamical system, see e.g. [5]. This model is related to PIC simulations. In theoretical studies, it is often used when considering the electromagnetic field as given (external) in order to

Hamiltonian systems, see [53, 99, 107, 109, 141] for instance.

study the behaviour of the basic components of the plasma, which are particles, each of which has negligible effects on the electromagnetic field (and also on the plasma). For instance, this model justifies the guiding-center approximation, the various drifts (grad-B drift, curvature drift, ExB drift, Baños drift, etc.), the Larmor radius and frequency, the bounce frequency, etc. It will be at the core of Part 1 of the present dissertation.

This model can be typified as a Lagrangian description of a phase-space fluid. It is not completely satisfactory, since it does not fit with the natural point of view, where particles cannot be followed individually, and what is measured is rather particle densities in a fixed space.

It is why the second model is just an Eulerianization of the first one. It replaces the collection of particles, characterized by their phase-space coordinates evolving in time, by a particle density, called *Vlasov density*, evolving on a fixed phase space. Then the dynamics is given by *Vlasov-Maxwell* equations. It is often designated as the *kinetic description* of plasma dynamics, and will be at the core of Part 2 of the present dissertation. Here it is considered in the Klimontovitch approach as a complete classical deterministic description of the plasma. In the BBGKY approach [120], it corresponds to taking the lowest-order equation. For a Hamiltonian formulation of this dynamical system, see e.g. [5].

The issue with the second model is that it is too complete in some way. Indeed, the plasma density is defined over a six-dimensional space, and in addition the elementary time step is generally very small. This makes numerical simulations for a three-dimensional configuration space far out of reach of present-day High Performance Computers.

The third model replaces the kinetic description (i.e. fluid description in phase space) by a *fluid description* in configuration space, by taking the velocity moments of the Vlasov density and truncating the resulting infinite series of equations by a closure assumption. This approximation is often justified by a collision argument, with its local thermodynamical equilibrium, and the model includes only the zeroth and first moment, i.e. the plasma densities (in configuration space) for mass and momentum (or velocity), together with the entropy density or pressure.

The corresponding model is the well-known fluid description of the plasma, with the *Euler-Maxwell* equations. It can be called *multi-fluid model* because in this model, as in the previous ones, each species (electrons, ions, etc.) is described by its own field variables (either the Vlasov phase-space density for the kinetic model, or the position-space densities in mass, momentum, pressure, etc. for the fluid model). This model is most used in numerical simulations, because fields are defined over a three dimensional grid, which is quite tractable for present-day computers. For a Hamiltonian formulation of this dynamical system, see e.g. [107]. It will be at the core of Part 3 of the present dissertation, together with the next one.

The last model, which is most used as well, both in theory and in simulations, is called *magnetohydrodynamics* (MHD). It comes because in many physical applications, at the scales of interest, the plasma can be considered as composed of mainly two species, electroneutral, non-relativistic and obeying a kind of Ohm's law. These assumptions allow for a description where the plasma is treated as a single fluid (in configuration space), i.e. only by three fields, e.g. total density, total momentum density and total entropy density. This model is thus a huge simplification compared to the Vlasov-Maxwell system, and it allows for a much simplified treatment, both for theory and numerical simulations, but it retains most aspects of the basic large-scale behaviour of a plasma and in many application, its assumptions are quite accurately satisfied [6, 46]. For a Hamiltonian formulation of this dynamical system, see e.g. [104].

All the same, there are many cases where it is not accurate enough for the problem under consideration or it does not include some physical effects, crucial for the purpose but eliminated as a result of the assumptions adopted, for instance phenomena related to the presence of two fluids or to a kinetic behaviour (e.g. the two-stream instability or the Landau effect) [7, 36, 54, 120]. Then one may have to come back to the more complete multi-fluid or Vlasov-Maxwell models. Finally, in some cases, none of these models is satisfactory and one has to build and use another model,

e.g. a variation of one of them or an intermediate model. This is the case for gyrokinetics.

### Gyrokinetics

Gyrokinetics [17, 29, 80] can be characterized as the kinetic model with elimination of the (small) Larmor scales both in time and in space (see next subsection for a definition of Larmor scales). A motivation is that Vlasov-Maxwell dynamics involves much useless information when those scales are not needed, which obfuscates the essential phenomena; in addition, retaining Larmor scales is completely out of range of present-day numerical simulations. The broad validity comes because most physical phenomena occur at rather large scale, where the much smaller Larmor scales have only averaged effects. Practical solicitations can be found in astrophysical turbulence and in nuclear fusion by magnetic confinement (see e.g. [132]) and [51]).

In *astrophysical turbulence*, the dissipation range does not occur at large scales, standard fluid models are not satisfactory for an accurate description of the dissipation phenomena, and kinetic effects must be taken into account. One has to use intermediate models between Vlasov-Maxwell and Euler-Maxwell.

One way to do so is to refine fluid models by including some kinetic effects in the closure assumption [14, 55, 75, 147, 154, 155]. For instance, this idea gave rise to *FLR-Landau fluids*, which allowed the model to better describe the mirror instability of anisotropic plasmas, and especially its small-scale behaviour [147]. However, when a more complete account of the kinetic behaviour is needed, instead of refining fluid models it may be preferable to derive a truly kinetic, yet simplified model, by eliminating (averaging) from the Vlasov-Maxwell system the Larmor scales while retaining their effects over larger scales (called *Finite-Larmor-Radius effects, or FLR effects*).

This feature is still more crucial in *nuclear fusion by magnetic confinement*, which is presently a very active domain, as is emphasized by the worldwide project ITER (International Thermonuclear Experimental Reactor), with its *tokamak* device [12]. Then the plasma behaviour is deeply kinetic, because of the very high temperature and very low collisionality, which precludes the use of fluid models for an accurate description of the plasma, in the core of tokamaks for instance. At the same time, the magnetic field is very strong, and hence the Larmor scales are very small (compared to the other physically important scales), which both makes it impossible to retain these scales in numerical simulations and justifies an averaging process that eliminates these scales while retaining their effects over larger scales.

Thus, gyrokinetics can be viewed as an intermediate model between Vlasov-Maxwell and Euler-Maxwell; when the magnetic field is strong enough, it retains a kinetic behaviour although eliminating small Larmor scales; it is a key model in plasma microturbulence and especially in magnetic fusion physics [80].

In the case of gyrokinetics, the presence of Hamiltonian structures is crucial [17, 29, 64, 65, 80].

It originates from the parent models (particle dynamics and Vlasov-Maxwell dynamics in the Klimontovitch approach) being Hamiltonian. In the BBGKY approach, Vlasov-Maxwell is not Hamiltonian by itself, because of the collision term, but this last is small when the plasma is at high temperature, so that it can be considered as a small correction in the dynamics, which is mainly Hamiltonian.

In addition, the derivation of the model reached its present-day efficiency only when introducing (non-canonical) Hamiltonian perturbation methods, especially because in the averaging procedure they naturally implement the presence of the magnetic moment, which is an additional reduction for the resulting model, as will be explained with more details in a few lines.

Last, Hamiltonian methods are desirable in the derivation of the model because they guarantee the preservation of the Hamiltonian structure, whose conservative properties are crucial (e.g. in numerical simulations) when studying long-time behaviours, as is needed in nuclear fusion.

### The gyrokinetic reduction

Let us have a few words on the derivation of the gyrokinetic model [17], which most of the dissertation will look into with more details. It works in two stages.

The first stage is called *guiding-center theory* [29] and concerns a reduction of particle dynamics, where the electromagnetic field is considered as external and often the electric field is neglected.

When the magnetic field is strong, the particle trajectory is roughly a helix around the magnetic field lines, with a small (negligible) radius and a small (negligible) period, corresponding to the *Larmor radius and time*. The idea is to perform an averaging transformation that removes from the effective dynamical coordinates the fast angle around the helix, namely the *gyro-angle*. In addition, the radius of the helix can be removed from the dynamical variables, because it is related to a constant of motion, called *magnetic moment*. Thus guiding-center theory reduces the dynamics from six to four dimensions. In this transformation, the particle position is changed to another point, called the *guiding-center*, which can be considered as (sliding) the center of the helix.

This transformation is not performed in one single stroke, but in an iterative way through a perturbation theory corresponding to an expansion in a *small parameter*, related to the ratio between the Larmor radius and the other physical lengths, such as the characteristic length at which the magnetic field varies.

After guiding-center theory comes the second stage, which is the *true gyrokinetic stage*. It works not at the level of particle dynamics but at the level of field dynamics, and proceeds in four steps. First, it performs the guiding-center change of coordinates for the Vlasov density. This corresponds to a transformation for field variables. It is called a *"lifting"* of the guiding-center transformation [115], which is "lifted" from the particle level to the field level. Second, it restores the *coupling between the plasma and the electromagnetic field*, i.e. it makes the electromagnetic field not external any more, but evolving in time consistently with the Vlasov density through Maxwell's equations. This slightly spoils the reduction previously obtained, and the third step is to perform a second perturbative transformation, called *gyro-center transformation* [17], in order to restore the desired reduction in the presence of the coupling. Then, the characteristics of the (transformed) Vlasov density are such that the magnetic moment coordinate is not dynamical any more, and the gyro-angle coordinate, although dynamical, has no effect on the other dimensions, so that it can be dropped (averaged out), which constitutes the fourth step.

At the end of the reduction, the dynamics of the (reduced) Vlasov density is such that its characteristics follow the slow, four-dimensional gyro-center trajectories. As for the reduced Maxwell's equations, the dynamics of the electromagnetic field involves source terms given by the reduced Vlasov density, with *polarization and magnetization* terms being induced by the process, and more precisely by the transformation of particle coordinates.

Perhaps it is useful to emphasize that the gain for numerical simulation is huge, because for typical applications in magnetic fusion, the time step is reduced by a factor of the order of $10^3$. In addition, the dynamical dimensions are reduced from 6 to 4. Indeed, the gyro-angle is completely removed from the theory; as for the magnetic moment, it is not dynamical any more and remains only as an index for the Vlasov density; it contributes in the computation of the current (or charge density), but since it is not dynamical, a few values can be enough to get relevant results. Thus, for a discretization of the numerical grid with $10^2 - 10^3$ points per dimension, the size of the grid is reduced by a factor of the order of $10^5$.

Such a reduction is needed to make numerical simulations doable: even after the reduction, simulations can be done only on supercomputers and in the electrostatic limit.

### Electrostatic gyrokinetics

Most often, the *electrostatic* limit can be used in a first approximation. In fusion simulations for instance, gyrokinetics is ordinarily just electrostatic gyrokinetics [17, 42, 56, 65]. This is because a dynamical magnetic field would complicate simulations, whereas the strong magnetic field is

mainly not dynamical (not self-consistently generated by the plasma), so that the main effects of the dynamical electromagnetic field come from the electrostatic field.

In the electrostatic case, the guiding-center stage consists in setting the electric field to zero. As for the gyrokinetic stage, after lifting the guiding-center transformation, it consists in restoring the Poisson equation, with its non-zero electric field. As usual, this spoils the guiding-center reduction, which is remedied by the gyro-center transformation.

So, the overall *two-stage five-step reduction* becomes the following:

1) The guiding-center reduction of particle dynamics reduces the number of dynamical coordinates from six to four. The electric field is zero, the magnetic field is static. The method is to perform a perturbative coordinate transformation on the phase space such that the magnetic moment is a constant of motion and the gyro-angle has a skew decoupled dynamics and can be dropped (averaged out); *"skew decoupled"* means that it does not influence the other coordinates, but that it depends on them, it is a one-way dependence.

2) The gyrokinetic reduction of plasma dynamics reduces the dimension of the grid over which the Vlasov density evolves from six to four (the magnetic moment is kept just as a label, not as a dynamical dimension). The electric field is dynamical and evolves consistently with the Vlasov density, the magnetic field remains static. This second stage proceeds in four steps:

a) to "lift" the guiding-center reduction to the field level by performing the guiding-center transformation for the Vlasov density. This reduces the Vlasov dynamics when the electric field is zero.

b) to restore the coupling, by putting back the Poisson equation. This spoils the reduction of the Vlasov equation, because of the non-zero electric field.

c) to rectify the guiding-center transformation to take into account the presence of the electric field, by performing the gyro-center transformation. In a similar way as for the guiding-center transformation, the gyro-center transformation is done on particle dynamics and lifted to the Vlasov density. Then, the fast gyro-angle coordinate is indeed skew-decoupled for the characteristics of the (transformed) Vlasov density, but it has not been dropped yet.

d) to average the gyro-angle out, which removes the fast time scale from the dynamics, and actually reduces the dimension of the base space for the Vlasov density. Then the reduced dynamics is obtained, which is called (electrostatic) *gyrokinetics*.

The guiding-center and gyro-center transformations can be done simultaneously in Step 1 by including directly the presence of both a strong magnetic field and a weak electric field [128]; it is what we will do. Then, Step 2c is done in the same time as 1 and 2a. Only Steps 1, 2a, 2b, and 2d remain, because Step 2b does not spoil the reduction, since the electric field has already been taken into account in Step 1.

In this case, the difference between guiding-center and gyro-center somehow disappears, and the reduction of particle dynamics (Step 1) often remains called *guiding-center reduction*. If there is a need, the difference can be easily identified afterwards thanks to the order in the electric field.

In our Hamiltonian approach, the gyrokinetic reduction from the Vlasov-Poisson system (electrostatic gyrokinetics) has the same ingredients as the reduction from the full Vlasov-Maxwell system (electromagnetic gyrokinetics), the only difference is that it makes several aspects simpler to identify and sometimes to work out. Accordingly, most of the time, we will not specify whether we have in mind the electrostatic or the electromagnetic case, and we will often use a wording best suited to the standard Vlasov-Poisson framework, whereas most of our derivation will concern the more complete Vlasov-Maxwell system.

Notice that even if the word "Vlasov-Poisson" may evoke just the Vlasov density coupled with an electric field, here it will always implicitly have in addition a (strong static) magnetic field, as is usual in gyrokinetics.

# Thesis work

## Questions addressed

Although well established, the foundations of gyrokinetics are still worth some study, because some questions remain unanswered in its derivation. As a result of recent advances in physics, the model became more widely used, which made clearer and possibly crucial the need for improvements or clarifications about these questions.

1) For instance, the model is obtained through a perturbation theory, but previous results included only a part of the second order [29]. Recent works show that going *further in the expansion* is needed for a proper description of the conservation of angular momentum [26,127], which plays a crucial role in the intrinsic rotation of plasmas in tokamaks, an important phenomenon in magnetic fusion to stabilize turbulence and improve the confinement of energy [63].

2) In addition, when deriving the model, the *fast gyro-angle*, which is somehow the pivotal coordinate, *suffers from difficulties*. It cannot be defined without choosing arbitrarily an axis for the zero of the angle (see formula (1.2)), which corresponds to introducing a local gauge. Such a gauge does not exist globally for a general magnetic geometry [25, 79, 145, 146]. Thus, not only the coordinate system is gauge dependent, but also it is not globally defined in many physical systems. These difficulties about the gyro-gauge are especially important to clarify in order to address the higher-order reduction [26, 145]. An additional awkward feature about the gauge-dependent gyro-angle is that it implies an anholonomic phase in the momentum: when performing a loop in configuration space, at the end, some of the quantities of the theory do not recover their initial value [86, 90].

3) Another question concerns the *Hamiltonian structure of the model*. Indeed, among the two stages of the reduction, the guiding-center stage treats charged-particle dynamics, and it is here that the Hamiltonian structure (of particle dynamics) is essential to obtain the reduced model in an efficient way. In contrast, the second stage, the true gyrokinetic stage, is performed at the level of the equations of motion, which generally does not guarantee the preservation of the Hamiltonian structure.

This asks the question whether the resulting model is Hamiltonian, or whether some fake dissipation has been introduced in the reduction process [108]. Previous attempts showed that answering directly this question by guessing the gyrokinetic Poisson bracket can be difficult. It may be more efficient to study what happens at the level of the Hamiltonian structure (of Vlasov-Maxwell dynamics) through the lifting and the restoring of the coupling. In addition, this strategy has the advantage of relating the Hamiltonian structure of gyrokinetics to the one of Vlasov-Maxwell dynamics.

Identifying the Hamiltonian structure is needed not only to somehow validate the model, and to insure its conservative properties, useful when studying long-time behaviours, but also to clarify some questions about its conservation laws, such as the conservation of angular momentum [21,133], whose answer would be much clarified in the presence of a Hamiltonian system.

## Complementary question

The thesis work deals with those three questions, but a fourth complementary question is also addressed. In order to study the Hamiltonian structure of gyrokinetics, instead of lifting a reduction of particle dynamics, it may be more interesting to apply directly *Hamiltonian reduction methods* to the Vlasov-Maxwell system. This is an extension of the third question above, which aims at better understanding the Hamiltonian reduction from Vlasov-Maxwell to gyrokinetics.

In addition, this extension will be needed in order to perform Step 2d of the reduction process in a Hamiltonian framework. In fact, after the lifting process, in order to remove the gyro-angle dimension from the base space of the Vlasov density (averaging reduction), an additional Hamiltonian reduction at the level of Vlasov-Maxwell field dynamics will appear as necessary before

definitely concluding about the Hamiltonian structure of gyrokinetics.

Last, it could lead to an improvement of the reduction method for gyrokinetics, in the same way as what occurred for the guiding-center reduction, where Littlejohn's Hamiltonian method was so efficient that it replaced previous methods working on the equations of motion.

Several indications suggest that *Dirac's theory of constrained systems* [44] could be efficient for the purpose, and a recent result in the literature [144] somehow confirms these indications. Especially in some cases, applying Dirac constraints just results in a bracket truncation, i.e. the removal of the irrelevant terms from the Poisson bracket. However, this result is not observed in general, and in addition the gyrokinetic reduction does not seem to fit with a straightforward application of Dirac's theory from the Vlasov-Maxwell system as it worked in plasma models up to now [31–33, 118, 119], since the constraints rather regard the coordinates than the fields. In order to develop tools more suited to the purpose, in particular closer to a bracket truncation, the thesis work is interested in variations of Dirac's theory, and as a guideline, several reductions of plasma dynamics are investigated from the point of view of Hamiltonian reductions. This part of the work leaves the mere gyrokinetic model, broadens the perspective and pertains the more general framework of Hamiltonian reductions in plasma physics.

## Organization of the document

The document is organized in three parts, which we will call *"episodes"* [2], in order to avoid the word "part", which could be confusing since "a part" (of the work) can denote a section of the current chapter, or an episode of the whole manuscript, or even some less defined part of the thesis work.

The three episodes are independent of each other, in the sense that they pertain a completely different framework and proceed with completely different sets of methods. They can be considered as three successive stopovers towards the full gyrokinetic reduction, although the third episode touches many other reductions as well.

- In the first episode, we address the first two questions, because they concern the same part of the derivation, which is *guiding-center theory* (Step 1), with the reduction of particle dynamics in a strong magnetic field. We propose an *intrinsic formulation*, i.e. a formulation that avoids introducing a gyro-gauge (considered as extrinsic). We study how the reduction can be performed in this framework and *to arbitrary order* in the perturbation expansion. First, we consider simplified reductions concerned either with the magnetic moment, or with the averaging process. Second we turn to the full reduction. Last, we come back to the troubles about the traditional (non-intrinsic, gauge-dependent) approach, in the light of the intrinsic approach.

- In the second episode, we turn to Steps 2a-2c, i.e. the reduction of plasma dynamics in a strong magnetic field, and we address the third question, by studying how the *Hamiltonian structure* is affected by the lifting and the coupling. This is first done for a simplified transformation suggested by the magnetic-moment reduction, and second for a general change of particle coordinates for Vlasov-Maxwell dynamics, which we then apply to the guiding-center and gyro-center transformations considered by gyrokinetics. Finally, the two-stage gyrokinetic reduction is shown to fit with the Hamiltonian structure of Vlasov-Maxwell and to be proper for a field Hamiltonian perturbation theory.

- In the third episode, we turn to Step 2d, and address the fourth aforementioned question. We study *Dirac* (or more generally *Hamiltonian*) *reductions in plasma dynamics*, with the goal of obtaining the gyrokinetic model by applying Hamiltonian reductions on the Vlasov-Maxwell system. We first consider examples where Dirac's procedure can be simplified, and show that the simplifying phenomenon is related to a different, very efficient reduction method, based on bracket truncations and Poisson subalgebras. Several applications are obtained in plasma physics, both to kinetic and to fluid models, especially for the derivation of incompressible dynamics, of the Vlasov-Poisson system, of the MHD model and of Hamiltonian fluid closures. By the way, the

[2] The motivation for this term purely relies on practical reasons, and on the sake of clarity. We do not intend to originality or to humour at all, but we have not found an alternative word, which would be both more common for a thesis manuscript and unambiguous. Perhaps the word "episode" is unusual, but sure it will be clear.

gyrokinetic reduction is shown to fit with this kind of reduction; this provides a way to implement the averaging reduction (removal of the gyro-angle coordinate) in the Hamiltonian structure of the Vlasov-Maxwell system, which constitutes the last step (Step 2d) towards the Hamiltonian structure of gyrokinetics.

## Comments for the reader

Each episode is divided into several steps towards its final result. Each step is devoted a *chapter*, which aims at clarifying a specific aspect of the overall problem. In the process, each chapter investigates additional, incidental but interesting, questions touching the corresponding subject. In some places, these *complementary questions* can have great interest in themselves, and possibly become an essential motivation for the chapter, besides the initial one coming from the progression of the thesis work.

For instance, Chapters 1, 2, and 10 aim at obtaining more and more complete intrinsic guiding-center reductions, but they are also interested in using the derivation to identify structures in guiding-center expansions, to exploit the non-uniqueness of guiding-center transformations, or to obtain a maximal reduction. As for Chapter 3, it mainly investigates the intrinsic formulation of the guiding-center anholonomy and gauge arbitrariness, but it is also interested in using this formulation to identify the existence condition for a scalar gyro-angle and the link with the existence condition for a (global) gyro-gauge. Chapter 7 mainly explores a formulation of Dirac's reduction as a projection of derivatives, in order to better understand the simplifying phenomenon occurring in some Dirac reductions in plasmas, but it is also interested in using this formulation to improve Poisson brackets that suffer from a conditional Jacobi identity, and this can be considered as an essential motivation for the work reported in this chapter.

As a result, each chapter can be considered as an independent work, with its own framework, viewpoint, subject and focus. So, we have written each of them in a rather self-consistent way, and they will come with an abstract, announcing the specific goals and results. Only the transitions between chapters (beginning of the introduction and end of the conclusion) will position the work of the chapter in the general progression of the thesis work. This presentation will imply a few repetitions between chapters in some points, but they will remain limited and should make the reading easier, especially by dividing this rather long document into shorter elementary entities[3].

It is useful to notice that the work of all chapters do not have the same degree of completion. Some works have been fully developed, exploited, and shaped for publication. For instance, it is the case of Chapters 1-4. Other works remain closer to an exploratory level. Their point of view, content, and formulation are more abstract and more formal-focused. They have not been shaped towards applications and publication yet. Such are Chapters 6 and 12, for instance. Of course, many chapters stand at an intermediate level. For the present manuscript, we have retained all the chapters that we feel play a role and do contribute in the overall understanding of the thesis matter, each with its own present positioning.

In order to avoid excessive length for this thesis dissertation, we have sent into an *appendix episode* several chapters, named *"Appendix chapters"*, and we have kept in the main body only the essential steps of the work. Be careful that still two other kinds of appendix objects can be met in the manuscript. First, in some places, parts of the content of a chapter have been put as *appendix material* at the end of the chapter. Second, this document has *two appendices*, given for information. Appendix A on page 303 is a list of publications. As for Appendix B on page 305, it is a paper of ours whose content has not been detailed elsewhere in the manuscript. It analyses some aspects of the gyro-angle dynamics and is quite aside from the four main questions to be addressed.

Notice that for interested readers, the appendix chapters 10-14 should not be read at the end, after Chapters 1-9, but rather in their place according to the general progression of the investiga-

[3]The appendix chapters 12 and 13 are exceptions from this point of view. They behave somehow as sections of Chapter 8, where the notations are introduced once for the three chapters.

tion. For instance, Chapter 10 takes place at the beginning of Chapter 1, and Chapter 11 takes place at the end of Chapter 4. For each appendix chapter, the place where it should be read is clearly indicated in the main body of the document, as well as at the beginning and at the end of the corresponding appendix chapter.

Under the title of each chapter, collaborators concerned with the corresponding work are indicated. We want to express here our deep gratitude to them, and especially to Michel Vittot, Cristel Chandre, Phil Morrison, and Alain Brizard. They have much helped us make the thesis instructive and productive. As a result, some readers can need to know what our relative contribution is, and more precisely which parts of the dissertation are more specifically ours or theirs. So, we have to stress that the present document focuses on our personal work. The material retained is only the one for which we are the main contributor, and most often by far. The only places where it is not completely the case concern Secs. 4.1-4.5, Secs. 6.1-6.2, Sec. 7.2.2, as well as the framework of Chapter 10. These parts of the thesis have been included rather for pedagogical purpose. For the rest of the dissertation, the role of our collaborators was mainly to suggest directions to investigate, to verify the outcomes, to advise us how to complement the work, to obtain some results differently, or to offer formulations for publications. Nevertheless, we insist that they have played a determinant role in guiding us so wisely and so liberally throughout our investigations. We wish to deeply and heartily thank them.

# Episode I

# Intrinsic approach of the guiding-center reduction

# Introduction of the episode

In this first episode, we are interested in clarifying the first two questions about the foundations of gyrokinetics mentioned in the general introduction, which concern the higher-order reduction and the gyro-gauge issues. Rather than gyrokinetics itself, they affect guiding-center theory, which is the preliminary stage of gyrokinetics (denoted as Step 1 on page 16), and hence do not really pertain plasma dynamics but rather particle dynamics.

Let us remind the principle of the reduction with a few more details [29, 86, 88, 122], to help the reader expect the procedure which will be used. When the magnetic field is strong, the particle trajectory is roughly a helix around the magnetic field lines with a small radius compared to the other scales of interest and a frequency much faster than the time scales of interest. This justifies both a perturbation expansion in the corresponding small parameter $r_L/L$ (where $r_L$ is the Larmor radius and L indicates the other scales of interest, e.g. a characteristic length of the magnetic geometry) and an averaging reduction, which replaces the particle position by the center of the helix, called *guiding-center position*, whose motion is slow and averaged at lowest order. The same averaging can be done with other coordinates, except that the gyro-angle keeps a fast dynamics.

This removes the fast time scale from the theory, because the slow dynamics can be studied independently of the fast dynamics: it does not depend of the only fast coordinate, the gyro-angle. But this is valid only at lowest order. At first order, the dynamics of the guiding-center (and other coordinates) depends on the gyro-angle. Thus a first-order transformation must be performed to average also these terms. Then, the dynamics depends on the gyro-angle only for order two and higher. Iterating the process to infinity provides a perturbative change of coordinates that completely removes from the equations of motion the presence of the gyro-angle, the only fast variable. Since the gyro-angle affects only short time scales and small spatial scales, it can be ignored (averaged out) and the theory remains with only five coordinates and a much less stringent time scale.

In addition, one of the remaining coordinates can also be removed from the dynamical variables, by making it a constant of motion. Indeed, the magnetic flux through the helix is an adiabatic invariant [68, 122], which means that at lowest order, it is a constant of motion. At first order, its dynamics is not zero, but a first-order transformation can make it conserved up to first order. Iterating the process to infinity provides a constant of motion to all orders, called the *magnetic moment*.

Practically, the transformation for the averaging reduction and the one for the magnetic moment are done simultaneously at each order. Thus, guiding-center theory reduces the dynamical dimensions from 6 to 4 by identifying a suitable perturbative coordinate transformation.

The reduction was initially obtained by averaging the equations of motion and trying to combine the results to identify the magnetic moment [8, 122, 123], which became very difficult, if not impossible, for higher orders. Since Littlejohn's Hamiltonian approach [85, 86, 88], it is rather obtained by Lie-transforming the phase-space Lagrangian, because it is a very efficient method, especially for the magnetic moment. In addition, it naturally provides the reduced motion with a Hamiltonian structure.

The first point to be developed about this reduction is that in previous works, the reduction was performed only up to the order one and a half [29]. Recent results emphasize that going further can be needed for the physical relevance of the equations, especially for the conservation of angular

momentum [26,127]. This is essential in order to properly describe the intrinsic rotation of plasmas in tokamaks, which plays a crucial role in the stabilization of plasma turbulence, and hence in improving the heat confinement, which is the main goal in magnetic fusion.

So, in this thesis, we are interested in studying the reduction at higher orders. In the present episode, this point will be addressed by obtaining algorithmic reductions with explicit induction relations to arbitrary order in the perturbation expansion.

The second question comes from the definition of the gyro-angle, which is the angle determining the direction of the so-called *perpendicular momentum* (or velocity), the component of the momentum perpendicular to the magnetic field. Indeed, the definition of this angle implies to choose an axis (a unit vector $\mathsf{e}_1$) from which the gyro-angle is measured (see formula (1.2)). Since the vector $\mathsf{e}_1$ must be orthogonal to the magnetic field $\mathbf{B}(\mathbf{q})$, this choice is to be made at each point in configuration space.

This corresponds to fixing a gauge in the theory, the so-called *gyro-gauge*, because the value of the gyro-angle depends on the choice made for $\mathsf{e}_1$. A first aspect of the question concerns the need for this gauge dependence in the coordinate system, whereas the particle state and all the physics do not explain such a gauge dependence [86, 90]. A second, more important aspect is that the gyro-gauge cannot be defined globally for a general magnetic geometry [25,79,145,146]. Hence the coordinate system does not exist in the whole physical system in general. Last, awkward features are involved in the gyro-angle, with the presence of an anholonomic phase in the momentum [86,90].

These troubles come because this coordinate is not purely physical. In order to address them, a natural idea is to avoid introducing such non-physical quantities in the theory, and hence to remain as close to the physical coordinates as possible. This formulation will be named *"intrinsic"*, because it avoids introducing any arbitrary (extrinsic) gyro-gauge, and it is directly induced by the (intrinsic) state of the system. A first goal will be to identify the corresponding intrinsic gyro-angle coordinate, which suggests to identify it a posteriori instead of defining it a priori. So, we should start the work without defining it.

On another hand, as will be seen in the next episode, when tackling the third issue mentioned in the introduction, related to the Hamiltonian structure of gyrokinetics and the lifting procedure, some difficulties will be induced by the introduction of a guiding-center, i.e. by a change in the spatial point. So, we are interested in an approach that changes as little as possible the coordinates, and especially that does not change the spatial position.

This excludes the averaging reduction, and suggests to focus on a simplified reduction, which takes into account only the magnetic moment, and not the removal of the fast angle. Then only one of the coordinates has to be changed, e.g. the norm of the momentum, to be replaced by the magnetic moment. The other coordinates keep their initial definition, both the spatial position and the unit vector of the momentum. Especially, no gyro-gauge is defined for the gyro-angle.

So, the magnetic-moment reduction is both the best way to look for an intrinsic formulation of gyrokinetics and a good way to prepare the lifting procedure.

Especially, this simplified reduction should show how gyrokinetics can work with no gyro-gauge, because the magnetic moment is the action conjugated to the gyro-angle [81,86,88], and hence its derivation will be related to Larmor gyration. Then, equipped with the resulting intrinsic formalism, less simplified guiding-center reductions will be considered.

The intrinsic gyro-angle will turn out to be a constrained coordinate, namely the direction (unit vector) of the perpendicular momentum (as could be guessed a priori, actually), which is constrained to remain of unit norm and perpendicular to the magnetic field. This will introduce new features in the theory, which must be dealt with carefully, especially because they could contain counterparts of the issues associated with the traditional gyro-angle. So, before going to the full guiding-center reduction, we will have to consider a simplified, introductory form of the averaging reduction.

As a result, this episode of the thesis manuscript is composed of four chapters.

- In a preliminary scrutiny, we study a transformation that keeps the spatial position unchanged, as well as the unit vector of the momentum. Its goal is just to obtain the magnetic moment reduction. It is not a guiding-center transformation (since there is no guiding-center), but it will introduce an intrinsic formalism for gyrokinetics, especially for the basic gyro-operators (gyro-angle derivation, gyro-integral, averaging and fluctuating operators), as well as an intrinsic coordinate for the gyro-angle, which will appear as a constrained coordinate. In order to avoid excessive length for the main body of the manuscript, this work will be put among the appendix chapters, in Chapter 10.

- Then, in order to identify how the guiding-center reduction works with this constrained coordinate, we investigate the opposite limit, i.e. a transformation that just averages the motion of the four slow reduced coordinates. This minimal guiding-center transformation will be the topic of Chapter 1.

- Next, in Chapter 2, we will generalize the intrinsic formulation by studying how the constrained coordinate can be used for the full guiding-center reduction, which is more involved because it works on the phase-space Lagrangian, not just on the equations of motion, and it can be viewed as a maximal reduction. In this reduction, as in the previous ones, in addition to using a gauge-independent gyro-angle, we will be interested in the higher orders, and will propose an algorithmic induction procedure.

So, for the methods used, our work will parallel and reconsider the historical path: the full reduction will be more efficiently obtained by working on the phase-space Lagrangian whereas the minimal averaging reduction will be more efficiently obtained by working on the equations of motion, but instead of an averaging procedure, a Lie-transform will make the scheme clearer and more straightforward.

- Last, after obtaining all the results of standard guiding-center reductions within the gauge-independent framework, it will be interesting to revisit the troubles associated with the traditional gyro-angle coordinate in the light of the intrinsic approach, in order to see what intrinsic features underlie these troubles, and to emphasize more clearly what the essential novelties of this approach are. This will be the purpose of Chapter 3.

# Chapter 1

# A gyro-gauge independent minimal guiding-center reduction by Lie-transforming the velocity vector field

in collaboration with Michel Vittot
Paper published in Phys. Plasmas **20**, 082505 (2013)

*A*bstract: We introduce a gyro-gauge independent formulation for a simplified guiding-center reduction, which removes the fast time-scale from particle dynamics by Lie-transforming the velocity vector field. This is close to Krylov-Bogoliubov method of averaging the equations of motion, although more geometric.
At leading order, the Lie-transform consists in the generator of Larmor gyration, which can be explicitly inverted, while working with gauge-independent coordinates and operators, by using the physical gyro-angle as a (constrained) coordinate. This brings both the change of coordinates and the reduced dynamics of the minimal guiding-center reduction order by order in a Larmor radius expansion.
The procedure is algorithmic and the reduction is systematically derived up to full second order, in a more straightforward way than when Lie-transforming the phase-space Lagrangian or averaging the equations of motion. The results write up some structures in the guiding-center expansion. Extensions and limitations of the method are considered.

## Introduction

As announced in the introduction of the episode, the preliminary step towards our intrinsic formulation of the guiding-center reduction focuses on a simplified reduction taking into account only the magnetic moment, but not the removal of the fast time scale. It is reported in the appendix Chapter 10.

⟶ See the appendix chapter 10

It shows that the physical variable playing the role of the gyro-angle is the unit vector of the perpendicular momentum. Unlike the standard gyro-angle used in the literature, it is not extrinsic, in the sense that no arbitrary gyro-gauge is involved in its definition, so that it is directly induced by the particle state.

The simplified reduction of Chapter 10 only constituted a small introductory part of the guiding-center reduction, since it did not remove the fast time scale from the theory, and had no guiding-center. As the next step towards our intrinsic approach, we now turn to a true guiding-center reduction, i.e. we aim at eliminating the fast time scale from the dynamics while using the physical gauge-independent coordinate for the gyro-angle.

To make this chapter clear for the reader, and somehow self-consistent, let us remind briefly the framework and purpose.

The basic idea of guiding-center theory is that particle dynamics in a strong static magnetic field implies a separation of scales, with the existence of a fast Larmor rotation, slower drifts, and an adiabatic invariant. This allows for a reduction, the so-called guiding-center reduction, which removes the fast time scale and the fast angle from the dynamics, builds a constant of motion, the magnetic moment, and thus provides a slow motion reduced by two dimensions [29]. The fast angle is the gyro-angle, which is an angle measuring the Larmor gyration. The main principle of the reduction is to remove the presence of the gyro-angle in the dynamics of the other coordinates, by building suitable coordinates order by order in a small parameter expansion, related to the Larmor radius. The averaging transformation is usually identified by Lie-transforming the phase-space Lagrangian, because it naturally provides the constant of motion and a Hamiltonian structure for the slow reduced motion.

The usual coordinate for the gyro-angle suffers from several issues, both from a mathematical and from a physical point of view [25,79,90,145,146]: it is gauge dependent, does not exist globally for a general magnetic geometry, and implies an anholonomic phase in the momentum. The gauge dependence induces some subtlety about gauge invariance, while the non-global existence and the anholonomy are still more puzzling. In order to avoid these issues and to give a more intrinsic framework to the theory, we consider performing the reduction using the initial physical gauge-independent coordinate for the gyro-angle, as suggested by the work reported in Chapter 10. Thus, not only will the theory be gauge invariant and the physical results be gauge independent, but all the theory itself will be gauge independent and defined globally, with all of its quantities.

The corresponding coordinate is constrained, which makes the scheme more involved. As a first step towards the full reduction, we study in this chapter a simplified guiding-center reduction, where we focus on the averaging reduction (i.e. we do not take into account the magnetic moment, which is the opposite limit compared to the previous step, considered in Chapter 10). The advantage is that only the four slow reduced coordinates have to be changed, since only they have to be gyro-averaged. For the remaining coordinates one can rely on the initial physical coordinates, which are gauge independent. The scheme also is strongly simplified, since one can work directly on the equations of motion, which makes it easier to identify how to deal with the gyro-angle using an intrinsic but constrained coordinate.

This simplified guiding-center transformation is obtained here by Lie-transforming directly the velocity vector field, which is the geometric point of view of how a coordinate transformation affects the equations of motion [27, 87]. As a result, it is much more efficient than previous methods either Lie-transforming the phase-space Lagrangian [29, 88] or averaging the equations of motion [8, 9, 81, 122, 123]. This makes it easier to go to higher order in the reduction and to identify structures in guiding-center expansions. The whole second-order reduction is obtained in a direct computation, despite the complicated expressions of the second-order transformation, and the algorithm can be easily implemented in a computer to reach any order.

The method is not intended to replace the traditional derivation by Lie-transforming the phase-space Lagrangian. Indeed, with some additional refinements, it can also provide the constant of motion as in Chapter 10, but it does not guarantee a Hamiltonian structure for the slow reduced coordinates, with all the associated conservation properties. All the same, it induces a complementary and faster derivation to get additional information on the reduction, and especially it is the first step towards a gauge-independent approach of guiding-center theory.

When applied to the velocity vector field, the Lie-transform method is close to Krylov-Bogoliubov method [82], but more geometric and more straightforward. Bogoliubov's procedure [8,9] also relied on a near-identity change of coordinates in the equations of motion and achieved greater efficiency than other methods working on the equations of motion, but it relied on substitutions and chain rules, whose process remained rather involved, so that it was not implemented further than the first-order transformation. The Lie-transform method condensates all the procedure into one single operation, related to the transformation generator.

It also comprises an interesting approach of the guiding-center arbitrariness. Indeed, the guiding-center reduction is not unique, as is again emphasized by some recent results [23, 128]. Such an arbitrariness caused troubles in some works dealing with the equations of motion, where it was unclear what the natural choices were to be at second order [123]. Hamiltonian approaches [85, 86, 88] did not really address the arbitrariness in the averaging procedure, because they average not just the dynamics of the four reduced coordinates, but the whole reduced Lagrangian, with its seven components, including the Hamiltonian and the Poisson bracket. Now, when Lie-transforming the equation of motion, the arbitrariness is distinct and the natural choice consists in setting to zero all arbitrary terms in the transformation generator, exactly as it was done at first order to define the guiding-center position. This also agrees with Bogoliubov's works [8, 9].

The resulting guiding-center reduction is unique and minimal. The requirements are just tailored to reduce the dynamical dimensions, which means just to extract a slow dynamics for the four reduced coordinates, the guiding-center position and the parallel velocity or the pitch-angle (the angle between the magnetic field and the particle velocity). Then, the method is forthright from the requirements to the related equations to be solved; it just corresponds to making a near-identity change of coordinates, i.e. a Lie-transform of the velocity vector field. Computations are straightforward and fully algorithmic, since the equations are solved just by an explicit inversion of the Larmor gyration generator. Last, the result is the minimal transformation for the desired reduction, in the sense that it is the only four-dimensional generator with zero average that allows for the reduction, and that the reduction cannot be performed without this minimum needed. This does not preclude the possible existence of other minimal guiding-center reductions when different minimization criteria are used, but it seems the criterion considered here is best suited to the main goal and procedure of guiding-center theory, concerned with the averaging reduction.

The chapter is organized as follows. In Sec. 1.1, the goals of the reduction are introduced and written as equations for the transformation generator. It is shown that these equations can be solved explicitly at any order in the Larmor radius, and the iteration mechanism of the reduction is identified. In Sec. 1.2, it is shown how the reduction is computed and the resulting algorithm is given. In Sec. 1.3, the result is written to full second order in the Larmor radius, obtained in a straightforward computation; a comparison is done with previous guiding-center reductions, and some insights are brought into the general structure of guiding-center results.

In Sec. 1.4, a few possible extensions and limitations of the method are studied, such as the presence of an electric field or non minimal guiding-center reductions. For instance, averaging also the dynamics of $\theta$ is considered, as well as including the magnetic moment in the reduced coordinates or providing the reduced dynamics with a Hamiltonian structure.

## 1.1 Lie-transforming the velocity vector field

The dynamical system is simply a charged particle with position $\mathbf{q}$, momentum $\mathbf{p}$, mass $m$ and charge $e$, under the influence of a static inhomogeneous magnetic field $\mathbf{B}$. The motion is given by the Lorentz force

$$\begin{aligned}\dot{\mathbf{q}} &= \tfrac{\mathbf{p}}{m}\,,\\ \dot{\mathbf{p}} &= \tfrac{\mathbf{p}}{m}\times e\mathbf{B}\,.\end{aligned}$$

For the sake of clarity, we consider no electric field. The presence of an electric field satisfying the guiding-center ordering would not change the method (see Sec. 1.4).

When the magnetic field is strong, the motion implies a separation of time-scales. This is best seen by choosing convenient coordinates for the momentum space

$$\begin{aligned}p &:= \|\mathbf{p}\|\,,\\ \varphi &:= \arccos\left(\tfrac{\mathbf{p}\cdot\mathsf{b}}{\|\mathbf{p}\|}\right)\,,\\ \mathsf{c} &:= \tfrac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|}\,,\end{aligned} \tag{1.1}$$

where $\mathsf{b} := \frac{\mathbf{B}}{\|\mathbf{B}\|}$ is the unit vector of the magnetic field, and $\mathbf{p}_\perp := \mathbf{p} - (\mathbf{p}\cdot\mathsf{b})\mathsf{b}$ is the orthogonal projection of the momentum onto the plane perpendicular to the magnetic field. The coordinate $p$ is the norm of the momentum, $\varphi$ is the so-called pitch-angle, i.e. the angle between the velocity and the magnetic field. Following Littlejohn's notations [85, 86, 88], the vector $\mathsf{c}$ is the unit vector of the perpendicular velocity.

To avoid misreadings, we insist that several conventional notations can be found in the literature. In the present manuscript, following many of our bibliographic references, we denote the gyro-angle by $\theta$ (see Eq. (1.2)), and the pitch-angle by $\varphi$, because $(p, \theta, \varphi)$ constitutes a spherical coordinate system for the momentum.

Differentiating Eqs. (1.1), the equations of motion are found as

$$\begin{aligned}
\dot{\mathbf{q}} &= \tfrac{\mathbf{p}}{m}\,,\\
\dot{p} &= 0\,,\\
\dot{\varphi} &= -\tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\,,\\
\dot{\mathsf{c}} &= -\tfrac{eB}{m}\mathsf{a} - \tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot(\mathsf{c}\mathsf{b} + \mathsf{a}\mathsf{a}\cot\varphi)\,,
\end{aligned}$$

where $\mathbf{p}$ is now a shorthand for $p(\mathsf{b}\cos\varphi + \mathsf{c}\sin\varphi)$, $B$ is the norm of the magnetic field and $\mathsf{a} := \mathsf{b}\times\mathsf{c}$ is the unit vector of the Larmor radius, following Littlejohn's notations, so that $(\mathsf{a}, \mathsf{b}, \mathsf{c})$ is a (rotating) right-handed orthonormal frame. The computation of $\dot{\mathsf{c}}$ is rather involved, but it can be avoided by using Eq. (1.12).

In the case of a strong magnetic field, the only fast term, the Larmor frequency $\omega_L := \frac{eB}{m}$ concerns only one coordinate, $\mathsf{c}$, the direction of the vector $\mathbf{p}_\perp$ in the 2-dimensional plane perpendicular to the magnetic field. This corresponds to an angle, the so-called gyro-angle, and measures the Larmor gyration of the particle momentum around the magnetic field.

To get a true scalar angle instead of the vector $\mathsf{c}$, one chooses at each point $\mathbf{q}$ in space a direction which will be considered as the reference axis $\mathbf{e}_1(\mathbf{q}) \in \mathbf{B}^\perp(\mathbf{q})$. Then, the gyro-angle $\theta$ is defined from the oriented angle between the chosen reference axis $\mathbf{e}_1(\mathbf{q})$ and the vector $\mathsf{c}$:

$$\mathsf{c} = -\sin\theta\mathbf{e}_1 - \cos\theta\mathbf{e}_2\,, \tag{1.2}$$

and the equation of motion for $\theta$ is

$$\dot{\theta} = \tfrac{eB}{m} + \cot\varphi\tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{a} + \tfrac{\mathbf{p}}{m}\cdot\nabla\mathbf{e}_1\cdot\mathbf{e}_2\,,$$

with $\mathbf{e}_2 := \mathsf{b}\times\mathbf{e}_1$ is the unit vector such that $(\mathsf{b}, \mathbf{e}_1, \mathbf{e}_2)$ is a fixed right-handed orthonormal frame.

To emphasize the fast term of the dynamics in strong $B$, the equations of motion can be expanded in $B^{-1}$. Writing $\mathbf{z} := (\mathbf{q}, p, \varphi, \theta)$, we get

$$\begin{aligned}
\dot{\mathbf{z}} &= \tfrac{eB}{m}\left[\begin{pmatrix}0\\0\\0\\1\end{pmatrix} + \tfrac{p\sin\varphi}{eB}\begin{pmatrix}\mathsf{b}\cot\varphi+\mathsf{c}\\ 0\\ -(\mathsf{b}\cot\varphi+\mathsf{c})\cdot\nabla\mathsf{b}\cdot\mathsf{c}\\ (\mathsf{b}\cot\varphi+\mathsf{c})\cdot(\cot\varphi\nabla\mathsf{b}\cdot\mathsf{a}+\nabla\mathbf{e}_1\cdot\mathbf{e}_2)\end{pmatrix}\right]\\
&= \dot{\mathbf{z}}_{-1} + \dot{\mathbf{z}}_0\,,\\
\text{with } \dot{\mathbf{z}}_{-1} &:= \begin{pmatrix}0\\0\\0\\ \frac{eB}{m}\end{pmatrix}\,,\\
\text{and } \dot{\mathbf{z}}_0 &:= \begin{pmatrix}\frac{\mathbf{p}}{m}\\ 0\\ -\frac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\\ \frac{\mathbf{p}}{m}\cdot(\cot\varphi\nabla\mathsf{b}\cdot\mathsf{a}+\nabla\mathbf{e}_1\cdot\mathbf{e}_2)\end{pmatrix}\,,
\end{aligned} \tag{1.3}$$

where the indices correspond to the order in $B^{-1}$. The term of order $B^1$ has dimension of $\frac{eB}{m}$, whereas all the terms of order $B^0$ have dimension of $\frac{\mathbf{p}}{m}\cdot\nabla$, if we use the correspondence between the vector field $\dot{\mathbf{q}}$ and the differential operator $\dot{\mathbf{q}}\cdot\nabla$.

The ordering parameter, that is the ratio of two consecutive terms in the perturbation expansion, is then $\varepsilon = \frac{\dot{\mathbf{z}}_0}{\dot{\mathbf{z}}_{-1}} = \frac{p\sin\varphi}{eB}\nabla$, which is the dimensionless quantity $r_L\nabla$, with $r_L := \frac{p\sin\varphi}{eB}$ the Larmor radius. This parameter is well-known as the magnetic inhomogeneity at the scale of the Larmor radius, because in $r_L\nabla$, the gradient always acts on the magnetic field, which is the only local property of the configuration space.

The ordering parameter is sometimes considered just as the Larmor radius $r_L$, or as $B^{-1}$, in agreement with an expansion in strong magnetic field. Also, it is often considered symbolically as $e^{-1}$, because it is an equivalent expansion but it avoids to deal with a space-dependent parameter involving $B$; this is symbolic, because $e^{-1}$ is not dimensionless. Notice that the ordering parameter $r_L\nabla$ is not just a scalar but an operator; it has only a kind of dimensional meaning: the terms of order $\varepsilon^2$ may not have a prefactor of $(r_L\nabla)^2$, but they will have the dimension of $\frac{p^2}{e^2B^2}\nabla^2$. This point will be illustrated by the results of Sec. 1.3.

The goal is to isolate the dynamics of the slow variables from the fast variable, i.e. to perform a change of coordinates $\tau : \mathbf{z} \to \bar{\mathbf{z}}$ such that the dynamics of the remaining coordinates $(\bar{\mathbf{q}}, \bar{\varphi}, p)$ does not depend on $\theta$. This requirement is already obtained for $p$, so one only has to change coordinates on $\mathbf{q}$ and $\varphi$. The coordinate transformation $\tau$ transfers to functions by duality through the "push-forward" operator [1] $\mathsf{T}^{-1}$, defined by the scalar invariance property [17]:

$$(\mathsf{T}^{-1}f)(\bar{\mathbf{z}}) = f(\tau^{-1}\bar{\mathbf{z}})\,.$$

At the lowest order in the Larmor radius $\varepsilon^{-1}$, the requirement is trivially satisfied. So, the transformation can be near-identity. It can be written as the exponential of a Lie-transform $\bar{\mathbf{z}} = e^{-\mathbf{X}}\mathbf{z}$, with $-\mathbf{X}$ a vector field, generator of the diffeomorphism $\tau$, which satisfies

$$\mathbf{X}_{p,\theta} = 0\,,$$

since only the coordinates $\mathbf{q}$ and $\varphi$ need to be changed. The index notation is used to indicate the components, e.g. $\mathbf{X}_\varphi$ denotes the component $\varphi$ of the vector $\mathbf{X}$, and $\mathbf{X}_{p,\theta}$ denotes a 2 dimensional vector, whose coordinates are the components $p$ and $\theta$ of $\mathbf{X}$. Through the transformation, the equations of motion become

$$\dot{\mathbf{z}} \to \dot{\bar{\mathbf{z}}} = e^{\mathsf{L}}\dot{\mathbf{z}}\,,$$

with $\mathsf{L} := \mathcal{L}_{\mathbf{X}}$ the Lie-transform along the vector field $\mathbf{X}$.

Now, the goal is that the equations of motion for $\mathbf{q}$ and $\varphi$ do not depend on $\theta$, which means that all their non-zero Fourier components (i.e. purely oscillatory terms) are zero:

$$\mathtt{osc}(\dot{\bar{\mathbf{z}}})_{\mathbf{q},\varphi} = 0\,,$$

where following Littlejohn's notations, $\mathtt{osc} = 1 - \mathtt{avg}$ is the projector onto gyro-fluctuations, with $\mathtt{avg}$ the complementary projector onto gyro-averages:

$$\mathtt{avg}(f) = \tfrac{1}{2\pi}\int_0^{2\pi} d\theta\; f$$

for any function $f$.

Last, $\mathsf{L}$ is expanded in series in the small parameter $\varepsilon$

$$0 = \mathtt{osc}\left(\dot{\bar{\mathbf{z}}}\right)_{\mathbf{q},\varphi} = \mathtt{osc}\left(e^{\mathsf{L}_1+\mathsf{L}_2+\cdots}\dot{\mathbf{z}}\right)_{\mathbf{q},\varphi}\,, \tag{1.4}$$

[1] Notice that some works [131] use the word "*pull-back*" to transform functions (and differential forms) both in the direct sense and in the reverse sense. They name $\mathsf{T}\bar{f}$ the pull-back of $\bar{f}$ (by $\tau$), and $\mathsf{T}^{-1}f$ the pull-back of $f$ (by $\tau^{-1}$). In this chapter, following references in the guiding-center literature [17, 87], we distinguish between direct and reverse transformations by naming $\mathsf{T}^{-1}f$ the *push-forward* of $f$, which emphasizes that the corresponding transformation is not backwards, but forwards.

Be careful that the word *push-forward* may refer to a different concept, namely transformations of vectors (or contravariant tensors), as opposed to forms (or covariant tensors). However, when $\tau$ is bijective, for instance for the guiding-center transformation, a unified treatment in the case of general tensors (mixed covariant and contravariant tensors) implies to consider the pull-back as the push-forward of the inverse transformation, and vice-versa, which agrees with the terminology used in the guiding-center literature.

where $\mathsf{L}_i$ is of order $\varepsilon^i$.

Eq. (1.4) is the equation to be solved for the change of coordinates $\mathsf{L}_{n+1}$. Expanding it in series in $\varepsilon$, we get an equation for each order:

$$\begin{aligned}
0 &= \mathtt{osc}\,(\dot{\mathbf{z}}_{-1})_{\mathbf{q},\varphi}\ , \\
0 &= \mathtt{osc}\,(\mathsf{L}_1\dot{\mathbf{z}}_{-1} + \dot{\mathbf{z}}_0)_{\mathbf{q},\varphi}\ , \\
0 &= \mathtt{osc}\left(\mathsf{L}_2\dot{\mathbf{z}}_{-1} + \tfrac{\mathsf{L}_1^2}{2}\dot{\mathbf{z}}_{-1} + \mathsf{L}_1\dot{\mathbf{z}}_0\right)_{\mathbf{q},\varphi}\ , \\
&\ldots
\end{aligned} \tag{1.5}$$

At each order $n \geqslant 1$ in $\varepsilon$, the highest-order unknown $\mathsf{L}_{n+1}$ is involved only in one term. Isolating it, the equation reads

$$-\,\mathtt{osc}(\mathsf{L}_{n+1}(\dot{\mathbf{z}}_{-1})_{\mathbf{q},\varphi}) = (\mathbf{N}_n)_{\mathbf{q},\varphi}\,, \tag{1.6}$$

where $\mathbf{N}_n$ is a shorthand for all the terms of Eq. (1.4) that are of order $n$ and that do not include the unknown $\mathsf{L}_{n+1}$, e.g. $\mathbf{N}_1 := \mathtt{osc}\left(\frac{\mathsf{L}_1^2}{2}\dot{\mathbf{z}}_{-1} + \mathsf{L}_1\dot{\mathbf{z}}_0\right)_{\mathbf{q},\varphi}$.

The operator to be inverted is

$$\begin{aligned}
&-\mathtt{osc}(\mathsf{L}_{n+1}(\dot{\mathbf{z}}_{-1}))_{\mathbf{q},\varphi} \\
&\quad = -\mathtt{osc}\,((\mathbf{X}_{n+1})_i\mathbf{p}_i(\dot{\mathbf{z}}_{-1})_{\mathbf{q},\varphi} - (\dot{\mathbf{z}}_{-1})_i\partial_i(\mathbf{X}_{n+1})_{\mathbf{q},\varphi}) \\
&\quad = \mathtt{osc}\,((\dot{\mathbf{z}}_{-1})_\theta\partial_\theta(\mathbf{X}_{n+1})_{\mathbf{q},\varphi}) \\
&\quad = \tfrac{eB}{m}\partial_\theta(\mathbf{X}_{n+1})_{\mathbf{q},\varphi}\,,
\end{aligned} \tag{1.7}$$

where, in the first equality we used the usual formula (1.18) for the Lie-transform of a vector field, in the second equality we used $(\dot{\mathbf{z}}_{-1})_{\mathbf{q},\varphi} = 0$, and in the third equality we used that $\mathtt{osc}\ \partial_\theta$ is just $\partial_\theta$, because $\partial_\theta$ takes its values in the gyro-fluctuations, i.e. the non-zero Fourier component in the variable $\theta$.

The operator $\omega_L\partial_\theta$ is the generator of Larmor gyration, with the Larmor frequency as a coefficient. Eq. (1.6) has a solution because the right-hand side $(\mathbf{N}_n)_{\mathbf{q},\varphi}$ is in the range of the operator $\partial_\theta$, since it is a pure gyro-fluctuation. This operator is easily inverted as

$$(\mathbf{X}_{n+1})_{\mathbf{q},\varphi} = \mathtt{avg}\left((\mathbf{X}_{n+1})_{\mathbf{q},\varphi}\right) + \tfrac{1}{\omega_L}\int d\theta\ (\mathbf{N}_n)_{\mathbf{q},\varphi}\,, \tag{1.8}$$

where $\mathtt{avg}(\mathbf{X}_{n+1})_{\mathbf{q},\varphi}$ is a free element in the kernel of $\partial_\theta$, that is a free gyro-averaged function. We defined $\int d\theta\ (\mathbf{N}_n)_{\mathbf{q},\varphi}$ as the primitive of $(\mathbf{N}_n)_{\mathbf{q},\varphi}$ with zero gyro-average, i.e. $\int d\theta\ (\mathbf{N}_n)_{\mathbf{q},\varphi} := \mathtt{osc}\,(\mathcal{N})$ with $\mathcal{N}$ any primitive of $(\mathbf{N}_n)_{\mathbf{q},\varphi}$.

As a result, Eq. (1.5) can be solved to arbitrary order in the small parameter $\varepsilon$ through Eq. (1.8). At each order, one only has to write Eq. (1.5) to the $n$-th order, group into $(\mathbf{N}_n)_{\mathbf{q},\varphi}$ all the terms that depend only on quantities that are already explicitly known, expand the result, and last invert the Larmor gyration operator through Eq. (1.8).

This is close to Kruskal's averaging procedure [81], whose principle was to expand the function defining the motion and average this expansion. Here, we focus on the vector field defining the motion and Lie-transform it so as to make it independent of the fast coordinate. Such a procedure actually does not rest on purely averaging methods so much; it could rather be related to normal form methods [57], which aim at giving a simplified form to a vector field, when studying the local dynamics near an equilibrium point, e.g. the stability and the possible bifurcations. Although distinct from purely averaging methods, normal form methods can be used in averaging reductions [57].

In the guiding-center transformation, at each order, the fluctuating part of $\mathbf{X}_{n+1}$ given by Eq. (1.8) is necessary and sufficient to solve Eq. (1.4). The averaged part is completely free, but also completely useless to solve Eq. (1.4). So, a natural choice is to put it to zero

$$\mathtt{avg}\,(\mathbf{X}_{n+1}) = 0\,. \tag{1.9}$$

This makes the transformation unique: the result is the minimal guiding-center reduction, i.e. the only transformation which gives the desired requirements and whose only non-zero components are the fluctuating part of $(\mathbf{X}_{n+1})_{\mathbf{q},\varphi}$.

In this guiding-center reduction, the transformation affects only the coordinates $(\mathbf{q},\varphi)$. The transformation generator has no gyro-angle component. Now, the corresponding coordinate $\theta$ is not intrinsic. It implies a non-trivial gauge fixing, whose global existence can fail [25, 145]. It is not purely physical: in all physical results, what appears is the physical quantity $\mathsf{c}$, as is clear in the literature [29, 86, 88]. The use of $\theta$ is a detour and should be avoidable.

So, we remain with the variable $\mathsf{c}$ with its initial definition: the unit vector of the perpendicular velocity

$$\mathsf{c} := \frac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|} = \frac{\mathbf{p}-(\mathbf{p}\cdot\mathsf{b})\mathsf{b}}{p\sin\varphi}\,. \tag{1.10}$$

With this variable, the coordinate space is constrained: the gyro-angle $\mathsf{c}$ is not independent of the spatial position, since $\mathsf{c}\in\mathsf{b}^\perp$, and $\mathsf{b}$ depends on $\mathbf{q}$. When the $\mathbf{q}$ coordinate is changed, the $\mathsf{c}$ coordinate cannot be kept unchanged, otherwise it may get out of $\mathsf{b}^\perp$. Differentiating relation (1.10) with respect to $\mathbf{q}$, we find

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot(\mathsf{cb}+\mathsf{aa}\cot\varphi)\,. \tag{1.11}$$

This formula can be obtained more easily by noticing that in the change of coordinates $(\mathbf{q},\mathbf{p})\longrightarrow(\mathbf{q},p,\varphi,\mathsf{c})$, we have

$$\begin{aligned}
&-\sin\varphi\nabla\varphi = \nabla\cos\varphi = \nabla\mathsf{b}\cdot\frac{\mathbf{p}}{p} = \nabla\mathsf{b}\cdot\mathsf{c}\sin\varphi\\
&\Rightarrow \nabla\varphi = -\nabla\mathsf{b}\cdot\mathsf{c}\\
&\Rightarrow 0 = \nabla\left(\frac{\mathbf{p}}{p}\right) = \nabla\mathsf{c}\sin\varphi + \nabla\mathsf{b}\cos\varphi + \nabla\varphi(-\mathsf{b}\sin\varphi+\mathsf{c}\cos\varphi)\\
&\qquad\qquad = (\nabla\mathsf{c}+\nabla\mathsf{b}\cdot\mathsf{cb})\sin\varphi + \nabla\mathsf{b}\cdot(1-\mathsf{cc})\cos\varphi\\
&\Rightarrow \nabla\mathsf{c} = -\nabla\mathsf{b}\cdot(\mathsf{cb}+\mathsf{aa}\cot\varphi)\,.
\end{aligned} \tag{1.12}$$

in which $\nabla$ means differentiation with respect to $\mathbf{q}$ while keeping $\mathbf{p}$ constant, and we used that $1=\mathsf{aa}+\mathsf{bb}+\mathsf{cc}$ and that $\mathsf{b}$ is a unit vector, which implies $\nabla\mathsf{b}\cdot\mathsf{b} = \nabla\left(\frac{\mathsf{b}^2}{2}\right) = 0$.

The action of $\nabla$ on the vector $\mathsf{c}$ must be taken into account through Eq. (1.11) when computing the action of $\dot{\mathbf{z}}_\mathbf{q}$ or $(\mathbf{X}_{n+1})_\mathbf{q}$, but also when defining the components of $\dot{\mathbf{z}}$: writing $\dot f = \dot{\mathbf{z}}\cdot\partial_\mathbf{z} f$ in coordinates $\mathbf{z} := (\mathbf{q},p,\varphi,\mathsf{c})$ with the property (1.11) implies

$$\begin{aligned}
\dot{\mathbf{z}} &= \frac{eB}{m}\left[\begin{pmatrix}0\\0\\0\\1\end{pmatrix} + \frac{p\sin\varphi}{eB}\begin{pmatrix}\mathsf{b}\cot\varphi+\mathsf{c}\\0\\-(\mathsf{b}\cot\varphi+\mathsf{c})\cdot\nabla\mathsf{b}\cdot\mathsf{c}\\0\end{pmatrix}\right]\\
&= \dot{\mathbf{z}}_{-1}+\dot{\mathbf{z}}_0\,,\\
\text{with } \dot{\mathbf{z}}_{-1} &:= \begin{pmatrix}0\\0\\0\\\frac{eB}{m}\end{pmatrix},\\
\text{and } \dot{\mathbf{z}}_0 &:= \begin{pmatrix}\frac{\mathbf{p}}{m}\\0\\-\frac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\\0\end{pmatrix}.
\end{aligned} \tag{1.13}$$

Notice that the $\mathsf{c}$ component of $\dot{\mathbf{z}}_0$ is zero, because the term $-\frac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot(\mathsf{cb}+\mathsf{aa}\cot\varphi)$ in the dynamics of $\mathsf{c}$ comes from $\dot{\mathbf{z}}_\mathbf{q}\cdot\nabla f$ with Eq. (1.11).

## 1.2 The reduction algorithm

In this section, it is shown how the computation proceeds for the minimal guiding-center transformation. At lowest order $\varepsilon^{-1}$, the equation to be solved (1.5) writes

$$0 = \mathsf{osc}\,(\dot{\mathbf{z}}_{-1})_{\mathbf{q},\varphi}\,.$$

From the definition (1.3) of $\mathbf{z}_{-1}$, this is trivially verified, and it is actually a condition for the near-identity Lie-transform to isolate the fast time-scale, which is possible only because at lowest order, the motion concerns only the fast variable $\mathsf{c}$.

At order 0, the equation to solve (1.5) is

$$0 = \mathtt{osc}\left(\mathsf{L}_1\dot{\mathbf{z}}_{-1} + \dot{\mathbf{z}}_0\right)_{\mathbf{q},\varphi} .$$

The solution for the first-order change of coordinates $\mathbf{X}_1$ is given by Eq. (1.8) with the choice (1.9)

$$(\mathbf{X}_1)_{\mathbf{q},\varphi} = \tfrac{1}{\omega_L}\int d\theta\ \mathtt{osc}(\dot{\mathbf{z}}_0)_{\mathbf{q},\varphi} . \tag{1.14}$$

The spatial component gives the lowest-order Larmor radius, which is often identified with the Larmor radius itself

$$\begin{aligned}(\mathbf{X}_1)_{\mathbf{q}} &= \tfrac{m}{eB}\int d\theta\ \mathtt{osc}(\dot{\mathbf{z}}_0)_{\mathbf{q}}\\ &= \tfrac{p}{eB}\int d\theta\ \mathtt{osc}(\mathsf{b}\cos\varphi + \mathsf{c}\sin\varphi)\\ &= \tfrac{p}{eB}\int d\theta\ \mathsf{c}\sin\varphi = \tfrac{p\sin\varphi}{eB}\mathsf{a} = r_L\mathsf{a} ,\end{aligned} \tag{1.15}$$

where the action of $\mathtt{osc}$ and $\int d\theta$ can be computed trivially by using Eq. (1.2). One can avoid this trick and make the computation with purely intrinsic operators using the results of Chapter 10.

On another hand, the component $\varphi$ of Eq. (1.14) gives the expression of the first-order change for the coordinate $\varphi$

$$\begin{aligned}(\mathbf{X}_1)_{\varphi} &= \tfrac{m}{eB}\int d\theta\ \mathtt{osc}(\dot{\mathbf{z}}_0)_{\varphi}\\ &= -\tfrac{p}{eB}\int d\theta\ \mathtt{osc}((\mathsf{b}\cos\varphi + \mathsf{c}\sin\varphi)\cdot\nabla\mathsf{b}\cdot\mathsf{c})\\ &= -\tfrac{p}{eB}\int d\theta\ \left(\cos\varphi\ \mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{c} + 2\sin\varphi\nabla\mathsf{b} : \bar{\bar{\mathbf{a}}}_2\right)\\ &= -\tfrac{p\sin\varphi}{eB}\left(\cot\varphi\ \mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{a} - \tfrac{1}{2}\nabla\mathsf{b} : \bar{\bar{\mathbf{a}}}_1\right) ,\end{aligned} \tag{1.16}$$

with $\bar{\bar{\mathbf{a}}}_1 := -\frac{\mathsf{ac}+\mathsf{ca}}{2}$ and $\bar{\bar{\mathbf{a}}}_2 := \frac{\mathsf{cc}-\mathsf{aa}}{4}$ the standard dyadic tensors of guiding-center works [29, 86]. Eqs. (1.15)-(1.16) agree with the usual fluctuating first-order generator of the guiding-center reduction [29, 86, 88].

At order one, Eq. (1.5) reads

$$0 = \mathtt{osc}\left(\mathsf{L}_2\dot{\mathbf{z}}_{-1} + \tfrac{\mathsf{L}_1^2}{2}\dot{\mathbf{z}}_{-1} + \mathsf{L}_1\dot{\mathbf{z}}_0\right)_{\mathbf{q},\varphi} .$$

As usual, this equation is already solved by Eq. (1.8) with condition (1.9)

$$(\mathbf{X}_2)_{\mathbf{q},\varphi} = \tfrac{1}{\omega_L}\int d\theta\ \mathtt{osc}\left(\tfrac{\mathsf{L}_1^2}{2}\dot{\mathbf{z}}_{-1} + \mathsf{L}_1\dot{\mathbf{z}}_0\right)_{\mathbf{q},\varphi} . \tag{1.17}$$

All we have to do is to make the left-hand side of Eq. (1.17) explicit. This is completely algorithmic. The first step is to compute the Lie derivatives using the standard formula

$$\begin{aligned}\mathsf{L}_n\mathbf{w} &= \mathcal{L}_{\mathbf{X}_n}(\mathbf{w}_k\partial_k)\\ &= \Big((\mathbf{X}_n)_j\partial_j\mathbf{w}_k - \mathbf{w}_j\partial_j(\mathbf{X}_n)_k\Big)\partial_k + (\mathbf{X}_n)_j\mathbf{w}_m[\partial_j,\partial_m] ,\end{aligned} \tag{1.18}$$

for any vector field $\mathbf{w}$. Einstein convention is used and repeated indices are implicitly summed. The derivative operator $\partial_k$ corresponding to the gyro-angle variable $\mathsf{c}$ is $\mathbf{p}\cdot\mathsf{b}\times\partial_{\mathbf{p}} = -\mathsf{a}\cdot\partial_{\mathsf{c}}$, since it is the generator of Larmor gyration (see Chapter 10), and is equal to $\partial_\theta$.

In Eq. (1.18), the commutator of derivatives $[\partial_j, \partial_m]$ appears because the coordinate $\mathsf{c}$ is constrained, it is space-dependent, and the corresponding connection involves the pitch-angle. So, the following commutators are non-zero:

$$\begin{aligned}
[\nabla, -\mathsf{a}\cdot\partial_{\mathsf{c}}] &= \cot\varphi\ (\nabla\mathsf{b}\cdot\mathsf{c})\ (\mathsf{a}\cdot\partial_{\mathsf{c}})\,,\\
[\nabla, \partial_\varphi] &= -(1+\cot^2\varphi)(\nabla\mathsf{b}\cdot\mathsf{a})\ (\mathsf{a}\cdot\partial_{\mathsf{c}})\,,\\
[\nabla_i, \nabla_j] &= (1+\cot^2\varphi)\ \Big[(\nabla_i\mathsf{b}\cdot\mathsf{c})\ (\nabla_j\mathsf{b}\cdot\mathsf{a})\,,\\
&\qquad - (\nabla_j\mathsf{b}\cdot\mathsf{c})\ (\nabla_i\mathsf{b}\cdot\mathsf{a})\Big]\ (\mathsf{a}\cdot\partial_{\mathsf{c}})\,.
\end{aligned}$$

Eq. (1.18) is systematically applied to $\mathsf{L}_1\dot{\mathbf{z}}_0$, then to $\mathsf{L}_1\dot{\mathbf{z}}_{-1}$, and last to $\frac{\mathsf{L}_1}{2}(\mathsf{L}_1\dot{\mathbf{z}}_{-1})$, which appear in the right-hand side of (1.17). Although uncomplicated, computations must be very orderly to remain tractable since the products and the Leibniz rule generate many terms from a single expression such as $\frac{\mathsf{L}_1^2}{2}\dot{\mathbf{z}}_{-1}$ .

The second step is to perform the action of the gyro-integral $\int d\theta\ \mathtt{osc}$. It acts only on the variables $\mathsf{c}$ and $\mathsf{a}$, involved in expressions such as

$$\mathsf{a}\cdot\nabla B\ \mathsf{c}\cdot\nabla\mathsf{b}\cdot\mathsf{a} = (\mathsf{aca})_{ijk}\ (\nabla B\ \nabla\mathsf{b})_{ijk}\,,$$

upon which $\int d\theta\ \mathtt{osc}$ operates on the first tensor in the left-hand side by mixing the $\mathsf{c}$ and $\mathsf{a}$.

A way to perform the action of $\int d\theta\ \mathtt{osc}$ is to compute its action on the basic tensors $\mathsf{c}$, $\mathsf{cc}$, $\mathsf{ccc}$, etc. and to deduce its action on the other tensors by using cross products with the magnetic field $(\mathsf{b}\times)_{ij} = \varepsilon_{ikj}\mathsf{b}_k$; this last operator is usually denoted by $\mathsf{b}_{ij}$. For instance, $(\mathsf{aca})$ can be written

$$(\mathsf{aca})_{ijk} = \mathsf{b}_{im}\ \mathsf{b}_{kn}\ (\mathsf{ccc})_{mjn}\,. \tag{1.19}$$

Last, the action of $\int d\theta\ \mathtt{osc}$ on the elementary tensors $\mathsf{c}^{\otimes N}$ can be computed by using the intrinsic calculus introduced in Chapter 10, or by introducing a local fixed basis $(\mathbf{e}_1, \mathbf{e}_2)$ through the change of coordinate:

$$\begin{aligned}
\mathsf{c} &= -\sin\theta\mathbf{e}_1 - \cos\theta\mathbf{e}_2\,,\\
\mathsf{a} &:= \cos\theta\mathbf{e}_1 - \sin\theta\mathbf{e}_2\,,
\end{aligned}$$

as is standard in guiding-center reductions. With the fixed basis, the action of $\int d\theta\ \mathtt{osc}$ is trivial to compute: $\mathtt{osc}$ just cancels the zeroth Fourier component (gyro-average), and $\int d\theta$ is an easy integral. If $N$ is large, the computation by hand may be tedious but it remains trivial with a computer. Then one can come back to the initial basis $(\mathsf{c}, \mathsf{a})$ to get $\int d\theta\mathtt{osc}(\mathsf{c}^{\otimes N})$.

For instance, the lowest-orders formulae are

$$\begin{aligned}
\int d\theta\mathtt{osc}\ (\mathsf{c}) &= \mathsf{a}\,,\\
\int d\theta\mathtt{osc}\ (\mathsf{cc}) &= \frac{\mathsf{ca}+\mathsf{ac}}{4}\,,\\
\int d\theta\mathtt{osc}\ (\mathsf{ccc}) &= \tfrac{1}{3}\big[\mathsf{acc}+\mathsf{cac}+\mathsf{cca}+2\mathsf{aaa}\big]\,,\\
\int d\theta\mathtt{osc}\ (\mathsf{cccc}) &= \tfrac{1}{32}\Big[5(\mathsf{accc}+\mathsf{cacc}+\mathsf{ccac}+\mathsf{ccca})\\
&\qquad + 3\big(\mathsf{caaa}+\mathsf{acaa}+\mathsf{aaca}+\mathsf{aaac}\big)\Big]\,.
\end{aligned} \tag{1.20}$$

These are all we need to get the second-order reduction. Higher harmonics could be as easily dealt with. In practical computations, the relations given above can often be simplified by the symmetries of the tensor that $\mathsf{c}^{\otimes N}$ is contracted with.

Notice that $\int d\theta\mathtt{osc}$ is a linear operator which preserves the order in the fast variable: the total number of factors $\mathsf{c}$ and $\mathsf{a}$ in each term on the right-hand side of Eq. (1.20) is the same as on its

left-hand side. It is a left-inverse for $\partial_\theta$. If we restrict the operator $\partial_\theta$ to the set of tensors that are harmonics of order $n$ in the fast variable, it becomes just a matrix with finite dimension $2^n$. If $n = 2N+1$ is odd, the matrix is invertible, and is easily obtained (on a computer for instance) and gives the action of $\int d\theta \mathtt{osc}$. As regards even harmonics $n = 2N$, the kernel of $\partial_\theta$ is not zero, but it is a complementary space to the range of $\partial_\theta$, as is obvious in Fourier series, for instance. So, after identifying the kernel and the range of this operator, one gets an invertible matrix by restricting the operator to its range. Then, the action of the operator $\int d\theta \mathtt{osc}$ is just this invertible matrix on Range$(\partial_\theta)$ and zero on ker$(\partial_\theta)$. It is actually a very efficient way to perform the action of the operator $\int d\theta \mathtt{osc}$.

After using Eqs. (1.19) and (1.20) on each term of $\frac{\mathsf{L}_1^2}{2}\dot{\mathbf{z}}_{-1} + \mathsf{L}_1\dot{\mathbf{z}}_0$, recombining the various terms and simplifying the result with formula

$$\mathsf{b} \cdot \nabla B = -B \nabla \cdot \mathsf{b}\,, \tag{1.21}$$

coming from $\nabla \cdot \mathbf{B} = 0$, one obtains the second-order generator of the minimal guiding-center transformation $\mathbf{X}_2$, which is given by Eqs. (1.22)-(1.23) in the next section.

The result is somehow intricate, especially for $(\mathbf{X}_2)_\varphi$, but this was expected for a formula at second order, which even was not reached in usual derivations of guiding-center reductions. Actually, it emphasizes how efficient Lie-transforming the velocity vector field is, since it provides by a straightforward computation such a complicated result.

The explicit computation at order 2 illustrates the method to perform the derivation to arbitrary order. At each order in the Larmor radius expansion, the algorithm consists in writing down Eq. (1.8) and making all the terms explicit. Only two special operations are involved: the Lie-derivatives in the computation of $(\mathbf{N}_n)_{\mathbf{q},\varphi}$, given by Eq. (1.18), and the gyro-integral $\int d\theta\ \mathtt{osc}$ computed by (1.20); these are automatic operations.

If computations soon become tedious, it is only because the numbers of terms rapidly increases, as a result of products in Lie derivatives, of the Leibniz rule for the action of gradients, and of the gyro-integral operator. This is what generates the complicated formula (1.23) for $(\mathbf{X}_2)_\varphi$. But this is no trouble, since computations are completely algorithmic and very direct: Eq. (1.23) for $(\mathbf{X}_2)_\varphi$ was obtained with no trouble by hand, and computations at higher order can be very easily done by computer-assisted symbolic calculus.

The fast growth of the number of terms with the order of expansion raises the question of the convergence of the series, but in perturbation expansions, convergence is not the first issue, even if it is important to be addressed [9, 87]. In addition, a factorial growth is quite standard and convergence is generally obtained only with methods of accelerated convergence [10]. When convergence is not guaranteed or when non-convergence is proven, the asymptotic expansion, with its truncated perturbation expansions, all the same allows for a strong reduction of the effect of the fast time-scale in the dynamics of $(\bar{\mathbf{q}}, \bar{\varphi})$, as emphasized in [81].

From this point of view, a simplified derivation, such as the minimal guiding-center reduction, is an interesting way to control the iteration process and to obtain more information on the asymptotic behaviour.

## 1.3 The result to second order

By applying the procedure introduced in last section, the second-order generator of the minimal guiding-center transformation is obtained

$$
\begin{aligned}
(\mathbf{X}_2)_{\mathbf{q}} := \left(\tfrac{p\sin\varphi}{eB}\right)^2 \Bigg[ & \mathsf{b}\left(-2\phi\bar{\mathsf{c}}\mathsf{b}'\mathsf{b} + \frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{8}\right) \\
& + \phi^2 \frac{\mathsf{c}\bar{\mathsf{c}} - \mathsf{a}\bar{\mathsf{a}}}{4}\mathsf{b}'\mathsf{b} + \tfrac{\phi}{8}\big(4\mathsf{c}\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} + 7\mathsf{a}\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - 9\mathsf{a}\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}\big)\Bigg] \,, \qquad (1.22)
\end{aligned}
$$

$$
\begin{aligned}
(\mathbf{X}_2)_{\varphi} := \left(\tfrac{p\sin\varphi}{eB}\right)^2 \Bigg[ & \tfrac{B'}{24B}\big(4\mathsf{c}\bar{\mathsf{c}}\mathsf{b}'\mathsf{c} - 4\mathsf{c}\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} + 5\mathsf{a}\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + 5\mathsf{a}\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}\big) \\
& + \phi^3\left\{\frac{\bar{\mathsf{a}}\bar{\mathsf{a}} - \bar{\mathsf{c}}\bar{\mathsf{c}}}{4}\mathsf{b}'\mathsf{b}\mathsf{b}'\mathsf{b}\right\} \qquad (1.23) \\
& + \phi^2\left\{-\bar{\mathsf{c}}\mathsf{b}''\mathsf{b}\mathsf{b} + \Big(-\bar{\mathsf{c}}\mathsf{b}'\frac{20\mathsf{c}\bar{\mathsf{c}} + 9\mathsf{a}\bar{\mathsf{a}}}{8} + \bar{\mathsf{a}}\mathsf{b}'\frac{7\mathsf{c}\bar{\mathsf{a}} - 16\mathsf{a}\bar{\mathsf{c}}}{8}\Big)\mathsf{b}'\mathsf{b}\right\} \\
& + \phi^1\left\{\frac{\bar{\mathsf{a}}\mathsf{b}''\mathsf{a} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{c}}{8}\mathsf{b} + (\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{c})\frac{9\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} + 7\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{32}\right. \\
& \qquad \left. - \frac{\bar{\mathsf{a}}\bar{\mathsf{a}} - \bar{\mathsf{c}}\bar{\mathsf{c}}}{8}\mathsf{b}'\mathsf{b}\mathsf{b}'\mathsf{b} + (\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c})\frac{3\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - 5\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{32}\right\} \\
& + \tfrac{\phi^0}{24}\Big\{8\bar{\mathsf{c}}\mathsf{b}''\mathsf{c}\mathsf{c} - 5\bar{\mathsf{c}}\mathsf{b}''\mathsf{a}\mathsf{a} + 11\bar{\mathsf{a}}\mathsf{b}''\mathsf{c}\mathsf{a} \\
& \quad + \big(-16\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}\bar{\mathsf{c}} + 10\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}\bar{\mathsf{a}} - 11\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}\bar{\mathsf{a}} - 11\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}\bar{\mathsf{c}}\big)\mathsf{b}'\mathsf{b}\Big\}\Bigg] \,,
\end{aligned}
$$

where $\phi$ is a shorthand for $\cot\varphi$. To make expressions easier to read, we used the primed notation for spatial gradients, and the over-bar over a vector $\mathsf{c}$ or $\mathsf{a}$ means matrix transpose, e.g. $\bar{\mathsf{a}}\mathsf{b}''\mathsf{a}\mathsf{b}'\mathsf{c}$ means $\mathsf{c}\cdot(\nabla\mathsf{b})\cdot[(\mathsf{a}\cdot\nabla)(\nabla\mathsf{b})]\cdot\mathsf{a}$. This notation is close to Littlejohn's notations ($\mathsf{abc}$) in [88], but it is more suited for higher-order derivatives and more explicit, as in Chapter 10.

In Eq. (1.22), the first line is the $\mathsf{b}$-component, parallel to the magnetic field, and the following two lines are perpendicular to $\mathsf{b}$. In Eq. (1.23), the first line contains the terms depending on $\nabla B$ and the following lines are organized as a polynomial of the pitch-angle, or rather its cotangent $\phi$. Note that the terms depending on $\mathsf{b}\cdot\nabla B$ do not appear in the first line, because they are rewritten using Eq. (1.21). The expressions involved in such a second-order result can be written in many different but equivalent ways, as already noticed by Northrop and Rome, and a rule for standardizing them is needed for the derivation to be efficient.

If there are many terms, this is especially because each term appears several times, with permutations of $\mathsf{a}$ and $\mathsf{c}$; these permutations are often condensed into one single tensor. For instance, the last line of Eq. (1.23) can be written just

$$\nabla\mathsf{b} : \Pi\cdot(\mathsf{b}'\mathsf{b}),$$

by defining the triadic tensor

$$\Pi := -16\mathsf{ccc} + 10\mathsf{aca} - 11\mathsf{caa} - 11\mathsf{aac}$$

Then the number of terms is strongly reduced, each order in $\phi$ has one or two terms, which shows that the result is not that complicated actually. However, we chose to avoid introducing such intermediate quantities, since they make formulae shorter but less explicit.

The second-order vector field gives the minimal guiding-center change of coordinates to second order

$$\bar{\mathbf{z}} = \left[1 - (\mathsf{L}_1) - (\mathsf{L}_2) + \left(\tfrac{\mathsf{L}_1^2}{2}\right) + O(\varepsilon^3)\right]\mathbf{z}\,,$$

where the vector fields $\mathbf{X}_1$ and $\mathbf{X}_2$ generating the Lie-transforms $\mathsf{L}_1$ and $\mathsf{L}_2$ are given by Eqs. (1.15), (1.16), (1.22) and (1.23).

So, to second order, the Larmor radius is

$$
\begin{aligned}
\mathbf{r}_L =& (\mathbf{z}-\bar{\mathbf{z}})_{\mathbf{q}} = (\mathbf{X}_1)_{\mathbf{q}} + (\mathbf{X}_2)_{\mathbf{q}} - \left(\frac{(\mathbf{X}_1)_{\mathbf{z}} \cdot \partial_{\mathbf{z}}(\mathbf{X}_1)_{\mathbf{q}}}{2}\right) + O(\varepsilon^3) \\
=& \tfrac{p\sin\varphi}{eB}\mathsf{a} \\
&+ \left(\tfrac{p\sin\varphi}{eB}\right)^2 \Bigg[\mathsf{b}\left(-2\phi\bar{\mathsf{c}}\mathsf{b}'\mathsf{b} + \frac{5\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{8}\right) \\
&\qquad + \mathsf{a}\tfrac{B'\mathsf{a}}{2B} + \phi^2\tfrac{\mathsf{b}'\mathsf{b}}{4} + \phi\mathsf{a}\big(\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}\big)\Bigg].
\end{aligned}
\tag{1.24}
$$

And to second order, the reduced pitch-angle is

$$
\begin{aligned}
\bar{\varphi} =& \left[1 - (\mathsf{L}_1) - (\mathsf{L}_2) + \left(\tfrac{\mathsf{L}_1^2}{2}\right) + O(\varepsilon^3)\right]\varphi \\
=& \varphi - \tfrac{p\sin\varphi}{eB}\left[-\phi\bar{\mathsf{a}}\mathsf{b}'\mathsf{b} - \frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4}\right] \\
&- \left(\tfrac{p\sin\varphi}{eB}\right)^2 \Bigg[\tfrac{B'}{12B}\Big\{-6\phi\mathsf{a}\bar{\mathsf{a}}\mathsf{b}'\mathsf{b} + 2\mathsf{c}\bar{\mathsf{c}}\mathsf{b}'\mathsf{c} \\
&\qquad - 2\mathsf{c}\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} + \mathsf{a}\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \mathsf{a}\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}\Big\} \\
&+ \phi^3\left\{\frac{\bar{\mathsf{a}}\bar{\mathsf{a}} - \bar{\mathsf{c}}\bar{\mathsf{c}}}{4}\mathsf{b}'\mathsf{b}\mathsf{b}'\mathsf{b}\right\} \\
&+ \phi^2\left\{-\bar{\mathsf{c}}\mathsf{b}''\mathsf{b}\mathsf{b} + \left(-5\bar{\mathsf{c}}\mathsf{b}'\frac{2\mathsf{c}\bar{\mathsf{c}} + \mathsf{a}\bar{\mathsf{a}}}{4} + 3\bar{\mathsf{a}}\mathsf{b}'\frac{\mathsf{c}\bar{\mathsf{a}} - 2\mathsf{a}\bar{\mathsf{c}}}{4}\right)\mathsf{b}'\mathsf{b}\right\} \\
&+ \phi^1\left\{\frac{5\bar{\mathsf{a}}(\mathsf{b}'\mathsf{b})'\mathsf{a} - \bar{\mathsf{c}}(\mathsf{b}'\mathsf{b})'\mathsf{c}}{8} + \frac{3\bar{\mathsf{a}}\bar{\mathsf{a}} + \bar{\mathsf{c}}\bar{\mathsf{c}}}{8}\mathsf{b}'\mathsf{b}\mathsf{b}'\mathsf{b}\right. \\
&\qquad \left. - \frac{3(\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{c})^2}{32} + (\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c})\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - 3\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{16}\right\} \\
&+ \tfrac{\phi^0}{12}\Big\{4\bar{\mathsf{c}}\mathsf{b}''\mathsf{c}\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{a}\mathsf{a} + 7\bar{\mathsf{a}}\mathsf{b}''\mathsf{c}\mathsf{a} \\
&\qquad + \big(-8\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}\bar{\mathsf{c}} + 5\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}\bar{\mathsf{a}} - 4\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}\bar{\mathsf{a}} - 7\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}\bar{\mathsf{c}}\big)\mathsf{b}'\mathsf{b}\Big\}\Bigg].
\end{aligned}
\tag{1.25}
$$

The ordering in the Larmor radius is obvious. The first line corresponds to the zeroth- and first-order terms, and all the following lines are the second-order terms, which are organized as in Eq. (1.23).

We wrote all these formulae exactly as they are yielded by the procedure, because it illustrates both the mechanism of the derivation and the structure of the resulting change of coordinates.

Indeed, all the above formulae rely only on a very restricted alphabet of entities: $B$, $\mathsf{b}$, $\mathsf{c}$, $\mathsf{a}$, $\nabla$ and the variable $\varphi$, or more precisely $\phi$, if we discard the $\sin\varphi$ occurring in the pre-factor (Larmor radius) and the quantities $p$ and $e$, which are mute parameters in the derivation.

Each formula is a series in the Larmor radius $r_L = \frac{p\sin\varphi}{eB}$, or more precisely in the ordering parameter of guiding-center reduction $\varepsilon = r_L\nabla$, i.e. the magnetic inhomogeneity at the scale of the Larmor radius. This dimensionless parameter is not just a number, it is to be understood in the sense that the only dimensional quantities involved are the Larmor radius, appearing as a pre-factor to the power given by the expansion order, and the gradients (acting on the magnetic field $B$ and $\mathsf{b}$), with the same order as the expansion order.

At each order in $\varepsilon$, formulae are polynomials in each of the variables (except $B$ whose disposition obeys trivial dimensional rules). This fact is useful to make the derivation easier: finally, the iteration at each order consists in one single formula (1.8) with two elementary operators (derivatives and gyro-integral), applied on terms composed of very few elementary entities, and those terms are just polynomials.

As the order in the small parameter $\varepsilon$ grows up, the polynomial increases its order in each of the variables. From the iteration mechanism, a rough estimate shows that the $n$-th-order generator $\mathbf{X}_n$ should be a polynomial of order at most $\phi^{2n-1}\mathsf{b}^{3n-1}\mathsf{c}^{2n}$, in addition to being a monomial in $r_L^n\nabla^n$.

Such features in the reduction transformation are useful to consider when the asymptotic behaviour is addressed. For instance, a hypothetical convergence condition would clearly involve the expected condition on the magnetic inhomogeneity at the scale of the Larmor radius, but the role of $\phi$ suggests conditions on the pitch-angle as well. This means that in the guiding-center reduction, the direction of the particle velocity must not be too close to the direction of the magnetic field. This is in complete agreement with the physical intuition, but the polynomial behaviour can help make this intuitive statement more precise.

Another interesting feature is well emphasized by the formulae as they are written: there is a link of parity between the order in $\phi$, the order in $r_L$ and the fast-angle harmonic (which corresponds to the order in $\mathsf{c}$ or $\mathsf{a}$). All non-zero terms of order $r_L^i\phi^j$ are harmonics of parity $(-1)^{i+j}$ in $\mathsf{c}$ for $\mathbf{X}_{\mathbf{q}}$, and harmonics of parity $(-1)^{i+j+1}$ in $\mathsf{c}$ or $\mathsf{a}$ for $\mathbf{X}_\varphi$. This parity relation could already be observed in the first-order results (1.15)-(1.16), although it was not so obvious, because few terms were present.

It seems that the polynomiality in $\phi$ and the parity relation were first noticed in [62] (which is the work reported in Chapter 10), where they were observed in the derivation of the magnetic moment series, and originated from the structure of the operator to be inverted for the secular equation. Here, they are obtained in the guiding-center reduction, and appear as originating directly from the equations of motion (1.3) or (1.13), and from the action of derivatives, especially $\partial_\varphi\phi$ and $\nabla\mathsf{c}$.

This polynomiality is related to the coordinate $\phi$. Previous works used the parallel velocity $v_\parallel$ as a coordinate instead of $\phi$. Their results were not polynomials (see e.g. [8, 29, 88]), but when expressed with the coordinate $\phi$, they become also polynomials.

The second-order results can be expressed using well-known quantities in guiding-center works. For instance, in the term of order $\phi^1$ in Eq. (1.25), $\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}$ and $\bar{\mathsf{c}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{a}$ can be recognized, commonly written $-2\nabla\mathsf{b} : \bar{\bar{\mathbf{a}}}_1$ and $4\nabla\mathsf{b} : \bar{\bar{\mathbf{a}}}_2$ where the dyadic tensors $\bar{\bar{\mathbf{a}}}_1 = -\frac{\mathsf{ac}+\mathsf{ca}}{2}$ and $\bar{\bar{\mathbf{a}}}_2 = \frac{\mathsf{cc}-\mathsf{aa}}{4}$ are well-known in guiding-center works [29, 86], and they were already met in first-order results (1.16).

Also, in the last line of Eq. (1.25), each of the terms is harmonic of order 3 in the fast angle $\mathsf{c}$ or $\mathsf{a}$. The sum can be rewritten

$$\big(-3\tau_m\bar{\mathsf{a}} + \bar{\mathsf{a}}\mathsf{b}'(\mathsf{b}\times) - 8\nabla\cdot\mathsf{b}\;\bar{\mathsf{c}} + 2\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}\;\bar{\mathsf{a}}\big)\mathsf{b}'\mathsf{b}\,, \tag{1.26}$$

where three out of four terms have been combined to obtain a harmonic of order 1 in the gyro-angle, which is the maximum we can do. Then, in the coefficients, the divergence of magnetic field lines $\nabla\cdot\mathsf{b} = \bar{\mathsf{c}}\mathsf{b}'\mathsf{c} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{a}$ is recovered, as well as their twist

$$\tau_m := \mathsf{b}\cdot\nabla\times\mathsf{b} = (\mathsf{b}\times\nabla)\cdot\mathsf{b} = \bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a}\,,$$

which quantities are most used in guiding-center theory.

Such an expression as (1.26) is useful to get a physical intuition of the terms involved, but it does not emphasize as much the mechanism of the derivation nor the polynomial structure of the results. Furthermore, it is not unique, e.g. it could equally be written

$$\big(-5\tau_m\bar{\mathsf{a}} - \bar{\mathsf{a}}\mathsf{b}'(\mathsf{b}\times) - 6\nabla\cdot\mathsf{b}\;\bar{\mathsf{c}} - 2\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}\;\bar{\mathsf{c}}\big)\mathsf{b}'\mathsf{b}\,,$$

or even

$$\big(-13\tau_m\bar{\mathsf{a}} - 9\bar{\mathsf{a}}\mathsf{b}'(\mathsf{b}\times) - 8\bar{\mathsf{c}}\mathsf{b}' + 2\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}\bar{\mathsf{c}}\big)\mathsf{b}'\mathsf{b}\,,$$

or (at least) nine other equivalent ways of writing this expression. In all of them, three out of four terms have been combined to obtain harmonics of order 1 in the gyro-angle, and only the last term remains a harmonic of order 3.

So, we preferred to give the expressions as they come out from the procedure, decomposed in elementary terms in a unique, standardized way, as is needed for an algorithmic procedure, especially when it can generate very many terms.

Anyway, it is quite common in perturbation theory and well known in guiding-center theory that when the order increases, formulae become messy and complicated, and the physical interpretation of each term loses some of its relevance [123]. The main point is to standardize the derivation

and the involved expressions, to make things tractable and as clear as possible.

Compared to the literature, our fluctuating first-order results are exactly identical to the usual results of guiding-center derivations, e.g. in [8,29,86,123]. This is because this part of the reduction is unique for all guiding-center reductions, since the arbitrariness is only in the average transformation generator. Our second-order results are almost identical to previous results. For instance, the Larmor radius (1.24) is exactly identical to the standard result of [86], except for the term $\phi^2$, which was absent from Littlejohn's result. As for the pitch-angle, the first-order term in (1.25) is exactly the same as in [86], except the averaged term, which is absent from our result.

A difference with previous results is not a surprise because here the minimal guiding-center reduction is considered, and the averaged part of the transformation generator $\mathbf{X}$ have been set to zero. It is just an effect of the guiding-center reduction non-uniqueness, which was well emphasized in [123] and [23]: even between the classical derivations by Lie-transforming the phase-space Lagrangian, such as [23, 86, 88, 128], various changes can be identified in the choices adopted for the second-order reduction.

In previous publications, the second-order pitch-angle generator $(\mathbf{X}_2)_\varphi$ was not computed. The works aiming at averaging the equations of motion were not able to reach the second-order transformation [8, 123]. On another hand, the works using a Lie-transform of the Lagrangian, although more efficient, required quite a lot of algebra to get $(\mathbf{X}_2)_\varphi$ and did not reach the full second-order transformation [29, 86]. Even the recent work [128] explicitly chose not to compute it, even though it aimed at improving the second-order terms and already made quite a lot of computations. Similarly, Ref. [23] uses a special property between the second- and first-order terms to identify the second-order reduced dynamics without computing the second-order transformation $\mathbf{X}_2$.

These difficulties can be partially explained because when Lie-transforming the Lagrangian, there is some mixing between the orders in $\varepsilon$ of the various quantities [29, 88]. Thus, some components of the first-order transformation generator $\mathbf{X}_1$ are identified at the first-order analysis of the reduced Lagrangian $\bar{\Gamma}_1$; other components of $\mathbf{X}_1$ are determined at second-order analysis $\bar{\Gamma}_2$, at the same time as some of the components of $\mathbf{X}_2$; and the last components of $\mathbf{X}_1$ are determined at third order $\bar{\Gamma}_3$, at the same time as some components of $\mathbf{X}_2$ and of $\mathbf{X}_3$. It clearly makes the scheme more involved.

This is very different from what happens when Lie-transforming the equations of motion, where there is no mixing between the order in $\varepsilon$ of the various quantities, and the non-trivial expression for $(\mathbf{X}_2)_\varphi$ was obtained by a direct derivation to order two with very limited algebra.

Let us turn now to the guiding-center equations of motion. Eqs. (1.4) and (1.5) show that to second order, the drift equations are given by

$$\begin{aligned}\dot{\bar{\mathbf{z}}} &= e^{\mathsf{L}}\dot{\mathbf{z}} = \mathcal{R}\left(e^{\mathsf{L}}\dot{\mathbf{z}}\right)\\ &= \mathcal{R}\left[1 + (\mathsf{L}_1) + (\mathsf{L}_2) + \left(\tfrac{\mathsf{L}_1^2}{2}\right) + O(\varepsilon^3)\right]\dot{\bar{\mathbf{z}}}\,. \end{aligned} \tag{1.27}$$

This is just the zero Fourier component of $e^{\mathsf{L}}\dot{\mathbf{z}}$, which was put to zero when computing the action of $\mathtt{osc}$ in (1.4). So, this term was already computed in the derivation.

In this sense as well, Lie-transforming the equations of motion is more straightforward because one actually directly derives the reduced motion, whereas when Lie-transforming the Lagrangian, one derives the reduced Lagrangian; the reduced equations of motion must be obtained in a second step, for instance by Eq. (1.27), which involves some algebra because of the lengthy expressions at order 2.

Also, it should be noticed that Eq. (1.27), together with Eq. (1.5), implies that the $n$-th order step of the algorithm provides the $n$-th order transformation, but the $(n-1)$-th order reduced motion. This lag originates from the fact that the fast lowest-order motion is of order $-1$ compared to the main time scale of the dynamics.

So, the reduced equations of motion are obtained here with no additional computation as

$$
\begin{aligned}
(\dot{\bar{\mathbf{z}}})_{\mathbf{q}} :=& \tfrac{p\sin\varphi}{m}\Big[\mathsf{b}\phi + \tfrac{p\sin\varphi}{eB}\Big(\tfrac{B'}{B}\tfrac{\mathsf{c}\mathsf{a}-\mathsf{a}\mathsf{c}}{2} \\
&\qquad + \mathsf{b}\tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2} + \phi^2\big(\mathsf{a}\bar{\mathsf{c}}-\mathsf{c}\bar{\mathsf{a}}\big)\mathsf{b}'\mathsf{b}\Big)\Big] \qquad (1.28)\\
=& \tfrac{p\sin\varphi}{m}\Big[\mathsf{b}\phi + r_L\Big(\tfrac{\mathsf{b}\times\nabla B}{2B} + \mathsf{b}\tfrac{\mathsf{b}\cdot\nabla\times\mathsf{b}}{2} + \phi^2\mathsf{b}\times\mathsf{b}'\mathsf{b}\Big)\Big],\\
(\dot{\bar{\mathbf{z}}})_{\varphi} :=& \tfrac{p\sin\varphi}{m}\Big[-\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}+\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2} \\
&\qquad + \tfrac{p\sin\varphi}{eB}\phi\Big(\tfrac{B'}{B}\tfrac{\mathsf{a}\bar{\mathsf{c}}-\mathsf{c}\bar{\mathsf{a}}}{2}\mathsf{b}'\mathsf{b} + \tfrac{\bar{\mathsf{a}}(\mathsf{b}'\mathsf{b})'\mathsf{c}-\bar{\mathsf{c}}(\mathsf{b}'\mathsf{b})'\mathsf{a}}{2}\Big)\Big] \qquad (1.29)\\
=& \tfrac{p\sin\varphi}{m}\Big[-\tfrac{\nabla\cdot\mathsf{b}}{2} + r_L\phi\Big(\tfrac{B'}{2B}\mathsf{b}\times\mathsf{b}'\mathsf{b} + \tfrac{\mathsf{b}\cdot\nabla\times(\mathsf{b}'\mathsf{b})}{2}\Big)\Big].
\end{aligned}
$$

For comparisons with the literature, remind that $\varphi$ is the pitch-angle, and $\phi$ is its cotangent, so that $p\sin\varphi = \|\mathbf{p}_\perp\|$ is the scalar perpendicular momentum, whereas $p\phi\sin\varphi = p_\parallel$ is the parallel momentum.

In these equations, we first gave the expression provided by the algorithm and then rewrote it using physical quantities, in order to show that the equations are indeed gyro-averages, and do not depend on the fast variables $\mathsf{c}$ and $\mathsf{a}$.

The right-hand side of Eqs. (1.28)-(1.29) is expressed in the reduced variables (e.g. the field is evaluated at the guiding-center position $\bar{\mathbf{q}}$ and the pitch-angle is actually the reduced pitch-angle $\bar{\varphi}$), but we dropped the bars for simplicity. Notice that the right-hand side of Eqs. (1.28)-(1.29) is expressed in the initial variables. It is why Eqs. (1.22)-(1.25) indeed define a coordinate transformation, whereas Eqs. (1.28)-(1.29) indeed define the guiding-center dynamics [86].

In results (1.28)-(1.29), the lowest-order term is again exactly identical to the usual guiding-center reductions, and the second-order drifts are almost identical. The guiding-center dynamics (1.28) contains exactly the four expected terms [8, 123]: the parallel motion along the magnetic field lines, the grad-B drift, the curvature drift, and the Baños drift. But the Baños term $\mathsf{b}\frac{\mathsf{b}\cdot\nabla\times\mathsf{b}}{2}$ was not in the reduced dynamics of [86]. This is again because the minimal reduction does not aim at minimizing the average dynamics, but at minimizing the change of variables. To make it short, it can be said that the reduction of [86] puts the Baños term into the change of variables $\mathbf{X}$ whereas the minimal reduction lets it in the equations of motion. This difference is mentioned in [8, 123] and further explored in the next section.

## 1.4 On non-minimal guiding-center reductions

In case there is an electric field $\mathbf{E}$ in addition to the strong magnetic field, the initial equations of motion (1.13) become

$$
\dot{\mathbf{z}} = \frac{eB}{m}\left[\begin{pmatrix}0\\0\\0\\1\end{pmatrix} + \frac{p\sin\varphi}{eB}\begin{pmatrix}\mathsf{b}\cot\varphi+\mathsf{c}\\0\\-(\mathsf{b}\cot\varphi+\mathsf{c})\cdot\nabla\mathsf{b}\cdot\mathsf{c}\\0\end{pmatrix} + \frac{m}{Bp\sin\varphi}\begin{pmatrix}0\\\mathbf{E}\cdot\mathbf{p}\sin\varphi\\\mathbf{E}\cdot\Big(\cos\varphi\,\frac{\mathbf{p}}{p}-\mathsf{b}\Big)\\(\mathbf{E}\cdot\mathsf{a})\mathsf{a}\end{pmatrix}\right],
$$

As for the variable $\theta$, its dynamics becomes

$$
\dot{\theta} = \frac{eB}{m} + \cot\varphi\frac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{a} + \frac{\mathbf{p}}{m}\cdot\nabla\mathbf{e}_1\cdot\mathbf{e}_2 - \frac{e\mathbf{E}\cdot\mathsf{a}}{p\,\sin\varphi}.
$$

In these time evolutions, the two common small parameters of guiding-center theory appear: in strong magnetic field, only one term implies fast dynamics, it is of order $\omega_L$; among the other terms, the ones that do not depend on the electric field are of order $\frac{p\nabla}{m}$, and the ones that do depend on $\mathbf{E}$ are of order $\frac{e\,\mathbf{E}}{p\,\sin\varphi}$. Thus, two ordering parameters are now involved, the magnetic parameter $\varepsilon := \frac{p\nabla}{\omega_L m} = \frac{p\nabla}{eB}$ is the same as in the case of a pure magnetic field; the second parameter

$\varepsilon_E := \frac{e\ \mathbf{E}}{\omega_L p\ \sin\varphi} = \frac{E}{v_\perp B}$ is induced by the electric field. It is the electric force over the magnetic force, and is assumed to be small in guiding-center theories [17, 29, 128].

Under this assumption, the procedure for the minimal guiding-center reduction is unchanged; the presence of an electric field only causes some additional contributions in the term $\mathbf{N}_n$ of Eq. (1.8), and makes the transformation include one more coordinate: as the norm of the particle momentum is no more conserved, this coordinate must be changed exactly in the same way as the coordinates $(\mathbf{q}, \varphi)$ in the previous sections. Notice that this minimal guiding-center reduction will imply $\bar{p}$ to have slow dynamics, but not to be a constant of motion. We turn to this point in a few lines.

In the previous sections, the minimal guiding-center reduction appeared as a natural choice when Lie-transforming the velocity vector field. Other possibilities are available.

For instance, one can think of treating the gyro-angle $\theta$ as the other coordinates, and changing coordinate $\theta \to \bar{\theta}$ to gyro-average its dynamics as well. Thus, all of the reduced dynamics would involve only slow variables. Especially, the reduced gyro-angle $\bar{\theta}$ would have fast dynamics, but would depend only on the slow dynamics, since it would be independent of $\bar{\theta}$ itself: $\dot{\bar{\theta}} = \dot{\bar{\theta}}(\bar{\mathbf{q}}, \bar{\varphi}, p)$. So, once the motion of the slow variables $(\bar{\mathbf{q}}(t), \bar{\varphi}(t))$ is known, the fast dynamics would be trivial to integrate:

$$\bar{\theta}(t) = \bar{\theta}(t_0) + \int_{t_0}^{t} dt\ \dot{\bar{\theta}}\Big(\bar{\mathbf{q}}(t), \bar{\varphi}(t), p\Big)\,.$$

The corresponding requirement relies on Eq. (1.4) written for the gyro-angle component

$$0 = \mathtt{osc}\left(e^{\mathsf{L}_1+\mathsf{L}_2+\dots}\dot{\mathbf{z}}\right)_\theta\,, \tag{1.30}$$

which is to be solved for $\mathsf{L}_\theta$ (or rather $\mathbf{X}_\theta$). The operator to be inverted is exactly the same as in the previous sections, since

$$\begin{aligned}
&-\mathtt{osc}(\mathsf{L}_{n+1}(\dot{\mathbf{z}}_{-1}))_\theta \\
&\qquad = -\mathtt{osc}\Big((\mathbf{X}_{n+1})_i\partial_i(\dot{\mathbf{z}}_{-1})_\theta - (\dot{\mathbf{z}}_{-1})_\theta\partial_\theta(\mathbf{X}_{n+1})_\theta\Big) \\
&\qquad = \tfrac{eB}{m}\partial_\theta(\mathbf{X}_{n+1})_\theta + o.t.\,,
\end{aligned} \tag{1.31}$$

where *o.t.* means other terms that are already known: they do not contain the unknown $(\mathbf{X}_{n+1})_\theta$ because $(\dot{\mathbf{z}}_{-1})_\theta$ does not depend on $\theta$. So, just as in the previous sections, a solution of Eq. (1.30) can be found at arbitrary order in the small parameter, and the dynamics of $\theta$ can be averaged as well as the dynamics of $(\mathbf{q}, \varphi)$, as is generally done in guiding-center reductions [8, 29, 86].

Notice that the averaging process for the reduced coordinates $(\mathbf{q}, \varphi)$ is possible without having to deal with resonances [57, 117] because there is no small divisor since the fast part of the dynamics involves only one angle $\theta$, as illustrated by Eq. (1.8). This is a rather broad property. Now, for the dynamics of $\theta$, the averaging process is possible only because the lowest-order term in the motion does not depend on $\theta$ itself, as is illustrated by Eq. (1.31). This is a more specific property.

For instance, after the guiding-center reduction has averaged the motion over the gyro-angle, the bounce-angle can be used to bounce-average the motion [16, 29]. For this reduction, the procedure of section 1.1 can be applied to remove the bounce time-scale from the dynamics of the three other components, but averaging the dynamics of the bounce-angle coordinate $\theta_b$ as well is guaranteed only if the lowest-order bounce-angle dynamics is independent of the bounce-angle, which is a condition over the definition of $\theta_b$.

A second natural requirement concerns the average part of the change of variables $\mathtt{avg}(\mathbf{X})$. In the minimal derivation, it is trivially put to zero. In the case of a non minimal reduction, it appears as an additional freedom in the reduction process, which allows for additional requirements, to be chosen.

For instance, the transformation being a pure gyro-fluctuation $\mathtt{avg}(\bar{\mathbf{z}} - \mathbf{z}) = 0$ would be interesting: once the effects of the fast coordinate averaged, then the transformation would become zero; so, the reduced coordinates would remain in some sense close to the initial ones.

In the minimal guiding-center reduction, this result is almost obtained, since the transformation generator $\mathbf{X}$ is already a pure gyro-fluctuation, but the whole transformation is not a pure fluctuation, because the non-linear terms in $\mathbf{X}$ have non-zero gyro-averages (see, e.g. Eqs. (1.24)-(1.25)).

This can be corrected by defining a non-zero average $\mathtt{avg}(\mathbf{X})$ to cancel the average of the non-linear terms. Indeed, the requirement

$$0 = \mathtt{avg}\Big(\bar{\mathbf{z}} - \mathbf{z}\Big) = \mathtt{avg}\left(e^{-\mathsf{L}}\mathbf{z} - \mathbf{z}\right)$$

implies

$$\mathtt{avg}\Big(\mathsf{L}\mathbf{z}\Big) = \mathtt{avg}\left(e^{-\mathsf{L}}\mathbf{z} - (1-\mathsf{L})\mathbf{z}\right) .$$

Now, in the equation at order $n$

$$\begin{aligned}\mathtt{avg}\Big(\mathbf{X}_n\Big) &= \mathtt{avg}\Big(\mathsf{L}\mathbf{z}\Big)_n = \mathtt{avg}\Big(e^{-\mathsf{L}}\mathbf{z} - (1-\mathsf{L})\mathbf{z}\Big)_n \\ &= \mathtt{avg}\left(\sum_{i=2}^{\infty} \tfrac{(-\mathsf{L})^i}{i!}\mathbf{z}\right)_n \\ &= \mathtt{avg}\left(\tfrac{\mathsf{L}_{n-1}\mathsf{L}_1 + \mathsf{L}_{n-2}\mathsf{L}_2 + \dots}{2}\mathbf{z} - \tfrac{\mathsf{L}_{n-2}\mathsf{L}_1\mathsf{L}_1 + \dots}{3!}\mathbf{z} + \dots\right) ,\end{aligned} \tag{1.32}$$

the right-hand side depends only on generators of orders lower than $n$. So Eq. (1.32) is no more an equation, but just a definition of $\mathtt{avg}(\mathbf{X})$. Up to second order, Eq. (1.32) writes

$$\begin{aligned}&\mathtt{avg}(\mathbf{X}_1) := 0 , \\ &\mathtt{avg}(\mathbf{X}_2) := \mathtt{avg}\left(\tfrac{\mathsf{L}_1^2}{2}\mathbf{z}\right) .\end{aligned}$$

The right-hand side was already computed in the minimal derivation, so that the additional requirement does not imply an additional computation.

Another possible requirement, commonly considered, is to include the magnetic moment $\bar{\mu}$ among the reduced coordinates as a conserved quantity instead of $p$ (see Eq. (2.22)). Indeed, the basic conserved quantity $p$ is not an adiabatic invariant in general when an electric field is present, whereas $\mu$ is so [122]. The requirement is to put to zero not only the fluctuating part of the reduced motion $\mathtt{osc}(\dot{\bar{\mathbf{z}}})_\mu$, as in Eq. (1.4), but its averaged part as well $\mathtt{avg}(\dot{\bar{\mathbf{z}}})_\mu = 0$, which simplifies the reduced dynamics in a drastic way.

For the other coordinates, the same requirement cannot be asked, because there is no constant of motion independent of $p$ and $\bar{\mu}$, since otherwise the Hamiltonian motion with 3 degrees of freedom and 3 independent constants of motion would be integrable. For other coordinates, one cannot ask so strong a simplification of the reduced dynamics $\mathtt{avg}(\dot{\bar{\mathbf{z}}})$ for all orders in $\varepsilon$, but one can ask it for the orders higher than two for instance. In this way, the reduced motion would be exactly known to all orders even before computing the change of variables. Of course, this is not guaranteed, since it is possible only if the chosen reduced dynamics is equivalent to the initial particle dynamics [57].

As an example, the minimal guiding-center reduction lets an average contribution (1.29) in the equations of motion for the reduced pitch-angle. Perhaps a stronger reduction could cancel it at least for the orders higher than two, by choosing a convenient (non-zero) averaged part $\underline{\mathbf{X}}$ for the transformation generator. More precisely, in the equation at lowest order, the unknown $\underline{\mathbf{X}}_1$ has no influence on $\dot{\bar{\mathbf{z}}}_0$, which means that $\bar{\varphi}$ cannot be made a constant of motion, as was expected. But at the following order, $\underline{\mathbf{X}}_1$ now contributes, since at that order, the equation $0 = \mathtt{avg}\left(e^{\mathsf{L}_1 + \mathsf{L}_2 + \dots}\dot{\bar{\mathbf{z}}}\right)_\varphi$ writes

$$\mathtt{avg}\left(\frac{\mathcal{L}_{\underline{\mathbf{X}}_1}\mathcal{L}_{\underline{\mathbf{X}}_1} + \mathcal{L}_{\widetilde{\mathbf{X}}_1}\mathcal{L}_{\underline{\mathbf{X}}_1} + \mathcal{L}_{\underline{\mathbf{X}}_1}\mathcal{L}_{\widetilde{\mathbf{X}}_1}}{2}\dot{\bar{\mathbf{z}}}_{-1} + \mathcal{L}_{\underline{\mathbf{X}}_1}\dot{\bar{\mathbf{z}}}_0\right)_\varphi = o.t. , \tag{1.33}$$

where we used that $\mathtt{avg}\left(\mathcal{L}_{\mathbf{X}_2}\dot{\mathbf{z}}_{-1}\right)_{\varphi} = 0$; $\underline{\mathbf{X}}$ and $\widetilde{\mathbf{X}}$ denote respectively the averaged and the fluctuating part of $\mathbf{X}$; and the symbol *o.t.* is a shorthand for all the other terms, which do not contain $\underline{\mathbf{X}}_1$.

Eq. (1.33) shows that, with the additional requirements, the operator to be inverted for $\underline{\mathbf{X}}$ can be quadratic. Even when it is not, e.g. at the next order, the linear operator may not be trivial; a part of it will be given by $\mathcal{L}_{\underline{\mathbf{X}}_1}\dot{\mathbf{z}}_0 = (\underline{\mathbf{X}}_1)_j\partial_j\dot{\mathbf{z}}_0 - (\dot{\mathbf{z}}_0)_j\partial_j\underline{\mathbf{X}}_1$, whose coefficients have the non-trivial expression given by (1.3). Such an operator may not be invertible: if the right-hand side is not in its range. This is the classical problem of secular terms in perturbation theory. Even when the operator is invertible, an explicit inverse may not be obvious to get.

Less strong requirements can be considered than putting to zero the higher-order reduced dynamics. For instance, the paper [123] mentions the possible requirement $\mathsf{b}\cdot\dot{\bar{\mathbf{q}}} = 0$, i.e. the guiding-center motion is a pure drift across the magnetic field lines, whereas the paper [8] mentions a requirement relating the parallel motion of the guiding-center and the reduced parallel velocity $\bar{v_{\|}}$ (where the parallel velocity $v_{\|}$ is used as a coordinate corresponding to the pitch-angle). In any case, care must be taken about additional requirements. They imply differential equations that are not so simple to deal with, and that even can be impossible to solve. It reminds as well that the minimal guiding-center reduction is very nice with its trivial operator $\frac{eB}{m}\partial_\theta$ or equivalently $-\frac{eB}{m}\mathsf{a}\cdot\partial_{\mathsf{c}}$.

The requirement aiming at the magnetic moment plays a special role. Its existence can be viewed as a consequence of the averaging reduction, as shown by Kruskal [81]. The corresponding secular differential equation can be solved at any order, but the procedure is more involved than the minimal one considered in this chapter, as appeared in Chapter 10.

In the approach based on Lie-transforming the equations of motion, the additional requirements do not change the equation (1.4) that gives the fluctuating components of the transformation; they generally aim at a further simplification of the reduced dynamics and imply additional differential equations such as (1.33), which impose the averaged components of the transformation. Those equations are not easily solved, and from this point of view, the method of Lie-transforming the phase-space Lagrangian is more powerful, because it does not rely on differential equations, but on algebraic equations, easier to study.

For instance, the idea to put to zero the averaged components of the reduced motion (at least at orders higher than 2 or 3 in the ordering parameter) is included in usual guiding-center derivations, which obtain that it is possible for six components of the Lagrangian, out of seven. This kind of result would not be so easy to get by Lie-transforming the velocity vector field.

Another additional requirement regards the Hamiltonian structure of guiding-center dynamics. The initial motion $\dot{\mathbf{z}}$ is Hamiltonian, and the reduced motion should be Hamiltonian. Indeed, the full reduced motion $\dot{\bar{\mathbf{z}}}$ is Hamiltonian, even if it has a different Poisson bracket as the initial motion, since it is just given by a (non-canonical) change of coordinates.

However the true reduced motion involves only the slow variables $(\bar{\mathbf{q}}, \bar{\varphi})$, and the corresponding dynamical system $(\dot{\bar{\mathbf{q}}}, \dot{\bar{\varphi}})$ is not guaranteed to be Hamiltonian, because it is given by a truncation of the full reduced dynamics, and truncations do not preserve the Hamiltonian structure in general.

The preservation of the Hamiltonian structure for the 4-components guiding-center dynamics can be considered as an additional requirement. It is hard to obtain by Lie-transforming the equations of motion, because when deriving reduced models by working on the equations of motion, the Hamiltonian character is not worked on, it is observed a posteriori as preserved or not. On the contrary, a Hamiltonian 4-components guiding-center dynamics is easily obtained when working on the phase-space Lagrangian, for instance by Lie-transforming it in such a way that the reduced Poisson bracket is quarter-canonical.

Achieving this in a gauge-independent framework makes the scheme much more involved than the one considered in this chapter. The reduction mechanism is more elaborated, all of the coordinates have to be changed, and the formalism involves not only derivative operators, but also differential forms, which must be dealt with carefully because of the constrained coordinate $\mathsf{c}$. This will be considered in the future.

## Conclusion

A gauge-independent minimal guiding-center reduction can be performed at any order in the Larmor radius expansion by Lie-transforming the velocity vector field. The procedure is very efficient and systematic: it just writes the minimal requirements for the reduction, expands the equation in the Larmor radius, and inverts the generator of Larmor gyration to get both the change of coordinates and the reduced equations of motion order by order. The full second-order reduction was straightforwardly obtained, in contrast with previous derivations of the guiding-center reduction, and the algorithm will be easily implemented in a computer to reach higher-order results, e.g. second-order drifts.

The corresponding transformation generator is a pure gyro-fluctuation and only four of its components are non-zero. This is the bare minimum since exactly the fluctuating part of the slow reduced motion involves a fast time-scale that has to be removed. Thus, the reduction is minimal and unique. All the arbitrary components of the transformation generator are set to zero.

The results bring insights into the structure of the guiding-center formulae, which show up a polynomiality in $\mathsf{b}$, $\mathsf{c}$, $\mathsf{a}$, $r_L$, $\nabla$ and $\phi$, and a parity relation between the orders in $r_L$ , $\phi$ and the fast angle $\mathsf{c}$ or $\mathsf{a}$. This polynomiality makes easier the algorithm, which consists in applying at each order two operations (derivatives, and gyro-integration) onto a polynomial of a very restricted alphabet of entities.

The method can be applied to perform the bounce-average reduction. Also, when an additional electric field is present, the procedure is exactly unchanged even if formulae have additional terms.

In the derivation, the gyro-gauge is never used nor introduced (in this chapter, it was mentioned only for analogy or pedagogical purpose). The gyro-angle is defined and dealt with as usual, but with the intrinsic, physical coordinate $\mathsf{c}$. Accordingly, the whole theory, with all its quantities and results, cannot be but gauge-independent and globally defined, with no use of tricks or conditions, and not only its main results (e.g. guiding-center dynamics), but also all other results (e.g. gyro-angle coordinate and dynamics).

These features could not be aspired at in the standard formalism, as was explicitly stressed in Littlejohn's papers and confirmed in subsequent works, where the theory could not get rid of the gyro-gauge dependence of several quantities, and the main results were made gauge independent only by requiring special conditions, for instance on the structure of the reduced Lagrangian.

The absence of a gyro-gauge removes from the theory the issues associated with the traditional gyro-angle $\theta$, and could bring interesting contributions to clarify these issues in previous results. Indeed, a recent work [25] was lead to a similar orientation when studying the questions raised by the non-existence of a global gauge to define a gyro-angle. This point will be further explored in Chapter 3.

Thus, this approach of the guiding-center reduction, with its minimal procedure and results as well as its gauge independence, can contribute to a better understanding of the guiding-center reduction by giving a simplified and more intrinsic point of view.

All the same, it is only the first step towards a gauge-independent guiding-center theory. Standard guiding-center reductions impose additional (non-minimal) requirements for the reduction.

The magnetic moment is commonly included in the reduced coordinates. This can be done using the gauge-independent gyro-angle and working on the equations of motion by solving the corresponding secular differential equation, but the derivation is more involved, as emphasized in Chapter 10.

Some averaged terms can be transferred from the reduced dynamics to the change of variables, in order to get a stronger guiding-center reduction. This means having non-zero gyro-averaged components in the transformation generator, and these components are again identified by solving secular differential equations.

Such non-minimal guiding-center reductions are more efficiently obtained by Lie-transforming

the phase-space Lagrangian, especially because it mainly relies on algebraic equations. Also, it guarantees a Hamiltonian structure for the 4-components guiding-center dynamics. Introducing the gauge-independent coordinate in this framework is the next step of the work. It makes the scheme more subtle for the constrained coordinate $\mathsf{c}$, especially because of the presence of differential forms. It will be considered in the next chapter.

# Chapter 2

# Gyro-gauge independent formulation of the guiding-center reduction to arbitrary order in the Larmor radius

in collaboration with Michel Vittot
Paper published in Plasma Phys. Control. Fusion **55**, 105001 (2013)

*A*bstract: The guiding-center reduction is studied using gyro-gauge independent coordinates. The Lagrangian 1-form of charged particle dynamics is Lie transformed without introducing a gyro-gauge, but using directly the unit vector of the component of the velocity perpendicular to the magnetic field as the coordinate corresponding to Larmor gyration. The reduction is shown to provide a maximal reduction for the Lagrangian and to work to all orders in the Larmor radius, following exactly the same procedure as when working with the standard gauge-dependent coordinate.
The gauge dependence is removed from the coordinate system by using a constrained variable for the gyro-angle. The closed 1-form $d\theta$ is replaced by a more general non-closed 1-form, which is equal to $d\theta$ in the gauge-dependent case. The gauge vector is replaced by a more general connection in the definition of the gradient, which behaves as a covariant derivative, in perfect agreement with the circle-bundle picture. This explains some results of previous works, whose gauge-independent expressions did not correspond to a gauge fixing but indeed correspond to a connection fixing.
In addition, some general results are obtained for the guiding-center reduction. The expansion is polynomial in the cotangent of the pitch-angle as an effect of the structure of the Lagrangian, preserved by Lie derivatives. The induction for the reduction is shown to rely on the inversion of a matrix which is the same for all orders higher than three. Explicit induction relations are obtained to arbitrary order in the perturbation expansion. The Hamiltonian and symplectic representations of the guiding-center reduction are recovered, but conditions for the symplectic representation at each order are emphasized.

## Introduction

In the quest for an intrinsic formulation of the guiding-center reduction, the work reported in Chapter 10 studied a first aspect, with the magnetic-moment reduction, while the previous chapter focused on the second aspect, with the averaging transformation. We now turn to the full guiding-center reduction.

Indeed, in the previous chapter, the idea was to use a simplified reduction method, by Lie-transforming directly the equations of motion instead of Lie transforming the Lagrangian, as is usually done [27, 87]. Especially, it provided the minimal guiding-center reduction which concerned only four coordinates (instead of six): the transformation generator had no gyro-angle component.

So, for the gyro-angle, the initial gauge-independent coordinate suggested by Chapter 10 could be used, and no gauge fixing was needed. This physical coordinate is the unit vector $\mathsf{c}$ of the component of the velocity orthogonal to the magnetic field, which defines the direction of the perpendicular velocity.

A limitation of that approach was that it was not suited to non-minimal guiding-center reductions, for which the method relying on Lie transforming the Lagrangian appeared as necessary, or at least much more efficient.

For instance, it is interesting to have the slow reduced motion Hamiltonian, but the Hamiltonian structure of the reduced model is hard to deal with when working on the equations of motion. In addition, the magnetic moment is usually taken as one of the reduced coordinates; this can be done by working on the equations of motion but it is not so straightforward, as shown in Chapter 10. In a deeper way, the freedom involved in the gyro-averaged part of the coordinate change can be employed for the reduced motion to be as strongly reduced as possible, to make the guiding-center dynamics as simplified as possible.

Obtaining such a maximal guiding-center reduction by Lie transforming the equations of motion is far from simple, especially because it implies to solve non-trivial secular differential equations. On the contrary, Lie transforming the Lagrangian 1-form basically relies on algebraic equations, and the requirements for a maximal reduction are not much more difficult to get than the minimal ones. This method also guarantees that the 4-dimensional slow reduced motion is Hamiltonian, by working on a quarter-canonical structure in the Poisson bracket.

So, the goal is to use the physical gauge-independent coordinate also when Lie transforming the Lagrangian, in order to consider a gauge-independent maximal guiding-center reduction.

The introduction of a vectorial quantity $\mathsf{c}$ for the gyro-angle coordinate raises some questions, because the coordinate system becomes constrained: the variable $\mathsf{c}$ has to remain normalized and perpendicular to the magnetic field. Changing the spatial position $\mathbf{q}$ implies to change the coordinate $\mathsf{c}$ at the same time. This induces a connection for a covariant derivative on a space-dependent circle, which is related to the circle-bundle picture underlying in the gyro-angle coordinate [25, 81]. It was already present in Chapter 1 when Lie transforming the equations of motion, but it is more involved to deal with for the full guiding-center reduction, because the coordinate $\mathsf{c}$ will be changed as well, and not only derivatives or vector fields are involved, but also differential forms.

The resulting reduction will naturally provide gauge-independent results, whereas in the usual approach, they were obtained only for a part of the reduced quantities. This fact can shed interesting light on previous guiding-center results, especially those related to gauge invariance. For instance, in the usual approach, the gradient is not gauge-independent, and the reduced Poisson bracket involves a gauge-independent corrected gradient. A comparison with the results of the gauge-independent formulation is a way to get an intrinsic interpretation for this corrected gradient.

The results of the gauge-independent formulation can also be used to explore other questions about the gauge-dependent approach, for instance related to gauge arbitrariness and anholonomy [86, 90]. To avoid confusion, they will be the topic of the distinct chapter 3. Here, we show how a maximal guiding-center reduction can be derived in a gauge-independent formulation to arbitrary order in the Larmor radius.

The proof relies on explicit induction relations to all orders, because the induction can be written as a matrix product, with some coefficients being differential operators. Through inversion of this matrix, a maximal reduction can be studied, towards a more complete reduction and a more general viewpoint on special reductions considered in previous works, such as the so-called Hamiltonian and symplectic representations identified in [23].

In the derivation, the cotangent of the pitch-angle is used as the coordinate corresponding to the parallel velocity (component of the velocity parallel to the magnetic field), since this coordinate simplified computations for the minimal guiding-center reduction in Chapter 1 and made all formulae polynomials. This will clarify why this polynomiality can be observed in the results of the full guiding-center reduction as well.

Taking care of the reduction at higher order is interesting, if not needed, for two reasons. On

the first hand, it is necessary to validate the gauge-independent approach. Indeed, the gauge issues became more sensitive when addressing the second-order reduction [26,145]. An acceptable solution to these issues, which is the goal of the present chapter, needs to be appropriate for higher orders.

On the other hand, it is motivated because standard works proceeded only up to part of the second-order reduction, but recent results emphasize the importance of higher-order terms, for instance because they are involved in the conservation of angular momentum [26,127] and are crucial for a proper description of intrinsic rotation of tokamak plasmas, a key phenomenon to stabilize turbulence and increase the energy confinement time, which is the main goal of magnetic fusion.

The chapter is organized as follows. In Sec. 2.1, a few facts are reminded about the initial dynamics, the choice of coordinates for the gyro-angle, the method of Lie transforming the Lagrangian 1-form, and the hierarchy of requirements involved in the guiding-center reduction.

For the sake of completeness and clarity, the mechanism of Lie transforming the Lagrangian through an expansion in a small parameter is described in an appendix, with emphasis on the three steps it involves: an initialization for the lowest orders, whose choices are the key to make the reduction work and possibly be optimal; an algorithm which applies for higher orders, is purely mechanical and can be applied to study the reduction to arbitrary order; and an intermediate step in between.

In Sec. 2.2, the method is applied to the guiding-center reduction in case the gyro-angle coordinate is chosen as the physical variable $\mathsf{c}$. The derivation is written in matrix form, which emphasizes both the lowest-order choices that allow the reduction to work and to be maximal for the Lagrangian, and the algorithmic character of the procedure at higher orders. The full derivation is explained because we are interested in the reduction at arbitrary order in the Larmor radius, which implies to use all the ingredients of the detailed mechanism at work. For orders lower than 3, the procedure follows the same lines as when working with the gauge-dependent coordinate, but formulae have to be used in their intrinsic version, for instance because the basis of 1-forms involves non-closed 1-forms. Finally, for orders higher than 2, explicit formulae are given for the induction relations, allowing to go to arbitrary order, and to give a unified framework where recent results may seem to discord somehow with each other [23,128].

In Sec. 2.3, the results are compared with previous works, either Lie transforming the equations of motion, or using a gauge-dependent gyro-angle.

For the sake of simplicity, we consider the special case where there is no electric field, but the generalization for a non-zero electric field is straightforward, as will be shown in Subsec. 2.3.1.

## 2.1 Coordinates, method and requirements

The dynamical system is simply a charged particle with position $\mathbf{q}$, momentum $\mathbf{p}$, mass $m$ and charge $e$, under the influence of a static inhomogeneous magnetic field $\mathbf{B}$. The motion is given by the Lorentz force

$$\begin{aligned} \dot{\mathbf{q}} &= \tfrac{\mathbf{p}}{m}\,, \\ \dot{\mathbf{p}} &= \tfrac{\mathbf{p}}{m} \times e\mathbf{B}\,. \end{aligned} \tag{2.1}$$

When the magnetic field is strong, the motion implies a separation of time-scales. This is best seen by choosing convenient coordinates for the momentum space, for instance

$$\begin{aligned} p &:= \|\mathbf{p}\|\,, \\ \varphi &:= \arccos\left(\tfrac{\mathbf{p}\cdot\mathsf{b}}{\|\mathbf{p}\|}\right), \\ \mathsf{c} &:= \tfrac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|}\,, \end{aligned}$$

where $\mathsf{b} := \frac{\mathbf{B}}{\|\mathbf{B}\|}$ is the unit vector of the magnetic field, and $\mathbf{p}_\perp := \mathbf{p} - (\mathbf{p}\cdot\mathsf{b})\mathsf{b}$ is the so-called *perpendicular momentum*, i.e. the orthogonal projection of the momentum onto the plane $\mathbf{B}^\perp$ perpendicular to the magnetic field. The coordinate $p$ is the norm of the momentum; the coordinate

$\varphi$ is the so-called *pitch-angle*, i.e. the angle between the velocity and the magnetic field. The last coordinate $\mathsf{c}$ is the unit vector of the perpendicular velocity.

Then, the equations of motion are

$$\begin{aligned}
\dot{\mathbf{q}} &= \tfrac{\mathbf{p}}{m}\,,\\
\dot{p} &= 0\,,\\
\dot{\varphi} &= -\tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\,,\\
\dot{\mathsf{c}} &= -\tfrac{eB}{m}\mathsf{a} - \tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot(\mathsf{c}\mathsf{b} + \mathsf{a}\mathsf{a}\cot\varphi)\,,
\end{aligned} \tag{2.2}$$

where $\mathbf{p}$ is now a shorthand for $p(\mathsf{b}\cos\varphi + \mathsf{c}\sin\varphi)$, the norm of the magnetic field $\|\mathbf{B}\|$ is denoted by $B$, and following Littlejohn's notations [85,86,88], the vector $\mathsf{a} := \mathsf{b}\times\mathsf{c}$ is the unit vector of the Larmor radius, so that $(\mathsf{a},\mathsf{b},\mathsf{c})$ is a right-handed orthonormal frame (rotating with the momentum).

Notice that we have adopted above the spherical coordinates $(p,\varphi,\mathsf{c})$ for the momentum as is natural at first. Later on, $p$ and $\varphi$ will be replaced by other coordinates, which will appear more convenient for the guiding-center reduction (see Eqs. (2.25) and (2.26)).

In the case of a strong magnetic field, the only fast term is the Larmor frequency $\omega_L := \frac{eB}{m}$. Writing the dynamics as $\dot{\mathbf{z}}\cdot\partial_{\mathbf{z}}$, all other terms as $\omega_L$ appear to be of order $\frac{p}{m}\nabla$, which means that the small parameter of the theory is of order $\frac{p}{m\omega_L}\nabla = \frac{p}{eB}\nabla$. A more detailed study (e.g. [29,59,86,122]) shows that it is rather

$$\varepsilon := \tfrac{p\sin\varphi}{eB}\nabla\,. \tag{2.3}$$

It is an operator, but the gradient $\nabla$ has only a meaning for orderings; it acts on the magnetic field and can be given a more precise meaning by relations such as $\nabla \approx \frac{\|\nabla^{n+1}\mathbf{B}\|}{\|\nabla^n\mathbf{B}\|}$. The ordering parameter $\varepsilon$ is related to the magnetic inhomogeneity at the scale of the Larmor radius $r_L := \frac{p\sin\varphi}{eB}$. By abuse of language, it is often considered as just the Larmor radius, or as the inverse charge $e^{-1}$ [29,86,122].

### 2.1.1 Choice for the gyro-angle coordinate

The Larmor frequency $\omega_L := \frac{eB}{m}$ concerns only one coordinate, $\mathsf{c}$, the direction of the vector $\mathbf{p}_\perp$ in the 2-dimensional plane perpendicular to the magnetic field. This corresponds to an angle, the so-called *gyro-angle*, and measures the Larmor gyration of the particle momentum around the magnetic field.

To get a true scalar angle instead of the vector $\mathsf{c}$, one chooses at each point $\mathbf{q}$ in space a direction which will be considered as the reference axis $\mathsf{e}_1(\mathbf{q}) \in \mathbf{B}^\perp(\mathbf{q})$. This corresponds to fixing a gauge in the theory, the so-called *gyro-gauge*. Then, the gyro-angle $\theta$ is defined from the oriented angle between the chosen reference axis $\mathsf{e}_1(\mathbf{q})$ and the vector $\mathsf{c}$ through the following relation:

$$\mathsf{c} = -\sin\theta\mathsf{e}_1 - \cos\theta\mathsf{e}_2\,. \tag{2.4}$$

The equation of motion for $\theta$ is

$$\dot{\theta} = \tfrac{eB}{m} + \cot\varphi\tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{a} + \tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{e}_1\cdot\mathsf{e}_2\,,$$

with $\mathsf{e}_2 := \mathsf{b}\times\mathsf{e}_1$ the unit vector such that $(\mathsf{b},\mathsf{e}_1,\mathsf{e}_2)$ is a (fixed) right-handed orthonormal frame [86].

The coordinate $\theta$ is not intrinsic, it depends on the chosen gauge $\mathsf{e}_1(\mathbf{q})$, which raised some questions about the gauge invariance of the theory, about the failure of global existence for $\mathsf{e}_1$, as well as about the presence of an anholonomic phase in the coordinate system [22,25,86,90,91]. All these difficulties originate because $\theta$ is not given by the physics, it is not a purely intrinsic coordinate. For physical results, what is needed is not $\theta$ but only $\mathsf{c}$. This is clearly illustrated by all the results of guiding-center theory, e.g. [29,86,88,122], where $\theta$ intervenes only through the quantity $\mathsf{c}$ everywhere (except in its own definition and in subsequent relations). So, we will avoid this coordinate and keep the corresponding initial variable, $\mathsf{c}$, as in Chapter 1. The quantity $\theta$ will be used only with a symbolic meaning for the fast angle, or when making comparison with the

gauge-dependent approach.

The use of a unit vector avoids having to fix a gauge for the zero of the angle, and it allows to work with a physical quantity: the unit vector of the perpendicular velocity

$$\mathsf{c} := \frac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|} = \frac{\mathbf{p}-\mathsf{b}(\mathsf{b}\cdot\mathbf{p}\ )}{p\sin\varphi}\,, \tag{2.5}$$

which indeed corresponds to the direction of the perpendicular velocity, and is a coordinate measuring the Larmor gyration. It is an angle, since it is a unit vector in a plane, namely the plane $\mathbf{B}^\perp$, orthogonal to the local magnetic field. But this unit vector is immersed into $\mathbb{R}^3$, which means it is in $\mathbb{S}^1(\mathbf{q})$. This spatial dependence implies that the coordinate space is constrained: the gyro-angle $\mathsf{c}$ is not independent of the spatial position.

When the coordinate $\mathbf{q}$ is changed, the coordinate $\mathsf{c}$ cannot be kept unchanged, otherwise it may get out of $\mathbf{B}^\perp$. Differentiating Eq. (2.5) with respect to $\mathbf{q}$ gives

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot(\mathsf{cb} + \mathsf{aa}\cot\varphi)\,. \tag{2.6}$$

This formula can be obtained more easily by noticing that in the change of coordinates $(\mathbf{q},\mathbf{p}) \longrightarrow (\mathbf{q}, p, \varphi, \mathsf{c})$, the following relations hold

$$\begin{aligned}
&-\sin\varphi\nabla\varphi = \nabla\cos\varphi = \nabla\mathsf{b}\cdot\frac{\mathbf{p}}{p} = \nabla\mathsf{b}\cdot\mathsf{c}\sin\varphi\,,\\
&\quad\Rightarrow \nabla\varphi = -\nabla\mathsf{b}\cdot\mathsf{c}\,,\\
&\quad\Rightarrow 0 = \nabla\left(\frac{\mathbf{p}}{p}\right)\\
&\qquad\quad = \nabla\mathsf{c}\sin\varphi + \nabla\mathsf{b}\cos\varphi + \nabla\varphi(-\mathsf{b}\sin\varphi + \mathsf{c}\cos\varphi)\\
&\qquad\quad = (\nabla\mathsf{c} + \nabla\mathsf{b}\cdot\mathsf{cb})\sin\varphi + \nabla\mathsf{b}\cdot(1-\mathsf{cc})\cos\varphi\,,\\
&\quad\Rightarrow \nabla\mathsf{c} = -\nabla\mathsf{b}\cdot(\mathsf{cb} + \mathsf{aa}\cot\varphi)\,,
\end{aligned} \tag{2.7}$$

in which $\nabla$ means differentiation with respect to $\mathbf{q}$ while keeping $\mathbf{p}$ constant, and we used that $1 = \mathsf{aa} + \mathsf{bb} + \mathsf{cc}$ and that $\mathsf{b}$ is a unit vector, which implies $\nabla\mathsf{b}\cdot\mathsf{b} = \nabla\left(\frac{\mathsf{b}^2}{2}\right) = 0$.

Eq. (2.6) is not well defined where $\varphi = 0 \pmod{\pi}$, i.e. where $\mathbf{p}$ is parallel to $\mathbf{B}$. But this is no trouble, since it fits in with a usual limitation of guiding-center theories. For instance, guiding-center transformations are not defined where $\varphi$ is zero, since they involve many $\sin\varphi$ as denominators [29, 86, 88, 122]. At the points where $\varphi = 0 \pmod{\pi}$, the vector $\mathsf{c}$ itself is not defined, neither is the coordinate $\theta$. It is an implicit assumption in all gyrokinetics and guiding-center works that those points are excluded from the theory.

If the coordinate $\theta$ was used, Eq. (2.6) would be replaced by

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot\mathsf{cb} + \nabla\mathsf{e}_1\cdot\mathsf{e}_2\ \mathsf{a}\,, \tag{2.8}$$

where $\nabla\mathsf{e}_1\cdot\mathsf{e}_2$ is the so-called *gauge vector*, which is usually denoted by $\mathbf{R}$, depends only of the position $\mathbf{q}$, and is related to the choice of gauge.

To fit in with both coordinates, we define a more general function $\mathbf{R}_g(\mathbf{q},\mathbf{p})$:

$$\mathbf{R}_g := \nabla\mathsf{c}\cdot\mathsf{a}\,.$$

Then, in any coordinates, the previous formulae (2.6) and (2.8) write

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot\mathsf{cb} + \mathbf{R}_g\mathsf{a}\,. \tag{2.9}$$

The value $\mathbf{R}_g = -\nabla\mathsf{b}\cdot\mathsf{aa}\cot\varphi$ corresponds to the physical definition of $\mathsf{c}$, and the corresponding Eq. (2.6). The value $\mathbf{R}_g = \nabla\mathsf{e}_1\cdot\mathsf{e}_2 = \mathbf{R}$ will be linked with the usual case relying on the gauge-dependent coordinate $\theta$, according to Eq. (2.8).

The spatial dependence of the vector $\mathsf{c}$ through Eq. (2.9) must be taken into account each time a gradient acts on a function that depends on the fast angle $\mathsf{c}$, e.g. in total derivatives, in Lie transforms, and when computing the action of a spatial component of the change of variables.

A more detailed study of the coordinate system in Chapter 3 shows that $\mathbf{R}_g$ is the connection associated to the gyro-angle circle bundle, which indeed can be any function of the phase space, and that $\nabla$ is the (spatial part of the) corresponding covariant derivative.

### 2.1.2 Lie transforming the Lagrangian 1-form

The goal is to isolate the slow part of the dynamics from the fast angle $\theta$ (or $\mathsf{c}$), by performing a near-identity change of coordinates such that the dynamics of the remaining coordinates does not depend on $\theta$; this averaging procedure is the primary requirement for the guiding-center reduction. We showed in Chapter 1 how it can be derived by Lie transforming the equations of motion when keeping the physical coordinate $\mathsf{c}$. It appeared to be very straightforward for the minimal guiding-center reduction, but the additional requirements (e.g. the use of the magnetic moment as a coordinate or a further simplification of the reduced dynamics) were not so easy to obtain. They are more efficiently obtained by Lie transforming the Lagrangian 1-form, which relies on the Hamiltonian structure of this dynamical system.

Let us remind that a dynamical system with coordinates $\mathbf{z}$ is Hamiltonian when its dynamics is given by Hamilton's equation[1]:

$$\dot{\mathbf{z}} = \{\mathbf{z}, H\}\,, \tag{2.10}$$

where $H(\mathbf{z})$ is a function, which is conserved in time and called the Hamiltonian function (or just the *Hamiltonian*). The operation $\{\cdot,\cdot\}$ is called a *Poisson bracket*; it takes two functions as its arguments and the result is a function. It is antisymmetric, bilinear, verifies the Leibniz rule, and satisfies the Jacobi identity, which respectively mean

$$\begin{aligned}
\{g, f\} &= -\{f, g\}\,,\\
\{f, g+h\} &= \{f, g\} + \{f, h\}\,,\\
\{f, gh\} &= \{f, g\}h + \{f, h\}g\,,\\
\{f, \{g, h\}\} &= \{g, \{f, h\}\} + \{h, \{g, f\}\}\,,
\end{aligned}$$

for any functions $f$, $g$, and $h$.

In order to define a *Hamiltonian field dynamics*, take the definition above while replacing the finite dimensional space $\mathbf{z}$ by a set of functions over a continuous space (see e.g. Eqs. (4.2)-(4.4)). For an introduction about Hamiltonian systems, see e.g. [53, 99, 107, 109, 141].

For the dynamics of a charged particle in a magnetic field, the coordinates $(\mathbf{p}, \mathbf{q})$ can be grouped in a single phase-space vector $\mathbf{z} := (\mathbf{p}, \mathbf{q})$. The Hamiltonian function is just the particle kinetic energy

$$H := \tfrac{p^2}{2m}\,. \tag{2.11}$$

The Poisson bracket is non-canonical, it contains the gyro-magnetic coupling term

$$\{F, G\} = \partial_{\mathbf{q}}F\cdot\partial_{\mathbf{p}}G - \partial_{\mathbf{p}}F\cdot\partial_{\mathbf{q}}G - \partial_{\mathbf{p}}F\cdot e\mathbf{B}\times\partial_{\mathbf{p}}G\,. \tag{2.12}$$

Together, this Poisson bracket and this Hamiltonian induce the equations of motion through Hamilton's equations (2.10), which just produce the traditional Lorentz-force motion (2.1), or the equations of motion (2.2) when the coordinates $(\mathbf{q}, p, \varphi, \mathsf{c})$ are used.

Instead of working with the Poisson bracket and the Hamiltonian, it is easier to work on the *Poincaré-Cartan 1-form* $\Gamma$ [88]. This last is usually called just a Lagrangian 1-form, or simply a *Lagrangian*. It concentrates all the information on the Hamiltonian structure into one single quantity, which in addition is much less constrained than a Poisson bracket. It is a 1-form, defined over a 7-dimensional space $\mathbf{y} := (\mathbf{p}, \mathbf{q}, t)$ by [29, 87, 88]

$$\Gamma := (e\mathbf{A} + \mathbf{p})\cdot d\mathbf{q} - H\,dt\,, \tag{2.13}$$

which yields a variational formulation of the dynamics with the action [86]:

$$\mathcal{A} := \int \Gamma\,.$$

[1] Notice that the convention $\dot{\mathbf{z}} = \{H, \mathbf{z}\}$ is equivalently used, by inverting the sign of the Poisson bracket.

The so-called symplectic Lagrangian [29, 88] is the contribution of $\Gamma$ that corresponds to its phase-space components. It is a 1-form $\Gamma_s = \Gamma^i_s dz^i := (e\mathbf{A} + \mathbf{p})\cdot d\mathbf{q}$ in the usual 6-dimensional phase space, which gives the Lagrange 2-form through exterior derivative

$$\omega_s := d\Gamma_s = (\partial_i \Gamma^j_s - \partial_j \Gamma^i_s)\, dz^i \otimes dz^j \,, \tag{2.14}$$

where $\otimes$ means tensorial product, and Einstein's convention is used: there is an implicit summation over repeated indices. In turn, $\omega_s$ is invertible, and its inverse gives the Poisson bracket

$$\mathbb{J} := \omega_s^{-1} \,,$$

with $\mathbb{J}$ the bivector defined by the relation $\{F, G\} = \partial_i F \mathbb{J}^{ij} \partial_j G$.

In the presence of a strong magnetic field, the small parameter $\varepsilon$ allows for a perturbation expansion of all quantities. The Lagrangian $\Gamma$ from Eq. (2.13) can be written

$$\Gamma = \Gamma_{-1} + \Gamma_0 \,,$$

where the index refers to the order in the magnetic field, (or in $e^{-1}$, following Northrop's work [121, 122])

$$\begin{aligned} \Gamma_{-1} &:= e\mathbf{A}\cdot d\mathbf{q} \,, \\ \Gamma_0 &:= \tfrac{\mathbf{p}}{m}\cdot d\mathbf{q} - \tfrac{p^2}{2m} dt \,, \end{aligned} \tag{2.15}$$

whose ratio is indeed of order $\varepsilon$. The lowest order being $-1$ reminds that the guiding-center is a singular perturbation theory.

The Lagrangian is transformed by the exponential of a Lie transform

$$\Gamma \longrightarrow \overline{\Gamma} := e^{\mathcal{L}_{\mathbf{X}}} \Gamma \,,$$

where $\mathcal{L}_{\mathbf{X}}$ is the Lie derivative along the vector field $\mathbf{X}$, whose inverse is the generator of the near-identity coordinate transformation:

$$\mathbf{y} \longrightarrow \overline{\mathbf{y}} := e^{-\mathbf{X}} \mathbf{y} \,.$$

The generator can be expanded $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + ...$, where $\mathbf{X}_n$ is the term of $\mathbf{X}$ that is of order $n$ in $\varepsilon$. From the point of view of formal expansions, it is equivalent but simpler to replace the single transformation $e^{\mathcal{L}_{\mathbf{X}_1} + \mathcal{L}_{\mathbf{X}_2} + \cdots}$ with a complicated generator by a series of transformations with a simple generator for each of them

$$\mathbf{y} \longrightarrow \overline{\mathbf{y}} := ... e^{-\mathbf{G}_2} e^{-\mathbf{G}_1} \mathbf{y} \,,$$

where $-\mathbf{G}_n$ is the vector field generating the $n$-th transformation, and is purely of order $\varepsilon^n$. The Lagrangian then transforms as

$$\Gamma \longrightarrow \overline{\Gamma} := ... e^{\mathsf{L}_2} e^{\mathsf{L}_1} \Gamma \,,$$

where $\mathsf{L}_n := \mathcal{L}_{\mathbf{G}_n}$ is the Lie derivative along the vector field $\mathbf{G}_n$. Notice that $\mathcal{L}_{\mathbf{X}} \neq \mathsf{L}_1 + \mathsf{L}_2 + ...$ because $\mathbf{G}_n \neq \mathbf{X}_n$.

As the Lagrangian is time-independent, it is interesting to use time-independent perturbation theory, by imposing a transformation that does not depend on time and does not affect the time coordinate: $\mathbf{G}$ is constant in time and $\mathbf{G}^t = 0$.

This implies that the Hamiltonian will transform as a scalar function. Indeed, for any time-independent vector field $\mathbf{G}$ and any 1-form,

$$\begin{aligned} (\mathcal{L}_{\mathbf{G}}\Gamma)_t &= [i_{\mathbf{G}}\cdot d\Gamma]_t + [d(i_{\mathbf{G}}\Gamma)]_t \\ &= \left[\mathbf{G}^i \partial_{\mathbf{y}^i}(\Gamma_j) d\mathbf{y}^j - \mathbf{G}^i \partial_{\mathbf{y}^j}(\Gamma_i) d\mathbf{y}^j\right]_t + \partial_t(\mathbf{G}\cdot\Gamma) \\ &= \mathbf{G}\cdot\partial_{\mathbf{y}}\Gamma_t + \partial_t(\mathbf{G})\cdot\Gamma = \mathbf{G}\cdot\partial_{\mathbf{y}}\Gamma_t \,, \end{aligned}$$

where Eq. (2.17) was used for Lie derivatives, and Eq. (2.14) for exterior derivatives (here in the 7-dimensional space).

Another consequence is that at each order $n$ in $\varepsilon$, there will be seven requirements (see next subsection), one for each component of the reduced Lagrangian $\overline{\Gamma}_n$, and only six freedoms, one for each component of the time-independent transformation generator $\mathbf{G}_n$. One freedom is missing in $\mathbf{G}_n$ and must be looked for elsewhere. Now, the Lagrangian is defined only to within a total derivative, since only its exterior derivative has a physical meaning, and $d(\Gamma + dS) = d\Gamma + d^2 S = d\Gamma$ for any function $S$, which is called a *gauge function* [29]. Be careful, this gauge has nothing to see with the gyro-gauge nor the gyro-angle, it is just an arbitrariness in the definition of the Lagrangian. Rather, it corresponds to the electromagnetic gauge, since it can be absorbed in a redefinition of the potential $\mathbf{A}$ (together with $\Phi$ when there is a non-zero electric field, see page 81): the Lagrangian is expressed in terms of the potential, and does depend on the electromagnetic gauge, but the dynamics, as well as all physical quantities, are electromagnetic-gauge invariant, they depend only on the electromagnetic field.

The freedom embodied in $S$ is needed to obtain a maximal reduction, since it gives the expected seventh freedom. Then the reduced Lagrangian is :

$$\overline{\Gamma} := \left(...e^{\mathsf{L}_2} e^{\mathsf{L}_1}\right) \left(\Gamma_{-1} + \Gamma_0\right) + \left(dS_{-1} + dS_0 + ...\right).$$

It will be determined order by order in $\varepsilon^n$:

$$\begin{aligned}
\overline{\Gamma}_{-1} &= \Gamma_{-1} + dS_{-1} \\
\overline{\Gamma}_0 &= \mathsf{L}_1 \Gamma_{-1} + \Gamma_0 + dS_0 \\
\overline{\Gamma}_1 &= \left(\mathsf{L}_2 + \tfrac{\mathsf{L}_1^2}{2}\right) \Gamma_{-1} + \mathsf{L}_1 \Gamma_0 + dS_1 \\
\overline{\Gamma}_2 &= \left(\mathsf{L}_3 + \mathsf{L}_2 \mathsf{L}_1 + \tfrac{\mathsf{L}_1^3}{6}\right) \Gamma_{-1} + \left(\mathsf{L}_2 + \tfrac{\mathsf{L}_1^2}{2}\right) \Gamma_0 + dS_2 \\
\overline{\Gamma}_3 &= \left[\mathsf{L}_4 + \mathsf{L}_3 \mathsf{L}_1 + \mathsf{L}_2 \left(\tfrac{\mathsf{L}_2}{2} + \tfrac{\mathsf{L}_1^2}{2}\right) + \tfrac{\mathsf{L}_1^4}{24}\right] \Gamma_{-1} \\
&\qquad + \left[\mathsf{L}_3 + \mathsf{L}_2 \mathsf{L}_1 + \tfrac{\mathsf{L}_1^3}{6}\right] \Gamma_0 + dS_3 \\
&\ ...
\end{aligned} \tag{2.16}$$

In principle, these are differential equations for $\mathbf{G}$, because the action of a Lie derivative $\mathcal{L}_{\mathbf{G}}$ over a 1-form $\gamma$ writes

$$\mathcal{L}_{\mathbf{G}} \gamma = (i_{\mathbf{G}} d + d\, i_{\mathbf{G}}) \gamma\,, \tag{2.17}$$

where the operator $i_{\mathbf{G}}$ is the interior product, e.g.

$$i_{\mathbf{G}} \gamma = \mathbf{G} \cdot \gamma = \gamma(\mathbf{G})\,. \tag{2.18}$$

So, the first term in (2.17) is algebraic in $\mathbf{G}$, but the second one is differential in $\mathbf{G}$.

However, the differential operators can be avoided by the following argument. The last term in (2.17) involves an exterior derivative, and can be removed by redefining the gauge function $S$. In addition, Eq. (2.17) together with the property $d^2 = 0$ imply that the exterior derivative and the Lie derivative commute:

$$\mathcal{L}_{\mathbf{G}} d = d \mathcal{L}_{\mathbf{G}} = d\, i_{\mathbf{G}} d\,.$$

This means that for any vector fields $\mathbf{X}$ and $\mathbf{G}$

$$\begin{aligned}
\mathcal{L}_{\mathbf{G}} \mathcal{L}_{\mathbf{X}} \Gamma &= (i_{\mathbf{G}} d + d i_{\mathbf{G}}) \mathcal{L}_{\mathbf{X}} \Gamma = i_{\mathbf{G}} d (i_{\mathbf{X}} d + d i_{\mathbf{X}}) \Gamma + d(i_{\mathbf{G}} \mathcal{L}_{\mathbf{X}} \Gamma) \\
&= i_{\mathbf{G}} d i_{\mathbf{X}} d\Gamma + d(i_{\mathbf{G}} \mathcal{L}_{\mathbf{X}} \Gamma)\,.
\end{aligned}$$

In computations for $\overline{\Gamma}$, the last term can again be removed by redefining the gauge function $S$. By induction, it is now easy to see that in Eq. (2.16), exponentials of Lie derivatives can be considered

as just exponentials of interior products provided the gauge function $S_n$ is defined in a convenient way at each order $\varepsilon^n$ to absorb all the exterior derivatives involved in Eq. (2.16):

$$\mathcal{L}_{\mathbf{G}_{n_1}}\mathcal{L}_{\mathbf{G}_{n_2}}...\mathcal{L}_{\mathbf{G}_{n_k}}\Gamma + dS \\ = (i_{\mathbf{G}_{n_1}}d)(i_{\mathbf{G}_{n_2}}d)...(i_{\mathbf{G}_{n_k}}d)\Gamma + dS' \,.$$

For the following, we will redefine $S$ according to this rule, but for simplicity, we drop the prime and write $S$ for $S'$.

Using this rule and the fact that in the equation for $\overline{\Gamma}_n$, the $\overline{\Gamma}_{i<n}$ are already known, Eqs. (2.16) can be written

$$\begin{aligned}
\overline{\Gamma}_{-1} &= \Gamma_{-1} + dS_{-1} \\
\overline{\Gamma}_0 &= \mathbf{G}_1\cdot\omega_{-1} + \Gamma_0 + dS_0 \\
\overline{\Gamma}_1 &= \mathbf{G}_2\cdot\omega_{-1} + \tfrac{\mathbf{G}_1}{2}\cdot(\omega_0 + \overline{\omega}_0) + dS_1 \\
\overline{\Gamma}_2 &= \mathbf{G}_3\cdot\omega_{-1} + \mathbf{G}_2\cdot\overline{\omega}_0 + \tfrac{\mathbf{G}_1}{6}\cdot d\big[\mathbf{G}_1\cdot(2\omega_0 + \overline{\omega}_0)\big] + dS_2 \\
\overline{\Gamma}_3 &= \mathbf{G}_4\cdot\omega_{-1} + \mathbf{G}_3\cdot\overline{\omega}_0 + \mathbf{G}_2\cdot\overline{\omega}_1 - \tfrac{(\mathbf{G}_2\cdot d)^2}{2}\omega_{-1} \\
&\qquad + \tfrac{(\mathbf{G}_1\cdot d)^2}{24}\mathbf{G}_1\cdot(3\omega_0 + \overline{\omega}_0) + dS_3 \\
&\ldots
\end{aligned} \tag{2.19}$$

where the notation $\mathbf{G}\cdot := i_{\mathbf{G}}$ is used for the interior product, as in Eq. (2.18). In addition, the $n$-$th$-order Lagrange 2-form was defined in the natural way:

$$\omega_n := d\Gamma_n \,.$$

In the next subsection, we study the properties that are wished for $\overline{\Gamma}$. In the next section, the unknowns $\mathbf{G}_n$ and $S_n$ will be determined such that $\overline{\Gamma}_n$ has those desired properties.

### 2.1.3 The hierarchy of requirements

A) The primary requirement for the guiding-center reduction is to isolate the slow dynamics of the coordinates $(\overline{\mathbf{q}}, \overline{\varphi})$ from the fast gyro-angle $\overline{\theta}$. From the point of view of the Lagrangian $\overline{\Gamma}$, it may be obtained by making $\overline{\Gamma}$ independent of $\overline{\theta}$. This is actually stronger than the strict minimal requirement, since it implies to average the dynamics of $\overline{\theta}$ as well. However, in the Lagrangian approach, contrary to when working on the equations of motion, the minimal requirement would be difficult to get, if not impossible, and it is quite easier to average all the reduced dynamics.

So, the goal is that the reduced Lagrangian does not depend on the reduced gyro-angle, which means that all its non-zero Fourier components (i.e. purely oscillatory terms) are zero:

$$\mathtt{osc}(\overline{\Gamma}_n) = 0 \,, \tag{2.20}$$

where following Littlejohn's notations, $\mathtt{osc} = 1 - \mathtt{avg}$ is the projector onto gyro-fluctuations, with $\mathtt{avg}$ the complementary projector onto gyro-averages:

$$\mathtt{avg}(f) = \tfrac{1}{2\pi}\int_0^{2\pi} d\theta \; f \,,$$

for any function $f$. This average can be computed without introducing any gauge: either using the intrinsic calculus introduced in Chapter 2, either using the matrix calculus introduced in Chapter 1. Also, the coordinate $\theta$ can be used as an intermediate quantity for this computation, which is made at constant $\mathbf{q}$, so that the presence of a gauge (only for the intermediate computation) is of no consequence.

B) Averaging the motion or the Lagrangian does not determine the average components of the coordinate change, as is clear in Chapter 1, for instance. This lets some freedom in the procedure

and suggests to impose stronger requirements for the reduction. The basic idea is to use the available freedoms to make the reduced dynamics as simplified as possible.

A natural prospect is to make trivial the reduced dynamics

$$\dot{\overline{\mathbf{z}}}^j = 0\,,$$

for some components $j$, by including constants of motion in the reduced coordinates. For the remaining coordinates, one can consider putting their reduced dynamics to zero just for orders higher than 2 or 3 for instance:

$$\dot{\overline{\mathbf{z}}}^j_n = 0\,,$$

for all higher orders, where the index $n$ refers to the order in $\varepsilon^n$ and the exponent $j$ indicates the component of the vector $\dot{\overline{\mathbf{z}}}$.

When this is achieved, the reduced dynamics is given just by lowest-order terms; it is exactly known after the lowest orders have been derived, without computing the reduction at higher orders, which are useful only to determine the transformation.

In the procedure working on the Lagrangian 1-form, the "components" are not the ones of the reduced equations of motion, but the ones of the reduced Lagrangian

$$\overline{\Gamma}^j_n = 0\,, \tag{2.21}$$

for higher $n$. For differential forms, we use the same convention as for vectors: the index $n$ indicates the order in $\varepsilon^n$ and the exponent $j$ refers to the component. It departs from the usual notation $\overline{\Gamma}_n = (\overline{\Gamma}_n)_j d\mathbf{z}^j$, but it avoids excessive use of parentheses.

