HAL open science

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

Loïc de Guillebon, Michel Vittot

To cite this version:

Loïc de Guillebon, Michel Vittot. Gyro-gauge-independent formulation of the guiding-center reduction to arbitrary order in the Larmor radius. Plasma Physics and Controlled Fusion, 2013, 55, pp.105001. hal-00812829v2

HAL Id: hal-00812829

https://hal.science/hal-00812829v2

Submitted on 5 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.

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

L. de Guillebon* and M. Vittot

Centre de Physique Théorique, Aix-Marseille Université, CNRS, UMR 7332, 13288 Marseille, France

Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France

Abstract: 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. It is inverted and explicit induction relations are obtained to go 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.

Keywords: guiding center, Lie transform, Lagrangian 1-form, maximal reduction, explicit induction, gauge-independent coordinates, gyro-gauge, connection, covariant derivative, principal circle bundle.

*Electronic address: `de-guillebon@cpt.univ-mrs.fr`

I. INTRODUCTION

The dynamics of a charged particle in a strong magnetic field shows up a separation of time-scales, with the existence of a fast Larmor rotation, slower drifts, and an adiabatic invariant. This can be used to isolate completely the slow part of the dynamics from the single fast coordinate, the gyro-angle, and to build a true constant of motion, the magnetic moment [1, 2]. Thus the guiding-center reduction brings a slow reduced motion and decreases by two the effective dimension of the dynamics. It is the starting point of gyrokinetics, which is a kinetic description of plasma dynamics in a strong magnetic field, and a key model in the study of plasma micro-turbulence [3]. The most efficient results for the guiding-center reduction are obtained by Lie transforming the Lagrangian 1-form in an expansion in a small parameter related with the magnetic inhomogeneity at the scale of the Larmor radius [2, 4–6].

In this paper, we clarify some aspects of guiding-center theory, especially some difficulties about the definition of the gyro-angle, by showing that the reduction can be done at arbitrary order in the Larmor radius using gyro-gauge-independent coordinates.

Indeed, the introduction of a scalar coordinate for the angle measuring the Larmor rotation, namely the *gyro-angle*, imposes to choose at each point in the configuration space a direction, an axis, in the plane perpendicular to the magnetic field, defined as the zero from which the angle is measured. This corresponds to fixing a gauge in the theory, the so-called gyro-gauge, and the presence of this gauge in the theory raised some non-trivial questions [7–11], even the bare existence of a global choice of gauge can fail. So, it is interesting to consider the reduction from a more intrinsic point of view.

In a previous work [12], we proposed a guiding-center reduction which avoided to introduce a gyro-gauge. The idea was to Lie transform directly the equations of motion instead of Lie transforming the Lagrangian, as is usually done [13, 14]. This brought a much simplified reduction. 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 could be used, and no gauge fixing was needed. This physical coordinate is the unit vector 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 [15]. 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 c for the gyro-angle coordinate raises some questions, because the coordinate system becomes constrained: the variable c has to remain normalized and perpendicular to the magnetic field. Changing the spatial position $\mathbf{q}$ implies to change the coordinate 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 [11, 16]. It was already present in [12] when Lie transforming the equations of motion, but it is more involved to deal with for the full guiding-center reduction, because the coordinate 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 [5, 7]. To avoid confusion, they will be the topic of another paper [17]. 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 [18].

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 [12] 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 [8, 19]. An acceptable solution to these issues, which is the goal of the present paper, 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 [19, 20] 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 paper is organized as follows. In sect. 2, 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 sect. 3, the method is applied to the guiding-center reduction in case the gyro-angle coordinate is chosen as the physical variable 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 [18, 21].

In sect. 4, 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 subsect. IV A.

II. 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

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

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 φ is the so-called *pitch-angle*, i.e. the angle between the velocity and the magnetic field. The last coordinate c is the unit vector of the perpendicular velocity.

Then, the equations of motion write

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

where $\mathbf{p}$ is now a shorthand for $p(\mathsf{b}\cos\varphi + \mathsf{c}\sin\varphi)$, the norm of the magnetic field $\|\mathbf{B}\|$ is denoted by B, and following Littlejohn's notations [4–6], 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 $\omega_L := \frac{eB}{m}$. Writing the dynamics as $\dot{\mathbf{z}}\cdot\partial_{\mathbf{z}}$, all other terms as ω_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. [1, 2, 5, 12]) shows that it is rather

$$\epsilon := \frac{p\sin\varphi}{eB}\nabla\,. \tag{1}$$

It is an operator, but the gradient ∇ 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 ϵ 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} [1, 2, 5].

A. Choice for the gyro-angle coordinate

The Larmor frequency $\omega_L := \frac{eB}{m}$ concerns only one coordinate, 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 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 θ is defined from the oriented angle between the chosen reference axis $\mathsf{e}_1(\mathbf{q})$ and the vector c through the following relation:

$$\mathsf{c} = -\sin\theta\mathsf{e}_1 - \cos\theta\mathsf{e}_2\,. \tag{2}$$

The equation of motion for θ is

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

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 [5].

The coordinate θ 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 e_1, as well as about the presence of an anholonomic phase in the coordinate system [5, 7, 11, 22, 23]. All these difficulties originate because θ is not given by the physics, it is not a purely intrinsic coordinate. For physical results, what is needed is not θ but only c. This is clearly illustrated by all the results of guiding-center theory, e.g. [1, 2, 5, 6], where θ intervenes only through the quantity c everywhere (except in its own definition and in subsequent relations). So, we will avoid this coordinate and keep the corresponding initial variable, c, as in [12]. The quantity θ 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

$$\mathsf{c} := \frac{\mathbf{p}_\perp}{\|\mathbf{p}_\perp\|} = \frac{\mathbf{p} - \mathsf{b}(\mathsf{b}\cdot\mathbf{p}\ )}{p\sin\varphi}\,, \tag{3}$$

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 c is not independent of the spatial position.

When the coordinate $\mathbf{q}$ is changed, the coordinate c cannot be kept unchanged, otherwise it may get out of $\mathbf{B}^\perp$. Differentiating relation (3) with respect to $\mathbf{q}$ gives

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot(\mathsf{cb} + \mathsf{aa}\cot\varphi)\,. \tag{4}$$

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

$$\begin{aligned}
&-\sin\varphi\nabla\varphi = \nabla\cos\varphi = \nabla\mathsf{b}\cdot\frac{\mathbf{p}}{p} = \nabla\mathsf{b}\cdot\mathsf{c}\sin\varphi\,,\\
&\quad\Rightarrow \nabla\varphi = -\nabla\mathsf{b}\cdot\mathsf{c}\,,\\
&\quad\Rightarrow 0 = \nabla\left(\frac{\mathbf{p}}{p}\right)\\
&\qquad\quad = \nabla\mathsf{c}\sin\varphi + \nabla\mathsf{b}\cos\varphi + \nabla\varphi(-\mathsf{b}\sin\varphi + \mathsf{c}\cos\varphi)\\
&\qquad\quad = (\nabla\mathsf{c} + \nabla\mathsf{b}\cdot\mathsf{cb})\sin\varphi + \nabla\mathsf{b}\cdot(1-\mathsf{cc})\cos\varphi\,,\\
&\quad\Rightarrow \nabla\mathsf{c} = -\nabla\mathsf{b}\cdot(\mathsf{cb} + \mathsf{aa}\cot\varphi)\,,
\end{aligned} \tag{5}$$

in which ∇ 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 b is a unit vector, which implies $\nabla\mathsf{b}\cdot\mathsf{b} = \nabla\left(\frac{\mathsf{b}^2}{2}\right) = 0$.

Formula (4) 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 φ is zero, since they involve many $\sin\varphi$ as denominators [1, 2, 5, 6]. At the points where $\varphi = 0 \pmod{\pi}$, the vector c itself is not defined, neither is the coordinate θ. It is an implicit assumption in all gyro-kinetics and guiding-center works that those points are excluded from the theory.

If the coordinate θ was used, formula (4) would be replaced by

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot\mathsf{cb} + \nabla\mathsf{e}_1\cdot\mathsf{e}_2\ \mathsf{a}\,, \tag{6}$$

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 (4) and (6) write

$$\nabla\mathsf{c} = -\nabla\mathsf{b}\cdot\mathsf{c}\mathsf{b} + \mathbf{R}_g\mathsf{a}\,. \tag{7}$$

The value $\mathbf{R}_g = -\nabla\mathsf{b}\cdot\mathsf{a}\mathsf{a}\cot\varphi$ corresponds to the physical definition of c, and the corresponding formula (4). 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 θ, according to formula (6).

The spatial dependence of the vector c through formula (7) must be taken into account each time a gradient acts on a function that depends on the fast angle 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 [17] 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 ∇ is the (spatial part of the) corresponding covariant derivative.

B. Lie transforming the Lagrangian 1-form

The goal is to isolate the slow part of the dynamics from the fast angle θ (or c), by performing a near-identity change of coordinates such that the dynamics of the remaining coordinates does not depend on θ; this averaging procedure is the primary requirement for the guiding-center reduction. We showed in [12] how it can be derived by Lie transforming the equations of motion when keeping the physical coordinate 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.

The corresponding Hamiltonian function is just the particle kinetic energy

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

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

$$\{F,G\} = \partial_{\mathbf{q}}F\cdot\partial_{\mathbf{p}}G - \partial_{\mathbf{p}}F\cdot\partial_{\mathbf{q}}G - \partial_{\mathbf{p}}F\cdot e\mathbf{B}\times\partial_{\mathbf{p}}G\,. \tag{8}$$

Together, the Poisson bracket and the Hamiltonian induce the equations of motion through Hamilton's equations

$$\dot{\mathbf{z}} = \{\mathbf{z},H\}\,, \tag{9}$$

where all the coordinates were grouped in a single phase-space vector $\mathbf{z} := (\mathbf{p}, \mathbf{q})$.