Eq. (2.21) gives additional (i.e. non minimal) requirements for the reduced Lagrangian $\overline{\Gamma}$ and it is used to determine the averaged transformation generators. When it cannot be obtained completely, the goal is to obtain it for as many components $j$ as possible, to get what can be considered as a maximal reduction.

C) When studying the Lie transform of the Lagrangian 1-form, one variable, namely the *magnetic moment* $\overline{\mu}$, plays a key role, as the variable conjugated to the gyro-angle.

Basically, including $\overline{\mu}$ among the reduced coordinates is a way to obtain a more efficient reduction process by making one of the components of the reduced Lagrangian trivial $\overline{\mu} := \overline{\Gamma}^\theta$. Indeed, when the derivation is performed with the variable $p$, then the $\theta$-component of the Lagrangian is given by a whole series, which is the magnetic moment. Changing the variable $p$ in such a way that the new variable absorbs this series is a way to have the reduced $\theta$-component trivial, just given by a coordinate. This also simplifies the reduction algorithm, by providing a simpler expression for $\overline{\omega}$, which will play a key role in the derivation.

Second, the resulting variables $\overline{\theta}$ and $\overline{\mu}$ are conjugated, which implies that the magnetic moment is a constant of motion besides the norm of the momentum $p$. And this conserved quantity is preferable to the variable $p$ because, unlike $p$, it remains an adiabatic invariant in the presence of a wide class of electric fields [122].

The variable $\overline{\mu}$ is a whole series in the Larmor radius, and its lowest-order term is the well-known adiabatic invariant $\mu$ often confounded with $\overline{\mu}$

$$\overline{\mu} \approx \mu := \tfrac{(p\sin\varphi)^2}{2mB}\,. \tag{2.22}$$

So, an interesting additional requirement is to include the magnetic moment in the reduced coordinates. For the equation to be solved, this requirement is expressed by

$$\overline{\Gamma}^\theta := \overline{\mu}\,. \tag{2.23}$$

Indeed, then the reduced Poisson bracket verifies

$$\overline{\mathbb{J}}^{\mu\theta} = 1, \text{ and } \overline{\mathbb{J}}^{\mu i} = 0\,, \tag{2.24}$$

for $i \neq \theta$, which means that the dynamics of $\overline{\mu}$ is zero

$$\dot{\overline{\mu}} = \{\overline{H}, \overline{\mu}\} = -\partial_\theta \overline{H} = 0\,,$$

since the Hamiltonian $\overline{H}$ does not depend on $\overline{\theta}$, as a consequence of (2.20).

There is a third reason for including the magnetic moment among the reduced coordinates and requiring $\overline{\Gamma}^\theta = \overline{\mu}$. It is concerned with the Hamiltonian structure of guiding-center dynamics. The 6-dimensional reduced motion $\dot{\overline{\mathbf{z}}}$ is Hamiltonian, since it is just the transform of the Hamiltonian motion $\dot{\mathbf{z}}$. But the true reduced guiding-center motion is the 4-dimensional slow motion $(\dot{\overline{\mathbf{q}}}, \dot{\overline{\varphi}})$. It is the truncation of the full dynamics $\dot{\overline{\mathbf{z}}}$, but truncations of a Hamiltonian dynamics are in general not Hamiltonian. However, in some cases, truncations are automatically Hamiltonian, and a special case is the quarter-canonical structure of the Poisson bracket, defined by conditions (2.24). As will be shown in Sec. 8.4.2, this can be seen by imposing Dirac's constraints $(\overline{\mu}, \overline{\theta})$ to the reduced dynamics, or by verifying that the truncated bracket is actually just given by starting from the initial Lie algebra of all functions of the phase space $f(\overline{\mathbf{z}})$, and taking the subalgebra of functions that do not depend on the gyro-angle $f(\overline{\mathbf{q}}, \overline{\varphi}, \overline{\mu})$. Thus, including $\overline{\mu}$ in the reduced coordinates is a way to guarantee the reduced slow motion to be Hamiltonian.

The requirement on the magnetic moment fixes one of the freedoms involved in the average components of the coordinate change. The other freedoms are used to make the reduced dynamics as simple as possible, by putting to zero as many average components of $\overline{\Gamma}_n$ as possible for higher $n$, as indicated by Eq. (2.21).

D) To sum it up, the guiding-center reduction involves a hierarchy of requirements: The primary requirement (minimal reduction, with an averaged reduced dynamics) is to remove the fast time-scale by averaging the Lagrangian over the gyro-angle; the corresponding equation is (2.20). The secondary requirement (intermediate reduction, with a constant of motion and a Hamiltonian slow reduced motion) is to include the magnetic moment among the reduced coordinates by the quarter-canonical structure; the corresponding equation is (2.23). The third optional requirement (maximal reduction, with a simplified reduced dynamics) is to use the remaining freedoms to make the reduced Lagrangian as simplified as possible; the corresponding equation is (2.21).

This makes seven requirements at each order in $\varepsilon$, one for each component of the reduced Lagrangian $\overline{\Gamma}_n$ in Eqs. (2.16), and seven freedoms are needed. For a time-independent transformation, those are $\mathbf{G}_n^{\mathbf{z}}$ and $S_n$, as announced in the previous subsection.

## 2.2 Derivation of the reduction

Let us now turn to the guiding-center reduction. The details of the procedure as well as the practical computations may seem intricate and they hide somehow that the basic ideas of the reduction are very elementary. It is why the principles and general lines of the procedure are presented in the appendix, to give a clear view of the reduction process.

In this section, the three stages of the method presented in the appendix are shown to work with the coordinate $\mathsf{c}$ in a similar way as with the standard approach relying on a gyro-gauge. The transformation at lowest orders is computed for comparison with previous works, and it is shown how the reduction can be performed to arbitrary order in the Larmor radius by obtaining explicit induction relations.

Each order of the derivation can be given several numbers. For instance, what is usually called the first order is the order just after the lowest order. For the Lagrangian, it corresponds to $\Gamma_0$ (since the lowest order corresponds to $\Gamma_{-1}$), which is rather considered here as the order 0. In addition, the order in the various quantities will be mixed up: for instance, in the derivation, the order involving $\Gamma_2$ will be the equation for $\mathbf{G}_3^{\mathsf{a},\mathsf{c}}$, as well as for $\mathbf{G}_2^{\mathsf{b},\phi}$ and also for $\mathbf{G}_1^\theta$. For the sake of clarity, we will always consider the order $n$ as the one corresponding to $\Gamma_n$ (or rather $\overline{\Gamma}_n$), and we will often use the expression "at order $\Gamma_n$", instead of "at the order corresponding to $\Gamma_n$".

### 2.2.1 Preliminary transformation and initial setting

The goal is to solve Eq. (2.19) for the guiding-center reduction, with the Lagrangian (2.15), and with the requirements (2.20)-(2.23) for the averaging reduction, for the magnetic moment reduction, and for the maximal reduction.

First of all, the change of coordinates from the norm of the momentum $p$ to the magnetic moment $\overline{\mu}$ is not near identity, as is clear in Eq. (2.22). Before beginning the reduction, a preliminary change of coordinates must be done, so that all the remaining transformation will be near identity. A suitable preliminary change of coordinates is

$$(\mathbf{q}, \varphi, p, \mathsf{c}) \longrightarrow (\mathbf{q}, \varphi, \mu, \mathsf{c}) \,,$$

where $\mu$ is the zeroth-order magnetic moment

$$\mu = \tfrac{(p \sin\varphi)^2}{2mB} \,, \tag{2.25}$$

which is a well-known adiabatic invariant, and often confounded with $\overline{\mu}$. In the new variables, the Lagrangian becomes

$$\Gamma := \left[e\mathbf{A} + \sqrt{2\mu m B}(\mathsf{b}\cot\varphi + \mathsf{c})\right]\cdot d\mathbf{q} - \mu B(1 + \cot^2\varphi)dt \,.$$

Interestingly, the pitch-angle $\varphi$ intervenes only through its cotangent, which was mentioned in Chapter 1 as making all quantities polynomials. Here, this feature is obvious in the Lagrangian, and it is preserved by derivatives, so that it will be preserved throughout all of the derivation. Actually, the magnetic moment makes the polynomiality still more accurate than in Chapter 1, where the variables $\varphi$ and $p$ were used, and the Larmor-radius prefactor $r_L = p\sin\varphi$ was not polynomial in $\cot\varphi$. Here, the magnetic moment $\mu$ absorbs the $p\sin\varphi$ and all formulae will be purely polynomials in $\cot\varphi$ and monomials in $\sqrt{\mu}$, which is useful to simplify computations.

So, we actually choose to change coordinates according to

$$(\mathbf{q}, \varphi, p, \mathsf{c}) \longrightarrow (\mathbf{q}, \phi, \mu, \mathsf{c}) \,,$$

with

$$\phi := \cot\varphi \tag{2.26}$$

the variable that makes all formulae polynomials.

Also, the structure of the Lagrangian shows that one can make the coefficients $e$ and $m$ disappear by noticing that the magnetic field $B$ appears only through $eB$, provided $\mu$ is considered as appearing only through $\mu m/e$, and $dt$ appears only through $dt/m$.

The particle charge $e$ is usually kept in guiding-center works because the order in $e^{-1}$ indicates the order in $\varepsilon$ [121, 122]. Here, it is useless since the order in $\varepsilon$ is already indicated by the order in other quantities: $\Gamma_n$ and $\overline{\Gamma}_n$ are of order $\frac{m\mu}{e}\left(\sqrt{\frac{m\mu}{e^2B}}\nabla\right)^n$. The reason is that all quantities will be series in $r_L\nabla = \sqrt{\frac{2m\mu}{e^2B}}\nabla$, as a result of the structure of the Lagrangian, and as will be confirmed by the derivation, e.g. Eqs. (2.37), (2.48)-(2.50), (2.60)-(2.62), (2.63)-(2.64), etc. Hence the order can be readily obtained by the overall order in $\sqrt{B}$ or $\nabla$. These last two quantities have the drawback of being an operator, or a space-dependent function; in addition, their order is only dimensional, e.g. $\frac{\nabla B \nabla B}{\sqrt{B}}$ is of order $\nabla^2 B^{3/2}$. For a readily control of the order, it is useful to have a scalar parameter, which was considered as $e^{-1}$ in previous works. Here, this role can be played by $\sqrt{\mu}$, since all quantities will be monomial in it. Thus, keeping $e$ to indicate the order of expansion is indeed not necessary.

Thus, we make the scaling

$$\begin{aligned} A &\longrightarrow \underline{A} := eA \\ B &\longrightarrow \underline{B} := eB \\ \mu &\longrightarrow \underline{\mu} := \tfrac{\mu m}{e} = \tfrac{(p\sin\varphi)^2}{2eB} \\ t &\longrightarrow \underline{t} := \tfrac{t}{m} \,, \end{aligned} \tag{2.27}$$

which avoids unnecessary coefficients in the derivation. It agrees with the physics, where the effect of the magnetic field on particle dynamics always includes the coupling constant $e$. For simplicity, we will drop the underline, e.g. we will write $B$ for $\underline{B}$. The Lagrangian becomes

$$\Gamma := \left[\mathbf{A} + \sqrt{2\mu B}(\mathsf{b}\phi + \mathsf{c})\right]\cdot d\mathbf{q} - \mu B(1+\phi^2)dt\,.$$

The derivation starts from the Lagrangian with the expansion (2.15)

$$\Gamma := \Gamma_{-1} + \Gamma_0\,,$$

with

$$\begin{aligned}\Gamma_{-1} &= \mathbf{A}\cdot d\mathbf{q}\,,\\ \Gamma_0 &= \sqrt{2\mu B}\;(\mathsf{b}\phi + \mathsf{c})\cdot d\mathbf{q} - \mu B(1+\phi^2)dt\,.\end{aligned} \tag{2.28}$$

It can be divided into its average and fluctuating part

$$\begin{aligned}\mathtt{avg}(\Gamma_{-1}) &= \Gamma_{-1} = \mathbf{A}\cdot d\mathbf{q}\\ \mathtt{avg}(\Gamma_0) &= \sqrt{2\mu B}\;\mathsf{b}\phi\cdot d\mathbf{q} - \mu B(1+\phi^2)dt\\ \mathtt{osc}(\Gamma_0) &= \sqrt{2\mu B}\;\mathsf{c}\cdot d\mathbf{q}\,.\end{aligned} \tag{2.29}$$

The previous section showed that the process involves the Lagrange 2-form $\omega_n$. It can be split in three basic terms

$$\omega = d\Gamma = \omega_{-1} + \widetilde{\omega}_0 + \overline{\omega}_0\,, \tag{2.30}$$

with

$$\begin{aligned}\omega_{-1} &:= d\Gamma_{-1} = d(\mathbf{A})\cdot \wedge d\mathbf{q}\\ \widetilde{\omega}_0 &:= d\mathtt{osc}(\Gamma_0) = d(\sqrt{2\mu B}\mathsf{c})\cdot \wedge d\mathbf{q}\\ \overline{\omega}_0 &:= d\mathtt{avg}(\Gamma_0) = d(\sqrt{2\mu B}\mathsf{b}\phi)\cdot \wedge d\mathbf{q} - d\big(\mu B(1+\phi^2)\big) \wedge dt\,,\end{aligned}$$

where the symbol $\wedge$ denotes the antisymmetry operator

$$da.b. \wedge dc = da.b.dc - dc.b.da\,,$$

for any matrix $b$ and any vectors $a$ and $b$.

The first term in (2.30) is the lowest-order Lagrange 2-form, related to Larmor gyrations. The second (resp. third) term in (2.30) is the exterior derivative of the oscillating (resp. average) zeroth-order Lagrangian. Be careful, this is not the oscillating (resp. average) zeroth-order Lagrange 2-form; for instance

$$\widetilde{\omega}_0 := d\big(\mathtt{osc}(\Gamma_0)\big) \neq \mathtt{osc}(\omega_0) = \mathtt{osc}(d\Gamma_0)\,,$$

because the exterior derivative does not preserve gyro-fluctuations.

The contributions (2.30) to the Lagrange 2-form are explicitly given by

$$\begin{aligned}\omega_{-1} :&= d\Gamma_{-1} = d\mathbf{q}\cdot\nabla(\mathbf{A})\cdot \wedge d\mathbf{q} = d\mathbf{q}\cdot(-\mathbf{B}) \times d\mathbf{q}\\ \overline{\omega}_0 :&= d\mathtt{avg}(\Gamma_0) = d(\sqrt{2\mu B}\mathsf{b}\phi)\cdot \wedge d\mathbf{q} - d(\mu B(1+\phi^2)) \wedge dt\\ &= \sqrt{2\mu B}\left\{d\phi\mathsf{b} + d\mathbf{q}\cdot\nabla\mathsf{b} + \left[\tfrac{d\mathbf{q}\cdot\nabla B}{2B} + \tfrac{d\mu}{2\mu}\right]\phi\mathsf{b}\right\}\cdot \wedge d\mathbf{q}\\ &\qquad - \left\{(1+\phi^2)(Bd\mu + \mu dB) + 2\mu B d\phi\right\} \wedge \tfrac{dt}{m}\\ \widetilde{\omega}_0 :&= d\mathtt{osc}(\Gamma_0) = d(\sqrt{2\mu B}\mathsf{c})\cdot \wedge d\mathbf{q}\\ &= \sqrt{2\mu B}\left\{d\mathsf{c} + \left[\tfrac{d\mathbf{q}\cdot\nabla B}{2B} + \tfrac{d\mu}{2\mu}\right]\mathsf{c}\right\}\cdot \wedge d\mathbf{q}\,.\end{aligned} \tag{2.31}$$

Now, $d\mathsf{c}$ involves two contributions: one corresponding purely to the gyro-angle, in which the variable $\mathsf{c}$ is changed at constant $\mathbf{q}$, and a second one coming from a change in the coordinate $\mathbf{q}$, which comes because the gyro-angle is a constrained coordinate, as mentioned about Eqs. (2.6) and (2.9). The second part is given by

$$d\mathbf{q}\cdot\nabla\mathsf{c} = -d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\mathsf{b} + d\mathbf{q}\cdot\mathbf{R}_g\mathsf{a}\,.$$

The first one is then $(d\mathsf{c} - d\mathbf{q}\cdot\nabla\mathsf{c})$, but it is written more precisely as $(d\mathsf{c} - d\mathbf{q}\cdot\nabla\mathsf{c})\cdot\mathsf{a}\mathsf{a}$ since when $\mathbf{q}$ is constant, the variation of $\mathsf{c}$ can only be in the direction of $\mathsf{a}$, i.e. along the circle $\mathbb{S}^1(\mathbf{q})$. On the whole

$$d\mathsf{c} = -d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\mathsf{b} + d\mathbf{q}\cdot\mathbf{R}_g\mathsf{a} + (d\mathsf{c} - d\mathbf{q}\cdot\nabla\mathsf{c})\cdot\mathsf{a}\mathsf{a}\,.$$

This shows that, in the gauge-independent approach, the natural form corresponding to $d\theta$ is

$$\delta\theta := -(d\mathsf{c} - d\mathbf{q}\cdot\nabla\mathsf{c})\cdot\mathsf{a}\,,$$

where the minus sign comes to agree with the usual convention (2.4) on the orientation of the gyro-angle. Unlike $d\theta$, this form is not closed in general; it is why it is written $\delta\theta$. In the gauge-dependent case, it is closed and indeed corresponds to $d\theta$, as a consequence of (2.8).

As the reduction process mainly relies on inversions of $\omega = d\Gamma$ (or rather $\overline{\omega}$), which can be viewed as a matrix inversion, a matrix notation is well suited and makes the discussion clearer. This implies to choose a basis of 1-forms; the derivation of $\omega$ above shows that a natural basis is

$$\Big(\mathsf{c}\cdot d\mathbf{q}, \mathsf{a}\cdot d\mathbf{q}, \mathsf{b}\cdot d\mathbf{q} \mid d\phi, d\mu, \delta\theta \mid -dt\Big)\,, \tag{2.32}$$

where a vertical dash $|$ is put to separate the space-, the momentum- and the time-components. The choice of $-dt$ makes the corresponding coordinate $H$ instead of $-H$.

Those 1-forms are not closed, unlike the standard $dz^i$, but it is not needed, provided one is careful of using intrinsic definitions for the operations involved in the procedure (see Eqs. (2.35) and (2.55) for instance).

In this basis, Eqs. (2.31) become

$$\omega_{-1} = B\left(\begin{array}{ccc|c|c} 0 & 1 & 0 & & \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & & \\ \hline & 0 & & 0 & 0 \\ \hline & 0 & & 0 & 0 \end{array}\right) \tag{2.33}$$

$$\overline{\omega}_0 = \left(\begin{array}{ccc|ccc|c} 0 & -J & I & 0 & 0 & 0 & \mathsf{c}\cdot\nabla H_0 \\ J & 0 & -K & 0 & 0 & 0 & \mathsf{a}\cdot\nabla H_0 \\ -I & K & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & \mathsf{b}\cdot\nabla H_0 \\ \hline 0 & 0 & \sqrt{2\mu B} & & & & \partial_\phi H_0 \\ 0 & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & & 0 & & \partial_\mu H_0 \\ 0 & 0 & 0 & & & & 0 \\ \hline -\mathsf{c}\cdot\nabla H_0 & -\mathsf{a}\cdot\nabla H_0 & -\mathsf{b}\cdot\nabla H_0 & -\partial_\phi H_0 & -\partial_\mu H_0 & 0 & 0 \end{array}\right)$$

$$\widetilde{\omega}_0 = \left(\begin{array}{ccc|ccc|c} 0 & -\tilde{J} & \tilde{I} & 0 & -\frac{\sqrt{2\mu B}}{2\mu} & 0 & \\ \tilde{J} & 0 & -\tilde{K} & 0 & 0 & \sqrt{2\mu B} & 0 \\ -\tilde{I} & \tilde{K} & 0 & 0 & 0 & 0 & \\ \hline 0 & 0 & 0 & & & & \\ \frac{\sqrt{2\mu B}}{2\mu} & 0 & 0 & & 0 & & 0 \\ 0 & -\sqrt{2\mu B} & 0 & & & & \\ \hline & 0 & & & 0 & & 0 \end{array}\right),$$

where for visual purpose, vertical and horizontal lines are used to separate the position-, momentum- and time-components. The spatial part of the matrix $\overline{\omega}_0$ (resp. $\tilde{\omega}_0$) is just a vector product $\mathbf{v}\times$,

with the vector $\mathbf{v} := I\mathsf{a} + J\mathsf{b} + K\mathsf{c}$ (resp. $\mathbf{v} := \tilde{I}\mathsf{a} + \tilde{J}\mathsf{b} + \tilde{K}\mathsf{c}$), where the coefficients are

$$\begin{aligned}
I &:= -\mathsf{a}\cdot\nabla\times(\sqrt{2\mu B}\phi\mathsf{b}) = \sqrt{2\mu B}\ \phi\Big(\tfrac{\mathsf{c}\cdot\nabla B}{2B} - \mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\Big)\\
J &:= -\mathsf{b}\cdot\nabla\times(\sqrt{2\mu B}\phi\mathsf{b}) = \sqrt{2\mu B}\ \phi\Big(\mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{c} - \mathsf{c}\cdot\nabla\mathsf{b}\cdot\mathsf{a}\Big)\\
K &:= -\mathsf{c}\cdot\nabla\times(\sqrt{2\mu B}\phi\mathsf{b}) = \sqrt{2\mu B}\ \phi\Big(\mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{a} - \tfrac{\mathsf{a}\cdot\nabla B}{2B}\Big)\\
\tilde{I} &:= -\mathsf{a}\cdot\nabla\times(\sqrt{2\mu B}\mathsf{c}) = \sqrt{2\mu B}\ \tfrac{\mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{a}-\mathsf{c}\cdot\nabla\mathsf{b}\cdot\mathsf{c}}{2}\\
\tilde{J} &:= -\mathsf{b}\cdot\nabla\times(\sqrt{2\mu B}\mathsf{c}) = \sqrt{2\mu B}\ \big(\tfrac{\mathsf{a}\cdot\nabla B}{2B} - \mathsf{c}\cdot\mathbf{R}_g\big)\\
\tilde{K} &:= -\mathsf{c}\cdot\nabla\times(\sqrt{2\mu B}\mathsf{c}) = \sqrt{2\mu B}\ \big(\mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{c} + \mathsf{b}\cdot\mathbf{R}_g\big)\ .
\end{aligned}$$

### 2.2.2 Immediate orders $-1$ and $0$

At lowest order $n = -1$, the reduced Lagrangian (2.19) writes

$$\overline{\Gamma}_{-1} = \Gamma_{-1} + dS_{-1}\,.$$

The change of variable has no effect at this order, since it is near-identity. But the averaging condition is trivially verified; it is actually a condition for the near-identity Lie transform to remove the fast time-scale, which is possible only when at lowest order, the Lagrangian is already gyro-averaged. Now, the only freedom involved $S_{-1}$ cannot be useful and is set to zero

$$S_{-1} = 0\,.$$

At the following order $n = 0$, the reduced Lagrangian is

$$\overline{\Gamma}_0 = \mathbf{G}_1\cdot\omega_{-1} + \Gamma_0 + dS_0\,. \tag{2.34}$$

To write vectors in matrix form, a basis has to be chosen for vector fields also. It is most convenient to choose the dual of the basis (2.32) for 1-forms, so that the coupling $\mathbf{G}_1\cdot\omega_{-1}$ is computed as a standard matrix product. The desired basis is easily identified as (using the natural isomorphism between vector fields and differential operators)

$$\big(\mathsf{c}\cdot\nabla, \mathsf{a}\cdot\nabla, \mathsf{b}\cdot\nabla \mid \partial_\phi, \partial_\mu, -\mathsf{a}\cdot\partial_{\mathsf{c}} \mid -\partial_t\big)\,, \tag{2.35}$$

because $\nabla$ can be written $\partial_{\mathbf{q}|\mathsf{c}} + \partial_{\mathbf{q}}\mathsf{c}\cdot\partial_{\mathsf{c}|\mathbf{q}}$. The operator $-\mathsf{a}\cdot\partial_{\mathsf{c}}$ is the generator of Larmor gyrations, as shown in Chapter 2. It just an intrinsic definition of the usual $\partial_\theta$. The set (2.35) is actually the natural basis for vector fields, which confirms the relevance of the basis (2.32) for 1-forms.

Be careful that the chosen basic 1-forms are not averages and they must be taken into account when computing averages or fluctuations of a quantity. For instance, it could seem that a vector (resp. a 1-form) with components $(1,0,0 \mid ,0,0,0 \mid 0)$ is averaged, whereas it is not, since it is equal to $\mathsf{c}\cdot\nabla$ (resp. $\mathsf{c}\cdot d\mathbf{q}$).

Then, Eq. (2.34) is easily computed in matrix form

$$\begin{aligned}
\overline{\Gamma}_0 &= \mathbf{G}_1\cdot\omega_{-1} + \Gamma_0 + dS_0 \qquad\qquad (2.36)\\
&= \ (\ -B\mathbf{G}_1^{\mathsf{a}} \quad B\mathbf{G}_1^{\mathsf{c}} \quad 0 \ \mid\ 0 \quad 0 \quad 0 \ \mid\ 0\ )\\
&\quad + (\ \sqrt{2\mu B} \quad 0 \quad \sqrt{2\mu B}\phi \ \mid\ 0 \quad 0 \quad 0 \ \mid\ \mu B(1+\phi^2)\ )\,,\\
&\quad + (\ \mathsf{c}\cdot\nabla S_0 \quad \mathsf{a}\cdot\nabla S_0 \quad \mathsf{b}\cdot\nabla S_0 \ \mid\ \partial_\phi S_0 \quad \partial_\mu S_0 \quad -\mathsf{a}\cdot\partial_{\mathsf{c}} S_0 \ \mid\ 0\ )
\end{aligned}$$

where, as usual, 1-forms are written as $1 * 7$ matrices: each column is an equation to be solved for the freedoms $\mathbf{G}_1$ and $S_0$ in such a way that $\overline{\Gamma}_0$ satisfies the desired requirements.

The lowest-order Lagrange 2-form $\omega_{-1}$ is linked to Larmor gyration, as appears in (2.31). It is not invertible. Only two components of $\mathbf{G}_1$ are involved in the equation: $\mathbf{G}_1^{\mathsf{c}}$ and $\mathbf{G}_1^{\mathsf{a}}$; and only two components of $\overline{\Gamma}$ can be controlled by these freedoms: $\overline{\Gamma}_1^{\mathsf{c}}$ and $\overline{\Gamma}_1^{\mathsf{a}}$. As announced in the appendix, the inversion is possible only under some conditions on the term $\mathsf{R}_0 = -\overline{\Gamma}_0 + \Gamma_0 + dS_0$, and the

solution, if it exists, is not unique.

More precisely, for the averaging requirement (2.20), the condition of the right-hand side of Eq. (2.91) being in the range of $\omega_{-1}$ is satisfied, since the only fluctuating terms are in $\overline{\Gamma}_1^{\mathsf{c},\mathsf{a}}$, which can precisely be controlled by the freedoms $\mathbf{G}_1^{\mathsf{c},\mathsf{a}}$. The solution imposes the fluctuating part of $\mathbf{G}_1^{\mathsf{c},\mathsf{a}}$, which is given by

$$\mathsf{osc}(\mathbf{G}_1^{\mathsf{c}}\mathsf{c} + \mathbf{G}_1^{\mathsf{a}}\mathsf{a}) := \tfrac{\sqrt{2\mu B}}{B}\mathsf{a}\,. \tag{2.37}$$

Notice that the oscillating and averaged parts of the components $\mathbf{G}_1^{\mathsf{c}}$ and $\mathbf{G}_1^{\mathsf{a}}$ must be dealt with together through the combination $\mathbf{G}_1^{\mathsf{c}}\mathsf{c} + \mathbf{G}_1^{\mathsf{a}}\mathsf{a}$. This is easily illustrated with an example vector $\mathbf{X} := \mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\mathsf{c}\cdot\nabla$: it is not a pure fluctuation, whereas its $\mathsf{c}$-component $\mathbf{X}^{\mathsf{c}} := \mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{c}$ is a pure fluctuation, and its fluctuating part is $\frac{1}{2}\mathsf{b}\cdot\nabla\mathsf{b}\cdot(\mathsf{c}\mathsf{c} - \mathsf{a}\mathsf{a})\cdot\nabla$, which mixes up the components $\mathbf{X}^{\mathsf{c}}$ and $\mathbf{X}^{\mathsf{a}}$.

At this point, the average part of $\overline{\Gamma}_1^{\mathsf{c},\mathsf{a}}$ remains free and must be identified by the secondary and tertiary requirements (2.21) and (2.23). The magnetic moment requirement (2.23) is not concerned here, since $\overline{\Gamma}^{\theta}$ is automatically zero. Last, for the requirement for a maximal reduction (2.21), only the components $\overline{\Gamma}_1^{\mathsf{c},\mathsf{a}}$ can be controlled, and setting them to zero imposes the average part of $\mathbf{G}_1^{\mathsf{c},\mathsf{a}}$

$$\mathsf{avg}(\mathbf{G}_1^{\mathsf{c}}\mathsf{c} + \mathbf{G}_1^{\mathsf{a}}\mathsf{a}) := 0\,. \tag{2.38}$$

Eqs. (2.37)-(2.38) give the traditional (lowest-order) Larmor radius, which is usually confused with the exact Larmor radius $r_L := \mathbf{q} - \bar{\mathbf{q}}$:

$$(r_L)_1 := (\mathbf{q} - \bar{\mathbf{q}})_1 = \mathbf{G}_1^{\mathsf{c}}\mathsf{c} + \mathbf{G}_1^{\mathsf{a}}\mathsf{a} + \mathbf{G}_1^{\mathsf{b}}\mathsf{b} = \tfrac{\sqrt{2\mu B}}{B}\mathsf{a} = \tfrac{p\sin\varphi}{B}\mathsf{b}\times\mathsf{c} = \tfrac{\mathbf{B}\times\mathbf{p}}{B^2}\,,$$

where the index denotes the order, and Eqs. (2.77) and (2.48) were implicitly used.

In a similar way, what is usually called guiding center in the literature actually corresponds to the first-order guiding center:

$$\bar{q} := q - \tfrac{\mathbf{B}\times\mathbf{p}}{B^2} + O(\epsilon^2)$$

where we recover the traditional formula for the (first-order) guiding center, since we remind that here $\mathbf{B}$ stands for $\underline{\mathbf{B}} = e\mathbf{B}$ (see Eq. (2.27)).

All other components of the first-order transformation generator $\mathbf{G}_1^{\mathsf{b},\phi,\mu,\theta}$ still remain undetermined. They embody the non-uniqueness of the matrix inverse $(\omega_{-1})^{-1}$ and the corresponding freedom will be useful for the solvability conditions at the next order. As a consequence, there will be some order mixing: the components $\mathsf{b}, \phi, \mu, \theta$ of $\mathbf{G}_1$ will be determined at higher order, at the same time as the components $\mathsf{c}, \mathsf{a}$ of $\mathbf{G}_2$.

Last, the gauge function $S_0$ cannot be useful and is set to zero

$$S_0 = 0\,.$$

The final zeroth-order reduced Lagrangian (2.36) writes

$$\overline{\Gamma}_0 = \sqrt{2\mu B}\mathsf{b}\phi\cdot d\mathbf{q} - \mu B(1+\phi^2)dt\,. \tag{2.39}$$

It is just the average of the zeroth-order initial Lagrangian $\Gamma_0$: $\overline{\Gamma}_0 = \mathsf{avg}(\Gamma_0)$. The zeroth-order reduced 2-form $\overline{\omega}_0$ is then

$$d\overline{\Gamma}_0 = \overline{\omega}_0\,,$$

which justifies the notation $\overline{\omega}_0$ introduced in (2.30).

It is easy to see that the matrix $\overline{\omega}_{-1} + \overline{\omega}_0$ is not invertible, for instance all its gyro-angle components are zero. So, the pivotal matrix for high orders $\mathsf{M}_\infty$ mentioned in the appendix (see page 90) is not identified yet.

### 2.2.3 Turning point: Order 1

At the following order $n = 1$, the reduced Lagrangian is given by (2.19), which is rather written with the unknown $\mathbf{G}_{1,2}$ on the left-hand side

$$\mathbf{G}_2 \cdot \omega_{-1} + \tfrac{\mathbf{G}_1}{2} \cdot (\widetilde{\omega}_0 + 2\overline{\omega}_0) = \overline{\Gamma}_1 - dS_1 \,.$$

As announced in the appendix, the pivotal matrix is the set of $\omega_{-1}$ and $(\widetilde{\omega}_0 + 2\overline{\omega}_0)/2$, acting on the set of unknown components of $(\mathbf{G}_2, \mathbf{G}_1)$. Again, it is not invertible; this corresponds to the case where the inversion is possible only under some integrability conditions on the right-hand side $\overline{\Gamma}_1 - dS_1$, to which one must add the set of $-\omega_{-1}$ and $-(\omega_0 + \overline{\omega}_0)/2$ acting on the set of components of $(\mathbf{G}_2, \mathbf{G}_1)$ that are already known, i.e. on $\mathbf{G}_1^{\mathsf{c},\mathsf{a}}$. And the solution, if it exists, is not unique.

More precisely, using a matrix notation and grouping in the left-hand side only the terms with unknown components of $(\mathbf{G}_2, \mathbf{G}_1)$ gives

$$\begin{aligned}
&\begin{pmatrix} \mathbf{G}_2^{\mathsf{c}} \\ \mathbf{G}_2^{\mathsf{a}} \\ \mathbf{G}_2^{\mathsf{b}} \\ \hline \mathbf{G}_2^{\phi} \\ \mathbf{G}_2^{\mu} \\ \mathbf{G}_2^{\theta} \\ \hline 0 \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c}
0 & B & 0 & & & & \\
-B & 0 & 0 & & 0 & & 0 \\
0 & 0 & 0 & & & & \\ \hline
& & & & & & \\
& 0 & & & 0 & & 0 \\
& & & & & & \\ \hline
& & & & & & \\
& 0 & & & 0 & & 0 \\
& & & & & & \\
\end{array}\right) \\
&+ \begin{pmatrix} 0 \\ 0 \\ \mathbf{G}_1^{\mathsf{b}} \\ \hline \mathbf{G}_1^{\phi} \\ \mathbf{G}_1^{\mu} \\ \mathbf{G}_1^{\theta} \\ \hline 0 \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c}
0 & -J_{21} & I_{21} & 0 & -\frac{\sqrt{2\mu B}}{4\mu} & 0 & \partial_c H_0 \\
J_{21} & 0 & -K_{21} & 0 & 0 & \frac{\sqrt{2\mu B}}{2} & \partial_a H_0 \\
-I_{21} & K_{21} & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & \partial_b H_0 \\ \hline
0 & 0 & \sqrt{2\mu B} & & & & \partial_\phi H_0 \\
\frac{\sqrt{2\mu B}}{4\mu} & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & & 0 & & \partial_\mu H_0 \\
0 & -\frac{\sqrt{2\mu B}}{2} & 0 & & & & 0 \\ \hline
-\partial_c H_0 & -\partial_a H_0 & -\partial_b H_0 & -\partial_\phi H_0 & -\partial_\mu H_0 & 0 & 0
\end{array}\right) \\
&= - \begin{pmatrix} \mathbf{G}_1^{\mathsf{c}} \\ \mathbf{G}_1^{\mathsf{a}} \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ \hline 0 \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c}
0 & -J_{21} & I_{21} & 0 & -\frac{\sqrt{2\mu B}}{4\mu} & 0 & \partial_c H_0 \\
J_{21} & 0 & -K_{21} & 0 & 0 & \frac{\sqrt{2\mu B}}{2} & \partial_a H_0 \\
-I_{21} & K_{21} & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & \partial_b H_0 \\ \hline
0 & 0 & \sqrt{2\mu B} & & & & \partial_\phi H_0 \\
\frac{\sqrt{2\mu B}}{4\mu} & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & & 0 & & \partial_\mu H_0 \\
0 & -\frac{\sqrt{2\mu B}}{2} & 0 & & & & 0 \\ \hline
-\partial_c H_0 & -\partial_a H_0 & -\partial_b H_0 & -\partial_\phi H_0 & -\partial_\mu H_0 & 0 & 0
\end{array}\right) \\
&+ \qquad \left( \begin{array}{ccc|ccc|c} \overline{\Gamma}_1^{\mathsf{c}} & \overline{\Gamma}_1^{\mathsf{a}} & \overline{\Gamma}_1^{\mathsf{b}} & \overline{\Gamma}_1^{\phi} & \overline{\Gamma}_1^{\mu} & \overline{\Gamma}_1^{\theta} & \overline{H}_1 \end{array} \right) \\
&- \qquad \left( \begin{array}{ccc|ccc|c} \partial_c S_1 & \partial_a S_1 & \partial_b S_1 & \partial_\phi S_1 & \partial_\mu S_1 & \partial_\theta S_1 & 0 \end{array} \right) ,
\end{aligned} \tag{2.40}$$

where the exponent $^T$ indicates matrix transpose. The coefficient $I_{21}$ is defined by $I_{21} := \frac{2I+1\tilde{I}}{2}$, and $J_{21}$ and $K_{21}$ are defined the same way. Last, for shortness, we used the short-hands

$$\begin{aligned}
\partial_c &:= \mathsf{c} \cdot \nabla \,, \\
\partial_a &:= \mathsf{a} \cdot \nabla \,, \\
\partial_b &:= \mathsf{b} \cdot \nabla \,, \\
\partial_\theta &:= -\mathsf{a} \cdot \partial_{\mathsf{c}} \,.
\end{aligned}$$

Be careful, $\partial_c$ is different from $\partial_{\mathsf{c}}$.

Let us have a word on the graphical presentation for these matrix products, because it may seem surprising at first glance and it will be frequently used. In principle, 1-forms are row-matrices, vectors are column-matrices, so that the pairing between them is just given by the matrix product. Now, if 2-forms are presented as a $2 * 2$ matrix, then the vector implied in the left-pairing must be written as a row-matrix, for the pairing to be just the usual matrix product. In this chapter, in order to save room in the manuscript, we wrote this row-matrix as the transpose of a column-matrix. By the way, it makes formulae easier to read, because each component of the vector is facing precisely the row of the matrix which it multiplies: for instance in the second term of Eq. (2.40), the third row $\mathbf{G}_1^{\mathsf{b}}$ of the vector $\mathbf{G}_1$ multiplies the third row $(-I_{21}, K_{21}, 0 \mid -\sqrt{2\mu B}, -\frac{\phi\sqrt{2\mu B}}{2\mu}, 0 \mid \partial_b H_0)$ of

the matrix $2\overline{\omega}_0 + \widetilde{\omega}_0$ which is on its right. So, this graphical presentation seems to be well-suited.

Eq. (2.40) can be simplified by removing the components that do not contribute

$$
\begin{aligned}
& \begin{pmatrix} \mathbf{G}_2^{\mathsf{c}} \\ \mathbf{G}_2^{\mathsf{a}} \end{pmatrix}^T \cdot \left( \begin{array}{ccc|ccc|c} 0 & B & 0 & 0 & 0 & 0 & 0 \\ -B & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \\
+ & \begin{pmatrix} \mathbf{G}_1^{\mathsf{b}} \\ \mathbf{G}_1^{\phi} \\ \mathbf{G}_1^{\mu} \\ \mathbf{G}_1^{\theta} \end{pmatrix}^T \cdot \left( \begin{array}{ccc|ccc|c} -I_{21} & K_{21} & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & H_0 \frac{\partial_b B}{B} \\ \hline 0 & 0 & \sqrt{2\mu B} & & & & 2\mu B\phi \\ \frac{\sqrt{2\mu B}}{4\mu} & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & & 0 & & \frac{H_0}{\mu} \\ 0 & -\frac{\sqrt{2\mu B}}{2} & 0 & & & & 0 \end{array} \right) \\
= - & \begin{pmatrix} \mathbf{G}_1^{\mathsf{c}} \\ \mathbf{G}_1^{\mathsf{a}} \end{pmatrix}^T \cdot \left( \begin{array}{ccc|ccc|c} 0 & -J_{21} & I_{21} & 0 & -\frac{\sqrt{2\mu B}}{4\mu} & 0 & H_0 \frac{\partial_c B}{B} \\ J_{21} & 0 & -K_{21} & 0 & 0 & \frac{\phi\sqrt{2\mu B}}{2} & H_0 \frac{\partial_a B}{B} \end{array} \right) \\
+ & \left( \begin{array}{ccc|ccc|c} \overline{\Gamma}_1^{\mathsf{c}} & \overline{\Gamma}_1^{\mathsf{a}} & \overline{\Gamma}_1^{\mathsf{b}} & \overline{\Gamma}_1^{\phi} & \overline{\Gamma}_1^{\mu} & \overline{\Gamma}_1^{\theta} & \overline{H}_1 \end{array} \right) \\
- & \left( \begin{array}{ccc|ccc|c} \partial_c S_1 & \partial_a S_1 & \partial_b S_1 & \partial_\phi S_1 & \partial_\mu S_1 & \partial_\theta S_1 & 0 \end{array} \right) .
\end{aligned}
\tag{2.41}
$$

This formula illustrates the typical form for $\overline{\Gamma}_n$ announced in the appendix. In the left-hand side, a matrix product involves the unknown components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$. In the right-hand side, there are two kinds of terms: the terms with the known components of $(\mathbf{G}_n, \mathbf{G}_{n-1}, ...)$, and the terms involving the freedoms to be determined by the integrability conditions, i.e. $\overline{\Gamma}_n$ and $S_n$.

The requirements concern the seven components of the reduced Lagrangian $\overline{\Gamma}_n^i$, which are ideally put to zero. The freedoms are embodied in six components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$ and in the gauge function $S_n$. More precisely, the primary requirement (averaging reduction) means that the fluctuating part of the reduced Lagrangian must be zero: $\mathtt{osc}(\overline{\Gamma}_n^i) = 0$; actually, the chosen basis for 1-forms is not averaged, so that each component cannot be averaged separately, e.g. neither $\mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{a}\ \mathsf{a}\cdot d\mathbf{q}$ nor $\mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{c}\ \mathsf{c}\cdot d\mathbf{q}$ are averages, but the sum of them $\mathsf{b}\cdot\nabla\mathsf{b}\cdot d\mathbf{q}$ is an average. The secondary requirement (magnetic moment reduction) means that the gyro-angle component of the Lagrangian must be the magnetic moment: $\overline{\Gamma}_1^\theta = \overline{\mu}$, and $\overline{\Gamma}_n^\theta = 0$ for $n \neq 1$. The tertiary requirement (maximal reduction) means that the average part of the reduced Lagrangian should be zero as well: $\mathtt{avg}(\overline{\Gamma}_n^i) = 0$ (except $\overline{\Gamma}_1^\theta$). When integrability conditions cannot be solved for so strong requirements, the tertiary requirement is released, but as little as possible, and one sets to zero as many components $\overline{\Gamma}_n^i$ as possible.

Let us solve Eq. (2.41). The pivotal matrix is not invertible. It has six columns and seven rows; the columns $\phi$ and $\mu$ are not linearly independent; in addition, its column for $\theta$ is zero, which means that $\overline{\Gamma}_1^\theta$ cannot be controlled by the freedoms $(\mathbf{G}_2, \mathbf{G}_1)$. So, the existence of solution of Eq. (2.41) is submitted to integrability conditions (by which the freedoms $S_1$ and $\overline{\Gamma}_1^i$ can be constrained), and then the existing solutions are not unique.

The integrability conditions are automatically satisfied. The overall equation for the column $\phi$ will be automatically zero, because the procedure will imply $\mathbf{G}_1^{\mathsf{b}} = 0$, and $S_1 = 0$.

As for the column $\theta$, it writes:

$$0 = -\tfrac{\sqrt{2\mu B}}{2}\mathbf{G}_1^{\mathsf{a}} + \overline{\Gamma}_1^\theta - \partial_\theta S_1 \,.$$

The averaging requirement means that $\partial_\theta S_1 = 0$, which implies

$$\mathtt{osc}(S_2) = 0 \,.$$

The average part of the equation then writes

$$\overline{\Gamma}_1^\theta = \tfrac{\sqrt{2\mu B}}{2}\mathbf{G}_1^{\mathsf{a}} = \mu \,, \tag{2.42}$$

which makes the automatically verified for this component. In fact, if $\mu$ had not been chosen as a preliminary coordinate, it is here that it could be identified.

As a side comment, it might seem that the right-hand side of Eq. (2.42) should be $\overline{\mu}$ instead of $\mu$, because of Eq. (2.23). However, in the derivation, all expressions for $\overline{\Gamma}$ are functions of $\mu$.

In the reduced system, the corresponding expressions will be the same functions evaluated on the reduced coordinate $\overline{\mu}$, as is well emphasized in [86]. So, Eq. (2.42) means that the first-order reduced Lagrangian $\overline{\Gamma}_1$ will actually have its $\overline{\theta}$ component equal to $\overline{\mu}$. It is why Eq. (2.42) verifies the requirement (2.23) on the magnetic moment. We will not insist more on this point.

With the result (2.42), the reduced Lagrangian will contain the term $\mu\delta\theta$, implying the presence of the vector $\mathbf{R}_g$, which is not determined. This feature already appeared in the computation of $\omega_0$, but the vector $\mathbf{R}_g$ involved in $\delta\theta = -(d\mathsf{c} - d\mathbf{q}\cdot\nabla\mathsf{c})\cdot\mathsf{a}$ came from a term $-d\mathsf{c}\cdot\mathsf{a}$, so that the overall contribution of the vector $\mathbf{R}_g$ actually cancelled. Just the same way now, the term $\mu\delta\theta$ is required to come from a total contribution

$$-\mu\mathsf{a}\cdot d\mathsf{c} = \mu(\delta\theta - d\mathbf{q}\cdot\mathbf{R}_g)\,.$$

This imposes a non-zero contribution for $\overline{\Gamma}^{\mathbf{q}}$:

$$\overline{\Gamma}^{\mathbf{q}} := \overline{\Gamma}'^{\mathbf{q}} - \mu d\mathbf{q}\cdot\mathbf{R}_g\,.$$

For a maximal reduction, minimizing $\overline{\Gamma}^{\mathbf{q}}$ then means minimizing $\overline{\Gamma}'^{\mathbf{q}}$ (and ideally setting it to zero).

For the remaining 5 columns, the pivotal $5*5$ matrix is invertible for the five unknowns $\left(\mathbf{G}_2^{\mathsf{c}}, \mathbf{G}_2^{\mathsf{a}}, \mathbf{G}_1^{\mathsf{b}}, \mathbf{G}_1^{\phi}, \mathbf{G}_1^{\mu}\right)$, as is clear through the following argument:

- The freedom $\mathbf{G}_1^{\mathsf{b}}$ controls the column for $\overline{\Gamma}_1^{\phi}$, since the coefficient $-\sqrt{2\mu B} = -p\sin\varphi$ is invertible and no other unknown component of $(\mathbf{G}_2, \mathbf{G}_1)$ appears in this column. This does not determines fully $\mathbf{G}_1^{\mathsf{b}}$, since the freedom $\mathsf{avg}(S_1)$ appears in the column, and $\mathsf{avg}(S_1)$ is to be identified by the column $\overline{\Gamma}_1^{\mu}$, in which $\mathbf{G}_1^{\mathsf{b}}$ appears again. Thus, the set of $\left(\overline{\Gamma}_1^{\phi}, \overline{\Gamma}_1^{\mu}\right)$ can be considered as a coupled set of equations for $\left(\mathbf{G}_1^{\mathsf{b}}, \mathsf{avg}(S_1)\right)$. But it is solvable; it implies that

$$\mathsf{avg}(S_1) = \mathcal{K}(\phi\sqrt{\mu}, \mathbf{q})\,, \tag{2.43}$$

is an arbitrary function of $\phi\sqrt{\mu}$ and $\mathbf{q}$, and that

$$\mathbf{G}_1^{\mathsf{b}} = \tfrac{1}{\sqrt{2\mu B}}\partial_\phi \mathsf{avg}(S_1)\,. \tag{2.44}$$

- In a similar way, the set of $\left(\overline{\Gamma}_1'^{\mathsf{b}}, \overline{\Gamma}_1^{t}\right)$ is a coupled set of equations for $\left(\mathbf{G}_1^{\phi}, \mathbf{G}_1^{\mu}\right)$, which is solvable.

- Indeed, the freedom $\mathbf{G}_1^{\phi}$ controls the column for $\overline{\Gamma}_1'^{\mathsf{b}}$, since the coefficient $\sqrt{2\mu B}$ is invertible. The solution for $\mathbf{G}_1^{\phi}$ is then parametrized by $\mathbf{G}_1^{\mu}$, which is still unknown, but appears in this column.

- Then the freedom $\mathbf{G}_1^{\mu}$ controls the column for $\overline{\Gamma}_1^{t} = \overline{H}_1$, because when inserting the solution for $\mathbf{G}_1^{\phi}$, the coefficient of $\mathbf{G}_1^{\mu}$ becomes just $B$, which is invertible.

- The freedom $\mathbf{G}_2^{\mathsf{a}}$ controls the column for $\overline{\Gamma}_1'^{\mathsf{c}}$, since the coefficient $-B$ is invertible and no other unknown component of $(\mathbf{G}_2, \mathbf{G}_1)$ appears in this column (because now $\mathbf{G}_1^{\mathsf{b}}$ and $\mathbf{G}_1^{\mu}$ are not unknowns any more).

- The freedom $\mathbf{G}_2^{\mathsf{c}}$ controls the column for $\overline{\Gamma}_1'^{\mathsf{a}}$, since the coefficient $B$ is invertible. The solution for $\mathbf{G}_2^{\mathsf{c}}$ is then parametrized by $\mathbf{G}_1^{\theta}$, which is still unknown, but appears in this column.

- Then, all the components of $\Gamma_1'$ remain free. They can be used to explore the various guiding-center representations [26]. Here, we are interested in a maximal reduction, which means to set them to zero, so that the optimal requirements are fulfilled. The freedom $\mathcal{K}$ in $\mathsf{avg}(S_1)$ (which is a parameter in the formulae obtained for $\mathbf{G}_1^{\mathsf{b}}$, $\mathbf{G}_1^{\phi}$, $\mathbf{G}_1^{\mu}$, $\mathbf{G}_2^{\mathsf{a}}$, and $\mathbf{G}_2^{\mathsf{c}}$) cannot be useful to improve anything and is set to zero:

$$\mathsf{avg}(S_1) = 0\,. \tag{2.45}$$

- The freedom $\mathbf{G}_1^{\theta}$ is still undetermined, it embodies the non-uniqueness of the solution implied by the pivotal matrix being not invertible. This freedom will be useful for the requirements at the following order. As a consequence, the order mixing will not include only two orders, but three of them, and the pivotal matrix will act on components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$. In addition, the orders

will not be solved independently, since the unknown $\mathbf{G}_1^\theta$ will be identified at order $\overline{\Gamma}_2$, whereas it already appeared in the equations at order $\overline{\Gamma}_1$, so that it is a parameter in the expression computed for $\mathbf{G}_2^{\mathsf{c}}$.

This procedure shows that the average and fluctuating parts of the equations are dealt with the same way, because the equations for $\mathbf{G}_n$ is algebraic. This is very different from the minimal guiding-center reduction by Lie transforming the equation of motion, whose equation relies on the operator $\partial_\theta$, and which easily controls the fluctuating part of the equation, but involves secular differential equations for the average part of the equation, as happened in Chapter 1.

At that point, it seems that the order $\overline{\Gamma}_1$ has been completed: it is indeed completely satisfactory in itself, since all of the requirements are perfectly fulfilled, with the resulting reduced Lagrangian $\overline{\Gamma}_1 = -\mu\mathsf{a}\cdot d\mathsf{c}$. However, the procedure will have to be slightly changed, because at the following order the secondary requirement for $\overline{\Gamma}_2^\theta$ can be controlled by no higher-order freedom; it can be controlled only by $\mathtt{avg}(\mathbf{G}_1^\mu)$, which was already determined above by the tertiary requirement for $\overline{\Gamma}_1^t$. These two requirements cannot be simultaneously fulfilled and one of them has to be dropped.

It is here that the requirements are not dealt with in the same way: as one of them must be dropped, the choice is imposed by the hierarchy and the secondary requirement must be preferred to the tertiary one.

As a consequence, $\mathtt{avg}(\mathbf{G}_1^\mu)$ must be let free at first order. It remains a parameter in $\mathbf{G}_2^{\mathsf{a}}$ and in $\overline{H}_1$. Another consequence is that the tertiary requirement for $\overline{\Gamma}_1^t$ has been lost, and $\overline{H}_1$ has a non-zero value. Actually, the equation for $\overline{H}_1$ was coupled with the one for $\overline{\Gamma}_1'^{\mathsf{b}}$, and the non-zero term can be put in either of these components of $\overline{\Gamma}$.

One can consider recovering a zero value for this term by using the freedom $\mathcal{K}$ available in $\mathtt{avg}(S_1)$, which had been arbitrarily fixed to zero in the process above in Eq. (2.45). But the corresponding equation for $\mathcal{K}$ has no solution. Indeed, requiring $\mathtt{avg}(\overline{H}_1) = 0$ is an equation for $\mathcal{K}$, which is the only available freedom

$$\left[\phi\partial_b + \tfrac{\nabla\cdot\mathsf{b}}{2}(1+\phi^2)\partial_\phi\right]\mathcal{K} = \tfrac{B}{\sqrt{2\mu B}}\mathtt{avg}(\mathbf{G}_1^\mu). \tag{2.46}$$

With the expression (2.60) for $\mathtt{avg}(\mathbf{G}_1^\mu)$, it can be studied by expansion in series $\mathcal{K}(\phi\sqrt{\mu}) = \sum_k \mathcal{K}_k(\phi\sqrt{\mu})^k$. Expanding the right- and left-hand side of the equation and equating the coefficient of the same orders in $\phi$ and $\mu$ gives a non-solvable equation. For instance, the coefficient of order $\phi^1\mu^1$ implies the following equation

$$\nabla\cdot\mathsf{b}\ \mathcal{K}_2 = -\mathsf{b}\cdot\nabla\times\mathsf{b}\,, \tag{2.47}$$

which has no solution for a general magnetic geometry. So, the only available freedom cannot be used to obtain the full reduction $\overline{\Gamma}_1'^{\mathsf{b}} = 0$ and $\overline{H}_1 = 0$. One of those components has to be non-zero.

Computing explicitly the expressions for the solution at order $\overline{\Gamma}_1$ according to the procedure identified above gives the following results

$$\mathtt{osc}(S_1) = \mathbf{G}_1^{\mathsf{b}} = 0 \tag{2.48}$$

$$\begin{aligned}
\mathtt{osc}(\mathbf{G}_1^\phi) &= \tfrac{\sqrt{2\mu B}}{B}(1+\phi^2)\left[\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} + \phi\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\right]\\
\mathtt{osc}(\mathbf{G}_1^\mu) &= \mu\tfrac{\sqrt{2\mu B}}{B}\left[-\tfrac{B'\mathsf{a}}{B} - \phi\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2} - 2\phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\right]
\end{aligned} \tag{2.49}$$

$$\begin{aligned}
\mathtt{avg}(\mathbf{G}_1^\phi) &= \tfrac{1}{\sqrt{2\mu B}}\left[\overline{\Gamma}_1'^{\mathsf{b}} + \mu\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\right] - \tfrac{\phi}{2\mu}\mathtt{avg}(\mathbf{G}_1^\mu)\\
\overline{H}_1 &= \sqrt{2\mu B}\phi\left[\overline{\Gamma}_1'^{\mathsf{b}} + \mu\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\right] + B\mathtt{avg}(\mathbf{G}_1^\mu)\\
\mathbf{G}_2^{\mathsf{a}} &= \tfrac{\mu}{B}\left[\phi\tfrac{7\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-9\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} - \phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\right] + \tfrac{1}{2\sqrt{2\mu B}}\mathtt{avg}(\mathbf{G}_1^\mu)\\
\mathbf{G}_2^{\mathsf{c}} &= -\tfrac{\sqrt{2\mu B}}{B}\tfrac{\mathbf{G}_1^\theta}{2} - \tfrac{\mu}{B}\bar{\mathsf{a}}\mathsf{c}'\mathsf{a}\,,
\end{aligned} \tag{2.50}$$

where following Littlejohn, a condensed notation is used for gradients: the curved prime is used to indicate gradients of the magnetic field, and the short overbar over a vector $\mathsf{c}$ or $\mathsf{a}$ indicates the matrix transpose (for the euclidean scalar product), so that $\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} := \mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{c}$ and $B'\mathsf{a} := \mathsf{a}\cdot\nabla B$. This notation is similar to the one used in [88], with a slightly adaptation to make it more explicit, in order to fit in with higher-order expressions as in Chapters 1 and 10. The context (the presence of $\mathsf{c}$ and $\mathsf{a}$, together with a contraction with other vectors) should avoid any confusion with other uses of primes and overbars. For safety, in this chapter, we emphasize the difference by the notation: the straight prime and the long overline do not indicate gradients or matrix transpose, e.g. in $\overline{\Gamma}'$.

Finally, the first-order reduced Lagrangian writes

$$\overline{\Gamma}_1 = -\mu\mathsf{a}\cdot d\mathsf{c} + \overline{\Gamma}_1'^{\mathsf{b}}\mathsf{b}\cdot d\mathbf{q} - \overline{H}_1 dt\,. \tag{2.51}$$

In the results above, the term $\overline{\Gamma}_1'^{\mathsf{b}}$ was kept free in order to include both of the choices considered above. As $\overline{\Gamma}_1'^{\mathsf{b}}$ is free, the natural choice for a maximal reduction is $\overline{\Gamma}_1'^{\mathsf{b}} = 0$; then the reduced Hamiltonian is non-zero:

$$\overline{H}_1 = \mu\sqrt{2\mu B}\phi\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2} + B\mathsf{avg}(\mathbf{G}_1^{\mu})\,.$$

Alternatively, $\overline{H}_1$ can be set to zero, by choosing

$$\overline{\Gamma}_1'^{\mathsf{b}} = -\frac{B}{\sqrt{2\mu B}\phi}\mathsf{avg}(\mathbf{G}_1^{\mu}) - \mu\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\,. \tag{2.52}$$

This last choice is possible only if the inversion of $\phi$ does not cause a singularity, i.e. if $\mathsf{avg}(\mathbf{G}_1^{\mu})$ has no overall contribution of order zero in $\phi$, which will have to be verified when $\mathsf{avg}(\mathbf{G}_1^{\mu})$ is identified (see Eq. (2.60)).

The first-order Lagrangian (2.51) induces the following first-order reduced Lagrange matrix, which will be part of the pivotal matrix at higher orders:

$$\overline{\omega}_1 = \left(\begin{array}{ccc|ccc|c} & \overline{\omega}_1^{\mathbf{q};\mathbf{q}} & & 0 & \mathbf{R}_g\cdot\mathsf{c} & 0 & \partial_c\overline{H}_1 \\ & & & 0 & \mathbf{R}_g\cdot\mathsf{a} & 0 & \partial_a\overline{H}_1 \\ & & & -\partial_\phi\overline{\Gamma}_1'^{\mathsf{b}} & -\partial_\mu\overline{\Gamma}_1'^{\mathsf{b}} + \mathbf{R}_g\cdot\mathsf{b} & 0 & \partial_b\overline{H}_1 \\ \hline 0 & 0 & \partial_\phi\overline{\Gamma}_1'^{\mathsf{b}} & 0 & 0 & 0 & \partial_\phi\overline{H}_1 \\ -\mathbf{R}_g\cdot\mathsf{c} & -\mathbf{R}_g\cdot\mathsf{a} & -\mathbf{R}_g\cdot\mathsf{b} + \partial_\mu\overline{\Gamma}_1'^{\mathsf{b}} & 0 & 0 & 1 & \partial_\mu\overline{H}_1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hline -\partial_c\overline{H}_1 & -\partial_a\overline{H}_1 & -\partial_b\overline{H}_1 & -\partial_\phi\overline{H}_1 & -\partial_\mu\overline{H}_1 & 0 & 0 \end{array}\right), \tag{2.53}$$

where the matrix $\overline{\omega}_1^{\mathbf{q};\mathbf{q}}$ is defined by

$$d\mathbf{q}\cdot\overline{\omega}_1^{\mathbf{q};\mathbf{q}}\cdot d\mathbf{q} := d\mathbf{q}\cdot\Big[-\mu\nabla\mathsf{b}\cdot\mathsf{a}\wedge\bar{\mathsf{c}}\mathsf{b}' - \Big(\nabla\times\big(\overline{\Gamma}_1'^{\mathsf{b}}\mathsf{b}\big)\Big)\times\Big]d\mathbf{q}\,.$$

At the end of the first-order analysis, three freedoms remain: $\mathbf{G}_1^{\theta}$, which is a parameter in $\mathbf{G}_2^{\mathsf{c}}$; $\mathbf{G}_1^{\mu}$, which is a parameter in $\mathbf{G}_1^{\phi}$ , in $\mathbf{G}_2^{\mathsf{a}}$ and in $\overline{H}_1$; and either $\overline{\Gamma}_1'^{\mathsf{b}}$ or $\overline{H}_1$, which is a parameter in the other one.

The results (2.48)-(2.50) have physical implications. For instance, $\mathbf{G}_2^{\mathsf{c},\mathsf{a}}$ determines (together with $\mathbf{G}_1$) the perpendicular component of the second-order Larmor radius through the formula $(\mathbf{q}-\overline{\mathbf{q}})_2 = \mathbf{G}_2^{\mathbf{q}} - \frac{1}{2}\mathbf{G}_1^{\mathbf{z}}\cdot\partial_{\mathbf{z}}\mathbf{G}_1^{\mathbf{q}}$. Its averaged contribution will imply that the Larmor radius is not a pure fluctuation, and later on, a non-zero $\mathbf{G}_2^{\mathsf{b}}$ will be obtained, which will imply that the Larmor radius is not purely transverse to the magnetic field.

Those facts are well known in guiding-center works, and we do not insist on them. Here, we focus on the mechanism of the reduction, to show how it can be performed to arbitrary order in the Larmor radius using gauge-independent coordinates for the gyro-angle, and why it can be considered as a maximal reduction.

This point is the turning point of all the reduction: the matrix $\overline{\omega}_{-1} + \overline{\omega}_0 + \overline{\omega}_1$ is invertible.

Thus, in the induction for high orders, the pivotal matrix $\mathsf{M}_\infty$ will be the set of $\big(\overline{\omega}_{-1}, \overline{\omega}_0, \overline{\omega}_1\big)$ acting on the set of unknown components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$.

It means that with respect to the reduction procedure announced in 2.3.5, the first stage of the reduction is achieved, and the order at which $\overline{\omega}_{-1}+\overline{\omega}_0+\ldots+\overline{\omega}_{n_b}$ becomes invertible is $n_b := 1$. So, the critical order at which the pivotal matrix becomes the same at each order is $n_c \leqslant 2n_b + 2 = 4$ (see the appendix for a more detailed analysis of the reduction procedure, with especially the role of the orders $n_b$ and $n_c$, introduced on page 90). Accordingly, the algorithm for the derivation at high orders (third stage mentioned in the appendix) can be identified by now, but it will be efficient only for orders $n \geqslant 4$. The intermediate orders (named "second stage" in the appendix) must be studied separately.

In order to introduce the derivation order by order, we first go through the second stage and postpone the third stage, but it is important to notice that this last is independent of the second stage and could be studied before. Especially, all the high-orders algorithm relies on the matrix $\mathsf{M}_\infty$, together with the differential operators involved in $dS_n$. They are already known by now, and are determined by the choices that have been made previously, and mainly by the choices at order 1.

### 2.2.4 Core of the second stage: order 2

At the following order $n = 2$, the reduced Lagrangian is given by (2.19)

$$\overline{\Gamma}_2 = \mathbf{G}_3\cdot\omega_{-1} + \mathbf{G}_2\cdot\overline{\omega}_0 + \tfrac{\mathbf{G}_1}{6}\cdot d\big[\mathbf{G}_1\cdot(3\overline{\omega}_0 + 2\tilde{\omega}_0)\big] + dS_2\,. \tag{2.54}$$

At the previous orders, it appeared that for the algebraic part of the equations, the requirements on the average Lagrangian were dealt with exactly the same way as the requirements on the fluctuating Lagrangian. So, they will not be studied separately. Only the integrability condition will restore a difference between them for some of the unknowns.

At this point, the scheme for the unknown $\mathbf{G}_i$ is not purely algebraical: $\mathbf{G}_1$ is still not completely known, and it is involved in a first-order differential non-linear equation because of the term $i_{\mathbf{G}_1}d\lambda$, with

$$\lambda := \tfrac{1}{6}\mathbf{G}_1\cdot(3\overline{\omega}_0 + 2\tilde{\omega}_0)\,.$$

In Eq. (2.54), the exterior derivative must be computed for the 1-form $\lambda$ that is not explicitly known yet. Unlike in previous subsections, an explicit computation is not possible, and an abstract formula must be used; care must be taken that the usual formula (2.14) for exterior derivative cannot be used, because it is valid only when the basis is composed of closed 1-forms. Otherwise, it is replaced by the more general formula:

$$\begin{aligned} d\gamma = d(\gamma^j\mathbf{e}_j) &= d(\gamma^j)\wedge\mathbf{e}_j + \gamma^j d\mathbf{e}_j \\ &= \mathbf{e}_i(\partial_i\gamma^j - \partial_j\gamma^i)\mathbf{e}_j + \gamma^j d\mathbf{e}_j\,, \end{aligned} \tag{2.55}$$

for any 1-form $\gamma$, with components $\gamma^j$ in the basis $\mathbf{e}_j$.

In Eq. (2.55), two operations are to be identified: the action of the exterior derivative on scalar functions $d\gamma^j$ expressed in the chosen basis for 1-forms, and the exterior derivative of the basis $d\mathbf{e}_j$.

For the exterior derivative on scalar functions $d\gamma^j$, the contribution writes as usual

$$d(\gamma^j)\wedge\mathbf{e}_j = \mathbf{e}_i(\partial_i\gamma^j - \partial_j\gamma^i)\mathbf{e}_j\,,$$

provided the differential operators $\partial_i$ are given by the dual basis (2.35) to the chosen basis (2.32) of 1 forms $\mathbf{e}_i$.

As for the exterior derivatives of the basis $d\mathbf{e}_i$, they are easily computed as:

$$\begin{aligned} d(d\phi) &= d(d\mu) = 0 \\ d(\mathsf{b}\cdot d\mathbf{q}) &= -d\mathbf{q}\cdot(\nabla\times\mathsf{b})\times d\mathbf{q} \\ d(\mathsf{c}\cdot d\mathbf{q}) &= -d\mathbf{q}\cdot(\nabla\times\mathsf{c})\times d\mathbf{q} - \delta\theta\wedge\mathsf{a}\cdot d\mathbf{q} \\ d(\mathsf{a}\cdot d\mathbf{q}) &= -d\mathbf{q}\cdot(\nabla\times\mathsf{a})\times d\mathbf{q} + \delta\theta\wedge\mathsf{c}\cdot d\mathbf{q} \\ d(\delta\theta) &= -\bar{\mathsf{a}}\mathsf{b}'d\mathbf{q}\wedge\bar{\mathsf{c}}\mathsf{b}'d\mathbf{q} + d(\bar{\mathsf{a}}\mathsf{c}')\wedge d\mathbf{q}\,. \end{aligned}$$

In $d\delta\theta$, the derivative $\mathsf{a}\cdot\mathsf{c}'$ appears. Such terms should be avoided, since they are not determined, as shown by (2.9). But there is no trouble here, since in all the 1-forms in this derivation, the term $\mathbf{e}_\theta = -\mathsf{a}\cdot(d\mathsf{c} - d\mathbf{q}\cdot\nabla\mathsf{c})$ comes from $\mathsf{a}\cdot d\mathsf{c} = -\mathbf{e}_1 + d\mathbf{q}\cdot\nabla\mathsf{c}\cdot\mathsf{a}$. The exterior derivative of the second term will generates $d(\mathsf{a}\cdot\mathsf{c}')\wedge d\mathbf{q}$, which will automatically cancel the term $\mathbf{R}_g$ coming from $d\delta\theta$.

As a result, there is no need to compute those terms. Computing the exterior derivative of a 1-form $\gamma$ can be made according to the following procedure: write $\gamma = \gamma' - \gamma^\theta \mathbf{R}_g\cdot d\mathbf{q}$, where now the spatial components $\gamma'^{\mathbf{q}}\cdot d\mathbf{q}$ do not involve $\mathsf{a}\mathsf{c}'$. Then apply Eq. (2.55) to $\gamma'$ and to $-\gamma'^\theta\mathbf{R}_g\cdot d\mathbf{q}$. The cancellation of the terms containing $d(\mathbf{R}_g)\cdot\wedge d\mathbf{q}$ gives the resulting formula:

$$d\gamma = \mathbf{e}_i(\partial_i\gamma'^j - \partial_j\gamma'^i)\mathbf{e}_j + \gamma'^i d'\mathbf{e}_i - \mathbf{e}_i\partial_i\gamma'^\theta \wedge \mathbf{R}_g\cdot d\mathbf{q}\,, \tag{2.56}$$

in which $d'\mathbf{e}_i = d\mathbf{e}_i - \delta^{i\theta} d(\mathsf{a}\cdot\mathsf{c}')\wedge d\mathbf{q}$, where $\delta$ is the Kronecker delta, which means that $d'\mathbf{e}_i$ is exactly $d\mathbf{e}_i$ but without the problematic term in $d(\delta\theta)$.

With Eq. (2.56), the quadratic term in $\mathbf{G}_1$ is found to have its momentum components linear in the unknowns $\left(\mathbf{G}_1^\theta, \mathsf{avg}(\mathbf{G}_1^\mu)\right)$, and differential only for $\mathbf{G}_1^\theta$:

$$\begin{aligned}
6\big(i_{\mathbf{G}_1} d\lambda\big)^\phi &= -\mathbf{G}_1^{\mathsf{a}}\partial_\phi\big[2\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 2\mathbf{G}_1^\theta\big] \\
6\big(i_{\mathbf{G}_1} d\lambda\big)^\mu &= -\mathbf{G}_1^{\mathsf{a}}\partial_\mu\big[2\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 2\mathbf{G}_1^\theta\big] + 4\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 4\mathbf{G}_1^\theta \\
6\big(i_{\mathbf{G}_1} d\lambda\big)^\theta &= -\mathbf{G}_1^{\mathsf{a}}\partial_\theta\big[2\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 2\mathbf{G}_1^\theta\big] + 6\mathbf{G}_1^\mu + \mathbf{G}_1^{\mathsf{a}}.6J_{322}\,,
\end{aligned} \tag{2.57}$$

where the coefficient $J_{322}$ is defined by

$$J_{322} := \tfrac{\mu}{3}\big(3J + 2\tilde{J} + 2\mathbf{R}_g\cdot\mathsf{a}\big) = \mu\left[\phi(\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}) + \tfrac{B'\mathsf{a}}{3B}\right].$$

Notice that it does not depend on $\mathbf{R}_g$, precisely because the term $2\mathsf{a}\mathsf{c}'\mathsf{a}$ coming from Eq. (2.55) cancels the corresponding term in $\tilde{J}$.

The differential equation for $\mathbf{G}_1^\theta$ does not make things much more complicated, since it can be easily solved, e.g. by expansion in $\sqrt{\mu}$ and $\phi$. Alternatively, a trick can be used to make the scheme purely algebraical [24]; it is not essential for the derivation, but we will use it because it simplifies much the explanations.

The idea is to notice that in this case, the only differential operators involved come from the second term of the exterior derivative $(\partial_i\lambda^j - \partial_j\lambda^i)\,d\mathbf{y}^i\otimes d\mathbf{y}^j$, i.e. they are involved in expressions that write $-\mathbf{G}_1^i\partial_j\lambda^i d\mathbf{y}^j = -\mathbf{G}_1^i d\lambda^i$. In addition, in the sum over the index $i$, only one of the terms involves a differential operator $d$ acting on the unknown $\mathbf{G}_1^\theta$, namely the term with $i = \mathsf{a}$. An integration by parts over $-\mathbf{G}_1^{\mathsf{a}} d\lambda^{\mathsf{a}}$ can transfer the differential operator over the pre-factor $\mathbf{G}_1^{\mathsf{a}}$, which is already known. The resulting equation is algebraic.

The integration by parts is justified by the gauge function. The equation for $\overline{\Gamma}_2$ can be added a total derivative, which can be chosen $d(\mathbf{G}_1^{\mathsf{a}}\lambda^{\mathsf{a}})$ and extracted from $dS_2$ by $dS_2 = dS_2' + d(\mathbf{G}_1^{\mathsf{a}}\lambda^{\mathsf{a}})$:

$$-\mathbf{G}_1^{\mathsf{a}} d\lambda^{\mathsf{a}} + dS_2 = -\mathbf{G}_1^{\mathsf{a}} d\lambda^{\mathsf{a}} + dS_2' + d(\mathbf{G}_1^{\mathsf{a}}\lambda^{\mathsf{a}}) = dS_2' + \lambda^{\mathsf{a}} d\mathbf{G}_1^{\mathsf{a}}\,,$$

and the right-hand side is not differential any more, but algebraic for $\mathbf{G}_1^\theta$, which is contained in $\lambda^{\mathsf{a}}$.

Then, Eq. (2.54) becomes

$$\overline{\Gamma}_2 = \mathbf{G}_3\cdot\omega_{-1} + \mathbf{G}_2\cdot\overline{\omega}_0 + \Lambda + dS_2'\,, \tag{2.58}$$

with

$$\begin{aligned}
\Lambda :=& i_{\mathbf{G}_1} d\lambda + d(\mathbf{G}^{\mathsf{a}}\lambda^{\mathsf{a}}) \\
=& \left(\begin{array}{ccc|ccc|c} \Lambda_{\mathsf{c}} & \Lambda_{\mathsf{a}} & \Lambda_{\mathsf{b}} & 0 & 0 & 0 & \Lambda_t \end{array}\right) \\
&+ \begin{pmatrix} \mathbf{G}_1^{\mathsf{a}} \\ \mathbf{G}_1^\mu \\ \mathbf{G}_1^\theta \end{pmatrix}^T \cdot \left(\begin{array}{c|ccc|c} & 0 & \bar{\mathsf{a}}\mathsf{c}'\mathsf{a} & J_{322} & \\ 0 & 0 & 0 & 1 & 0 \\ & 0 & -1 & 0 & \end{array}\right),
\end{aligned}$$

where the momentum components are written in matrix form because they will determine the unknowns $\left(\mathbf{G}_1^\theta, \mathtt{avg}(\mathbf{G}_1^\mu)\right)$.

Now, the induction relation just relies on a pivotal matrix $\mathsf{M}_2$, which appears in the set of

$$\left(\mathbf{G}_3{\cdot}\omega_{-1}, \mathbf{G}_2{\cdot}\overline{\omega}_0, \Lambda\right),$$

in which a linear algebraic operator (a matrix) acts on the unknown components of $(\mathbf{G}_3, \mathbf{G}_2, \mathbf{G}_1)$.

The pivotal matrix in Eq. (2.58) is invertible in the sense that it determines six unknown components of $(\mathbf{G}_3, \mathbf{G}_2, \mathbf{G}_1)$, which is the maximum that can be done at each order. Remember the pivotal matrix cannot be fully invertible, since the transformation is time-independent, so that $\mathbf{G}_n$ is 6-dimensional, whereas the matrix has value in the 7-dimensional space $(\mathbf{q}, \mathbf{p}, t)$. In addition, the 7-dimensional pivotal matrix is anti-symmetric, hence not invertible. The seventh requirement is to be provided by the gauge function $S_2'$, which is the only integrability condition involved at this order.

More precisely, removing all the coefficients that do not contribute, as was done in (2.41), Eq. (2.58) becomes

$$\begin{aligned}
&\begin{pmatrix}\mathbf{G}_3^{\mathsf{c}}\\ \mathbf{G}_3^{\mathsf{a}}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & B & 0 & 0 & 0 & 0 & 0\\ -B & 0 & 0 & 0 & 0 & 0 & 0\end{array}\right)\\
&+\begin{pmatrix}\mathbf{G}_2^{\mathsf{b}}\\ \mathbf{G}_2^{\phi}\\ \mathbf{G}_2^{\mu}\\ \mathbf{G}_2^{\theta}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|cc|c} -I & K & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & H_0\frac{\partial_b B}{B}\\ \hline 0 & 0 & \sqrt{2\mu B} & & & & 2\mu B\phi\\ 0 & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & & 0 & & \frac{H_0}{\mu}\\ 0 & 0 & 0 & & & & 0\end{array}\right)\\
&+\begin{pmatrix}\mathbf{G}_1^{\mu}\\ \mathbf{G}_1^{\theta}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & -1 & 0 & 0\end{array}\right)\\
&=-\begin{pmatrix}\mathbf{G}_2^{\mathsf{c}}\\ \mathbf{G}_2^{\mathsf{a}}\\ \mathbf{G}_1^{\mathsf{a}}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & -J & I & 0 & 0 & 0 & H_0\frac{\partial_c B}{B}\\ J & 0 & -K & 0 & 0 & 0 & H_0\frac{\partial_a B}{B}\\ 0 & 0 & 0 & 0 & \bar{\mathsf{a}}\mathsf{c}\text{'}\mathsf{a} & J_{322} & 0\end{array}\right)\\
&-\left(\begin{array}{ccc|ccc|c} \Lambda_{\mathsf{c}} & \Lambda_{\mathsf{a}} & \Lambda_{\mathsf{b}} & 0 & 0 & 0 & \Lambda_t\end{array}\right)\\
&+\left(\begin{array}{ccc|ccc|c} \overline{\Gamma}_2^{\mathsf{c}} & \overline{\Gamma}_2^{\mathsf{a}} & \overline{\Gamma}_2^{\mathsf{b}} & \overline{\Gamma}_2^{\phi} & \overline{\Gamma}_2^{\mu} & \overline{\Gamma}_2^{\theta} & \overline{H}_2\end{array}\right)\\
&-\left(\begin{array}{ccc|ccc|c} \partial_c S_2' & \partial_a S_2' & \partial_b S_2' & \partial_\phi S_2' & \partial_\mu S_2' & \partial_\theta S_2' & 0\end{array}\right),
\end{aligned} \tag{2.59}$$

where the left-hand side contains just the terms involved in the matrix inversion to determine the unknown components of $\left(\mathbf{G}_3, \mathbf{G}_2, \mathbf{G}_1\right)$.

Eq. (2.59) is similar to (2.41), and the same comments can be done as at the previous order, in the two paragraphs following formula (2.41).

To solve the equation, an analysis similar to the one at the previous order leads to the procedure summarized in the following table, where each row corresponds to one of the equations. The component $\overline{\Gamma}_2^i$ of the reduced Lagrangian involved in the corresponding equation is indicated in the first column; the unknown which controls the equation and permits $\overline{\Gamma}_2^i = 0$ is indicated in the second column; the coefficient to be inverted is indicated in the third column.

| Equation | Unknown | Coefficient |
|---|---|---|
| $\mathtt{avg}(\overline{\Gamma}_2^\theta)$ | $\mathtt{avg}(\mathbf{G}_1^\mu)$ | $1$ |
| $\mathtt{osc}(\overline{\Gamma}_2^\theta)$ | $\mathtt{osc}(S_2')$ | $\partial_\theta$ |
| $\overline{\Gamma}_2^\phi$ | $\mathbf{G}_2^{\mathsf{b}}$ | $-\sqrt{2\mu B}$ |
| $\overline{\Gamma}_2^\mu$ | $\mathbf{G}_1^\theta$ | $-1$ |
| $\overline{\Gamma}_2^{\mathsf{b}}$ | $\mathbf{G}_2^\phi$ | $\sqrt{2\mu B}$ |
| $\overline{H}_2$ | $\mathbf{G}_2^\mu$ | $B$ |
| $\overline{\Gamma}_2^{\mathsf{c}}$ | $\mathbf{G}_3^{\mathsf{a}}$ | $-B$ |
| $\overline{\Gamma}_2^{\mathsf{a}}$ | $\mathbf{G}_2^{\mathsf{c}}$ | $B$ |

A few comments are in place. As had been announced in the derivation of Eq. (2.49), the new feature is that the equation for $\mathtt{avg}(\overline{\Gamma}_2^\theta)$ can be controlled only by the average first-order magnetic moment $\mathtt{avg}(\mathbf{G}_1^\mu)$. It is why it was not available at the previous order.

The coefficient $\partial_\theta$ is an operator, but it is invertible over gyro-fluctuations.

The set of $\left(\overline{\Gamma}_2^\phi, \overline{\Gamma}_2^t\right)$ is a coupled set of equations for $\left(\mathbf{G}_2^\phi, \mathbf{G}_2^\mu\right)$, but each of the unknowns can be assigned to one of the equation because the system can be solved in the following way. The freedom $\mathbf{G}_2^\phi$ controls the column for $\overline{\Gamma}_2^{\mathsf{b}}$, since the coefficient $\sqrt{2\mu B}$ is invertible. The solution for $\mathbf{G}_2^\phi$ is parametrized by $\mathbf{G}_2^\mu$, which is still unknown, but appears in this equation. Then, the freedom $\mathbf{G}_2^\mu$ controls the column for $\overline{\Gamma}_2^t = \overline{H}_1$, because when inserting the solution for $\mathbf{G}_2^\phi$, the coefficient of $\mathbf{G}_2^\mu$ becomes $B$, which is invertible.

At the end, the reduced Lagrangian $\overline{\Gamma}_2$ is free and can be set to zero, as required for a maximal reduction. The freedom $\mathtt{avg}(S_2')$ (which is a parameter in the formulae obtained for $\mathbf{G}_1^\theta$, $\mathbf{G}_2^{\mathsf{b}}$, $\mathbf{G}_2^\phi$, $\mathbf{G}_2^\mu$, $\mathbf{G}_3^{\mathsf{a}}$, and $\mathbf{G}_3^{\mathsf{c}}$) cannot be useful and is set to zero.

In a similar way as at the previous order, at that point, the order $\overline{\Gamma}_2$ is completely satisfactory in itself, since the reduced Lagrangian has been fully reduced $\overline{\Gamma}_2 = 0$.

However, the procedure will have to be changed, because at the following order the secondary requirement for $\overline{\Gamma}_3^\theta$ can be controlled only by $\mathtt{avg}(\mathbf{G}_2^\mu)$, which is therefore not available to get the tertiary requirement for $\overline{\Gamma}_2^t$. So, this last requirement has to be dropped.

Accordingly, $\mathtt{avg}(\mathbf{G}_2^\mu)$ remains free at this order, and it is a parameter in $\mathbf{G}_3^{\mathsf{a}}$ and in $\overline{\Gamma}_2^t$. Thus, $\overline{\Gamma}_2^t$ has a non-zero value, but its equation was coupled with the one for $\overline{\Gamma}_2^{\mathsf{b}}$, and the non-zero term can be put in either of these components of $\overline{\Gamma}_2$.

One can consider recovering a zero value for both $\overline{\Gamma}_2^t$ and $\overline{\Gamma}_2^{\mathsf{b}}$ by using the freedom of $\mathtt{avg}(S_2')$, whose value had been arbitrarily fixed to zero in the process above.

Indeed, if $\mathtt{avg}(S_n)$ is let as a free parameter in $\mathbf{G}_2^{\mathsf{b}}$, $\mathbf{G}_1^\theta$ and $\mathbf{G}_2^\phi$, then when replacing these variables by their expression, the requirement $\mathtt{avg}(\overline{H}_2) = 0$ becomes an equation for $\mathtt{avg}(S_2')$. Unfortunately, this equation is not easily studied.

As a first attempt, one can assume that $S_2'$ does not affect the equation for $\mathbf{G}_1^\theta$, as was observed in Eqs. (2.43)-(2.44). Then the equation for $\mathtt{avg}(S_2')$ has the same structure as Eq. (2.46). This equation may not be integrable, as it was the case for Eq. (2.46). In this case, one should relax the assumption for $\mathtt{avg}(S_2')$ not to affect the equation for $\mathbf{G}_1^\theta$. The differential equation for $\mathtt{avg}(S_2')$ then becomes more difficult to study, because $\mathbf{G}_1^\theta$ is involved in the 1-form $\Lambda$ in Eq. (2.59) in a rather intricate way.

Applying the procedure identified above gives

$$\mathtt{avg}(\mathbf{G}_1^\mu) = \mu\tfrac{\sqrt{2\mu B}}{B}\phi(\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a}) \tag{2.60}$$

$$\begin{aligned}
\mathtt{osc}(S_2') &= \mu\tfrac{\sqrt{2\mu B}}{B}\Big[-\tfrac{2B'\mathsf{c}}{3B} + \phi\tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{4} - 2\phi^2\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\Big]\\
\mathbf{G}_2^{\mathsf{b}} &= \tfrac{\mu}{B}\Big[\tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{4} - 4\phi\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\Big] + \tfrac{\partial_\phi\mathtt{avg}(S_2')}{\sqrt{2\mu B}}\\
\mathbf{G}_1^\theta &= \tfrac{\sqrt{2\mu B}}{B}\Big[\mathsf{a}\mathsf{c}'\mathsf{a} - \tfrac{B'\mathsf{c}}{B} + \phi\tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{4} - \phi^2\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\Big]\\
&\qquad + \partial_\mu\mathtt{avg}(S_2')
\end{aligned} \tag{2.61}$$

$$\begin{aligned}
\mathbf{G}_3^{\mathsf{a}} &= \tfrac{1}{B}\Big[\sqrt{2\mu B}\Big\{J\mathbf{G}_2^{\mathsf{a}} - I\mathbf{G}_2^{\mathsf{b}}\Big\} + S_2'\mathsf{c} + \Lambda^{\mathsf{c}}\Big]\\
\mathbf{G}_3^{\mathsf{c}} &= \tfrac{1}{B}\Big[\sqrt{2\mu B}\Big\{J\mathbf{G}_2^{\mathsf{c}} - K\mathbf{G}_2^{\mathsf{b}}\Big\} - S_2'\mathsf{a} - \Lambda^{\mathsf{a}}\Big]\\
\mathtt{osc}(\mathbf{G}_2^\mu) &= \tfrac{\mathtt{osc}}{-B}\Big[\mathsf{R}_2^t + \mathbf{G}_2^{\mathsf{b}}(1+\phi^2)\mu B'\mathsf{b} - \sqrt{2\mu B}\phi\mathsf{R}_2^{\mathsf{b}}\Big]\\
\mathtt{osc}(\mathbf{G}_2^\phi) &= \tfrac{\mathtt{osc}}{2\mu B}\Big[\phi\mathsf{R}_2^t + (1+\phi^2)\Big\{\mathbf{G}_2^{\mathsf{b}}\phi\mu B'\mathsf{b} - \sqrt{2\mu B}\mathsf{R}_2^{\mathsf{b}}\Big\}\Big]\\
\mathtt{avg}(\mathbf{G}_2^\phi) &= \tfrac{\mathtt{avg}}{2\mu B}\Big[-B\phi\mathbf{G}_2^\mu + \sqrt{2\mu B}\{\overline{\Gamma}_2^{\mathsf{b}} - \mathsf{R}_2^{\mathsf{b}}\}\Big]\\
\overline{H}_2 &= \mathtt{avg}\Big[\mathsf{R}_2^t + \mathbf{G}_2^{\mathsf{b}}(1+\phi^2)\mu B'\mathsf{b}\\
&\qquad + B\mathbf{G}_2^\mu + \sqrt{2\mu B}\phi\Big\{\overline{\Gamma}_2^{\mathsf{b}} - \mathsf{R}_2^{\mathsf{b}}\Big\}\Big],
\end{aligned} \tag{2.62}$$

where

$$\mathsf{R}_2 := \Lambda + dS_2' + \begin{pmatrix} \mathbf{G}_2^{\mathsf{c}} \\ \mathbf{G}_2^{\mathsf{a}} \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & -J & I & 0 & 0 & 0 & H_0 \frac{\partial_c B}{B} \\ J & 0 & -K & 0 & 0 & 0 & H_0 \frac{\partial_a B}{B} \end{array}\right).$$

With the results (2.60)-(2.62), the parameters involved in Eqs. (2.48)-(2.50) can be made explicit:

$$\mathsf{avg}(\mathbf{G}_1^{\phi}) = \frac{1}{\sqrt{2\mu B}} \left[ \overline{\Gamma}_1^{\mathsf{b}} + \mu \frac{1+2\phi^2}{2} (\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}) \right] \tag{2.63}$$

$$\begin{aligned} \overline{H}_1 &= \sqrt{2\mu B}\phi \left[ \overline{\Gamma}_1^{\mathsf{b}} - \mu \frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2} \right] \\ \mathbf{G}_2^{\mathsf{a}} &= \frac{\mu}{B} \left[ \phi \frac{5\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - 7\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} - \phi^2 \bar{\mathsf{a}}\mathsf{b}'\mathsf{b} \right] \\ \mathbf{G}_2^{\mathsf{c}} &= \frac{\mu}{B} \left[ -\frac{B'\mathsf{c}}{B} + \phi \frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{4} - \phi^2 \bar{\mathsf{c}}\mathsf{b}'\mathsf{b} \right] + \frac{\sqrt{2\mu B}}{B} \frac{\partial_\mu \mathsf{avg}(S_2')}{2} . \end{aligned} \tag{2.64}$$

The reduced first-order Hamiltonian for the choice $\overline{\Gamma}_2^{\mathsf{b}} = 0$ is

$$\overline{H}_1 = \mu \sqrt{2\mu B} \phi \frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2} .$$

The reverse choice $\overline{H}_1 = 0$ is possible, since its existence condition in Eq. (2.52) is satisfied, as is clear in (2.60). It corresponds to a component $\overline{\Gamma}_1^{\mathsf{b}}$ of the reduced Lagrangian given by

$$\overline{\Gamma}_1'^{\mathsf{b}} = \mu \frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2} , \tag{2.65}$$

which is regular in $\phi = 0$, as expected.

In Eqs. (2.60)-(2.64), the components $\mathbf{G}_1$ and $\mathbf{G}_2^{\mathsf{q}}$ have been completely computed and simplified for comparison with previous results, because it is the point where usual derivations stop.

For the other components $\mathbf{G}_2^{\phi}$, $\mathbf{G}_3^{\mathsf{c}}$, $\mathbf{G}_3^{\mathsf{a}}$, and $\overline{H}_3$, Eqs. (2.60)-(2.64) are explicit solutions. Their right-hand side involves only known quantities (or quantities that are free parameters for these relations), but it has not been not expanded and simplified. This can can be done in a straightforward way just by computing explicitly the terms involved, but we will not pursue in that direction, since the calculation for the $\Lambda_{\mathsf{c,a,b}}$ is lengthy, and useless for our purpose, which is just to show how the procedure can be performed to arbitrary order. In addition, they are the topic of a work by the authors of [23], and were already partly introduced in [128].

In the results, the term $\overline{\Gamma}_2^{\mathsf{b}}$ was kept free to include both of the choices considered above: setting the non-zero term in $\overline{H}_2$ just means choosing $\overline{\Gamma}_2^{\mathsf{b}} = 0$. The other choice $\overline{H}_2 = 0$ corresponds to

$$\overline{\Gamma}_2^{\mathsf{b}} = \mathsf{avg} \left[ \mathsf{R}_2^{\mathsf{b}} - \frac{1}{\sqrt{2\mu B}\phi} \left\{ B\mathbf{G}_2^{\mu} + \mathsf{R}_2^{t} + \mathbf{G}_2^{\mathsf{b}}(1+\phi^2)\mu B'\mathsf{b} \right\} \right] .$$

This last choice is possible only if the term inside the curled parentheses has no overall contribution of order zero in $\phi$.

Also $\mathsf{avg}(S_2')$ was kept free, because one can consider using it to obtain the full reduction with both $\overline{\Gamma}_2^{\mathsf{b}} = 0$ and $\overline{H}_2 = 0$: it will imply a differential equation for $\mathsf{avg}(S_2')$, which might not be integrable, just as happened at the previous order in Eq. (2.46), but it might be integrable, or partly integrable, and then provide the full reduction, or at least a stronger reduction. Otherwise, it can be set to zero.

At the end of the analysis at order $n = 2$, it remains one unknown, one freedom, and a binary choice. The unknown is $\mathsf{avg}(\mathbf{G}_2^{\mu})$; it is a parameter in $\mathbf{G}_2^{\phi}$ and in $\overline{H}_2$ (or $\overline{\Gamma}_2^{\mathsf{b}}$); it will be determined at the following order. The freedom is $\mathsf{avg}(S_2')$, which remains free, but could not be used to improve the reduction. The binary choice is that either $\overline{\Gamma}_2^{\mathsf{b}}$ or $\overline{H}_2$ is set to zero, the other is computed accordingly. This is very similar to what had occurred at the previous order, but now, the unknown $\mathbf{G}_2^{\theta}$ does not appear as a parameter, since it is not at all involved at this order, as is clear in Eq. (2.59).

### 2.2.5 End of the second stage: order 3

Let us turn now to the following order $\overline{\Gamma}_3$. The equation writes (2.19)

$$\overline{\Gamma}_3 = \mathbf{G}_4{\cdot}\omega_{-1} + \mathbf{G}_3{\cdot}\overline{\omega}_0 + \mathbf{G}_2{\cdot}\overline{\omega}_1 - \tfrac{(\mathbf{G}_2{\cdot}d)^2}{2}\Gamma_{-1} + \tfrac{(\mathbf{G}_1{\cdot}d)^2}{24}\mathbf{G}_1{\cdot}(3\omega_0 + \overline{\omega}_0) + dS_3\,.$$

The unknowns are components of $(\mathbf{G}_4, \mathbf{G}_3, \mathbf{G}_2)$. The pivotal matrix is not $\mathsf{M}_\infty$, i.e. not just given by $(\omega_{-1}, \overline{\omega}_0, \overline{\omega}_1)$, because of the correcting term $-\frac{(\mathbf{G}_2{\cdot}d)^2}{2}\Gamma_{-1}$. This term might make the derivation more difficult, since it is not algebraic, but non-linear and differential for $\mathbf{G}_2$:

$$-\tfrac{\mathbf{G}_2}{2}{\cdot}d(\mathbf{G}_2{\cdot}d\Gamma_{-1}) = -\tfrac{\mathbf{G}_2}{2}{\cdot}\big[d\big\{-B(\mathbf{G}_2^{\mathsf{a}}\mathsf{c}) + B(\mathbf{G}_2^{\mathsf{c}}\mathsf{a})\big\}{\cdot}\wedge d\mathbf{q}\big]\,.$$

It can be written in matrix form

$$-\tfrac{1}{2}\begin{pmatrix}\mathbf{G}_2^{\mathsf{c},\mathsf{a},\mathsf{b}}\\ \hline \mathbf{G}_2^{\phi,\mu,\theta}\\ \hline 0\end{pmatrix}^T \cdot \begin{pmatrix} M_{11} & M_{12} & 0\\ M_{21} & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\,,$$

where the matrices $M_{ij}$ have obvious definitions, and are independent of $\mathbf{G}_2^{\mu,\theta}$.

This is enough to show that this additional term can be transferred into the right-hand side, i.e. it involves only terms that are already known at each step of the computation.

When computing $\overline{\Gamma}_3^{\mathbf{p}}$ for the unknowns $\mathbf{G}_2^{\theta}$, $\mathtt{avg}(\mathbf{G}_2^{\mu})$, $\mathbf{G}_3^{\mathsf{b}}$ and $\mathtt{osc}(S_3)$, the only components of $\mathbf{G}_2$ involved in the correcting term are in $\mathbf{G}_2^{\mathbf{q}}$, which is already known at that point. Then, when computing $\overline{\Gamma}_3^{\mathbf{q},t}$ for the unknowns $\mathbf{G}_3^{\phi}$, $\mathbf{G}_4^{\mathsf{c}}$, $\mathbf{G}_4^{\mathsf{a}}$, and $\mathtt{osc}(\mathbf{G}_3^{\mu})$ all the components of $\mathbf{G}_2$ are involved in the correcting term, but they are all known at that point.

As a consequence, the correcting term can be put in the right-hand side of the equation, and the pivotal matrix is actually $\mathsf{M}_\infty$, i.e. the set of $(\omega_{-1}, \overline{\omega}_0, \overline{\omega}_1)$ acting on the unknown components of $(\mathbf{G}_4, \mathbf{G}_3, \mathbf{G}_2)$. This means that the critical order at which the pivotal matrix becomes the same at each order is $n_c := 3$, and the order $\overline{\Gamma}_3$ can be included in the third stage, with all higher orders, which is studied in the following section.

### 2.2.6 Third stage: algorithmic orders 4 and higher

Now, the second stage of the method mentioned in appendix is ended and the third stage is beginning, which means that the matrix to be inverted is always the same at any order $n \geqslant 3$, and it is indeed invertible. So, the reduction can be performed to arbitrary order. The only possible complication comes from the integrability condition for the gauge function $S_n$ (and possibly $\overline{\Gamma}_n$), but after settling it, the reduction process becomes fully algorithmic and unique.

#### Equation and algorithm

This is proven by induction. Let us suppose that at some order $n \geqslant 3$, the set of unknowns are

$$\mathbf{g}_n := \Big(\overline{\Gamma}_n, S_n, \mathbf{G}_{n-1}^{\theta}, \mathtt{avg}(\mathbf{G}_{n-1}^{\mu}), \mathtt{osc}(\mathbf{G}_n^{\mu}), \mathbf{G}_n^{\mathsf{b}}, \mathbf{G}_n^{\phi}, \mathbf{G}_{n+1}^{\mathsf{c}}, \mathbf{G}_{n+1}^{\mathsf{a}}\Big)\,, \tag{2.66}$$

which means that before that order, all lower-order quantities $\mathbf{g}_{i<n}$ are already determined, and that after that order, all higher-order quantities $\mathbf{g}_{i>n}$ will remain free parameters. This assumption is verified at order $n = 3$, which initializes the induction. As announced in the appendix, we have included the reduced Lagrangian $\overline{\Gamma}_n$ in the vector $\mathbf{g}_n$, because some components of $\overline{\Gamma}_n$ cannot be set to zero and have to be computed in the process.

The reduced Lagrangian is given by Eq. (2.95)

$$\mathbf{G}_{n+1}{\cdot}\omega_{-1} + \mathbf{G}_n{\cdot}\overline{\omega}_0 + \mathbf{G}_{n-1}{\cdot}\overline{\omega}_1 = \overline{\Gamma}_n - \mathsf{R}_n - dS_n\,,$$

where $\mathsf{R}_n$ indicates all terms of $...e^{\mathbf{G}_2}e^{\mathbf{G}_1}\Gamma$ that are of order $n$ but do not involve $\mathbf{G}_{n+1}$, $\mathbf{G}_n$ or $\mathbf{G}_{n-1}$:

$$\mathsf{R}_n := [...e^{\mathbf{G}_2}e^{\mathbf{G}_1}\Gamma]_n - [\mathbf{G}_{n+1}\cdot\omega_{-1} + \mathbf{G}_n\cdot\overline{\omega}_0 + \mathbf{G}_{n-1}\cdot\overline{\omega}_1]\,,$$

in which the index $n$ indicates the term of order $n$.

Then, the pivotal matrix $\mathsf{M}_\infty$ is the set of

$$\left(\omega_{-1}, \overline{\omega}_0, \overline{\omega}_1\right),$$

acting on the unknown components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$.

It is invertible in the sense that it determines six unknown components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$, which is the maximum that can be done at each order. The last requirement is provided by the integrability condition for the gauge function $S_n$.

More precisely, grouping as usual in the left-hand side only the terms with unknown components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$, and removing all the coefficients that do not contribute, the induction relation for $\overline{\Gamma}_n$ writes

$$\begin{aligned}
&\begin{pmatrix}\mathbf{G}^{\mathsf{c}}_{n+1}\\ \mathbf{G}^{\mathsf{a}}_{n+1}\end{pmatrix}^T \cdot\left(\begin{array}{ccc|ccc|c} 0 & B & 0 & 0 & 0 & 0 & 0\\ -B & 0 & 0 & 0 & 0 & 0 & 0\end{array}\right)\\
&+\begin{pmatrix}\mathbf{G}^{\mathsf{b}}_{n}\\ \mathbf{G}^{\phi}_{n}\\ \mathbf{G}^{\mu}_{n}\end{pmatrix}^T \cdot\left(\begin{array}{ccc|ccc|c} -I & K & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & H_0\frac{\partial_b B}{B}\\ 0 & 0 & \sqrt{2\mu B} & 0 & 0 & 0 & 2\mu B\phi\\ 0 & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & 0 & 0 & \frac{H_0}{\mu}\end{array}\right)\\
&+\begin{pmatrix}\mathsf{avg}(\mathbf{G}^{\mu}_{n-1})\\ \mathbf{G}^{\theta}_{n-1}\end{pmatrix}^T \cdot\left(\begin{array}{ccc|ccc|c} 0 & 0 & \partial_\mu\overline{\Gamma}^{\prime\mathsf{b}}_1 & 0 & 0 & 1 & \partial_\mu\overline{H}_1\\ 0 & 0 & 0 & 0 & -1 & 0 & 0\end{array}\right)\\
&= -\,\mathsf{R}'_n\\
&+\qquad\left(\begin{array}{ccc|ccc|c}\overline{\Gamma}^{\mathsf{c}}_n & \overline{\Gamma}^{\mathsf{a}}_n & \overline{\Gamma}^{\mathsf{b}}_n & \overline{\Gamma}^{\phi}_n & \overline{\Gamma}^{\mu}_n & \overline{\Gamma}^{\theta}_n & \overline{H}_n\end{array}\right)\\
&-\qquad\left(\begin{array}{ccc|ccc|c}\partial_c S_n & \partial_a S_n & \partial_b S_n & \partial_\phi S_n & \partial_\mu S_n & \partial_\theta S_n & 0\end{array}\right),
\end{aligned} \tag{2.67}$$

where the terms involving known components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$ have been grouped with $\mathsf{R}_n$:

$$\begin{aligned}
&\mathsf{R}'_n = \mathsf{R}_n\\
&-\begin{pmatrix}\mathbf{G}^{\mathsf{c}}_{n}\\ \mathbf{G}^{\mathsf{a}}_{n}\end{pmatrix}^T \cdot\left(\begin{array}{ccc|ccc|c} 0 & -J & I & 0 & 0 & 0 & H_0\frac{B'_{\mathsf{c}}}{B}\\ J & 0 & -K & 0 & 0 & 0 & H_0\frac{B'_{\mathsf{a}}}{B}\end{array}\right)\\
&-\left(\begin{array}{c}\mathbf{G}^{\mathsf{c}}_{n-1}\\ \mathbf{G}^{\mathsf{a}}_{n-1}\\ \mathbf{G}^{\mathsf{b}}_{n-1}\\ \hline \mathbf{G}^{\phi}_{n-1}\\ \mathsf{osc}(\mathbf{G}^{\mu}_{n-1})\\ 0\\ \hline 0\end{array}\right)^T \cdot\left(\begin{array}{ccc|ccc|c} & & & 0 & 0 & 0 & \partial_c\overline{H}_1\\ & \overline{\omega}_1^{\mathbf{q};\mathbf{q}} & & 0 & 0 & 0 & \partial_a\overline{H}_1\\ & & & -\partial_\phi\overline{\Gamma}^{\prime\mathsf{b}}_1 & -\partial_\mu\overline{\Gamma}^{\prime\mathsf{b}}_1 & 0 & \partial_b\overline{H}_1\\ \hline 0 & 0 & \partial_\phi\overline{\Gamma}^{\prime\mathsf{b}}_1 & 0 & 0 & 0 & \partial_\phi\overline{H}_1\\ 0 & 0 & \partial_\mu\overline{\Gamma}^{\prime\mathsf{b}}_1 & 0 & 0 & 1 & \partial_\mu\overline{H}_1\\ 0 & 0 & 0 & 0 & -1 & 0 & 0\\ \hline -\partial_c\overline{H}_1 & -\partial_a\overline{H}_1 & -\partial_b\overline{H}_1 & -\partial_\phi\overline{H}_1 & -\partial_\mu\overline{H}_1 & 0 & 0\end{array}\right).
\end{aligned} \tag{2.68}$$

Again, the same comments as the ones after Eq. (2.41) are in place. Also, when computing $\mathsf{R}_n$, Eq. (2.56) is to be used to account for the derivative of the chosen basis of 1-forms and for the cancellation of the derivatives of $\mathbf{R}_g$.

When solving the induction relation, the mechanism is the same as at order $\overline{\Gamma}_0$ for $\overline{\Gamma}^{\mathsf{c},\mathsf{a}}_n$, and the same as at order $\overline{\Gamma}_1$ for $\overline{\Gamma}^{\mathsf{b},\phi,t}_n$ and $\mathsf{osc}(\overline{\Gamma}^{\theta}_n)$. In addition, the new feature is the presence of $\overline{\omega}_1$ for $\overline{\Gamma}^{\mu}_n$ and $\mathsf{avg}(\overline{\Gamma}^{\theta}_n)$, but the mechanism is similar to what happens at order $\overline{\Gamma}_2$, in the sense that the pivotal coefficients are the same. Mainly, the procedure relies on three conjugation-like relations: $(\mathsf{c},\mathsf{a})$ are conjugated for $\omega_{-1}$; $(\mu,\theta)$ are half-conjugated for $\overline{\omega}_1$, in the sense that the structure is quarter-canonical; and $(\mathsf{b},\phi)$ are half-conjugated for the symplectic part of $\overline{\omega}_0$.

The same procedure can be applied, which is reminded in the following table.

| Equation | Unknown | Coefficient |
|---|---|---|
| $\mathtt{avg}(\overline{\Gamma}^{\theta}_n)$ | $\mathtt{avg}(\mathbf{G}^{\mu}_{n-1})$ | $1$ |
| $\mathtt{osc}(\overline{\Gamma}^{\theta}_n)$ | $\mathtt{osc}(S'_n)$ | $\partial_\theta$ |
| $\overline{\Gamma}^{\phi}_n$ | $\mathbf{G}^{\mathsf{b}}_n$ | $-\sqrt{2\mu B}$ |
| $\overline{\Gamma}^{\mu}_n$ | $\mathbf{G}^{\theta}_{n-1}$ | $-1$ |
| $\overline{\Gamma}^{\mathsf{b}}_n$ | $\mathbf{G}^{\phi}_n$ | $\sqrt{2\mu B}$ |
| $\overline{H}_n$ | $\mathbf{G}^{\mu}_n$ | $B$ |
| $\overline{\Gamma}^{\mathsf{c}}_n$ | $\mathbf{G}^{\mathsf{a}}_{n+1}$ | $-B$ |
| $\overline{\Gamma}^{\mathsf{a}}_n$ | $\mathbf{G}^{\mathsf{c}}_{n+1}$ | $B$ |

Then, all the components of $\Gamma_n$ remain free. This can be used to explore the various guiding-center representations at higher orders [26]. Here, we are interested in a maximal reduction, which means to set them to zero, so that the optimal requirements are fulfilled. The freedom $\mathtt{avg}(S_n)$, which is a parameter in the formulae obtained for $\mathbf{G}^{\theta}_{n-1}$, $\mathbf{G}^{\mathsf{b}}_n$, $\mathbf{G}^{\phi}_n$, $\mathbf{G}^{\mu}_n$, $\mathbf{G}^{\mathsf{a}}_{n+1}$, and $\mathbf{G}^{\mathsf{c}}_{n+1}$, cannot be useful and can be set to zero.

In the same way as at orders 1 and 2, $\mathtt{avg}(\mathbf{G}^{\mu}_n)$ must remain free at that order, because it will be needed to solve $\overline{\Gamma}^{\theta}_{n+1}$ at the following order, just as $\mathtt{avg}(\mathbf{G}^{\mu}_{n-1})$ is needed here to solve $\overline{\Gamma}^{\theta}_n$. So, one cannot have the reduced Lagrangian fully simplified $\overline{\Gamma}_n = 0$. One of its component remains uncontrolled, either $\overline{\Gamma}^{t}_n$ or $\overline{\Gamma}^{\mathsf{b}}_n$.

One can consider recovering a zero value for this component by using the freedom $\mathtt{avg}(S_n)$, whose value had been arbitrarily fixed to zero in the process above. Then, when computing $\mathbf{G}^{\theta}_{n-1}$, $\mathbf{G}^{\mathsf{b}}_n$ and $\mathbf{G}^{\phi}_n$, the average gauge function $\mathtt{avg}(S_n)$ remains a free parameter. When replacing these variables by their expressions, the equation $\mathtt{avg}(\overline{H}_n) = 0$ becomes a differential equation for $\mathtt{avg}(S_n)$, whose structure is[2]

$$\left[\phi\partial_b + \tfrac{\nabla\cdot\mathsf{b}}{2}(1+\phi^2)\partial_\phi\right]\mathtt{avg}(S_n) = o.t.\,, \tag{2.69}$$

where $o.t.$ means other terms that can be explicitly computed. This resembles Eq. (2.46), but here, the condition (2.43) has not been required in the process, contrary to what happened in previous orders. The reason is that now $\overline{\Gamma}^{\mu}_n$ is controlled by $\mathbf{G}^{\theta}_{n-1}$, which has no effect on $\overline{H}_n$. The integrability of Eq. (2.69) will depend on the right-hand side and must be studied at each order; a priori, it is not guaranteed, since obstructions such as (2.47) are possible.

So, a systematic procedure cannot use the freedom $\mathtt{avg}(S_n)$ to get the additional requirement $\overline{H}_n = 0$, which must be dropped. Then the freedom $\mathtt{avg}(S_n)$ is useless and can be set to zero.

At the end of the $n$-th-order analysis, exactly all of the unknowns $\mathbf{g}_n$ have been determined. All the components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$ that remain unknown are in $\mathbf{g}_{i>n}$. Yet, this does not allow us to conclude that the induction is proven, because the unknown $\mathtt{avg}(\mathbf{G}^{\mu}_n)$ already appeared as a parameter in $\mathbf{G}^{\mathsf{a}}_{n+1}$ and in $\overline{H}_1$; hence it is not completely free, whereas the induction assumes it is free (independent of the quantities $\mathbf{g}_{i\leqslant n}$), since it is in $\mathbf{g}_{n+1}$; it will be determined at the next order, and this could imply coupled equations, whose solvability is to be verified.

However, $\mathtt{avg}(\mathbf{G}^{\mu}_n)$ will be computed in equation for $\mathtt{avg}(\overline{\Gamma}^{\theta}_{n+1})$, which corresponds to the column $\overline{\Gamma}^{\theta}_n$ in Eq. (2.67) at the next order, and does not involve any of the parameter-dependent quantities. Thus, there is no coupled equations, and the solutions are indeed explicit. This terminates the proof of the induction: the reduction can be performed to arbitrary order in $\varepsilon$.

Notice that here, the induction relation is considered from the point of view of $\overline{\Gamma}_n$; this caused an interlocking between the orders, where in the solution at each order, a parameter is involved, which will be identified at the next order, when computing $\overline{\Gamma}^{\mathbf{p}}_{n+1}$. To avoid this interlocking phenomenon, it is possible to consider the induction relation from the point of view of $\left(\overline{\Gamma}^{\mathbf{q},t}_n, \overline{\Gamma}^{\mathbf{p}}_{n+1}\right)$. The drawback

[2] Interestingly, the operator to be inverted is the same as for the secular differential equation for the magnetic moment, see e.g. Eq. (10.17).

would be that when solving the equations for $\overline{\Gamma}^{\mathbf{p}}_{n+1}$, one would begin the heavy computations for $\mathsf{R}'_{n+1}$, which are involved in $\overline{\Gamma}^{\mathbf{q}}_{n+1}$, hence at the next order. In computations by hands, this can be a trouble, but when using computer-assisted computations, this is no trouble and it would probably be a more relevant choice.

### Explicit induction relations

The argument above emphasizes the distinction to be made between four kinds of quantities. First, some of the quantities are already known before the computation at order $n$, namely $\mathbf{g}_{k<n}$.

Second, for $i \notin \{\mathsf{b}, t\}$ the components $\overline{\Gamma}^i_n$ have not been computed yet, but they can be excluded both from the unknowns and from the parameters, since the algorithm shows that before any computation, they are already known to be zero for all $n \geqslant 3$, to fulfil the requirements (2.20)-(2.23) for $\overline{\Gamma}$.

Third, some quantities are not known yet, and will be determined after the matrix inversion, namely

$$(\mathbf{g}_n)_\infty := \Big(\mathtt{avg}(\mathbf{G}^\mu_{n-1}), \mathtt{osc}(S_n) \mid \mathbf{G}^{\mathsf{b}}_n, \mathbf{G}^\theta_{n-1}, \mathbf{G}^{\mathsf{a}}_{n+1}, \mathbf{G}^{\mathsf{c}}_{n+1} \mid \mathtt{avg}(\mathbf{G}^\phi_n), \mathtt{osc}(\mathbf{G}^\phi_n), \mathtt{osc}(\mathbf{G}^\mu_n), \overline{H}_n\Big)\,, \tag{2.70}$$

in which a vertical dash | was written at the places where a vertical line will be written in the matrix $\mathsf{M}_\infty$ below.

Last, other quantities are not known and will remain free after the matrix inversion, namely $\overline{\Gamma}^{\mathsf{b}}_n$, $\mathtt{avg}(S_n)$ and $\mathbf{g}_{k>n}$; the variables $\mathbf{g}_{k>n}$ will be determined at higher order, but one of its component, $\mathtt{avg}(\mathbf{G}^\mu_n)$, is already involved in the equations at order $n$ and behaves as a parameter in this matrix inversion. So, the parameters are

$$(\mathbf{g}_n)_\alpha := \Big(\mathtt{avg}(\mathbf{G}^\mu_n), \mathtt{avg}(S_n), \overline{\Gamma}^{\mathsf{b}}_n\Big)\,. \tag{2.71}$$

Notice that $\mathtt{avg}(\mathbf{G}^\mu_n)$ is included in the parameters $(\mathbf{g}_n)_\alpha$ even if it is not an element of $\mathbf{g}_n$ but of $\mathbf{g}_{n+1}$.

With the procedure above, the left-hand side of Eq. (2.67) can be written as just a matrix product, provided the pivotal matrix is extended, to act on all the quantities $(\mathbf{g}_n)_\infty$ to be computed at this order, even the reduced Hamiltonian $\overline{H}_n$ and the gauge function $S_n$. To include the gauge function $S_n$ in the vector which is acted upon by the matrix, some coefficients in the matrix must be operators, and the equation will be transposed, so that the operators act on their right. For clarity, the order of the columns is chosen to fit with the steps of the algorithm

$$\Big(\mathtt{avg}(\overline{\Gamma}^\theta_n), \mathtt{osc}(\overline{\Gamma}^\theta_n) \mid \overline{\Gamma}^\phi_n, \overline{\Gamma}^\mu_n, \overline{\Gamma}^{\mathsf{c}}_n, \overline{\Gamma}^{\mathsf{a}}_n \mid \mathtt{avg}(\overline{\Gamma}^{\mathsf{b}}_n), \mathtt{osc}(\overline{\Gamma}^{\mathsf{b}}_n), \mathtt{osc}(\overline{\Gamma}^t_n), \mathtt{avg}(\overline{\Gamma}^t_n)\Big)\,,$$

so that the equations are solved one after the other in order. A vertical dash | was written at the places where a horizontal line will be written in the matrix $\mathsf{M}_\infty$ below.

With this order for the rows and for the columns, the equation becomes

$$\mathsf{M}_\infty \cdot (\mathbf{g}_n)^T_\infty + \mathsf{M}_\alpha \cdot (\mathbf{g}_n)^T_\alpha + \mathsf{R}'^T = 0\,. \tag{2.72}$$

Here, the rows are the equations to be solved, corresponding to the (re-ordered) columns of Eq. (2.67). The first term involves exactly the unknown quantities to be identified at this order. The second term involves exactly the parameters involved at this order but which will remain free at the end of this order. The third term involves only quantities that are already known at the beginning of this order.

The matrices are given by

$$\mathsf{M}_\infty := \left(\begin{array}{cc|ccc|cccc}
1 & & & & & & & & \\
 & \partial_\theta & & & & & & & \\
\hline
 & \partial_\phi & -\sqrt{2\mu B} & & & & & & \\
 & \partial_\mu & -\phi\sqrt{\frac{B}{2\mu}} & -1 & & & & & \\
 & \partial_c & -I & & -B & & & & \\
 & \partial_a & K & & B & & & & \\
\hline
\mathtt{avg}\partial_\mu\overline{\Gamma}_1^{\mathsf{b}} & \partial_b & & & & \sqrt{2\mu B} & & & \\
\mathtt{osc}\partial_\mu(\overline{\Gamma}_1^{\mathsf{b}}) & \partial_b & & & & & \sqrt{2\mu B}\phi & \sqrt{\frac{B}{2\mu}} & \\
\mathtt{osc}\partial_\mu(\overline{H}_1) & & -H_0\nabla\cdot\mathsf{b} & & & & 2\mu B\phi & \frac{H_0}{\mu} & \\
\mathtt{avg}\partial_\mu(\overline{H}_1) & & -H_0\nabla\cdot\mathsf{b} & & & 2\mu B\phi & & & -1
\end{array}\right), \tag{2.73}$$

and

$$\mathsf{M}_\alpha := \left(\begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & 0 \\
\hline
 & \partial_\phi & \\
 & \partial_\mu & \\
 & \partial_c & \\
 & \partial_a & \\
\hline
\phi\sqrt{\frac{B}{2\mu}} & \partial_b & -1 \\
 & \partial_b & \\
0 & 0 & 0 \\
\frac{H_0}{\mu} & &
\end{array}\right),$$

in which the zeros were written only in the empty rows, for clarity, and we used that $\nabla\cdot\mathbf{B} = 0$ implies $\frac{\partial_b B}{B} = -\nabla\cdot\mathsf{b}$.

In the matrix $\mathsf{M}_\infty$, grouping together the eighth and ninth rows and columns produces a $9*9$ lower triangular matrix, whose coefficients on the diagonal are invertible, since the operator $-\mathsf{a}\cdot\partial_{\mathsf{c}}$ is invertible on gyro-fluctuations. The eighth and ninth rows and columns have been grouped together because they constitute an invertible $2*2$ matrix on the diagonal. A convenient way to invert the resulting $9*9$ matrix is to separate its diagonal terms:

$$\mathsf{M} = \mathsf{M}' + \mathsf{D}\,,$$

where $\mathsf{M}'$ has null diagonal and $\mathsf{D}$ is purely diagonal. Then the equation can be written

$$-\mathsf{D}\cdot(\mathbf{g}_n)_\infty^T = \mathsf{M}'_\infty\cdot(\mathbf{g}_n)_\infty^T + \mathsf{M}_\alpha\cdot(\mathbf{g}_n)_\alpha^T + \mathsf{R}'^T\,.$$

The solution of Eq. (2.72) is then

$$(\mathbf{g}_n)_\infty^T = (-\mathsf{D})^{-1}\cdot\Big[\mathsf{M}'_\infty\cdot(\mathbf{g}_n)_\infty^T + \mathsf{M}_\alpha\cdot(\mathbf{g}_n)_\alpha^T + \mathsf{R}'^T\Big]\,, \tag{2.74}$$

where $(-\mathsf{D})^{-1}$ is a diagonal matrix with coefficients

$$(-\mathsf{D})^{-1} := \mathrm{Diag}\Big(-1, -(\partial_\theta)^{-1} \mid \tfrac{1}{\sqrt{2\mu B}}, 1, \tfrac{1}{B}, -\tfrac{1}{B} \mid -\tfrac{1}{\sqrt{2\mu B}}, -\mathsf{D}_2^{-1}, 1\Big)\,,$$

in which $\mathsf{D}_2^{-1}$ is the inverse matrix for the coupled system (the eighth diagonal term of the 9*9 matrix mentioned above)

$$-\mathsf{D}_2^{-1} := -\begin{pmatrix} \sqrt{2\mu B} & \phi\sqrt{\frac{B}{2\mu}} \\ 2\mu B\phi & B(1+\phi^2) \end{pmatrix}^{-1} = \begin{pmatrix} -\frac{1+\phi^2}{\sqrt{2\mu B}} & \frac{\phi}{2\mu B} \\ \phi\sqrt{\frac{2\mu}{B}} & -\frac{1}{B} \end{pmatrix}.$$

The operator $\partial_\theta^{-1}$ is the gyro-integral operator. It can be computed without introducing any gyro-gauge, with the intrinsic calculus introduced in Chapter 2, or with the matrix calculus introduced in Chapter 1. Also, the coordinate $\theta$ can be used as an intermediate quantity for this computation, which is made at constant $\mathbf{q}$, so that the presence of a gauge (only for the intermediate computation) is of no consequence; then $\partial_\theta^{-1}$ is the primitive with respect to $\theta$ such that its gyro-average is zero. Over the Fourier modes $k \neq 0$ (i.e. over gyro-fluctuations), it is the operator $\frac{1}{ik}$.

Even if $(\mathbf{g}_n)_\alpha$ appears in its right-hand side, Eq. (2.74) is an explicit solution for the induction relation: in the matrix $\mathsf{M}'_\infty$, all the coefficients on the diagonal or above it are zero, so that when computing the unknowns one after the other starting from the left, each of them is computed as a function of previously computed quantities, i.e. the right-hand side contains only known quantities or parameters, but none of the remaining unknowns. Alternatively, the induction can be solved using a standard matrix inverse

$$(\mathbf{g}_n)^T_\infty = -\mathsf{M}^{-1}_\infty \cdot \left[\mathsf{M}_\alpha \cdot (\mathbf{g}_n)^T_\alpha + \mathsf{R}'^T\right], \tag{2.75}$$

with $\mathsf{M}^{-1}_\infty$ easily computed from (2.73), but the coefficients are more complicated, and practical computations usually follow the procedure of Eq. (2.74).

For the solution $\mathbf{g}_n$, Eq. (2.74) or (2.75) must be completed by the following relations for the trivial components of $\overline{\Gamma}_n$:

$$\overline{\Gamma}^i_n = 0 \quad \text{for all } i \notin \{\mathsf{b}, t\},$$

and by the determination of the parameters $(\mathbf{g}_n)_\alpha$.

The first parameter $\mathtt{avg}(\mathbf{G}^\mu_n)$ will be determined at the following order, in an equation that does not involve $\mathbf{G}^\phi_n$ and $\overline{H}_n$, so that there are not coupled equations between the orders.

The second parameter $\overline{\Gamma}^{\mathsf{b}}_n$ is either put to zero or adjusted so as to make the reduced Hamiltonian $\overline{H}_n$ zero by the relation

$$\begin{aligned}\overline{\Gamma}^{\mathsf{b}}_n =& \mathtt{avg}\Big[\mathsf{R}'^{\mathsf{b}}_n + \mathbf{G}^\mu_{n-1}\partial_\mu \overline{\Gamma}'^{\mathsf{b}}_1 + \partial_b S_n \\ &- \tfrac{1}{\sqrt{2\mu B}\phi}\Big\{\mathbf{G}^\mu_{n-1}\partial_\mu \overline{H}_1 - \mathbf{G}^{\mathsf{b}}_n H_0 \nabla\cdot\mathsf{b} + B\mathbf{G}^\mu_n + \mathsf{R}'^t_n\Big\}\Big].\end{aligned} \tag{2.76}$$

This last choice is possible only if the term inside the parentheses has no overall contribution of order zero in $\phi$.

The last parameter $\mathtt{avg}(S_n)$ is determined by Eq. (2.69), in order to make both $\overline{\Gamma}^{\mathsf{b}}_n$ and $\overline{H}_n$ zero, when the equation is integrable. Otherwise, it is determined to cancel as many terms of $\overline{\Gamma}^{\mathsf{b}}_n$ and $\overline{H}_n$ as possible, or it can be set to zero for simplicity.

Accordingly, at the end of each order $n \geqslant 3$, the situation is the same as at the end of order 2, with the presence of one parameter $\mathtt{avg}(\mathbf{G}^\mu_n)$, of one binary choice between $\overline{\Gamma}^{\mathsf{b}}_n$ and $\overline{H}_n$, and of one free variable $\mathtt{avg}(S_n)$. When the integrability condition can be satisfied, the reduction of the Lagrangian $\overline{\Gamma}_n$ is complete and the solution is defined to within an element in the kernel of the operator (2.69). Otherwise, $\mathtt{avg}(\mathbf{G}^\mu_n)$ is generally set to zero, and after the binary choice the transformation is unique, but on the whole there are two maximal reductions. As announced in the appendix, the unicity of the transformation is determined by the integrability condition for $S_n$, and possibly by an additional criterion for simplicity.

## 2.3 Comparison with previous works

### 2.3.1 Agreeing results

Computations of the previous section for the guiding-center transformation and reduced Hamiltonian can be summarized by Eqs. (2.37), (2.48)-(2.50) and (2.60)-(2.64). As for the reduced

symplectic Lagrangian, it is exactly known, and is given by Eqs. (2.39) and (2.51), together with the prescription that all other terms $\overline{\Gamma}_n^j$ are zero, except $\overline{\Gamma}_{n\geqslant 2}^{\mathsf{b}}$ (as well as $\overline{\Gamma}_1'^{\mathsf{b}}$), which can be freely chosen, for instance it can be chosen zero, or such that it makes the reduced Hamiltonian $\overline{H}_n$ zero by Eq. (2.76).

These results agree with the standard results of the literature, provided the connection vector is defined as $\mathbf{R}_g := \mathbf{R}$, which corresponds to the traditional gauge-dependent framework. For instance, in the paper [88] Littlejohn made the choice $\overline{\Gamma}_2 = 0$, and accounting for this choice, our formulae agree with his ones. In the paper [29], the choice is $H_2 = 0$, and again, accounting for this choice through Eq. (2.65) for $\overline{\Gamma}_1'^{\mathsf{b}}$, our results agree with the ones of [29]. Thus, the procedure of the previous section succeeds in obtaining the standard guiding-center reduction without introducing any gyro-gauge and using purely intrinsic coordinates.

The explicit induction relation (2.74) shows that the reduction is possible to arbitrary orders, but it also gives an explicit formula to practically compute the transformation generator order by order. At any order in the Larmor radius, all that remains to do is to develop the Lie derivatives involved in the term $\mathsf{R}_n'$. Only the number of terms generated by the Leibniz rule makes the process difficult to compute by hand at higher order, but the explicit induction involves few basic operations (just exterior derivatives and matrix products) and can easily be implemented to higher orders on a computer. Actually, as the series is a polynomial, the derivation does not rely on formal calculus but just on symbolic calculus, which is still easier to implement.

Once the generators $\mathbf{G}_n$ are obtained, the guiding-center coordinate transformation is given by

$$\mathbf{z} \longrightarrow \overline{\mathbf{z}} := \ldots e^{-\mathbf{G}_2} e^{-\mathbf{G}_1} \mathbf{z}\,. \tag{2.77}$$

The reduced Lagrangian 1-form is

$$\begin{aligned}\overline{\Gamma} &= \mathsf{avg}(\Gamma) - \mu \mathsf{a}\cdot d\mathsf{c} + \Big[\overline{\Gamma}_1'^{\mathsf{b}} + \sum_{n\geqslant 2} \overline{\Gamma}_n^{\mathsf{b}}\Big] \mathsf{b}\cdot d\mathbf{q} - \sum_n \overline{H}_n dt \\ &= (\mathbf{A} + \overline{\Gamma}^{\mathsf{b}} \mathsf{b})\cdot d\mathbf{q} - \mu \mathsf{a}\cdot d\mathsf{c} - \overline{H} dt\,,\end{aligned} \tag{2.78}$$

where the Hamiltonian terms $\overline{H}_n$ are provided by (2.74), whereas the parallel Lagrangian terms $\overline{\Gamma}_1'^{\mathsf{b}}$ and $\overline{\Gamma}_n^{\mathsf{b}}$ are chosen freely at each order $n$, and can be chosen zero.

The reduced dynamics is obtained the usual way, by computing the Lagrange matrix $\overline{\omega}_s := d\overline{\Gamma}_s$, then inverting it to get the Poisson matrix $\overline{\mathbb{J}} := \overline{\omega}_s^{-1}$, and last computing Hamilton's equations $\dot{\overline{\mathbf{z}}}^i := \mathbb{J}^{ij}\partial_j \overline{H}$. Alternatively, the reduced equations of motion can be obtained by Lie transforming directly the velocity vector field

$$\dot{\mathbf{z}} \longrightarrow \dot{\overline{\mathbf{z}}} := \ldots e^{\mathsf{L}_2} e^{\mathsf{L}_1} \dot{\mathbf{z}}\,.$$

Here, performing these computations is useless, since the reduced Lagrangian (2.78) completely agrees with previous results in the literature in the gauge-dependent case, and so will the reduced motion.

We only indicate the guiding-center Poisson bracket, in which the effects of the gauge independence and of the higher-order corrections are interesting. It is computed from the Lagrange 2-form, easily obtained from Eq. (2.78) in a matrix form:

$$\overline{\omega} := d\overline{\Gamma} = \left(\begin{array}{ccc|ccc|c} 0 & -\mathtt{B} & \mathtt{A} & 0 & \mathtt{x} & 0 & \mathsf{c}\cdot\nabla\overline{H} \\ \mathtt{B} & 0 & -\mathtt{C} & 0 & \mathtt{y} & 0 & \mathsf{a}\cdot\nabla\overline{H} \\ -\mathtt{A} & \mathtt{C} & 0 & -\mathtt{D} & \mathtt{z}-\mathtt{E} & 0 & \mathsf{b}\cdot\nabla\overline{H} \\ \hline 0 & 0 & \mathtt{D} & 0 & 0 & 0 & \partial_\phi\overline{H} \\ -\mathtt{x} & -\mathtt{y} & \mathtt{E}-\mathtt{z} & 0 & 0 & 1 & \partial_\mu\overline{H} \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hline -\mathsf{c}\cdot\nabla\overline{H} & -\mathsf{a}\cdot\nabla\overline{H} & -\mathsf{b}\cdot\nabla\overline{H} & -\partial_\phi\overline{H} & -\partial_\mu\overline{H} & 0 & 0 \end{array}\right),$$

with

$$
\begin{aligned}
\mathtt{A} &:= -\mathsf{a}\cdot\nabla\times(\mathbf{A}+\overline{\Gamma}^{\mathsf{b}}\mathsf{b}) - \mu\mathsf{c}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'\mathsf{b} = I + O(B^0)\,, & \mathtt{x} &:= \mathbf{R}_g\cdot\mathsf{c}\,,\\
\mathtt{B} &:= -\mathsf{b}\cdot\nabla\times(\mathbf{A}+\overline{\Gamma}^{\mathsf{b}}\mathsf{b}) - \mu\mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'\mathsf{c} = -B + J + O(B^0)\,, & \mathtt{y} &:= \mathbf{R}_g\cdot\mathsf{a}\,,\\
\mathtt{C} &:= -\mathsf{c}\cdot\nabla\times(\mathbf{A}+\overline{\Gamma}^{\mathsf{b}}\mathsf{b}) - \mu\mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'\mathsf{a} = K + O(B^0)\,, & \mathtt{z} &:= \mathbf{R}_g\cdot\mathsf{b}\,,\\
\mathtt{D} &:= \partial_\phi\overline{\Gamma}^{\mathsf{b}} = \sqrt{2\mu B} + O(B^0)\,, \\
\mathtt{E} &:= \partial_\mu\overline{\Gamma}^{\mathsf{b}} = \phi\tfrac{\sqrt{2\mu B}}{2\mu} + O(B^0)\,,
\end{aligned}
\tag{2.79}
$$

where we used the fact that the order in $\sqrt{B}$ indicates the expansion order, as mentioned about Eq. (2.27).

Then the Poisson bracket in matrix form writes:

$$
\overline{\mathbb{J}} := (\overline{\omega}^{\mathbf{z};\mathbf{z}})^{-1} = \left(\begin{array}{ccc|ccc}
0 & \frac{1}{\mathtt{B}} & 0 & \frac{\mathtt{C}}{\mathtt{BD}} & 0 & \frac{\mathtt{y}}{\mathtt{B}} \\
-\frac{1}{\mathtt{B}} & 0 & 0 & \frac{\mathtt{A}}{\mathtt{BD}} & 0 & -\frac{\mathtt{x}}{\mathtt{B}} \\
0 & 0 & 0 & \frac{\mathtt{B}}{\mathtt{BD}} & 0 & 0 \\
\hline
-\frac{\mathtt{C}}{\mathtt{BD}} & -\frac{\mathtt{A}}{\mathtt{BD}} & -\frac{\mathtt{B}}{\mathtt{BD}} & 0 & 0 & -\alpha \\
0 & 0 & 0 & 0 & 0 & -1 \\
-\frac{\mathtt{y}}{\mathtt{B}} & \frac{\mathtt{x}}{\mathtt{B}} & 0 & \alpha & 1 & 0
\end{array}\right),
$$

with

$$
\alpha := -\frac{\mathtt{E}}{\mathtt{D}} + \frac{\mathtt{Cx}+\mathtt{Ay}+\mathtt{Bz}}{\mathtt{BD}} = -\frac{\mathtt{E}}{\mathtt{D}} + \frac{\mathbf{B}_*\cdot\mathbf{R}_g}{\mathtt{BD}}\,,
$$

where the reduced magnetic field $\mathbf{B}_*$ is defined as usual by Eq. (2.82). The coefficients $\mathtt{B}$ and $\mathtt{D}$ are invertible since Eq. (2.79) shows that they are small corrections from $-B$ and $\sqrt{2\mu B}$, which are invertible.

As a result, between two arbitrary functions of the reduced phase space $f(\overline{z})$ and $g(\overline{z})$, the Poisson bracket is

$$
\{f,g\} = -\nabla_* f\cdot\frac{\mathsf{b}}{\mathtt{B}}\times\nabla_* g + \frac{\mathbf{B}_*}{\mathtt{BD}}\cdot\nabla_* f\wedge\partial_\phi g + \frac{\mathtt{E}}{\mathtt{D}}\partial_\phi f\wedge\partial_\theta g + \partial_\theta f\wedge\partial_\mu g\,, \tag{2.80}
$$

where the symbol $\nabla_*$ is a shorthand for the operator

$$
\nabla_* := \nabla + \mathbf{R}_g\partial_\theta\,. \tag{2.81}
$$

This operator does not depend of the free function $\mathbf{R}_g$, since for any choice of $\mathbf{R}_g$ the definition (2.81) gives the same result, which is equal to the covariant derivative $\nabla$ when the gauge vector is chosen zero, i.e. $(\mathbf{R}_g)_* = \nabla_*\mathsf{c}\cdot\mathsf{a} = 0$.

The Poisson bracket (2.80) agrees with the literature, e.g. [17, 29, 86, 88, 128]. Especially, the traditional ordering is patent: the last two terms are of order $\mu^{-1}$, and correspond to the fast gyro-angle dynamics; the second term is of order $\frac{\nabla}{\sqrt{\mu B}} = \mu^{-1}\epsilon$, and corresponds to the intermediate motion along the magnetic field lines; as for the first term, it is of order $B^{-1}\nabla^2 = \mu^{-1}\epsilon^2$ , and corresponds to the slow drifts across the magnetic field lines. Remind $\epsilon := r_L\nabla = \sqrt{\frac{\mu}{B}}\nabla$, as defined in (2.3), is the usual small parameter of guiding-center theory.

In previous works, the reduced Poisson bracket often had only three coefficients (the ones for $\mathbf{B}_*$, namely $\mathtt{A}$, $\mathtt{B}$, and $\mathtt{C}$), either because of incidental lowest-order simplifications, or because of the choice they had performed for $\overline{\Gamma}^{\mathsf{b}}$. For a general result about maximal reductions at arbitrary order, the Poisson bracket is given by Eq. (2.80), where the higher-order correction $\overline{\Gamma}^{\mathsf{b}}_{n\geqslant 1}$ to the Lagrangian impacts five coefficients $\mathtt{A}$, $\mathtt{B}$, $\mathtt{C}$, $\mathtt{D}$, and $\mathtt{E}$ (and hence $\mathbf{B}_*$), through the definitions (2.79), i.e. as an effect of the five derivative operators $\partial_{\mathsf{c}}$, $\partial_{\mathsf{a}}$, $\partial_{\mathsf{b}}$, $\partial_\phi$, $\partial_\mu$, acting on the seminal term $\overline{\Gamma}^{\mathsf{b}}_{n\geqslant 1}$.

The only difference compared to previous results, besides the coordinate system and the term $\overline{\Gamma}^{\mathsf{b}}_{n\geqslant 1}$ being let free, is that the gauge vector $\mathbf{R} = \nabla\mathsf{e}_1\cdot\mathsf{e}_2$ is replaced by the general connection $\mathbf{R}_g = \nabla\mathsf{c}\cdot\mathsf{a}$. This difference could impact the reduced magnetic field $\mathbf{B}_*$, given by

$$
\mathbf{B}_* := -\mathtt{A}\mathsf{a} - \mathtt{B}\mathsf{b} - \mathtt{C}\mathsf{c} = \nabla\times\left(\mathbf{A}+\mathsf{b}\overline{\Gamma}^{\mathsf{b}}\right) + \mathbf{V}\,, \tag{2.82}
$$

where the vector $\mathbf{V}$ is generated by the term $-\mu\mathsf{a}\cdot d\mathsf{c}$ in the Lagrangian $\overline{\Gamma}$, which implies for the Lagrange 2-form the presence of the term

$$\begin{aligned}
-\mu\ d\mathsf{a}\cdot\wedge d\mathsf{c} &= -\mu\ d\mathsf{a}\cdot(\mathsf{aa}+\mathsf{bb}+\mathsf{cc})\cdot\wedge d\mathsf{c}\\
&= -\mu\ d\mathsf{a}\cdot\mathsf{b}\wedge\mathsf{b}\cdot d\mathsf{c}\\
&= -\mu\ d\mathsf{b}\cdot\mathsf{a}\wedge\mathsf{c}\cdot d\mathsf{b}\\
&= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{a}\wedge\mathsf{c}\cdot\mathsf{b}'d\mathbf{q}\\
&= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'d\mathbf{q}\\
&=: -d\mathbf{q}\cdot\mathbf{V}\times d\mathbf{q}\,. \qquad (2.83)
\end{aligned}$$

The first equality comes by inserting the identity $(\mathsf{aa}+\mathsf{bb}+\mathsf{cc})$ beside the wedge symbol. The second and third equalities come because $(\mathsf{a},\mathsf{b},\mathsf{c})$ is orthonormal. The fact that the vector $\mathsf{b}$ depends only on $\mathbf{q}$ implies the fourth equality, which can be rewritten to get the final answer.

Eq. (2.83) shows why the vector $\mathbf{V}$ (and hence also the reduced Poisson bracket) is indeed independent of the free function $\mathbf{R}_g$. In the reduced Lagrange 2-form, the concerned term (2.83) is the counterpart in the gauge-independent approach of the term $d\mathbf{q}\cdot(\mu\nabla\times\mathbf{R})\times d\mathbf{q}$ involving the gauge vector in the traditional approach. Actually, the curl of the gauge vector $\nabla\times\mathbf{R}$ is also gauge-independent, and the corresponding term is explicitly given by

$$\begin{aligned}
-\mu\ d(d\mathbf{q}\cdot\nabla\mathsf{e}_1\cdot\mathsf{e}_2) &= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{e}_2\cdot\wedge\mathsf{e}_1'd\mathbf{q}\\
&= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'d\mathbf{q}\,, \qquad (2.84)
\end{aligned}$$

where again, the second equality comes by inserting the identity $(\mathsf{bb}+\mathsf{e}_1\mathsf{e}_1+\mathsf{e}_2\mathsf{e}_2)\cdot$ beside the wedge symbol and by using the fact that $(\mathsf{b},\mathsf{e}_1,\mathsf{e}_2)$ is orthonormal.

So, in the reduced magnetic field $\mathbf{B}_*$, the term with $\mathbf{R}_g$ exactly fits with the corresponding term with $\mathbf{R}$ of previous results. This is a good illustration of how the approach using the coordinate $\mathsf{c}$ completely agrees with the gauge-independent part of the usual results, whereas it proceeds in a different way and never introduces the gauge $\mathsf{e}_1$. Indeed, in the usual approach, $\nabla\mathsf{e}_1\cdot\mathsf{e}_2$ is assumed to depend only on the position, which implies the formula above. On the contrary, in the gauge-independent approach, the velocity is present in the first lines of (2.83), but it comes out from the computation that the result is naturally just a purely spatial term in the Lagrange matrix.

By the way, the argument above shows that $\mathbf{V}$ is indeed the curl of $-\mu\mathbf{R}_g=-\mu\nabla\mathsf{c}\cdot\mathsf{a}$, but it is not surprising since the spatial part of a term $df\wedge dg$ is given by the cross product with the curl of $-f\nabla g$:

$$\begin{aligned}
d\mathbf{q}\cdot(df\wedge dg)^{\mathbf{q};\mathbf{q}}\cdot d\mathbf{q} &= d\mathbf{q}\cdot\nabla f\wedge g'd\mathbf{q}\\
&= d\mathbf{q}\cdot(\nabla(fg'))\wedge d\mathbf{q}\\
&= -d\mathbf{q}\cdot(\nabla\times(fg'))\times d\mathbf{q}\,.
\end{aligned}$$

Thus, Eq. (2.82) can be rewritten

$$\mathbf{B}_* := \nabla\times\left[\mathbf{A}+\mathsf{b}\left\{\sqrt{2\mu B}\phi+\overline{\Gamma}_1'^{\mathsf{b}}+\sum_{n\geqslant 2}\overline{\Gamma}_n^{\mathsf{b}}-\mu\mathbf{R}_g\right\}\right],$$

which is manifestly divergenceless.

Here, for clarity and following previous works, we considered no electric field, in order to focus on the reduction mechanism. The extension for a non-zero electric field is straightforward, because the Lagrangian is changed only by the addition of $-e\Phi dt$, with $\Phi(\mathbf{q})$ the electric potential [17,29,88]. It affects the spatial part of $\partial_{\mathbf{z}}H$, which plays no pivotal role in the derivation: it always appears in the right-hand side of the equations, i.e. in terms that are already known. As a consequence, the presence of this term does not change the procedure at all. In the results, each term of order $n$ becomes a polynomial in $\Phi$ (or rather $\nabla\Phi$), which introduces a second parameter $\varepsilon_E$ in the theory, which corresponds to $\frac{mE}{pB}$ or to $\frac{mE\nabla}{eB^2}$. A more detailed study shows that the momentum in denominators is only the perpendicular momentum $\|\mathbf{p}_\perp\|$, and that at the lowest orders the perpendicular electric field can be one order higher than the parallel electric field [17, 86, 88]. The series remains perturbative provided the associated parameter is small $\varepsilon_E << 1$, as usual in guiding-center reductions.

### 2.3.2 Polynomiality in the momentum coordinates

In the results above, the usual expansions are recovered, but the choice of the coordinates $\mathsf{c}$ and $\phi$ makes all quantities polynomial in the used coordinates and monomial in $\sqrt{\mu}$ and $\sqrt{B}$. This is useful to simplify the derivation, which can be considered as a symbolic-calculus algorithm based on just two operations acting on words (polynomials) composed from a very restricted alphabet. Such structures can also be useful when going beyond the formal expansions and considering them as asymptotic series. For instance, the polynomiality in the cotangent of the pitch-angle is important to control the loss of accuracy of the expansion in the domains where the direction of the particle momentum is close to the direction of the magnetic field.

Previous results were obviously polynomial in the variable $\mathsf{c}$, but they did not use it as a coordinate, since they replaced it by the variable $\theta$. In addition, the expansion in the usual small parameter $\varepsilon := \sqrt{2\mu/B}\nabla := r_L\nabla$ was also present, but it did not correspond to an expansion in $\mu$ nor in $B$, but only in $\nabla$, which is not a scalar quantity; it is why keeping the quantity $e^{-1}$ was useful to indicate the expansion order. Here, the order is directly indicated by the power in $\sqrt{\mu}$ or $\sqrt{B}$, since each order is a monomial in those quantities.

The monomiality in $\sqrt{\mu}$ is especially interesting, because the orders in the operator $\nabla$ and in the function $B$ have only a dimensional meaning: the term of order $\varepsilon^n$ will involve terms like $\left(\frac{\nabla B}{B}\right)^n$, but also terms like $\frac{\nabla^{n-k}B}{B}\frac{\nabla^k B}{B}$, or $\nabla^n\mathsf{b}$. In a similar way, the order in $B$ is given by the order of the prefactor $r_L^n = \sqrt{2\mu/B}^n$, but the variable $B$ appears in other places when it is acted upon by gradients; then, it is compensated by a $B$ in the denominator, which means that gradients act only on the variable $\ln B$.

The use of the variable $\phi$ instead of the usual $v_\parallel$ is crucial for the results at each order to be both polynomial in the coordinates and monomial in $\sqrt{\mu}$ and $\sqrt{B}$. It is a generalization of Chapter 1, which considered only the minimal guiding-center reduction. It seems it had not been noticed in previous works on the full guiding-center reduction [29, 86, 88, 128].

When using the standard variable $v_\parallel$ instead of $\phi$, the monomiality in $\sqrt{\mu}$ and $\sqrt{B}$ is not verified: for instance, in $\mathbf{G}_1^\theta$, the term of order $\phi^2$ writes $-\frac{v_\parallel^2}{B\sqrt{2\mu B}}\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}$, which is not a polynomial in $\sqrt{\mu}$ and which is not of the same order in $\sqrt{\mu}$ nor in $\sqrt{B}$ as the term of order $\phi^1$, since this last writes $\frac{v_\parallel}{B}\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{4}$, as is confirmed in [29, 88], for instance.

On the contrary, when using the variable $\phi$, the polynomiality is verified, because of the structure of the Lagrangian, of the action of derivatives, especially (2.9), and of the coefficients to be inverted for the matrix inversions involved in the derivation. Notice that the induction procedure also guarantees that all formulae will be polynomial in the variable $\mathsf{b}$.

As for the monomiality, it is easily understood from a dimensional analysis: only three independent dimensional quantities are involved in the results, e.g. $\nabla$, $\mathbf{B}$ and $\mathbf{p}$. When using the momentum coordinates $(\phi, p, \mathsf{c})$ or $(\phi, \mu, \mathsf{c})$, two of them are dimensionless, and only one of them can generate the dimension of $p$, namely $p$ or $\sqrt{\mu}$. For an expansion in $r_L\nabla = \sqrt{\frac{\mu}{B}}\nabla$, the variable $\mu$ can be involved only in the pre-factor of each order, otherwise, it could not be compensated to generate a dimensionless quantity.

On the other hand, when using the momentum coordinates $(v_\parallel, p, \mathsf{c})$ or $(v_\parallel, \mu, \mathsf{c})$, there is a redundancy in dimension between $v_\parallel$ and $p$ (or $\sqrt{\mu B}$), which means that ratios of them are expected in order to get dimensionless quantities. In an expansion in $\varepsilon$, if the term of order $(r_L\nabla)^n$ is a polynomial in $v_\parallel$, it has to be actually a polynomial in $\frac{v_\parallel}{p}$ or $\frac{v_\parallel}{\sqrt{\mu B}}$. As a consequence, formulae will be sums of terms $(\frac{p\nabla}{B})^n\left(\frac{v_\parallel}{p}\right)^j$, which is not a monomial in $p$ (or equivalently in $\sqrt{\mu}$ and $\sqrt{B}$); by the way, it is not a polynomial either, because of the terms where $j > n$.

Last, the polynomiality in $\phi$ means that each term writes $(\frac{p\sin\varphi\nabla}{B})^n\left(\frac{p\cos\varphi}{p\sin\varphi}\right)^j = \left(\frac{p\nabla}{B}\right)^n\cos^j\varphi\sin^{n-j}\varphi$, which agrees with the idea that the entities $\cos\varphi$ and $\sin\varphi$ come from expansions of the momentum $\mathbf{p}$, or rather the corresponding dimensionless vector $\frac{\mathbf{p}}{p}$.

### 2.3.3 A two-fold maximal reduction

In addition to the averaging reduction and the inclusion of the magnetic moment among the reduced coordinates, the goal was to obtain a reduced dynamics as strongly reduced as possible. So, a complete achievement is obtained when all the components of $\overline{\Gamma}_n$ are zero, at least for higher orders. For the procedure, it means using the seven unknowns to solve the seven corresponding equations, or rather twice these numbers, if one considers the average and the fluctuating parts as different variables. It did work for all the requirements but one, which could not be satisfied and had to be dropped.

The obstruction for a complete reduction $\overline{\Gamma}_{n\geqslant 3} = 0$ comes from the requirement (2.23): it imposes to obtain $\overline{\Gamma}_n^\theta$ by fixing the freedom $\mathtt{avg}(\mathbf{G})$, which was the only freedom available for the equation $\overline{\Gamma}_t = 0$ and is no more available for it. It is why one of the requirements has to be dropped. Then, it remains more unknowns than requirements. So, the transformation is not unique. Especially, while it naturally appeared in $\overline{H}$, the non-zero component of $\overline{\Gamma}$ can be transferred to $\overline{\Gamma}^{\mathsf{b}}$. Thus two maximal reductions can be considered.

The first alternative (called the *Hamiltonian* representation in [23]) sets $(\overline{\Gamma}_s)_{n\geqslant 2} = 0$. Then, the reduced Poisson bracket is completely known before computing the transformation to higher orders; it is given by the lowest three orders of the symplectic Lagrangian $(\overline{\Gamma}_s)_{n\leqslant 1}$. The reduced Hamiltonian is not exactly known; it is given by a whole series in $\varepsilon$ and must be computed order by order. The reduced dynamics is a Hamiltonian perturbation of the guiding-center equations of motion at order 2.

The second alternative (called the *symplectic* representation in [23]) is to set $\overline{H}_{n\geqslant 1} = 0$. Then, the Hamiltonian is completely known, and the structure of the Poisson bracket is also known. The only unknown information on the reduced dynamics is concentrated in the component of the reduced Lagrangian parallel to the magnetic field $\overline{\Gamma}_n^{\mathsf{b}}$, which is given by a whole series in $\varepsilon$. The Poisson bracket includes a kind of reduced magnetic field $\mathbf{B}_*$, induced by the higher-order terms of the Lagrangian.

The choice of symplectic or Hamiltonian representation can be made at each order in the derivation, but it seems more convenient to be consistent and to make the same choice for all orders, as suggested in [23].

These two maximal reductions give a unified view of various choices that can be found in the literature, and they anticipate what will happen at higher orders. Indeed, even in the standard non-canonical Hamiltonian approach of the guiding-center reduction introduced by Littlejohn, several transformations can be found, often related by differences of choice related to this two-fold maximal reduction. For instance, Littlejohn's initial guiding-center reduction [86] corresponded to the second possibility above at order $n = 1$, but at higher order, it is unclear whether the procedure provided a maximal reduction, or if some terms could remain both in the Hamiltonian and in the Poisson bracket. The seminal reduction by Lie transforming the Lagrangian [88] corresponded to a maximal reduction with the first choice ("Hamiltonian representation").

Later papers by Lie transforming the Lagrangian turned to the other choice, e.g. [29] used a maximal reduction (at order $n = 1$) but with the second possibility. Recently, while improvements in the second order were addressed [23, 128], interest was renewed in the first possibility. The paper [128] actually corresponds to a mixed choice, where the second possibility is used at order $n = 1$ and the first one is used at order $n = 2$. The work [23] introduced the designation of *Hamiltonian* and *symplectic* representation to differentiate between the two choices, and studied an "equivalence relation" between the two choices (when the same choice is made for all orders), which is a way to go from one representation to the other by a redefinition of the reduced parallel momentum $\overline{p}_\parallel$ (or equivalently of $\overline{\phi}$).

The algorithm at higher order, with the condition (2.76) for the symplectic representation, had not been studied in detail. In the previous section, the Hamiltonian representation appeared as indeed guaranteed at arbitrary order, and as naturally provided by the procedure as a maximal

guiding-center reduction, whereas the symplectic representation appeared as submitted to a condition at each order in $\varepsilon$: in Eq. (2.76), the term inside the bracket must have no overall contribution of order 0 in $\phi$. This could explain why first papers addressing both the first and the second order in $\varepsilon$ systematically used the Hamiltonian representation; the symplectic representation was used only later, when the condition was observed as verified.

The equivalent relation introduced in [23] relies on a relationship between these representations, which allows to go from one to the other. Hence it might seem that it guarantees the existence of the symplectic representation, but it is not the case.

The underlying idea (see Eq. (17) in [23]) is the following. Start from the reduced Lagrangian written in symplectic representation (we use here the variable $p_\parallel$ instead of $\phi$, and the symbol $(\Pi_\parallel)_n$ for the higher-order terms of $\overline{\Gamma}^{\mathsf{b}}$ in order to better agree with the notation used in [23]):

$$\begin{aligned}\overline{H} &= \mu B + \tfrac{\overline{p}_\parallel^2}{2}\,,\\ \overline{\Gamma}^{\mathsf{b}} &= \overline{p}_\parallel + \sum_{n\geqslant 1}(\Pi_\parallel)_n\,.\end{aligned}$$

Then, redefine the reduced parallel momentum so as to absorb all the series $\overline{\Gamma}^{\mathsf{b}}$ in it:

$$\overline{p}'_\parallel := \overline{p}_\parallel + \sum_{n=1}(\Pi_\parallel)_n\,. \tag{2.85}$$

With this coordinate, the symplectic part of the Lagrangian is fully reduced:

$$\overline{\Gamma}^{\mathsf{b}} = \overline{p}'_\parallel\,.$$

To obtain the reduced Hamiltonian with this coordinate, one just inverts the series (2.85):

$$\overline{p}_\parallel := \overline{p}'_\parallel - \sum_{n\geqslant 1}(\Pi_\parallel)_n = \overline{p}'_\parallel + \sum_{n\geqslant 1}(\Pi'_\parallel)_n\,, \tag{2.86}$$

with $(\Pi'_\parallel)_n$ some coefficients easily obtained by inserting iteratively the first equality in the occurrences of $\overline{p}_\parallel$ in $\sum_{n=1}(\Pi_\parallel)_n$, as is standard to invert a near-identity series.

Then the reduced Hamiltonian in the new coordinate writes

$$\overline{H} = \mu B + \tfrac{1}{2}\big[\overline{p}'_\parallel + \sum_{n\geqslant 1}(\Pi'_\parallel)_n\big]^2\,.$$

It is a full series in $\varepsilon$, which corresponds to the Hamiltonian representation. This is a constructive procedure showing that when the symplectic representation exists, then the Hamiltonian representation exists and is easily obtained.

Now, what is actually needed is to go in the reverse direction, since the derivation of the guiding-center reduction shows that the Hamiltonian representation is natural and guaranteed to exist, whereas the symplectic is suspected of having existence conditions.

It turns out that the procedure in the reverse direction can break down. Start from the reduced Lagrangian written in Hamiltonian representation:

$$\begin{aligned}\overline{H} &= \mu B + \tfrac{\overline{p}_\parallel^2}{2} + \sum_{n\geqslant 1}\overline{H}_n\,,\\ \overline{\Gamma}^{\mathsf{b}} &= \overline{p}_\parallel\,.\end{aligned} \tag{2.87}$$

Then, redefine the reduced parallel momentum so as to absorb all the higher-order terms $\sum_{n\geqslant 1}\overline{H}_n$ in the term with $\overline{p}_\parallel$:

$$\overline{p}'^2_\parallel := \overline{p}_\parallel^2 + 2\sum_{n\geqslant 1}\overline{H}_n\,. \tag{2.88}$$

With this coordinate, the Hamiltonian part of the Lagrangian is fully reduced:

$$\overline{H} = \mu B + \frac{\overline{p}_{\|}'^2}{2} .$$

To obtain the reduced symplectic Lagrangian with this coordinate, one just inverts the near-identity transformation (2.88) by writing it as:

$$\overline{p}_{\|}' := \pm\sqrt{\overline{p}_{\|}^2 + 2\sum_{n\geqslant 1}\overline{H}_n} = \overline{p}_{\|}\sqrt{1 + \frac{2\sum_{n\geqslant 1}\overline{H}_n}{\overline{p}_{\|}^2}} , \tag{2.89}$$

and then by expanding the term $\sqrt{1+\varepsilon}$. This assumes that the ratio $\frac{2\sum_{n=1}\overline{H}_n}{\overline{p}_{\|}^2}$ is small. The point is that this condition is not guaranteed a priori, even if the series is near-identity.

Indeed, the derivations here are formal. "Near-identity" has only a dimensional meaning. It means that the ratio between the first and zeroth-order term is of order $\varepsilon = r_L\nabla$, but only in dimension, its value might not to be small if it is multiplied by a large dimensionless factor such as $1/\cos\varphi$, as in (2.89). Thus, a division by $\overline{p}_{\|}$ can cause a singularity.

When going from the symplectic to the Hamiltonian representation, no such a division was needed, since the series inversion (2.86) just consisted in composing series. On the contrary, when starting from the Hamiltonian representation, the series inversion (2.89) involves a division by $\overline{p}_{\|}^2$. This causes a singularity if $\overline{H}_n$ contains a term of order 0 or 1 in $\overline{p}_{\|}$.

It is interesting to see that difficulties arise here at $\overline{p}_{\|} = 0$ (or equally at $\phi = 0$), which is precisely where they appeared in the guiding-center reduction in the previous section. This suggests that singularities in $\phi = 0$ are indeed a difficulty for the symplectic representation. Accordingly, at each order, it can be used only when the absence of singularity in (2.76) is verified.

### 2.3.4 Gyro-gauge independence

The intrinsic formulation of the guiding-center reduction was motivated by questions about the traditional gyro-angle variable $\theta$. The derivation with the coordinate $\mathsf{c}$ shows that it does succeed in shedding light on those questions.

First, the traditional coordinate was a detour. In all guiding-center works, the variable $\theta$ never appears in itself (except in its own definition and subsequent relations); for instance, it does not appear explicitly in one of the components of $\mathbf{G}$ or $\overline{\Gamma}$, which all depend on $\theta$ only through the corresponding physical quantity $\mathsf{c}$ (or $\mathsf{a} := \mathsf{b}\times\mathsf{c}$); even the gyro-angle component of the generator $\mathbf{G}^\theta$ verifies this statement. The detour is not given by the physics, since it imposes to fix arbitrarily a gauge $\mathsf{e}_1(\mathbf{q})$, which is not related to the physics of the problem. The role of the intrinsic approach was to avoid this detour, and it achieves its goal since it obtains the full guiding-center results without introducing any gauge and by working purely with $\mathsf{c}$.

From a mathematical point of view also, the use of the variable $\theta$ was not completely satisfactory, because the gyro-angle corresponds to a circle bundle [25, 81, 90]. The traditional coordinate $\theta$ makes this structure somehow disappear, because the manifold trivially becomes $\mathbb{R}^3\times\mathbb{S}^1$. It is why the variable $\theta$ does not have a global existence in a general magnetic geometry [25]. On the contrary, when using the physical variable $\mathsf{c}$, the circle bundle naturally arises: as $\mathsf{c}$ is defined on a space-dependent circle, spatial displacements imply a variation of $\mathsf{c}$, so that a covariant derivative is involved, which encodes the circle-bundle geometry for the gyro-angle [25] and does not imply some restricted class of circle bundle. A more detailed study of the coordinate system is outside the scope of the present chapter, and will be reported in Chapter 3.

In some way, the relevance of this coordinate is obvious, since it just results from keeping the initial coordinate, in which all the circle-bundle picture was included. From this point of view, performing the derivation with this variable is a way to see how it globally agrees with the physics and the mathematics of the problem, and to make intrinsic definitions arise naturally for all the quantities involved in the process.

Indeed, the previous section shows that the reduction follows the same procedure with the vectorial constrained coordinate $\mathsf{c}$ as with the scalar coordinate $\theta$, but that there are slight changes

in the quantities used. The gauge vector $\mathbf{R}$ came as naturally replaced by the connection $\mathbf{R}_g$ for the covariant derivative. The generator of Larmor gyrations $\partial_\theta$ came with an intrinsic definition $-\mathsf{a}\cdot\partial_\mathsf{c}$. The basic 1-form for the gyro-angle $d\theta$ appeared as replaced by a non-closed 1-form $\delta\theta$, which agrees with the fact that $\theta$ considers the circle bundle trivial, whereas it should not. This implied to use more intrinsic definitions for the operations used, such as (2.55) and (2.56) for exterior derivatives. Also, this implied to be careful on how the basis of 1-forms and of vector fields are chosen, but the natural ones were found to agree with each other.

Thus, the formalism with $\mathsf{c}$ is slightly more involved, but it perfectly fits both with the physics and the mathematics of the problem, which correspond to a non-trivial circle bundle.

From a formal point of view, the results with the physical variable $\mathsf{c}$ mainly correspond to replacing the gauge vector $\mathbf{R}$ by the connection term $\mathbf{R}_g$. Thus, they include the standard gauge-dependent results as a special case, but they are more general: in the usual approach $\mathbf{R}(\mathbf{q})$ depends only on the position and cannot be chosen freely (e.g. $\mathbf{R} = 0$ is not possible [25, 86, 90]); here, $\mathbf{R}_g(\mathbf{q}, \mathbf{p})$ can be any function of the position and momentum. Especially, the physical definition of $\mathsf{c}$ corresponds to the function $\mathbf{R}_g := -\phi\nabla\mathsf{b}\cdot\mathsf{c}$, which depends also on the momentum and preserves the polynomiality in $\phi$.

Other gauge-dependent quantities are interesting to consider. In previous works, the coordinate transformation $\theta \longrightarrow \overline{\theta}$ was gauge dependent (see for instance Eq. (30c) in [88], or in [17] the solution for the generator $\mathbf{G}_1^\zeta$ below Eq. (5.45)), as well as the definition of the coordinate $\theta$, and also the gradient $\partial_{\mathbf{q}|\theta}$. It is why the gauge vector $\mathbf{R}$ was involved in some of the resulting formulae, e.g. the Poisson bracket, in such a way as to make all the physical or geometrical (intrinsic) quantities gauge-independent. For instance, in the Poisson bracket, gradients appear only in the combination $\nabla_* := \nabla + \mathbf{R}\partial_\theta$ [17]. It would be interesting to interpret it as the gradient corresponding to a special gauge, because it would remove the appearance of the gauge vector in all the derivation, and would simplify computations. The issue is that it is not possible because it would correspond to fix the gauge in such a way that $\nabla\mathsf{e}_1\cdot\mathsf{e}_2 = 0$, which is not possible even locally [86, 90].

In the gauge-independent approach, all the coordinates, including the gyro-angle $\mathsf{c}$, are gauge-independent, as well as the transformation $\mathsf{c} \longrightarrow \overline{\mathsf{c}}$: at first order, it is not transformed by $\mathsf{a}\mathbf{G}_1^\theta$, because the covariant derivative must be taken into account, which means that it is given by $\mathsf{a}\mathbf{G}_1^\theta - \mathbf{G}_1^\mathbf{q}\cdot\nabla\mathsf{c}$. This last quantity is indeed independent of the connection vector $\mathbf{R}_g := \nabla\mathsf{c}\cdot\mathsf{a}$, as can be verified in (2.61). In the same way, at higher order, all the transformed coordinates $\overline{\mathbf{z}} = ...e^{-\mathbf{G}_2}e^{-\mathbf{G}_1}\mathbf{z}$ will be independent of $\mathbf{R}_g$, where $\mathbf{z} := (\mathbf{q}, \phi, \mu, \mathsf{c})$ are physical coordinates.

Gradients are also involved in combinations involving $\mathbf{R}_g$ (see Eq. (2.80), for instance). This is no surprise, since the connection on the fiber bundle involves some arbitrariness, but the combinations can always be written $\nabla_* := \nabla + \mathbf{R}_g\partial_\theta$, which is connection-independent. In addition, it can be interpreted as the covariant derivative associated with the trivial connection $\mathbf{R}_g = 0$. Thus, when working with the coordinate $\mathsf{c}$, this choice can be used to simplify computations and to make them connection-independent.

### 2.3.5 Maximal vs. minimal reduction

The derivation procedure confirms the respective interests of Lie transforming the velocity vector field and the Lagrangian 1-form.

As with concerns the minimal requirements for the guiding-center transformation, working on the equation of motion is much more efficient, since it systematically obtains the fluctuating part of the reduced motion just by inverting the operator $\partial_\theta$.

The procedure with the Lagrangian is much more involved, as can be seen in previous sections, especially because the order mixing makes the scheme more elaborated and because the algorithmic stage begins only at higher order: the induction matrix mixes up the orders, changes at each order for $n \leqslant 3$, and involves differential operators in some coefficients. It is why in this chapter, as in previous works, only a part of $\mathbf{G}_2$ is explicitly computed, whereas Chapter 1 directly obtained the full second-order generator $\mathbf{G}_2$.

In addition, the minimal guiding-center reduction can hardly be obtained by working on the Lagrangian, because going from the Lagrangian to the motion mixes the components up. To guarantee an averaged slow reduced motion for the four coordinates $(\overline{\mathbf{q}}, \overline{\phi})$, one would need to average all of the seven components of the Lagrangian, which is not a minimal transformation.

As with concerns the additional requirements for the slow dynamics, working on the equation of motion is not efficient, because the equations to be solved are secular differential equations that are not simple to deal with, as shown in Chapter 1. Working on the Lagrangian is more efficient, because it essentially consists in algebraic equations, which deals the same way with gyro-averages as with gyro-fluctuations. This makes it easy to identify good choices for the averaged transformation generator $\mathtt{avg}(\mathbf{G}_n)$ in order to obtain a reduced Lagrangian as strongly reduced as possible. Thus, it provides a maximal guiding-center reduction almost as simply as the minimal one.

Also, working on the Lagrangian 1-form makes it easy to impose requirements on the reduced Hamiltonian structure, for instance to obtain a quarter-canonical structure for the coordinates $(\overline{\mu}, \overline{\theta})$, which both provides a constant of motion $\overline{\mu}$ and a Hamiltonian sub-dynamics for the 4-dimensional reduced motion $(\dot{\overline{\mathbf{q}}}, \dot{\overline{\phi}})$.

## Conclusion

The full guiding-center reduction can be performed to arbitrary order in the Larmor-radius expansion by Lie transforming the Lagrangian 1-form while keeping physical gyro-gauge independent variables as coordinates, following the same procedure as when working with the standard gauge-dependent gyro-angle.

For higher orders, the procedure was shown to be completely algorithmic. The pivotal role is played by the inverse of the lowest-order Lagrange matrix $\overline{\omega}_{-1} + \overline{\omega}_0 + \overline{\omega}_1$, together with a differential equation for the function $S_n$. An extended matrix was defined and used to explicitly solve the induction equation to arbitrary order in the Larmor radius.

The results exactly agree with previous works, but they were obtained without introducing any gyro-gauge, and working purely with the physical coordinate $\mathsf{c}$ as the gyro-angle coordinate. In addition, the choice of the cotangent of the pitch-angle as a coordinate for the parallel velocity made the results purely polynomial in the coordinates and monomial in $\sqrt{\mu}$ and $\sqrt{B}$.

Compared to the method by Lie transforming the equations of motion, the process is much more elaborated, especially because of the order mixing, but it easily obtains a much stronger result. It does not rely on differential equations for the reduced motion, but on algebraic equations for the reduced Lagrangian. A quarter-canonical reduced Hamiltonian structure provides a constant of motion $\overline{\mu}$ and a Hamiltonian sub-dynamics for the 4-dimensional slow reduced motion $(\dot{\overline{\mathbf{q}}}, \dot{\overline{\phi}})$. In addition, the procedure makes the reduced dynamics trivial not only in the gyro-fluctuating components of the Lagrangian, but also in six of the averaged components out of seven.

As a result, all but one of the components of the reduced Lagrangian 1-form are put to zero for all orders higher than two. Only one of them cannot be made exact, and is given by a whole series. The two canonical choices are recovered: either to enclose the series into the Hamiltonian (Hamiltonian representation), then the reduced Poisson bracket is exact, or to enclose the series into the spatial component of the Lagrangian parallel to the magnetic field (symplectic representation), then the Hamiltonian is exact and the uncertainty of the reduced motion is traduced by five coefficients, and especially a kind of reduced magnetic field $\mathbf{B}_*$.

The Hamiltonian representation appeared as naturally induced by the reduction process, whereas the symplectic representation is subjected to a condition at each order, to avoid a singularity in $\phi = 0$, i.e. at the bounce points of particle trajectories.

These representations make the reduction maximal because for a general magnetic field, the procedure cannot get a stronger reduction for which even the last component of the Lagrangian would be zero. The obstruction originates from the special role of the magnetic moment; in the

Hamiltonian representation, this can be viewed because the magnetic-moment component of the transformation generator can remove only the fluctuating part of the reduced Hamiltonian function, since the averaged part is imposed by the requirement of adiabatic invariance.

The use of gauge-independent coordinates had little effect on the reduction procedure. All the ingredients of the standard reduction with gauge-dependent coordinates were found to be present, but they naturally arose with an intrinsic definition or they were replaced by a different intrinsic object playing a similar role.

It was observed to fit in with both the physics and the mathematics of the system, by restoring the general circle-bundle framework, which practically disappeared with the coordinate $\theta$, and by making the coordinates directly induced by the physical state of the system.

For instance, the gauge removal introduced a vectorial quantity $\mathsf{c}$ for the gyro-angle coordinate. This caused the coordinate system to be constrained and implied a connection for the covariant derivative on a space-dependent circle, which is directly linked to the circle-bundle structure underlying in the gyro-angle coordinate and which replaced the gauge vector of the gauge-dependent approach.

The closed 1-form $d\theta$ was replaced by a non-closed 1-form $\delta\theta$, which is related to the non-triviality of the circle bundle for a general magnetic geometry.

In previous works relying on the coordinate $\theta$, the gauge independence of the physical results implied that gradients were systematically involved in special expressions, which were not related to derivative operators because no gauge fixing were suited to them. These expressions were found to be related to covariant derivatives corresponding to suitable connections.

Unlike the gauge fixing for the coordinate $\theta$, the connection fixing for the coordinate $\mathsf{c}$ depends not only on the position, but on the momentum as well. This is all the more convenient as the physical definition of $\mathsf{c}$ and its associated connection depend on both the position and the momentum. In addition, this removes one of the assumptions causing the presence of anholonomy in the gyro-angle motion.

So, the intrinsic gyro-angle coordinate $\mathsf{c}$ is a way to tackle some of the questions involved in the guiding-center anholonomy and gauge- (or connection-) arbitrariness. These questions will be the topic of the next chapter.

In the present chapter, we focused on the formal derivation of perturbation series, as is usual in guiding-center works, and as is the standard first step in perturbation theory [87]. A possible continuation for future works will be to relate these formal expansions with asymptotic series, in a similar way as what was done in [4] for Kruskal's work, and what is beginning being done about Littlejohn's works [47,48]. Also, as usual in perturbation theory, convergence of the guiding-center series is an interesting question to investigate, probably with methods of accelerated convergence [10]. In these attempts, the structures of the expansion series, such as the polynomiality induced by the cotangent of the pitch-angle, are expected to play a role.

## Appendix: Mechanism of the reduction

In this appendix, we introduce the mechanism at work when Lie transforming the Lagrangian 1-form. Indeed, the basic ideas of the derivation are very elementary, but they are hidden by the details of the procedure, which are rather involved because of some order mixing and other subtleties between algebraic and differential integrability conditions. In addition, practical computations in the case of the guiding-center are somehow intricate. All the same, the method is very efficient and has quite a wider domain of application than just the guiding-center reduction. So, it seems useful to give a general overview of the method for people not familiar with it.

### Fundamental ingredients

A) The goal is to solve Eqs. (2.16) for the unknowns $\mathbf{G}_n$ and $S_n$, with the requirements (2.20), (2.21) and (2.23) identified above. Ideally, the maximal reduction sets $\overline{\Gamma}_n = 0$ for all higher orders $n$, as a result of (2.21).

The solution is built order by order in the Larmor radius. Each order implies to solve the equation

$$\overline{\Gamma}_n = \mathbf{G}_{n+1}\cdot\omega_{-1} + \mathsf{R}_n + dS_n\,,$$

where $\mathsf{R}_n$ is a shorthand for all other terms, that do not contain the highest-order generator $\mathbf{G}_{n+1}$.

The very basic idea of the reduction is that $\mathbf{G}_{n+1}$ is involved only through a matrix product. So, at any order, the solution is just given by a matrix inversion

$$\mathbf{G}_{n+1} = (\omega_{-1})^{-1}\cdot\big[\overline{\Gamma}_n - \mathsf{R}_n - dS_n\big]\,, \tag{2.90}$$

provided $\omega_{-1}$ is invertible; then $\overline{\Gamma}_n$ can be chosen zero, and $S_n$ is not useful and can also be set to zero. This idea of matrix inverse is the key ingredient of the underlying mechanism, even if the corresponding basic picture is not true at the lowest orders, and at higher orders, it is slightly complicated by some order mixing and integrability conditions, especially for $S_n$.

B) As a matter of fact, the matrix $\omega_{-1}$ is usually not invertible, since it corresponds to the fast part of the dynamics, here the Larmor gyration, which does not concern all the phase-space coordinates.

At zeroth order, under the requirement $\overline{\Gamma}_n = 0$, Eq. (2.16) writes

$$\mathbf{G}_1\cdot\omega_{-1} = -\Gamma_0 - dS_0\,.$$

It has a solution only if the right-hand side $-\Gamma_0 - dS_0$ is in the range of the matrix $\omega_{-1}$ (solvability condition); this is a necessary condition for the corresponding reduction to exist. Usually, it is not verified for so strong a requirement as $\overline{\Gamma}_0 = 0$, and the reduced Lagrangian $\overline{\Gamma}_0$ has to be used as a softening parameter. Then Eq. (2.16) writes

$$\mathbf{G}_1\cdot\omega_{-1} = \overline{\Gamma}_0 - \Gamma_0 - dS_0\,. \tag{2.91}$$

One has to check that, with the freedoms $S_0$ and $\overline{\Gamma}_0$, the solvability condition can be satisfied at least for the minimal requirement (2.20) and if possible for the intermediate requirement (2.23); then, the reduction is possible, and a maximal requirement (2.21) can be considered by trying to remain as close as possible to the condition $\overline{\Gamma}_0 = 0$.

At that point, the solution exists, but it is not unique; it is defined to within an element of the kernel of $\omega_{-1}$. The choice of this element may be free at this stage of the reduction, but care must be taken that it may be constrained by the solvability conditions at the following order.

At the next order $n = 1$, Eq. (2.16) writes

$$\mathbf{G}_2\cdot\omega_{-1} + \tfrac{\mathbf{G}_1}{2}\cdot(\omega_0 + \overline{\omega}_0) = \overline{\Gamma}_1 - dS_1\,. \tag{2.92}$$

Now, the pivotal matrix $\mathsf{M}_1$ to be inverted is the set of $\omega_{-1}$ and $\frac{\omega_0+\overline{\omega}_0}{2}$, acting on the set of unknown components of $(\mathbf{G}_2, \mathbf{G}_1)$. Its rank is greater than (or equal to) the rank of $\omega_{-1}$. Care must be taken that some of the coordinates of $\mathbf{G}_1$ are already determined. This introduces some order mixing, where some components of $\mathbf{G}_n$ are determined at order $\overline{\Gamma}_{n-1}$, others are computed at order $\overline{\Gamma}_n$ at the same time as some of the components of $\mathbf{G}_{n+1}$.

Notice that $\mathsf{M}_1$ cannot be invertible on the unknown components of $\mathbf{G}_2$, because of the non-trivial kernel of $\omega_{-1}$. In the same way as at zeroth order, this kernel has to be excluded when studying the invertibility of $\mathsf{M}_1$, because it will be involved only at the following order.

Then, if the pivotal matrix $\mathsf{M}_1$ is invertible for the ideal requirement $\overline{\Gamma}_1 = 0$, then the solution exists and is unique. Otherwise, there is again both a solvability condition and a non-uniqueness

of the solution. More precisely, the solvability condition means that there is a solution only if the right-hand side is in the range of the matrix to be inverted, and that in order to fulfil this condition, the reduced Lagrangian $\overline{\Gamma}_1$ may have be chosen non-zero, but having as many null components as possible. The non-uniqueness means that the solution is determined only to within an element of the kernel of the pivotal matrix $\mathsf{M}_1$. The choice of this element may be free, but care must be taken that it may be constrained by the solvability condition at the following order.

At the next orders, the same process goes on. The pivotal matrix $\mathsf{M}_n$ evolves at each order and its rank increases to determine more and more of the unknowns. At high orders, it becomes of constant rank, and actually it becomes the same at each order. The critical value of $n$ at which this occurs will be denoted by $n_c$, and the corresponding pivotal matrix will be denoted by $\mathsf{M}_\infty$. So, for $n \geqslant n_c$, the pivotal matrix verifies $\mathsf{M}_n = \mathsf{M}_\infty$, whereas for $n = n_c - 1$, it verifies $\mathsf{M}_n \neq \mathsf{M}_\infty$.

This can be explained as follows: Eq. (2.19) generically (i.e. at high orders) writes

$$\begin{aligned}\overline{\Gamma}_n = &\left[(\mathbf{G}_{n+1}\cdot d) + (\mathbf{G}_n\cdot d)(\mathbf{G}_1\cdot d) + ... + \tfrac{(\mathbf{G}_1\cdot d)^{n+1}}{n+1!}\right]\Gamma_{-1}\\ &+\left[(\mathbf{G}_n\cdot d) + (\mathbf{G}_{n-1}\cdot d)(\mathbf{G}_1\cdot d) + ... + \tfrac{(\mathbf{G}_1\cdot d)^{n}}{n!}\right]\Gamma_0 + dS_n\,.\end{aligned} \tag{2.93}$$

In this analysis, low orders are excluded because there would be some additional coefficients coming from the exponential series: for instance for $n = 1$, the term $(\mathbf{G}_n\cdot d)(\mathbf{G}_1\cdot d)$ has a factor $1/2$ and is confounded with the last term $\frac{(\mathbf{G}_1\cdot d)^2}{2!}$.

Denoting $\mathbf{G}_n\cdot d$ by $\mathsf{G}_n$, and grouping together the highest-order Lie derivatives, which contain the unknowns (which are some of the components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$), the previous equation becomes

$$\begin{aligned}\overline{\Gamma}_n = \;& \mathsf{G}_{n+1}\Gamma_{-1}\\ &+\mathsf{G}_n\big(\mathsf{G}_1\Gamma_{-1}+\Gamma_0\big)\\ &+\mathsf{G}_{n-1}\left[\big(\mathsf{G}_2+\tfrac{\mathsf{G}_1^2}{2}\big)\Gamma_{-1}+\mathsf{G}_1\Gamma_0\right]\\ &+...\\ &+\mathsf{G}_1\left[\tfrac{(\mathsf{G}_1)^n}{n+1!}\Gamma_{-1}+\tfrac{(\mathsf{G}_1)^{n-1}}{n!}\Gamma_0\right]+dS_n\,.\end{aligned} \tag{2.94}$$

Using (2.16) for the lowest-orders reduced Lagrangian $\overline{\Gamma}_k$, that are already known, the previous formula can be rewritten

$$\begin{aligned}\overline{\Gamma}_n &= \mathsf{G}_{n+1}\big(\overline{\Gamma}_{-1}-dS_{-1}\big)+\mathsf{G}_n\big(\overline{\Gamma}_0-dS_0\big)+\mathsf{G}_{n-1}\big(\overline{\Gamma}_1-dS_1\big)\\ &\qquad +...+\mathsf{G}_1\left[\tfrac{(\mathsf{G}_1)^n}{n+1!}\Gamma_{-1}+\tfrac{(\mathsf{G}_1)^{n-1}}{n!}\Gamma_0\right]+dS_n\\ &= \mathbf{G}_{n+1}\cdot\overline{\omega}_{-1}+\mathbf{G}_n\cdot\overline{\omega}_0+...+dS_n+\mathsf{R}_n\,,\end{aligned} \tag{2.95}$$

where $\mathsf{R}_n$ indicates all other terms, that are already known, since they do not involve $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$.

Eq. (2.95) shows that the pivotal matrix $\mathsf{M}_n$ is given by the set of matrices $\overline{\omega}_{-1}$, $\overline{\omega}_0$, etc., acting on the set of unknown components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$. The matrix $\mathsf{M}_n$ is the same at any (high) order. It is exactly given by the reduced Lagrange matrix at lowest orders.

As a consequence,the pivotal matrix $\mathsf{M}_\infty$ for all high orders is identified as soon as the set of $(\overline{\omega}_{-1}, \overline{\omega}_0, ...)$ is observed to be invertible on the set of unknown components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$, i.e. as soon as $\overline{\omega}_{-1}+\overline{\omega}_0+...$ becomes invertible. We will call $n_b$ the order such that $\overline{\omega}_{-1}+\overline{\omega}_0+...+\overline{\omega}_{n_b}$ is invertible, whereas $\overline{\omega}_{-1}+\overline{\omega}_0+...+\overline{\omega}_{n_b-1}$ is not.

At that point, the basic picture of A) has become a simple picture B), which includes two stages. At low orders it consists in dealing with non invertible pivotal matrices changing at each order, and in choosing $\overline{\Gamma}_n$ such that it both fulfils the requirement and leads to an interesting invertible matrix $\overline{\omega}_{-1}+\overline{\omega}_0+...+\overline{\omega}_{n_b}$. Then at high orders, the induction becomes just a matrix inversion $\mathsf{M}_\infty^{-1}$, as in the initial basic picture A).

C) The simple picture B) has to be refined. Between these two stages, an intermediate stage takes place, since usually $n_c > n_b + 1$. For $n \in \{n_b+1, n_b+2, ..., n_c-1\}$, the higher-orders pivotal matrix is already known but not yet efficient.

The reason is that Eqs. (2.93)-(2.95) hold only for orders that are high enough, because of the coefficients generated by expanding the exponentials. If the factors involving $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$ have some coefficients non unity, then Eq. (2.95) does not hold, which spoils the conclusion.

But this concerns only low orders. Eq. (2.94) shows that the coefficients will be unity as soon as $n > 2n_b + 1$, which means that $n_c \leqslant 2n_b + 2$.

As an example, consider $\overline{\omega}_n$ for $n = 0$. When computing the next order $\overline{\Gamma}_{n+1} = \overline{\Gamma}_1$, then $\overline{\Gamma}_0$ is already known, but it is not yet efficient: in Eq. (2.95)

$$\begin{aligned}\overline{\Gamma}_1 &= \mathsf{G}_2\Gamma_{-1} + \mathsf{G}_1\big(\tfrac{1}{2}\mathsf{G}_1\Gamma_{-1} + \Gamma_0\big) + dS_1 \\ &\neq \mathbf{G}_2{\cdot}\omega_{-1} + \mathbf{G}_1{\cdot}\overline{\omega}_0 + dS_1\,,\end{aligned}$$

the operator to be inverted is not just the set of $(\overline{\omega}_{-1}, \overline{\omega}_0)$ because of the factor 1/2 in the first line, which comes because the generator $\mathbf{G}_1$ outside the parenthesis has the same order as the generator $\mathbf{G}_1$ inside the parenthesis. For all higher orders, this will not happen, as is illustrated by the next order

$$\begin{aligned}\overline{\Gamma}_2 &= \mathsf{G}_3\Gamma_{-1} + \mathsf{G}_2\big(\mathsf{G}_1\Gamma_{-1} + \Gamma_0\big) + o.t. \\ &= \mathbf{G}_2{\cdot}\omega_{-1} + \mathbf{G}_1{\cdot}\overline{\omega}_0 + o.t.\,,\end{aligned}$$

where $o.t.$ is used for "other terms", in order to avoid writing uninteresting terms.

D) The order mixing can also slightly complicate the picture of C), by spoiling the linear algebraic framework, mainly at order $n = 1$. Indeed, the first equation to be solved for $\mathbf{G}_2$ is (2.92). However, if some of the components of $\mathbf{G}_1$ are still not determined at that point (this is fairly general as $\omega_{-1}$ is usually not invertible), then they can be involved in a differential equation. Indeed, $\overline{\Gamma}_0$ can be undetermined at that point, and Eq. (2.92) must be let under its initial form (2.16)

$$\overline{\Gamma}_1 = \Big(\mathbf{G}_2 + \tfrac{\mathbf{G}_1{\cdot}d}{2}\mathbf{G}_1\Big){\cdot}\omega_{-1} + \mathbf{G}_1{\cdot}\omega_0 + dS_1\,,$$

which is now a differential equation for $\mathbf{G}_1$, and may even be non-linear in the unknown components of $\mathbf{G}_1$. This can make the scheme much more complicated: even solvability conditions may be difficult to identify.

E) Finally, one last point has to be taken into account as well and still makes the scheme more elaborated than the picture D) above. The pivotal matrix $\mathsf{M}_n$ determines the unknown components of the generator $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$, but this can generate non-zero time-component $\mathbf{G}_n^t$ for the generator.

For a time-independent transformation, the requirement $\mathbf{G}_n^t = 0$, reduces the dimension of the effective generator $\mathbf{G}_n$. Then the pivotal matrix can be inverted only if some integrability conditions are fulfilled. Another way of saying it is that $\overline{\Gamma}_n$ has seven components (seven requirements) whereas $\mathbf{G}_n$ has only six freedoms. The additional freedom comes from the gauge function $S_n$.

Actually, the presence of this integrability condition for the pivotal matrix $\mathsf{M}_\infty$ is completely general and comes because $\mathsf{M}_\infty$ is antisymmetric. It is not invertible on the 7-dimensional space $(\mathbf{q}, \mathbf{p}, t)$, and can be invertible only on a subspace, e.g. on the 6-dimensional phase space $(\mathbf{q}, \mathbf{p})$. For a symplectic Hamiltonian system, the high-orders pivotal matrix $\mathsf{M}_\infty$ is indeed invertible when restricted to the phase space, since the Lagrange 2-form $\omega_s$ is invertible, and so is $\overline{\omega}_s$.

So, the gauge function $S_n$ is not determined by the algebraic matrix inversion, but by the solvability condition for the matrix inversion. Furthermore, it appears in a differential equation. Existence of solution for this differential equation can involve other integrability conditions. For instance, in an equation such as

$$\partial_\theta S_n = f_n\,,$$

inverting $\partial_\theta S$ implies the function $f_n$ to have no gyro-average.

As a result, both the algebraic and the differential integrability conditions must be played with so as to make $\overline{\Gamma}_n = 0$. If it is not possible, one has to choose a non-zero reduced Lagrangian

$\overline{\Gamma}_n \neq 0$. This means playing with the requirements also, and releasing them slightly, so as to make the integrability conditions fulfilled and at the same time to keep $\overline{\Gamma}$ as strongly reduced as possible.

All these features do not spoil the algorithmic character of the reduction for high orders, because the differential scheme is very simple (the operators are just $\partial_{z^k}$) and in addition, at all $n \geqslant n_c$, the algebraic scheme for $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$ is fixed, which makes it possible to conclude about the differential scheme for $S_n$ so as to make the resulting reduction maximal.

At the end, the induction relations can be written in matrix form provided some coefficients of the matrix are differential operators. By such a redefinition of the matrix $\mathsf{M}_\infty$, the induction relation for high orders $n \geqslant n_c$ just relies on a matrix inverse $\mathsf{M}_\infty^{-1}$. Then, the basic picture of Eq. (2.90) becomes efficient: Eqs. (2.16) are solved at arbitrary order through a formula completely analogous to (2.90), even if the framework is much more elaborated. We want to stress this fact because the order mixing and the presence of integral operators may hide the triviality of the induction mechanism.

### Resulting procedure in three stages

The previous subsection shows that the reduction is performed in three stages. The first stage corresponds to the first few orders $n \leqslant n_b$. The work consists in verifying that the freedoms can be used both to make the solvability conditions satisfied and to get an interesting invertible matrix $\mathsf{M}_\infty$. At the end $n = n_b$, the invertible high-order pivotal matrix $\mathsf{M}_\infty$ becomes identified, and the first stage is ended.

The second stage corresponds to a transition stage. The pivotal matrix for high order is identified, but it is still not efficient at that order. The goal is only to check that the solvability conditions can be satisfied at these intermediate orders.

The third stage begins at order $n = n_c$, i.e. as soon as the matrix to be inverted becomes $\mathsf{M}_\infty$. From that order on, it is sure that the reduction can be performed to any order in the Larmor radius. As the matrix is now invertible, the solution exists and is unique to each order, and the process becomes fully algorithmic.

In order to get a formula analogous to (2.90), the pivotal matrix must be extended to include the gauge function $S_n$, and some coefficients of the inverted matrix $\mathsf{M}_\infty^{-1}$ are then integral operators. In addition, in order to deal with the order mixing, some intermediate quantities must be introduced to isolate the components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, \mathbf{G}_{n-1})$ that are already known from the ones that are not identified yet.

For example, if the pivotal matrix $\mathsf{M}_\infty$ involves only $\overline{\omega}_{-1}$ and $\overline{\omega}_0$. Then, Eq. (2.16) or (2.95) writes

$$\overline{\Gamma}_n = \mathbf{G}_{n+1}\cdot\overline{\omega}_{-1} + \mathbf{G}_n\cdot\overline{\omega}_0 + dS_n + \mathsf{R}_n\,, \tag{2.96}$$

where $\mathsf{R}_n$ indicates all terms of (2.16), that do not depend on the unknowns, which are the gauge function $S_n$ and some components of $(\mathbf{G}_{n+1}, \mathbf{G}_n)$. These last quantities can be grouped into one single vector

$$\mathbf{g}_n := \big(\mathbf{G}_{n+1}, \mathbf{G}_n, S_n\big)\,. \tag{2.97}$$

The pivotal matrix $\mathsf{M}_n$ is then extended to act on all $\mathbf{g}_n$ (including the gauge function) in (2.96) and is defined by

$$\mathsf{M}_n\cdot\mathbf{g}_n := \mathbf{G}_{n+1}\cdot\overline{\omega}_{-1} + \mathbf{G}_n\cdot\overline{\omega}_0 + dS_n\,.$$

As announced, some of its coefficients (the ones acting on the component $S_n$) are differential operators. With these conventions, Eq. (2.96) writes

$$\overline{\Gamma}_n = \mathsf{M}_n\cdot\mathbf{g}_n + \mathsf{R}_n\,.$$

Now, some of the components of $(\mathbf{G}_{n+1}, \mathbf{G}_n)$ are already identified at that order. Let us denote them by the index $a$, and the remaining components of $\mathbf{g}$, which are not identified are denoted by the index $\infty$:

$$\mathbf{g} = \big((\mathbf{g}_n)_a; (\mathbf{g}_n)_\infty\big)\,,$$

with $(\mathbf{g}_n)_a$ fully identified and all terms of $(\mathbf{g}_n)_\infty$ fully unknown. The Lagrangian writes

$$\overline{\Gamma}_n = (\mathsf{M}_n)_a \cdot (\mathbf{g}_n)_a + (\mathsf{M}_n)_\infty \cdot (\mathbf{g}_n)_\infty + \mathsf{R}_n \,, \tag{2.98}$$

with obvious definitions for the linear operators $(\mathsf{M}_n)_a$ and $(\mathsf{M}_n)_\infty$. By assumption, the quantities $\mathsf{R}_n$ and $(\mathsf{M}_n)_a(\mathbf{g}_n)_a$ are known; in addition, $(\mathsf{M}_n)_\infty = \mathsf{M}_n = \mathsf{M}_\infty$ is known and invertible. As a consequence, the induction relation writes

$$(\mathbf{g}_n)_\infty = \mathsf{M}_\infty^{-1} \cdot \left[\overline{\Gamma}_n - (\mathsf{M}_n)_a \cdot (\mathbf{g}_n)_a - \mathsf{R}_n\right] \,. \tag{2.99}$$

It is explicit and makes the basic picture (2.90) apply to all orders $n \geqslant n_c$. Some coefficients of $\mathsf{M}_\infty^{-1}$ are integral operators, since in the inverse matrix $\mathsf{M}_\infty$ some coefficients are differential operators.

A few comments are in place. First, some components of $\mathbf{G}_{n+1}$ remain non-identified after the order $n$; they must be excluded from $(\mathbf{g}_n)_\infty$ to get an invertible matrix, because they are elements of the kernel of $\omega_{-1}$ and will be determined at the next order; this is well illustrated by (2.66) and (2.70)-(2.71). Second, the components $(\mathbf{g}_n)_a$ can be extracted from $\mathbf{g}_n$ and its term $(\mathsf{M}_n)_a(\mathbf{g}_n)_a$ can be grouped with $\mathsf{R}_n$ (see Eq. (2.68)), which plays the same role. Last, the reduced Lagrangian $\overline{\Gamma}_n$ is in principle taken to be zero, but it was kept free because integrability conditions for $S_n$ can make it necessary to choose some of its components non-zero; then, it can be included in the vector $\mathbf{g}_n$, as is done in (2.66) and (2.70)-(2.71).

The final algorithm to be iterated for the $n$-th-order term is trivial: in Eq. (2.99), replace the lowest orders terms by their expression, already known, then compute the Lie derivatives involved in the term $\mathsf{R}_n$, and last apply the matrix product with $\mathsf{M}_\infty^{-1}$. The mechanism involves just two kinds of operations, derivatives and a matrix product, which can be easily implemented to arbitrary order using computer-assisted formal calculus.

The basic idea shown in (2.90) and (2.99) explains why Lie transforming the Lagrangian 1-form has the advantage of algebraic equations, which makes it easy to reduce also the averaged part of the reduced motion, and thus to get non-minimal guiding-center reductions. Indeed, computations for the non-minimal requirements are treated the same way as for the minimal ones, the only difference concerns the priority: if all requirements cannot be satisfied, then the order of priority may impose the requirements to be preferred and the ones to be released. This is an essential advantage of Lie transforming the Lagrangian.

But the overall process is much more involved than the method relying on a Lie transform of the equations of motion. This last has the essential advantage of relying on just a gyro-integral, which makes it much more efficient to work on the fluctuating part of the reduced dynamics and to perform the minimal guiding-center reduction, as is clear in Chapter 1.

In both cases, the reduction relies on explicit induction relations, but when working with the Lagrangian, the algorithmic stage (third stage introduced above) is efficient only for higher orders. For lowest orders, the reduction is not systematic at all, the choices are crucial to make the reduction work or not, but they must be guessed rather than derived. In addition, many solvability conditions appear in the process, and there is no a-priori guarantee that they can be satisfied.

In the case of the guiding-center reduction, good choices appear rather naturally, solvability conditions come as easily satisfied, and the reduction can be considered as rather straightforward, but two specificities must be taken into account.

Indeed, as expected, at each order, the fluctuating part of $\mathbf{G}_n$ is imposed by the minimal requirement (2.20), which means to put to zero the gyro-fluctuating part of the Lagrangian; and the averaged part is imposed by the other requirements (2.21) and (2.23), which mean to put to zero the averaged part of the Lagrangian as well (except that $\overline{\Gamma}_1^\theta = \overline{\mu}$).

However, one of the components of $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$ that remains not identified is already present in Eq. (2.96): $\mathtt{avg}(\Gamma_n^\mu)$ remains as a parameter in the right-hand side of (2.99).

Furthermore, the integrability conditions on $S_n$ cannot be fully satisfied, one of the optional requirements (2.21) must be dropped; so, the average component $\overline{\Gamma}^{\mathsf{b}}$ (or alternatively $\overline{\Gamma}^t$) is not

zero but used to make the integrability condition satisfied. Accordingly, one of the freedoms (the average gauge function $\mathtt{avg}(S_n)$) remains undetermined for the maximal reduction. To determine it, a prescription must be added. For the simplest maximal reduction, it is put to zero.

All these features will suggest to define and decompose the vector $\mathbf{g}_n$ in a different way as in (2.97) and (2.98), by including in this vector only the unknowns that are involved at that order (see Eq. (2.66)), and by distinguishing between the unknowns that will be identified and the ones that will remain parameters (see Eqs. (2.70)-(2.71)).

# Chapter 3

# On an intrinsic approach of the guiding-center anholonomy and gyro-gauge arbitrariness

in collaboration with Michel Vittot
Paper submitted to Phys. Plasmas

*A*bstract: In guiding center theory, the standard gyro-angle coordinate is associated with gyro-gauge dependence, the global existence problem for unit vectors perpendicular to the magnetic field, and the notion of anholonomy, which is the failure of the gyro-angle to return to its original value after being transported around a loop in configuration space. We analyse these three intriguing topics through the lens of a recently proposed, global, gauge-independent gyro-angle. This coordinate is constrained, and therefore necessitates the use of a covariant derivative. It also highlights the intrinsic meaning and physical content of gyro-gauge freedom and anholonomy. There are, in fact, many possible covariant derivatives compatible with the intrinsic gyro-angle, and each possibility corresponds to a different notion of gyro-angle transport. This observation sheds new light on Littlejohn's notion of gyro-angle transport and suggests a new derivation of the recently-discovered global existence condition for unit vectors perpendicular to the magnetic field. We also discuss the relationship between Cartesian position-momentum coordinates and the intrinsic gyro-angle.

## Introduction

Chapters 1 and 10 showed that both aspects of guiding-center theory (averaging reduction, and presence of the magnetic moment) can be addressed while using an alternative global, physical coordinate for the gyro-angle, which is the unit vector of the component of the momentum perpendicular to the magnetic field. No gauge fixing was needed. Then, it was shown in Chapter 2 that this intrinsic coordinate can be used also in the standard procedure for the guiding-center reduction, which provides a Hamiltonian structure for the guiding-center dynamics and a maximal reduction for the guiding-center Lagrangian. All the results of the literature can thus be obtained using this physical but constrained coordinate.

In the light of this approach, we now come back to the questions about the usual coordinate. This can clarify whether these questions are related to intrinsic properties of the physics and mathematics of the guiding-center system or to artificial, physically meaningless constructions. In the former case, it is interesting to study how these properties can be observed in the gauge-independent formulation, i.e. what the counterparts of the traditional questions are. Indeed, the guiding-center derivations in Chapters 1, 2, and 10 showed that, in the gauge-independent approach, all the features of the standard approach seemed to be present, including a kind of generalized gauge vector. This makes it necessary to clarify whether the gauge-independent coordinate actually resolves the questions or only transfers them into other questions.

An additional advantage of the present study is to make clear the essential differences between

the intrinsic gyro-angle and the traditional gyro-angle, which opens the way to variations of the intrinsic approach. For instance, this study could indicate how to eliminate the presence of a constrained coordinate while working with only intrinsically defined quantities, or how to wisely choose a scalar gyro-angle coordinate.

Thus, the purpose of this paper is twofold: first, the intrinsic counterparts of the guiding-center anholonomy and gauge arbitrariness are identified, and, second, we investigate whether it is possible to remove the intricacies of the intrinsic framework induced by its constrained coordinate system.

The present chapter is organized as follows. In Sec. 3.1, we review the standard and the gauge-independent gyro-angles. This will emphasize that, when using the latter variable, the questions associated with the gyro-angle have disappeared from the coordinate system. Especially, the main difficulty, the non-global existence, will not have a counterpart in the intrinsic approach.

In Sec. 3.2, we turn our attention to the intrinsic counterpart of the gauge freedom. It is present in the intrinsic approach because the gauge-independent gyro-angle is treated as a constrained coordinate, which is not completely independent of the spatial position: when the spatial position is changed, the gyro-angle has to be changed in order to remain perpendicular to the magnetic field. This is similar to what happens to tangent vectors on a curved space, e.g. in general relativity. Thus, gradients are actually covariant derivatives. Their definition involves a free term, called the *connection* of the covariant derivative, which embodies the choice of a parallel transport for the gyro-angle. In the principal-circle-bundle picture [25, 81], it corresponds to the connection 1-form.

In Sec. 3.3, we investigate the anholonomy question. In the intrinsic framework, the anholonomy will result from the curvature of the coordinate space encoded in covariant derivatives, or more precisely in commutators between them. In the circle-bundle picture, it corresponds to the curvature 2-form.

Secs. 3.1-3.3 are devoted to our first goal, namely the remnants of the three questions about the intrinsic coordinate system, while Secs. 3.4-3.5 are devoted to our second goal. These last two sections aim at simplifying the gauge-independent formalism by removing the presence of anholonomy and covariant derivatives along with their associated intricacies.

In Sec. 3.4, we consider using the freedom embodied in the connection in order to remove the anholonomy by making covariant derivatives commute. This will provide an interesting approach to the existence condition for a scalar gyro-angle coordinate.

Lastly, in Sec. 3.5, we consider the flaws associated with the presence of a constrained coordinate. We show how they can be eliminated by avoiding the splitting between the gyro-angle and the pitch-angle in the coordinate system.

## 3.1 A global gauge-independent coordinate for the gyro-angle

The physical system under consideration is a charged particle with position $\mathbf{q}$, momentum $\mathbf{p}$, mass $m$, and charge $e$, under the influence of an electromagnetic magnetic field $(\mathbf{E}, \mathbf{B})$. The motion is given by the Lorentz force

$$\begin{aligned} \dot{\mathbf{q}} &= \tfrac{\mathbf{p}}{m}\,, \\ \dot{\mathbf{p}} &= \tfrac{\mathbf{p}}{m} \times e\mathbf{B} + e\mathbf{E}\,. \end{aligned}$$

When the magnetic field is strong, the motion implies a separation of time scales. This is best seen by choosing convenient coordinates for the momentum space, for instance as in Chapter 1

$$\begin{aligned} p &:= \|\mathbf{p}\|\,, \\ \varphi &:= \arccos\left(\tfrac{\mathbf{p}\cdot\mathsf{b}}{\|\mathbf{p}\|}\right)\,, \\ \mathsf{c} &:= \tfrac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|}\,, \end{aligned} \tag{3.1}$$

where $\mathsf{b} := \frac{\mathbf{B}}{\|\mathbf{B}\|}$ is the unit vector of the magnetic field, and $\mathbf{p}_\perp := \mathbf{p} - (\mathbf{p}\cdot\mathsf{b})\mathsf{b}$ is the so-called *perpendicular momentum*, i.e. the orthogonal projection of the momentum onto the plane perpendicular to the magnetic field. The coordinate $p$ is the norm of the momentum. The coordinate $\varphi$

is the so-called *pitch-angle*, i.e. the angle between the momentum and the magnetic field. The last variable $\mathsf{c}$ is the unit vector of the perpendicular momentum.

Then, the equations of motion write

$$\begin{aligned}
\dot{\mathbf{q}} &= \tfrac{\mathbf{p}}{m}\,,\\
\dot{p} &= \tfrac{e\mathbf{E}\cdot\mathbf{p}}{p}\,,\\
\dot{\varphi} &= -\tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{c} + \tfrac{e\mathbf{E}}{p\sin\varphi}\cdot\left(\cos\varphi\ \tfrac{\mathbf{p}}{p} - \mathsf{b}\right)\,,\\
\dot{\mathsf{c}} &= -\tfrac{eB}{m}\mathsf{a} - \tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot(\mathsf{c}\mathsf{b} + \mathsf{a}\mathsf{a}\cot\varphi) + \tfrac{e\mathbf{E}\cdot\mathsf{a}}{p\ \sin\varphi}\mathsf{a}\,,
\end{aligned} \tag{3.2}$$

where $\mathbf{p}$ is now a shorthand for $p(\mathsf{b}\cos\varphi+\mathsf{c}\sin\varphi)$, $B$ is the norm of the magnetic field and, following Littlejohn's notations [86,88], the vector $\mathsf{a} := \mathsf{b}\times\mathsf{c}$ is the unit vector of the Larmor radius, so that $(\mathsf{a},\mathsf{b},\mathsf{c})$ is a right-handed orthonormal frame (rotating with the momentum).

In the case of a strong magnetic field, the only fast term is the Larmor frequency $\omega_L := \frac{eB}{m}$. It corresponds to the gyration of the particle momentum around the magnetic field. It concerns only one coordinate, namely $\mathsf{c}$, the direction of the perpendicular momentum $\mathbf{p}_\perp$ in the 2-dimensional plane perpendicular to the magnetic field. This coordinate corresponds to the gyro-angle.

To get a scalar angle instead of the vector $\mathsf{c}$, one chooses at each point $\mathbf{q}$ in space a direction $\mathsf{e}_1(\mathbf{q}) \in \mathbf{B}^\perp(\mathbf{q})$ in the plane perpendicular to the magnetic field, which will be considered as the reference axis. Then, the angle $\theta$ is defined as the oriented angle between the chosen reference axis $\mathsf{e}_1(\mathbf{q})$ and the vector $\mathsf{c}$ through the following relation :

$$\mathsf{c} = -\mathsf{e}_1\sin\theta - \mathsf{e}_2\cos\theta\,, \tag{3.3}$$

with $\mathsf{e}_2 := \mathsf{b}\times\mathsf{e}_1$ the unit vector such that $(\mathsf{b},\mathsf{e}_1,\mathsf{e}_2)$ is a (fixed) right-handed orthonormal frame [86]. The angle $\theta$ is the usual coordinate for the gyro-angle [8,29,86,122]. Its equation of motion is

$$\dot{\theta} = \tfrac{eB}{m} + \cot\varphi\tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{b}\cdot\mathsf{a} + \tfrac{\mathbf{p}}{m}\cdot\nabla\mathsf{e}_1\cdot\mathsf{e}_2 - \tfrac{e\mathbf{E}\cdot\mathsf{a}}{p\ \sin\varphi}\,. \tag{3.4}$$

From the initial dynamics $(\dot{\mathbf{q}},\dot{p},\dot{\varphi},\dot{\theta})$, guiding-center reductions perform a change of coordinates $(\mathbf{q},p,\varphi,\theta) \longrightarrow (\bar{\mathbf{q}},\bar{p},\bar{\varphi},\bar{\theta})$ in order to obtain a reduced dynamics with suitable properties, mainly a constant of motion $\dot{\bar{p}} = 0$ and a slow reduced motion $(\dot{\bar{\mathbf{q}}},\dot{\bar{\varphi}})$ that is both independent of the fast coordinate $\bar{\theta}$ and Hamiltonian [29,60]. The reduced position $\bar{\mathbf{q}}$ is the *guiding-center*. The constant of motion $\bar{p}$ is the *magnetic moment*. It is close to the well-known adiabatic invariant $\mu := \frac{\mathbf{p}_\perp^2}{2mB}$, and is usually written $\bar{\mu}$ instead of $\bar{p}$.

In the definition of $\theta$, the necessary introduction of $\mathsf{e}_1(\mathbf{q})$ implies important and awkward features in the theory.

First, the choice of $\mathsf{e}_1(\mathbf{q})$ is arbitrary, which induces a local gauge in the theory. The coordinate system is gauge dependent since the value of $\theta$ depends of the chosen $\mathsf{e}_1(\mathbf{q})$. For a general reduction procedure, the guiding-center dynamics can end up being gauge dependent. For instance, the maximal reduction by Lie-transforming the phase-space Lagrangian is gauge dependent in Chapter 2.

Guiding-center reductions have to use prescriptions in order to avoid such unphysical results. For instance, in the reduced Lagrangian [29, 88], the 1-form $\mathrm{d}\theta$ must appear only through the quantity $d\theta - (d\mathbf{q}\cdot\nabla\mathsf{e}_1)\cdot\mathsf{e}_2$. These gauge-dependence questions emphasize that the gyro-angle is artificial, it is not given by the physics and its meaning is restricted.

Second and more substantial, a continuous choice of $\mathsf{e}_1(\mathbf{q})$ does not exist globally in a general magnetic geometry [25, 146]. The reason is that the possible values for $\mathsf{e}_1(\mathbf{q})$ define a principal circle bundle over the configuration space [25, 81]. A specific choice $\mathsf{e}_1(\mathbf{q})$ is a global section of the bundle. It corresponds to a trivialization of the bundle, whose coordinate system is precisely $(\mathbf{q},\theta)$. But such a trivialization does not exist globally for a general circle bundle. In the case of the guiding-center, this global non-existence can be proven by using the theory of principal bundles and characteristic classes [25]. Thus, the gyro-angle does not exist in the whole physical system in general. It does not capture the mathematical description of the system, only a strongly simplified description that is valid only in the trivial case.

It was mentioned in Ref. [25] that the local descriptions are consistent with a global description provided the change of local descriptions satisfy some relations. For instance, the 1-form $d\theta$ must appear only through the combination $d\theta - (d\mathbf{q}\cdot\nabla\mathsf{e}_1)\cdot\mathsf{e}_2$. So, working with $\theta$ is not meaningless. For instance, it can provide a slow guiding-center dynamics $(\dot{\bar{\mathbf{q}}}, \dot{\bar{\varphi}})$ that is globally defined. Nevertheless, this does not provide the local coordinate $\theta$ with a global meaning, neither does it explain what the global description of the gyro-angle is.

Last, even the local description is not completely regular, because it involves a non-holonomic phase in the gyro-angle. When a loop $\gamma$ is performed in position space (while keeping the momentum coordinates $(p, \varphi, \theta)$ constant), at the end of the process all the coordinates and all the physical quantities have recovered their value, but the variation of the gyro-angle involves a non-zero partial contribution [86, 90, 145]. This last is related to the third term in the right-hand side of Eq. (3.4) and writes

$$\Delta\theta_g := \oint_\gamma (d\mathbf{q}\cdot\nabla\mathsf{e}_1)\cdot\mathsf{e}_2\,. \tag{3.5}$$

It is called the geometric phase. This is the well-known anholonomy associated with to the gyro-angle coordinate.

A similarity with Berry's phase and more generally with Hannay's phase was often pointed out [86, 90, 91], but there are significant differences, as mentioned in Ref. [90]. Especially, these phases are related to adiabaticity with a single path in parameter space followed by the system, which makes them physically determinable. On the contrary, guiding-center anholonomy is related to path dependence in configuration space, with all paths coexisting simultaneously. This precludes any definite value for this phase. It raises questions whether this phase is physically meaningful or if it affects only non-physical quantities concerning the extrinsic coordinate system.

All these questions come because $\theta$ is only an artificial quantity. They motivated to keep the primitive gauge-independent coordinate $\mathsf{c}$ instead of introducing $\theta$. Even if this quantity does not have scalar values, it embodies an angle since it is a unit vector in a plane and hence belongs to a circle $\mathbb{S}^1$.

The vector $\mathsf{c}$ represents the physical quantity corresponding to the gyro-angle; $\theta$ never appears alone in the theory (e.g. in guiding-center transformations), except in its own definition and subsequent relations. What appears everywhere is the vector $\mathsf{c}$ (see e.g. Refs. [29, 86, 122]). Even for the correction to the gyro-angle $(\bar{\theta} - \theta)$ in guiding-center reductions, $\theta$ never appears alone but only as the argument of the vector $\mathsf{c}$. An example can be found in the equations of motion (3.2) and (3.4), which also illustrate the gauge dependence or independence.

In addition, the coordinate $\mathsf{c}$ is globally defined, since the perpendicular momentum is well defined everywhere in the guiding-center system. It is useful to remind that the points where the momentum is parallel to the magnetic field are always implicitly excluded from guiding-center theory, even in the local description using the scalar coordinate $\theta$. Indeed, at those points, the angle $\theta$ cannot be defined. In addition, guiding-center expansions involves some $\sin\varphi$ in denominators[1], which implies to exclude the mentioned points.

The variations of $\mathsf{c}$ do not include the gauge contribution $(d\mathbf{q}\cdot\nabla\mathsf{e}_1)\cdot\mathsf{e}_2$, with its anholonomic geometric phase. In addition, no extrinsic effect can be observed in the coordinate system. So, one can expect the questions about the anholonomic geometrical phase to disappear from the theory. The underlying idea is basically true, as will be confirmed by the next sections, but some subtleties should be taken into account. As is emphasized in Eq. (3.5), anholonomy does not affect the coordinate system itself, but rather some contributions to the gyro-angle after one loop in phase space (or in configuration space with a given parallel transport for the gyro-angle). The analysis in the next sections will show that the coordinate system is not enough to define the corresponding contributions, but that they can indeed be defined with a zero geometric phase. All the same, anholonomy will not disappear from the system. It will be induced by the structure of

[1] For instance, at order 1, Eq. (30c) in Ref. [88] shows that $\bar{\theta}$ is singular in $p\sin\varphi = 0$, as is also confirmed by Ref. [29]. In addition, at order 2, Eq. (23) in Chapter 1 shows that $\bar{\varphi}$ is also singular in $p\sin\varphi = 0$. At higher orders, this singularity is expected to affect all the coordinates, because it is related to the order in the cotangent of the pitch-angle compared to the order in the Larmor radius, as is emphasized in Chapters 1, 2 and 10.

the coordinate system, and especially by the magnetic geometry. Accordingly, the coordinate $\mathsf{c}$ does not remove anholonomy but it does remove the anholonomy question: with the coordinate $\theta$ the puzzling point was not anholonomy in itself, but rather the fact that in principle anholonomy is absent when the coordinate system is trivial.

Last, the coordinate $\mathsf{c}$ agrees with the mathematical description of the system. For any magnetic geometry, it induces a circle bundle [25, 81, 90]. Indeed, the circle for $\mathsf{c}$ is position-dependent since $\mathsf{c}$ is perpendicular to the magnetic field. So, $\mathsf{c}$ is not just in $\mathbb{S}^1$ but in $\mathbb{S}^1(\mathbf{q})$. A few consequences will be studied in the next sections. In the traditional coordinate system, this picture is absent because $\theta$ is independent of the position. The circle bundle rather concerns the vector $\mathsf{e}_1$, but the corresponding bundle is different from the intrinsic bundle for the gyro-angle. It is not defined by the whole phase space but confined to the four-dimensional space $(\mathbf{q}, \mathsf{e}_1)$. In addition, the global section $\mathsf{e}_1$ assumes the topology of the bundle trivial.

Accordingly, the use of the gyro-angle $\mathsf{c}$ removes from the coordinate system all of the questions involved in the standard gyro-angle $\theta$. It indeed provides the intrinsic description of the physics and mathematics of the system.

## 3.2 Intrinsic counterpart of the gauge arbitrariness

The previous section showed that in the intrinsic framework, the gauge arbitrariness disappears from the coordinate system. In this section, we show that some arbitrariness remains in the theory and that it corresponds to the intrinsic counterpart of the gyro-gauge arbitrariness

The origin is that, when the coordinate $\mathsf{c}$ is used, the spatial dependence of $\mathbb{S}^1$ causes the coordinate space to be constrained, i.e. the coordinates are not completely independent of each other. When the position $\mathbf{q}$ is changed, the coordinate $\mathsf{c}$ cannot be kept unchanged, otherwise it may get out of $\mathsf{b}^\perp$:

$$\nabla\mathsf{c} \neq 0\,. \tag{3.6}$$

Differentiating Eq. (3.1) with respect to $\mathbf{q}$, one finds (as in Chapter 1, but this formula is reminded here for consistency of the current chapter)

$$\partial_{\mathbf{q}|\mathbf{p}}\mathsf{c} = -\nabla\mathsf{b}\cdot(\mathsf{c}\mathsf{b} + \mathsf{a}\mathsf{a}\cot\varphi)\,. \tag{3.7}$$

The right-hand side is well defined everywhere, since the points where $\cot\varphi = \pm\infty$, i.e. where $\mathbf{p}$ is parallel to the magnetic field $\mathbf{B}$, are excluded from the theory.

Eq. (3.7) must not be given a completely intrinsic meaning, because the two terms in its right-hand side play a very different role through coordinate change: the first term is always unchanged, whereas the second one is generally changed. For instance, if one uses the scalar angle $\theta$ as a local coordinate for $\mathsf{c}$, Eq. (3.7) becomes

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot\mathsf{c}\;\mathsf{b} + \mathbf{R}\;\mathsf{a}\,, \tag{3.8}$$

where

$$\mathbf{R} := \nabla\mathsf{e}_1\cdot\mathsf{e}_2 \tag{3.9}$$

is the so-called *gauge vector*. $\mathbf{R}$ is a function of the position $\mathbf{q}$, and it is not unique: it depends of the choice of gauge $\mathsf{e}_1(\mathbf{q})$.

The reason for this difference of role is that the definition space $\mathbb{S}^1(\mathbf{q})$ for the coordinate $\mathsf{c}$ exactly imposes the first term in Eq. (3.7) but gives no constraints on the second term. Indeed, the gyro-angle $\mathsf{c}$ is a free 1-dimensional coordinate, but it is at the same time a vector immersed in $\mathbb{R}^3$. The two remaining dimensions are fixed by the condition for $\mathsf{c}$ to have unit norm $\mathsf{c}\cdot\mathsf{c} = 1$ and to be transverse to the magnetic field $\mathsf{c}\cdot\mathsf{b} = 0$. This implies

$$\nabla\mathsf{c}\cdot\mathsf{c} = 0 \quad\text{and}\quad \nabla\mathsf{c}\cdot\mathsf{b} = -\nabla\mathsf{b}\cdot\mathsf{c}\,. \tag{3.10}$$

Thus, in $\nabla\mathsf{c}$, only the component parallel to $\mathsf{a}$ is not imposed by intrinsic properties linked with $\mathbb{S}^1(\mathbf{q})$. Practically, it is induced by the specific definition chosen for the gyro-angle coordinate.

From an intrinsic point of view, it is completely free:

$$\nabla \mathsf{c} = -\nabla \mathsf{b}\cdot\mathsf{c}\ \mathsf{b} + \mathbf{R}_g\ \mathsf{a}\,, \tag{3.11}$$

where

$$\mathbf{R}_g := \nabla \mathsf{c}\cdot\mathsf{a} \tag{3.12}$$

is a free phase-space function.

The geometric picture of this freedom is the following. In the gradient $\nabla\mathsf{c}$, i.e. in the effects of an infinitesimal spatial displacement on $\mathsf{c}$, one of the terms, $-\nabla\mathsf{b}\cdot\mathsf{c}\ \mathsf{b}$, is mandatory: it is necessary and sufficient for $\mathsf{c}$ to remain inside its definition domain in the process of spatial transportation. This is easily seen on a diagram. The other term $\nabla\mathsf{b}\cdot\mathsf{a}\ \mathsf{a}$ is only optional. It corresponds to a rotation of $\mathsf{c}$ around $\mathsf{b}$ (hence a gyration around the circle) accompanying the spatial displacement. However, it could be removed or given a different value. It is not imposed by intrinsic properties and has to be arbitrarily chosen, in a similar way as when the target set of a projection is determined, but not the kernel. It encodes the way points in a circle $\mathbb{S}^1(\mathbf{q}_1)$ at the position $\mathbf{q}_1$ are "projected" (more precisely connected) to points in the circle $\mathbb{S}^1(\mathbf{q}_1 + \delta\mathbf{q}_1)$ at a neighbouring position.

This phenomenon is exactly the same as what occurs to tangent vectors on a curved space, for instance on the surface of the Earth or in general relativity. When the spatial position is moved, the tangent vectors from the initial point have to be moved into the tangent space of the final point. Otherwise they are not tangent vectors any more. So, the gradient is actually a covariant derivative, that is an infinitesimal operator that not only changes the position, but also all other quantities in such a way that they remain inside their definition domain throughout the displacement. This phenomenon has to appear as soon as the coordinate space is constrained, i.e. when the definition domain of some coordinate (or parameter) depends on other coordinates.

As a result, when using the coordinate $\mathsf{c}$, the generator of spatial displacements is not a standard gradient operator, such as the gradient with the gauge-dependent gyro-angle coordinate or the one with the Cartesian position-momentum coordinate. In order to emphasize the difference, the covariant-derivative gradient will be denoted with an over-bar $\bar{\nabla}$. When acting on functions depending only on $\mathbf{q}$, the operators $\nabla$ and $\bar{\nabla}$ are the same, and the over-bar will not be used.

In the definition of the covariant derivative, there is always some freedom, since the definition domain at the initial point can be connected to the definition domain at the final point in an arbitrary way. The free term in the covariant derivative corresponds to its *connection*. It determines the way objects are parallel transported. In the example of tangent vectors on a curved surface, the so-called affine connection corresponds to the well-known Christoffel symbols. For the gyro-angle $\mathsf{c}$, the connection freedom is embodied in $\mathbf{R}_g$.

More generally, the connection freedom might not affect only gradients $\nabla$. Other derivative operators could also become covariant derivatives impacting the gyro-angle, i.e.

$$\partial_p \mathsf{c} = f_1\ \mathsf{a} \quad \text{and} \quad \partial_\varphi \mathsf{c} = f_2\ \mathsf{a}\,, \tag{3.13}$$

with $f_1$ and $f_2$ arbitrary functions of the phase space. However, this additional refinement will not be used here, because it does not seem to be useful nor natural: the definition space for the gyro-angle $\mathsf{c}$ depends only on the magnetic field, and hence of the spatial coordinates.

All these features agree with the circle-bundle picture and emphasize its relevance, which was hidden and even suppressed from the coordinate system when using the gauge-dependent gyro-angle. Indeed, instead of viewing the phase space as a constrained coordinate system $(\mathbf{q}, p, \varphi, \mathsf{c})$ as we did above, one can equivalently consider it as a principal $\mathbb{S}^1$-bundle with a base space given by $(\mathbf{q}, p, \varphi)$ and a fiber $\mathbb{S}^1$. This setting offers rigorous mathematical structures. Then, the covariant derivative is defined in terms of the parallel transport associated to a chosen connection 1-form, i.e. a 1-form on the fiber bundle that evaluates to 1 on the canonical infinitesimal generator of gyro-rotation (see e.g. Ref. [99] for an introduction to the subject). In the present paper, we prefer to avoid focusing on this more technical viewpoint. Its terminology and concepts could seem unfamiliar to some readers, and they are not indispensable to investigate the guiding-center system, which can be considered in the more basic picture of a constrained coordinate system.

In either formalism, the conclusion is that, although the gyro-gauge with its arbitrariness is absent from the intrinsic approach, some arbitrariness is present, not in the coordinate system itself but in the choice of a connection for the covariant derivative.

The comparison with the gauge-dependent approach emphasizes the relationship between the two arbitrarinesses. Eq. (3.8) shows that the connection is then embodied in the gauge vector $\mathbf{R}$. This quantity is called gauge vector [86, 90] because it exactly embodies the information on the local gauge. Changing the gauge $\mathsf{e}_1(\mathbf{q}) \longrightarrow \mathsf{e}_1'(\mathbf{q})$ just means adding to the value of $\theta$ a scalar function $\psi(\mathbf{q})$, whose value is the angle between $\mathsf{e}_1$ and $\mathsf{e}_1'$ at $\mathbf{q}$. Under this change, the gauge vector is changed by

$$\mathbf{R}' = \mathbf{R} + \nabla\psi\,. \tag{3.14}$$

The term $\nabla\psi$ exactly embodies the purely local gauge. The remaining part of the gauge corresponds to a global gauge, i.e. the value of $\psi$ at a reference point $\mathbf{q}_1$. Changing it corresponds to a global rotation by the same angle $\psi(\mathbf{q}_1)$ for all points $\mathbf{q}$. It is not specifically a (local) gauge, but rather a common change of angular coordinate.

As a side comment, which will be useful for the following, Eq. (3.14) shows that, even if the gauge vector is gauge dependent, its curl is not. The curl [86]

$$\nabla \times \mathbf{R} = \mathbf{N}\,, \tag{3.15}$$

$$\text{with}\quad \mathbf{N} := \tfrac{\mathsf{b}}{2}\big(\mathrm{Tr}(\nabla\mathsf{b}\cdot\nabla\mathsf{b}) - (\nabla\cdot\mathsf{b})^2\big) + (\nabla\cdot\mathsf{b})\mathsf{b}\cdot\nabla\mathsf{b} - \mathsf{b}\cdot\nabla\mathsf{b}\cdot\nabla\mathsf{b} \tag{3.16}$$

is related to intrinsic properties of the system. Accordingly, the gauge vector cannot be a free function of the position space, but it must satisfy the integrability condition (3.15). Especially, the choice $\mathbf{R} = 0$ is not available in general, even locally.

With the above interpretation for $\mathbf{R}$, Eq. (3.8) means that the local gauge freedom is exactly the counterpart of the connection freedom $\mathbf{R}_g$. The geometric reason is that $\theta$ is the angle between $\mathsf{e}_1$ and $\mathsf{c}$. Through spatial transportation $\delta\mathbf{q}$ with $\theta$ unchanged, the gauge vector $\mathbf{R}$ means that $\mathsf{e}_1$ is rotated around $\mathsf{b}$ by an angle $\delta\mathbf{q}\cdot\mathbf{R}$. It implies that $\mathsf{c}$ undergoes the same rotation. Thus, the connection for $\mathsf{e}_1$ is at the same time a connection for $\mathsf{c}$.

This discussion shows that the gyro-gauge freedom originates from an inherent property of the guiding-center system. It also clarifies the precise meaning and physical content of this intrinsic arbitrariness in two ways. First, it indicates that the arbitrariness is not associated with a choice of coordinate system, but rather a choice of covariant derivative. Second, it implies not only a restricted class of functions of the configuration space but a free function of the phase space.

The first point clarifies how in the intrinsic formulation the arbitrariness is present but does not affect the coordinates at all, whereas in the standard formulation it affects also the coordinate $\theta$. A first illustration is provided by the guiding-center transformation. For the gyro-angle $\mathsf{c}$, the transformation is connection-independent, as was clear in Chapters 1 and 2. For the coordinate $\theta$, the transformation $\bar\theta = \cdot\cdot e^{\mathbf{G}_2} e^{\mathbf{G}_1}\theta$ is gauge dependent, but in such a way as to make the induced transformation $\bar{\mathsf{c}} = \cdot\cdot e^{\mathbf{G}_2} e^{\mathbf{G}_1}\mathsf{c}$ for $\mathsf{c}$ gauge independent, with $\mathbf{G}_n$ the vector field generating the $n$-th order transformation. For instance, the first-order reduced gyro-angle $\bar\theta$ is given by $\mathbf{G}_1^\theta$, which is gauge dependent. As for the first-order $\bar{\mathsf{c}}$, it is given by $(\mathbf{G}_1^{\mathbf{q}}\cdot\nabla + \mathbf{G}_1^\theta\ \partial_\theta)\mathsf{c}$, which is gauge-independent [29, 86, 88].

A second illustration is given by the guiding-center Poisson bracket. When using the coordinate $\theta$, its expression is gauge dependent [86, 88] because of the presence of the gauge vector $\mathbf{R}$. Nevertheless, it was noticed [17] that this presence can be combined with the gradient through the combination $\nabla + \mathbf{R}\ \partial_\theta$. Actually, this is exactly the connection-independent gradient for the coordinate $\mathsf{c}$. Indeed, formally the covariant derivative $\bar\nabla$ can be written $\partial_{\mathbf{q}|\mathsf{c}} + \bar\nabla\mathsf{c}\cdot\partial_{\mathsf{c}|\mathbf{q}}$, which is connection-dependent, but in such a way as to make the quantity $\bar\nabla - \bar\nabla\mathsf{c}\cdot\partial_{\mathsf{c}|\mathbf{q}} = \bar\nabla + \nabla\mathsf{b}\cdot\mathsf{c}\ \mathsf{b}\cdot\partial_{\mathsf{c}} - \mathbf{R}_g\mathsf{a}\cdot\partial_{\mathsf{c}}$ connection-independent. This last quantity is not a gradient because the second term in the right-hand side brings $\mathsf{c}$ out of its definition space. Thus, this term has to be removed to obtain the connection-independent gradient $\bar\nabla_* := \bar\nabla + \mathbf{R}_g\ \partial_\theta$. It is the minimal connection, since it formally

writes $\bar{\nabla}_* = \partial_{\mathbf{q}|\mathsf{c}} - \nabla\mathsf{b}\cdot\mathsf{c}\ \mathsf{b}\cdot\partial_{\mathsf{c}|\mathbf{q}}$, which corresponds to $\bar{\nabla}_*\mathsf{c} = -\nabla\mathsf{b}\cdot\mathsf{c}\ \mathsf{b}$, or equivalently $(\mathbf{R}_g)_* = 0$. This minimal connection corresponds to the orthogonal projection, in the picture mentioned above where the connection is viewed as a projection from the circle $\mathbb{S}^1(\mathbf{q})$ to the circle $\mathbb{S}^1(\mathbf{q}+\delta\mathbf{q})$.

Also the content of the true intrinsic arbitrariness can have practical interest. For instance, the physical connection (3.7) is not available when using the coordinate $\theta$. It does not correspond to a gauge fixing (since it depends on the momentum variables). Neither does the connection $\mathbf{R} = 0$, which would simplify computations and remove the arbitrary terms from the theory. As a result, there is no natural gauge fixing in this case. It is why previous works tried to identify a natural gauge fixing based on the magnetic geometry [90,145]. On the contrary, with the gauge-independent coordinate $\mathsf{c}$, the physical choice $\mathbf{R}_g = -\cot\varphi\nabla\mathsf{b}\cdot\mathsf{a}$ does correspond to a connection fixing. As for the simplifying choice $\mathbf{R}_g = 0$, it is just the minimal connection. It is also the natural geometrical connection, in the sense that it is induced by the definition of $\mathbb{S}^1(\mathbf{q})$ and that it corresponds to the connection-independent gradient, as appeared in the previous paragraph.

Thus, the intrinsic approach is not just an optional reformulation of the theory. It emphasizes the intrinsic properties underlying the gauge arbitrariness, which concern the connection of the covariant derivative $\bar{\nabla}$ rather than the coordinate system. It also makes available some relevant gradients that were inaccessible in the gauge-dependent formulation.

## 3.3 Intrinsic counterpart of the anholonomy

We now turn to the remnants of the traditional guiding-center anholonomy in the intrinsic approach. Anholonomy means that, after performing one closed path in some space, some partial contribution does not sum up to zero or some quantity defined in another space with a given parallel transport does not recover its initial value. The guiding-center anholonomy initially concerns the gauge $\mathsf{e}_1$ and can be considered from two complementary points of view [86,90].

In the first point of view, one performs a closed path $\gamma$ in configuration space with $\mathsf{e}_1$ parallel transported along the connection $\mathbf{R}$. At the end of the loop, the vector $\mathsf{e}_1$ recovers its value but the sum of its infinitesimal rotations around the magnetic field corresponds to a non-zero angle [86,90]

$$\Delta\theta_g := \oint_\gamma (d\mathbf{q}\cdot\nabla\mathsf{e}_1)\cdot\mathsf{b}\times\mathsf{e}_1 = \int_S \nabla\times\mathbf{R}\cdot d\mathbf{S} \neq 0\,, \tag{3.17}$$

where $S$ is a surface with boundary $\partial S = \gamma$. In the second point of view, the loop $\gamma$ is performed while preventing $\mathsf{e}_1$ to rotate around the magnetic field (i.e. it is moved in such a way that $\delta\mathbf{q}\cdot\nabla\mathsf{e}_1\cdot\mathsf{e}_2 = 0$). Then at the end of the process, the vector $\mathsf{e}_1$ does not recover its value: it has rotated by a non-zero angle, which is given by $-\Delta\theta_g$.

Notice that Eq. (3.17) assumes that the loop $\gamma$ is contractible. In this paper, we will always consider that it is the case (e.g. the space is simply connected). Otherwise $S$ does not exist and there is no relation equivalent to the second equality in (3.17). This simply-connected assumption is enough to study the local structure of the system, since locally in a three-dimensional domain any loop is contractible. In order to study also global aspects, the assumption must be released. For instance in a tokamak not all loops are contractible. This should not cause a problem because the results of the next section will agree with the results of Ref. [25], where the assumption on contractible loops is not used. Thus this assumption plays no essential role. It is useful only to simplify the argument.

The non-zero angle (3.17) impacts the coordinate $\theta$, whose variations do not depend only on the state of the particle but also on the gauge fixing $\mathsf{e}_1$. The partial contribution due to the gyro-gauge is called the *geometric phase*, denoted by $\Delta\theta_g$. It is anholonomic because of Eq. (3.17). This anholonomy term cannot be made zero by a choice of gauge, because its integrand is given by Eq. (3.16), which is gauge-independent. This fact suggests that the anholonomic quantity (3.17) is related to an intrinsic property of the system. It is why the anholonomy was considered as unavoidable in guiding-center coordinates[2] [86,90].

[2] In principle, anholonomy in a principal fiber bundle is not related to a coordinate system, but induced by a choice of connection. However, in the standard description of the guiding-center system, the fiber bundle is induced

With the gyro-angle $\mathsf{c}$, the anholonomy does not concern the coordinate system, since the coordinates are defined directly from the physical state. There is no extrinsic quantity (such as $\mathsf{e}_1$) implied in the definition of this gyro-angle to generate anholonomy.

However, this does not preclude the possible presence of anholonomy in the intrinsic framework. To investigate this point, we will first identify the intrinsic counterparts of the quantities involved in the traditional guiding-center anholonomy, mainly the total variation of the gyro-phase $d\theta$ and the geometric phase $\Delta\theta_g$. Then the question of anholonomy in the gauge-independent framework will become clear.

To begin with, the gauge-independent gyro-phase $\mathsf{c}$ is not a scalar angle. It is a vector, and its infinitesimal variations are two-dimensional, as shown in Eq. (3.11). For an intrinsic description of the phenomenon at work in Eq. (3.17), it is convenient to identify a scalar quantity for the variation of the gyro-angle $\mathsf{c}$.

The change of $\mathsf{c}$ in the direction $\mathsf{b}$ is not relevant, since it just corresponds to maintaining $\mathsf{c}$ in its definition domain through spatial displacement. Thus, the effective variation of the gyro-angle is only in the direction of $\mathsf{a}$. In addition, the true change of $\mathsf{c}$ is obtained after removing the contribution coming from the spatial displacement. From this point of view,

$$\delta\Theta := -\mathsf{a}\cdot(d\mathsf{c} - d\mathbf{q}\cdot\bar{\nabla}\mathsf{c}) = -\mathsf{a}\cdot d\mathsf{c} + d\mathbf{q}\cdot\mathbf{R}_g \tag{3.18}$$

is the quantity measuring the Larmor gyration. The minus sign in the prefactor is a convention in order to agree with the usual orientation for the gyro-angle $\theta$.

The relevance of the scalar variation $\delta\Theta$ is emphasized by the fact that the set of 1-forms

$$\Big(d\mathbf{q}, dp, d\varphi, \delta\Theta\Big) \tag{3.19}$$

is dual to the natural derivative operators of the theory

$$\Big(\bar{\nabla}, \partial_p, \partial_\varphi, \partial_\theta := -\mathsf{a}\cdot\partial_{\mathsf{c}}\Big)\,. \tag{3.20}$$

Here we insist that the generator of Larmor gyration is written $\partial_\theta$ but its definition does not depend on the gauge.

In addition, when the local gauge-dependent description for the gyro-angle is used (implicitly chosen as usual such that inside the local description there is a globally defined $\mathsf{e}_1$ and $\mathbf{R}_g$ is defined by $\mathbf{R}$), it is readily checked that $\delta\Theta = d\theta$. This observation confirms that $\delta\Theta$ is the intrinsic (global) quantity corresponding to $d\theta$. As a corollary, it explains why the 1-form $d\theta$ is gauge dependent and the associated gauge-independent 1-form is $d\theta - d\mathbf{q}\cdot\mathbf{R}$. Indeed, the reason is that $\delta\Theta$ depends on $\mathbf{R}_g$ and the associated connection-independent quantity is $-\mathsf{a}\cdot d\mathsf{c} = \delta\Theta - d\mathbf{q}\cdot\mathbf{R}_g$. An essential difference compared to $d\theta$ is that $\delta\Theta$ is not closed:

$$d(\delta\Theta) = -(d\mathbf{q}\cdot\nabla\mathsf{b})\cdot(\mathsf{b}\times\mathsf{b}'d\mathbf{q}) + d\mathbf{R}_g\cdot\wedge\ d\mathbf{q}\,, \tag{3.21}$$

where for notational convenience, the primed notation is used for gradients acting on their left: $\mathsf{b}'d\mathbf{q} = d\mathbf{q}\cdot\nabla\mathsf{b}$. The wedge symbol $\wedge$ indicates antisymmetry: $a.\wedge\ b = a.b - b.a$.

In Eq. (3.21), the magnetic term can be rewritten by using Eq. (3.29), in agreement with Littlejohn's results [90]. However, Littlejohn derived the curvature of $\delta\Theta$ in the special case where $\mathbf{R}_g = 0$. He did not consider a more general connection, nor did he indicate that there was any freedom in selecting the connection. The interest of Eq. (3.21) is that it applies to any connection $\mathbf{R}_g$.

---

by $(\mathbf{q}, \mathsf{e}_1)$, and the choice of a connection on this bundle is implied in the choice of a gyro-gauge, which in turn corresponds to the choice of a coordinate $\theta$. It is why there is a direct correspondence between the anholonomy and the standard gyro-angle coordinate system, although the anholonomy does not concern $\theta$ itself, but the vector $\mathsf{e}_1$ through the parallel transport defined by the connection.

Interestingly, when using the intrinsic gyro-angle $\mathsf{c}$, this correspondence disappears. As it should, the anholonomy is not related to the coordinate system any more, but only induced by the connection or the covariant derivative.

For completeness, let us mention that, with the most general connection for the gyro-angle $\mathsf{c}$, Eq. (3.18) is written as

$$\delta\Theta = -\mathsf{a}\cdot d\mathsf{c} + d\mathbf{q}\cdot\mathbf{R}_g + f_1 dp + f_2 d\varphi\,, \tag{3.22}$$

where $f_1$ and $f_2$ are arbitrary phase-space functions. Then Eq. (3.21) would become

$$d(\delta\Theta) = -(d\mathbf{q}\cdot\nabla\mathsf{b})\cdot(\mathsf{b}\times\mathsf{b}'d\mathbf{q}) + d\mathbf{R}_g\cdot\wedge\ d\mathbf{q} + df_1\wedge dp + df_2\wedge d\varphi\,. \tag{3.23}$$

Non-zero $f_1$ and $f_2$ would imply that the connection and the covariant derivative concern not only spatial displacements (i.e. variations of the coordinate $\mathbf{q}$), but also variations of the two other non-gyro-angle coordinates $p$ and $\varphi$, as in Eq. (3.13). Since this refinement is useless here, for the following we will set $f_1 = f_2 = 0$ and remain with only the connection $\mathbf{R}_g$.

The variation of $\Theta$ is defined along a path $\gamma$ by $\Delta\Theta = \int_\gamma \delta\Theta$. After performing one closed path that is the boundary of some surface $S$, $\Delta\Theta$ is given by:

$$\Delta\Theta = \oint_\gamma \delta\Theta = \int_S d(\delta\Theta) \neq 0\,, \tag{3.24}$$

which is non-zero in general. Thus, $\Theta$ is not a holonomic quantity.

More precisely, using Eq. (3.21), the anholonomy (3.24) of the scalar angle can be written

$$\Delta\Theta = \Delta\Theta_B + \Delta\Theta_c\,. \tag{3.25}$$

The first contribution comes from the magnetic geometry

$$\Delta\Theta_B = -\int_S (d\mathbf{q}\cdot\nabla\mathsf{b})\cdot(\mathsf{b}\times\mathsf{b}'d\mathbf{q})\,. \tag{3.26}$$

The second contribution comes from the choice of connection

$$\Delta\Theta_c = \int_S d\mathbf{R}_g\cdot\wedge\ d\mathbf{q}\,, \tag{3.27}$$

and it is expected to be the intrinsic counterpart of the geometric phase.

When the local gauge-dependent coordinate $\theta$ is used for the gyro-angle, the connection vector $\mathbf{R}_g = \mathbf{R} = \nabla\mathsf{e}_1\cdot\mathsf{e}_2$ depends only on $\mathbf{q}$. Then the integrand of the contribution (3.27) exactly compensates the magnetic contribution (3.26). Indeed, it writes

$$\begin{aligned} d\mathbf{R}\cdot\wedge\ d\mathbf{q} &= (d\mathbf{q}\cdot\nabla\mathsf{e}_2)\cdot\wedge\ (d\mathbf{q}\cdot\nabla\mathsf{e}_1) \\ &= (d\mathbf{q}\cdot\nabla\mathsf{b})\cdot(\mathsf{e}_2\mathsf{e}_1 - \mathsf{e}_1\mathsf{e}_2)\cdot(\mathsf{b}'d\mathbf{q}) \\ &= (d\mathbf{q}\cdot\nabla\mathsf{b})\cdot\mathsf{b}\times(\mathsf{b}'d\mathbf{q})\,, \end{aligned} \tag{3.28}$$

where the first equality comes from the antisymmetry, and the second comes by inserting the identity matrix $(\mathsf{b}\mathsf{b} + \mathsf{e}_1\mathsf{e}_1 + \mathsf{e}_2\mathsf{e}_2)\cdot$ and by using the relations

$$\begin{aligned} (d\mathbf{q}\cdot\nabla\mathsf{e}_1)\cdot\mathsf{e}_1 &= 0\,, \\ (d\mathbf{q}\cdot\nabla\mathsf{e}_2)\cdot\mathsf{e}_2 &= 0\,, \\ (d\mathbf{q}\cdot\nabla\mathsf{e}_1)\cdot\mathsf{b} &= -(d\mathbf{q}\cdot\nabla\mathsf{b})\cdot\mathsf{e}_2\,, \end{aligned}$$

which come because $(\mathsf{b}, \mathsf{e}_1, \mathsf{e}_2)$ is an orthonormal basis.

For a comparison with the anholonomic phase (3.17), the integrand of the connection contribution (3.27) can be written by using that it is only a function of the position:

$$\begin{aligned} d\mathbf{R}\cdot\wedge\ d\mathbf{q} &= d\mathbf{q}\cdot(\nabla\mathbf{R} - \mathbf{R}')\ d\mathbf{q} \\ &= -d\mathbf{q}\cdot(\nabla\times\mathbf{R})\times d\mathbf{q}\,, \end{aligned}$$

which agrees with (3.17).

These results can be shown to agree with Littlejohn's expression (3.16) by using the antisymmetry of the matrix $\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'$ in order to write it as a cross product:

$$\begin{aligned}(d\mathbf{q}\cdot\nabla\mathsf{b})\cdot(\mathsf{b}\times\mathsf{b}'d\mathbf{q}) &= d\mathbf{q}^i\nabla^i\mathsf{b}\cdot\mathsf{b}\times\nabla^j\mathsf{b}\ d\mathbf{q}^j \\ &= d\mathbf{q}^i d\mathbf{q}^j\tfrac{1}{2}(\delta^{ik}\delta^{jl}-\delta^{il}\delta^{jk})\nabla_k\mathsf{b}\cdot\mathsf{b}\times\nabla_l\mathsf{b} \\ &= d\mathbf{q}^i d\mathbf{q}^j\tfrac{1}{2}\varepsilon_{ijA}\varepsilon_{Akl}\nabla_k\mathsf{b}\cdot\mathsf{b}\times\nabla_l\mathsf{b} \\ &= \varepsilon_{iAj}d\mathbf{q}^i\left(-\tfrac{1}{2}\right)\varepsilon_{Akl}\nabla_k\mathsf{b}\cdot\mathsf{b}\times\nabla_l\mathsf{b}\ d\mathbf{q}^j \\ &= d\mathbf{q}\cdot\mathbf{N}\times d\mathbf{q}\,,\end{aligned} \tag{3.29}$$

with

$$\begin{aligned}\mathbf{N}_A &:= \left(-\tfrac{1}{2}\right)\varepsilon_{Akl}\varepsilon_{\alpha\beta\gamma}\nabla_k\mathsf{b}_\alpha\mathsf{b}_\beta\nabla_l\mathsf{b}_\gamma \\ &= \tfrac{\mathsf{b}_A}{2}\left(\mathrm{Tr}(\nabla\mathsf{b}\cdot\nabla\mathsf{b})-(\nabla\cdot\mathsf{b})^2\right) \\ &\quad + (\nabla\cdot\mathsf{b})\ (\mathsf{b}\cdot\nabla\mathsf{b})_A-(\mathsf{b}\cdot\nabla\mathsf{b}\cdot\nabla\mathsf{b})_A\,.\end{aligned}$$

This is exactly Littlejohn's expression (3.16). In the computation (3.29), the first equality comes from the antisymmetry of the matrix $\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'$, and all the other equalities are properties of the Levi-Civitá symbol. An alternative (more direct but heavier) way to prove the result is to insert the identity matrix $(\mathsf{b}\mathsf{b}+\mathsf{e}_1\mathsf{e}_1+\mathsf{e}_2\mathsf{e}_2)\cdot$ everywhere in $(d\mathbf{q}\cdot\nabla\mathsf{b})\cdot(\mathsf{b}\times\mathsf{b}'d\mathbf{q})$. Then expanding and simplifying the formula gives the expected result.

So, in the local gauge-dependent case, the connection contribution is exactly the geometric phase

$$\Delta\Theta_c = \Delta\theta_g\,, \tag{3.30}$$

which confirms that $\Delta\Theta_c$ is the intrinsic counterpart of the geometric phase. In addition, Eq. (3.28) shows that, when the coordinate $\theta$ is used, the gauge contribution is exactly the inverse of the anholonomic magnetic contribution (3.26):

$$\Delta\Theta_c = -\Delta\Theta_B\,. \tag{3.31}$$

This equality explains both that the corresponding gyro-angle $\theta$ is holonomic $d(\delta\Theta)=d^2\theta=0$, and that the associated geometric phase $\Delta\theta_g$ is anholonomic.

A first consequence can be noticed by now: in the intrinsic approach, the counterpart of the geometric phase is arbitrary and it can be made holonomic. Even its integrand can be set to zero, since $\mathbf{R}_g$ can be chosen freely. So, the anholonomic geometric phase observed in the gauge-dependent approach is not intrinsic in itself. On another hand, choosing the geometric phase zero makes the total phase $\Delta\Theta$ anholonomic, since in that case it is exactly given by the anholonomic term (3.26) due to the magnetic geometry. This suggests that in the intrinsic framework anholonomy is not only present but inherent to the structure of the system.

To investigate this point, let us remind that the introduction of a scalar variation for $\mathsf{c}$ was used only to identify the correspondence between the local and global descriptions. From an intrinsic point of view, what we have obtained is just that $\delta\Theta$ is not closed. Hence $\Theta$ is not a proper coordinate, but it is not needed since the gyro-angle coordinate is the vectorial quantity $\mathsf{c}$. All the same, anholonomy effects can be viewed even in this framework from two complementary (dual) points of view.

The first of them considers the properties of the basic 1-forms (3.19) of the theory. The previous investigations showed that the basic differential form for the gyro-angle is $\delta\Theta$, and that it is not closed. This implies anholonomy, as appeared in Eq. (3.24) and as can be found in textbooks, e.g. in Ref. [97]. In the fiber-bundle approach, this conclusion is still clearer, since the anholonomy is given by the curvature 2-form[3], here $d(\delta\Theta)$.

[3] More precisely, the anholonomy is measured by the variation of the angle after a closed loop $\gamma$ in the base space while parallel transported along the connection. This variation is computed by integrating the connection (here $\delta\Theta$) along $\gamma$. When the loop can be used to define a two-dimensional surface $S$ with boundary $\partial S=\gamma$, as is the case when the base space is simply connected (i.e. when it has no non-contractible loops), then the result can be expressed as the integral of the curvature 2-form on $S$.

The second point of view more basically considers the properties of the elementary vector fields of the theory. Indeed, the effects of an infinitesimal loop in configuration space (with the parallel transport defined by the connection) are evaluated with the commutator of gradients, as is confirmed in Ref. [99] or [84] for instance. Here, the commutator is

$$[\bar{\nabla}_i, \bar{\nabla}_j] = \Big( \nabla_i \mathsf{b}\cdot\mathsf{b} \times \nabla_j \mathsf{b} - \bar{\nabla}_i (\mathbf{R}_g)_j + \bar{\nabla}_j (\mathbf{R}_g)_i \Big) \, \partial_\theta \,, \tag{3.32}$$

which is dual to Eq. (3.21).

The non-commutation of gradients in Eq. (3.32) is a consequence of the presence of a constrained coordinate system, with its associated non-zero connection. It means that the action of gradients does not fit directly with the coordinate system. After a closed path in the sense of the gradients, i.e. of the sum of infinitesimal variations, the coordinates do not recover their initial value. Conversely, after one loop in coordinate space, the coordinates recover their initial value, but the sum of infinitesimal changes is not zero. Thus, not only is anholonomy inherent to the introduction of a scalar angle, which generalizes the conclusion of Refs. [86, 90], but it is an intrinsic feature of the space of particle states $(\mathbf{q}, p, \varphi, \mathsf{c})$. This feature, although absent from trivial coordinate systems, is not an issue. It is quite common in spaces with non-zero curvature, e.g. in general relativity.

The issue comes with the gauge-dependent approach when requiring a coordinate system that fits with the action of gradients. This is possible only when the geometry of the bundle is trivial. In addition, in the resulting trivialized space the unavoidable presence of non-trivial anholonomy effects become puzzling. So, the scalar coordinate $\theta$ for the gyro-angle has holonomic (commuting) gradients but all the same involves anholonomic (unphysical) phases. In addition, it is valid only locally because it loses (makes trivial) the geometry of the coordinate space. On the contrary, the global gyro-angle $\mathsf{c}$ has anholonomic gradients rather than anholonomic phases, but it retains all the geometry of the guiding-center coordinate system.

This necessary alternative comes from Eq. (3.21) or (3.32), which shows that anholonomy is unavoidable and intrinsically related to the magnetic geometry. The anholonomic term $\nabla\mathsf{b}\cdot\mathsf{b} \times \mathsf{b}'$ has to be put either as a non-zero commutator of gradients or as an anholonomic phase for the gyro-angle.

## 3.4 Towards a scalar intrinsic gyro-angle

The previous three sections were concerned with the first goal of this paper. They showed how the intrinsic approach clarifies the questions caused by the presence of a gyro-gauge. They also emphasized the true intrinsic properties that were underlying in the traditional guiding-center anholonomy and gauge arbitrariness. We now turn to the second goal of this paper, which is to investigate how the intricacies caused by the presence of anholonomy and of covariant derivatives can be eliminated from the intrinsic approach. The former is studied in this section, while the latter will be considered in the next section.

So, we are presently interested in removing the anholonomy effects observed in the previous section. The basic idea is to use the freedom embodied in the connection in order to make the commutator of gradients zero. This question comes very timely since one point is to be clarified about our previous results. When the connection is given its physical expression (3.7), the gyro-angle $\mathsf{c}$ is just the perpendicular velocity. It should be holonomic (i.e. it should not have non-zero commutators of gradients), since it is directly given by the physical momentum and the magnetic field, both of which are holonomic.

The reason for this anholonomy is that the connection was defined through the physical definition of $\mathsf{c}$, but not the physical definition of the whole momentum. In order to take into account the whole momentum, a more general connection should be used, affecting also the pitch-angle $\varphi$.

The variable $\varphi$ is not a constrained coordinate, since it is defined over an independent space $\mathbb{R}^1$. Unlike the variable $\mathsf{c}$, it does not have to change value through spatial transportation, but it is allowed to. A flat (zero) connection is possible but it is only the trivial choice, analogous to the choice $\mathbf{R}_g = 0$ for the coordinate $\mathsf{c}$. In the same way as the free term $\mathbf{R}_g$, the coordinate $\varphi$

can have an arbitrary connection. Especially, its definition from the physical momentum through Eq. (3.1) induces a non-zero connection

$$\partial_{\mathbf{q}|\mathbf{p}}\varphi = -\nabla \mathsf{b}\cdot\mathsf{c}\,. \tag{3.33}$$

Notice that in this argument two different gradients are implied: the one in the initial coordinates $\nabla = \partial_{\mathbf{q}|\mathbf{p}}$ and the one in the final coordinates $\bar{\nabla}$. This last is roughly $\partial_{\mathbf{q}|p,\varphi,\mathsf{c}}$ but it takes into account the necessary connection for $\mathsf{c}$ and the possible connection for $\varphi$. When acting on functions of $\mathbf{q}$ only, e.g. in the right-hand side of Eq. (3.33), they are equal, but in general they are not. What we call the *physical connection* is the one that makes them equal. For instance, the action of the associated covariant derivative on the quantity $\varphi$ is defined by

$$\bar{\nabla}\varphi = \partial_{\mathbf{q}|\mathbf{p}}\varphi\,. \tag{3.34}$$

Notice also that within the constrained-coordinate picture the presence of a non-zero connection for $\varphi$ causes no complication. In the fiber-bundle picture, it would imply to change the framework, because the coordinate $\varphi$ would have to be considered in the fiber, not in the base space.

The physical relevance of this connection can be viewed in the components $\mathbf{V}_i$ of the velocity vector field. They are defined by the relation $\dot{f} = \mathbf{V}_i\cdot\partial_i f$ for any function $f$ of the phase space. Because of the non-trivial connection, they are different from the components of the velocity $\dot{\mathbf{z}}_i$, where the vector $\mathbf{z} = (\mathbf{q}, p, \varphi, \mathsf{c})$ combines all the coordinates. The relation between $\dot{\mathbf{z}}_i$ and $V_i$ is given by

$$\dot{\mathbf{z}}_i = \tfrac{d}{dt}\mathbf{z}_i = \sum_j \mathbf{V}_j\cdot\partial_j\mathbf{z}_i = \mathbf{V}_i + \mathbf{K}_i\,, \tag{3.35}$$

where $\mathbf{K}_i := \sum_j (1-\delta_{ij})\mathbf{V}_j\cdot\partial_j\mathbf{z}_i$ is a connection term for it does not contribute when $\partial_j\mathbf{z}_i = 0$ for $i \neq j$. In the specific case we are considering, the connection is involved only when a gradient acts on the coordinates $\mathsf{c}$ or $\varphi$. So, Eq. (3.35) can be interpreted as

$$\tfrac{d}{dt}\mathsf{c} = (\partial_t + \dot{\mathbf{q}}\cdot\bar{\nabla})\mathsf{c}\,, \tag{3.36}$$

together with the same formula with $\mathsf{c}$ replaced by $\varphi$.

With the physical connections (3.33) and (3.7) for the pitch angle and the gyro-angle, the components of the velocity vector field are given by

$$\mathbf{V}_i := \begin{pmatrix} \frac{\mathbf{p}}{m} \\ \frac{e\mathbf{E}\cdot\mathbf{p}}{p} \\ \frac{e\mathbf{E}}{p\sin\varphi}\cdot\left(\cos\varphi\ \frac{\mathbf{p}}{p} - \mathsf{b}\right) \\ -\frac{eB}{m}\mathsf{a} + \frac{e\mathbf{E}\cdot\mathsf{a}}{p\ \sin\varphi}\mathsf{a} \end{pmatrix}.$$

They perfectly agree with the physical force, which is just the Lorentz force. Especially, the limit where there is no electric field $\mathbf{E} = 0$ is expressive: there remain only the velocity $\mathbf{V}_\mathbf{q} = \mathbf{p}/m$ and the Larmor gyration $\mathbf{V}_\mathsf{c} = -eB\mathsf{a}/m$. All the additional terms in the components $\varphi$ and $\mathsf{c}$ of equations (3.2), which do not come from the physical dynamics but from the magnetic geometry, are absorbed in the connection. This is satisfactory since the role of the connection is precisely to encode the change of the momentum coordinates through spatial displacement as a result of the magnetic geometry.

When using the coordinate $\theta$, the geometric contributions in Eq. (3.4) cannot be absorbed in a connection contribution. In fact, the scalar coordinate $\theta$ is precisely introduced to make the coordinate system trivial, and hence to have flat connection. Providing $\theta$ with the corresponding connection would amount to using the intrinsic approach, with an additional detour by the gauge $\mathsf{e}_1$.

The dynamics of the gyro-angle can be reinterpreted in this light, in relation with Refs. [22,91]. In the dynamics (3.4) of the gyro-angle $\theta$, only the first and last terms are contributions due to the physical dynamics. The second term corresponds to the so-called "adiabatic phase" in the case considered by Ref. [91]. It comes from the magnetic term in the physical connection (3.7),

related to the definition for $\mathsf{c}$ to be physically the unit vector of the perpendicular momentum. It is induced by the change of the projection as a result of the change of the magnetic field (through spatial displacement). Thus it concerns also the intrinsic gyro-angle $\mathsf{c}$, e.g. in Eq. (3.2) or Eq. (3.7). In addition, it is expected to be adiabatic only for the specific case considered by Ref. [91], but not for a general (inhomogeneous) strong magnetic field. This is confirmed by Ref. [22]. As for the "geometric phase", i.e. the third term in Eq. (3.4), it is actually a gauge phase since it is purely related to the choice of gauge. It is absent from the intrinsic dynamics (3.2) or connection (3.7).

With the full physical connection given by (3.7) and (3.33), although the coordinates $\mathsf{c}$ and $\varphi$ are not independent of the variable $\mathbf{q}$, they behave exactly as the components of a vector $\hat{\mathbf{v}} := \mathsf{b}\cos\varphi + \mathsf{c}\sin\varphi$ that is independent of $\mathbf{q}$, i.e. that has flat connection:

$$\begin{aligned}\bar{\nabla}\hat{\mathbf{v}} &= \bar{\nabla}(\cos\varphi\mathsf{b} + \sin\varphi\mathsf{c}) \\ &= \cos\varphi\nabla\mathsf{b} + \sin\varphi\ \bar{\nabla}\mathsf{c} + \bar{\nabla}\varphi\ (-\mathsf{b}\sin\varphi + \mathsf{c}\cos\varphi) = 0\,.\end{aligned}$$

Actually, this computation shows that the flat connection $\bar{\nabla}\hat{\mathbf{v}} = 0$ is obtained if and only if the connection is the full physical one. The vector $\hat{\mathbf{v}}$ stands for the unit vector of the momentum

$$\hat{\mathbf{v}} := \tfrac{\mathbf{p}}{p}\,. \tag{3.37}$$

As a consequence and as expected from physical intuition, the commutator of gradients with this connection is zero:

$$[\bar{\nabla}_i, \bar{\nabla}_j] = 0\,, \tag{3.38}$$

as is easily verified by direct computation. It traduces that, after one loop in configuration space (with $\mathbf{p}$ constant), both the momentum and the magnetic field come back to their initial value.

The physical connection achieves a part of our goal by making the commutators of gradients zero, but other commutators have to be considered as well. Indeed, in the gauge-independent approach the space defined by the bundle is not the space of all $(\mathbf{q}, \mathsf{c})$ but the whole phase space $(\mathbf{q}, p, \varphi, \mathsf{c})$. So, all the above conclusions apply after replacing the gradients $\bar{\nabla}$ by the complete set of basic derivative operators $\partial_{\mathbf{z}} = (\bar{\nabla}, \partial_p, \partial_\varphi, \partial_\theta)$. Between these operators, even with the physical connection, the non-triviality of the fiber bundle for a general magnetic geometry should imply non-zero commutators. This is confirmed in Eq. (3.39), where non-trivial commutators of the basic derivative operators are given for a general connection, both for the gyro-angle $\mathbf{R}_g := \bar{\nabla}\mathsf{c}\cdot\mathsf{a}$ and for the pitch-angle $\mathbf{R}_\varphi := \bar{\nabla}\varphi$. This includes all of the four choices of connection previously mentioned as special cases: the connection (3.8) for the gauge-dependent case with coordinate $\theta$; the physical connection (3.7) for $\mathsf{c}$; the general connection (3.11) for $\mathsf{c}$; and the full physical connection (3.7) and (3.33) for $\mathsf{c}$ and $\varphi$.

$$\begin{aligned}[\bar{\nabla}_i, \bar{\nabla}_j] = &\ \nabla_i\mathsf{b}\cdot\mathsf{b}\times\nabla_j\mathsf{b}\ \partial_\theta \\ &- \Big(\bar{\nabla}_i(\mathbf{R}_g)_j - \bar{\nabla}_j(\mathbf{R}_g)_i\Big)\ \partial_\theta \\ &+ \Big(\bar{\nabla}_i(\mathbf{R}_\varphi)_j - \bar{\nabla}_j(\mathbf{R}_\varphi)_i\Big)\ \partial_\varphi\,, \\ [\partial_p, \bar{\nabla}] = &\ -\partial_p\mathbf{R}_g\ \partial_\theta + \partial_p\mathbf{R}_\varphi\ \partial_\varphi\,, \\ [\partial_\varphi, \bar{\nabla}] = &\ -\partial_\varphi\mathbf{R}_g\ \partial_\theta + \partial_\varphi\mathbf{R}_\varphi\ \partial_\varphi\,, \\ [\partial_\theta, \bar{\nabla}] = &\ -\partial_\theta\mathbf{R}_g\ \partial_\theta + \partial_\theta\mathbf{R}_\varphi\ \partial_\varphi\,.\end{aligned} \tag{3.39}$$

In practical cases the second commutator in Eq. (3.39) is zero because $\mathbf{R}_g$ does not depend on $p$. The reason is that the gyro-angle comes from the splitting of the coordinate $\hat{\mathbf{v}}$ into the pitch-angle and the gyro-angle *via* the magnetic geometry $\mathbf{B}(\mathbf{q})$. The coordinate $p$ plays no role in the process.

Eq. (3.39) clearly emphasises the crucial role of the anholonomic magnetic term $\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b}'$, which is the only affine term in the connection. For the minimal connection this magnetic term is the only non-zero term, which indeed simplifies computations. As for the full physical connection, it cancels the magnetic term, and also the whole commutator of gradients. However, the two commutators

$[\partial_\varphi, \bar{\nabla}]$ and $[\partial_\theta, \bar{\nabla}]$ become non-zero. For instance $\partial_\varphi \mathbf{R}_g \neq 0$ and $\partial_\theta \mathbf{R}_\varphi \neq 0$.

With the general setting considered in Eq. (3.39), one can look for a connection that would make all commutators zero. This would provide a splitting of the vector $\hat{\mathbf{v}}$ into proper scalar coordinates for the pitch-angle and the gyro-angle, i.e. coordinates that fit with the action of commuting derivative operators. These coordinates would be defined from the value of the quantities $\varphi$ and $\mathsf{c}$ at one point in phase space through parallel transportation by the commuting derivative operators[4].

A solution is expected not to be generally possible since a scalar coordinate for the gyro-angle means that the circle bundle is trivial. The goal is to identify existence condition for the desired coordinate system. Indeed, the free 4-dimensional connection function of phase space $\big(\mathbf{R}_g(\mathbf{z}), \mathbf{R}_\varphi(\mathbf{z})\big)$ opens new possibilities. One can consider using this larger freedom to obtain more complete results than with the gauge-dependent framework, whose freedom corresponds only to the 1-dimensional gauge function of position space $\psi(\mathbf{q})$ in Eq. (3.14).

The last three rows in Eq. (3.39) imply that a solution $(\mathbf{R}_g, \mathbf{R}_\varphi)$ must not depend on $\varphi$, nor $p$, nor $\mathsf{c}$, hence it must be purely position-dependent. In addition, because of the third row in Eq. (3.39), $\mathbf{R}_\varphi$ must be curl-free. Actually, $\mathbf{R}_\varphi$ is useless and can be set to zero. As for the first two rows, they imply that $\mathbf{R}_g$ must cancel the anholonomy term, which is purely position-dependent. Thus, the equation for the desired connection is

$$0 = \nabla_i \mathsf{b} \cdot \mathsf{b} \times \nabla_j \mathsf{b} - \bar{\nabla}_i (\mathbf{R}_g)_j + \bar{\nabla}_j (\mathbf{R}_g)_i \,, \tag{3.40}$$

which can be rewritten

$$\bar{\nabla} \times \mathbf{R}_g = \mathbf{N} \,. \tag{3.41}$$

This equation is reminiscent of the usual relation (3.15) in the gauge-dependent approach, but they are different both in their origin and in their meaning. On the one hand, Eq. (3.41) is obtained without appealing to the idea of a gyro-gauge nor its associated gyro-angle $\theta$. Starting from the structure of the manifold defined by the physical coordinates $(\mathbf{q}, p, \varphi, \mathsf{c})$, we are looking for an arbitrary scalar gyro-angle coordinate. The solution might not be related to a choice of gauge. More precisely we are looking for a connection corresponding to this coordinate. On the other hand, Eq. (3.41) is a necessary and sufficient condition for the existence of a scalar gyro-angle, whereas in previous works, the analogous relation (3.15) only appeared as a consequence of the existence of a gauge.

When a solution $\mathbf{R}_g := \mathbf{R}_s$ of Eq. (3.41) exists, the scalar gyro-angle coordinate is defined from parallel transportation with the derivative operators defined by the associated connection $\mathbf{R}_s$. This parallel transportation results in a (trivializing) global section of the circle bundle, which in turn provides a zero for measuring the gyro-angle. Now, the important point in the above analysis is that $\mathbf{R}_s$ depends only on the position. The parallel transportation actually results in a section of the restricted circle bundle over the position space. This property means that a scalar coordinate always defines a gyro-gauge. It is the reciprocal of the property that a gyro-gauge provides a scalar gyro-angle, which was considered in previous works and in the previous sections.

As a consequence, condition (3.41) is also a necessary and sufficient condition for a global gauge to exist. Here, it is obtained in a direct argument on commuting derivatives (but under the assumption of contractible loops). This is very different from the work [25] where the existence of a global gauge was studied: the proof for the necessary condition used an auxiliary property and the proof for sufficiency required "a lengthy digression into the theory of principal bundles and characteristic classes". Finally, their condition for the existence of a gauge is slightly different from ours, but they are equivalent. Their condition states that through the boundary $S$ of any hole inside the spatial domain the vector field $\mathbf{N}$ has no net flux:

$$\oint_S \mathbf{N} \cdot d\mathbf{S} = 0 \,, \tag{3.42}$$

[4] More precisely, making all derivative operators commute only guaranties local existence for the desired scalar coordinates. For a global existence, obstructions could come from non-contractible loops, which could generate a discontinuity or a multi-valuation for the desired coordinate. However, here the base space is simply connected, so that the local criterion is enough to provide global existence.

and this is the boundary condition for the solvability of Eq. (3.41), since $\nabla\cdot\mathbf{N} = 0$.

Our derivation assumed contractible loops. It does not apply when the base space contains non-contractible loops, e.g. in a tokamak geometry. In this case, the work [25] is needed to conclude about the existence of a gauge. In turn, the existence of a scalar angle is implied by the presence of a gauge. Since the existence condition is the same in any case and regards only non-contractible spheres, the assumption on non-contractible loops plays no essential role and can be released. A complete understanding of the origin of this fact requires a more detailed study, which is outside the scope of the present paper.

## 3.5 An intrinsic formalism with no covariant derivative

We now turn to the last aspect of the second goal of the paper, which is to simplify the formalism by removing from the intrinsic coordinate system the presence of covariant derivatives.

Indeed, Indeed, Secs. 3.1-3.3 showed that the questions of the standard coordinate $\theta$ were related to properties of the basic derivative operators involved in guiding-center theory, which do not fit with a standard coordinate system. In this sense, the philosophy of the gauge-independent approach is to consider separately the coordinate system and the basic derivative operators. It avoids putting in the coordinate system some properties that actually concern the derivative operators.

However, the resulting formalism may seem more complicated than expected, because of the constrained coordinate system, with the associated subtleties about covariant derivatives, non-zero commutators, non-closed basis of 1-forms, etc. One can consider going one step further in the development of this intrinsic approach and removing these subtleties from the coordinate system. In one way or another, they are unavoidable in the theory since they result from properties of the non-trivial circle bundle implied by the fast guiding-center coordinate. But they concern the derivative operators, not the coordinate system, so that one can consider making the coordinate system both gauge-independent and unconstrained. To do so, it is natural to try to identify a scalar gyro-angle coordinate, as in the previous section. In such a coordinate system, covariant derivatives would automatically disappear. However, Sec. 3.4 showed that a global scalar coordinate for the gyro-angle does not exist in general. Thus, the removal of covariant derivatives must be looked for by other means.

Since the constrained coordinate system comes from the gyro-angle coordinate with its $\mathbb{S}^1$ fiber-bundle, a possible idea is again to come back to more primitive coordinates and to avoid the splitting of the momentum into the pitch-angle and the gyro-angle. Then, the unit vector of the momentum $\hat{\mathbf{v}} := \frac{\mathbf{p}}{p}$ is kept as a single two-dimensional coordinate, as was approached by the results of the previous section.

Separating the pitch-angle from the gyro-angle is necessary at the end of the guiding-center reduction, when the gyro-angle is removed from the dynamics to obtain the slow reduced dynamics $(\dot{\bar{\mathbf{q}}}, \dot{\bar{\varphi}})$. But at that point the fiber-bundle and the constrained coordinate $\mathsf{c}$ are also removed. In the course of the reduction process the splitting is not needed. What is needed is only a basis of 1-forms and a basis of derivative operators that fit with the separation of scales: for instance to decompose the transformation of the vector $\hat{\mathbf{v}}$ between the contribution for the fast gyro-angle and the one for the slow pitch-angle; or to decompose the change of spatial coordinate into its components transverse and parallel to the magnetic field.

As a result, the method of using intrinsic coordinates and defining derivative operators adapted to the purpose can be applied. The interesting point is that the definition space for the coordinate $\hat{\mathbf{v}}$ is the sphere $\mathbb{S}^2$, whose immersion in $\mathbb{R}^3$ is independent of the spatial position. So, a trivial connection available, and this minimal connection is also the physical one since the definition (3.37) for $\hat{\mathbf{v}}$ does not depend on the position.

Notice that, since this coordinate is a two-dimensional vector immersed in $\mathbb{R}^3$, its variations are constrained. The operator $\partial_{\hat{\mathbf{v}}}$ and the 1-form $d\hat{\mathbf{v}}$ are purely transverse: $\hat{\mathbf{v}}\cdot\partial_{\hat{\mathbf{v}}} = \hat{\mathbf{v}}\cdot d\hat{\mathbf{v}} = 0$. But the coordinate system is not constrained any more: the coordinates are independent of each other. The basic differential operators $(\partial_{\mathbf{q}}, \partial_p, \partial_{\hat{\mathbf{v}}})$ and 1-forms $(d\mathbf{q}, dp, d\hat{\mathbf{v}})$ behave trivially (i.e. practical calculations are similar to those using standard coordinates).

For the purpose of the guiding-center reduction, the splitting between the pitch-angle and the gyro-angle is implemented in the basis of vector fields and 1-forms: $\partial_{\hat{\mathbf{v}}}$ and $d\hat{\mathbf{v}}$ are decomposed to distinguish their contributions in the azimuthal direction $\mathsf{a}$ (corresponding to the variable $\theta$) and in the elevation direction $\mathsf{a} \times \hat{\mathbf{v}}$ (corresponding to the variable $\varphi$). For instance, the operator $\partial_{\hat{\mathbf{v}}}$ can be decomposed as $\left(\hat{\mathbf{v}} \times \mathsf{b}\cdot\partial_{\hat{\mathbf{v}}}, \frac{(\mathsf{b}\times\hat{\mathbf{v}})\times\hat{\mathbf{v}}}{\sin\varphi}\cdot\partial_{\hat{\mathbf{v}}}\right)$, in order to agree with the traditional operators $(\partial_\theta, \partial_\varphi)$. Alternatively, the second operator can be chosen as $-\sin\varphi\,(\mathsf{b}\times\hat{\mathbf{v}})\times\hat{\mathbf{v}}\cdot\partial_{\hat{\mathbf{v}}} = \partial_\phi$, in order to fit with the variable $\phi := \cot\varphi$ that made formulae polynomials in Chapters 1, 2, and 10. In a simpler way, it can be chosen just $(\mathsf{b}\times\hat{\mathbf{v}})\times\hat{\mathbf{v}}\cdot\partial_{\hat{\mathbf{v}}}$. This arbitrariness in the choice of a basis for vector fields is similar to the connection freedom in previous sections, but it is different. It concerns the splitting of the operator $\partial_{\hat{\mathbf{v}}}$ rather than the definition of the gradient operator.

This splitting procedure is only a generalization of what is commonly done for the position space. The three-dimensional quantity $\mathbf{q}$ is usually kept as a coordinate but the gradient $\nabla$ and the differential form $d\mathbf{q}$ are split into scalar components suited to the derivation, namely their components parallel to $\mathsf{a}$, $\mathsf{b}$, and $\mathsf{c}$.

It is straightforward to verify that guiding-center reductions work as usual within this formalism. In a similar way as what occurred when going from the coordinate $\theta$ to the coordinate $\mathsf{c}$, the introduction of the coordinate $\hat{\mathbf{v}}$ removes all the intricacies from the coordinate system and confines them to the basis of vector fields of the theory. The only difference is that here the basis is not induced by the coordinate system, whose associated derivative operators are now trivial. It is induced by the purposes of the guiding-center reduction, such as the separation of scales. All the subtleties mentioned in the previous sections remain present but they are encoded in the properties of the chosen basis. For instance, when the basis is chosen with $\left(\hat{\mathbf{v}} \times \mathsf{b}\cdot\partial_{\hat{\mathbf{v}}}, \frac{(\mathsf{b}\times\hat{\mathbf{v}})\times\hat{\mathbf{v}}}{\sin\varphi}\cdot\partial_{\hat{\mathbf{v}}}\right)$, the formalism is the same as in the previous section with the "full physical connection" (3.7) and (3.33).

One could consider going one step further and keeping all the momentum coordinates $(p, \hat{\mathbf{v}})$ as a single coordinate $\mathbf{p}$. This is unsure to be relevant, since the coordinate $p$ actually plays no role in the introduction of the gyro-angle, as we mentioned previously.

In addition, keeping the coordinate $\mathbf{p}$ does not seem convenient for practical computations. When the norm of the momentum $p$ is kept in the reduced coordinates, as in Chapter 1, it is unchanged by the guiding-center transformation. Then it is useless to combine it with $\hat{\mathbf{v}}$, which is changed by the transformation. On another hand, most often the coordinate $p$ is replaced by the constant of motion conjugated to the gyro-angle, the magnetic moment $\bar{\mu}$. Then the coordinate $p$ is usually changed to $\mu$ in a preliminary step, in order for the remaining transformation to be near-identity. In this case, the splitting of the coordinate $\mathbf{p}$ into $p$ and $\hat{\mathbf{v}}$ is essential.

Last, separating the variables $p$ and $\hat{\mathbf{v}}$ is interesting for dimensional reasons. It implies that only one of the momentum coordinates is non-dimensionless, which can be very convenient for the derivation of the guiding-center reduction (see Chapter 2, for instance).

These considerations do not mean that using the coordinate $\mathbf{p}$ is to be excluded. For instance, the recent work [26] proposed an algorithm for guiding-center reductions where the coordinate $\mathbf{p}$ was kept in a first near-identity transformation. Then the magnetic moment was identified and it could be adopted as a coordinate in a second transformation, not near-identity.

## Conclusion

The gauge-independent approach of guiding-center theory clarifies the questions associated with the usual gyro-angle coordinate. The use of the physical gyro-angle as a coordinate removes the non-global existence, the gauge dependence and the anholonomy from the coordinate system. The corresponding coordinate system agrees both with the physical description and with the mathematical structure of the system, a non-trivial circle bundle.

This physical gyro-angle is constrained and position-dependent, which implies the presence of a covariant derivative encoding the geometry of the bundle. The induced connection involves a freedom, which is the intrinsic counterpart of the gauge arbitrariness but is much larger and very

different. Especially, this freedom does not affect the coordinate system and embodies a choice in the basis of vector fields of the theory.

Because of the larger freedom, relevant choices become available. For instance, the connection can be chosen so as to fit with the physical definition of the gyro-angle. Alternatively, it can be set to zero. This minimal connection simplifies computations and removes the arbitrary terms from the theory. It was found to be underlying in previous results and can be considered as the natural geometrical connection. Each of these two choices does not correspond to a gauge fixing but to a connection fixing, which shows that the intrinsic formulation is needed for the description to fit with the physics and the mathematics of the guiding-center system. This is also emphasized by the fact that the question about non-global existence not only disappears but also has no counterparts in the intrinsic formulation.

Both the physical and the minimal covariant derivatives have non-zero commutators, which are the counterparts of the anholonomy of the gauge-dependent approach. Again, they do not concern the coordinate system but the basic derivative operators. A third choice of connection was identified, by giving to the pitch-angle the connection induced by its definition from the physical momentum. Then, covariant derivatives $\bar{\nabla}$ do commute, but other non-zero commutators appear in phase space, which traduce the non-triviality of the circle bundle defined by the gyro-angle.

In this framework, existence conditions for a splitting of the momentum into scalar coordinates for the pitch-angle and for the gyro-angle can be studied. The perspective is broader and complementary compared to the study of the existence of a gyro-gauge. The resulting condition is just the invertibility of a curl on a divergenceless vector field. It corresponds to boundary conditions on this vector field, in agreement with previous results. In this paper, the analysis used contractible loops, but this restriction appeared as optional. A detailed investigation of this point could bring additional information on the structure of the guiding-center system.

As a result, the gauge-independent formulation exhibits the intrinsic properties underlying the questions about the standard gyro-angle. The formulation mainly replaces puzzling aspects of the coordinate system by non-trivial, yet regular, properties of the basic derivative operators.

The questions originated from the requirement for the coordinates and the basic derivative operators to behave trivially. This requirement can be obtained only when the circle bundle is trivial, i.e. only locally for a general magnetic geometry. In addition, it generates both the gauge dependence and the anholonomy question in the coordinate system. To agree with the physical system, the basic derivative operators must have non commuting properties. In turn, it means that there do not exist trivializing coordinates and that a constrained coordinate has to be used for the gyro-angle.

An idea underlying the intrinsic approach is that perturbation theory needs adapted derivative operators but not necessarily adapted coordinates. So, the coordinate system can be chosen as it is intrinsically, i.e. as it comes primitively. This idea can be generalized to avoid introducing the gyro-angle coordinate, with the intricacies caused by the constrained coordinate system. The gyro-angle and the pitch-angle can be kept as the single initial coordinate, which is the unit vector of the momentum. Then the $\mathbb{S}^1$-bundle is replaced by an $\mathbb{S}^2$-bundle. The corresponding structure is trivial and the formalism becomes elementary, i.e. with no covariant derivatives (trivial connection) nor any non-zero commutator. The resulting coordinate system is both intrinsic and unconstrained. For guiding-center perturbation theory, an adapted basis of derivative operators is defined, which encodes all the intrinsic properties of the circle bundle associated to the gyro-angle.

About possible future works, let us remind that our analysis about the existence condition for a scalar gyro-angle used contractible loops, whereas the existence condition for a gyro-gauge only concerns the holes (non-contractible spheres) in the magnetic domain. Thus, contractible loops are only optional. A detailed investigation of this point will be interesting, and could bring additional information on the structure of the guiding-center coordinate system.

# Conclusion of the episode

The intrinsic formulation clarified the troubles involved by the standard gyro-angle, showed how the difficulties can be made to disappear completely, and thus enhanced somehow the validity of the theory. It had practical consequences on the derivation and in the theory, but for numerical simulations, it seems it should not have strong impacts.

Indeed using a vector instead of a scalar gyro-angle coordinate would complicate simulations, and it may be preferable to decompose the physical domain in several areas in each of which a global choice of gauge exists. In the case of magnetic fusion, for most magnetic geometries (e.g. no hole in the spatial domain, such as for tokamaks) a global choice of gauge is possible, so that the local description corresponding to the traditional gyro-angle is actually a global one.

It is not the case of all magnetic geometries, and the intrinsic formulation keeps its relevance for general or non-trivial magnetic geometries, where it offers a global description, and hence intrinsically insures the consistency of the local descriptions.

Another possible application can be found in the investigation of the collision operator for gyrokinetics in this intrinsic framework, because replacing the trivial gauge-dependent gyro-angle by a non-trivial intrinsic variable could impact the description of collisions occurring in the gyro-angle dimension. More technical, although not less interesting, questions can also be considered: for instance the quantification of the flux of the vector field $\mathbf{N} := \nabla \times \mathbf{R}$ and its meaning in the intrinsic formulation; or the condition about contractible loops for the existence of a scalar gyro-angle, which should be optional since it is not needed for the gyro-gauge; or the use of the gauge-independent vectorial gyro-angle in the Vlasov-Maxwell field theory, with especially the definition of the measure for functionals.

In the previous chapters, various methods were used, each with its own purpose. Perhaps their role can usefully be reviewed here. In order to identify a constant of motion, Chapter 10 used expansions of differential operators, with the common issue about secular terms. For a minimal averaging reduction, Chapter 1 showed the efficiency of Lie-transforming the equations of motion. In order to clarify questions about the coordinate system, Chapter 3 was interested in more geometric structures, especially with connections, commutators of vector fields or non-closures of differential forms. Last, for a more complete reduction, Chapter 2 used a Lie transform of the phase-space Lagrangian. It is the only chapter that was directly concerned with Hamiltonian methods. We want to point out how its Hamiltonian approach increased the efficiency of guiding-center reductions. Also, it emphasized how Hamiltonian systems encode the dynamics in the structure, and replace the differential framework by an algebraic one, where structures in the dynamics are more easily identified. This suggests that Hamiltonian approaches could also increase the efficiency of the gyrokinetic reduction, which will be the topic of Chapters 6 and 8.

In order to address higher-order reductions, we obtained explicit induction formulae for the perturbative procedure at higher orders, both for the minimal and for the full guiding-center reduction. Although they involve few basic operations, the Leibniz rule applied at each order implies lengthy expressions and the derivation would be more relevantly addressed by introducing computer-assisted symbolic calculus.

Such a method is applied by [26] to a different reduction procedure. They use Cartesian position-velocity coordinates and include the magnetic moment in the coordinates only as a second step. Also, compared to our approach, the symplectic representation is handled in a different way. It can

be interesting to investigate the efficiency and results of our procedure compared to theirs.

Especially, higher-order results offer many possible representations. It will be interesting to explore them further, to identify or confirm how they can be used for a suitable purpose, and possibly to identify optimal choices for some relevant criteria. For instance, as was suggested by [26], various representations could have a difference in the simplicity of the results, or in the time validity of the transformation truncated at some order. here, we rather focused on the idea of a maximal reduction, related to the simplicity of the reduced dynamics, but the time validity would be interesting to explore as well.

Answering such questions would involve numerical simulations for several test representations, as is suggested in [26]. At the same time, optimal choices can be looked for by analytical studies, as is rather pursued in the present document. Both kinds of works should be developed since they are complementary. Notice that the numerical simulations considered here are easy ones, because they concern six-dimensional particle dynamics (or guiding-center dynamics), not gyrokinetic field dynamics.

The role of the electric field was mentioned in several places, but it can be considered with more details. Explicit formulae for non-zero electric field are interesting to compute and study.

In addition, this first episode was developed in a similar way as what is typically done in the literature, with especially a static magnetic and electric fields. A possible extension would be to consider dynamical electrostatic or electromagnetic field (slowly varying in time, for this effect to be only perturbative). This should not pose a problem, since it would not affect the pivotal coefficients of the reduction procedure. It only implies additional terms in the right-hand side of the equations at each order, which complicates the formulae, but does not affect the procedure at all. This extension will be necessary for gyrokinetics as is emphasized in the next episode, and especially in Chapters 5-6.

So, at this point, the intrinsic formulation of guiding-center theory, as well as the higher-order reduction, are well established in their foundations, but much more remains to be done in order to exploit the results. In some way, this episode was only the beginning of many extensions to be explored and developed, all the more as we only mentioned the immediate extensions, but many other ones can be considered, such as including the bounce-angle reduction [16, 29].

Perhaps the most important continuation concerns the conservation of angular momentum, since this important question was a motivation for developing the higher-order reduction [26,127]. Again, this work will probably imply analytical studies, together with numerical analyses, verifications or explorations.

With regard to the numerical studies, they could be somehow complicated, since they imply plasma simulations (gyrokinetic field dynamics). Present-day simulations are already very heavy for High Perfomance Computers [42, 56], and including additional (second-order) terms would significantly complexify the simulation.

As a first exploration, a possible process is to evaluate a posteriori in present-day numerical simulations whether the second-order neglected terms are indeed much smaller (or at least induce smaller effects) than the first-order terms, and also than the second-order terms that are kept. If this is not clearly verified, it can be useful to study the regions in phase space (or in tokamaks) where it is verified or not, in order to identify the possible domains of relevance or limitation of present-day simulations and models. This should not be heavy task, since it only means to post-process the data coming from standard simulations.

As for the analytical studies, they will be best addressed with the Hamiltonian formulation of gyrokinetics, which is the topic of the next episodes of this dissertation.

## Selected bibliography


[8] N. N. Bogoliubov and D. N. Zubarev, "An asymptotic approximation method for a system with rotating phases and its application to the motion of a charged particle in a magnetic field", Physical Research Laboratory, Space Technology Laboratories, Los Angeles, 1960. Original paper: Ukrain. Matem. Zh., **7**, 1 (1955).

[17] A. J. Brizard and T. S. Hahm, "Foundations of nonlinear gyrokinetic theory", Rev. Mod. Phys. **79**, 421 (2007).

[23] A. J. Brizard and N. Tronko, "Equivalent representations of higher-order Hamiltonian guiding-center theory", arXiv:1205.5772 [physics.plasm-ph] (2012).

[25] J. W. Burby and H. Qin, "Gyrosymmetry: Global considerations", Phys. Plasmas **19**, 052106 (2012).

[26] J. W. Burby, J. Squire, and H. Qin, "Automation of the guiding center expansion", PPPL report, http://bp.pppl.gov/pub_report//2013/PPPL-4861.pdf (2013).

[27] J. R. Cary, "Lie transformations and their use in Hamiltonian perturbation theory", Phys. Reports **79**, 129 (1981).

[29] J. R. Cary and A. J. Brizard, "Hamiltonian theory of guiding-center motion", Rev. Mod. Phys. **81**, 693 (2009).

[59] L. de Guillebon and M. Vittot, "A gyro-gauge independent minimal guiding-center reduction by Lie transforming the velocity vector field", accepted in Phys. Plasmas (2013) arXiv:1211.5792.

[60] L. de Guillebon and M. Vittot, "Gyro-gauge independent formulation of the guiding-center reduction to arbitrary order in the Larmor radius", submitted, arXiv:1304.3727 (2013).

[61] L. de Guillebon and M. Vittot, "On an intrinsic formulation of the guiding-center anholonomy and gyro-gauge arbitrariness", submitted, arXiv:1304.4563 (2013).

[79] J. A. Krommes, "Comment on “Guiding-center plasma models in three dimensions“", Phys. Plasmas **16**, 084701 (2009).

[81] M. Kruskal, "Asymptotic theory of Hamiltonian and other systems with all solutions nearly periodic", J. Math. Phys. **3**, 806 (1962).

[86] R. G. Littlejohn, "Hamiltonian formulation of guiding-center motion", Phys. Fluids **24**, 1730 (1981).

[87] R. G. Littlejohn, "Hamiltonian perturbation theory in noncanonical coordinates", J. Math. Phys. **23**, 742 (1982).

[88] R. G. Littlejohn, "Variational principles of guiding-center motion", J. Plasm. Phys. **29**, 111 (1983).

[90] R. G. Littlejohn, "Phase anholonomy in the classical adiabatic motion of charged particles", Phys. Rev. A **38**, 6034 (1988).

[91] J. Liu and H. Qin, "Geometric phase of the gyromotion for charged particles in a time-dependent magnetic", field", Phys. Plasmas **18**, 072505 (2011).

[122] T. G. Northrop, *The adiabatic motion of charged particles*, Wiley, New York, 1963.

[123] T. G. Northrop and J. A. Rome, "Extensions of guiding center motion to higher order", Physics of Fluids. **21**, 384 (1978).

[127] F. I. Parra and P. J. Catto, "Transport of momentum in full f gyrokinetics", Phys. Plasmas **17**, 056106 (2010).

[128] F. I. Parra and I. Calvo, "Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry", Plasma Phys. Control. Fusion **53**, 045001 (2011).

[145] L. E. Sugiyama, "Guiding-center plasma models in three dimensions", Phys. Plasmas **15**, 092112 (2008).

[146] L. E. Sugiyama, "Answer to “Comment on ’Guiding-center plasma models in three dimensions’ “ ", Phys. Plasmas **16**, 084702 (2009).

# Episode II

# From particle guiding-center to Vlasov-Maxwell gyro-center

# Introduction of the episode

After its preliminary stage, with the guiding-center reduction of particle dynamics in an external electromagnetic field, gyrokinetics proceeds in its main stage, which is the reduction of plasma dynamics coupled with Maxwell's equations (or with Poisson equation) [17, 49, 65]. It can be viewed as the transfer (lifting) of guiding-center reduction from particle dynamics to plasma dynamics, together with inclusion of the coupling between the plasma and the electromagnetic field.

Let us remind that it proceeds in four steps.

The first one (denoted as Step 2a on page 16) keeps the magnetic field external (and the electric field zero in principle, as in the first stage). It just consists in transferring ("lifting" [115]) the change of coordinates from the particle level to the field level, by performing the guiding-center change of coordinates for the Vlasov density. This gives the reduction for the Vlasov equation in the case of a strong external magnetic field.

The second step (Step 2b) is to restore self-consistency in the dynamics, i.e. the coupling between the plasma and the electromagnetic field. In the electrostatic case, it consists in "plugging" besides the Vlasov equation the Poisson equation, where the source term is expressed in guiding-center coordinates, which generates an additional polarization term.

In principle, as the electric field was not accounted for in the guiding-center reduction, the presence of a non-zero electric field spoils the reduction of the Vlasov equation, as can be viewed directly in particle dynamics. So, the third step (Step 2c) is to perform a second transformation (gyro-center transformation) on particle dynamics and to lift it to the Vlasov density, in order to restore the reduction by taking into account the presence of an electric field. For the reduced dynamics, this generates additional terms in the reduced Vlasov equation and in the polarization.

Here, the presence of an electric field was already taken into account in the guiding-center reduction, so that Step 2c disappears. More precisely, it is done at the same time as Steps 1 and 2a. For the second stage, only the first two steps 2a and 2b are left: the plasma reduction is obtained by lifting the particle change of coordinates to the Vlasov density and then the coupling between the plasma and the electric field is obtained just by restoring the Poisson equation, with the source term expressed in gyro-center coordinates.

After Step 2c, the characteristics of the reduced Vlasov density occur in subspaces of constant magnetic moment, and their fast gyro-angle dimension does not influence their other dimensions. The last step of the reduction (Step 2d) is then to drop this superfluous dimension, by averaging it out.

This is how the usual derivation of gyrokinetic equations works [17]. The issue is that it proceeds at the level of the equations of motion, and this gives no information on what the Hamiltonian structure becomes in the process.

As a result, the Hamiltonian structure of gyrokinetics is not identified yet, whereas it would have various advantages (see e.g. [38,70,92,99,107,108,159]), for instance in making available Hamiltonian perturbation methods, or energy-Casimir methods for equilibria and their stability, or in making the Hamiltonian structure usable in numerical simulations and schemes. More importantly, it would provide an efficient framework for conservation laws and especially for the question about the conservation of angular momentum [21, 127, 133].

On another hand, in the present situation, one cannot completely exclude that the present-day reduced dynamics may not be Hamiltonian. Identifying the Hamiltonian structure is a way to check

that no fake dissipation has been introduced in the derivation [108], and to validate the equations or to identify the changes to be done in order to make them Hamiltonian, in a similar way as what occurred around 1980 for the guiding-center reduction, for instance with Littlejohn's works [86,88].

More precisely, our goal is twofold: to obtain the Hamiltonian structure of gyrokinetics and to identify how it is induced by the derivation, which is mainly a lifting of a change of particle coordinates to the Vlasov-Maxwell field dynamics. Indeed, we will avoid trying to directly guess the Hamiltonian structure of gyrokinetics, especially because previous attempts showed that it can be difficult to find the answer, and especially to prove the Jacobi identity with an expectedly complicated bracket.

Instead, we will study how the Vlasov-Maxwell Poisson bracket is affected by the reduction process. This will automatically guarantee that the Jacobi identity is verified by the reduced bracket. So, we will follow the standard procedure, but instead of working on the equations of motion, we will work at the level of the Hamiltonian structure.

In this context, the coupling is automatically included in the structure, and more precisely in the Vlasov-Maxwell non-canonical Poisson bracket. So, nothing has to be done for Step 2b, already taken into account by the Hamiltonian framework. Only Steps 2a (the lifting) and 2d (the removal of the gyro-angle) remain to be done. In this episode, we focus on the lifting method (Step 2a), since it is the main ingredient in the Stage 2 of the gyrokinetic reduction. The subtleties involved in a Hamiltonian approach of the averaging reduction (Step 2d) are quite different and will be addressed in the next episode.

In principle, the lifting is expected to correspond to a change of field coordinates, and hence to a chain rule for functional derivatives in the Poisson bracket, which should naturally provide the transformed Hamiltonian structure, associated with the correct equations of motion. However, practically, the "chain rule" is not just a usual chain rule.

First the change of phase-space coordinates is not implemented in functions defined over the phase space, as it should be for a chain rule, but on functionals defined over a set of functions defined over the phase space.

Second, when the spatial point is changed, additional complications and possible obstructions can take place. The definition of the transformed fields depends not only on the initial fields, but also on the transformed field. This is because the point where the field is evaluated is changed as well, and it hence depends on the transformed field. This could imply an inconsistency in the theory, and a possible obstruction to define properly the transformation for fields, even if the transformation for particle coordinates is well defined.

Third, other subtleties come from specificities of the base space of the Vlasov-Maxwell fields: the Vlasov density is defined over the phase space, which undergoes a coordinate transformation both in configuration space and in momentum space, whereas the electromagnetic field is defined over a configuration space which does not undergo any change of coordinates. This implies to deal with several spaces, and to distinguish between the particle configuration space and the electromagnetic configuration space.

Because of all these features, in order to address the Hamiltonian lifting procedure for gyrokinetics, we will proceed with successive deepenings, by studying first a simplified case, and then more and more complete cases, which should clarify successively the above mentioned possible complications.

- First, in Chapter 4, we consider a simplified transformation where the last two difficulties are removed, and especially the possible obstruction. This is done by considering a transformation that does not affect the spatial position, with the example of the magnetic moment reduction (introduced in Chapter 10). This removes the above mentioned second and third difficulties, because the transformation depends only on the electromagnetic field, which is defined over the configuration space. Thus, only the first difficulty is addressed, i.e. we can focus on the lifting mechanism and phenomena.

- Second, we use the resulting lifting mechanism to address the second difficulty (possible

obstruction), while discarding the last difficulty. This is done by considering a system where all the fields are defined over the same space, i.e. by departing from the Vlasov-Maxwell system. For instance, this could apply to a fluid dynamics where magnetic coordinates are adopted for the configuration space. This work is reported in the appendix chapter 11.

- Third, in Chapter 5, we will use the conclusions about the first two difficulties and will address the last difficulty, by coming back to the Vlasov-Maxwell system. This should provide the mechanism and the results of the guiding-center and gyro-center transformations from the point of view of the Hamiltonian structure.

- Last, in Chapter 6, the Hamiltonian structure of Vlasov-Maxwell is investigated, and it is shown to be exactly suited to the two-stage gyrokinetic reduction. More precisely, its second stage (gyro-center reduction) is given by a Hamiltonian perturbation of the first one. This Hamiltonian perturbation is explored and compared to the traditional reduction relying on a lift of a reduction of particle dynamics.

It is important to notice that the work of this episode relates particle Hamiltonian perturbation theory and field Hamiltonian perturbation theory. Indeed, perturbation theory in the context of Hamiltonian dynamics has proven to be unquestionably useful in many contexts, ranging from celestial mechanics (e.g. [142]), to atomic physics (e.g. [11]), to plasma physics (e.g. [81]). The superconvergent expansions of the Kolmogorov-Arnold-Moser theorem (e.g. [1, 78, 92]) and the techniques of adiabatic invariance (e.g. [28,52,68,150]) all are aspects of perturbation theory in the Hamiltonian context.

Although such techniques are well-developed and well-known for finite-dimensional systems, this is not the case for such perturbation theories for partial differential equations. This is particularly true for Hamiltonian systems with noncanonical Poisson brackets of the form of those given in [102,104,107] for plasma systems. A main effect of the present episode is to provide tools for such perturbation theory using the Poisson bracket for Vlasov-Maxwell equations [95,101,104,114,156] in situations with a short time scale introduced by the presence of a strong magnetic field.

Derivations of gyrokinetic theories have proceeded directly from the Vlasov-Maxwell equations of motion as in the nonlinear development of [49], they have been based on Hamiltonian particle orbit perturbation theory that is lifted up to the kinetic level as in the linear development of [89], or they have incorporated both particle orbit and kinetic perturbations to arrive at a nonlinear theory [19]. (See [17,29] for review.) None of these procedures parallels that for finite-dimensional Hamiltonian systems that has historically achieved such great success. It is why none of these theories obtain an infinite-dimensional Hamiltonian form as a consequence of their method of derivation, and at present it is not known if nonlinear gyrokinetics has Hamiltonian form, the form possessed by all of the important systems of plasma physics when dissipative terms are neglected. An alternative approach was recently introduced in [144], one which uses a kind of kinetic action principle (see, e.g., [158]), rather than the Hamiltonian structure of the Vlasov-Maxwell dynamics, and one that incorporates an approximately self-consistent electromagnetic field.

This is where the lifting appears. To effect an infinite-dimensional Hamiltonian gyrokinetic-like perturbation theory requires changes of particle coordinates, motivated by the guiding-center reduction, that induce changes in the field Hamiltonian dynamics. Thus, this perturbation theory involves a sequence of particle coordinate changes that depends both on the dynamical (field) variables and on their arguments, which are not variables but labels or indices from the point of view of the Hamiltonian structure. This complicates matters significantly and care must be taken when performing transformations, most notably with the chain rule. It is why we will be very careful in establishing the lifting method, especially in Chapters 4 and 11.

In order to emphasize the positioning of the work reported in this episode, notice also that the lifting in this context appears as the process of determining the form of field dynamics, induced by particle orbit dynamics. It is a natural relative of the lifting treated in Ref. [114], which treats the lifting of microscopic particle dynamics up to the field level: the purpose of the present episode is distinct from that of Ref. [114], but the framework is closely related. Indeed, on the one hand, the paper [114] is about a general prescription for the Hamiltonian structure of Vlasov-Maxwell

theories with polarization and magnetization, but often such effects come from a transformation of the particle phase space. On the other hand, in the present episode, we study how to lift a change of particle coordinates to the Hamiltonian structure of Vlasov-Maxwell dynamics; as a side effect, this can generate polarization and magnetization. Thus, these are two complementary aspects of one and the same question.

# Chapter 4

# Lifting particle coordinate changes of magnetic-moment type to Vlasov-Maxwell Hamiltonian dynamics

in collaboration with Phil J. Morrison and Michel Vittot
Paper published in Phys. Plasmas **20**, 032109 (2013)

*A*bstract: Techniques for coordinate changes that depend on both dependent and independent variables are developed and applied to the Vlasov-Maxwell Hamiltonian theory. Particle coordinate changes with a new velocity variable dependent on the magnetic field, with spatial coordinates unchanged, are lifted to the field theoretic level, by transforming the noncanonical Poisson bracket and Hamiltonian structure of the Vlasov-Maxwell dynamics.
Several examples are given including magnetic coordinates, where the velocity is decomposed into components parallel and perpendicular to the local magnetic field, and the case of spherical velocity coordinates. An example of the lifting procedure is performed to obtain a simplified version of gyrokinetics, where the magnetic moment is used as a coordinate and the dynamics is reduced by elimination of the electric field energy in the Hamiltonian.

## Introduction

The goal of this chapter is to start addressing the Hamiltonian structure of gyrokinetics, and more precisely the lifting of the guiding-center reduction to the Hamiltonian structure of Vlasov-Maxwell field dynamics, by dealing with a very simplified case, in which the position coordinate is not changed. Thus, we consider in this chapter just the lifting of the magnetic-moment reduction, whose transformation was studied in Chapter 10. This simplifies matters significantly because the transformation depends only on the fields $\mathbf{E}$ and $\mathbf{B}$, which are not affected at all by the transformation, since they are defined only over the configuration space. It is a way to focus just on the phenomena at work in the lifting procedure, while avoiding any additional intricacy.

The organization of the chapter is the following. In Sec. 4.1 the framework of the subsequent development is made precise, with especially the class of transformation considered and the Hamiltonian structure of Vlasov-Maxwell dynamics to be transformed. It is followed by four sections. Each of them considers a specific transformation, and it is shown how to lift these coordinate transformations, which are tailored to particle orbit dynamics, up to the level of fields, by detailing how to transform the Vlasov-Maxwell Poisson bracket into the new coordinates. In the process, four features associated with steps of lifting must be considered; these are progressively taken into account to reach the final complete transformation for the magnetic moment $\bar{\mu}$. The four features/steps are

Step A A chain rule for functions in the particle bracket $[\cdot,\cdot]$, because the transformation affects particle coordinates.

Step B The presence of the Jacobian both in functionals and in functional derivatives.

Step C A chain rule for functionals in the field bracket $\{\cdot,\cdot\}$, because the transformation depends of the magnetic field, which is a dynamical field.

Step D The possible presence of gradients of the magnetic field in the transformation, which implies the presence of differential operators in the chain rule.

Section 4.2 considers magnetic coordinates, where the particle velocity coordinate is projected parallel and perpendicular to a space-dependent dynamic magnetic field; this introduces the features A and B mentioned above. Next, in Sec. 4.3, spherical velocity coordinates are considered. Here the velocity coordinates are chosen as the unit vector of the velocity (independent of the spatial coordinates) and a coordinate in one-to-one correspondence with the norm of the velocity. This transformation introduces a new feature in that the Jacobian determinant of the transformation is no longer unity, but the transformation does not depend of the magnetic field. This provides a simple example for the features A and C. In Sec. 4.4, we turn to a case that is both more complete and closer to that needed for gyrokinetics with the magnetic moment coordinate: the change of coordinates depends on the local value of the magnetic field, which implies the presence of the features A, B and C. Next, Sec. 4.5 considers the physically important situation, where the change of coordinates involves spatial derivatives of the magnetic field to arbitrary order, i.e., as given by Eq. (4.1) below; this involves all of the four features A, B, C, and D. Finally, with the techniques of the previous four sections in hand, in Sec. 4.6 we treat an example where the reduced coordinate is indeed the magnetic moment and explicitly transform the Hamiltonian form of the Vlasov-Maxwell equation into the new coordinates.

Notice that the first example differs from the others in that the last four sections 4.3-4.6 are about the same ultimate transformation, with each section bringing us closer to the magnetic moment transformation: the velocity coordinates are first chosen as spherical coordinates; then the norm of the velocity is taken as a general local function of the magnetic field; next, it is taken as a general non-local function of the magnetic field; last, it is chosen as the magnetic moment. Also, for the sake of conciseness, the equations of motion of the transformed Vlasov-Maxwell system will be studied only in the example of Sec. 4.6, but it is clear that the conclusions are general and hold for other examples of lifting.

## 4.1 Framework of the present lifting

From a general point of view, a main purpose of this chapter is to transform the Vlasov-Maxwell Hamiltonian structure when the phase-space variables $(\mathbf{q},\mathbf{v})$ are changed to the following new coordinates that depend on the magnetic field and all of its derivatives:

$$\bar{\mathbf{q}} = \mathbf{q}\,, \qquad \bar{\mathbf{v}} = \bar{\mathbf{v}}(\mathbf{q},\mathbf{v};\mathbf{B},\nabla\mathbf{B},\dots)\,. \tag{4.1}$$

This process is called a "lifting" of the transformation (4.1) from the particle phase space (and dynamics) to the Vlasov-Maxwell Hamiltonian structure (and dynamics).

For the noncanonical Hamiltonian structure of Vlasov-Maxwell dynamics, the observables are the set of all functionals of the magnetic field $\mathbf{B}(\mathbf{q})$, the electric field $\mathbf{E}(\mathbf{q})$, and the phase-space density $f(\mathbf{q},\mathbf{v})$, where the time variable has been suppressed. The Poisson bracket is [95, 101, 104, 156]:

$$\begin{aligned}\{F,G\} = &\int d^3q d^3v\, f\, [F_f, G_f] \\ &+ e\int d^3q d^3v\; f\,(G_{\mathbf{E}}\cdot\partial_{\mathbf{v}}F_f - F_{\mathbf{E}}\cdot\partial_{\mathbf{v}}G_f) \\ &+ \int d^3q\, \left(F_{\mathbf{E}}\cdot\nabla\times G_{\mathbf{B}} - G_{\mathbf{E}}\cdot\nabla\times F_{\mathbf{B}}\right),\end{aligned} \tag{4.2}$$

where subscripts are used for functional derivatives, $F_f := \delta F/\delta f$, $F_{\mathbf{E}} := \delta F/\delta\mathbf{E}$, etc., and the particle bracket is $[g,h] = \nabla g\cdot\partial_{\mathbf{v}}h - \nabla h\cdot\partial_{\mathbf{v}}g + e\mathbf{B}\cdot\partial_{\mathbf{v}}g\times\partial_{\mathbf{v}}h$, with $\nabla g = \partial g/\partial\mathbf{q}$ and $\partial_{\mathbf{v}}h = \partial h/\partial\mathbf{v}$, for any functions of the phase space $g(\mathbf{q},\mathbf{v})$ and $h(\mathbf{q},\mathbf{v})$. For the sake of simplicity physical constants

have been scaled away as usual, but a dimensionless charge variable $e$ that indicates the coupling term has been retained (see [114] for a dimensional form of this bracket). The variable $e$ becomes the charge ratios when (4.2) is generalized by summing over multiple species.

The Hamiltonian functional is

$$H[\mathbf{E},\mathbf{B},f] = \tfrac{1}{2}\int d^3q d^3v\, \|\mathbf{v}\|^2 f + \tfrac{1}{2}\int d^3q\, \left(\|\mathbf{E}\|^2 + |\mathbf{B}\|^2\right)\,, \tag{4.3}$$

which is the sum of the kinetic energy of the plasma and the energy of the electromagnetic field. The relativistic model is obtained by replacing $\|\mathbf{v}\|^2$ in the kinetic energy term with $\sqrt{1+\|\mathbf{v}\|^2}$, where in the latter case $\mathbf{v}$ is the scaled relativistic momentum. The coupling between the plasma and electromagnetic field is included in the noncanonical Poisson bracket (4.2). The Hamiltonian (4.3) together with the Poisson bracket generates the motion through Hamilton's equations expressed as

$$\dot{F}=\{F,H\}\,,$$

for any observable $F$. In particular, if $F$ denotes the field variables the bracket induces Vlasov-Maxwell equations as follows:

$$\begin{aligned}
\dot{\mathbf{B}}&=\{\mathbf{B},H\} = -\nabla\times\mathbf{E}\,,\\
\dot{\mathbf{E}}&=\{\mathbf{E},H\} = \nabla\times\mathbf{B} - e\int d^3v\; f\mathbf{v}\,,\\
\dot{f}&=\{f,H\} = -\mathbf{v}\cdot\nabla f - e\,\left(\mathbf{E}+\mathbf{v}\times\mathbf{B}\right)\cdot\partial_{\mathbf{v}} f\,.
\end{aligned} \tag{4.4}$$

As noted in the introduction of this chapter, in order to transform the Hamiltonian structure to facilitate the separation or removal of fast time scales (as in oscillating-center, guiding-center, and gyrokinetic theories) care must be taken because such a change of coordinates involves both the dependent and independent variables, i.e., the spatial observation points of the field. A simple case of this is treated in the next section.

## 4.2 Lifting magnetic velocity coordinates

As a first case of lifting, consider velocity coordinates based on a decomposition of the velocity using the magnetic field. This transformation of the spatial coordinate is unchanged, but the velocity $\mathbf{v}$ is transformed as follows:

$$\mathbf{v} = \mathbf{v}(\bar{\mathbf{v}};\mathbf{B}) = \mathbf{v}(v_{||},\mathbf{v}_\perp;\mathbf{B}) = \mathsf{b}v_{||} + \mathbf{v}_\perp\,,$$

where $\mathsf{b} = \mathbf{B}/\|\mathbf{B}\|$ is the unit vector the direction of the magnetic field,

$$v_{||} = \mathsf{b}\cdot\mathbf{v}$$

is the (scalar) component of the velocity parallel to the magnetic field, and

$$\mathbf{v}_\perp = \mathbf{v} - \mathsf{b}\mathsf{b}\cdot\mathbf{v} = \bar{\bar{I}}_\perp\cdot\mathbf{v}$$

is the (vectorial) component of the velocity perpendicular to the magnetic field, with

$$\bar{\bar{I}}_\perp := \bar{\bar{I}} - \mathsf{b}\mathsf{b} \tag{4.5}$$

being the orthogonal projector onto the plane perpendicular to the magnetic field.

As a side comment, notice that the coordinate $\mathbf{v}_\perp$ is constrained. It is not completely independent of the spatial position, since moving the position $\mathbf{q}$ implies a change in $\mathbf{B}(\mathbf{q})$, and then $\mathbf{v}_\perp$ may have to be changed in order to remain orthogonal to the magnetic field, as it should (see e.g. Eq. (3.11)).

There are two chain rules to consider: that for functions, considered next, and that for functionals, such as the energy expression of (4.3), which will follow.

### 4.2.1 Function chain rule

The transformation of the field Poisson bracket of (4.2) requires the transformation of the particle bracket,

$$[g,h] = \frac{\partial g}{\partial \mathbf{q}} \cdot \frac{\partial h}{\partial \mathbf{v}} - \frac{\partial h}{\partial \mathbf{q}} \cdot \frac{\partial g}{\partial \mathbf{v}} + e\mathbf{B} \cdot \left( \frac{\partial g}{\partial \mathbf{v}} \times \frac{\partial h}{\partial \mathbf{v}} \right) , \tag{4.6}$$

into the new coordinates, $(\mathbf{q}, \mathbf{v}) \to (\mathbf{q}, v_{||}, \mathbf{v}_\perp)$. The following abbreviations are convenient:

$$\nabla := \frac{\partial}{\partial \mathbf{q}} , \quad \partial_i := \frac{\partial}{\partial q_i} , \quad \partial_{||} := \frac{\partial}{\partial v_{||}} , \quad \partial_\perp := \frac{\partial}{\partial \mathbf{v}_\perp} .$$

Note the last operator acts only in the plane perpendicular to $\mathbf{B}$, which implies the following properties:

$$\partial_\perp \bar{g} \cdot \bar{\bar{I}}_\perp = \partial_\perp \bar{g} \qquad \text{and} \qquad \mathsf{b} \cdot \partial_\perp \bar{g} = 0 .$$

Here and everywhere in the chapter, the overbar over a function will indicate a function of the phase space expressed in new coordinates $\bar{g}(\mathbf{q}, \bar{\mathbf{v}})$.

Total variations of $g(\mathbf{q}, \mathbf{v}) = \bar{g}(\mathbf{q}, v_{||}, \mathbf{v}_\perp)$ are given by

$$\delta g = \frac{\partial g}{\partial \mathbf{q}} \cdot \delta \mathbf{q} + \frac{\partial g}{\partial \mathbf{v}} \cdot \delta \mathbf{v} = \nabla \bar{g} \cdot \delta \mathbf{q} + \partial_{||} \bar{g} . \delta v_{||} + \partial_\perp \bar{g} \cdot \delta \mathbf{v}_\perp , \tag{4.7}$$

while variations of the initial and final coordinates are related by

$$\begin{aligned} \delta v_{||} &= \mathsf{b} \cdot \delta \mathbf{v} + (\delta \mathbf{q} \cdot \nabla \mathsf{b}) \cdot \mathbf{v} , \\ \delta \mathbf{v}_\perp &= \bar{\bar{I}}_\perp \cdot \delta \mathbf{v} + \delta \bar{\bar{I}}_\perp \cdot \mathbf{v} \\ &= \bar{\bar{I}}_\perp \cdot \delta \mathbf{v} - (\delta \mathbf{q} \cdot \nabla \mathsf{b}) \, (\mathsf{b} \cdot \mathbf{v}) \\ &\qquad - \mathsf{b} \, (\delta \mathbf{q} \cdot \nabla \mathsf{b}) \cdot \mathbf{v} . \end{aligned} \tag{4.8}$$

For the function chain rule the field $\mathbf{B}$ is assumed to be a fixed function with the coordinates $(\mathbf{q}, \mathbf{v})$ changing.

Inserting (4.8) into (4.7) implies the chain rule relations

$$\frac{\partial g}{\partial \mathbf{v}} = \mathsf{b} \, \partial_{||} \bar{g} + \partial_\perp \bar{g} \cdot \bar{\bar{I}}_\perp = \mathsf{b} \, \partial_{||} \bar{g} + \partial_\perp \bar{g} , \tag{4.9}$$

$$\frac{\partial g}{\partial q_i} = \partial_i \bar{g} + (\mathbf{v} \cdot \partial_i \mathsf{b}) \, \partial_{||} \bar{g} - (\mathsf{b} \cdot \mathbf{v}) \, \partial_\perp \bar{g} \cdot \partial_i \mathsf{b} , \tag{4.10}$$

and using (4.9) and (4.10) in (4.6) gives the particle bracket in the magnetic coordinates

$$\begin{aligned} [\bar{g}, \bar{h}] = & \mathsf{b} \cdot \left( \nabla \bar{g} \, \partial_{||} \bar{h} - \nabla \bar{h} \, \partial_{||} \bar{g} \right) \\ & + \left( \nabla \bar{g} \cdot \partial_\perp \bar{h} - \nabla \bar{h} \cdot \partial_\perp \bar{g} \right) \\ & + \mathbf{a} \cdot \left( \partial_\perp \bar{h} \, \partial_{||} \bar{g} - \partial_\perp \bar{g} \, \partial_{||} \bar{h} \right) \\ & + \partial_\perp \bar{g} \cdot \bar{\bar{\mathbf{b}}} \cdot \partial_\perp \bar{h} + e\mathbf{B} \cdot \left( \partial_\perp \bar{g} \times \partial_\perp \bar{h} \right) , \end{aligned} \tag{4.11}$$

with

$$a_i = \mathbf{v} \cdot \partial_i \mathsf{b} + (\mathsf{b} \cdot \mathbf{v}) \, \mathsf{b} \cdot \nabla \mathsf{b}_i \quad \text{and} \quad \bar{\bar{b}}_{ij} = (\mathsf{b} \cdot \mathbf{v}) \left( \partial_i \mathsf{b}_j - \partial_j \mathsf{b}_i \right) .$$

In all these relations, recall that $\partial_\perp \bar{g} = \partial_\perp \bar{g} \cdot \bar{\bar{I}}_\perp$. This is important because, for instance, the component of $\nabla \bar{g}$, $\mathbf{a}$ or $\bar{\bar{\mathbf{b}}}$ parallel to $\mathsf{b}$ are non-zero, but vanish when contracted with $\partial_\perp \bar{g}$.

### 4.2.2 Jacobian

In general care must be taken with the Jacobian determinant $\mathcal{J}$ when defining functional derivatives, but here Step B is trivial, since the Jacobian is unity

$$\mathcal{J} := \frac{\partial(\mathbf{q}, v_{||}, \mathbf{v}_\perp)}{\partial(\mathbf{q}, \mathbf{v})} = 1 .$$

This follows because rotations have unit Jacobians and at any time there exists a rotation to a cartesian coordinate system with one of the $\mathbf{v}$ axes aligned with $\mathsf{b}$. Thus

$$dz := d^3q d^3v = d^3q dv_{||} d^2v_\perp =: dq dv\,.$$

Because the volume integral is ultimately independent of how it is calculated, $dz$ can be assumed to be independent of $\mathbf{B}$, e.g. when calculating functional derivatives with respect to $\mathbf{B}$, the topic considered next.

### 4.2.3 Functional chain rule

For the functional chain rule, the transformation of the fields must be made definite, Here,

$$\begin{aligned}\mathbf{E}(\mathbf{q})&=\bar{\mathbf{E}}(\mathbf{q})\,, \qquad \mathbf{B}(\mathbf{q}) = \bar{\mathbf{B}}(\mathbf{q})\,,\\ f(\mathbf{q},\mathbf{v})&=\bar{f}(\mathbf{q}, v_{||}, \mathbf{v}_\perp) = \bar{f}(\mathbf{q}, \mathsf{b}\cdot\mathbf{v}, \bar{\bar{I}}_\perp\cdot\mathbf{v})\\ &=f(\mathbf{q}, \mathsf{b}v_{||} + \mathbf{v}_\perp)\,,\end{aligned}$$

where now the coordinates $(\mathbf{q},\mathbf{v})$ are fixed and the field $\mathsf{b}$ varies.

Variation of a transformed functional, $F[f,\mathbf{B},\mathbf{E}] = \bar{F}[\bar{f},\bar{\mathbf{B}},\bar{\mathbf{E}}]$, gives

$$\begin{aligned}\delta F&=\int dz\ F_f\delta f + \int dq\ \left(F_{\mathbf{B}}\cdot\delta\mathbf{B} + F_{\mathbf{E}}\cdot\delta\mathbf{E}\right)\\ &=\int dz\ \bar{F}_{\bar{f}}\delta\bar{f} + \int dq\ \left(\bar{F}_{\bar{\mathbf{B}}}\cdot\delta\bar{\mathbf{B}} + \bar{F}_{\bar{\mathbf{E}}}\cdot\delta\bar{\mathbf{E}}\right)\,.\end{aligned} \tag{4.12}$$

With the variations of the initial and final fields related by

$$\delta\mathbf{E} = \delta\bar{\mathbf{E}}\,, \quad \delta\mathbf{B} = \delta\bar{\mathbf{B}}\,, \quad \text{and} \quad \delta f = \delta\bar{f} + \partial_{||}\bar{f}\ (\mathbf{v}\cdot\delta\mathsf{b}) + \partial_\perp\bar{f}\cdot\delta\bar{\bar{I}}_\perp\cdot\mathbf{v}\,, \tag{4.13}$$

expressions relating functional derivatives of new and old variables can be obtained. Using

$$\delta\bar{\bar{I}}_\perp = -\frac{1}{\|\mathbf{B}\|}\left(\mathsf{b}\ \bar{\bar{I}}_\perp\cdot\delta\mathbf{B} + \bar{\bar{I}}_\perp\cdot\delta\mathbf{B}\ \mathsf{b}\right)\,,$$

and after some work the last equation of (4.13) becomes

$$\delta f = \delta\bar{f} + \frac{(\mathbf{v}_\perp\cdot\delta\mathbf{B})}{\|\mathbf{B}\|}\ \partial_{||}\bar{f} - \frac{v_{||}}{\|\mathbf{B}\|}\ \delta\mathbf{B}\cdot\partial_\perp\bar{f}\,.$$

Inserting this and the other two equations of (4.13) into (4.12), and then equating coefficients, gives the functional chain rule relations

$$\begin{aligned}\frac{\delta F}{\delta f}&=\frac{\delta\bar{F}}{\delta\bar{f}}\,, \quad \frac{\delta F}{\delta\mathbf{E}} = \frac{\delta\bar{F}}{\delta\bar{\mathbf{E}}}\,,\\ \frac{\delta F}{\delta\mathbf{B}}&=\frac{\delta\bar{F}}{\delta\bar{\mathbf{B}}} + \frac{1}{\|\mathbf{B}\|}\int dv\ \frac{\delta F}{\delta\bar{f}}\ \partial^*_{\mathbf{v}}\bar{f}\,,\end{aligned} \tag{4.14}$$

where

$$\partial^*_{\mathbf{v}} := \mathbf{v}_\perp\partial_{||} - v_{||}\partial_\perp\,. \tag{4.15}$$

Finally, the Vlasov-Maxwell bracket expressed in these magnetic coordinates is

$$\begin{aligned}\{F,G\}=&\int dz\ f\ [F_f, G_f]\\ &+e\int dz\ f\,(G_{\mathbf{E}}\cdot\partial_{\mathbf{v}}F_f - F_{\mathbf{E}}\cdot\partial_{\mathbf{v}}G_f)\\ &+\int d^3q\ \left(F_{\mathbf{E}}\cdot\nabla\times\left[G_{\mathbf{B}} + \frac{1}{\|\mathbf{B}\|}\int dv\ G_f\ \partial^*_{\mathbf{v}}f\right]\right.\\ &\left.-G_{\mathbf{E}}\cdot\nabla\times\left[F_{\mathbf{B}} + \frac{1}{\|\mathbf{B}\|}\int dv\ F_f\ \partial^*_{\mathbf{v}}f\right]\right)\,,\end{aligned} \tag{4.16}$$

where the 'bars' have been dropped, $[\,,\,]$ means the bracket of (4.6) rewritten in the new coordinates as (4.11), and $\partial_{\mathbf{v}} = \mathsf{b}\ \partial_{||} + \partial_\perp$ is a shorthand as in (4.9). Note, $\partial^*_{\mathbf{v}}v^2 = 0$.

## 4.3 Lifting spherical velocity coordinates $\mathbf{v} = V\hat{\mathbf{v}}$

Now turn to the new coordinates considered for intrinsic gyrokinetics (used in Chapter 10), which changes only one of the velocity coordinates to get the magnetic moment $\bar{\mu}$. The two other velocity coordinates are usually chosen as the unit vector of the velocity. So, a preliminary change of coordinates consists in adopting spherical coordinates for the velocity space: $\mathbf{v} = V\hat{\mathbf{v}}$ where $V := \|\mathbf{v}\| \in \mathbb{R}_+$ is the norm of the velocity and $\hat{\mathbf{v}} := \mathbf{v}/\|\mathbf{v}\| \in S^2$ is the unit vector of the velocity. This transformation is considered in this section, but later the change $V \to \bar{\mu}$ will be considered.

Step A: The transformation $\mathbf{v} \leftrightarrow (\hat{\mathbf{v}}, V)$ is clearly invertible. For the chain rule the following are needed:

$$\delta V = \hat{\mathbf{v}} \cdot \delta\mathbf{v} \quad \text{and} \quad \delta\hat{\mathbf{v}} = \bar{\bar{\mathcal{I}}}_\perp \cdot \frac{\delta\mathbf{v}}{V}\,,$$

where

$$\bar{\bar{\mathcal{I}}}_\perp = \bar{\bar{I}} - \hat{\mathbf{v}}\hat{\mathbf{v}}$$

is the orthogonal projector onto the plane perpendicular to the velocity. Note $\bar{\bar{\mathcal{I}}}_\perp$ is different from the magnetic projector $\bar{\bar{I}}_\perp$ of (4.5) used in Sec. 4.2.

As in Sec. 4.2 the above are used to calculate the function chain rule, giving

$$\frac{\partial g}{\partial \mathbf{v}} = \frac{1}{V}\frac{\partial \bar{g}}{\partial \hat{\mathbf{v}}} \cdot \bar{\bar{\mathcal{I}}}_\perp + \frac{\partial \bar{g}}{\partial V}\,\hat{\mathbf{v}}\,, \tag{4.17}$$

$$\nabla g = \frac{\partial g}{\partial \mathbf{q}} = \frac{\partial \bar{g}}{\partial \mathbf{q}} = \nabla \bar{g}\,. \tag{4.18}$$

Inserting (4.17) and (4.18) into (4.6) and, after some manipulations, the particle bracket expressed in spherical coordinates is obtained

$$\begin{aligned}[g,h] = &\frac{1}{V}\left(\nabla g \cdot \bar{\bar{\mathcal{I}}}_\perp \cdot \partial_{\hat{\mathbf{v}}} h - \nabla h \cdot \bar{\bar{\mathcal{I}}}_\perp \cdot \partial_{\hat{\mathbf{v}}} g\right)\\ &+\hat{\mathbf{v}} \cdot (\nabla g\, \partial_V h - \nabla h\, \partial_V g)\\ &+\frac{e\mathbf{B}}{V^2} \cdot \left(\partial_{\hat{\mathbf{v}}} g \cdot \bar{\bar{\mathcal{I}}}_\perp\right) \times \left(\partial_{\hat{\mathbf{v}}} h \cdot \bar{\bar{\mathcal{I}}}_\perp\right)\\ &+\frac{e\mathbf{B} \times \hat{\mathbf{v}}}{V} \cdot (\partial_V g\, \partial_{\hat{\mathbf{v}}} h - \partial_{\hat{\mathbf{v}}} g\, \partial_V h)\,,\end{aligned} \tag{4.19}$$

where, for convenience, the 'bars' have been dropped and the abbreviations

$$\frac{\partial g}{\partial \hat{\mathbf{v}}} =: \partial_{\hat{\mathbf{v}}} g \quad \text{and} \quad \frac{\partial g}{\partial V} =: \partial_V g\,,$$

have been employed.

Step B: This step is not trivial here, since the Jacobian for this special case is not unity

$$dz = V^2 dV d\Omega d^3q = \mathcal{J} dV d\Omega d^3q =: \mathcal{J} d\eta d^3q =: \mathcal{J} dw\,, \tag{4.20}$$

because the integration measures are changed from $d^3v$ and $dz$ to $d\eta$ and $dw$, which are defined by Eq. (4.20).

Step C: Turning to the functional chain rule, notice that the change of coordinates does not depend on the fields, and the Step C is simplified here, but it is not completely trivial, because the Jacobian has to be taken into account

$$\begin{aligned}\delta F &= \int dz\; F_f \delta f + \int d^3q\; \left(F_{\mathbf{B}} \cdot \delta\mathbf{B} + F_{\mathbf{E}} \cdot \delta\mathbf{E}\right)\\ &= \int dw\; \bar{F}_{\bar{f}} \delta \bar{f} + \int d^3q\; \left(\bar{F}_{\bar{\mathbf{B}}} \cdot \delta\bar{\mathbf{B}} + \bar{F}_{\bar{\mathbf{E}}} \cdot \delta\bar{\mathbf{E}}\right)\,.\end{aligned} \tag{4.21}$$

Inserting (4.20) into (4.21) gives

$$F_f = \mathcal{J}^{-1}\bar{F}_{\bar{f}}\,, \quad F_{\mathbf{B}} = \bar{F}_{\bar{\mathbf{B}}}\,, \quad \text{and} \quad F_{\mathbf{E}} = \bar{F}_{\bar{\mathbf{E}}}\,. \tag{4.22}$$

Note, in (4.22) the new functional derivative is defined with respect to the *bare* measure $dw$.