Instead of working with the Poisson bracket and the Hamiltonian, it is easier to work on the *Poincaré-Cartan 1-form* Γ [6]. 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 [2, 6, 14]

$$\Gamma := (e\mathbf{A} + \mathbf{p})\cdot d\mathbf{q} - H dt \,, \tag{10}$$

which yields a variational formulation of the dynamics with the action [5]:

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

The symplectic part of Γ is $\Gamma_s := (e\mathbf{A} + \mathbf{p})\cdot d\mathbf{q}$ [2, 6]. It is a 1-form $\Gamma_s = \Gamma_s^i dz^i$ in the usual 6-dimensional phase-space, which gives the Lagrange 2-form through exterior derivative

$$\omega_s := d\Gamma_s = (\partial_i \Gamma_s^j - \partial_j \Gamma_s^i)\; dz^i \otimes dz^j \,, \tag{11}$$

where $\otimes$ means tensorial product, and Einstein's convention is used: there is an implicit summation over repeated indices. In turn, ω_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 ϵ allows for a perturbation expansion of all quantities. The Lagrangian Γ from formula (10) 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 [1])

$$\begin{aligned} \Gamma_{-1} &:= e\mathbf{A}\cdot d\mathbf{q} \,, \\ \Gamma_0 &:= \tfrac{\mathbf{p}}{m}\cdot d\mathbf{q} - \tfrac{p^2}{2m} dt \,, \end{aligned} \tag{12}$$

whose ratio is indeed of order ϵ. 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 ϵ. In fact, 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 ϵ^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 formula (14) was used for Lie derivatives, and formula (11) for exterior derivatives (here in the 7-dimensional space).

Another consequence is that at each order n in ϵ, 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 $d(\Gamma + dS) = d\Gamma + d^2S = d\Gamma$ for any function S, which is called a *gauge function* [2]. 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 Φ when there is a non-zero electric field, see page 48): 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 the maximal reduction, since it gives the expected seventh freedom. Then the reduced Lagrangian is :

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

It will be determined order by order in ϵ^n:

$$\begin{aligned}
\overline{\Gamma}_{-1} &= \Gamma_{-1} + dS_{-1} \qquad (13)\\
\overline{\Gamma}_0 &= \mathsf{L}_1\Gamma_{-1} + \Gamma_0 + dS_0\\
\overline{\Gamma}_1 &= \left(\mathsf{L}_2 + \tfrac{\mathsf{L}_1^2}{2}\right)\Gamma_{-1} + \mathsf{L}_1\Gamma_0 + dS_1\\
\overline{\Gamma}_2 &= \left(\mathsf{L}_3 + \mathsf{L}_2\mathsf{L}_1 + \tfrac{\mathsf{L}_1^3}{6}\right)\Gamma_{-1} + \left(\mathsf{L}_2 + \tfrac{\mathsf{L}_1^2}{2}\right)\Gamma_0 + dS_2\\
\overline{\Gamma}_3 &= \left[\mathsf{L}_4 + \mathsf{L}_3\mathsf{L}_1 + \mathsf{L}_2\left(\tfrac{\mathsf{L}_2}{2} + \tfrac{\mathsf{L}_1^2}{2}\right) + \tfrac{\mathsf{L}_1^4}{24}\right]\Gamma_{-1}\\
&\qquad + \left[\mathsf{L}_3 + \mathsf{L}_2\mathsf{L}_1 + \tfrac{\mathsf{L}_1^3}{6}\right]\Gamma_0 + dS_3\\
&...
\end{aligned}$$

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

$$\mathcal{L}_{\mathbf{G}}\gamma = (i_{\mathbf{G}}d + d\ i_{\mathbf{G}})\gamma\,, \qquad (14)$$

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

$$i_{\mathbf{G}}\gamma = \mathbf{G}\cdot\gamma = \gamma(\mathbf{G})\,. \qquad (15)$$

So, the first term in (14) 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 (14) involves an exterior derivative, and can be removed by redefining the gauge function S. In addition, formula (14) 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 $\mathbf{G}$

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

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 formula (13), 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 ϵ^n to absorb all the exterior derivatives involved in equation (13):

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

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, equations (13) can be written

$$\begin{aligned}
\overline{\Gamma}_{-1} &= \Gamma_{-1} + dS_{-1} \qquad (16)\\
\overline{\Gamma}_0 &= \mathbf{G}_1\cdot\omega_{-1} + \Gamma_0 + dS_0\\
\overline{\Gamma}_1 &= \mathbf{G}_2\cdot\omega_{-1} + \tfrac{\mathbf{G}_1}{2}\cdot(\omega_0 + \overline{\omega}_0) + dS_1\\
\overline{\Gamma}_2 &= \mathbf{G}_3\cdot\omega_{-1} + \mathbf{G}_2\cdot\overline{\omega}_0 + \tfrac{\mathbf{G}_1}{6}\cdot d\big[\mathbf{G}_1\cdot(2\omega_0 + \overline{\omega}_0)\big] + dS_2\\
\overline{\Gamma}_3 &= \mathbf{G}_4\cdot\omega_{-1} + \mathbf{G}_3\cdot\overline{\omega}_0 + \mathbf{G}_2\cdot\overline{\omega}_1 - \tfrac{(\mathbf{G}_2\cdot d)^2}{2}\omega_{-1}\\
&\qquad + \tfrac{(\mathbf{G}_1\cdot d)^2}{24}\mathbf{G}_1\cdot(3\omega_0 + \overline{\omega}_0) + dS_3\\
&\ ...
\end{aligned}$$

where the notation $\mathbf{G}\cdot := i_{\mathbf{G}}$ is used for the interior product, as in formula (15). 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.

C. 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:

$$\texttt{osc}(\overline{\Gamma}_n) = 0 \,, \qquad (17)$$

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 [15], either using the matrix calculus introduced in [12]. Also, the coordinate θ 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 [12], 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 ϵ^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

$$\overline{\Gamma}^j_n = 0 \,, \tag{18}$$

for higher n. For differential forms, we use the same convention as for vectors: the index n indicates the order in ϵ^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.

Equation (18) 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 the 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 θ-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 θ-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.

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

$$\overline{\mu} \approx \mu := \tfrac{(p \sin\varphi)^2}{2mB} \,. \tag{19}$$

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

$$\overline{\Gamma}^{\theta} := \overline{\mu} \,. \tag{20}$$

Indeed, then the reduced Poisson bracket verifies

$$\overline{\mathbb{J}}^{\mu\theta} = 1, \text{ and } \overline{\mathbb{J}}^{\mu i} = 0 \,, \tag{21}$$

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 (17).

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 (21). 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 sub-algebra 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 formula (18).

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 (17). 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 (20). 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 (18).

This makes seven requirements at each order in ϵ, one for each component of the reduced Lagrangian $\overline{\Gamma}_n$ in formulae (13), 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.

III. 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 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 Γ_0 (since the lowest order corresponds to Γ_{-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 Γ_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 Γ_n (or rather $\overline{\Gamma}_n$), and we will often use the expression "at order Γ_n", instead of "at the order corresponding to Γ_n".

A. Preliminary transformation and initial setting

The goal is to solve equation (16) for the guiding-center reduction, with the Lagrangian (12), and with the requirements (17)-(20) 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 equation (19). 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 μ is the zeroth-order magnetic moment

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

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

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

Interestingly, the pitch-angle φ intervenes only through its cotangent, which was mentioned in [12] 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 [12], where the variables φ and p were used, and the Larmor-radius prefactor $r_L = p\sin\varphi$ was not polynomial in $\cot\varphi$. Here, the magnetic moment μ 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

$$\phi := \cot\varphi$$

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 μ 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 ϵ [1]. Here, it is useless since the order in ϵ is already indicated by the order in other quantities: Γ_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. formulae (32), (43)-(45), (55)-(57), (58)-(59), etc. Hence the order can be readily obtained by the overall order in $\sqrt{B}$ or ∇. 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

$$\begin{aligned} A &\longrightarrow \underline{A} := eA \\ B &\longrightarrow \underline{B} := eB \\ \mu &\longrightarrow \underline{\mu} := \tfrac{\mu m}{e} = \tfrac{(p\sin\varphi)^2}{2eB} \\ t &\longrightarrow \underline{t} := \tfrac{t}{m}\,, \end{aligned} \tag{22}$$

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 (12)

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

with

$$\begin{aligned} \Gamma_{-1} &= \mathbf{A}\cdot d\mathbf{q}\,, \\ \Gamma_0 &= \sqrt{2\mu B}\;(\mathsf{b}\phi + \mathsf{c})\cdot d\mathbf{q} - \mu B(1+\phi^2)dt\,. \end{aligned} \tag{23}$$

It can be divided into its average and fluctuating part

$$\begin{aligned} \mathsf{avg}(\Gamma_{-1}) &= \Gamma_{-1} = \mathbf{A}\cdot d\mathbf{q} \\ \mathsf{avg}(\Gamma_0) &= \sqrt{2\mu B}\;\mathsf{b}\phi\cdot d\mathbf{q} - \mu B(1+\phi^2)dt \\ \mathsf{osc}(\Gamma_0) &= \sqrt{2\mu B}\;\mathsf{c}\cdot d\mathbf{q}\,. \end{aligned} \tag{24}$$

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

$$\omega = d\Gamma = \omega_{-1} + \widetilde{\omega}_0 + \overline{\omega}_0 \,, \tag{25}$$

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 (25) is the lowest-order Lagrange 2-form, related to Larmor gyrations. The second (resp. third) term in (25) 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 (25) to the Lagrange 2-form are explicitly given by

$$\begin{aligned}
\omega_{-1} :&= d\Gamma_{-1} = d\mathbf{q}\cdot\nabla(\mathbf{A})\cdot \wedge d\mathbf{q} = d\mathbf{q}\cdot(-\mathbf{B}) \times d\mathbf{q} \\
\overline{\omega}_0 :&= d\mathtt{avg}(\Gamma_0) = d(\sqrt{2\mu B}\mathsf{b}\phi)\cdot \wedge d\mathbf{q} - d(\mu B(1+\phi^2)) \wedge dt \\
&= \sqrt{2\mu B}\left\{d\phi\mathsf{b} + d\mathbf{q}\cdot\nabla\mathsf{b} + \left[\tfrac{d\mathbf{q}\cdot\nabla B}{2B} + \tfrac{d\mu}{2\mu}\right]\phi\mathsf{b}\right\}\cdot \wedge d\mathbf{q} \\
&\qquad - \left\{(1+\phi^2)(Bd\mu + \mu dB) + 2\mu B d\phi\right\} \wedge \tfrac{dt}{m} \\
\widetilde{\omega}_0 :&= d\mathtt{osc}(\Gamma_0) = d(\sqrt{2\mu B}\mathsf{c})\cdot \wedge d\mathbf{q} \\
&= \sqrt{2\mu B}\left\{d\mathsf{c} + \left[\tfrac{d\mathbf{q}\cdot\nabla B}{2B} + \tfrac{d\mu}{2\mu}\right]\mathsf{c}\right\}\cdot \wedge d\mathbf{q} \,.
\end{aligned} \tag{26}$$

Now, $d\mathsf{c}$ involves two contributions: one corresponding purely to the gyro-angle, in which the variable 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 formulae (4) and (7). 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 $(d\mathsf{c} - d\mathbf{q}\cdot\nabla\mathsf{c})\cdot\mathsf{aa}$ since when $\mathbf{q}$ is constant, the variation of c can only be in the direction of a, i.e. along the circle $\mathbb{S}^1(\mathbf{q})$. On the whole

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

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) 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 (6).