So, the first term of the Vlasov-Maxwell bracket transforms as

$$\{F,G\}_1 := \int dz\; f\,[F_f, G_f] = \int dw\; \mathcal{J}\bar{f}\,\left[\mathcal{J}^{-1}\bar{F}_{\bar{f}}, \mathcal{J}^{-1}\bar{G}_{\bar{f}}\right] = \{\bar{F},\bar{G}\}_1\,, \tag{4.23}$$

with the bracket of the second equality above given by (4.19).

In practical computations with Vlasov-like Poisson brackets (especially when computing the equations of motion), a very useful property is the usual '$f$-$g$-$h$' identity, viz. $\int dz\, f[g,h] = -\int dz\, g[f,h]$, for canonical brackets. Here, because of the Jacobian, this identity is changed and replaced by:

$$\int dw\; \mathcal{J} f\,\left[\mathcal{J}^{-1}g, h\right] = -\int dw\; g\,[f,h]\;, \tag{4.24}$$

and in terms of the bare measure

$$\frac{\delta f(w)}{\delta f(w')} = \delta(w-w')\,. \tag{4.25}$$

With this identity (4.24), it is straightforward to verify that the bracket of (4.23) produces the correct current for the Vlasov-Maxwell (and also Vlasov-Poisson) equations of motion.

Now consider the coupling terms of the bracket

$$\begin{aligned}\{F,G\}_2 :=& e\int dz\; f\,(G_{\mathbf{E}}\cdot\partial_{\mathbf{v}}F_f - F_{\mathbf{E}}\cdot\partial_{\mathbf{v}}G_f)\\ =& e\int dw\; \mathcal{J}\bar{f}\,\left(G_{\bar{\mathbf{E}}}\cdot\partial_{\mathbf{v}}\mathcal{J}^{-1}\bar{F}_{\bar{f}} - F_{\bar{\mathbf{E}}}\cdot\partial_{\mathbf{v}}\mathcal{J}^{-1}\bar{G}_{\bar{f}}\right)\;,\end{aligned}$$

where $\partial_{\mathbf{v}}$ is a shorthand for the expression of (4.17). When generating Maxwell's equations, the Hamiltonian gives

$$\bar{H}_{\bar{f}} = \mathcal{J}\|\mathbf{v}\|^2/2\,,$$

which gives the correct expression for the current density $\mathbf{J} = \int d\eta\, \mathcal{J} f \mathbf{v}$.

Finally, the pure field terms of the Vlasov-Maxwell bracket are unchanged and, thus, the Vlasov-Maxwell bracket in these spherical coordinates becomes

$$\begin{aligned}\{F,G\} =& \int dw\; \mathcal{J} f\,\left[\mathcal{J}^{-1}F_f, \mathcal{J}^{-1}G_f\right]\\ &+ e\int dw\; \mathcal{J} f\,\left(G_{\mathbf{E}}\cdot\partial_{\mathbf{v}}\mathcal{J}^{-1}F_f - F_{\mathbf{E}}\cdot\partial_{\mathbf{v}}\mathcal{J}^{-1}G_f\right)\\ &+ \int d^3q\;\left(F_{\mathbf{E}}\cdot\nabla\times G_{\mathbf{B}} - G_{\mathbf{E}}\cdot\nabla\times F_{\mathbf{B}}\right),\end{aligned}$$

where the 'bars' have been dropped, and $[\,,\,]$ means the bracket of (4.6) rewritten in the new coordinates as (4.19).

## 4.4 Inclusion of a local dependence on B

To include the magnetic moment in the coordinates, the next step is to investigate the coordinate transformation $V \leftrightarrow A$, where $A$ is a coordinate in one-to-one correspondence with the coordinate $V$ of Sec. 4.3, but in this section it is assumed to have local dependence on the magnetic field, i.e, it depends on **B** but not its derivatives. Explicitly, the transformation is $(\mathbf{q}, V, \hat{\mathbf{v}}) \leftrightarrow (\bar{\mathbf{q}}, A, \mathsf{v})$ where

$$\mathbf{q} = \bar{\mathbf{q}}\,, \quad \hat{\mathbf{v}} = \mathsf{v}\,, \quad \text{and} \quad V = V(A, \mathsf{v}, \mathbf{B})\,.$$

For an example of an allowable function $V$, see Sec. 4.6. Clearly, invertibility requires $V_A := \partial V/\partial A \neq 0$, which is the only assumption on this function, beside the absence of gradients of the

magnetic field $\mathbf{B}$. Since the first two equations above are identities, eventually $\hat{\mathbf{v}}$ will be used for $\mathsf{v}$ and $\mathbf{q}$ for $\bar{\mathbf{q}}$.

Step A: Now the chain rule is effected on functions analogous to (4.9)-(4.10) and (4.17)-(4.18) and on functionals analogous to (4.14) and (4.22). Varying $g(\mathbf{q}, V, \hat{\mathbf{v}}) = \bar{g}(\bar{\mathbf{q}}, A, \mathsf{v})$ in the label (coordinates) dependence, and then equating as above, gives

$$\frac{\partial g}{\partial \mathbf{q}} = \frac{\partial \bar{g}}{\partial \bar{\mathbf{q}}} - \frac{V_{B_i}}{V_A}\frac{\partial B_i}{\partial \mathbf{q}}\frac{\partial \bar{g}}{\partial A}\,, \tag{4.26}$$

$$\frac{\partial g}{\partial V} = \frac{1}{V_A}\frac{\partial \bar{g}}{\partial A}\,, \tag{4.27}$$

$$\frac{\partial g}{\partial \hat{\mathbf{v}}} = \frac{\partial \bar{g}}{\partial \mathsf{v}} - \frac{1}{V_A}\frac{\partial V}{\partial \hat{\mathbf{v}}}\frac{\partial \bar{g}}{\partial A}\,. \tag{4.28}$$

Inserting (4.27) and (4.28) into (4.17) gives the chain rule on functions

$$\begin{aligned} D_*\bar{g} = \frac{\partial g}{\partial \mathbf{v}} &= \frac{1}{V}\left(\frac{\partial \bar{g}}{\partial \mathsf{v}} - \frac{1}{V_A}\frac{\partial \bar{g}}{\partial A}\frac{\partial V}{\partial \mathsf{v}}\right)\cdot \bar{\bar{\mathcal{I}}}_\perp + \frac{\mathsf{v}}{V_A}\frac{\partial \bar{g}}{\partial A} \\ &= \frac{1}{V}\partial_{\mathsf{v}}\bar{g}\cdot\bar{\bar{\mathcal{I}}}_\perp + \frac{\partial_A \bar{g}}{V_A}\,\mathsf{v} - \frac{\partial_A \bar{g}}{VV_A}\,\partial_{\mathsf{v}}V\cdot\bar{\bar{\mathcal{I}}}_\perp\,, \end{aligned} \tag{4.29}$$

while (4.26) gives

$$\nabla_*\bar{g} = \frac{\partial g}{\partial \mathbf{q}} = \frac{\partial \bar{g}}{\partial \bar{\mathbf{q}}} - \frac{V_{B_i}}{V_A}\frac{\partial B_i}{\partial \bar{\mathbf{q}}}\frac{\partial \bar{g}}{\partial A} = \bar{\nabla}\bar{g} - \frac{V_{B_i}}{V_A}\,\bar{\nabla}B_i\,\partial_A\bar{g}\,. \tag{4.30}$$

Then, inserting (4.29) and (4.30) into (4.19) gives the following complicated expression for the particle bracket [ , ] in the new coordinates:

$$\begin{aligned} [\bar{g}, \bar{h}] = {} & \nabla_*\bar{g}\cdot D_*\bar{h} - \nabla_*\bar{h}\cdot D_*\bar{g} + e\mathbf{B}\cdot\left(D_*\bar{g}\times D_*\bar{h}\right) \\ = {} & \frac{1}{V}\left(\bar{\nabla}\bar{g}\cdot\bar{\bar{\mathcal{I}}}_\perp\cdot\partial_{\mathsf{v}}\bar{h} - \bar{\nabla}\bar{h}\cdot\bar{\bar{\mathcal{I}}}_\perp\cdot\partial_{\mathsf{v}}\bar{g}\right) \\ & + \frac{\mathsf{v}}{V_A}\cdot\left(\bar{\nabla}\bar{g}\,\partial_A\bar{h} - \bar{\nabla}\bar{h}\,\partial_A\bar{g}\right) \\ & + \frac{\partial_{\mathsf{v}}V\cdot\bar{\bar{\mathcal{I}}}_\perp}{VV_A}\cdot\left(\bar{\nabla}\bar{h}\,\partial_A\bar{g} - \bar{\nabla}\bar{g}\,\partial_A\bar{h}\right) \\ & + \frac{V_{B_i}}{VV_A}\,\bar{\nabla}B_i\cdot\bar{\bar{\mathcal{I}}}_\perp\cdot\left(\partial_{\mathsf{v}}\bar{g}\,\partial_A\bar{h} - \partial_{\mathsf{v}}\bar{h}\,\partial_A\bar{g}\right) \\ & + \frac{e\mathbf{B}}{V^2}\cdot\left(\partial_{\mathsf{v}}\bar{g}\cdot\bar{\bar{\mathcal{I}}}_\perp\right)\times\left(\partial_{\mathsf{v}}\bar{h}\cdot\bar{\bar{\mathcal{I}}}_\perp\right) \\ & + \frac{e\mathbf{B}\times\mathsf{v}}{VV_A}\cdot\left(\partial_{\mathsf{v}}\bar{h}\,\partial_A\bar{g} - \partial_{\mathsf{v}}\bar{g}\,\partial_A\bar{h}\right) \\ & - \frac{e\mathbf{B}}{V^2V_A}\times\left(\bar{\bar{\mathcal{I}}}_\perp\cdot\partial_{\mathsf{v}}V\right) \\ & \qquad\cdot\bar{\bar{\mathcal{I}}}_\perp\cdot\left(\partial_{\mathsf{v}}\bar{h}\,\partial_A\bar{g} - \partial_{\mathsf{v}}\bar{g}\,\partial_A\bar{h}\right)\,. \end{aligned} \tag{4.31}$$

Step B: The Jacobian for this transformation is now

$$\begin{aligned} dz &= V^2 dV d\Omega d^3q = V^2 V_A dA d\Omega d^3q \\ &= \mathcal{J} dA d\Omega d^3q =: \mathcal{J} d\eta d^3q =: \mathcal{J} dw\,, \end{aligned}$$

which define the Jacobian $\mathcal{J}$ and the integration measures $d\eta$ and $dw$. Note that these are not the same as those of Sec. 4.3, even though the same symbols are used. Furthermore, $\mathcal{J}$ now depends on $\mathbf{B}$ and, hence, $\mathbf{q}$. Also, $d\Omega$ contains a portion of the Jacobian from cartesian coordinates, but one that is independent of $\mathbf{q}$.

Step C: Now consider the functional chain rule as above,

$$\begin{aligned}\delta F &= \int dz\ F_f \delta f + \int d^3q\ \left(F_{\mathbf{B}} \cdot \delta \mathbf{B} + F_{\mathbf{E}} \cdot \delta \mathbf{E}\right) \\ &= \int dw\ \bar{F}_{\bar{f}} \delta \bar{f} + \int d^3q\ \left(\bar{F}_{\bar{\mathbf{B}}} \cdot \delta \bar{\mathbf{B}} + \bar{F}_{\bar{\mathbf{E}}} \cdot \delta \bar{\mathbf{E}}\right)\ ,\end{aligned} \tag{4.32}$$

Functionally varying $f(\mathbf{q}, V, \hat{\mathbf{v}}) = \bar{f}(\bar{\mathbf{q}}, A, \mathsf{v})$ gives

$$\delta f = \delta \bar{f} + \frac{\partial \bar{f}}{\partial A} \frac{\partial A}{\partial \mathbf{B}} \cdot \delta \mathbf{B}\,, \tag{4.33}$$

while $\delta \mathbf{B} = \delta \bar{\mathbf{B}}$ and $\delta \mathbf{E} = \delta \bar{\mathbf{E}}$. Whence, upon substitution of (4.33) into (4.32), the chain rule on functionals is obtained,

$$\begin{aligned}\frac{\delta F}{\delta f} &= \frac{1}{\mathcal{J}} \frac{\delta \bar{F}}{\delta \bar{f}}\,, \\ \frac{\delta F}{\delta \mathbf{E}} &= \frac{\delta \bar{F}}{\delta \bar{\mathbf{E}}}\,, \\ \frac{\delta F}{\delta \mathbf{B}} &= \frac{\delta \bar{F}}{\delta \bar{\mathbf{B}}} - \int d\eta\ \frac{\partial A}{\partial \mathbf{B}} \frac{\partial \bar{f}}{\partial A} \frac{\delta \bar{F}}{\delta \bar{f}}\,,\end{aligned} \tag{4.34}$$

where the last expression of (4.34) can be written in a more convenient way as

$$\frac{\delta F}{\delta \mathbf{B}} = \frac{\delta \bar{F}}{\delta \bar{\mathbf{B}}} + \int d\eta\ \frac{V_{\bar{\mathbf{B}}}}{V_A} \frac{\partial \bar{f}}{\partial A} \frac{\delta \bar{F}}{\delta \bar{f}}\,.$$

This follows from

$$\frac{\partial A}{\partial \mathbf{B}} = -\frac{V_{\bar{\mathbf{B}}}}{V_A}\,,$$

which comes about because the change in $A$ induced by a change in $\mathbf{B}$ at fixed $V$ and $\mathsf{v}$, satisfies $0 = \delta V = V_A \delta A + V_{\bar{B}_i} \delta \bar{B}_i$.

Finally, the Vlasov-Maxwell bracket in the coordinates $(\mathbf{q}, A, \hat{\mathbf{v}})$ is given by

$$\begin{aligned}\{F, G\} &= \int d\eta d^3q\ \mathcal{J} f \left[\mathcal{J}^{-1} F_f, \mathcal{J}^{-1} G_f\right] \\ &+ e \int d\eta d^3q\ \mathcal{J} f \left(G_{\mathbf{E}} \cdot D_* \mathcal{J}^{-1} F_f - F_{\mathbf{E}} \cdot D_* \mathcal{J}^{-1} G_f\right) \\ &+ \int d^3q\ \left(F_{\mathbf{E}} \cdot \nabla \times \left[G_{\mathbf{B}} + \int d\eta\ \frac{V_{\mathbf{B}}}{V_A} \frac{\partial f}{\partial A} \frac{\delta G}{\delta f}\right]\right. \\ &\qquad \left. - G_{\mathbf{E}} \cdot \nabla \times \left[F_{\mathbf{B}} + \int d\eta\ \frac{V_{\mathbf{B}}}{V_A} \frac{\partial f}{\partial A} \frac{\delta F}{\delta f}\right]\right)\,,\end{aligned} \tag{4.35}$$

where the particle bracket [ , ] is given by (4.31), $D_*$ is the operator defined by (4.29), and the bars have been dropped.

## 4.5 Inclusion of a nonlocal dependence on B

In order to include the physical coordinates where $A$ is the magnetic moment $\bar{\mu}$, the last step is to consider the case where the coordinate transformation involves derivatives of the magnetic field. This is important because perturbative reductions, such as those based on Lie-transforms [27,60,87] or mixed variable generating functions [40], often involve derivatives to arbitrary high order in the fields.

So, a more general transformation to new coordinates $(\mathbf{q}, V, \hat{\mathbf{v}}) \leftrightarrow (\bar{\mathbf{q}}, A, \mathsf{v})$ is considered:

$$\mathbf{q} = \bar{\mathbf{q}}, \quad \hat{\mathbf{v}} = \mathsf{v}\,, \quad \text{and} \quad V = V[A, \mathsf{v}, \mathbf{B}]\,, \tag{4.36}$$

where now $V[A, \mathsf{v}, \mathbf{B}]$ means a transformation that depends on $\mathbf{B}$ and, possibly, all its derivatives. For an example of an allowable function $V$, see Sec. 4.6. Clearly, invertibility requires $V_A := \partial V/\partial A \neq 0$, which is the only assumption on this function. Since the first two equations above are identities, as before eventually $\hat{\mathbf{v}}$ will be used for $\mathsf{v}$ and $\mathbf{q}$ for $\bar{\mathbf{q}}$.

Step B: The Jacobian for this transformation is again

$$dz = V^2 V_A dA d\Omega d^3q = \mathcal{J} dA d\Omega d^3q =: \mathcal{J} d\eta d^3q =: \mathcal{J} dw\,.$$

but now $\mathcal{J}$ depends on $\mathbf{q}$ through $\mathbf{B}$ and its derivatives.

Steps A and C: For the chain rule on functions or functionals, $g(\mathbf{q}, V, \hat{\mathbf{v}}) = \bar{g}(\bar{\mathbf{q}}, A, \mathsf{v})$ and $h(\mathbf{q}, V, \hat{\mathbf{v}}) = \bar{h}(\bar{\mathbf{q}}, A, \mathsf{v})$ are varied as in Sec. 4.4, and all terms are the same as before, except some slight changes in the relations involving derivatives with respect to the magnetic field (Step D), which we study in the next paragraph.

Step D: Indeed, the Fréchet derivative with respect to $\mathbf{B}$ is now a differential operator, and care must be taken with the order of terms. For instance, Eq. (4.26) becomes

$$\frac{\partial g}{\partial \mathbf{q}} = \frac{\partial \bar{g}}{\partial \bar{\mathbf{q}}} - \frac{\partial \bar{g}}{\partial A}\,\frac{1}{V_A}\,V_{B_i}\,\frac{\partial B_i}{\partial \mathbf{q}}\,, \tag{4.37}$$

where $V_{B_i}$ is now a differential operator acting on $\partial B_i/\partial \mathbf{q}$. Eqs. (4.30)-(4.31) must be changed accordingly.

As for Eqs. (4.33)-(4.34), variation is performed slightly differently this time as follows:

$$\delta \bar{f} = \delta f + f_V\, V_{\bar{\mathbf{B}}} \cdot \delta \bar{\mathbf{B}}\,,$$

where $V_{\bar{\mathbf{B}}}$ is the Fréchet derivative operating on $\delta \bar{\mathbf{B}}$. Thus the chain rule for functional derivatives gives

$$\frac{\delta F}{\delta \mathbf{B}} = \frac{\delta \bar{F}}{\delta \bar{\mathbf{B}}} + \int d\eta\, V_{\mathbf{B}}^{\dagger}\left(\frac{\partial f}{\partial V}\frac{\delta \bar{F}}{\delta \bar{f}}\right) = \frac{\delta \bar{F}}{\delta \bar{\mathbf{B}}} + \int d\eta\, V_{\mathbf{B}}^{\dagger}\left(\frac{\bar{F}_{\bar{f}}}{V_A}\frac{\partial \bar{f}}{\partial A}\right)\,,$$

where the adjoint $\dagger$ is done with respect to $dw$.

Finally, the Vlasov-Maxwell bracket (4.35) in these coordinates becomes

$$\begin{aligned}\{F, G\} &= \int d\eta d^3q\; \mathcal{J} f\, [\mathcal{J}^{-1} F_f, \mathcal{J}^{-1} G_f] \\ &+ e \int d\eta d^3q\; \mathcal{J} f\, (G_{\mathbf{E}} \cdot D_* \mathcal{J}^{-1} F_f - F_{\mathbf{E}} \cdot D_* \mathcal{J}^{-1} G_f) \\ &+ \int d^3q\; \left(F_{\mathbf{E}} \cdot \nabla \times \left[G_{\mathbf{B}} + \int d\eta\, V_{\mathbf{B}}^{\dagger}\left(\frac{G_f}{V_A}\frac{\partial f}{\partial A}\right)\right]\right. \\ &\qquad \left. - G_{\mathbf{E}} \cdot \nabla \times \left[F_{\mathbf{B}} + \int d\eta\, V_{\mathbf{B}}^{\dagger}\left(\frac{F_f}{V_A}\frac{\partial f}{\partial A}\right)\right]\right)\,.\end{aligned} \tag{4.38}$$

## 4.6 Application to the magnetic moment

With the transformed bracket (4.38), the first thing to be checked is whether the dynamics agrees with the conservation of the magnetic moment, when appropriate, since this is what suggested the reduction in the first place. To this end, suppose the coordinate $A$ is the magnetic moment, $A := \bar{\mu}(\mathbf{q}, \mathbf{v})$, which to lowest order is given by $A = \mu = \|\mathbf{v}_\perp\|^2/\|\mathbf{B}\|$. To get a true conserved quantity, small corrections must be added to all orders in the Larmor radius, including derivatives of all orders in the magnetic field [29, 62], as is clear in Chapters 10 and 2.

This illustrates an example of an allowable function $V$. For the case studied in Sec. 4.4, one can choose the coordinate $V(A)$ to be the inverse of the zeroth-order magnetic moment relation $A = \mu = \|\mathbf{v}_\perp\|^2/\|\mathbf{B}\| = V^2\|\hat{\mathbf{v}} \times \mathsf{b}\|^2/\|\mathbf{B}\|$:

$$V_0 := \frac{\sqrt{A\|\mathbf{B}\|}}{\|\hat{\mathbf{v}} \times \mathsf{b}\|} = \frac{\sqrt{A}(\mathbf{B}^2)^{3/4}}{\sqrt{(\hat{\mathbf{v}} \times \mathbf{B})^2}}\,,$$

which implies for the function $V_{\mathbf{B}}$ of Eq. (4.35)

$$(V_0)_{\mathbf{B}} = \frac{3}{4}\frac{V}{\mathbf{B}^2}(\mathbf{B}^2)_{\mathbf{B}} - \frac{1}{2}\frac{V}{(\hat{\mathbf{v}}\times\mathbf{B})^2}((\hat{\mathbf{v}}\times\mathbf{B})^2)_{\mathbf{B}} = \frac{V}{2B}\left(3\mathsf{b} - 2\frac{\bar{\bar{\mathcal{I}}}_\perp\cdot\mathsf{b}}{(\hat{\mathbf{v}}\times\mathsf{b})^2}\right).$$

For the case studied in Sec. 4.5, one can choose the coordinate $V(A)$ to be the inverse of the following expression for the magnetic moment to first order [29, 62] (see Chapter 10):

$$A = \frac{\|\mathbf{v}_\perp\|^2}{\|\mathbf{B}\|} - 2\frac{\|\mathbf{v}_\perp\|^3}{\|\mathbf{B}\|^2}\left[\frac{\mathsf{a}\cdot\nabla B}{2B} + \phi\frac{3\mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{c} - \mathsf{c}\cdot\nabla\mathsf{b}\cdot\mathsf{a}}{4} + \phi^2\mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{a}\right], \tag{4.39}$$

where $\phi := (\mathsf{b}\cdot\hat{\mathbf{v}})/(\sqrt{(\mathsf{b}\cdot\hat{\mathbf{v}})^2})$, and the vectors $\mathsf{a} := (\mathsf{b}\times\hat{\mathbf{v}})/\|\hat{\mathbf{v}}\times\hat{\mathbf{v}}\|$ and $\mathsf{c} := \mathsf{a}\times\mathsf{b}$ are the unit vectors of the perpendicular velocity and of the Larmor radius, respectively. In Eq. (4.39), the second term is assumed to be small compared to the first. Then, the inverse function $V(A)$ is given by

$$V_1 := \frac{1}{\|\mathbf{v}_\perp\|}\left\{\sqrt{AB} + A\left[\frac{\mathsf{a}\cdot\nabla B}{2B} + \phi\frac{3\mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{c} - \mathsf{c}\cdot\nabla\mathsf{b}\cdot\mathsf{a}}{4} + \phi^2\mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{a}\right]\right\}. \tag{4.40}$$

Here, we are not interested in computing the complete quantity $V_{\mathbf{B}}$, but only in showing how it can be an operator when the coordinate $V$ involves derivatives of the magnetic field. As an example, consider the term $\frac{A}{\|\mathbf{v}_\perp\|}\frac{\mathsf{a}\cdot\nabla B}{2B}$ in (4.40). When computing $V_{\mathbf{B}}$, the contribution coming from $B$, which is acted upon by the gradient, is given by the first variation in $\delta\mathbf{B}$ of

$$\frac{A\mathsf{a}\cdot\nabla}{2B\|\mathbf{v}_\perp\|}\sqrt{(\mathbf{B}+\delta\mathbf{B})^2} - \frac{A\mathsf{a}\cdot\nabla}{2B\|\mathbf{v}_\perp\|}\sqrt{\mathbf{B}^2},$$

which is just

$$\frac{A\mathsf{a}\cdot\nabla}{2B\|\mathbf{v}_\perp\|}(\mathsf{b}\cdot\delta\mathbf{B}) = \frac{A}{2B\|\mathbf{v}_\perp\|}\left(\mathsf{a}\cdot\nabla\mathsf{b}\cdot + \mathsf{b}\cdot(\mathsf{a}\cdot\nabla)\right)\delta\mathbf{B}.$$

The corresponding contribution to $V_{\mathbf{B}_i}$ is

$$\frac{A}{2B\|\mathbf{v}_\perp\|}\left(\mathsf{a}\cdot\nabla\mathsf{b}_i + \mathsf{b}_i\ \mathsf{a}\cdot\nabla\right).$$

As expected, it is a first-order differential operator.

For a more precise expression of the magnetic moment, higher order corrections could be added, and $V_{\mathbf{B}}$ would be a differential operator of higher order.

Let us now check the conservation of the magnetic moment $A := \bar{\mu} = \|\mathbf{v}_\perp\|^2/\|\mathbf{B}\| + O(\epsilon)$, which is exactly defined as solution of the following equation

$$0 = \dot{\bar{\mu}} = \mathbf{v}\cdot\nabla\bar{\mu} + e\mathbf{v}\times\mathbf{B}\cdot\partial_{\mathbf{v}}\bar{\mu}.$$

At the field level, the conservation of the magnetic moment corresponds to the conservation of the functional

$$M := \int dz\ f\bar{\mu}$$

for any particle distribution $f$. In the transformed coordinates, this is

$$\bar{M} := \int dw\ \mathcal{J}\bar{f}\bar{\mu}.$$

To investigate the conservation of $\bar{M}$, note that a static magnetic field corresponds to elimination of the electric field term in the Hamiltonian functional, since this eliminates the $\nabla\times\mathbf{E}$ term in the Maxwell-Faraday equation (see Chapter 6). In this case

$$\begin{aligned}\dot{\bar{M}} = \{\bar{M},\bar{H}\} &= \int d\eta d^3q\ \mathcal{J}\bar{f}\left[\bar{\mu},\mathcal{J}^{-1}\bar{H}_{\bar{f}}\right]\\ &= \tfrac{1}{2}\int d\eta d^3q\ \mathcal{J}\bar{f}\left(\nabla_*\bar{\mu}\cdot D_*V^2 + e\mathbf{B}\cdot D_*\bar{\mu}\times D_*V^2\right)\\ &= \int d^3v d^3q\ f\left(\mathbf{v}\cdot\nabla\bar{\mu} + \mathbf{v}\times e\mathbf{B}\cdot\partial_{\mathbf{v}}\bar{\mu}\right) = 0,\end{aligned}$$

as was expected.

Accordingly, the transformed bracket (4.38) is expressed in coordinates adapted to the conserved magnetic moment. As is ususal in gyrokinetics, the electromagnetic field dynamics spoils the conservation of the magnetic moment. This is why the feed-back of the plasma dynamics onto the electromagnetic field dynamics needs to be restored as a perturbation, i.e., a perturbed magnetic moment must be defined that is conserved [17].

Consider now the transformed Vlasov-Maxwell equations of motion generated by the bracket (4.38). In this bracket, most of the terms are actually identical to those of the initial bracket (4.2), even though their formal expressions look different because they are expressed in the reduced coordinates $(\bar{\mathbf{q}}, A, \mathsf{v})$, e.g. through Eqs. (4.29) and (4.31). The only new terms are

$$\int d^3q\ \bar{F}_{\bar{\mathbf{E}}} \cdot \nabla \times \int d\eta\ V^\dagger_{\bar{\mathbf{B}}} \left( \frac{\bar{G}_f}{V_A} \frac{\partial \bar{f}}{\partial A} \right) ,$$

and one obtained by permuting $\bar{F}$ and $\bar{G}$ (and with a minus sign for bracket antisymmetry).

In the equations of motion, this new bracket term generates an additional term in Maxwell-Ampere equation, viz.

$$\begin{aligned} \dot{\bar{\mathbf{E}}} = & \nabla \times \bar{H}_{\bar{\mathbf{B}}} - e \int d\eta\ \mathcal{J} \bar{f}\ D_* \left( \mathcal{J}^{-1} \bar{H}_{\bar{f}} \right) \\ & + \nabla \times \int d\eta\ V^\dagger_{\bar{\mathbf{B}}} \left( \frac{\bar{H}_f}{V_A} \frac{\partial \bar{f}}{\partial A} \right) . \end{aligned} \tag{4.41}$$

At first glance this additional term looks like a new magnetization current. But, one must remember that the usual $\nabla \times \mathbf{B}$ term has itself another additional contribution $\nabla \times \delta \bar{H}_{kin} / \delta \bar{\mathbf{B}}$, because in the reduced variables, the plasma kinetic energy depends on the magnetic field $\bar{H}_{kin} := \int dw \mathcal{J} \bar{f} V^2/2$ that is not constant in $\bar{\mathbf{B}}$ (both because of $\mathcal{J}$ and $V$). And, it turns out that this last additional contribution exactly cancels the "magnetization" term in (4.41):

$$\begin{aligned} \frac{\delta \bar{H}_{kin}}{\delta \bar{\mathbf{B}}(\mathbf{x})} = & \tfrac{1}{2} \int dw\ \bar{f} \left( \mathcal{J} V^2 \right)_{\bar{\mathbf{B}}} \delta(\mathbf{q} - \mathbf{x}) \\ = & \tfrac{1}{2} \int dw\ \bar{f} \left( \partial_A V \cdot V^4 \right)_{\bar{\mathbf{B}}} \delta(\mathbf{q} - \mathbf{x}) \\ = & -\tfrac{1}{10} \int dw\ \partial_A \bar{f} \left( V^5 \right)_{\bar{\mathbf{B}}} \delta(\mathbf{q} - \mathbf{x}) \\ = & -\tfrac{1}{10} \int dw\ \partial_A \bar{f} \left( V^5 \right)_V V_{\bar{\mathbf{B}}} \delta(\mathbf{q} - \mathbf{x}) \\ = & -\tfrac{1}{2} \int dw\ \partial_A \bar{f}\ \mathcal{J} V^2 V_A\ V_{\bar{\mathbf{B}}} \delta(\mathbf{q} - \mathbf{x}) \\ = & - \int dw\ \partial_A \bar{f}\ \tfrac{\bar{H}_{\bar{f}}}{V_A}\ V_{\bar{\mathbf{B}}} \delta(\mathbf{q} - \mathbf{x}) \\ = & - \int d\eta\ V^\dagger_{\bar{\mathbf{B}}} \left( \tfrac{\bar{H}_{\bar{f}}}{V_A}\ \partial_A \bar{f} \right) . \end{aligned} \tag{4.42}$$

This cancellation was to be expected, since the electric field $\mathbf{E} = \bar{\mathbf{E}}$ is not affected by the change of velocity coordinates, and the current term has not been changed either, but only expressed in the new variables:

$$-e \int d\eta\ \mathcal{J} \bar{f}\ D_* \left( \mathcal{J}^{-1} \bar{H}_{\bar{f}} \right) = -\frac{e}{2} \int d^3v\ f\ \partial_{\mathbf{v}} \left( \mathcal{J}^{-1} \mathcal{J} \|\mathbf{v}\|^2 \right) = -\mathbf{J} .$$

Finally, the additional term in the transformed bracket (4.38) generates another additional term in the equation of motion: the dynamics of the Vlasov phase-space density $\dot{f}$ has an additional force term

$$-\frac{1}{V_A} \frac{\partial \bar{f}}{\partial A}\ V_{\bar{\mathbf{B}}} \cdot \nabla \times \bar{\mathbf{E}} .$$

This term is not cancelled by any other term. It can be rewritten as

$$-\frac{\partial f}{\partial V}\, V_{\mathbf{B}} \cdot \nabla \times \mathbf{E} = \frac{\partial f}{\partial V}\, V_{\mathbf{B}} \cdot \dot{\mathbf{B}}\,,$$

which is exactly the expected contribution when applying the chain rule for the time derivative of the transformed fields. It comes about because the change of coordinates is time-dependent when the magnetic field is not static.

## Conclusion

In this chapter techniques for transforming the Vlasov-Maxwell Poisson bracket to new coordinates, when the transformation law mixes dependent and independent variables, have been developed. Four transformations were considered, each of which considered a new feature needed for understanding the more general transformation of (4.36). In Sec. 4.2 a transformation that mixed the independent velocity variable with the magnetic field was considered and the associated function and functional chain rules were described. In Sec. 4.3, spherical velocity coordinates were treated and here it was seen how a nontrivial Jacobian determinant influences a transformation. In Sec. 4.4 a class of transformations that mixes the dependent and independent variables by having dependence on $\mathbf{B}$ and in addition possesses a nontrivial Jacobian was considered. Finally, in Sec. 4.5, the nonlocal transformation of (4.36) was effected, the most general transformation of this chapter that results in the transformed noncanonical Poisson bracket of (4.38). This final form of the Poisson bracket was seen to contain additional terms that appear to be magnetization-like contributions. However, these bracket terms were shown to produce no magnetization term in the equations of motion, since the electromagnetic fields are not affected by the change of field coordinates. Only the dynamics of the Vlasov density obtained an additional term, a term that results from the change of field coordinates being time-dependent through $\mathbf{B}$.

The transformations of Secs. 4.2–4.5 paved the way for the simple example of Sec. 4.6. Here the dynamics was reduced by dropping the electric field energy from the Hamiltonian, resulting in the magnetic moment being conserved by a reduced dynamics that must have a static magnetic field. However, when restoring the feed-back of the plasma dynamics onto the electromagnetic field dynamics, the magnetic moment was seen to be no longer conserved and must be perturbatively changed to be conserved. This will be considered with more details in Chapter 6.

In all the cases considered, the lifting was eased because the change of coordinates only concerned a new particle velocity that depends on the magnetic field, but no change was made in the spatial coordinate. If Eq. (4.1) is generalized by adding dependence on the electric field and all its derivatives, then results similar to those presented are immediate. However, if the new spatial variable has velocity and field dependence, then the situation becomes considerably more complex. Such transformations are of interest for some oscillation-center, guiding-center, and gyrokinetic theory development, and the same methods of function and functional chain rule can be used, but some additional effects will show up, e.g., non-zero polarization and magnetization terms like those of [114]. This more general lifting will be considered in Chapter 5.

However, as mentioned in the introduction of the episode, the lifting for the Vlasov-Maxwell system will involve two additional difficulties: a transformation depending on fields defined over a space affected by the transformation, and the presence of two different spaces over which some fields are defined. Before tackling a lifting with those two difficulties, it is more convenient to study the first difficulty alone, which is the next step of the work and is reported in the appendix chapter 11.

⟶ See the appendix chapter 11

# Chapter 5

# Lifting guiding-center particle coordinate changes to Vlasov-Maxwell Hamiltonian dynamics

in collaboration with Alain J. Brizard, Phil J. Morrison and Michel Vittot

*A*bstract: General changes of particle coordinates depending on the electromagnetic field are lifted to the Vlasov-Maxwell Hamiltonian dynamics. Compared to previous works, in addition to a change of all the particle coordinates, the presence of two different spaces over which fields are defined is taken into account. This implies to distinguish between particle position and field position, and generates polarization and magnetization terms, whereas they were absent from previous examples of lifting.
The results are applied to the guiding-center and gyro-center reductions. It is shown that they indeed give rise to a Hamiltonian reduced field dynamics, which includes the effects of the plasma-field coupling and the time dependence of the transformation in a more complete way than standard approaches, with striking consequences on gyrokinetic theory. Correction terms are identified in the reduced Vlasov equation. For the magnetic moment to be conserved, the gyro-center transformation has to depend not only on the magnetic and electric fields, but also on all the moments of the Vlasov density. Also, the removal of the gyro-angle from the theory is not trivial, because it may spoil the Hamiltonian structure.
Last, in order to agree with the possible presence of several species in gyrokinetic theory, or to include electron spin effects, extensions of the method are considered towards a general field theory where there are several fields, each defined over its own space.

## Introduction

Chapters 4 and 11 studied the basic phenomena at work when lifting a field-dependent coordinate transformation to the corresponding Hamiltonian field dynamics. We now turn to the application that motivated the study, by lifting the guiding-center transformation to the Vlasov-Maxwell system.

Indeed, as a first step towards Hamiltonian gyrokinetics and more precisely towards the lifting of the guiding-center reduction to the Vlasov-Maxwell Hamiltonian structure, Chapter 4 focused on the introductory case where the coordinate $\mathbf{q}$ was not affected by the transformation. It aimed at identifying the essential mechanisms involved in the lifting, while avoiding questions caused by a transformation depending on fields that are defined over a space which is itself affected by the transformation.

For gyrokinetics, this phenomenon must be taken into account, since transformations affect all

the coordinates $\mathbf{z} = (\mathbf{q}, \mathbf{p})$, even the one over which the electromagnetic field is defined. In order to focus on the main mechanisms at work in such a case, the appendix chapter 11 studied a framework simpler than the Vlasov-Maxwell system. It avoided additional and non-essential difficulties coming because all the Vlasov-Maxwell fields are not defined over the same space, since the Vlasov density is defined over the particle phase space, whereas the electromagnetic field is defined over the configuration space, and those fields do not transform the same way. Thus Chapter 11 studied coordinate transformations for a general fluid-like theory, where all the fields are defined over the same space.

Equipped with these results, we can now turn to the physical problem that initially motivated the study: lifting a transformation of particle dynamics, such as the guiding-center transformation, to the Hamiltonian structure of the Vlasov-Maxwell dynamics. The results of Chapters 4 or 11 cannot be used directly, because the particle position is affected by the transformation and all the fields are not defined on the same space. But the principles and methods can be applied or adapted to get similar results. All the same, the outcomes are expected to have original features, because the liftings considered up to now could not generate non-zero polarization and magnetization, which should be present in the gyrokinetic equations of motion.

The organization of the chapter is the following. In Sec. 5.1, we will make more precise the specificities of the lifting we are interested in, compared to the examples previously considered. In Sec. 5.2, the lifting will be performed for a general transformation on particle phase space, depending on the electromagnetic field. In Sec. 5.3, the results will be analysed and compared with a related work in the literature. Last, in Sec. 5.4, we will specify to the guiding-center and gyro-center transformations. In an appendix, the results obtained for the Vlasov-Maxwell system will be extended to a generic system with several sets of fields, each one defined over its own space.

## 5.1 Specificities of the Vlasov-Maxwell system

Compared to the lifting of Chapter 4, we consider now a transformation for all the phase-space coordinates

$$\mathbf{z} \longrightarrow \bar{\mathbf{z}}[\mathbf{z}; \mathbf{E}, \mathbf{B}] = \bar{\mathbf{z}}\Big(\mathbf{z}, \mathbf{E}(\mathbf{z}), \mathbf{B}(\mathbf{z}), \nabla\mathbf{E}(\mathbf{z}), \nabla\mathbf{B}(\mathbf{z}), \nabla\nabla\mathbf{E}(\mathbf{z}), ...\Big)\,,$$

where, as usual, the square parentheses (brackets) indicate a non-local dependence in the fields, because the transformation depends not only of the value of the electromagnetic field, but possibly on all their derivatives, as occurs for theories like guiding-center or gyrokinetics.

Compared to Chapter 11, changes come because the fields are the Vlasov density $f(\mathbf{q}, \mathbf{p})$, defined over the particle phase space $\mathbf{z} := (\mathbf{q}, \mathbf{p})$, and the electromagnetic field $\Big(\mathbf{E}(\mathbf{q}), \mathbf{B}(\mathbf{q})\Big)$, defined over the particle configuration space. The presence of those two different spaces is already a new feature of this system.

In addition, the transformed fields are defined by:

$$\bar{f}(\bar{\mathbf{z}}) = f(\mathbf{z})\,, \qquad \bar{\mathbf{E}}(\bar{\mathbf{q}}) = \mathbf{E}(\bar{\mathbf{q}})\,, \qquad \bar{\mathbf{B}}(\bar{\mathbf{q}}) = \mathbf{B}(\bar{\mathbf{q}})\,, \tag{5.1}$$

where the transformed Vlasov density is defined by the scalar invariance property, whereas the electromagnetic field is not changed. This choice intends to avoid introducing the velocity in the argument of the electromagnetic field: otherwise the scalar invariance property would have implied $\bar{\mathbf{E}}(\bar{\mathbf{q}}) = \mathbf{E}(\mathbf{q}) = \mathbf{E}(\mathbf{q}[\bar{\mathbf{z}}, \bar{\mathbf{E}}, \bar{\mathbf{B}}])$.

This defines a change of field coordinates, which can be studied in a way analogous to Chapters 4 and 11, but it does not correspond to a lifting, where the change of field coordinates is induced by the change of particle coordinates through the scalar invariance property. To treat this transformation as a lifting, one should consider the definition space $\mathbf{x}$ for the electromagnetic field $\Big(\mathbf{E}(\mathbf{x}), \mathbf{B}(\mathbf{x})\Big)$ as different from the configuration space $\mathbf{q}$.

Physically, the spaces $\mathbf{x}$ and $\mathbf{q}$ are identical; it is traduced by a Dirac delta functions $\delta(\mathbf{q} - \mathbf{x})$ everywhere in the theory in the initial system, but not in the final system, since the transformation does not verify this property, as is clear in Eq. (5.11) for instance.

The specific case treated in Chapter 4 corresponded to the case where this issue was avoided: when the coordinate transformation does not change the position $\bar{\mathbf{q}} = \mathbf{q}$, then the transformation for fields (5.1) agrees with the scalar invariance property. Thus, the lifting can be done without distinguishing the electromagnetic field space $\mathbf{x}$ from the particle space $\mathbf{q}$.

## 5.2 Lifting for a general transformation

The coordinate space is $\Big(\mathbf{z}, \mathbf{x}\Big)$, the Vlasov density is defined over $\mathbf{z}$ and the electromagnetic field is defined over $\mathbf{x}$. The particle transformation $\tau$ is

$$\Big(\mathbf{z}, \mathbf{x}\Big) \longrightarrow \Big(\bar{\mathbf{z}}, \bar{\mathbf{x}}\Big) = \tau\Big(\mathbf{z}, \mathbf{x}\Big) = \Big(\bar{\mathbf{z}}[\mathbf{z}, \mathbf{E}, \mathbf{B}], \mathbf{x}\Big) \,. \tag{5.2}$$

The space $\mathbf{x}$ is unchanged, which implies that the transformation for fields is just given by the scalar invariance

$$\bar{f}(\bar{\mathbf{z}}) = f(\mathbf{z}) \,, \qquad \bar{\mathbf{E}}(\bar{\mathbf{x}}) = \mathbf{E}(\mathbf{x}) \,, \qquad \bar{\mathbf{B}}(\bar{\mathbf{x}}) = \mathbf{B}(\mathbf{x}) \,, \tag{5.3}$$

so that now the transformation is a lifting and the same method as in Chapters 4 and 11 can be applied. A chain-rule formula is obtained by computing the total derivative of the defining equations (5.3); transformed functional derivatives give the same chain-rule formula, provided a lemma analogous to (11.25) is used. We do not give the details of the derivation, but just write the results. The main difference is that now there are two coordinate spaces over which fields are defined, so care must be taken over which variable the operators act and over which spaces integrals are taken or not. This makes the derivation slightly more elaborate. Notice that from a conceptual point of view the transformation (5.3) simplifies the framework compared to the general lifting of Chapter 11, because the fields $\mathbf{E}$ and $\mathbf{B}$ involved in the coordinate transformation are not changed. This is an effect of the distinction in Eq. (5.2) between the Vlasov position $\mathbf{q}$, which is changed, and the Maxwell position $\mathbf{x}$, which remains unchanged.

The Hamiltonian of the initial system is just the sum of the kinetic energy of the plasma and of the energy of the electromagnetic field

$$H = \int d\mathbf{z} \; f \tfrac{\mathbf{p}^2}{2} + \int d\mathbf{x} \; \tfrac{\mathbf{E}^2 + \mathbf{B}^2}{2} \,, \tag{5.4}$$

where the physical coefficient $m$, $c$ and $\mu_0$ have been scaled out. Even the charge $e$ will be scaled out in this chapter, for simplicity. The Hamiltonian has a specific form (separate contributions from Vlasov density and from the electromagnetic field). More generally functionals have the form

$$G = \int d\mathbf{z} d\mathbf{x} \; g[\mathbf{z}, \mathbf{x}; f, \mathbf{E}, \mathbf{B}] \,. \tag{5.5}$$

Physically $\mathbf{x} = \mathbf{q}$, so that physical functionals have the form

$$\begin{aligned} G_{phys} &= \int d\mathbf{z} d\mathbf{x} \; \delta(\mathbf{x} - \mathbf{q}) \; g_{phys}[\mathbf{z}, \mathbf{x}; f, \mathbf{E}, \mathbf{B}] \\ &= \int d\mathbf{z} \; g_{phys}[\mathbf{z}, \mathbf{q}; f, \mathbf{E}, \mathbf{B}] \,, \end{aligned} \tag{5.6}$$

and $g_{phys}[\mathbf{z}, \mathbf{q}; f, \mathbf{E}, \mathbf{B}]$ can be considered as just $g_{phys}[\mathbf{z}; f, \mathbf{E}, \mathbf{B}]$. The special functionals that depend only on the electromagnetic field and on the position $\mathbf{x} = \mathbf{q}$ can be considered as included in the forms (5.5) and (5.6) by using a Dirac delta functon $\delta(\mathbf{p} - 0)$, or any positive $L^1$-normalized function in the variable $\mathbf{p}$, which eliminates the $\int d\mathbf{p}$ from (5.5)-(5.6) in a regular way; this trick is rather artificial, and hence not very elegant, but it provides a unified treatment of functionals.

The Poisson bracket is non-canonical, it contains the coupling between the plasma and the electromagnetic field

$$\begin{aligned} \{F, G\} = \int d\mathbf{z} \; \int d\mathbf{x} \; \delta(\mathbf{x} - \mathbf{q}) f \Big\{ [F_f, G_f] + \Big( \partial_{\mathbf{p}} F_f \cdot G_{\mathbf{E}} - \partial_{\mathbf{p}} G_f \cdot F_{\mathbf{E}} \Big) \Big\} \\ + \int d\mathbf{x} \; \Big( F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} - G_{\mathbf{E}} \cdot \nabla \times F_{\mathbf{B}} \Big) \,, \end{aligned} \tag{5.7}$$

where [ , ] is the Poisson bracket of particle dynamics

$$[g,h] := \nabla f \cdot \partial_{\mathbf{p}} g - \partial_{\mathbf{p}} f \cdot \nabla g + \mathbf{B} \cdot \partial_{\mathbf{p}} f \times \partial_{\mathbf{p}} g \tag{5.8}$$

for any functions $f(\mathbf{z}), g(\mathbf{z})$.

The transformation for particle coordinates (5.2) induces a transformation for functions and fields

$$\begin{aligned} f(\mathbf{z}) &\longrightarrow T^{-1}[f](\bar{\mathbf{z}}) := \bar{f}(\bar{\mathbf{z}}) := f(\tau^{-1}(\bar{\mathbf{z}})) = f(\mathbf{z})\,, \\ \mathbf{E}(\mathbf{x}) &\longrightarrow T^{-1}[\mathbf{E}](\bar{\mathbf{x}}) := \bar{\mathbf{E}}(\bar{\mathbf{x}}) := \mathbf{E}(\tau^{-1}(\bar{\mathbf{x}})) = \mathbf{E}(\mathbf{x})\,, \\ \mathbf{B}(\mathbf{x}) &\longrightarrow T^{-1}[\mathbf{B}](\bar{\mathbf{x}}) := \bar{\mathbf{B}}(\bar{\mathbf{x}}) := \mathbf{B}(\tau^{-1}(\bar{\mathbf{x}})) = \mathbf{B}(\mathbf{x})\,. \end{aligned} \tag{5.9}$$

In turn, it induces a transformation for functionals

$$\begin{aligned} F[f,\mathbf{E},\mathbf{B}] \longrightarrow \big(\mathbb{T}F\big)[\bar{f},\bar{\mathbf{E}},\bar{\mathbf{B}}] &:= \bar{F}[\bar{f},\bar{\mathbf{E}},\bar{\mathbf{B}}] \\ &:= F[T(\bar{f},\bar{\mathbf{E}},\bar{\mathbf{B}})] = F[f,\mathbf{E},\mathbf{B}]\,. \end{aligned} \tag{5.10}$$

To obtain the transformed Hamiltonian structure, first the Hamiltonian functional (5.4) just transforms as any functional

$$\begin{aligned} \bar{H}[\bar{f},\bar{\mathbf{E}},\bar{\mathbf{B}}] = H[f,\mathbf{E},\mathbf{B}] &= \int d\mathbf{z}\ f\tfrac{\mathbf{p}^2}{2} + \int d\mathbf{x}\ \tfrac{\mathbf{E}^2+\mathbf{B}^2}{2} \\ &= \int d\bar{\mathbf{z}}\ \mathcal{J}\bar{f}\tfrac{\mathbf{p}^2[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}]}{2} + \int d\bar{\mathbf{x}}\ \tfrac{\bar{\mathbf{E}}^2+\bar{\mathbf{B}}^2}{2}\,, \end{aligned} \tag{5.11}$$

where $\mathcal{J}$ is the Jacobian determinant of the transformation $\mathbf{z} \longrightarrow \bar{\mathbf{z}}$, and we used that the Jacobian of $\mathbf{x} \longrightarrow \bar{\mathbf{x}}$ is 1.

For a more general functional (5.5), the transformation would be

$$\begin{aligned} \bar{G}[\bar{\psi}] = G[\psi] &= \int d\mathbf{z}d\mathbf{x}\ g[\mathbf{z},\mathbf{x};f,\mathbf{E},\mathbf{B}] \\ &= \int d\bar{\mathbf{z}}d\bar{\mathbf{x}}\mathcal{J}\ g[\mathbf{z}[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}],\bar{\mathbf{x}};\bar{f},\bar{\mathbf{E}},\bar{\mathbf{B}}]\,, \end{aligned}$$

where $\psi := (f,\mathbf{E},\mathbf{B})$ is a vectorial shorthand for all the fields. And physical functionals (5.6) transform as

$$\begin{aligned} \bar{G}[\bar{\psi}] = G_{phys}[\psi] &= \int d\mathbf{z}\ g_{phys}[\mathbf{z};f,\mathbf{E},\mathbf{B}] \\ &= \int d\bar{\mathbf{z}}\mathcal{J}\ g_{phys}[\mathbf{z}[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}];\bar{f},\bar{\mathbf{E}},\bar{\mathbf{B}}]\,. \end{aligned}$$

The Poisson bracket (5.7) transforms as a functional, but functional derivatives have to be transformed as well

$$\begin{aligned} \bar{\{}\bar{F},\bar{G}\bar{\}} = \mathbb{T}\{\mathbb{T}^{-1}\bar{F},\mathbb{T}^{-1}\bar{G}\} = \int d\bar{\mathbf{x}}\ \int d\bar{\mathbf{z}}\ \mathcal{J}\ \delta(\bar{\mathbf{x}}-\mathbf{q}[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}])\ \bar{f} \\ \Big\{\bar{\Big[}T^{-1}\tfrac{\delta}{\delta f}\mathbb{T}^{-1}\bar{F}, T^{-1}\tfrac{\delta}{\delta f}\mathbb{T}^{-1}\bar{G}\bar{\Big]} \\ + \left(T^{-1}\,\partial_{\mathbf{p}}\,T\right)\left(T^{-1}\tfrac{\delta}{\delta f}\mathbb{T}^{-1}\bar{F}\right)\cdot T^{-1}\tfrac{\delta}{\delta\mathbf{E}}\mathbb{T}^{-1}\bar{G} \\ - \left(T^{-1}\,\partial_{\mathbf{p}}\,T\right)\left(T^{-1}\tfrac{\delta}{\delta f}\mathbb{T}^{-1}\bar{G}\right)\cdot T^{-1}\tfrac{\delta}{\delta\mathbf{E}}\mathbb{T}^{-1}\bar{F}\Big\} \\ + \int d\bar{\mathbf{x}}\ \Big\{\left(T^{-1}\tfrac{\delta}{\delta\mathbf{E}}\mathbb{T}^{-1}\bar{F}\right)\cdot\bar{\nabla}\times\left(T^{-1}\tfrac{\delta}{\delta\mathbf{E}}\mathbb{T}^{-1}\bar{G}\right) \\ - \left(T^{-1}\tfrac{\delta}{\delta\mathbf{E}}\mathbb{T}^{-1}\bar{G}\right)\cdot\bar{\nabla}\times\left(T^{-1}\tfrac{\delta}{\delta\mathbf{E}}\mathbb{T}^{-1}\bar{F}\right)\Big\}\,, \end{aligned} \tag{5.12}$$

where $\bar{[\ ,\ ]}$ is the transformed Poisson bracket of particle dynamics:

$$\begin{aligned}\bar{[}\bar{g},\bar{h}\bar{]} := &\left(T^{-1}\,\nabla\, T\bar{g}\right)\cdot\left(T^{-1}\,\partial_{\mathbf{p}}\, T\bar{h}\right)\\ &-\left(T^{-1}\,\partial_{\mathbf{p}}\, T\bar{g}\right)\cdot\left(T^{-1}\,\nabla\, T\bar{h}\right)\\ &+\mathbf{B}\cdot\left(T^{-1}\,\partial_{\mathbf{p}}\, T\bar{g}\right)\times\left(T^{-1}\,\partial_{\mathbf{p}}\, T\bar{h}\right)\end{aligned} \tag{5.13}$$

for any functions $\bar{g}(\bar{\mathbf{z}})$ and $\bar{h}(\bar{\mathbf{z}})$.

Transformed function derivatives are obtained through the usual chain rule:

$$\begin{aligned}T^{-1}\,\partial_{\mathbf{p}}\, T &= \partial_{\mathbf{p}}\bar{\mathbf{z}}\cdot\partial_{\bar{\mathbf{z}}}\\ T^{-1}\,\nabla\, T &= \nabla\bar{\mathbf{z}}\cdot\partial_{\bar{\mathbf{z}}}\end{aligned} \tag{5.14}$$

Functional derivatives are transformed using the method of Chapter 11, and using a lemma analogous to (11.25), which here gives the formula

$$\begin{aligned}0 &= \int d\mathbf{z}\ \left[\bar{g}\mathcal{J}_{\mathbf{E}}^{-1}+\mathcal{J}^{-1}\bar{g}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}+\mathcal{J}^{-1}\bar{g}_{\bar{f}}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}\right]\delta(\mathbf{q}-\mathbf{x})\\ &= \int d\mathbf{p}\ \left[\bar{g}\mathcal{J}_{\mathbf{E}}^{-1}+\mathcal{J}^{-1}\bar{g}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}+\mathcal{J}^{-1}\bar{g}_{\bar{f}}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}\right]^{\dagger}\ ,\end{aligned}$$

for any function $\bar{g}(\bar{\mathbf{z}})$, and the same formula but replacing the electric field $\mathbf{E}$ by the magnetic field $\mathbf{B}$.

Then transformed functional derivatives are found equal to the chain-rule formula obtained by studying total derivation of the defining equation (5.10) for the transformed fields

$$\begin{aligned}F_f &= T^{-1}\tfrac{\delta}{\delta\bar{f}}\mathbb{T}^{-1}\bar{F} = \mathcal{J}^{-1}\bar{F}_{\bar{f}}\,,\\ F_{\mathbf{E}} &= T^{-1}\tfrac{\delta}{\delta\bar{\mathbf{E}}}\mathbb{T}^{-1}\bar{F} = \bar{F}_{\bar{\mathbf{E}}}-\int d\mathbf{z}\mathcal{J}^{-1}\bar{F}_{\bar{f}}\left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}\right)\ \delta\Big(\mathbf{q}-\mathbf{x}\Big)\\ &= \bar{F}_{\bar{\mathbf{E}}}-\int d\mathbf{p}\ \left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}\right)^{\dagger}\left(\mathcal{J}^{-1}\bar{F}_{\bar{f}}\right),\\ F_{\mathbf{B}} &= T^{-1}\tfrac{\delta}{\delta\bar{\mathbf{B}}}\mathbb{T}^{-1}\bar{F} = \bar{F}_{\bar{\mathbf{B}}}-\int d\mathbf{z}\mathcal{J}^{-1}\bar{F}_{\bar{f}}\left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{B}}\right)\ \delta\Big(\mathbf{q}-\mathbf{x}\Big)\\ &= \bar{F}_{\bar{\mathbf{B}}}-\int d\mathbf{p}\ \left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{B}}\right)^{\dagger}\left(\mathcal{J}^{-1}\bar{F}_{\bar{f}}\right),\end{aligned} \tag{5.15}$$

where we insist that $\bar{\mathbf{z}}_{\mathbf{E}}$ and $\bar{\mathbf{z}}_{\mathbf{B}}$ are operators in general, and the adjoint $\dagger$ is with respect to $d\mathbf{z}$, which is equivalent to $d\mathbf{q}$ in these formulae.

Compared to the initial Hamiltonian structure, the transformed Poisson bracket retains all the terms of the initial bracket, as shown by the first term in each of the formulae (5.15). However, they are written in the transformed coordinates, as in Eqs. (5.12)and (5.14). In addition, the transformed Poisson bracket has a few (exactly eight) additional terms, coming from the last term in the third and fifth lines of (5.15). More precisely, it involves two additional self-couplings for the Vlasov density $\bar{f}$, an additional coupling between the Vlasov density $\bar{f}$ and the electric field $\bar{\mathbf{E}}$, and more importantly it shows up a coupling between the Vlasov density $\bar{f}$ and the magnetic field $\bar{\mathbf{B}}$ (each of these four couplings appears twice, as a result of the antisymmetry of the bracket).

Notice that in the transformed dynamics, not all of these additional terms will contribute. For instance, at first glance, it could seem that the second term in the fifth line will generate a magnetization term for the transformed Maxwell-Ampère equation of motion, but it is not the case. This magnetization-like term will be cancelled by terms coming from additional dependences on the fields in the Hamiltonian functional, which appear in Eq. (5.11). This is shown by a computation similar to Eq; (4.42) in Chapter 4. It is a general feature, evidenced in the derivation leading to Eq. (11.28) in Chapter 11, which implies that at least half of the additional terms in the Poisson bracket will cancel as a result of property (11.27).

This phenomenon explains why the transformed Poisson bracket, given by Eqs. (5.12) and (5.15), was very difficult (and probably impossible) to guess just by considering the gyrokinetic

equations of motion, since all the induced cancellations could hardly be predicted from the equations of motions.

Indeed, together, the transformed Hamiltonian functional (5.11) and Poisson bracket (5.12) with Eq. (5.15) for the transformed functional derivatives give the following transformed dynamics

$$\begin{aligned}
\dot{\bar{f}} &= -\dot{\bar{\mathbf{z}}} \cdot \partial_{\bar{\mathbf{z}}} \bar{f} + \bar{f}_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\mathbf{B}} \cdot \bar{\nabla} \times \bar{\mathbf{E}} \\
&\qquad + \bar{f}_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\mathbf{E}} \cdot \Big( -\bar{\nabla} \times \bar{\mathbf{B}} + \int d\bar{\mathbf{z}}\ \mathcal{J} \bar{f}\ \tfrac{\mathbf{p}[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}]}{m}\ \delta\Big(\mathbf{q}[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}] - \bar{\mathbf{x}}\Big)\Big)\,, \\
\dot{\bar{\mathbf{E}}} &= \bar{\nabla} \times \bar{\mathbf{B}} - \int d\bar{\mathbf{z}}\ \mathcal{J} \bar{f}\ \tfrac{\mathbf{p}[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}]}{m}\ \delta\Big(\mathbf{q}[\bar{\mathbf{z}};\bar{\mathbf{E}},\bar{\mathbf{B}}] - \bar{\mathbf{x}}\Big)\,, \\
\dot{\bar{\mathbf{B}}} &= -\bar{\nabla} \times \bar{\mathbf{E}}\,,
\end{aligned} \tag{5.16}$$

where we remind that the phase space $\bar{\mathbf{z}}$ is considered as fixed (as usual in field theory). The symbol $\dot{\bar{\mathbf{z}}}$ is a shorthand for the usual particle dynamics, corresponding to the Lorentz force, but expressed in the reduced coordinates $\dot{\bar{\mathbf{z}}} = \dot{\mathbf{z}} \cdot d_{\mathbf{z}} \bar{\mathbf{z}}$.

The transformed dynamics (5.16) confirms that no magnetization is induced by the lifting itself, as is clear by the case $\bar{\mathbf{q}} = \bar{\mathbf{x}}$, i.e. when the transformation does not change the particle position $\bar{\mathbf{q}} = \mathbf{q}$. This was expected from the results of Chapter 4. Thus, the magnetization-like term in the Poisson bracket mentioned above was indeed cancelled. From the eight additional terms in the transformed Poisson bracket, only three terms were generated in the transformed equations of motion, and all of them concern the dynamics of the Vlasov density.

## 5.3 Interpretation of the results

To interpret this dynamics in the light of (11.28), we use the property

$$\bar{f}_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\mathbf{E}} = -f_{\mathbf{z}} \cdot \mathbf{z}_{\bar{\mathbf{E}}}\,, \tag{5.17}$$

which comes because when transforming $\mathbf{z}$ forwards and then backwards, the result $\mathbf{z}[\bar{\mathbf{z}}[\mathbf{z};\mathbf{E},\mathbf{B}];\bar{\mathbf{E}},\bar{\mathbf{B}}] = \mathbf{z}$ (or equivalently $\tau^{-1} \circ \tau = 1$) is independent of $\mathbf{E} = \bar{\mathbf{E}}$. The same property as (5.17) holds when replacing $\mathbf{E}$ by $\mathbf{B}$.

With the property (5.17), Eq. (5.16) can be rewritten

$$\begin{aligned}
\dot{\bar{f}} &= \dot{f} + f_{\mathbf{z}} \cdot \mathbf{z}_{\bar{\mathbf{E}}} \cdot \dot{\bar{\mathbf{E}}} + f_{\mathbf{z}} \cdot \mathbf{z}_{\bar{\mathbf{B}}} \cdot \dot{\bar{\mathbf{B}}}\,, \\
\dot{\bar{\mathbf{E}}} &= \dot{\mathbf{E}}\,, \\
\dot{\bar{\mathbf{B}}} &= \dot{\mathbf{B}}\,.
\end{aligned} \tag{5.18}$$

So, the transformed dynamics just results from the initial dynamics through a chain-rule formula, which expresses that the electromagnetic field is unchanged by the transformation, whereas the Vlasov density $f$ is changed by the coordinate transformation (5.2) on $\mathbf{z}$, which is implicitly time-dependent when the electro-magnetic field is evolving in time. A posteriori it is obvious that no magnetization was to appear in Eq. (5.16) because the coordinate transformation does not change the dynamics of the electric field nor the current term in Maxwell-Ampère equation.

All the same, the presence of the Dirac delta term in Eq. (5.16) has consequences on how the current is computed. In order to express the current in a way well suited to the transformed system, it is convenient to extract from the current contribution the "reduced" current $\int d\bar{\mathbf{z}}\ \mathcal{J} \bar{f}\ \frac{\mathbf{p}}{m}$ evaluated at $\bar{\mathbf{q}} = \bar{\mathbf{x}}$, then the remaining contribution introduces polarization and magnetization[1]. This completely agrees with the example of the gyrokinetic dynamics.

[1] Here we will not be concerned with details about polarization and magnetization [17,18,77]. For completeness, let us remind that the polarization $\mathbf{P}$ corresponds to the contribution to the charge density $\rho := \int d\mathbf{p} f$ which is not included in the reduced charge density $\bar{\rho} := \int d\bar{\mathbf{p}} \mathcal{J} \bar{f}$, through the relation $-\nabla \cdot \mathbf{P} := \rho_{pol} := \rho - \bar{\rho}$. It induces a contribution in the current $\mathbf{J}_{pol} := \partial_t \mathbf{P}$. The magnetization corresponds to the contribution to the current $\mathbf{J} := \int d\mathbf{p} f \mathbf{p}$ that is not included in the reduced current $\bar{\mathbf{J}} := \int d\bar{\mathbf{p}} \mathcal{J} \bar{f} \bar{\mathbf{p}}$ nor in the polarization-induced current $\mathbf{J}_{pol}$, through the relation $\nabla \times \mathbf{M} := \mathbf{J}_{mag} := \mathbf{J} - \bar{\mathbf{J}} - \mathbf{J}_{pol}$.

Thus, a non-zero magnetization originates from the difference between the transformed particle point $\bar{\mathbf{q}}$ and the transformed field point $\bar{\mathbf{x}}$, which means from the presence of two points physically identical $\mathbf{x} = \mathbf{q}$ but transformed in different ways $\bar{\mathbf{x}} \neq \bar{\mathbf{q}}$, and which perfectly agrees with the usual picture of the Maxwell dynamics in non-vacuum media [77]. Also, it evidences that the appearance of two different coordinate spaces, as considered in Eq. (5.2), is needed to obtain such magnetization terms when lifting a change of particle coordinates to the field dynamics. It is why the work of Chapters 4 and 11 could not yield a non-zero magnetization.

Notice that even when the magnetization is zero, the transformed dynamics does not have the same expression as the initial dynamics, because additional terms come from the (implicit) time dependence in the transformation.

Before the transformation, the Jacobi identity for the Vlasov-Maxwell bracket is induced by the Jacobi identity for the particle bracket (5.8), and is thus submitted to the condition $\nabla \cdot \mathbf{B} = 0$ [104, 114]. The transformed bracket shares this property, since it is just given by a change of coordinates for the Vlasov-Maxwell bracket. This can also be shown directly:

$$\begin{aligned}
\sum_{cycl} \bar{\{}\bar{F}, \bar{\{}\bar{G}, \bar{H}\bar{\}}\bar{\}} &= \sum_{cycl} \mathbb{T}\{\mathbb{T}^{-1}\bar{F}, \mathbb{T}^{-1}\mathbb{T}\{\mathbb{T}^{-1}\bar{G}, \mathbb{T}^{-1}\bar{H}\}\} \\
&= \mathbb{T}\Big[\sum_{cycl}\{F, \{G,H\}_V + \{G,H\}_O\}_V \\
&\qquad\qquad + \sum_{cycl}\{F, \{G,H\}_V + \{G,H\}_O\}_O\Big] \\
&= \mathbb{T}\sum_{cycl}\{F, \{G,H\}_V\}_V = \mathbb{T}\int d\mathbf{z}\; f \sum_{cycl}[F_f, [G_f, H_f]] \\
&= \mathbb{T}\int d\mathbf{z}\; f\; \nabla\cdot\mathbf{B}\; \partial_{\mathbf{p}}F_f \times \partial_{\mathbf{p}}G_f \cdot \partial_{\mathbf{p}}H_f = 0\,,
\end{aligned}$$

where the symbol $F$ is a shorthand for $\mathbb{T}^{-1}\bar{F}$, and the same for $G$ and $H$. The index $V$ (resp. $O$) indicates the Vlasov part (resp. all other terms) of the Vlasov-Maxwell bracket.

The transformed Poisson bracket completely agrees with the results of [114]. This paper obtained the Hamiltonian structure of a class of transformed Vlasov-Maxwell dynamics by studying the expected motion (and especially Maxwell's equations in non-vacuum media), guessing a suitable form for the Hamiltonian functional and for the Poisson bracket, and then adding a closure assumption, which provided the final bracket through a functional chain rule. Especially, polarization and magnetization were induced by the chosen closure assumption.

On the other hand, in the present chapter, the transformed Hamiltonian functional and Poisson bracket are obtained just from the ones of the Vlasov-Maxwell system by a change of coordinates on the particle phase space. So, polarization and magnetization are just induced by the coordinate transformation, they do not come from a closure assumption.

The difference between these approaches is illustrated by the Hamiltonian functional: it is taken as arbitrary even in the initial field coordinates $(f, \mathbf{E}, \mathbf{B})$ and particle coordinates $(\mathbf{q}, \mathbf{v})$ in Eqs. (10) and (20) of [114]. On the contrary, in our approach, it is exactly given by the initial Hamiltonian functional of the Vlasov-Maxwell dynamics. Only when the coordinates $\bar{\mathbf{z}}$ and fields $(\bar{f}, \bar{\mathbf{E}}, \bar{\mathbf{B}})$ are changed, the Hamiltonian functional is changed and its change is exactly induced by the transformation. A similar difference in point of view can be observed in the Poisson bracket.

So, the work [114] rather performed a phenomenological study of Hamiltonian transformed Vlasov-Maxwell systems, whereas here we rather develop an ab-initio study of how a transformation of particle coordinates induces a transformed Hamiltonian Vlasov-Maxwell dynamics. These are the two complementary approaches of the electromagnetism in non-vacuum media, respectively bottom-up and top-down, and the coherence of the results is a mutual confirmation.

## 5.4 Case of the guiding-center transformation

The previous sections considered a general transformation $\bar{\mathbf{z}}[\mathbf{z}, \mathbf{E}, \mathbf{B}]$. We now specify to the transformations considered by gyrokinetics, the guiding-center and gyro-center transformation, mainly derived in Chapter 2. Here, the goal is not to compute explicitly the transformed Poisson bracket at lowest orders. This would require quite elaborated computations and will be explored later. Before doing so, the goal is to identify the main features of the transformed system (5.16), and especially to compare it with previous works in the literature.

The lifting was studied above for the Vlasov-Maxwell system, but in the case of electrostatic gyrokinetics, it can be applied the same way to the Vlasov-Poisson system, with similar results: only the derivatives with respect to the magnetic field are absent from the theory. So, in the present section we do not distinguish between these two systems.

### 5.4.1 Agreements with standard gyrokinetics

Transformed[2] Maxwell's equations (5.16) are exactly identical to previous work [17]. It is not a surprise since they are not affected by the transformation. Accordingly, when expressing the current term as a function of the transformed current $\int d\bar{\mathbf{z}}\ \mathcal{J}\bar{f}\ \frac{\mathbf{p}}{m}$ evaluated at $\bar{\mathbf{q}} = \bar{\mathbf{x}}$, the induced polarization and magnetization terms will agree with the literature [17]. It is useless to compute them here.

As for Vlasov's equation, the usual term $-\dot{\bar{\mathbf{z}}}\cdot\partial_{\bar{\mathbf{z}}}\bar{f}$ is recovered. In the case of electrostatic gyrokinetics, the corresponding characteristics are given by the transformed particle Hamiltonian dynamics:

$$-\dot{\bar{\mathbf{z}}}\cdot\partial_{\bar{\mathbf{z}}}\bar{f} = -\bar{[}\bar{f}, \bar{h}\bar{]}\,, \tag{5.19}$$

where $\bar{[}\cdot,\cdot\bar{]}$ is the transformed particle Poisson bracket (5.13), and $\bar{h}$ is the transformed particle Hamiltonian function

$$\bar{h} := T^{-1}(h) := T^{-1}\big(\tfrac{\mathbf{p}^2}{2} + \Phi\big)\,,$$

with $\Phi := -\Delta^{-1}\nabla\cdot\mathbf{E}$ the electric potential.

In the case of Vlasov-Maxwell, the same comments apply, except that the solenoidal part of the electric field implies that particle dynamics is not Hamiltonian, in the sense that it cannot be written under the form (5.19). Especially, there is no conserved mechanical (kinetic+potential) energy, because the force does not derive from a potential. Thus in this case, the system (plasma+Maxwell) is still more intrinsically a coupled system, only the sum is Hamiltonian, but particle dynamics is not Hamiltonian by itself.

### 5.4.2 Difference with standard gyrokinetics

All the same, in Eqs. (5.16), the transformed Vlasov equation is different from previous results, because of the two additional terms $f_{\mathbf{z}}\cdot\mathbf{z}_{\mathbf{E}}\cdot\dot{\mathbf{E}} + f_{\mathbf{z}}\cdot\mathbf{z}_{\mathbf{B}}\cdot\dot{\mathbf{B}}$ in Eq. (5.18), the ones that are caused by the time dependence of the transformation. So, either our derivation is mistaken or previous derivations miss some effects. In the rest of the section, we will analyse in details this difference, because the conclusion will mean that some aspects of the gyrokinetic reduction should be rectified. We want to make the reasons and the consequences clear for readers.

In order to conclude about the additional terms in the Vlasov equation, one can consider an arbitrary function of the initial phase space, but it is probably more enlightening to use a simple

[2] In principle, there is a distinction between a transformation and a reduction. The former changes things (e.g. coordinates) but loses no information on the system, whereas the latter extracts a reduced system, which loses some information, considered as non-essential. For instance, in guiding-center theory, the change of coordinates $\mathbf{z} \longrightarrow \bar{\mathbf{z}}$ is a transformation, which gives to the dynamics a simplified form but loses no information. It is not a true reduction. Thus $\bar{\mathbf{z}}$ is not a reduced system, but rather just a transformed system. The reduced system is $(\bar{\mathbf{q}}, \bar{\varphi})$, because the reduction occurs when extracting the important part of the dynamics, which is given by the slow 4-dimensional motion $(\dot{\bar{\mathbf{q}}}, \dot{\bar{\varphi}})$. Often the distinction is not maintained and the transformed system (e.g. $\bar{\mathbf{z}}$) is considered as a reduced system, because it has much simplified the system, and made it ready for the true reduction.

explicit function, such as $g[\mathbf{z},\mathbf{E}] := \mathbf{E}(\mathbf{q})\cdot\mathbf{p}$. It is related to the time evolution of the kinetic energy density, but we are not interested in its physical meaning: it is used only for illustrative purpose.

In order to make all the time dependences explicit, one should remind that $(\mathbf{p}(t),\mathbf{q}(t))$ are the particle's coordinates, which evolve in time, but that when $\mathbf{q}$ is fixed, the electric field all the same evolves in time, since it is a dynamical field. It has both an implicit and an explicit time dependence $\mathbf{E}=\mathbf{E}(\mathbf{q}(t),t)$. Then the time evolution of $g$ is

$$\dot{g} = \mathbf{E}\cdot\dot{\mathbf{p}} + \dot{\mathbf{q}}\cdot\nabla\mathbf{E}\cdot\mathbf{p} + \partial_t\mathbf{E}\cdot\mathbf{p} = \dot{\mathbf{z}}\cdot\partial_{\mathbf{z}}g + g_{\mathbf{E}}\cdot\dot{\mathbf{E}} = [g,h] + g_{\mathbf{E}}\cdot\dot{\mathbf{E}}\,, \tag{5.20}$$

where for particle variables $\mathbf{z}=(\mathbf{p},\mathbf{q})$, the dot refers to the evolution caused by particle dynamics, whereas for the field variables, the dot refers to the time evolution at fixed position[3] , i.e. the explicit time dependence $\dot{\psi}(\mathbf{z}(t),t) = \partial_t\psi = \frac{d}{dt}\psi - \dot{\mathbf{z}}\cdot\partial_{\mathbf{z}}\psi$. Eq. (5.20) confirms that, with regard to the time evolution, a field-dependent function must not be considered as just a function, but one should take into account the time evolution of the fields, which are not just functions, but explicitly time-dependent functions.

A similar analysis can be applied to the reduced Vlasov density $\bar{f}(\bar{\mathbf{z}}) = f(\mathbf{z}[\bar{\mathbf{z}},\mathbf{B}(t),\mathbf{E}(t)])$, where $\bar{\mathbf{z}}$ is not the particle position but a fixed coordinate for the base space of the field $\bar{f}$. The result writes

$$\frac{d}{dt}\bar{f}(\bar{\mathbf{z}}) = \dot{f}(\mathbf{z}) + f_{\mathbf{z}}\cdot\frac{d\mathbf{z}[\bar{\mathbf{z}},\bar{\mathbf{B}}(t),\bar{\mathbf{E}}(t)]}{dt} = -f_{\mathbf{z}}\cdot\dot{\mathbf{z}} + f_{\mathbf{z}}\cdot(\mathbf{z}_{\mathbf{E}}\cdot\dot{\mathbf{E}} + \mathbf{z}_{\mathbf{B}}\cdot\dot{\mathbf{B}}) = -\bar{f}_{\bar{\mathbf{z}}}\cdot\dot{\bar{\mathbf{z}}} + f_{\mathbf{z}}\cdot(\mathbf{z}_{\mathbf{E}}\cdot\dot{\mathbf{E}} + \mathbf{z}_{\mathbf{B}}\cdot\dot{\mathbf{B}})\,, \tag{5.21}$$

where we follow the standard notation, and use $\dot{\mathbf{z}}$ or $\dot{\bar{\mathbf{z}}}$ only as a shorthand for the particle's Hamiltonian dynamics, as in (5.19), because we do not consider the coordinates as dynamical (see the previous footnote). Eq. (5.21) confirms the validity of Eq. (5.16).

Thus, our result is correct and previous works missed the time evolution induced by the field dependence of the transformation. The essential difference relies in the presence of a functional chain rule, not just a function chain rule. As a corollary, this conclusion explains in a more complete way why previous attempts to guess the Hamiltonian structure of gyrokinetics starting from the equations of motion failed: they could not find the answer since their equations of motion were incomplete. Especially, it implied an incoherence in the guessed Poisson bracket $\mathbb{J}$, since its right-hand part, which acts on the Hamiltonian, needed corrective terms to cancel the electromagnetic-field dependence in the kinetic energy, whereas its left-hand part, generating the Vlasov equation, needed no such corrective terms. This fact enhances the relevance of the lifting method in order to transfer the coordinate transformation to the Hamiltonian structure of Vlasov-Maxwell in a proper way. Also, it emphasizes the efficiency of Hamiltonian approaches, where the time evolution is encoded in an algebraic and geometric structure.

### 5.4.3 Consequences for particle dynamics

A first consequence is that the characteristics of the transformed Vlasov density, given by the vector $-\bar{[}\bar{\mathbf{z}},\bar{h}\bar{]} + \bar{\mathbf{z}}_{\mathbf{E}}\cdot\dot{\mathbf{E}} + \bar{\mathbf{z}}_{\mathbf{B}}\cdot\dot{\mathbf{B}}$, are not given by a pure Hamiltonian system[4]. The origin is that particle dynamics is actually Hamiltonian, but also time dependent. As soon as a function depends not only on the coordinates but also on the fields, the dynamics becomes:

$$g = [g,h] + \partial_t g \neq [g,h]\,. \tag{5.22}$$

As a result, when using the coordinates $(\mathbf{p},\mathbf{q})$, then the difference is not noticeable in particle dynamics $\dot{\mathbf{z}}$, which is Hamiltonian, but when working with the coordinates $\bar{\mathbf{z}}$, it is not the case.

[3] In standard field theory, the base space over which fields are defined is assumed fixed (Eulerian description). This means that in the Vlasov equation, the use of the symbol $\dot{\mathbf{z}}$ is misleading, and possibly incorrect if not properly understood. The coordinates $\mathbf{z}$ are assumed fix, hence $\dot{\mathbf{z}}=0$. But the Vlasov equation writes $\dot{f} = \mathbf{w}\cdot\partial_{\mathbf{z}}f$, and its characteristics correspond to the motion of particle dynamics $\mathbf{w} = -\partial_t\mathcal{Z}(\mathbf{z}_0,t)$, where $\mathcal{Z}$ is just the particle coordinates in phase space (which evolve in time ), and $\mathbf{z}_0$ is the initial condition. It is common to use the same symbol for the particle position-momentum $\mathcal{Z}$ and for the phase-space coordinates $\mathbf{z}$, but they are different objects.

[4] In such statements, remind that a Hamiltonian system is understood here as a system where the dynamics is given by a Poisson bracket and a Hamiltonian function according to Eq. (5.19)). It does not preclude the system to be implicitly time dependent, such as particle dynamics in the Vlasov-Poisson system, but it excludes systems where the dynamics is given by Eq. (5.22)), such as particle dynamics for the Vlasov-Maxwell system. With this definition, it is obvious that an explicitly time-dependent transformation makes a Hamiltonian system non-Hamiltonian.

Even the dynamics of the coordinates is not given just by $\bar{[\bar{\mathbf{z}}, \bar{h}]}$ any more. From an abstract point of view, it is no trouble, but for practical purposes, it is not good news. For instance for theoretical studies, as well as for numerical simulations, the presence of purely Hamiltonian characteristics was quite useful, e.g. for the conservation of the phase-space volume.

A second consequence is that the guiding-center and gyro-center transformations do not fulfil their main requirements. They do not provide the desired reduction. Indeed, they used the Hamiltonian structure of particle dynamics in order to make both the magnetic moment a constant of motion and the 4-dimensional reduced motion independent of the fast gyro-angle:

$$0 = [\bar{\mu}, h]\,, \tag{5.23}$$
$$0 = \partial_\theta([\bar{\mathbf{y}}, h])\,, \tag{5.24}$$

with $\bar{\mathbf{y}} := \bar{\mathbf{z}}^{1-4} = (\bar{\mathbf{q}}, \bar{\varphi})$. But this overlooked the time evolution induced by the field dependence. The reduced dynamics actually verifies the relations

$$\begin{aligned} \tfrac{d\bar{\mu}}{dt} &= \partial_t \bar{\mu} = \bar{\mu}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \bar{\mu}_{\mathbf{B}} \cdot \dot{\mathbf{B}} \neq 0\,, \\ \partial_\theta\big(\tfrac{d\bar{\mathbf{y}}}{dt}\big) &= \partial_\theta\big(\partial_t \bar{\mathbf{y}}\big) = \partial_\theta(\bar{\mathbf{y}}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \bar{\mathbf{y}}_{\mathbf{B}} \cdot \dot{\mathbf{B}}) \neq 0\,. \end{aligned}$$

Thus the magnetic moment is not conserved and the decoupling of the fast time-scale is not obtained.

All the same, the spoilage is small, since the change of coordinates $\bar{\mathbf{z}}[\mathbf{z}; \mathbf{E}, \mathbf{B}]$ is near identity. So, $\bar{\mathbf{z}}_{\mathbf{E}}$ and $\bar{\mathbf{z}}_{\mathbf{B}}$ are at least first-order quantities. This suggests to perform a third near-identity transformation in order to restore the reduction by taking into account the time dependence of the guiding- and gyro-center transformations. These last transformations relied on expansions in the small parameters $\varepsilon$ and $\varepsilon_E$, respectively related to the magnetic inhomogeneity at the scale of the Larmor radius and the electric force over the magnetic force. They can also be related to the time-scale ratios respectively between the particle velocity $\dot{\mathbf{q}} \cdot \nabla$ and the Larmor frequency, and between the electric acceleration $\mathbf{E} \cdot \partial_{\mathbf{p}}$ and the Larmor frequency. The expansion parameter for the third transformation $\varepsilon_t$ is related to the time-scale ratio between the electromagnetic field dynamics and the Larmor frequency.

This third transformation can be easily obtained when Lie-transforming the Lagrangian 1-form as in Chapter 2, since it works in the 7-dimensional space $(\mathbf{z}, t)$, and the method is geometric [87]. The time is a coordinate as much as the particle position and momentum. When the parameter $\varepsilon_t$ is small, the time dependence of the transformation will not have much impact on the method, since it will not affect the pivotal coefficients, and will only imply additional terms in the transformation. Then the desired requirements for gyrokinetics will be obtained:

$$\begin{aligned} \tfrac{d\bar{\mu}}{dt} &= [\bar{\mu}, h] + \bar{\mu}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \bar{\mu}_{\mathbf{B}} \cdot \dot{\mathbf{B}} = 0\,, \\ \partial_\theta\big(\tfrac{d\bar{\mathbf{y}}}{dt}\big) &= \partial_\theta([\bar{\mathbf{y}}, h]) + \partial_\theta(\bar{\mathbf{y}}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \bar{\mathbf{y}}_{\mathbf{B}} \cdot \dot{\mathbf{B}}) = 0\,. \end{aligned}$$

### 5.4.4 Consequences for the gyro-center transformation

The main consequence of the third transformation is that the additional terms generated by the time dependence will imply $\dot{\mathbf{E}} = \nabla \times \mathbf{B} - \mathbf{J}$, i.e. the current. At the next order, it will generate terms with $\dot{\mathbf{J}} = -\nabla . \int d\mathbf{p} f \mathbf{p} \otimes \mathbf{p} - \mathbf{B} \times \int d\mathbf{p} f \mathbf{p} + \mathbf{E} \int d\mathbf{p} f$, and iteratively all the moments of the distribution function will be involved in the transformation.

This conclusion is still clearer in the case of the Vlasov-Poisson system: its Hamiltonian structure is given in Sec. 7.4. Let us remind it here:

$$\begin{aligned} \{F, G\} &= \int d\mathbf{z}\, f [F_f, G_f] + \int d\mathbf{z}\, f \Big(\partial_{\mathbf{p}} F_f \cdot \nabla \Delta^{-1} \nabla \cdot G_{\mathbf{E}} - \partial_{\mathbf{p}} G_f \cdot \nabla \Delta^{-1} \nabla \cdot F_{\mathbf{E}}\Big)\,, \\ H &= \int d\mathbf{z}\, f \frac{\mathbf{p}^2}{2} + \int d\mathbf{q}\, \frac{\mathbf{E}^2}{2}\,, \end{aligned} \tag{5.25}$$

where the dynamical fields are $f$ and $\mathbf{E}$. Actually, the quantities $\nabla \times \mathbf{E}$ and $\nabla \cdot \mathbf{E}$ are not true independent variables, because they are related to the Casimir invariants (constants of motion) $\phi_1 := \nabla \times \mathbf{E}$ and $\phi_2 := \nabla \cdot \mathbf{E} - \int d\mathbf{p} f$. Their value is fixed by the initial conditions, and physically they are to be chosen zero, which implies $\mathbf{E} := -\nabla\Phi$ with $\Phi = -\Delta^{-1} \int d\mathbf{p} f$ the electric potential. Changing the field variables $(f, \mathbf{E}) \longrightarrow (f, \phi_1, \phi_2)$, the Poisson bracket becomes

$$\{F, G\} = \int d\mathbf{z}\, f [F_f, G_f] \,.$$

It is just the Vlasov part of the Poisson bracket (5.25). It emphasizes that only the Vlasov density is an independent dynamical variable. As for the Hamiltonian, it becomes

$$H = \int d\mathbf{z}\, f \Big( \frac{\mathbf{p}^2}{2} - \frac{1}{2} \Delta^{-1} \int d\mathbf{p}\, f \Big) \,.$$

Then, in the gyro-center transformation, when the electric field appears, it has to be replaced by the quantity $\nabla\Delta^{-1} \int d\mathbf{p} f$. Thus the gyro-center transformation explicitly depends on the zeroth-order moment of the Vlasov density. Since this last is dynamical, the transformation is time dependent and will depend also on $\dot{\mathbf{E}} = \frac{d}{dt} \nabla\Delta^{-1} \int d\mathbf{p} f = -\nabla\Delta^{-1}\nabla \cdot \int d\mathbf{p} f \frac{\mathbf{p}}{m} = -\nabla\Delta^{-1}\nabla \cdot \mathbf{J}$. At the next orders, the higher order moments of $f$ will be implied.

This strongly changes the framework, since the reduction of particle dynamics will depend not only on the electric and magnetic field, but also on the particle distribution itself. So, the transformation is not just a reduction of particle dynamics, with each particle being independently submitted to an electromagnetic field. Particles cannot be reduced independently of each other, as it was the case in previous works. Their dynamics are coupled with each other through the distribution function $f$. In some way, it is not a surprise, since the overall dynamics is not composed of independent particles, but it is fully coupled through the electromagnetic field. This fact gives a better view of the viewpoint adopted in previous works. It evidences that they did not really achieve the reduction of plasma dynamics, but implicitly considered that the electromagnetic field was external, not self-consistently generated by the plasma.

More precisely, previous works implemented only half of the coupling, as is easily seen by the following argument, formulated in the case of electrostatic gyrokinetics for simplicity (see the description of steps 2b-2c of the gyrokinetic reduction on page 16). Restoring the coupling does not mean just to take into account the presence of electric field in particle dynamics, because this is only a one-way coupling, with the action of the electric field on the plasma. For the coupling to be complete, the reverse coupling should be taken into account, with the feed-back of the plasma on the electric field, which is traduced by the electric field being not static but dynamical, with a dynamics depending on the plasma. It was neglected in previous works, as is confirmed by the absence of the terms with $\dot{\mathbf{E}}$ in the transformed Vlasov equation, and also by the absence of the moments of the Vlasov density in the gyro-center transformation, spoiling the guiding-center reduction.

As a result, the usual gyro-center transformation is actually a "*particle gyro-center transformation*", in which the electromagnetic field is external. The third transformation introduced above can be considered as a "*plasma gyro-center transformation*" (or "*Vlasov-Maxwell gyro-center transformation*"), because it really considers that the electromagnetic field is self-consistently generated by the plasma.

It was mentioned in the first episode (see also page 16) that the first two transformations can be derived simultaneously by taking directly into account the presence of a dynamical electric field in the Hamiltonian. Then the overall transformation is commonly called guiding-center transformation. The same way, for the third transformation, it would be more natural to perform it at the same time as the second one (and hence possibly at the same time as the first one), by taking directly into account the time dependence of the electric field. The overall transformation could hardly be called guiding-center transformation, since it is no more a reduction of particle dynamics in an electromagnetic field. It would be better called "*gyrokinetic transformation*", since it provides the ultimate coordinate transformation for the gyrokinetic reduction.

In order to take into account the third transformation, i.e. the time-dependence of the guiding- and gyro-center transformations, the Lie-transform of the Lagrangian of particle dynamics, described in Chapter 2 will have to be rewritten in the time dependent framework. In addition, the lifting considered in the previous sections will have to be extended in order to include a dependence of the moments of the distribution function in the coordinate transformation. There should not be a trouble, since the lifting method can be applied also in this case. In fact, it will make the framework closer to the general lifting considered in Chapter 11, since the transformation of a coordinate space $\mathbf{z}$ will depend on a field $f$ defined over this space. This was somehow avoided in the previous sections, because the transformation of $\mathbf{z} = (\mathbf{p}, \mathbf{q})$ depended only an the electromagnetic field, which is defined over the space $\mathbf{x}$, considered as different from the particle position space $\mathbf{q}$.

### 5.4.5 Complementary comments

The terms generated by the time dependence are expected to be small in the cases considered by gyrokinetics. In the Vlasov equation for instance, they will be of order $f_{\mathbf{z}} \cdot \mathbf{z}_{\mathbf{E}} \cdot \dot{\mathbf{E}}$, with $\dot{\mathbf{E}} = -\mathbf{J}$ in the electrostatic case (or also $-\nabla\dot{\Phi} = -\nabla\Delta^{-1}\nabla \cdot \mathbf{J}$), with $\Phi$ the electric potential. When studying lowest-order effects, they are negligible, but when including higher-order terms in the equations, then they might have non-negligible contributions. In addition, even when the terms are negligible in the equations of motion, removing them can spoil the Hamiltonian structure. In order to avoid such a difficulty, the removal should be considered by working at the level of the Hamiltonian structure, for instance by using methods developed in Episode 3.

To make things clearer for some readers, a last question is useful to address. It may seem that $\bar{f}$ has not the same advection equation as $f$, whereas the Lagrangian advection of the Vlasov equation had strong physical motivations: particles are just moving, and they are conserved. After a change of coordinates, this should also be true.

Let us consider first the argument on moving particles. After the transformation, particles are described by the guiding-center coordinates $\bar{\mathbf{z}}$, they are still moving, and they still induce the Vlasov density to be advected, since Eq. (5.21) can be rewritten

$$\tfrac{d}{dt}\bar{f}(\bar{\mathbf{z}}) = -\mathbf{w} \cdot \partial_{\bar{\mathbf{z}}}\bar{f}\,,$$

with

$$\mathbf{w} := \dot{\bar{\mathbf{z}}} + \bar{\mathbf{z}}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \bar{\mathbf{z}}_{\mathbf{B}} \cdot \dot{\mathbf{B}} = [\bar{\mathbf{z}}, h] + \partial_t\bar{\mathbf{z}} = \tfrac{d}{dt}\bar{\mathbf{z}}[\mathbf{z}; \mathbf{E}, \mathbf{B}]\,. \tag{5.26}$$

The only difference is that the advection does not follow the pure Hamiltonian dynamics $[\bar{\mathbf{z}}, h]$, but the explicitly time-dependent Hamiltonian dynamics $[\bar{\mathbf{z}}, h] + \partial_t\bar{\mathbf{z}}$.

With respect to the argument about conserved particles, the number of particles is not given by the 0-form (the function) $f$, but by the 6-form $f d\mathbf{z}$. After a change of coordinates, it becomes the 6-form $\bar{f}\mathcal{J}d\bar{\mathbf{z}}$, which corresponds to the function $\bar{\bar{f}} := \bar{f}\mathcal{J}$. This quantity is easily verified to be conserved, because

$$\begin{aligned}\tfrac{d}{dt}\big(\mathcal{J}\bar{f}\big) = \tfrac{d}{dt}\Big(\mathcal{J}[\bar{\mathbf{z}}; \mathbf{E}, \mathbf{B}]f\big(\mathbf{z}[\bar{\mathbf{z}}; \mathbf{E}, \mathbf{B}]\big)\Big) &= \bar{f}\mathcal{J}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \bar{f}\mathcal{J}_{\mathbf{B}} \cdot \dot{\mathbf{B}} + \mathcal{J}\dot{f} + \mathcal{J}f_{\mathbf{z}} \cdot \Big(\mathbf{z}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \mathbf{z}_{\mathbf{B}} \cdot \dot{\mathbf{B}}\Big) \\ &= -\partial_{\bar{\mathbf{z}}^i}\Big(\bar{\bar{f}}\bar{\mathbb{J}}^{ij}\partial_{\bar{\mathbf{z}}^j}\bar{h}\Big) - \partial_{\bar{\mathbf{z}}^i}\Big(\bar{\bar{f}}\,\big(\bar{\mathbf{z}}^i_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \bar{\mathbf{z}}^i_{\mathbf{B}} \cdot \dot{\mathbf{B}}\big)\Big)\,.\end{aligned} \tag{5.27}$$

To obtain the last equality, the divergence form of the particle Poisson bracket has been used, and the Jacobian has been transferred into the divergence by changing to transformed coordinates $\bar{\mathbf{z}}$:

$$\begin{aligned}\mathcal{J}\dot{f} &= -\mathcal{J}[f, h] = -\mathcal{J}\partial_{\mathbf{z}^i}f\ \mathbb{J}^{ij}\partial_{\mathbf{z}^j}h = -\mathcal{J}\partial_{\mathbf{z}^i}\Big(f\ \mathbb{J}^{ij}\partial_{\mathbf{z}^j}h\Big) \\ &= -\partial_{\bar{\mathbf{z}}^k}\Big(\bar{\mathbf{z}}^k_{\mathbf{z}^i}\mathcal{J}f\ \mathbb{J}^{ij}\bar{\mathbf{z}}^l_{\mathbf{z}^j}\partial_{\bar{\mathbf{z}}^l}\bar{h}\Big) = -\partial_{\bar{\mathbf{z}}^k}\Big(\bar{\bar{f}}\bar{\mathbb{J}}^{kl}\partial_{\bar{\mathbf{z}}^l}\bar{h}\Big)\,.\end{aligned}$$

In addition, the terms with $\dot{\mathbf{E}}$ or $\dot{\mathbf{B}}$ have been combined using formulae such as

$$\bar{f}\mathcal{J}_{\mathbf{E}} \cdot \dot{\mathbf{E}} + \mathcal{J}f_{\mathbf{z}} \cdot \mathbf{z}_{\mathbf{E}} \cdot \dot{\mathbf{E}} = -\partial_{\bar{\mathbf{z}}^i}\Big(\bar{\bar{f}}\ \bar{\mathbf{z}}^i_{\mathbf{E}} \cdot \dot{\mathbf{E}}\Big)\,,$$

which is shown in a similar way as Eq. (11.23).

In Eq. (5.27), the first term corresponds to an advection by a Hamiltonian system. The second term spoils this fact, but the density $\bar{\bar{f}}$ remains conserved. Comparing with Eq. (5.26), it appears that its flux corresponds to an advection by $\bar{\bar{f}}\frac{d\bar{\mathbf{z}}}{dt}$. So, as expected, the density $\bar{\bar{f}}$ follows particle dynamics, even if this last is not Hamiltonian any more.

The field coordinate $\bar{\bar{f}} := \mathcal{J}\bar{f}$ can be used in the Vlasov-Maxwell Hamiltonian structure instead of $\bar{f}$. Then a standard chain rule for field coordinates gives[5]:

$$\begin{aligned}
\bar{F}_{\bar{f}} &= \mathcal{J}\bar{\bar{F}}_{\bar{\bar{f}}}\,,\\
\bar{F}_{\mathbf{E}} &= \bar{\bar{F}}_{\mathbf{E}} + \int d\mathbf{z}\bar{\bar{F}}_{\bar{\bar{f}}}\left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}\right)\ \delta\Big(\mathbf{q}-\mathbf{x}\Big) + \int d\bar{\mathbf{z}}\bar{\bar{f}}\partial_{\bar{\mathbf{z}}}\Big(\bar{\bar{F}}_{\bar{\bar{f}}}\Big)\cdot\bar{\mathbf{z}}_{\mathbf{E}}\ \delta\Big(\mathbf{q}-\mathbf{x}\Big)\\
&= \bar{\bar{F}}_{\mathbf{E}} + \int d\mathbf{p}\ \left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}\right)^{\dagger}\Big(\bar{\bar{F}}_{\bar{\bar{f}}}\Big) + \int d\mathbf{p}\ (\bar{\mathbf{z}}^i_{\mathbf{E}})^{\dagger}\Big(\bar{f}\partial_{\bar{\mathbf{z}}^i}\bar{\bar{F}}_{\bar{\bar{f}}}\Big)\,,\\
\bar{F}_{\mathbf{B}} &= \bar{\bar{F}}_{\mathbf{B}} + \int d\mathbf{z}\bar{\bar{F}}_{\bar{\bar{f}}}\left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{B}}\right)\ \delta\Big(\mathbf{q}-\mathbf{x}\Big) + \int d\bar{\mathbf{z}}\bar{\bar{f}}\partial_{\bar{\mathbf{z}}}\Big(\bar{\bar{F}}_{\bar{\bar{f}}}\Big)\cdot\bar{\mathbf{z}}_{\mathbf{B}}\ \delta\Big(\mathbf{q}-\mathbf{x}\Big)\\
&= \bar{\bar{F}}_{\mathbf{B}} + \int d\mathbf{p}\ \left(\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{B}}\right)^{\dagger}\Big(\bar{\bar{F}}_{\bar{\bar{f}}}\Big) + \int d\mathbf{p}\ (\bar{\mathbf{z}}^i_{\mathbf{B}})^{\dagger}\Big(\bar{f}\partial_{\bar{\mathbf{z}}^i}\bar{\bar{F}}_{\bar{\bar{f}}}\Big)\,.
\end{aligned} \tag{5.28}$$

Thus compared to the transformed Poisson bracket (5.12), the change from $\bar{f}$ to $\bar{\bar{f}}$ exactly absorbs all the Jacobians and all the additional terms with $\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{E}}$ and $\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\mathbf{B}}$ coming from the field dependence of the transformation. This makes it easier to work on the transformed Hamiltonian structure, as will be confirmed in Chapter 8. All the same, another additional term remains $\int d\mathbf{p}\ (\bar{\mathbf{z}}^i_{\mathbf{E}})^{\dagger}\big(\bar{f}\partial_{\bar{\mathbf{z}}^i}\bar{\bar{F}}_{\bar{\bar{f}}}\big)$ or $\int d\mathbf{p}\ (\bar{\mathbf{z}}^i_{\mathbf{B}})^{\dagger}\big(\bar{f}\partial_{\bar{\mathbf{z}}^i}\bar{\bar{F}}_{\bar{\bar{f}}}\big)$. Their effect on the Hamiltonian functional is to cancel the action of the derivatives $\frac{\delta}{\delta\mathbf{E}}$ and $\frac{\delta}{\delta\mathbf{B}}$ on the particle kinetic energy $\frac{1}{2}\mathbf{p}^2[\bar{\mathbf{z}};\mathbf{E},\mathbf{B}]$. On the equations of motion, their effect is to generate the last term in the Vlasov equation (5.27), originating from the presence of the Jacobian and the time dependence of the transformation.

## Conclusion

For a lifting of a coordinate transformation to the Hamiltonian structure of the Vlasov-Maxwell dynamics, we have applied the method introduced in Chapters 4 and 11.

The presence of two different coordinate spaces changed the framework, but the general method could be applied. One of the spaces was physically embedded in the other whereas the two spaces do not undergo the same transformation. It implied to distinguish the electromagnetic field point from the particle position point. This appeared as essential to generate polarization and magnetization terms in the transformed motion, whereas the lifting itself did not generate any magnetization.

The results were first obtained for a general transformation, then applied to lift the guiding-center and gyro-center transformations to the Vlasov-Maxwell field dynamics. They confirmed that both transformations naturally induces a Hamiltonian structure for the resulting field dynamics.

However, this did not provide the expected reduction, and the failure was shown to come from the time evolution of the electromagnetic field. Actually, previous procedures had not implemented the coupling in a completely proper way. They had taken into account the action of the electromagnetic field over the plasma, but not completely the feed-back of the plasma on the electromagnetic field. This meant that additional terms were present in the resulting dynamics, which spoilt the expected reduction.

In order to restore the expected reduction, the dynamics of the electromagnetic field had to be taken into account in the gyro-center transformation, which deeply changed the framework of this reduction. Especially, the Hamiltonian structure of particle dynamics becomes time-dependent, and the Lie-transform of the phase-space Lagrangian must be time-dependent. A striking effect

[5] It is obtained after some algebra. It is easier to compute directly the change from $f$ to $\bar{\bar{f}}$ rather than following the sequence $f \longrightarrow \bar{f} \longrightarrow \bar{\bar{f}}$.

is that the characteristics of the transformed Vlasov equation are not given any more by just the Hamiltonian dynamics of the transformed particle Poisson bracket. In addition, the transformation must depend not only on the electromagnetic field, but also on the moments of the Vlasov density. This is not an artefact of our method, but it is needed for a proper account of the coupling between the plasma and the electromagnetic field. It will be further investigated in the next chapter.

For completeness of the theory, a necessary extension of the lifting considered in the present chapter will include in the coordinate transformation not only a dependence on the electromagnetic field, but also a dependence on the moments of the Vlasov density. This will change somehow the framework, but the extension should be straightforward, and it will make the scheme closer to the one of Chapter 11, where the transformation on a space depended on fields defined over this space.

On another hand, the Hamiltonian structure of the transformed dynamics did not completely answer the question of the Hamiltonian structure of gyrokinetics. Indeed, gyrokinetics is not just the transformed system (5.16), because in addition, it drops the dependence in the gyro-angle, which could spoil the Hamiltonian character. This will be the topic of the next episode, and especially of Chapter 8.

In the case of gyrokinetic with several species, or in case electron spin effects are taken into account [41,83,93], the method has to be extended to include the presence of several sets of fields, each one defined over its own coordinate space. This is considered in the appendix. This extension also makes the framework closer to transformations considered by particle-physics field theory [66,152]. Magnetization-like terms appear to be quite a general feature, even when no spaces are embedded in one another, because the transformation of one space generally depends on the fields defined over other spaces, which implies that integrals and Dirac delta functions remain in the transformed dynamics, Poisson brackets, and functional derivatives.

From a more general point of view, the results of this chapter interestingly explain why previous approaches and attempts at guessing the Hamiltonian structure of gyrokinetics by starting from the equations of motion had failed. On the one hand, in the relation between the transformed Poisson bracket and the associated equations of motions, many cancellations occur, which makes it very difficult to guess the Hamiltonian structure just by observing the equations of motion. On the other hand, since the equations of motion used in the literature missed some terms, the result was definitely impossible to obtain with such an approach.

So, the lifting method was crucial in order to derive the transformed Hamiltonian structure, to better understand the reduction process, and especially to identify the corrections needed for the transformed dynamics. Also, it confirms the relevance of the idea to revisit dynamical reductions in the framework of Hamiltonian reductions: instead of deriving models at the level of the equations of motion and then guessing the Hamiltonian structure (when feasible) or the corrections needed to recover the Hamiltonian character (hardly feasible), it is much more enlightening to work at the level of the Hamiltonian structure, to derive the reduced structure, and then to deduce (i.e. recover and confirm, or invalidate and rectify) the reduced equations of motion. This idea will be at the core of Episode 3.

## Appendix: Generalization to multiple coordinate spaces

In the previous sections, two sets of fields were involved, each one defined over its own "coordinate space", or "particle space"; a specific feature was that one of the spaces, $\mathbf{x}$, was naturally embedded in the other one through the relation $\mathbf{x} = \mathbf{q}$. In this section, we generalize the results to a general field theory, where the "particle coordinate space" involves several spaces, each field is defined over one of these spaces, and there is no specific embedding of one space into another one.

For instance, if the gyrokinetic approximation holds for electrons but not for ions, a convenient description of the corresponding plasma dynamics can rely on the electromagnetic field $(\mathbf{E}(\mathbf{q}), \mathbf{B}(\mathbf{q}))$, the ion density in phase space $f_i(\mathbf{q}, \mathbf{p})$, and the electron density in the guiding-center transformed space $\bar{f}_e(\bar{\mathbf{q}}, \bar{\varphi})$, where $\bar{\varphi}$ is the transformed pitch-angle.

More importantly, if the gyrokinetic approximation holds for ions and electrons, the transformed space for electrons is not the same as for ions, because the guiding-center or gyro-center transformation depends on the particle charge, for instance, which is species-dependent:

$$\bar{\mathbf{z}}[\mathbf{z}, q_j; \mathbf{E}, \mathbf{B}] \,.$$

Another example is when the spin effects of electrons are important, then the particle space can be chosen as the product of the position space $\mathbf{q}$, the momentum space $\mathbf{p}$ and the electron spin space $\mathbf{s}$ [114]. The fields can be the Vlasov density for ions $f_i(\mathbf{q}, \mathbf{p})$, the Vlasov density for electrons $f_e(\mathbf{q}, \mathbf{p}, \mathbf{s})$, and the electromagnetic field $(\mathbf{E}(\mathbf{q}), \mathbf{B}(\mathbf{q}))$. If the field-dependent particle-coordinate transformation concerns also the spin coordinate, then the results of this section apply.

Compared to previous sections, the presence of several coordinate spaces changes the framework, but the lifting can be derived in a similar way as in previous sections. A chain-rule formula is obtained by a total derivative of the defining equations (5.3), and transformed functional derivatives give the same chain-rule formula, provided a lemma analogous to (11.25) is used. Again, we do not give the details of the derivation, but just write the results. The only differences are that the setting is more abstract and that since there are several "particle spaces" over which fields are defined, care must be taken over which variable the operators act, and over which spaces integrals are taken or not; in addition, the lack of embedding between spaces implies to be careful both in the definitions to be chosen and in computations.

We consider a general field-theoretic Hamiltonian dynamical system, with a set of fields $\psi^\alpha(\mathbf{z}^\alpha)$, where each field $\psi^\alpha$ is defined over its own space $\mathbf{z}^\alpha$ and can be vectorial (i.e. composed of several scalar fields). This does not exclude a possible identification of several of these spaces by Dirac delta function $\delta(\mathbf{z}^\alpha - \mathbf{z}^{\alpha'})$ everywhere in the initial (non-transformed) theory, and especially in the Hamiltonian and in the Poisson bracket. This can be needed if several fields defined over the same space do not transform the same way, as it was the case for the Vlasov-Maxwell system. We will denote by $\mathbf{y} := (\{\mathbf{z}^\alpha\}_\alpha)$ a vector grouping together all the coordinate spaces, and $\phi := (\{\psi^\alpha\}_\alpha)$ a vector grouping together all the fields.

The Hamiltonian is given by

$$H = \int d\mathbf{y}\ h[\mathbf{y}; \phi] \,.$$

In a similar way as for the Vlasov-Maxwell system, the Hamiltonian may involve some terms involving no $\mathbf{z}^\alpha$ nor $\int d\mathbf{z}^\alpha$ for some value of $\alpha$. As in the previous section, such terms are included by using an artificial delta function $\delta(\mathbf{z}^\alpha - 0)$ or any positive $L^1$-normalized function $f(\mathbf{z}^\alpha)$, which absorbs the $\int d\mathbf{z}^\alpha$ in a regular way.

The Poisson bracket writes

$$\{F, G\} = \int d\mathbf{y}\ F_{\psi^\alpha} \mathbb{J}^{\alpha,\beta}[\mathbf{y}; \phi]\ G_{\psi^\beta} \,.$$

The coordinate transformation in the "particle" spaces $\mathbf{z}^\alpha$ is given by:

$$\mathbf{z}^\alpha \longrightarrow \tau_\alpha(\mathbf{z}^\alpha) := \bar{\mathbf{z}}^\alpha := \bar{\mathbf{z}}^\alpha[\mathbf{z}^\alpha; \phi] \,, \tag{5.29}$$

where the change of each coordinate $\mathbf{z}^\alpha$ can depend on all the fields, $\psi^{\{\gamma\}}$, even for $\gamma \neq \alpha$, as was the case for the Vlasov-Maxwell system. All the same, we assume each of the transformation $\tau_\alpha$ to be a change of coordinates for $\mathbf{z}^\alpha$, hence to be one-to-one, as is needed for a lifting, i.e. for the fields to transform through the scalar invariance property.

In addition, if $\tau_\alpha$ depended on $\mathbf{z}^\gamma$ with $\gamma \neq \alpha$, then generically a functional $F = \int d\mathbf{z}^\alpha f[\mathbf{z}^\alpha; \phi]$ would transform as

$$\int d\mathbf{z}^\alpha\ f[\mathbf{z}^\alpha; \phi] = \int d\bar{\mathbf{z}}^\alpha \mathcal{J}_\alpha\ f[\mathbf{z}[\bar{\mathbf{z}}^{\alpha,\gamma}; \bar{\phi}], \bar{\phi}] \,,$$

which would make the transformed functional position-dependent. It is why we assumed in Eq. (5.29) that the transformation $\tau_\alpha$ is independent of $\mathbf{z}^\gamma$ for $\gamma \neq \alpha$. In specific cases, this

requirement can be made less strong, because what is needed is only to introduce no position dependence for transformed functionals.

Last, for $\gamma \neq \alpha$, the presence of $\psi^\gamma$ but no $\mathbf{z}^\gamma$ may imply[6] the presence of a map $\pi^\gamma_\alpha(\mathbf{z}^\alpha) = \mathbf{z}^\gamma[\mathbf{z}^\alpha;\phi]$ between the spaces $\mathbf{z}^\alpha$ and $\mathbf{z}^\gamma$. For $\pi^\gamma_\alpha$, the $\alpha$ is just an index, which reminds that the map is different for different values of $\alpha$. Actually, for each value of $\gamma$ and $\alpha$, several different such functions $\pi^\gamma_\alpha$ could be involved. The maps $\pi$ explain and generalize the embedding $\mathbf{x} = \mathbf{q}$ used in the coordinate transformation of the Vlasov-Maxwell system, but there is no need for the $\pi$ to correspond to embeddings, nor to be related to the structure of the initial system, as was the case for the Vlasov-Maxwell system.

The coordinate transformation (5.29) induces a transformation for functions of $\mathbf{z}^\alpha$ and for the fields $\psi^\alpha$:

$$\psi^\alpha(\mathbf{z}^\alpha) \longrightarrow T^{-1}_\alpha[\psi^\alpha](\bar{\mathbf{z}}^\alpha) := \bar{\psi}^\alpha(\bar{\mathbf{z}}^\alpha) := \psi^\alpha(\tau^{-1}_\alpha(\bar{\mathbf{z}}^\alpha)) = \psi^\alpha(\mathbf{z}^\alpha)\,. \tag{5.30}$$

Last, it induces a transformation for functionals of $\psi$:

$$F[\{\psi^\alpha\}_\alpha] \longrightarrow (\mathbb{T}F)[\{\bar{\psi}^\alpha\}_\alpha] := \bar{F}[\{\bar{\psi}^\alpha\}_\alpha] := F[\{T_\alpha\bar{\psi}^\alpha\}_\alpha] = F[\{\psi^\alpha\}_\alpha]\,.$$

To obtain the transformed Hamiltonian structure, first the Hamiltonian functional just transforms as any functional

$$\bar{H}[\bar{\phi}] = H[\phi] = \int d\mathbf{y}\; h[\mathbf{y};\phi] = \int d\bar{\mathbf{y}}\mathcal{J}\; h\Big[\big\{\mathbf{z}^\alpha[\bar{\mathbf{z}}^\alpha;\bar{\phi}]\big\}_\alpha;\bar{\phi}\Big]\,, \tag{5.31}$$

where $\mathcal{J} := \Pi_\alpha\mathcal{J}_\alpha$ is the Jacobian of the overall transformation for $\mathbf{y}$, with $\mathcal{J}_\alpha$ the Jacobian determinant of the transformation $\tau_\alpha$.

The Poisson bracket also transforms as a functional, but functional derivatives have to be transformed as well

$$\bar{\{}\bar{F},\bar{G}\bar{\}} = \int d\bar{\mathbf{y}}\mathcal{J}\;\left(T^{-1}\tfrac{\delta}{\delta\bar{\psi}^\alpha}\mathbb{T}^{-1}\bar{F}\right)\;\mathbb{J}^{\alpha,\beta}\;\left[\{\mathbf{z}^\gamma[[\bar{\mathbf{z}}^\gamma;\bar{\phi}]]\}_\gamma;\bar{\phi}\right]\;\left(T^{-1}\tfrac{\delta}{\delta\bar{\psi}^\beta}\mathbb{T}^{-1}\bar{G}\right)\,. \tag{5.32}$$

Using the method of the previous sections, transformed functional derivatives are obtained as

$$T^{-1}\tfrac{\delta}{\delta\bar{\psi}^\alpha(\mathbf{z}^\alpha)}\mathbb{T}^{-1}\bar{F} = \sum_\gamma\int d\mathbf{z}^\gamma\Big(\mathcal{J}^{-1}_\gamma\bar{F}_{\bar{\psi}^\gamma}\Big)\Big[1^\gamma_\alpha\delta(\mathbf{z}^\alpha-\mathbf{z}^\gamma) - \bar{\psi}^\gamma_{\bar{\mathbf{z}}^\gamma}\left(\bar{\mathbf{z}}^\gamma_{\psi^\alpha}\right)^{\dagger_\alpha}\delta(\mathbf{z}^\alpha-\pi^\alpha_\gamma)\Big]\,, \tag{5.33}$$

where the adjoint $\dagger_\alpha$ is with respect to the measure $d\mathbf{z}^\alpha$. Eq. (5.33) agrees with the chain-rule formula obtained by computing the total derivative of the defining formula (5.30) for the transformed fields

$$\delta\psi^\gamma(\mathbf{z}^\gamma) = \delta\bar{\psi}^\gamma(\bar{\mathbf{z}}^\gamma[\mathbf{z}^\gamma;\phi]) + \sum_\alpha\bar{\psi}^\gamma_{\bar{\mathbf{z}}^\gamma}\cdot\bar{\mathbf{z}}^\gamma_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z}^\alpha[\mathbf{z}^\gamma;\phi])\,. \tag{5.34}$$

This means that

$$\tfrac{\delta\bar{\psi}^\gamma(\bar{\mathbf{z}}^\gamma)}{\delta\psi^\alpha(\mathbf{z}'^\alpha)} = 1^\gamma_\alpha\delta(\mathbf{z}'^\alpha-\mathbf{z}^\gamma[\bar{\mathbf{z}}^\gamma;\bar{\phi}]) - \bar{\psi}^\gamma_{\bar{\mathbf{z}}^\gamma}\cdot\left(\bar{\mathbf{z}}^\gamma_{\psi^\alpha}\right)^{\dagger_\alpha}\delta(\mathbf{z}'^\alpha-\mathbf{z}^\alpha[\mathbf{z}^\gamma;\phi])\,,$$

where the adjoint $\dagger_\alpha$ is with respect to $d\mathbf{z}'^\alpha$, and in the last term $\mathbf{z}^\gamma$ is actually $\mathbf{z}^\gamma[\bar{\mathbf{z}}^\gamma;\bar{\phi}]$.

Notice that $\mathbf{z}^\gamma[\bar{\mathbf{z}}^\gamma;\bar{\phi}]$ is the transformation $(\tau_\gamma)^{-1}$, whereas $\mathbf{z}^\alpha[\mathbf{z}^\gamma;\phi]$ is the function $\pi^\alpha_\gamma$, so they look similar but they are completely different objects. Furthermore, the function $\pi$ implies the presence of a Dirac delta function $\delta(\mathbf{z}^\alpha-\pi^\alpha_\gamma)$, which cannot be integrated (as is clear in (5.33) for instance), unlike in previous examples of lifting, where $\pi$ did not exist, or was trivial. In the case

[6] More precisely it has to imply the presence of a map $\pi^\gamma_\alpha$ unless there is an integral acting on the $\psi^\gamma$ (functional dependence). Notice that the map $\pi^\gamma_\alpha$ can anyway be considered as a functional dependence and be absorbed in an integral, viz. $\phi^\gamma\big(\pi^\gamma_\alpha(\mathbf{z}^\alpha)\big) = \int d\mathbf{z}^\gamma\phi^\gamma(\mathbf{z}^\gamma)\delta(\mathbf{z}^\gamma-\pi^\gamma_\alpha(\mathbf{z}^\alpha))$.

of the Vlasov-Maxwell system, the natural embedding of $\mathbf{x}$ into $\mathbf{z}$ avoided its explicit appearance in the final formulae for the transformed functional derivatives (see (5.15), for instance).

Together, the transformed Hamiltonian functional (5.31) and Poisson bracket (5.32) with the transformed functional derivative (5.33) give the transformed dynamics, which is found to be related with the time evolution of the initial dynamics through a chain rule

$$\begin{aligned}\dot{\bar{F}} = \bar{\{}\bar{F}, \bar{H}\bar{\}} &= \int d\mathbf{y} \int d\mathbf{z}'^{\lambda} \left(\mathcal{J}_{\lambda}^{-1} \bar{F}_{\bar{\psi}^{\lambda}}\right) \left[1_{\alpha}^{\lambda} \delta(\mathbf{z}^{\alpha} - \mathbf{z}'^{\lambda}) - \bar{\psi}_{\bar{\mathbf{z}}^{\lambda}}^{\lambda} \cdot \left(\bar{\mathbf{z}}_{\psi^{\alpha}}^{\lambda}\right)^{\dagger_{\alpha}} \delta(\mathbf{z}^{\alpha} - \mathbf{z}^{\alpha}[\mathbf{z}'^{\lambda}; \phi])\right] \mathbb{J}^{\alpha,\beta} \; H_{\psi^{\beta}} \\ &= \int d\mathbf{y} \int d\mathbf{z}'^{\lambda} \left(\mathcal{J}_{\lambda}^{-1} \bar{F}_{\bar{\psi}^{\lambda}}\right) \tfrac{\delta \bar{\psi}^{\lambda}(\bar{\mathbf{z}}'^{\lambda})}{\delta \psi^{\alpha}(\mathbf{z}^{\alpha})} \mathbb{J}^{\alpha,\beta} \; H_{\psi^{\beta}} = \int d\bar{\mathbf{z}}'^{\lambda} \bar{F}_{\bar{\psi}^{\lambda}} \{\bar{\psi}^{\lambda}(\bar{\mathbf{z}}'^{\lambda}), H\} = \int d\bar{\mathbf{z}}'^{\lambda} \bar{F}_{\bar{\psi}^{\lambda}} \dot{\bar{\psi}}^{\lambda}(\bar{\mathbf{z}}'^{\lambda}) \,, \end{aligned} \tag{5.35}$$

as expected for a change of field variables. In this computation, the cancellation mentioned about (11.27) was useful again.

The same comments as in previous sections hold: the two terms involved in transformed functional derivatives (5.33) transfer to the transformed Poisson bracket and dynamics: the first term just gives to the transformed Poisson bracket and dynamics the same expression as the initial ones. It corresponds to considering that the scalar invariance (5.30) implies that the transformed fields evolve in a similar way as the initial one. As for the second and third terms, they traduce the field dependence in the coordinate transformation. They imply additional terms in the transformed dynamics and Poisson bracket compared to the corresponding expressions in the initial system. For the dynamics, they traduce the effects of the time dependence in the coordinate transformation.

More precisely, the term with $\pi_{\alpha}^{\alpha}$ (i.e. $\gamma = \alpha$) corresponds to the direct time dependence in the coordinate transformation $\tau_{\alpha}$ and was already present previously, whereas the terms with $\gamma \neq \alpha$ correspond to the time dependence in the map $\pi_{\alpha}^{\gamma}$ involved in the transformation needed to evaluate a field $\psi^{\gamma}$ in a point $\mathbf{z}^{\gamma}$ depending only on another space $\mathbf{z}^{\alpha}$, in an analogous way as what is done for the Vlasov-Maxwell system, when evaluating the electromagnetic field at a field position $\mathbf{x}$ parametrized by the particle position $\mathbf{q}$; this term was not present in previous cases of lifting, where the map $\pi_{\alpha}^{\gamma}$ was absent or trivial.

In a similar way as the previous section, it is clear that magnetization-like terms are present when two of the particle spaces are physically identical, hence coupled in all physical functionals by Dirac delta functions $\delta(\mathbf{z}^{\alpha} - \mathbf{z}^{\gamma})$, and these spaces do not undergo the same transformation $\bar{\mathbf{z}}^{\alpha} \neq \bar{\mathbf{z}}^{\gamma}$. But now the presence of the map $\pi_{\alpha}^{\gamma}$ also introduces a kind of coupling between spaces $\mathbf{z}^{\alpha}$ and $\mathbf{z}^{\gamma}$. So, one can expect magnetization-like terms when these spaces undergo a change of coordinates that does not preserve the map $\pi_{\alpha}^{\gamma}$. This is confirmed by the result (5.35).

# Chapter 6

# Gyrokinetics as a Hamiltonian perturbation theory

in collaboration with Michel Vittot

*A*bstract: The Vlasov-Maxwell dynamics in the limit of a strong magnetic field and weak electric field is formulated as a Hamiltonian perturbation of particle dynamics in an external magnetic field. At lowest order, the reduction of particle dynamics can be directly lifted to the plasma level. Then the perturbation, together with the coupling between the plasma and the electromagnetic field, is restored with the electric term in the Hamiltonian.
This field perturbation theory is applied to a simplified version of the gyrokinetic reduction, the magnetic-moment reduction. The guiding-center transformation is shown to provide the lowest-order reduction. At higher order, the scheme fits with a canonical transformation, whose derivation at lowest orders is explored through a formal computation.
A non-canonical approach is considered as well. It induces an equation for the magnetic-moment function, which exaclty agrees with the gyro-center transformation and can be solved by using standard methods of particle perturbation theory. The results emphasize that the gyro-center transformation is implicitly time dependent, and the magnetic moment depends on the moments of the Vlasov density.

## Introduction

In the previous chapters, guiding-center theory and gyrokinetics started by reducing particle dynamics in a strong magnetic field, with possibly also a weak electric field. Next, the reduction was lifted to the Vlasov-Maxwell system, but then it appeared that the particle reduction did not work at the field level, basically because particles cannot be considered as independent of each other, i.e. only submitted to a given electromagnetic field. The system is fully coupled, and the plasma must be dealt with as a whole, together with the Maxwell fields.

One way to take into account these effects is to use time-dependent particle perturbation theory, as was mentioned in the previous chapter, but it could be more convenient to work with the full dynamical system. This means to perform a perturbation analysis directly at the level of field dynamics, in the framework of the Vlasov-Maxwell system.

Such an approach is expected to be more elaborate than particle perturbation theory, especially because coefficients can be operators, hence not trivial to invert. However, it would be more suited to the system under consideration. It could give a better view of the reduction and improve the efficiency of the derivation, in a similar way as what occurred with the introduction of non-canonical Hamiltonian perturbation theory in guiding-center theory [86, 88]. This last relied on the idea to benefit from Hamiltonian reduction methods, while keeping natural physical coordinates, instead of canonical coordinates. Initially it could have seemed more elaborate than non-Hamiltonian perturbation theory (see e.g. Chapter 1 compared to Chapter 2), and also than canonical perturbation

theory, but finally it has proven much more efficient.

In the present chapter, we explore this approach of the gyrokinetic reduction considered directly in the framework of the Vlasov-Maxwell system. The goal is twofold: to probe the tools and methods involved in this field perturbation theory, and to identify the essential differences with the standard approach of gyrokinetics, working at the particle level.

A first step is to establish the framework of the perturbation theory. Here, it will be induced by a splitting in the Hamiltonian with a first part corresponding to a "simplified" dynamics, and a remaining part which is just a small correction and can be treated as a perturbation. The splitting will fit with the gyrokinetic reduction, where the first stage (the guiding-center transformation) keeps the fields external, whereas the second stage (the gyro-center transformation) restores the coupling.

Then the second step is to solve the Hamiltonian perturbation theory, i.e. to compute the gyro-center transformation, which will provide the reduction in the presence of the coupling. For simplicity here, it will not be applied to the full gyrokinetic reduction, but rather to a simplified version, concerned with the magnetic-moment reduction. It is a way to avoid intricacies obfuscating the main mechanisms at work.

Care should be taken on the fact that for the field dynamics, the magnetic moment is a coordinate on a fixed space, not a dynamical variable. So, the framework is very different from particle perturbation theory. It can end up with a mere reformulation of the reduction giving similar results, and then the main point will be to compare the efficiency of the methods. Alternatively, it can provide different results, and then the main point will be to clarify their differences and their relative relevance.

The organization of the chapter is the following. In Sec. 6.1, the framework of the gyrokinetic reduction from Vlasov-Maxwell dynamics is established, by showing that plasma dynamics in a strong magnetic field can be considered as a Hamiltonian perturbation of a simplified dynamics, which corresponds to a static magnetic field with zero electric field. In Sec. 6.2, the criterion for the magnetic-moment reduction is formulated at the field level. Then the initialization of the reduction is studied, and the guiding-center reduction is shown to provide the solution, i.e. to satisfy the criterion at lowest order. In Sec. 6.3, in order to solve the reduction at higher orders, a canonical approach will be explored by using formal expansions. In Sec. 6.4, we will turn to a non-canonical transformation, which will make the connection with the standard approach of gyrokinetics, with its particle perturbation theory.

Here, the argument is developed in the case of the more complete Vlasov-Maxwell system, but it can be applied in a similar way to the Vlasov-Poisson system in the case of electrostatic gyrokinetics.

## 6.1 Hamiltonian perturbation framework

For the consistency of the chapter, let us remind the Vlasov-Maxwell non-canonical Hamiltonian system. The observables are the set of all functionals of the magnetic field $\mathbf{B}(\mathbf{q})$, the electric field $\mathbf{E}(\mathbf{q})$, and the phase-space density $f(\mathbf{q},\mathbf{p})$, with $\mathbf{q}$ and $\mathbf{p}$ the phase-space coordinates (resp. for the particle position and momentum). Between two functionals $F$ and $G$, the Poisson bracket is [95, 101, 104, 156]:

$$\{F,G\} = \int d\mathbf{q}d\mathbf{p}\, f\, \Big([F_f,G_f] + G_{\mathbf{E}}\cdot\partial_{\mathbf{p}}F_f - F_{\mathbf{E}}\cdot\partial_{\mathbf{p}}G_f\Big) + \int d\mathbf{q}\, \left(F_{\mathbf{E}}\cdot\nabla\times G_{\mathbf{B}} - G_{\mathbf{E}}\cdot\nabla\times F_{\mathbf{B}}\right), \quad (6.1)$$

where the particle bracket is

$$[g,h] = \nabla g\cdot\partial_{\mathbf{p}}h - \nabla h\cdot\partial_{\mathbf{p}}g + \mathbf{B}\cdot\partial_{\mathbf{p}}g\times\partial_{\mathbf{p}}h\,,$$

for any functions of the phase space $g$ and $h$. For the sake of simplicity physical constants such as $e$, $m$, or $c$ are scaled away.

The Hamiltonian functional is

$$H[\mathbf{E},\mathbf{B},f] = \int d\mathbf{q}d\mathbf{p}\,\tfrac{\mathbf{p}^2}{2}f + \int d\mathbf{q}\,\tfrac{\mathbf{E}^2+\mathbf{B}^2}{2}\,, \tag{6.2}$$

which is the sum of the kinetic energy of the plasma and the energy of the electromagnetic field. The coupling between the plasma and electromagnetic field is included in the non-canonical Poisson bracket (6.1). The Hamiltonian (6.2) together with the Poisson bracket generates the motion through Hamilton's equations expressed as

$$\dot{F} = \{F,H\}\,,$$

for any observable $F$. In particular, $F$ can be chosen as the field variables $(\mathbf{E},\mathbf{B},f)$):

$$\begin{aligned}
\dot{\mathbf{E}} &= \{\mathbf{E},H\} = \nabla\times\mathbf{B} - \int d\mathbf{p}\;f\mathbf{p}\,,\\
\dot{\mathbf{B}} &= \{\mathbf{B},H\} = -\nabla\times\mathbf{E}\,,\\
\dot{f} &= \{f,H\} = -\mathbf{p}\cdot\nabla f - (\mathbf{E}+\mathbf{p}\times\mathbf{B})\cdot\partial_{\mathbf{p}}f\,,
\end{aligned} \tag{6.3}$$

which is the Vlasov-Maxwell dynamics.

In the case considered by gyrokinetics, the main effects of the coupling between the plasma and the dynamical Maxwell field come from the electric field, and precisely, the electric term in the Hamiltonian is small and compared to the magnetic term. This suggests to write

$$\begin{aligned}
H &= H_0 + \varepsilon H_1\,,\\
\text{with } H_0 &:= \int d\mathbf{q}d\mathbf{p}\,f\tfrac{\mathbf{p}^2}{2} + \int d\mathbf{q}\,\tfrac{\mathbf{B}^2}{2}\,,\\
\text{and } H_1 &:= \int d\mathbf{q}\,\tfrac{\mathbf{E}^2}{2}\,,
\end{aligned}$$

where $\varepsilon$ is a small parameter. Thus, a Hamiltonian perturbation framework is natural. The motion is mainly given by a simplified Hamiltonian $H_0$, which corresponds to the following dynamics:

$$\begin{aligned}
\dot{\mathbf{E}} &= \{\mathbf{E},H_0\} = \nabla\times\mathbf{B} - \int d\mathbf{p}\;f\mathbf{p}\,,\\
\dot{\mathbf{B}} &= \{\mathbf{B},H_0\} = 0\,,\\
\dot{f} &= \{f,H_0\} = -\mathbf{p}\cdot\nabla f - (\mathbf{p}\times\mathbf{B})\cdot\partial_{\mathbf{p}}f\,.
\end{aligned} \tag{6.4}$$

The dynamics of the electric is given by Maxwell-Ampère's equation, as usual, but it is skew decoupled from the dynamics of the other fields, since it does not influence them. The Vlasov density and the magnetic field evolve as if there were no electric field: $\mathbf{B}$ is static and $f$ is given by the Lorentz force with a static (external) magnetic field. It implies that the dynamics of the particles are independent of each other.

It is a huge reduction, because it makes the infinite dimensional Vlasov dynamics equivalent to a (continuous) set of identical, independent, 6-dimensional particle dynamics. It exactly agrees with the framework of Episode 1 (Steps 1 and 2a of the gyrokinetic reduction, introduced on page 16). The dynamics of all particles can be reduced independently and identically by using particle perturbation theory. The reduction directly transfers to the Vlasov field by a lifting procedure, which is equivalent to a change of coordinates for the Vlasov density in this case.

Then the coupling (Step 2b) is restored by putting back the electric field in the Hamiltonian. It restores the effect of the electric field both on the plasma and on the magnetic field. Then the guiding-center reduction obtained previously is spoilt, and the gyro-center transformation must be performed to correct perturbatively the guiding-center transformation and recover the reduction.

This Hamiltonian perturbative reduction exactly corresponds to the two-stage gyrokinetic reduction, with especially the distinction between the guiding-center and gyro-center transformations,

and between particle reduction and plasma reduction.

The rationale above focused on the Hamiltonian functional, as is usual in Hamiltonian perturbation. However, here the Poisson bracket is non-canonical, and its contribution must be taken into account in the argument. Otherwise the conclusion might not be valid, as is illustrated by the following observation. Since the strong term in the Hamiltonian functional is the magnetic term, one could be tempted to consider also the kinetic energy as a small perturbation, and to exclude it from the simplified Hamiltonian:

$$\begin{aligned} H &= H_0 + \varepsilon H_1 \,, \\ \text{with } H_0 &:= \int d\mathbf{q}\, \tfrac{\mathbf{B}^2}{2} \,, \\ \text{and } H_1 &:= \int d\mathbf{q} d\mathbf{p}\, f \tfrac{\mathbf{p}^2}{2} + \int d\mathbf{q}\, \tfrac{\mathbf{E}^2}{2} \,. \end{aligned} \tag{6.5}$$

Then one would end up with the following simplified motion:

$$\begin{aligned} \dot{\mathbf{E}} &= \{\mathbf{E}, H_0\} = \nabla \times \mathbf{B} \,, \\ \dot{\mathbf{B}} &= \{\mathbf{B}, H_0\} = 0 \,, \\ \dot{f} &= \{f, H_0\} = 0 \,. \end{aligned}$$

Clearly, it is not close to the physical motion. Especially, the main term in the physical motion, the gyromagnetic term $-(\mathbf{p} \times \mathbf{B}) \cdot \partial_{\mathbf{p}} f$, is missing. The reason is that in this term the magnetic field does not come from the Hamiltonian functional, but from the Poisson bracket.

More precisely, the strong magnetic field is mainly external (only a small part of the magnetic field is self-consistently generated by the plasma). It should be decomposed as $\mathbf{B} := \mathbf{B}_0 + \varepsilon \tilde{\mathbf{B}}$, with $\mathbf{B}_0$ external (given) and large, and with $\varepsilon \tilde{\mathbf{B}}$ self-consistent with the plasma and small. This can be viewed as a translational change of field coordinates provided $\nabla \cdot \tilde{\mathbf{B}} = 0$. In the Poisson bracket, the functional derivatives are not affected, since $F_{\mathbf{B}}$ just becomes $F_{\tilde{\mathbf{B}}}$. The only effect concerns the gyromagnetic term, which becomes split into two contributions by $\mathbf{B} := \mathbf{B}_0 + \varepsilon \tilde{\mathbf{B}}$.

As for the Hamiltonian, its magnetic term becomes

$$\int d\mathbf{q}\, \tfrac{(\mathbf{B}_0 + \varepsilon \tilde{\mathbf{B}})^2}{2} = \int d\mathbf{q}\, \tfrac{\mathbf{B}_0^2}{2} + 2 \int d\mathbf{q}\, \tfrac{\mathbf{B}_0 \cdot \varepsilon \tilde{\mathbf{B}}}{2} + \int d\mathbf{q}\, \tfrac{(\varepsilon \tilde{\mathbf{B}})^2}{2} \,.$$

This illustrates the observation below: the first term is large, but its gradients (in field dynamics) $\frac{\delta}{\delta \psi}$ are small (even zero), whereas the second term is not large, but its gradient is large. Thus what is important is not the Hamiltonian $H$, but rather its gradients $\partial_{\mathbf{z}} H$ (or $\frac{\delta H}{\delta \boldsymbol{\psi}}$ for continuous media), and more precisely its symplectic gradient $\mathbb{J} \cdot \nabla H$. For canonical systems, it is equivalent to $\nabla H$, but not for non-canonical systems.

To take into account this subtlety in the rationale above, one can use an ordering in the Poisson bracket in addition to the ordering in the Hamiltonian. Alternatively, one can just consider that the electric field is weak, hence only functionals with no electric field contribute. It corresponds to a subalgebra of the initial system, where the magnetic field becomes static, but the reduced Poisson bracket retains the large gyro-magnetic term. The dynamics is exaclty the one given by Eqs. (6.4). This fact also explains why the same Hamiltonian perturbation as in (6.5) can be applied to the Vlasov-Poisson system, although the magnetic term is absent from the Hamiltonian.

## 6.2 The magnetic-moment reduction at lowest order

This section starts exploring the gyrokinetic reduction in the framework of the field perturbation theory identified in the previous section. We begin by studying a simplified reduction, in order to better focus on the main aspects of the procedure. Afterwards, it will be easier to address the full reduction, and especially to see how the method can be applied or adapted to it. Accordingly, we only consider the magnetic-moment reduction, which is simpler to handle than the averaging reduction when working at the level of the Vlasov-Maxwell dynamics.

Indeed, the preliminary step of the procedure is to specify the reduction criteria, i.e. the goal aimed at. The criteria for the reduction of particle dynamics cannot be used directly. For instance, for particle dynamics, the magnetic-moment reduction means that one of the coordinates is conserved. It is denoted by $\bar{\mu}$ (see Chapter 2). This is meaningless in field dynamics, where the coordinates are fixed, since they constitute the base space over which the fields evolve (Eulerian description).

Alternatively, one could be tempted to adopt the criterion that the Vlasov density evolves inside leaves of constant magnetic moment. Again, it is not satisfactory, since it corresponds to a Lagrangian description rather than an Eulerian description. All the same, Lagrangian markers (or the characteristics of the Vlasov density) establish a connection between the two descriptions. In case the Vlasov density is a particle, it writes $f(\mathbf{z}) = \delta(\mathbf{z} - \mathcal{Z})$, and the observable (functional) related to the conserved magnetic moment is

$$\bar{\mathcal{A}} := \int d\mathbf{z}\ \bar{\mu}(\mathbf{z})\ \delta(\mathbf{z} - \mathcal{Z})\,.$$

Here, we distinguish between the phase space coordinate $\mathbf{z}$, which is fixed, and the particle position (usually also denoted by $\mathbf{z}$, but here denoted by $\mathcal{Z}$) in phase space which is a time-dependent function taking its value in the phase space

$$\mathcal{Z} : M \times \mathbb{R} \to M : (\mathbf{z}_0, t) \to \mathcal{Z}(\mathbf{z}_0, t)\,,$$

where $M$ is the phase space (with coordinate $\mathbf{z}$), and $\mathbf{z}_0$ is the position at time $t = 0$.

Thus, in the Eulerian description, a candidate criterion for the magnetic moment is that the observable

$$\int d\mathbf{z} f \bar{\mu} \tag{6.6}$$

is conserved. Interestingly, it is the total magnetic moment. Notice that in principle, for the magnetic moment coordinate to become superfluous and removable, what is needed is not only the total magnetic moment to be conserved, because it implies only a global conservation, not a local in $\mu$, so that the characteristics of the Vlasov density might have varying magnetic moment, provided there is a global compensation between the variations. This is clear in the result of Eq. (8.31) for instance. A more precise requirement would be the conservation of the observable $\int d\mathbf{z} f g(\bar{\mu})$ for any function $g$. Then when choosing $g$ as an approximation of the Dirac delta function peaking on any specific value of the magnetic moment $\bar{\mu} = \bar{\mu}_0$, the corresponding observable would be exactly the number of particles with magnetic moment equal to $\bar{\mu}_0$. However, the presence of $g$ is generally superfluous, because the conservation of $\int d\mathbf{z} f \bar{\mu}$ for any $f$ implies that $f$ can be chosen to be just a particle $f(\mathbf{z}) = \delta(\mathbf{z} - \mathcal{Z})$, which in turn ordinarily implies the conservation of the coordinate $\bar{\mu}$ along the characteristics. Only the presence of global coefficients replacing local ones can make the conclusion invalid, as was the case for Eq. (8.31).

On another hand, in principle, the magnetic-moment reduction would naturally take place in a coordinate system adapted to $\bar{\mu}$, e.g. $(\mathbf{q}, \hat{\mathbf{v}}, \bar{\mu})$, with $\hat{\mathbf{v}}$ the unit vector of the momentum (or velocity). Then the requirement (6.6) would mean that one of the coordinates is not dynamical (in the sense of the characteristics of the Vlasov density). In addition, in the spirit of a KAM-like Hamiltonian perturbation theory [92], adopting these coordinates could prove useful also in order to localize the "invariant manifold" for the conserved quantity outside "resonant regions" (this vocabulary is inspired by the analogy with finite-dimensional systems, although it does not strictly apply to continuous systems). However in the present case, it would complicate the formulae and obfuscate their meaning. Accordingly, we start with the usual coordinates $(\mathbf{p}, \mathbf{q})$. We will change to adapted coordinate when it becomes useful, e.g. when the conditions for the resonant regions appear.

Now the derivation can start. The goal is to find the conserved quantity $\bar{\mathcal{A}}$. At lowest order, $H = H_0$, the particles are decoupled and the guiding-center reduction should provide the expected result by taking $\mathcal{A}_0 := \int d\mathbf{z} f \bar{\mu}_{gc}$, where the index refers to the order in the perturbation expansion,

and $\bar{\mu}_{gc}$ is the guiding-center magnetic moment, that is the magnetic moment in the absence of electric field. This is easily confirmed

$$\begin{aligned}\dot{\mathcal{A}}_0 = \{\mathcal{A}_0, H_0\} &= \int d\mathbf{q}d\mathbf{p}\, f\, [(\mathcal{A}_0)_f, (H_0)_f] = \int d\mathbf{q}d\mathbf{p}\, f\, [\bar{\mu}_{gc}, \tfrac{\mathbf{p}^2}{2}] \\ &= \int d\mathbf{q}d\mathbf{p}\, f\, \Big(\mathbf{p}\cdot\nabla\bar{\mu}_{gc} + \mathbf{p}\times\mathbf{B}\cdot\partial_{\mathbf{p}}\bar{\mu}_{gc}\Big) = 0\,, \end{aligned} \tag{6.7}$$

where the last equality follows from the definition of the magnetic moment in a strong static magnetic field as the solution of the following equation:

$$0 = \dot{\bar{\mu}}_{gc} = \dot{\mathbf{q}}\cdot\nabla\bar{\mu}_{gc} + \dot{\mathbf{p}}\cdot\partial_{\mathbf{p}}\bar{\mu}_{gc} = \mathbf{p}\cdot\nabla\bar{\mu}_{gc} + \mathbf{p}\times\mathbf{B}\cdot\partial_{\mathbf{p}}\bar{\mu}_{gc}\,. \tag{6.8}$$

As a result, at order 0, one can choose

$$\mathcal{A}_0 := \int d\mathbf{z} f \bar{\mu}_{gc}\,.$$

## 6.3 Canonical transformation at first order

In case the electric field is non-zero, $\mathcal{A}_0$ is not conserved any more. However, since the electric field only introduces a perturbation in the Hamiltonian functional, one can look for a functional $\mathcal{S}$ generating a canonical transformation restoring the conserved quantity:

$$\bar{\mathcal{A}} := e^{\{\mathcal{S},\cdot\}}\mathcal{A}_0\,, \tag{6.9}$$

with $\mathcal{S}$ a functional to be identified in such a way as to make $\bar{\mathcal{A}}$ conserved

$$0 = \tfrac{d}{dt}\bar{\mathcal{A}} = \{e^{\{\mathcal{S},\cdot\}}\mathcal{A}_0, H\}\,. \tag{6.10}$$

The symbol $\{\mathcal{S},\cdot\}$ denotes the operator such that $\{\mathcal{S},\cdot\}F = \{\mathcal{S}, F\}$ for any functional $F$. Expanding the generating function $\mathcal{S}$ and the Hamiltonian $H$, Eq. (6.10) writes at first order

$$0 = \{\{\mathcal{S}_1, \mathcal{A}_0\}, H_0\} + \{\mathcal{A}_0, H_1\} = \{\{\mathcal{S}_1, H_0\}, \mathcal{A}_0\} + \{\mathcal{A}_0, H_1\} = \{\{\mathcal{S}_1, H_0\} - H_1, \mathcal{A}_0\}\,,$$

where the Jacobi identity was used in the second equality, as well as Eq. (6.7). This is equivalent to

$$\{\mathcal{S}_1, H_0\} = H_1 + H_1'\,, \tag{6.11}$$

where $H_1'$ is an element in the kernel of the operator $\{\mathcal{A}_0,\cdot\}$, to be chosen freely. It can be used to make Eq. (6.11) solvable. It plays a similar role as $\overline{\Gamma}$ in Chapter 2. Now, the strategy is clear: identify the freedom embodied in $H_1'$, which is the topic of Subsec. 6.3.1, and then solve the equation (6.11) for $\mathcal{S}_1$ and $H_1'$, which is the topic of Subsec. 6.3.2.

Notice that our exploration of this section have not provided complete results yet: we will only indicate the main features of the derivation.

### 6.3.1 On the reduced Hamiltonian $H_1'$

The goal is to identify $\ker\{\mathcal{A}_0,\cdot\}$, i.e. all the solutions $F$ of the following equation:

$$\begin{aligned} 0 &= \int d\mathbf{q}d\mathbf{p}\, f\, \Big[\Big(\big(\nabla\bar{\mu}_{gc} + \mathbf{B}\times\partial_{\mathbf{p}}\bar{\mu}_{gc}\big)\cdot\partial_{\mathbf{p}} - \partial_{\mathbf{p}}\bar{\mu}_{gc}\cdot\nabla\Big)F_f + \Big(\partial_{\mathbf{p}}\bar{\mu}_{gc} - (\bar{\mu}_{gc})_{\mathbf{B}}\times\nabla\Big)\cdot F_{\mathbf{E}}\Big] \\ &= \int d\mathbf{q}d\mathbf{p}\, f\, w_1^i\partial_i F_f + \int d\mathbf{q}\, \mathbf{w}_2\cdot F_{\mathbf{E}}\,, \end{aligned} \tag{6.12}$$

with $w_1^i\partial_i$ the scalar phase-space-dependent operator

$$w_1^i\partial_i := \big(\nabla\bar{\mu}_{gc} + \mathbf{B}\times\partial_{\mathbf{p}}\bar{\mu}_{gc}\big)\cdot\partial_{\mathbf{p}} - \partial_{\mathbf{p}}\bar{\mu}_{gc}\cdot\nabla\,,$$

and $\mathbf{w}$ the vectorial space-dependent operator

$$\mathbf{w}_2[\mathbf{q};\mathbf{B}] := \int d\mathbf{p}\, f \left(\partial_{\mathbf{p}}\bar{\mu}_{gc} - (\bar{\mu}_{gc})_{\mathbf{B}} \times \nabla\right).$$

As a preliminary analysis let us consider solutions making zero both operators in Eq. (6.12). An obvious solution is obtained when $F$ is a functional of $\mathbf{B}$. For a function depending only on $f$ and $\mathbf{B}$, the equation writes

$$0 = \int d\mathbf{q}d\mathbf{p}\, f\ w_1^i \partial_i F_f = \int d\mathbf{z}\, f\ [\bar{\mu}_{gc}, F_f] = \int d\bar{\mathbf{z}} \mathcal{J}\ \bar{f}\ \bar{[}\bar{\mu}_{gc}, \mathcal{J}^{-1}\bar{F}_{\bar{f}}\bar{]} = -\int d\bar{\mathbf{z}} \mathcal{J}\ \bar{f}\ \partial_\theta\big(\mathcal{J}^{-1}\bar{F}_{\bar{f}}\big)\,,$$

where we used $\bar{[}\bar{\mu}_{gc}, \cdot\bar{]} = -\partial_\theta$, and the overbars denote quantities transformed by the guiding-center change of coordinates, in a similar way as in Chapter 5. So, $\mathcal{J}^{-1}\bar{F}_{\bar{f}}$ can be any function of the variables $(\bar{\mathbf{q}}, \bar{\varphi}, \bar{\mu}_{gc})$, with $\bar{\mathbf{q}}$ the guiding-center (transformed position), $\bar{\varphi}$ the transformed pitch-angle (the angle between the momentum and the magnetic field), and $\bar{\theta}$ the transformed gyro-angle (the angle measuring the direction of the momentum in a plane perpendicular to the magnetic field). See Chapter 2 for more details about these coordinates. Especially, one recovers the special solution $F = H_0$, which can be verified directly in Eq. (6.12), using Eq. (6.8). Thus, for functionals independent of $\mathbf{E}$, the set of solutions is exactly given by

$$\int d\mathbf{z}\, f\, g_1\big[\bar{\mathbf{q}}, \bar{\varphi}, \bar{\mu}_{gc}; \mathbf{B}\big] + \int d\mathbf{q}\, g_2[\mathbf{q};\mathbf{B}]\,.$$

For functionals depending on $\mathbf{E}$, cancelling separately the operator $\int d\mathbf{q}\, \mathbf{w}_2$ implies the solution to depend also on $f$. And the solution should cancel the operator $\int d\mathbf{q}d\mathbf{p}\, f\ w_1^i \partial_i F_f$ at the same time. It would imply the solution to write

$$\int d\mathbf{z}\, f\, g_3\big[\bar{\mathbf{q}}, \bar{\varphi}, \bar{\mu}_{gc}; \mathbf{B}, \mathbf{E}\big]\,,$$

with $g_3$ chosen to satisfy $\int d\mathbf{q}\, \mathbf{w}_2 \cdot F_{\mathbf{E}} = 0$, i.e.

$$0 = \int d\mathbf{q}d\mathbf{p}\, f \left[\left(\partial_{\mathbf{p}}\bar{\mu}_{gc} - (\bar{\mu}_{gc})_{\mathbf{B}} \times \nabla\right) \cdot \int d\mathbf{p}\, (g_3)^\dagger_{\mathbf{E}}\, f\right].$$

This would imply a dependence in $\bar{\theta}$ in $g_3$, which contradicts the assumed form $g_3\big[\bar{\mathbf{q}}, \bar{\varphi}, \bar{\mu}_{gc}; \mathbf{B}, \mathbf{E}\big]$. So, a functional with non zero dependence in $\mathbf{E}$ cannot cancel separately the operators in Eq. (6.12). One has to deal with the full equation.

For a weak electric field, the solution can be looked for by expanding the solution in $\mathbf{E}$. Since the magnetic field is assumed to be large, one can also expand everything in $B^{-1}$. Then the equation becomes an induction equation. In order to solve it, the difficulty is that the coefficients to be inverted are operators. The solution can be built only when they are invertible.

The leading operator for the expansion in $B^{-1}$ is $\big(\mathbf{B} \times \partial_{\mathbf{p}}\bar{\mu}_{gc}\big) \cdot \partial_{\mathbf{p}}$. When making $\bar{\mu}_{gc}$ explicit, its leading order appears to be proportional to the operator $\partial_\theta$, which is invertible only on gyro-fluctuations. As for the gyro-averages, the leading operator of the expansion in $\mathbf{E}$ is $\mathbf{w}_2$, whose leading order in $B^{-1}$ is given by $\mathbf{W} := \int d\mathbf{p} f \partial_{\mathbf{p}}\bar{\mu}_{gc}$. It implies equations like

$$g_4 = \mathbf{W} \cdot F_{\mathbf{E}}\,, \tag{6.13}$$

with $g_4(\mathbf{q})$ a known function. At least at lowest orders, it can be solved. For instance, when $g_4$ is constant in $\mathbf{E}$, the solution is[1]

$$F = \int d\mathbf{q}\ g_4 \tfrac{\mathbf{W}\cdot\mathbf{E}}{\mathbf{W}^2}\,,$$

[1] Because of the kernel of the operator $\mathbf{W} \cdot \frac{\delta}{\delta \mathbf{E}}$, the solution can be added $\int d\mathbf{q}\ \mathbf{E} \cdot \mathbf{W} \times \mathbf{W}_1$, with $\mathbf{W}_1[\mathbf{q}, \mathbf{B}, f]$ an arbitrary vector field (such that the integral is meaningful). This fact is neglected here for pedagogical purpose: in this section, we rather indicate the method than derive complete results.

which requires $\mathbf{W}^2$ to be invertible (at least on $g_4\mathbf{W}\cdot\mathbf{E}$ and under the integral).

As a result, at order $E^1$, the series in $B^{-1}$ writes[2]

$$F_1 = -\int d\mathbf{q}\ \tfrac{\mathbf{E}\cdot\mathbf{W}}{\mathbf{W}^2}\sum_{n=0}^{\infty}\left[\int d\mathbf{p}\ f\bar{\mu}_{\mathbf{B}}\times\nabla\cdot\left(\tfrac{\mathbf{W}}{\mathbf{W}^2}\circ\right)\right]^n\int d\mathbf{p}\, f\ w_1^i\partial_i(F_0)_f\,,$$

where the lowest order of the solution $F_0[f,\mathbf{B}]$ is an arbitrary functional. The symbol $\circ$ means the composition of operators, i.e. the operator acts on all that is on its right, e.g. $\nabla\cdot(\mathbf{w}\circ)g = \nabla\cdot(\mathbf{w}g)$ for any function $g$ and any vector $\mathbf{w}$. The solution obtained is only a formal series. The freedom in $F_0[f,\mathbf{B}]$ is restricted by the requirement for the series to be meaningful. At lowest order in $B^{-1}$, the choice $F_0[f,\mathbf{B}] := \int d\mathbf{q}d\mathbf{p}\ fg_5$ regularizes the possible small divisor, with $g_5 = g_5[\mathbf{q},\mathbf{B}]$, but at higher order in $B^{-1}$ a suitable $F_0$ is not so obvious.

At second order in $E$, the lowest order in $B^{-1}$ can be formally written

$$F_{20} = -\int d\mathbf{q}\ \left(\mathbf{W}\cdot\tfrac{\delta}{\delta\mathbf{E}}\right)^{-1}\int d\mathbf{q}d\mathbf{p}\, f\ w_1^i\partial_i(F_1)_f\,,$$

where the first index refers to the order in $E$, and the second means that it is the lowest order in $B^{-1}$. The inversion of the operator $\mathbf{W}\cdot\frac{\delta}{\delta\mathbf{E}}$ implies equations like

$$\mathbf{W}\cdot\tfrac{\delta}{\delta\mathbf{E}}F_{20} = g_6\ \mathbf{W}_6\cdot\mathbf{E}\,,$$

with $\mathbf{W}_6$ and $g_6$ some known functions. The solution is easily found with formulae such as

$$F_{20} = \int d\mathbf{q}\ g_6\ \tfrac{(\mathbf{W}_6\cdot\mathbf{E})^2}{2\mathbf{W}_6\cdot\mathbf{W}}\,,$$

which requires $\mathbf{W}_6\cdot\mathbf{W}$ to be invertible, at least on $g_6(\mathbf{W}_6\cdot\mathbf{E})^2$ and under the integral. Then, the higher orders in $B^{-1}$, i.e. $F_{2i}$ with $i>0$, can be identified using the same method as for $F_1$.

At higher order in $E$, other obstructions could arise because $F_{\mathbf{E}}$ is symmetric in $\mathbf{E}$, whereas the function $g_4$ involved in Eq. (6.13) may not be so, for instance some terms occur involving $\mathbf{E}\cdot\mathbf{W}_7.\mathbf{E}\cdot\mathbf{W}_8$, with $\mathbf{W}_7\neq\mathbf{W}_8$. All the same, in this expansion, these terms always come in combinations such that the sum is symmetric in $\mathbf{E}$. So, it seems that the equation can be formally solved to all orders, but further investigation is required to definitely conclude about this point. More important, it will be necessary to explore the mathematical and physical meaning of the resulting series, which does not seem to converge towards a plain formula. As a consequence, it is unsure whether there exists a relevant solution depending on $\mathbf{E}$ in general.

### 6.3.2 Solving the first order

The next step is to solve Eq. (6.11) for $\mathcal{S}_1$ and $H_1'$, which writes

$$\begin{aligned}\int d\mathbf{q}\,\tfrac{\mathbf{E}^2}{2} + H_1' &= \{\mathcal{S}_1, H_0\}\\ &= \int d\mathbf{q}d\mathbf{p}\, f\ \Big(\mathbf{p}\times\mathbf{B}\cdot\partial_{\mathbf{p}} + \mathbf{p}\cdot\nabla\Big)(\mathcal{S}_1)_f + \int d\mathbf{q}\Big(\nabla\times\mathbf{B} - \int d\mathbf{p}\, f\ \mathbf{p}\Big)\cdot(\mathcal{S}_1)_{\mathbf{E}}\\ &= \int d\mathbf{q}d\mathbf{p}\, f\ w_1^i\partial_i(\mathcal{S}_1)_f + \int d\mathbf{q}\,\mathbf{w}_2\cdot(\mathcal{S}_1)_{\mathbf{E}}\,,\end{aligned}$$

where $H_1'$ is the general element of $\ker\{\mathcal{A}_0,\cdot\}$ studied in the previous subsection. The scalar phase-space-dependent operator $w_1^i\partial_i$ is now defined by

$$w_1^i\partial_i := \mathbf{p}\times\mathbf{B}\cdot\partial_{\mathbf{p}} + \mathbf{p}\cdot\nabla\,,$$

[2] In principle, the inversion of the operator $\int d\mathbf{p} f(\mathbf{B}\times\partial_{\mathbf{p}}\bar{\mu}_{gc})\cdot\partial_{\mathbf{p}}\frac{\delta}{\delta f}$ on gyro-fluctuations should be taken into account. For pedagogical purpose, we neglect it here. Its computation is rather straightforward, and its effects only complicate the explanation and the resulting formulae, but do not affect the main mechanism of the expansion in $E$.

and $\mathbf{w}_2$ is the space-dependent vector given by

$$\mathbf{w}_2[\mathbf{q};\mathbf{B}] := \nabla\times\mathbf{B} - \int d\mathbf{p}\, f\ \mathbf{p}\,.$$

As usual for linear equations, the resolution involves two parts: first to identify the freedom in the solution by studying $\ker\{H_0,\cdot\}$, second to identify a particular solution. Each of them proceeds a similar way as in the previous subsection, with especially a twofold analysis, by working first at the functional level, and then at the function (or functional density) level. This is because in the field perturbation theory, instead of dealing with differential equations $\sum_i g^i \frac{\partial}{\partial \mathbf{z}^i} x$ for a function $x(\mathbf{z})$ and some scalar coefficients $g^i(\mathbf{z})$, one deals with differential equations for functionals $\sum_\alpha \int d\mathbf{z} G^\alpha \frac{\delta}{\delta\psi^\alpha} X$, for the functional $X$ and some operators $G^\alpha$. In some way, the resolution implies to inverse both the "functional differential operator" $\frac{\delta}{\delta\psi^\alpha}$, and the coefficients $G^\alpha$, which are often differential operator. Rather than solving the complicated equation in one go, it is often more efficient to use expansions in small parameters, or formal expansions in the field variables. Then the equation becomes an induction equation and in some cases its leading coefficient can be inverted. Many obstructions can occur, because of either a non-invertible scalar product, or a non-invertible function differential operator, or a symmetry incompatibility.

In a similar way as in the previous subsection, a formal inductive solution can be built by perturbation expansion, but it requires further investigation, especially in order to conclude about its physical relevance. If the answer is negative, a detailed analysis will be needed in order to better understand the obstructions, and especially whether they can be related to some kinds of "resonances", which could perhaps be avoided by localizing the "invariant manifold" outside the "resonant regions", in an analogous way as in finite-dimensional KAM theory. Alternatively, one can turn to non-canonical Hamiltonian perturbation theory, which opens more possible transformations and could make solutions easier to find. The next section investigates the latter possibility, which will make a link with the standard approach of gyrokinetics.

## 6.4 Non-canonical approach

The previous section explored the field perturbation theory by using canonical transformations, which are the natural transformations when the perturbation is induced by the Hamiltonian. Many difficulties have come out, which suggests to make the requirements less stringent, and especially to drop the condition for the transformation to be canonical. So, we consider now an arbitrary near-identity transformation. Then Eq. (6.9) becomes

$$\bar{\mathcal{A}} := \mathcal{A}_0 + \mathcal{A}_1 + \dots\,,$$

where $\mathcal{A}_{i\geqslant 0}$ are functionals to be identified in such a way as to make $\bar{\mathcal{A}}$ conserved

$$0 = \tfrac{d}{dt}\bar{\mathcal{A}} = \{\mathcal{A}_0 + \mathcal{A}_1 + \dots, H_0 + H_1\}\,.$$

At lowest order, it is verified by $\mathcal{A}_0$, as shown in Sec. 6.2. At first order, it writes

$$\begin{aligned} 0 =& \{\mathcal{A}_0, H_1\} + \{\mathcal{A}_1, H_0\} \\ =& \int d\mathbf{q}d\mathbf{p}\ f\ \Big(\partial_{\mathbf{p}}\bar{\mu}_{gc}\cdot\mathbf{E} - (\bar{\mu}_{gc})_{\mathbf{B}}\cdot\nabla\times\mathbf{E}\Big) \\ &+ \int d\mathbf{q}d\mathbf{p}\, f\ \Big(\mathbf{p}\times\mathbf{B}\cdot\partial_{\mathbf{p}} + \mathbf{p}\cdot\nabla\Big)(\mathcal{A}_1)_f + \int d\mathbf{q}\Big(\nabla\times\mathbf{B} - \int d\mathbf{p}\, f\ \mathbf{p}\Big)\cdot(\mathcal{A}_1)_{\mathbf{E}}\,. \end{aligned} \tag{6.14}$$

In principle, the resolution for $\mathcal{A}_1$ follows the same lines as in the previous subsections. However, in this equation, the coefficients can be observed to fit exactly with the order in the electric field:

$$0 = \int d\mathbf{q}d\mathbf{p}\ \Big((\mathcal{A}_0)_f(\dot{f})_1 + \int d\mathbf{q}(\mathcal{A}_0)_{\mathbf{B}}(\dot{\mathbf{B}})_1 + \int d\mathbf{q}d\mathbf{p}\ (\mathcal{A}_1)_f(\dot{f})_0 + \int d\mathbf{q}\ (\mathcal{A}_1)_{\mathbf{E}}\cdot(\dot{\mathbf{E}})_0\,,$$

where the indices refer to the order in the electric field. This fact suggests to write $\mathcal{A}_1$ under the form

$$\mathcal{A}_1 = \int d\mathbf{z}\ f\ \bar{\mu}_1\,. \tag{6.15}$$

It is a restriction, but it will be enough to find a solution. In addition, it is analogous to (6.6), and this form is necessary in order to obtain a dimensional reduction. Finding a constant of motion $\bar{\mathcal{A}}$ for field dynamics would not be satisfactory: reducing by 1/2 the degrees of freedom of an infinite-dimensional system is not very useful. The goal is rather a dimensional reduction, i.e. to make "conserved" (in the sense of the characteristics) a dimension of the base space. If Eq. (6.15) holds at arbitrary order for all $f$, then specifying $f$ as a single particle will show that the characteristics occur in leaves of constant $\bar{\mu}_{gc}+\bar{\mu}_1+\bar{\mu}_2+...$, and this quantity will be interpreted as the gyro-center magnetic moment coordinate. Notice that here the index refers to the order in the electric field, not the order in the Larmor radius, as was the case in Chapter 2: all the series in the Larmor radius is now included in $\bar{\mu}_{gc}$.