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 ω above shows that a natural basis is

$$\Big(\mathsf{c}\cdot d\mathbf{q}, \mathsf{a}\cdot d\mathbf{q}, \mathsf{b}\cdot d\mathbf{q} \mid d\phi, d\mu, \delta\theta \mid -dt\Big)\,, \tag{27}$$

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 formulae (30) and (50) for instance).

In this basis, formulae (26) become

$$\omega_{-1} = B\left(\begin{array}{ccc|c|c} 0 & 1 & 0 & & \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & & \\ \hline & 0 & & 0 & 0 \\ \hline & 0 & & 0 & 0 \end{array}\right) \tag{28}$$

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

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

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

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

B. Immediate orders -1 and 0

At lowest order $n = -1$, the reduced Lagrangian (16) 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

$$\overline{\Gamma}_0 = \mathbf{G}_1\cdot\omega_{-1} + \Gamma_0 + dS_0\,. \tag{29}$$

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 (27) 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)

$$\big(\mathsf{c}\cdot\nabla, \mathsf{a}\cdot\nabla, \mathsf{b}\cdot\nabla \mid \partial_\phi, \partial_\mu, -\mathsf{a}\cdot\partial_{\mathsf{c}} \mid -\partial_t\big)\,, \tag{30}$$

because ∇ 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 [15]. It just an intrinsic definition of the usual ∂_θ. The set (30) is actually the natural basis for vector fields, which confirms the relevance of the basis (27) for 1-forms.

Be careful that the chosen bases 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, equation (29) is easily computed in matrix form

$$\begin{aligned}\overline{\Gamma}_0 &= \mathbf{G}_1\cdot\omega_{-1} + \Gamma_0 + dS_0 \\ &= \begin{pmatrix} -B\mathbf{G}_1^{\mathsf{a}} & B\mathbf{G}_1^{\mathsf{c}} & 0 & \mid & 0 & 0 & 0 & \mid & 0 \end{pmatrix} \\ &+ \begin{pmatrix} \sqrt{2\mu B} & 0 & \sqrt{2\mu B}\phi & \mid & 0 & 0 & 0 & \mid & \mu B(1+\phi^2) \end{pmatrix}, \\ &+ \begin{pmatrix} \mathsf{c}\cdot\nabla S_0 & \mathsf{a}\cdot\nabla S_0 & \mathsf{b}\cdot\nabla S_0 & \mid & \partial_\phi S_0 & \partial_\mu S_0 & -\mathsf{a}\cdot\partial_{\mathsf{c}} S_0 & \mid & 0 \end{pmatrix}\end{aligned} \tag{31}$$

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 ω_{-1} is linked to Larmor gyration, as appears in (26). 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 (17), the condition of the right-hand side of equation (A2) being in the range of ω_{-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

$$\mathsf{osc}(\mathbf{G}_1^{\mathsf{c}}\mathsf{c} + \mathbf{G}_1^{\mathsf{a}}\mathsf{a}) := \tfrac{\sqrt{2\mu B}}{B}\mathsf{a}\,. \tag{32}$$

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 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 (18) and (20). The magnetic moment requirement (20) is not concerned here, since $\overline{\Gamma}^{\theta}$ is automatically zero. Last, for the requirement for the maximal reduction (18), only the components $\overline{\Gamma}_1^{\mathsf{c},\mathsf{a}}$ can be controlled, and setting them to zero imposes the average part of $\mathbf{G}_1^{\mathsf{c},\mathsf{a}}$

$$\mathsf{avg}(\mathbf{G}_1^{\mathsf{c}}\mathsf{c} + \mathbf{G}_1^{\mathsf{a}}\mathsf{a}) := 0\,. \tag{33}$$

Formulae (32)-(33) 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 formulae (72) and (43) 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 formula (22)).

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 c, 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 (31) writes

$$\overline{\Gamma}_0 = \sqrt{2\mu B}\mathsf{b}\phi\cdot d\mathbf{q} - \mu B(1+\phi^2)dt\,. \tag{34}$$

It is just the average of the zeroth-order initial Lagrangian Γ_0: $\overline{\Gamma}_0 = \mathtt{avg}(\Gamma_0)$. 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 (25).

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 M_∞ mentioned in the appendix (see page 60) is not identified yet.

C. Turning point: Order 1

At the following order $n = 1$, the reduced Lagrangian is given by (16), 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 ω_{-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|c|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) \\
&+ \quad \left(\overline{\Gamma}_1^{\mathsf{c}} \quad \overline{\Gamma}_1^{\mathsf{a}} \quad \overline{\Gamma}_1^{\mathsf{b}} \;\middle|\; \overline{\Gamma}_1^{\phi} \quad \overline{\Gamma}_1^{\mu} \quad \overline{\Gamma}_1^{\theta} \;\middle|\; \overline{H}_1 \right) \\
&- \quad \left(\partial_c S_1 \quad \partial_a S_1 \quad \partial_b S_1 \;\middle|\; \partial_\phi S_1 \quad \partial_\mu S_1 \quad \partial_\theta S_1 \;\middle|\; 0 \right) ,
\end{aligned}
\tag{35}
$$

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, ∂_c is different from ∂_{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 paper, in order to save room for the editor, 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 equation (35), 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.

Equation (35) can be simplified by removing the components that do not contribute

$$
\begin{aligned}
&\begin{pmatrix} \mathbf{G}_2^{\mathsf{c}} \\ \mathbf{G}_2^{\mathsf{a}} \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & B & 0 & 0 & 0 & 0 & 0 \\ -B & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \\
&+ \begin{pmatrix} \mathbf{G}_1^{\mathsf{b}} \\ \mathbf{G}_1^{\phi} \\ \mathbf{G}_1^{\mu} \\ \mathbf{G}_1^{\theta} \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} -I_{21} & K_{21} & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & H_0 \frac{\partial_b B}{B} \\ \hline 0 & 0 & \sqrt{2\mu B} & & & & 2\mu B\phi \\ \frac{\sqrt{2\mu B}}{4\mu} & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & & 0 & & \frac{H_0}{\mu} \\ 0 & -\frac{\sqrt{2\mu B}}{2} & 0 & & & & 0 \end{array}\right) \\
&= -\begin{pmatrix} \mathbf{G}_1^{\mathsf{c}} \\ \mathbf{G}_1^{\mathsf{a}} \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & -J_{21} & I_{21} & 0 & -\frac{\sqrt{2\mu B}}{4\mu} & 0 & H_0 \frac{\partial_c B}{B} \\ J_{21} & 0 & -K_{21} & 0 & 0 & \frac{\phi\sqrt{2\mu B}}{2} & H_0 \frac{\partial_a B}{B} \end{array}\right) \\
&+ \quad \left(\overline{\Gamma}_1^{\mathsf{c}} \quad \overline{\Gamma}_1^{\mathsf{a}} \quad \overline{\Gamma}_1^{\mathsf{b}} \mid \overline{\Gamma}_1^{\phi} \quad \overline{\Gamma}_1^{\mu} \quad \overline{\Gamma}_1^{\theta} \mid \overline{H}_1 \right) \\
&- \quad \left(\partial_c S_1 \quad \partial_a S_1 \quad \partial_b S_1 \mid \partial_\phi S_1 \quad \partial_\mu S_1 \quad \partial_\theta S_1 \mid 0 \right) .
\end{aligned}
\tag{36}
$$

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 equation (36). The pivotal matrix is not invertible. It has six columns and seven rows; the columns ϕ and μ are not linearly independent; in addition, its column for θ 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 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 ϕ will be automatically zero, because the procedure will imply $\mathbf{G}_1^{\mathsf{b}} = 0$, and $S_1 = 0$.

As for the column θ, 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

$$\overline{\Gamma}_1^{\theta} = \tfrac{\sqrt{2\mu B}}{2}\mathbf{G}_1^{\mathsf{a}} = \mu \,, \tag{37}$$

which makes the automatically verified for this component. In fact, if μ 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 equation (37) should be $\overline{\mu}$ instead of μ, because of formula (20). However, in the derivation, all expressions for $\overline{\Gamma}$ are functions of μ. In the reduced system, the corresponding expressions will be the same functions evaluated on the reduced coordinate $\overline{\mu}$, as is well emphasized in [5]. So, formula (37) means that the first-order reduced Lagrangian $\overline{\Gamma}_1$ will actually have its $\overline{\theta}$ component equal to $\overline{\mu}$. It is why this result verifies the requirement (20) on the magnetic moment. We will not insist more on this point.

With the result (37), 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 ω_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 $\mathtt{avg}(S_1)$ appears in the column, and $\mathtt{avg}(S_1)$ 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 $\left(\mathbf{G}_1^{\mathsf{b}}, \mathtt{avg}(S_1)\right)$. But it is solvable; it implies that

$$\mathtt{avg}(S_1) = \mathcal{K}(\phi\sqrt{\mu}, \mathbf{q})\,, \tag{38}$$

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

$$\mathbf{G}_1^{\mathsf{b}} = \tfrac{1}{\sqrt{2\mu B}}\partial_\phi \mathtt{avg}(S_1)\,. \tag{39}$$

- 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 Γ_1' remain free. They can be used to explore the various guiding-center representations [19]. 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 $\mathtt{avg}(S_1)$ (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:

$$\mathtt{avg}(S_1) = 0\,. \tag{40}$$

- 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 ∂_θ, and which easily controls the fluctuating part of the equation, but involves secular differential equations for the average part of the equation [12].