Under the assumption (6.15), Eq. (6.14) becomes

$$0 = \int d\mathbf{q}d\mathbf{p}\ f\ \left[\left(\partial_{\mathbf{p}}\bar{\mu}_{gc}\cdot\mathbf{E} - (\bar{\mu}_{gc})_{\mathbf{B}}\cdot\nabla\times\mathbf{E}\right) + \left(\mathbf{p}\times\mathbf{B}\cdot\partial_{\mathbf{p}} + \mathbf{p}\cdot\nabla\right)\bar{\mu}_1 + (\bar{\mu}_1)_{\mathbf{E}}\left(\nabla\times\mathbf{B} - \int d\mathbf{p}\ f\ \mathbf{p}\right)\right].$$

It must hold for any $f$, whence

$$\begin{aligned} 0 &= \left(\partial_{\mathbf{p}}\bar{\mu}_{gc}\cdot\mathbf{E} - (\bar{\mu}_{gc})_{\mathbf{B}}\cdot\nabla\times\mathbf{E}\right) + \left(\mathbf{p}\times\mathbf{B}\cdot\partial_{\mathbf{p}} + \mathbf{p}\cdot\nabla\right)\bar{\mu}_1 + (\bar{\mu}_1)_{\mathbf{E}}\left(\nabla\times\mathbf{B} - \int d\mathbf{p}\ f\ \mathbf{p}\right) \\ &= \partial_{\mathbf{p}}\bar{\mu}_{gc}\cdot(\dot{\mathbf{p}})_1 + (\bar{\mu}_{gc})_{\mathbf{B}}\cdot(\dot{\mathbf{B}})_1 + \left((\dot{\mathbf{p}})_0\cdot\partial_{\mathbf{p}} + (\dot{\mathbf{q}})_0\cdot\nabla\right)\bar{\mu}_1 + (\bar{\mu}_1)_{\mathbf{E}}\cdot(\dot{\mathbf{E}})_0 \\ &= \left(\tfrac{d}{dt}\right)_1\bar{\mu}_{gc} + \left(\tfrac{d}{dt}\right)_0\bar{\mu}_1\,, \end{aligned} \tag{6.16}$$

where $\left(\frac{d}{dt}\right)_i$ is the $i$-th order in the variable $\mathbf{E}$ for the generator of the time evolution $\frac{d}{dt}g = \partial_{\mathbf{z}}g\cdot\dot{\mathbf{z}} + \partial_{\boldsymbol{\psi}}g\cdot\dot{\boldsymbol{\psi}}$, with $\boldsymbol{\psi} := (\mathbf{E},\mathbf{B})$ a vector grouping the fields involved in the transformation.

In equation (6.16), the unknown is not a functional, but the function $\bar{\mu}_1$. It is exactly the equation of order $E^1$ for the gyro-center magnetic moment in the standard approach of gyrokinetics. Thus in this case, the field perturbation theory and the particle perturbation theory give the same equation and the same results. The reason relies on the idea of dimensional reduction, with its associated assumption (6.15), which does not concern functionals but functions.

Accordingly, Eq. (6.16) can be solved by using the standard approach of the gyro-center transformation, by Lie-transforming the phase-space Lagrangian. The terms with $\dot{\boldsymbol{\psi}}$ traduce the time dependence of the transformation, and the presence of the current $\int d\mathbf{p} f\mathbf{p}$ in the coefficients means that the moments of the distribution function will be implied in the transformation. They confirm the conclusions of the previous chapter in a more direct way.

## Conclusion

Plasma dynamics in a strong magnetic field and weak electric field is a Hamiltonian perturbation of particle dynamics in a static magnetic field. The two steps of this perturbation theory exactly correspond to the two steps of gyrokinetics. The perturbation is the electric energy in the Hamiltonian, which is the term generating the coupling between the Vlasov density and the electromagnetic field.

When this term is negligible, the reduction is huge, because the infinite dimensional coupled system becomes equivalent to a set of uncoupled 6-dimensional systems. This legitimates the standard approach of the guiding-center reduction (Steps 1 and 2a of the gyrokinetic reduction on page 16), which works on particle dynamics, often with no electric field, and is directly lifted to the plasma level, as was done in the previous chapters.

The criteria for the field dynamical reduction are more subtle than the criteria for the particle dynamical reduction, but they are closely related. The results of the guiding-center reduction were shown to provide the lowest-order reduction for the Vlasov-Maxwell system.

At higher order, the self-consistency between the plasma and the electromagnetic field is restored by reinstating the electric field in the Hamiltonian functional and treating it as a perturbation. This is the common setting where a canonical transformation is used in order to restore the reduction in the presence of the perturbation. Roughly speaking, the resolution method has to first integrate

the functional derivatives, then to invert the coefficients, which are often vectors or operators. This implies strong constraints on the solutions, and many obstructions can appear. In the case of the magnetic-moment reduction, a perturbation series for the solution could be derived, but it still requires some study, especially to conclude about the suitability of the corresponding formal expansion.

A non-canonical transformation opened more possibilities. It emphasized that in order for the transformation to provide a dimensional reduction, the unknown to be identified was just a function of the phase space, bringing the framework to the particle level. This is somehow expected since the transformation concerns particle coordinates or functions rather than purely functionals; the goal is a dimensional reduction, not just a constant of motion. The resulting equation to be solved was shown to fit with particle perturbation theory, and it is most efficiently solved by the variational formulation of particle dynamics as in Chapter 2.

This legitimates the standard approach of the gyro-center transformation (Steps 2b and 2c of the gyrokinetic reduction), which works on particle dynamics and is then lifted to the field level. It enhances the relevance of Chapters 2 and 5. It also confirms that the time-dependent setting identified in the previous chapter must be used, with the plasma gyro-center transformation depending on the moments of the Vlasov density.

This chapter mainly started an application of field perturbation theory, confirmed the relevance of using particle perturbation theory in the case of the gyrokinetic reduction, and lightened several aspects in the reduction. Finally, the field perturbation theory is more proper than the particle approach, but it naturally shrinks into a particle perturbation, which is more efficient to find the solution. Thus the field perturbation approach in itself did not provide original results yet. This first exploration needs to be continued.

The argument was considered for a simplified version of gyrokinetics, with the magnetic-moment reduction, but it should apply to the full reduction, with probably more accuracy, since the averaging reduction is still more closely related to the coordinate system.

On another hand, the conclusion leading back to particle perturbation theory is somehow unsatisfactory, for the initial goal aimed at applying field Hamiltonian perturbation theory to the gyrokinetic reduction. To reach this goal, the idea of field Hamiltonian perturbation theory can be explored in other systems, while gyrokinetics can be studied with other field Hamiltonian reduction methods, as will be the topic of Episode 3, and especially Chapter 8. Another possible extension will be to consider the field Hamiltonian perturbation theory in a variational context, since it was the key ingredient to make the reduction really efficient in the case of particle perturbation theory, especially by replacing a differential framework by an algebraic one.

# Conclusion of the episode

This second episode clarified the three difficulties mentioned in its introduction about the lifting procedure, by obtaining a functional chain rule for the lifting transformation, by dealing with an implicit definition of the transformed fields, and by the presence of a Dirac delta function relating the particle position and the electromagnetic field position, considered as pertaining different spaces, which is crucial to generate polarization and magnetization terms through a lifting process.

This answered an important question for gyrokinetic theory, by providing a Hamiltonian structure for the transformed Vlasov-Maxwell system after the guiding- and gyro-center changes of particle coordinates. It emphasized that the time dependence in the transformation is crucial for a proper coupling between the plasma and the electromagnetic field. As a consequence, the transformation of particle dynamics depends not only on the electromagnetic field, but also on all the moments of the Vlasov density. In addition, the reduced Vlasov characteristics are not Hamiltonian by themselves.

Although overlooked in previous works, this result is actually just a consequence of the self-consistency of the field dynamics, since the electromagnetic field is generated by the plasma. It enhances how the Hamiltonian structure, although formal, perfectly encodes physical effects, such as time effects or coupling effects.

It has to be taken into account, otherwise the equations would be truncated. This would introduce a new phenomenon in the gyrokinetic reduction, which would depart from its clean setting of being just a change of coordinates with an ignorable fast gyro-angle. In addition, such a truncation would be legitimate only when the removed term is small, which would generate additional assumptions in the theory.

For the future, a first point to be clarified is the smallness of the removed term in physically relevant situations, especially because they could be comparable to other terms that are kept in some applications of gyrokinetics. In a similar way as in the previous episode, a first answer about this question could come from a post-processing of the out-coming data in present-day numerical simulations.

However, even when the assumptions for the truncation are verified, the Hamiltonian character is not guaranteed for the truncated system, because it is generally suppressed by truncations in the equations of motion.

This joins a more general question, which constitutes a natural extension of the work. In practical applications of gyrokinetics, and especially in numerical simulations [42, 56], not the full series is taken into account for the gyrokinetic equations or transformation. Only a few terms are kept, for instance all first-order terms and a part of the second-order terms.

From abstract arguments, one can guess that most of such truncations spoil the Hamiltonian structure, whereas only few ones respect it and do generate a Hamiltonian dynamics. This structure is important for gyrokinetic codes, which are interested in the long-term behaviour of plasma turbulence, and for which an important role is played by conservative structures, included in the Hamiltonian character.

The results of this episode can be pursued to answer such questions. They provide an interesting framework to identify good truncations, suited to the Hamiltonian structure of gyrokinetics.

A most important extension concerns the conservation laws for gyrokinetics, and especially the

conservation of angular momentum, since it was one of the motivations for the work [21, 127, 133]. Previous results already studied the question, but it is interesting to come back to it, because it should be enlightened by the results of this episode, since the Noether theorem makes conservation laws easy to study in Hamiltonian systems. Again, this extension should consider first the general results for gyrokinetic theory, together with a comparison with previous results on the subject, and then the consequences on practical applications, for instance on numerical codes.

From another point of view, this work was a good opportunity to develop reduction tools, with the exploration of both a lifting method and a Hamiltonian perturbation theory in field dynamics.

Concerning the lifting, it should be extended in order to include in the coordinate transformation also a dependence on the moments of the Vlasov density. On another hand, it would be interesting to apply it to other systems. We had in mind the adoption of magnetic coordinates for fluid models, but we were not sure it would be so interesting. Perhaps other kinds of Vlasov-like systems should be considered first, e.g. the case of the oscillating-center, or gyrokinetics with several species or with spin effects, as was initiated in Chapter 5.

As for field Hamiltonian perturbation theory, it provided new insights on the links between plasma dynamics and particle dynamics, which both legitimated from a Hamiltonian point of view the two-stage gyrokinetic reduction, and opened an interesting domain for KAM-like theories in infinite dimensional systems. However, first studies about this last application did not induce interesting effects compared to particle perturbation theory, and further explorations are needed.

More generally, this episode perfectly exemplifies the interest of reformulating dynamical reductions in the framework of Hamiltonian reductions, in order to better understand the reduced Hamiltonian structure and to validate or rectify the reduced dynamics. This will be the main idea of the next episode of the dissertation.

This episode will also be necessary to definitely conclude about the Hamiltonian structure of gyrokinetics, because a subtle question presently remains open. The lifting process used in the standard derivation of gyrokinetics was developed in this episode to show how the guiding-center and gyro-center transformations of particle dynamics indeed induce a Hamiltonian transformed structure for Vlasov-Maxwell, which verifies the requirements of the gyrokinetic reduction, both the averaging and the magnetic moment requirement (provided the time-dependence of the transformation is taken into account, for instance in the Lie-transform of the phase-space Lagrangian). This achieves Steps 2a-2c announced in the general introduction on page 16. As a result, the characteristics of the Vlasov equation occur in leaves of constant magnetic moment, and their fast time scale is isolated into the gyro-angle dimension, which does not influence the last four dimensions.

However, this is not exactly gyrokinetics. In order to completely remove the fast time scale and to benefit from the dimensional reduction, one needs to drop the gyro-angle dimension, and to consider the magnetic moment as just an index, not a dimension of the base space. It is why Step 2d announced on page 16 is needed in order to really obtain the gyrokinetic equations, which occur over the slow four-dimensional base-space of guiding-center dynamics $(\overline{\mathbf{q}}, \overline{\phi})$.

In usual derivations working at the level of the equations of motion, this step is trivial. From the point of view of Hamiltonian reductions (i.e. in order for the resulting dynamics to be Hamiltonian), it is not so obvious, especially because bracket truncations do not preserve the Hamiltonian structure in general [58]. It is why in the present document, it is treated as a definite step of the gyrokinetic reduction.

So, in order to conclude about the Hamiltonian structure of gyrokinetics, the removal of the two superfluous dimensions (gyro-angle and magnetic moment) from the base space has to be studied in details in the framework of Hamiltonian reductions. This is very similar to what occurred in guiding-center theory in Chapter 2, when going from the transformed six-dimensional dynamics to the slow four-dimensional dynamics, where the method of Dirac's constraints [44] was noticed as essential to preserve the Hamiltonian character. Accordingly, this method could play a role in the quest for the Hamiltonian structure of gyrokinetics. Its applications in plasma physics will be the topic of the next episode.

## Selected bibliography

[5] I. Bialynicki-Birula, J. C. Hubbard, and L. A. Turski, "Gauge-independent canonical formulation of relativistic plasma theory", Physica A **128**, 509 (1984).

[17] A. J. Brizard and T. S. Hahm, "Foundations of nonlinear gyrokinetic theory", Rev. Mod. Phys. **79**, 421 (2007).

[19] A. J. Brizard and A. Mischenko, "Guiding-center recursive Vlasov and Lie-transform methods in plasma physics", J. Plas. Phys. **75**, 675 (2009).

[21] A. J. Brizard and N. Tronko, "Exact momentum conservation laws for the gyrokinetic Vlasov-Poisson equations", Phys. Plasmas **18**, 082307 (2011).

[49] E. A. Frieman and L. Chen, "Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria", Phys. Fluids **25**, 502 (1982).

[64] T. S. Hahm, W. W. Lee, and A. J. Brizard, "Nonlinear gyrokinetic theory for finite-beta plasmas", Phys. Fluids **31**, 1940 (1988).

[65] T. S. Hahm, "Nonlinear gyrokinetic equations for tokamak microturbulence", Phys. Fluids **31**, 2670 (1988).

[89] R. G. Littlejohn, "Linear relativistic gyrokinetic equation", Phys. Fluids **27**, 976 (1984).

[93] M. Marklund and P. J. Morrison, "Gauge-free Hamiltonian structure of the spin Maxwell–Vlasov equations", Phys. Lett. A **375**, 23622365 (2011).

[95] J. E. Marsden and A. Weinstein, "The Hamiltonian structure of the Maxwell-Vlasov equations", Physica D **4**, 394 (1982).

[104] P. J. Morrison, "Poisson brackets for fluids and plasmas", AIP Conference Proceedings **88**, 13 (1982).

[114] P. J. Morrison, "A general theory for gauge-free lifting", Phys. Plasmas **20**, 012104 (2013).

[115] P. J. Morrison, M. Vittot, and L. de Guillebon, "Magnetic moment type of lifting from particle dynamics to Vlasov-Maxwell dynamics", Phys. Plasmas **20**, 032109 (2013).

[127] F. I. Parra and P. J. Catto, "Transport of momentum in full f gyrokinetics", Phys. Plasmas **17**, 056106 (2010).

[133] B. Scott and J. Smirnov, "Energetic consistency and momentum conservation in the gyrokinetic description of tokamak plasmas", Phys. Plasmas **17**, 112302 (2010).

[152] R. Utiyama, "Theory of invariant variation and the generalized canonical dynamics", Progr. Theor. Phys. Suppl. **9**, 19 (1959).

# Episode III

# Hamiltonian reductions in plasma physics

# Introduction of the episode

This last episode can be seen from two complementary points of view.

First, it is a natural extension of the previous episode, which showed that the Hamiltonian structure of gyrokinetics can be addressed by lifting the reduction of particle dynamics, but which also suggested that one can consider obtaining it directly from the Vlasov-Maxwell system with purely Hamiltonian reduction tools. The relevance of such an idea is enhanced by the fact that the lift did not exactly provide the gyrokinetic equations since its Vlasov density remained defined over a six-dimensional space; a Hamiltonian method is called for in the last step of this reduction (denoted as Step 2d on page 16), in order to actually remove the two superfluous dimensions (gyro-angle and magnetic moment) from the transformed Vlasov-Maxwell dynamics while preserving the Hamiltonian character.

Chapter 6 started to develop a Hamiltonian perturbation method for the gyrokinetic reduction, but it did not give straightforward results, and other methods can be considered at the same time. A particular attention must be drawn on Dirac's theory of constraints [44], since in some cases it produced just a bracket truncation, i.e. a removal of some contributions from the Poisson bracket. In addition, it plays an essential role in guiding-center theory (see Chapter 2), which is closely related to gyrokinetics. The idea is also supported by a recent work [144], which derives a simplified version of gyrokinetics using Dirac's constraints.

Now, the gyrokinetic reduction does not fit in with a direct application of Dirac's theory as it was previously involved in plasma physics, especially because it somehow concerns dimensions of the base space rather than a field coordinate. So, the idea is to look for extensions or variations of Dirac's reduction that could be applied to the gyrokinetic reduction, in particular by making the method closer to a bracket truncation. In order to identify such extensions, good ideas can be suggested by studying other instances of reduction. These last are easy to find in plasma physics, which is characterized by a wide variety of models and hence of dynamical reductions.

This leads us to the second point of view on this third episode, which departs from the mere gyrokinetic reduction, and rather takes place in the general framework of plasma models. Indeed, most derivations of plasma models work at the level of the equations of motion, which can generate fake dissipation in the derivation [108]. Usually, the Hamiltonian structure is verified to be preserved or not only after the derivation. As we mentioned in the general introduction and as was fully confirmed in the previous episode, it may be difficult to answer this question and to identify the corrections needed to make the model Hamiltonian. It is often more enlightening to derive the reduction using purely Hamiltonian reduction methods. In addition, even when the reduced model is known to be Hamiltonian, it is interesting to investigate the corresponding Hamiltonian reduction, because it provides a better understanding on the links between the models and their Hamiltonian structures, and it is a way to apply and develop Hamiltonian reduction methods, which will help derive new models whose Hamiltonian character will be guaranteed.

Here again, our work primarily focuses on Dirac's theory of constrained systems, because previous works [31–33,118,119] showed that they play a role in several Hamiltonian derivations of plasma models. Starting from a (Hamiltonian) physical system and imposing constraints suggested by the physics, a straightforward application of Dirac's theory provided exactly the expected physical reduced model.

Extensions or developments on the method are expected because this straightforward applica-

tion proceeded by a sequence of computations but it obtained a simple and intuitive result, just corresponding to removing the constrained variables from the dynamical field coordinates. It is interesting to study these examples of reduction, in order to identify the mechanism at work in this simplification, especially because it could generate an extension of the method, or at least make it softer to use.

After clarifying the simplifying/softening phenomena for some Hamiltonian reductions already identified in the literature, it will be interesting to consider other examples of reductions, and to see how they can be derived using Hamiltonian methods, and whether they fit with Dirac's reduction, or with its "softened" version.

As primary candidates, we will turn to the main models of plasma dynamics mentioned in the general introduction, Vlasov-Maxwell, Euler-Maxwell, magnetohydrodynamics, together with variations of these models. Especially, the reduction to Euler-Maxwell, which is called the fluid reduction, is interesting to consider because it relies on a closure assumption and to our knowledge, no Hamiltonian closure including second or higher moments of the Vlasov density as dynamical field variables exists in the literature.

There does not seem to be an obstruction for such a Hamiltonian model, which would include the pressure tensor as a dynamical variable. The relevance of such a purpose is confirmed by the papers [136,138]. More generally, it is enhanced in warm or hot plasmas, where kinetic effects mean that some moments higher than one have to be taken into account. Including the second moment is especially attractive since in this case the Hamiltonian functional would not be affected by the reduction process.

The work will proceed in four steps:

- First, in Chapter 7, the simplifying/softening phenomenon will be made more precise. When viewing Dirac's reduction as a projection of derivatives, then it will be shown that in some cases Dirac's projector can be strongly simplified and become just the trivial projector expected from intuition. Instances will be studied in the reduction to incompressible magnetohydrodynamics and to Vlasov-Poisson. By the way, this chapter will be concerned with a circumstantial question about Poisson brackets that satisfy the Jacobi identity only conditionally, because the use of projectors can remedy their difficulty.

- Second, the presence of the simplified projector will be explained by a quarter-canonical structure in the Poisson bracket, and then formulated as a bracket truncation, related to a subalgebra reduction. In addition to softening the method, this will impact and clarify some aspects of Dirac's procedure. Various examples will be considered about Vlasov-Maxwell or Euler-Maxwell. Finally, we will come back to the gyrokinetic reduction, whose step 2d will appear as a fine application of the developed methods.

This part of the work will be mainly reported in Chapter 8, but some of the results will be reported in the appendix chapters 12 and 13.

- Third, in Chapter 9, an application to the fluid reduction will be considered. The standard closure at order one for the Euler-Maxwell system turns out to be just given by a subalgebra, and closures including moments of order two or higher will be considered. Obstructions for such Hamiltonian closures will be identified, and a model from the literature, which was believed to be Hamiltonian, will turn out to be non-Hamiltonian.

- As an additional application, the reduction from Euler-Maxwell to magnetohydrodynamics will be studied in the appendix chapter 14. It will turn out to be a subalgebra reduction as well.

# Chapter 7

# On the use of projectors for Hamiltonian systems and their relationship with Dirac brackets

in collaboration with Cristel Chandre, Aurore Back, Emanuele Tassi, and Phil J. Morrison
Paper published in J. Phys. A **46**, 125203 (2013)

*A*bstract: The role of projectors associated with Poisson brackets of constrained Hamiltonian systems is analysed. Projectors act in two instances in the bracket: in the explicit dependence on the variables and in the computation of the functional derivatives. They are shown to play a role both in order to make the Poisson bracket satisfy the Jacobi identity unconditionally and in bracket reductions or extensions.
The projectors are investigated by using Dirac's theory of constrained Hamiltonian systems, whose reduction procedure can be formulated as a projection of functional derivatives. It is observed that the Dirac projector can often be replaced by a simplified projector, which just corresponds to extracting from the Poisson bracket the part that concerns the reduced variables.
The results are illustrated by examples taken from plasma physics in magnetohydrodynamics and in the Vlasov-Maxwell system.

## Introduction

As a first study about Dirac's reduction in plasma physics, we focus on previously identified examples of such reductions, related to the inclusion of incompressibility in magnetohydrodynamics and fluid dynamics. The interest is that these examples seem to involve some subtleties that could usefully be understood with more details. Indeed, after computing Dirac's bracket, what comes out is just an intuitive result, which corresponds to extracting the incompressible part of the Poisson bracket.

In this chapter, we characterize this phenomenon with a formulation of Dirac's constraints relying on a projection of functional derivatives. This is a first step towards identifying the properties that justify the phenomenon, which will be the purpose of the next chapter. In addition, this formulation already has interesting applications on other plasma models, for instance in the reduction to Vlasov-Poisson and in an extension of the Vlasov-Maxwell system.

On another hand, the presence of projectors can be useful to improve some Hamiltonian plasma models suffering from a conditional Jacobi identity. Actually, this can be viewed as another independent motivation for this chapter, and will be explained with more details in the next section.

The organization of the chapter is the following. In Sec. 7.1, the concept of conditional Jacobi identity for constrained Hamiltonian systems will be introduced. In Sec. 7.2, the Dirac method will be reminded and formulated as a projection of functional derivatives; semi-local constraints will be considered, which will be useful for the gyrokinetic reduction in the next chapter, for instance. In Sec. 7.3, an application of the method will be done for magnetohydrodynamics, in order to

untaint its bracket and to make the dynamics incompressible. In Sec. 7.4, we will turn to the Vlasov-Maxwell system, in order to untaint its bracket, derive the Vlasov-Poisson reduced model, and derive an extended model where two Casimir invariants of Vlasov-Maxwell system become dynamical.

## 7.1 Constrained and tainted brackets

We consider an arbitrary Poisson bracket of a Poisson algebra of functionals of field variables $\boldsymbol{\chi}(\mathbf{x})$ given by

$$\{F, G\} = \int d^n x \, F_{\boldsymbol{\chi}} \cdot \mathbb{J}(\boldsymbol{\chi}) \cdot G_{\boldsymbol{\chi}} \,, \tag{7.1}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\boldsymbol{\chi}\colon \mathbb{R}^n \to \mathbb{R}^d$, and $F_{\boldsymbol{\chi}} \cdot \mathbb{J} \cdot G_{\boldsymbol{\chi}} = F_{\chi^i} \, \mathbb{J}^{ij} \, G_{\chi^j}$ with repeated indices summed. By Poisson algebra we mean a Lie algebra realization on functionals with an associative product of functionals that satisfies the Leibniz law. Also, we assume that the resulting equations of motion given by $\dot{\boldsymbol{\chi}} = \{\boldsymbol{\chi}, H\}$, for some Hamiltonian functional $H[\boldsymbol{\chi}]$, possess a conservation law $\mathcal{Q}[\boldsymbol{\chi}] = 0$, where $\mathcal{Q}$ is a functional of the field variables and their derivatives. Here we address the specific case where these conservation laws are obtained regardless of the choice of Hamiltonian $H$, so $\mathcal{Q} = 0$ is an intrinsic property of the bracket of the Poisson algebra.

There are two ways to define such a constrained Poisson algebra. The usual way is to place a restriction on the set of field variables $\boldsymbol{\chi}$ in the Poisson algebra. However, this definition raises the question of how to appropriately compute the constrained functional derivatives $F_{\boldsymbol{\chi}}$. The second way is to define a Poisson algebra that does not include any constraint on the field variables and, consequently, there is no ambiguity in defining the functional derivatives – conservation laws such as $\mathcal{Q} = 0$ take the form of Casimir invariants.

In this chapter we investigate the links between these two ways of defining constrained Hamiltonian structures, and we propose a way to lift Poisson structures defined via the constrained field variables approach to ones that have the constraints as Casimir invariants. As can be expected, the difficulty resides in assuring the validity of the Jacobi identity. If we keep the same Poisson bracket but extended to the bigger algebra (the one without any constraint on the field variables), then in general, the Poisson structure is only obtained when the constraint is satisfied, i.e., the Jacobi identity is satisfied conditionally when $\mathcal{Q}[\boldsymbol{\chi}] = 0$. It turns out that one can remedy this limitation by modifying the bracket with the inclusion of suitable projectors that leave the functional derivatives unconstrained and guarantee the Jacobi identity unconditionally. We identify such projectors acting on the functional derivatives and on the explicit dependence of the bracket on $\boldsymbol{\chi}$. We discuss the various choices of projectors and highlight a particularly relevant one obtained from Dirac's theory of constrained Hamiltonian systems.

In order to illustrate our purpose, consider the relatively simple and common example, the vorticity equation of a compressible or incompressible fluid in $\mathbb{R}^3$. The vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{u}$, with $\mathbf{u}$ the velocity field, satisfies

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = \nabla \times (\mathbf{u} \times \boldsymbol{\omega}). \tag{7.2}$$

In terms of a commonly used Poisson bracket (see, e.g., Ref. [159]),

$$\{F, G\}_0 = \int d^3 x \, \boldsymbol{\omega} \cdot (\nabla \times F_{\boldsymbol{\omega}}) \times (\nabla \times G_{\boldsymbol{\omega}}) \,, \tag{7.3}$$

Eq. (7.2) has the from $\dot{F} = \{F, H\}_0$ with the Hamiltonian $H = \int d^3 x \, v^2/2$. Here and in what follows, we suppose that the boundary conditions are such that no surface terms appear in subsequent calculations which, e.g., would be the case on a periodic box or all space. If one forgets about the constraint on the vector fields $\boldsymbol{\omega}$ or if one wants to lift the algebra of functionals of divergence-free $\boldsymbol{\omega}$ to the algebra of functionals of any vector field $\boldsymbol{\omega}$, then the bracket (7.3) does not satisfy the Jacobi identity. This is easily seen by the following counterexample:

$$F_1 = \frac{1}{2} \int d^3 x \, \boldsymbol{\omega} \cdot \hat{\mathbf{x}} \, y^2 \,, \quad F_2 = \frac{1}{2} \int d^3 x \, \boldsymbol{\omega} \cdot \hat{\mathbf{y}} \, z^2 \,, \quad F_3 = \int d^3 x \, \boldsymbol{\omega} \cdot \hat{\mathbf{z}} \, x \,,$$

which yields

$$\{F_1, \{F_2, F_3\}_0\}_0 + \circlearrowleft = -\int d^3x\, \boldsymbol{\omega} \cdot \nabla(yz) \neq 0\,.$$

So, the bracket (7.3) satisfies the Jacobi identity only if $\nabla\cdot\boldsymbol{\omega} = 0$. We refer to such Poisson brackets that only satisfy the Jacobi identity conditionally as *tainted brackets*. One of the questions we address in this chapter is how to correct a tainted bracket so that it satisfies the Jacobi identity unconditionally. For this particular example, the correction is obtained by inserting a projection operator, following Ref. [33], given by[1] $\mathcal{P}_\perp = 1 - \nabla\Delta^{-1}\nabla\cdot$, so that it defines a new bracket

$$\{F, G\} = \int d^3x\, (\mathcal{P}_\perp \boldsymbol{\omega}) \cdot (\nabla \times F_{\boldsymbol{\omega}}) \times (\nabla \times G_{\boldsymbol{\omega}}).$$

It is rather straightforward (see Ref. [33]) to show that this bracket satisfies the Jacobi identity unconditionally. We notice that $\nabla\cdot\boldsymbol{\omega}$ is a Casimir invariant of the modified bracket, i.e. $\{\nabla\cdot\boldsymbol{\omega}, G\} = 0$ for any functional $G$. Here $\nabla\cdot\boldsymbol{\omega}$ is viewed as a functional using the formula $\nabla\cdot\boldsymbol{\omega}(\mathbf{x}) = \int d^3x' \nabla\cdot\boldsymbol{\omega}(\mathbf{x}')\delta(\mathbf{x}' - \mathbf{x})$.

As mentioned above, projectors are not only useful to lift algebras so as to satisfy the Jacobi identity, they are also involved in the way functional derivatives are computed when the field variables are constrained. As an illustration, we consider the incompressible Euler equation for the velocity field $\mathbf{u}(\mathbf{x}, t)$,

$$\dot{\mathbf{u}} = -\mathbf{u}\cdot\nabla\mathbf{u} - \nabla P,$$

where $P$ is determined by the constraint $\nabla\cdot\mathbf{u} = 0$. This equation has a Hamiltonian structure [2, 70, 73, 98] given by the Hamiltonian $H[\mathbf{u}] = \int d^3x\, v^2/2$ and the Poisson bracket

$$\{F, G\} = \int d^3x\, \mathbf{u}\cdot[F_{\mathbf{u}}, G_{\mathbf{u}}]_L,$$

where $F_{\mathbf{u}}$ are the functional derivatives of an observable $F$ with respect to the field variable $\mathbf{u}$ and the Lie bracket $[\mathbf{V}, \mathbf{W}]_L$ is given by

$$[\mathbf{V}, \mathbf{W}]_L = (\mathbf{W}\cdot\nabla)\mathbf{V} - (\mathbf{V}\cdot\nabla)\mathbf{W}.$$

It should be noted that the incompressible Euler equation cannot be directly obtained from $\dot{F} = \{F, H\}$ using unconstrained functional derivatives $F_{\mathbf{u}}$ since $\nabla\cdot\mathbf{u} = 0$ would not be conserved by the flow. One way of correcting the bracket is to use an orthogonal projector [70]. For divergence-free fields, this orthogonal projector is again given by $\mathcal{P}_\perp = 1 - \nabla\Delta^{-1}\nabla\cdot$ (see also Refs. [33, 159]). In other words, the constrained functional derivative $F_{\mathbf{u}}$ must be computed such that it satisfies $\nabla\cdot F_{\mathbf{u}} = 0$. However, the fundamental reason for this constraint on the functional derivative is unclear, even though it yields the correct equation of motion. For a more general constraint $\mathcal{Q}[\boldsymbol{\chi}] = 0$, is it still the orthogonal projector that has to be used for the constrained functional derivatives? In addition, this projector is in general not unique. It therefore raises natural questions such as which is the most relevant projector and how is it obtained in a systematic way?

In this chapter, we investigate two possible placements of a projectors: one is on the explicit dependence on the field variables, while the other is on the computation of the functional derivatives. We clarify the choice of the relevant projector by using Dirac's theory of constrained Hamiltonian systems. In order to prove the relevance of these projectors, we consider two examples taken from plasma physics. The first one is magnetohydrodynamics (MHD), both compressible and incompressible, the second one is about the Vlasov-Maxwell system.

The goal of this chapter is to present a general method which highlights the role of appropriate projectors, and identifies a particular projector using a reformulation of Dirac's theory. From this general method, we show that the tainted brackets can be corrected such that the new brackets satisfy the Jacobi identity unconditionally. In addition, we connect these corrected brackets to the ones obtained from Dirac's theory of constrained Hamiltonian systems.

[1] Note that the inversion of the Laplacian implies to specify the boundary conditions. For the projector $\nabla\Delta^{-1}\nabla\cdot$, the Laplacian stands between a divergence and a gradient, so the Fourier component $\mathbf{k} = 0$ is automatically zero.

## 7.2 Formulation of the general method

### 7.2.1 Projected functional derivatives

At the outset we assume that the bracket (7.1) is a Poisson bracket on the algebra of functionals of $\boldsymbol{\chi}$, where $\boldsymbol{\chi}$ denotes a $d$-tuple of fields such that $\mathcal{Q}[\boldsymbol{\chi}](\mathbf{x}) = 0$ and $\mathcal{Q}[\boldsymbol{\chi}]$ is function of $\boldsymbol{\chi}$ and its derivatives. These fields will be referred to as $\mathcal{Q}$-free fields. In this section, our aim is to get a corresponding Poisson bracket on the algebra of any functionals of $\boldsymbol{\chi}$, satisfying $\mathcal{Q}[\boldsymbol{\chi}](\mathbf{x}) = 0$ or not. The functional derivatives $\bar{F}_{\boldsymbol{\chi}}$ are defined in the following way:

$$\delta F = \int d^n x\, \bar{F}_{\boldsymbol{\chi}} \cdot \delta\boldsymbol{\chi}, \tag{7.4}$$

for all $\mathcal{Q}$-free $\delta\boldsymbol{\chi}$, which here means that $\mathbb{Q}\delta\boldsymbol{\chi} = 0$ where $\mathbb{Q}$ is the Fréchet derivative of $\mathcal{Q}$ defined by

$$\mathcal{Q}[\boldsymbol{\chi} + \delta\boldsymbol{\chi}](\mathbf{x}) - \mathcal{Q}[\boldsymbol{\chi}](\mathbf{x}) = \mathbb{Q}\delta\boldsymbol{\chi} + O(\|\delta\boldsymbol{\chi}\|^2).$$

This means that $\bar{F}_{\boldsymbol{\chi}}$ is not uniquely defined: it is arbitrary up to an element of Rg $\mathbb{Q}^\dagger$, since $\int d^n x\, \bar{F}_{\boldsymbol{\chi}} \cdot \delta\boldsymbol{\chi} = \int d^n x\, (\bar{F}_{\boldsymbol{\chi}} + \mathbb{Q}^\dagger \mathbf{g}) \cdot \delta\boldsymbol{\chi}$ where $\mathbf{g}$ is arbitrary. We define the constrained functional derivative $\bar{F}_{\boldsymbol{\chi}}$ from the unconstrained one $F_{\boldsymbol{\chi}}$ by the following equation:

$$\int d^n x\, \bar{F}_{\boldsymbol{\chi}} \cdot \delta\boldsymbol{\chi} = \int d^n x\, F_{\boldsymbol{\chi}} \cdot \delta\bar{\boldsymbol{\chi}}, \tag{7.5}$$

where now $\delta\bar{\boldsymbol{\chi}}$ is the constrained ($\mathcal{Q}$-free) variation and $\delta\boldsymbol{\chi}$ the unconstrained one. For the unconstrained variation $\delta\boldsymbol{\chi}$, we use a linear operator $\mathcal{P}$ acting as $\delta\bar{\boldsymbol{\chi}} = \mathcal{P}^\dagger \delta\boldsymbol{\chi}$ such that $\mathbb{Q}\mathcal{P}^\dagger = 0$. Moreover, the range of this operator $\mathcal{P}^\dagger$ should be Ker $\mathbb{Q}$ and, in addition, $\mathcal{P}^\dagger$ should act as the identity on Ker $\mathbb{Q}$. This is equivalent to requiring that $\mathcal{P}$ be a projector. Consequently, this leads to a condition on the possible projectors $\mathcal{P}$ such that $\bar{F}_{\boldsymbol{\chi}} = \mathcal{P} F_{\boldsymbol{\chi}}$, viz.

$$\text{Ker } \mathcal{P} = \text{Rg } \mathbb{Q}^\dagger. \tag{7.6}$$

Note that given this condition, $\mathcal{Q}[\boldsymbol{\chi}](\mathbf{x})$ is a Casimir invariant that is naturally preserved by the flow. Still this projector is not unique. In the literature (see, e.g., Ref. [70]), the functional derivative is chosen such that $\mathbb{Q}F_{\boldsymbol{\chi}} = 0$, so that the projector satisfies $\mathbb{Q}\mathcal{P} = 0$. This corresponds to the orthogonal projector

$$\mathcal{P}_\perp = 1 - \mathbb{Q}^\dagger (\mathbb{Q}\mathbb{Q}^\dagger)^{-1} \mathbb{Q}, \tag{7.7}$$

provided $\mathbb{Q}\mathbb{Q}^\dagger$ is invertible on Rg $\mathbb{Q}$. However it is not clear if it is the best choice for the projection. Other solutions satisfy

$$\mathcal{P}_\perp \mathcal{P} = \mathcal{P}_\perp,$$
$$\mathcal{P}\mathcal{P}_\perp = \mathcal{P},$$

which are needed in order to satisfy Eq. (7.6). Given a particular projector $\mathcal{P}$ the bracket (7.1) becomes

$$\{F, G\}_{\mathrm{t}} = \int d^n x\, (\mathcal{P} F_{\boldsymbol{\chi}}) \cdot \mathbb{J}(\boldsymbol{\chi}) \cdot (\mathcal{P} G_{\boldsymbol{\chi}}), \tag{7.8}$$

where now the functional derivatives are the unconstrained ones. We have released the constraint on the functional derivatives but, in general the Poisson bracket (7.8) does not satisfy the Jacobi identity for functionals of arbitrary $\boldsymbol{\chi}$, i.e., ones no longer restricted to $\mathcal{Q}$-free fields. This is because $\mathbb{J}(\boldsymbol{\chi})$ may give contributions that do not satisfy the Jacobi identify when $\mathcal{Q}[\boldsymbol{\chi}] \neq 0$. However, if the projector $\mathcal{P}$ does not depend on the field variables $\boldsymbol{\chi}$, as is the case for the examples we deal with in this chapter, then a bracket that satisfies the Jacobi identity for all functionals of $\boldsymbol{\chi}$, satisfying $\mathcal{Q}[\boldsymbol{\chi}] = 0$ or not, is given by

$$\{F, G\} = \int d^n x\, (\mathcal{P} F_{\boldsymbol{\chi}}) \cdot \mathbb{J}(\mathcal{P}\boldsymbol{\chi}) \cdot (\mathcal{P} G_{\boldsymbol{\chi}}). \tag{7.9}$$

In order to verify the Jacobi identity, we perform the change of variables $\boldsymbol{\chi}_P = \mathcal{P}\boldsymbol{\chi}$ and $\boldsymbol{\chi}_Q = \boldsymbol{\chi} - \mathcal{P}\boldsymbol{\chi}$ so that bracket (7.9) formally becomes bracket (7.8) with $\boldsymbol{\chi}_P$ instead of $\boldsymbol{\chi}$. Since $\boldsymbol{\chi}_P$ is by definition $\mathcal{Q}$-free, the Jacobi identity is satisfied. For the Poisson bracket (7.9), we notice that $\mathcal{Q}[\boldsymbol{\chi}](\mathbf{x})$ is a Casimir invariant, and that the equations of motion for $\boldsymbol{\chi}_P$ are identical to the ones given by the Poisson bracket (7.1) or (7.8).

### 7.2.2 Dirac brackets

#### Local constraints

In order to identify the most appropriate projector, we use Dirac's theory of constrained Hamiltonian systems [44, 112]. We begin with the following good Poisson bracket:

$$\{F, G\} = \int d^n x \, F_{\boldsymbol{\chi}} \cdot \mathbb{J}(\boldsymbol{\chi}) \cdot G_{\boldsymbol{\chi}} \tag{7.10}$$

and then impose the local constraint $\Phi(\mathbf{x}) := \mathcal{Q}[\boldsymbol{\chi}](\mathbf{x}) = 0$, where as before $\mathcal{Q}[\boldsymbol{\chi}](\mathbf{x})$ is a function of $\boldsymbol{\chi}(\mathbf{x})$ and its derivatives. The Dirac procedure begins with the computation of the matrix of Poisson brackets between the local constraints,

$$C(\mathbf{x}, \mathbf{x}') \equiv \{\Phi(\mathbf{x}), \Phi(\mathbf{x}')\} = \mathbb{Q}\mathbb{J}\mathbb{Q}^\dagger \delta(\mathbf{x} - \mathbf{x}').$$

We set $\mathbb{C} := \mathbb{Q}\mathbb{J}\mathbb{Q}^\dagger$ and we assume that this quantity is invertible. Then, the Dirac correction to the bracket (7.10) is given by

$$-\int\int d^n x \, d^n x' \, \{F, \Phi(\mathbf{x})\} D(\mathbf{x}, \mathbf{x}') \{\Phi(\mathbf{x}'), G\},$$

where $D(\mathbf{x}, \mathbf{x}') = \mathbb{C}^{-1}(\boldsymbol{\chi}(\mathbf{x}))\delta(\mathbf{x} - \mathbf{x}')$. Since $\{F, \Phi(\mathbf{x})\} = -\mathbb{Q}\mathbb{J} \cdot F_{\boldsymbol{\chi}}$, this contribution is equal to

$$-\int d^n x \, F_{\boldsymbol{\chi}} \cdot \mathbb{J}\mathbb{Q}^\dagger \mathbb{C}^{-1} \mathbb{Q}\mathbb{J} \cdot G_{\boldsymbol{\chi}}.$$

Therefore, the Dirac bracket is given by

$$\{F, G\}_* = \int d^n x \, F_{\boldsymbol{\chi}} \cdot \mathbb{J}_*(\boldsymbol{\chi}) \cdot G_{\boldsymbol{\chi}}, \tag{7.11}$$

where

$$\mathbb{J}_* = \mathbb{J} - \mathbb{J}\mathbb{Q}^\dagger \mathbb{C}^{-1} \mathbb{Q}\mathbb{J}. \tag{7.12}$$

It is straightforward to verify that $\mathbb{J}_*$ given by Eq. (7.12) is antisymmetric because $\mathbb{C}$ is antisymmetric. We notice that $\mathbb{Q}\mathbb{J}_* = 0$ (and therefore $\mathbb{J}_*\mathbb{Q}^\dagger = 0$). As a consequence, the constraint $\Phi$ is a Casimir invariant. We notice that a sufficient but not necessary condition to define the Dirac bracket (7.11) is that $\mathbb{C}$ is invertible on the range of $\mathbb{Q}$. If $\mathbb{C}$ is not invertible (neither globally nor in the range of $\mathbb{Q}$, the matrix $\mathbb{C}^{-1}$ has to be chosen according to the condition

$$\mathbb{J}\mathbb{Q}^\dagger(\mathbb{C}^{-1}\mathbb{C} - 1) = 0, \tag{7.13}$$

in order to obtain the constraints as Casimir invariants.

The Poisson brackets obtained by the Dirac procedure are Poisson brackets of the form (7.8) but untainted, i.e., they satisfy the Jacobi identity unconditionally even though they are not of the form (7.9) in general. This can be seen by considering a projector $\mathcal{P}$ as discussed in the previous section. Under the assumption that Ker $\mathcal{P} = \mathrm{Rg}\ \mathbb{Q}^\dagger$, we deduce that $\mathbb{J}_*(1 - \mathcal{P}) = 0$, and consequently:

$$\mathbb{J}_* = \mathcal{P}^\dagger \mathbb{J}_* \mathcal{P}.$$

With this equality, the Poisson bracket becomes

$$\{F, G\}_* = \int d^n x \, (\mathcal{P} F_{\boldsymbol{\chi}}) \cdot \mathbb{J}_*(\boldsymbol{\chi}) \cdot (\mathcal{P} G_{\boldsymbol{\chi}}).$$

The additional feature is that, a priori, the Poisson matrix $\mathbb{J}_*$ is a function of both $\mathcal{P}\boldsymbol{\chi}$ and $(1-\mathcal{P})\boldsymbol{\chi}$. However, it is straightforward to check that $(1 - \mathcal{P})\boldsymbol{\chi}$ is a Casimir invariant.

The Dirac procedure shows that among the possible projectors $\mathcal{P}$ satisfying Eq. (7.6), one turns out to be most convenient. The matrix $\mathbb{J}_*$ can be rewritten using the Dirac projector

$$\mathcal{P}_* = 1 - \mathbb{Q}^\dagger \mathbb{C}^{-1} \mathbb{Q}\mathbb{J}, \tag{7.14}$$

as

$$\mathbb{J}_* = \mathcal{P}_*^\dagger \mathbb{J} \mathcal{P}_*,$$

so that the Dirac bracket becomes the same as the original one (7.10) with the exception that the functional derivatives are projected using the Dirac projector,

$$\{F, G\}_* = \int d^n x \, (\mathcal{P}_* F_{\boldsymbol{\chi}}) \cdot \mathbb{J}(\boldsymbol{\chi}) \cdot (\mathcal{P}_* G_{\boldsymbol{\chi}}), \tag{7.15}$$

where we notice that the Poisson matrix is $\mathbb{J}$ and not $\mathbb{J}_*$. The main difference between the the orthogonal projector $\mathcal{P}_\perp$ and the Dirac projector $\mathcal{P}_*$ is that $\mathcal{P}_\perp$ is a purely geometric object since it only depends on the constraints, and $\mathcal{P}_*$ is a dynamical object since it involves the Poisson matrix.

*Remark:* We observe that the matrix corresponding to the Dirac bracket has the following property:

$$\mathbb{J}_* = \mathcal{P}_*^\dagger \mathbb{J} \mathcal{P}_* = \mathbb{J} \mathcal{P}_* = \mathcal{P}_*^\dagger \mathbb{J},$$

i.e., the Dirac bracket can be rewritten from Eq. (7.15) using only one Dirac projector, e.g.,

$$\{F, G\}_* = \int d^n x \, F_{\boldsymbol{\chi}} \cdot \mathbb{J}(\boldsymbol{\chi}) \cdot \mathcal{P}_* G_{\boldsymbol{\chi}}.$$

As a result, the computation of the Dirac bracket is made easier.

### Semi-local constraints

The calculation of Sec. 7.2.2 can be generalized to allow semi-local constraints in phase space. To this end we split the set of coordinates into two pieces, i.e., $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$ where $\mathbf{x}_1 \in \mathbb{R}^{n-m}$ and $\mathbf{x}_2 \in \mathbb{R}^m$. The semi-local constraints are given by

$$\Phi(\mathbf{x}_1) = \bar{\mathcal{Q}}[\boldsymbol{\chi}](\mathbf{x}_1) = \int d^m x_2 \, \mathcal{Q}[\boldsymbol{\chi}](\mathbf{x}),$$

where $\mathcal{Q}[\boldsymbol{\chi}](\mathbf{x})$ is a function of $\boldsymbol{\chi}(\mathbf{x})$ and its derivatives. The linear operator $\bar{\mathbb{Q}}$ is defined by the linear operator associated with the function $\mathcal{Q}$ by

$$\bar{\mathbb{Q}} = \int d^m x_2 \, \mathbb{Q}.$$

Since $\bar{\mathbb{Q}}$ acting on a function of $\mathbf{x}$ is only a function of $\mathbf{x}_1$, the linear operator $\bar{\mathbb{Q}}^\dagger$ is defined by the equation

$$\int d^{n-m} x_1 \, \bar{\mathbb{Q}} \boldsymbol{\chi} \cdot \mathbf{w}(\mathbf{x}_1) = \int d^n x \, \boldsymbol{\chi}(\mathbf{x}) \cdot \bar{\mathbb{Q}}^\dagger \mathbf{w}.$$

Consequently, $\bar{\mathbb{Q}}^\dagger$ is a linear operator acting on functions of $\mathbf{x}_1$ as $\mathbb{Q}^\dagger$, i.e., $\bar{\mathbb{Q}}^\dagger w(\mathbf{x}_1) = \mathbb{Q}^\dagger w(\mathbf{x}_1)$. In a manner similar to that of Sec. 7.2.2, the computation of the Dirac bracket shows that the operator

$$\mathbb{C} = \bar{\mathbb{Q}} \mathbb{J} \bar{\mathbb{Q}}^\dagger,$$

must be invertible. More explicitly, the linear operator $\mathbb{C}$ acts on functions of $\mathbf{x}_1$ as

$$\mathbb{C} w(\mathbf{x}_1) = \int d^m x_2 \mathbb{Q} \mathbb{J} \mathbb{Q}^\dagger w(\mathbf{x}_1).$$

The expression of the Dirac projector is given by

$$\mathcal{P}_* = 1 - \bar{\mathbb{Q}}^\dagger \mathbb{C}^{-1} \bar{\mathbb{Q}} \mathbb{J},$$

in a very similar way as the case of the local constraints. We notice that the linear operator $\mathbb{C}$ only needs to be invertible on $\operatorname{Rg} \bar{\mathbb{Q}}$. Another important projector is the orthogonal projector given by

$$\mathcal{P}_\perp = 1 - \bar{\mathbb{Q}}^\dagger (\bar{\mathbb{Q}} \bar{\mathbb{Q}}^\dagger)^{-1} \bar{\mathbb{Q}}.$$

As in the case of local constraints, these two projectors satisfy $\mathbb{J}_* = \mathcal{P}^\dagger \mathbb{J}_* \mathcal{P}$, along with the two properties $\mathcal{P}_\perp \mathcal{P}_* = \mathcal{P}_\perp$ and $\mathcal{P}_* \mathcal{P}_\perp = \mathcal{P}_*$. In addition, the Dirac projector satisfies $\mathbb{J}_* = \mathcal{P}_*^\dagger \mathbb{J} \mathcal{P}_* = \mathcal{P}_*^\dagger \mathbb{J} = \mathbb{J} \mathcal{P}_*$.

## 7.3 Example 1: magnetohydrodynamics

### 7.3.1 Compressible magnetohydrodynamics

A particularly interesting example is afforded by the Hamiltonian structure of magnetohydrodynamics. The equations for the velocity field $\mathbf{u}(\mathbf{x},t)$, the density $\rho(\mathbf{x},t)$, the magnetic field $\mathbf{B}(\mathbf{x},t)$, and the entropy $s(\mathbf{x},t)$ are given by

$$\begin{aligned}
&\dot{\rho} = -\nabla\cdot(\rho\mathbf{u}),\\
&\dot{\mathbf{u}} = -\mathbf{u}\cdot\nabla\mathbf{u} - \rho^{-1}\nabla(\rho^2 U_\rho) + \rho^{-1}(\nabla\times\mathbf{B})\times\mathbf{B},\\
&\dot{\mathbf{B}} = \nabla\times(\mathbf{u}\times\mathbf{B}),\\
&\dot{s} = -\mathbf{u}\cdot\nabla s,
\end{aligned}$$

where $U$ is the internal energy and $U_\rho$ here denotes the partial derivative of $U$ with respect to $\rho$. The dynamical variables are $\rho(\mathbf{x})$, $\mathbf{u}(\mathbf{x})$, $\mathbf{B}(\mathbf{x})$ and $s(\mathbf{x})$ where $\mathbf{x}$ belongs to $\mathbb{R}^3$. The observables of the system are functionals of these vector fields, denoted generically by $F(\rho,\mathbf{u},\mathbf{B},s)$. In these coordinates, this system has the following Hamiltonian

$$H(\rho,\mathbf{u},\mathbf{B},s) = \int d^3x\left(\frac{1}{2}\rho\mathbf{u}^2 + \rho U(\rho,s) + \frac{\mathbf{B}^2}{2}\right).$$

There are two slightly different Poisson brackets that have been proposed in Refs. [102–104]. A first one was given in Ref. [102],

$$\begin{aligned}
\{F,G\}=-\int d^3x\,&\left[F_\rho\nabla\cdot G_\mathbf{u} + F_\mathbf{u}\cdot\nabla G_\rho - \rho^{-1}(\nabla\times\mathbf{u})\cdot(F_\mathbf{u}\times G_\mathbf{u})\right.\\
&\left.+\rho^{-1}\nabla s\cdot(F_s G_\mathbf{u} - F_\mathbf{u} G_s)\right] + \{F,G\}_B,
\end{aligned}\tag{7.16}$$

where the magnetic part $\{F,G\}_B$ of the Poisson bracket is chosen as $\{F,G\}_B = \{F,G\}_{B,\mathrm{t}}$

$$\{F,G\}_{B,\mathrm{t}} = -\int d^3x\,\rho^{-1}\left(F_\mathbf{u}\cdot\mathbf{B}\times(\nabla\times G_\mathbf{B}) - G_\mathbf{u}\cdot\mathbf{B}\times(\nabla\times F_\mathbf{B})\right).\tag{7.17}$$

It was pointed out in Ref. [104] that this bracket satisfies the Jacobi identity only when $\nabla\cdot\mathbf{B}=0$, and also that $\nabla\cdot\mathbf{B}$ commutes with any other functionals, i.e., $\{F,\nabla\cdot\mathbf{B}\}=0$ for all $F$ (it is a Casimir-like property, even though we cannot call it a Casimir invariant since the Jacobi identity is only satisfied when $\nabla\cdot\mathbf{B}=0$). As was the case for the vorticity equation (7.2), the functional derivatives with respect to $\mathbf{B}$ must be divergence-free for coherence. However, we notice that here, since only $\nabla\times F_\mathbf{B}$ are involved in the expression of the magnetic part (7.16) of the Poisson bracket, it does not make any difference whether $F_\mathbf{B}$ is divergence-free or not.

In order to extend the definition of the Poisson bracket to functionals of any $\mathbf{B}$, ones not necessarily divergence-free, a second Poisson bracket was proposed in Refs. [103, 104]. There the magnetic part of the Poisson bracket (7.16) was replaced by

$$\begin{aligned}
\{F,G\}_{B,1}=-\int d^3x\,&\left[\left(\rho^{-1}F_\mathbf{u}\cdot[\nabla G_\mathbf{B}] - \rho^{-1}G_\mathbf{u}\cdot[\nabla F_\mathbf{B}]\right)\cdot\mathbf{B}\right.\\
&\left.+\mathbf{B}\cdot\left([\nabla\left(\rho^{-1}F_\mathbf{u}\right)]\cdot G_\mathbf{B} - [\nabla\left(\rho^{-1}G_\mathbf{u}\right)]\cdot F_\mathbf{B}\right)\right].
\end{aligned}$$

Here the notation $\mathbf{a}\cdot[M]\cdot\mathbf{b}$ is a scalar explicitly given by $\sum_{ij} a_i M_{ij} b_j$ for any vectors $\mathbf{a}$ and $\mathbf{b}$ and any matrix $[M]$. It was shown that this bracket satisfies the Jacobi identity for all functionals of $(\rho,\mathbf{u},s,\mathbf{B})$ regardless of the condition $\nabla\cdot\mathbf{B}=0$. The magnetic part of this Poisson bracket is rewritten as

$$\begin{aligned}
\{F,G\}_{B,1}=&-\int d^3x\,\rho^{-1}\left[F_\mathbf{u}\cdot\mathbf{B}\times(\nabla\times G_\mathbf{B}) - G_\mathbf{u}\cdot\mathbf{B}\times(\nabla\times F_\mathbf{B})\right]\\
&+\int d^3x\,\rho^{-1}\nabla\cdot\mathbf{B}\left(F_\mathbf{u}\cdot G_\mathbf{B} - F_\mathbf{B}\cdot G_\mathbf{u}\right).
\end{aligned}\tag{7.18}$$

The first line of the above bracket corresponds to the Poisson bracket introduced in Ref. [102] [see Eq. (7.17)]. With the additional terms (proportional to $\nabla\cdot\mathbf{B}$) the Jacobi identity is unconditionally satisfied for any functionals of $(\rho, \mathbf{u}, s, \mathbf{B})$. However, a property of the bracket (7.16) with the magnetic part (7.17) has been lost, $\nabla\cdot\mathbf{B}$ does not Poisson-commute with any functional, so it is not a Casimir invariant.

In order to have both the Jacobi identity unconditionally satisfied and $\nabla\cdot\mathbf{B}$ a Casimir invariant, we apply the prescription (7.9) on the magnetic part (7.17). At every instance in the Poisson bracket where $\mathbf{B}$ is explicitly mentioned, we replace $\mathbf{B}$ with $\bar{\mathbf{B}} = \mathbf{B} - \nabla\Delta^{-1}\nabla\cdot\mathbf{B}$. The magnetic part becomes

$$\{F, G\}_B = -\int d^3x\, \rho^{-1} \left(F_{\mathbf{u}} \cdot (\bar{\mathbf{B}} \times (\nabla \times G_{\mathbf{B}}) - G_{\mathbf{u}} \cdot (\bar{\mathbf{B}} \times (\nabla \times F_{\mathbf{B}}))\right),$$

and it is rewritten as

$$\begin{aligned}\{F, G\}_B = &-\int d^3x\, \rho^{-1} \left(F_{\mathbf{u}} \cdot (\mathbf{B} \times (\nabla \times G_{\mathbf{B}}) - G_{\mathbf{u}} \cdot (\mathbf{B} \times (\nabla \times F_{\mathbf{B}}))\right) \\ &+ \int d^3x\, \nabla\cdot\mathbf{B}\, \Delta^{-1}\nabla\cdot\left(\rho^{-1}F_{\mathbf{u}} \times (\nabla \times G_{\mathbf{B}}) - \rho^{-1}G_{\mathbf{u}} \times (\nabla \times F_{\mathbf{B}})\right). \end{aligned} \tag{7.19}$$

Here we notice that the correction term still contains terms proportional to $\nabla\cdot\mathbf{B}$ but is different from the one in Eq. (7.18). The main difference is that $\nabla\cdot\mathbf{B}$ is not a Casimir invariant for the Poisson bracket (7.18) whereas it is one for the Poisson bracket (7.19) since it only involves terms like $\nabla \times G_{\mathbf{B}}$.

### 7.3.2 Incompressible magnetohydrodynamics

For incompressible MHD we begin with the equations for compressible magnetohydrodynamics from Sec. 7.3.1 and apply constraints. The Poisson bracket given by Eqs. (7.16)-(7.19) is of the form (7.10) with

$$\mathbb{J} = \begin{pmatrix} 0 & -\nabla\cdot & 0 & 0 \\ -\nabla & -\rho^{-1}(\nabla\times\mathbf{u})\times & -\rho^{-1}\bar{\mathbf{B}}\times(\nabla\times) & \rho^{-1}\nabla s \\ 0 & -\nabla\times(\rho^{-1}\bar{\mathbf{B}}\times) & 0 & 0 \\ 0 & -\rho^{-1}\nabla s & 0 & 0 \end{pmatrix}.$$

We impose the following local constraints on the field variables $\boldsymbol{\chi} = (\rho, \mathbf{u}, \mathbf{B}, s)$,

$$\mathcal{Q}[\boldsymbol{\chi}](\mathbf{x}) = (\rho, \nabla\cdot\mathbf{u}).$$

The reduction to incompressible MHD using Dirac's theory has already been done in Ref. [33] and the reduction to incompressible Euler equation in Refs. [118, 119]. Here we propose a more compact way to present this reduction using the operators introduced in the previous sections. The expressions of the intermediate operators are

$$\mathbb{Q} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \nabla\cdot & 0 & 0 \end{pmatrix},$$

$$\mathbb{Q}^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & -\nabla \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\mathbb{C} = \begin{pmatrix} 0 & \Delta \\ -\Delta\nabla\cdot(\rho^{-1}(\nabla\times\mathbf{u})\times\nabla) & \end{pmatrix},$$

$$\mathbb{C}^{-1} = \begin{pmatrix} \Delta^{-1}\nabla\cdot(\rho^{-1}(\nabla\times\mathbf{u})\times\nabla) & -\Delta^{-1} \\ \Delta^{-1} & 0 \end{pmatrix}.$$

The orthogonal projector is given by Eq. (7.7) and its expression is

$$\mathcal{P}_\perp F_{\boldsymbol{\chi}} = (0, \bar{F}_{\mathbf{u}}, F_{\mathbf{B}}, F_s),$$

where $\bar{F}_{\mathbf{u}} = F_{\mathbf{u}} - \nabla\Delta^{-1}\nabla\cdot F_{\mathbf{u}}$. The Dirac projector, computed from the Poisson bracket (7.16) where $\mathbf{B}$ has been replaced by $\bar{\mathbf{B}} = \mathbf{B} - \nabla\Delta^{-1}\nabla\cdot\mathbf{B}$, is given by

$$\mathcal{P}_* F_{\boldsymbol{\chi}} = (F_*, \bar{F}_{\mathbf{u}}, F_{\mathbf{B}}, F_s),$$

where

$$F_* = \Delta^{-1}\nabla\cdot\left(\rho^{-1}\left((\nabla\times\mathbf{u})\times\bar{F}_{\mathbf{u}} - \bar{\mathbf{B}}\times(\nabla\times F_{\mathbf{B}}) - F_s\nabla s\right)\right).$$

We notice that the two projectors differ in the first component. Even though the two projectors $\mathcal{P}_\perp$ and $\mathcal{P}_*$ are different, both of these projectors satisfy the equation $\mathbb{J}_* = \mathcal{P}^\dagger\mathbb{J}\mathcal{P}$, which is always the case for the Dirac projector but not true in general for the orthogonal projector. Actually any projector $\mathcal{P}F_{\boldsymbol{\chi}} = (F_*(\bar{F}_{\mathbf{u}}, F_{\mathbf{B}}, F_s), \bar{F}_{\mathbf{u}}, F_{\mathbf{B}}, F_s)$ satisfies $\mathbb{J}_* = \mathcal{P}^\dagger\mathbb{J}\mathcal{P}$ for any function $F_*$. The first component is thus irrelevant, and consequently the orthogonal projector is the simplest projector to be used for constrained functional derivatives. From this projector, we compute the Dirac bracket from Eq. (7.15), and it gives the same bracket as that produced in Ref. [33]:

$$\begin{aligned}\{F,G\}_* = \int d^3x\,\rho^{-1}\Big(&(\nabla\times\mathbf{u})\cdot(\bar{F}_{\mathbf{u}}\times\bar{G}_{\mathbf{u}}) - \nabla s\cdot(F_s\bar{G}_{\mathbf{u}} - \bar{F}_{\mathbf{u}}G_s)\\ &+\bar{\mathbf{B}}\cdot(\bar{F}_{\mathbf{u}}\times(\nabla\times G_{\mathbf{B}}) + (\nabla\times F_{\mathbf{B}})\times\bar{G}_{\mathbf{u}})\Big),\end{aligned}$$

where $\bar{F}_{\mathbf{u}} = F_{\mathbf{u}} - \nabla\Delta^{-1}\nabla\cdot F_{\mathbf{u}}$.

## 7.4 Example 2: Vlasov-Maxwell equations

### 7.4.1 Vlasov-Maxwell modified bracket as a Dirac bracket

As a second example, we consider the Vlasov-Maxwell equations for the distribution of charged particles in phase space $f(\mathbf{x},\mathbf{v},t)$ and the electromagnetic fields $\mathbf{E}(\mathbf{x},t)$ and $\mathbf{B}(\mathbf{x},t)$ given by

$$\begin{aligned}\dot{f} &= -\mathbf{v}\cdot\nabla f - (\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot\partial_{\mathbf{v}} f,\\ \dot{\mathbf{E}} &= \nabla\times\mathbf{B} - \mathbf{J},\\ \dot{\mathbf{B}} &= -\nabla\times\mathbf{E},\end{aligned}$$

where $\mathbf{J} = \int d^3v\,\mathbf{v}f$. The Hamiltonian of this system is given by

$$H = \int d^6z\, f\frac{v^2}{2} + \int d^3x\,\frac{\mathbf{E}^2+\mathbf{B}^2}{2},$$

where we denote $\mathbf{z} = (\mathbf{x},\mathbf{v})$. The Poisson bracket between two functionals of $f(\mathbf{x},\mathbf{v})$, $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ is given by

$$\begin{aligned}\{F,G\}_{\mathrm{t}} = &\int d^6z\, f\left([F_f,G_f]_{\mathrm{c}} + [F_f,G_f]_{\mathrm{B}} + G_{\mathbf{E}}\cdot\partial_{\mathbf{v}}F_f - F_{\mathbf{E}}\cdot\partial_{\mathbf{v}}G_f\right)\\ &+\int d^3x\,\left(F_{\mathbf{E}}\cdot\nabla\times G_{\mathbf{B}} - \nabla\times F_{\mathbf{B}}\cdot G_{\mathbf{E}}\right),\end{aligned} \tag{7.20}$$

where the two brackets $[\cdot,\cdot]_{\mathrm{c}}$ and $[\cdot,\cdot]_{\mathrm{B}}$ are defined by

$$[f,g]_{\mathrm{c}} = \nabla f\cdot\partial_{\mathbf{v}}g - \partial_{\mathbf{v}}f\cdot\nabla g, \tag{7.21}$$

$$[f,g]_{\mathrm{B}} = \mathbf{B}\cdot(\partial_{\mathbf{v}}f\times\partial_{\mathbf{v}}g). \tag{7.22}$$

The Poisson bracket (7.20) was proposed in Ref. [101], except the second term given in Ref. [95] (see also Ref. [5]) which removed an obstruction to the Jacobi identity. However, it was pointed out in Ref. [104] that the Poisson bracket (7.20) only satisfies the Jacobi identity when $\nabla\cdot\mathbf{B} = 0$, which is to say that it does not satisfy the Jacobi identity for arbitrary functionals of $(f,\mathbf{E},\mathbf{B})$ (see Ref. [114] for the details of the direct proof of the Jacobi identity up to this condition). This problem is actually already present in the Lagrangian description (for the dynamics of charged

particles) since $[\cdot,\cdot]_{\rm c}+[\cdot,\cdot]_{\rm B}$ only satisfies the Jacobi identity for functions $\mathbf{B}$ such that $\nabla\cdot\mathbf{B}=0$, whereas, individually, $[\cdot,\cdot]_{\rm c}$ and $[\cdot,\cdot]_{\rm B}$ satisfy the Jacobi identity for an arbitrary function $\mathbf{B}$.

In order to remedy this problem, we modify the bracket $[\cdot,\cdot]_{\rm B}$ to take the form of (7.9),

$$[f,g]_{{\rm B}_P}=(\mathbf{B}-\nabla\Delta^{-1}\nabla\cdot\mathbf{B})\cdot(\partial_{\mathbf{v}}f\times\partial_{\mathbf{v}}g).$$

With this modified gyro-magnetic bracket, we readily check that $[\cdot,\cdot]_{\rm c}+[\cdot,\cdot]_{{\rm B}_P}$ satisfies the Jacobi identity. Next, we consider the modified Poisson bracket (7.20) obtained by replacing $[\cdot,\cdot]_{\rm B}$ by $[\cdot,\cdot]_{{\rm B}_P}$, i.e., we consider the Poisson bracket

$$\begin{aligned}\{F,G\}_{\rm VM}=&\int d^6z\, f\left([F_f,G_f]_{\rm c}+[F_f,G_f]_{{\rm B}_P}+G_{\mathbf{E}}\cdot\partial_{\mathbf{v}}F_f-F_{\mathbf{E}}\cdot\partial_{\mathbf{v}}G_f\right)\\&+\int d^3x\,\left(F_{\mathbf{E}}\cdot\nabla\times G_{\mathbf{B}}-\nabla\times F_{\mathbf{B}}\cdot G_{\mathbf{E}}\right),\end{aligned}\tag{7.23}$$

which satisfies the Jacobi identity unconditionally. This follows from the change of variable $\mathbf{B}_P=\mathbf{B}-\nabla\Delta^{-1}\nabla\cdot\mathbf{B}$ and $\mathbf{B}_Q=\nabla\Delta^{-1}\nabla\cdot\mathbf{B}$ where it should be noted that

$$\nabla\times G_{\mathbf{B}}=\nabla\times G_{\mathbf{B}_P},$$

since the operator $\mathcal{P}=1-\nabla\Delta^{-1}\nabla\cdot$ satisfies $\mathcal{P}\nabla\times=\nabla\times$.

Here it should be noticed that $\nabla\cdot\mathbf{B}$ is a Casimir invariant for the Poisson bracket (7.23). The untainted form of the Vlasov-Maxwell bracket (7.23) gives the Hamiltonian structure of the Vlasov-Maxwell equations in terms of physical fields without introducing the vector potential, i.e., without the restriction of $\nabla\cdot\mathbf{B}=0$.

In order to realize the link between brackets defined using projectors and Dirac brackets, we show below that the Poisson bracket (7.23) is a Dirac bracket of some parent bracket obtained using two constraints which, by definition, are Casimir invariants of the bracket (7.23)

$$\mathcal{Q}[f,\mathbf{E},\mathbf{B}](\mathbf{x})=(\nabla\cdot\mathbf{E}-\rho,\nabla\cdot\mathbf{B}),$$

where $\rho=\int d^3v\, f$. As expected there is an infinite number of solutions for the parent bracket. A family of solutions is given by

$$\{F,G\}=\{F,G\}_{\rm VM}+\int d^3x\,\left(\nabla\cdot F_{\mathbf{B}}\mathcal{D}\nabla\cdot G_{\mathbf{E}}-\nabla\cdot F_{\mathbf{E}}\mathcal{D}^\dagger\nabla\cdot G_{\mathbf{B}}\right),\tag{7.24}$$

where $\mathcal{D}$ is a linear operator independent of the field variables, so that the Jacobi identity is guaranteed by Morrison's lemma of Ref. [104]. This statement uses the fact that the Vlasov-Maxwell bracket has been made untainted; it would not be true if the original tainted Vlasov-Maxwell bracket (7.20) was considered instead of the Poisson bracket (7.23).

Now, if we apply the Dirac procedure on the extended Poisson bracket (7.24) with the primary constraint $\nabla\cdot\mathbf{E}-\rho$, we get the secondary constraint $\nabla\cdot\mathbf{B}$, and the reduced Dirac bracket is obtained from $\mathbb{J}_*=\mathcal{P}_*^\dagger\mathbb{J}\mathcal{P}_*$ where $\mathcal{P}_*$ is the Dirac projector (7.14). The Dirac projector can be explicitly computed. However, in order to further simplify the computation of the Dirac bracket, we use the orthogonal projector since, as in the case of incompressible MHD (see Sec. 7.3.2), it satisfies the same relation as the Dirac projector, i.e., $\mathbb{J}_*=\mathcal{P}_*^\dagger\mathbb{J}\mathcal{P}_*=\mathcal{P}_\perp^\dagger\mathbb{J}\mathcal{P}_\perp$, where

$$\mathcal{P}_\perp F_{\boldsymbol{\chi}}=(F_{\mathbf{B}}-\nabla\Delta^{-1}\nabla\cdot F_{\mathbf{B}},F_{\mathbf{E}},F_f).$$

This implies the expected result that the Vlasov-Maxwell bracket (7.23) is the Dirac bracket of the bracket (7.24) with Dirac constraints $(\nabla\cdot\mathbf{E}-\rho,\nabla\cdot\mathbf{B})$.

With the extended bracket (7.24), the Casimir invariants $(\nabla\cdot\mathbf{E}-\rho,\nabla\cdot\mathbf{B})$ of the Vlasov-Maxwell system now have dynamics given by

$$\begin{aligned}\frac{\partial}{\partial t}(\nabla\cdot\mathbf{E}-\rho)&=\Delta\mathcal{D}^\dagger\nabla\cdot\mathbf{B},\\ \frac{\partial}{\partial t}\nabla\cdot\mathbf{B}&=-\Delta\mathcal{D}\nabla\cdot\mathbf{E}.\end{aligned}$$

We notice that here $\nabla \cdot \mathbf{B}$ and $\nabla \cdot \mathbf{E} - \rho$ are no longer constant since they are no longer Casimir invariants of the extended bracket (7.24). However even though $\nabla \cdot \mathbf{E} - \rho$ is not zero, the total charge remains conserved (i.e., $\int d^6 z f$ is still a Casimir invariant). The above equations suggest two particularly interesting choices for our still undetermined operator $\mathcal{D}$. Defining $\mathcal{D} = \Delta^{-1}$ gives to $\nabla \cdot \mathbf{E} - \rho$ and $\nabla \cdot \mathbf{B}$ the dynamics of stationary waves when $\rho = 0$, whereas defining $\mathcal{D} = (-\Delta)^{-1/2}$ gives them the dynamics of propagating waves. We note that these operators always act on divergences of vector fields.

*Remark:* As a side note, we point out that the choice of $\mathcal{D} = (-\Delta)^{-1/2}$ naturally exhibits the operator $\nabla * := \nabla(-\Delta)^{-1/2}\nabla\cdot$ which corresponds to $\nabla\times$ for the compressible part of a vector field. Indeed, the operator

$$-\nabla * \Delta^{-1}\nabla * = \nabla\Delta^{-1}\nabla\cdot,$$

is the orthogonal projector onto the kernel of $\nabla\times$, just as $-\nabla \times \Delta^{-1}\nabla\times = 1 - \nabla\Delta^{-1}\nabla\cdot$ is the complementary projector onto the kernel of $\nabla *$. With this choice, the resulting dynamical equations associated with the Poisson bracket (7.24) for the solenoidal and the compressible parts of the electromagnetic fields become independent and similar:

$$\begin{aligned} \Box\mathbf{E}_S &= -\dot{\mathbf{J}}_S, & \Box\mathbf{E}_C &= -\dot{\mathbf{J}}_C, \\ \Box\mathbf{B}_S &= \nabla \times \ \mathbf{J}_S, & \Box\mathbf{B}_C &= \nabla * \ \mathbf{J}_C, \end{aligned}$$

where $\Box$ is the d'Alembert operator $\Box = \partial^2/\partial t^2 - \Delta$ and $\boldsymbol{\psi}_S$ is the solenoidal part of the vector field $\boldsymbol{\psi}$, i.e., $\boldsymbol{\psi}_S = -\nabla \times \Delta^{-1}\nabla \times \boldsymbol{\psi} = (1 - \nabla\Delta^{-1}\nabla) \cdot \boldsymbol{\psi}$ and $\boldsymbol{\psi}_C$ is its compressible part, which is $\boldsymbol{\psi}_C = -\nabla * \Delta^{-1}\nabla * \boldsymbol{\psi} = \nabla\Delta^{-1}\nabla \cdot \boldsymbol{\psi}$. In the absence of matter, the fields $\boldsymbol{\psi}_S$ and $\boldsymbol{\psi}_C$ propagate as independent free waves.

### 7.4.2 From Vlasov-Maxwell to Vlasov-Poisson equations

In order to obtain Vlasov-Poisson equations from the Vlasov-Maxwell equations we impose two constraints:

$$\mathcal{Q}[f, \mathbf{E}, \mathbf{B}](\mathbf{x}) = (\mathbf{B} - \mathbf{B}_0(\mathbf{x}), \nabla \times \mathbf{E}),$$

where $\mathbf{B}_0$ is a non-uniform background magnetic field. The operators $\mathbb{Q}$ and $\mathbb{Q}^\dagger$ are given by

$$\mathbb{Q} = \begin{pmatrix} 0 & \nabla\times & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and

$$\mathbb{Q}^\dagger = \begin{pmatrix} 0 & 0 \\ \nabla\times & 0 \\ 0 & 1 \end{pmatrix}.$$

The orthogonal projector $\mathcal{P}_\perp$ given by Eq. (7.7) is given by

$$\mathcal{P}_\perp = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \nabla\Delta^{-1}\nabla\cdot & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Contrary to the orthogonal projector, the expression of the Dirac projector depends on the dynamics, and in particular on the Poisson matrix $\mathbb{J}$ which is given by

$$\mathbb{J} = \begin{pmatrix} -[f,\cdot] & -\partial_{\mathbf{v}} f & 0 \\ -f\partial_{\mathbf{v}} & 0 & \nabla\times \\ 0 & -\nabla\times & 0 \end{pmatrix},$$

where the small bracket $[\cdot,\cdot]$ is given by $[\cdot,\cdot] = [\cdot,\cdot]_c + [\cdot,\cdot]_{B_P}$ with these two brackets given by Eqs. (7.21)-(7.22). The operator $\mathbb{C}$ is given by

$$\mathbb{C} = \begin{pmatrix} 0 & (\nabla\times)^2 \\ -(\nabla\times)^2 & 0 \end{pmatrix}.$$

The operator $\mathbb{C}$ is not invertible; however, the Dirac procedure still applies as explained in Sec. 7.2.2 with a choice for $\mathbb{C}^{-1}$ given by

$$\mathbb{C}^{-1} = \begin{pmatrix} 0 & \Delta^{-1} \\ -\Delta^{-1} & 0 \end{pmatrix},$$

so that Eq. (7.13) is satisfied. As a result, the Dirac projector is computed,

$$\mathcal{P}_* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \nabla\Delta^{-1}\nabla\cdot & 0 \\ -\Delta^{-1}\nabla\times f\partial_{\mathbf{v}} & 0 & \nabla\Delta^{-1}\nabla\cdot \end{pmatrix}.$$

We notice that both projectors $\mathcal{P}_\perp$ and $\mathcal{P}_*$ satisfy the equation $\mathbb{J}_* = \mathcal{P}^\dagger \mathbb{J}\mathcal{P}$ and the Poisson matrix of the Vlasov-Poisson equations is given by

$$\mathbb{J}_* = \begin{pmatrix} -[f,\cdot] & -\nabla\Delta^{-1}\nabla\cdot\partial_{\mathbf{v}} f 0 \\ -\nabla\Delta^{-1}\nabla\cdot(f\partial_{\mathbf{v}}) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It leads to the expression of the Poisson bracket,

$$\{F,G\}_* = \int d^6z\, f[F_f - \Delta^{-1}\nabla\cdot F_{\mathbf{E}}, G_f - \Delta^{-1}\nabla\cdot G_{\mathbf{E}}]. \tag{7.25}$$

Like for the incompressible MHD equations, even if the Dirac and orthogonal projectors are different, both of them can be used to compute the Dirac bracket from the Poisson matrix $\mathbb{J}$, the orthogonal projector being slightly simpler and more straightforward to compute.

Let us mention that the orthogonal projector is not always available, as is emphasized by the example of linear Vlasov with quasi-neutrality in [34], which also constitutes an interesting example of semi-local constraints. In this case, since the orthogonal projector does not exist, it cannot be a solution for the computation of constrained functional derivatives. A convenient choice is afforded by the Dirac projector.

## Conclusion

In this chapter, the Dirac reduction, viewed as a projection of derivatives, appeared as useful in order to correct tainted brackets and to make them satisfy the Jacobi identity unconditionally. This proceeded through an extension of the phase space and a projection of functional derivatives.

In addition to making the Vlasov-Maxwell bracket and the MHD bracket untainted, it found applications in Dirac reductions, such as the reduction for incompressible MHD and for Vlasov-Poisson. In addition, it suggested an extension of the Vlasov-Maxwell system that includes a dynamical $\nabla\cdot\mathbf{B}$ and $\nabla\cdot\mathbf{E}-\rho$.

More important for the purpose of this episode, this formulation made more definite the phenomenon that affects the Dirac reduction for incompressibility in fluid-like models, which consists in replacing the Dirac projector by a simplified (orthogonal) projector.

This observation is interesting both in order to avoid unnecessary computations in the derivation, and in our quest for extensions or variations of Dirac reductions. It is not specific to the reduction for incompressibility, since it was involved in other plasma models, for instance the Vlasov-Poisson reduction mentioned here. On the other hand, it is not completely general, as shown by the reduction including quasi-neutrality in a linear Vlasov system in [34], where the orthogonal projector cannot be used.

Here, the phenomenon was observed only a posteriori; in order to use it a priori in derivations of models and avoid useless computations, its origin, as well as its extent and meaning, should be identified in a clearer way. This will be the topic of the next chapter.

# Chapter 8

# On the role of quarter-canonical Poisson brackets in Dirac reductions and gyrokinetics

in collaboration with Cristel Chandre

*A*bstract: Reduced Poisson brackets derived by Dirac's theory of constraints are related to a projection of derivatives, but they can also often be obtained by another simplified projector, which makes the reduction equivalent to a mere bracket truncation. We show that this phenomenon is connected to a special structure of the initial bracket, which is quarter-canonical.
This structure is close to the canonical one, and also related to the Darboux theorem, but it is less stiff. It is frequent even in continuous Hamiltonian systems and is involved as soon as half of the constraints is not coupled with itself. It makes Dirac reductions more efficient by removing the need to compute the inverse of the matrix of constraints, and by enclosing the reduction in the simplified truncation projector. Consequences on the Dirac procedure and links with a subalgebra reduction are mentioned, as well as the role of the coordinates in the phenomenon.
Various applications are considered for reductions or extensions of Poisson brackets in plasma models. Especially, an important step of the gyrokinetic reduction, namely the removal of the gyro-angle dimension, is shown to fit with a quarter-canonical Dirac reduction, which seems to play a crucial role to guarantee a Hamiltonian structure for gyrokinetics.

## Introduction

In this chapter, we investigate the mechanism and the structures at work in the simplification observed in some Dirac reductions in the previous chapter. Then, we will come back to our initial purpose, the gyrokinetic reduction, and relate the removal of its superfluous dimensions (gyro-angle and magnetic moment) to a Hamiltonian reduction, and especially to a Dirac reduction.

Indeed, the previous chapter showed that when deriving Hamiltonian reduced models by using Dirac's theory of constraints [44, 66], the reduced bracket is obtained by a projector on derivatives given by the inverse of the matrix of constraints, but that afterwards, it often appears that some terms involved in the Dirac projector are irrelevant. Removing them induces the resulting projector to become greatly simplified and seemingly very intuitive, but the reduced bracket to be unchanged. This suggests to look for a reduction method based on this simplified projector instead of the Dirac projector.

With the simplified projector, one does not need to invert the matrix of constraints, which not only makes the computations easier but also could make the method more general by removing the condition about the invertibility of the matrix of constraints. However, this projector might

not exist, or its definition might not be so obvious. The question is to identify in what conditions it is available, and why it is so often available in Dirac reductions in plasmas. In this chapter, as well as in the complementary appendix chapters 12 and 13, we relate its existence to the presence of a quarter-canonical structure in the Poisson bracket, and we explore the simplified reduction methods induced by its existence.

The organization of this part of the thesis work is the following. Sec. 8.1 is devoted to an introductory example. In Sec. 8.2, the answer to the initial question is shown to be related to a quarter-canonical structure in the initial bracket. This structure is connected to the canonical and semi-canonical structures, but it has much weaker requirements. Sec. 8.3 will summarize some complementary results reported in the appendix chapters 12 and 13. The first of them will investigate the incidence of quarter-canonical coordinates. They will appear as frequent, and otherwise most often constructible. On another hand, they can be avoided by obtaining explicitly the simplified projector even in non-quarter-canonical coordinates. As for Chapter 13, it will study consequences of the quarter-canonical structure on the Dirac procedure, and will make links with another reduction method, based on a projection onto a Lie-subalgebra. Last, in the light of the previous results, we will come back to the gyrokinetic reduction in Sec. 8.4.

In this chapter, we will often consider Poisson brackets for continuous Hamiltonian systems. They will always be assumed to involve only local couplings as is usual in fluids and plasma models:

$$\{F,G\} = \int d\mathbf{x} \int d\mathbf{y}\, \delta(\mathbf{x}-\mathbf{y}) F_{\psi_i(\mathbf{x})}\, \mathbb{J}'^{ij}(\mathbf{x},\mathbf{y},\boldsymbol{\psi}(\mathbf{x}),\boldsymbol{\psi}(\mathbf{y}))\, G_{\psi_j(\mathbf{y})} = \int d\mathbf{x}\, F_{\psi_i(\mathbf{x})}\, \mathbb{J}^{ij}(\mathbf{x},\boldsymbol{\psi}(\mathbf{x}))\, G_{\psi_j(\mathbf{x})}\,,$$

where all the symbols are generic: $\boldsymbol{\psi}$ is a set of fields, defined over a space with coordinates $\mathbf{x}$ (or also $\mathbf{y}$). The symbol $\mathbb{J}$ denotes a matrix whose coefficients can depend on both the coordinates and the fields, it can be an operator, i.e. involve derivatives of the fields to any order. In addition, we identify the Poisson bracket $\{\ ,\ \}$ and the local-interaction matrix $\mathbb{J}$.

On another hand, the constraints, denoted by $\boldsymbol{\phi}$ will be local or semi-local (except in Sec. 8.4), i.e. they involve the fields at each point of the space:

$$f(\mathbf{x},\boldsymbol{\psi}(\mathbf{x}),\nabla\boldsymbol{\psi}(\mathbf{x}),...) = 0\,,$$

for some function $f$. For an example, see Eqs. (8.3)-(8.4), or Subsec. 8.2.4). For our purpose, it is the natural analogue in continuous media of the usual constraints in finite-dimensional systems.

## 8.1 Introductory example

Throughout the chapter, the purpose will be illustrated with various Hamiltonian reductions. The warm fluid model for plasmas will be used as a framework for most of the examples. The dynamical fields are the plasma mass density $\rho$, the plasma fluid velocity $\mathbf{u}$, the plasma entropy per unit mass $s$, the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$; the Hamiltonian function is simply the total energy

$$H[\rho,\mathbf{u},\mathbf{E},\mathbf{B}] = \int d^3\mathbf{x}\, \left(\frac{\rho\mathbf{u}^2}{2} + \rho U(\rho,s) + \frac{\mathbf{E}^2+\mathbf{B}^2}{2}\right), \tag{8.1}$$

where $U(\rho,s)$ is the internal energy per unit mass of the plasma. The Poisson bracket is given by (see Ref. [104])

$$\begin{aligned}\{F,G\} = \int d^3\mathbf{x}\, \Big[ & G_{\mathbf{u}}\cdot\nabla F_\rho - F_{\mathbf{u}}\cdot\nabla G_\rho + \tfrac{\nabla\times\mathbf{u}}{\rho}\cdot\ (F_{\mathbf{u}}\times G_{\mathbf{u}}) \\ & + G_s\, F_{\mathbf{u}}\cdot\tfrac{\nabla s}{\rho} - F_s\, G_{\mathbf{u}}\cdot\tfrac{\nabla s}{\rho} + \tfrac{\mathbf{B}}{\rho}\cdot F_{\mathbf{u}}\times G_{\mathbf{u}} \\ & + F_{\mathbf{u}}\cdot G_{\mathbf{E}} - G_{\mathbf{u}}\cdot F_{\mathbf{E}} + F_{\mathbf{E}}\cdot\nabla\times G_{\mathbf{B}} - G_{\mathbf{E}}\cdot\nabla\times F_{\mathbf{B}}\Big]\,,\end{aligned} \tag{8.2}$$

where the index notation is used for functional derivatives, e.g. $F_{\mathbf{u}} := \delta F/\delta\mathbf{u}$. The particle charge/mass ratio, the velocity of light and the vacuum permeability have been set to 1 for convenience: $e/m = c = \mu_0 = 1$. For the present work, the entropy $s$ plays a minor role, almost

similar to $\rho$, and even simpler. For simplicity, we will drop the entropy terms from the bracket and the term corresponding to the internal energy from the Hamiltonian. This corresponds to a pressureless (cold) plasma. Actually, the Hamiltonian will not really be needed in our investigation (except for the choice of a secondary constraint, see e.g. Eq. (8.4)). We will focus on the essential part of the Dirac reduction, which concerns the Poisson bracket.

For the consistency of this chapter, let us remind the usual procedure to get a reduced Poisson bracket using Dirac constraints, and illustrate it with the standard example of the reduction to incompressible-fluid dynamics [33]. In this chapter (and in the associated appendix chapters 12 and 13), we will come back several times to this reduction or to some of its variants. It will constitute a kind of common thread. For clarity here, we take the simplest fluid, without the electromagnetic field (nor any pressure), removing $\mathbf{E}$ and $\mathbf{B}$ from the variables. Then the bracket writes

$$\{F,G\} = \int d^3\mathbf{x}\; G_{\mathbf{u}}\cdot\nabla F_\rho - \; F_{\mathbf{u}}\cdot\nabla G_\rho + \tfrac{\nabla\times\mathbf{u}}{\rho}\cdot\;(F_{\mathbf{u}}\times G_{\mathbf{u}})\,.$$

Starting from this Poisson bracket, the primary local constraint is chosen as

$$\phi_1(\mathbf{x}) := \rho(\mathbf{x}) - \rho_0\,, \tag{8.3}$$

where $\rho_0$ is a homogeneous density. Since the matrix of constraint $\{\phi_1(\mathbf{x}),\phi_1(\mathbf{x}')\} = 0$ is not invertible, the Dirac procedure induces a (local) *secondary constraint*, given by the time evolution of $\phi_1(\mathbf{x})$:

$$\phi_2(\mathbf{x}) := \dot\phi_1(\mathbf{x}) = \{H,\phi_1(\mathbf{x})\} = \nabla\cdot(\rho_0\mathbf{u})(\mathbf{x})\,, \tag{8.4}$$

and as in most cases where the Dirac procedure works, it is enough for the matrix of constraints $C_{\alpha\beta}(\mathbf{x},\mathbf{x}') := \{\phi_\alpha(\mathbf{x}),\phi_\beta(\mathbf{x}')\}$ to be invertible. The reduced bracket is given by

$$\{F,G\}_* := \int d^3\mathbf{x}\int d^3\mathbf{x}'\;\{F,\phi_\alpha(\mathbf{x})\}\;C^{-1}_{\alpha\beta}(\mathbf{x},\mathbf{x}')\;\{\phi_\beta(\mathbf{x}'),G\} = \int d^3x\;\tfrac{\nabla\times\mathbf{u}}{\rho}\cdot(\bar F_{\mathbf{u}}\times\bar G_{\mathbf{u}})\,, \tag{8.5}$$

where $\bar F_{\mathbf{u}} = (1-\nabla\Delta^{-1}\nabla)\cdot F_{\mathbf{u}}$, and summation over repeated indices is implied.

It can be written as in Chapter 7

$$\mathbb{J}_* = \mathcal{P}_*^\dagger\mathbb{J}\mathcal{P}_*\,,$$

by defining the projector

$$\mathcal{P}_*F_{\boldsymbol{\psi}} = (F_*,\bar F_{\mathbf{u}})\,,$$

where

$$F_* = \Delta^{-1}\nabla\cdot\left(\tfrac{\nabla\times\mathbf{u}}{\rho}\times\bar F_{\mathbf{u}}\right)\,,$$

and $\boldsymbol{\psi} = (\rho,\mathbf{u})$ is a vector grouping together all the field variables.

Accordingly, the reduced bracket is just the initial one, but using constrained functional derivatives through the Dirac projector $\mathcal{P}_*$. As observed in Chapter 7, this projector involves terms that are irrelevant, because the reduced bracket is independent of the first component of the projector, which can be removed. This provides a simplified orthogonal projector

$$\mathcal{P}F_{\boldsymbol{\psi}} = (0,\bar F_{\mathbf{u}})\,.$$

This last projector is very interesting because the Dirac projector was not entirely satisfactory: it depended on the field variables and it had a complicated action on functional derivatives. The simplified projector is much more convenient; it is the most natural operator that keeps functional derivatives in the reduced space. The relation $\mathbb{J}_* = \mathcal{P}\mathbb{J}\mathcal{P}$ means that the reduced bracket is exactly the initial one with orthogonal projection of the functional derivatives onto the reduced space.

Another way to show the effect of the simplified projector is to change the field variables and to adopt coordinates adapted to the constraints, choosing for field coordinates $\big(\boldsymbol{\phi},\nabla\times(\rho_0\mathbf{u})\big)$, with

$$\boldsymbol{\phi} := (\phi_1,\phi_2) = \big(\rho-\rho_0,\nabla\cdot(\rho_0\mathbf{u})\big)$$

the vector of constraints. Then, the projector $\mathcal{P}$ simply writes

$$\mathcal{P} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) , \tag{8.6}$$

which shows that the reduced bracket is just a truncation of the initial bracket, obtained by removing derivatives along the constrained fields $\phi_1$ and $\phi_2$.

In this situation, to get the reduced bracket, all the computations involved in the Dirac procedure can be replaced by a much simpler method: just take the initial bracket and truncate it. The question is to know when this simplification of the Dirac procedure is available, i.e. when the Dirac reduction results in a bracket truncation, and how to identify the simplified projector without inverting the matrix of constraints.

## 8.2 The quarter-canonical structure in Dirac reductions

In this section, the presence of the simplified truncation projector is connected to a special structure in the Poisson bracket, which is quarter-canonical. In Subsec. 8.2.1, the condition for the truncation is identified. In Subsec. 8.2.2, the quarter canonical structure is introduced and shown to be involved in the truncation reduction. Subsecs. 8.2.3 and 8.2.4 will be devoted to an example of bracket extension and to an example of bracket reduction. Last, Subsec. 8.2.5 will be concerned with the case where the coordinates are not suited to the quarter-canonical structure.

### 8.2.1 Condition for a Dirac truncation

First of all, in order to study the mechanism at work in the presence of a simplified projector in Dirac reductions, we consider a general Hamiltonian system, with field variables $\boldsymbol{\psi}$, on which the (local) Dirac constraints $\boldsymbol{\phi}$ is imposed. The general form of Dirac's reduced bracket was given in (8.5). It can be put into matrix form

$$\mathbb{J}_* = \mathbb{J} - \mathbb{J}\mathbb{Q}^\dagger\mathbb{C}^{-1}\mathbb{Q}\mathbb{J} , \tag{8.7}$$

where $\mathbb{Q}$ and $\mathbb{C}$ are defined by

$$\begin{aligned} \tfrac{\delta\boldsymbol{\phi}(\mathbf{x})}{\delta\boldsymbol{\psi}(\mathbf{q})} &= \mathbb{Q}(\mathbf{x})\ \delta(\mathbf{x}-\mathbf{q}) = \mathbb{Q}^\dagger(\mathbf{q})\ \delta(\mathbf{q}-\mathbf{x}) , \\ \mathbb{C} &:= \mathbb{Q}\mathbb{J}\mathbb{Q}^\dagger , \end{aligned}$$

and $\mathbb{C}$ is assumed to be invertible, in order to fit with the situation where the Dirac procedure is applicable. The link with a bracket truncation is clearer when adapted coordinates are chosen: $\boldsymbol{\psi} = (\boldsymbol{\phi}, \boldsymbol{\chi})$, where $\boldsymbol{\phi}$ are the constraints and $\boldsymbol{\chi}$ are complementary coordinates. In these coordinates, we have

$$\mathbb{Q} := \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) .$$

The matrix of constraints $\mathbb{C}$ is one of the coefficients of the Poisson matrix

$$\mathbb{J} = \begin{pmatrix} \mathbb{C} & -\mathbb{B}^\dagger \\ \mathbb{B} & \mathbb{A} \end{pmatrix} ,$$

and the Dirac correction for the bracket is

$$\mathbb{J}\mathbb{Q}^\dagger\mathbb{C}^{-1}\mathbb{Q}\mathbb{J} = \begin{pmatrix} \mathbb{C} & -\mathbb{B}^\dagger \\ \mathbb{B} & -\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger \end{pmatrix} . \tag{8.8}$$

The constraints $\boldsymbol{\phi}$ have become Casimir invariants of the reduced bracket $\mathbb{J}_*$, since the correction is such as to put to zero the first three coefficents of the Poisson bracket:

$$\mathbb{J}_* = \mathbb{J} - \mathbb{J}\mathbb{Q}^\dagger\mathbb{C}^{-1}\mathbb{Q}\mathbb{J} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & \mathbb{A}+\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger \end{smallmatrix}\right) . \tag{8.9}$$

The remaining coefficient of the reduced bracket can now be compared with the one of the truncated bracket, which is

$$\mathbb{J}_t = \mathcal{P}^\dagger\mathbb{J}\mathcal{P} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & \mathbb{A} \end{smallmatrix}\right) .$$

Thus, the result of the Dirac procedure is in general different from a simple bracket truncation. They are equal if and only if $\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger = 0$.

It must be noticed that even in this case, the Dirac projector is not the simplified projector $\mathcal{P}$ of (8.6):

$$\mathcal{P}_* = 1 - \mathbb{Q}^\dagger\mathbb{C}^{-1}\mathbb{Q}\mathbb{J} = \left(\begin{smallmatrix} 0 & \mathbb{C}^{-1}\mathbb{B}^\dagger \\ 0 & 1 \end{smallmatrix}\right) , \tag{8.10}$$

which is not equal to $\mathcal{P}$ (except if $\mathbb{C}^{-1}\mathbb{B}^\dagger = 0$, i.e.if [1] $\mathbb{B} = 0$). But Eq. (8.9) shows that the Dirac projector can be replaced by the simplified projector $\mathcal{P}$ even when they are not equal.

### 8.2.2 The quarter-canonical structure

There is a special case where $\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger = 0$ is automatically zero, namely when the Poisson matrix is quarter-canonical. Let us remind that a canonical bracket writes

$$\mathbb{J}_c = \begin{pmatrix} M_c & 0 & & 0 \\ 0 & M_c & & 0 \\ & & \ddots & \\ 0 & 0 & & M_c \end{pmatrix} ,$$

with $M_c = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ ; all the coordinates are pairs of canonically conjugated variables. A semi-canonical bracket [85] is less demanding, it has only one pair of canonically conjugated coordinates:

$$\mathbb{J}_s = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & \mathbb{A} \end{pmatrix} . \tag{8.11}$$

A quarter-canonical bracket is still half less demanding

$$\mathbb{J}_q = \begin{pmatrix} 0 & -b^\dagger & 0 \\ b & a & -d^\dagger \\ 0 & d & \mathbb{A} \end{pmatrix} . \tag{8.12}$$

The first coordinate is "conjugated" to the second one in the sense that it is coupled only with it, but the second coordinate is not coupled only with the first one. This dissymmetry is an important feature, which will impact the following. So, a quarter-canonical structure is defined by the existence of coordinates $(\psi_1, \psi_2, \psi_{j\geqslant 3})$, with $\dim\psi_1 = \dim\psi_2$, such that

$$\{\psi_1, \psi_1\}=0 , \tag{8.13}$$

$$\{\psi_1, \psi_{j\geqslant 3}\}=0 . \tag{8.14}$$

It must be pointed out that all the forms of Poisson bracket introduced above are observed only when they are expressed in a suitable set of coordinates, which are qualified as "canonical", "semi-canonical", or "quarter-canonical". Changing coordinates generally breaks the form of the bracket (except if the change of coordinates is canonical). For finite-dimensional systems, the Darboux theorem [84,98] guarantees the existence of (local) coordinates such that the bracket is canonical. It also gives a constructive algorithm to isolate successively each pair of conjugated variables in a semi-canonical bracket (8.11), where $\mathbb{A}$ depends only on the remaining coordinates $\phi_i \geqslant 3$. The semi-canonical bracket is the result of the first iteration of the Darboux algorithm. The quarter-canonical bracket can be seen as the first step to get a semi-canonical bracket. Alternatively, given a first coordinate $\psi_1$, it can be considered as resulting from the semi-canonical bracket, and authorizing an arbitrary change of variables for $\psi_2$, and a partially free change of variables for $\psi_j$, $j \geqslant 3$:

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_j \end{pmatrix} \longrightarrow \begin{pmatrix} \psi_1 \\ \psi_2{}'(\psi_1,\psi_2,\psi_j) \\ \psi_j{}'(\psi_1,\psi_j) \end{pmatrix} . \tag{8.15}$$

This is easily seen by writing that through a coordinate transformation the Poisson bracket is transformed by the chain rule.

Notice that in this chapter, as well as in the companion chapters 12 and 13, parentheses are used for changes of field coordinates $\psi_2{}'(\psi_1, \psi_2, \psi_j)$ or for constraints $\phi(\boldsymbol{\psi})$, but the dependence is

[1] In the language of the next subsections, this corresponds to a semi-canonical Poisson bracket. More precisely, the coefficient $\mathbb{C}$ is not $M_c$ in general, but the bracket benefits from many properties of the semi-canonical structure.

nonlocal in general, i.e. the "function" actually involves derivatives of arbitrary order. For instance, $\phi(\boldsymbol{\psi})$ actually means

$$\boldsymbol{\phi}(\mathbf{x}) := f\Big(\boldsymbol{\psi}(\mathbf{x}), \nabla\boldsymbol{\psi}(\mathbf{x}), \nabla\nabla\boldsymbol{\psi}(\mathbf{x}), ...\Big)$$

for some function $f$. Accordingly, square parentheses $\phi[\boldsymbol{\psi}]$ should be used, in a similar way as in Chapters 4-5. However, this point will not be crucial here. For simplicity, we follow the common notation with parentheses.

When the bracket is quarter-canonical, and the constraints are $\phi_1 = \psi_1$, $\phi_2 = \psi_2$, then, the inverse of the matrix of constraints is

$$\mathbb{C}^{-1} = \begin{pmatrix} 0 & -b^\dagger \\ b & a \end{pmatrix}^{-1} = \begin{pmatrix} b^{-1}ab^{-1\dagger} & b^{-1} \\ -b^{-1\dagger} & 0 \end{pmatrix}^{-1} . \tag{8.16}$$

The coefficient $b$ has to be invertible for $\mathbb{C}$ to be so. The (complementary) Dirac projector is

$$\mathbb{Q}^\dagger\mathbb{C}^{-1}\mathbb{Q}\mathbb{J} = \begin{pmatrix} 1 & 0 & -b^{-1}d^\dagger \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \tag{8.17}$$

which is not the projector for the truncated bracket, as in (8.10). But the correction to the initial braket is

$$\mathbb{J}\mathbb{Q}^\dagger\mathbb{C}^{-1}\mathbb{Q}\mathbb{J} = \begin{pmatrix} \mathbb{C} & -\mathbb{B}^\dagger \\ \mathbb{B} & -\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger \end{pmatrix} = \begin{pmatrix} 0 & -b^\dagger & 0 \\ b & a & -d^\dagger \\ 0 & d & 0 \end{pmatrix} ,$$

which shows that $\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger = 0$, and the Dirac procedure results in a truncation of the initial bracket.

As for the reciprocal, the Dirac bracket is identical to the truncated bracket only if $\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger = 0$. If one chooses on the space of constraints $\boldsymbol{\phi}$ the coordinates $(\phi_A, \phi_B)$ with[2] $\phi_A = \ker\mathbb{B}$ and $\phi_B$ supplementary coordinates, then the matrix $\mathbb{B}$ writes

$$\mathbb{B} = \begin{pmatrix} 0 & \mathbb{B}_r \end{pmatrix} ,$$

with $\mathbb{B}_r$ injective. Then $\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger = \mathbb{B}_r\mathbb{C}^{-1}_{BB}\mathbb{B}_r^\dagger$, which is zero only if $0 = \mathbb{C}^{-1}_{BB}$. This implies $\dim\phi_B \leqslant \dim\phi_A$. This basically means that the quarter-canonical structure is involved: Eq. (8.14) holds. In addition, if $\dim\phi_B =$, the bracket is semi-canonical (more precisely it has the main property of the semi-canonical bracket, which is to split the Poisson bracket in two independent sub-algebras). On another hand, if $\dim\phi_B = \dim\phi_A$, then the Poisson bracket is exactly quarter-canonical. Between this extremal cases, the criterion $\dim\psi_1 = \dim\psi_2$ can most often be obtained by transferring a part of the $\phi_A$ to $\phi_B$ in such a way as $\mathbb{C}^{-1}_{BB}$ remains zero, for instance by using $\ker\{\phi_2, \cdot\}$ in $\phi_A$. Then, $0 = \mathbb{C}^{-1}_{BB}$ means $\{\phi_A, \phi_A\} = 0$, and the coordinates $(\phi_A, \phi_B, \boldsymbol{\chi})$ are quarter-canonical.

If one starts with a quarter-canonical Poisson bracket and changes coordinates $\boldsymbol{\psi} \longrightarrow \boldsymbol{\psi}'$, the bracket generally loses its quarter-canonical structure, and the Dirac reduced bracket is no more a truncation of the initial bracket. The same happens if one chooses other constraints than the first two coordinates. It is the reason why, in the following, when we speak about a "quarter-canonical bracket", it will mean that there exist (and we have adopted) coordinates $(\boldsymbol{\phi}, \boldsymbol{\chi})$, where $\boldsymbol{\phi} = (\phi_1, \phi_2)$ are the Dirac constraints, and $\boldsymbol{\chi}$ are complementary coordinates (that remain dynamical in the reduced bracket) such that the Poisson bracket writes (8.12).

From the point of view of the Dirac reduction, the choice of coordinates $\boldsymbol{\chi}$ is allowed because Dirac's reduced bracket is independent of the coordinates $\boldsymbol{\chi}$ used, as is clear by Eq. (8.5).

The case of a quarter-canonical structure expressed in coordinates that are not quarter-canonical is considered in the appendix chapter 12. Then the Dirac reduction is not a truncation of the initial bracket any more, but it is still obtainable by a projection of the functional derivatives, where the projector is just the equivalent of (8.6) but expressed in other coordinates: $\mathbb{J} = \mathcal{P}^\dagger\mathbb{J}\mathcal{P}$, with $\mathcal{P}$ given by formulas such as (12.13) or (12.11).

[2] This is an elliptic expression. More precisely, $\partial_{\phi_A}$ generates $\ker\mathbb{B}$.

### 8.2.3 Application to bracket extensions

The Dirac reduced bracket being always a truncated bracket for a quarter-canonical reduction is a powerful property because it can simplify many works on the link between Hamiltonian structures. We will speak much about Hamiltonian reductions, which are the most natural applications, but let us begin by a more original application, concerning bracket extensions.

Starting from a Poisson bracket $\mathbb{J}$ that has Casimir invariants $\boldsymbol{\phi} = (\phi_1, \phi_2)$, one can wonder if it comes from an extended bracket $\mathbb{J}_e$, in which $\boldsymbol{\phi}$ are dynamical, and from which the initial bracket $\mathbb{J}$ is obtained by imposing the Dirac constraints $\boldsymbol{\phi}$. A typical pattern is the Euler-Maxwell bracket (8.2) which has two (local) Casimir invariants $\boldsymbol{\phi} = (\nabla \cdot \mathbf{E} - \rho, \nabla \cdot \mathbf{B})$.

A simple answer to the problem is to write $\mathbb{J}_e$ as a quarter-canonical bracket in coordinates $(\boldsymbol{\phi}, \boldsymbol{\chi})$, with $\mathbb{A} = \mathbb{J}$, and the remaining coefficients in (8.12) are chosen freely in such a way that both $b$ is invertible and $\mathbb{J}_e$ satisfies the Jacobi identity. For instance when $\mathbb{J}$ does not depend on $\boldsymbol{\phi}$, the Jacobi identity is guaranteed provided the coefficients are all chosen constant in the field variables. Then, the matrix of constraints is invertible, and the quarter-canonical structure guarantees that imposing the Dirac constraints $\boldsymbol{\phi}$ will induce the reduced bracket to be exactly the initial bracket $\mathbb{J}$. This procedure was applied to an extension of the Vlasov-Maxwell bracket in Chapter 7, and it can be applied the same way to the Euler-Maxwell bracket.

### 8.2.4 Application to fluid and kinetic incompressibility

As a second application of the quarter-canonical structure, incompressibility can be applied to arbitrary order in fluid models, and even to the kinetic Vlasov-Maxwell model [95, 101].

The reason is that the kinetic Vlasov-Maxwell dynamics can be expressed using the fluid moments instead of the distribution function $f(\mathbf{x}, \mathbf{v})$. In the computations involved for Chapter 9 (but not reported in the present manuscript), we noticed that formula

$$\int d^3\mathbf{v} f \partial_{\mathbf{v}} (\widetilde{P}_n)_f = 0 \tag{8.18}$$

holds at arbitrary order in the fluid moments, where

$$\widetilde{P}_j := \int d^3\mathbf{v} f (\mathbf{v} - \mathbf{u})^{\otimes j}$$

is the $j$-th pressure-like moment of the distribution function $f$. This suggests to adopt the coordinates $\boldsymbol{\psi} = (\boldsymbol{\phi}, \boldsymbol{\chi})$ with

$$\boldsymbol{\phi} = (\rho, \nabla \cdot \mathbf{u}) \quad \text{and} \quad \boldsymbol{\chi} = (\nabla \times \mathbf{u}, \{\widetilde{P}_j\}_{j \geqslant 2}) \,.$$

Then the particle density $\rho$ is coupled only with $\nabla \cdot \mathbf{u}$, just as in the usual fluid bracket:

$$\{F, G\} = \int d^3\mathbf{x} \, \left[ G_{\mathbf{u}} \cdot \nabla F_\rho - F_{\mathbf{u}} \cdot \nabla G_\rho + F_{\phi_2} \, \mathbb{J}^{\phi_2 j} \, G_{\chi_j} - F_{\chi_j} \, \mathbb{J}^{j \phi_2} \, G_{\phi_2} + F_{\chi_i} \, \mathbb{J}^{ij} \, G_{\chi_j} \right] .$$

The detailed expression of $\mathbb{J}^{ij}$ is not given here, because it is quite complicated, and completely useless for our purpose. The important feature is that the bracket is quarter-canonical and the constraints $\big(\rho - \rho_0, \nabla \cdot (\rho_0 \mathbf{u})\big)$ can be imposed exactly as in the usual reduction to incompressible fluid. The reduced structure is just the truncated bracket, which means that all terms involving couplings with $\rho$ or $\nabla \cdot \mathbf{u}$ are put to zero, and all other terms are unchanged:

$$\{F, G\}_* = \int d^3\mathbf{x} \, F_{\chi_i} \, \mathbb{J}^{ij} \, G_{\chi_j} \,. \tag{8.19}$$

If the reduction is performed while keeping the coordinates $(\rho, \mathbf{u})$, instead of $(\rho, \nabla \cdot \mathbf{u}, \nabla \times \mathbf{u})$, the results of Chapter 12 can be used.

This example shows the efficiency of the quarter-canonical structure and the requirement to choose good coordinates. If we had used the usual fluid moments $P_j := \int d^3\mathbf{v} f \mathbf{v}^{\otimes j}$, the bracket would not have been quarter-canonical any more, and the result would not have been so obvious

to get.

The incompressibility to arbitrary order in the fluid moments suggests to perform the reduction directly in the kinetic model, by using the kinetic variable $f$ instead of the fluid moments. Then, the bracket is not quarter-canonical any more, and the reduction is not a truncation of the initial bracket. Besides, the constraints are not local but semi-local [34]. All the same, they are equivalent to local quarter-canonical constraints on the bracket expressed in moments of the distribution function. The Dirac reduction is independent of the chosen coordinates $\boldsymbol{\chi}$ complementary to the constraints and it is interesting to work with the coordinate $f$ rather than with the moments $\widetilde{P}_n$ because the kinetic bracket is much simpler than the corresponding fluid bracket.

The phase space is the set of $(f, \mathbf{E}, \mathbf{B})$, where $f(\mathbf{q}, \mathbf{v})$ is the distribution function, $\mathbf{E}(\mathbf{q})$ is the electric field and $\mathbf{B}(\mathbf{q})$ is the magnetic field. The Hamiltonian is the total energy

$$H = \int_{\mathbf{q},\mathbf{v}} f \tfrac{\mathbf{v}^2}{2} + \int_{\mathbf{q}} \tfrac{\mathbf{E}^2+\mathbf{B}^2}{2} \,.$$

The initial Poisson bracket is the Vlasov-Maxwell bracket [95, 101, 104, 156]

$$\begin{aligned}\{F,G\}_V = &\int d^3x d^3v \; f[\nabla F_f \cdot \partial_{\mathbf{v}} G_f - \partial_{\mathbf{v}} F_f \cdot \nabla G_f + e\mathbf{B} \cdot \partial_{\mathbf{v}} F_f \times \partial_{\mathbf{v}} G_f] \\ &+ e \int d^3x d^3v \; f\,(\partial_{\mathbf{v}} F_f \cdot G_{\mathbf{E}} - F_{\mathbf{E}} \cdot \partial_{\mathbf{v}} G_f) \\ &+ \int d^3x \; (F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} - F_{\mathbf{B}} \cdot \nabla \times G_{\mathbf{E}}) \,. \end{aligned} \tag{8.20}$$

The constraints are $\rho - \rho_0$ and $\nabla \cdot (\rho_0 \mathbf{u})$, where $\rho_0$ is a fixed homogeneous mass density. The variables

$$\rho = \int d^3v f \quad \text{and} \quad \mathbf{u} = \tfrac{1}{\rho} \int d^3v f \mathbf{v}$$

are respectively the fluid mass density and the fluid velocity. The matrix of constraints is the same as in the case of the fluid incompressibility, because the fluid Poisson bracket for $(\rho, \mathbf{u}, \mathbf{E}, \mathbf{B})$ is given by a subalgebra of the kinetic system (see Chapter 9) and thus $\{\phi_\alpha, \phi_\beta\}$ is the same when computed with the fluid and with the kinetic bracket. After some algebra, the reduced bracket is found to be

$$\{F,G\}_* = \{F,G\} + \int_{\mathbf{q}} [\int_{\mathbf{v}} PG_f] \cdot \int_{\mathbf{v}} \xi(F) - \int_{\mathbf{q}} [\int_{\mathbf{v}} PF_f] \cdot \int_{\mathbf{v}} \xi(G) - \int_{\mathbf{q}} [\int_{\mathbf{v}} PF_f] \cdot \int_{\mathbf{v}} \lambda \cdot [\int_{\mathbf{v}} PG_f] \,, \tag{8.21}$$

where

$$\begin{aligned} P &= \nabla \Delta^{-1} \nabla \cdot f \partial_{\mathbf{q}} \,, \\ \xi(F) &= \tfrac{\mathbf{u}-\mathbf{v}}{\rho} \nabla \cdot (f \partial_{\mathbf{v}} F_f) - \tfrac{f}{\rho} \nabla F_f - \tfrac{f}{\rho} e\mathbf{B} \times \partial_{\mathbf{v}} F_f + \tfrac{ef}{\rho} G_{\mathbf{E}} \,, \\ \lambda &= \tfrac{\mathbf{v}-\mathbf{u}}{\rho} \nabla \tfrac{f}{\rho} + \tfrac{f}{\rho} \nabla \tfrac{\mathbf{v}-\mathbf{u}}{\rho} + \tfrac{ef\mathbf{B}\times}{\rho^2} \,. \end{aligned}$$

In these expressions, the derivatives act on all that is on their right-hand side.

The reduced bracket may seem complicated, and the computations to obtain it are not elementary, but the existence of a quarter-canonical structure implies the reduced bracket to be simply given by the initial bracket with projected functional derivatives. The projector is the one involved in the truncation (8.19), but expressed in coordinates that are not suited to the truncation. Its expression can be obtained using Eqs. (12.10), (12.11) or (12.13) from Chapter 12. The last one is especially simple and computed in a few lines

$$\begin{aligned} \phi_{1_{\boldsymbol{\psi}}} &= (1,0,0) \,, \\ \phi_{2_{\boldsymbol{\psi}}}^{\dagger} &= (-\tfrac{\mathbf{v}-\mathbf{u}}{\rho} \cdot \rho_0 \nabla, 0, 0) \,, \\ \phi_{1_{\boldsymbol{\psi}}} \mathbb{J} \phi_{2_{\boldsymbol{\psi}}}^{\dagger} &= \nabla \cdot \rho_0 \nabla \,, \\ \phi_{2_{\boldsymbol{\psi}}}^{\dagger} (\phi_{1_{\boldsymbol{\psi}}} \mathbb{J} \phi_{2_{\boldsymbol{\psi}}}^{\dagger})^{-1} \phi_{1_{\boldsymbol{\psi}}} \mathbb{J} &= \begin{pmatrix} \frac{\mathbf{v}-\mathbf{u}}{\rho} \cdot \nabla \Delta^{-1} \nabla \cdot \int_{\mathbf{v}} f \partial_{\mathbf{v}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \,. \end{aligned}$$

Thus, the reduced bracket is simply

$$\{F,G\}_* = \int \partial_{\boldsymbol{\psi}} F \cdot \mathbb{J}_* \partial_{\boldsymbol{\psi}} G = \int \partial_{\boldsymbol{\psi}} F \cdot \mathcal{P}^\dagger \mathbb{J} \mathcal{P} \partial_{\boldsymbol{\psi}} G \,,$$

where

$$\mathcal{P} F_{\boldsymbol{\psi}} = \mathcal{P}(F_f, F_{\mathbf{E}}, F_{\mathbf{B}}) = (\bar{F}_f, F_{\mathbf{E}}, F_{\mathbf{B}}) \,,$$

with

$$\bar{F}_f = \left(1 - \tfrac{\mathbf{v}-\mathbf{u}}{\rho} \cdot \nabla \Delta^{-1} \nabla \cdot \int_{\mathbf{v}} f \partial_{\mathbf{v}}\right) F_f \,.$$

Thus Eq. (12.13) gives an efficient way to compute in a few lines the reduced bracket (8.21), whose derivation by applying Dirac's theory had been rather heavy.

### 8.2.5 Incidence, constructibility, and dispensability of quarter-canonical coordinates

The previous subsections showed the advantages of quarter-canonical Poisson brackets in Dirac reductions, especially by replacing the Dirac projector by a simplified truncation projector. However, they did not explain why this phenomenon often occurs when applying Dirac reductions in plasmas. This suggests that quarter-canonical coordinates are frequent, which calls for a closer analysis.

On another hand, the quarter-canonical structure is explicitly present only when suitable coordinates are adopted. Often, the natural coordinates are not quarter-canonical. Then, instead of applying Dirac's method, it can be convenient to adopt suited coordinates, in order to make the reduction become just a bracket truncation. This asks the question about the constructibility of such coordinates.

Last, instead of changing coordinates to make them suited to the quarter-canonical structure, one can wonder whether the simplified projector cannot be obtained directly in the initial coordinates, even when they are not quarter-canonical. This would imply an extension of the simplified-projector reduction to non-quarter-canonical cases.

This threefold question about quarter-canonical coordinates is addressed in the appendix chapter 12.

$\longrightarrow$ See the appendix chapter 12

This chapter shows that these coordinates are much more frequent than semi-canonical coordinates, they always exist for finite-dimensional systems when $\{\phi_1, \phi_1\} = 0$, and as soon as they exist, they are the only possible structure for Dirac reductions. This explains the frequent occurrence of simplified projectors in these reductions. In field theory, it is a common structure, especially in Dirac reductions. It must be thought of as soon as one half of the constraints has no coupling with itself $\mathbb{J}^{\phi_1\phi_1} = 0$. When quarter-canonical coordinates are not already present, they can most often be build by hand, at least when the Poisson bracket and the constraints contain linear dependences in the fields, which is fairly common. Many examples are shown in the ideal-fluid model. Finally, quarter-canonical coordinates are not indispensable in order to benefit from the existence of the simplified projector. This last can be explicitly identified even when the coordinates are not suited to the quarter-canonical structure, and even when such coordinates do not exist.

## 8.3 Consequences for Dirac procedure and truncation methods

The previous section focused on the use of the quarter-canonical structure in order to make the Dirac procedure become a bracket truncation. It did not investigate the meaning and the consequences of this fact for Hamiltonian reduction methods. Indeed, a bracket truncation is in principle a specific reduction method, which involves completely different structures and conditions than a

Dirac reduction. The coexistence of these two methods can impact the Dirac procedure and the bracket truncation method.

More precisely, truncations usually do not preserve the Hamiltonian character, and Dirac reductions usually require the matrix of constraints to be invertible, whereas in the quarter-canonical case, these difficulties seem to have disappeared. Clearly further investigation is needed. It will be the topic of the appendix chapter 13, where the methods will also be applied to several reductions in the ideal-fluid model.

$\longrightarrow$ See the appendix chapter 13

This chapter confirms that the quarter-canonical structure has strong impacts on reduction methods. For the Dirac procedure, it automatically implies that the matrix of constraints is invertible, which in turn means that the secondary constraints are determined by this structure. This gives a way to identify all the available secondary constraints. These *$J$-secondary constraints* are directly induced by the structure of the Poisson bracket, in contrast with the traditional *$H$-secondary constraints*, which are related to the Hamiltonian function. When taking into account both of them, the procedure to choose secondary constraints becomes more complete, more transparent, and more efficient.

As for the bracket truncation induced by the quarter-canonical structure, it is associated to the presence of a Lie-subalgebra naturally induced by the structure. This establishes a link with another Hamiltonian reduction method, based on a projection onto a subalgebra. In the case of a quarter-canonical structure, its applicability conditions are automatically satisfied, and it gives the same result as the Dirac method provided the direction of the projection is chosen accordingly. In other cases, the subalgebra reduction appears as less stiff, faster, and more often available than the Dirac method. It is not so often marked by traces of the canonical structure.

## 8.4 Application to the gyrokinetic reduction

After clarifying in the previous sections the truncation phenomena at work in Dirac reductions, we now come back to the initial question that motivated the study, which is the step 2d of the gyrokinetic reduction (see page 16). In Subsec. 8.4.1, the framework and the goal of the reduction is reminded. It consists in removing two dimensions from the base space of the Vlasov-Maxwell system. In Subsec. 8.4.2, the reduction in the case of particle dynamics is presented, because it provides the local-interaction matrix for the Vlasov bracket. In Subsec. 8.4.3, the case of plasma dynamics in an external electromagnetic field is considered. The reduction is shown to be related to a bracket truncation and to fit with Dirac reductions, but satisfactory constraints are difficult to identify. The results of this subsection are not completely conclusive yet, but they are worth some analysis because they will be crucial in order to address the reduction in the presence of the coupling. In Subsec. 8.4.4, the reduction is efficiently derived as a projection onto a subalgebra and complementary comments will be given about the reduction when the electromagnetic field is external. In Subsec. 8.4.5, the presence of the coupling is taken into account. It spoils the quarter-canonical structure, but not the application of the Dirac reduction.

### 8.4.1 Framework and goals of the reduction

The initial system is the Vlasov-Maxwell non-canonical Hamiltonian system. The observables are the set of all functionals of the magnetic field $\mathbf{B}(\mathbf{q})$, the electric field $\mathbf{E}(\mathbf{q})$, and the phase-space density $f(\mathbf{q},\mathbf{p})$, with $\mathbf{q}$ and $\mathbf{p}$ the phase-space coordinates (resp. for the particle position and momentum). The Hamiltonian functional is

$$H[\mathbf{E},\mathbf{B},f] = \tfrac{1}{2}\int d\mathbf{q}d\mathbf{p}\, f\, \|\mathbf{p}\|^2 + \tfrac{1}{2}\int d\mathbf{q}\, \left(\|\mathbf{E}\|^2 + |\mathbf{B}\|^2\right) . \tag{8.22}$$

For any observables $F$ and $G$, the Poisson bracket is [95, 101, 104, 156]:

$$\{F,G\} = \int d\mathbf{q}d\mathbf{p}\, f\, \left(\tfrac{1}{2}[F_f,G_f] + G_{\mathbf{E}}\cdot\partial_{\mathbf{p}}F_f\right) + \int d\mathbf{q}\; F_{\mathbf{E}}\cdot\nabla\times G_{\mathbf{B}} - (F\leftrightarrow G)\,, \tag{8.23}$$

where the symbol $-(F\leftrightarrow G)$ indicates that the previous terms are repeated, but with permutation of $F$ and $G$ and inversion of the signs (in order to fulfil the antisymmetry property of the Poisson bracket). The particle bracket $[\cdot,\cdot]$ is defined by

$$[g,h] = \nabla g\cdot\partial_{\mathbf{p}}h - \nabla h\cdot\partial_{\mathbf{p}}g + \mathbf{B}\cdot\partial_{\mathbf{p}}g\times\partial_{\mathbf{p}}h\,, \tag{8.24}$$

for any functions of the phase space $g$ and $h$. For the sake of simplicity physical constants have been scaled away.

The Vlasov density is defined over the 6-dimensional phase space of particle dynamics $(\mathbf{p},\mathbf{q})$, which is rather written in adapted coordinates $\mathbf{z} := (\mathbf{q},\varphi,\mu,\theta)$, where the variable $\varphi$ is the angle between the magnetic field and the particle momentum, the variable $\theta$ is an angle defining the direction of the momentum perpendicular to the magnetic field, and the variable $\mu := \frac{(\mathbf{p}\times\mathbf{B})^2}{2B^3}$ is an adiabatic invariant of particle dynamics in a strong magnetic field (see Chapter 2). Then, working on particle dynamics, the coordinates are changed through a near-identity transformation to the gyro-center coordinates

$$\mathbf{z} \longrightarrow \bar{\mathbf{z}} := (\bar{\mathbf{q}},\bar{\varphi},\bar{\mu},\bar{\theta})\,, \tag{8.25}$$

defined Chapter 2 for instance. Notice that the local gyro-angle $\theta$ is used here, as in Eq. (2.4) in order to avoid the vectorial coordinate $\mathsf{c}$. After the coordinate change (8.25), the transformed Vlasov-Maxwell Poisson bracket becomes similar to Eq. (5.12), but the additional terms coming from the dependence of $\bar{\mathbf{z}}$ on the moments of $f$ have to be taken into account as shown in Section 5.4.

The transformed Vlasov density $\bar{f}(\bar{\mathbf{z}})$ follows characteristics which do not depend on the gyro-angle coordinate and where the magnetic-moment coordinate is conserved:

$$\begin{aligned} \dot{\bar{f}} &= \mathbf{w}\cdot\partial_{\bar{\mathbf{z}}}\bar{f}\,,\\ \text{with } \mathbf{w}^{\mu} &= 0\,,\\ \text{and } \partial_{\theta}\mathbf{w} &= 0\,. \end{aligned}$$

The slow part of the Vlasov density can be studied independently: the Vlasov characteristics on the five dimensions $(\bar{\mathbf{q}},\bar{\varphi},\bar{\mu})$ constitutes a closed dynamical system with $\bar{\mu}$ conserved. Since the gyro-angle coordinate is concerned with spatio-temporal scales much smaller than most physical phenomena, it can be ignored. As a result, in the base space of the Vlasov density two dimensions are superfluous and can be removed: the gyro-angle coordinate disappears from the theory (for instance by averaging all quantities over it), and the magnetic moment becomes an index, a label, foliating the base space. It contributes only through a sum when computing the current.

The idea is to remove these two superfluous dimensions from the Hamiltonian structure of the Vlasov-Maxwell system. For the Hamiltonian functional, it is trivially obtained by averaging these quantities out. For the Poisson bracket, one has to be more careful. It roughly corresponds to removing all the derivatives $\partial_\theta$ and $\partial_\mu$, but it is not sure to be the proper method, and in addition it is not guaranteed to preserve the Hamiltonian structure, and more specifically the Jacobi identity. It is impossible to verify explicitly the Jacobi identity with such a complicated bracket (all the more as it is explicitly known only at lowest orders for most guiding-center representations). In addition, even if it were possible to verify the Jacobi identity, it would not be enlightening about the mechanism at work in this Hamiltonian reduction.

So, the goal is to achieve the removal with Hamiltonian reduction methods, which automatically preserve the Jacobi identity. It is here that the results of the previous sections play a crucial role: since the Dirac reduction generates a bracket truncation in some cases, it is natural to investigate whether the gyrokinetic reduction fits with these cases.

### 8.4.2 Introductory reduction of particle dynamics

As usual, we will begin by studying a simplified case, and then will go towards the complete case. It is an efficient way to isolate the basic phenomena at work, and to progressively identify more and more involved mechanisms. At the end, it allows for a clearer analysis of the full system, and especially it makes it easier to understand how the method can be applied or adapted to the specific case of interest, or to distinctly identify the obstructions.

Let us start by studying the corresponding reduction for particle dynamics. In coordinates $(\bar{\mathbf{q}}, \bar{\varphi}, \bar{\mu}, \bar{\theta})$, the particle Poisson bracket (8.24) writes in matrix form (see Chapter 2)

$$\bar{\mathbb{J}} = \begin{pmatrix} \bar{\mathbb{J}}^{\mathbf{qq}} & \bar{\mathbb{J}}^{\mathbf{q}\varphi} & 0 & \bar{\mathbb{J}}^{\mathbf{q}\theta} \\ \bar{\mathbb{J}}^{\varphi\mathbf{q}} & 0 & 0 & -\alpha \\ 0 & 0 & 0 & -1 \\ \bar{\mathbb{J}}^{\theta\mathbf{q}} & \alpha & 1 & 0 \end{pmatrix}, \tag{8.26}$$

where the coefficients are defined in Eq. (2.79). It is quarter-canonical for the constraints $\boldsymbol{\phi} := (\bar{\mu}, \bar{\theta})$. In the light of the previous sections, the simplified Dirac procedure can be trivially applied: the Dirac bracket is just the truncated bracket

$$\bar{\mathbb{J}}_* = \begin{pmatrix} \bar{\mathbb{J}}^{\mathbf{qq}} & \bar{\mathbb{J}}^{\mathbf{q}\varphi} & 0 & 0 \\ \bar{\mathbb{J}}^{\varphi\mathbf{q}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{8.27}$$

The reduction can also be obtained by taking the subalgebra of functions that do not depend on $\bar{\theta}$, i.e. the gyro-averaged functions. Thus the result is immediate in this case, which was actually the initial motivation suggesting that the Dirac method plays a role in the removal considered by gyrokinetics.

### 8.4.3 Dirac reduction of plasma dynamics

The next step is to go to the plasma level. In the simplest description, the electromagnetic field is static and external. The Hamiltonian is

$$\begin{aligned} H = \int d\mathbf{z}\, f \left(\frac{\mathbf{p}^2}{2} + \Phi(\mathbf{q})\right) &= \int d\bar{\mathbf{z}} \mathcal{J}\, \bar{f} \left(\frac{\mathbf{p}^2[\bar{\mathbf{z}}; \mathbf{B}, \Phi]}{2} + \Phi\big(\mathbf{q}[\bar{\mathbf{z}}; \mathbf{B}, \Phi]\big)\right) \\ &= \int d\bar{\mathbf{z}}\, \bar{\bar{f}} \left(\frac{\mathbf{p}^2[\bar{\mathbf{z}}; \mathbf{B}, \Phi]}{2} + \Phi\big(\mathbf{q}[\bar{\mathbf{z}}; \mathbf{B}, \Phi]\big)\right), \end{aligned} \tag{8.28}$$

where $\bar{\bar{f}} := \mathcal{J}\bar{f}$ is related to the Vlasov 6-form, with $\mathcal{J}$ the Jacobian of the transformation $\mathbf{z} \longrightarrow \bar{\mathbf{z}}$ (see Chapter 5). The Poisson bracket is just the Vlasov part of the Vlasov-Maxwell bracket

$$\{F, G\} = \int d\mathbf{z}\, f[F_f, G_f] = \int d\bar{\mathbf{z}} \mathcal{J}\, \bar{f}[\mathcal{J}^{-1}\bar{F}_{\bar{f}}, \mathcal{J}^{-1}\bar{G}_{\bar{f}}] = \int d\bar{\mathbf{z}}\, \bar{\bar{f}}[\bar{\bar{F}}_{\bar{\bar{f}}}, \bar{\bar{G}}_{\bar{\bar{f}}}], \tag{8.29}$$

which is the Lie-Poisson bracket of the particle bracket

$$[g, h] := \nabla g \cdot \partial_{\mathbf{p}} h - \nabla h \cdot \partial_{\mathbf{p}} g + \mathbf{B} \cdot \big(\partial_{\mathbf{p}} g \times \partial_{\mathbf{p}} h\big) = \partial_i g\, \mathbb{J}^{ij}\, \partial_j h\,,$$

where $\mathbb{J}$ writes as the matrix (8.26) when the coordinates $(\bar{\mathbf{q}}, \bar{\varphi}, \bar{\mu}, \bar{\theta})$ are adopted. The interest of viewing the Vlasov density as a 6-form is that it restores the Lie-Poisson structure, which was spoilt by the field coordinate $\bar{f}$. With this structure, the removal can be performed by applying the Dirac reduction on particle dynamics as for Eq. (8.26), and then by lifting the result through a Lie-Poisson bracket.

This gives the expected result, but the reduction actually works at the particle level, it is only artificially lifted afterwards. An essential ingredient is that the electromagnetic field is external,

and the method will not be applicable when the coupling is restored. So, the idea is to apply the Dirac reduction directly at the field level.

The quarter-canonical structure in the bracket (8.26) suggests that it should be doable, but the choice of constraints is not obvious in this case, because no field is to be removed by the reduction. Only a part of the field $f$ is constrained, and the constraints should act as dimensions of the base space.

Because of the Lie-Poisson structure, the removal of the two superfluous dimensions can expectedly be induced by the particle reduction (8.26). This analogy suggests that the constraints should satisfy

$$(\phi_1)_f = \mu\,, \qquad (\phi_2)_f = \theta\,.$$
$$\text{Hence} \qquad \phi_1 = \int d\mathbf{z} f\mu\,, \qquad \phi_2 = \int d\mathbf{z} f\theta\,, \tag{8.30}$$

where the overbars have been dropped for simplicity. Physically, these constraints seem to be well suited: if they are conserved for any Vlasov density $f$, then the special case where $f$ is a single particle suggests that the characteristics should indeed preserve the magnetic moment and the gyro-angle, which would achieve the goal to remove the two superfluous dimensions from the dynamical base space.

Then the matrix of constraints is given by

$$\mathbb{C} = \left(\begin{smallmatrix} 0 & -N \\ N & 0 \end{smallmatrix}\right)\,,$$

with $N := \int d\mathbf{z} f$ the total number of particles. The Dirac bracket is

$$\begin{aligned}\{F,G\}_* &= \tfrac{1}{2}\{F,G\} - \tfrac{1}{N}\int d\mathbf{z}\, f[F_f,(\phi_1)_f].\int d\mathbf{z}\, f[(\phi_2)_f,G_f] - (F\leftrightarrow G)\\ &= \tfrac{1}{2}\{F,G\} - \tfrac{1}{N}\int d\mathbf{z}\, f\partial_\theta F_f.\int d\mathbf{z}\, f\mathbb{J}^{\theta j}\partial_j G_f - (F\leftrightarrow G)\,,\end{aligned} \tag{8.31}$$

where we have used $\mathbb{J}^{i\mu} = \delta^{i\theta}$, with $\delta$ the Kronecker delta symbol.

It is readily checked that the constraints have become Casimir invariants. However, the bracket (8.31) does not correspond to a truncation of the initial bracket. Especially, it is not the Lie-Poisson bracket of the reduced particle bracket $\mathbb{J}_*$ in Eq. (8.27), and the reduced Vlasov equation involves additional terms with the coefficient $\frac{1}{N}\int d\mathbf{z}\, f\mathbb{J}^{\theta j}\partial_j H_f$ or $\frac{\mathbb{J}^{\theta j}}{N}\int d\mathbf{z}\, f\partial_\theta H_f$ traducing global properties, which does not correspond to extracting the slow dynamics of the Vlasov equation.

In the case of the guiding-center reduction, this last coefficient is zero, because $\partial_\theta H_f = 0$, and the former coefficient affects only the dynamics of the coordinate $\theta$, in such a way as to make the constraint $\phi_2$ in (8.30) conserved. So, the dynamics of the four slow dimensions remains given by a truncation of the Vlasov equation, but the point is that the dimensions $\theta$ actually remains dynamical for the characteristics. The Casimir $\phi_2$ implies a global conservation, but not a local one in $\theta$, in the sense that $\int d\mathbf{z}\, fg(\theta)$ is not conserved for any $g$. Hence the total number of particles with a specific value of the gyro-angle, e.g. $\theta = 0$, is not conserved. So, the dimension $\theta$ is not removed from the base space of the theory.

The reason for the non-local coefficients is that there is no common variable between the two integral factors in the second term of Eq. (8.31): the two integrals remain. To remedy this difficulty, the constraints must be local or at least semi-local. For instance, one can consider the constraints

$$\phi_1(\mu_\alpha) = \int d\mathbf{z}^4 d\theta f\mu_\alpha = \int d\mathbf{z} f\mu\delta(\mu-\mu_\alpha)\,, \qquad \phi_2(\theta_\beta) = \int d\mathbf{z}^4 d\mu f\theta_\beta = \int d\mathbf{z} f\theta\delta(\theta-\theta_\beta)\,,$$

where $d\mathbf{z}^4 = d\mathbf{q}d\varphi$ is the measure for the four slow dimensions. The indices $\alpha$ and $\beta$ indicate the (non-integrated) variables of the semi-local constraints. The matrix of constraints becomes

$$\mathbb{C} = \left(\begin{smallmatrix} 0 & -b \\ b & 0 \end{smallmatrix}\right)\,, \tag{8.32}$$
$$\text{with} \quad b := \{\phi_2(\theta_\beta),\phi_1(\mu_\alpha)\} = \int d\mathbf{z}\, f\partial_i\big(\theta.\delta(\theta-\theta_\beta)\big)\mathbb{J}^{ij}\partial_j\big(\mu.\delta(\mu-\mu_\alpha)\big) = \mu_\alpha\theta_\beta\partial_\mu\partial_\theta\int d\mathbf{z}^4 f\,.$$

Assuming that $b$ is invertible for exploratory purpose, the Dirac bracket is

$$\begin{aligned}\{F,G\}_* &= \tfrac{1}{2}\{F,G\} - \int d\mu_\alpha d\theta_\beta \left[ \int d\mathbf{z}^4 d\theta \left\{ \mu_\alpha \partial_\mu \big( f \partial_i F_f \mathbb{J}^{i\mu} \big) \right\} . \tfrac{1}{b} . \int d\mathbf{z}^4 d\mu \left\{ \theta_\beta \partial_\theta \big( f \mathbb{J}^{\theta j} \partial_j G_f \big) \right\} \right] - (F \leftrightarrow G) \\ &= \tfrac{1}{2}\{F,G\} - \int d\mu_\alpha d\theta_\beta \left[ \tfrac{1}{\int d\mathbf{z}^4 \partial_\mu \partial_\theta f} . \int d\mathbf{z}^4 d\theta \left\{ \partial_\mu \big( f \partial_\theta F_f \big) \right\} . \int d\mathbf{z}^4 d\mu \left\{ \partial_\theta \big( f \mathbb{J}^{\theta j} \partial_j G_f \big) \right\} \right] - (F \leftrightarrow G) \,,\end{aligned} \tag{8.33}$$

where in the square parentheses the non-integrated variables $\mu$ and $\theta$ are implicitly $\mu_\alpha$ and $\theta_\beta$. The bracket (8.33) is complicated and does not provide the expected result, but we indicate it because it shows that although the quarter-canonical reduction is usually straightforward, in this case some care and some investigations can be needed before identifying the precise reduction. Especially, even if the bracket (8.33) is not relevant in itself, it shows that the variables present in the definition of the constraints $\mu_\alpha$ and $\theta_\beta$ are precisely those which remove the integral in the two factors of the second term of equation (8.31), as well as in the coefficient of $\mathbb{C}$.

Accordingly, in order to avoid global (integrated) coefficients in the Vlasov equation, purely local constraints are needed. They will remove both integrals in (8.31), and will imply the presence of Dirac delta functions in the matrix of constraints, which will couple the variables in the two integrated factors of (8.31).

The issue is that then the matrix of constraints becomes not invertible. Indeed, writing the constraints as

$$\phi_1(\mathbf{z}_\alpha) = f\mu_\alpha = \int d\mathbf{z} f \mu \delta(\mathbf{z} - \mathbf{z}_\alpha) \,, \qquad \phi_2(\mathbf{z}_\beta) = f\theta_\beta = \int d\mathbf{z} f \theta \delta(\mathbf{z} - \mathbf{z}_\beta) \,, \tag{8.34}$$

then the coefficient $b$ in (8.32) is given by

$$\begin{aligned} b := \{\phi_2(\mathbf{z}_\beta), \phi_1(\mathbf{z}_\alpha)\} &= \int d\mathbf{z}\, f \partial_i \big( \theta \delta(\mathbf{z} - \mathbf{z}_\beta) \big) \mathbb{J}^{ij} \partial_j \big( \mu \delta(\mathbf{z} - \mathbf{z}_\alpha) \big) \\ &= - \int d\mathbf{z}\, \theta \delta(\mathbf{z} - \mathbf{z}_\beta) \partial_i \Big( f \mathbb{J}^{ij} \partial_j \big( \mu \delta(\mathbf{z} - \mathbf{z}_\alpha) \big) \Big) = \theta \partial_i \Big( f \mathbb{J}^{ij} \partial_j \big( \mu \delta(\mathbf{z}_\beta - \mathbf{z}_\alpha) \big) \Big) \,, \end{aligned}$$

where the operator acting on $\delta(\mathbf{z}_\beta - \mathbf{z}_\alpha)$ is a function of $\mathbf{z}_\beta$. The point is that this operator is not invertible in general. In addition, it has no relation with the operators $\partial_\theta$ and $\partial_\mu$:

$$\partial_j \big( \mu \delta(\mathbf{z}_\beta - \mathbf{z}_\alpha) \big) \neq \delta_{j\mu} \partial_\mu \big( \mu \delta(\mathbf{z}_\beta - \mathbf{z}_\alpha) \big) \,,$$

so that the presence of $\mu$ in $\partial_j \big( \mu \delta(\mathbf{z}_\beta - \mathbf{z}_\alpha) \big)$ is not efficient to extract the coefficients $j = \mu$ in the particle Poisson bracket $\mathbb{J}^{ij}$.

At this point, the goal seems difficult to reach, because local constraints are necessary, but they imply a non-invertible matrix of constraints, and they spoil the action of the quarter-canonical structure in $\mathbb{J}^{ij}$.

The discussion can be made clearer by using the simplified projector (12.13). In fact, the reduction does not fit with the previous sections, because the constraints do not concern some of the fields of the theory, but a part of the only field involved in the theory. In addition, the condition $\{\phi_1, \phi_1\} = 0$ is not sure to be verified. However, if the reduction is induced by the quarter-canonical structure of $\mathbb{J}^{ij}$, one can expect a property related to $\{\phi_1, \phi_1\} = 0$ and a projector related to (12.13).

So, the Dirac bracket should be related to the simplified projector $\mathcal{P}$, defined by

$$1 - \mathcal{P} := (\phi_2)^\dagger_{\boldsymbol{\psi}} \{\phi_1, \phi_2\}^{-1} \{\phi_1, \cdot\} \,. \tag{8.35}$$

The analysis in Chapter 12 shows that in order for $\mathcal{P}$ to provide the truncation of the quarter-canonical particle bracket, $1 - \mathcal{P}$ should be related to the trivial projector on $\partial_\theta$:

$$\int d\mathbf{z}_\alpha \int d\mathbf{z}_\beta \Big\{ \partial_k (\phi_2)_f \{\phi_1, \phi_2\}^{-1} \int d\mathbf{z}\, f \partial_i (\phi_1)_f \mathbb{J}^{ij} \partial_j \Big\} = \delta_{k\theta} \partial_\theta \,. \tag{8.36}$$

Then the solution or the obstruction for the existence of $(\phi_1, \phi_2)$ can be identified by integrating successively the operators. The condition $\partial_j = \partial_\theta$ implies $i = \mu$ in $\partial_i(\phi_1)_f$. It is the requirement $\partial_i(\phi_1)_f = \partial_\mu(\phi_1)_f$ identified above, which suggests that $(\phi_1)_f$ is only a function of $\mu$. On another hand, the integral in $\int d\mathbf{z}$ has to be inverted, which implies $\phi_1$ to be local, i.e. to be a function of the 6-dimensional variable $\mathbf{z}_\alpha$, which in turn prevents $(\phi_1)_f$ to be just a function of $\mu$. To summarize, the constraint $\phi_1$ must be local, and in such a way that $\partial_i(\phi_1)_f$ generates only $\delta_{i\mu}\partial_\mu$. It seems this condition is impossible to fulfil, because it would imply to find two functions $g(\mu)$ and $h(\mathbf{z})$, with $h$ explicitly dependent on $(\mathbf{q}, \varphi, \theta)$, such that $\partial_i(gh) = h\partial_i g$. Thus, the requirement on global invertibility must be released. The solution should use a local invertibility near the surface of constraints.

With this feature, a solution can be inspired from the obstruction observed with the local constraints (8.34), which spoiled the condition $\partial_i(\phi_1)_f = \delta_{i\mu}\partial_\mu(\phi_1)_f$, but only when the derivative acted on the Dirac delta symbol. Then the prefactor $\mu$ can be used to make the corresponding term zero. Thus, instead of inverting the matrix of constraints globally, one inverts it only locally, at the vicinity of the surface of constraints. The constraints are then taken

$$\phi_1(\mathbf{z}_\alpha) = f.(\mu_\alpha - \mu_0) = \int d\mathbf{z} f \varepsilon_\mu \delta(\mathbf{z} - \mathbf{z}_\alpha)\,, \qquad \phi_2(\mathbf{z}_\beta) = f.(\theta_\beta - \theta_0) = \int d\mathbf{z} f \varepsilon_\theta \delta(\mathbf{z} - \mathbf{z}_\beta)\,, \tag{8.37}$$

where $\mu_0$ and $\theta_0$ are the value of the magnetic moment and of the gyro-angle on the surface of constraints. The interpretation of the constraints is that on the surface $\mu = \mu_0$ and $\theta = \theta_0$ the dynamics of $f$ is free, whereas outside the surface it is zero. The symbol $\varepsilon_\mu$ is used for $(\mu - \mu_0)$ to indicate that it is small when $\mu$ is near $\mu_0$, and the same when replacing $\mu$ by $\theta$.

Then the basic term $\partial_i(\phi_1)_f$ for the Poisson bracket $\{\phi_1, \cdot\}$ is given by

$$\partial_i(\phi_1)_f = \partial_i\big((\mu - \mu_0)\delta(\mathbf{z} - \mathbf{z}_\alpha)\big) = \delta_{i\mu}\delta(\mathbf{z} - \mathbf{z}_\alpha) + \varepsilon_\mu \partial_i \delta(\mathbf{z} - \mathbf{z}_\alpha)\,. \tag{8.38}$$

The first term in (8.38) exactly fulfils all the requirements evidenced by the projector (8.36) and which seemed difficult to combine: it is local (it depends on the 6-dimensional phase-space coordinates $\mathbf{z}_\alpha$), but the derivative $\partial_i$ acts on a function depending only on the variable $\mu$. The trick to combine those two antagonistic properties is the presence of the small additional term $\varepsilon_\mu \partial_i \delta(\mathbf{z} - \mathbf{z}_\alpha)$. The counterpart is that the projector will be only local, near the surface of constraints.

As a result, at lowest order in $\varepsilon_\mu$ and $\varepsilon_\theta$, in the projector (8.36), the integral on $\mathbf{z}$ is inverted and the operator $\partial_\theta$ is obtained as expected

$$\int d\mathbf{z}\, f \partial_i(\phi_1)_f \mathbb{J}^{ij}\partial_j = f\mathbb{J}^{\mu\theta}\partial_\theta + \varepsilon_1\,,$$

where the relation $\mathbb{J}^{\mu j} = \mathbb{J}^{\mu\theta}\delta_{\theta j}$ was taken into account, and symbols $\varepsilon_n$ will be used to indicate quantities of order $O(\varepsilon_\mu)$ or $O(\varepsilon_\theta)$. The same way, the term $\partial_k(\phi_2)_f^\dagger$ is a small perturbation of the expected projection $\delta_{k\theta}$:

$$\partial_k(\phi_2)_f = \partial_k\big((\theta - \theta_0)\delta(\mathbf{z} - \mathbf{z}_\beta)\big) = \delta_{k\theta}\delta(\mathbf{z} - \mathbf{z}_\beta) + \varepsilon_\theta \partial_k \delta(\mathbf{z} - \mathbf{z}_\beta)\,.$$

Last, the operator $\{\phi_1, \phi_2\}$ to be inverted is obtained as:

$$\begin{aligned}\int d\mathbf{z}\, f \partial_i(\phi_1)_f \mathbb{J}^{ij}\partial_j(\phi_2)_f &= \int d\mathbf{z}\, f\Big(\delta_{i\mu}\delta(\mathbf{z} - \mathbf{z}_\alpha) + \varepsilon_\mu\partial_i\delta(\mathbf{z} - \mathbf{z}_\alpha)\Big)\mathbb{J}^{ij}\Big(\delta_{j\theta}\delta(\mathbf{z} - \mathbf{z}_\beta) + \varepsilon_\theta\partial_j\delta(\mathbf{z} - \mathbf{z}_\beta)\Big)\\ &= f\mathbb{J}^{\mu\theta}\delta(\mathbf{z}_\beta - \mathbf{z}_\alpha) + \varepsilon_2\,.\end{aligned} \tag{8.39}$$

Combining all these results, the projector (8.36) writes

$$\int d\mathbf{z}_\alpha \int d\mathbf{z}_\beta \left\{\Big(\delta_{k\theta}\delta(\mathbf{z} - \mathbf{z}_\beta) + \varepsilon_4\Big).\left(\frac{\delta(\mathbf{z}_\beta - \mathbf{z}_\alpha)}{f\mathbb{J}^{\mu\theta}} + \varepsilon_3\right).\Big(f\mathbb{J}^{\mu\theta}\partial_\theta + \varepsilon_1\Big)\right\} = \delta_{k\theta}\partial_\theta + \varepsilon_5\,.$$

On the surface of constraints, all the terms with $\varepsilon_n$ become zero, and the expected truncation projector is obtained. This means that the reduced bracket $\{F, G\}_*$ should be the same as the

initial bracket $\{F, G\}$, with removal in $\mathbb{J}$ of the row and the column corresponding to the coordinate $\theta$. It corresponds to $\mathbb{J}_*$ given in Eq. (8.27). So, the Dirac bracket should be its Lie-Poisson bracket, where the dimensions $\mu$ and $\theta$ have indeed been made non-dynamical

$$\{F, G\}_* = \int d\mathbf{z} f [F_f, G_f]_* \,,$$

with $[\cdot, \cdot]_*$ the particle bracket associated to $\mathbb{J}_*$.

Since the framework does not fit exactly with the previous sections, the Dirac bracket has to be computed in order to verify that it results in a bracket truncation. Unfortunately, it turns out that integrations by parts spoil the perturbative expansions used above, and the resulting bracket is not just the truncated bracket. Even at lowest order the matrix of constraints does not fit with the quarter-canonical picture:

$$\mathbb{C} = -\begin{pmatrix} \varepsilon_\mu \partial_i f \mathbb{J}^{ij} \partial_j \varepsilon_\mu & \varepsilon_\mu \partial_i f \mathbb{J}^{ij} \partial_j \varepsilon_\theta \\ \varepsilon_\theta \partial_i f \mathbb{J}^{ij} \partial_j \varepsilon_\mu & \varepsilon_\theta \partial_i f \mathbb{J}^{ij} \partial_j \varepsilon_\theta \end{pmatrix} \delta(\mathbf{z}_\beta - \mathbf{z}_\alpha) \,,$$

where the derivatives act on all that is on their right. So, the constraints (8.37) are not satisfactory. Additional investigations are needed to identify the constraints to be chosen. Although a quarter-canonical structure is clearly involved, the final answer is not so straightforward to find out.

### 8.4.4 Complementary comments

In the previous section, it appeared that truncation reductions could often be efficiently formulated as a projection onto a subalgebra. In the case of a quarter-canonical structure, the subalgebra is obtained by removing the coordinate for the second constraint $\phi_2$, here related to the gyro-angle. Here it should correspond to the set of functionals $F[\omega]$ of the gyro-averaged Vlasov density $\omega := \int d\theta f$. Between two such functionals $F$ and $G$, the Poisson bracket writes

$$\{F, G\}_s = \int d\mathbf{z} \, (\omega + \tilde{f}) \; \partial_i F_f \mathbb{J}^{ij} \partial_j G_f = \int d\mathbf{z} \, \omega \; \partial_i F_f \mathbb{J}^{ij} \partial_j G_f \,, \tag{8.40}$$

where $\tilde{f} := f - \omega = \mathtt{osc}(f) = \phi_2$. The last equality comes because $\mathbb{J}^{ij}$ is a gyro-average, and in this case the functional derivatives $F_f$ and $G_f$ are also gyro-averages, so that the term with $\tilde{f}$ is a pure gyro-fluctuation, and its integral over $\theta$ is zero.

As a result, the set of $F[\omega]$ is a subalgebra. In the reduced Poisson bracket, the derivatives with respect to $\theta$ zero. Thus in $\mathbb{J}_s^{ij}$, all the terms with $i = \theta$ or $j = \theta$ do not contribute. This removes the row and the column of the matrix (8.26), which provides the truncated bracket, i.e. $\mathbb{J}_s = \mathbb{J}_*$:

$$\{F, G\}_s = \int d\mathbf{z} \, \omega \; (1 - \delta_{i\theta}) \partial_i F_f \mathbb{J}^{ij} (1 - \delta_{j\theta}) \partial_j G_f = \int d\mathbf{z} \, \omega \; \partial_i F_f \mathbb{J}_*^{ij} \partial_j G_f = \{F, G\}_* \,. \tag{8.41}$$

This subalgebra reduction is again much more straightforward than the corresponding Dirac reduction. Here it just consists in considering that the Vlasov density is gyro-averaged. As usual, it explicitly removes only the constraint $\phi_2$, but the quarter-canonical structure implies that $\phi_1$ is removed at the same time, because it was coupled only with $\phi_2$.

To conclude about the case of external electromagnetic field, in the gyrokinetic reduction the removal of the two superfluous dimensions (step 2d on page 16) fits with an application of Dirac constraints viewed as a bracket truncation, induced by the presence of a quarter-canonical structure. This structure originates from the particle Poisson bracket and transfers to the Vlasov bracket through a Lie-Poisson structure. Thus the removal obtained for particle dynamics seems to be available at the level of the fields, as is somehow confirmed by the presence of a subalgebra. The results of the previous sections are useful to provide markers but not precise indications, because the framework is quite different.

Indeed, gyrokinetics is a subtle application, because one single field is constrained, and the constraints are not obvious to identify. They can involve integral operators and mix up local and

semi-local constraints. In addition, a non-global invertibility seems to be needed in order to obtain the Dirac bracket.

The work is still in progress, because several aspects of the derivation require further investigation. First of all, the reduction has been completely obtained by a lift of particle reduction or by a projection onto a subalgebra, but suitable Dirac constraints have not been identified yet. In addition, the formal analysis above suggested that constraints making the invertibility global do not exist, but a more detailed study would be useful to definitely conclude about this point.

Perhaps it is possible to make the quarter-canonical structure explicit, because in the derivation above the presence of a single field hides it somehow. More precisely, it would be convenient to identify a set of field coordinates $(\phi_1, \phi_2, \boldsymbol{\chi})$ such that the matrix of local interactions is quarter-canonical. Here, since the reduced Vlasov density is to be independent of the removed dimensions $(\mu, \theta)$, adapted coordinates should perhaps be considered by distinguishing between averaged and fluctuating quantities, for instance by

$$\phi_1 := \int d\mu \left(1 - \int d\theta\right) f\,, \qquad \phi_2 := \int d\theta \left(1 - \int d\mu\right) f\,, \qquad \boldsymbol{\chi}' := \int d\theta d\mu f\,.$$

However, the Poisson bracket is not quarter-canonical in such coordinates, and the subalgebra is not related to them. So, much remains to be clarified in this reduction.

### 8.4.5 The reduction in the presence of the coupling

The next step is to study a plasma where the Maxwell fields are not external but self-consistently generated. Let us consider first the Vlasov-Poisson system. Its Hamiltonian structure can be obtained in two different ways. The first one is to keep the Poisson bracket (8.29) and add the coupling in the Hamiltonian

$$H = \int d\mathbf{z}\, f \left(\frac{\mathbf{p}^2}{2} - \frac{1}{2}\Delta^{-1} \int d\mathbf{p}\, f\right).$$

Then the same methods as above can be used since they concerned only the Poisson bracket. However, it assumes that the electric field in the transformation is external, it is not the same as the electric field generated by the Vlasov density $\mathbf{E} := \nabla\Delta^{-1} \int d\mathbf{p}\, f$. Otherwise the transformation depends on $f$, and this dependence is acted upon by functional derivatives $(\phi_1)_f$, which leads to the same approach as when considering the electric field as a dynamical variable, described in the following lines.

The second Hamiltonian structure for Vlasov-Poisson is closer to the Vlasov-Maxwell system. Its coupling is in the Poisson bracket. It can be obtained by considering that in the previous structure, two independent field coordinates are present $(f, \nabla\cdot\mathbf{E} - f)$, even if $\nabla\cdot\mathbf{E} - f$ is not dynamical. It is a Casimir invariant because the derivatives $\frac{\delta}{\delta f}$ are at constant $\nabla\cdot\mathbf{E} - f$. Then changing coordinates according to

$$(f, \nabla\cdot\mathbf{E} - f) \longrightarrow (f, \mathbf{E})$$

gives the Poisson bracket

$$\begin{aligned}\{F, G\} &= \int d\mathbf{z}\, f[F_f - \Delta^{-1}\nabla\cdot F_{\mathbf{E}}, G_f - \Delta^{-1}\nabla\cdot G_{\mathbf{E}}] \\ &= \int d\mathbf{z}\, f\Big([F_f, G_f] + \partial_{\mathbf{p}} F_f \cdot \nabla\Delta^{-1}\nabla\cdot G_{\mathbf{E}} - \partial_{\mathbf{p}} G_f \cdot \nabla\Delta^{-1}\nabla\cdot F_{\mathbf{E}}\Big)\,,\end{aligned} \tag{8.42}$$

where now the derivatives $\frac{\delta}{\delta f}$ are at constant $\mathbf{E}$. This bracket agrees with Eq. (7.25).

The bracket (8.42) is not the Lie-Poisson bracket of particle dynamics, so that the particle reduction cannot be lifted as it was the case for the Poisson bracket (8.29). In addition, for

$\omega := \oint d\theta f$, the set of $F[\omega, \mathbf{E}]$ is not a subalgebra any more, because between two such functionals, the Poisson bracket is

$$\begin{aligned}\{F, G\} &= \int d\mathbf{z}\,(\omega + \tilde{f})\left(\tfrac{1}{2}[F_f, G_f] + \partial_{\mathbf{p}} F_f \cdot \nabla\Delta^{-1}\nabla \cdot G_{\mathbf{E}}\right) - (F \leftrightarrow G) \\ &= \int d\mathbf{z}\,\omega\ \tfrac{1}{2}[F_f, G_f] + \int d\mathbf{z}\,\omega\ \partial_{\mathbf{p}} F_f \cdot \nabla\Delta^{-1}\nabla \cdot G_{\mathbf{E}} \\ &\qquad + \int d\mathbf{z}\,\mathtt{avg}\Big(\tilde{f}\partial_{\mathbf{p}}\Big) F_f \cdot \nabla\Delta^{-1}\nabla \cdot G_{\mathbf{E}} - (F \leftrightarrow G)\,,\end{aligned}$$

which is not a functional of $(\omega, \mathbf{E})$, because $\tilde{f}$ contributes in the last term. Notice that $\partial_{\mathbf{p}}$ is not gyro-averaged: it is actually $\partial_{\mathbf{p}}\bar{\mathbf{z}} \cdot \partial_{\bar{\mathbf{z}}}$ since the coordinates used are $\bar{\mathbf{z}}$, because Eqs. (8.40)-(8.41) show that the subalgebra is obtained with the coordinates $(\bar{\mathbf{q}}, \bar{\varphi}, \bar{\mu}, \bar{\theta})$. Incidentally, this fact also implies in the argument of the operator $\mathtt{avg}$ the additional presence of the non-gyro-averaged factor $\oint d\mathbf{x}\delta(\mathbf{q}(\bar{\mathbf{z}}) - \mathbf{x})$, because the electric field in $G_{\mathbf{E}}$ is evaluated at $\mathbf{x}$, in a similar way as in Chapter 5. Here we do not insist on such subtleties in order to avoid diverting from the main argument.

As a corollary, the absence of a subalgebra suggests that the structure is not quarter-canonical any more, and the truncation projector cannot be used either. Even the simplified projector (8.35) will not be available any more, since the constraints will be coupled with themselves, because of the electric field dependence in $\bar{\mu}$.

As for the Dirac reduction, it can be applied all the same, because the structure of the Poisson bracket is only a small perturbation of the one of the previous subsections. Thus the matrix of constraints is close to the invertible matrix obtained in Subsec. 8.4.3 (assuming that suitable constraints can be found in that case). Expanding it in series, it can be inverted to get the Dirac reduced bracket.

It is why in the previous subsection, we insisted on the Dirac reduction even if at that point it was less efficient than the Lie-Poisson and subalgebra methods. In addition, the example (8.34) was useful to introduce a perturbative approach, and the example (8.30) showed that when the Dirac reduction can be applied, the resulting reduced bracket may not be exaclty the expected one, although it is often close to it.

Notice that instead of working in the initial coordinates $(\mathbf{p}, \mathbf{q})$, and considering $\bar{\mu}$ and $\bar{\theta}$ as a function of them (and a non-local function of $\mathbf{E}$ and $\mathbf{B}$), one can change to the gyro-center coordinates $\bar{\mathbf{z}}$. Then the results of Chapter 5 apply. The departure from the Lie-Poisson structure and from the subalgebra is still clearer in the Poisson bracket given by Eqs. (5.12) together with (5.15) or (5.28). For the Dirac reduction, the perturbative series is provided by terms such as the one with $\bar{f}_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\mathbf{E}}$ in the bracket (5.12).

Last, one can turn to the full Vlasov-Maxwell bracket (8.23). It has no essential differences compared to the Vlasov-Poisson system. The same reduction can be applied. The only difference is that the dependence on $\mathbf{B}$ in the guiding-center transformation will generate additional terms in the results, such as the one with $\bar{f}_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\mathbf{B}}$ in the bracket (5.12). A similar conclusion also holds when taking into account the presence of the moments of the Vlasov density in the gyro-center transformation. The computation and a more detailed analysis of the Dirac reduction in the case of the Vlasov-Poisson and Vlasov-Maxwell systems will be done in the future.