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 equation (40). 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

$$\left[\phi\partial_b + \tfrac{\nabla\cdot\mathsf{b}}{2}(1+\phi^2)\partial_\phi\right]\mathcal{K} = \tfrac{B}{\sqrt{2\mu B}}\mathtt{avg}(\mathbf{G}_1^\mu). \tag{41}$$

With the expression (55) 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 ϕ and μ gives a non-solvable equation. For instance, the coefficient of order $\phi^1\mu^1$ implies the following equation

$$\nabla\cdot\mathsf{b}\ \mathcal{K}_2 = -\mathsf{b}\cdot\nabla\times\mathsf{b}\,, \tag{42}$$

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

$$\mathtt{osc}(S_1) = \mathbf{G}_1^{\mathsf{b}} = 0 \tag{43}$$

$$\begin{aligned}
\mathtt{osc}(\mathbf{G}_1^{\phi}) &= \tfrac{\sqrt{2\mu B}}{B}(1+\phi^2)\Big[\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} + \phi\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\Big] \\
\mathtt{osc}(\mathbf{G}_1^{\mu}) &= \mu\tfrac{\sqrt{2\mu B}}{B}\Big[-\tfrac{B'\mathsf{a}}{B} - \phi\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}+\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2} - 2\phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\Big]
\end{aligned} \tag{44}$$

$$\begin{aligned}
\mathtt{avg}(\mathbf{G}_1^{\phi}) &= \tfrac{1}{\sqrt{2\mu B}}\Big[\overline{\Gamma}_1'^{\mathsf{b}} + \mu\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\Big] - \tfrac{\phi}{2\mu}\mathtt{avg}(\mathbf{G}_1^{\mu}) \\
\overline{H}_1 &= \sqrt{2\mu B}\phi\Big[\overline{\Gamma}_1'^{\mathsf{b}} + \mu\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\Big] + B\mathtt{avg}(\mathbf{G}_1^{\mu}) \\
\mathbf{G}_2^{\mathsf{a}} &= \tfrac{\mu}{B}\Big[\phi\tfrac{7\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-9\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} - \phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\Big] + \tfrac{1}{2\sqrt{2\mu B}}\mathtt{avg}(\mathbf{G}_1^{\mu}) \\
\mathbf{G}_2^{\mathsf{c}} &= -\tfrac{\sqrt{2\mu B}}{B}\tfrac{\mathbf{G}_1^{\theta}}{2} - \tfrac{\mu}{B}\bar{\mathsf{a}}\mathsf{c}'\mathsf{a}\,,
\end{aligned} \tag{45}$$

where following Littlejohn, a condensate notation is used for gradients: the curved prime is used to indicate gradients of the magnetic field, and the short overbar over a vector c or 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 comes from [6], slightly adapted to make it more explicit, in order to fit in with higher-order expressions [12]. Be careful, the straight prime and the long overline are different and do not indicate gradients or matrix transpose, e.g. in $\overline{\Gamma}'$.

Finally, the first-order reduced Lagrangian writes

$$\overline{\Gamma}_1 = -\mu\mathsf{a}\cdot d\mathsf{c} + \overline{\Gamma}_1'^{\mathsf{b}}\mathsf{b}\cdot d\mathbf{q} - \overline{H}_1 dt\,. \tag{46}$$

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:

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

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

$$\overline{\Gamma}_1'^{\mathsf{b}} = -\tfrac{B}{\sqrt{2\mu B}\phi}\mathtt{avg}(\mathbf{G}_1^{\mu}) - \mu\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\,. \tag{47}$$

This last choice is possible only if the inversion of ϕ does not cause a singularity, i.e. if $\mathtt{avg}(\mathbf{G}_1^{\mu})$ has no overall contribution of order zero in ϕ, which will have to be verified when $\mathtt{avg}(\mathbf{G}_1^{\mu})$ is identified (see equation (55)).

The first-order Lagrangian (46) induces the following first-order reduced Lagrange matrix, which will be part of the pivotal matrix at higher orders:

$$\overline{\omega}_1 = \left(\begin{array}{ccc|ccc|c}
\multicolumn{3}{c|}{\overline{\omega}_1^{\mathbf{q};\mathbf{q}}} & 0 & \mathbf{R}_g\cdot\mathsf{c} & 0 & \partial_c\overline{H}_1 \\
 & & & 0 & \mathbf{R}_g\cdot\mathsf{a} & 0 & \partial_a\overline{H}_1 \\
 & & & -\partial_\phi\overline{\Gamma}_1'^{\mathsf{b}} & -\partial_\mu\overline{\Gamma}_1'^{\mathsf{b}} + \mathbf{R}_g\cdot\mathsf{b} & 0 & \partial_b\overline{H}_1 \\
\hline
0 & 0 & \partial_\phi\overline{\Gamma}_1'^{\mathsf{b}} & 0 & 0 & 0 & \partial_\phi\overline{H}_1 \\
-\mathbf{R}_g\cdot\mathsf{c} & -\mathbf{R}_g\cdot\mathsf{a} & -\mathbf{R}_g\cdot\mathsf{b} + \partial_\mu\overline{\Gamma}_1'^{\mathsf{b}} & 0 & 0 & 1 & \partial_\mu\overline{H}_1 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 \\
\hline
-\partial_c\overline{H}_1 & -\partial_a\overline{H}_1 & -\partial_b\overline{H}_1 & -\partial_\phi\overline{H}_1 & -\partial_\mu\overline{H}_1 & 0 & 0
\end{array}\right), \tag{48}$$

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 (43)-(45) 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 $(\mathbf{q}-\overline{\mathbf{q}})_2 = \mathbf{G}_2^{\mathbf{q}} - \frac{1}{2}\mathbf{G}_1^{\mathbf{z}}\cdot\partial_{\mathbf{z}}\mathbf{G}_1^{\mathbf{q}}$. 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 features 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 M_∞ 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 A, the first stage of the reduction is achieved, and the order at which $\overline{\omega}_{-1}+\overline{\omega}_0+...+\overline{\omega}_{n_b}$ 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 in pages 60 and 61). 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 M_∞, 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.

D. Core of the second stage: order 2

At the following order $n = 2$, the reduced Lagrangian is given by (16)

$$\overline{\Gamma}_2 = \mathbf{G}_3{\cdot}\omega_{-1} + \mathbf{G}_2{\cdot}\overline{\varpi}_0 + \tfrac{\mathbf{G}_1}{6}{\cdot}d\big[\mathbf{G}_1{\cdot}(3\overline{\varpi}_0 + 2\tilde{\omega}_0)\big] + dS_2\,. \tag{49}$$

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

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

In equation (49), the exterior derivative must be computed for the 1-form λ 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 (11) 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:

$$\begin{aligned} d\gamma = d(\gamma^j\mathbf{e}_j) &= d(\gamma^j)\wedge\mathbf{e}_j + \gamma^j d\mathbf{e}_j \\ &= \mathbf{e}_i(\partial_i\gamma^j - \partial_j\gamma^i)\mathbf{e}_j + \gamma^j d\mathbf{e}_j\,, \end{aligned} \tag{50}$$

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

In formula (50), 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 $d\gamma^j$, it 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 ∂_i are given by the dual basis (30) to the chosen basis (27) 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 (7). 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 γ 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 $\gamma'^{\mathsf{q}}\cdot d\mathbf{q}$ do not involve ac'. Then apply formula (50) to γ' 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:

$$d\gamma = \mathbf{e}_i(\partial_i\gamma'^j - \partial_j\gamma'^i)\mathbf{e}_j + \gamma'^i d'\mathbf{e}_i - \mathbf{e}_i\partial_i\gamma'^\theta\wedge\mathbf{R}_g\cdot d\mathbf{q}\,, \tag{51}$$

in which $d'\mathbf{e}_i = d\mathbf{e}_i - \delta^{i\theta}d(\mathsf{a}\cdot\mathsf{c}')\wedge d\mathbf{q}$, where δ 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 formula (51), 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 $\mathbf{G}_1^\theta$:

$$\begin{aligned}
6\big(i_{\mathbf{G}_1}d\lambda\big)^\phi &= -\mathbf{G}_1^{\mathsf{a}}\partial_\phi\big[2\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 2\mathbf{G}_1^\theta\big] \\
6\big(i_{\mathbf{G}_1}d\lambda\big)^\mu &= -\mathbf{G}_1^{\mathsf{a}}\partial_\mu\big[2\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 2\mathbf{G}_1^\theta\big] + 4\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 4\mathbf{G}_1^\theta \\
6\big(i_{\mathbf{G}_1}d\lambda\big)^\theta &= -\mathbf{G}_1^{\mathsf{a}}\partial_\theta\big[2\mathbf{G}_1^{\mathsf{a}}\mathbf{R}_g\cdot\mathsf{a} - 2\mathbf{G}_1^\theta\big] + 6\mathbf{G}_1^\mu + \mathbf{G}_1^{\mathsf{a}}.6J_{322}\,,
\end{aligned} \tag{52}$$

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 $2\mathsf{ac}'\mathsf{a}$ coming from formula (50) 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 ϕ. 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 this term can transfer the differential operator over the pre-factor $\mathbf{G}_1^{\mathsf{a}}$, which is already known. This is a way to make the equation algebraic.

This 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 this formula is not differential any more, but algebraic for $\mathbf{G}_1^{\theta}$, which is contained in λ^{a}.

Then, equation (49) becomes

$$\overline{\Gamma}_2 = \mathbf{G}_3{\cdot}\omega_{-1} + \mathbf{G}_2{\cdot}\overline{\omega}_0 + \Lambda + dS_2'\,, \tag{53}$$

with

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

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 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 $(\mathbf{G}_3, \mathbf{G}_2, \mathbf{G}_1)$.

It 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 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 $(\mathbf{q}, \mathbf{p}, t)$; and by the way, the 7-dimensional 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 (36), equation

(53) becomes

$$
\begin{aligned}
&\begin{pmatrix}\mathbf{G}_3^{\mathsf{c}}\\ \mathbf{G}_3^{\mathsf{a}}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & B & 0 & 0 & 0 & 0 & 0\\ -B & 0 & 0 & 0 & 0 & 0 & 0\end{array}\right)\\
+&\begin{pmatrix}\mathbf{G}_2^{\mathsf{b}}\\ \mathbf{G}_2^{\phi}\\ \mathbf{G}_2^{\mu}\\ \mathbf{G}_2^{\theta}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} -I & K & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & H_0\frac{\partial_b B}{B}\\ \hline 0 & 0 & \sqrt{2\mu B} & & & & 2\mu B\phi\\ 0 & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & & 0 & & \frac{H_0}{\mu}\\ 0 & 0 & 0 & & & & 0\end{array}\right)\\
+&\begin{pmatrix}\mathbf{G}_1^{\mu}\\ \mathbf{G}_1^{\theta}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & -1 & 0 & 0\end{array}\right)\\
=-&\begin{pmatrix}\mathbf{G}_2^{\mathsf{c}}\\ \mathbf{G}_2^{\mathsf{a}}\\ \mathbf{G}_1^{\mathsf{a}}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & -J & I & 0 & 0 & 0 & H_0\frac{\partial_c B}{B}\\ J & 0 & -K & 0 & 0 & 0 & H_0\frac{\partial_a B}{B}\\ 0 & 0 & 0 & 0 & \bar{\mathsf{a}}\mathsf{c}'\mathsf{a} & J_{322} & 0\end{array}\right)\\
-&\quad\left(\begin{array}{ccc|ccc|c}\Lambda_{\mathsf{c}} & \Lambda_{\mathsf{a}} & \Lambda_{\mathsf{b}} & 0 & 0 & 0 & \Lambda_t\end{array}\right)\\
+&\quad\left(\begin{array}{ccc|ccc|c}\overline{\Gamma}_2^{\mathsf{c}} & \overline{\Gamma}_2^{\mathsf{a}} & \overline{\Gamma}_2^{\mathsf{b}} & \overline{\Gamma}_2^{\phi} & \overline{\Gamma}_2^{\mu} & \overline{\Gamma}_2^{\theta} & \overline{H}_2\end{array}\right)\\
-&\quad\left(\begin{array}{ccc|ccc|c}\partial_c S_2' & \partial_a S_2' & \partial_b S_2' & \partial_\phi S_2' & \partial_\mu S_2' & \partial_\theta S_2' & 0\end{array}\right),
\end{aligned}
\tag{54}
$$

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

Formula (54) is similar to (36), and the same comments can be done as at the previous order, in the two paragraphs following formula (36).

To solve the equation, an analysis similar to the one at the previous order leads to the procedure summarized in the following tabular, 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
$\mathtt{osc}(\overline{\Gamma}_2^{\theta})$	$\mathtt{osc}(S_2')$	∂_θ
$\overline{\Gamma}_2^{\phi}$	$\mathbf{G}_2^{\mathsf{b}}$	$-\sqrt{2\mu B}$
$\overline{\Gamma}_2^{\mu}$	$\mathbf{G}_1^{\theta}$	-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 formula (44), 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 ∂_θ 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 the 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 $\overline{\Gamma}_2^{t}$. This last 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 this term by using the freedom $\mathtt{avg}(S_2')$, 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, S_2' can be chosen such that it is absent from $\mathbf{G}_1^{\theta}$, as it was done in equations (38)-(39). Then it is easy to see that the equation for $\mathtt{avg}(S_2')$ will have the same structure as equation (41). This equation may not be integrable, as it was the case for equation (41). In this case, one should relax the condition for $\mathtt{avg}(S_2')$ to be absent from $\mathbf{G}_1^{\theta}$, and the differential equation for $\mathtt{avg}(S_2')$ could be more difficult to study, because $\mathbf{G}_1^{\theta}$ is involved in the 1-form Λ in equation (54) in a rather intricate way.

Applying the procedure identified above gives

$$\begin{aligned}
\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}) \qquad (55)\\
\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') \qquad (56)\\
\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]\,, \qquad (57)
\end{aligned}$$

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 (55)-(57), the parameters involved in formulae (43)-(45) can be made explicit:

$$\begin{aligned}
\mathtt{avg}(\mathbf{G}_1^{\phi}) &= \tfrac{1}{\sqrt{2\mu B}}\Big[\overline{\Gamma}_1^{\mathsf{b}} + \mu\tfrac{1+2\phi^2}{2}(\bar{\mathsf{c}}\mathsf{b}'\mathsf{a} - \bar{\mathsf{a}}\mathsf{b}'\mathsf{c})\Big] \qquad (58)\\
\overline{H}_1 &= \sqrt{2\mu B}\phi\Big[\overline{\Gamma}_1^{\mathsf{b}} - \mu\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\Big]\\
\mathbf{G}_2^{\mathsf{a}} &= \tfrac{\mu}{B}\Big[\phi\tfrac{5\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-7\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{4} - \phi^2\bar{\mathsf{a}}\mathsf{b}'\mathsf{b}\Big]\\
\mathbf{G}_2^{\mathsf{c}} &= \tfrac{\mu}{B}\Big[-\tfrac{B'\mathsf{c}}{B} + \phi\tfrac{\bar{\mathsf{a}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{c}}\mathsf{b}'\mathsf{c}}{4} - \phi^2\bar{\mathsf{c}}\mathsf{b}'\mathsf{b}\Big] + \tfrac{\sqrt{2\mu B}}{B}\tfrac{\partial_\mu\mathtt{avg}(S_2')}{2}\,. \qquad (59)
\end{aligned}$$

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

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

The reverse choice $\overline{H}_1 = 0$ is possible, since its existence condition in equation (47) is satisfied, as is clear in (55). It corresponds to a component $\overline{\Gamma}_1^{\mathsf{b}}$ of the reduced Lagrangian given by

$$\overline{\Gamma}_1'^{\mathsf{b}} = \mu\tfrac{\bar{\mathsf{c}}\mathsf{b}'\mathsf{a}-\bar{\mathsf{a}}\mathsf{b}'\mathsf{c}}{2}\,, \qquad (60)$$

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

In formulae (55)-(59), 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 $\overline{H}_3$, formulae (55)-(59) 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 [18], and were already partly introduced in [21].

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

$$\overline{\Gamma}_2^{\mathsf{b}} = \mathtt{avg}\Big[\mathsf{R}_2^{\mathsf{b}} - \tfrac{1}{\sqrt{2\mu B\phi}}\Big\{B\mathbf{G}_2^\mu + \mathsf{R}_2^t + \mathbf{G}_2^{\mathsf{b}}(1+\phi^2)\mu B\mathsf{b}\Big\}\Big]\,.$$

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

Also $\mathtt{avg}(S_2')$ 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 $\mathtt{avg}(S_2')$, which might not be integrable, just as happened at the previous order in equation (41), 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 $\mathtt{avg}(\mathbf{G}_2^\mu)$; 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 $\mathtt{avg}(S_2')$, 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 equation (54).

E. End of the second stage: order 3

Let us turn now to the following order $\overline{\Gamma}_3$. The equation writes (16)

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

The unknowns are components of $(\mathbf{G}_4, \mathbf{G}_3, \mathbf{G}_2)$. The pivotal matrix is not M_∞, 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

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

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 M_∞, 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.

F. 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.

1. *Equation and algorithm*

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

$$\mathbf{g}_n := \Big(\overline{\Gamma}_n, S_n, \mathbf{G}_{n-1}^{\theta}, \mathtt{avg}(\mathbf{G}_{n-1}^{\mu}), \mathtt{osc}(\mathbf{G}_n^{\mu}), \mathbf{G}_n^{\mathsf{b}}, \mathbf{G}_n^{\phi}, \mathbf{G}_{n+1}^{\mathsf{c}}, \mathbf{G}_{n+1}^{\mathsf{a}}\Big)\,, \tag{61}$$

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 equation (A6)

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

where 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 M_∞ is the set of

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

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

$$\begin{aligned}
&\begin{pmatrix}\mathbf{G}^{\mathsf{c}}_{n+1}\\ \mathbf{G}^{\mathsf{a}}_{n+1}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & B & 0 & 0 & 0 & 0 & 0\\ -B & 0 & 0 & 0 & 0 & 0 & 0\end{array}\right)\\
&+\begin{pmatrix}\mathbf{G}^{\mathsf{b}}_{n}\\ \mathbf{G}^{\phi}_{n}\\ \mathbf{G}^{\mu}_{n}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} -I & K & 0 & -\sqrt{2\mu B} & -\frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & H_0\frac{\partial_b B}{B}\\ 0 & 0 & \sqrt{2\mu B} & 0 & 0 & 0 & 2\mu B\phi\\ 0 & 0 & \frac{\phi\sqrt{2\mu B}}{2\mu} & 0 & 0 & 0 & \frac{H_0}{\mu}\end{array}\right)\\
&+\begin{pmatrix}\mathtt{avg}(\mathbf{G}^{\mu}_{n-1})\\ \mathbf{G}^{\theta}_{n-1}\end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & 0 & \partial_\mu\overline{\Gamma}'^{\mathsf{b}}_1 & 0 & 0 & 1 & \partial_\mu\overline{H}_1\\ 0 & 0 & 0 & 0 & -1 & 0 & 0\end{array}\right)\\
=&-\mathsf{R}'_n\\
&+\qquad\left(\begin{array}{ccc|ccc|c}\overline{\Gamma}^{\mathsf{c}}_n & \overline{\Gamma}^{\mathsf{a}}_n & \overline{\Gamma}^{\mathsf{b}}_n & \overline{\Gamma}^{\phi}_n & \overline{\Gamma}^{\mu}_n & \overline{\Gamma}^{\theta}_n & \overline{H}_n\end{array}\right)\\
&-\qquad\left(\begin{array}{ccc|ccc|c}\partial_c S_n & \partial_a S_n & \partial_b S_n & \partial_\phi S_n & \partial_\mu S_n & \partial_\theta S_n & 0\end{array}\right)\,,
\end{aligned} \tag{62}$$

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

$$
\begin{aligned}
\mathsf{R}'_n = \mathsf{R}_n & \\
- & \begin{pmatrix} \mathbf{G}^{\mathsf{c}}_n \\ \mathbf{G}^{\mathsf{a}}_n \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} 0 & -J & I & 0 & 0 & 0 & H_0 \frac{B'_{\mathsf{c}}}{B} \\ J & 0 & -K & 0 & 0 & 0 & H_0 \frac{B'_{\mathsf{a}}}{B} \end{array}\right) \\
- & \begin{pmatrix} \mathbf{G}^{\mathsf{c}}_{n-1} \\ \mathbf{G}^{\mathsf{a}}_{n-1} \\ \mathbf{G}^{\mathsf{b}}_{n-1} \\ \hline \mathbf{G}^{\phi}_{n-1} \\ \mathtt{osc}(\mathbf{G}^{\mu}_{n-1}) \\ \hline 0 \end{pmatrix}^T \cdot \left(\begin{array}{ccc|ccc|c} & & & 0 & 0 & 0 & \partial_c \overline{H}_1 \\ & \overline{\omega}_1^{\mathbf{q};\mathbf{q}} & & 0 & 0 & 0 & \partial_a \overline{H}_1 \\ & & & -\partial_\phi \overline{\Gamma}'^{\mathsf{b}}_1 & -\partial_\mu \overline{\Gamma}'^{\mathsf{b}}_1 & 0 & \partial_b \overline{H}_1 \\ \hline 0 & 0 & \partial_\phi \overline{\Gamma}'^{\mathsf{b}}_1 & 0 & 0 & 0 & \partial_\phi \overline{H}_1 \\ 0 & 0 & \partial_\mu \overline{\Gamma}'^{\mathsf{b}}_1 & 0 & 0 & 1 & \partial_\mu \overline{H}_1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hline -\partial_c \overline{H}_1 & -\partial_a \overline{H}_1 & -\partial_b \overline{H}_1 & -\partial_\phi \overline{H}_1 & -\partial_\mu \overline{H}_1 & 0 & 0 \end{array}\right).
\end{aligned}
\tag{63}
$$

Again, the same comments as the ones after equation (36) are in place. Also, when computing R_n, formula (51) 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 $\mathtt{osc}(\overline{\Gamma}^{\theta}_n)$. In addition, the new feature is the presence of $\overline{\omega}_1$ for $\overline{\Gamma}^{\mu}_n$ and $\mathtt{avg}(\overline{\Gamma}^{\theta}_n)$, 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: (c, a) are conjugated for ω_{-1}; (μ, θ) are half-conjugated for $\overline{\omega}_1$, in the sense that the structure is quarter-canonical; and (b, ϕ) are half-conjugated for the symplectic part of $\overline{\omega}_0$.