## Conclusion

To conclude this chapter, together with its associated appendix chapters 12 and 13, the quarter-canonical structure appears as fairly common in Poisson brackets, even for field theory. For Dirac reductions in finite-dimensional systems, it is the only available structure when $\{\phi_1, \phi_1\} = 0$, e.g. when starting with a scalar constraint.

It is related to the canonical and semi-canonical structures, but the requirements are a lot less strong, which often allows to identify explicitly quarter-canonical coordinates when they are not already present. Those variables are conjugated in a weakened sense, because the first coordinate

is conjugated to the second one, but the converse is not verified.

This structure is strongly involved in the invertibility of the matrix of constraints, which is the condition for the Dirac procedure to work. When quarter-canonical coordinates exist, then the matrix of constraint is invertible if and only if the constraints are such coordinates. It is why the relevant secondary constraints are determined by quarter-canonical coordinates. These $J$-secondary constraints make much clearer the role of secondary constraints, and they constitute an interesting complement to the usual procedure relying on $H$-secondary constraints.

In most of Dirac reductions for fluids and plasmas, the quarter-canonical structure plays an underlying role. It is a generic case where the Dirac projector can be replaced by a simplified projector, corresponding to a bracket truncation, which simplifies the Dirac method. Even when quarter-canonical coordinates are not adopted or not identified, and even when they do not exist, the benefits of the structure associated to the condition $\{\phi_1, \phi_1\} = 0$ can be used in Dirac reductions, with especially the existence of an explicit simplified projector, although it cannot be completely interpreted as a truncation projector in this case.

The justification of the truncation method automatically associated to quarter-canonical brackets relies on the presence of a subalgebra, underlying this structure. It is a consequence of the Jacobi identity, and a property shared with semi-canonical brackets, but in a weakened form, since the subalgebra involves only one of the two constraints. For quarter-canonical structures, the existence conditions both for a Dirac reduction and for a subalgebra reduction are automatically satisfied, and the two methods are equivalent in their result, because the matrix of constraints is always invertible, the reduced bracket is always independent of $\phi_2$, and $\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger$ in (8.8) is always zero.

Even for quarter-canonical brackets, the flexibility of subalgebra reductions often offer complementary possibilities, for instance to impose one single constraint, or to get several possible reduced structures. This is because subalgebra reductions are less marked by traces of the canonical structure than Dirac reductions, even if they are still influenced by them.

Practically, the method of Dirac's constraints and the reduction by a subalgebra, although very different, are often related and the quarter-canonical structure is a key ingredient in this relation. This implies a softening of the Dirac reduction, but also, and above all, a link with other reduction methods. Many applications were observed, mainly taken from the ideal-fluid model. In particular, subalgebra reductions appeared to be very soft and pliable. A natural extension of the work is to investigate other examples of Hamiltonian reductions in plasmas in the light of these results. It is the topic of chapters 9 and 14.

An important application was identified in the removal of the two superfluous dimensions (gyro-angle and magnetic moment) in the Hamiltonian approach of the gyrokinetic reduction. This removal was shown to be closely related to a quarter-canonical framework. Its essential mechanism can be viewed as a bracket truncation induced by a projector, whose constraints are related to the magnetic moment and the gyro-angle. It is justified by a subalgebra argument, but only in a restricted case where the electromagnetic field is external. The complete implementation to the gyrokinetic reduction will require a more detailed study, especially in order to better understand the impact of the field dependence of the particle transformation but also to confirm that Dirac constraints can be found for the reduction with external electromagnetic field.

# Chapter 9

# Hamiltonian structure of reduced fluid models for plasmas obtained from a kinetic description

in collaboration with Cristel Chandre
Paper published in Phys. Lett. A **376**, 3172 (2012)

*A*bstract: We consider the Hamiltonian structure of reduced fluid models obtained from the kinetic description of collisionless plasmas by Vlasov-Maxwell equations. This is done by investigating the possibility of finding Poisson subalgebras associated with fluid models starting from the Vlasov-Maxwell Poisson algebra. In this way, we show that the only possible Poisson subalgebra involves the moments of zeroth and first order of the Vlasov distribution, meaning the fluid density and the fluid velocity, which corresponds to the standard ideal-fluid dynamics.
As with concerns possible closures for a Hamiltonian fluid model including moments of order two and higher, obstructions are identified, and the bracket derived in [Phys. Rev. Lett. **93**, 175002 (2004)], which involves moments of order two, is shown not to be a Poisson bracket since it does not satisfy the Jacobi identity.

## Introduction

As an application of the Hamiltonian reduction methods considered in previous chapters, this chapter is concerned with the Hamiltonian reduction from the Vlasov-Maxwell model to fluid models, and especially with including in the ideal-fluid model some moments of the Vlasov density of order 2 and higher.

This choice of application is justified because improvements are expected about this reduction. Indeed, for Hamiltonian systems, the fluid closure is usually implemented at the first moment of the distribution function, with the closure assumption relying on a local thermodynamical equilibrium. This corresponds to the standard fluid model, with the plasma field variables being the plasma density, velocity (or momentum) and entropy.

When the system is not collisional enough, the closure assumption may not hold, and higher-order moments can have important contributions. There does not seem to be a special reason why only a first-order model would be Hamiltonian, and one can look for a Hamiltonian model at each order in the moment expansion. Especially, as the temperature increases, the collisionality diminishes, and including moments of order two or higher may be needed for a physically relevant description of the plasma.

This is corroborated by a recently proposed model including moments of order two [136]. So, the idea here is to examine the mechanism at work in the derivation of this model compared to the standard fluid model, and to try to generalize the result by including higher-order moments.

The chapter is organized as follows. After reminding the principles of the fluid reduction in Sec. 9.1, we consider in Sec. 9.2 the reduction to the standard fluid model, which will appear as a truncation of the initial kinetic Poisson bracket. Bracket truncations usually do not produce a proper Poisson bracket, but the presence of a subalgebra remedies this issue, which offers a very practical and useful method for deriving reduced fluid models by using subalgebras. Applying it to the Vlasov-Maxwell Hamiltonian system provides the dynamics of the usual ideal-fluid model composed by the continuity equation for the fluid density and the momentum equation for the fluid velocity.

Then, in Sec. 9.3, we consider applying the method for Hamiltonian fluid closures including moments of order two and higher. However it will appear that when the closure occurs at those orders, the reduced structure is no longer a Poisson subalgebra associated with the parent structure. As a consequence it is not possible to obtain reduced fluid models containing higher order moments by just applying a Poisson subalgebra argument.

Last, in Sec. 9.4, we will turn to pressure-like variables, instead of moments of the Vlasov density, in order to better agree with the physics. This also aims at better understanding the reduction method involved in the results of Refs. [136, 138]. They introduce a fluid model for low-temperature relativistic plasmas, called the warm fluid model, which involves second-order moments of the Vlasov distribution. This model was suggested by the interaction between a strong laser pulse and a low-density plasma. In Ref. [136] a conserved quantity has been constructed based on the Vlasov-Maxwell Hamiltonian, and a bracket was proposed. It has been shown that this model conserves energy and entropy. However, based on the results explained in Sec. 9.3, the Hamiltonian property of this model has to be scrutinized since it involves higher-order moments of the Vlasov distribution. In Sec. 9.4 we exhibit a counterexample for the Jacobi identity for the bracket proposed in Ref. [136]. Therefore, this warm fluid model is not a Hamiltonian system.

## 9.1 Principles of the reduction

Let us begin by reminding the principle of fluid reductions, and the interest of Hamiltonian fluid reductions.

In the kinetic framework, the dynamics of collisionless plasmas is provided by the Vlasov-Maxwell equations which are dynamical equations for the phase-space density $f(\mathbf{x}, \mathbf{v})$ of the charged particles (also called Vlasov density) and the electromagnetic fields $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ where $\mathbf{x}$ and $\mathbf{v}$ belong to $\mathbb{R}^3$. Here, we work with only one species of charged particles of unit mass and charge $e$ for the sake of simplicity, but the generalization to several species is straightforward (see, e.g., Ref. [143]). The dynamical equations are

$$\dot{f} = -\mathbf{v} \cdot \nabla f - e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \partial_{\mathbf{v}} f, \tag{9.1}$$

$$\dot{\mathbf{E}} = \nabla \times \mathbf{B} - e \int d^3v f \mathbf{v}, \tag{9.2}$$

$$\dot{\mathbf{B}} = -\nabla \times \mathbf{E}, \tag{9.3}$$

where the dot designates the time derivative. Given nowadays computer capability, integrating numerically this kinetic model is too demanding for realistic laboratory plasmas. Moreover the dynamics is not easily analysed in such a kinetic framework since it provides dynamical information at temporal and spatial scales which might be irrelevant. In addition, the plasma dynamics is more conveniently analysed in configuration space rather than in particle phase space. As a consequence, there is a need for reduction in order to eliminate the irrelevant parts of these equations so as to obtain a much more tractable model. One way of doing this is to consider a fluid reduction which is obtained by considering the moments of the Vlasov distribution, e.g., the fluid density, the fluid momentum, the pressure tensor, etc...

Starting from the dynamics of the zeroth order moment, a series of dynamical equations for the higher order moments is built. This is a priori an infinite set of differential equations. The reduction is obtained by truncating this set and closing the system of equations through a closure assumption, often based on a collisional argument or on a local thermodynamical equilibrium. According to

the physical situation of interest, the choice of an order of truncation and of a closure assumption brings us a corresponding fluid model. For instance, starting from the parent model, i.e., the Vlasov-Maxwell equations, many physically interesting fluid models have been derived (see, e.g., Refs. [7, 55, 107, 120, 138]). A particularly important property of the parent model is that it possesses a Hamiltonian structure, i.e., the equations of motion can be rewritten using a Hamiltonian functional $H$ and a Poisson bracket $\{\cdot,\cdot\}$ as $\dot{F} = \{F, H\}$. Truncating and closing the set of equations might not conserve this property. This introduces dissipative or non-Hamiltonian terms in the equations of motion whether they have a physical origin (as described by phenomenological constants like diffusion constant, magnetic diffusivity or kinematic viscosity) or not. If these terms do not have a physical origin, they have been coined mutilations (Ref. [108]). If one considers only the ideal part of the equations which is the part where the phenomenological constants characterizing the dissipative terms are set to zero, one should recover a Hamiltonian system as an inheritance of the Hamiltonian parent model [3, 32, 33, 107–109].

There are two ways to proceed. The first one is to work with the equations of motion and check a posteriori that the resulting set of reduced equations possesses a Hamiltonian structure (by finding an appropriate conserved quantity and a Poisson bracket). This can be tedious as one needs to check that the bracket which has been guessed satisfies all the properties of a Poisson bracket, and in particular the Jacobi identity. The second method is to work directly on the Hamiltonian structure of the parent model by performing the reduction on the Hamiltonian and on the bracket; it is our goal here. Of course the main difficulty resides in verifying that the reduced bracket still satisfies the Jacobi identity. This puts some restrictions on what is allowed to do on a Poisson bracket in the course of the reduction. The gain is significant since the preservation of the Hamiltonian structure is ensured, i.e., there is no fake dissipation or mutilation, and it allows one to keep track of the conserved quantities throughout the derivation and apply all the techniques already available for Hamiltonian systems (like perturbation theory, energy-Casimir methods for equilibria, Lie transforms, etc...).

## 9.2 Reduction to the ideal-fluid model

The starting point is the Hamiltonian structure of the Vlasov-Maxwell equations [95, 101, 104, 156]. The Hamiltonian functional is the total energy :

$$H(f, \mathbf{E}, \mathbf{B}) = \int d^3v d^3x \; f \; \frac{\mathbf{v}^2}{2} + \int d^3x \; \frac{\mathbf{E}^2 + \mathbf{B}^2}{2}.$$

The Poisson bracket acts on the Poisson algebra of observables, that is the set of functionals of the field variables $f(\mathbf{x}, \mathbf{v})$, $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ :

$$\begin{aligned}\{F, G\}_V = &\int d^3x d^3v \; f[F_f, G_f] + e \int d^3x d^3v \; f \left(\partial_{\mathbf{v}} F_f \cdot G_{\mathbf{E}} - F_{\mathbf{E}} \cdot \partial_{\mathbf{v}} G_f\right) \\ &+ \int d^3x \; \left(F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} - F_{\mathbf{B}} \cdot \nabla \times G_{\mathbf{E}}\right), \end{aligned} \tag{9.4}$$

where $F_\psi$ indicates the functional derivative with respect to the field variable $\psi$ and the bracket $[\cdot,\cdot]$ is given by

$$[f, g] = \nabla f \cdot \partial_{\mathbf{v}} g - \partial_{\mathbf{v}} f \cdot \nabla g + e\mathbf{B} \cdot \partial_{\mathbf{v}} f \times \partial_{\mathbf{v}} g,$$

and $\partial_{\mathbf{v}}$ designates the partial derivatives with respect to $\mathbf{v}$. The dynamics of a functional $F$ of the Poisson algebra is given by $\dot{F} = \{F, H\}$. In particular we recover Eqs. (9.1)-(9.3) if we choose $F = f$, $F = \mathbf{E}$ or $F = \mathbf{B}$. We notice that we have considered here a non-relativistic plasma. However the discussion which follows is unchanged in the case of a relativistic plasma since the changes only occur in the Hamiltonian, not in the Poisson bracket [5].

Fluid models rely on the idea of replacing the Vlasov density $f$ by the series of its moments [159]

$$P_n^{i_1 \cdots i_n} = \int d^3v f v_{i_1} \cdots v_{i_n}, \tag{9.5}$$

for $n \in \mathbb{N}$. The chain rule yields

$$F_f = \sum_{n=0}^{\infty} F_{P_n{}^{i_1 i_2 \cdots i_n}}\ v_{i_1} v_{i_2} \cdots v_{i_n},$$

where we have used Einstein's convention of implicit summation over repeated indices $i_k$. The expression of the Poisson bracket in these new variables involves the derivatives with respect to $\mathbf{v}$ given by

$$\partial_{v_l} F_f = \sum_{n=0}^{\infty} \sum_{k=1}^{n} F_{P_n^{i_1 \cdots i_n}}\ v_{i_1} \cdots v_{i_{k-1}} \delta^l_{i_k} v_{i_{k+1}} \cdots v_{i_n}.$$

We use the following symmetrization of the tensor $P_n$

$$F_{P_n^{(i_1 \cdots i_n)}} = \frac{1}{n} \sum_{k=0}^{n-1} F_{P_n^{i_{n-k+1} \cdots i_n i_1 \cdots i_{n-k}}}.$$

In such a way, the derivative with respect to $\mathbf{v}$ becomes

$$\partial_{v_l} F_f = \sum_{n=0}^{\infty} n F_{P_n^{(i_1 \cdots i_n)}}\ v_{i_1} \cdots v_{i_{n-1}} \delta^l_{i_n}.$$

In these variables, the Hamiltonian becomes

$$H(\{P_n\}, \mathbf{E}, \mathbf{B}) = \int d^3x\ \frac{P_2^{ii}}{2} + \int d^3x\ \frac{\mathbf{E}^2 + \mathbf{B}^2}{2}, \tag{9.6}$$

and the expression of the bracket becomes

$$\begin{aligned}\{F, G\}_V = &\sum_{m,n} \int d^3x\ m P^{\alpha\beta}_{n+m-1} \partial_k F_{P_n^\alpha} G_{P_m^{(\beta k)}} + \sum_{m,n} \int d^3x\ nm P^{\alpha\beta}_{n+m-2} \frac{eB_{ij}}{2} F_{P_n^{(\alpha i)}} G_{P_m^{(\beta j)}} \\ &+ \sum_n \int d^3x\ en P^\alpha_{n-1} F_{P_n^{(\alpha i)}} G_{E_i} + \int d^3x\ F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} - (F \leftrightarrow G),\end{aligned} \tag{9.7}$$

where $B_{ij} = \varepsilon_{ijk} B^k$ with $\varepsilon_{ijk}$ the Levi-Civita tensor, and $(F \leftrightarrow G)$ indicates that the terms obtained by inverting $F$ and $G$ in the summation have to be subtracted (in order to fulfill the antisymmetry property of the Poisson bracket). Here and in what follows, $\partial_k$ designates the partial derivative with respect to $x_k$; it acts on the next term, e.g. $\partial_k fg = (\partial_k f)g$. The Greek indices $\alpha$ and $\beta$ in Eq. (9.7) denote a set of indices so as to complete the summation. For instance, in $P_n^{(\alpha i)}$, the indices $\alpha$ is a set of $n-1$ indices $\alpha = (i_1, \ldots, i_{n-1})$ so that $P_n^{(\alpha i)} = P_n^{(i_1 \ldots i_{n-1} i)}$.

The commonly used fluid model corresponds to a truncation of the moments at order one, keeping as field variables only the fluid density $\rho = P_0$ and the momentum density $\mathbf{M} = P_1$ (or equivalently the fluid velocity defined as $\mathbf{M}/\rho$). The bracket (9.7) has the particular property that the set of all functionals of the reduced field variables $(\rho, \mathbf{M}, \mathbf{E}, \mathbf{B})$ is a *subalgebra*, i.e. it is invariant, in the sense that the Poisson bracket (9.7) of two functionals $F(\rho, \mathbf{M}, \mathbf{E}, \mathbf{B})$ and $G(\rho, \mathbf{M}, \mathbf{E}, \mathbf{B})$ is again a functional of $(\rho, \mathbf{M}, \mathbf{E}, \mathbf{B})$. For this subset of functionals, the bracket reduces to

$$\begin{aligned}\{F, G\}_1 = &\int d^3x\ \left( (P_0 \partial_k F_{P_0} + P_1^j \partial_k F_{P_1^j}) G_{P_1^k} + eP_0 \left( B_{ij} F_{P_1^i} G_{P_1^j}/2 + F_{P_1^i} G_{E_i} \right) + F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} \right) \\ &- (F \leftrightarrow G),\end{aligned} \tag{9.8}$$

which is indeed again a functional of $(\rho, \mathbf{M}, \mathbf{E}, \mathbf{B})$. At this stage, the reduction of the dynamics has not yet been performed because the Hamiltonian $H$ given by Eq. (9.6) depends on $P_2$ so it does not belong to the subalgebra of functionals of $(\rho, \mathbf{M}, \mathbf{E}, \mathbf{B})$. In order to perform the reduction, one needs to make an assumption on the Hamiltonian. For example, in the so-called cold plasma model [7], the reduced Hamiltonian is $H^* = \int d^3x\ (\mathbf{M}^2/(2\rho) + (\mathbf{E}^2 + \mathbf{B}^2)/2)$, i.e., the original kinetic part of the Hamiltonian $\int d^3x\ P_2^{ii}/2$ has been replaced by $\int d^3x\ \mathbf{M}^2/(2\rho)$. Practically, the choice

is inspired by the physics of the system under consideration. In the presence of scalar pressure terms and in the absence of heat flux (adiabatic closure), the entropy $s$ has to be included following Ref. [7]. To account for this, $s$ has to be considered as an independent scalar field advected by the fluid. This is achieved by including the advection term given by

$$\int d^3x \; G_s F_{\mathbf{M}} \cdot \nabla s - (F \leftrightarrow G),$$

into the bracket. The addition of these terms preserves the Jacobi identity, since it only corresponds to a passive scalar advected by the Hamiltonian fluid. The last step is to include the pressure term in the Hamiltonian which may be any function $\rho U(\rho, s)$. Accordingly, $U(\rho, s)$ stands for the thermodynamical internal energy per unit mass, and the pressure is $\rho^2 \partial_\rho U$. The model we get is the most common Hamiltonian fluid model (see Ref. [104]) with the Hamiltonian

$$H[\rho, \mathbf{M}, \mathbf{E}, \mathbf{B}] = \int d^3x \; \left( \frac{\mathbf{M}^2}{2\rho} + \rho U(\rho, s) + \frac{\mathbf{E}^2 + \mathbf{B}^2}{2} \right),$$

and the bracket

$$\begin{aligned} \{F, G\} = \int d^3x \; \Big[ & \rho \; G_{\mathbf{M}} \cdot \nabla F_\rho + \mathbf{M} \cdot \; (G_{\mathbf{M}} \cdot \nabla) F_{\mathbf{M}} + G_s \; F_{\mathbf{M}} \cdot \nabla \; s + e\rho \mathbf{B} \cdot F_{\mathbf{M}} \times G_{\mathbf{M}}/2 \\ & + e\rho F_{\mathbf{M}} \cdot G_{\mathbf{E}} + F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} \Big] - (F \leftrightarrow G). \end{aligned}$$

The introduction of $U$ in the Hamiltonian is a way of keeping some information that is enclosed in the pressure tensor. As a result, this model is almost a reduction at order 1 in the moment series, but it is related to a closure at order 2 as well.

## 9.3 Fluid reduction at orders higher than 1

As it was shown above, the Hamiltonian fluid reduction at order 1 is given by a Poisson subalgebra of the parent Hamiltonian structure. This results in a reduced bracket that is just a truncation of the initial bracket (9.7). However, in many cases such a fluid model with only the first two moments is not rich enough to account for all the physics of interest, and one has to include kinetic effects and retain higher order moments in the fluid model. In this section we consider the Hamiltonian derivation of such higher order models. It should be pointed out that including the second order moments would be particularly interesting since the Hamiltonian belongs to the reduced algebra, and the closure does not affect the Hamiltonian.

The goal is to consider generalizing the results of the previous section to get fluid models of order higher than 1 by truncating the initial bracket. This method seems natural in the case of the fluid reduction, which relies on the idea that all the physical information is contained in the set $\{P_n\}_{n \leqslant N}$ of the $N$ first moments of $f$. The remaining moments will have to be expressed as functionals of them, $P_{i>N} = \Phi_i(\{P_n\}_{n \leqslant N})$. Here we consider the simplest closure assumption $P_n = 0$ for $n > N$.

It should be noticed that it would actually be best suited to consider moments defined by quantities like $\Pi_n = \int d^3 \; f(\mathbf{v} - \mathbf{M}/\rho)^{\otimes n}/\rho$. However it makes the discussion slightly more complicated, so for the sake of clarity, we have kept the moments as defined by Eq. (9.5) in this section. We consider reduced brackets defined from the moments $\Pi_n$ in Sec. 9.4.

We look for the reduced bracket by removing the undesired moments from the initial Poisson bracket (9.7), i.e., we consider the subset of functionals $F(f, \mathbf{E}, \mathbf{B}) = \bar{F}(\{P_n\}_{n \leq N}, \mathbf{E}, \mathbf{B})$ and we truncate the Poisson bracket (9.7) acting on two functionals of this subset by removing all terms proportional to $P_n$ for $n > N$. This bracket does not involve functional derivatives $F_{P_n}$ for $n > N$ since these terms vanish when acting of an element of the subset. We denote this truncated bracket $\{F, G\}_N$. Since $\{F, G\}_N = \{F, G\}_V$ for all functionals $F$ and $G$ of this subset by enforcing $P_n = 0$ for $n > N$, the reduced bracket is automatically algebraically closed and it retains from the initial bracket $\{\cdot, \cdot\}_V$ the bilinearity, the antisymmetry and the Leibniz rule. The only property one

has to check to have a Hamiltonian structure is the Jacobi identity, since this property does not automatically transfer to truncated brackets.

In order to be more specific, we inspect more closely the case $N = 2$. The reduced bracket is given by

$$\{F, G\}_2 = \{F, G\}_1 + \{F, G\}_2',$$

where the bracket $\{F, G\}_1$ is given by Eq. (9.8) and $\{F, G\}_2'$ is a bracket which only involves $F_{P_2}$ or $G_{P_2}$. This bracket is given by

$$\{F, G\}_2' = \int d^3x \ \Big(2(P_1^j \partial_k F_{P_0} + P_2^{ij} \partial_k F_{P_1^i}) G_{P_2^{(jk)}} + P_2^{ij} \partial_k F_{P_2^{ij}} G_{P_1^k} + 2e P_1^i F_{P_2^{(ij)}} G_{E_j}$$
$$+ 2e B_{ij} (P_1^k F_{P_2^{(ik)}} G_{P_1^j} + P_2^{kl} F_{P_2^{(ik)}} G_{P_2^{(jl)}})\Big) - (F \leftrightarrow G).$$

The bracket given by Eq. (9.7) between functionals of $(\rho, \mathbf{M}, P_2, \mathbf{E}, \mathbf{B})$ involves only one term that is not a functional of $(\rho, \mathbf{M}, P_2, \mathbf{E}, \mathbf{B})$, and which is proportional to $P_3$, and it is given by

$$\{F, G\}_2'' = 2 \int d^3x \ P_3^{ijk} \partial_l F_{P_2^{ij}} G_{P_2^{(kl)}} - (F \leftrightarrow G).$$

The reduction involves $P_3 = 0$, it is why this term has been dropped from the reduced bracket $\{\cdot, \cdot\}_2$. However we cannot conclude that the truncated bracket $\{\cdot, \cdot\}_2$ is Hamiltonian since the Jacobi identity has to be checked a posteriori. We show that in fact some terms not proportional to $P_3$ are generated in the Jacobi identity by the contribution $\{\cdot, \cdot\}_2''$. These contributions originate from the terms in the bracket (9.7) that are proportional to $P_n$ for $n \leq 2$ and which involve $F_{P_3}$ or $G_{P_3}$. These terms are

$$\{F, G\}_3^c = 3 \int d^3x \ \Big(P_2^{ij} \partial_k F_{P_0} G_{P_3^{(ijk)}} + e B_{ij} P_2^{kl} F_{P_1^i} G_{P_3^{(jkl)}} + e P_2^{ij} F_{P_3^{(ijk)}} G_{E_k}\Big) - (F \leftrightarrow G)$$

If we restrict these additional contributions to the first term, i.e., in absence of magnetic and electric field, the non-trivial contribution in the Jacobi identity comes from the bracket $\{\{F, G\}_2'', H\}_3^c$ which includes terms of the form

$$-2 \int d^3x \ P_2^{ij} \partial_l F_{P_2^{ij}} G_{P_2^{(kl)}} \partial_k H_{P_0},$$

and as well two other terms, that have the same expression, but with circular permutation of $(ijk)$. This contribution to the Jacobi identity of the bracket (9.7) suggests a counterexample for the failure of the Jacobi identity for the truncated bracket $\{\cdot, \cdot\}_2$. We select $F$ as a functional of $P_2$, $G$ as a functional of $P_2$ and $H$ as a functional of $P_0$. For instance, we choose $F = \int d^3x \ (P_2^{11})^2/2$, $G = \int d^3x \ P_2^{22}$ and $H = P_0$ as a counterexample of the Jacobi identity. It leads to

$$\{\{F, G\}_2, H\}_2 + \circlearrowleft = -\partial_2^2 \left((P_2^{11})^2\right) - 4\partial_1 \left(P_2^{12} \partial_2 P_2^{11}\right).$$

The same reasoning can be performed at order $N$. The number of families of possible counterexamples to the Jacobi identity increases as $N$ increases. However a convenient one is inspired from the case $N = 2$ and involves a functional of $P_N$, a functional of $P_2$ and a functional of $P_0$. It follows that the truncated bracket $\{\cdot, \cdot\}_N$ neglects terms proportional to $P_{N+1}$ which provides a non-vanishing contribution to the Jacobi identity for $\{\cdot, \cdot\}_N$ even in the subset of functionals of $\{P_n\}_{n \leq N}$, $\mathbf{E}$ and $\mathbf{B}$.

Therefore, starting from $N = 2$, the truncated brackets $\{\cdot, \cdot\}_N$ do not satisfy the Jacobi identity.

We notice that the reduced bracket $\{\cdot, \cdot\}_N$ contains most of the terms of the initial bracket $\{\cdot, \cdot\}_V$ when acting on functionals of the reduced variables $\{P_n\}_{n \leq N}$, $\mathbf{E}$ and $\mathbf{B}$. It is then more efficient to study the few removed terms than the many kept ones. This suggests a simpler way of verifying the Jacobi identity by only considering the truncated terms. It is explained in the Appendix.

## 9.4 Fluid brackets expressed with pressure-like moments

In Refs. [136–138], a warm fluid model for a collisionless low temperature relativistic plasma has been introduced. It gives dynamical equations for moments of the kinetic (Vlasov) distribution up to order two, i.e., it involves $P_0$ (fluid density), $\Pi_1 = P_1/P_0$ (fluid velocity) and

$$\Pi_2^{ij} = \frac{1}{P_0}\int d^3v\ (v_i - \Pi_1^i)(v_j - \Pi_1^j) f.$$

It is argued in Refs. [136, 138] that this model is Hamiltonian and the bracket is given by

$$\{F, G\}_2 = \{F, G\}_1 + \{F, G\}_2', \tag{9.9}$$

where

$$\{F, G\}_1 = \int d^3x \left(\partial_k F_{P_0} G_{\Pi_1^k} + \frac{1}{2P_0}(\partial_k \Pi_1^l - \partial_l \Pi_1^k + eB_{kl}) F_{\Pi_1^k} G_{\Pi_1^l} + eF_{\Pi_1^k} G_{E_k} + F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}}\right) - (F \leftrightarrow G).$$

and

$$\begin{aligned}\{F, G\}_2' = \int d^3x \ \Bigg(&\frac{\partial_k \Pi_2^{rs}}{P_0} F_{\Pi_1^k} G_{\Pi_2^{(rs)}} + 2\Pi_2^{rs}\partial_k\left(\frac{F_{\Pi_1^s}}{P_0}\right) G_{\Pi_2^{(rk)}} \\ &+ \frac{2\Pi_2^{rs}}{P_0}(\partial_k \Pi_1^l - \partial_l \Pi_1^k + eB_{kl}) F_{\Pi_2^{(kr)}} G_{\Pi_2^{(ls)}}\Bigg) - (F \leftrightarrow G).\end{aligned}$$

This model is derived in a very similar way as in the previous section, except that, instead of going from $f$ to $P_n$, the change of variables is from $f$ to $\Pi_n$ where $\Pi_0 = P_0$, $\Pi_1 = P_1/P_0$, and for $n \geqslant 2$, $\Pi_n = \int d^3v\ (\mathbf{v} - \Pi_1)^{\otimes n}/P_0$. The reduction at order $N$ corresponds to setting $\Pi_n = 0$ for $n > N$. In particular, in the derivation of the model given in Ref. [136], the contributions in the bracket which are proportional to $\Pi_3$ have been neglected given a specific assumption on the Vlasov distribution. However, as in the previous section, these neglected terms contribute to satisfying the Jacobi identity. Given the counterexample

$$\begin{aligned} F &= P_0 \Pi_2^{11}, \\ G &= \int d^3x\ P_0 \Pi_2^{22}, \\ H &= \int d^3x\ P_0 \Pi_2^{33}, \end{aligned}$$

the bracket(9.9) fails to satisfy the Jacobi identity since

$$\{\{F, G\}, H\} + \circlearrowleft = 8P_0 \Pi_2^{12} \partial_2 \left(\frac{1}{P_0}\partial_3 (P_0 \Pi_2^{13})\right) + 4\partial_2 \Pi_2^{11} \partial_3 (P_0 \Pi_2^{23}) - (2 \leftrightarrow 3),$$

where $(2 \leftrightarrow 3)$ indicates the same two terms where $\Pi_2^{12}$ has been exchanged with $\Pi_2^{13}$, and $\partial_2$ with $\partial_3$.

## Conclusion

The derivation of reduced fluid models from parent kinetic models (like Vlasov-Maxwell equations) is a delicate task since most often the Hamiltonian structure is lost by truncating the hierarchy of dynamical equations. The truncation at order one, thus consisting of dynamical equations for the density and the fluid velocity, does not pose any problem since the resulting model has a valid Hamiltonian structure. However starting at order two, and thus involving dynamical equations for the pressure tensor, the truncation might not lead to a Hamiltonian structure since in general the Jacobi identity is not satisfied. We have shown this statement by considering a fluid example taken from Ref. [136]. The rationale goes as follows : In Ref. [136], it is mentioned that the reduction "is exact: for any functionals $F[n, \mathbf{P}, \mathbf{\Pi}]$ and $G[n, \mathbf{P}, \mathbf{\Pi}]$ we have $\{F, G\}_M = \{F, G\}_V$.

As a consequence we see that the moment bracket inherits the Jacobi identity (as well as all other properties) from the full bracket." Here the bracket $\{\cdot,\cdot\}_M$ refers to the reduced bracket $\{\cdot,\cdot\}_2$ given by Eq. (9.9) and $\{\cdot,\cdot\}_V$ refers to the Vlasov-Maxwell bracket (9.4). The first part of the sentence is correct. This is why the reduced bracket inherits all the properties of the Vlasov-Maxwell bracket that are linked to the values of the bracket[1]: bilinearity, antisymmetry and Leibniz rule. However the Jacobi identity involves not only the values of the bracket but also their gradients, and in general,

$$\{\{F,G\}_M,H\}_M \neq \{\{F,G\}_V,H\}_V,$$

even if $\{F,G\}_M = \{F,G\}_V$ (after some assumption on the Vlasov distribution which is not preserved by the flow, or after neglecting higher order terms), so the reduced bracket can fail the Jacobi identity. We have shown that the model introduced in Ref. [136] involving moments up to order two does not satisfy the Jacobi identity, by exhibiting a counterexample.

To our knowledge, there is no Hamiltonian fluid models involving the pressure tensor as a dynamical field variable independent from the fluid density and fluid velocity. One of the few attempts was precisely Ref. [136]. In order to derive Hamiltonian fluid models involving higher order moments, like the pressure tensor, instead of truncating the Poisson bracket, it would be worth considering reduction methods which preserve the Hamiltonian character. In this chapter, we rather focused on subalgebra methods, but perhaps other methods would be more suited, for example other kinds of Hamiltonian truncations, such as the ones associated to the lowest order of an expansion in a small parameter. We have already started to explore this direction, but the results are not conclusive yet, and we have not had the time to report them here. Alternatively, Dirac's theory of constrained Hamiltonian systems constitutes a good candidate, even if we have not considered it as a first line and it does not come as the most natural method, since instead of dealing with a finite number of fluid field variables, it would deal with an infinite number of constrained variables.

On another hand, this chapter mainly aimed at investigating Hamiltonian reductions in plasma physics by applying the methods developed in previous chapters to specific instances of reductions. A natural prolongation is to consider other instances as well. One of them is the reduction from the fluid model to the MHD model, which will be investigated in the next (and last) part of the work. To avoid excessive length for the main text of this manuscript, we have put it in the appendix, in Chapter 14.

⟶ See the appendix chapter 14

## Appendix: Verifying the Jacobi identity for truncated brackets

In this appendix, we derive a criterion for a truncated bracket to satisfy the Jacobi identity. For the sake of clarity, we only consider systems with a finite number of variables. The extension to field variables is straightforward. First, we consider a Poisson bracket written as $\{F,G\} = F_{z_i}\mathbb{J}^{ij}G_{z_j}$ where $F_{z_i}$ is the derivative of $F(\mathbf{z})$ with respect to $z_j$. The Poisson matrix $\mathbb{J}$ satisfies the Jacobi identity rewritten as

$$\sum_{(ijk)} \mathbb{J}^{il}\partial_{z_l}\mathbb{J}^{jk} = 0\,, \tag{9.10}$$

where $\sum_{(ijk)}$ means summation over circular permutations of the indices $i$, $j$, $k$. For the reduction we split the set of dynamical variables in two subsets $\mathbf{z} = (\mathbf{Z},\boldsymbol{\zeta})$, where $\mathbf{Z}$ are the reduced variables and $\boldsymbol{\zeta}$ are the constrained variables we would like to get rid of. With this splitting of variables, we

[1] Notice that the corresponding reduction method has the strong advantage of guaranteeing a conserved energy, as well as conservations laws and symmetries associated to Casimir invariants, in relation to the initial structure, and also a consistent ordering in the reduction process. The only thing it loses is the Jacobi identity, and hence the Hamiltonian character, with the associated conservative structures (Liouville theorem, Poincaré invariants, etc.) and tools (Darboux theorem, Birkhoff theorem, KAM theory, etc. ).

rewrite the Poisson matrix in block form :

$$\mathbb{J} = \begin{pmatrix} \mathbb{A} & \mathbb{B} \\ -\mathbb{B}^T & \mathbb{C} \end{pmatrix},$$

where, e.g., the matrix $\mathbb{A}$ brings contributions of the form $F_{Z_i}\mathbb{A}^{ij}G_{Z_j}$ in the Poisson bracket. Between two reduced functions $F(\mathbf{Z})$ and $G(\mathbf{Z})$, the reduced bracket is defined as the truncated one

$$\{F, G\}_* := \lim_{\boldsymbol{\zeta}\to 0}\{F, G\}\,, \tag{9.11}$$

which means that the reduced Poisson matrix is simply the truncation of $\mathbb{J}$ obtained by removing terms involving $\partial_{\boldsymbol{\zeta}}$ and taking the limit $\boldsymbol{\zeta} \to 0$, i.e.,

$$\mathbb{J}_* = \lim_{\boldsymbol{\zeta}\to 0}\mathbb{A}\,.$$

By its definition, it is bilinear, antisymmetric and satisfies the Leibniz rule, because these properties involve only the value of the bracket. As for the Jacobi identity, it involves derivatives in the bracket as well, and is therefore not guaranteed. In order to get a criterion for its validity, we use the fact that $\mathbb{J}$ satisfies the Jacobi identity. First, we Taylor expand its components in $\boldsymbol{\zeta}$ :

$$\begin{aligned} \mathbb{A} &= \mathbb{A}_0 + \mathbb{A}_1 + \mathbb{A}_{\geq 2}, \\ \mathbb{B} &= \mathbb{B}_0 + \mathbb{B}_{\geq 1}, \\ \mathbb{C} &= \mathbb{C}_0 + \mathbb{C}_{\geq 1}, \end{aligned}$$

where the subscript denotes the order in the $\boldsymbol{\zeta}$ variables. Between two reduced functions $F(\mathbf{Z})$ and $G(\mathbf{Z})$, the Jacobi identity (9.10) then writes

$$\sum_{(ijk)} \mathbb{A}^{il}\partial_{Z_l}\mathbb{A}^{jk} + \sum_{(ijk)} \mathbb{B}^{il}\partial_{\zeta_l}\mathbb{A}^{jk} = 0. \tag{9.12}$$

This relation is verified at each order in $\boldsymbol{\zeta}$. As a result of Eq. (9.11), the Jacobi identity for the reduced Poisson matrix $\mathbb{J}_*$ is involved in the zeroth order of Eq. (9.12) which writes

$$\sum_{(ijk)} \mathbb{A}_0^{il}\partial_{Z_l}\mathbb{A}_0^{jk} + \sum_{(ijk)} \mathbb{B}_0^{il}\partial_{\zeta_l}\mathbb{A}_1^{jk} = 0.$$

The first term corresponds to the Jacobi identity for $\mathbb{J}_* = \mathbb{A}_0$. So, a necessary and sufficient condition for $\mathbb{J}_*$ to satisfy the Jacobi identity is

$$\sum_{(ijk)} \mathbb{B}_0^{il}\partial_{\zeta_l}\mathbb{A}_1^{jk} = 0\,.$$

In the example of Sec. 9.3, $\mathbb{B}_0$ corresponds to $\{\cdot,\cdot\}_3^c$, and $\mathbb{A}_1$ to $\{\cdot,\cdot\}_2''$.
There are cases where this criterion is automatically satisfied, e.g., because of an ordering argument or a Lie subalgebra. This last case, for instance, corresponds to the trivial condition $\mathbb{A}_{\geq 1} = 0$. For the validity of the reduction, this condition can be generalized to the case $\mathbb{A}_1 = 0$ which is sufficient to verify the Jacobi identity. However this assumption is not often verified since brackets generally contain linear terms in the variables [159].

# Conclusion of the episode

This third episode was interested in Hamiltonian reductions for plasma physics, and especially in variations of Dirac reductions. It explained why several instances of Dirac's theory of constraints ended up with a reduced Poisson bracket much simpler than the computation required by the procedure in order to obtain it. A formulation of the Dirac reduction as a projection of derivatives showed that the phenomenon was related to the presence of a simplified projector, originating from a special structure in the initial Poisson bracket, which was quarter-canonical.

This structure was seen to be fairly common in previous Dirac reductions. It clarified some aspects of the procedure, especially about the choice of secondary constraints, which both explained several features in previously identified reductions and made the method more efficient for future applications.

Variations from the Dirac reduction were obtained because in the presence of a quarter-canonical bracket, the Dirac reduction can be equivalently considered as a bracket truncation, justified by a subalgebra, whose computations are elementary and whose procedure replaces the invertibility condition of the Dirac reduction by a subalgebra criterion, often readily checked.

This opened a wide scope in the sense that the applicability condition is much softer. In addition, given a set of constraints, several reduced structures can be generated, which was not the case for the Dirac method.

Various applications in plasma physics were considered, beginning with the quarter-canonical examples of an extension from the Vlasov-Maxwell system and of the reductions to the Vlasov-Poisson system and to incompressible models, fluid or kinetic. On another hand, the reduction from the kinetic to the fluid model was found to be given by a subalgebra projection, as well as the reduction from the fluid model to magnetohydrodynamics.

Extensions are available, if not needed, because the MHD model has several variations, whose Hamiltonian structure could be significantly different from standard MHD. As for the fluid model, we were interested in Hamiltonian closures including moments of order two or higher. Obstructions and counterexamples were found, but no example solution. So, further explorations are desirable in order to identify suitable closures and to better understand the mechanisms at work.

More importantly for our purpose, the developed tools could be applied to the last step of the gyrokinetic reduction, the removal of the fast gyro-angle (and also the magnetic moment) from the base space of the theory. This should provide the Hamiltonian structure of gyrokinetics, although additional investigations is needed for a complete understanding of this Hamiltonian averaging process, especially when taking into account the time dependence of the guiding-center coordinate transformation, but also more basically in order to confirm that suitable Dirac constraints do exist in the time-independent setting. Contrary to the work in the previous episode, it does not seem to have effects on the gyrokinetic equations of motion, but rather on admissible transformations, with conclusions agreeing with guiding-center results.

The role of the Dirac reduction together with the quarter-canonical structure in the local-interaction matrix is to make the reduction of particle dynamics available at the field level. So, in a similar way as in the previous episode, the process mixes up field reductions and particle reductions; here the origin relies on the Lie-Poisson structure of the Vlasov bracket. A subalgebra reduction is not so efficient in this case, because it can be performed in a more restricted case where the

electromagnetic field is external.

In this reduction, the role of the quarter-canonical structure in the guiding-center Poisson bracket is important not only to confirm that the gyrokinetic equations of motion are indeed Hamiltonian, but also to make future variations in the model possible by emphasizing the conditions that ensure the Hamiltonian character for the reduced dynamics.

For instance, it will be useful when studying truncation questions for numerical simulations, as was mentioned in the conclusion of the previous episode. In addition, it will be crucial in order to identify how the conservation laws for the Vlasov-Maxwell system transfer into the conservation laws for gyrokinetics [21, 133]. These are two natural extensions of the analysis, since they constituted practical motivations for the work, and more generally for a better understanding of the Hamiltonian aspects of the gyrokinetic reduction.

As a result, this episode succeeded in developing variations of Dirac's reduction method and in starting an application to the step 2d of the gyrokinetic reduction. It made links with other Hamiltonian reduction methods, enlightened several reductions in plasma physics, and opened prospects about these reductions. So, an obvious extension, in addition to finalizing the gyrokinetic reduction, is to investigate other examples of Hamiltonian reductions in plasmas, since also they could be clarified or improved by a Hamiltonian approach.

In this episode, the work was mainly interested in studying examples where both the initial and the final model were known to be Hamiltonian, in order to get a better understanding of the reduction methods. In this direction, there is still much to be done, since the work only focused on a few reductions about the four main models for plasmas.

However, a more attractive work will be to apply the developed tools to reductions where only the initial model is known to be Hamiltonian, and the reduced model is either identified but unsure to be Hamiltonian, or even not definitely identified yet. In such works, in the same way as in the case of guiding-center or gyrokinetic theory, the efficiency of Hamiltonian reduction tools will be still more useful and better emphasized. For instance, it might be interesting to consider Hamiltonian closures starting from the gyrokinetic model, in order to identify Hamiltonian fluid models including FLR effects, and to compare them with present-day gyro-fluid models, e.g. [14, 55, 75, 147, 154, 155].

On another hand, from the point of view of Hamiltonian reduction theory, an obvious extension is to investigate other reduction methods than the ones used here, which mainly concerned Dirac, subalgebra, or ordering reductions. Perhaps the most natural one would regard the moment map [94, 99, 124], which can be viewed as a generalization of the Dirac reduction and has already proven efficient in plasma physics, since it was involved in the foundations of the Hamiltonian structure of Vlasov-Maxwell [95] and of the BBGKY hierarchy equations [96], for instance.

The relabelling symmetry [125, 126] could provide an adequate example of application, and it should bring additional light on an important step when going from particle descriptions to kinetic models. It would be all the more attractive as among the sequence of the four main models of plasma dynamics, it is the only Hamiltonian reduction which has not been investigated in the present work. This reduction is roughly related to a duality relation between the particle phase space and the set of functions defined over this phase space [5], but much more is hidden behind this duality, and above all a quotient reduction by particles' indiscernibility, with the associated relabelling symmetry.

## Selected bibliography


[5] I. Bialynicki-Birula, J. C. Hubbard, and L. A. Turski, "Gauge-independent canonical formulation of relativistic plasma theory", Physica A **128**, 509 (1984).

[31] C. Chandre, P. J. Morrison, and E. Tassi, "Hamiltonian derivation of the Charney-Hasegawa-Mima equation", Phys. Plasmas **16**, 082301 (2009).

[32] C. Chandre, P. J. Morrison,and E. Tassi, "Derivation of reduced two-dimensional fluid models via Dirac's theory of constrained Hamiltonian systems", Phys. Plasmas **17**, 042307 (2010).

[33] C. Chandre, P. J. Morrison, and E. Tassi, "On the Hamiltonian formulation of incompressible ideal fluids and magnetohydrodynamics via Dirac's theory of constraints", Phys. Lett. A **376**, 737 (2012).

[34] C. Chandre, L. de Guillebon, A. Back, E. Tassi, and P. J. Morrison, "On the use of projectors for Hamiltonian systems and their relationship with Dirac brackets", J. Phys. A: Math. Theor. **46**, 125203 (2013).

[44] P.A.M. Dirac, "Generalized hamiltonian dynamics", Can. J. Math. **2**, 129 (1950).

[58] L. de Guillebon and C. Chandre, "Hamiltonian structure of reduced fluid models for plasmas obtained from a kinetic description", Phys. Lett. A **376**, 3172 (2012).

[66] A. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Systems*, Accademia Nazionale dei Lincei, Roma, 1976.

[72] D. D. Holm and B. A. Kupperschmidt, "Hamiltonian theory of relativistic magnetohydrodynamics with anisotropic pressure", Phys. Fluids **29**, 3889 (1986).

[95] J. E. Marsden and A. Weinstein, "The Hamiltonian structure of the Maxwell-Vlasov equations", Physica D **4**, 394 (1982).

[102] P. J. Morrison and J. M. Greene, "Noncanonical Hamiltonian Density Formulation of Hydrodynamics and Ideal Magnetohydrodynamics", Phys. Rev. Lett. **45**, 790 (1980); **48**, 569 (1982).

[103] P. J. Morrison and J. M. Greene, "Errata for "Noncanonical Hamiltonian Density Formulation of Hydrodynamics and Ideal Magnetohydrodynamics"" Phys. Rev. Lett. **48**, 569 (1982).

[104] P. J. Morrison, "Poisson brackets for fluids and plasmas", AIP Conference Proceedings **88**, 13 (1982).

[107] P. J. Morrison, "Hamiltonian description of the ideal fluid", Rev. Mod. Phys. **70**, 467 (1998).

[118] S. Nguyen and L. A. Turski, "Canonical description of incompressible fluid: Dirac brackets approach", Physica A **272**, 48 (1999).

[119] S. Nguyen and L. A. Turski, "Examples of the Dirac approach to dynamics of systems with constraints", Physica A **290**, 431 (2001).

[136] B. A.Shadwick, G. M. Tarkenton, and E. H. Esarey, "Hamiltonian description of low-temperature relativistic plasmas", Phys. Rev. Lett. **93**, 175002 (2004).

[137] B. A. Shadwick, G. M. Tarkenton, E. H. Esarey, and C. B. Schroeder, "Fluid and Vlasov models of low-temperature, collisionless, relativistic plasma interactions", Phys. Plasmas **12**, 056710 (2005).

[138] B. A.Shadwick, G. M. Tarkenton, E. H. Esarey, and F. M. Lee "Hamiltonian reductions for modeling relativistic laser-plasma interactions", Commun. Nonlinear Sci. Numer. Simul. **17**, 2153 (2012).

[143] R. G. Spencer and A. N. Kaufman, "Hamiltonian structure of two-fluid plasma dynamics", Phys. Rev. A **25**, 2437 (1982).

[144] J. Squire, H. Qin, W. M. Tang, and C. Chandre, "The Hamiltonian structure and Euler-Poincaré formulation of the Vlasov-Maxwell and gyrokinetic systems", Phys. Plasmas in press (2013).

# General conclusion

At the end of this dissertation, it is time to take a step back, and to have a more general point of view both on the road travelled and on the horizon. This will avoid redundancy with the conclusions of the episodes, where more detailed results and prospects were summarized and considered.

## Positioning of the matter

From a general point of view, the works reported in the present document emphasize and develop the strong mutual interests between plasma physics and dynamical systems, especially between plasma models and Hamiltonian reductions. On the one hand, they remind that plasma physics involves many models and especially many Hamiltonian models, and that it is interested in reduction tools in order to improve its models or derive new models. On the other hand, they show that dynamical systems and reduction methods are interested in disposing of a wide set of models closely related to each other, such as plasma models, because it constitutes a rich domain for applying reduction tools, and applications often generate new developments in the methods.

The subject mainly focused on the gyrokinetic reduction, because of its role in present-day plasma physics, and because it constitutes by itself a rich example of reduction for dynamical systems. It involves reductions by a constant of motion and by an averaging procedure, both in finite and infinite dimensional systems, together with exchanges between the two through the lifting and the plasma-field coupling. It took time to become well-established and several foundational questions are only in the process of being clarified.

In order to answer such questions, and also to better understand gyrokinetics as well as other plasma models, various reduction tools have proven useful, applied either just to the equations of motion, or to the Hamiltonian or variational structure: averaging methods, Lie-transforms, expansions of differential equations, lifting procedure, Hamiltonian perturbation, constrained systems, Dirac and subalgebra reductions, etc. Especially, in several places Hamiltonian reduction tools appeared as very convenient in order to make a reduction more efficient, or to elucidate subtleties or issues about it, or even to identify necessary corrections.

## Main points of the work

The previous chapters bring contributions to clarify the initial four questions, but, at the same time, there is still much to be done for a more complete exploitation of the results.

The higher-order reduction was addressed by obtaining explicit induction relations to compute the reduction to arbitrary order, which gave a better understanding of the reduction mechanism and of possible representations for a maximal reduction. However, practical implementations for higher orders (up to full second order for instance) have not been done yet, both because other collaborators were interested in deriving them and because we feel it will be best done by introducing computer-assisted symbolic calculus, which was not our priority and has been delayed up to now.

The troubles associated to the traditional gyro-angle were clarified by introducing a gauge-independent but constrained coordinate system directly induced by the physics. Especially, the traditional questions about guiding-center anholonomy and gauge arbitrariness were shown to be related to intrinsic, yet regular properties of the physical system, and that the difficulties can be made to disappear completely. However, we did not address complementary questions, such as the role of contractible loops in the existence condition for a scalar gyro-angle, which should be only

optional, or the quantification of the flux of the vector field $\mathbf{N} := \nabla \times \mathbf{R}$ [90, 91], which seems to have deep, yet abstract foundations and wide perspectives.

The Hamiltonian approach of the gyrokinetic reduction was addressed with a lifting of the particle reduction to the Hamiltonian structure of the Vlasov-Maxwell field dynamics, together with an application of Dirac's theory of constraints. This developed each of the steps involved in the reduction process by providing it with a Hamiltonian approach. Especially, it implemented the coupling between the plasma and the electromagnetic field in a more consistent way, and the moments of the Vlasov density were shown to be involved in the gyro-center transformation. However, the Hamiltonian removal of the gyro-angle has not been fully completed yet and consequences on the conservation laws of gyrokinetics have not been addressed yet.

Variations of Dirac's theory and their application to plasma models were developed, with especially the role of projections, bracket truncations, quarter-canonical structures, or subalgebras reductions, and with an investigation of fluid closures and of the MHD reduction. However, a fluid closure including moments of order two or higher has not been obtained yet.

## For the future

In addition to these fundamental extensions of the thesis work, other possible continuations naturally appeared in the process, such as implementing also the bounce-angle reduction, or exploring Hamiltonian fluid closures from the gyrokinetic model. So, even if the thesis provided many results about the four initial questions and contributed to a better understanding of the underlying theory, it also lead to wide perspectives to be explored in the future, with many possible complementary questions, extensions or research directions. In some way, it may be considered as providing more questions than answers, since each of the results typically opened several possible continuations, both about theoretical and about more applied concerns.

Perhaps the most important points concern the higher-order computations and the Hamiltonian structure of gyrokinetics, with their possible impact on gyrokinetic theory. The primary extension is to more completely understand the Hamiltonian removal of the gyro-angle and thus the Hamiltonian structure of the gyrokinetic equations. Then the consequences of the results will be worth studying, especially for conservation laws [21, 127, 133] and for the Hamiltonian structure of the equations implemented in gyrokinetic codes [42, 56].

This would be all the more interesting as the work focused on foundational questions in the theory. Hence it naturally calls for a complementary part, concerned with more practical aspects, and especially with applications of the results.

On another hand, a wide domain to be explored comes because the methods worked on are not specific to gyrokinetics. They can be applied to various models, and clarify issues in other fields of physics. To mention only a few examples, Hamiltonian systems play an important role in particle, quantum, atomic, and statistical physics, algebraic and geometrical structures are often crucial for dynamical systems, and structure-preserving numerical methods are an active research, as is again emphasized in [13], [50] and [134], for instance.

Last, more theoretical extensions are far from being uninteresting as well, either for themselves or for the many practical applications they carry along in germ since they develop tools with fairly broad range. From this point of view, the thesis touched only a very restricted domain, many related fields were left aside and would constitute interesting complements, such as other Hamiltonian reduction methods [99], KAM theory [92], control theory [38, 39, 153], stability analysis [70, 116, 148, 149], nonlinear waves [159], variational formulations of plasma models [18, 110, 113, 129, 130, 158, 159], metriplectic dissipation [100, 105, 106, 111], etc. just to randomly mention a few of them incidentally met during the thesis, some of which are directly or possibly related to gyrokinetics. For example, variational formulations have proven very efficient in the reduction of particle dynamics, as was well exemplified by Chapter 2, and they have already obtained interesting results for gyrokinetics [18, 129, 130]. As for the metriplectic dissipation, it could provide an algebraic structure for collisional Vlasov-Maxwell and gyrokinetics.

As a result, this thesis work fulfils its role as a piece of research in theoretical physics: "theoretical physics", because it is closely related both to applied physical questions and to more abstract mathematical concerns; "piece of research", because it clarifies many issues, but calls for many other questions or complementary investigations.

# Episode IV

# Appendix chapters

# Chapter 10

# Magnetic-moment reduction without a guiding center nor a gyro-gauge

in collaboration with Natalia Tronko, Michel Vittot, and Philippe Ghendrih

*A*bstract: We perform the dynamical reduction for charged particles in a strong external magnetic field using the adiabatic invariant of gyrokinetics or guiding-center theory, namely the magnetic moment, but without introducing a gyro-gauge nor involving a guiding center.
For the gyro-angle, our approach does not require to choose an axis orthogonal to the magnetic field to define the zero of the angle, with its questions about gyro-gauge and non-global existence. Intrinsic (gauge-independent and global) definitions are obtained for the operators of gyrokinetics, such as gyro-averaging and gyro-integral.
We build the normal form and compute the Poisson bracket for this Hamiltonian system (with a non-canonical structure). The results are well suited to the Vlasov-Maxwell dynamics, where they can be lifted directly thanks to the absence of a virtual guiding center in the reduction.

## Introduction

As announced in the general introduction (and in the introduction of the first episode as well), this appendix chapter takes place before Chapter 1, as the starting point when addressing the issues about the gyro-gauge by developing an intrinsic (i.e. gyro-gauge independent) formulation of the guiding-center reduction. The idea is to begin by studying a simplified reduction that initially avoids introducing a gyro-angle, with its associated gyro-gauge.

In addition, we are interested in a reduction that does not change the spatial position, because this avoids many difficulties when lifting the transformation from particle dynamics to field dynamics (see Chapters 4 and 5). This will be needed in order to take into account the coupling between the plasma and the fields in a Hamiltonian way, in the long run of achieving the Hamiltonian structure of gyrokinetic (or guiding-center) Vlasov-Maxwell equations.

Thus, in the present chapter, we focus on the reduction of particle dynamics in a strong magnetic field, but we exclude the introduction of the guiding center, and hence the averaging reduction (removal of the fast time scale from the dynamics). Only the reduction associated with the magnetic moment $\bar{\mu}$ is considered.

Indeed, including this constant of motion in the coordinates implies to change only one coordinate, for instance the norm of the momentum, whereas the five other coordinates (the position and the unit vector of the momentum, for instance) can be kept unchanged. Especially, no gyro-angle needs to be defined a priori.

Notice that the norm of the momentum also is a constant of motion, but it is usually replaced by the magnetic moment because it does not remain an adiabatic invariant in the presence of a weak electric field [122].

As for the magnetic moment, it is associated with Larmor gyration, since it is conjugated to the gyro-angle [86, 88]. A consequence, most important for gyrokinetic theory [56, 67], is that it remains an adiabatic invariant even in the presence of a weak electric field. Another consequence, more important for our purpose, is that its derivation is bound to be related to Larmor gyration, and hence to introduce a gauge-independent formulation of the basic elements of gyrokinetics: a coordinate for the gyro-angle, a generator of Larmor gyration, a gyro-integral operator, etc.

Particle dynamics is Hamiltonian, which means it is given by a Hamiltonian function and a Poisson bracket. So, the reduced dynamics is obtained not by computing the magnetic moment itself, but rather by computing the expression of the Poisson bracket and of the Hamiltonian function when the magnetic moment is used as a coordinate. This is also what will be needed for the Vlasov-Maxwell dynamics in Chapter 4.

Accordingly, the work proceeds in two sections.

In Sec. 10.1, we study how to derive the normal form, i.e. the Hamiltonian written as a function of the magnetic moment. We compute the solution to second order in the magnetic moment (or rather its square root); this is an expansion in strong magnetic field because it can be considered as an expansion in $B^{-1/2}$, or as the inverse of the particle charge $e^{-1}$, which is commonly used as an expansion parameter in strong magnetic field [29, 122]. Our results are shown to be consistent with the standard results concerning the magnetic moment [29, 86, 88, 123].

In Sec. 10.2, in order to completely define the reduced dynamics, we derive the Poisson bracket.

These results are obtained without using Lie transforms to average the Lagrangian 1-form, as is usually done, but by solving directly the equation for the normal form.

Notice that, contrary to all other chapters of the present dissertation, this chapter was first written in the relativistic framework and distinction was made between vectors and covectors, these last being denoted by an overbar (see e.g. Eq. (10.2)). We have chosen to keep these specificities because the relativistic framework does not make the derivation more complicated at all. As for the use of overbars, it is different from the notational convention used in all other chapters, where the overbar rather indicates transformed coordinates or fields (see e.g. Eqs. (5.2) and (5.3)). But this does not cause any ambiguity since here only one coordinate is changed, related to the magnetic moment $\bar{\mu}$, which is the only symbol in this chapter where the overbar will not indicate matrix transpose. Incidentally, let us announce that the changed coordinate will rather be denoted by the symbol $A$ in order to emphasize that it is an arbitrary constant of motion (preferably independent of the Hamiltonian), not a priori restricted to the magnetic moment (see Eq. (2.22)).

## 10.1 Building the normal form

In this section, we study the normal form for particle dynamics in an external magnetic field. This will be done in four steps. In Subsec. 10.1.1, we establish the defining equation for the normal form and transform it into an induction relation. In Subsec. 10.1.2, we study the initialization of the induction, from which the expression for the lowest-order constant of motion (the adiabatic invariant) is derived. In Subsec. 10.1.3, we compute the normal form to second order in the constant of the motion. Last, in Subsec. 10.1.4, we study the structure of the general term and show that the small parameter of the normal form expansion is related to the small parameter of gyrokinetics.

### 10.1.1 Equation for the normal form

The system under consideration is just a charged particle with position $\mathbf{q}$, momentum $\mathbf{p}$, mass $m$, and charge $e$, submitted to an external, static but inhomogeneous, magnetic field $\mathbf{B}$. The presence of an electric field is excluded here in order to make the scheme simpler for this introductory work about an intrinsic formulation of the guiding-center reduction; it will be taken into account in

Chapter 2. The motion is induced by the Lorentz force

$$\begin{aligned}\dot{\mathbf{q}} &= \frac{\mathbf{p}}{m\gamma} = \{H, \mathbf{q}\}\,,\\ \dot{\mathbf{p}} &= -e\mathbf{B}\times\frac{\mathbf{p}}{m\gamma} = \{H, \mathbf{p}\}\,.\end{aligned} \tag{10.1}$$

It is a Hamiltonian system, where the Hamiltonian is the particle's relativistic energy

$$H = m\gamma = \sqrt{\mathbf{p}^2 + m^2}\,.$$

Among the phase-space coordinates, it depends only on the norm of the momentum

$$h := \|\mathbf{p}\|\,.$$

So, $h$ can be viewed as a Hamiltonian, after a scaling on the time coordinate or on the Poisson bracket. For simplicity, we choose units such that the speed of light $c$ is equal to 1.

The Poisson bracket (suited to the Hamiltonian $H$) is given by[1] [29]

$$\{f, g\} = \partial_{\mathbf{p}} f.\partial_{\overline{\mathbf{q}}} g - \partial_{\mathbf{q}} f.\partial_{\overline{\mathbf{p}}} g + \partial_{\mathbf{p}} f.e\mathbf{B}\times\partial_{\overline{\mathbf{p}}} g\,, \tag{10.2}$$

for two arbitrary functions $f$ and $g$.

Our goal is to reduce the dynamics of Eqs. (10.1), which means to adopt a constant of motion $A$ as one of the coordinates in phase space. So, among the six coordinates, only five will remain dynamical. A constant of motion is linked to a continuous symmetry according to Noether theorem, but the existence of such a constant (independent of the Hamiltonian $m\gamma$) has not been proven yet. It is believed to exist for strong magnetic fields. So, rather than an exact constant, we are looking for an asymptotic series, which comes from the adiabatic invariance of the lowest-order magnetic moment, denoted by $\mu$ in order to distinguish it from the full series for the magnetic moment, denoted by $\bar{\mu}$.

The Hamiltonian expressed as a function of the constant of motion is called the *"normal form"*; in fact, the normal form will be only partial here, since we will have only one constant of motion for three degrees of freedom. The change of coordinates is performed in two steps, by expressing first the momentum $\mathbf{p}$ in polar coordinates $(h, \mathsf{v})$, and then introducing a coordinate $A$ defined as constant of motion:

$$(\mathbf{p}, \mathbf{q}) \longrightarrow (h, \mathsf{v}, \mathbf{q}) \longrightarrow (A, \mathsf{v}, \mathbf{q}) \tag{10.3}$$

where $A$ is constant of motion in bijection with $h$, and

$$\mathsf{v} := \tfrac{\mathbf{p}}{\|\mathbf{p}\|} \in \mathbb{S}^2$$

is the unit vector of the momentum. In these new variables, the equation of motion becomes

$$\begin{aligned}\dot{\mathbf{q}} &= \frac{\mathbf{p}}{m\gamma} = \frac{h\mathsf{v}}{m\gamma}\,,\\ \dot{\mathsf{v}} &= \frac{d}{dt}\frac{\mathbf{p}}{h} = \frac{\dot{\mathbf{p}}}{h} = -e\mathbf{B}\times\frac{\mathbf{p}}{hm\gamma} = -e\mathbf{B}\times\frac{\mathsf{v}}{m\gamma}\,,\\ \dot{A} &= 0\,,\end{aligned} \tag{10.4}$$

because $\mathbf{p} = h\mathsf{v}$, and $h$ is a constant of motion as well as $A$.

Let us notice that $A$ is defined only through two of its properties: it must allow for a change of coordinates (which implies bijectivity with $h$) and it must be a constant of motion. In fact, it can be different from the usual magnetic moment $\bar{\mu}$ (see Chapter 2) and it is the reason why

[1] The overbar denotes the transpose of a vector into a linear form; thus, $\partial_{\mathbf{p}}$ is a covector and $\partial_{\overline{\mathbf{p}}}$ is a vector. Be careful, this notational convention for the overbar is different from all other chapters, where the overbar usually indicates transformed coordinates. In this chapter, only $\bar{\mu}$ contains an overbar which does not mean matrix transpose.

we did not use the symbol $\bar{\mu}$. Nonetheless, in the absence of any symmetry in the magnetic-field geometry [15, 45], this Hamiltonian system is generically non integrable, which means that there exist only two independent constants of motion: $\bar{\mu}$ and $h$. Hence, if $A$ is not a pure function of $h$, it is expected to be linked to the magnetic moment $\bar{\mu}$.

Furthermore, a consequence of the non-canonical structure should be noted. Even if $A$ is a constant of motion, the Hamiltonian will not be independent of the fast part of the dynamics, corresponding to the gyro-angle $\theta$ in the guiding-center approach. Indeed, we do not have the action-angle correspondence, because this last exists only in a canonical (or at least quarter-canonical) structure:

$$0 = \dot{A} = \{H, A\} \neq -\partial_\theta H \,. \tag{10.5}$$

Now, we can get the equation for the normal form, using that $h$ and $A$ are constants of motion :

$$0 = \dot{h}(A, \mathsf{v}, \mathbf{q}) = (\partial_A h).\dot{A} + (\partial_{\mathsf{v}} h).\dot{\mathsf{v}} + (\partial_{\mathbf{q}} h).\dot{\mathbf{q}} = -(\partial_{\mathsf{v}} h).e\mathbf{B} \times \frac{\mathsf{v}}{m\gamma} + (\partial_{\mathbf{q}} h).\frac{h\mathsf{v}}{m\gamma} \,.$$

$$\text{Hence } -(\partial_{\mathsf{v}} h).e\mathbf{B} \times \mathsf{v} + (\partial_{\mathbf{q}} h) h\mathsf{v} = 0 \,. \tag{10.6}$$

where the derivatives are understood on functions of $(A, \mathsf{v}, \mathbf{q})$. It is the equation we must solve for the normal form $h(A)$, in an analogous way to the "diagonalization" of the Hamiltonian in the Hamilton-Jacobi method, or in Hamiltonian perturbation methods [27, 53].

Grouping the magnetic field with the particle charge, and decomposing it into its norm $B := \|e\mathbf{B}\|$ and its unit vector $\mathsf{b} := \frac{e\mathbf{B}}{\|e\mathbf{B}\|}$, Eq. (10.6) becomes

$$-B\mathcal{D}h = h\bar{\mathsf{v}}\nabla h \,,$$

where the gradient is just $\nabla := \partial_{\bar{\mathbf{q}}}$, and the linear operator $\mathcal{D}$ is defined by

$$\mathcal{D} := \bar{\mathsf{v}}.\mathsf{b} \times \partial_{\bar{\mathsf{v}}} \,. \tag{10.7}$$

The operator $\mathcal{D}$ is the generator of rotations of the momentum around the magnetic field and induces the cyclotronic gyration. Some of its properties are presented in App. A at the end of the chapter on page 243. In particular, it has a left-inverse $\mathcal{G}$, that is an operator which is zero on $\ker(\mathcal{D})$ and $\mathcal{D}^{-1}$ on $\mathrm{Rg}\,(\mathcal{D})$. The projectors associated with these last subspaces are precisely the gyro-fluctuating and the gyro-averaging operators. All these operators are involved in the computation of the normal form. Their definition is the following

$$\begin{aligned}
\mathcal{G} &:= \frac{\pi\mathcal{D} - \sinh(\pi\mathcal{D})}{\pi\mathcal{D}^2} = \oint d\theta(\theta - \pi \text{ sign } \theta)\, e^{\theta\mathcal{D}} \text{ is the left-inverse of } \mathcal{D}, \\
\mathtt{osc} &:= \mathcal{G}\mathcal{D} = \mathcal{D}\mathcal{G} \text{ is the gyro-fluctuating operator}, \\
\mathtt{avg} &:= 1 - \mathtt{osc} = \oint d\theta e^{\theta\mathcal{D}} = \frac{\sinh(\pi\mathcal{D})}{\pi\mathcal{D}} \text{ is the gyro-averaging operator.}
\end{aligned}$$

The next step is to isolate the unknown function $h$ in the left-hand side of Eq. (10.6). For this purpose, let us multiply by $\mathcal{G}$ on the left

$$B\ \mathtt{osc}\,(h) = -\mathcal{G}h\bar{\mathsf{v}}\nabla h \,. \tag{10.8}$$

Thus, the fluctuating part of $h$ is easily isolated in the left-hand side. As for the information about its averaged part, it is obtained by multiplying Eq. (10.6)by $\mathtt{avg}$ on the left. Then $\mathtt{avg}\,\mathcal{D} = 0$ implies

$$0 = \mathtt{avg}\,(h\bar{\mathsf{v}}\nabla h) \,. \tag{10.9}$$

It is the secular equation for this constant of motion, which is more technical to deal with, as is standard in perturbation theory.

Notice that, in order to avoid extensive use of parentheses, we adopt the convention that the operators $\mathcal{D}$, $\mathcal{G}$, $\mathtt{avg}$ and $\mathtt{osc}$ act on the whole expression on their right-hand side. Together, Eqs.

(10.8) and (10.9) are equivalent to the initial equation (10.6). So, the inversion of $\mathcal{D}$ in Eq. (10.6) produces two equations. This is due to $\mathcal{D}$ being only left-invertible, so that in order to invert it, we must keep the condition that the right-hand side of (10.6) has to be in the range of $\mathcal{D}$, which is the information contained in Eq. (10.9).

The inversion of $\mathcal{D}$ gives a different role to the average and fluctuating parts of $h$. It suggests to define

$$\alpha := \mathtt{avg}\,(h) \text{ and } \eta := \tfrac{h}{\alpha}\,.$$

The function $\alpha$ is the averaged energy of the particle, the average being taken over the direction of the perpendicular momentum $\mathbf{p}_\perp$. It is important to notice that it is not constant in the motion, even if it results from an average of the constant of motion $h$. The reason is that the average is not taken along a trajectory.

As for $\eta := \frac{h}{\alpha}$, it is the fluctuating part of $h$, in a multiplicative sense, and without dimension. By construction, we have $\mathtt{avg}\,(\alpha) = \alpha$, $\mathtt{osc}\,(\alpha) = 0$ and $\mathtt{avg}\,(\eta) = 1$. Expressing Eqs. (10.8) and (10.9) with $\alpha$ and $\eta$ and dividing them by $\alpha$ gives

$$0 = \mathtt{avg}\left(\eta\bar{\mathtt{v}}\nabla(\alpha\eta)\right), \tag{10.10}$$

$$\eta = 1 - B^{-1}\mathcal{G}\eta\bar{\mathtt{v}}\nabla(\alpha\eta)\,. \tag{10.11}$$

The last step is to transform these equations into induction relations, since the solution is expected as a formal series (related to an expansion in strong magnetic field). We expand $\alpha$ and $\eta$ in series in $A^{1/2}$:

$$\alpha = \sum_1^\infty \alpha_n A^{n/2} \quad \text{and} \quad \eta = \sum_0^\infty \eta_n A^{n/2}\,. \tag{10.12}$$

It is only for practical reasons that we choose to expand in $A^{1/2}$ instead of $A$; it makes clearer that our result agrees with the usual case $A := \bar{\mu}$. In another hand, it is convenient for the following to choose $\alpha_0 := 0$, i.e. to begin the series for $\alpha$ at order 1. This hypothesis seems reasonable for physical cases, in which $A := \bar{\mu} \approx \mu = \frac{(h\sin\varphi)^2}{2mB}$, so that $h \approx \frac{\sqrt{2mAB}}{\sin\varphi}$, where $\varphi$ is defined by

$$\varphi := \arccos\left(\frac{\mathbf{p}\cdot\mathbf{B}}{\|\mathbf{p}\|\|\mathbf{B}\|}\right), \tag{10.13}$$

i.e. it is the angle $\widehat{(\mathtt{b},\mathbf{p})} \in [0;\pi]$ between the particle momentum and the magnetic field, commonly called *pitch-angle*.

In a more fundamental way, this hypothesis comes from a symmetry in the defining equation for the normal form, which is left unchanged through the simultaneous change $h \to kh$ and $B \to kB$ for any $k$ independent of $\mathbf{q}$ and $\mathtt{v}$, hence for any $k = k(A)$. This allows for the special choice $k := A^{1/2}$, which means $\alpha_0 = 0$.

With formulas (10.12), Eqs. (10.10)-(10.11) become induction equations. Isolating the highest-order term into the left-hand side, we get

$$\mathtt{avg}\left(\bar{\mathtt{v}}\nabla(\alpha_N\eta_0)\eta_0\right) = -\mathtt{avg}\sum_{i=0}^{N-1}\sum_{j=0}^{N-i}\bar{\mathtt{v}}\nabla(\alpha_i\eta_j)\eta_{N-i-j}\,, \tag{10.14}$$

$$\eta_N = \delta_{N0} - B^{-1}\mathcal{G}\sum_{i=0}^{N}\sum_{j=0}^{N-i}\bar{\mathtt{v}}\nabla(\alpha_i\eta_j)\eta_{N-i-j}\,, \tag{10.15}$$

where $\delta_{N0}$ is the Kronecker delta symbol.

Now, we will be able to solve Eqs. (10.14)-(10.15) by induction after the initialization condition is verified : if $\alpha_m$ and $\eta_m$ are known up to $m = N-1$, Eq. (10.14) will give $\alpha_N$ and, next, Eq. (10.15) will give $\eta_N$.

### 10.1.2 The adiabatic invariant

In this subsection, we perform the initialization of the induction for the normal form. First, we obtain the equation for the lowest-order term. Next, we study the general solution of this equation. Lastly, we explore some interesting solutions, before choosing one of them as our lowest-order normal form.

Let us begin by identifying the initializing condition. From $\alpha_0 = 0$, Eq. (10.15) trivially gives $\eta_0 = 1$. So, the initializing condition relies on the lowest-order term of $h = \alpha_1 \eta_0 A^{1/2} = \alpha_1 A^{1/2}$, whose equation is given by (10.14) as

$$\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right) = 0\,. \tag{10.16}$$

This means that the lowest-order term can be chosen as any function in the intersection of the kernels of $\mathtt{avg}\,\bar{\mathsf{v}}\nabla$ and of $\mathcal{D}$ (because $\alpha_1$ has to be a gyro-averaged quantity). To characterize $\alpha_1$, we need to study the operator $\mathtt{avg}\,\bar{\mathsf{v}}\nabla$. This is done in App. B at the end of the chapter on page 246, where we show that Eq. (10.16) writes

$$\cos\varphi\ \partial_*\alpha_1 = 0\,, \tag{10.17}$$

$$\text{with}\quad \partial_* := \bar{\mathsf{b}}\nabla_\star + \frac{\overline{\nabla}\mathsf{b}}{2}\frac{1+\phi^2}{\phi}\partial_\phi\,, \tag{10.18}$$

where the operator $\nabla_\star$ was defined in Eq. (10.59), and the variable $\mathsf{v}$ is decomposed into two components. The component parallel to the magnetic field is the cotangent of the pitch-angle:

$$\phi := \frac{\bar{\mathsf{b}}\mathsf{v}}{\bar{\mathsf{c}}\mathsf{v}} = \cot\varphi \in \mathbb{R}\,.$$

The component perpendicular to $\mathbf{B}$ is the unit vector of the perpendicular momentum:

$$\mathsf{c} := \frac{(1-\mathsf{b}\bar{\mathsf{b}})\mathsf{v}}{\|(1-\mathsf{b}\bar{\mathsf{b}})\mathsf{v}\|} \in \mathbb{S}^1\,.$$

In this chapter, the components $\phi$ and $\mathsf{c}$ are considered as intermediate quantities for the computations, not as true coordinates. The corresponding coordinate remains $\mathsf{v}$, as indicated in Eq. (10.3).

The second step is to solve Eq. (10.17) for the lowest-order term $\alpha_1$ of the normal form. A formal constraint on the elements of the kernel of the operator $\partial_*$ is obtained by expanding $\alpha_1$ in $\phi$. Here, we are going to use a simpler but more qualitative method, which gives the same results. Separating the terms depending on $\phi$ from those depending on $\mathbf{q}$, Eq. (10.17) writes:

$$\frac{1+\phi^2}{2\phi}\partial_\phi\alpha_1 = -\frac{1}{\overline{\nabla}\mathsf{b}}\bar{\mathsf{b}}\nabla_\star\alpha_1\,. \tag{10.19}$$

In this equation, the operator $\frac{1+\phi^2}{2\phi}\partial_\phi$ must generate no $\phi$, but only a purely spatial operator $-\frac{1}{\overline{\nabla}\mathsf{b}}\bar{\mathsf{b}}\nabla_\star$. This suggests the following structure for Eq. (10.19)

$$\frac{f}{\partial_\phi f}\partial_\phi f^n = n f^n\,, \tag{10.20}$$

valid for any function $f(\phi)$, with the correspondence $f^n := \alpha_1$, $2n := -\frac{2}{\overline{\nabla}\mathsf{b}}\bar{\mathsf{b}}\nabla_\star$, and $f := 1+\phi^2$. Therefore, for any function $g(\mathbf{q})$ we can expect to have

$$\frac{1+\phi^2}{2\phi}\partial_\phi\left\{(1+\phi^2)^{-\frac{1}{\overline{\nabla}\mathsf{b}}\bar{\mathsf{b}}\nabla_\star}g\right\} = -\frac{1}{\overline{\nabla}\mathsf{b}}\bar{\mathsf{b}}\nabla_\star\left\{(1+\phi^2)^{-\frac{1}{\overline{\nabla}\mathsf{b}}\bar{\mathsf{b}}\nabla_\star}g\right\}\,. \tag{10.21}$$

Let us notice that, when acting on functions depending only on $\mathbf{q}$, the operators $\nabla_\star$ and $\nabla$ are the same.

In order to verify that Eq. (10.21) indeed holds, let us define the following differential operator acting on functions of $\mathbf{q}$:

$$\mathcal{T} := (1+\phi^2)^{-\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla_\star}\,.$$

Then, for any $g(\mathbf{q})$, the function $\mathcal{T}g$ is expected to be a formal solution of Eq. (10.19), i.e. $\mathcal{T}$ takes values in $\ker\partial_*$:

$$\forall g(\mathbf{q}), \qquad \partial_*\mathcal{T}g = 0\,. \tag{10.22}$$

To verify it with more details, we express $\mathcal{T}$ as the exponential of something $e^{-\ln(1+\phi^2).\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla_\star}$, which can be developed in series. The commutation of $\partial_\phi$ with $\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla_\star$ implies that we indeed have a relation analogous to Eq. (10.20), which in turn indeed provide the expected relation (10.19).

Thus, $\ker\partial_\star$ formally contains the functions $\mathcal{T}g(\mathbf{q})$, with $g(\mathbf{q})$ arbitrary, which is a considerable freedom for the choice of $\alpha_1$. It is related with the freedom in the choice of a second constant of motion: if $h$ and $A$ are independent constants of motion, then $F(h,A)$ is again a constant of motion, for any function $F$. Actually, the freedom contained in the function $g$ may seem too large, but it is going to be reduced by the requirement for $\mathcal{T}g$ to be meaningful and not only formal.

Now, the last step is to investigate the interesting solutions and to choose one of them for our lowest-order normal form. First, we are going to consider three particular families of solutions. Next, we will show that they are the only available choices in the absence of any symmetry in the magnetic field. Last, we will fix our choice for the lowest-order normal form.

First, the trivial choice $g(\mathbf{q}) := k$, with $k \in \mathbb{R}$ constant is not entirely useless to consider; in that case, $\alpha_1$ is

$$\alpha_1 = \mathcal{T}g(\mathbf{q}) = \sum_{i=0}^{\infty} \frac{\left(-\ln(1+\phi^2).\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla_\star\right)^i}{i!}k = k\,.$$

So, $h = \alpha\eta = (kA^{1/2}+...)(1+...) = kA^{1/2}+....$ This means choosing $A_1 := \frac{h^2}{k^2}$, where the index means that $A_1$ is the lowest-order constant of motion. It is functionally equivalent to choosing $A_1 := h$. This corresponds to adopting $A := h$ as the constant of motion, which is actually a possible choice, but not very attractive.

The second simplest choice is $g := B^n$ monomial in $B$, then:

$$\alpha_1 = \mathcal{T}g = \sum_{0}^{\infty} \frac{\left(-\ln(1+\phi^2).\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla_\star\right)^k}{k!}B^n = \sum_{0}^{\infty} \frac{\left(\ln(1+\phi^2).n\right)^k}{k!}B^n = e^{n\ln(1+\phi^2)}B^n = \left(\frac{B}{\sin^2\varphi}\right)^n,$$

where we used $\frac{\overline{\mathsf{b}}\nabla_\star B}{\overline{\nabla}\mathsf{b}} = -B$, which comes from $\overline{\nabla}\mathbf{B} = 0$. To lowest order

$$h \approx \alpha_1 A^{1/2} = \left(\frac{B}{\sin^2\varphi}\right)^n A^{1/2} \quad \text{i.e.} \quad A_1 = h^2\left(\frac{\sin^2\varphi}{B}\right)^{2n}.$$

And we recover the usual choice for th emagnetic moment $A_1 = \mu$, which corresponds to the case $n := \frac{1}{2}$ (within a multiplication factor $2m$).

Third, a more general choice is $g := f(B)$ for any function $f$. Then, using the second case, we get $\alpha_1 = f\left(\frac{B}{\sin^2\varphi}\right)$, and

$$A_1^{1/2} = \frac{h}{f\left(\frac{B}{\sin^2\varphi}\right)} = \frac{h}{f\left(\frac{h^2}{(2m\mu)^2}\right)},$$

which is a function of $h$ and $\mu$, as expected.

Other choices can be explored for the function $g$, but one can argue that they are of less interest. Indeed, the formal operator $\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla$ is meaningful only when the denominator is compensated by the

numerator. The only natural relation concerning the geometry comes from the magnetic constraint $\overline{\nabla}\mathbf{B} = 0$, which writes $\overline{\mathsf{b}}\nabla B = -B\overline{\nabla}\mathsf{b}$. So, $B$ is the quantity to be used in order to make sure that $\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla$ has a meaning.

As a result, the usual adiabatic invariant $\mu$ for the motion of a charged particle in a strong magnetic field results from the presence of the operators $\mathcal{D}$ and $\overline{\mathsf{v}}\nabla$ in the defining equation (10.6) for the normal form of this Hamiltonian system, which implies for the lowest-order term to be in the kernels of both $\mathcal{D}$ and $\mathsf{avg}\,\overline{\mathsf{v}}\nabla$; with this requirement (together with a functional independence of $h$), all the choices for the lowest order are functionnally equivalent, and one of the simplest is $\mu$, which has the physical meaning of the (lowest-order) magnetic moment.

Two remarks must be added here. First, we have not yet proven that $A_1$, the lowest-order constant of motion, is an adiabatic invariant, because our formal series $h(A)$ might not be a perturbative series, in which the lowest-order term dominates and the following ones are only corrections (see e.g. [74]). The answer about this point will appear later in our analysis: the ratio $\frac{h_i}{h_{i-1}}$ is small under the condition of "strong magnetic field", i.e. when the variations of $\mathbf{B}$ are small at the scale of the Larmor radius. Under this assumption, the first term of the series dominates, and the usual lowest-order magnetic moment is an adiabatic invariant, which, in turn causes the relation between $h$ and $A$ to be bijective, since the lowest-order term dominates and is bijective.

Second, we remain with a large choice for $\alpha_1$. Nevertheless, if we want $g$ to be a monomial in $B$ and to have the dimension of a magnetic moment, then we must make exactly the physical choice $A_1 := C\mu = C\frac{(h\sin\varphi)^2}{2mB}$, with $C$ a free coefficent, i.e. $h = \sqrt{\frac{2mAB}{C\sin^2\varphi}}$ at first order. We adopt, from now on, this physical choice, with $C = 2m$ for simplicity. So, the lowest-order term for the normal form is defined as

$$\alpha_1 := \frac{\sqrt{B}}{\sin\varphi}\,. \tag{10.23}$$

### 10.1.3 The normal form to second order

The induction has been initialized, we can compute the normal form in an iterative way. In this subsection, we compute it to second order in $A$. This is done in three steps. First, we factorize the $\sin\varphi$ coming from the choice we have made for $\alpha_1$. Second, we compute $\eta_1$, the first correction to the fluctuation of the normal form. Third, we turn to the first correction to the average part of the normal form, $\alpha_2$. We will then have the normal form $h(A, \mathsf{v}, \mathbf{q})$ to second order

$$h = \alpha\eta = (\alpha_1 + \alpha_2 ...)(\eta_0 + \eta_1 + ...) = \alpha_1 + \alpha_2 + \alpha_1\eta_1 + O(A^{3/2})\,. \tag{10.24}$$

First of all, to simplify computations, it is useful to factorize the $\sin\varphi$ coming from $\alpha_1 := \frac{\sqrt{B}}{\sin\varphi}$. In this way, all the $\sin\varphi$ and $\cos\varphi$ will disappear from the equations, and our expressions will be purely polynomial in $\phi$. So, let us define $a_i$ so that

$$\alpha_i = \frac{a_i}{\sin\varphi}\,.$$

For $\alpha_1$ given by Eq. (10.23), the initializing condition is $a_1 = \sqrt{B}$. As for the induction relations, replacing $\alpha_i$ in favor of $\frac{a_i}{\sin\varphi}$, the right-hand side of Eq. (10.14) becomes

$$\mathsf{avg}\,\overline{\mathsf{v}}\nabla\left(\frac{a_N}{\sin\varphi}\right) = \phi\left(\overline{\mathsf{b}}\nabla_\star + \frac{\overline{\nabla}\mathsf{b}}{2}(1 + \frac{1+\phi^2}{\phi}\partial_\phi) - \overline{\mathsf{b}}(\nabla\overline{\mathsf{b}})\mathsf{c}\;\overline{\mathsf{b}}\partial_{\overline{\mathsf{c}}}\right)a_N =: \phi\partial_{\mathsf{a}}a_N\,, \tag{10.25}$$

in which the operator $\partial_{\mathsf{a}}$ plays for $a_i$ the same role as did $\partial_*$ for $\alpha_i$. Here we have kept the operator with its complete form (10.66), rather than its simplified form (10.67), because it will be more convenient to work with formulae as they come out, i.e. containing some $\mathsf{a}$ and $\mathsf{c}$, rather than to reexpress them in such a way as to remove the $\mathsf{a}$ and $\mathsf{c}$. Its kernel is simply the set of functions that write $\sin\varphi\,\mathcal{T}g(\mathbf{q}) = (1+\phi^2)^{\frac{-1}{2}-\frac{1}{\overline{\nabla}\mathsf{b}}\overline{\mathsf{b}}\nabla}g(\mathbf{q})$.

Let us turn to the right-hand side of Eqs. (10.14) and (10.15). Their main term is

$$\bar{\mathsf{v}}\nabla(\alpha_i\eta_j)\eta_k = \bar{\mathsf{v}}\nabla\left(\frac{a_i\eta_j}{\sin\varphi}\right)\eta_k = \bar{\mathsf{v}}\nabla(a_i\eta_j)\eta_k + (a_i\eta_j)\eta_k\phi\bar{\mathsf{v}}(\nabla\bar{\mathsf{b}})\mathsf{c}\,, \tag{10.26}$$

where we have defined the vector

$$\mathsf{v} := \tfrac{\mathsf{v}}{\sin\varphi} = \phi\mathsf{b} + \mathsf{c}\,.$$

In addition, in order to avoid extensive use of parentheses, we adopt the convention that $\nabla$ acts only on the following term, i.e. for any functions $a$, $b$ and $c$ we have $\nabla abc = (\nabla a)bc$.

Combining Eqs. (10.25)-(10.26), we get the induction relations for $a_N$ and $\eta_N$

$$\phi\partial_{\mathrm{a}}a_N = -\mathsf{avg}\sum_{i=1}^{N-1}\sum_{j=0}^{N-i}\left\{\bar{\mathsf{v}}\nabla(a_i\eta_j)\eta_{N-i-j} + (a_i\eta_j)\eta_{N-i-j}\phi\bar{\mathsf{v}}\nabla\bar{\mathsf{b}}\mathsf{c}\right\}\,, \tag{10.27}$$

$$\eta_N = \delta_{N0} - B^{-1}\mathcal{G}\sum_{i=1}^{N}\sum_{j=0}^{N-i}\left\{\bar{\mathsf{v}}\nabla(a_i\eta_j)\eta_{N-i-j} + (a_i\eta_j)\eta_{N-i-j}\phi\bar{\mathsf{v}}\nabla\bar{\mathsf{b}}\mathsf{c}\right\}\,. \tag{10.28}$$

Eqs. (10.25)-(10.28) show that the $\cos\varphi$ and $\sin\varphi$ have indeed been replaced by polynomials in $\phi$, as announced.

Now, the second step is to compute $\eta_1$, the first fluctuating correction to the lowest-order normal form. Indeed, by hypotheses $\alpha_0 := 0$ and $\eta_0 = 1$. We have chosen $a_1 := \sqrt{B}$. It remains to compute $\eta_1$ and $\alpha_2$.

For $\eta_1$, the computation is straightforward, using Eqs. (10.56)-(10.57) from Appendix A,

$$\begin{aligned}\eta_1 &= -B^{-1}\mathcal{G}\left\{\eta_0(a_1\eta_0)'\mathsf{v} + (a_1\eta_0)\eta_0\phi\bar{\mathsf{c}}\mathsf{b}'\mathsf{v}\right\}\\ &= -B^{-1/2}\left\{\frac{1}{2B}B'\mathcal{G}(\phi\mathsf{b}+\mathsf{c}) + \phi\mathcal{G}\bar{\mathsf{c}}\mathsf{b}'\mathsf{c} + \phi^2\mathcal{G}\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\right\} = -B^{-1/2}\left\{\frac{B'\mathsf{a}}{2B} + \phi\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} + \phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\right\}\,,\end{aligned} \tag{10.29}$$

where we remind that $\mathsf{a} := \mathsf{b}\times\mathsf{c}$. For convenience in the computations to come, we use the primed notation for the spatial derivative acting on its left and contracted to its right: for any vector $\mathbf{w}$,

$$f'\mathbf{w} := \bar{\mathbf{w}}\nabla f\,. \tag{10.30}$$

Now, the third step is to computed $\alpha_2$, the average first correction to the lowest-order normal form. It is computed from Eq. (10.27) as

$$\phi\partial_{\mathrm{a}}a_2 = -\mathsf{avg}\left\{2\eta_1\sqrt{B}'\mathsf{v} + \sqrt{B}(\eta_1)'\mathsf{v} + 2(\sqrt{B}\eta_1)\phi\bar{\mathsf{c}}\mathsf{b}'\mathsf{v}\right\}\,. \tag{10.31}$$

The right-hand side is computed in App. C at the end of the chapter on page 248, and it writes

$$-\frac{B'\mathsf{b}}{2B}\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}''\mathsf{bc} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{ba}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a}}{2}\right)\,.$$

To solve Eq. (10.31), we expand $\phi\partial_{\mathrm{a}}a_2$ in $\phi$

$$\begin{aligned}&-\frac{B'\mathsf{b}}{2B}\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}''\mathsf{bc} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{ba}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a}}{2}\right)\\ &\quad= \sum_{k=0}^{\infty}\phi\left(\bar{\mathsf{b}}\nabla_\star + \frac{\overline{\nabla}\mathsf{b}}{2}(1 + \frac{1+\phi^2}{\phi}\partial_\phi) - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\,\bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\right)a_{2,k}\phi^k\\ &\quad= \sum_{j=-1}^{\infty}\phi^j\left\{\tfrac{\overline{\nabla}\mathsf{b}}{2}(j+1)a_{2,j+1}\chi_{j+1\in\mathbb{N}} + \left[\tfrac{\overline{\nabla}\mathsf{b}}{2}j + \bar{\mathsf{b}}\nabla_\star - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\,\bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\right]a_{2,j-1}\chi_{j-1\in\mathbb{N}}\right\}\,,\end{aligned} \tag{10.32}$$

where $\chi$ is the characteristic function. It is again an induction equation, and it can be solved directly.

The initializing condition for $a_{2,0}$ concerns $j = -1$. It is obviously satisfied for any $a_{2,0}$. We make the simplest choice $a_{2,0} = 0$, which induces $a_{2,k}$ to vanish for all even orders in $k$. This is because in Eq. (10.32), the left-hand side contains only even orders in $\phi^k$, and the right-hand side contains $a_{2,j+1}$ and $a_{2,j-1}$, which have inverted parity compared to $j$. Thus

$$a_{2,k} := 0 \text{ for any even } k. \tag{10.33}$$

The following order $j = 0$ gives the initializing condition for the odd orders in $k$:

$$-\tfrac{B'\mathsf{b}}{2B}\tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2}\phi^0 = \phi^0\left\{\tfrac{\overline{\nabla}\mathsf{b}}{2}(1)a_{2,1} + \left[\tfrac{\overline{\nabla}\mathsf{b}}{2}0 + \bar{\mathsf{b}}\nabla_\star - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\,\bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\right]a_{2,0-1}\right\} = \phi^0\tfrac{\overline{\nabla}\mathsf{b}}{2}a_{2,1}\,, \tag{10.34}$$

where we used $a_{2,-1} = 0$, which comes because $a_2$ must not be singular in $\phi = 0$, since this condition corresponds to the so-called bounce points in the particle trajectory. Eq. (10.34) can be simplified because of Eq. (10.53), which implies

$$a_{2,1} = \tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2}\,. \tag{10.35}$$

For the next order $j = 2$, Eq. (10.32) writes

$$\begin{aligned}
&\phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}''\mathsf{b}\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{b}\mathsf{a}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a}}{2}\right) = \phi^2\left\{\tfrac{\overline{\nabla}\mathsf{b}}{2}(3)a_{2,3} + \left[\tfrac{\overline{\nabla}\mathsf{b}}{2}2 + \bar{\mathsf{b}}\nabla_\star - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\,\bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\right]a_{2,1}\right\}\,,\\
&\text{wence}\qquad \tfrac{\overline{\nabla}\mathsf{b}}{2}3a_{2,3} = -\left[\overline{\nabla}\mathsf{b} + \bar{\mathsf{b}}\nabla_\star - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\,\bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\right]a_{2,1} + \frac{\bar{\mathsf{a}}\mathsf{b}''\mathsf{b}\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{b}\mathsf{a}}{2} + \frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a}}{2}\,.
\end{aligned} \tag{10.36}$$

Evaluating each term, we get

$$\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\,\bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}a_{2,1} = -\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\tfrac{\bar{\mathsf{b}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}}{2} = \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}}{2}\,, \tag{10.37}$$

$$-\bar{\mathsf{b}}\nabla_\star a_{2,1} = +\tfrac{\bar{\mathsf{c}}\mathsf{b}''\mathsf{b}\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}''\mathsf{b}\mathsf{c}}{2} - \tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}}{2}\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\,. \tag{10.38}$$

The two terms from Eq. (10.38) cancel exactly both the term from (10.37) and the third term in the right-hand side of Eq. (10.36). Finally, Eq. (10.36) writes

$$\begin{aligned}
&\tfrac{\overline{\nabla}\mathsf{b}}{2}3a_{2,3} = -\overline{\nabla}\mathsf{b}.a_{2,1} + \frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a}}{2}\\
&\Rightarrow a_{2,3} = 0\,,
\end{aligned} \tag{10.39}$$

where the last equality follows from

$$\begin{aligned}
\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a} &= \bar{\mathsf{a}}\mathsf{b}'(\mathsf{b}\bar{\mathsf{b}} + \mathsf{c}\bar{\mathsf{c}} + \mathsf{a}\bar{\mathsf{a}})\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'(\mathsf{b}\bar{\mathsf{b}} + \mathsf{c}\bar{\mathsf{c}} + \mathsf{a}\bar{\mathsf{a}})\mathsf{b}'\mathsf{a}\\
&= (\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a})(\bar{\mathsf{c}}\mathsf{b}'\mathsf{c} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{a}) = \overline{\nabla}\mathsf{b}.2a_{2,1}\,.
\end{aligned}$$

This puts an end to the induction for $a_2$: For all the higher odd orders, we get $a_{2,k} = 0$ because the left-hand side of Eq. (10.32) is zero and the right-hand side depends only on $a_{2,k-2}$, which is zero, by induction. Together, Eqs. (10.33), (10.35), and (10.39) show that the solution for $a_2$ writes exactly

$$a_2 = -\phi\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2} = \phi\tfrac{\bar{\mathsf{b}}.\nabla\times\mathsf{b}}{2}\,. \tag{10.40}$$

As expected, $a_2$ does not depend on the fast variables $\mathsf{c}$, $\mathsf{a}$ and can be written without using them.

Now, Eqs. (10.24), (10.29), and (10.32) give the desired result, the normal form up to second order in $A$:

$$\begin{aligned} h = \alpha\eta &= \alpha_1 A^{1/2} + (\alpha_1\eta_1 + \alpha_2)A^{2/2} + O(A^{3/2}) \\ &= \frac{1}{\sin\varphi}\left\{\sqrt{B}A^{1/2} - \left[\frac{B'\mathsf{a}}{2B} + \phi\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} + \phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b} + \phi\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\right]A^{2/2}\right\} + O(A^{3/2})\,. \end{aligned} \tag{10.41}$$

Let us verify the consistency with the literature, which was more interested in the magnetic moment than in the normal form. But Eq. (10.41) implies, as a corollary, an expression for the magnetic moment by inverting the series:

$$A = \frac{(h\sin\varphi)^2}{B} + 2\frac{(h\sin\varphi)^3}{B^2}\left[\frac{B'\mathsf{a}}{2B} + \phi\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} + \phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b} + \phi\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\right] + O(h^4)\,. \tag{10.42}$$

This expression is written with the convenient variables to solve the equation for the normal form. But the literature rather uses other variables, related with physical quantities (restoring physical units):

| | | | |
|---|---|---|---|
| $h\sin\varphi$ | is expressed with | $\mu := \frac{\mathbf{p}_\perp^2}{2mB} = \frac{h^2\sin^2\varphi}{2mB}$, | the lowest-order magnetic moment; |
| $\phi$ | is expressed with | $p_\parallel := h\cos\varphi = \phi h\sin\varphi$, | the parallel momentum; |
| $\mathsf{a}$ | is expressed with | $\mathbf{r_L} := \frac{\lVert\mathbf{p}_\perp\rVert}{eB}\mathsf{a} = \frac{h\sin\varphi}{eB}\mathsf{a}$, | the Larmor radius; |
| $\mathsf{b}$ | is expressed with | $\hat{\mathbf{b}} := \frac{e\mathsf{b}}{\lvert e\rvert}$, | the unit vector of $\mathbf{B}$; |
| $\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}$ | is expressed with | $\tau := \bar{\mathsf{b}}\nabla\times\mathsf{b}$, | the torsion coefficient of $\mathsf{b}$; |
| $\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}$ | is expressed with | $\mathbf{a}_1 := \mathsf{ac} + \mathsf{ca}$, | a dyad often used in gyrokinetics. |

With these quantities, Eq. (10.42) becomes

$$A = m\mu\left[2 + \frac{p_\parallel}{eB}\left(\mathbf{a}_1 : \nabla\hat{\mathbf{b}} - 2\tau\right)\right] + 2\mathbf{r_L}\cdot\left[\mu\nabla\ln B + \frac{p_\parallel^2}{mB}\hat{\mathbf{b}}\cdot\nabla\hat{\mathbf{b}}\right] + \ldots \tag{10.43}$$

in which the double dot ":" denotes the double contraction of the dyad $\mathbf{a}_1$. Eq. (10.43) agrees with the previous results in the literature [29, 86].

The variables used in Eq. (10.43) are more expressive to identify the physical meaning of the terms involved in the magnetic moment, because they are linked to various drifts of the guiding-center dynamics. On the contrary, our coordinates $\phi$, $\mathsf{c}$ were naturally provided by the derivation, and have proven to be especially suited to the action of the operators involved in the process. As a consequence, they make the derivation easier to compute, and make more obvious structures in the results, e.g. the polynomiality in $B$ and $\phi$, and the systematic dependence in $h$ through $h\sin\varphi$. This will be developed in the next subsection.

### 10.1.4 The normal form to higher order

In this subsection, we study the structure of the general term of the normal form. It will exhibit the small parameter of the normal form's expansion, and it will emphasize the symmetry of the equation for the normal form linked to the existence of the constant of motion $A$.

To compute the third-order term of the normal form, one should iterate the method. However, it is outside the scope of the present document and will be investigated later. In most cases, the explicit expressions of the terms are not essential. What is more important is the structure of the formulae involved: the orders in $B$, the number of fast variables $\mathsf{c}$ and $\mathsf{a}$, etc.

The structure of the induction relations shows that $a_N$ and $\eta_N$ are monomials in $B$ and $\nabla$, with

$$a_N \propto B^{1-\frac{N}{2}}\nabla^{N-1} \quad \text{and} \quad \eta_N \propto B^{-\frac{N}{2}}\nabla^N\,,$$

where products are understood in a tensorial sense: for instance $B^2\nabla^2$ may represent $\nabla B\nabla B$, or $B^2\nabla^2 B/B$, or even $B^2\nabla^2 f$ with $f$ some function involving no $B$ and no $\nabla$.

Furthermore, the operators involved in the computations are polynomials in the variables we have chosen. So, the resulting expressions will be polynomials in $\phi$, $\mathsf{c}$ and $\mathsf{b}$, and the orders of the polynomials will be linear in $N$. A rough estimate of the upper bound leads to the following structure for $h_N$

$$h_N = \tfrac{1}{\sin\varphi}\sum_{n=1}^{N} A^{N/2} a_n \eta_{N-n} = \tfrac{1}{\sin\varphi}\sum_{i=0}^{2(N-1)}\sum_{j=0}^{3(N-1)}\sum_{k=0}^{2(N-1)} A^{N/2} C_{Nijk}\phi^i \mathsf{b}^j \mathsf{u}^k B^{1-\frac{N}{2}}\nabla^{N-1}\,, \tag{10.44}$$

with $C_{Nijk}$ a pure (tensorial) coefficient depending on none of the variables $B$, $\mathsf{b}$, $\mathsf{u}$, $\phi$, and $A$. In these tensorial expressions, $\mathsf{u}$ stands for the fast variable, hence for either $\mathsf{c}$ or $\mathsf{a}$; this is because these variables are mixed by $\mathcal{G}$ and $\mathsf{avg}$ and it is uneasy to follow each of them. Of course, many terms are zero in the sum, as can already be seen in $h_1$; for instance, there is a symmetry in the defining equation which guarantees that $C_{Nijk}$ is zero when $N+i+k$ is even.

From the general term, we can deduce some properties of the normal form. First, it has the physical dimension $B^1\nabla^1$ of energy, as expected. More interesting, the dimensionless increment term is just $\left(\frac{A}{B}\right)^{\frac{N}{2}}\nabla^N \sim r_L^N\nabla^N$, where $r_L := \sqrt{\frac{A}{B}}$ is the Larmor radius, and $\nabla$ acts only on $B$ and $\mathsf{b}$, because they contain all the geometry of the configuration space. So, the increment term in the series for the normal form corresponds to the variations of the magnetic field $\mathsf{b}$ and $B$ at the scale of the Larmor radius. It corresponds to the usual parameter of gyrokinetics $\varepsilon_L = r_L/L_B$, defined as the ratio of the Larmor radius over the characteristic length scale for the non-uniformity of $\mathbf{B}$. This explains why our result to second order in $A$ corresponds to an expansion at first order in the Larmor radius.

At first order, it implies the small parameter $\varepsilon_B := r_L\nabla\ln B$ for the norm of $\mathbf{B}$, but also terms with $\nabla\mathsf{b}$ (such as the torsion coefficient of $\mathsf{b}$, for instance), and the related parameter is $\varepsilon_{\mathsf{b}} := r_L\|\nabla\mathsf{b}\|$. Of course, these terms are small when the direction and the norm of the magnetic field change little at the scale of the Larmor radius $r_L$.

For higher-order terms, things are a little bit more subtle, the "small parameter" involves $n$th-order derivatives, which are not exactly equivalent to $\varepsilon_B^n$ or $\varepsilon_{\mathsf{b}}^n$.

Another interesting property of Eq. (10.44) is that the constant of motion $A$ only makes a scaling on both $h$ and $B$. It may be absorbed in these quantity through the scaling $\frac{1}{B}\to\frac{A}{B}=r_L^2$ and $h\to\frac{h}{A}$. This is related to the symmetry of the defining equation for the normal form (10.6): $h\to kh$ and $B\to kB$ for any $k$ independent of $\mathbf{q}$ and $\mathsf{v}$, hence for any $k=k(A)$, which allows for the special choice $k:=1/A$.

So, from one particular solution in strong field $h(1/B)$, this symmetry in the equation generates the family of solutions $h(A/B)/A$. The result (10.44) shows that this symmetry is enough to get all the possible normal forms. This agrees with the existence of only one independent constant of the motion besides $h$.

Last, the role of $A$ in $\frac{1}{B}\to\frac{A}{B}$ shows that $A$ acts as a scaling of the magnetic field exactly as a small parameter in an expansion in strong magnetic field, it is why the normal form as a formal series in $A$ was to be a perturbative series in strong $\mathbf{B}$.

## 10.2 The reduced Poisson bracket

In the previous section, we computed the Hamiltonian $h(\mathbf{q},\mathsf{v},A)$ for the reduced dynamics of a charged particle in a strong magnetic field. Now, we turn to the corresponding reduced Poisson bracket, which induces the dynamics together with the Hamiltonian. This will be performed in three steps. First, we establish a lemma we need to go back and forth between our Poisson brackets and our symplectic 2-forms. Second, we use the lemma to get the symplectic 2-form for the initial dynamics and we compute the 2-form for the reduced dynamics. Last, using the lemma again,

we invert the 2-form to get the Poisson bracket for the reduced dynamics. This strategy is used because it is simpler to change coordinates for the 2-form than for the Poisson bracket (a similar method is used in [86]). The first two steps are done in the first subsection, whereas the last step is done in the second subsection.

### 10.2.1 Change of variables for the symplectic 2-form

Let us begin by proving the following lemma to connect our Poisson brackets and our symplectic 2-forms:

*If $M$ is an antisymmetric endomorphism in $\mathbb{R}^N$ and, on a $2N$-dimensional manifold, $\mathbf{D}$ and $\partial$ are two $N$-dimensional vectors of vectors fields, and $\mathbf{P}$ and $\mathbf{Q}$ two $N$-dimensional vectors of linear-form fields, and if we have the following orthogonality-completeness relations:*

$$\begin{aligned} \mathbf{P}\cdot\overline{\mathbf{D}} &= \mathrm{Id}_{\mathbb{R}^N} \\ \mathbf{Q}\cdot\overline{\mathbf{D}} &= 0 \\ \mathbf{P}\cdot\overline{\partial} &= 0 \\ \mathbf{Q}\cdot\overline{\partial} &= \mathrm{Id}_{\mathbb{R}^N} \\ \overline{\mathbf{D}}.\mathbf{P}+\overline{\partial}.\mathbf{Q} &= \mathrm{Id} \end{aligned} \tag{10.45}$$

*then the 2-form $\sigma := \overline{\mathbf{P}}.\mathbf{Q}-\overline{\mathbf{Q}}.\mathbf{P}-\overline{\mathbf{Q}}.M\mathbf{Q}$*
*and the bivector $\pi := \overline{\mathbf{D}}.\partial-\overline{\partial}.\mathbf{D}+\overline{\mathbf{D}}.M\mathbf{D}$ are inverse of each other.*

To illustrate our definitions, let us notice that $\mathbf{D}$ and $\partial$ play the role of $\partial_{\overline{\mathbf{p}}}$ and $\partial_{\overline{\mathbf{q}}}$, whereas $\mathbf{P}$ and $\mathbf{Q}$ play the role of $d\mathbf{p}$ and $d\mathbf{q}$, but here we do not assume our forms to be exact. On another hand, in the completeness relation $\overline{\mathbf{D}}.\mathbf{P}+\overline{\partial}.\mathbf{Q}=\mathrm{Id}$, the operator Id is the identity on both the space of vector fields and the space of linear form fields, in a similar way as in the relation $\partial_{\mathbf{p}}\otimes d\mathbf{p}+\partial_{\mathbf{q}}\otimes d\mathbf{q}=\mathrm{Id}$.

*Proof of the lemma:* The proof just consists in expanding the contraction $\sigma_{\alpha B}\pi^{B\gamma}$ and using Eqs. (10.45). A first simplification occurs:

$$\begin{aligned} \sigma\cdot\pi &= \big(\overline{\mathbf{P}}.\mathbf{Q}-\overline{\mathbf{Q}}.\mathbf{P}-\overline{\mathbf{Q}}.M\mathbf{Q}\big)\cdot\big(\overline{\mathbf{D}}.\partial-\overline{\partial}.\mathbf{D}+\overline{\mathbf{D}}.M\mathbf{D}\big) \\ &= \overline{\mathbf{P}}.(\mathbf{Q}\cdot\overline{\partial}).\mathbf{D}+\overline{\mathbf{Q}}.(\mathbf{P}\cdot\overline{\mathbf{D}}).\partial+\overline{\mathbf{Q}}.(\mathbf{P}\cdot\overline{\mathbf{D}}).M\mathbf{D}-\overline{\mathbf{Q}}.M(\mathbf{Q}\cdot\overline{\partial}).\mathbf{D}\,, \end{aligned}$$

because $\mathbf{Q}\cdot\overline{\mathbf{D}}=0$ and $\mathbf{P}\cdot\overline{\partial}=0$. To avoid ambiguities, we write the overbars for the contraction between the $N$-dimensional vectors, whereas we write $\cdot$ for the contraction between a form field and a vector field on the $2N$-dimensional manifold.

Last, using $\mathbf{P}\cdot\overline{\mathbf{D}}=\mathrm{Id}_{\mathbb{R}^N}$ and $\mathbf{Q}\cdot\overline{\partial}=\mathrm{Id}_{\mathbb{R}^N}$ for a first step, and then $\overline{\mathbf{D}}.\mathbf{P}+\overline{\partial}.\mathbf{Q}=\mathrm{Id}_{\mathbb{R}^{2N}}$ for a second step, we get

$$\sigma\cdot\pi=\overline{\mathbf{P}}.\mathbf{D}+\overline{\mathbf{Q}}.\partial+\overline{\mathbf{Q}}.M\mathbf{D}-\overline{\mathbf{Q}}.M\mathbf{D}=\overline{\mathbf{D}}.\mathbf{P}+\overline{\partial}.\mathbf{Q}=\mathrm{Id}\,,$$

which completes the proof of the lemma. The antisymmetry of $M$ is not used in the proof, it is required only for the 2-form to be antisymmetric.□

Now, we can apply the lemma to obtain the 2-form "inverse" to our bivector (10.2), since it is just of the form concerned by the lemma, with $N=3$ and

$$\begin{aligned} \mathbf{D} &:= \partial_{\overline{\mathbf{p}}}|_{\mathbf{q}} & \qquad \mathbf{P} &:= d\mathbf{p} \\ \partial &:= \partial_{\overline{\mathbf{q}}}|_{\mathbf{p}} & \qquad \mathbf{Q} &:= d\mathbf{q} \\ M &:= e\mathbf{B}\times \end{aligned}$$

Thus, in the original variables $(\mathbf{p},\mathbf{q})$, the inverse of the Poisson bi-vector is

$$\sigma := d\overline{\mathbf{p}}.d\mathbf{q}-d\overline{\mathbf{q}}.d\mathbf{p}-d\overline{\mathbf{q}}.e\mathbf{B}\times d\mathbf{q}\,.$$

It is called the Lagrange 2-form. Let us remind incidentally that the Jacobi identity for the Poisson bracket is directly related to the closure of the Lagrange 2-form $\sigma$, which writes $0 = d\sigma = -e\ d\overline{\mathbf{q}}.d\mathbf{B} \times d\mathbf{q} = -e \text{ div } \mathbf{B}\ d\mathbf{q} \wedge d\mathbf{q} \wedge d\mathbf{q}$. This emphasizes that the Jacobi identity is equivalent to div $\mathbf{B} = 0$.

Next, it is straightforward to change this 2-form through the change of variables (10.3), which we remind for the reader's convenience

$$(\mathbf{p}, \mathbf{q}) \longrightarrow (h, \mathsf{v}, \mathbf{q}) \longrightarrow (A, \mathsf{v}, \mathbf{q})\,.$$

The vector $\mathsf{v} = \frac{\mathbf{p}}{\|\mathbf{p}\|}$ is the unit vector of the momentum, while the norm of the momentum is given by $h = \|\mathbf{p}\| = h(\mathbf{q}, \mathsf{v}, A)$, where $h(\mathbf{q}, \mathsf{v}, A)$ is any function in bijection with $A$. It is important to point out that in this section, we will not use the fact that $A$ is a constant of motion.

Our task is simplified because only $\mathbf{p}$ is concerned by the change

$$\begin{aligned} d\mathbf{p} &= \frac{\partial h\mathsf{v}}{\partial A} dA + \frac{\partial (h\mathsf{v})}{\partial \mathsf{v}} d\mathsf{v} + \frac{\partial h\mathsf{v}}{\partial \mathbf{q}} d\mathbf{q} = \mathbf{P}_\circ + \mathsf{v}\overline{\nabla} h.d\mathbf{q}\,, \\ &\text{with } \mathbf{P}_\circ := \mathsf{v}\partial_A h dA + (h + \mathsf{v}\partial_{\mathsf{v}} h) d\mathsf{v}\,. \end{aligned} \tag{10.46}$$

Let us insert this result[2] in $\sigma$

$$\sigma := \overline{\mathbf{P}}_\circ \wedge d\mathbf{q} + \overline{(\mathsf{v}\overline{\nabla} h.d\mathbf{q})} \wedge d\mathbf{q} - d\overline{\mathbf{q}}.e\mathbf{B} \times d\mathbf{q}\,. \tag{10.47}$$

$$\begin{aligned} &\text{Now, } \overline{(\mathsf{v}\overline{\nabla} h.d\mathbf{q})} \wedge d\mathbf{q} = \overline{(\mathsf{v}\overline{\nabla} h.d\mathbf{q})}.d\mathbf{q} - d\overline{\mathbf{q}}.(\mathsf{v}\overline{\nabla} h.d\mathbf{q}) = d\overline{\mathbf{q}}.M_\circ d\mathbf{q}\,, \\ &\text{where } M_\circ := \nabla h\overline{\mathsf{v}} - \mathsf{v}\overline{\nabla} h = (\mathsf{v} \times \nabla h)\times \text{ is an antisymmetric matrix.} \end{aligned}$$

So, we can add this term to the magnetic part and define $e\mathbf{B}_\circ := e\mathbf{B} - \mathsf{v} \times \nabla h$. Finally, for the reduced dynamics, the symplectic 2-form writes

$$\sigma := \overline{\mathbf{P}}_\circ \wedge d\mathbf{q} - d\overline{\mathbf{q}}.e\mathbf{B}_\circ \times d\mathbf{q}\,.$$

### 10.2.2 Change of variables for the Poisson bracket

Transforming the 2-form was the first step. Now, we can get the transformed Poisson bracket by inversion of the symplectic 2-form, by using the lemma again, this time in order to go from the 2-form to the bivector. We first have to find $\mathbf{D}$ and $\partial$ verifying the Eqs. (10.45) with respect to the forms $\mathbf{P}_\circ$ and $d\mathbf{q}$. Next, we will apply the lemma to reach the Poisson bracket. In a last stage, the bracket will be rewritten in a more convenient way.

Let us begin by identifying $\mathbf{D}$ and $\partial$. Since $d\mathbf{q}$ remains unchanged in the new variables and does not enter in $\mathbf{P}_\circ$, then $\partial := \partial_{\mathbf{q}}|_{\mathsf{v},A}$ is "unchanged"[3]. It only remains to identify the vector field corresponding for this particular case to the generic operator $\mathbf{D}$ of the lemma; we will write it $\mathbf{D}_\circ$. We expect to have

$$\mathbf{D}_\circ := \mathbf{w}\partial_A + N\partial_{\overline{\mathsf{v}}} \quad \text{where } \mathbf{w} \text{ is a vector and } N \text{ is a matrix.} \tag{10.48}$$

The matrix $N$ is simply

$$N = \partial_{\overline{\mathbf{p}}}\overline{\mathsf{v}} = \partial_{\overline{\mathbf{p}}}\frac{\overline{\mathbf{p}}}{h} = \frac{1}{h} - \partial_{\overline{\mathbf{p}}} h \frac{\overline{\mathbf{p}}}{h^2} = \frac{1 - \mathsf{v}.\overline{\mathsf{v}}}{h} = -\frac{\mathcal{W}^2}{h}\,,$$

in which we used $\mathcal{W}^2 := \mathsf{v} \times (\mathsf{v} \times \cdot) = \mathsf{v}\overline{\mathsf{v}} - \overline{\mathsf{v}}\mathsf{v} = \mathsf{v}\overline{\mathsf{v}} - 1$, according to the vectorial triple product formula $a \times (b \times c) = (\overline{a}c)b - (\overline{a}b)c = (b\overline{a} - \overline{a}b)c$.

[2] We use the notation $\overline{\mathbf{u}} \wedge \mathsf{v}$ for $\overline{\mathbf{u}}.\mathsf{v} - \overline{\mathsf{v}}.\mathbf{u}$.

[3] To be more precise, $\partial_{\mathbf{q}}$ is slightly changed: initially we had $\partial_{\mathbf{q}}|_{\mathsf{v},h}$, whereas finally we have $\partial_{\mathbf{q}}|_{\mathsf{v},A}$. To avoid confusion, the symbol $\nabla$ is used only for $\partial_{\mathbf{q}}|_{\mathsf{v},A}$.

To identify $\mathbf{w}$, the equation to be solved is

$$\begin{aligned}
\mathrm{Id} = \mathbf{P}_\circ \cdot \overline{\mathbf{D}}_\circ &= (\mathsf{v}\partial_A h).\bar{\mathbf{w}} + (h + \mathsf{v}\partial_\mathsf{v} h).\frac{1-\mathsf{v}\bar{\mathsf{v}}}{h}\,. \\
\text{Thus } (\mathsf{v}\partial_A h).\bar{\mathbf{w}} &= 1 - (h + \mathsf{v}\partial_\mathsf{v} h).\frac{1-\mathsf{v}\bar{\mathsf{v}}}{h} = \mathsf{v}\left\{\bar{\mathsf{v}} - \frac{1}{h}\partial_\mathsf{v} h(1-\mathsf{v}\bar{\mathsf{v}})\right\}\,, \\
\text{so that } \mathbf{w} := &\frac{\mathsf{v} - \frac{1}{h}(1-\mathsf{v}\bar{\mathsf{v}})\partial_{\bar{\mathsf{v}}} h}{\partial_A h} = \frac{\mathsf{v} + \frac{1}{h}\mathcal{W}^2\partial_{\bar{\mathsf{v}}} h}{\partial_A h}\,.
\end{aligned} \tag{10.49}$$

We assumed the change of variable to be bijective, which agrees with $\partial_A h \neq 0$.

Now, we have identified the operators $\mathbf{D}$ and $\partial$. The second step is to invoke the lemma to obtain the Poisson bracket for the reduced dynamics. The assumptions of the lemma are verified[4] with

$$\begin{aligned}
\mathbf{D} &:= \mathbf{D}_\circ & \mathbf{P} &:= \mathbf{P}_\circ \\
\partial &:= \nabla & \mathbf{Q} &:= d\mathbf{q} \\
M &:= e\mathbf{B}_\circ\times
\end{aligned}$$

Indeed, the only assumptions which are not trivial are the one proven by Eq. (10.49), and the momentum part of the completeness relation, which is proven in the same way as the calculus for Eq. (10.49):

$$\overline{\mathbf{D}}_\circ.\mathbf{P}_\circ = \left(\frac{\bar{\mathsf{v}} + \frac{1}{h}\partial_\mathsf{v} h\mathcal{W}^2}{\partial_A h}\partial_A - (\partial_\mathsf{v})\frac{\mathcal{W}^2}{h}\right).(\mathsf{v}\partial_A h dA + (h + \mathsf{v}\partial_\mathsf{v} h)d\mathsf{v}) = (\partial_A)dA + (\partial_\mathsf{v})d\mathsf{v}\,.$$

Thus, we can apply the lemma, and the Poisson bivector for the reduced dynamics writes

$$\pi := \overline{\mathbf{D}}_\circ \wedge \nabla + \overline{\mathbf{D}}_\circ.e\mathbf{B}_\circ \times \mathbf{D}_\circ\,. \tag{10.50}$$

This Poisson bracket strongly resembles the initial bracket: the operator $\mathbf{D}$ was only replaced by $\mathbf{D}_\circ$, and the magnetic field $\mathbf{B}$ replaced by $\mathbf{B}_\circ$ (and $\partial_{\overline{\mathbf{q}}}$ replaced by the corresponding $\nabla$).

Our last task is to rewrite the Poisson bracket in a more convenient way. Indeed, instead of modifying $\mathbf{B}$, we can merge with $\nabla$ the correction due to the presence of $d\overline{\mathbf{q}}.M_\circ d\mathbf{q}$ in Eq. (10.47). Indeed

$$\overline{\mathbf{D}}_\circ.e(\mathbf{B}_\circ - \mathbf{B}) \times \mathbf{D}_\circ = -\overline{\mathbf{D}}_\circ.M_\circ\mathbf{D}_\circ = -\overline{\mathbf{D}}_\circ \wedge (\nabla h\bar{\mathsf{v}})\mathbf{D}_\circ\,,$$

and this term can be merged with $\overline{\mathbf{D}}_\circ \wedge \nabla$, defining $\nabla_\circ$

$$\overline{\mathbf{D}}_\circ \wedge (\nabla - (\nabla h\bar{\mathsf{v}})\mathbf{D}_\circ) = \overline{\mathbf{D}}_\circ \wedge \nabla_\circ \quad \text{with} \quad \nabla_\circ := \nabla - \frac{\nabla h}{\partial_A h}\partial_A\,.$$

We end up with the bivector

$$\pi := \overline{\mathbf{D}}_\circ \wedge \nabla_\circ + \overline{\mathbf{D}}_\circ.e\mathbf{B} \times \mathbf{D}_\circ\,, \tag{10.51}$$

which has the same structure as the initial Poisson bracket, but $\nabla$ has been redefined, instead of $\mathbf{B}$. The advantage of this choice is that $\nabla_\circ$ and $\mathbf{D}_\circ$ become "conjugated"

$$\mathbf{D}_\circ h = \frac{\mathsf{v} + \frac{1}{h}\mathcal{W}^2\partial_{\bar{\mathsf{v}}} h}{\partial_A h}\partial_A h - \frac{\mathcal{W}^2}{h}\partial_{\bar{\mathsf{v}}} h = \mathsf{v} \quad \text{and} \quad \nabla_\circ h = \nabla h - \frac{\nabla h}{\partial_A h}\partial_A h = 0\,,$$

just as $\partial_{\overline{\mathbf{p}}}$ and $\partial_{\overline{\mathbf{q}}}$ were in the initial variables

$$\partial_{\mathbf{p}} h = \partial_{\overline{\mathbf{p}}}\sqrt{\overline{\mathbf{p}}\mathbf{p}} = \mathsf{v} \quad \text{and} \quad \partial_{\overline{\mathbf{q}}} h = \partial_{\overline{\mathbf{q}}}|_{\mathbf{p}}\sqrt{\overline{\mathbf{p}}\mathbf{p}} = 0\,.$$

[4] The 1-form $\mathbf{P}_\circ$ is not closed, and it is why we pointed out that this assumption was unnecessary for the lemma.

With this choice, we reach our final result, the Poisson bracket for the reduced dynamics, acting on two functions $f$ and $g$ with variables $\mathbf{q}$, $A$ and $\mathsf{v}$:

$$\begin{aligned} \{f,g\} &= \mathbf{D}_\circ f \cdot \nabla_\circ g - \mathbf{D}_\circ g \cdot \nabla_\circ f + \mathbf{D}_\circ f \cdot e\mathbf{B} \times \mathbf{D}_\circ g\,, \\ \text{with } \nabla_\circ &:= \nabla - \tfrac{\nabla h}{\partial_A h}\partial_A \\ \text{and } \mathbf{D}_\circ &:= \frac{\mathsf{v} - \frac{\mathcal{P}_\perp}{h}\partial_{\mathsf{v}} h}{\partial_A h}\partial_A + \frac{\mathcal{P}_\perp}{h}\partial_{\mathsf{v}}\,. \end{aligned} \tag{10.52}$$

where $\mathcal{P}_\perp := 1 - \mathsf{v}\overline{\mathsf{v}}$ is the orthogonal projector onto the plane perpendicular to $\mathbf{p}$.

It is important to notice that the bracket we obtained is valid for any change of coordinates $(\mathbf{q},\mathbf{p}) \to (\mathbf{q},\mathsf{v},h) \to (\mathbf{q},\mathsf{v},A)$, the only requirement is for $A$ to be in bijection with $h$ and $\partial_A h \neq 0$. For instance, we could choose the change of coordinates given by the normal form truncated to second order $h = h_1 A^{1/2} + h_2 A^{2/2}$. Indeed, even if the normal form $h(\mathbf{q},\mathsf{v},A)$ can be explicitly computed only to some orders in $A$, the result (10.52) is exact to any order.

## Conclusion

For the dynamics of a charged particle in a strong magnetic field, this intrinsic approach provides an iterative computation for the normal form and an exact formula for the corresponding Poisson bracket as functions of the constant of motion $A$, without using any gyro-gauge, and lying at the particle position. The result is obtained directly from the definition of $A$ as a constant of motion through Eq. (10.6) and looking only for the normal form as a formal series in $A$. It turned out to be a series in $r_L^N \nabla^N$, which implies that the resulting expansion is perturbative when the magnetic field is strong, i.e. when its variations at the scale of the Larmor radius are small. We have shown that for the lowest orders this agrees with the previous results.

As expected by the magnetic moment being conjugated to the gyro-angle, the usual operators of gyrokinetics, such as gyro-average and gyro-fluctuation, appear in the process, but the initial absence of a gyro-gauge make them naturally defined in an intrinsic way.

By the way, the arbitrariness involved in the choice of the usual lowest-order adiabatic invariant was clarified: it corresponds to the choice of an element in the intersection of the kernel of two operators, which have to be inverted to get the normal form.

The convergence of the series was not considered, but it clearly implies a condition of strong $\mathbf{B}$, and it involves a domain in $(\frac{A}{B}, \varphi, \mathsf{b})$ with conditions not only on $\mathbf{B}$, but on all its derivatives as well, and also on the initial momentum.

This intrinsic reduction of particle dynamics in a strong magnetic field offers several possible follow-ups. Two most natural extensions are to compute the normal form at the next order, as was done in the literature [26, 157], and to take into account the presence of an electric field, which would make the secular equation more involved. However, such extensions would deviate from our purpose. Here we have considered the magnetic moment only because it provided a simplified version of the guiding-center reduction. The goal was to identify an intrinsic formulation for the gyro-angle in guiding-center theory. The most important continuation is to use this gauge-independent approach also in the other (and the main) part of the guiding-center reduction, i.e. when averaging the motion on the fast time scale. It is the next step of the work and is reported in Chapter 1.

On another hand, one of the main features of the present chapter is that it did not change the position coordinate. This will fit with a simplified framework for the lifting procedure, which will be used to transfer the reduction of particle dynamics to the Vlasov-Maxwell dynamics, whose space of indices will benefit from the reduction given by the constant of motion. This will be studied in Chapter 4.

In the presence of a dynamical electromagnetic field, the reduction for the magnetic moment has to be extended in order to take into account the coupling between the plasma and the electromagnetic field. This will be the topic of Chapters 5 and 6.

As a complementary comment about the work reported in the present chapter, we want to emphasize that the presentation follows our initial exploration about the magnetic moment, because the chapter was written in a specific context. Later on, several aspects could be simplified, but we could never take the time to write the upgraded derivation. We just indicate here the three main simplifying aspects.

First and main point, it is much more convenient to look first for the magnetic moment $A(\mathbf{q}, \mathsf{v}, h)$, instead of the normal form $h(\mathbf{q}, \mathsf{v}, A)$; then the normal form is obtained afterwards by inverting the series. In doing so, the equations to be solved become linear (instead of quadratic), which makes computations easier and the derivation more transparent. Especially, the induction relation depends only on terms of the previous order. It allows one to get an explicit formula for the secular equation as a function of only lowest-order average terms, whereas Eqs. (10.27) and (10.28) for instance are obfuscated by sums and convolutions, as well as by the presence of lowest-order fluctuating terms. Especially, this opens the way to solve the secular equation explicitly, in a similar way as the non-secular equation (10.15) or (10.28).

Second, the secular equation (10.14) relies on the inversion of the operator $\partial_*$, defined in Eqs.(10.66)-(10.67). It can be rewritten $B^{-\gamma}(\mathsf{b} \cdot \nabla_\star)B^\gamma$, with $\gamma = -\frac{1+\phi^2}{2\phi}\partial_\phi$. This fact makes the algorithm to solve the secular equation more transparent and more straightforward. It also explains the origin of the adiabatic invariant in a simpler way.

Third, the reduced Poisson bracket can actually be obtained by applying directly a chain rule, which is much faster than the method used in the present version.

In a word, this chapter is interesting for the methods used and for the results, but it is far from its optimized version.

# App. A: Gyro-rotations and related operators

In this appendix, we study four operators involved in our derivation. They are defined in a first subsection, and some of their properties useful for our computations are presented in a second subsection.

## The operators $\mathcal{D}$, $\mathcal{G}$, `avg` and `osc`

In this subsection, we first show that $\mathcal{D}$ is the generator of gyro-rotations. Next, we define its left-inverse $\mathcal{G}$. Then, the projectors onto the range and onto the kernel of $\mathcal{D}$ will turn out to be the gyro-fluctuating and the gyro-averaging operator. These four operators are involved in our procedure, which is not surprising since they are the most common operators of gyrokinetics.

First, in order to make $\mathcal{D}$ explicit, let us define a local right-handed orthonormal basis suited to the action of $\mathcal{D}$:

• $\mathsf{b} := \frac{e\mathbf{B}}{\|e\mathbf{B}\|}$, unit vector of the magnetic field (grouped with the particle charge $e$), whose norm is designated by $B := \|e\mathbf{B}\|$, so that $e\mathbf{B} = B\mathsf{b}$. With this notation, condition $\mathrm{div}\mathbf{B} = 0$ becomes $0 = \overline{\nabla}(B\mathsf{b}) = (\overline{\nabla}B)\mathsf{b} + B\overline{\nabla}\mathsf{b}$, which is most often used in the form

$$\overline{\nabla}\mathsf{b} = -\frac{B'\mathsf{b}}{B}\,. \tag{10.53}$$

• $\mathsf{c} := \frac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|} = \frac{(1-\mathsf{b}\bar{\mathsf{b}})\mathsf{v}}{\|(1-\mathsf{b}\bar{\mathsf{b}})\mathsf{v}\|}$, unit vector of $\mathbf{p}_\perp$, the part of the momentum perpendicular to $\mathbf{B}(\mathbf{q})$. We remind that $\|\mathbf{p}_\perp\| = \sin\varphi$, where $\varphi$ is the pitch-angle $\widehat{(\mathsf{b}, \mathbf{p})} \in [0; \pi]$;

• $\mathsf{a} := \mathsf{b} \times \mathsf{c} = \mathsf{b} \times \frac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|}$, unit vector of the Larmor radius, completing the direct orthonormal triad $(\mathsf{b}, \mathsf{c}, \mathsf{a})$.

With these notations, we can write $\bar{\mathsf{v}}.e\mathbf{B} \times \partial_{\bar{\mathsf{v}}} = B\bar{\mathsf{v}}.\mathsf{b} \times \partial_{\bar{\mathsf{v}}} =: B\mathcal{D}$, using the adimensional, scalar differential operator

$$\mathcal{D} := \bar{\mathsf{v}}.\mathsf{b} \times \partial_{\bar{\mathsf{v}}} = -\bar{\mathsf{a}}\partial_{\bar{\mathsf{c}}}\,.$$

It is the generator of the rotations around $e\mathbf{B}$ in the momentum space, because for any function of $\mathsf{v}$,

$$\left(e^{\Theta\mathcal{D}} f\right)(\mathsf{v}) = f(e^{-\Theta\mathcal{B}}\mathsf{v}), \quad \text{with } \mathcal{B} := \mathsf{b}\times, \tag{10.54}$$

where $e^{\Theta\mathcal{B}}$ is the rotation around the axis $\mathsf{b}$ through an angle $\Theta$, since $\mathcal{B}^3 = -\mathcal{B}$ implies

$$e^{\Theta\mathcal{D}} = \sum_{k=1}^{\infty} \frac{\mathcal{B}^k}{k!} = 1 + \cos\Theta\ \mathcal{B}^2 + \sin\Theta\ \mathcal{B}\,.$$

Eq. (10.54) is proven by expanding the function in $\mathsf{v}$: $f(\mathsf{v}) = \sum f_n \mathsf{v}^n$. The linearity of $\mathcal{D}$ leads to consider only monomials, for which the proof can be obtained by induction, or simply from the fact that $\mathcal{D}$ is a derivative, hence its exponential is a ring homomorphism: $e^{\mathcal{D}}(fg) = (e^{\mathcal{D}}f)(e^{\mathcal{D}}g)$, which leads us to consider only the elementary monomial $\mathsf{v}$, for which the property is obvious:

$$\mathcal{D}\overline{\mathsf{v}} = \overline{\mathsf{v}}\mathcal{B}, \quad \Rightarrow \quad \mathcal{D}\mathsf{v} = -\mathcal{B}\mathsf{v} = -\mathsf{b}\times\mathsf{v}\,.$$

Since $\mathcal{D}$ is the generator of rotations around $\mathsf{b}$, it plays exactly the role of $\partial_\Theta$, where $\Theta$ would correspond the gyro-angle of the guiding-center approach. On a Fourier series $f(\Theta) = f_0 + \sum_{n\neq 0} f_n \cos(n\Theta + \psi_n)$, the action of $\mathcal{D}$ is $\mathcal{D}f(\Theta) = -\sum_{n\neq 0} n f_n \sin(n\Theta + \psi_n)$.

Let us remind that the traditional gyro-angle $\theta$ is the angle indicating the direction of $\mathsf{c}$ (or $\mathsf{a}$) in $\mathbf{B}^\perp$ (the plane perpendicular to $\mathbf{B}$) with respect to some direction (more precisely a unit vector $\mathsf{e}_1$, called the "gyro-gauge", see Eq. (1.2)) chosen arbitrarily as a reference axis in $\mathbf{B}^\perp$. In our derivation, an important point to be noted is that we use the variable $\mathsf{c}$ and we do not define the gyro-angle $\theta$. This avoids the need for a gyro-gauge. We write $\theta$ only for a dummy variable in the gyro-averaging operator, and $\Theta$ for analogy, to identify the fast part of the dynamics, which explains why we have chosen not to write it $\theta$.

The kernel of $\mathcal{D}$ consists of all functions that do not depend on $\Theta$, i.e. on the direction of the vector $\mathsf{c}$ (or $\mathsf{a}$); $\mathcal{D}$ takes values in the functions with zero average in $\Theta$, which is a supplementary space to $\ker\mathcal{D}$.

This allows to define a left-inverse $\mathcal{G} =$ "$\mathcal{D}^{-1}$", i.e. a function with value $\mathcal{D}^{-1}$ on $\mathrm{Rg}\,(\mathcal{D})$, and with value zero on the complementary space $\ker(\mathcal{D})$. The operator $\mathcal{G}$ is such that

$\mathtt{osc} := \mathcal{D}\mathcal{G} = \mathcal{G}\mathcal{D}$ is the (non-resonant) projector onto $(\mathrm{Rg}\,(\mathcal{D}))$ parallel to $\ker(\mathcal{D})$,

$\mathtt{avg} := 1 - \mathtt{osc}$ is the complementary gyro-averaging (resonant) operator.

On the previous Fourier series, the effect of $\mathcal{G}$ cancels almost the effect of $\mathcal{D}$:

$$\mathcal{G}\mathcal{D}f = \mathcal{G}\left(-\sum_{n\neq 0} n f_n \sin(n\Theta + \psi)\right) = \sum_{n\neq 0} f_n \cos(n\Theta + \psi) = f - f_0 = \mathtt{osc}\,(f)\,.$$

It is the operator of gyro-fluctuation, whose effect is to remove the gyro-average. The operator $\mathcal{G}$ can be seen as a kind of $\int^\theta d\theta'$, but which vanishes on the functions that do not depend on $\theta$, and which implies no gyro-gauge in its definition. As for $\mathtt{avg}$, on the contrary, it gives the gyro-averaged part of a function: $\mathtt{avg}\, f = f_0$, so that $\mathtt{avg} = \oint d\theta e^{\theta\mathcal{D}}$, where $\oint d\theta := \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta$ is the loop integral in $\mathbb{S}^1$. Thus

$$\begin{aligned}
\mathtt{avg} &= \oint d\theta e^{\theta\mathcal{D}} = \frac{\sinh(\pi\mathcal{D})}{\pi\mathcal{D}}\,,\\
\mathtt{osc} &= 1 - \mathtt{avg} = \frac{\pi\mathcal{D} - \sinh(\pi\mathcal{D})}{\pi\mathcal{D}}\,,\\
\mathcal{G} &= \frac{\mathtt{osc}}{\mathcal{D}} = \frac{\pi\mathcal{D} - \sinh(\pi\mathcal{D})}{\pi\mathcal{D}^2} = \oint d\theta(\theta - \pi\ \mathrm{sign}\ \theta)\ e^{\theta\mathcal{D}}\,.
\end{aligned}$$

One can check that the effect of these operators is as expected. For instance, $\mathtt{avg} = 1$ when $\mathcal{D}$ is 0, which corresponds to the gyro-averaged functions, whereas for functions with zero average we have $\mathtt{avg} = 0$, because $\mathcal{D} \neq 0$ and $\sinh \pi\mathcal{D} = 0$ since

$$2\sinh(\pi\mathcal{D})f(\mathbf{p}) = (e^{\pi\mathcal{D}} - e^{-\pi\mathcal{D}})f = f(e^{-\pi\mathcal{B}}\mathbf{p}) - f(e^{\pi\mathcal{B}}\mathbf{p}) = 0\,.$$

Thus, in the present intrinsic approach, the gyro-averaging $\mathtt{avg}$ and gyro-fluctuating $\mathtt{osc}$ operators, as well as the gyroderivative $\mathcal{D}$ and gyrointegral $\mathcal{G}$ operators appear naturally from the presence of $\bar{\mathsf{v}}.\mathbf{B} \times \partial_{\bar{\mathsf{v}}}$ in the equation for the normal form (10.6).

### Computational properties of $\mathcal{D}$, $\mathcal{G}$, $\mathtt{avg}$ and $\mathtt{osc}$

In this subsection, we summarize a few properties we need for our computations. The first ones come from the range and the kernel of the operators. The second ones are some kinds of integration by parts formulas. The third ones make explicit the action of the four operators for the cases we meet in our calculations.

First of all, the definitions of the four operators immediately imply

$$\mathtt{avg}\,\mathcal{D} = \mathtt{avg}\,\mathcal{G} = \mathtt{avg}\,\mathtt{osc} = 0 = \mathcal{D}\mathtt{avg} = \mathcal{G}\mathtt{avg} = \mathtt{osc}\,\mathtt{avg}\,.$$

This property is to be remembered, because it simplifies many computations. For instance in Eq. (10.69), the second term $\mathtt{avg}\left(\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4}\right)$ is zero because it is $\mathtt{avg}\,\mathcal{G}\left(\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}\right)$. We remind that the operators $\mathcal{D}$, $\mathcal{G}$, $\mathtt{avg}$ and $\mathtt{osc}$ act on the whole expression on their right-hand side.

Next, the previous property, together with Leibniz formula for $\mathcal{D}$, induces several useful properties. For instance, some kinds of integration by parts formulas can be obtained:

$$\begin{aligned}
&\mathtt{avg}\left(g\mathcal{D}f\right) = -\mathtt{avg}\left(f\mathcal{D}g\right),\\
&\mathtt{avg}\left(f\mathcal{D}f\right) = 0\,,\\
&\mathtt{avg}\left(f\mathcal{G}g\right) = -\mathtt{avg}\left(g\mathcal{G}f\right),\\
&\mathtt{avg}\left(f\mathcal{G}f\right) = 0\,.
\end{aligned} \tag{10.55}$$

These results simplify many formulas in the computation of $a_2$. Especially, they allow us not to deal with harmonics of order 4, because all those we get can be written $\mathtt{avg}\left(f\mathcal{G}f\right)$ for some function $f$; for instance, the second term in Eq. (10.71) contains $\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4}\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}$, and it writes $\mathtt{avg}\left(f\mathcal{G}f\right)$ with $f = \bar{\mathsf{c}}\mathsf{b}'\mathsf{c}$.

Now, let us now compute explicitly the action of $\mathcal{D}$. As a preliminary matter, an important point to notice is that, when acting on a (tensorial) product of order $k$ in the fast variables[5], $\mathcal{D}$ takes values in the same space: the set $D_k$ of harmonics $k$ in the fast variables. Hence, we split the set of all functions on the phase space into the direct sum of the spaces $D_k$ of all functions of equal order $k$ in the fast variables, and we study the action of $\mathcal{D}$ inside each of these spaces. We will denote by $\mathcal{D}_k$, $\mathtt{avg}_k$, $\mathcal{G}_k$ and $\mathtt{osc}_k$ the restriction of the corresponding operators onto the space of harmonics $k$ in the fast variables.

For harmonics of order zero $k = 0$, the results are trivial: $\mathcal{D}_0 = 0$ because all functions are constant in $\mathsf{c}$. Thus $\ker \mathcal{D}_0 = D_0$, $\mathcal{G}_0 = 0$, $\mathtt{osc}_0 = 0$, $\mathtt{avg}_0 = 1$.

For harmonics of order $k = 1$, we have $\mathcal{D}\mathsf{c} = -\bar{\mathsf{a}}\partial_{\bar{\mathsf{c}}}\mathsf{c} = -\mathsf{a}$ and $\mathcal{D}\mathsf{a} = -\bar{\mathsf{a}}\partial_{\bar{\mathsf{c}}}(\mathsf{b}\times\mathsf{c}) = -(\mathsf{b}\times\mathsf{a}) = \mathsf{c}$. Thus, in the basis $(\mathsf{c},\mathsf{a})$, the matrix of the endomorphism $\mathcal{D}_1$ is $M := \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$

$$\ker \mathcal{D}_1 = 0,\ \mathtt{avg}_1 = 0,\ \mathtt{osc}_1 = 1 \text{ and } \mathcal{G}_1 = \mathcal{D}_1^{-1} = -\mathcal{D}_1\,. \tag{10.56}$$

[5]That means that there are $k$ factors $\mathsf{c}$ and/or $\mathsf{a}$; such a product is called a *"harmonic of order $k$"*, or more simply "harmonics $k$".

For harmonics of order $k = 2$, a natural basis of $D_2$ is $(\mathsf{cc}, \mathsf{aa}, \mathsf{ca}, \mathsf{ac})$. The meaning of this basis is the following: $\mathsf{ca}$ stands for any function that depends tensorially on the fast variables through an ordered product $\mathsf{ca}$, for instance $\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}$, or $B'\mathsf{c}B'\mathsf{a}$.

$$\begin{aligned}
\text{Now, } \mathcal{D}\mathsf{cc} &= (\mathcal{D}\mathsf{c})\mathsf{c} + \mathsf{c}\mathcal{D}\mathsf{c} = -(\mathsf{ac} + \mathsf{ca})\,,\\
\mathcal{D}\mathsf{aa} &= (\mathcal{D}\mathsf{a})\mathsf{a} + \mathsf{a}\mathcal{D}\mathsf{a} = (\mathsf{ca} + \mathsf{ac})\,,\\
\mathcal{D}\mathsf{ca} &= (\mathcal{D}\mathsf{c})\mathsf{a} + \mathsf{c}\mathcal{D}\mathsf{a} = -\mathsf{aa} + \mathsf{cc}\,,\\
\mathcal{D}\mathsf{ac} &= (\mathcal{D}\mathsf{a})\mathsf{c} + \mathsf{a}\mathcal{D}\mathsf{c} = \mathsf{cc} - \mathsf{aa}\,.
\end{aligned}$$

In this basis, the matrix of the endomorphism $\mathcal{D}_2$ is $M := \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix}$.

So, $\operatorname{Rg}\mathcal{D}_2$ is the space generated by $(\mathbf{u}_1, \mathbf{u}_2) := (\mathsf{cc} - \mathsf{aa}, \mathsf{ca} + \mathsf{ac})$. And $\ker \mathcal{D}_2$ is the space generated by $(\mathbf{u}_3, \mathbf{u}_4) := (\mathsf{cc} + \mathsf{aa}, \mathsf{ca} - \mathsf{ac})$. Using this basis, and coming back to the natural basis, it is straightforward to get

$$\mathtt{avg}_2 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \qquad \mathtt{osc}_2 = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \qquad \mathcal{G}_2 = \frac{1}{4}\begin{pmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}. \tag{10.57}$$

With these formulas, one can compute directly the action of any of the four operators on any function of harmonic $k \leqslant 2$ in the fast variables. For instance, $\mathcal{G}\mathsf{c} = \mathsf{a}$ and $\mathcal{G}\bar{\mathsf{c}}\mathsf{b}'\mathsf{c} = \frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4}$ (which was used to simplify Eq. (10.29)) can be read in the previous formulas. And the term $\mathtt{avg}\left(\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}\right)$, which appears in Eq. (10.65), reads as $\mathtt{avg}\left(\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}\right) = \overline{\nabla\mathsf{b}}\ \mathtt{avg}\left(\mathsf{cc}\right) = \frac{1}{2}\overline{\nabla\mathsf{b}}(\mathsf{cc} + \mathsf{aa}) = \frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{c} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{a}}{2} = \frac{\overline{\nabla\mathsf{b}}}{2}$, where the last equality follows from $\bar{\mathsf{b}}\mathsf{b}' = 0$.

We could compute the action of $\mathcal{D}_k$, $\mathtt{avg}_k$, $\mathtt{osc}_k$ and $\mathcal{G}_k$ for the following harmonics $k \geqslant 3$, but they are not needed here. The only property we need is the action of $\mathtt{avg}$ on harmonics 3, but the result is zero because every odd harmonic is a pure fluctuation, as is obvious using a Fourier decomposition in the angle $\Theta$.

Thus, on the space of odd harmonic functions

$$\mathtt{avg}_{odd} = 0, \quad \mathtt{osc}_{odd} = 1, \quad \mathcal{D}_{odd} \text{ is invertible}, \quad \text{and } \mathcal{G}_{odd} = \mathcal{D}_{odd}^{-1}\,. \tag{10.58}$$

The property $\mathtt{avg}_{odd} = 0$ is especially useful in our derivation because it cancels many terms in the computations for $a_2$, for instance in the calculations leading to Eq. (10.69).

## App. B: Explicit expression for the operator $\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)$

In this appendix, we obtain an explicit expression for $\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)$, which is involved in the equation for the lowest-order term of the normal form $\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right) = 0$. Since $\alpha_1$ is an average, the operator $\mathtt{avg}$ acts only on $\bar{\mathsf{v}}\nabla$; our aim is to perform the result of this action. First, we will express the operator $\nabla$ in coordinates suited to the action of $\mathtt{avg}$, then we will be able to reach an explicit expression for $\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)$, i.e. an expression where the operator $\mathtt{avg}$ is absent.

The operator $\mathtt{avg}$ performs an average over the direction of the perpendicular momentum, and it is convenient to decompose the variable $\mathsf{v} \longrightarrow (\phi, \mathsf{c})$, into two components: its perpendicular (non-averaged) part, and its parallel (averaged) part. As parallel component, we choose the pitch-angle $\varphi \in [0; \pi]$ defined in Eq. (10.13), or rather its cotangent

$$\phi := \frac{\bar{\mathsf{b}}\mathsf{v}}{\bar{\mathsf{c}}\mathsf{v}} \in \mathbb{R}\,,$$

which is in bijection with $\varphi$ and will exhibit a polynomial structure in the normal form expansion. As for the perpendicular part, the simplest choice is the unit vector of the component of $\mathsf{v}$ perpendicular to $\mathbf{B}$:

$$\mathsf{c} := \frac{(1 - \mathsf{b}\bar{\mathsf{b}})\mathsf{v}}{\|(1 - \mathsf{b}\bar{\mathsf{b}})\mathsf{v}\|} \in \mathbb{S}^1\,.$$

This avoids a gyro-angle which would imply to introduce a gyro-gauge. In App. A, more details are given about the local frame $(\mathsf{b},\mathsf{c},\mathsf{a})$.

In the variables $(\mathbf{q},\mathsf{c},\phi,A)$, the operator $\mathcal{D}$ writes $\bar{\mathsf{c}}\mathcal{B}\partial_{\bar{\mathsf{c}}} = -\bar{\mathsf{a}}\partial_{\bar{\mathsf{c}}}$; so, it does not affect the variables $\phi$, $A$, and $\mathbf{q}$. Its effect is only to replace $\mathsf{c}$ by $-\mathsf{a}$, and therefore $\mathsf{a}=\mathsf{b}\times\mathsf{c}$ by $-\mathsf{b}\times\mathsf{a}=\mathsf{c}$.

As for the operator $\nabla$, it writes

$$\begin{aligned}\nabla :&= \nabla|_{A,\mathsf{v}} = \nabla|_{A,\phi,\mathsf{c}} + (\nabla|_{A,\mathsf{v}}A)\partial_A|_{\phi,\mathsf{c},\mathbf{q}} + (\nabla|_{A,\mathsf{v}}\phi)\partial_\phi|_{A,\mathsf{c},\mathbf{q}} + (\nabla|_{A,\mathsf{v}}\bar{\mathsf{c}})\partial_{\bar{\mathsf{c}}}|_{A,\phi,\mathbf{q}}\\ &= \nabla_\star + \nabla\phi\partial_\phi + \nabla\bar{\mathsf{c}}\partial_{\bar{\mathsf{c}}}\,,\end{aligned} \tag{10.59}$$

which is a definition for the operators $\nabla_\star$, $\partial_\phi$, and $\partial_{\bar{\mathsf{c}}}$.

To compute $\nabla\phi$ and $\nabla\bar{\mathsf{c}}$, we obtain successively

$$\begin{aligned}&\nabla(\cos\varphi) = \nabla(\bar{\mathsf{b}}\mathsf{v}) = \nabla(\bar{\mathsf{b}})\mathsf{v} = \nabla\bar{\mathsf{b}}(\cos\varphi\ \mathsf{b} + \sin\varphi\ \mathsf{c}) = \nabla\bar{\mathsf{b}}\sin\varphi\mathsf{c}\,,\\ &\nabla(\sin\varphi) = \nabla(\sqrt{1-\cos^2\varphi}) = -\frac{\cos\varphi}{\sin\varphi}\nabla(\cos\varphi) = -\phi\sin\varphi\nabla\bar{\mathsf{b}}\mathsf{c} = -\cos\varphi\nabla\bar{\mathsf{b}}\mathsf{c}\,,\\ &\nabla\phi = \nabla\left(\frac{\cos\varphi}{\sin\varphi}\right) = \frac{\nabla\bar{\mathsf{b}}\mathsf{c}}{\sin^2\varphi} = \nabla\bar{\mathsf{b}}\mathsf{c}(1+\phi^2)\,,\\ &\nabla\bar{\mathsf{c}} = -\frac{\nabla\bar{\mathsf{b}}(\sin\varphi\mathsf{c}\bar{\mathsf{b}} + \cos\varphi\mathsf{a}\bar{\mathsf{a}})}{\sin\varphi} = -\nabla\bar{\mathsf{b}}(\mathsf{c}\bar{\mathsf{b}} + \phi\mathsf{a}\bar{\mathsf{a}})\,.\end{aligned} \tag{10.60}$$

The last line comes from $0=\nabla\bar{\mathsf{v}} = \nabla(\cos\varphi\ \bar{\mathsf{b}} + \sin\varphi\ \bar{\mathsf{c}})$. We finally get the expected expression for $\nabla$ in variables $(\mathbf{q},\mathsf{c},\phi,A)$

$$\nabla = \nabla_\star + \nabla\phi\partial_\phi + \nabla\bar{\mathsf{c}}\partial_{\bar{\mathsf{c}}} = \nabla_\star + \nabla\bar{\mathsf{b}}\mathsf{c}(1+\phi^2)\partial_\phi - \nabla\bar{\mathsf{b}}(\mathsf{c}\bar{\mathsf{b}} + \phi\mathsf{a}\bar{\mathsf{a}})\partial_{\bar{\mathsf{c}}}\,. \tag{10.61}$$

Now, from this expression for $\nabla$, we can compute $\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)$ explicitly. As $\alpha_1$ is gyro-averaged, $0=\mathcal{D}\alpha_1 = -\bar{\mathsf{a}}\partial_{\bar{\mathsf{c}}}\alpha_1$, and the last term in equation(10.61) does not contribute on $\alpha_1$:

$$\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right) = \mathtt{avg}\left(\bar{\mathsf{v}}\nabla_\star\alpha_1\right) + \mathtt{avg}\left(\bar{\mathsf{v}}\nabla\bar{\mathsf{b}}\mathsf{c}(1+\phi^2)\partial_\phi\alpha_1\right) - \mathtt{avg}\left(\bar{\mathsf{v}}\nabla\bar{\mathsf{b}}\mathsf{c}\right)\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\alpha_1\,. \tag{10.62}$$

Notice that $\nabla$ acts only on the following term, whereas $\mathtt{avg}$ acts on the whole expression on its right-hand side. The first and third terms in Eq. (10.62) write

$$\mathtt{avg}\left(\cos\varphi\ \bar{\mathsf{b}} + \sin\varphi\ \bar{\mathsf{c}}\right)\nabla_\star\alpha_1 = \cos\varphi\ \bar{\mathsf{b}}\ \mathtt{avg}\left(\nabla_\star\alpha_1\right)\,, \tag{10.63}$$

$$\text{and } -\mathtt{avg}\left(\cos\varphi\ \bar{\mathsf{b}} + \sin\varphi\ \bar{\mathsf{c}}\right)\nabla\bar{\mathsf{b}}\mathsf{c}\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\alpha_1 = -\cos\varphi\ \mathtt{avg}\left(\bar{\mathsf{b}}\nabla\bar{\mathsf{b}}\mathsf{c}\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\alpha_1\right)\,. \tag{10.64}$$

Now, $-\bar{\mathsf{b}}\nabla\bar{\mathsf{b}}\mathsf{c}\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\alpha_1$ is not gyro-averaged, nor is $\bar{\mathsf{b}}\nabla_\star\alpha_1$, because of the action of $\bar{\mathsf{b}}\nabla_\star$ on $\mathsf{a} := \mathsf{b}\times\mathsf{c}$:

$$\bar{\mathsf{b}}\nabla_\star\mathsf{a} = (\bar{\mathsf{b}}\nabla_\star\mathsf{b})\times\mathsf{c} = \bar{\mathsf{b}}\nabla\bar{\mathsf{b}}(\mathsf{ba}-\mathsf{ab}) = -\bar{\mathsf{b}}\nabla\bar{\mathsf{b}}\mathsf{ab}\,.$$

However, the last formula shows that $\bar{\mathsf{b}}\nabla_\star\alpha_1$ generates such a fluctuating part on the $\mathsf{a}$ present in $\alpha_1$ as to correspond exactly to the fluctuating part coming from the action of $-\bar{\mathsf{b}}\nabla\bar{\mathsf{b}}\mathsf{c}\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\alpha_1$ on the $\mathsf{c}$ present in $\alpha_1$. Thus, the sum of these terms is a gyro-average. Hence, together the first and third terms in Eq. (10.62) give $\cos\varphi\ (\bar{\mathsf{b}}\nabla_\star - \bar{\mathsf{b}}\nabla\bar{\mathsf{b}}\mathsf{c}\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}})\alpha_1$.

As for the second term in Eq. (10.62), using formulas (10.56) and (10.57) from App. A, it writes

$$\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\bar{\mathsf{b}}\mathsf{c}\right)(1+\phi^2)\partial_\phi\alpha_1 = \mathtt{avg}\left((\cos\varphi\ \bar{\mathsf{b}} + \sin\varphi\ \bar{\mathsf{c}})\nabla\bar{\mathsf{b}}\mathsf{c}\right)(1+\phi^2)\partial_\phi\alpha_1 = \sin\varphi\ \frac{\overline{\nabla}\mathsf{b}}{2}(1+\phi^2)\partial_\phi\alpha_1\,. \tag{10.65}$$

Finally, Eqs. (10.63) and (10.65) allow us to write $\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)$ as

$$\mathtt{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right) = \cos\varphi\ \bar{\mathsf{b}}\nabla_\star\alpha_1 + \sin\varphi\ \frac{\overline{\nabla}\mathsf{b}}{2}(1+\phi^2)\partial_\phi\alpha_1 - \cos\varphi\bar{\mathsf{b}}\nabla\bar{\mathsf{b}}\mathsf{c}\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}\alpha_1\,. \tag{10.66}$$

This expression for $\mathsf{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)$ is explicit, but it can be simplified. As was previously mentioned, the exact role of $-\bar{\mathsf{b}}\nabla\bar{\mathsf{b}}\mathsf{c}\ \bar{\mathsf{b}}\partial_{\bar{\mathsf{c}}}$ is to compensate the non-averaged effect of $\bar{\mathsf{b}}\nabla_\star$; these terms are necessary only when $\alpha_1$ is expressed as a function of $\mathsf{c}$ and $\mathsf{a}$. Nevertheless, as $\alpha_1$ is gyro-averaged, it is always possible to write it without using $\mathsf{c}$ and $\mathsf{a}$: For instance, $\bar{\mathsf{c}}\nabla\bar{\mathsf{b}}\mathsf{c}+\bar{\mathsf{a}}\nabla\bar{\mathsf{b}}\mathsf{a}$ is gyro-averaged, and it can be written $\overline{\nabla}\mathsf{b}$. In that case[6], the last term does not contribute in Eq. (10.66) and we get our final expression for $\mathsf{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)$

$$\mathsf{avg}\left(\bar{\mathsf{v}}\nabla\alpha_1\right)=\cos\varphi\ \left(\bar{\mathsf{b}}\nabla_\star+\frac{\overline{\nabla}\mathsf{b}}{2}\frac{1+\phi^2}{\phi}\partial_\phi\right)\alpha_1\,. \tag{10.67}$$

It is important to notice that the derivation leading to Eq. (10.67) relied on the hypothesis that $\alpha_1$ is a pure average. It is valid for $\alpha_i$ to any order $i$, but it is not valid for $\eta_i$: the action of $\mathsf{avg}\,\bar{\mathsf{v}}\nabla$ on fluctuating functions such as $\eta_i$ is not easy to obtain, because the action of $\mathsf{avg}$ on $\bar{\mathsf{v}}\nabla$ can hardly be distinguished from its action on the function $\eta_i$.

## App. C: Computation of the right-hand side of Equation (10.31)

In this appendix, we compute the right-hand side of Eq. (10.31), which we remind here

$$\phi\partial_{\mathrm{a}}a_2=-\mathsf{avg}\left\{2\eta_1\sqrt{B}'\mathsf{v}+\sqrt{B}(\eta_1)'\mathsf{v}+2(\sqrt{B}\eta_1)\phi\bar{\mathsf{c}}\mathsf{b}'\mathsf{v}\right\}\,. \tag{10.68}$$

Its first term writes

$$-2\mathsf{avg}\left(\eta_1\frac{1}{2\sqrt{B}}B'\mathsf{v}\right)=\frac{1}{B}\left\{\frac{1}{2B}B'B'\mathsf{avg}\,(\mathsf{ca})+\phi^2B'\mathsf{b}\mathsf{avg}\left(\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4}\right)+\phi^2B'\mathsf{avg}\,(\mathsf{c}\bar{\mathsf{a}})\mathsf{b}'\mathsf{b}\right\}\,, \tag{10.69}$$

because every term with odd number of $\mathsf{c}$ and/or $\mathsf{a}$ is a pure fluctuation and cancels under the action of $\mathsf{avg}$ (see Eq. (10.58) in the appendix).
Now, the first term in Eq. (10.69) gives $\mathsf{avg}\left(\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4}\right)=\mathsf{avg}\,\mathcal{G}\left(\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}\right)$, which is zero because $\mathsf{avg}\,\mathcal{G}=0$. The second term contains $B'B'\mathsf{avg}\,(\mathsf{ca})=B'B'(\frac{\mathsf{ca}-\mathsf{ac}}{2})$ (see Eq. (10.57) in the appendix); but this is zero because $B'B'$ is symmetric whereas $\mathsf{ca}-\mathsf{ac}$ is antisymmetric. Last, for the third term $\mathsf{avg}\,(\mathsf{c}\bar{\mathsf{a}})=\frac{\mathsf{c}\bar{\mathsf{a}}-\mathsf{a}\bar{\mathsf{c}}}{2}$. Finally,

$$-\mathsf{avg}\left(\eta_1\frac{1}{\sqrt{B}}B'\mathsf{v}\right)=\phi^2\frac{B'}{B}\frac{\mathsf{c}\bar{\mathsf{a}}-\mathsf{a}\bar{\mathsf{c}}}{2}\mathsf{b}'\mathsf{b}\,. \tag{10.70}$$

The third term in the right-hand side of Eq. (10.68) writes

$$-\mathsf{avg}\left(2\sqrt{B}\eta_1\phi\bar{\mathsf{c}}\mathsf{b}'(\phi\mathsf{b}+\mathsf{c})\right)=2\phi\mathsf{avg}\left\{\phi\frac{B'\mathsf{a}}{2B}\ \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}+\phi\frac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4}\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}+\phi^3\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\ \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\right\}\,, \tag{10.71}$$

where again, in the first equality, we discarded all the odd terms in $\mathsf{c}$ and/or $\mathsf{a}$, because their gyro-average is zero. The last term in Eq. (10.71) gives zero because $\mathsf{avg}\left(\bar{\mathsf{c}}\bar{\mathsf{a}}\right)=\frac{\bar{\mathsf{c}}\bar{\mathsf{a}}-\bar{\mathsf{a}}\bar{\mathsf{c}}}{2}$ is antisymmetric and it is multiplied by a symmetric factor. The second term is also zero, because, with $f=\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}$, it writes $\mathsf{avg}\,(f\mathcal{G}f)$, which is zero by Eq. (10.55). Finally, we get

$$-\mathsf{avg}\left(2\sqrt{B}\eta_1\phi\bar{\mathsf{c}}\mathsf{b}'(\phi\mathsf{b}+\mathsf{c})\right)=2\phi^2\frac{B'}{2B}\mathsf{avg}\,(\mathsf{a}\ \bar{\mathsf{c}})\mathsf{b}'\mathsf{b}=\phi^2\frac{B'}{B}\frac{\mathsf{a}\bar{\mathsf{c}}-\mathsf{c}\bar{\mathsf{a}}}{2}\mathsf{b}'\mathsf{b}\,, \tag{10.72}$$

so that in the right-hand side of Eq. (10.68), the third term, which is (10.72), cancels exactly the first term, which is (10.70).