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

Equation	Unknown	Coefficient
$\mathtt{avg}(\overline{\Gamma}^{\theta}_n)$	$\mathtt{avg}(\mathbf{G}^{\mu}_{n-1})$	1
$\mathtt{osc}(\overline{\Gamma}^{\theta}_n)$	$\mathtt{osc}(S'_n)$	∂_θ
$\overline{\Gamma}^{\phi}_n$	$\mathbf{G}^{\mathsf{b}}_n$	$-\sqrt{2\mu B}$
$\overline{\Gamma}^{\mu}_n$	$\mathbf{G}^{\theta}_{n-1}$	-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 Γ_n remain free. This can be used to explore the various guiding-center representations at higher orders [19]. 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

$$\big[\phi\partial_b + \tfrac{\nabla\cdot\mathsf{b}}{2}(1+\phi^2)\partial_\phi\big]\mathtt{avg}(S_n) = o.t.\,, \tag{64}$$

where $o.t.$ means other terms that can be explicitly computed. This resembles equation (41), but here, the condition (38) 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 equation (64) will depend on the right-hand side and must be studied at each order; a priori, it is not guaranteed, since obstructions such as (42) 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 equation (62) 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 ϵ.

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 would be that when solving the equations for $\overline{\Gamma}^{\mathbf{p}}_{n+1}$, one would begin the heavy computations for 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.

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 (17)-(20) for $\overline{\Gamma}$.

Third, some quantities are not known yet, and will be determined after the matrix inversion, namely

$$\begin{aligned}(\mathbf{g}_n)_\infty := \Big(&\mathtt{avg}(\mathbf{G}^\mu_{n-1}),\mathtt{osc}(S_n) \mid \\ &\mathbf{G}^{\mathsf{b}}_n, \mathbf{G}^\theta_{n-1}, \mathbf{G}^{\mathsf{a}}_{n+1}, \mathbf{G}^{\mathsf{c}}_{n+1} \mid \\ &\mathtt{avg}(\mathbf{G}^\phi_n), \mathtt{osc}(\mathbf{G}^\phi_n), \mathtt{osc}(\mathbf{G}^\mu_n), \overline{H}_n\Big)\,,\end{aligned} \tag{65}$$

in which a vertical dash | was written at the places where a vertical line will be written in the matrix M_∞ 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

$$(\mathbf{g}_n)_\alpha := \Big(\mathtt{avg}(\mathbf{G}^\mu_n), \mathtt{avg}(S_n), \overline{\Gamma}^{\mathsf{b}}_n\Big)\,. \tag{66}$$

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 equation (62) 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

$$\begin{aligned}\Big(&\mathtt{avg}(\overline{\Gamma}^\theta_n), \mathtt{osc}(\overline{\Gamma}^\theta_n) \mid \overline{\Gamma}^\phi_n, \overline{\Gamma}^\mu_n, \overline{\Gamma}^{\mathsf{c}}_n, \overline{\Gamma}^{\mathsf{a}}_n \mid \\ &\mathtt{avg}(\overline{\Gamma}^{\mathsf{b}}_n), \mathtt{osc}(\overline{\Gamma}^{\mathsf{b}}_n), \mathtt{osc}(\overline{\Gamma}^t_n), \mathtt{avg}(\overline{\Gamma}^t_n)\Big)\,,\end{aligned}$$

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 M_∞ below.

With this order for the rows and for the columns, the equation becomes

$$\mathsf{M}_\infty\cdot(\mathbf{g}_n)^T_\infty + \mathsf{M}_\alpha\cdot(\mathbf{g}_n)^T_\alpha + \mathsf{R}'^T = 0\,. \tag{67}$$

Here, the rows are the equations to be solved, corresponding to the (re-ordered) columns of equation (62). 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

$$\mathsf{M}_\infty := \left(\begin{array}{cc|ccc|ccc}
1 & & & & & & & \\
 & \partial_\theta & & & & & & \\
\hline
 & \partial_\phi & -\sqrt{2\mu B} & & & & & \\
 & \partial_\mu & -\phi\sqrt{\frac{B}{2\mu}} & -1 & & & & \\
 & \partial_c & -I & & -B & & & \\
 & \partial_a & K & & B & & & \\
\hline
\mathtt{avg}\partial_\mu\overline{\Gamma}^{\mathsf{b}}_1 & \partial_b & & & & \sqrt{2\mu B} & & \\
\mathtt{osc}\partial_\mu(\overline{\Gamma}^{\mathsf{b}}_1) & \partial_b & & & & & \sqrt{2\mu B}\phi\sqrt{\frac{B}{2\mu}} & \\
\mathtt{osc}\partial_\mu(\overline{H}_1) & & -H_0\nabla\cdot\mathsf{b} & & & & 2\mu B\phi & \frac{H_0}{\mu} \\
\mathtt{avg}\partial_\mu(\overline{H}_1) & & -H_0\nabla\cdot\mathsf{b} & & & 2\mu B\phi & & -1
\end{array}\right), \tag{68}$$

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 M_∞, 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 M' has null diagonal and D is purely diagonal. Then the equation can be written

$$-\mathsf{D}\cdot(\mathbf{g}_n)^T_\infty = \mathsf{M}'_\infty\cdot(\mathbf{g}_n)^T_\infty + \mathsf{M}_\alpha\cdot(\mathbf{g}_n)^T_\alpha + \mathsf{R}'^T\,.$$

The solution of equation (67) is then

$$(\mathbf{g}_n)^T_\infty = (-\mathsf{D})^{-1}\cdot\Big[\mathsf{M}'_\infty\cdot(\mathbf{g}_n)^T_\infty + \mathsf{M}_\alpha\cdot(\mathbf{g}_n)^T_\alpha + \mathsf{R}'^T\Big]\,, \tag{69}$$

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 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 ∂_θ^{-1} is the gyro-integral operator. It can be computed without introducing any gyro-gauge, with the intrinsic calculus introduced in [15], or with the matrix calculus introduced in [12]. Also, the coordinate θ 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 ∂_θ^{-1} is the primitive with respect to θ 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, formula (69) is an explicit solution for the induction relation: in the matrix M'_∞, 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

$$(\mathbf{g}_n)^T_\infty = -\mathsf{M}_\infty^{-1}\cdot\Big[\mathsf{M}_\alpha\cdot(\mathbf{g}_n)^T_\alpha + \mathsf{R}'^T\Big]\,, \tag{70}$$

with M_∞^{-1} easily computed from (68), but the coefficients are more complicated, and practical computations usually follow the procedure of formula (69).

For the solution $\mathbf{g}_n$, formula (69) or (70) must be completed by the following relations for the trivial components of $\overline{\Gamma}_n$:

$$\overline{\Gamma}_n^i = 0 \quad \text{for all } i \notin \{\mathsf{b}, t\}\,,$$

and by the determination of the parameters $(\mathbf{g}_n)_\alpha$.

The first parameter $\mathtt{avg}(\mathbf{G}_n^\mu)$ will be determined at the following order, in an equation that does not involve $\mathbf{G}_n^\phi$ and $\overline{H}_n$, so that there are not coupled equations between the orders.

The second parameter $\overline{\Gamma}_n^{\mathsf{b}}$ is either put to zero or adjusted so as to make the reduced Hamiltonian $\overline{H}_n$ zero by the relation

$$\begin{aligned}\overline{\Gamma}_n^{\mathsf{b}} =& \mathtt{avg}\Big[\mathsf{R}_n'^{\mathsf{b}} + \mathbf{G}_{n-1}^\mu \partial_\mu \overline{\Gamma}_1'^{\mathsf{b}} + \partial_b S_n \\ &- \tfrac{1}{\sqrt{2\mu B\phi}}\Big\{\mathbf{G}_{n-1}^\mu \partial_\mu \overline{H}_1 - \mathbf{G}_n^{\mathsf{b}} H_0 \nabla\cdot\mathsf{b} + B\mathbf{G}_n^\mu + \mathsf{R}_n'^t\Big\}\Big]\,.\end{aligned} \tag{71}$$

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

The last parameter $\mathtt{avg}(S_n)$ is determined by equation (64), in order to make both $\overline{\Gamma}_n^{\mathsf{b}}$ and $\overline{H}_n$ zero, when the equation is integrable. Otherwise, it is determined to cancel as many terms of $\overline{\Gamma}_n^{\mathsf{b}}$ 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 $\mathtt{avg}(\mathbf{G}_n^\mu)$, of one binary choice between $\overline{\Gamma}_n^{\mathsf{b}}$ 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 (64). Otherwise, $\mathtt{avg}(\mathbf{G}_n^\mu)$ 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.

IV. COMPARISON WITH PREVIOUS WORKS

A. Agreeing results

Computations of the previous section for the guiding-center transformation and reduced Hamiltonian can be summarized by formulae (32), (43)-(45) and (55)-(59). As for the reduced symplectic Lagrangian, it is exactly known, and is given by formulae (34) and (46), 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 formula (71).