[6] In the general case, the term can also be removed, by using a redefinition of $\nabla_\star$. The resulting operator is precisely the reference covariant derivative $\nabla_*$ of Chapters 2-3.

Eq. (10.68) for $a_2$ then becomes

$$\phi\partial_{\mathrm{a}}a_2 = -\mathsf{avg}\left(\sqrt{B}\eta_1'\mathsf{v}\right). \tag{10.73}$$

Our last task is to compute $\eta_1'\mathsf{v}$. The derivation is straightforward, but it is too long a task for all the details to be given here. We only sketch the method, which has already been illustrated on the derivation of Eqs. (10.69)–(10.72). First, one computes the derivative of $\eta_1$, using Eq. (10.60) for $\phi'$ and $\mathsf{c}'$, so that they both give terms in which the only derivative is in $\mathsf{b}'$. The derivative of $\mathsf{a}$ are expressed in the same way as

$$\mathsf{a}' = (\mathsf{b}\times\mathsf{c})' = \mathsf{b}\times\mathsf{c}' - \mathsf{c}\times\mathsf{b}' = (\phi\mathsf{c}\bar{\mathsf{a}} - \mathsf{b}\bar{\mathsf{a}})\mathsf{b}'. \tag{10.74}$$

So, only the derivatives of $B$ and of $\mathsf{b}$ are kept, there is no general way to simplify them. Second, one uses $\mathsf{v} = \phi\mathsf{b}+\mathsf{c}$ and expands the products, discarding all terms with odd number of $\mathsf{c}$ and/or $\mathsf{a}$; some simplifications occur because of the orthonormality relations between $\mathsf{b}$, $\mathsf{c}$ and $\mathsf{a}$ and because $\bar{\mathsf{b}}\mathsf{b}' = 0$ since $\mathsf{b}$ is a unit vector. Then, one computes the action of $\mathsf{avg}$ on each term, using formulas from the appendix, and especially Eq. (10.57). Many simplifications occur because many terms are products between a symmetric and an antisymmetric factor, or they can be written $\mathsf{avg}\,\mathcal{G}f$ or $\mathsf{avg}\left(f\mathcal{G}f\right)$ for some function $f$. Finally, one gets only three terms

$$\phi\partial_{\mathrm{a}}a_2 = \mathsf{avg}\left\{-\frac{B'\mathsf{b}}{2B}\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} + \phi^2(\bar{\mathsf{a}}\mathsf{b}''\mathsf{b}\mathsf{c} + \bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c})\right\},$$

in which the double prime means double spatial derivative: $\bar{\mathsf{a}}\mathsf{b}''\mathsf{b}\mathsf{c}$ is $(\mathsf{b}\mathsf{c} : \nabla\nabla\mathsf{b})\cdot\mathsf{a}$, where ":" denotes double contraction.
The first term comes from the action of $\bar{\mathsf{v}}\nabla$ upon the $\mathsf{a}$ in $\frac{B'\mathsf{a}}{2B}$, the last two ones come from its action upon the $\mathsf{b}'$ and $\mathsf{b}$ in $\phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}$. All other terms have cancelled. After computing the action of $\mathsf{avg}$ with Eq. (10.57), the right-hand side of Eq. (10.73) becomes

$$-\frac{B'\mathsf{b}}{2B}\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}''\mathsf{b}\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{b}\mathsf{a}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a}}{2}\right).$$

Finally, Eq. (10.31) writes

$$\phi\partial_{\mathrm{a}}a_2 = -\frac{B'\mathsf{b}}{2B}\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{a}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}''\mathsf{b}\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}''\mathsf{b}\mathsf{a}}{2}\right) + \phi^2\left(\frac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}'\mathsf{c} - \bar{\mathsf{c}}\mathsf{b}'\mathsf{b}'\mathsf{a}}{2}\right).$$

# Chapter 11

# Lifting coordinate changes to a fluid-like field Hamiltonian dynamics

in collaboration with Phil J. Morrison, Alain J. Brizard, and Michel Vittot

*A*bstract: The lifting of a general coordinate transformation to the Hamiltonian structure of the corresponding field dynamics is studied, with its effects on the Hamiltonian functional, the Poisson bracket, the functional derivatives, and the dynamics. It applies to any Hamiltonian field dynamics where all the fields are defined over the same space, which undergoes a coordinate transformation that depends on the fields. A concrete example is an ideal-fluid dynamics (e.g. ideal MHD) when studied in magnetic coordinates. An essential motivation (but less immediate application) can be found in gyrokinetic theory.
The effects for the functional chain rule expected from the simplified magnetic-moment-like lifting are recovered, but additional subtleties are introduced by releasing the simplifying assumptions previously considered, because the transformation now depends on fields defined over a space which is affected by the transformation. For instance, this implies the field point to be moved by the transformation, which precludes a trivial composition of derivative operators. This also makes clearer that the transformation affects functional derivatives, although it is mute for functionals.

## Introduction

This appendix chapter takes place between Chapters 4 and 5. It is interested in the lifting method in a case more involved than for the magnetic-moment reduction, since all the coordinates can now be affected by the transformation, but not as involved as for the guiding-center reduction, since it will be restricted to the case where all the fields are defined over the same coordinate space.

Indeed, as a first step towards the lifting of the guiding-center and gyro-center reductions of particle dynamics to the Hamiltonian structure of the Vlasov-Maxwell field theory, Chapter 4 treated the introductory case where the transformation affected only the velocity coordinate; the particle position was not changed. This restriction was useful to simplify the scheme because the transformation depends only on the magnetic (and possibly electric) field, which is defined over the position space. It aimed at identifying the essential mechanism involved in the lifting, while avoiding questions caused by a transformation that would depend on fields that are themselves affected by the transformation.

For the full guiding-center transformation, the previous restriction must be dropped. So we now turn to a general lifting where all of the coordinates can be changed, even those of the definition space for the fields involved in the transformation. This quite general lifting will apply to many other field dynamics than the Vlasov-Maxwell system.

This last system is more involved than many other field theories, for instance fluid dynamics,

because different coordinate spaces are involved: the Vlasov density $f$ is defined over the phase space $(\mathbf{q}, \mathbf{p})$, whereas the electromagnetic field $(\mathbf{E}, \mathbf{B})$ is defined over the configuration space $\mathbf{q}$. For the sake of simplicity, we first study the lifting mechanism in details for a general fluid-like system, i.e. in the special case where all the fields are defined over the same coordinate space (denoted by $\mathbf{z}$).

This applies to any Hamiltonian fluid theory, when the coordinate space (which is the usual configuration space $\mathbf{z} := \mathbf{q}$) undergoes a field-dependent coordinate transformation. For instance, in ideal-fluid dynamics or in ideal magnetohydrodynamics, when magnetic coordinates are adopted while the magnetic field is dynamical, then the Hamiltonian structure of the corresponding field dynamics is affected by the coordinate change and the treatment below applies. Adaptation of the lifting to the Vlasov-Maxwell system will be investigated as the next part of the work, in Chapter 5.

The organization of the chapter is the following. After introducing the framework and transformation operators in Sec. 11.1, we will turn to the transformed Hamiltonian functional and Poisson bracket in Sec. 11.2. In order to transform functional derivatives, the chain rule will be applied in Sec. 11.3. The explicit computation of the transformed functional derivatives will provide different results, which will be investigated with more details in Sec. 11.4. Last, the transformed dynamics will be obtained in Sec. 11.5.

Several aspects of the full lifting considered here may be quite confusing and counterintuitive at the beginning. In order to give a clear understanding of the mechanisms at work and to verify the substance of the method, we scrutinize the subtleties involved throughout the chapter, e.g. the ones caused by releasing the simplifying assumption of the magnetic-moment-like lifting or the differences compared to a coordinate change for functions. This results in a rather detailed analysis. Perhaps some readers might think we insist too much on some points, but we feel it is useful to make things clear for other readers less familiar with the formalism and its associated subtleties.

## 11.1 Framework, operators and meta-operators

The framework is a generic Hamiltonian field dynamics. The field coordinate system $\boldsymbol{\psi}$ may be vectorial with components $\psi^\alpha$. It is defined over a coordinate space $\mathbf{z}$, which we will sometimes call the "particle space", by analogy with the particle phase space over which the Vlasov density is defined. Note that actually this space does not need to correspond to a particle space, i.e. there is no need to exist a dynamics over this space, whereas in the case of the Vlasov density, the coordinate space $\mathbf{z}$ really corresponds to a particle coordinate, it has a Hamiltonian dynamics, whose Lie-Poisson bracket is just the Vlasov part of the Vlasov-Maxwell Poisson bracket, and whose trajectories just correspond to the characteristics of the Vlasov density.

The phase space is the set of functionals of the fields $\boldsymbol{\psi}$. For any functional $F[\boldsymbol{\psi}]$, the dynamics is given by Hamilton's equations

$$\dot{F} := \{F, H\} ,$$

where the functional $H[\boldsymbol{\psi}] = \int d\mathbf{z}\ h[\mathbf{z}; \boldsymbol{\psi}]$ is given, it is the Hamiltonian functional.

The Poisson bracket $\{F, G\}$ acts on two arbitrary functionals $F[\boldsymbol{\psi}]$ and $G[\boldsymbol{\psi}]$ through formula

$$\{F, G\} := \int d\mathbf{z}\ F_{\psi^\alpha} \mathbb{J}^{\alpha\beta} G_{\psi^\beta} ,$$

where the operator-valued matrix $\mathbb{J}^{\alpha\beta}$ is given and we use the index notation for functional derivatives: $F_{\boldsymbol{\psi}} := \frac{\delta F}{\delta \boldsymbol{\psi}}$. By definition, the Poisson bracket is bilinear, antisymmetric, verifies the Leibniz rule, and satisfies the Jacobi identity.

For the sake of generality, we consider an arbitrary Poisson bracket. Readers interested in examples of such Hamiltonian systems can see [104, 107, 159] for instance (or other chapters in this document, e.g. Chapters 4, 7, or 14). For simplicity, we assume local interactions (i.e. $\mathbb{J}^{\alpha\beta}$ is an operator-valued matrix evaluated at the very point $\mathbf{z}$ implied in $F_{\psi^\alpha}$ and $G_{\psi^\beta}$), as is most common in field theory and especially in fluid dynamics. It makes the scheme simpler, but it is not essential, and the generalization of the results to non-local interactions is straightforward.

In addition, we study here the lifting mechanism in details assuming that all the fields $\psi^\alpha$ are defined over the same space $\mathbf{z}$, but the generalization to the case where this assumption does not hold is rather straightforward, although it makes formulae somehow more involved, as will appear in the Chapter 5.

The lifting starts from a coordinate transformation in the particle space $\mathbf{z}$:

$$\mathbf{z} \longrightarrow \tau(\mathbf{z}) := \bar{\mathbf{z}} := \bar{\mathbf{z}}[\mathbf{z};\boldsymbol{\psi}] = \bar{\mathbf{z}}\Big(\mathbf{z};\boldsymbol{\psi}(\mathbf{z}),\nabla\boldsymbol{\psi}(\mathbf{z}),\nabla\nabla\boldsymbol{\psi}(\mathbf{z}),...\Big)\,,$$

where the notation $\bar{\mathbf{z}}[\mathbf{z};\boldsymbol{\psi}]$ means that $\bar{\mathbf{z}}$ is not just a function, because on $\boldsymbol{\psi}$ it acts as an operator, with possibly derivatives at all orders, and the operator is evaluated at the position $\mathbf{z}$.

By duality, this induces the transformation for functions and for the fields (scalar invariance property):

$$\boldsymbol{\psi}(\mathbf{z}) \longrightarrow T^{-1}[\boldsymbol{\psi}](\bar{\mathbf{z}}) := \bar{\boldsymbol{\psi}}(\bar{\mathbf{z}}) := \boldsymbol{\psi}(\tau^{-1}(\bar{\mathbf{z}})) = \boldsymbol{\psi}(\mathbf{z})\,. \tag{11.1}$$

The operator $T^{-1}$ is called pull-back of $\tau^{-1}$, or sometimes push-forward of $\tau$.

Following the common convention for guiding-center and gyrokinetic theories, our operator $T$ is defined by $T\bar{f} = \bar{f}\circ\tau$, whereas another convention can be found $Tf = f\circ\tau^{-1}$, which would have the advantage that $T$ would be forwards and $T^{-1}$ would be backwards, as the intuition would expect, and the composition law would also be more intuitive.

Last, the coordinate change induces a transformation for functionals of $\boldsymbol{\psi}$ by duality:

$$F[\boldsymbol{\psi}] \longrightarrow \big(\mathbb{T}F\big)[\bar{\boldsymbol{\psi}}] := \bar{F}[\bar{\boldsymbol{\psi}}] := F[T\bar{\boldsymbol{\psi}}] = F[\boldsymbol{\psi}]\,. \tag{11.2}$$

The goal is to identify the transformed Hamiltonian structure, i.e. the transformed Hamiltonian functional $\bar{H}[\bar{\boldsymbol{\psi}}]$ and the transformed Poisson bracket $\bar{\{\cdot,\cdot\}}$, such that the transformed dynamics writes

$$\dot{\bar{F}} := \bar{\{}\bar{F},\bar{H}\bar{\}}\,,$$

where the transformed bracket acts on two transformed functionals $\bar{F}$ and $\bar{G}$.

## 11.2 Transformed Hamiltonian functional and Poisson bracket

The Hamiltonian functional transforms just as other functionals, so that it is given by

$$\bar{H}[\bar{\boldsymbol{\psi}}] := H[\boldsymbol{\psi}] = \int d\mathbf{z}\; h[\mathbf{z};\boldsymbol{\psi}] = \int \mathcal{J}d\bar{\mathbf{z}}\; h[\mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]);\bar{\boldsymbol{\psi}}]\,, \tag{11.3}$$

where the inverse transformation $\mathbf{z} = \mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]$ is used, as well as the scalar invariance property, and $\mathcal{J}$ is the Jacobian determinant of the transformation $\mathbf{z} \longrightarrow \bar{\mathbf{z}}$.

The Poisson bracket is functional-valued, and also transforms as a functional. However, it involves functional derivatives, so that each of its parts must be transformed separately.

$$\begin{aligned}\{F,G\}[\boldsymbol{\psi}] &= \{F,G\}[T\bar{\boldsymbol{\psi}}] = \mathbb{T}\Big(\{F,G\}\Big)[\bar{\boldsymbol{\psi}}] = \mathbb{T}\Big(\{\mathbb{T}^{-1}\bar{F},\mathbb{T}^{-1}\bar{G}\}\Big)[\bar{\boldsymbol{\psi}}]\\ &= \mathbb{T}\int d\mathbf{z}\quad \Big(\tfrac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F})\Big)\;\; \mathbb{J}^{\alpha\beta}\;\; \Big(\tfrac{\delta}{\delta\psi^\beta}\mathbb{T}^{-1}\bar{G}\Big)\\ &= \int \mathcal{J}d\bar{\mathbf{z}}\;\; \Big(T^{-1}\tfrac{\delta}{\delta\bar{\psi}^\alpha}\mathbb{T}^{-1}(\bar{F})\Big)\;\; \Big(T^{-1}\mathbb{J}^{\alpha\beta}\Big)\;\; \Big(T^{-1}\tfrac{\delta}{\delta\bar{\psi}^\beta}(\mathbb{T}^{-1}\bar{G})\Big)\,,\end{aligned} \tag{11.4}$$

where $T^{-1}\big(\mathbb{J}^{\alpha\beta}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]\big) = \mathbb{J}^{\alpha\beta}[\mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}];\bar{\boldsymbol{\psi}}]$ just undergoes a change of coordinates, using the transformation and the scalar invariance property.

The transformed bracket is Hamiltonian, since it is just given by an invertible transformation from the initial Poisson bracket. Especially, it is bilinear, antisymmetric verifies the Leibniz rule and satisfies the Jacobi identity. The first three properties are obvious to see, the last one is easily

verified using that the initial bracket is known to satisfy the Jacobi identity: for three arbitrary functionals $\bar{F}$, $\bar{G}$, and $\bar{H}$ (of $\bar{\psi}$), we have

$$\begin{aligned}\sum_{cycl} \bar{\{}\bar{F}, \bar{\{}\bar{G}, \bar{H}\bar{\}}\bar{\}} &= \sum_{cycl} \mathbb{T}\Big\{\mathbb{T}^{-1}\bar{F}, \mathbb{T}^{-1}\mathbb{T}\{\mathbb{T}^{-1}\bar{G}, \mathbb{T}^{-1}\bar{H}\}\Big\} \\ &= \mathbb{T}\sum_{cycl}\{F, \{G, H\}\} = 0\,,\end{aligned}$$

where $\sum_{cycl}$ means sum over cyclic permutations of the arguments, e.g. $\bar{F}$, $\bar{G}$ and $\bar{H}$. In this computation, the symbol $F$ is considered as just a shorthand for $\mathbb{T}^{-1}\bar{F}$, and the same for $G$ and $H$.

## 11.3 Functional chain-rule formula

It should be noticed that $J^{\alpha\beta}$ generically involves operators $\partial_{\mathbf{z}}$, and the functional density $h$ may involve derivatives $\partial_{\mathbf{z}}$ as well; in the transformed formulae, those operators are to be transformed according to the usual chain rule on functions (see Chapter 4):

$$\begin{aligned}T^{-1}\big(\partial_{\mathbf{z}} f\big) &= T^{-1}\partial_{\mathbf{z}}\big(TT^{-1}f\big) = T^{-1}\partial_{\mathbf{z}}\big(T\bar{f}\big) = T^{-1}\partial_{\mathbf{z}}\big[\bar{f}(\tau\mathbf{z})\big] \\ &= T^{-1}\partial_{\mathbf{z}}\big[\bar{f}(\bar{\mathbf{z}}(\mathbf{z}))\big] = T^{-1}\big(\partial_{\mathbf{z}}\bar{\mathbf{z}}\big)\cdot\partial_{\bar{\mathbf{z}}}\bar{f}\,,\end{aligned} \tag{11.5}$$

for any function $f(\mathbf{z})$. The operator $T^{-1}$ remaining in the last line just means that $\partial_{\mathbf{z}}\bar{\mathbf{z}}$ must be expressed in the transformed coordinates $\bar{\mathbf{z}}$, whereas initially it comes as a function of the initial coordinate $\mathbf{z}$. It is often considered as implicit.

Now, all that remains is to transform the functional derivatives, i.e. to write $T^{-1}\frac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F})$ as a function of $\bar{F}_{\bar{\boldsymbol{\psi}}}$. In this way, the transformed Poisson bracket will have the same form as the initial Poisson bracket:

$$\bar{\{}\bar{F}, \bar{G}\bar{\}}[\bar{\boldsymbol{\psi}}] = \int d\bar{\mathbf{z}}\ \bar{F}_{\bar{\psi}^\alpha}\ \ \bar{\mathbb{J}}^{\alpha\beta}\ \ \bar{G}_{\bar{\psi}^\beta}\,, \tag{11.6}$$

for some matrix $\bar{\mathbb{J}}^{\alpha\beta}$, so that the Hamiltonian structure will have been transformed properly.

This is obtained by a functional chain-rule formula relating $F_{\boldsymbol{\psi}}$ and $F_{\bar{\boldsymbol{\psi}}}$. The presence of the two chain rules, the first one for functions and the second one for functionals is well emphasized in Chapter 4; in this work, the function chain rule was all the more important as the particle coordinates defined a Hamiltonian system involved in the definition of the Vlasov bracket, and the function chain rule was needed to transform the particle Poisson bracket.

It turns out that the chain rule for functionals involves some subtleties. It is the reason why Chapter 4 studied a restricted class of transformations. In this section, we show that using a method similar to Chapter 4, a general chain-rule formula can be obtained without restricting the class of transformation. In the next section, we will come back to the transformed functional derivative $T^{-1}\frac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F})$ and show that it agrees with this chain-rule formula.

This section is organized in three subsections. In the first subsection, for the sake of clarity and consistency with previous works, we show that when the restriction of Chapter 4 is released, a bare chain rule (composition of derivative operators) cannot be used directly. In the second subsection, we show that the definition of the transformed fields all the same allows for a chain-rule formula, in the same way as in Chapter 4. Last, in the third subsection, we come back to the initial difficulties in the light of the results obtained.

### 11.3.1 Difficulty for a bare chain rule

At first, because the transformation is just a change of variable also in the field coordinates, one can consider applying directly the standard functional chain rule

$$F_{\psi^\alpha} = \Big(\bar{\psi}^\beta_{\psi^\alpha}\Big)^\dagger\,\mathcal{J}^{-1}\bar{F}_{\bar{\psi}^\beta}\,. \tag{11.7}$$

One could think that the definition (11.2) for $\bar{\boldsymbol{\psi}}$ implies that $\bar{\psi}^\beta_{\psi^\alpha} = 1^\beta_\alpha$, but this is not correct, because $\left(\bar{\psi}^\beta_{\psi^\alpha}\right)^\dagger$ is actually $\frac{\delta\bar{\psi}^\beta(\bar{\mathbf{z}})}{\delta\psi^\alpha(\mathbf{z})}$, i.e. the variation of $\bar{\boldsymbol{\psi}}$ keeping $\bar{\mathbf{z}}$ constant as a function of the variation of $\boldsymbol{\psi}$ keeping $\mathbf{z}$ constant. The issue is that the definition (11.2) for $\bar{\boldsymbol{\psi}}$ can be written in a more explicit way

$$\psi(\mathbf{z}) = \bar{\psi}(\bar{\mathbf{z}}) = \bar{\psi}(\bar{\mathbf{z}}[\mathbf{z}, \psi])\,, \tag{11.8}$$

which shows that if $\boldsymbol{\psi}$ is varied with $\mathbf{z}$ constant, then $\bar{\mathbf{z}}$ is not constant.

From another point of view, one can consider the reciprocal relation

$$\bar{\psi}(\bar{\mathbf{z}}) = \psi(\mathbf{z}) = \psi(\mathbf{z}[\bar{\mathbf{z}}, \bar{\psi}])\,, \tag{11.9}$$

but then the issue is that Eq. (11.9) is not explicit for $\bar{\boldsymbol{\psi}}$, since it is still present in the right-hand side. So, an implicit function is involved in this relation.

All these troubles were avoided in the work of Chapter 4 because only the velocity coordinate was transformed, and the transformation depended only on the electromagnetic field, which is independent of the velocity. Thus, the definition of the transformed fields implied that

$$\begin{aligned} \bar{f}(\mathbf{q}, \bar{\mathbf{p}}) &= f(\mathbf{q}, \mathbf{p}) = f\Big(\mathbf{q}, \mathbf{p}[\mathbf{q}, \bar{\mathbf{p}}; \bar{\mathbf{E}}, \bar{\mathbf{B}}]\Big) = f\Big(\mathbf{q}, \mathbf{p}[\mathbf{q}, \bar{\mathbf{p}}; \mathbf{E}, \mathbf{B}]\Big)\,, \\ \bar{\mathbf{E}}(\mathbf{q}) &= \mathbf{E}(\mathbf{q})\,, \\ \bar{\mathbf{B}}(\mathbf{q}) &= \mathbf{B}(\mathbf{q})\,, \end{aligned} \tag{11.10}$$

and no transformed field was left in the right-hand side. Then the chain-rule formula $\bar{\boldsymbol{\psi}}_{\boldsymbol{\psi}}$ could be used directly, which made the setting simpler for the lifting.

Let us add some comments about Eq. (11.8). When the transformation is just a coordinate change for $\mathbf{z}$ (it could alternatively be viewed as a transformation between different spaces), then from the point of view of the coordinate space, the inverse of the transformation $\bar{\mathbf{z}}[\mathbf{z}; \boldsymbol{\psi}] = \bar{\mathbf{z}}(\mathbf{z})[\boldsymbol{\psi}]$ is given by the inverse function $\mathbf{z}(\bar{\mathbf{z}})[\boldsymbol{\psi}] = \mathbf{z}[\bar{\mathbf{z}}; \boldsymbol{\psi}]$, where the field is $\boldsymbol{\psi}$, not $\bar{\boldsymbol{\psi}}$. Then the field $\bar{\boldsymbol{\psi}}$ can be defined explicitly as $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}}) = \boldsymbol{\psi}(\mathbf{z}) = \boldsymbol{\psi}(\mathbf{z}[\bar{\mathbf{z}}; \boldsymbol{\psi}])$.

However, this is not satisfactory because in $\mathbf{z}[\bar{\mathbf{z}}; \boldsymbol{\psi}]$, the function $\boldsymbol{\psi}$ is evaluated at $\bar{\mathbf{z}}$ whereas the operators acting on it are still the initial ones $\partial_{\mathbf{z}}$. In addition, in order for the transformation to be completely consistent (reversible) even at the level of the fields, the reverse transformation should be defined reciprocally compared to the direct one, i.e. the field should be the transformed one $\bar{\boldsymbol{\psi}}$. So, one should express everything with the transformed quantities: the operator $\partial_{\mathbf{z}}$ must become a function of $\partial_{\bar{\mathbf{z}}}$, and the field $\boldsymbol{\psi}$ a function of $\bar{\boldsymbol{\psi}}$.

For the operator, the answer is easily provided by the chain rule $\partial_{\mathbf{z}} = \partial_{\mathbf{z}}\bar{\mathbf{z}} \cdot \partial_{\bar{\mathbf{z}}}$. For the fields, things are more subtle and it is here that obstructions can come. As an illustrative example, let us consider the transformation

$$\bar{\mathbf{z}} := f(\mathbf{z} + \varepsilon\boldsymbol{\psi}(\mathbf{z}))\,,$$

with $f$ some invertible function (independent of $\boldsymbol{\psi}$) and $\varepsilon$ some small parameter; note that the formula $\mathbf{z} + \varepsilon\boldsymbol{\psi}(\mathbf{z})$ assumes that in the example the field dimension is the same as the dimension of the coordinate space. The transformation can be written $\bar{\mathbf{z}} = (f \circ f_2)(\mathbf{z})$, with $f_2 := 1 + \varepsilon\boldsymbol{\psi}$, which is invertible, since it is a small perturbation of an invertible function. Then the reverse transformation is given by

$$\mathbf{z} := (f_2^{-1} \circ f^{-1})(\bar{\mathbf{z}}) = \big((1 + \varepsilon\boldsymbol{\psi})^{-1} \circ f^{-1}\big)(\bar{\mathbf{z}}) = \left(\Big[\sum_n (-\varepsilon\boldsymbol{\psi})^n\Big] \circ f^{-1}\right)(\bar{\mathbf{z}})\,.$$

Thus, the initial field $\boldsymbol{\psi}$ is evaluated at a point related to the transformed position $\bar{\mathbf{z}}$ (e.g. take $f = 1$). This mixes up initial and transformed objects.

For completeness, one needs the transformation to be revertible also for fields, so that $\boldsymbol{\psi}$ can be replaced by $\bar{\boldsymbol{\psi}}$. The point is that there is no guarantee in general for such a relation to exist,

and this can induce an obstruction to the lifting. It can happen that the transformation is proper (reversible) for coordinates but not for fields.

For instance, in order to establish a relation between $\boldsymbol{\psi}$ and $\bar{\boldsymbol{\psi}}$, one could take Eq. (11.8) and relate $\bar{\mathbf{z}}$ to $\mathbf{z}$, by replacing $\bar{\mathbf{z}}$ by $\mathbf{z}+\mathbf{x}$, with $\mathbf{x} := \bar{\mathbf{z}} - \mathbf{z}$. Then Taylor expanding $\bar{\boldsymbol{\psi}}$ in series in $\mathbf{x}$ at the position $\mathbf{z}$ would provide a relation such as

$$\boldsymbol{\psi}(\mathbf{z}) = \bar{\boldsymbol{\psi}}(\mathbf{z}) + g_1(\mathbf{x})[\bar{\boldsymbol{\psi}}] = \bar{\boldsymbol{\psi}}(\mathbf{z}) + g_1(\bar{\mathbf{z}} - \mathbf{z})[\bar{\boldsymbol{\psi}}]\,,$$

where $g_1$ is some function given by the expansion. In order to get rid of $\bar{\mathbf{z}}$, one could iterate the process to infinity, which would provide a relation such as

$$\boldsymbol{\psi}(\mathbf{z}) = \bar{\boldsymbol{\psi}}(\mathbf{z}) + g_\infty(\mathbf{z})[\bar{\boldsymbol{\psi}}]\,.$$

This would provide an explicit relation for $\boldsymbol{\psi}$ as a function of $\bar{\boldsymbol{\psi}}$. However, such a procedure is not valid in general, because it implicitly assumes that the contribution of $\bar{\mathbf{z}}$ diminishes at each iteration, which is not guaranteed. In addition, the Taylor expansion assumes that $\bar{\mathbf{z}}$ is close enough to $\mathbf{z}$. From a more abstract point of view, a kind of self-consistency is assumed for the fields between the position $\mathbf{z}$ and the position $\bar{\mathbf{z}}$.

All the same, for a formal near-identity transformation (such as the example above), which are most common in perturbation theory, and especially in guiding-center theory and gyrokinetics, the procedure proposed above works, and there is no obstruction to the lifting.

### 11.3.2 Derivation of a chain-rule formula

To circumvent the issue mentioned in the previous subsection, the idea is to start from the defining relation (11.8), and to study how it relates the variations of $\boldsymbol{\psi}(\mathbf{z})$ and $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$, in a similar way as what was done in Chapter 4. In the variations, one has to choose which space is considered as fixed, since $\mathbf{z}$ and $\bar{\mathbf{z}}$ cannot be kept fixed simultaneously. Here, we choose the initial coordinate space $\mathbf{z}$ as fixed, which implies a displacement of $\bar{\mathbf{z}}$ in the process.

Performing total variation of Eq. (11.8) results in

$$\delta\psi^\alpha(\mathbf{z}) = \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) + \bar{\psi}^\alpha_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\psi^\beta}\delta\psi^\beta(\mathbf{z})\,, \tag{11.11}$$

where we remind that the index notation is used for any kind of derivatives, i.e. first-order variations, and especially for derivatives of function(e.g. $\bar{\psi}_{\bar{\mathbf{z}}}$), for derivatives of functionals (e.g. $H_\psi$), and for derivatives of operator-function (e.g. $\bar{\mathbf{z}}_\psi$). In this third case, the result is an operator in general; for an explicit example, see page 133.

Now, Eq. (11.11) implies

$$\left(1^\alpha_\beta - \bar{\psi}^\alpha_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\psi^\beta}\right)\delta\psi^\beta(\mathbf{z}) = \delta\bar{\psi}^\alpha(\bar{\mathbf{z}})\,, \tag{11.12}$$

which in turn gives the chain-rule formula

$$\frac{\delta\bar{\psi}^\alpha(\bar{\mathbf{z}})}{\delta\psi^\beta(\mathbf{z}')} = \left(1^\alpha_\beta - \bar{\psi}^\alpha_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\psi^\beta}\right)^\dagger \delta(\mathbf{z}' - \mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}])\,, \tag{11.13}$$

where we wrote $\mathbf{z}'$ to emphasize that $\mathbf{z}'$ is an independent variable, whereas $\mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]$ is not: it is a function of the independent variable $\bar{\mathbf{z}}$. The adjoint $\dagger$ is with respect to $d\mathbf{z}$. In all these formulae notice that $\bar{\mathbf{z}}_{\psi^\beta}$ is generically a differential operator. For instance this implies that the order of the factors cannot be changed in the expressions with $\bar{\psi}^\alpha_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\psi^\beta}$ in Eqs. (11.11)-(11.13).

Using the language of transformation operators gives a more abstract point of view on the

mechanism at work:

$$\begin{aligned}
\delta\psi^\alpha(\mathbf{z}) &= \delta\left(T\bar{\psi}^\alpha\right)(\mathbf{z}) \\
&= T\delta\bar{\psi}^\alpha(\mathbf{z}) + \left([\delta T]\bar{\psi}^\alpha\right)(\mathbf{z}) \\
&= T\delta\bar{\psi}^\alpha(\mathbf{z}) + \Big(\big(T[\boldsymbol{\psi}+\delta\boldsymbol{\psi}] - T[\boldsymbol{\psi}]\big)\bar{\psi}^\alpha\Big)(\mathbf{z}) \\
&= \delta\bar{\psi}^\alpha(\tau\mathbf{z}) + \bar{\psi}^\alpha(\tau[\mathbf{z};\boldsymbol{\psi}+\delta\boldsymbol{\psi}]) - \bar{\psi}^\alpha(\tau[\mathbf{z};\boldsymbol{\psi}]) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) + \bar{\psi}^\alpha(\bar{\mathbf{z}}[\mathbf{z};\boldsymbol{\psi}+\delta\boldsymbol{\psi}]) - \bar{\psi}^\alpha(\bar{\mathbf{z}}[\mathbf{z};\boldsymbol{\psi}]) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) + \bar{\psi}^\alpha(\bar{\mathbf{z}} + \bar{\mathbf{z}}_{\psi^\beta}\delta\psi^\beta) - \bar{\psi}^\alpha(\bar{\mathbf{z}}) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) + \bar{\psi}^\alpha_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\beta}\delta\psi^\beta(\mathbf{z})\,,
\end{aligned} \tag{11.14}$$

where the equalities are only up to first order. In the first and second equalities, we used the definition (11.2) and the linearity of the operator $T$. In the fourth equality, we used the property (11.17) between the operators $T$ and $\tau$. In the fifth equality, the definition (11.1) of the operator $\tau$ was used. Finally, in the third and last two equalities, infinitesimal variations were made explicit and a first order expansion was performed.

The result (11.14) exactly agrees with Eq. (11.11). In these formulae, the right-hand side involves both $\delta\boldsymbol{\psi}$ and $\delta\bar{\boldsymbol{\psi}}$. But the use of operators provides an efficient way to remedy this by using the inverse operator, whose variations obey the property

$$0 = \delta(TT^{-1}) = \delta T T^{-1} + T\delta T^{-1} \Rightarrow \delta T = -T\delta T^{-1}T\,.$$

Applying this property to the second line of Eq. (11.14) gives

$$\begin{aligned}
\delta\psi^\alpha(\mathbf{z}) &= T\delta\bar{\psi}^\alpha(\mathbf{z}) - \left(T\delta T^{-1}T\bar{\psi}^\alpha\right)(\mathbf{z}) \\
&= \delta\bar{\psi}^\alpha(\tau\mathbf{z}) - \left(\delta T^{-1}\psi^\alpha\right)(\tau\mathbf{z}) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) - \Big(\big(T^{-1}[\bar{\boldsymbol{\psi}}+\delta\bar{\boldsymbol{\psi}}] - T^{-1}[\bar{\boldsymbol{\psi}}]\big)\psi^\alpha\Big)(\bar{\mathbf{z}}) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) - \psi^\alpha(\tau^{-1}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}+\delta\bar{\boldsymbol{\psi}}]) + \psi^\alpha(\tau^{-1}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) - \psi^\alpha(\mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}+\delta\bar{\boldsymbol{\psi}}]) + \psi^\alpha(\mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) - \psi^\alpha(\mathbf{z} + \mathbf{z}_{\bar{\psi}^\beta}\delta\bar{\psi}^\beta]) + \psi^\alpha(\mathbf{z}) \\
&= \delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) - \psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}\delta\bar{\psi}^\beta(\bar{\mathbf{z}})\,,
\end{aligned} \tag{11.15}$$

where the derivation followed the development of Eq. (11.14). The result exactly agrees with (the reciprocal of) Eq. (11.12).

With Eqs. (11.12)-(11.13), the chain-rule formula follows from Eq. (11.7). Alternatively, it can be obtained by using the standard method of total variation of an arbitrary functional, in a similar way as in Chapter 4:

$$\begin{aligned}
\delta F &= \int d\mathbf{z}\; F_{\psi^\alpha}\delta\psi^\alpha = \int d\mathbf{z}\; F_{\psi^\alpha}\left(1^\alpha_\beta - \psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}\right)\delta\bar{\psi}^\beta(\bar{\mathbf{z}}) \\
&= \int \mathcal{J}d\bar{\mathbf{z}}\; F_{\psi^\alpha}\left(1^\alpha_\beta - \psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}\right)\delta\bar{\psi}^\beta(\bar{\mathbf{z}}) \\
&= \int d\bar{\mathbf{z}}\; \delta\bar{\psi}^\beta(\bar{\mathbf{z}})\left(1^\alpha_\beta - \psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}\right)^{\bar{\dagger}}\mathcal{J}F_{\psi^\alpha} \\
&= \int d\bar{\mathbf{z}}\; \bar{F}_{\bar{\psi}^\alpha}\delta\bar{\psi}^\alpha = \delta\bar{F}\,,
\end{aligned}$$

where the reciprocal of Eq. (11.12) was used in the second equality, the change of integration variable was performed in the third equality, and in the fourth equality the adjoint $\bar{\dagger}$ is computed with respect to $d\bar{\mathbf{z}}$. So, the chain-rule formula writes

$$\bar{F}_{\bar{\psi}^\alpha} = \left(1^\beta_\alpha - \psi^\beta_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\alpha}\right)^{\bar{\dagger}}\mathcal{J}F_{\psi^\beta}\,,$$

which agrees with (11.7). Notice that in this relation $F_{\psi^\alpha}$ is actually $T^{-1}F_{\psi^\alpha}$, in a similar way as for the function chain rule: when expressing $\bar{f}_{\bar{\mathbf{z}}}$ as a function of $f_{\mathbf{z}}$, the traditional formula is $f_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\mathbf{z}}}$, but it actually means $T^{-1}\big(f_{\mathbf{z}}\big)\cdot\mathbf{z}_{\bar{\mathbf{z}}}$ (as also appears in Eq. (11.5)). The reciprocal relation writes

$$F_{\psi^\alpha} = \Big(1^\beta_\alpha - \bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\Big)^\dagger \mathcal{J}^{-1}\bar{F}_{\bar{\psi}^\beta} \tag{11.16}$$

and it is the one that gives the transformed Poisson bracket, as we will see in the next section.

### 11.3.3 Interpretation: implicit function and moving point

In the light of the chain-rule formula obtained above, the initial issue about the bare chain rule can be interpreted in two directions.

First, the implicit function involved in Eq. (11.8) prevented a straightforward evaluation of $\frac{\delta\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})}{\delta\boldsymbol{\psi}(\mathbf{z})}$; in the derivation, it is traduced by the presence of both $\delta\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$ and $\delta\boldsymbol{\psi}(\mathbf{z})$ in the right-hand side of the total variation (11.11) of the defining formula (11.2), which does not happen when a definition is explicit. In the final chain rule, it is traduced by the minus sign in the right-hand side of (11.13), which means that in the total variations (11.11), one term appeared in the wrong side of the equation, or rather that the inverse transformation was involved in the derivation (11.15), as expected for an implicit function.

Another way of tackling the definition (11.1) (or rather its detailed meaning (11.9)) is to make it explicit by a perturbative procedure. When the transformation is near-identity (and the fields regular enough), then at lowest order $\mathbf{z}\approx\bar{\mathbf{z}}$ and $\boldsymbol{\psi}(\mathbf{z})\approx\boldsymbol{\psi}(\bar{\mathbf{z}})$, which implies that $\boldsymbol{\psi}(\mathbf{z})=\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$ is close to $\boldsymbol{\psi}(\bar{\mathbf{z}})$. Then, to lowest order, one can consider that $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}}) := \boldsymbol{\psi}(\mathbf{z}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}])\approx\boldsymbol{\psi}(\mathbf{z}[\bar{\mathbf{z}};\boldsymbol{\psi}])$, i.e. $\bar{\boldsymbol{\psi}} := \boldsymbol{\psi}\circ\tau^{-1}_{[\bar{\boldsymbol{\psi}}]}\approx\boldsymbol{\psi}\circ\tau^{-1}_{[\boldsymbol{\psi}]}$, which makes the definition explicit. Then, the bare chain rule can be applied, and gives

$$\delta\bar{\psi}^\alpha(\bar{\mathbf{z}})\approx\delta\psi^\alpha(\mathbf{z}[\bar{\mathbf{z}};\boldsymbol{\psi})+\psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}\delta\psi^\beta(\bar{\mathbf{z}})\,.$$

Now, this identity is valid only at lowest order. To first order, the lowest order approximation must be inserted in the place of the index $\bar{\boldsymbol{\psi}}$, which means that $\bar{\boldsymbol{\psi}} := \boldsymbol{\psi}\circ\tau^{-1}_{[\bar{\boldsymbol{\psi}}]}\approx\boldsymbol{\psi}\circ\tau^{-1}_{\left[\boldsymbol{\psi}\circ\tau^{-1}_{[\boldsymbol{\psi}]}\right]}$. Iterating the process to infinity yields

$$\delta\bar{\psi}^\alpha(\bar{\mathbf{z}}) = \sum_{n=0}^{\infty}\Big(\psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}\Big)^n\delta\psi^\beta(\mathbf{z}) = \Big(1-\psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}\Big)^{-1}\delta\psi^\beta(\mathbf{z})\,,$$

where we assumed the transformation is close enough to identity so that both the procedure and the series converge. Especially, we assumed $\psi^\alpha_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\beta}<<1$. The previous relation is more conveniently written

$$\delta\psi^\beta(\mathbf{z}) = \Big(1-\psi^\beta_{\mathbf{z}}\cdot\mathbf{z}_{\bar{\psi}^\alpha}\Big)\delta\bar{\psi}^\beta(\bar{\mathbf{z}})\,.$$

This gives exactly the general formula (11.12), valid even when the transformation is not near-identity, although the intermediate step of the argument above may not be meaningful if the series does not converge. This development clearly shows that the definition (11.9) for $\bar{\boldsymbol{\psi}}$ is not explicit, it defines $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$ as a function of $\boldsymbol{\psi}$ and $\bar{\mathbf{z}}$ as expected, but $\bar{\boldsymbol{\psi}}$ itself is involved in the relation, as is needed for a relation where the coordinate is $\bar{\mathbf{z}}$.

Second, since the transformation starts from the initial field theory, where the space $\mathbf{z}$ is fixed, another point of view is not the implicit function, but rather the displacement of $\bar{\mathbf{z}}$ when varying $\boldsymbol{\psi}$ with $\mathbf{z}$ constant. In the derivation, it is obvious that the additional term in (11.11) corresponds to a displacement of the observation point $\bar{\mathbf{z}}$. In the resulting chain rule (11.12) or (11.13), the ratio $\frac{\delta\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})}{\delta\boldsymbol{\psi}(\mathbf{z})}$ is not 1, and the correction term is such that it just cancels the displacement of $\bar{\mathbf{z}}$.

Indeed, when field-varying $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}}[\mathbf{z};\boldsymbol{\psi}])$, two effects are present: the variation of $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$ at constant $\mathbf{z}$, which is equal to the variation of $\boldsymbol{\psi}(\mathbf{z})$ as a result of (11.2), but in addition, the point $\bar{\mathbf{z}}$ is moved, the induced variation of $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$ is just given by $\bar{\boldsymbol{\psi}}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\boldsymbol{\psi}}$. As a consequence, the variation of $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$)

at constant $\bar{\mathbf{z}}$ is just obtained by removing this excess contribution. This is a common feature in transformations for field theory, see e.g. [66, 152] among many others.

This phenomenon is analogous to what happens when computing the term $\partial_t f$ for the dynamics of the Vlasov density: $f$ is advected by particle dynamics $\dot{\mathbf{z}}$, hence during the time evolution, the variation $\delta f(\mathbf{z})$ is zero provided the point $\mathbf{z}$ is advected with particle dynamics. To obtain the dynamics of $f$ at constant $\mathbf{z}$, one has to remove the contribution corresponding to the displacement of the advected $\mathbf{z}$, which is given by $dt(\dot{\mathbf{q}} \cdot \partial_{\mathbf{q}} f + \dot{\mathbf{p}} \cdot \partial_{\mathbf{p}} f)$. And minus this last quantity just gives the expected Vlasov equation, according to formula

$$\tfrac{\partial}{\partial t} f = \tfrac{d}{dt} f - \dot{\mathbf{z}} \cdot \partial_{\mathbf{z}} f = -\dot{\mathbf{z}} \cdot \partial_{\mathbf{z}} f \,.$$

Another way of stating it is to say that when $\mathbf{z}$ is a fixed point, then $\boldsymbol{\psi}(\mathbf{z})$ is an Eulerian field, whereas $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$ is a Lagrangian field, since the point $\bar{\mathbf{z}}$ is advected by the transformation.

The specific case treated in Chapter 4 corresponded to the case where all these subtleties were avoided: when the coordinate transformation does not change the position $\bar{\mathbf{q}} = \mathbf{q}$, and depends only of the electromagnetic field $\mathbf{E}(\mathbf{q}), \mathbf{B}(\mathbf{q})$, then the transformation for fields is just given by Eq. (11.10), which can equivalently be considered as:

$$\begin{aligned}
\bar{f}(\bar{\mathbf{z}}) &= f(\mathbf{z}) = f(\mathbf{z}[\bar{\mathbf{z}}; \bar{\mathbf{E}}, \bar{\mathbf{B}}]) = f(\mathbf{z}[\bar{\mathbf{z}}; \mathbf{E}, \mathbf{B}]) \,, \\
\bar{\mathbf{E}}(\bar{\mathbf{q}}) &= \mathbf{E}(\mathbf{q}) = \mathbf{E}(\bar{\mathbf{q}}) \,, \\
\bar{\mathbf{B}}(\bar{\mathbf{q}}) &= \mathbf{B}(\mathbf{q}) = \mathbf{B}(\bar{\mathbf{q}}) \,,
\end{aligned}$$

where all the definitions are explicit, and there is no issue about the moving (Lagrangian) point, because definitions involve the transformed point $\bar{\mathbf{z}}$, but the initial fields $\mathbf{E}, \mathbf{B}$. Because of this, the bare chain rule (11.7) could be used directly.

## 11.4 Transformed functional derivatives

To have the transformed bracket (11.6) at our disposal, it remains to compute $T^{-1} \frac{\delta}{\delta \psi^\alpha} \mathbb{T}^{-1}(\bar{F})$ involved in the transformed bracket (11.4). In this section, we show that it is just given by the chain-rule formula (11.16).

From an abstract point of view, this equality is trivial, but in practical computations, it is not so obvious. In a similar way as in the previous section, some subtleties are involved and have to be clarified, because this is a not a chain rule for functions but for functionals. The integration variable plays a role, unlike indices when a chain rule is applied to a function defined as the scalar product of two vectorial quantities.

Interestingly, the difference between the two chain rules is already evidenced by the expression for the transformed functional derivative $T^{-1} \frac{\delta}{\delta \psi^\alpha} \mathbb{T}^{-1}(\bar{F})$, because the transformation operators $T^{-1}$ and $\mathbb{T}^{-1}$ are not defined between the same spaces, whereas for a transformed function derivative the expression would be $T^{-1} \partial_{\mathbf{z}} T \bar{f}$ (see Eq. (11.5)), and the two operators $T^{-1}$ and $T$ are defined between the same spaces.

The difference can also be viewed in the definition of the transformation operators: the transformation for functions is just a composition with the particle transformation:

$$(T^{-1} f)(\bar{\mathbf{z}}) = f(\tau^{-1}(\bar{\mathbf{z}})) \quad \text{whence} \quad T^{-1} f = f \circ \tau^{-1} \,, \tag{11.17}$$

whereas rigorously there is no analogous definition for the transformation operator $\mathbb{T}$, because what is to be transformed is not a function, but a functional, i.e. an integral; so, two operations are contained in the operator $\mathbb{T}$: not only a change in the fields, but also a change of the integration variable; the point is that these two operations cannot be done separately, because the transformed fields are defined over the transformed space. More precisely, writing $F[\boldsymbol{\psi}] = \bar{F}[\bar{\boldsymbol{\psi}}]$ does not mean just $\mathbb{T}F = F \circ T$, because it also implies that the first functional $F$ is written as an integral over $d\mathbf{z}$ whereas the second one $\bar{F}$ is written as an integral over $d\bar{\mathbf{z}}$. This can be expressed by writing

$$\mathbb{T} \int d\mathbf{z} \; g[\mathbf{z}; \boldsymbol{\psi}] = \int T^{-1} d\mathbf{z} \; T^{-1} g[\mathbf{z}; \boldsymbol{\psi}] = \int d\bar{\mathbf{z}} \mathcal{J} \; g[\mathbf{z}[\bar{\mathbf{z}}; \bar{\boldsymbol{\psi}}]; \bar{\boldsymbol{\psi}}] \,,$$

where the operator $T^{-1}$ is used also to transform the measure $d\mathbf{z}$ and the density $g$.

This section is organized as follows. In the first subsection, the transformed functional derivative $T^{-1}\frac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F})$ is computed, but it comes out in an unsatisfactory form. In the second subsection, the origin of the difficulty is explored and an intermediate lemma is established. In the third subsection, the lemma is used to re-express the result of $T^{-1}\frac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F})$ and obtain it in a satisfactory form.

### 11.4.1 Transformed functional derivatives: unexpected form

In the case of function derivatives, the expression $T^{-1}\partial_{\mathbf{z}}T\bar{f}$ just gives the chain rule for functions (11.5), i.e. the same result as when computing the total derivative of the defining formula (11.1). On the contrary, computing the expression $T^{-1}\frac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F})$ gives a very different result from the one expected from the chain rule for functionals (11.16), obtained from the total derivative of the defining formula (11.2). This is what we are going to see in this subsection.

Indeed, changing the integration variable and then functionally differentiating the result implies a field derivative of the coordinate transformation and of the Jacobian:

$$\begin{aligned}
T^{-1}\frac{\delta}{\delta\psi^\alpha(\mathbf{y})}\mathbb{T}^{-1}(\bar{F}) &= T^{-1}\frac{\delta}{\delta\psi^\alpha(\mathbf{y})}\mathbb{T}^{-1}\left(\int d\bar{\mathbf{z}}\ \bar{f}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]\right)\\
&= T^{-1}\frac{\delta}{\delta\psi^\alpha(\mathbf{y})}\left(\int \mathcal{J}^{-1}d\mathbf{z}\ \bar{f}[\bar{\mathbf{z}}[\mathbf{z};\boldsymbol{\psi}];\boldsymbol{\psi}]\right)\\
&= T^{-1}\int d\mathbf{z}\left(\mathcal{J}^{-1}\bar{f}_{\psi^\alpha}+\mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}+\bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)\delta(\mathbf{z}-\mathbf{y})\\
&= T^{-1}\left(\mathcal{J}^{-1}\bar{f}_{\psi^\alpha}+\mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}+\bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)^\dagger[\mathbf{y};\boldsymbol{\psi}]\\
&= \left(\mathcal{J}^{-1}\bar{f}_{\psi^\alpha}+\mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}+\bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)^\dagger[\mathbf{y}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}];\bar{\boldsymbol{\psi}}]\,.
\end{aligned}\tag{11.18}$$

In the first equality, the functional was made explicit by introducing $\bar{f}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]$ as the density of the functional $\bar{F}$, which may involve derivatives of the fields $\bar{\boldsymbol{\psi}}$. In the second equality, the action of the operator $\mathbb{T}^{-1}$ was performed. In the third equality, the action of the functional derivative $\frac{\delta}{\delta\psi^\alpha}$ was computed. In the fourth equality, the Dirac delta was integrated; the adjoint $\dagger$ is related to $d\mathbf{z}$ and the symbol $[\mathbf{y};\boldsymbol{\psi}]$ was written to remind that the function is evaluated at $\mathbf{z}=\mathbf{y}$, and that it depends of $\boldsymbol{\psi}$ and possibly its derivatives. In the fifth equality, the action of the operator $T^{-1}$ was performed. Note that here the overall expression is a function, not an operator: the adjoint function involved in the last line is actually

$$\left(\mathcal{J}^{-1}\bar{f}_{\psi^\alpha}+\mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}+\bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)^\dagger 1\,,\tag{11.19}$$

i.e. the adjoint operator $\left(\mathcal{J}^{-1}\bar{f}_{\psi^\alpha}+\mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}+\bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)^\dagger$ acting on the constant function 1.

The result (11.18) is very different from what would be obtained using Eq. (11.16):

$$\begin{aligned}
T^{-1}\frac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F}) \overset{?}{=} F_{\psi^\alpha(\mathbf{y})} &= \left(1^\beta_\alpha-\bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\right)^\dagger\mathcal{J}^{-1}\bar{F}_{\bar{\psi}^\beta(\bar{\mathbf{z}}[\mathbf{y};\boldsymbol{\psi}])}\\
&= \left(1^\beta_\alpha-\bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\right)^\dagger\mathcal{J}^{-1}\frac{\delta}{\delta\bar{\psi}^\beta(\bar{\mathbf{z}}[\mathbf{y};\boldsymbol{\psi}])}\int d\bar{\mathbf{z}}\ \bar{f}[\bar{\mathbf{z}};\bar{\boldsymbol{\psi}}]\\
&= \left(1^\beta_\alpha-\bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\right)^\dagger\mathcal{J}^{-1}\bar{f}_{\bar{\psi}^\beta}[\bar{\mathbf{z}}[\mathbf{y};\boldsymbol{\psi}]];\boldsymbol{\psi}]\,.
\end{aligned}\tag{11.20}$$

The first term of Eqs. (11.18) and (11.20) agree, but the remaining terms are different. On another hand, Eq. (11.18) is not satisfactory, because it is not clear at all whether it is a functional derivative, i.e. whether it verifies the Leibniz rule, whereas it should, since it is just the transformed object from a functional derivative. A property must exist that justifies the Leibniz rule, i.e. that allows one to rewrite Eq. (11.18) under the form $\bar{F}_{\bar{\psi}^\alpha}=\mathbf{M}^\beta_\alpha F_{\psi^\beta}$, with $\mathbf{M}$ some matrix to be

identified. Note that it is a matrix in the infinite dimensional phase space of this field theory, which means it can involve integrals and differential operators. The desired property is established in the following subsections; it will clarify some differences between functional chain rule and function chain rule, and at the same time, it will show that actually (11.18) and (11.20) are equal.

### 11.4.2 Intermediate lemma

The failure for (11.18) to give an obvious derivative over the functional $\bar{F}$ comes from its terms where the functional derivative acts on the coordinate transformation $\bar{\mathbf{z}}_{\boldsymbol{\psi}}$ and on the Jacobian $\mathcal{J}_{\psi}^{-1}$. Precisely, those terms are surprising because the integration variable is mute for functionals: it is analogous to the mute index $i$ in the function $\sum_i \mathbf{V}^i(\mathbf{z})\mathbf{W}^i(\mathbf{z})$, where $V^i(\mathbf{z})$ and $W^i(\mathbf{z})$ are vectors. Indeed, the change of variable under the integral does not affect the value of the functional, and therefore should not have an effect on variations or derivations of the functional.

This is a special feature of the functional chain rule for a lifting: contrary to chain rules for functions, where the index $i$ is not affected by the transformation, here the integration variable, which is a kind of mute index for the fields, is affected by the transformation and it is involved in the computation of functional derivatives.

In order to avoid this trouble, one would need to keep the variable $\mathbf{z}$, but this is not possible when using the field $\bar{\psi}$, since this field is defined over the transformed space $\bar{\mathbf{z}}$. Thus, the two effects of the transformation (on $\psi$ and on $\mathbf{z}$) cannot be separated, as we announced above. This is reminiscent of the issues in the previous section. So, in a similar way, we aim here at isolating the effect of the change $\boldsymbol{\psi} \longrightarrow \bar{\boldsymbol{\psi}}$ from the change $\mathbf{z} \longrightarrow \bar{\mathbf{z}}$. The goal is to identify one of these effect, since afterwards the other one can be deduced by substraction from the total effect given by (11.18). This task is not so easy, because it means for instance studying the field $\bar{\boldsymbol{\psi}}$ while remaining in the space $\mathbf{z}$ without introducing the field $\boldsymbol{\psi}$.

A trick to get the resulting formula relies on functionals that are constant in the fields. This is a restricted class of functionals (indeed!) but a general formula can be obtained by keeping track of all the arbitrariness involved in the process.

Let us consider a functional constant in the field, which we denote by $\bar{G}$, in order to avoid confusion with the general functional $\bar{F}$ considered in the previous subsection:

$$\bar{G} = \int d\bar{\mathbf{z}}\ \bar{g}(\bar{\mathbf{z}})\,.$$

Its value is completely determined, it does not depend on $\bar{\boldsymbol{\psi}}$, nor of $\boldsymbol{\psi}$. Now, after changing to coordinates $\mathbf{z}$, a dependence in the field $\boldsymbol{\psi}$ seems to appear, but this overall contribution must be zero:

$$\begin{aligned}
\delta\bar{G} = 0 = \delta G &= \int d\mathbf{y} G_{\psi^\alpha(\mathbf{y})}\delta\psi^\alpha(\mathbf{y}) \\
&= \int d\mathbf{y} \tfrac{\delta}{\delta\psi^\alpha(\mathbf{y})}\left(\int \mathcal{J}^{-1} d\mathbf{z}\ \bar{g}(\bar{\mathbf{z}}[\mathbf{z};\boldsymbol{\psi}])\right)\delta\psi^\alpha(\mathbf{y}) \\
&= \int d\mathbf{y} \int d\mathbf{z}\ \left(\mathcal{J}^{-1}\bar{g}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \bar{g}\mathcal{J}_{\psi^\alpha}^{-1}\right)\delta(\mathbf{z}-\mathbf{y})\delta\psi^\alpha(\mathbf{y}) \\
&= \int d\mathbf{z}\ \left(\mathcal{J}^{-1}\bar{g}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \bar{g}\mathcal{J}_{\psi^\alpha}^{-1}\right)\delta\psi^\alpha(\mathbf{z}) \\
&= \int d\mathbf{z}\ \bar{g}\ \left(-\mathcal{J}^{-1}d_{\bar{\mathbf{z}}}\cdot\left(\bar{\mathbf{z}}_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z})\right) + \mathcal{J}_{\psi^\alpha}^{-1}\delta\psi^\alpha(\mathbf{z})\right),
\end{aligned} \tag{11.21}$$

where in the first four lines, we applied the same calculation as for Eq. (11.18), and in the last line, an integration by parts was performed to have $\bar{g}$ factorized.

In Eq. (11.21), we use the symbol $d_{\bar{\mathbf{z}}}$ to indicate the total derivative with respect to $\bar{\mathbf{z}}$, in order to emphasize that every occurrence of $\bar{\mathbf{z}}$ is to be differentiated. The distinction with the partial derivative $\partial_{\bar{\mathbf{z}}}$ appears in some notations. For instance when using the notation $\bar{g}[\bar{\mathbf{z}},\bar{\boldsymbol{\psi}}] = \bar{g}(\bar{\mathbf{z}})[\bar{\boldsymbol{\psi}}]$, then the operators are the same $d_{\bar{\mathbf{z}}} = \partial_{\bar{\mathbf{z}}}$. However, in some cases, a notation such

as $\bar{g}(\bar{\mathbf{z}}, \bar{\boldsymbol{\psi}}(\bar{\mathbf{z}}), \partial_{\bar{\mathbf{z}}}\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}}), ...)$ is used; for instance, it is natural in the case of a local dependence in $\bar{\boldsymbol{\psi}}(\bar{\mathbf{z}})$. Then the operators are different, the operator $\partial_{\bar{\mathbf{z}}}$ is a kind of partial derivative, and the corresponding total derivative is $d_{\bar{\mathbf{z}}}$; they are related by $d_{\bar{\mathbf{z}}}\bar{g} = \partial_{\bar{\mathbf{z}}}\bar{g} + \partial_{\bar{\psi}^\alpha}\bar{g}\ \partial_{\bar{\mathbf{z}}}\bar{\psi}^\alpha$. The distinction is important for commutation properties for instance: $\partial_{\bar{\mathbf{z}}}$ commutes with $\partial_{\bar{\boldsymbol{\psi}}}$, whereas $d_{\bar{\mathbf{z}}}$ does not. Notice that a definition of $\partial_{\bar{\mathbf{z}}}$ relying on first-order variations may not be most convenient, because $\bar{g}[\bar{\mathbf{z}}, \bar{\boldsymbol{\psi}}]$ is actually $\bar{g}(\bar{\mathbf{z}})[\bar{\boldsymbol{\psi}}]$, i.e. a function in the variable $\bar{\mathbf{z}}$ and a functional (yet often just an operator) in the field $\bar{\boldsymbol{\psi}}$. First-order variations of $\bar{g}$ with respect to $\bar{\mathbf{z}}$ gives $d_{\bar{\mathbf{z}}}$, then $\partial_{\bar{\mathbf{z}}}$ is defined by $\partial_{\bar{\mathbf{z}}}\bar{g} = d_{\bar{\mathbf{z}}}\bar{g} - \partial_{\bar{\psi}^\alpha}\bar{g}\ \partial_{\bar{\mathbf{z}}}\bar{\psi}^\alpha$.

Now, Eq. (11.21) holds for any $\bar{g}$ and any $\delta\boldsymbol{\psi}$. This implies that

$$0 = -\mathcal{J}^{-1} d_{\bar{\mathbf{z}}} \cdot \left(\bar{\mathbf{z}}_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z})\right) + \mathcal{J}^{-1}_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z}) \tag{11.22}$$

holds for any $\delta\boldsymbol{\psi}$.

Formula (11.22) can be confirmed by proving it directly with usual properties of derivatives of determinants. Indeed, it can be written

$$d_{\bar{\mathbf{z}}^j}\mathbf{z} \cdot d_{\mathbf{z}}\ \left(\bar{\mathbf{z}}^j_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z})\right) = \mathcal{J}\mathcal{J}^{-1}_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z})\,. \tag{11.23}$$

Now, the right-hand side can be rewritten using the usual property of variations of a determinant

$$\begin{aligned}
\mathcal{J}\mathcal{J}^{-1}_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z}) &= \mathrm{Det}^{-1}(M)\delta\mathrm{Det}(M) = \mathrm{Trace}(M\cdot\delta M)\\
&= d_{\bar{\mathbf{z}}^j}\mathbf{z}^k\delta\left(d_{\mathbf{z}^k}\bar{\mathbf{z}}^j\right)\\
&= d_{\bar{\mathbf{z}}^j}\mathbf{z}^k\partial_{\psi^\alpha}\left(d_{\mathbf{z}^k}\bar{\mathbf{z}}^j\right)\delta\psi^\alpha\\
&= d_{\bar{\mathbf{z}}^j}\mathbf{z}^k\partial_{\psi^\alpha}\left(\partial_{\mathbf{z}^k}\bar{\mathbf{z}}^j + \partial_{\psi^\beta}\bar{\mathbf{z}}^j\partial_{\mathbf{z}^k}\psi^\beta\right)\delta\psi^\alpha\\
&= d_{\bar{\mathbf{z}}^j}\mathbf{z}^k\Big(\partial_{\psi^\alpha}\partial_{\mathbf{z}^k}\bar{\mathbf{z}}^j + \partial_{\psi^\alpha}\left(\partial_{\psi^\beta}\bar{\mathbf{z}}^j\right)\partial_{\mathbf{z}^k}\psi^\beta\\
&\qquad\qquad + \partial_{\psi^\beta}\bar{\mathbf{z}}^j\partial_{\psi^\alpha}\left(\partial_{\mathbf{z}^k}\psi^\beta\right)\Big)\delta\psi^\alpha\\
&= d_{\bar{\mathbf{z}}^j}\mathbf{z}^k\left(\partial_{\mathbf{z}^k}\partial_{\psi^\alpha}\bar{\mathbf{z}}^j + \partial_{\psi^\beta}\left(\partial_{\psi^\alpha}\bar{\mathbf{z}}^j\right)\partial_{\mathbf{z}^k}\psi^\beta + \partial_{\psi^\alpha}\bar{\mathbf{z}}^j\partial_{\mathbf{z}^k}\right)\delta\psi^\alpha\\
&= d_{\bar{\mathbf{z}}^j}\mathbf{z}^k\left(d_{\mathbf{z}^k}\partial_{\psi^\alpha}\bar{\mathbf{z}}^j + \partial_{\psi^\alpha}\bar{\mathbf{z}}^j\partial_{\mathbf{z}^k}\right)\delta\psi^\alpha\\
&= d_{\bar{\mathbf{z}}^j}\mathbf{z}^k d_{\mathbf{z}^k}\left(\partial_{\psi^\alpha}\bar{\mathbf{z}}^j\delta\psi^\alpha\right),
\end{aligned}$$

which is exactly the left-hand side of Eq. (11.23). Here, we used $\mathrm{Det}(M) = \mathcal{J}^{-1}$, and $M = d_{\mathbf{z}}\bar{\mathbf{z}}$. In the first and second lines, we used the property of variations of a determinant, in the third and fourth lines, variations were made explicit, as well as the total derivative. In the fifth and sixth lines, derivatives were commutated. In the last lines, all terms were recombined into a single term.

Let us turn back to the case of a general functional (not constant in $\bar{\boldsymbol{\psi}}$), which we denote by

$$\bar{F}[\bar{\boldsymbol{\psi}}] = \int d\bar{\mathbf{z}}\ \bar{f}[\bar{\mathbf{z}}; \bar{\boldsymbol{\psi}}] = F[\boldsymbol{\psi}] = \int d\mathbf{z}\ \bar{f}[\bar{\mathbf{z}}[\mathbf{z}; \boldsymbol{\psi}]; \boldsymbol{\psi}]$$

as in the previous subsection, in order to avoid confusion with the previous functional $\bar{G}$ that was constant in $\bar{\boldsymbol{\psi}}$. Eq. (11.22) implies

$$\begin{aligned}
0 &= \int d\mathbf{z}\ \bar{f}[\bar{\mathbf{z}}[\mathbf{z}; \boldsymbol{\psi}]; \boldsymbol{\psi}]\left(-\mathcal{J}^{-1}d_{\bar{\mathbf{z}}}\cdot\left(\bar{\mathbf{z}}_{\psi^\alpha}\delta\boldsymbol{\psi}(\mathbf{z})\right) + \mathcal{J}^{-1}_{\psi^\alpha}\delta\psi^\alpha(\mathbf{z})\right)\\
&= \int d\mathbf{z}\ \left(\mathcal{J}^{-1}\bar{f}_{\bar{\psi}^\beta}\bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)\delta\psi^\alpha(\mathbf{z})\\
&= \int d\mathbf{z}\ \delta\psi^\alpha(\mathbf{z})\left(\mathcal{J}^{-1}\bar{f}_{\bar{\psi}^\beta}\bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)^\dagger,
\end{aligned} \tag{11.24}$$

where in the second line, an integration by part was performed in the reverse direction compared to Eq. (11.21), and in the last line, another integration by parts was performed to have the function $\delta\boldsymbol{\psi}$ factorized. The adjoint $\dagger$ is related to the measure $d\mathbf{z}$, and here it is the adjoint function, i.e. the adjoint operator acting on the constant function 1, in the same way as in Eq. (11.19).

Eq. (11.24) holds for any $\delta\boldsymbol{\psi}$, which implies

$$0 = \left(\mathcal{J}^{-1}\bar{f}_{\bar{\psi}^\beta}\bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)^\dagger . \tag{11.25}$$

This is the lemma we were seeking for, since it relates the field derivative of the Jacobian $\mathcal{J}^{-1}_{\bar{\boldsymbol{\psi}}}$ and of the coordinate transformation $\bar{\mathbf{z}}_{\boldsymbol{\psi}}$ with the field derivative of the functional density $\bar{f}_{\bar{\boldsymbol{\psi}}}$, which is related with the functional derivative $\bar{F}_{\bar{\boldsymbol{\psi}}}$.

### 11.4.3 Transformed functional derivatives in the good form

Using the previous lemma (11.25), Eq. (11.18) can be rewritten in a convenient way, where derivatives of the Jacobian and of the coordinate transformation do not appear any more:

$$\begin{aligned} T^{-1}\tfrac{\delta}{\delta\psi^\alpha(\mathbf{y})}\mathbb{T}^{-1}(\bar{F}) &= \left(\mathcal{J}^{-1}\bar{f}_{\psi^\alpha} + \mathcal{J}^{-1}\bar{f}_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha} + \bar{f}\mathcal{J}^{-1}_{\psi^\alpha}\right)^\dagger \\ &= \left(\mathcal{J}^{-1}\bar{f}_{\psi^\alpha} - \mathcal{J}^{-1}\bar{f}_{\bar{\psi}^\beta}\bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\right)^\dagger \\ &= \left(1^\beta_\alpha - \bar{\psi}^\beta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\right)^\dagger \mathcal{J}^{-1}\bar{F}_{\bar{\psi}^\beta} . \end{aligned} \tag{11.26}$$

This formula for the transformed functional derivative is now in a convenient form $\bar{F}_{\bar{\psi}^\alpha} = \mathbf{M}^\beta_\alpha F_{\psi^\beta}$. As expected, the result agrees with the chain-rule formula (11.16), as is confirmed by (11.20).

## 11.5 Transformed dynamics

Now, all the Hamiltonian structure has been expressed in the transformed system: the transformed Hamiltonian functional in (11.3), and the transformed Poisson bracket in (11.4), with the transformed functional derivative in (11.26). So, the transformed dynamics can be computed:

$$\begin{aligned} \dot{\bar{F}} &= \bar{\{}\bar{F},\bar{H}\bar{\}} \\ &= \int \mathcal{J}d\bar{\mathbf{z}}\ \left(1^\gamma_\alpha - \bar{\psi}^\gamma_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\right)^\dagger \left(\mathcal{J}^{-1}\bar{F}_{\bar{\psi}^\gamma}\right) \\ &\qquad\qquad T^{-1}\mathbb{J}^{\alpha\beta}\ \left(1^\delta_\beta - \bar{\psi}^\delta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\beta}\right)^\dagger \left(\mathcal{J}^{-1}\bar{H}_{\bar{\psi}^\delta}\right) . \end{aligned}$$

To interpret this dynamics and compare with the expected formula, let us come back to the initial variable $\mathbf{z}$

$$\begin{aligned} \dot{\bar{F}} = \int d\mathbf{z}\ T\left(1^\gamma_\alpha - \bar{\psi}^\gamma_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\alpha}\right)^\dagger \left(\mathcal{J}^{-1}\bar{F}_{\bar{\psi}^\gamma}\right) \\ \mathbb{J}^{\alpha\beta}\ T\left(1^\delta_\beta - \bar{\psi}^\delta_{\bar{\mathbf{z}}}\cdot\bar{\mathbf{z}}_{\psi^\beta}\right)^\dagger \left(\mathcal{J}^{-1}\bar{H}_{\bar{\psi}^\delta}\right) . \end{aligned}$$

The contribution of $\bar{H}$ can be rewritten

$$TT^{-1}\tfrac{\delta}{\delta\psi^\beta}\mathbb{T}^{-1}(\bar{H}) = \tfrac{\delta}{\delta\psi^\beta}H . \tag{11.27}$$

This abstract property is obvious but important to keep in mind, because it explains some cancellations, such as the magnetization term in Chapter 4, which can be highly non-trivial in practical computations when the change of coordinates is more involved than the one considered in Chapter 4.

The contribution of $\bar{F}$ can also be rewritten

$$TT^{-1}\tfrac{\delta}{\delta\psi^\alpha}\mathbb{T}^{-1}(\bar{F}) = \tfrac{\delta}{\delta\psi^\alpha}F .$$

Inserting the previous two relations, the dynamics becomes

$$\dot{\bar{F}} = \int d\mathbf{z}\; F_{\psi^\alpha} \mathbb{J}^{\alpha\beta} H_{\psi^\beta} = \dot{F}\,.$$

This only confirms the consistency of the definitions we adopted, but it does not show what explains the form of the transformed dynamics. It is more interesting to keep the transformed field variable $\bar{\boldsymbol{\psi}}$ in the functional $\bar{F}$:

$$\begin{aligned}\dot{\bar{F}} &= \int d\mathbf{z}\; \mathcal{J}^{-1} \bar{F}_{\bar{\psi}^\gamma} \left(1^\gamma_\alpha - \bar{\psi}^\gamma_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\psi^\alpha}\right) \mathbb{J}^{\alpha\beta} H_{\psi^\beta} \\ &= \int d\bar{\mathbf{z}}\; \bar{F}_{\bar{\psi}^\gamma} \int d\mathbf{z}\; \tfrac{\delta \bar{\psi}^\gamma(\bar{\mathbf{z}})}{\delta \psi^\alpha(\mathbf{z})} \mathbb{J}^{\alpha\beta} H_{\psi^\beta} = \int d\bar{\mathbf{z}}\; \bar{F}_{\bar{\psi}^\gamma} \{\bar{\psi}^\gamma(\bar{\mathbf{z}}), H\} = \int d\bar{\mathbf{z}}\; \bar{F}_{\bar{\psi}^\gamma} \dot{\bar{\psi}}^\gamma(\bar{\mathbf{z}})\,, \end{aligned} \tag{11.28}$$

using Eq. (11.13) to get the second equality. This shows that the transformation completely agrees with a change of field coordinates, and the transformed dynamics both agrees with the initial dynamics and is just given by a chain rule at the field level.

An additional remark is that the two terms involved in transformed functional derivatives, e.g. (11.13) and (11.16), transfer to the transformed Poisson bracket and dynamics.

The first term just gives to the transformed Poisson bracket and dynamics the same expression as the initial ones. It corresponds to considering that the scalar invariance (11.2) implies that the transformed fields evolve in a similar way as the initial one.

As for the second term, it traduces the field dependence in the coordinate transformation. For a near-identity transformation, it is of higher order compared to the first term. It implies additional terms in the transformed dynamics and Poisson bracket compared to the corresponding expressions in the initial system. For the dynamics, it traduces the effects of the time dependence in the coordinate transformation:

$$\dot{\bar{\psi}}^\gamma(\bar{\mathbf{z}}) = \dot{\psi}^\gamma(\mathbf{z}) - \bar{\psi}^\gamma_{\bar{\mathbf{z}}} \cdot \bar{\mathbf{z}}_{\psi^\alpha} \dot{\psi}^\alpha(\mathbf{z})\,,$$

which is just a particular way to write the time evolution of Eq. (11.8). These remarks were clearly illustrated in the simplified lifting considered in Chapter 4, and here they are confirmed for a general lifting.

## Conclusion

For a lifting of a coordinate transformation to the Hamiltonian structure of the corresponding fluid-like field dynamics, we have generalized the method used in Chapter 4 to an arbitrary transformation, even when the transformation depends on fields defined over a space that is changed by the transformation. For the Hamiltonian structure, as expected, it mainly corresponds to a chain rule, but interesting features are involved about this chain rule because transformed fields are defined through an implicit function and at a moving point. This is because either the initial point or the transformed one can be kept fixed, but not both, and when the coordinate used is the transformed one, the field used has to be the transformed one.

In addition, we have shown how the chain rule can be obtained from the transformed functional derivatives. This emphasized the deep difference between function chain rule and functional chain rule, because the integration variable is mute for functionals but sensitive to computations of functional derivatives, which is not the case of mute indices for functions (e.g. for scalar functions defined as the scalar product of two vectorial coordinates).

After deriving the transformed Hamiltonian structure, the transformed dynamics was obtained as agreeing with a change of field coordinate, with a transformation that implicitly depends on time.

Here, the purpose was to study the lifting of transformations that affect all the coordinates. For simplicity, and in order to better focus on the essential mechanisms at work, the case was limited to a fluid-like theory, where all the fields are defined over the same space. For the Vlasov-Maxwell system, an additional feature has to be taken into account: all the fields are not defined over the same space. The lifting in such a framework is the next step of the work, and will be reported in Chapter 5.

# Chapter 12

# Incidence, constructibility, and dispensability of quarter-canonical coordinates for Dirac truncations

in collaboration with Cristel Chandre

*A*bstract: In the presence of a quarter-canonical Poisson bracket, Hamiltonian reductions by Dirac's theory of constraints, formulated as a projection of derivatives, can be simplified by using a projector related to a mere bracket truncation. The stiffness of the condition on the coordinates is investigated. Quarter-canonical coordinates are observed to be fairly common, and especially in Dirac reductions. In addition, they can often be identified by hand when they are not already present.
On another hand, such coordinates are not necessary to benefit from the simplified projector, which can be investigated even in case the coordinates used are not quarter-canonical ones. The formula for the projector is obtained first for an explicit generic bracket suggested by examples from the ideal-fluid model for plasmas, and then in a more abstract setting.
This brings about a generalization of the simplified projector to non-quarter-canonical cases by providing it with a coordinate-independent expression even when quarter-canonical coordinates and their associated truncation projector do not exist. The results emphasize that the simplified projector only requires half of the constraints not to be coupled with itself. Relations with the quarter-canonical case and with orthogonal projectors are observed.

## Introduction

This appendix chapter takes place just before Sec. 8.4 in Chapter 8. Indeed, the first few sections of Chapter 8 showed the importance of the quarter-canonical structure for Dirac's theory of constraints, formulated as a projection of derivatives. It induces the presence of a simplified projector and makes the reduction closer to bracket truncations. But this structure is explicitly present only when adapted coordinates are used. This asks a question about the incidence of such coordinates, because they are interesting only if they are common enough.

On another hand, it often occurs that the natural variables are different. One can wonder whether in such cases, the corresponding quarter-canonical coordinates can be identified, in order to adopt them and make the Dirac reduction become just given by the simplified truncation projector. Alternatively, it can be more interesting to identify the simplified projector directly in the natural, non-quarter-canonical coordinates.

The present chapter addresses these questions. The first two sections are interested in the incidence and constructibility of quarter-canonical coordinates. Sec. 12.1 shows that the quarter-canonical structure is indeed fairly frequent in Dirac reductions, even in infinite-dimensional systems. Sec. 12.2 studies how quarter-canonical coordinates can be constructed when they are not

present. Then, for the rest of the chapter, we will come to the second aspect of the problem, i.e. how the corresponding bracket truncation and simplified projector appear when non-quarter-canonical coordinates are kept, in a similar way as when the reduction to an incompressible fluid is performed while keeping the natural field coordinates $\rho$ and $\mathbf{u}$. The quarter-canonical structure exists, but it is not explicit in the Poisson bracket. Sec. 12.3 investigates its effects by considering the explicit form of the Poisson bracket. Sec. 12.4 adopts a more abstract point of view and goes beyond the quarter-canonical framework.

After clarifying those points about simplified projectors in non-quarter-canonical coordinates, we will turn to the consequences of the quarter-canonical structure on Hamiltonian reduction methods in the next Chapter. Then, in the light of the results, we will come back to the gyrokinetic reduction and study how to use the Dirac truncation method in order to remove the gyro-angle dimension from the base space of the Vlasov density. This will be the topic of Sec. 8.4 in Chapter 8.

For the notations, the present chapter behaves as a section of Chapter 8, where the common notations are introduced once for the three companion chapters Chapter 8, 12, and 13.

## 12.1 Incidence of the quarter-canonical structure

In this section, the commonness of quarter-canonical coordinates is investigated. In Subsec. 12.1.1, it is shown that they are always involved in Dirac reductions for finite dimensional systems when half of the constraints is not coupled with itself, i.e. $\{\phi_1, \phi_1\} = 0$. Then, in Subsec. 12.1.2, more qualitative reasons are given for their frequency in continuous media, where a general proof is not available, and examples are given in the ideal-fluid model.

### 12.1.1 Fundamental case of existence

First, let us consider the case where the Dirac reduction starts from a scalar primary coordinate $\phi_1$. Then the Darboux algorithm implies the existence of (local) semi-canonical coordinates $(\phi_1, \Phi_2, \boldsymbol{\chi})$, i.e. coordinates such that the Poisson matrix becomes semi-canonical (8.11). Thus, the semi-canonical structure (which is quarter-canonical as well) always exists in this case.

There are more complicated situations where the first constraint $\phi_1$ is a vectorial quantity, and then the coefficients in (8.16) are actually matrices. Then, there is no Darboux theorem to guarantee the existence of a semi-canonical structure, and there are explicit counterexamples, as will be illustrated in Sec. 13.4.1. Nonetheless, it is still guaranteed when $\{\phi_1, \phi_1\} = 0$. The Darboux algorithm allows to iteratively construct variables $\Phi_2^i$ canonically conjugated to $\phi_1^i$, for each value of $i$. The condition $\{\phi_1, \phi_1\} = 0$ implies the independence of the variables $\phi_1^i$ and $\Phi_2^j$. One ends up with the variables $(\phi_1, \Phi_2, \chi)$ that are semi-canonical even if $\phi_1$ is a vector.

To be more precise, the Darboux theorem gives only local coordinates, but in most of our examples, the coordinates happen to be global ones. More generally, we avoid to consider here technical points which can generate obstructions to the argument in general but do not in many applications. For instance, when inverting the Laplacian, we implicitly assume that proper boundary conditions have been chosen. When going from vector fields to coordinates (e.g. when defining the Casimir invariant coordinates $\psi_C$ by $\ker(\mathbb{J})$), we assume no obstruction occurs. The examples we mention emphasize that such assumptions are often verified. They simplify much the framework, and imply to work with the coordinate changes for the Poisson bracket (even for plasmas, i.e. continuous media) in a similar way as in the case of matrices in a finite dimensional vector space. Also, a property that helps much is that the Poisson bracket is linear in the field variables, but this is fairly common in fluid and plasma dynamics [159].

Thus, the semi-canonical structure (and therefore the quarter-canonical) always exists when $\{\phi_1, \phi_1\} = 0$. Because of this conclusion, the requirements for the semi canonical structures can be shrunk to the only condition $\{\phi_1, \phi_1\} = 0$. This condition is not very stringent, since it often appears in Dirac reductions.

Now, we come to the point where semi-canonical coordinates lose their relevance, and the role of the quarter-canonical arises. Indeed, even if the semi-canonical structure exists, it is not necessarily

implied in the reduction, because the second constraint does not have to be $\Phi_2$. The semi-canonical form for the Poisson bracket is preserved with only a restricted number of constraints: $\Phi_2'(\phi_1, \Phi_2)$. It is why it is seldom involved in Dirac reductions. On the contrary, the quarter-canonical bracket is obtained with much more of them: $\phi_2'(\phi_1, \Phi_2, \boldsymbol{\chi})$, i.e. the second coordinate[1] can be any function in bijection with $\Phi_2$ (because it has to define a good change of coordinates). It is exactly the same criterion that the second Dirac constraint has to satisfy: it has to be a function in bijection with $\Phi_2$, otherwise the matrix of constraints is not invertible, because of (8.16), with $b = -(\phi_2')_{\Phi_2}$. So, not only quarter-canonical reductions are much more frequent than semi-canonical ones, but they are actually the only possible Poisson bracket for Dirac reductions when $\{\phi_1, \phi_1\} = 0$.

### 12.1.2 Reasons for the occurrence

The argument of the previous subsection do not apply to continuous Hamiltonian systems, because of the absence of a Darboux theorem for instance. In addition, for finite dimensional systems, the argument can fail because of the possible non-global existence of the coordinates defined by the Darboux theorem. So, in the present subsection, more qualitative reasons are given for the incidence of quarter-canonical coordinates in Dirac reductions.

From an abstract point of view, the occurrence of quarter-canonical coordinates compared to semi-canonical ones is as follows. Starting from a coordinate $\phi_1$, semi-canonical coordinates imply for $\Phi_2$ and $\boldsymbol{\chi}$ to satisfy conditions

$$\{\phi_1, \Phi_2\} = 1\,, \qquad \{\phi_1, \boldsymbol{\chi}\} = 0\,, \qquad \{\Phi_2, \boldsymbol{\chi}\} = 0\,. \tag{12.1}$$

It means integrating those equations "along the characteristics", which is guaranteed only locally for finite-dimensional systems, as in the Darboux algorithm. Quarter-canonical coordinates drop the requirements on $\Phi_2$. As a result, instead of relying on the Darboux theorem, the quarter-canonical structure only relies on the Frobenius theorem [84, 98].

Indeed, this structure is defined by the two conditions (8.13)-(8.14). The second one means the existence of $\dim(\boldsymbol{\psi}) - 2.\dim(\psi_1)$ solutions to equation $\{\phi_1, F\} = 0$. By the first condition, one only needs to find $\dim(\psi) - \dim(\psi_1)$ of them. Now, the Frobenius theorem states that such coordinates exist (at least locally) if and only if the operator $\{\phi_1^\alpha, \{\phi_1^\beta, \cdot\}\} - \{\phi_1^\beta, \{\phi_1^\alpha, \cdot\}\}$ is inside the linear space generated by the operators $\{\phi_1^\alpha, \cdot\}$. But the Jacobi identity implies

$$\{\phi_1^\alpha, \{\phi_1^\beta, \cdot\}\} - \{\phi_1^\beta, \{\phi_1^\alpha, \cdot\}\} = \{\{\phi_1^\alpha, \phi_1^\beta\}, \cdot\}\,, \tag{12.2}$$

and the first condition $\{\phi_1, \phi_1\} = 0$ is enough for the Frobenius condition to be satisfied under its most trivial form.

Like the Darboux theorem, the Frobenius theorem gives only local existence, but it is already more general. So, obstructions to global existence are less frequent. For continuous media, such as fluid models, these theorems do not apply but the absence of canonical coordinates can all the same be linked to topological obstructions in the change of coordinates to Clebsch potentials [159], and the topological requirements for the existence of quarter-canonical coordinates is much less strong.

A practical reason for the frequent occurrence of the quarter-canonical structure is that Hamiltonian reduction methods often rely on special properties of the semi-canonical structure, but the quarter-canonical structure is most often enough to get them, while it is much less stringent. For instance, we saw that it was enough for the Dirac reduced bracket to be a truncation of the initial bracket. In Hamiltonian perturbation theory, it is enough to guarantee the existence of a conserved quantity $\phi_1$ (the "action") when the Hamiltonian does not depend on the coordinate $\phi_2$ (the "conjugated" angle).

In order to illustrate the arguments above, a well-known example is the work of [85, 86] about the guiding-center reduction. It is the reduction of particle dynamics in a strong magnetic field,

[1] Notice that also for the coordinates $\boldsymbol{\chi}$, the quarter-canonical structure is much more frequent than semi-canonical one: instead of allowing only the coordinate change $\boldsymbol{\chi}'(\boldsymbol{\chi})$, it allows for any transformation $\boldsymbol{\chi}'(\phi_1, \boldsymbol{\chi})$.

which was also the topic of Chapters 1 and 2. The phase space is given by $(\mathbf{q}, \mathbf{v})$, where $\mathbf{q}$ and $\mathbf{v}$ are the particle's position and velocity. The Hamiltonian is the particle kinetic energy

$$H = \frac{\mathbf{v}^2}{2}\,,$$

with the particle mass chosen unity $m = 1$ for simplicity. The Poisson bracket is

$$\{F, G\} = \partial_{\mathbf{v}} F \cdot \partial_{\mathbf{q}} G - \partial_{\mathbf{q}} F \cdot \partial_{\mathbf{v}} G + \partial_{\mathbf{v}} F . e\mathbf{B} \times \partial_{\mathbf{v}} G\,.$$

The principle of the reduction is to get a conserved quantity $\bar{\mu}$, the magnetic moment, and to isolate the fast dynamics into the conjugated coordinate, the so-called gyro-angle $\bar{\theta}$. The Darboux algorithm was used in [85, 86] to build semi-canonical coordinates $(\bar{\mu}, \bar{\theta}, \bar{\boldsymbol{\chi}})$ and get the desired Hamiltonian reduction. The definition of the gyro-angle involved a local gauge, which could be removed from the dynamics of the slow variables only by adopting new coordinates $(\bar{\mu}, \bar{\theta}, \bar{\boldsymbol{\chi}}')$, and by removing the requirement of the Poisson matrix being semi-canonical. Thus in this case, the quarter-canonical structure is necessary and sufficient to get the desired Hamiltonian reduction with a gauge-independent slow reduced dynamics.

This example has other interesting aspects than the requirements on $\bar{\mu}$ and $\bar{\theta}$. A third requirement was desirable, that is to have a Hamiltonian sub-dynamics for $\bar{\boldsymbol{\chi}}'$. This corresponds to a Hamiltonian truncation. As shown in Chapter 8, this truncation is provided by the Dirac reduction, but only if the structure is quarter canonical. Actually, in this case, the need for quarter-canonical constraints is still more fundamental: if it is not obtained, not only the Dirac reduction is not a truncation any more, but in addition the Dirac reduction cannot be applied at all. Indeed, starting from the constraint on the magnetic moment, a secondary constraint is needed, and $\phi_2$ must correspond to the gyro-angle, otherwise the matrix of constraint is not invertible. This situation is standard for Dirac reductions starting from a primary scalar constraint, as appeared in the previous subsections. This point will be further developed in Chapter 13.

Let us turn to examples of quarter-canonical brackets in continuous Hamiltonian systems. In the fluid model for plasmas (8.2) for instance, the mass density is coupled only with the compressible part of the velocity, the magnetic field is coupled only with the (solenoidal part of the) electric field, the compressible part of the electric field is coupled only with the compressible part of the velocity, etc. but each of these pairing is non symmetric. Actually, only the velocity is not the coordinate $\phi_1$ of a quarter-canonical bracket.

This observation is not a surprise, because most Poisson brackets are often built from canonical brackets, and adding interaction terms with one of the initially conjugated variables, which spoil both the canonical and the semi-canonical structure, but not the quarter-canonical structure. Also, most Hamiltonian systems for continuous media are Lie algebras of semidirect product type [69,95].

As with concerns Dirac reductions for fluids and plasmas, most of them reveal a quarter-canonical structure, e.g. the reduction to incompressible fluid, the reduction to electrostatic field, the reduction to the dipolar approximation (uniform electric field), etc.

The quarter-canonical structure is not the only available structure for Dirac reductions in continuous media even for scalar constraints. It is easily seen by the cases where both $\phi_1$ and $\phi_2$ are coupled with themselves. For scalar constraints, it is a specificity of infinite-dimensional systems: the coefficients of the Poisson bracket are differential operators, and the diagonal components can be non zero. For finite-dimensional systems, it can happen when the constraints $\phi_1$ and $\phi_2$ are vectorial. Two examples of such cases are the reduction to the Charney-Hasegawa-Mima model [32], in which the constraints are related to $\nabla \cdot \mathbf{u}$ and $\rho - f(\nabla \times \mathbf{u})$, with $f$ some given function, and a reduction related to the ideal-MHD model, in which the constraints are $\mathbf{E} + \mathbf{u} \times \mathbf{B}$ and $\mathbf{J} - \nabla \times \mathbf{B}$.

In practice, this is not typical. On the contrary, Dirac reductions in fluids and plasmas are most often associated with a quarter-canonical structure, especially when $\{\phi_1, \phi_1\} = 0$, as is emphasized by the various examples mentioned throughout this chapter, as well as Chapter 13. Probably one of the reasons is that the invertibility condition is easy to control in quarter-canonical structures, as will be emphasized in the next chapter, whereas it is much more involved to find invertible matrices of constraints in other situations, and especially invertible matrices that give a reduced bracket not

too complicated to deal with. For instance, one way to invert the matrix $\begin{pmatrix} c & -b^\dagger \\ b & a \end{pmatrix}$ is to invert not only $b$, but $b + ab^{-1\dagger}c$ as well. On the contrary, the matrix $\begin{pmatrix} 0 & -b^\dagger \\ b & a \end{pmatrix}$ is straightforwardly inverted as soon as $b^{-1}$ is known.

## 12.2 On the identification of suited coordinates

The previous section showed that the quarter-canonical structure is very common in Dirac reductions. This is most convenient since Chapter 8 showed that in this case Dirac's reduction procedure can be strongly simplified, by becoming just a trivial projection of derivatives in the Poisson bracket, which results in a mere bracket truncation. Now, this works only when the coordinates are suited to the quarter-canonical structure. Most often, it is not the case in the initial coordinates. In such cases, in order to benefit from the simplified truncation procedure, quarter-canonical coordinates should be adopted, which essentially implies to identify them. This section investigates how this can be done. More precisely, it is the topic of Subsec. 12.2.1. Then, the case of semi-canonical coordinates together with a few examples will be considered in Subsec. 12.2.2. Last, in Subsec. 12.2.3 we will mention weakenings of the quarter-canonical requirements which could be interesting to explore.

### 12.2.1 Case of quarter-canonical coordinates

In order to obtain quarter-canonical coordinates, one should solve equations (8.13)-(8.14). As a first illustration from the warm fluid model, the first constraint can be taken as $\phi_1 = \nabla \times \mathbf{E}$. Then the bracket is not quarter-canonical in the initial coordinates, since the solenoidal part of the electric field is not coupled with only one field variable, but with two of them, the magnetic field and the solenoidal part of the velocity, through

$$G_{\mathbf{E}} \cdot (F_{\mathbf{u}} - \nabla \times F_{\mathbf{B}}) - F_{\mathbf{E}} \cdot (G_{\mathbf{u}} - \nabla \times G_{\mathbf{B}}) \,.$$

But these can be combined in one variable $\phi_2$ by a change of variable from $\mathbf{B}$ (or $\mathbf{u}$) to $\nabla \times \mathbf{u} + \mathbf{B}$. Then a quarter-canonical structure is obtained with

$$\phi_1 = \nabla \times \mathbf{E} \,, \qquad \boldsymbol{\chi} = (\nabla \cdot \mathbf{u}, \nabla \times \mathbf{u} + \mathbf{B}, \nabla \cdot \mathbf{E}, \rho) \,,$$

and $\phi_2$ is taken as a supplementary coordinate, e.g. the simple variable $\nabla \times \mathbf{u}$, or the variable $\nabla \times \mathbf{u} - \mathbf{B}$ orthogonal to $\boldsymbol{\chi}$.

This idea can be developed to fit in with the frequent situations where there are coordinates $(\psi_1, \psi_2)$ on $\boldsymbol{\psi}$ such that

$$\{\phi_1, \cdot\} = f(\psi_1) \cdot \partial_{\psi_2} \,, \tag{12.3}$$

and $\psi_1$ contains $\phi_1$, which insures $\mathbb{J}^{\phi_1\phi_1} = 0$. This situation concerns most of the examples in this chapter. Then the desired quarter-canonical coordinates can be obtained explicitly. First, restrict the space on which $\mathbb{J}^{\phi_1\boldsymbol{\psi}}$ acts so as to make it injective. Then for each of the derivatives $\partial_{\psi_{2i}}$, identify its range $R_i = \mathrm{Rg}\, f_i$, and define the coordinates $\phi_{1i}$ as the minimal non-zero intersections[2] of these $R_i$. Then, define $K_j$ as the kernels of the transpose of the restriction of $\mathbb{J}^{\boldsymbol{\psi}\phi_1}$ to $R_i$. Finally, set $\boldsymbol{\chi} = \cap_j K_j$, and $\phi_2$ as complementary coordinates to $(\phi_1, \boldsymbol{\chi})$ in $\boldsymbol{\psi}$. The procedure allows to compute the corresponding quarter-coordinates.

More generally, starting from a given coordinate $\phi_1$, quarter-canonical coordinates are obtained by choosing $\boldsymbol{\chi}$ as free coordinates complementary to $\phi_1$ in $\ker\{\phi_1, \cdot\} = \ker b^\dagger$ [3] , and $\phi_2$ a free coordinate complementary to $(\phi_1, \boldsymbol{\chi})$ in $\boldsymbol{\psi}$. For instance one standard way to define $\phi_2$ is to take it as $\mathrm{Rg}\, b$, which corresponds to the coordinates orthogonal to $(\phi_1, \boldsymbol{\chi})$. Even if this method works in some cases, the point is that $\ker\{\phi_1, \cdot\}$ is not sure to define good coordinates. The Frobenius

[2] For an illustration, see the example with $\phi_1 = \mathbf{E}$ in the next subsection.
[3] The operator $b$ is $\{\phi_2, \phi_1\}$, as in Eq. (12.7) for instance.

condition is needed in finite-dimensional systems to conclude locally when $\mathbb{J}^{\phi_1\phi_1} = 0$, and it does not apply to infinite-dimensional systems.

As an example where it is difficult to identify the corresponding coordinates, the first constraint can be taken as $\phi_1 = \nabla\cdot\mathbf{u}$. Then the bracket is not quarter-canonical in the initial coordinates. For a first step, let us make the bracket constant in the field variables[4], excluding the gyro-magnetic term (proportional to $\mathbf{B}/\rho$ and the self-coupling of the velocity (proportional to $\nabla\times\mathbf{u}/\rho$). Then the compressible part of the velocity is coupled with the mass density and the compressible part of the electric field, through the terms

$$\nabla\cdot F_{\mathbf{u}}(G_\rho + \Delta^{-1}\nabla\cdot G_{\mathbf{E}})\,, \tag{12.4}$$

and these can be combined in $\rho - \nabla\cdot\mathbf{E}$. Then a quarter-canonical structure is obtained for this part of the bracket, with

$$\phi_1 = \nabla\cdot\mathbf{u}\,, \qquad \boldsymbol{\chi} = (\rho - \nabla\cdot\mathbf{E}, \mathbf{B}, \nabla\times\mathbf{E})\,,$$

and $\phi_2$ is taken as a supplementary coordinate, e.g. $\rho$, or $\nabla\cdot\mathbf{u}$, or rather $\rho+\nabla\cdot\mathbf{E}$ for an orthogonal change of variables. In this example, there is an additional feature, because (12.4) are the only terms involving derivatives with respect to the variables $\rho$ or $\nabla\cdot\mathbf{E}$, so that after the change of variable, $\chi_1 = \rho - \nabla\cdot\mathbf{E}$ is a Casimir invariant, but this is only due to the special form of the bracket.

Now, if we come back to the true fluid bracket, restoring the terms depending on the field variables, the bracket previously obtained is not quarter-canonical. The system of equations for $\ker\{\phi_1,\cdot\}$ trivially satisfies the Frobenius condition, but quarter-canonical coordinates are not easy to find. They imply to identify the kernel of $\mathbb{J}^{\nabla\cdot\mathbf{u}\boldsymbol{\psi}}\cdot\partial_{\boldsymbol{\psi}}F$ including the terms $\nabla\cdot(\frac{\nabla\times\mathbf{u}+\mathbf{B}}{\rho}\times F_{\mathbf{u}})$ involving $\nabla\times\mathbf{u}+\mathbf{B}$. It implies a nonlinear differential equation, which may not be easy to solve, even if it defines good global coordinates (which is not sure a priori).

In physically interesting situations, such difficulties are often avoided. This is confirmed by the fact that Casimir invariants are generically easily identified as good coordinates, whereas they should be concerned with the same kind of difficulties [159].

### 12.2.2 Case of semi-canonical coordinates

Let us turn to the case of semi-canonical coordinates. They are still more involved to define, because the coordinates $\Phi_2$ are more restricted. First, they must satisfy $\mathbb{J}^{\Phi_2\Phi_2} = 0$. This is guaranteed for finite-dimensional systems when $\Phi_2$ is a scalar, and then by induction it carries over to the case where $\phi_1$ and $\Phi_2$ are vectors. In infinite-dimensional systems, the diagonal coefficients can be non-zero, the argument cannot be applied, and the semi-canonical structure often does not exist [159], even when the structure is quarter-canonical. Second, after a convenient coordinate $\Phi_2$ can be identified, then semi-canonical coordinates are obtained only after identifying in addition coordinates $\boldsymbol{\chi}'$ on $\ker\{\phi_1,\cdot\}\cap\ker\{\Phi_2,\cdot\}$. One way to proceed is to look for coordinates $\boldsymbol{\chi}''$ on $\ker\{\Phi_2,\cdot\}$, and then to define the expected coordinates $\boldsymbol{\chi}'$ by $(\phi_1,\boldsymbol{\chi}'')\cap(\Phi_2,\boldsymbol{\chi})$.

On another hand, the condition $\{\Phi_2,\Phi_2\}$ is not essential. If it is removed, then $\boldsymbol{\chi}'$ are defined as coordinates on $\ker\{\phi_1,\cdot\}\cap\ker\{\Phi_2,\cdot\}$, with the condition that this space has dimension equal to $\dim\boldsymbol{\chi}$. For instance, if $\ker\{\Phi_2,\cdot\}$ has codimension $\dim\phi_1$, (i.e. the operator satisfies the Frobenius condition), then $\ker\{\phi_1,\cdot\}\cap\ker\{\Phi_2,\cdot\}$ verifies all the requirements: it has codimension $2\dim\phi_1$, it is a supplementary to $\phi_1$ and to $\Phi_2$. When the condition is not satisfied, one can look for $\boldsymbol{\chi}' = \ker\{\phi_2,\cdot\}$ in $(\phi_1,\boldsymbol{\chi})$, and the Frobenius condition involves the operator only on this subspace. In any case, this condition restricts the available choices for $\Phi_2$.

In coordinates $(\phi_1,\phi_2,\boldsymbol{\chi}')$, the bracket writes

$$\mathbb{J}_s = \begin{pmatrix} 0 & -b^\dagger & 0 \\ b & d & 0 \\ 0 & 0 & \mathbb{A} \end{pmatrix}\,. \tag{12.5}$$

This loose semi-canonical bracket is more general because from a true semi-canonical bracket, it authorizes a change of coordinates $\Phi_2 \longrightarrow \Phi_2'(\phi_1,\Phi_2)$, where $\Phi_2'$ can depend on $\phi_1$. It retains the

[4]This guarantees that it satisfies the Jacobi identity.

important features of the semi-canonical coordinates because there is no coupling between $(\phi_1, \Phi_2)$ and $\boldsymbol{\chi}'$, and because $\phi_1$ is coupled only with $\phi_2$. The loose semi-canonical bracket is a way to isolate the couplings of $(\phi_1, \Phi_2)$ from $\boldsymbol{\chi}'$, but not requiring to identify coordinates $\Phi_2$ such that $\mathbb{J}^{\Phi_2,\Phi_2} = 0$, which may be pointless, and also impossible to get. In a similar way as for semi-canonical brackets, the Jacobi identity implies that both $(\phi_1, \phi_2)$ and $\boldsymbol{\chi}'$ are subalgebras, which agrees with the presence of the Frobenius condition in the definition of $\Phi_2$ (see Sec. 12.2.3).

It practice, this additional information about subalgebras is quite helpful in order to identify suitable coordinates. In some cases, it suggests the desired coordinates. For instance, for a quarter-canonical Poisson bracket, the coefficients of $\mathbb{J}^{\boldsymbol{\phi}\boldsymbol{\phi}}$ generally depend on $\boldsymbol{\chi}$, and the coefficients of $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ depend on $\boldsymbol{\phi}$; the change to semi-canonical coordinates has to absorb these dependencies in the change of variables, which can be a precise indication.

Let us consider a few examples now, and begin by the example $\phi_1 = \mathbf{B}$, which is already quarter-canonical in initial variables with $\Phi_2 = \nabla \times \mathbf{E}$. It is not semi-canonical, because the terms involving $\nabla \times \mathbf{E}$ and the magnetic field are

$$\tfrac{\nabla\times\mathbf{u}+\mathbf{B}}{2\rho} \cdot F_{\mathbf{u}} \times G_{\mathbf{u}} - F_{\mathbf{E}} \cdot G_{\mathbf{u}} + F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} - (F \leftrightarrow G)\,, \tag{12.6}$$

where the symbol $-(F \leftrightarrow G)$ indicates antisymmetry, i.e. all the previous terms are written again, with permutation of $F$ and $G$ and with inversion of the signs.

To get a semi-canonical bracket, the condition $\{\Phi_2, \Phi_2\} = 0$ is already verified, and

$$\ker\{\Phi_2, \cdot\} = (\nabla \times \mathbf{u} + \mathbf{B}, \mathbf{E}, \rho, \nabla \cdot \mathbf{u})\,.$$

Then, the complementary coordinates are defined by

$$\boldsymbol{\chi} = \ker\{\phi_1, \cdot\} \cap \ker\{\Phi_2, \cdot\} = (\nabla \times \mathbf{u} + \mathbf{B}, \rho, \nabla \cdot \mathbf{u})\,.$$

The resulting structure is semi-canonical. It is simply the well-known structure obtained using the potential vector, and the canonical momentum instead of the kinetic momentum. Another way to identify this structure could have been to try to combine all the terms (12.6) of $\mathbb{J}^{\nabla\times\mathbf{E}\boldsymbol{\psi}}\partial_{\boldsymbol{\psi}}$ in only one term $\nabla \times \partial_{\mathbf{B}}$, which indicates the change of coordinates

$$\mathbf{u} \longrightarrow (\nabla \cdot \mathbf{u}, \nabla \times \mathbf{u})\ , \ \text{ and then } \ \nabla \times \mathbf{u} \longrightarrow \chi_1 = \nabla \times \mathbf{u} + \mathbf{B}\,.$$

Last, a third method was available. It relies on the comment above about subalgebras: the semi-canonical bracket has to induce a subalgebra for $\boldsymbol{\chi}$. Eq. (12.6) shows that the variables $(\mathbf{u}, \mathbf{B})$ must be combined $\chi_1 = \nabla \times \mathbf{u} + \mathbf{B}$, and this is enough to get the semi-canonical coordinates in this case.

Next, one can start from the other constraint $\phi_1 = \nabla \times \mathbf{E}$, as in the example in Sec. 12.2.1. It is more involved because the initial bracket is not quarter-canonical. From

$$\ker\{\phi_1, \cdot\} = (\rho, \mathbf{E}, \nabla \times \mathbf{u} + \mathbf{B})\,,$$

intermediate quarter-canonical coordinates

$$\boldsymbol{\chi}' = (\nabla \times \mathbf{u} + \mathbf{B}, \rho, \nabla \cdot \mathbf{E})$$

are defined. The secondary constraint $\phi_2$ can be chosen freely as a supplementary coordinate. It is natural to choose it as a supplementary to $\nabla \times \mathbf{u} + \mathbf{B}$ in $(\nabla \times \mathbf{u}, \mathbf{B})$, e.g. $\phi_2 = \nabla \times \mathbf{u}$, and one gets a quarter-canonical bracket, where the terms involving $\phi_1$ or $\phi_2$ are

$$\tfrac{\chi'^1}{2\rho} \cdot (\nabla \times F_{\phi_2} + \nabla \times F_{\chi'^1}) \times (\nabla \times G_{\phi_2} + \nabla \times G_{\chi'^1}) + F_{\mathbf{E}} \cdot \nabla \times G_{\phi_2} - (F \leftrightarrow G)\,.$$

The freedom can be used to look for semi-canonical coordinates $(\Phi_2, \boldsymbol{\chi})$. A first way is to use the condition $\{\Phi_2, \Phi_2\} = 0$. With $\Phi_2 = \Phi_2(\nabla \times \mathbf{u}, \mathbf{B})$, the condition writes $\frac{\chi'^1}{\rho} \cdot \Phi_{2\mathbf{u}} \times \Phi_{2\mathbf{u}} = 0$, which indicates $\Phi_2 = \mathbf{B}$ as the simplest choice. Then the standard procedure can be used to get semi-canonical coordinates $\boldsymbol{\chi}$, but in this case, this is useless, because the coordinates $\boldsymbol{\chi}'$ already convene.

A second way is to use the subalgebra condition; it indicates that $\nabla\times\partial_{\phi_2}+\nabla\times\partial_{\boldsymbol{\chi}'^1}$ must be absorbed in $\partial_{\boldsymbol{\chi}'^1}$, which means $\Phi_2=\phi_2-\boldsymbol{\chi}'^1$, and this is again enough to get a semi-canonical bracket.

As a third example, the primary constraint is chosen as $\phi_1=\mathbf{E}$. It illustrates the procedure when the matrix $\mathbb{J}^{\phi_1\boldsymbol{\chi}}$ involves several independent quarter-canonical sub-structures. Then $\{\phi_1,\cdot\}=-\partial_{\mathbf{u}}+\nabla\times\partial_{\mathbf{B}}$ and $\ker\mathbb{J}^{\mathbf{E}\boldsymbol{\psi}}$ evidently includes $(\rho,\mathbf{E},\nabla\cdot\mathbf{B})$. It remains to study the action of the matrix $\mathbb{J}^{\mathbf{E}\boldsymbol{\psi}}$ on the subspace $(\mathbf{u},\nabla\times\mathbf{B})$. The term $\partial_{\mathbf{u}}$ is a bijection between $\mathbf{E}$ and $\mathbf{u}$. And the term $\nabla\times\partial_{\mathbf{B}}$ is a bijection between $\nabla\times\mathbf{E}$ and $\nabla\times\mathbf{B}$. Thus, the space $\mathbf{E}$ is divided into two: $R_1=\nabla\times\mathbf{E}$ and $R_2=\nabla\cdot\mathbf{E}$. On the first space, $\mathbb{J}^{\mathbf{E}\boldsymbol{\psi}\dagger}$ is a bijection onto $\nabla\times(\partial_{\mathbf{u}}-\nabla\times\partial_{\mathbf{B}})$, and the kernel of $\mathbb{J}^{\mathbf{E}\boldsymbol{\psi}}$ is $K_1=\nabla\times(\partial_{\mathbf{u}}+\nabla\times\partial_{\mathbf{B}})$. On the second space, $\mathbb{J}^{\mathbf{E}\boldsymbol{\psi}\dagger}$ is a bijection onto $\nabla\cdot\partial_{\mathbf{u}}$, and the kernel $K_2$ of $\mathbb{J}^{\mathbf{E}\boldsymbol{\psi}}$ is zero. Thus, adopting coordinates

$$(\phi_{1a},\phi_{1b})=(\nabla\cdot\mathbf{E},\nabla\times\mathbf{E})\ ,\quad(\phi_{2a},\phi_{2b})=(\nabla\cdot\mathbf{u},\nabla\times\mathbf{u})\ \text{ and }\ \boldsymbol{\chi}=(\nabla\times(\mathbf{u}+\nabla\times\mathbf{B}),\rho,\nabla\cdot\mathbf{B})\,,$$

one gets a quarter-canonical structure both for $(\phi_{1a},\phi_{2a})$ and for $(\phi_{1b},\phi_{2b})$.

As for the semi-canonical structure, the one corresponding to $(\phi_{1b},\phi_{2b})$ was already presented in the previous paragraphs. The one corresponding to $(\phi_{1a},\phi_{2a})$ is not easy to get, nor sure to exist. Neither the condition $\{\Phi_2,\Phi_2\}=0$, nor the condition $\{\Phi_2,\boldsymbol{\chi}\}=0$, nor the subalgebra argument give a straightforward result.

This reminds that quarter-canonical coordinates are much easier to find than semi-canonical ones. Actually, in this chapter, quarter-canonical coordinates are presented starting from any of the initial coordinates, except $\phi_1=\mathbf{u}$ (because it is self-coupled, and in addition, the self-coupling coefficient involves $\mathbf{u}$ itself). On the contrary, the semi-canonical structure is obtained only for the couple $(\nabla\times\mathbf{E},\nabla\times\mathbf{B})$, but the result was already known in the literature.

### 12.2.3 Towards weakened requirements for the quarter-canonical structure

Among the two defining properties (8.13)-(8.14) of the quarter-canonical structure, the second one turned out to be only optional. The hypothesis $\{\phi_1,\phi_1\}=0$ was enough to guarantee quarter-canonical coordinates, for instance by applying the Darboux theorem. Now, one can consider releasing also this hypothesis by replacing it by a less stringent property, which would be enough to obtain it. The reason is that for the quarter-canonical structure, it is more efficient to work with the Frobenius theorem, and the role of the condition $\{\phi_1,\phi_1\}=0$ was only to make the Frobenius condition trivial. So strong a requirement is not needed. It is enough to have the Frobenius condition, there is no need for it to be trivial.

So, it would be interesting to identify an hypothesis allowing a self-coupling for $\phi_1$, but at the same time maintaining the Frobenius theorem available to identify quarter-canonical coordinates. In this case, the Frobenius condition would recover its full form, which, for Hamiltonian systems is related to the operator (12.2). This would mean that the matrix $\mathbb{J}^{\phi_1\phi_1}$ corresponds to a subalgebra (to within a free presence of the Casimir invariants, and free diagonal components). This subalgebra criterion may not be satisfied and a simple illustration from the fluid model is given by $\phi_1=\nabla\times\mathbf{u}$. Then, the operator involved in $\{\phi_1,\phi_1\}$ is $\nabla\times(\frac{\nabla\times\mathbf{u}+\mathbf{B}}{\rho}\times\nabla\times)$, and the Frobenius condition is not verified.

With this subalgebra hypothesis for $\phi_1$ (which is much less stringent than the condition $\{\phi_1,\phi_1\}=0$), the self-coupling can be isolated to get the corresponding quarter-canonical structure, provided it also satisfies the subalgebra criterion. The principle is to adopt the coordinates $(\phi_{1a},\phi_{1b})$ on $\phi_1$, where $\phi_{1b}=\ker\{\phi_1,\cdot\}\cap\phi_1$, and $\phi_{1a}$ are complementary coordinates. Then, $\boldsymbol{\chi}$ are chosen as complementary coordinates to $\phi_{1b}$ in $\ker\{\phi_1,\cdot\}$. Next, $\phi_2$ is chosen as a complementary coordinate to $(\phi_{1b},\boldsymbol{\chi})$ in $\ker\{\phi_{1a},\cdot\}$. Finally, in coordinates $(\phi_{1a},\phi_{1b},\phi_2,\boldsymbol{\chi})$, the bracket writes

$$\mathbb{J}=\begin{pmatrix}M&0&0&0\\0&0&-b''^{\dagger}&0\\0&b''&a&-d^{\dagger}\\0&0&d&\mathbb{A}\end{pmatrix}\,.$$

It is quarter-canonical in $(\phi_{1b},\phi_2,\boldsymbol{\chi}')$ with $\boldsymbol{\chi}'=(\boldsymbol{\chi},\phi_{1a})$, and weakened-semi-canonical in coordinates $(\phi_{1a},\boldsymbol{\chi}'')$ with $\boldsymbol{\chi}''=(\phi_{1b},\phi_2,\boldsymbol{\chi})$.

Notice however that with the subalgebra hypothesis, the equations to identify the coordinates might be difficult to solve in practice. Indeed, in the examples, the difficulties to identify $\ker\{\phi_1, \cdot\}$ often arose because of the presence of variables $\phi_i$ factorized with the corresponding derivative $\partial_{\phi_i}$, i.e. when one departs from the setting of (12.3).

## 12.3 Investigation of an explicit generic bracket

The two previous sections were interested in obtaining quarter-canonical coordinates, in order to benefit from the simplified truncation projector for Dirac reduction. One can think the other way round. Instead of transforming the system to fit with the action of the projector, one can try to identify the projector in the initial coordinate system. It is the topic of this section and the next one. More precisely, this section investigates the question by working on the explicit form of the Poisson bracket, whereas the next one will be interested in a more abstract, coordinate independent setting. Both sections proceed by first identifying the simplified projector, and then turn to examples, consequences, or interpretations of the results.

As an illustration and a motivation for the following, one can want to perform the reduction to an incompressible fluid, but keeping $\rho$, $\mathbf{u}$ as the field coordinates. The quarter-canonical structure exists, as was shown in Chapter 8, but it is not explicit in the Poisson bracket.

In order to describe this kind of situations, we consider a general setting where both constraints involve operators. The field variables are $\boldsymbol{\psi}$. The constraints are $\boldsymbol{\phi} := \Big(\hat{g}(\psi_1), \hat{h}(\psi_2)\Big)$, where $\hat{g}$ and $\hat{h}$ are functions of the field and their derivatives. The field derivative of the constraints are

$$(g, h) := (\phi_{1\boldsymbol{\psi}}, \phi_{2\boldsymbol{\psi}}) \,,$$

where the operators $\phi_{i\boldsymbol{\psi}}$ are defined by

$$\tfrac{\delta\phi_i(\mathbf{x})}{\delta\boldsymbol{\psi}(\mathbf{y})} = \phi_{i\boldsymbol{\psi}}(\mathbf{x})\delta(\mathbf{x}-\mathbf{y}) = \phi_{i\boldsymbol{\psi}}^{\dagger}(\mathbf{y})\delta(\mathbf{y}-\mathbf{x}) \,,$$

for $i \in \{1, 2\}$. The particular case of an incompressible fluid will corresponds to

$$\hat{g}(\psi_1) = \rho \;, \quad \hat{h}(\psi_2) = \nabla\cdot\mathbf{u} \;, \quad \psi_1 = \rho - \rho_0 \;, \quad \psi_2 = \mathbf{u} \,.$$

The existence of a quarter-canonical structure corresponds to two conditions: the first one is $\{\hat{g}(\psi_1), \hat{g}(\psi_1)\} = 0$, and the second one is the existence of coordinates $\boldsymbol{\chi}$ complementary to $\hat{g}(\psi_1), \hat{h}(\psi_2)$ such that $\{\hat{g}(\psi_1), \boldsymbol{\chi}\} = 0$. Those are not simple to explicit in the initial coordinates, and we simplify the requirements: we rather assume that $\{\psi_1, \psi_1\} = 0$. So, in coordinates $\boldsymbol{\psi}$, the Poisson bracket writes

$$\mathbb{J} = \begin{pmatrix} 0 & -b^{\dagger} & -v^{\dagger} \\ b & a & -d^{\dagger} \\ v & d & \mathbb{A} \end{pmatrix} . \tag{12.7}$$

This is more general than the quarter-canonical structure, in the sense that it does not imply $\{\phi_1, \boldsymbol{\chi}\} = 0$, but it is more restrictive than the quarter-canonical structure in the sense that requiring $\{\psi_1, \psi_1\} = 0$ is stronger than requiring $\{\phi_1, \phi_1\} = 0$.

These assumptions agree with most of the examples considered in Chapters 8 and 13, when they are studied in the natural coordinates $(\rho, \mathbf{u}, \mathbf{E}, \mathbf{B})$. For instance, when the initial (natural) variables are kept for the incompressible reduction, the electrostatic reduction, or the dipolar approximation (uniform electric field), then operators are involved in the constraints. The quarter-canonical structure is not exactly present in the initial coordinates $\psi_1, \phi_2$, but there is no self-coupling for $\psi_1$.

### 12.3.1 Simplified projector

In order to perform the Dirac reduction on this system, the matrix of constraints

$$\mathbb{C} = \begin{pmatrix} 0 & -gb^{\dagger}h^{\dagger} \\ hbg^{\dagger} & hah^{\dagger} \end{pmatrix}$$

must be invertible, i.e. $hbg^\dagger$ must be invertible. The invertibility is only modulo the kernel of the operator $g^\dagger$, since it is the effective space for the constraint $\phi_1$. This corresponds to the standard (strong) invertibility condition. Actually, for the Dirac bracket to provide the expected reduction, what is needed is just that the constraints become Casimir invariants [35]. For instance for $\phi_1$, it writes $bg^\dagger(hbg^\dagger)^{-1}(hbg^\dagger) = bg^\dagger$. For this weak invertibility, the condition is just $\ker(h) \cap \mathrm{Rg}\,(bg^\dagger) = 0$, and it is also the condition for the projector $g^\dagger(hbg^\dagger)^{-1}hb$ to be unique.

Then the inverse of the matrix of constraints is

$$\mathcal{D} = \mathbb{C}^{-1} = \begin{pmatrix} \gamma & \alpha \\ -\alpha^\dagger & 0 \end{pmatrix},$$

with $\alpha = (hbg^\dagger)^{-1}$ and $\gamma = \alpha hah^\dagger\alpha^\dagger$. The Dirac projector is given by

$$1 - \mathcal{P}_* = \boldsymbol{\phi}_{\boldsymbol{\psi}}^\dagger \mathcal{D} \boldsymbol{\phi}_{\boldsymbol{\psi}} \mathbb{J} = \begin{pmatrix} \alpha hb & \alpha ha(1-h^\dagger\alpha^\dagger gb^\dagger) & \alpha h(d^\dagger + ah^\dagger\alpha^\dagger gv^\dagger) \\ 0 & \alpha^\dagger gb^\dagger & \alpha^\dagger gv^\dagger \\ 0 & 0 & 0 \end{pmatrix}. \tag{12.8}$$

The Dirac bracket $\mathbb{J}_* = \mathcal{P}_*^\dagger \mathbb{J} \mathcal{P}_*$ is complicated to write. Instead, it can be written with a simplified projector, which makes its meaning clearer:

$$\mathbb{J}_* = \mathcal{P}^\dagger \mathbb{J}'_* \mathcal{P}, \tag{12.9}$$

$$\text{with } \mathcal{P} = \begin{pmatrix} 1-g^\dagger\alpha hb & 0 & 0 \\ 0 & 1-h^\dagger\alpha^\dagger gb^\dagger & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and } \mathbb{J}'_* = \mathbb{J} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -d^\dagger - ah^\dagger\alpha^\dagger gv^\dagger \\ 0 & d + vg^\dagger\alpha ha^\dagger & vg^\dagger\alpha hd^\dagger - dh^\dagger\alpha^\dagger gv^\dagger + vg^\dagger\gamma gv^\dagger \end{pmatrix}. \tag{12.10}$$

Notice that $\mathcal{P}$ is the simplified projector corresponding to a bracket truncation. The reduced bracket is not given by a truncation of the initial bracket, because $\mathbb{J}'_* \neq \mathbb{J}$. But the presence of the projector $\mathcal{P}$ guarantees that the constraints are Casimir invariants of the reduced bracket, since $(1 - g^\dagger\alpha hb)g^\dagger = 0$ and $(1 - h^\dagger\alpha^\dagger gb^\dagger)h^\dagger = 0$.

When $v = 0$ (or when $d = a = 0$, but it is more strictly restrictive), then $\mathbb{J}'_* = \mathbb{J}$ and $\mathcal{P}$ is a truncation projector. It precisely corresponds to the case where the initial bracket writes as a quarter-canonical matrix[5]. For the following, we will suppose it is the case: $v = 0$.

In the incompressible reduction, for instance, this condition is fulfilled and the simplified projector in the initial coordinates is

$$\mathcal{P}_* = \begin{pmatrix} 1-g^\dagger\alpha hb & 0 & 0 \\ 0 & 1-h^\dagger\alpha^\dagger gb^\dagger & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1-\nabla\Delta^{-1}\nabla\cdot & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

### 12.3.2 Relation with the quarter-canonical structure

In the previous subsection, the presence of the operators $\hat{g}$ and $\hat{h}$ in the definition of the constraints means that the reduced bracket is not exactly quarter-canonical. All the same, in Eqs. (12.8)-(12.10), the vicinity of a quarter-canonical structure in initial coordinates is traduced by the reduced bracket being given by the simplified projector $\mathcal{P}$. The reduced bracket is a projection of the initial bracket for the parts of the bracket that are already put into quarter-canonical form. This generalizes the result about simplified projectors for Dirac brackets.

Notice that because of the absence of quarter-canonical coordinates, the projector involves terms (namely caused by $g$ and $h$) on the left of the bracket coefficient $b$. It does not corresponds exactly to a bracket truncation. Thus, the generalization only concerns Dirac brackets being given by simplified projectors, but not Dirac brackets being given by bracket truncation.

Even when the reduction is quarter-canonical, i.e. there exist coordinates $\boldsymbol{\chi}$ such that $\{\phi_1, \boldsymbol{\chi}\} = 0$, the use of non quarter-canonical coordinates is traduced by some coefficients being not 1 but a projector, e.g. $1 - h^\dagger\alpha^\dagger gb^\dagger$. This projector is just the projector onto truncated quarter-canonical coordinates, but it is obtained remaining in natural variables. This is an advantage of the Dirac reduction over the subalgebra reduction (see Chapter 13): it provides a projector when adapted coordinates cannot be guessed. As usual, let us illustrate it by the incompressible reduction: the variable complementary to the constraint $\phi_2$ is provided by the procedure:

$$\mathbf{u}_S = (1 - h^\dagger\alpha^\dagger gb^\dagger)\mathbf{u} = (1 - \nabla\Delta^{-1}\nabla\cdot)\mathbf{u},$$

[5] Notice that this does not mean that the bracket is quarter-canonical, since $\dim\psi_1 \neq \dim\psi_2$ in general. In addition, the coordinates are $(\psi_1, \psi_2)$, i.e. they are different from the constraints $(\phi_1, \phi_2)$).

it is not chosen as an a priori coordinate.

As a result, quarter-canonical coordinates can be defined a posteriori by the projectors $\mathcal{P}_2 = h^\dagger\alpha^\dagger g b^\dagger$ and $1-\mathcal{P}_2$. Then the bracket becomes quarter-canonical on the surface $\psi_2 = 0$, and in the whole space when $\mathcal{P}_2$ does not depend of $\psi_2$ or at least $\{\phi_1, \mathcal{P}_2\} = 0$. Indeed the field coordinates can be taken as $(\psi_{1a}, \psi_{1b}, \psi_{2a}, \psi_{2b}, \psi_{i\geqslant 3})$, with $\psi_{1a} = \phi_1$, $\psi_{1b}$ any supplementary coordinate on $\psi_1$, $\psi_{2a} = \mathcal{P}_2\psi_2$ and $\psi_{2b} = (1-\mathcal{P}_2)\psi_2$. Then, $\{\psi_{1a}, \psi_1\} = 0$ and $\{\psi_{1a}, \psi_{i\geqslant 3}\} = 0$ were already guaranteed, and now one has (at least on the surface $\psi_2 = 0$) $\{\psi_{1a}, \psi_{2a}\} = -gb^\dagger\mathcal{P}_2 = -gb^\dagger$, and thus $\{\psi_{1a}, \psi_{2b}\} = -gb^\dagger(1-\mathcal{P}_2) = 0$. As the matrix of constraints is invertible, one has $\dim \psi_{1a} = \psi_{1b}$, and the matrix is indeed quarter-canonical.

### 12.3.3 A projection by changing coordinates twice

In the results of the previous subsections, expressions such as $h^\dagger\alpha^\dagger g b^\dagger$ show up in the projector. They can be interpreted as traducing a projection (or a truncation) by changing coordinates forwards and then backwards. In adapted coordinates $(\phi_1, \phi_2, \boldsymbol{\chi})$, the simplified projector $\mathcal{P}$ is given by the matrix (8.6). In other coordinates $\boldsymbol{\psi}$, the projector $\mathcal{P}$ is given by the same matrix expressed in a different coordinate system. Its expression is exactly

$$1 - \mathcal{P} = (\partial_{\boldsymbol{\psi}}\phi_i)^\dagger(\partial_{\phi_i}\boldsymbol{\psi})^\dagger\,. \tag{12.11}$$

When the projector $h^\dagger\alpha^\dagger g b^\dagger$ is studied from this point of view, the factor $\alpha^\dagger g b^\dagger$ can be viewed as a change (and a choice) of coordinates adapted to the constraint $\phi_2$, and the factor $h^\dagger$ corresponds to the change of coordinates back to $\psi_2$. Other interpretations are possible, although less natural. For instance, $gb^\dagger$ can be considered as the transformation forwards, then $h^\dagger\alpha^\dagger$ corresponds to the transformation backwards. There is an arbitrariness in the choice of the second space (the range of the transformation forwards). An intrinsic choice is to take it as identical to the initial space (or rather a subspace immersed in the initial space). In this case, $h^\dagger\alpha^\dagger g b^\dagger$ is interpreted as the factor for the transformation forwards, i.e. as a coordinate for the constraint $\phi_2$, that lies in the same space as $\psi_2$. Then the factor corresponding to the change of coordinate back to $\psi_2$ is just a factor 1 (it is an immersion).

### 12.3.4 Orthogonal projector

Another interesting feature of this example is that it explains why the Dirac projector is often an orthogonal projector, but it confirms that it may not be so. An orthogonal projector is obtained when the constraints are such that $h^\dagger\alpha^\dagger g b^\dagger$ and $g^\dagger\alpha h b$ are self-adjoint, e.g. for the incompressible reduction, where $\alpha$ is self-adjoint, $h = gb^\dagger$ and $g = hb$.

More generally, this case just corresponds to the natural situation for quarter-canonical structures, i.e. the natural $J$-secondary constraint. Indeed, starting from a constraint $\phi_1 = \hat{g}(\psi_1)$, $J$-secondary constraints are defined by the condition that $\{\phi_1, F\} = -gb^\dagger\phi_{2\psi_2}^\dagger$ is an invertible operator in the weak sense, i.e. $\mathrm{Rg}\,\phi_{2\psi_2}^\dagger \cap \ker gb^\dagger = 0$. A natural choice is $\phi_{2\psi_2} = gb^\dagger$. If $gb^\dagger$ is independent of $\psi_2$, it corresponds to $\phi_2 = gb^\dagger\psi_2$. In any case, it means that $\{\phi_1, \phi_2\} = (gb^\dagger)(gb^\dagger)^\dagger$ is self-adjoint, and the projectors (e.g. $(gb^\dagger)^\dagger\alpha^\dagger(gb^\dagger)$) are orthogonal. Another point of view is that $(gb^\dagger)^\dagger\alpha^\dagger(gb^\dagger)$ is 1 on $\mathrm{Rg}\,(gb^\dagger)^\dagger$, which means that $\phi_2$ is defined such that $\mathbb{J}^{\phi_1\phi_2} = 1$ on the space of constraints. This is indeed the most natural $J$-secondary constraint.

It is obvious but useful to remind that this natural case is not the only available choice, and the simplified projector is not forced to be orthogonal. For instance, in the reduction to an incompressible fluid, the choice $\phi_2 = \nabla\cdot\mathbf{u}$ corresponds to the orthogonal projector $\nabla\Delta^{-1}\nabla\cdot$, whereas the choice $\phi_2 = \nabla\cdot(\rho\mathbf{u})$ corresponds to the projector $\nabla(\nabla\cdot\rho\nabla)^{-1}\nabla\cdot\rho$ [33], that is not orthogonal.

## 12.4 Coordinate-independent formulation

In the last section, the explicit form of the Poisson bracket was used to study how quarter-canonical effects can indeed be observed in coordinates that are not quite quarter-canonical. The drawback of

the method is that it did not deal exactly with a quarter-canonical structure, and it was restricted, because its results can be applied only to situations that fit with its form of Poisson bracket.

On the other hand, the results emphasized the distinction between the truncation projector, which exists only for a quarter-canonical structure, and the simplified projector, which exists in a more general context. This suggests to identify in a more abstract perspective the conditions and effects of the simplified projector in contrast with the truncation projector.

So, we get rid of both the coordinate system and the explicit form of the Poisson bracket. In Subsec. 12.4.1, the expression of the simplified projector is obtained, and the following subsections will study examples or consequences of the result.

Arbitrary coordinates are used, and the condition $\{\phi_1, \boldsymbol{\chi}\} = 0$ (which is coordinate dependent) is released. We only assume that $\{\phi_1, \phi_1\} = 0$. This condition is coordinate independent, since it depends only on the constraints themselves, and more precisely on the primary constraint, which is fixed at the beginning of the Dirac reduction. As usual, the Dirac projector is defined by

$$\begin{aligned} 1 - \mathcal{P}_* =& (\partial_{\boldsymbol{\psi}}\phi_\alpha)^\dagger \mathcal{D}^{\alpha\beta}\partial_{\boldsymbol{\psi}}\phi_\beta \mathbb{J} \\ =& (\partial_{\boldsymbol{\psi}}\phi_1)^\dagger \mathcal{D}^{11}\partial_{\boldsymbol{\psi}}\phi_1 \mathbb{J} + (\partial_{\boldsymbol{\psi}}\phi_1)^\dagger \mathcal{D}^{12}\partial_{\boldsymbol{\psi}}\phi_2 \mathbb{J} \\ &+ (\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{21}\partial_{\boldsymbol{\psi}}\phi_1 \mathbb{J} + (\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{22}\partial_{\boldsymbol{\psi}}\phi_2 \mathbb{J}\,, \end{aligned} \tag{12.12}$$

where $\mathcal{D}^{\alpha\beta} = (\mathcal{C}^{-1})^{\alpha\beta}$ is the inverse of the matrix of constraints.

### 12.4.1 Simplified projector

The role of $\mathcal{P}_*$ is to give the Dirac reduced bracket through $\mathbb{J}_* = \mathcal{P}_*^\dagger \mathbb{J} = \mathbb{J}\mathcal{P}_* = \mathcal{P}_*^\dagger \mathbb{J}\mathcal{P}_*$. On another hand, in the specific case where the condition $\{\phi_1, \phi_1\} = 0$ is verified, the following relations hold:

$$D^{22} = 0\,, \qquad \mathcal{D}^{12} = -\mathcal{D}^{21\dagger} = \mathcal{C}_{21}^{-1}\,, \qquad \mathcal{D}^{11} = -\mathcal{D}^{12}\mathcal{C}^{22}\mathcal{D}^{21} = -\mathcal{D}^{12}\partial_{\boldsymbol{\psi}}\phi_2 \mathbb{J}(\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{21}\,.$$

They emphasize that the matrix of constraints is invertible if and only if $\{\phi_1, \phi_2\}$ is invertible. Inserting these relation in Eq. (12.12), the Dirac bracket $\mathbb{J}\mathcal{P}_*$ can be rewritten as

$$\begin{aligned} \mathcal{P}_*^\dagger \mathbb{J}\mathcal{P}_* = \mathbb{J}\mathcal{P}_* =& \mathbb{J} - \mathbb{J}(\partial_{\boldsymbol{\psi}}\phi_1)^\dagger (\mathcal{D}^{21})^\dagger \partial_{\boldsymbol{\psi}}\phi_2 \mathbb{J}(\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{21}\partial_{\boldsymbol{\psi}}\phi_1 \mathbb{J} \\ &+ \mathbb{J}(\partial_{\boldsymbol{\psi}}\phi_1)^\dagger (\mathcal{D}^{21})^\dagger \partial_{\boldsymbol{\psi}}\phi_2 \mathbb{J} - \mathbb{J}(\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{21}\partial_{\boldsymbol{\psi}}\phi_1 \mathbb{J} \\ =& \Big(1 - (\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{21}\partial_{\boldsymbol{\psi}}\phi_1 \mathbb{J}\Big)^\dagger \mathbb{J} \Big(1 - (\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{21}\partial_{\boldsymbol{\psi}}\phi_1 \mathbb{J}\Big) = \mathcal{P}^\dagger \mathbb{J}\mathcal{P}\,, \end{aligned}$$

where $\mathcal{P}$ is the simplified projector, defined by

$$1 - \mathcal{P} = (\partial_{\boldsymbol{\psi}}\phi_2)^\dagger \mathcal{D}^{21}\partial_{\boldsymbol{\psi}}\phi_1 \mathbb{J}\,. \tag{12.13}$$

Thus the condition $\{\phi_1, \phi_1\} = 0$ is enough for the existence of a simplified projector, and in this case a coordinate-independent expression for this projector is provided by Eq. (12.13). This is much less restrictive than in Sec. (12.7), and also than the quarter-canonical case. Notice that this simplified projector does not involve the inverse of the matrix of constraint, but only the inverse of the operator $\{\psi_1, \psi_2\}$, as expected. In the case of the reduction for incompressibility, the operator $\{\psi_1, \psi_2\}$ is the Laplacian, involved in the projector $\nabla\Delta^{-1}\nabla\cdot$.

### 12.4.2 Example of application

To illustrate the interest of this formula for the simplified projector, a generic Poisson bracket with $\{\phi_1, \phi_1\} = 0$ is written in coordinates $(\phi_1, \phi_2, \boldsymbol{\chi})$, with $\boldsymbol{\chi}$ arbitrary

$$\mathbb{J} = \begin{pmatrix} 0 & -b^\dagger & -v^\dagger \\ b & a & -d^\dagger \\ v & d & \mathbb{A} \end{pmatrix}\,. \tag{12.14}$$

Be careful that the same symbols as in Eq.(12.7) are used for the coefficients of the matrix, but now the coordinates are suited to the constraints. The Dirac projector is easily obtained as

$$1 - \mathcal{P}_* = \begin{pmatrix} 1 & 0 & b^{-1}(vb^{-1}a - d)^\dagger \\ 0 & 1 & b^{-1\dagger}v^\dagger \\ 0 & 0 & 0 \end{pmatrix}\,.$$

The Dirac bracket writes

$$\mathbb{J}_* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbb{A}+vb^{-1}ab^{-1\dagger}v^\dagger+vb^{-1}d^\dagger-db^{-1\dagger}v^\dagger \end{pmatrix} .$$

As for the projector given by (12.13), it is quite simplified compared to $\mathcal{P}_*$ :

$$1-\mathcal{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & b^{-1\dagger}v^\dagger \\ 0 & 0 & 0 \end{pmatrix} .$$

But it is easy to check that it actually gives the same reduced Poisson bracket. The use of non-quarter-canonical coordinates is traduced by the simplified projector being different from a mere truncation projection. So, the projector (12.13) can be used to simplify the computations even when the bracket is not quarter-canonical.

### 12.4.3 Case of quarter-canonical coordinates

In the specific case where the bracket is quarter-canonical (i.e. $v = 0$ in Eq. (12.14)), all the results are simplified, but the simplified bracket (12.13) still keeps its interest. The Dirac projector $\mathcal{P}_*$ is easily obtained with Eq. (8.17), and it provides the Dirac bracket $\mathbb{J}_*$, which results in a bracket truncation in this case:

$$1-\mathcal{P}_* = \begin{pmatrix} 1 & 0 & -b^{-1}d^\dagger \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \qquad \mathbb{J}_* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbb{A} \end{pmatrix} .$$

So, the Dirac projector can be replaced by the truncation projector $\mathcal{P}_t$. Now, the projector $\mathcal{P}$ given by (12.13) is still simpler, since they are given by

$$1-\mathcal{P}_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \quad \text{and} \quad 1-\mathcal{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{12.15}$$

Thus, the projector $\mathcal{P}$ from formula (12.13) is the most interesting, because of the simplicity of its result, because it is available in arbitrary coordinates, and because it relies on the structure of the bracket itself: it corresponds to taking the subalgebra without $\phi_2$ (see Chapter 13).

### 12.4.4 When some coefficients are operators

We now come back to the case considered in the previous section. Notice that in the general case (where $v \neq 0$), we had not identified a simplified projector to replace the Dirac bracket. The result of the previous subsections show that actually the simplified projector (12.13) exists. It writes

$$1-\mathcal{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & h^\dagger\alpha^\dagger gb^\dagger & h^\dagger\alpha^\dagger gv^\dagger \\ 0 & 0 & 0 \end{pmatrix} .$$

It is indeed much simpler than the Dirac projector (12.8). When the bracket is quarter-canonical, it becomes trivial

$$1-\mathcal{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & h^\dagger\alpha^\dagger gb^\dagger & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

It is simplier both than the Dirac projector (given by (12.8) with $v = 0$), and than the truncation projector obtained in (12.10). Its action is just to project the component $\psi_2$ in such a way as to make $\phi_2$ become a Casimir invariant. For example, for the incompressible reduction, it straightforwardly gives the projector $\nabla\Delta^{-1}\nabla\cdot$ whether the coordinates used are adapted ones or not.

### 12.4.5 A few consequences

Another interest of formula (12.13) is that it shows that the range of the simplified projector is $\mathrm{Rg}\,({\phi_2}_{\boldsymbol{\psi}}^\dagger)$ and its kernel is $\ker({\phi_1}_{\boldsymbol{\psi}}\mathbb{J})$. The projector is orthogonal exactly when those spaces are orthogonal. The (strong) invertibility condition implies the inverse of ${\phi_1}_{\boldsymbol{\psi}}\mathbb{J}{\phi_2}_{\boldsymbol{\psi}}^\dagger$ to be unique between $\mathrm{Rg}\,({\phi_1}_{\boldsymbol{\psi}}\mathbb{J})$ and a complementary set to $\ker({\phi_2}_{\boldsymbol{\psi}}^\dagger)$. This means that $\mathrm{Rg}\,({\phi_2}_{\boldsymbol{\psi}}^\dagger)$ and $\ker({\phi_1}_{\boldsymbol{\psi}}\mathbb{J})$ are complementary, i.e. $\mathrm{Rg}\,({\phi_1}_{\boldsymbol{\psi}}\mathbb{J}{\phi_2}_{\boldsymbol{\psi}}^\dagger) = \mathrm{Rg}\,({\phi_1}_{\boldsymbol{\psi}}\mathbb{J})$. The weak invertibility condition means that

the inverse does not have to be unique, but only unique modulo $\ker(\mathbb{J}\phi_{2\boldsymbol{\psi}}^{\dagger})$, which corresponds to $\mathrm{Rg}\,(\mathbb{J}\phi_{2\boldsymbol{\psi}}^{\dagger}) \cap \ker(\phi_{1\boldsymbol{\psi}}) = \{0\}$.

In a similar way as in Subsec. 12.3.2, in case $\{\phi_1, \mathcal{P}\} = 0$, e.g. when $\mathcal{P}$ does not depend on the field variables, then this projector can be used to identify $\phi_2$ and build quarter-canonical coordinates $(\phi_1, \phi_2, \boldsymbol{\chi})$, with $\phi_2 = (1 - \hat{\mathcal{P}})\boldsymbol{\psi}$, and $(\phi_1, \boldsymbol{\chi}) = \hat{\mathcal{P}}\boldsymbol{\psi}$. Indeed, in these coordinates, the Poisson bracket is quarter-canonical since

$$\{\phi_1, \hat{\mathcal{P}}\boldsymbol{\psi}\} = \phi_{1\boldsymbol{\psi}}\mathbb{J}\left(1 - \phi_{2\boldsymbol{\psi}}^{\dagger}(\phi_{1\boldsymbol{\psi}}\mathbb{J}\phi_{2\boldsymbol{\psi}}^{\dagger})^{-1}\phi_{1\boldsymbol{\psi}}\mathbb{J}\right) = 0\,,$$
$$\text{and}\quad \{\phi_1, (1 - \hat{\mathcal{P}})\boldsymbol{\psi}\} = \phi_{1\boldsymbol{\psi}}\mathbb{J}\phi_{2\boldsymbol{\psi}}^{\dagger}(\phi_{1\boldsymbol{\psi}}\mathbb{J}\phi_{2\boldsymbol{\psi}}^{\dagger})^{-1}\phi_{1\boldsymbol{\psi}}\mathbb{J} = \phi_{1\boldsymbol{\psi}}\mathbb{J}$$

is (weakly) invertible. This is another way to see that the quarter-canonical structure is implied as soon as $\{\phi_1, \phi_1\} = 0$, even if computations can be done with other coordinates, as was the goal of this section.

## Conclusion

This chapter investigated the stiffness of the condition on the coordinates to benefit from the simplified truncation projector for Dirac reductions in the presence of a quarter-canonical structure. The condition appeared as easily satisfied as soon as $\{\phi_1, \phi_1\} = 0$. Quarter-canonical coordinates are then very common, and especially in Dirac reductions. In finite dimensional systems, it is even the only possible structure, as a result of the Darboux theorem, or rather of the Frobenius theorem. More generally, the same conclusions are most often observed also in field theory. A practical motivation for the high incidence of the quarter-canonical structure is that it is much less stringent than the semi-canonical one, while it retains its main advantages.

On another hand, when they are not already present in the Poisson bracket, quarter-canonical coordinates can most often be identified by hand. This is because linearity is most common in Poisson brackets for fluids and plasmas, and also in the definition of constraints. Many sources of obstructions can appear: algebraic, topological, computational, related to operator invertibility, to integrability, etc. But in most practical cases, such pathological effects are absent and the derivation works in a similar way as for matrix calculus.

Changing to quarter-canonical coordinates is not necessary. The effects of the quarter-canonical structure can be identified in Dirac reductions while keeping non-quarter-canonical coordinates. In this chapter, the formula for the simplified projector was obtained for a specific case where the functional derivatives involved in the quarter-canonical structure are acted upon by operators, as is suggested by frequent examples of Dirac reductions in plasmas.

More, even the existence of quarter-canonical coordinates is not needed for a simplified projector to be available, even if in this case it cannot be completely interpreted as a truncation projector. Using only the requirement that half of the constraints is not coupled with itself, an explicit coordinate-independent expression was obtained for the Dirac simplified projector, which is thus formulated in a quite general setting. This extends its domain and standardizes its application.

As the next exploration about the quarter-canonical structure, we will turn to its impacts on the Dirac procedure and its relations with the reduction method based on bracket truncations. This is the topic of the following chapter 13.

Then, we will come back to the gyrokinetic reduction and will consider the removal of its two superfluous dimensions (gyro-angle and magnetic moment) in the light of quarter-canonical reductions. This will be the topic of Sec. 8.4 in Chapter 8.

# Chapter 13

# Impacts of quarter-canonical structures on Dirac and subalgebra reduction methods

in collaboration with Cristel Chandre

*A*bstract: Consequences of the presence of a quarter-canonical structure in a Poisson bracket on Hamiltonian reduction methods are investigated. For Dirac's theory of constraints, this structure is shown to have an crucial impact on the reduction procedure, because it guarantees the inversion of the matrix of constraints and it is determinant for the available choices for secondary constraints. This induces a more complete method for identifying suitable secondary constraints.
In addition, the quarter-canonical structure gives rise to an alternative procedure based on a mere bracket truncation. It is legitimated by the presence of a Lie-subalgebra, and is best viewed in the framework of another Hamiltonian reduction method, corresponding to a projection onto a subalgebra, which offers complementary possibilities compared to the Dirac method, since it is often computationally little involved and offers much pliancy.
The results are illustrated with several examples, mainly taken from the ideal-fluid model for plasmas, for instance the reduction for an incompressible or for an electrostatic dynamics.

## Introduction

Similarly as Chapter 12, the present appendix chapter takes place just before Sec. 8.4 in Chapter 8. Indeed, the first few sections of Chapter 8 showed that the quarter-canonical structure replaced the Dirac projector by a simplified projector, related to a mere bracket truncation. We now analyse the impacts of these facts on Hamiltonian reduction methods.

Indeed, truncations usually do not preserve the Hamiltonian character, and Dirac reductions usually require the matrix of constraints to be invertible, whereas such difficulties seem to have disappeared in this case. This clearly requires further investigation. Especially, it suggests that for Dirac reductions, the quarter-canonical structure impacts the invertibility of the matrix of constraints, and therefore the choice of a secondary constraint. On the other hand, for reductions based on bracket truncations, it suggests that some structure (which will appear to be a subalgebra) is underlying in order to legitimate the truncation method from a Hamiltonian point of view; its relations with the quarter-canonical structure and with the Dirac method will be worth considering.

After clarifying those points about quarter-canonical structures in Hamiltonian reductions, we will come back to the gyrokinetic reduction and consider the removal of its two superfluous dimensions from this point of view. This will be the topic of Sec. 8.4 in Chapter 8.

The organization of this chapter is the following. In Sec. 13.1, the invertibility of the matrix

of Dirac's constraints is considered in the case of a quarter-canonical bracket. In Sec. 13.2, the effects of this structure on the available choices for the secondary constraints are investigated. A procedure shows up, linked to the structure of the bracket and of the Dirac reduction. Especially, it allows to study all the possible secondary constraints in a way that is complementary to the usual procedure involving the Hamiltonian function. In Sec. 13.3, we turn to the simplified reduction method based on the use of the truncation projector instead of the Dirac projector. It is shown to consist in a projection onto a subalgebra, which is equivalent to the Dirac method in the case of quarter-canonical reductions. In Sec. 13.4, a few aspects of the pliancy of the subalgebra method are presented, and the traces of the canonical structure in subalgebra reductions are analysed.

With regard to notational conventions, the present chapter behaves as a section of Chapter 8, where the common notations are introduced once for the three companion chapters Chapter 8, 12, and 13.

## 13.1 On the invertibility of the matrix of constraints

For quarter-canonical brackets, the Dirac procedure can give only one very simple reduced bracket: the truncated bracket, which can be identified without inverting the matrix of constraints. But in principle, the procedure needs the matrix of constraints to be invertible. In this section, we investigate this point, especially in order to clarify what would happen with the truncated bracket if the matrix was not invertible. In Subsec. 13.1.1, it is shown that for quarter-canonical brackets, the invertibility condition is automatically satisfied. This will have effects on the choice of a secondary constraint, for which strong requirements arise from the quarter-canonical structure. This will be the topic of Subsec. 13.1.3. In Subsecs. 13.1.2 and 13.1.4, a few examples will be considered.

### 13.1.1 Quarter-canonical matrices of constraints are invertible

Let us begin as usual by considering a finite-dimensional system with scalar primary constraint $\phi_1$. Then, the antisymmetry of $\mathbb{J}$ implies $\{\phi_1, \phi_1\} = 0$, hence the matrix of constraints is not invertible and a secondary constraint $\phi_2$ is needed.

The Darboux algorithm provides semi-canonical coordinates $(\phi_1, \Phi_2, \boldsymbol{\chi})$ (at least locally...). Then, the matrix of constraints is exactly given by

$$\mathbb{C} = \left(\begin{smallmatrix} 0 & -b \\ b & 0 \end{smallmatrix}\right) ,$$

with $b = -\partial\phi_2/\partial\Phi_2$. The matrix $\mathbb{C}$ is invertible if and only if $b$ is non-zero. It means that the constraint $\phi_2$ is in bijection with $\Phi_2$. So, $\Phi_2 \longrightarrow \phi_2$ is a change of coordinates. With this coordinate, the Poisson bracket writes exactly (8.12), with suitable coefficents. Thus, the quarter-canonical structure is the only available structure for the Dirac reduction with two scalar constraints.

A remark about terminology is in place. The necessary and sufficient condition above concerns only the constraints $(\phi_1, \phi_2)$. The remaining coordinates $\boldsymbol{\chi}$ can be chosen arbitrarily, since they do not affect Dirac reductions. So, the expression of the bracket is not quarter-canonical in general, and quarter-canonical Poisson brackets **are not** the only brackets available for the reduction. Nevertheless, the quarter-canonical structure **is** present, as shown in Chapter 12. This emphasizes that the *quarter-canonical structure*, which is coordinate independent, must be distinguished from the *quarter-canonical Poisson bracket*, which is coordinate dependent. Sometimes, by abuse of language, we do not distinguish between them in our phrasing, but with regard to the concepts it is important to keep the distinction in mind. For the sake of completeness, let us mention that we call *quarter-canonical reduction* a reduction associated to a quarter-canonical structure, i.e. a reduction with constraints $(\phi_1, \phi_2)$ such that there exist quarter-canonical coordinates $(\phi_1, \phi_2, \boldsymbol{\chi})$. Last, the constraints associated to some quarter-canonical reduction are called *quarter-canonical constraints*.

Let us turn now to the case where $\phi_1$ is vectorial. Then the quarter-canonical structure is no more guaranteed, since $\{\phi_1, \phi_1\}$ can be non zero, and any antisymmetric matrix. However, when

quarter-canonical coordinates exist (e.g. when $\{\phi_1, \phi_1\} = 0$), the criterion for invertibility remains simple, since it corresponds to the invertibility of the matrix $b$.

One can go beyond: even when $b$ is not invertible, the effective matrix of constraints can always be inverted. Indeed, when $b$ is not invertible, it means that a part of the constraint $\phi_1$ is a Casimir invariant, which should not be considered as a constraint (since is is already constrained).
More precisely, the coordinate $\phi_1$ can be decomposed in $(\phi_{1a}, \phi_{1b})$, where $\phi_{1a} = \ker\{\phi_2, \cdot\} \cap \phi_1$ is the Casimir invariant and $\phi_{1b}$ is a complementary coordinate with $\phi_{1b} \cap \ker\{\phi_2, \cdot\} = 0$. The fact that $\phi_{1a}$ is a Casimir invariant appears because the assumed quarter-canonical structure implies $\{\phi_{1a}, \phi_1\} = 0$, and $\{\phi_{1a}, \boldsymbol{\chi}\} = 0$, so that $\{\phi_{1a}, \boldsymbol{\psi}\} = 0$. In coordinates $(\phi_{1a}, \phi_{1b}, \phi_2, \boldsymbol{\chi})$, the bracket writes

$$\mathbb{J} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -b'^{\dagger} & 0 \\ 0 & b' & a & -d^{\dagger} \\ 0 & 0 & d & \mathbb{A} \end{pmatrix} .$$

From the point of view of the Dirac reduction, the matrix $b$ being not invertible is no trouble, because $\phi_{1a}$ is already a Casimir invariant, and one has to impose as a constraint only the dynamical part of $\phi_1$, that is $\phi_{1b}$. For the corresponding reduction, the bracket is not quarter-canonical yet, because $\dim \phi_2 \neq \dim \phi_{1b}$. The reason is that $\phi_2$ has been taken too large, not so many secondary constraints are needed for $\phi_1$, and the superfluous part of $\phi_2$ should be dropped from the constraints. More precisely, the method of the previous chapter can be used to decompose $\phi_2$ in $(\phi_{2a}, \phi_{2b})$ such that the bracket is quarter-canonical in coordinates $(\phi_{1b}, \phi_{2a}), \boldsymbol{\chi}'$, with $\boldsymbol{\chi}' = (\phi_{2b}, \boldsymbol{\chi}, \chi_{1a})$. Finally, the bracket writes

$$\mathbb{J} = \begin{pmatrix} 0 & -b''^{\dagger} & 0 & 0 & 0 \\ b'' & a_{11} & a_{12} & -d_a^{\dagger} & 0 \\ 0 & a_{21} & a_{22} & -d_b^{\dagger} & 0 \\ 0 & d_a & d_b & \mathbb{A} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} ,$$

where $b''$ is bijective by construction. The constraints $(\phi_{1b}, \phi_{2a})$ gives an invertible matrix of constraints, together with a reduced bracket where $\phi_1$ is a Casimir invariant, as was desired.

This development evidences a subtlety about quarter-canonical structures. Through the above procedure, the quarter-canonical structure has been reduced to its minimal size, in the sense that the dimension of the constraint $\phi_1$ has been removed its component that is already a Casimir invariant. The corresponding structure will be called *minimal quarter-canonical*. We will ordinarily assume that the quarter-canonical structures have been made minimal. This, in case the quarter-canonical structure exists, its matrix of constraints is always invertible, i.e. the Dirac procedure always works.

Reciprocally, it is straightforward to show that in this case, the matrix is invertible only if the reduction is quarter-canonical. The argument follows exactly the same line as in the case of two scalar constraints.

In conclusion, even for vectorial constraints $\phi_1$ and $\phi_2$, when the quarter-canonical structure exists (i.e. at least locally provided $\{\phi_1, \phi_1\} = 0$), the matrix of constraint is invertible if and only if the constraints are quarter-canonical.

In the case of continuous media, the same argument can be considered. As usual, the difficulty is that there is no general proof. So, $\ker\{\phi_{1b}, \cdot\}$ is not sure to define good coordinates $\phi_{2b}$. All the same, in practical examples, it is often the case.

### 13.1.2 Examples: guiding-center and Euler-Poisson

As an illustration from finite-dimensional systems, we already mentioned in Sec. 12.1.2 the guiding-center reduction, for which the Dirac constraint $\phi_1 = \mu$ can be imposed only with the second constraint $\phi_2 = \theta$ (or a function in bijection with $\theta$), so as to make the matrix of constraints invertible. Then the reduction is quarter-canonical, and the reduced bracket is just the truncated one, as expected for guiding-center dynamics.

As for an example concerning the fluid model, if the Dirac constraint are $(\phi_1, \phi_2) = (\mathbf{B}, \mathbf{E})$, the matrix $b$ (in the natural coordinates) is $b = \nabla\times$. It is not invertible whereas the Poisson bracket is quarter-canonical. This is because $\mathbf{B}$ is not the true dynamical variable: $\phi_{1a} = \nabla \cdot \mathbf{B}$ is a Casimir invariant, the dynamical part of $\mathbf{B}$ corresponds to $\phi_{1b} = \nabla \times \mathbf{B}$, and it is exactly coupled

with $\phi_{2a} = \nabla \times \mathbf{E}$, i.e. the coupling is bijective between these variables. With the coordinates $(\nabla \times \mathbf{B}, \nabla \times \mathbf{E})$, the bracket is minimal quarter-canonical and the matrix of constraints is invertible. Then the Dirac method gives exactly the reduction to the electrostatic field (with external magnetic field) (see Chapter 7). In this example, none of the choices $\rho$, $\nabla \cdot \mathbf{E}$, $\nabla \cdot \mathbf{u}$, $\nabla \times \mathbf{u}$, $s$ nor any function of them would give an invertible matrix of constraints: the matrix of constraint is invertible if and only if $\phi_2$ is in bijection with $\nabla \times \mathbf{E}$, which is the variable quarter-conjugated with the (solenoidal part of the) magnetic field.

### 13.1.3 Consequence on the choice of secondary constraints

The previous subsections showed that when the quarter-canonical structure exists, the invertibility of the matrix of constraints was equivalent to choosing for the secondary constraint a quantity quarter-conjugated to $\phi_1$. An immediate consequence is that the available secondary constraints are dictated by the quarter-canonical structure.

Once the primary constraint is chosen and an arbitrary "conjugated" variable $\phi_2'$ is identified, then the available secondary constraints are exactly known: they are all the variables $\phi_2(\phi_1, \phi_2', \boldsymbol{\chi})$ in bijection with $\phi_2'$, i.e. all the variables that preserve the quarter-canonical structure.

Notice that the procedure does not imply to build quarter-canonical coordinates for $\boldsymbol{\chi}$. In addition, even the need to first identify a "conjugated" variable $\phi_2'$ can be removed. Indeed, once the primary constraint $\phi_1$ is chosen (such that $\{\phi_1, \phi_1\} = 0$, in order to ensure the presence of a quarter-canonical structure), the condition for the matrix of constraints to be invertibility exactly shrinks to the condition that $\{\phi_1, \phi_2\}$ be invertible. Thus, the possible solutions for $\phi_2$ can be read directly in the functional formula $\{\phi_1, F\}$, considered as an operator acting on $F$ which is to be made invertible by a suitable choice for $F$.

These possible solutions for $\phi_2$ will be called *J-secondary constraints*, because they are defined only from the Poisson bracket $\mathbb{J}$, they have nothing to do with the Hamiltonian. They are different from the traditional secondary constraints, since the usual procedure consists in adopting as the secondary constraints just the time evolution of the primary constraint $\phi_2 = \dot{\phi}_2 = \{\phi_1, H\}$. This (unique) last choice will be called *H-secondary constraint*, because its definition depends on the Hamiltonian function. The links between the two kinds of secondary constraints will be explored with more details in the next section.

### 13.1.4 The Euler-Poisson by constraining the electric field

As an illustration taken from the fluid model, let us come back to the previous example, but inverting the roles of the constraints. If the first Dirac constraint is the solenoidal part of the electric field $\phi_1 := \nabla \times \mathbf{E}$, then the matrix of constraints writes $\begin{pmatrix} 0 & \{\phi_1,\phi_2\} \\ \{\phi_2,\phi_1\} & \{\phi_2,\phi_2\} \end{pmatrix}$, its invertibility relies on the condition that the operator $\{\phi_1, \phi_2\} = \nabla \times \nabla \times \phi_{2\mathbf{B}} + \nabla \times \phi_{2\mathbf{u}}$ be invertible. For instance, it is not invertible if $\phi_2$ is chosen to be $\mathbf{u}$, or $\rho\mathbf{u}$, or $\nabla \times (\rho\mathbf{u})$ for instance. On the contrary, it is invertible if $\phi_2 = \nabla \times \mathbf{B}$, which gives the electrostatic reduction, or if $\phi_2 = \nabla \times \mathbf{u}$, or any variable in bijection with any combination of them. This example evidences that the bracket does not need to be quarter-canonical in the used coordinates $\boldsymbol{\chi}$: here both $\nabla \times \mathbf{E}$ and $\nabla \times \mathbf{u}$ are coupled with several fields, not just with each other. Only the condition $\{\phi_1, \phi_1\} = 0$ is needed. But when the coordinates are not quarter-canonical, the reduction is not just a bracket truncation. The Dirac bracket must be computed, unless the simplified projector (12.13) derived in Chapter 12 is used.

All these $J$-secondary constraints were obtained without the Hamiltonian (8.1). Now, the $H$-secondary constraint is $\phi_2 = \nabla \times (-\rho\mathbf{u} + \nabla \times \mathbf{B})$. It is only one of the possible solutions (since it is in bijection with $\nabla \times \mathbf{B}$), but it is not one of the simplest ones, and it is not the one corresponding to the electrostatic reduction.

## 13.2 On the procedure for secondary constraints

The previous section showed that the quarter-canonical canonical structure dictates the available choices of secondary constraints. It provided a procedure to identify the possible secondary con-

straints, called $J$-secondary constraints, in contrast with the traditional method, which provided one single constraint, called $H$-secondary constraints. Let us study a bit more the links between $J$- and $H$-secondary constraints. In Subsec. 13.2.1 the differences between them will be emphasized. In Subsec. 13.2.2, it is shown why they sometimes give the same result, but why they often do not. Last, Subsec. 13.2.3 will illustrate these concepts and conclusions through a detailed study of the reduction for incompressibility.

### 13.2.1 Natural $H$-constraints, but structural $J$-constraints

The procedure identified above for $J$-secondary constraints exactly indicates what secondary constraints make the Dirac reduction work or not. It relies on the structure of the Poisson bracket, and has nothing to do with the Hamiltonian function $H$. This is comfortable since the invertibility of the matrix of constraints is purely related to the Poisson bracket.

On the contrary, the usual procedure is to compute the secondary constraint, using the Hamiltonian according to the formula $\phi_2 := \dot{\phi}_1 = \{\phi_1, H\}$. It seems quite trustful, because it evokes some consistency in the model: in order for $\phi_1$ to be constant, it seems obviously needed to make constant its time variation as well.

This is but an appearance, this kind of consistency is neither necessary nor sufficient for the Dirac reduction. It is not sufficient because often it gives a set of constraints $(\phi_1, \phi_2)$ that does not make the matrix of constraints invertible, even when there exist some choices that are both physically relevant and would make the matrix invertible. And the consistency is not necessary either because most often there is no need to impose a third constraint, whereas a true consistency should require to impose $\{\phi_2, H\}$ as a constraint as well. This idea of consistency does not tell why most often the consistency implies only the second constraint. An additional awkward feature confirms that the procedure is not so consistent at all: it is most often not reversible. If $\phi_2 = \{\phi_1, H\}$ is taken as the first constraint, the associated $H$-secondary constraint $\phi_1' := \{\phi_2, H\}$ is completely different from $\phi_1$ in general, as will be illustrated in Subsec. 13.2.3. Nothing in the process tells which constraint is to be used as the primary one.