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 [6] Littlejohn made the choice $\overline{\Gamma}_2 = 0$, and accounting for this choice, our formulae agree with his ones. In the paper [2], the choice is $H_2 = 0$, and again, accounting for this choice through formula (60) for $\overline{\Gamma}_1'^{\mathsf{b}}$, our results agree with the ones of [2]. 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 (69) 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 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

$$\mathbf{z} \longrightarrow \overline{\mathbf{z}} := ...e^{-\mathbf{G}_2}e^{-\mathbf{G}_1}\mathbf{z}\,. \tag{72}$$

The reduced Lagrangian 1-form is

$$\begin{aligned}\overline{\Gamma} &= \mathsf{avg}(\Gamma) - \mu\mathsf{a}\cdot d\mathsf{c} + \Big[\overline{\Gamma}_1'^{\mathsf{b}} + \sum_{n\geqslant 2}\overline{\Gamma}_n^{\mathsf{b}}\Big]\mathsf{b}\cdot d\mathbf{q} - \sum_n \overline{H}_n dt \\ &= (\mathbf{A} + \overline{\Gamma}^{\mathsf{b}}\mathsf{b})\cdot d\mathbf{q} - \mu\mathsf{a}\cdot d\mathsf{c} - \overline{H}dt\,,\end{aligned} \tag{73}$$

where the Hamiltonian terms $\overline{H}_n$ are provided by (69), 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

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

Here, performing these computations is useless, since the reduced Lagrangian (73) 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. (73) 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

$$\begin{aligned} \mathtt{A} &:= -\mathsf{a}\cdot\nabla\times(\mathbf{A}+\overline{\Gamma}^{\mathsf{b}}\mathsf{b}) - \mu\mathsf{c}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b'b} = I + O(B^0)\,, & \mathtt{x} &:= \mathbf{R}_g\cdot\mathsf{c}\,, \\ \mathtt{B} &:= -\mathsf{b}\cdot\nabla\times(\mathbf{A}+\overline{\Gamma}^{\mathsf{b}}\mathsf{b}) - \mu\mathsf{a}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b'c} = -B + J + O(B^0)\,, & \mathtt{y} &:= \mathbf{R}_g\cdot\mathsf{a}\,, \\ \mathtt{C} &:= -\mathsf{c}\cdot\nabla\times(\mathbf{A}+\overline{\Gamma}^{\mathsf{b}}\mathsf{b}) - \mu\mathsf{b}\cdot\nabla\mathsf{b}\cdot\mathsf{b}\times\mathsf{b'a} = K + O(B^0)\,, & \mathtt{z} &:= \mathbf{R}_g\cdot\mathsf{b}\,, \\ \mathtt{D} &:= \partial_\phi\overline{\Gamma}^{\mathsf{b}} = \sqrt{2\mu B} + O(B^0)\,, \\ \mathtt{E} &:= \partial_\mu\overline{\Gamma}^{\mathsf{b}} = \phi\tfrac{\sqrt{2\mu B}}{2\mu} + O(B^0)\,, \end{aligned} \tag{74}$$

where we used the fact that the order in $\sqrt{B}$ indicates the expansion order, as mentioned about Eq. (22).

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 formula (77). The coefficients $\mathtt{B}$ and $\mathtt{D}$ are invertible since Eq. (74) 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

$$\{f,g\} = -\nabla_* f\cdot\frac{\mathsf{b}}{\mathtt{B}}\times\nabla_* g + \frac{\mathbf{B}_*}{\mathtt{BD}}\cdot\nabla_* f\wedge\partial_\phi g + \frac{\mathtt{E}}{\mathtt{D}}\partial_\phi f\wedge\partial_\theta g + \partial_\theta f\wedge\partial_\mu g\,, \tag{75}$$

where the symbol ∇_* is a shorthand for the operator

$$\nabla_* := \nabla + \mathbf{R}_g \partial_\theta \,. \tag{76}$$

This operator does not depend of the free function $\mathbf{R}_g$, since for any choice of $\mathbf{R}_g$ the definition (76) gives the same result, which is equal to the covariant derivative ∇ when the gauge vector is chosen zero, i.e. $(\mathbf{R}_g)_* = \nabla_* \mathsf{c}\cdot\mathsf{a} = 0$.

The Poisson bracket (75) agrees with the literature, e.g. [2, 3, 5, 6, 21]. Especially, the traditional ordering is patent: the last two terms are of order μ^{-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 (1), 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. (75), 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 (74), i.e. as an effect of the five derivative operators ∂_{c}, ∂_{a}, ∂_{b}, ∂_ϕ, ∂_μ, 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

$$\mathbf{B}_* := -\mathtt{A}\mathsf{a} - \mathtt{B}\mathsf{b} - \mathtt{C}\mathsf{c} = \nabla \times \left(\mathbf{A} + \mathsf{b}\overline{\Gamma}^{\mathsf{b}}\right) + \mathbf{V}\,, \tag{77}$$

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

$$\begin{aligned}
-\mu\ d\mathsf{a}\cdot \wedge d\mathsf{c} &= -\mu\ d\mathsf{a}\cdot(\mathsf{aa} + \mathsf{bb} + \mathsf{cc})\cdot \wedge d\mathsf{c} \\
&= -\mu\ d\mathsf{a}\cdot\mathsf{b} \wedge \mathsf{b}\cdot d\mathsf{c} \\
&= -\mu\ d\mathsf{b}\cdot\mathsf{a} \wedge \mathsf{c}\cdot d\mathsf{b} \\
&= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{a} \wedge \mathsf{c}\cdot\mathsf{b}' d\mathbf{q} \\
&= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{b} \times \mathsf{b}' d\mathbf{q} \\
&=: -d\mathbf{q}\cdot\mathbf{V} \times d\mathbf{q}\,.
\end{aligned} \tag{78}$$

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 b depends only on $\mathbf{q}$ implies the fourth equality, which can be rewritten to get the final answer.

Eq. (78) 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 (78) 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

$$\begin{aligned} -\mu\ d(d\mathbf{q}\cdot\nabla\mathsf{e}_1\cdot\mathsf{e}_2) &= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{e}_2\cdot \wedge \mathsf{e}_1\text{'}d\mathbf{q} \\ &= -\mu\ d\mathbf{q}\cdot\nabla\mathsf{b}\cdot\mathsf{b} \times \mathsf{b}\text{'}d\mathbf{q}\,, \end{aligned} \tag{79}$$

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 c completely agrees with the gauge-independent part of the usual results, whereas it proceeds in a different way and never introduces the gauge 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 (78), 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 $-f\nabla g$:

$$\begin{aligned} d\mathbf{q}\cdot(df \wedge dg)^{\mathbf{q};\mathbf{q}}\cdot d\mathbf{q} &= d\mathbf{q}\cdot\nabla f \wedge g\text{'}d\mathbf{q} \\ &= d\mathbf{q}\cdot(\nabla(fg\text{'})) \wedge d\mathbf{q} \\ &= -d\mathbf{q}\cdot(\nabla \times (fg\text{'})) \times d\mathbf{q}\,. \end{aligned}$$

Thus, formula (77) can be rewritten

$$\mathbf{B}_* := \nabla \times \left[\mathbf{A} + \mathsf{b}\left\{\sqrt{2\mu B}\phi + \overline{\Gamma}_1^{\prime\mathsf{b}} + \sum_{n\geqslant 2}\overline{\Gamma}_n^{\mathsf{b}} - \mu\mathbf{R}_g\right\}\right],$$

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 [2, 3, 6]. 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 Φ (or rather $\nabla\Phi$), which introduces a second parameter ϵ_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 [3, 5, 6]. The series remains perturbative provided the associated parameter is small $\epsilon_E << 1$, as usual in guiding-center reductions.

B. Polynomiality in the momentum coordinates

In the results above, the usual expansions are recovered, but the choice of the coordinates c and ϕ 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 c, but they did not use it as a coordinate, since they replaced it by the variable θ. In addition, the expansion in the usual small parameter $\epsilon := \sqrt{2\mu/B}\nabla := r_L\nabla$ was also present, but it did not correspond to an expansion in μ nor in B, but only in ∇, 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 ∇ and in the function B have only a dimensional meaning: the term of order ϵ^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 ϕ 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 [12], which considered only the minimal guiding-center reduction. It seems it had not been noticed in previous works on the full guiding-center reduction [2, 5, 6, 21].

When using the standard variable $v_\parallel$ instead of ϕ, the monomiality in $\sqrt{\mu}$ and $\sqrt{B}$ is not verified: for instance, in $\mathbf{G}_1^\theta$, the term of order ϕ^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 ϕ^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 [2, 6], for instance.

On the contrary, when using the variable ϕ, the polynomiality is verified, because of the structure of the Lagrangian, of the action of derivatives, especially (7), 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 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. ∇, $\mathbf{B}$ and $\mathbf{p}$. When using the momentum coordinates (ϕ, p, c) or (ϕ, μ, 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 μ 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 ϵ, 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 ϕ 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 $\frac{\mathbf{p}}{p}$.

C. 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 (20): 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 [18]) 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 ϵ 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 [18]) 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 ϵ. 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 [18].

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 [5] 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 [6] 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. [2] used a maximal reduction (at order $n = 1$) but with the second possibility. Recently, while improvements in the second order were addressed [18, 21], interest was renewed in the first possibility. The paper [21] 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 [18] 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 $\overline{p}_{\|}$ (or equivalently of $\overline{\phi}$).

The algorithm at higher order, with the condition (71) 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 ϵ: in formula (71), the term inside the bracket must have no overall contribution of order 0 in ϕ. This could explain why first papers addressing both the first and the second order in ϵ systematically used the Hamiltonian representation; the symplectic representation was used only later, when the condition was observed as verified.

The equivalent relation introduced in [18] 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 equation (17) in [18]) is the following. Start from the reduced Lagrangian written in symplectic representation (we use here the variable $p_\parallel$ instead of ϕ, in order to agree with the notation used in [18]):

$$\overline{H} = \mu B + \tfrac{\overline{p}_\parallel^2}{2}\,,$$
$$\overline{\Gamma}^{\mathsf{b}} = \overline{p}_\parallel + \sum_{n\geqslant 1}(\Pi_\parallel)_n\,.$$

Then, redefine the reduced parallel momentum so as to absorb all the series $\overline{\Gamma}^{\mathsf{b}}$ in it:

$$\overline{p}'_\parallel := \overline{p}_\parallel + \sum_{n=1}(\Pi_\parallel)_n\,. \tag{80}$$

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 (80):

$$\overline{p}_\parallel := \overline{p}'_\parallel - \sum_{n\geqslant 1}(\Pi_\parallel)_n = \overline{p}'_\parallel + \sum_{n\geqslant 1}(\Pi'_\parallel)_n\,, \tag{81}$$

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 ϵ, 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:

$$\overline{H} = \mu B + \tfrac{\overline{p}_\parallel^2}{2} + \sum_{n\geqslant 1} \overline{H}_n \,,$$
$$\overline{\Gamma}^{\mathsf{b}} = \overline{p}_\parallel \,. \tag{82}$$

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 $\overline{p}_\parallel$:

$$\overline{p}_\parallel'^2 := \overline{p}_\parallel^2 + 2\sum_{n\geqslant 1} \overline{H}_n \,. \tag{83}$$

With this coordinate, the Hamiltonian part of the Lagrangian is fully reduced:

$$\overline{H} = \mu B + \tfrac{\overline{p}_\parallel'^2}{2} \,.$$

To obtain the reduced symplectic Lagrangian with this coordinate, one just inverts the near-identity transformation (83) by writing it as:

$$\overline{p}_\parallel' := \pm\sqrt{\overline{p}_\parallel^2 + 2\sum_{n\geqslant 1} \overline{H}_n} = \overline{p}_\parallel \sqrt{1 + \tfrac{2\sum_{n\geqslant 1} \overline{H}_n}{\overline{p}_\parallel^2}} \,, \tag{84}$$

and then by expanding the term $\sqrt{1+\epsilon}$. This assumes that the ratio $\frac{2\sum_{n=1} \overline{H}_n}{\overline{p}_\parallel^2}$ 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 $\epsilon = r_L \nabla$, 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 (84). Thus, a division by $\overline{p}_\parallel$ can cause a singularity.