On the contrary, the procedure relying on $J$-secondary constraints gives both necessary and sufficient conditions[1] on $\phi_2$ for the set $(\phi_1, \phi_2)$ to make the matrix of constraints invertible. In addition, it often indicates several elementary possible choices. It explains why the consistency most often involves exactly two constraints and relates it with the canonical structure or with its traces, which are not really absent, even in continuous media [159]. Last, it is reversible: if $\phi_2$ is a $J$-secondary constraint associated to the primary constraint $\phi_1$, then $\phi_1$ is a $J$-secondary constraint associated to the primary constraint $\phi_2$.

All the same, most often the method for $J$-secondary constraints misses some information. Contrary to the method for $H$-constraints, it offers several possible solutions, but cannot indicate which of them have physical relevance. This choice should be guided by the physics of the problem. It is at this point that the $H$-secondary constraint may be useful. But this $H$-secondary constraint gives only an indication, not a necessary condition, even with respect to the physical relevance of the constraint $\phi_2$ (see the electrostatic reduction in Sec. 13.1.4, and the local incompressibility in Sec. 13.2.3).

### 13.2.2 Frequent equivalence

There is an additional point to be explained: the reason why $H$-secondary constraint are often acceptable, and why they also often give absurd results as well. The answer comes from the special form of the Hamiltonian function, which is often quadratic in variables that are quarter-canonical. This special case implies $\{\phi_1, H\} \propto (\delta/\delta\phi_2)(\phi_2)^2 \propto \phi_2$, and the $H$-secondary constraint is just the most natural $J$-secondary constraint. On another hand, when the Hamiltonian is not exactly quadratic in $\phi_2$, it is often close to it. For instance, $\rho\mathbf{u}^2$ is almost quadratic in $\mathbf{u}$ (but not in $\rho$...).

The example of Sec. 13.1.4 is a perfect illustration: in the $H$-constraint, the term with $\mathbf{B}$ succeeds in giving an invertible matrix of constraints, and it corresponds to a quadratic term in the Hamiltonian. On the contrary, the term with $\rho\mathbf{u}$ has an undesirable factor $\rho$, which comes from

[1] Remind that for this statement, the assumption $\{\phi_1, \phi_1\} = 0$ is important.

this excess factor in the Hamiltonian. The inadequateness of this factor is confirmed by Sec. 13.2.3.

To conclude with a more general result, in general the $H$-secondary constraint is exactly $\delta H/\delta\phi_2$, which can be any variable. For instance, if the Hamiltonian function is $\int \phi_2\omega$, with $\omega$ independent of $\phi_2$, then the $H$-secondary constraint is $\omega$, which can be anything but has no link with $\phi_2$. Accordingly, contrary to the $J$-secondary constraints, the $H$-secondary constraint has no link with the structure of Dirac reductions.

### 13.2.3 Application to fluid incompressibility

As an illustration, one could come back to a few of the previous examples ($\phi_1 = \nabla\times\mathbf{B}$, $\phi_1 = \nabla\times\mathbf{E}$, $\phi_1 = \nabla\cdot\mathbf{E}$, etc.). They confirm that each time, the $H$-constraint is not adequate unless the Hamiltonian function is quadratic in a variable conjugated to $\phi_1$. This also explains both why the procedure of $H$-secondary constraints is not reversible, and which of the two constraints has to be chosen as the primary one. Instead of coming back to these examples with details, let us study a bit more carefully the reduction to an incompressible fluid.

Starting from the natural constraint for incompressibility $\phi_1 = \nabla\cdot(\rho\mathbf{u})$, the matrix of constraints $\{\phi_1,\phi_1\}$ involves the operator

$$\mathcal{F} = \nabla\cdot\rho\nabla\mathbf{u}\cdot\nabla + \nabla\cdot\mathbf{u}\nabla\cdot\rho\nabla + \nabla\cdot\rho(\nabla\times\mathbf{u}+\mathbf{B})\times\nabla\,,$$

which is not invertible, and a secondary constraint is needed. The natural $H$-constraint is

$$\phi_2 = \{\phi_1,H\} = -\nabla\cdot\big(\rho\nabla(\tfrac{\mathbf{u}^2}{2})\big) - \nabla\cdot\Big(\mathbf{u}\big(\nabla\cdot(\rho\mathbf{u})\big)\Big) - \nabla\cdot((\nabla\times\mathbf{u}+\mathbf{B})\times\rho\mathbf{u}) + \nabla\cdot\rho\mathbf{E}\,,$$

that does not improve the situation, and fails to give the reduction to incompressible fluid. This is because the important term in the Hamiltonian is $\rho\mathbf{u}^2$, which is almost quadratic in $\mathbf{u}$, and hence far from quadratic in the desired secondary constraint $\rho$.

So, the trick used in [33] is to start from another constraint. In the reduction for incompressibility, a second variable is removed from the independent dynamical variables, the density[2]. It can be used as the primary constraint $\phi_1 = \rho - \rho_0$, then again the matrix of constraint $\{\phi_1,\phi_1\} = 0$ is not invertible, and a secondary constraint is again needed. The $H$-secondary constraint is

$$\phi_2 = \{\phi_1,H\} = -\nabla\cdot(\rho_0\mathbf{u})\,.$$

Then the matrix of constraints is the matrix

$$\mathbb{C} = \begin{pmatrix} 0 & \nabla\cdot\rho_0\nabla \\ -\nabla\cdot\rho_0\nabla & \mathcal{F} \end{pmatrix}\,,$$

with the convention that here, the derivatives act on all that is on their right hand side. The procedure works almost perfectly, because it is invertible when $\nabla\rho_0 = 0$ (remind that the invertibility is required only in the reduced space, where $\rho = \rho_0$). As a side comment, let us notice that if the field $\rho_0$ is arbitrary, then the operator $\nabla\cdot\rho_0\nabla$ is not invertible, which means the Dirac reduction cannot be applied. When $\rho_0$ is a physical density, it is positive and then the operator is invertible (since it has only positive eigenvalues[3]), but the reduction is not quarter-canonical, and it does not seem to have a physical relevance [33].

This exemplifies the above conclusions on the defects of the $H$-secondary constraint. Especially, it may give a set of constraints that does not give the desired Dirac reduction, even if such a set exists. The reason why the $H$-constraints are not good in this case appears as coming from two things: for $\phi_1 = \nabla\cdot(\rho\mathbf{u})$, the bracket is not quarter-canonical; for $\phi_1 = \rho - \rho_0$, it is quarter-canonical, but the Hamiltonian is not quadratic in the variable conjugated to $\rho$, there is an excess

[2]Notice that the natural constraint for incompressibility concerns the compressible part of the velocity. The trick to start from the variable $\rho$ is awkward because it assumes that one already has some information on the reduced model, and also because it does not take as primary constraint the essential one. This awkward feature disappears in the method working with $J$-secondary constraints, except that the specific choice of secondary constraint must be guided by some physical reason or intuition.

[3]Emanuele Tassi, private communication. But this fact is obvious to see actually.

factor of $\rho$ in $H$, which is present in the $H$-secondary constraint. Thus, the resulting reduction is almost quarter-canonical, and it becomes exactly quarter-canonical when $\nabla\rho_0 = 0$, since then $\phi_1$ becomes conjugated to the $H$-secondary constraint.

Let us turn now to the procedure using $J$-secondary constraints. Starting from the constraint $\nabla \cdot (\rho\mathbf{u})$, it gives the condition that

$$-\nabla \cdot (\rho\nabla G_\rho) - \nabla \cdot (\mathbf{u}\nabla \cdot G_\mathbf{u}) - \nabla \cdot ((\nabla \times \mathbf{u} + \mathbf{B}) \times G_\mathbf{u}) + \nabla \cdot (\rho G_\mathbf{E})$$

must be an invertible operator. One possible choice is $\rho - \rho_0$ with the condition $\nabla\rho_0 = 0$, and the usual reduction is recovered. On another hand, starting from the constraint $\rho - \rho_0$, one gets the condition that $\nabla \cdot G_\mathbf{u}$ must be invertible, which again indicates $\nabla \cdot (\rho_0\mathbf{u})$ as an acceptable secondary constraint when $\nabla\rho_0 = 0$. So, contrary to the case of $H$-secondary constraint, starting from any of the constraints, the desired reduction is obtained as one of the natural choices of $J$-secondary constraints.

One can get more, because the $J$-secondary constraints for $\phi_1 = \rho - \rho_0$ include a secondary constraint that is acceptable whatever $\rho_0$ may be. This because a variable conjugated to $\rho$ is indicated by the formula $\nabla \cdot G_\mathbf{u}$. So, the set of constraints $(\rho - \rho_0, \nabla \cdot \mathbf{u})$ gives an invertible matrix of constraints and therefore a good Dirac reduction, under no condition on $\nabla\rho_0$.

With regard to the physical relevance of these reductions, the $H$-secondary constraint $\phi_2 = \nabla \cdot (\rho_0\mathbf{u})$, which works when $\nabla\rho_0 = 0$, corresponds to an Eulerian incompressibility. But it is not the only physically interesting secondary constraint. The $J$-secondary constraint $\nabla \cdot \mathbf{u}$, which is valid for arbitrary $\rho_0$, corresponds to a local Lagrangian incompressibility [159]. In some way, this is not a surprise because Lagrangian descriptions are closer to the canonical structure than Eulerian descriptions.

This example enhances the conclusion of the previous subsection: the procedure to fix the secondary constraint is clearer and more efficient when it relies first on $J$-secondary constraints, purely related to the structure of the bracket. The usual $H$-secondary constraint is interesting as an additional information suggesting a typical choice that is bound to have some physical relevance.

## 13.3 A simplified subalgebra reduction

The two previous sections focused on the consequences of the quarter-canonical structure on the Dirac procedure, more precisely on the invertibility of the matrix of constraints and on the choice of a secondary constraint. Now, we turn to the link of the quarter-canonical structure with another reduction method, the truncation method. Chapter 8 showed that for quarter-canonical reductions, the Dirac method may be simplified because it results in a mere bracket truncation. This opens the door to a much simpler procedure, but at the same time quite a different reduction method. A natural question is what kind of Hamiltonian reduction is involved, and whether it is justified only by the Dirac procedure or if it may be related to other Hamiltonian reduction methods, closer to truncations.

It is useful to clarify this point in order to identify the conditions for the truncation method to be applicable. This could make it possible to use it even when the Dirac method does not work, e.g. when the structure is not quarter-canonical and the matrix of constraints is not invertible. The question is twofold: What are the conditions for the truncation method to give a valid Poisson bracket? Why is the truncated bracket always a valid bracket for quarter-canonical reductions?

This section is organized as follows. Subsec. 13.3.1 will be devoted to some reminders about bracket truncations. This will suggest that the bracket truncation for quarter-canonical Dirac reductions is related to a projection onto a subalgebra. Subsec. 13.3.2 will show that the subalgebra is not the one generated by the reduced variables $\boldsymbol{\chi}$, but the one generated by $(\phi_1, \boldsymbol{\chi})$. Last, Subsec. 13.3.3 will give a few examples.

### 13.3.1 Bracket truncations and subalgebras

First of all, let us remind that a bracket truncation generally does not yield a Hamiltonian bracket. The Jacobi identity is not preserved by such a method in general. More precisely, consider a Hamiltonian system, with field variables $\boldsymbol{\psi}$, as usual. Then choose a set of fields $\boldsymbol{\chi}$, considered as the reduced coordinates. Define the reduced (truncated) bracket between two functional $F[\boldsymbol{\chi}]$, $G[\boldsymbol{\chi}]$ by

$$\{F, G\}_t := \{F, G\}\,,$$

which means that $\mathbb{J}_t := \mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$. Here we call "bracket truncation" only the bracket truncations that corresponds to removing functional derivatives with respect to the constrained variables. Other kinds of brackets truncations exist [58, 136, 138], but they will not be considered in this chapter.

The reduction procedure above seems reasonable from a non-Hamiltonian point of view. But the point is that the truncated Poisson bracket loses the Jacobi identity in general, i.e. it is not Hamiltonian any more. The reason is that between three functionals of the reduced fields $F[\boldsymbol{\chi}]$, $G[\boldsymbol{\chi}]$, and $H[\boldsymbol{\chi}]$, the following relation does not hold

$$\{K, \{F, G\}\} \neq \{K, \{F, G\}_t\}_t\,. \tag{13.1}$$

The reason is the following: $\{F, G\} = \{F, G\}_t =: K'$, but this quantity generally depends not only on $\boldsymbol{\chi}$, but also on the complementary coordinates $\boldsymbol{\phi}$ (the constraints). Then $\{K, K'\}$ contains the terms $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$, but also the terms $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\phi}}$, whereas these last terms are not taken into account in $\{K, K'\}_t$.

In fact, all other properties of a Poisson bracket but the Jacobi identity are preserved by bracket truncation. Thus, a truncated bracket is Hamiltonian if and only if it verifies the Jacobi identity, which can be written $\sum_{(ijk)} \mathbb{J}^{\chi_i\chi_m}\partial_{\chi_m}\mathbb{J}^{\chi_j\chi_k} = 0$, where $\sum_{(ijk)}$ means circular permutation of $i$ ,$j$, $k$.

So, the truncated bracket is exactly $\mathbb{J}_t = \mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$, and the previous relation writes

$$\sum_{(ijk)} \mathbb{J}^{\chi_i\chi_m}\partial_{\chi_m}\mathbb{J}^{\chi_j\chi_k} = 0\,.$$

Using that the initial bracket is Hamiltonian, the relation can be written

$$\sum_{(ijk)} \mathbb{J}^{\chi_i\phi_m}\partial_{\phi_m}\mathbb{J}^{\chi_j\chi_k} = 0\,.$$

This condition for such a truncated bracket to be Hamiltonian is not automatically guaranteed and must be checked case by case. All the same, there are situations where it is automatically fulfilled.

The most trivial case is when $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ is constant in $\boldsymbol{\phi}$. In the argument following Eq. (13.1), it corresponds to the case where $K'$ is only a functional of the reduced fields. Also, it means that the reduced set of variables is closed for the Poisson bracket: the initial bracket between two functionals $F[\boldsymbol{\chi}]$ and $G[\boldsymbol{\chi}]$ of the fields $\boldsymbol{\chi}$ is again a functional of $\boldsymbol{\chi}$. The set of functionals $F[\boldsymbol{\chi}]$ defines a proper Poisson algebra, which is a subalgebra of the initial structure.

Another frequent case is when an ordering parameter is present and $\mathbb{J}^{\chi_i\chi_m}\partial_{\chi_m}\mathbb{J}^{\chi_j\chi_k}$ is not of the same order as $\mathbb{J}^{\chi_i\phi_m}\partial_{\phi_m}\mathbb{J}^{\chi_j\chi_k}$, e.g. $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\phi}} \propto (\boldsymbol{\phi})^\alpha$ and $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}} \propto (\boldsymbol{\phi})^\beta$ with $\beta \neq \alpha - 1$.

In the special case where a Dirac reduction is applied to a quarter-canonical bracket, then the Dirac bracket is just the truncated bracket. This last is Hamiltonian. So, there should be a general reason justifying the preservation of the Hamiltonian structure through these bracket truncations. The quarter-canonical structure has no ordering limitation, but it has a structure adapted to the bracket operation. So, it might be linked to a subalgebra method, which is the truncation method that is related to the structure of the Hamiltonian system.

### 13.3.2 Counterexample but proper justification

This suspicion is not correct in general. It is shown by a counterexample in the quarter-canonical reduction to an electrostatic field, given in Secs. 13.1.2 and 13.1.4. In the initial fluid bracket if one

takes two functionals $F[\boldsymbol{\chi}]$ and $G[\boldsymbol{\chi}]$ that do not depend on $\mathbf{B}$ and $\nabla \times \mathbf{E}$, then the initial bracket between them is not a functional of $\boldsymbol{\chi}$, since it will depend on the magnetic field as well, because of the presence of the gyro-magnetic term $\frac{\mathbf{B}}{\rho} \cdot F_{\mathbf{u}} \times G_{\mathbf{u}}$.

All the same, some link between subalgebras and the structure of the Poisson bracket was evidenced in Chapter 8. For instance, in finite-dimensional systems, a subalgebra is indeed involved in semi-canonical reductions. This is confirmed by the Darboux theorem, which states, as mentioned in Sec. 8.2.2, that the matrix $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}} = \mathbb{A}$ in (8.11) is constant in $\boldsymbol{\phi}$ (but this is actually just a consequence of the Jacobi identity).

As a result, for the quarter-canonical reduction, the departure from a subalgebra reduction is exactly achieved by the departure from semi-canonical coordinates. This departure was identified in the coordinate change (8.15). The change $\Phi_2 \longrightarrow \phi_2{}'(\phi_1, \Phi_2, \boldsymbol{\chi})$ does not affect the reduced bracket, but the change of coordinates $\boldsymbol{\chi} \longrightarrow \boldsymbol{\chi}'(\phi_1, \boldsymbol{\chi})$ modifies the bracket and makes it depend on $\phi_1$. This prevents the reduced bracket $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ from being given just by a subalgebra of the initial bracket, which explains the counterexample observed above. Nevertheless, the reduced bracket remains independent of $\phi_2$.

So, instead of considering a subalgebra without $(\phi_1, \phi_2)$ as was the case for semi-canonical coordinates, one should consider a subalgebra without $\phi_2$ only. Then the subalgebra suspicion becomes confirmed. Indeed the independence of the reduced bracket $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ with respect to $\phi_2$ is a consequence of the Jacobi identity. If $F$ and $G$ do not depend on $\phi_2$ then $\{G, \phi_1\} = 0$ and $\{\phi_1, F\} = 0$, which implies

$$\{\phi_1, \{F, G\}\} = -\{F, \{G, \phi_1\}\} - \{G, \{\phi_1, F\}\} = 0\,. \tag{13.2}$$

Thus the bracket $\mathbb{J}_t = \mathbb{A} = \{\boldsymbol{\chi}, \boldsymbol{\chi}\}$ cannot depend on the variable $\phi_2$ that is (minimal) conjugated to $\phi_1$, and the set of functionals $F(\phi_1, \boldsymbol{\chi})$ is a subalgebra. This conclusion relies just on the Jacobi identity. Thus it is general and applies both in finite dimensions and in field theory, both locally and globally. Compared to the semi-canonical case, quarter-canonical brackets retains only half of the requirements, which induces that only one of $(\phi_1\phi_2)$ remains concerned by the subalgebra property.

Notice that the conjugation must be minimal, i.e. bijection is needed in $\mathbb{J}^{\phi_1\phi_2}$, in order to avoid Casimir invariants in the components of $\phi_1$, which would let $\phi_2$ have the dimension of $\phi_1$ but have some components $\phi_{2b}$ such that $\{\phi_1, \phi_{2b}\} = 0$. In such a case, Eq. (13.2) would not mean that $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ is independent of $\phi_2$.

Last step, it remains to verify that the subalgebra reduced bracket is the same as the truncated bracket obtained by the quarter-canonical Dirac reduction. Actually, the subalgebra $(\phi_1, \boldsymbol{\chi})$ does not correspond to the Poisson bracket $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$, since also $\phi_1$ is retained. But the point is that $\phi_1$ is conjugated only with $\phi_2$. When $\phi_2$ is absent, then $\phi_1$ becomes a Casimir invariant. Thus the subalgebra bracket is exactly the truncated bracket $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$, which the same as the quarter-canonical Dirac bracket.

### 13.3.3 Examples

The presence of a subalgebra reduction can be verified and illustrated by any of the quarter-canonical examples previously mentioned. For instance, in the electrostatic reduction from the warm fluid model, the constraints are $\boldsymbol{\phi} = (\nabla \times \mathbf{E}, \nabla \times \mathbf{B})$, and the set of reduced observables $F(\nabla \cdot \mathbf{E}, \rho, \mathbf{u})$ is not a subalgebra, but the set of $F(\nabla \cdot \mathbf{E}, \mathbf{B}, \rho, \mathbf{u})$ is a subalgebra, and its reduced bracket is exactly Dirac's truncated bracket.

As a consequence, when the bracket is (minimal) quarter-canonical the Dirac reduction gives exactly the same result as the reduction to a subalgebra $F(\phi_1, \boldsymbol{\chi})$. It is the reason why it results in a truncated bracket and why the Hamiltonian character is preserved by such a bracket truncation. Accordingly, quarter-canonical Dirac reductions can always be shortened and replaced by a truncation method. This makes the procedure quite elementary: one only has to write the bracket

in adapted coordinates $(\boldsymbol{\phi}, \boldsymbol{\chi})$ and remove all the terms involving derivatives with respect to the constraint $\phi_2$.

Of course, the reduction can be performed in arbitrary coordinates. Then the reduced bracket is no more a truncation of the initial bracket, but the simplified projector still exists, as was emphasized in the previous chapter. Especially, in the light of the present section, the coordinate-independent projector (12.13), which is equivalent to (12.15), appears with its complete meaning: it is exactly the projector onto the subalgebra without $\boldsymbol{\phi}_2$. This confirms that it corresponds to the simplest possible projector, and that it is adapted to the structure of the Poisson bracket.

## 13.4 Flexibility and quarter-canonical effects in subalgebra reductions

To summarize the result of the previous section, together with Chapters 8 and 12, it came out that Dirac reductions are often related to a quarter-canonical structure, which is automatically related to a subalgebra reduction, which can then equivalently replace the Dirac reduction. Now, we can reverse the viewpoint, and wonder how much subalgebra reductions are linked to quarter-canonical structures, and whether they can be replaced by Dirac reductions, or if they are concerned by reductions for which the Dirac method is not available. The first question will be addressed in Subsec. 13.4.4, but beforehand Subsecs. 13.4.1-13.4.3 will investigate the second question, based on three kinds of examples.

### 13.4.1 Example with a single constraint

For quarter-canonical reductions, the two methods are equivalent, but even in this case the subalgebra method shows up greater flexibility. With the primary constraint $\phi_1$, it is able to give a reduced dynamics without $\phi_2$ becoming constrained.

An example can again be found in the reduction to incompressible fluid. With the Dirac primary constraint $\phi_1 = \rho - \rho_0$, the matrix of constraints is $\{\phi_1, \phi_1\} = 0$ is not invertible, the secondary constraint is needed, and it may only be chosen as a "conjugated" variable, which induces that the reduction must be quarter-canonical, as shown in the previous sections.

Using subalgebra reductions, the same quarter-canonical reduction can be obtained by taking the subalgebra defined by $(\rho, \nabla \times (\rho_0 \mathbf{u}), \mathbf{E}, \mathbf{B})$, but the reduction with one single constraint $\phi_1$ is possible as well. The set of observables depending only on $(\rho\mathbf{u}, \mathbf{E}, \mathbf{B})$ is a subalgebra, making $\rho - \rho_0$ become a Casimir invariant, but letting $\nabla \cdot (\rho\mathbf{u})$ dynamical.

As a side comment, this example is a counterexample to the Darboux theorem for the local-interaction matrix for Hamiltonian systems in continuous media: if it was verified, there would exist field coordinates $(\rho, \Phi_2, \boldsymbol{\chi})$ making the bracket semi-canonical. Then the subalgebra without $\rho$ would have an additional Casimir invariant $\Phi_2$, which is not the case actually.

### 13.4.2 Presence of multiple subalgebras

Let us go on with the quarter-canonical case. Even when both constraints $(\phi_1, \phi_2)$ are applied, the subalgebra reduction opens interesting possibilities. Indeed, in this case, there is only one Dirac reduction available, because the Dirac reduction does not depend on the chosen complementary coordinates $\boldsymbol{\chi}$. On the contrary, the subalgebra method depends on $\boldsymbol{\chi}$. With given constraints $(\phi_1, \phi_2)$, several reduced structures can be obtained, one of which is the Dirac one, and it corresponds to choosing $\boldsymbol{\chi}$ as any coordinates that keep the bracket quarter-canonical. But by taking a subalgebra defined by some other coordinates $\boldsymbol{\chi}'$, one can get other reduced structures.

As an example, if

$$(\phi_1, \phi_2) = (\nabla \cdot \mathbf{E}, \nabla \cdot \mathbf{u}),$$

the structure is quarter-canonical, and the Dirac reduction can be performed. It is independent of the chosen complementary coordinates

$$\boldsymbol{\chi} = (\rho, \nabla \times \mathbf{u}, \nabla \times \mathbf{E}, \mathbf{B}).$$

It gives the reduced dynamics corresponding to the subalgebra without $\phi_2 = \nabla \cdot \mathbf{u}$ in these coordinates, or, equivalently, to the subalgebra without $\boldsymbol{\phi} = (\nabla \cdot \mathbf{E}, \nabla \cdot \mathbf{u})$. Now, let us perform on $\boldsymbol{\chi}$ the change of variables

$$(\nabla \times \mathbf{u}, \mathbf{B}) \longrightarrow (\nabla \times \mathbf{u} + f, \mathbf{B} - f)\,,$$

where $f = \nabla \times (\overrightarrow{k}\nabla \cdot \mathbf{u} + g)$ is a divergence free vector linearly depending on $\phi_2$, with

$$g(\rho, \nabla \times \mathbf{E}, \nabla \cdot \mathbf{E} - \rho, \nabla \times \mathbf{u} + \mathbf{B})$$

an arbitrary vectorial function, and $\overrightarrow{k}$ a vector field independent of the field variables. In the new coordinates, the initial bracket is no more quarter-canonical, but the Dirac reduction gives the same reduced bracket as before the change of variables (even if it is not given by a truncation of the bracket in these coordinates). On the contrary, the set of observables without $\boldsymbol{\phi}$ remains a subalgebra, whose reduced bracket is as always given by a truncation of the initial bracket. So, it is different from the reduced bracket before the change of variable, and especially it is different from the Dirac bracket.

A generic example is easier found starting from canonical coordinates $\mathbf{z} = (\mathbf{v}, \mathbf{q})$. Then a sub algebra is obtained if

$$\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}} = \boldsymbol{\chi}_{\mathbf{v}^i} \cdot \boldsymbol{\chi}_{\mathbf{q}^i} - \boldsymbol{\chi}_{\mathbf{q}^i} \cdot \boldsymbol{\chi}_{\mathbf{v}^i}$$

depends only on $\boldsymbol{\chi}$. It means that

$$\partial_{\mathbf{z}^j}(\boldsymbol{\chi}_{\mathbf{v}^i} \cdot \boldsymbol{\chi}_{\mathbf{q}^i} - \boldsymbol{\chi}_{\mathbf{q}^i} \cdot \boldsymbol{\chi}_{\mathbf{v}^i}) = 0$$

is orthogonal to $\partial_{\boldsymbol{\phi}}\mathbf{z}^j$. Especially, any set of variables linear in the canonical coordinates $\mathbf{z}$ induce a subalgebra reduction.

Semi-canonical coordinates $(\phi_1, \Phi_2, \boldsymbol{\chi})$ can be considered as well. In the same way as for canonical coordinates, the corresponding reduction is sure to be a subalgebra without $\boldsymbol{\phi}$. Then, perform a change of coordinates

$$\boldsymbol{\chi} \longrightarrow \boldsymbol{\chi}' = \hat{F}(\boldsymbol{\chi}) + M \cdot \boldsymbol{\phi}\,,$$

where $M$ is a linear operators independent of the fields and $\hat{F}(\boldsymbol{\chi})$ is independent of $\boldsymbol{\phi}$ (and bijective, in order to define a proper change of coordinates). Then the bracket becomes

$$\mathbb{J} = \begin{pmatrix} M_c & M_c \cdot M^\dagger \\ M \cdot M_c & M' \end{pmatrix}\,, \tag{13.3}$$

where $M_c$ is the canonical matrix and

$$M' = F \cdot \mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}} \cdot F^\dagger + M \cdot M_c \cdot M^\dagger\,,$$

with $F := \hat{F}_{\boldsymbol{\psi}}$. Now, $M'$ is independent of $\boldsymbol{\phi}$, and the situation is the same as in the previous paragraph. The subalgebra after change of coordinates gives a reduced dynamics that is different from the one obtained before the change of variables (which is the Dirac truncated bracket).

The reduction to incompressible fluid again provides an interesting illustration. With the primary constraint $\rho - \rho_0$, we already saw that the $H$-secondary constraint $\nabla \cdot (\rho\mathbf{u})$ induced troubles for the Dirac method when $\nabla\rho_0 \neq 0$, because the structure is not exactly quarter-canonical. Now, a subalgebra reduction can be performed all the same: take the set of observables $F[\nabla \times (\rho\mathbf{u}), \mathbf{E}, \mathbf{B}]$. It is a Lie subalgebra and induces the constraints $\rho - \rho_0$ and $\nabla \cdot (\rho\mathbf{u})$ to become Casimir invariants. For the purely kinetic part (i.e. when dropping the electric and magnetic fields), the reduced dynamics corresponds to the Lie-Poisson structure for a solenoidal vector field

$$\{F, G\} = \int d^3\mathbf{x}\; \mathbf{M}' \cdot (\nabla \times F_{\mathbf{M}'} \times \nabla \times G_{\mathbf{M}'})\,,$$

where $\mathbf{M}' = \nabla \times (\rho\mathbf{u})$ is the solenoidal part of the fluid momentum density. This reduced bracket is meaningful whatever $\nabla\rho_0$ is.

All these examples emphasize that subalgebra reductions indeed open new reduction possibilities, which can include physically relevant reductions. Two noteworthy examples are the reduction from the kinetic dynamics to the fluid model, and the reduction from the (multi) fluid model to magnetohydrodynamics. They will be the topic of Chapters 9 and 14.

The key point about subalgebra reductions is that the applicability condition refers to the coordinates $\boldsymbol{\chi}$, it is independent of the constraints $\boldsymbol{\phi}$. Especially, the matrix of constraints does not have to be invertible, as is the case for Dirac reductions. So, given acceptable $\boldsymbol{\chi}$, the constraints are completely free. They must only be complementary to $\chi$, in order to define proper coordinates. In the reverse direction, given a set of constraints $\boldsymbol{\phi}$, the freedom in the coordinates $\boldsymbol{\chi}$ can be adjusted to satisfy the applicability condition. It is why the subalgebra method opens much more possible reductions.

### 13.4.3 Application to bracket extensions

The question of bracket extensions also constitutes an expressive instance. Just as in Sec. 8.1 of Chapter 8, starting from a Poisson bracket that has Casimir invariants $\boldsymbol{\phi} = (\phi_1, \phi_2)$, one can wonder if it comes from an extended bracket where the constraints $\boldsymbol{\phi}$ are dynamical. We saw in Chapter 8 that the Dirac reduction method together with the quarter-canonical structure provided a powerful tool to achieve the result.

If the reduced bracket is independent of $\boldsymbol{\phi}$, just as the Euler-Maxwell bracket is independent of

$$\boldsymbol{\phi} = (\nabla \cdot \mathbf{E} - \rho, \nabla \cdot \mathbf{B}) \,.$$

A subalgebra method is still more powerful: it allows one to choose $\mathbb{J}_e$ as

$$\mathbb{J}_e = \begin{pmatrix} e & -b^\dagger & -f^\dagger \\ b & a & -d^\dagger \\ f & d & \mathbb{A} \end{pmatrix} \,, \tag{13.4}$$

with $\mathbb{A} = \mathbb{J}$, and the remaining coefficients are freely chosen so that $\mathbb{J}_e$ satisfies the Jacobi identity, e.g. they are all chosen constant in the field variables. Then starting from the extended bracket $\mathbb{J}_e$, the initial bracket $\mathbb{J}$ is just given by a sub-algebra reduction, with the subalgebra being defined by

$$\boldsymbol{\chi} = (\rho, \mathbf{u}, \nabla \times \mathbf{E}, \nabla \times \mathbf{B}) \,.$$

A comparison between formulae (8.12) and (13.4) clearly shows that the subalgebra method gives much more possibilities than the Dirac method: the quarter-canonical structure is no more mandatory and all the extension coefficients can now be non-zero. This example also emphasizes the difference between the methods. The extension in the Dirac approach by formula (8.12) implies an even number of initial Casimir invariants, whereas the extension in the subalgebra approach by formula (13.4) has no such requirement, but it implies the reduced bracket to be independent of the initial Casimir invariants $\boldsymbol{\phi}$.

### 13.4.4 Quarter-canonical effects on subalgebras

This section was interested in two questions: how often subalgebra reductions can be replaced by a Dirac reduction, and whether subalgebra reductions are often related to a quarter-canonical structure. The three previous subsections showed that subalgebra reductions generally offer much more possibilities than Dirac reductions. This answered the first question. We now turn to the second question.

If quarter-canonical structures automatically come with a subalgebra, the contrary is not so true. The closure criterion for a subalgebra $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}} = constant(\boldsymbol{\phi})$ concerns the functional dependence of the coefficients in the Poisson bracket. It is independent of the quarter-canonical criterion, which regards the presence or absence of the coupling between fields. The last two examples of subalgebra reductions are expressive from that point of view, since they are not quarter-canonical at all. Thus, subalgebra reductions are not so often related to quarter-canonical structures as Dirac reductions.

Still, the traces of the canonical structure have an influence on subalgebra reductions. For instance, in order to impose a scalar constraint in a finite-dimensional system, if one takes a

subalgebra without $\phi_A$, then, the reduced bracket is defined on an odd dimensional space, and must have an additional Casimir invariant $\phi_B$. In coordinates $(\phi_A, \phi_B, \boldsymbol{\chi})$, the initial bracket writes

$$\mathbb{J} = \begin{pmatrix} a & -b^\dagger & -d^\dagger \\ b & 0 & 0 \\ d & 0 & \mathbb{A} \end{pmatrix},$$

which is just quarter-canonical with $(\phi_1, \phi_2, \boldsymbol{\chi}) = (\phi_B, \phi_A, \boldsymbol{\chi})$.

For a reduction with two constraints $\phi_1$, $\phi_2$, if one takes a subalgebra $F(\boldsymbol{\chi})$ where $\boldsymbol{\chi}$ are coordinates complementary to $\boldsymbol{\phi}$, the conclusion is not so simple. All the same, the Darboux theorem allows one to find coordinates $\boldsymbol{\psi} = (\phi_1, \Phi_2, \boldsymbol{\chi}')$, such that $\mathbb{J}$ is semi-canonical (8.11), and $\mathbb{J}^{\boldsymbol{\chi}'\boldsymbol{\chi}'}$ does not depend on $(\phi_1, \phi_2)$. Then, adopting coordinates suited to the reduction

$$(\phi_1, \Phi_2, \boldsymbol{\chi}') \longrightarrow (\phi_1, \Phi_2, \boldsymbol{\chi}),$$

the reduced bracket obtained is

$$\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}} = \mathbb{J}^{ij}\partial_{\psi_i}\boldsymbol{\chi}\partial_{\psi_j}\boldsymbol{\chi} = \partial_{\phi_1}\boldsymbol{\chi}\partial_{\Phi_2}\boldsymbol{\chi} - \partial_{\phi_2}\boldsymbol{\chi}\partial_{\Phi_1}\boldsymbol{\chi} + \mathbb{J}^{\boldsymbol{\chi}'_i\boldsymbol{\chi}'_j}\partial_{\boldsymbol{\chi}'_i}\boldsymbol{\chi}\partial_{\boldsymbol{\chi}'_j}\boldsymbol{\chi}. \tag{13.5}$$

If the change of coordinates is arbitrary, the coefficients of $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ will most often depend on $(\phi_1, \phi_2)$, and $\boldsymbol{\chi}$ will not be a Lie subalgebra. So, the possible subalgebra reductions are limited by the traces of the canonical structure of the Hamiltonian system. This fact already appeared in the examples about formula (13.3): as soon as one departs from a linear change of coordinates and from a bracket constant in the field variables, the conditions for a subalgebra are not so easily obtained. For instance, if only one of the coordinates is involved in the change from canonical coordinates, then the only allowed change of variable is linear in $\boldsymbol{\phi}$. If several coordinates are involved in the change from canonical coordinates, then the coordinates given by a non-linear relation must be such that in $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ all their contributions proportional to $\boldsymbol{\phi}^\alpha$ exactly cancel for each $\alpha \neq 0$.

For continuous systems, the canonical structure may not exist and the argument cannot be applied exactly. However, in Poisson brackets for fluids and plasmas, the canonical structure is not completely absent [159]. We saw that quarter-canonical coordinates can often be identified even in natural variables. Then the argument can be applied. Once $\phi_1$ is chosen, the available subalgebras do not so easily depart from the quarter-canonical structure.

By the way, notice that even the quarter-canonical structure does not correspond exactly to a subalgebra reduction, since $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ is independent on $\phi_2$, but not on $\phi_1$. It is why the generic examples about Eq. (13.5) and in Sec. 13.4.2 involved semi-canonical coordinates; and the example

$$(\nabla \times \mathbf{u}, \mathbf{B}) \longrightarrow (\nabla \times \mathbf{u} + f(\boldsymbol{\phi}), \mathbf{B} - f(\boldsymbol{\phi}))$$

based on true quarter-canonical coordinates was not very natural. Actually, when coming back to all of the previous examples of quarter-canonical structures, true subalgebra reductions can be verified to be not so common, especially if the coordinates linear in semi- or quarter- canonical coordinates are excluded. This is also confirmed by the fact that when looking for semi-canonical coordinates, the condition of $\mathbb{J}^{\boldsymbol{\chi}\boldsymbol{\chi}}$ being subalgebra is a good indication, which means it is rather seldom. In the case of $\phi_1 = \nabla \times \mathbf{E}$ for instance, it is enough to catch such coordinates. All these observations show that subalgebra reductions remain influenced by traces of the canonical structure, whether they be semi-canonical or quarter-canonical.

As a result, Dirac reductions are often linked with the quarter-canonical structure and with a subalgebra reduction, but the reciprocal is not so often verified. subalgebra reductions are much more independent of the traces of the canonical structure, such as quarter-canonical structures, even though they remain influenced by them all the same.

In the particular case of a quarter-canonical structure, the two methods are equivalent provided the reduction is adapted to the quarter-canonical structure, but otherwise the subalgebra method opens up many interesting possibilities.

## Conclusion

The quarter-canonical structure is strongly involved in the invertibility of the matrix of constraints, which is the condition for the Dirac procedure to work. When minimal quarter-canonical coordinates exist, then the matrix of constraint is invertible if and only if the constraints are such coordinates. It is why the relevant secondary constraints are determined by quarter-canonical coordinates. These $J$-secondary constraints make much clearer the role of secondary constraints, and they constitute an interesting complement of the usual procedure relying on $H$-secondary constraints.

The justification of the truncation method in the case of quarter-canonical brackets relies on the presence of a subalgebra, automatically associated with the quarter-canonical structure. It is a consequence of the Jacobi identity, and a property shared with semi-canonical brackets, but in a weakened form, where the subalgebra involves only one of the two constraints. For quarter-canonical structures, the existence conditions both for a Dirac reduction and for a subalgebra reduction are automatically satisfied, and the two methods are equivalent in their result, because the matrix of constraints is always invertible, the reduced bracket is always independent of $\phi_2$, and $\mathbb{B}\mathbb{C}^{-1}\mathbb{B}^\dagger$ in (8.8) is always zero.

Even for quarter-canonical brackets, the flexibility of subalgebra reductions often offers complementary possibilities, for instance to impose one single constraint, or to get several possible reduced structures. This is because subalgebra reductions are less marked by traces of the canonical structure than Dirac reductions, even if they are still influenced by them.

Practically, the method of Dirac's constraints and the reduction by a subalgebra, although very different, are often related and the quarter-canonical structure is a key ingredient in this relation.

Equipped with these results about Dirac and subalgebra reductions, it is interesting to come back to the gyrokinetic reduction, where the removal of the superfluous dimensions evokes a bracket truncation. It is the next step of the work, which is reported in Sec. 8.4 in Chapter 8.

Another extension of the work will be to investigate more involved Hamiltonian reductions in plasmas than the ones considered in the present chapter. It will especially be interesting to observe the possible role of Dirac, subalgebras, and quarter-canonical reductions. This will be the topic of Chapters 9 and 14, where the subalgebra method will prove very efficient for the reduction from Vlasov-Maxwell to Euler-Maxwell, and then from Euler-Maxwell to magnetohydrodynamics.

# Chapter 14

# Hamiltonian derivation of magnetohydrodynamics from multi-fluid dynamics

in collaboration with Cristel Chandre and Emanuele Tassi

*A*bstract: The Hamiltonian structure of ideal magnetohydrodynamics is derived from the ideal two-fluid dynamics, using Hamiltonian reduction methods. The reduced Poisson bracket is obtained from the two-fluid Euler-Maxwell bracket by imposing constraints on the current, the electric field and the charge density, through a projection onto a Lie-subalgebra. The reduced Hamiltonian functional results from the limit of a non-relativistic, low-density, large-scale regime. The preservation of total momentum is found to be determinant in the coupling between the plasma and the electric field. The case of relativistic magnetohydrodynamics is considered, as well as the presence of more than two fluids. A key role is played by the Poincaré subalgebra, and by the $\mathbf{E} \times \mathbf{B}$ plasma velocity.

## Introduction

This appendix chapter takes place after Chapter 9. Indeed, the Hamiltonian methods developed in Chapters 8, 12 and 13 called for applications. As a first example, Chapter 9 investigated Hamiltonian fluid closures. We now turn to another application, the Hamiltonian reduction from ideal-fluid dynamics to the ideal-MHD dynamics, which relates the structures of two most used models for plasmas.

Indeed, as was mentioned in the general introduction, in many cases, even the fluid model for plasmas is more precise than needed. Much physical information already relies in the behaviour of the plasma as a whole, considered as a single charged fluid coupled with the electromagnetic field [6,46]. This motivated the magnetohydrodynamics (MHD) description. Combining the motion of the fluids, and adding some assumptions on the regime involved (essentially non-relativistic and electroneutral), the dynamical variables are reduced to four[1]: the magnetic field $\mathbf{B}$, the total plasma (mass) density $\rho$, entropy density $\sigma$, and momentum density $\mathbf{M}$. The dynamics is given by

$$\begin{aligned} \dot{\rho} &= -\nabla \cdot \mathbf{M}\,, & \dot{\mathbf{B}} &= \nabla \times \left(\tfrac{\mathbf{M}\times\mathbf{B}}{\rho}\right), \\ \dot{\sigma} &= -\nabla \cdot \left(\sigma \tfrac{\mathbf{M}}{\rho}\right), & \dot{\mathbf{M}} &= -\nabla \cdot \left(\tfrac{\mathbf{M}\mathbf{M}}{\rho}\right) + (\nabla \times \mathbf{B}) \times \mathbf{B} - \nabla P\,, \end{aligned} \tag{14.1}$$

with $P(\rho, \sigma)$ the pressure.

In a similar way as for the fluid reduction, the derivation of the MHD model was initially done by working at the level of the equations of motion, and only later on, the MHD model revealed as Hamiltonian [101,102,104,107,143,156] provided the dissipation coefficients are put to zero, which

[1] Among the four fields, two are three-dimensional vectors, so there are actually eight scalar fields.

corresponds to the so-called ideal-MHD model. This is a remarkable property, because the Hamiltonian character is very specific and involves strong requirements on the structure of the dynamics. As the parent model is Hamiltonian, it is true that the reduced model has to be Hamiltonian if no dissipation is added in the reduction process. But the point is that an arbitrary reduction process generally breaks the Hamiltonian character of the model [58, 136, 138, 156].

When the reduced model is Hamiltonian, it usually means that its Hamiltonian structure comes from the structure of the parent model through a Hamiltonian reduction process [3, 31–34, 58]. For instance, the ideal-fluid dynamics is essentially induced by a Lie subalgebra from the Hamiltonian structure of the kinetic dynamics, as shown in Chapter 9; and the (collisionless) Vlasov-Maxwell dynamics [101, 156] may be obtained using Hamiltonian methods associated to gauge groups [95], or it may be viewed as coming from a lift of the Hamiltonian structure of particle dynamics [5].

Accordingly, we are interested here in identifying the Hamiltonian reduction that generates the ideal-MHD model from the two-fluid system, in order to improve our understanding of this reduction and to clarify the links between the Hamiltonian structures of the two models. In addition, it permits to keep track of all the conserved quantities in the whole process, and shows how the model has to be modified if one wants to include additional effects, e.g. to retain higher order effects or to drop one of the assumptions of the model. Last, it develops Hamiltonian reduction methods, which can be useful to derive other models.

Because of the previous chapters, here we are especially interested in identifying whether the reduction fits with a Dirac or a subalgebra reduction, or with variations of them. Alternatively, it could rely on an another Hamiltonian reduction method, which would be as interesting to explore.

The chapter is organized as follows. In Sec. 14.1, we change the field coordinates in order to describe the total plasma rather than each fluid separately. In Sec. 14.2, we reduce the Poisson bracket to describe the plasma only as a single-fluid. We show that it provides the Poisson bracket of (ideal) MHD. It is given by a Poisson subalgebra of the two-fluid dynamics, and by imposing the usual relations of MHD as constraints. The momentum cannot be the plasma momentum, it has to be the total momentum; the coupling between the plasma and the magnetic field appears to be essentially linked to the preservation of total momentum by the constraint on the electric field. In Sec. 14.3, we show that, with the constraints, the reduced Hamiltonian functional and the corresponding motion are not exactly the ones of MHD, but they are a Hamiltonian perturbation of them. The two ordering parameters are consistent with the physical framework. In Sec. 14.4, the generalization to the multi-fluid, relativistic case with anisotropic pressure is considered, and the structure of MHD is related to the subalgebra of Poincaré generators. Notice that here we always consider ideal MHD and ideal fluids, even if for brevity, we often drop the word "ideal".

## 14.1 Field coordinates for the total plasma

The starting point of our derivation is the Hamiltonian structure of multi-fluid dynamics [143]. Each fluid is described by the following field variables : its mass density $\rho_s(\mathbf{x})$, its momentum density $\mathbf{M}_s(\mathbf{x})$ and its entropy density $\sigma_s(\mathbf{x})$, where the index $s$ indicates the species. Here we consider two species, e.g., electrons and ions: $s \in \{1; 2\}$. The two fluids are interacting self-consistently with the electromagnetic fields $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$. The observables are functionals of the field variables $\boldsymbol{\psi} = (\rho_1, \rho_2, \mathbf{M}_1, \mathbf{M}_2, \sigma_1, \sigma_2, \mathbf{E}, \mathbf{B})$. The Poisson bracket acts on two observables $F$ and $G$ as

$$\{F, G\} = \int d^3\mathbf{x} \left[ \sum_s \left[ \rho_s \, G_{\mathbf{M}_s} \cdot \nabla F_{\rho_s} + \mathbf{M}_s \cdot \, (G_{\mathbf{M}_s} \cdot \nabla) F_{\mathbf{M}_s} + \frac{\alpha_s \rho_s \mathbf{B}}{2} \cdot F_{\mathbf{M}_s} \times G_{\mathbf{M}_s} \right.\right.$$
$$\left.\left. + \sigma_s \, G_{\mathbf{M}_s} \cdot \nabla \, F_{\sigma_s} + \alpha_s \rho_s F_{\mathbf{M}_s} \cdot G_{\mathbf{E}} \right] + F_{\mathbf{E}} \cdot \nabla \times G_{\mathbf{B}} \right] - (F \leftrightarrow G) \quad (14.2)$$

where we use the index notation for functional derivatives, e.g., $F_{\mathbf{E}} = \delta F / \delta \mathbf{E}$, the coefficient $\alpha_s = q_s / m_s$ is the ratio of the charge and mass for the species $s$, and the symbol $(F \leftrightarrow G)$ indicates that the previous terms are repeated, but with permutation of $F$ and $G$ and inverting the signs (in

order to fulfil the antisymmetry property of the Poisson bracket). The Hamiltonian functional is the total energy

$$H = \int d^3\mathbf{x} \left[ \sum_s \left( \frac{\mathbf{M}_s^2}{2\rho_s} + U_s(\rho_s, \sigma_s) \right) + \frac{\mathbf{E}^2 + \mathbf{B}^2}{2} \right] , \tag{14.3}$$

where $U_s$ is the internal energy density of the fluid $s$. The dynamics of an observable $F$ is obtained from the Hamiltonian and the Poisson bracket as $\dot{F} = \{F, H\}$, and in particular the equations of motion for the field variables are obtained by taking $F = \psi_i$. It should be noted that the two-fluid model can be derived in a Hamiltonian way from Vlasov-Maxwell equations using a subalgebra argument, as shown in Chapter 9. The local observables $\nabla \cdot \mathbf{E} - \sum_s \alpha_s \rho_s$ and $\nabla \cdot \mathbf{B}$ are conserved quantity, since they are Casimir invariants of the bracket (14.2), and physically, they have to be chosen equal to zero in the initial conditions for the field variables. Rigorously speaking , $\nabla \cdot \mathbf{B}$ is a Casimir only when the $\mathbf{B}$ in the bracket is replaced by $(1 - \nabla\Delta^{-1}\nabla\cdot)\mathbf{B}$ (see Chapter 7).

The basic idea of the MHD reduction is to describe the plasma as a single charged fluid. Accordingly, we adopt as new plasma field variables

$$\rho = \rho_1 + \rho_2 , \tag{14.4}$$
$$\rho_\mathrm{e} = \alpha_1\rho_1 + \alpha_2\rho_2 , \tag{14.5}$$
$$\sigma = \sigma_1 + \sigma_2 , \tag{14.6}$$
$$\sigma_\mathrm{e} = \alpha_1\sigma_1 + \alpha_2\sigma_2 , \tag{14.7}$$
$$\mathbf{J} = \alpha_1\mathbf{M}_1 + \alpha_2\mathbf{M}_2 , \tag{14.8}$$
$$\mathbf{M} = \mathbf{M}_1 + \mathbf{M}_2 . \tag{14.9}$$

The new variable $\rho$ is the total mass density, $\rho_\mathrm{e}$ is the total charge density, $\mathbf{M}$ is the total momentum density, $\mathbf{J}$ is the total current density, $\sigma$ is the total entropy density, and $\sigma_\mathrm{e}$ is the analogue of $\rho_\mathrm{e}$ for the entropy, i.e. it is the total "charge-weighted" entropy density.

In these variables, the Hamiltonian functional (14.3) becomes

$$H = \int d^3\mathbf{x} \left[ \frac{1}{2(\alpha_1\rho-\rho_\mathrm{e})(\alpha_2\rho-\rho_\mathrm{e})} \left(2\rho_\mathrm{e}\mathbf{M}\cdot\mathbf{J} - \rho_\alpha\mathbf{M}^2 - \rho\mathbf{J}^2\right) + U + \frac{\mathbf{E}^2 + \mathbf{B}^2}{2} \right] , \tag{14.10}$$

where the coefficient $\rho_\alpha$ is defined in Eq. (14.12). The total internal energy density is given by $U = U_1(\rho_1, \sigma_1) + U_2(\rho_2, \sigma_2)$, where the initial variables must be considered as functions of the new variables, e.g. $\rho_1 = \frac{\alpha_2\rho-\rho_\mathrm{e}}{\alpha_2-\alpha_1}$.

As for the Poisson bracket, it is obtained by computing the functional derivatives in the new field variables from the scalar invariance $F[\rho_1, \rho_2, \mathbf{M}_1, \mathbf{M}_2, \sigma_1, \sigma_2, \mathbf{E}, \mathbf{B}] = \bar{F}[\rho, \rho_\mathrm{e}, \mathbf{M}, \mathbf{J}, \sigma, \sigma_\mathrm{e}, \mathbf{E}, \mathbf{B}]$. The functional derivatives are given by

$$\begin{aligned} F_{\rho_s} &= \bar{F}_\rho + \alpha_s\bar{F}_{\rho_\mathrm{e}} , \\ F_{\mathbf{M}_s} &= \bar{F}_\mathbf{M} + \alpha_s\bar{F}_\mathbf{J} , \\ F_{\sigma_s} &= \bar{F}_\sigma + \alpha_s\bar{F}_{\sigma_\mathrm{e}} , \end{aligned}$$

and the functional derivatives with respect to $\mathbf{E}$ and $\mathbf{B}$ are unchanged. By inserting these functional derivatives in the bracket (14.2), it becomes

$$\begin{aligned} \{F, G\} = \int d^3\mathbf{x} \Big[ & \rho G_\mathbf{M}\cdot\nabla F_\rho + \rho_\mathrm{e}(G_\mathbf{M}\cdot\nabla F_{\rho_\mathrm{e}} + G_\mathbf{J}\cdot\nabla F_\rho) + \rho_\alpha G_\mathbf{J}\cdot\nabla F_{\rho_\mathrm{e}} \\ & + \mathbf{M}\cdot(G_\mathbf{M}\cdot\nabla)F_\mathbf{M} + \mathbf{J}\cdot[(G_\mathbf{M}\cdot\nabla)F_\mathbf{J} + (G_\mathbf{J}\cdot\nabla)F_M) + \mathbf{M}_\alpha\cdot(G_\mathbf{J}\cdot\nabla)F_\mathbf{J} \\ & + \frac{\rho_\mathrm{e}\mathbf{B}}{2}\cdot F_\mathbf{M}\times G_\mathbf{M} + \rho_\alpha\mathbf{B}\cdot F_\mathbf{M}\times G_\mathbf{J} + \frac{\rho_\beta\mathbf{B}}{2}\cdot F_\mathbf{J}\times G_\mathbf{J} \\ & + \sigma G_\mathbf{M}\cdot\nabla F_\sigma + \sigma_\mathrm{e}(G_\mathbf{M}\cdot\nabla F_{\sigma_\mathrm{e}} + G_\mathbf{J}\cdot\nabla F_\sigma) + \sigma_\alpha G_\mathbf{J}\cdot\nabla F_{\sigma_\mathrm{e}} \\ & + \rho_\mathrm{e}F_\mathbf{M}\cdot G_\mathbf{E} + \rho_\alpha F_\mathbf{J}\cdot G_\mathbf{E} + F_\mathbf{E}\cdot\nabla\times G_\mathbf{B} \Big] - (F \leftrightarrow G) , \end{aligned} \tag{14.11}$$

where we have dropped the bars on the functionals. The quantities $\rho_\alpha$, $\rho_\beta$, $\mathbf{M}_\alpha$, and $\sigma_\alpha$ are defined by

$$\begin{aligned}
\rho_\alpha &= (\alpha_1+\alpha_2)\rho_\mathrm{e} - \alpha_1\alpha_2\rho\,, \\
\rho_\beta &= (\alpha_1^2+\alpha_1\alpha_2+\alpha_2^2)\rho_\mathrm{e} - \alpha_1\alpha_2(\alpha_1+\alpha_2)\rho\,, \\
\mathbf{M}_\alpha &= (\alpha_1+\alpha_2)\mathbf{J} - \alpha_1\alpha_2\mathbf{M}\,, \\
\sigma_\alpha &= (\alpha_1+\alpha_2)\sigma_\mathrm{e} - \alpha_1\alpha_2\sigma\,.
\end{aligned} \tag{14.12}$$

Notice that they are not new variables but only shorthands.

We notice that $\mathbf{M}$ is the total momentum density of the plasma. However the invariant associated to spatial translation symmetry is the total momentum density of both the plasma and the field, i.e. $\mathbf{M}+\mathbf{E}\times\mathbf{B}$. This suggests a new change of variables

$$\mathbf{M}' = \mathbf{M} + \mathbf{E}\times\mathbf{B}\,, \tag{14.13}$$

following Ref. [72]. Then, the total fluid variables are just the densities involved in the conservation laws for mass, entropy and momentum.

In the new coordinate system, the expression of the Hamiltonian is obtained by replacing $\mathbf{M}$ by $\mathbf{M}'-\mathbf{E}\times\mathbf{B}$ in Eq. (14.10). As for the Poisson bracket, it is obtained by replacing $F_\mathbf{M}$ by $F_{\mathbf{M}'}$, $F_\mathbf{E}$ by $F_\mathbf{E}-F_{\mathbf{M}'}\times\mathbf{B}$ and $F_\mathbf{B}$ by $F_\mathbf{B}+F_{\mathbf{M}'}\times\mathbf{E}$ in Eq. (14.11). The bracket becomes

$$\{F,G\} = \{F,G\}_1 + \{F,G\}_2\,, \tag{14.14}$$

where the first term $\{F,G\}_1$ has the same expression as in Eq. (14.11), but with $\mathbf{M}'$ replacing $\mathbf{M}$ and $\{F,G\}_2$ is given by

$$\begin{aligned}
\{F,G\}_2 = \int d^3\mathbf{x}\,\Big[&-(\mathbf{E}\times\mathbf{B})\cdot(G_{\mathbf{M}'}\cdot\nabla)F_{\mathbf{M}'} + \alpha_1\alpha_2(\mathbf{E}\times\mathbf{B})\cdot(G_\mathbf{J}\cdot\nabla)F_\mathbf{J} \\
&+\rho_\mathrm{e}F_{\mathbf{M}'}\cdot\mathbf{B}\times G_{\mathbf{M}'} + \rho_\alpha F_\mathbf{J}\cdot\mathbf{B}\times G_{\mathbf{M}'} + (\mathbf{B}\times F_{\mathbf{M}'})\cdot\nabla\times G_\mathbf{B} \\
&-F_\mathbf{E}\cdot\nabla\times(\mathbf{E}\times G_{\mathbf{M}'}) - (\mathbf{B}\times F_{\mathbf{M}'})\cdot\nabla\times(\mathbf{E}\times G_{\mathbf{M}'})\Big] - (F\leftrightarrow G)\,.
\end{aligned} \tag{14.15}$$

We rewrite the last term of Eq. (14.15) together with the corresponding antisymmetric term as

$$\begin{aligned}
&-\int d^3\mathbf{x}\,(\mathbf{B}\times F_{\mathbf{M}'})\cdot\nabla\times(\mathbf{E}\times G_{\mathbf{M}'}) - (F\leftrightarrow G) \\
&= \int d^3\mathbf{x}\,\Big[(\mathbf{E}\times\mathbf{B})\cdot(G_{\mathbf{M}'}\cdot\nabla)F_{\mathbf{M}'} - (\mathbf{E}\times\mathbf{B})\cdot(F_{\mathbf{M}'}\cdot\nabla)G_{\mathbf{M}'} \\
&- (\nabla\cdot\mathbf{B})\;\mathbf{E}\cdot F_{\mathbf{M}'}\times G_{\mathbf{M}'} + (\nabla\cdot\mathbf{E})\;\mathbf{B}\cdot F_{\mathbf{M}'}\times G_{\mathbf{M}'}\Big]\,.
\end{aligned} \tag{14.16}$$

The first two terms in Eq. (14.16) compensate exactly the first term in the bracket (14.15). The third term in Eq. (14.16) is proportional to $\nabla\cdot\mathbf{B}$ which vanishes on the physical shell. The last term in Eq. (14.16) is combined with the gyro-magnetic term $\rho_\mathrm{e}F_{\mathbf{M}'}\cdot\mathbf{B}\times G_{\mathbf{M}'}$ (repeated twice with the corresponding antisymmetric term) in the bracket (14.15) to give a contribution proportional to $\nabla\cdot\mathbf{E}-\rho_\mathrm{e}$ which is a Casimir invariant and vanishes on the physical shell. So, the contribution of Eq. (14.15) to the Poisson bracket (14.14) is

$$\begin{aligned}
\{F,G\}_2 = \int d^3\mathbf{x}\,\Big[&\alpha_1\alpha_2(\mathbf{E}\times\mathbf{B})\cdot(G_\mathbf{J}\cdot\nabla)F_\mathbf{J} - \frac{\rho_\mathrm{e}\mathbf{B}}{2}\cdot F_{\mathbf{M}'}\times G_{\mathbf{M}'} \\
&+\rho_\alpha F_\mathbf{J}\cdot\mathbf{B}\times G_{\mathbf{M}'} + (\mathbf{B}\times F_{\mathbf{M}'})\cdot\nabla\times G_\mathbf{B} \\
&-F_\mathbf{E}\cdot\nabla\times(\mathbf{E}\times G_{\mathbf{M}'})\Big] - (F\leftrightarrow G)\,.
\end{aligned} \tag{14.17}$$

We notice that the gyro-magnetic term $\rho_\mathrm{e}\mathbf{B}\cdot F_{\mathbf{M}'}\times G_{\mathbf{M}'}$ automatically vanishes in the bracket (14.14) due to opposite contributions in $\{\cdot,\cdot\}_1$ and $\{\cdot,\cdot\}_2$. Up to now, no approximations were made. The bracket (14.14) given by Eqs. (14.11) and (14.17) is equivalent to the initial bracket (14.2) expressed in the relevant field variables $(\rho,\rho_\mathrm{e},\mathbf{M}',\mathbf{J},\sigma,\sigma_\mathrm{e},\mathbf{E},\mathbf{B})$ on the physical shell ($\nabla\cdot\mathbf{B}=0$ and $\nabla\cdot\mathbf{E}=\rho_\mathrm{e}$).

## 14.2 Reduction to a single-fluid Poisson bracket

Now, we consider reducing the dynamics to the MHD model, which excludes from the independent dynamical variables the fields $(\rho_{\rm e}, \sigma_{\rm e}, \mathbf{J}, \mathbf{E})$, as a result of Eqs. (14.19)-(14.20). There remain only the variables $(\rho, \sigma, \mathbf{M}', \mathbf{B})$. Interestingly, in the Poisson bracket (14.14), this exactly corresponds to a subalgebra . Indeed, if we consider two functionals $F$ and $G$ of the field variables $(\rho, \sigma, \mathbf{M}', \mathbf{B})$. The expression of the bracket (14.14) for these two functionals is given by

$$\{F, G\} = \int d^3\mathbf{x} \Big[\rho\, G_{\mathbf{M}'} \cdot \nabla F_\rho + \mathbf{M}' \cdot (G_{\mathbf{M}'} \cdot \nabla) F_{\mathbf{M}'} \\ + \sigma\, G_{\mathbf{M}'} \cdot \nabla F_\sigma + (\mathbf{B} \times F_{\mathbf{M}'}) \cdot \nabla \times G_{\mathbf{B}}\Big] - (F \leftrightarrow G)\,. \tag{14.18}$$

The important point is that this bracket is only a functional of $(\rho, \mathbf{M}', \sigma, \mathbf{B})$. Therefore the subset of functionals of $(\rho, \mathbf{M}', \sigma, \mathbf{B})$ constitutes a Poisson subalgebra of the full algebra for the two-fluid dynamics.

This provides a reduction of the original Poisson bracket, which should correspond to the motion of a single-fluid description. This is confirmed by the literature, since this bracket is actually just the Poisson bracket of ideal magnetohydrodynamics [102]. Thus, MHD is given by a subalgebra of the two-fluid system.

For a more detailed understanding of the reduction mechanism, the subalgebra is best viewed as a projection of functional derivatives in order to fulfil the physical constraints, in a very similar way as the reduction method by Dirac's constraints.

For MHD, the constraints fix the charge density, the current density and the electric field, because in the MHD limit, they are given by [7]

$$\rho_{\rm e} = 0\,, \qquad \mathbf{J} = \nabla \times \mathbf{B}\,, \qquad \mathbf{E} = \frac{\mathbf{B} \times \mathbf{M}}{\rho}\,. \tag{14.19}$$

The constraint on $\rho_{\rm e}$ translates into the global electroneutrality at scales larger than the Debye length. The constraint on the charge current $\mathbf{J}$ is a non-relativistic regime in the Maxwell-Ampère equation. The constraint on the electric field $\mathbf{E}$ relies on the assumption that the temporal and spatial scales of interest are much larger than the Larmor time and radius, and that the Hall term can be neglected.

The last constraint to be applied involves the pressure tensor, because in the equations of motion, even if each of the two fluids has scalar pressure, the total pressure tensor is not scalar. This constraint can also be related to the charge-weighted entropy density $\sigma_{\rm e}$, which is the last field eliminated in the reduction. In the case of a subalgebra reduction, it corresponds to imposing $\sigma_{\rm e}$ as a prescribed function of the reduced variables:

$$\sigma_{\rm e} = f(\rho, \sigma, \mathbf{M}, \mathbf{B})\,, \tag{14.20}$$

but more specifically $f$ usually depends only on $(\rho, \sigma)$.

In order to apply the subalgebra projection, let us decompose our field variables into two subsets

$$\boldsymbol{\psi}' := (\boldsymbol{\phi}, \boldsymbol{\chi})\,, \tag{14.21}$$

where the first one corresponds to the reduced fields and the second subset gathers the constraints

$$\boldsymbol{\chi} := (\rho, \sigma, \mathbf{M}', \mathbf{B})\,, \qquad \boldsymbol{\phi} := (\rho_{\rm e}, \sigma_{\rm e} - f, \mathbf{J} - \nabla \times \mathbf{B}, \mathbf{E} - \tfrac{\mathbf{B}\times\mathbf{M}}{\rho})\,.$$

The principle of a projection reduction is to consider that the dynamics occurs in a space of constant $\boldsymbol{\phi}$. Then, in the Poisson bracket the functional derivatives $\frac{\delta}{\delta\boldsymbol{\phi}}$ should be projected in order to become tangent to this space. Physically, it means that the variations of the quantities $\boldsymbol{\phi}$ have negligible influence on the observables. This lets free the direction of the projection. For instance the Dirac reduction computes the projector by using the inverse of the matrix of constraints. The

principle of a subalgebra reduction is precisely to project the derivatives parallel to the spaces of constant $\boldsymbol{\chi}$. So, in the coordinates $(\boldsymbol{\phi}, \boldsymbol{\chi})$, the projector is trivially $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ , i.e. all derivatives $\frac{\delta}{\delta\boldsymbol{\phi}}$ are set to zero. This projection just results in a bracket truncation: from the initial Poisson bracket, only the $\mathbf{J}^{\chi\chi}$ part remains, and the constraints have become Casimir invariants.

In order to get the reduced Poisson bracket, we should change variables according to (14.21) before truncating the bracket. But this is actually useless because such a coordinate change does not affect the functional expression of the $\mathbf{J}^{\chi\chi}$. So, the reduced bracket is indeed given by (14.18), which exaclty agrees with the Poisson bracket of ideal magnetohydrodynamics.

It could seem that the argument on projectors is useless, since the reduced bracket was already identified in Eq. (14.18)). However, it is important by showing how the reduction works in the global space $\boldsymbol{\psi}$ (or equivalently $\boldsymbol{\psi}'$), not just in the reduced space $\boldsymbol{\chi}$. Especially, it emphasizes that the derivative operators involved in the reduced bracket are not at constant $(\rho_\mathrm{e}, \sigma_\mathrm{e}, \mathbf{J}, \mathbf{E})$ (as suggested by formula (14.18)), but rather at constant $\boldsymbol{\phi} = (\rho_\mathrm{e}, \sigma_\mathrm{e} - f, \mathbf{J} - \nabla\times\mathbf{B}, \mathbf{E} - \frac{\mathbf{B}\times\mathbf{M}}{\rho})$. This justifies the way the Hamiltonian is reduced in the next section.

In addition, it explains the importance of a subtle point in the derivation. In principle, MHD corresponds to the description of the plasma as a single fluid. The natural variables would be the total fluid mass, entropy, and momentum densities $(\rho, \sigma, \mathbf{M})$; for instance it is the notation used in the introduction. But using these variables in the procedure above does not provide an interesting Poisson bracket, for instance because it results in a static magnetic field[2]. The reason is clear in the projection mechanism described above: practically, the reduced dynamics is projected onto the surface of the constraints, but the projection has a direction, i.e. it changes the value of some variables. If the direction is not properly chosen, then the structure of the dynamics may be spoilt. In the derivation, when taking the constraints, it is important to preserve the total (fluid+field) momentum $\mathbf{M}'$, not just the total plasma momentum $\mathbf{M}$. Accordingly, the momentum density in the introduction should be $\mathbf{M}'$, not $\mathbf{M}$. The reason why they are often not distinguished is that they are almost equal in the MHD regime, as will appear in the next section.

Actually, the coupling between the MHD momentum and the magnetic field comes in an essential way from this choice of momentum, together with the constraint on the electric field. Indeed, the MHD bracket can be obtained from the Poisson bracket of a single non-charged fluid. Then $\rho_\mathrm{e} = 0$ and $\mathbf{J} = 0$, and the dynamics corresponds to a free fluid and a free electromagnetic field

$$\{F, G\} = \int d^3\mathbf{x} \left[\rho\, G_\mathbf{M}\cdot\nabla F_\rho + \mathbf{M}\cdot\ (G_\mathbf{M}\cdot\nabla)F_\mathbf{M} + \sigma\, G_\mathbf{M}\cdot\nabla\ F_\sigma + F_\mathbf{E}\cdot\nabla\times G_\mathbf{B}\right] - F \leftrightarrow G\,.$$

At this stage there is no coupling between the electromagnetic field $(\mathbf{E}, \mathbf{B})$ and the fluid $(\rho, \mathbf{M}, \sigma)$. Then changing variable $\mathbf{M} \to \mathbf{M}'$ and imposing a constraint on the electric field gives exactly the MHD bracket (14.18). So, the coupling field-plasma in MHD results because we impose a constraint on $\mathbf{E}$ in such a way that it preserves the total momentum.

## 14.3 Reduction of the Hamiltonian functional

The reduced Hamiltonian functional is obtained from the initial one (14.10) by changing to the variables (14.21) and then setting $\boldsymbol{\phi} = 0$. Alternatively, in a similar way as for the Poisson bracket, one can just adopt the coordinates[3] (14.4)-(14.9) and take the constraints (14.19)-(14.20) into account. Then the Hamiltonian (14.10) becomes

$$H = \int d^3\mathbf{x}\ \left[\tfrac{1}{2\alpha_1\alpha_2\rho}\left(\alpha_1\alpha_2\mathbf{M}^2 - \mathbf{J}^2\right) + \tfrac{\mathbf{E}^2+\mathbf{B}^2}{2} + U\right] = \int d^3\mathbf{x}\ \left[\tfrac{\mathbf{M}^2}{2\rho} - \tfrac{(\nabla\times\mathbf{B})^2}{2\alpha_1\alpha_2\rho} + \tfrac{(\mathbf{B}\times\mathbf{M})^2}{2\rho^2} + \tfrac{\mathbf{B}^2}{2} + U\right]. \tag{14.22}$$

[2] In addition, it does not provides a subalgebra, unless an additional ordering argument on the electric charge $\rho_\mathrm{e}$ is applied.

[3] In principle, the coordinate change (14.13) should also be taken into account, but for pedagogical purpose, we will do it later. See Eq. (14.29).

The total internal energy density is given by $U = U_1 + U_2$, with for each fluid $U_i(\rho_i, \sigma_i)$ expressed as a function of the final variables $\boldsymbol{\psi}'$, e.g. $\rho_1 = \frac{\alpha_2\rho - \rho_e}{\alpha_2 - \alpha_1}$. Using the constraints $\rho_e = 0$ and $\sigma_e = f(\rho, \sigma)$, it writes as a functional of the variables $(\rho, \sigma)$, i.e. belonging to the reduced subalgebra, and acts as the internal energy of a single fluid, as expected. For this term of the Hamiltonian, the reduction is just a consequence of the constraints.

This is not the case for the whole Hamiltonian, which has two additional terms compared to the Hamiltonian of MHD

$$H_{MHD} = \int d^3\mathbf{x} \left[\frac{\mathbf{M}^2}{2\rho} + \frac{\mathbf{B}^2}{2} + U\right].$$

One can check that these terms induce a motion different from the MHD one

$$\dot{\rho} = -\nabla\cdot\mathbf{M} + \nabla\cdot\left(\frac{\mathbf{M}\times\mathbf{B}}{\rho}\times\mathbf{B}\right), \tag{14.23}$$

$$\dot{\sigma} = -\nabla\cdot(\sigma\frac{\mathbf{M}}{\rho}) + \nabla\cdot\left(\sigma\frac{\mathbf{M}\times\mathbf{B}}{\rho^2}\times\mathbf{B}\right), \tag{14.24}$$

$$\dot{\mathbf{B}} = \nabla\times(\frac{\mathbf{M}\times\mathbf{B}}{\rho}) - \nabla\times(\mathbf{B}\times\mathbf{B}\times\frac{\mathbf{M}\times\mathbf{B}}{\rho^2}), \tag{14.25}$$

$$\dot{\mathbf{M}} = -\nabla\cdot(\frac{\mathbf{M}\mathbf{M}}{\rho}) + (\nabla\times\mathbf{B})\times\mathbf{B} - \nabla\times\left(\frac{\mathbf{M}\times(\mathbf{M}\times\mathbf{B})}{\rho^2}\right)\times\mathbf{B} - \nabla P \tag{14.26}$$

$$-\nabla\cdot\left(\mathbf{B}\times\frac{\mathbf{M}\times\mathbf{B}}{\rho^2}\,\mathbf{M}\right) + \nabla\left(\frac{(\mathbf{M}\times\mathbf{B})^2}{\rho^3}\right)\rho + \nabla\left(\frac{(\mathbf{M}\times\mathbf{B})\times\mathbf{B}}{\rho^2}\right)\cdot\mathbf{M} \tag{14.27}$$

$$-\nabla\times\left(\nabla\times\frac{\nabla\times\mathbf{B}}{\alpha_1\alpha_2\rho}\right)\times\mathbf{B} - \nabla\left(\frac{(\nabla\times\mathbf{B})^2}{2\alpha_1\alpha_2\rho^2}\right)\rho, \tag{14.28}$$

where the pressure is given by the usual formula $P = \left(\rho\frac{\partial U}{\partial\rho} + \sigma\frac{\partial U}{\partial\sigma} - U\right)$. In each equation, the terms of the MHD dynamics (14.1) are present, but there are additional terms, whose contribution can be evaluated by two ordering parameters, which are small in the range of application of MHD.

Indeed, in equations (14.23), (14.24), and (14.25), the additional terms are of order (the units are restored here) $\epsilon_1 = \frac{\mathbf{B}^2}{\rho\mu_0 c^2}$. The same is obtained when the additional terms in (14.26) and (14.27) are compared to the MHD term $-\nabla\cdot(\frac{\mathbf{M}\mathbf{M}}{\rho})$. All these additional terms result from the presence of $\frac{(\mathbf{B}\times\mathbf{M})^2}{2\rho^2}$ in the Hamiltonian (14.22), which is of order $\epsilon_1$ compared to the MHD kinetic energy $\frac{\mathbf{M}^2}{2\rho}$.

On the other hand, the additional term in Eq. (14.26) can be compared to the MHD term $(\nabla\times\mathbf{B})\times\mathbf{B}$, and then it is of order $\epsilon_{1b} = \frac{\mathbf{u}^2}{c^2}$, where $\mathbf{u} = \frac{\mathbf{M}}{\rho}$ is the fluid velocity. So, $\epsilon_{1b}$ is small in the non-relativistic regime, which is an assumption of standard MHD; for instance, it is invoked to get the constraint $\mathbf{J} = \nabla\times\mathbf{B}$, by excluding the term $\partial_t\mathbf{E}$ from Maxwell-Ampère's equation.

The parameter $\epsilon_1$ is small when the mass energy of the plasma is much larger than the magnetic energy, which means when the plasma is dense enough, or the magnetic field strength is not too high. The parameters $\epsilon_1$ and $\epsilon_{1b}$ are related to the parameter $\beta = \frac{\mathbf{B}^2}{\mu_0\rho\mathbf{u}^2} = \frac{\epsilon_1}{\epsilon_{1b}}$. It is the ratio of the magnetic over kinetic energy (or pressure), and is of order 1 in usual MHD, since otherwise one of the terms in $H_{MHD}$ can be dropped.

Another parameter is present, because of the additional terms (14.28), which are of order $\epsilon_2 = \frac{\nabla^2}{\alpha_1\alpha_2\rho\mu_0}$, where $\nabla$ corresponds to the (inverse of the) spatial scale of variations of the magnetic field. This is a result of the contribution $-\frac{(\nabla\times\mathbf{B})^2}{2\alpha_1\alpha_2\rho}$ in the Hamiltonian (14.22), which is of order $\epsilon_2$ compared to the magnetic energy $\frac{\mathbf{B}^2}{2}$.

The ordering parameter $\epsilon_2$ induces a large-scale criterion. It is small for large enough magnetic scales $L_{\mathbf{B}} := \left(\frac{\|\nabla\mathbf{B}\|}{\|\mathbf{B}\|}\right)^{-1}$, since $\epsilon_2 << 1$ means that $L_{\mathbf{B}} >> L_c$, with the critical scale given, by $L_c = (\alpha_1\alpha_2\mu_0\rho)^{-1/2}$. Using that the Larmor frequency for the fluid $s$ is $f_{Ls} = \|\alpha_s\mathbf{B}\|$, the parameter $L_c$ can be related to $\epsilon_1$ through the relation $L_c = \frac{c\sqrt{\epsilon_1}}{f_L}$, defining the mean Larmor frequency $f_L := \sqrt{f_{L1}f_{L2}}$. Last, defining a typical Larmor radius by $r_L := \|\mathbf{u}\|/f_L$, we get the relation $L_c = r_L.\sqrt{\beta}$. So, the critical scale is of the order of the typical Larmor radius, and this scale is excluded from the MHD description, because at this scale, even the multi-fluid description is no more convenient since kinetic effects are important.

For simplicity, the argument above was made using the coordinate $\mathbf{M}$. Actually, as a result of the variables defining the reduced subalgebra, the reduced Hamiltonian (14.22) must be taken as a function of $\mathbf{M}'$, instead of $\mathbf{M}$. But those variables are linked by the relation

$$\mathbf{M}' = \mathbf{M} + \mathbf{E} \times \mathbf{B} = \mathbf{M} - \tfrac{\mathbf{M}\times\mathbf{B}}{\rho} \times \mathbf{B} = \mathbf{M}_{\|} + (1+\epsilon_1)\mathbf{M}_{\perp}\,, \tag{14.29}$$

where $\mathbf{M}_{\|} := \frac{\mathbf{B}\mathbf{B}\cdot\mathbf{M}}{\|\mathbf{B}\|^2}$ is the component of the momentum parallel to the magnetic field, and $\mathbf{M}_{\perp} := \frac{1-\mathbf{B}\mathbf{B}}{\|\mathbf{B}\|^2}\cdot\mathbf{M}$ is its component perpendicular to $\mathbf{B}$. Computationally, it complicates much the equations of motion (14.23)-(14.28), but the ordering parameter involved is again $\epsilon_1$, and all the conclusions above still apply.

Interestingly, formula (14.29) shows that when $\epsilon_1$ is small, the variables $\mathbf{M}$ and $\mathbf{M}'$ are close to each other. Then the momentum used in MHD can be considered to be the plasma momentum, since the momentum density of the electromagnetic field is much smaller than the plasma momentum density.

As a result, the reduced motion is a Hamiltonian perturbation from the MHD dynamics. It is ordered by small parameters, which all come from the Hamiltonian functional:

$$H = H_{MHD} + \epsilon_1 H_1 + \epsilon_2 H_2\,,$$

with $H_1 = \frac{(\mathbf{B}\times\mathbf{M})^2}{2\rho^2}$ and $H_2 = -\frac{(\nabla\times\mathbf{B})^2}{2\alpha_1\alpha_2\rho}$. A consequence is that the specific form of the constraints on $\mathbf{J}$ and $\mathbf{E}$ are not essential in the derivation. The only requirement is that their associated contributions in the Hamiltonian must be small.

In this framework, the parameters $\epsilon_1$ and $\epsilon_2$ define the range of validity of MHD. They correspond to the size of the terms that the MHD dynamics neglects compared to the complete reduced Hamiltonian (14.22). The MHD motion is valid only when they are small. In the limit where they are not negligible, the Hamiltonian (14.22) cannot be simplified. However, in this case, care should be taken to the consistency of the approximations made, because the reduction of the Poisson bracket implicitly used the non-relativistic regime when imposing the constraint $\mathbf{J} - \nabla\times\mathbf{B}$, and the large-scale regime when excluding kinetic effects. These hypotheses are not compatible with large values of the parameters $\epsilon_1$ and $\epsilon_2$.

A last point can be considered. We saw that the subalgebra reduction, which concerns the bracket is exact, whereas the reduction of the Hamiltonian functional involved an ordering approximation. So, the reduction we performed lead us from the two-fluid dynamics to a dynamics that is close to the MHD system, but not exactly identical. One can wonder if the link cannot be made exact. This question relies on the observation that in the subalgebra reduction, the constraints can be arbitrary functions of the reduced variables, whence the idea to look for other constraints which would give exactly $H = H_{MHD}$.

The corresponding constraints should make the two excess terms in formula (14.22) disappear. The order in the parameters $\alpha_i$ already shows that the constraint involving $\mathbf{J}$ must be $\mathbf{J} = 0$. With $\rho_{\mathrm{e}} = 0$, these constraints means that we are dealing with the non charged fluid, mentioned at the end of the previous section. Now, the constraint on the electric field must be chosen so that

$$\tfrac{\mathbf{M}^2}{2\rho} + \tfrac{\mathbf{E}^2}{2} = \tfrac{\mathbf{M}'^2}{2\rho}\,. \tag{14.30}$$

Using (14.29), we straightforwardly get

$$\mathbf{E} = \tfrac{2}{1-\epsilon_1}\,\tfrac{\mathbf{B}\times\mathbf{M}}{\rho}\,. \tag{14.31}$$

Thus, the Hamiltonian structure of MHD is exactly given by the dynamics of a non charged fluid ($\rho_{\mathrm{e}} = 0$ and $\mathbf{J} = 0$) coupled with the electromagnetic field by imposing the constraint (14.31) while keeping the total momentum conserved. This gives an exact link between the Hamiltonian structures of MHD and of fluid dynamics. Notice that in the limit $\epsilon_1 = 0$, the constraint (14.31) does not converge towards the physical constraint $\mathbf{E} = \frac{\mathbf{B}\times\mathbf{M}}{\rho}$, between them there is a ratio of 2. This has no effect on the reduced dynamics, but it impacts the electric field, which is a dynamical variable (although not independent).

## 14.4 Extension to multi-fluid, relativistic and anisotropic-pressure cases

If the system has more than two fluids, the reduction to the MHD Poisson bracket can be performed in a similar way as well, but the role of the single-fluid description is clearer. Indeed, starting from a multi-fluid description, let us include the total-fluid densities into the fluid coordinates $\boldsymbol{\xi}_1 \to \boldsymbol{\xi}_T$, with $\boldsymbol{\xi} = (\rho, \mathbf{M}, \sigma)$, and $\boldsymbol{\xi}_T = \sum_i \boldsymbol{\xi}_i$. Then, the constraints on $\sigma_\mathrm{e}$ and $\mathbf{J}$ can be imposed, because the set of variables $(\rho, \mathbf{M}, \sigma, \rho_\mathrm{e}, \mathbf{E}, \mathbf{B})$ define a subalgebra, as a result of the linearity of the Poisson bracket (14.2). This brings to a sub-system of Sec. 14.1, and the last two constraints can be imposed just the same way as previously.

Since there are more initial variables, additional constraints must be imposed. From a formal point of view, they makes the reduction softer. One of them can compensate the excess "pressure" term $\sum_i \rho_i \delta\mathbf{u}_i^2$ in the kinetic part of the Hamiltonian:

$$\sum_i \frac{\mathbf{M}_i^2}{2\rho_i} = \frac{\mathbf{M}^2}{2\rho} + \frac{\sum_i \rho_i \delta\mathbf{u}_i^2}{2}\,,$$

with $\delta\mathbf{u}_i = \mathbf{u}_i - \mathbf{u}$, and $\mathbf{u} = \frac{\mathbf{M}}{\rho}$. Then the reduction of the Hamiltonian is the same as in the previous section. Another additional constraint can compensate the ordering terms coming from the change of variable $\mathbf{M} \to \mathbf{M}'$, to get the same result as in (14.30), but without constraining the electric field as was done in (14.31).

In a deeper way, the Hamiltonian structure of MHD even in the multi-fluid context can be viewed as a consequence of the presence of the Poincaré group of transformations. As the Galilean group does not let Maxwell's equations invariant, we consider a relativistic framework, whose bracket is the same as in the non-relativistic case, but with the variables defined in the laboratory frame [69, 71, 72]. Indeed, in the relativistic Vlasov-Maxwell system, the Poincaré group is a subgroup of the canonical transformations [5, 43]. The corresponding subalgebra spreads to the multi-fluid dynamics, because this last is just a subalgebra of the parent kinetic model and includes the densities of the generators of the Poincaré group. Besides the energy $H$, these generators are the total momentum $\overrightarrow{M}$ and angular momentum $\overrightarrow{L}$, as well as the generator of Lorentz transformations $\overrightarrow{N}$. They are defined by

$$H = \int d^3\mathbf{x}\ \mathcal{H}\,, \qquad \overrightarrow{M} = \int d^3\mathbf{x}\ \mathbf{M}'\,, \qquad \overrightarrow{L} = \int d^3\mathbf{x}\ \mathbf{x}\times\mathbf{M}'\,, \qquad \overrightarrow{N} = \int d^3\mathbf{x}\ \left(\mathbf{x}\mathcal{H} - t\mathbf{M}'\right),$$

where $\mathcal{H}$ is the relativistic expression for the plasma kinetic and internal energy density [69, 72]

$$\mathcal{H} = (\gamma^2 - 1)(\rho + P) + \gamma^2 U + \frac{\mathbf{E}^2 + \mathbf{B}^2}{2}\,, \qquad \text{with } \gamma = (1 - \mathbf{u}^2)^{-1/2} \text{ the relativistic factor.}$$

When there is only one species, the fluid model is given exactly by the subalgebra corresponding to the densities $\rho$, $\sigma$, $\mathbf{M}'$, $\mathbf{E}$, and $\mathbf{B}$, that define these generators ($\sigma$ is involved in the definition of the Hamiltonian through the internal energy $U$). When there are several species, the multi-fluid model is larger than this subalgebra, but contains it in an own subalgebra, which generates the corresponding single-fluid description. Especially, this enlightens the role of the total momentum instead of the fluid momentum, and explains why we noticed that the reduction to a single-fluid model is such a cornerstone of the reduction to MHD.

More precisely, the subalgebra generated by the densities of the Poincaré generators contains also the charge density $\rho_\mathrm{e}$, as shown by the first term in the last line of Eq. (14.11). The reason is that even if $\rho_\mathrm{e}$ is not involved in the global conservation laws (since it only implies exchanges between the plasma momentum and the electromagnetic-field momentum), all the same it is involved in the local conservation laws. So, in order to consider not the Poincaré subalgebra but the associated densities, $\rho_\mathrm{e}$ must be kept as a dynamical variable. Now, in the single-fluid model defined by $(\rho, \rho_\mathrm{e}, \sigma, \mathbf{M}', \mathbf{E}, \mathbf{B})$, the only effect of the charge density is to generate a coupling between the dynamics of the electric field and the one of the momentum density. The MHD model is obtained by considering that at large scale, the plasma is neutral, and the coupling just makes the electric field

a dependent variable, e.g. $\mathbf{E} = -\mathbf{u} \times \mathbf{B}$, while preserving the conservation of the total momentum. So, the constraint on $\mathbf{E}$ plays a more important role than the constraint on the current.

This agrees with the role of the constraint for $\mathbf{J}$, which is also related to the question of relativistic regime or not. In previous sections, the derivation of MHD relied on a non-relativistic assumption, and especially for the constraint $\mathbf{J} = \nabla \times \mathbf{B}$. On the other hand, the above argument takes place in the relativistic framework, and it does not rely explicitly on this constraint for $\mathbf{J}$. The common point is that the Poisson bracket is the same between the relativistic and non-relativistic dynamics, and a subalgebra reduction relies more essentially on the choice of the reduced variables than on the choice of constraints.

For the sake of simplicity, we considered standard isotropic-pressure MHD. When the pressure tensor is considered anisotropic, e.g. as a result of the Chew-Goldberger-Low equations [37,72,104], the Poisson bracket can be reduced the same way as above since the changes only concern the Hamiltonian. Its internal energy term acquires a dependence in the magnetic field $\mathbf{B}$, and in the relativistic case, it has an additional term involving the pressure tensor and the electric field [72]. This generates additional terms in the reduced Hamiltonian (14.22), and an analysis similar to Sec. 14.3 must be performed to verify under what conditions they can be neglected.

## Conclusion

The Poisson bracket of magnetohydrodynamics is given by a Lie subalgebra of the Hamiltonian structure of the two-fluid dynamics, imposing as constraints the MHD relations for the electric field, the current density, and the charge density. It can be obtained from a single non-charged fluid coupled with the fields only through the constraint on the electric field. In the reduction process, the constraints must be applied keeping constant not the plasma momentum density but the total momentum density.

The reduced Hamiltonian functional is a perturbation of the Hamiltonian for MHD. Two parameters are involved, the first one can be viewed as a non-relativistic or as a low magnetic field condition, and the second one as a large-scale regime. When they are small, the Hamiltonian and the dynamics become the one of MHD. Then, the plasma momentum density and the total momentum density become equal.

The two ordering parameters of the Hamiltonian perturbation are consistent with the usual framework of MHD, but they are not additional requirements. They were already considered small as a result of the fluid framework and of the imposed constraints.

The reduction also applies to the multi-fluid, relativistic case with anisotropic pressure tensor. Then, the reduction to the MHD bracket more clearly appears to be linked to the subalgebra of Poincaré generators, with a constraint on the electric field, while keeping the total momentum conserved. The reduction to the MHD Hamiltonian functional involves more precise requirements, but more constraints are available to get them.

An interesting extension of the work will be to consider variations of the MHD model, and especially Hall-MHD, whose Hamiltonian structure seems to involve significant differences with the Hamiltonian structure of MHD. Whether it is given by a subalgebra reduction or not, it should be an interesting application and could suggest developments for Hamiltonian reduction methods.

As a complementary investigation, we have studied whether the Hamiltonian reduction derived in this chapter (starting from the two-fluid dynamics) could be addressed also by using Dirac's theory of constraints, with the natural MHD constraints (14.19)-(14.20). The corresponding reduction is much more intricate than the natural subalgebra method presented above. Obtaining the matrix of constraints already requires some computations, and trying to invert it is still heavier. Especially, it is an $8 \times 8$ matrix, whose coefficients are operators. In addition to the algebraic manipulations, much care must be taken to all the arbitrariness involved when inverting operators. Unfortunately, this finally drove us to the conclusion that the matrix was not invertible, which confirms that the Dirac method is not relevant for the MHD reduction.

# Appendix A: List of papers

## Published papers:

[22] A. J. Brizard and L. de Guillebon, "Comment on “Geometric phase of the gyromotion for charged particles in a time-dependent magnetic field [Phys. Plasmas 18, 072505 (2011)]“", Phys. Plasm. **19**, 094701 (2012).

[58] L. de Guillebon and C. Chandre, "Hamiltonian structure of reduced fluid models for plasmas obtained from a kinetic description", Phys. Lett. A **376**, 3172 (2012).

[34] C. Chandre, L. de Guillebon, A. Back, E. Tassi, and P. J. Morrison, "On the use of projectors for Hamiltonian systems and their relationship with Dirac brackets", J. Phys. A: Math. Theor. **46**, 125203 (2013).

[115] P. J. Morrison, M. Vittot, and L. de Guillebon, "Magnetic moment type of lifting from particle dynamics to Vlasov-Maxwell dynamics", Phys. Plasmas **20**, 032109 (2013).

[59] L. de Guillebon and M. Vittot, "A gyro-gauge independent minimal guiding-center reduction by Lie transforming the velocity vector field", Phys. Plasmas **20**, 082505 (2013).

[60] L. de Guillebon and M. Vittot, "Gyro-gauge independent formulation of the guiding-center reduction to arbitrary order in the Larmor radius", accepted in Plasma Phys. Control. Fusion (2013), Plasma Phys. Control. Fusion **55**, 105001 (2013).

## Submitted papers:

[61] L. de Guillebon and M. Vittot, "On an intrinsic formulation of the guiding-center anholonomy and gyro-gauge arbitrariness", submitted, arXiv:1304.4563 (2013).

## In preparation:

[a] L. de Guillebon, N. Tronko, M. Vittot, and Ph. Ghendrih, "Dynamical reduction for charged particles in a strong magnetic field without guiding-center", in preparation.

[b] L. de Guillebon and C. Chandre, "On the role of quarter-canonical Poisson brackets in Dirac and subalgebra reductions", in preparation.

[c] L. de Guillebon and C. Chandre, "Hamiltonian derivation of magnetohydrodynamics from multi-fluid dynamics", in preparation.

[d] L. de Guillebon, P. J. Morrison, A. J. Brizard, and M. Vittot, "Lifting coordinate transformations to the Hamiltonian structure of the corresponding field dynamics", in preparation.

[e] A. J. Brizard, P. J. Morrison, L. de Guillebon, and M. Vittot, "Lifting of the Vlasov-Maxwell Bracket by Lie-transform Method", in preparation.

[f] L. de Guillebon and C. Chandre, "On the role of Dirac's theory of constraints in a Hamiltonian derivation of gyrokinetics", in preparation.

# Appendix B: Complementary paper

## Comment on "Geometric phase of the gyromotion for charged particles in a time-dependent magnetic field" [Phys. Plasmas 18, 072505 (2011)]

Alain J. Brizard[1] and Loïc de Guillebon[2]
[1] *Department of Physics, Saint Michael's College, Colchester, VT 05439, USA*
[2] *Centre de Physique Théorique Aix-Marseille Université, CNRS (UMR 7332), 13288 Marseille cedex 09 France*

The geometric analysis of the gyromotion for charged particles in a time-dependent magnetic field by J. Liu and H. Qin [Phys. Plasmas **18**, 072505 (2011)] is reformulated in terms of the spatial angles that represent the instantaneous orientation of the magnetic field. This new formulation, which includes the equation of motion for the pitch angle, clarifies the decomposition of the gyroangle-averaged equation of motion for the gyrophase into its dynamic and geometric contributions.

In a recent paper [1], Liu and Qin studied the dynamics of the gyrophase of a charged particle moving in a time-dependent (but uniform) magnetic field $\mathbf{B}(t) \equiv B_x(t)\,\widehat{\mathsf{x}} + B_y(t)\,\widehat{\mathsf{y}} + B_z(t)\,\widehat{\mathsf{z}}$. Liu and Qin then obtained an expression for the equation of motion for the gyrophase in terms of the Cartesian components $(B_x, B_y, B_z)$ and their time derivatives. By using this Cartesian formulation, Liu and Qin demonstrated the existence of an anholonomic (i.e., path-dependent) geometric contribution to the gyrophase shift that depends on the evolution history of the magnetic field during a gyro-period.

In related previous work, Littlejohn [2, 3] showed that the Hamiltonian theory associated with guiding-center motion in an inhomogeneous magnetic field [4] displays an anholonomic geometric phase associated with the gyrogauge vector $\mathbf{R} \equiv \nabla\widehat{\mathsf{e}}_1 \cdot \widehat{\mathsf{e}}_2$ constructed from two basis unit vectors $\widehat{\mathsf{e}}_1$ and $\widehat{\mathsf{e}}_2 \equiv \widehat{\mathsf{b}} \times \widehat{\mathsf{e}}_1$ that span the plane that is locally perpendicular to the local magnetic unit vector $\widehat{\mathsf{b}}(t) \equiv \mathbf{B}(t)/B(t)$. Liu and Qin [1], on the other hand, showed that anholonomic geometric phases also exist in the problem of the gyromotion of charged particles (i.e., not guiding-centers) in a time-dependent magnetic field.

In the present Comment, we show how the analysis of Liu and Qin can be greatly simplified by using the polar angle $\theta$ and the azimuthal angle $\varphi$ describing the orientation of the spatially-uniform magnetic field [3], where $\cos\theta(t) \equiv B_z(t)/B(t)$ and $\tan\varphi(t) \equiv B_y(t)/B_x(t)$. We also clarify the role played by the pitch angle in the evolution of the gyrophase. We now reformulate the analysis of Liu and Qin [1] concerning the case of time-dependent (but uniform) magnetic field $\mathbf{B}(t)$.

First, we introduce the decomposition of the particle velocity $\mathbf{v}$ in terms of its pitch angle $\lambda$ (i.e., $\mathbf{v}\cdot\widehat{\mathsf{b}} = |\mathbf{v}|\cos\lambda$) and gyroangle $\zeta$ (i.e., $\partial\mathbf{v}/\partial\zeta = \mathbf{v}\times\widehat{\mathsf{b}}$):

$$\mathbf{v} \equiv v\left(\cos\lambda\,\widehat{\mathsf{b}} + \sin\lambda\,\widehat{\mathsf{c}}\right), \tag{1}$$

where the unit vector $\widehat{\mathsf{c}} \equiv \widehat{\mathsf{a}} \times \widehat{\mathsf{b}} \equiv \partial\widehat{\mathsf{a}}/\partial\zeta$ depends explicitly on $\zeta$ and the speed $v = |\mathbf{v}|$ is a constant of the motion.

Next, we obtain the equations of motion for $\lambda$ and $\zeta$ by introducing the *magnetic* unit vectors [3]

$$\left.\begin{aligned} \widehat{\mathsf{b}} &= \cos\theta\,\widehat{\mathsf{z}} + \sin\theta\,\widehat{\rho} \\ \widehat{\theta} &= -\sin\theta\,\widehat{\mathsf{z}} + \cos\theta\,\widehat{\rho} = \partial\widehat{\mathsf{b}}/\partial\theta \\ \widehat{\varphi} &= -\sin\varphi\,\widehat{\mathsf{x}} + \cos\varphi\,\widehat{\mathsf{y}} = \partial\widehat{\rho}/\partial\varphi \end{aligned}\right\}. \tag{2}$$

Here, the "radial" vector $\widehat{\mathsf{b}}$ points in the direction of the magnetic field at each point in space (for a time-dependent uniform field, this direction is only a function of time). Since the unit vectors (2) satisfy $\widehat{\mathsf{b}} \equiv \widehat{\theta} \times \widehat{\varphi}$, the unit vectors $\widehat{\theta} \equiv \widehat{\mathsf{e}}_1$ and $\widehat{\varphi} \equiv \widehat{\mathsf{e}}_2$ provide a possible choice of unit vectors $(\widehat{\mathsf{e}}_1, \widehat{\mathsf{e}}_2)$ in the plane perpendicular to the unit magnetic vector $\widehat{\mathsf{b}}$ (i.e., tangent to the unit sphere).

### I. GYROGAUGE GEOMETRY

With the choice (2) for the unit vectors $(\widehat{\mathsf{b}}, \widehat{\mathsf{e}}_1, \widehat{\mathsf{e}}_2)$, we construct the *magnetic* one-forms

$$\left.\begin{aligned} \mathsf{d}\widehat{\mathsf{b}}\cdot\widehat{\mathsf{e}}_1 &= \mathsf{d}\theta \\ \mathsf{d}\widehat{\mathsf{b}}\cdot\widehat{\mathsf{e}}_2 &= \sin\theta\,\mathsf{d}\varphi \\ \mathsf{d}\widehat{\mathsf{e}}_1\cdot\widehat{\mathsf{e}}_2 &= \cos\theta\,\mathsf{d}\varphi \end{aligned}\right\}, \tag{3}$$

where $\mathsf{d}$ denotes an exterior derivative [5, 6]. Since the vectors $(\widehat{\mathsf{b}}, \widehat{\mathsf{e}}_1, \widehat{\mathsf{e}}_2)$ are orthogonal, we also have $\mathsf{d}\widehat{\mathsf{e}}_i \cdot \widehat{\mathsf{b}} = -\,\mathsf{d}\widehat{\mathsf{b}} \cdot \widehat{\mathsf{e}}_i$ and $\mathsf{d}\widehat{\mathsf{e}}_j \cdot \widehat{\mathsf{e}}_i = -\,\mathsf{d}\widehat{\mathsf{e}}_i \cdot \widehat{\mathsf{e}}_j$, for $i, j = 1, 2$. These definitions are not unique, however, since a rotation of the perpendicular unit vectors $(\widehat{\theta}, \widehat{\varphi}) \to (\widehat{\theta}', \widehat{\varphi}')$ about the $\widehat{\mathsf{b}}$-axis generated by the gyrogauge angle $\psi$:

$$\begin{pmatrix} \widehat{\theta}' \\ \widehat{\varphi}' \end{pmatrix} \equiv \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \cdot \begin{pmatrix} \widehat{\theta} \\ \widehat{\varphi} \end{pmatrix} \tag{4}$$

leads to the new one-forms

$$\left.\begin{aligned} \mathsf{d}\widehat{\mathsf{b}} \cdot \widehat{\mathsf{e}}_1' &= \cos\psi\, \mathsf{d}\theta + \sin\theta \sin\psi\, \mathsf{d}\varphi \equiv \omega_2\, \mathsf{d}t \\ \mathsf{d}\widehat{\mathsf{b}} \cdot \widehat{\mathsf{e}}_2' &= -\sin\psi\, \mathsf{d}\theta + \sin\theta \cos\psi\, \mathsf{d}\varphi \equiv -\omega_1\, \mathsf{d}t \\ \mathsf{d}\widehat{\mathsf{e}}_1' \cdot \widehat{\mathsf{e}}_2' &= \cos\theta\, \mathsf{d}\varphi + \mathsf{d}\psi \equiv \omega_3\, \mathsf{d}t \end{aligned}\right\}. \tag{5}$$

Equation (5) shows the deep connection between the magnetic one-forms and the Eulerian angular frequencies $(\omega_1, \omega_2, \omega_3)$, which are defined in terms of the Euler angles $(\varphi, \theta, \psi - \pi/2)$. We can thus imagine an infinitesimaly thin symmetric top spinning about its axis of symmetry $\widehat{\mathsf{b}} \equiv \widehat{\mathsf{e}}_3$ (at constant $\theta$ and $\varphi$ with $d\psi/dt \neq 0$), which also undergoes precession (at constant $\theta$ with $d\varphi/dt \neq 0$) and nutation (at constant $\varphi$ with $d\theta/dt \neq 0$).

If we define the gyrogauge one-form

$$\mathcal{R} \equiv \cos\theta\, \mathsf{d}\varphi, \tag{6}$$

and we denote its gyrogauge transformation as

$$\mathcal{R}' \equiv \mathcal{R} + \mathsf{d}\psi, \tag{7}$$

then the property of gyrogauge invariance

$$\mathsf{d}\mathcal{R}' = \mathsf{d}\mathcal{R} \tag{8}$$

is guaranteed by the identity $\mathsf{d}^2 \equiv 0$ (corresponding to the vector identity $\nabla \times \nabla \equiv 0$). The gyrogauge one-form (6) has a simple geometrical interpretation in terms of the solid-angle two-form [5]

$$\Phi \equiv \mathsf{d}\theta \wedge \sin\theta\, \mathsf{d}\varphi = -\,\mathsf{d}\Big(\cos\theta\, \mathsf{d}\varphi\Big) = -\,\mathsf{d}\mathcal{R}. \tag{9}$$

Hence, according to Eq. (8), the solid-angle two-form (9) is a gyrogauge-invariant. In addition, using Stokes' Theorem [5], the solid angle $\Omega \equiv \int_{\mathcal{D}} \Phi$ defined by the open surface $\mathcal{D}$ on the unit sphere is also expressed as

$$\Omega = -\int_{\mathcal{D}} \mathsf{d}\mathcal{R} \equiv -\oint_{\partial\mathcal{D}} \mathcal{R}, \tag{10}$$

where $\partial\mathcal{D}$ denotes the boundary of $\mathcal{D}$.

## II. PITCH-ANGLE AND GYROANGLE DYNAMICS

Next, we introduce the gyration unit vectors $\widehat{\mathsf{b}} \equiv \widehat{\mathsf{c}} \times \widehat{\mathsf{a}}$ in the plane perpendicular to $\widehat{\mathsf{b}}$:

$$\widehat{\mathsf{a}} \equiv \cos\zeta\, \widehat{\theta} - \sin\zeta\, \widehat{\varphi}, \tag{11}$$

$$\widehat{\mathsf{c}} \equiv -\sin\zeta\, \widehat{\theta} - \cos\zeta\, \widehat{\varphi}. \tag{12}$$

These definitions are gyrogauge-invariant (i.e., $\widehat{\mathsf{a}}' = \widehat{\mathsf{a}}$ and $\widehat{\mathsf{c}}' = \widehat{\mathsf{c}}$) under the transformation $\zeta' = \zeta + \psi$ and Eq. (4). Using Eq. (1), the equation of motion $d\mathbf{v}/dt = \omega_{\mathrm{c}}\, \mathbf{v} \times \widehat{\mathsf{b}}$ becomes

$$\begin{aligned} 0 = \omega_{\mathrm{c}} \sin\lambda\, \widehat{\mathsf{a}} &+ \cos\lambda \left( \frac{d\widehat{\mathsf{b}}}{dt} + \frac{d\lambda}{dt}\, \widehat{\mathsf{c}} \right) \\ &+ \sin\lambda \left( \frac{d\widehat{\mathsf{c}}}{dt} - \frac{d\lambda}{dt}\, \widehat{\mathsf{b}} \right), \end{aligned} \tag{13}$$

where $\omega_{\mathrm{c}}(t) \equiv qB(t)/m$ denotes the time-dependent cyclotron frequency and, using Eqs. (11)-(12), we find

$$\begin{aligned} \frac{d\widehat{\mathsf{b}}}{dt} &= \frac{d\theta}{dt}\, \widehat{\theta} + \sin\theta\, \frac{d\varphi}{dt}\, \widehat{\varphi} \\ &= \left( \cos\zeta\, \frac{d\theta}{dt} - \sin\zeta \sin\theta\, \frac{d\varphi}{dt} \right) \widehat{\mathsf{a}} \\ &\quad - \left( \sin\zeta\, \frac{d\theta}{dt} + \cos\zeta \sin\theta\, \frac{d\varphi}{dt} \right) \widehat{\mathsf{c}}, \end{aligned} \tag{14}$$

$$\begin{aligned} \frac{d\widehat{\mathsf{c}}}{dt} &= -\left( \frac{d\zeta}{dt} - \cos\theta\, \frac{d\varphi}{dt} \right) \widehat{\mathsf{a}} \\ &\quad + \left( \sin\zeta\, \frac{d\theta}{dt} + \cos\zeta \sin\theta\, \frac{d\varphi}{dt} \right) \widehat{\mathsf{b}}. \end{aligned} \tag{15}$$

We note that these expressions can also be expressed as $d\widehat{\mathsf{b}}/dt = \omega_1\, \widehat{\mathsf{a}} - \omega_2\, \widehat{\mathsf{c}}$ and $d\widehat{\mathsf{c}}/dt = \omega_3\, \widehat{\mathsf{a}} + \omega_2\, \widehat{\mathsf{b}}$ in terms of the Eulerian angular frequencies $(\omega_1, \omega_2, \omega_3)$ defined in Eq. (5) with the substitution $\psi \to \pi/2 - \zeta$ (i.e., gyromotion occurs as a counter-rotation about $\widehat{\mathsf{b}}$). With these new definitions for the Eulerian frequencies, we easily recover the standard relations $\partial\omega_1/\partial\zeta = -\omega_2$, $\partial\omega_2/\partial\zeta = \omega_1$, and $\partial\omega_3/\partial\zeta = 0$.

Equation (13) can be divided into two separate equations of motion:

$$0 = \frac{d\lambda}{dt} + \frac{d\widehat{\mathsf{b}}}{dt} \cdot \widehat{\mathsf{c}}, \tag{16}$$

$$0 = \omega_{\mathrm{c}} + \frac{d\widehat{\mathsf{c}}}{dt} \cdot \widehat{\mathsf{a}} + \cot\lambda\, \frac{d\widehat{\mathsf{b}}}{dt} \cdot \widehat{\mathsf{a}}. \tag{17}$$

By using Eq. (14), Eq. (16) yields the equation of motion for the pitch angle $\lambda$:

$$\frac{d\lambda}{dt} = \sin\zeta\, \frac{d\theta}{dt} + \cos\zeta \sin\theta\, \frac{d\varphi}{dt} \equiv \omega_2. \tag{18}$$

While this equation is not considered by Liu and Qin [1], it plays an important role in the evolution of the gyrophase [see Eq. (20) below]. We note that, since the angular velocities $d\theta/dt$ and $d\varphi/dt$ are gyroangle-independent (i.e., they represent the rate of change of the orientation of the magnetic field), the pitch-angle equation (18) satisfies $\langle d\lambda/dt \rangle = 0$ (which is valid for a uniform magnetic field), where $\langle \cdots \rangle$ denotes a gyroangle-average.

By using Eqs. (14)-(15), on the other hand, Eq. (17) yields the equation of motion for the gyroangle $\zeta$:

$$\begin{aligned}\frac{d\zeta}{dt} &= \omega_{\rm c} + \cos\theta\,\frac{d\varphi}{dt} \\ &\quad + \cot\lambda\left(\cos\zeta\,\frac{d\theta}{dt} - \sin\zeta\,\sin\theta\,\frac{d\varphi}{dt}\right) \\ &\equiv \omega_{\rm c} + \cos\theta\,\frac{d\varphi}{dt} + \omega_1\,\cot\lambda. \end{aligned} \tag{19}$$

Equation (19) corresponds exactly to Eq. (7) of the paper [1] by Liu and Qin, with the substitution $\zeta \to -\zeta - \pi/2$ and

$$\frac{B_z}{B}\left(\frac{B_x\,\dot{B}_y - B_y\,\dot{B}_x}{B_x^2 + B_y^2}\right) = \cos\theta\,\frac{d\varphi}{dt}.$$

In Eq. (19), $\omega_{\rm c}$ is described as the *dynamical* term by Liu and Qin, while $\cos\theta\ d\varphi/dt$, which is clearly related to the gyrogauge one-form (6), is described as the *geometric* term. The last term on the right side of Eq. (19) is described by Liu and Qin as an *adiabatic* term (because it is shown in the Appendix of Ref. [1] to be one order higher than the geometric term when the magnetic field evolves slowly compared to the gyration period). Instead, we use the pitch-angle equation (18) to write $\partial(d\lambda/dt)/\partial\zeta = \omega_1$ (once again valid for a uniform magnetic field), so that Eq. (19) is written as

$$\frac{d\zeta}{dt} = \omega_{\rm c} + \cos\theta\,\frac{d\varphi}{dt} + \frac{\partial}{\partial\zeta}\left(\cot\lambda\,\frac{d\lambda}{dt}\right). \tag{20}$$

The third term in Eq. (19) therefore appears as an exact gyroangle derivative, which disappears when Eq. (19) is gyroangle-averaged:

$$\left\langle\frac{d\zeta}{dt}\right\rangle = \omega_{\rm c} + \cos\theta\,\frac{d\varphi}{dt}. \tag{21}$$

We note that Eq. (21) is gyrogauge-invariant since, under a gyrogauge transformation generated by $\psi$, we have

$$\left(\left\langle\frac{d\zeta'}{dt}\right\rangle,\ \frac{d\widehat{\theta}'}{dt}\cdot\widehat{\varphi}'\right) = \left(\left\langle\frac{d\zeta}{dt}\right\rangle + \frac{d\psi}{dt},\ \cos\theta\,\frac{d\varphi}{dt} + \frac{d\psi}{dt}\right),$$

which follows from the gyrogauge transformation (7). Upon gyroangle-averaging, the equation of motion (21) for the gyroangle is therefore decomposed in terms of dynamical and geometric terms only.

## III. GEOMETRICAL CONTRIBUTIONS TO THE GYROPHASE

We now follow Ref. [1] and use Eq. (21) to calculate the averaged gyrophase shift, denoted $\langle\Delta\zeta\rangle$, in one gyration period $T \equiv 2\pi/\omega_c(t)$:

$$\langle\Delta\zeta\rangle \equiv \int_t^{t+T}\left\langle\frac{d\zeta}{dt'}\right\rangle dt' = \langle\Delta\zeta_{\rm d}\rangle + \langle\Delta\zeta_{\rm g}\rangle[\mathcal{C}]. \tag{22}$$

The first term on the right side of Eq. (22) denotes dynamical gyrophase shift

$$\langle\Delta\zeta_{\rm d}\rangle \equiv \int_t^{t+T}\omega_{\rm c}(t')\,dt', \tag{23}$$

which equals $2\pi$ for a time-independent magnetic field. The second term on the right side of Eq. (22), on the other hand, denotes the geometrical gyrophase shift

$$\begin{aligned}\langle\Delta\zeta_{\rm g}\rangle[\mathcal{C}] &\equiv \int_t^{t+T}\cos\theta(t')\,\frac{d\varphi}{dt'}\,dt' \\ &= \int_{\mathcal{C}}\cos\theta\,d\varphi = \int_{\mathcal{C}}\mathcal{R}, \end{aligned} \tag{24}$$

where the path $\mathcal{C}$ moves on the unit sphere from the initial point at $\theta(t)$ and $\varphi(t)$, to the final point at $\theta(t+T)$ and $\varphi(t+T)$. We note that the geometrical gyrophase shift (24) is path-dependent (i.e., it is anholonomic) since, by constructing the closed contour $\partial\mathcal{D} = \mathcal{C}_1 - \mathcal{C}_2$ from two paths with identical end points, we find

$$\begin{aligned}\langle\Delta\zeta_{\rm g}\rangle[\mathcal{C}_1] - \langle\Delta\zeta_{\rm g}\rangle[\mathcal{C}_2] &= \oint_{\partial\mathcal{D}}\mathcal{R} = \int_{\mathcal{D}}\mathsf{d}\mathcal{R} \\ &= -\,\Omega \neq 0, \end{aligned} \tag{25}$$

where $\Omega$ denotes the solid angle enclosed by the open surface $\mathcal{D}$ on the unit sphere.

Lastly, we introduce the time-scale ordering on the evolution of the magnetic field

$$\epsilon \equiv T\left|\frac{d\ln B}{dt}\right| \sim T\left|\frac{d\varphi}{dt}\right| \sim T\left|\frac{d\theta}{dt}\right| \ll 1, \tag{26}$$

where the magnitude and direction of the magnetic field are assumed to change on the same slow time scale compared to the gyro-period $T$. By inserting this ordering in Eqs. (23)-(24), we find the dynamical gyrophase shift

$$\langle\Delta\zeta_{\rm d}\rangle \simeq 2\pi + \pi\left(T\,\frac{d\ln B}{dt}\right) \tag{27}$$

and the geometrical gyrophase shift

$$\begin{aligned}\langle\Delta\zeta_{\rm g}\rangle[\mathcal{C}] &\simeq \frac{d\varphi}{dt}\int_0^T\cos[\theta(t+\tau)]\,d\tau \\ &\simeq \cos\theta\left(T\,\frac{d\varphi}{dt}\right), \end{aligned} \tag{28}$$

which is just $T$ times the instantaneous value of the geometric term in Eq. (21), and is of order $\epsilon$. Finally, Liu and Qin compute the average value of the adiabatic term in Eq. (20) and show that it is of order $\epsilon^2$. In our view, this contribution disappears upon gyroaveraging.

## IV. SUMMARY

The formulation of the gyromotion in terms of the spatial angles relates the gyroangle dynamics with the

motion of a spinning rigid body, through a natural appearance of the Euler frequencies. In addition, it emphasizes the role of the gyrogauge one-form $\mathcal{R}$ in the geometric interpretation of the gyromotion anholonomy. Lastly, the inclusion of the pitch-angle dynamics in our formulation shows that it is related to the adiabatic contribution to the gyroangle dynamics and it explains its adiabaticity.

We conclude this Comment with a few remarks concerning a general magnetic field that is space-time-dependent (details will be presented elsewhere). In this case, Eq. (21) is replaced with

$$\left\langle \frac{d\zeta}{dt} \right\rangle = \omega_{\rm c} + \sigma + \left( \tau \, \frac{ds}{dt} + \frac{1}{2} \, \frac{d\chi}{dt} \right), \qquad (29)$$

where $\mathsf{d}\widehat{\mathsf{e}}_1 \cdot \widehat{\mathsf{e}}_2 \equiv \sigma \, \mathsf{d}t + \mathbf{R} \cdot \mathsf{d}\mathbf{x}$ [2]. Here, two additional contributions appear. The first one involves the (Frenet-Serret) torsion of the magnetic-field line $\tau \equiv \widehat{\mathsf{b}} \cdot \mathbf{R} = (\partial \widehat{\mathsf{e}}_1 / \partial s) \cdot \widehat{\mathsf{e}}_2$, with $v_\parallel \equiv ds/dt$ used in Eq. (29). The second one involves the twist of the magnetic-field lines $\tau_{\rm m} = \widehat{\mathsf{b}} \cdot \nabla \times \widehat{\mathsf{b}} \equiv d\chi/ds$, defined as the rate of rotation (denoted by the angle $\chi$) of a nearby field line about the magnetic field line represented by $\widehat{\mathsf{b}}$ [7, 8], with $v_\parallel \, \tau_{\rm m} = d\chi/dt$ used in Eq. (29). We note that the torsion and the magnetic twist can be comparable in some magnetic geometries [9]. The torsion contributes to an anholonomic (path-dependent) geometric gyrophase shift $\langle \Delta\zeta_\tau \rangle [\mathcal{C}] \equiv \int_{\mathcal{C}} \tau \, ds$ while the magnetic twist contributes an holonomic (path-independent) geometric gyrophase shift $\langle \Delta\zeta_\chi \rangle [\mathcal{C}] \equiv \frac{1}{2} \int_{\mathcal{C}} d\chi = \frac{1}{2} \, \Delta\chi$.


[1] J. Liu and H. Qin, Phys. Plasmas **18**, 072505 (2011).
[2] R. G. Littlejohn, Phys. Fluids **24**, 1730 (1981).
[3] R. G. Littlejohn, Phys. Rev. A **38**, 6034 (1988).
[4] R. G. Littlejohn, J. Plasma Phys. **29**, 111 (1983).
[5] H. Flanders, *Differential Forms with Applications to the Physical Sciences* (Dover, New York, 1989).
[6] H. Cartan, *Differential Forms* (Dover, New York, 2006).
[7] H. K. Moffatt and R. L. Ricca, Proc. R. Soc. London A **439**, 411 (1992).
[8] M. A. Berger and C. Prior, J. Phys. A: Math. Gen. **39**, 8321 (2006).
[9] For example, in a screw-pinch cylindrical geometry $(r, \theta, z)$ described by $\widehat{\mathsf{b}}_0 = \cos\Theta \, \widehat{\mathsf{z}} + \sin\Theta \, \widehat{\theta}$, $\widehat{\mathsf{b}}_1 = -\widehat{\mathsf{r}}$, and $\widehat{\mathsf{b}}_2 = -\cos\Theta \, \widehat{\theta} + \sin\Theta \, \widehat{\mathsf{z}}$, where $\Theta \equiv \Theta(r)$, we find the magnetic curvature $\kappa = r^{-1} \sin^2\Theta$, the Frenet-Serret torsion $\tau = r^{-1} \cos\Theta \, \sin\Theta$, and the magnetic torsion (twist) $\tau_{\rm m} = \Theta' + \tau$.

# Bibliography


[1] V. I. Arnold, "Proof of a theorem of A.N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian", Russ. Math. Surv. **18**, 9 (1963). Original paper: Uspehi Mat. Nauk **18**, 13 (1963).

[2] V. I. Arnold, "Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits", Ann. Inst. Fourier **16**, 319 (1966).

[3] R. Bachelard , C. Chandre, and M. Vittot, "Hamiltonian description of a self-consistent interaction between charged particles and electromagnetic waves", Phys. Rev. E **78**, 036407 (2008).

[4] J. Berkowitz and C. S. Gardner, "On the asymptotic series expansion of the motion of a charged particle in slowly-varying fields", Institute of mathematical sciences, Magneto-fluid dynamics division, NYO-7975, New- York University, 1957.

[5] I. Bialynicki-Birula, J. C. Hubbard, and L. A. Turski, "Gauge-independent canonical formulation of relativistic plasma theory", Physica A **128**, 509 (1984).

[6] D. Biskamp, *Nonlinear Magnetohydrodyamics*, Cambridge University Press, 1997.

[7] J. A. Bittencourt, *Fundamentals of Plasma Physics*, Springer, 2004.

[8] N. N. Bogoliubov and D. N. Zubarev, "An asymptotic approximation method for a system with rotating phases and its application to the motion of a charged particle in a magnetic field", Physical Research Laboratory, Space Technology Laboratories, Los Angeles, 1960. Original paper: Ukrain. Matem. Zh., **7**, 1 (1955).

[9] N. N. Bogoliubov and Y. A. Mitropolski, *Asymptotic methods in the theory of non-linear oscillations*, Gordon-Breach, New-York, 1961.

[10] N. N. Bogoliubov, J. A. Mitropolski, and A. M. Samoilenko, *Methods of accelerated convergence in non-linear mechanics*, Springer-Verlag, Berlin, 1976.

[11] M. Born, *The Mechanics of the Atom*, F. Ungar, New York, 1967.

[12] A. H. Boozer, "Physics of magnetically confined plasmas", Rev. Mod. Phys. **76**, 1071 (2005).

[13] F. Bouchet and A. Venaille, "Statistical mechanics of two-dimensional and geophysical flows", Phys. Rep. **515**, 227 (2012).

[14] A. J. Brizard, "Nonlinear gyrofluid description of turbulent magnetized plasmas", Phys. Fluids B **4**, 1213 (1992).

[15] A. J. Brizard, "Nonlinear gyrokinetic Vlasov equation for toroidally rotating axisymmetric tokamaks", Phys. Plasmas **2**, 459 (1995).

[16] A. J. Brizard, "Nonlinear bounce-center Hamiltonian dynamics in general magnetic geometry", Phys Plasmas **7**, 3238 (2000).

[17] A. J. Brizard and T. S. Hahm, "Foundations of nonlinear gyrokinetic theory", Rev. Mod. Phys. **79**, 421 (2007).

[18] A. J. Brizard, "Variational principles for reduced plasma physics", Jour. Phys. Conf. Ser. **169**, 012003 (2009).

[19] A. J. Brizard and A. Mischenko, "Guiding-center recursive Vlasov and Lie-transform methods in plasma physics", J. Plas. Phys. **75**, 675 (2009).

[20] A. J. Brizard, "Exact energy conservation laws for full and truncated nonlinear gyrokinetic equations", Phys. Plasmas **17**, 042303 (2010).

[21] A. J. Brizard and N. Tronko, "Exact momentum conservation laws for the gyrokinetic Vlasov-Poisson equations", Phys. Plasmas **18**, 082307 (2011).

[22] A. J. Brizard and L. de Guillebon, "Comment on "Geometric phase of the gyromotion for charged particles in a time-dependent magnetic field [Phys. Plasmas 18, 072505 (2011)]"", Phys. Plasm. **19**, 094701 (2012).

[23] A. J. Brizard and N. Tronko, "Equivalent representations of higher-order Hamiltonian guiding-center theory", arXiv:1205.5772 [physics.plasm-ph] (2012).

[24] A. J. Brizard, private communication about a non-published document to celebrate the $30^{th}$ anniversary of Littlejohn's derivation of Hamiltonian guiding-center dynamics by Lie-transform perturbation analysis (2009).

[25] J. W. Burby and H. Qin, "Gyrosymmetry: Global considerations", Phys. Plasmas **19**, 052106 (2012).

[26] J. W. Burby, J. Squire, and H. Qin, "Automation of the guiding center expansion", Phys. Plasmas **20**, 072105 (2013).

[27] J. R. Cary, "Lie transformations and their use in Hamiltonian perturbation theory", Phys. Reports **79**, 129 (1981).

[28] J. R. Cary, D. F. Escande, and J. L. Tennyson, "Adiabatic-invariant change due to separatrix crossing", Phys. Rev. A **34**, 4256 (1986).

[29] J. R. Cary and A. J. Brizard, "Hamiltonian theory of guiding-center motion", Rev. Mod. Phys. **81**, 693 (2009).

[30] C. Chandre, M. Vittot, G. Ciraolo, Ph. Ghendrih, and R. Lima, "Control of stochasticity in magnetic field lines", Nucl. Fusion **46**, 33 (2006)

[31] C. Chandre, P. J. Morrison, and E. Tassi, "Hamiltonian derivation of the Charney-Hasegawa-Mima equation", Phys. Plasmas **16**, 082301 (2009).

[32] C. Chandre, P. J. Morrison,and E. Tassi, "Derivation of reduced two-dimensional fluid models via Dirac's theory of constrained Hamiltonian systems", Phys. Plasmas **17**, 042307 (2010).

[33] C. Chandre, P. J. Morrison, and E. Tassi, "On the Hamiltonian formulation of incompressible ideal fluids and magnetohydrodynamics via Dirac's theory of constraints", Phys. Lett. A **376**, 737 (2012).

[34] C. Chandre, L. de Guillebon, A. Back, E. Tassi, and P. J. Morrison, "On the use of projectors for Hamiltonian systems and their relationship with Dirac brackets", J. Phys. A: Math. Theor. **46**, 125203 (2013).

[35] C. Chandre, "Casimir invariants and the Jacobi identity in Dirac's theory of constrained Hamiltonian systems", arXiv:1303.1989 (2013).

[36] F. F. Chen, *Introduction to plasma physics and controlled fusion*, Springer Verlag, 1984.

[37] G. F. Chew, M. L. Goldberger, and F. E. Low, "The Boltzmann equation and the one-fluid hydromagnetic equations in the absence of particle collisions", Proc. R. Soc. London Ser. A **236**, 112 (1956).

[38] G. Ciraolo, C. Chandre, R. Lima, M. Vittot, and M. Pettini, "Control of chaos in Hamiltonian systems", Celest. Mech. Dyn. Astron. **90**, 3 (2004).

[39] G. Ciraolo, F. Briolle, C. Chandre, E. Floriani, R. Lima, M. Vittot, M. Pettini, C. Figarella, and Ph. Ghendrih, "Control of Hamiltonian chaos as a possible tool to control anomalous transport in fusion plasmas", Phys. Rev. E **69**, 056213 (2004).

[40] D. Correa-Restrepo and D. Pfirsch, "Poisson brackets for guiding-centre and gyrocentre theories", J. Plasma Phys. **71**, 1 (2005).

[41] S. C. Cowley, R. M. Kulsrud, and E. Valeo, "A kinetic equation for spin-polarized plasmas", Phys. Fluids **29**, 430 (1986).

[42] T. Dannerta and F. Jenko, "Gyrokinetic simulation of collisionless trapped-electron mode turbulence", Phys. Plasmas **12**, 072309 (2005).

[43] P. A. M. Dirac, "Forms of relativistic dynamics", Rev. Mod. Phys. **21**, 392 (1949).

[44] P.A.M. Dirac, "Generalized hamiltonian dynamics", Can. J. Math. **2**, 129 (1950).

[45] A. J. Dragt and J. M. Finn, "Normal form for mirror machine Hamiltonians", J. Math. Phys. **20**, 2649 (1979).

[46] J. P. Freidberg, "Ideal magnetohydrodynamic theory of magnetic fusion systems", Rev. Mod. Phys. **54**, 801 (1982).

[47] E. Frenod and M. Lutz, "The gyro-kinetic Approximation. An attempt at explaining the method based on Darboux algorithm and Lie transform", arXiv:1111.1510 (2011).

[48] E. Frenod and M. Lutz, "On the geometrical gyro-ginetic theory", arXiv:1306.5639 (2013).

[49] E. A. Frieman and L. Chen, "Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria", Phys. Fluids **25**, 502 (1982).

[50] S. Froehlich and P. Cvitanović, "Reduction of continuous symmetries of chaotic flows by the method of slices", Commun. Nonlinear Sci. Numer. Simul. **17**, 2074 (2012).

[51] X. Garbet, Y. Idomura, L. Villard, and T.H. Watanabe, "Gyrokinetic simulations of turbulent transport", Nucl. Fusion **50**, 043002 (2009).

[52] C. S. Gardner, "Adiabatic invariants of periodic classical systems", Physical Review**115**, 791 (1959).

[53] H. Goldstein, *Classical Mechanics*, Addison-Wesley Reading, Mass., 1980.

[54] R. J. Goldston and P. H. Rutherford, *Introduction to Plasma Physics*, Plasma Physics Series, CRC Press, 2010.

[55] P. Goswami, T. Passot, and P. L. Sulem, "A Landau fluid model for warm collisionless plasmas", Phys. Plasmas **12**, 102109 (2005).

[56] V. Grandgirard, Y. Sarazin, P. Angelino, A. Bottino, N. Crouseilles, G. Darmet, G. Dif-Pradalier, X. Garbet, Ph. Ghendrih, S. Jolliet, G. Latu, E. Sonnendrücker, and L Villard, "Global full-f gyrokinetic simulations of plasma turbulence", Plasma Phys. Control. Fusion **49**, B173 (2007).

[57] J. Guckheimer and P. Holmes, *Nonlinear oscillations, Dynamical systems and Bifurcations of vector fields*, Appl. Math. Sci. **42** (1983), Springer-Verlag, New-York.

[58] L. de Guillebon and C. Chandre, "Hamiltonian structure of reduced fluid models for plasmas obtained from a kinetic description", Phys. Lett. A **376**, 3172 (2012).

[59] L. de Guillebon and M. Vittot, "A gyro-gauge independent minimal guiding-center reduction by Lie transforming the velocity vector field", Phys. Plasmas **20**, 082505 (2013).

[60] L. de Guillebon and M. Vittot, "Gyro-gauge independent formulation of the guiding-center reduction to arbitrary order in the Larmor radius", Plasma Phys. Control. Fusion **55**, 105001 (2013).

[61] L. de Guillebon and M. Vittot, "On an intrinsic formulation of the guiding-center anholonomy and gyro-gauge arbitrariness", submitted, arXiv:1304.4563 (2013).

[62] L. de Guillebon, N. Tronko, M. Vittot, and Ph. Ghendrih, "Dynamical reduction for charged particles in a strong magnetic field without guiding-center", in preparation.

[63] Ö. D. Gürcan, P. H. Diamond, T. S. Hahm, and R. Singh, "Intrinsic rotation and electric field shear", Phys. Plasmas **14**, 042306 (2007).

[64] T. S. Hahm, W. W. Lee, and A. J. Brizard, "Nonlinear gyrokinetic theory for finite-beta plasmas", Phys. Fluids **31**, 1940 (1988).

[65] T. S. Hahm, "Nonlinear gyrokinetic equations for tokamak microturbulence", Phys. Fluids **31**, 2670 (1988).

[66] A. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Systems*, Accademia Nazionale dei Lincei, Roma, 1976.

[67] R. J. Hastie, J. B. Taylor, and F. A. Haas, "Adiabatic invariants and the equilibrium of magnetically trapped particles", Ann. Phys. **41**, 302 (1967).

[68] J. Henrard, *The adiabatic invariant in classical dynamics*, Springer, Berlin, 1993.

[69] D. D. Holm and B. A. Kupperschmidt, "Relativistic fluid dynamics as a Hamiltonian system" Phys. Lett. **101A**, 23 (1984).

[70] D. D. Holm, J. E. Marsden, T. S. Ratiu, and A. Weinstein, "Nonlinear stability of fluid and plasma equilibria", Phys. Rep. **123**, 1 (1985).

[71] D. D. Holm, "Hamiltonian formalism for general-relativistic adiabatic fluids", Physica D **17**, 1 (1985).

[72] D. D. Holm and B. A. Kupperschmidt, "Hamiltonian theory of relativistic magnetohydrodynamics with anisotropic pressure", Phys. Fluids **29**, 3889 (1986).

[73] D. D. Holm, J. E. Marsden, T. S. Ratiu, "The Euler–Poincaré Equations and Semidirect Products with Applications to Continuum Theories", Adv. Math. **137**, 1 (1998).

[74] J. E. Howard, "Nonadiabatic particle motion in cusped magnetic fields", Phys. Fluids **14**, 2378 (1971).

[75] O. Izacard, C. Chandre, E. Tassi, and G. Ciraolo, "Gyromap for a two-dimensional Hamiltonian fluid model derived from Braginskii's closure for magnetized plasmas", Phys. Plasmas **18**, 062105 (2011).

[76] O. Izacard, N. Tronko, C. Chandre, G. Ciraolo, M. Vittot, and Ph. Ghendrih, "Transport barrier for the radial diffusion due to the E x B drift motion of guiding centers in cylindrical confinement geometry", Plasma Phys. Control. Fusion **53**, 125008 (2011).

[77] J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, New York, 1962.

[78] A. N. Kolmogorov, "On conservation of conditionally periodic motions for a small change in Hamilton's function", Los Alamos Scientific Laboratory translation from Dokl. Akad. Nauk SSSR (N.S.) **98**, 527 (1954).

[79] J. A. Krommes, "Comment on "Guiding-center plasma models in three dimensions"", Phys. Plasmas **16**, 084701 (2009).

[80] J. A. Krommes, "The Gyrokinetic Description of Microturbulence in Magnetized Plasmas", Ann. Rev. of Fluid Mech. **44**, 175 (2012).

[81] M. Kruskal, "Asymptotic theory of Hamiltonian and other systems with all solutions nearly periodic", J. Math. Phys. **3**, 806 (1962).

[82] N. M. Krylov and N. N. Bogoliubov, *Introduction to non-linear mechanics*, Ann. of Math. Stud. **11** (1949), Princeton University Press.

[83] R. M. Kulsrud, E. J. Valeo, and S. C. Cowley, "Physics of spin-polarized plasmas", Nucl. Fusion **26**, 1443 (1986).

[84] S. Lang, *Differential and Riemannian manifolds*, Springer-Verlag, 1995.

[85] R. G. Littlejohn, "A guiding-center Hamiltonian: a new approach", J. Math. Phys. **20**, 2445 (1979).

[86] R. G. Littlejohn, "Hamiltonian formulation of guiding-center motion", Phys. Fluids **24**, 1730 (1981).

[87] R. G. Littlejohn, "Hamiltonian perturbation theory in noncanonical coordinates", J. Math. Phys. **23**, 742 (1982).

[88] R. G. Littlejohn, "Variational principles of guiding-center motion", J. Plasm. Phys. **29**, 111 (1983).

[89] R. G. Littlejohn, "Linear relativistic gyrokinetic equation", Phys. Fluids **27**, 976 (1984).

[90] R. G. Littlejohn, "Phase anholonomy in the classical adiabatic motion of charged particles", Phys. Rev. A **38**, 6034 (1988).

[91] J. Liu and H. Qin, "Geometric phase of the gyromotion for charged particles in a time-dependent magnetic", field", Phys. Plasmas **18**, 072505 (2011).

[92] R. de la Llave, *Introduction to KAM theory*, in Computational Physics, World Scientic, River Edge, NJ, 1992.

[93] M. Marklund and P. J. Morrison, "Gauge-free Hamiltonian structure of the spin Maxwell–Vlasov equations", Phys. Lett. A **375**, 23622365 (2011).

[94] J. E. Marsden and A. Weinstein, "Reduction of symplectic manifolds with symmetry", Rep. Math. Phys. **5**, 121 (1974).

[95] J. E. Marsden and A. Weinstein, "The Hamiltonian structure of the Maxwell-Vlasov equations", Physica D **4**, 394 (1982).

[96] J. E. Marsden, P. J. Morrison, and A. Weinstein, "The Hamiltonian structure of the BBGKY equations", Contemp. Math. **28**, 115 (1984).

[97] J. E. Marsden, R. Montgomery, and T. S. Ratiu, "Reduction, symmetry and phases in mechanics", Memoirs of the American Mathematical Society, **88**, 436 (1990).

[98] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, Springer-Verlag, Berlin, 2002.

[99] J. E. Marsden, G. Misiolek, J.-P. Ortega, M. Perlmutter, and T. S. Ratiu, *Hamiltonian reduction by stages*, Lecture Notes in Mathematics **1913** (2007).

[100] M. Materassi and E. Tassi, "Metriplectic framework for dissipative magneto-hydrodynamics", Physica D **241**, 729 (2012).

[101] P. J. Morrison, "The Maxwell-Vlasov equations as a continuous Hamiltonian system", Phys. Lett. **80A**, 383 (1980).

[102] P. J. Morrison and J. M. Greene, "Noncanonical Hamiltonian Density Formulation of Hydrodynamics and Ideal Magnetohydrodynamics", Phys. Rev. Lett. **45**, 790 (1980); **48**, 569 (1982).

[103] P. J. Morrison and J. M. Greene, "Errata for "Noncanonical Hamiltonian Density Formulation of Hydrodynamics and Ideal Magnetohydrodynamics"" Phys. Rev. Lett. **48**, 569 (1982).

[104] P. J. Morrison, "Poisson brackets for fluids and plasmas", AIP Conference Proceedings **88**, 13 (1982).

[105] P. J. Morrison, "Bracket formulation for irreversible classical fields", Phys. Lett. A **100**, 423 (1984).

[106] P. J. Morrison, Physica D: Nonlinear Phenomena "A paradigm for joined Hamiltonian and dissipative systems", Physica D **18**, 410 (1986).

[107] P. J. Morrison, "Hamiltonian description of the ideal fluid", Rev. Mod. Phys. **70**, 467 (1998).

[108] P. J. Morrison, "Hamiltonian and action principle formulations of plasma physics", Phys. Plasmas **12**, 058102 (2005).

[109] P. J. Morrison, *Hamiltonian fluid mechanics*, in Encyclopedia of Mathematical Physics (Elsevier, Amsterdam), **2**, 593 (2006).

[110] P. J. Morrison, "On Hamiltonian and action principle formulations of plasma dynamics", in AIP Conf. Proc. **1188**, 329 (2009).

[111] P. J. Morrison, "Thoughts on brackets and dissipation: old and new", J. Phys. Conf. Ser. **169**, 012006 (2009).

[112] P. J. Morrison, N. Lebovitz, and J. Biello, "The Hamiltonian description of incompressible fluid ellipsoids", Annals of Physics **324**, 1747(2009).

[113] P. J. Morrison, "On the Hamilton-Jacobi variational formulation of the Vlasov equation", Math-for-Industry Lecture Series **39**, 64 (2012).

[114] P. J. Morrison, "A general theory for gauge-free lifting", Phys. Plasmas **20**, 012104 (2013).

[115] P. J. Morrison, M. Vittot, and L. de Guillebon, "Magnetic moment type of lifting from particle dynamics to Vlasov-Maxwell dynamics", Phys. Plasmas **20**, 032109 (2013).

[116] P. J. Morrison, E. Tassi, and N. Tronko, "Stability of compressible reduced magnetohydrodynamic equilibria – Analogy with magnetorotational instability", Phys. Plasmas **20**, 042109 (2013).

[117] A. I. Neishtadt and A. A. Vasiliev, "Destruction of adiabatic invariance at resonances in slow–fast Hamiltonian systems", Nucl. Instr. Meth. Phys.Research A **561**, 2, 158 (2006).

[118] S. Nguyen and L. A. Turski, "Canonical description of incompressible fluid: Dirac brackets approach", Physica A **272**, 48 (1999).

[119] S. Nguyen and L. A. Turski, "Examples of the Dirac approach to dynamics of systems with constraints", Physica A **290**, 431 (2001).

[120] D. R. Nicholson, *Introduction to Plasma Theory*, Wiley, New York, 1983.

[121] T. G. Northrop, "The guiding center approximation to charged particle motion", Ann. Phys. **1**, 79 (1961).

[122] T. G. Northrop, *The adiabatic motion of charged particles*, Wiley, New York, 1963.

[123] T. G. Northrop and J. A. Rome, "Extensions of guiding center motion to higher order", Physics of Fluids. **21**, 384 (1978).

[124] J.-P. Ortega and T. S. Ratiu, "Momentum maps and Hamiltonian reduction", Progr. Math. **222**, Birkäuser, Boston (2004).

[125] N. Padhye and P. J. Morrison, "Fluid element relabeling symmetry", Phys. Lett. A **219**, 287 (1996).

[126] N. Padhye and P. J. Morrison, "Relabeling Symmetries in Hydrodynamics and Magnetohydrodynamics", Plasma Phys. Rep. **22**, 869 (1996).

[127] F. I. Parra and P. J. Catto, "Transport of momentum in full f gyrokinetics", Phys. Plasmas **17**, 056106 (2010).

[128] F. I. Parra and I. Calvo, "Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry", Plasma Phys. Control. Fusion **53**, 045001 (2011).

[129] D. Pfirsch, "New variational formulation of Maxwell-Vlasov and guiding center theories local charge and energy conservation laws", Zeitschrift fuer Naturforschung, **39a**, 1 (1984).

[130] D. Pfirsch and P. J. Morrison, "Local conservation laws for the Maxwell-Vlasov and collisionless kinetic guiding-center theories", Phys. Rev. A **32**, 1714 (1985).

[131] H. Qin and W. M. Tang, "Pullback transformations in gyrokinetic theory ", Phys. Plasmas **11**, 1052 (2004).

[132] A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, E. Quataert, and T.Tatsno, "Astrophysical gyrokinetics: Kinetic and fluid turbulent cascades in magnetized", Astrophys. J. Suppl. Ser. **182**, 310 (2009).

[133] B. Scott and J. Smirnov, "Energetic consistency and momentum conservation in the gyrokinetic description of tokamak plasmas", Phys. Plasmas **17**, 112302 (2010).

[134] B. A. Shadwick, J. C. Bowman, and P.J. Morrison, "Exactly conservative integrators", SIAM J. Appl. Math. **59**, 1112 (1999).

[135] B. A. Shadwick and J. S. Wurtele, "General moment model of beam transport", in: Proceedings of the 1999 Particle Accelerator Conference, vol. 4, 2888 (1999).

[136] B. A.Shadwick, G. M. Tarkenton, and E. H. Esarey, "Hamiltonian description of low-temperature relativistic plasmas", Phys. Rev. Lett. **93**, 175002 (2004).

[137] B. A. Shadwick, G. M. Tarkenton, E. H. Esarey, and C. B. Schroeder, "Fluid and Vlasov models of low-temperature, collisionless, relativistic plasma interactions", Phys. Plasmas **12**, 056710 (2005).

[138] B. A.Shadwick, G. M. Tarkenton, E. H. Esarey, and F. M. Lee "Hamiltonian reductions for modeling relativistic laser-plasma interactions", Commun. Nonlinear Sci. Numer. Simul. **17**, 2153 (2012).

[139] T. G. Shepherd, "Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics", Adv. Geophys. **32**, 287 (1990).

[140] T. G. Shepherd, "A unified theory of available potential energy", Atm.-Ocean **31**, 1 (1993).

[141] T. G. Shepherd, *Hamiltonian dynamics* in Encyclopedia of Atmospheric Sciences, J. R. Holton et al. Eds., Academic Press, 2003.

[142] C. L. Siegel and J. K. Moser, *Lectures on Celestial Mechanics*, Springer Verlag, Berlin, 1971.

[143] R. G. Spencer and A. N. Kaufman, "Hamiltonian structure of two-fluid plasma dynamics", Phys. Rev. A **25**, 2437 (1982).

[144] J. Squire, H. Qin, W. M. Tang, and C. Chandre, "The Hamiltonian structure and Euler-Poincaré formulation of the Vlasov-Maxwell and gyrokinetic systems", Phys. Plasmas **20**, 022501 (2013).

[145] L. E. Sugiyama, "Guiding-center plasma models in three dimensions", Phys. Plasmas **15**, 092112 (2008).

[146] L. E. Sugiyama, "Answer to "Comment on 'Guiding-center plasma models in three dimensions' " ", Phys. Plasmas **16**, 084702 (2009).

[147] P. L. Sulem and T. Passot, "FLR–Landau fluids for collisionless plasmas", Commun. Nonlinear Sci. Numer. Simul. **13**, 189 (2008).

[148] E. Tassi, P. J. Morrison, F. L. Waelbroeck, and D. Grasso, "Hamiltonian formulation and analysis of a collisionless fluid reconnection model", Plasma Phys. Control. Fusion **50**, 085014 (2008).

[149] E. Tassi and P. J. Morrison, "Mode signature and stability for a Hamiltonian model of electron temperature gradient turbulence", Phys. Plasmas **18**, 032115 (2011).

[150] J. L. Tennyson, J. R. Cary, and D. F. Escande, "Change of the adiabatic invariant due to separatrix crossing", Phys. Rev. Lett. **56**, 2117 (1986).

[151] N. Tronko, M. Vittot, C. Chandre, Ph. Ghendrih, and G. Ciraolo, "Barriers for the reduction of transport due to the E x B drift in magnetized plasmas", J. Phys. A **42**, 085501 (2009).

[152] R. Utiyama, "Theory of invariant variation and the generalized canonical dynamics", Progr. Theor. Phys. Suppl. **9**, 19 (1959).

[153] M. Vittot, C. Chandre, G. Ciraolo, and R. Lima, "Localized control for non-resonant Hamiltonian systems", Nonlinearity **18**, 423 (2005).

[154] F. L. Waelbroeck, R. D. Hazeltine, and P. J. Morrison, "A Hamiltonian electromagnetic gyrofluid model", Phys. Plasmas **16**, 032109 (2009).

[155] F. L. Waelbroeck and E. Tassi, "A compressible Hamiltonian electromagnetic gyrofluid model", Communications in Nonlinear Science and Numerical Simulation, **17**, 2171 (2012).

[156] A. Weinstein and P. J. Morrison, "Comments on "The Maxwell-Vlasov equations as a continuous Hamiltonian system"", Phys. Lett. A **86**, 235 (1981).

[157] B. Weyssow and R. Balescu, "Hamiltonian theory of guiding centre motion revisited", J. Plasma Phys. **35**, 449 (1986).

[158] H. Ye and P. J. Morrison, "Action Principles for the Vlasov Equation", Phys. Fluids B **4**, 771 (1992).

[159] V. E. Zakharov and E. A. Kuznetsov, "Hamiltonian formalism for nonlinear waves", Phys. Usp. **40**, 1087 (1997).

# Index


Adiabatic invariant, **23**, 43, 56, 227, **232**, 232–234
Anholonomy, 17, 24, 98, **102**, 102–106, 108, 111
Arbitrariness, 24, 29, 40, 45, 54, 86, 88, 96, 97, 99–102, 107, 111, 242

Closure assumption, 13, **208**, *see also* Fluid model
Connection, 48, 51, 79, 86, 96, **100**, 99–112
Constrained system, 25, 28, 33, 48, 51, 60, 88, **99**, 95–112, 125, *see also* Dirac constraint
Covariant derivative, 51, 80, 88, 96, **100**, 99–112

Dirac constraints, 18, 57, **179**, 173–205, 217, 265–292

Electric field, **16**, 29, 41, 49, 56, 81, 114, 119, 133, 135, 157, 204, 227, 268, 269, 282, 297, 301
Electrostatic, 268, 273, 282, 286, 287, *see also* Vlasov-Poisson
Euler-Maxwell, 13, *see also* Fluid model

Fiber bundle, 48, 51, 85, 97, **99**, 100, 105, 107, 110, 112
Field dynamics, 52
Fluid model, 13, 18, **188**, 173–219, 222, 265–302, *see also* Magnetohydrodynamics (MHD)

Guiding center, 15, 16, **23**, **55**, 27–94, 113, 119, 137, 144, 221, 267
Gyro-angle, 15, 19, 23, 25, **30**, **50**, 27–113, 168, 173, 197, 217, 221, 227, 230, 242, 244
Gyro-center, 15, 16, 147, 164, *see also* Guiding-center
Gyro-gauge, 17, 18, 24, 25, 27, 33, 45, 48, **50**, 87, 95, **97**, 99, 111, 113, 221, 227, 242
Gyrokinetics, **14**, 15, 119, 144, 150, 156, 167, 168, 173, 196, 217, 221

H-secondary constraints, *see* J-secondary constraints
Hamiltonian, 12, **52**, 145

Incompressible, 18, 174, 176, 177, **182**, 182–183, 186, **189**, 188–195, 273–275, 277, 284–285, 288, 289
Induction relations, 24, 32, 48, 77, 93, 231
Intrinsic approach, 24, *see also* Gyro-gauge

J-secondary constraints, 196, 275, **282**, 282–285, *see also* Secondary constraint
Jacobi identity, 120, 143, 174–186, 212–215, **254**, 267, 271

Kinetic, 13, *see also* Vlasov-Maxwell

Lagrangian 1-form, 23, 25, 40, 44, 48, **52**, 86, 113
Larmor, 14, **15**, 23, 30, 31, 34, 38, **50**, 62, 228, 299
Lie-transform, 23, 25, 28, **31**, **53**, 113, 131, 149
Lifting, 15, 16, 18, 24, **124**, 119–169, 222, 251–264

Magnetic moment, 15, 16, **23**, 25, 43, **56**, 66, 71, 119, 120, 123, 132, 158–164, 168, 173, 197, 227–249, 268
Magnetization, 15, 122, 134, 138, 141, **142**, 142, 144, 149, 153, 263
Magnetohydrodynamics (MHD), 13, 18, 174, **181**, 181–183, 217, 252, 268, 293–302
Maximal reduction, 25, **56**, **83**, 47–94
Minimal transformation, 25, **33**, 27–47, 55, 57, 86

Perturbation expansion, 15, 23, 31, 81, 114, 163, 234, 238, 300
Poisson bracket, 52
Polarization, **142**, 144, *see also* Magnetization
Projector, 100, 102, 174–195, 217, 265–278

Quarter-canonical, 57, 174, **191**, 188–205, 217, 222, 265–292

Secondary constraint, 184, **189**, 205, 268, 271, 285, 288, 289, *see also* J-secondary constraints
Secular equation, **44**, 48, 87, 230
Subalgebra, 18, 57, 174, 188, 205, 208, **210**, 214, 279, 285–292, 294, 297, 301, 302

Truncation, 18, 44, 57, 167, 168, 173, 174, **190**, 189–196, 205, 207–215, 217, 265–292

Vlasov-Maxwell, 13, 16, **125**, 119–169, 183, 184, 194, 196, 208, 209
Vlasov-Poisson, 16, 18, 119, 144, 147, 148, 156, 174, 185–186, **203**, 217, 268, 273, 281, 282

# List of main symbols

In principle, we have tried to make each chapter self-consistent. In each of them the various symbols involved are defined. This was all the more adequate as different chapters often have quite different frameworks and subjects. We summarize here the main symbols used in the document, i.e. the ones used for at least a few pages.

Perhaps it is useful to mention that the various chapters were initially written in different contexts, and with different notational conventions. We have tried to homogenize them. We hope to have avoided omissions and mistakes, but if some were to remain in the present version, we would be happy to correct them in the next versions. Your remarks and suggestions are welcome.

## Notational conventions

As is rather standard, bold-faced characters (e.g. $\mathbf{B}$) usually denote vectorial quantities, while the associated scalar quantities are indicated by normal-font characters (e.g. $B$). Sans-serif fonts or hatted vectors for lower-case characters (e.g. $\mathsf{b}$ or $\hat{\mathbf{v}}$) ordinarily mean unit vectors.

As usual also, exponents indicate the power (e.g. $\mathbf{B}^2$) or the component of a vector or of a differential form (e.g. $\mathbf{X}^i$ or $\Gamma^i$). A tilde denotes gyro-fluctuation or related quantities, e.g. $\widetilde{\omega}_0$ on page 59.

Indices have three possible meanings: they can refer to the component of a vector or of a differential form (e.g. $\partial_i$), or to the expansion order (e.g. $\mathbf{G}_1$), or to a derivative. Be careful that three distinct kinds of derivatives (i.e. first-order variations) appear, namely for functions (e.g. $f_{\mathbf{z}} = \frac{\partial f}{\partial \mathbf{z}}$, with $f$ a function with argument $\mathbf{z}$), for functionals (e.g. $H_{\mathbf{B}} = \frac{\delta H}{\delta \mathbf{B}}$, with $H$ a Hamiltonian functional with argument $\mathbf{B}$), and for operator-functions. This last case is exemplified by the magnetic moment $\bar{\mu}(\mathbf{p}, \mathbf{q}) = \bar{\mu}[\mathbf{p}, \mathbf{B}(\mathbf{q})]$, which depends on all the derivatives of the magnetic field; it can be viewed as an operator acting on the magnetic field, and $\bar{\mu}_{\mathbf{B}}$ is an operator (see page 133 for a detailed example). The indices "parallel" and "perpendicular" fairly often refer to the magnetic field (e.g. $v_{\parallel}$ and $\mathbf{v}_{\perp}$).

Overbars can have two meanings. Most often, they indicate a transformed quantity (e.g. $\bar{\mathbf{q}}$ or $\bar{\mu}$). However, when they appear over $\mathsf{c}$ or $\mathsf{a}$ followed by another vector, they indicate the matrix transpose for vectors. The only possible ambiguity concerns the vector coordinate $\mathsf{c}$, but the meaning of the overbar should be clear by the context, especially because the overbar to transpose $\mathsf{c}$ is used when it is contracted with another vector. In addition, only Chapter 2 is concerned with both uses of the overbar over the symbol $\mathsf{c}$, and for safety, we have distinguished the long, thick overline ($\overline{\mathsf{c}}$) for the transformed value of $\mathsf{c}$ from the shorter, thin overbar ($\bar{\mathsf{c}}$) for the transpose of $\mathsf{c}$.

About the overbars, Chapter 10 uses a specific notational convention: overbars always indicate matrix transpose, and they can be used over any vector (not just on $\mathsf{c}$ and $\mathsf{a}$). Indeed, in this chapter, there is no transformed quantity, except the constant of motion, which is denoted by $A$.

The primed notation mainly indicates a twin object, as is standard (e.g. in $\delta(\mathbf{x}-\mathbf{x}')$). However, in order to make some formulae easier to read (see e.g. Eq. (1.23)), we also have used a primed quantity followed by a vector to indicate a gradient acting on its left: $f'\mathbf{B} = \mathbf{B} \cdot \nabla f$. In principle there should be no confusion, since the meaning is clear by the context.

## List of main symbols

*Note: the number appearing just after each symbol indicates a page in the manuscript where a typical use (or better a definition whenever suited) of the symbol appears.*

$\mathbf{A}$ : 52 magnetic vector potential

$\mathbb{A}$ : 190 matrix, esp. truncated matrix from $\mathbb{J}$

$\mathcal{A}$ : 160 magnetic-moment-related functional

$A$ : 229 constant of motion (Chap. 10 and 4)

$a$ : 273 generic coefficient (possibly operator)

$\mathsf{a}$ : 30 unit vector of the Larmor radius

$\mathsf{avg}$ : 31 gyro-average operator

$\mathbf{B}$ : 29 magnetic field vector

$\mathbf{B}^{\perp}$ : 30 plane perpendicular to $\mathbf{B}$

$\mathbf{B}_*$ : 80 modified magnetic field vector

$\mathbb{B}$ : 190 generic matrix, possibly operator

$\mathcal{B}$ : 244 matrix operator $\mathsf{b}\times$

$B$ : 30 norm of the magnetic field

$\mathsf{b}$ : 29 unit vector of $\mathbf{B}$ (or rather of $e\mathbf{B}$)

$b$ : 273 generic coefficient (possibly operator)

$\mathbb{C}$ : 179 operator for matrix of local constraints

$C$ : 179 matrix of Dirac constraints

$\mathsf{c}$ : 29 unit vector of the perpendicular velocity

$c$ : 269 generic coefficient (possibly operator)

$\mathcal{D}$ : 184 generic operator
230 esp. intrinsic $\partial_\theta$

$d$ : 68 exterior derivative on differential forms
61 esp. total derivative on functions
273 or generic coefficient (possibly operator)

$d_{\mathbf{z}}$ : 261 total derivative with respect to $\mathbf{z}$

$\mathbf{E}$ : 124 electric field (vector)

$E$ : 81 electric field (norm)

$e$ : 29 particle charge
53 or Euler's number for exponential

$\mathsf{e}_1$ : 30 unit vector for the gyro-gauge

$\mathsf{e}_2$ : 30 unit vector for $\mathsf{b}\times\mathsf{e}_1$

$F$ : 143 generic functional

$f$ : 52 generic function
124 esp. Vlasov density

$\mathbf{G}$ : 54 generic vector field
53 esp. generator of transformations

$G$ : 143 generic functional

$\mathcal{G}$ : 230 left-inverse of $\mathcal{D}$

$\mathbf{g}$ : 73 vector combining several quantities

$g$ : 52 generic function
273 or generic coefficient (possibly operator)

$H$ : 143 Hamiltonian function
181 or generic functional
52 esp. Hamiltonian functional

$h$ : 52 generic function
229 esp. function related to the Hamiltonian
273 or generic coefficient (possibly operator)

$\bar{\bar{I}}$ : 125 identity operator

$\bar{\bar{I}}_{\perp}$ : 125 orthogonal projector onto $\mathbf{B}^{\perp}$

$\bar{\bar{\mathcal{I}}}_{\perp}$ : 128 projector perpendicular to the velocity

$i_{\mathbf{X}}$ : 54 interior product with the vector field $\mathbf{X}$

$\mathbf{J}$ : 183 Electric current

$\mathbb{J}$ : 53 Poisson matrix in finite dimension
179 or matrix of local field interactions

$\mathbb{J}_*$ : 179 the specific $\mathbb{J}$ for Dirac brackets

$\mathcal{J}$ : 128 Jacobian

$\mathcal{K}$ : 65 function to be fixed

$\mathsf{L}$ : 53 shorthand about a Lie derivative

$\mathcal{L}$ : 53 Lie derivative

$\mathbf{M}$ : 210 fluid momentum density

$\mathsf{M}$ : 77 matrix
89 esp. pivotal matrix for Chapt. 2

$\mathsf{M}_\infty$: 77 pivotal matrix for higher orders

$m$ : 29 particle mass

$\mathbf{N}$ : 101 vector field for $\nabla \times \mathbf{R}$
32 or remaining terms in an equation

$n_b$ : 90 order from which $\mathsf{M}_n$ is invertible

$n_c$ : 90 order from which $\mathsf{M}_\infty$ is active

$\mathtt{osc}$ : 31 gyro-fluctuation operator

$P$ : 299 pressure

$\mathcal{P}$ : 178 generic projector

$\mathcal{P}_\perp$ : 178 orthogonal projector

$\mathcal{P}_*$ : 179 Dirac projector

$P_n$ : 209 n-th moment of the Vlasov density

$\mathbf{p}$ : 29 particle momentum (vector)

$\mathbf{p}_\perp$ : 29 perpendicular momentum (perp. to $\mathbf{B}$)

$p$ : 29 particle momentum (scalar)

$p_\parallel$ : 84 parallel momentum

$\mathbb{Q}$ : 178 derivative of $\mathcal{Q}$

$\mathcal{Q}$ : 178 constraint

$\mathbf{q}$ : 29 particle position

$q$ : 294 particle charge

$\mathbf{R}$ : 99 gauge vector

$\mathbf{R}_g$ : 100 generalized gauge vector

$\mathbf{R}_\varphi$ : 108 (scalar) connection coefficient for $\varphi$

$\mathsf{R}$ : 74 remaining terms in an equation

$\mathrm{Rg}$ : 178 range of an operator

$r_L$ : 31 Larmor radius

$S$ : 54 gauge function for the Lagrangian

$s$ : 181 fluid specific entropy

$\mathbb{T}$ : 253 transformation for functionals

$\mathsf{T}$ : 253 transformation for functions (pull-back)

$U$ : 181 internal energy density

$\mathbf{u}$ : 188 fluid velocity

$\mathbf{V}$ : 177 generic vector
107 or specific vectors in several contexts

$V$ : 128 particle velocity (scalar)

$\mathbf{v}$ : 181 particle velocity (vector)

$\mathbf{v}_\perp$ : 125 perpendicular velocity (vectorial)

$\hat{\mathbf{v}}$ : 128 unit vector of $\mathbf{v}$

$\mathsf{v}$ : 129 other symbol for $\hat{\mathbf{v}}$ used for clarity

$\mathsf{v}$ : 235 specific vector parallel to $\mathbf{v}$

$v$ : 273 norm of the particle velocity
183 or generic coefficient (possibly operator)

$v_\perp$ : 42 perpendicular velocity (scalar)

$v_\parallel$ : 125 parallel velocity

$\mathbf{W}$ : 177 generic vector or vector field

$\mathbf{w}$ : 34 generic vector

$\mathbf{X}$ : 54 generic vector field
53 esp. generator of transformations

$\mathbf{x}$ : 294 position in configuration space
179 or generic coordinate

$\mathbf{y}$ : 52 phase-space-time coordinate $(\mathbf{p}, \mathbf{q}, t)$
261 or generic coordinate

$\mathbf{z}$ : 138 phase-space coordinate $(\mathbf{p}, \mathbf{q})$
52 or generic coordinate

$\alpha$ : 231 gyro-average of $A(\mathbf{p}, \mathbf{q})$
294 or charge-to-mass ratio

$\Gamma$ : 52 Lagrangian 1-form

$\gamma$ : 68 generic 1-form
102 or generic path (or loop)

$\Delta$ : 182 Laplace operator
102 or finite variation

$\delta$ : 105 Kronecker delta symbol
177 or Dirac delta distribution
126 or infinitesimal variation
60 esp. for a 1-form (non-closed)
127 or for functional differentiation

$\epsilon$ : 299 generic small parameter
31 esp. the one for guiding-center theory

$\varepsilon$ : 105 Levi-Cività symbol

$\eta$ : 231 gyro-fluctuation of $A(\mathbf{p}, \mathbf{q})$

$\Theta$ : 103 angle for analogy with $\theta$

$\theta$ : 30 gyro-angle

$\Lambda$ : 69 specific 1-form

$\lambda$ : 68 specific 1-form

$\mu$ : 56 adiabatic invariant (lowest-order) for $\bar{\mu}$

$\bar{\mu}$ : 56 magnetic moment

$\Pi_n$ : 213 pressure-like moment of $f$

$\pi_\gamma^\alpha$ : 152 map between coordinate spaces

$\rho$ : 181 fluid mass or charge density

$\rho_e$ : 295 charge density

$\sigma$ : 294 fluid entropy density

$\tau$ : 253 transformation for coordinates

$\Phi$ : 81 electric potential
266 or semi-canonical coordinate

$\phi$ : 58 cotangent of $\varphi$
190 or constrained field

$\varphi$ : 29 pitch-angle

$\boldsymbol{\chi}$ : 176 generic field
190 or coordinate complementary to $\phi$

$\boldsymbol{\psi}$ : 252 generic field

$\omega$ : 55 Lagrange 2-form
176 or vorticity

$\omega_L$ : 30 Larmor frequency

$\omega_s$ : 53 symplectic (phase-space) part of $\omega$

$\dagger$ : 182 adjoint

$\partial_{\mathrm{a}}$ : 234 specific differential operator

$\partial_a$ : 63 shorthand for $\mathsf{a} \cdot \nabla$

$\partial_b$ : 63 shorthand for $\mathsf{b} \cdot \nabla$

$\partial_c$ : 63 shorthand for $\mathsf{c} \cdot \nabla$

$\partial_\theta$ : 63 shorthand for $-\mathsf{a} \cdot \partial_{\mathsf{c}}$

$\nabla_*$ : 80 reference choice for the covariant derivative
130 or gradient expressed in other coordinates

$\nabla_\star$ : 247 specific gradient-like derivative operator

$\{\cdot, \cdot\}$ : 52 Poisson bracket

$\{\cdot, \cdot\}_*$ : 179 Dirac Poisson bracket

$[\cdot, \cdot]$ : 124 Poisson bracket (esp. for particles)
108 or commutator
131 or similar to $[\cdot]$ but with several dependences

$[\cdot]$ : 254 functional dependence
131 esp. dependence through operators