When going from the symplectic to the Hamiltonian representation, no such a division was needed, since the series inversion (81) just consisted in composing series. On the contrary, when starting from the Hamiltonian representation, the series inversion (84) involves a division by $\overline{p}_\parallel^2$. This causes a singularity if $\overline{H}_n$ contains a term of order 0 or 1 in $\overline{p}_\parallel$.

It is interesting to see that difficulties arise here at $\overline{p}_\parallel = 0$ (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 (71) is verified.

D. Gyro-gauge independence

The intrinsic formulation of the guiding-center reduction was motivated by questions about the traditional gyro-angle variable θ. The derivation with the coordinate 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 θ 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 θ only through the corresponding physical quantity 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 c.

From a mathematical point of view also, the use of the variable θ was not completely satisfactory, because the gyro-angle corresponds to a circle bundle [7, 11, 16]. The traditional coordinate θ makes this structure somehow disappear, because the manifold trivially becomes $\mathbb{R}^3 \times \mathbb{S}^1$. It is why the variable θ does not have a global existence in a general magnetic geometry [11]. On the contrary, when using the physical variable c, the circle bundle naturally arises: as c is defined on a space-dependent circle, spatial displacements imply a variation of c, so that a covariant derivative is involved, which encodes the circle-bundle geometry for the gyro-angle [11] 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 paper, and will be reported in [17].

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 c as with the scalar coordinate θ, 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 ∂_θ 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 θ considers the circle bundle trivial, whereas it should not. This implied to use more intrinsic definitions for the operations used, such as (50) and (51) 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 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 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 [5, 7, 11]); here, $\mathbf{R}_g(\mathbf{q}, \mathbf{p})$ can be any function of the position and momentum. Especially, the physical definition of 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 ϕ.

Other gauge-dependent quantities are interesting to consider. In previous works, the coordinate transformation $\theta \longrightarrow \overline{\theta}$ was gauge dependent (see for instance equation (30c) in [6], or in [3] the solution for the generator $\mathbf{G}_1^\zeta$ below equation (5.45)), as well as the definition of the coordinate θ, 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 $\nabla_* := \nabla + \mathbf{R}\partial_\theta$ [3]. 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 [5, 7].

In the gauge-independent approach, all the coordinates, including the gyro-angle 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 (56). 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. (75), for instance). This is no surprise, since the connection on the fibre 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 c, this choice can be used to simplify computations and to make them connection-independent.

E. 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 ∂_θ.

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 paper, as in previous works, only a part of $\mathbf{G}_2$ is explicitly computed, whereas the work [12] 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 [12]. 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 $(\dot{\overline{\mathbf{q}}}, \dot{\overline{\phi}})$.

V. 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 $\mathtt{c}$ 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 θ, and by making the coordinates directly induced by the physical state of the system.

For instance, the gauge removal introduced a vectorial quantity 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 θ, 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 θ, the connection fixing for the coordinate c depends not only on the position, but on the momentum as well. This is all the more convenient as the physical definition of 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 c is a way to tackle some of the questions involved in the guiding-center anholonomy and gauge- (or connection-) arbitrariness. These questions are outside the scope of this paper and will be reported elsewhere [17].

In this paper, 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 [14]. A next step will be to relate these formal expansions with asymptotic series, in a similar way as what was done in [25] for Kruskal's work, and what is beginning being done about Littlejohn's works [26]. Also, as usual in perturbation theory, convergence of the guiding-center series is an interesting question to investigate, probably with methods of accelerated convergence [27]. 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.

Acknowledgement

We acknowledge financial support from the Agence Nationale de la Recherche (ANR GYPSI). This work was also supported by the European Community under the contract of Association between EURATOM, CEA, and the French Research Federation for fusion study. The views and opinions expressed herein do not necessarily reflect those of the European Commission. The authors also acknowledge fruitful discussions with Alain Brizard, Phil Morrison, Mathieu Lutz and with the Équipe de Dynamique Nonlinéaire of the Centre de Physique Théorique of Marseille.

Appendix A: 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.

1. Fundamental ingredients

A) The goal is to solve equations (13) for the unknowns $\mathbf{G}_n$ and S_n, with the requirements (17), (18) and (20) identified above. Ideally, the maximal reduction sets $\overline{\Gamma}_n = 0$ for all higher orders n, as a result of (18).

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 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

$$\mathbf{G}_{n+1} = (\omega_{-1})^{-1}\cdot\big[\overline{\Gamma}_n - \mathsf{R}_n - dS_n\big] \,, \tag{A1}$$

provided ω_{-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 ω_{-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 $\overline{\Gamma}_n = 0$, equation (13) 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 ω_{-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 equation (13) writes

$$\mathbf{G}_1 \cdot \omega_{-1} = \overline{\Gamma}_0 - \Gamma_0 - dS_0 \,. \tag{A2}$$

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 (17) and if possible for the intermediate requirement (20); then, the reduction is possible, and the maximal requirement (18) 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 ω_{-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 $n = 1$, equation (13) writes

$$\mathbf{G}_2 \cdot \omega_{-1} + \tfrac{\mathbf{G}_1}{2} \cdot (\omega_0 + \overline{\omega}_0) = \overline{\Gamma}_1 - dS_1 \,. \tag{A3}$$

Now, the pivotal matrix M_1 to be inverted is the set of ω_{-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 ω_{-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 M_1 cannot be invertible on the unknown components of $\mathbf{G}_2$, because of the non-trivial kernel of ω_{-1}. In the same way as at zeroth order, this kernel has to be excluded when studying the invertibility of M_1, because it will be involved only at the following order.

Then, if the pivotal matrix 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 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 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 M_∞. 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: Equation (16) generically (i.e. at high orders) writes

$$\begin{aligned}\overline{\Gamma}_n = &\left[(\mathbf{G}_{n+1}\cdot d) + (\mathbf{G}_n\cdot d)(\mathbf{G}_1\cdot d) + ... + \tfrac{(\mathbf{G}_1\cdot d)^{n+1}}{n+1!}\right]\Gamma_{-1} \\ &+ \left[(\mathbf{G}_n\cdot d) + (\mathbf{G}_{n-1}\cdot d)(\mathbf{G}_1\cdot d) + ... + \tfrac{(\mathbf{G}_1\cdot d)^{n}}{n!}\right]\Gamma_0 + dS_n\,. \end{aligned} \tag{A4}$$

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 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

$$\begin{aligned}\overline{\Gamma}_n = \;& \mathsf{G}_{n+1}\Gamma_{-1} \\ &+ \mathsf{G}_n\big(\mathsf{G}_1\Gamma_{-1} + \Gamma_0\big) \\ &+ \mathsf{G}_{n-1}\left[\big(\mathsf{G}_2 + \tfrac{\mathsf{G}_1^2}{2}\big)\Gamma_{-1} + \mathsf{G}_1\Gamma_0\right] \\ &+ ... \\ &+ \mathsf{G}_1\left[\tfrac{(\mathsf{G}_1)^n}{n+1!}\Gamma_{-1} + \tfrac{(\mathsf{G}_1)^{n-1}}{n!}\Gamma_0\right] + dS_n\,. \end{aligned} \tag{A5}$$

Using (13) for the lowest-orders reduced Lagrangian $\overline{\Gamma}_k$, that are already known, the previous formula can be rewritten

$$\begin{aligned}\overline{\Gamma}_n &= \mathsf{G}_{n+1}\big(\overline{\Gamma}_{-1} - dS_{-1}\big) + \mathsf{G}_n\big(\overline{\Gamma}_0 - dS_0\big) + \mathsf{G}_{n-1}\big(\overline{\Gamma}_1 - dS_1\big) \\ &\qquad + ... + \mathsf{G}_1\left[\tfrac{(\mathsf{G}_1)^n}{n+1!}\Gamma_{-1} + \tfrac{(\mathsf{G}_1)^{n-1}}{n!}\Gamma_0\right] + dS_n \\ &= \mathbf{G}_{n+1}\cdot\overline{\omega}_{-1} + \mathbf{G}_n\cdot\overline{\omega}_0 + ... + dS_n + \mathsf{R}_n\,, \end{aligned} \tag{A6}$$

where R_n indicates all other terms, that are already known, since they do not involve $(\mathbf{G}_{n+1}, \mathbf{G}_n, ...)$.

Equation (A6) shows that the pivotal matrix 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 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 M_∞ 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 M_∞^{-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 formulae (A4)-(A6) 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 formula (A6) does not hold, which spoils the conclusion.

But this concerns only low orders. Formula (A5) 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 formula (A6)

$$\begin{aligned}\overline{\Gamma}_1 &= \mathsf{G}_2\Gamma_{-1} + \mathsf{G}_1\big(\tfrac{1}{2}\mathsf{G}_1\Gamma_{-1} + \Gamma_0\big) + dS_1 \\ &\neq \mathbf{G}_2{\cdot}\omega_{-1} + \mathbf{G}_1{\cdot}\overline{\omega}_0 + dS_1\,,\end{aligned}$$

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 (A3). However, if some of the components of $\mathbf{G}_1$ are still not determined at that point (this is fairly general as ω_{-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 equation (A3) must be let under its initial form (13)

$$\overline{\Gamma}_1 = \Big(\mathbf{G}_2 + \tfrac{\mathbf{G}_1{\cdot}d}{2}\mathbf{G}_1\Big){\cdot}\omega_{-1} + \mathbf{G}_1{\cdot}\omega_0 + dS_1\,,$$

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 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 M_∞ is completely general and comes because M_∞ is antisymmetric. It is not invertible on the 7-dimensional space $(\mathbf{q}, \mathbf{p}, t)$, and can be invertible only on a sub-space, e.g. on the 6-dimensional phase-space $(\mathbf{q}, \mathbf{p})$. For a symplectic Hamiltonian system, the high-orders pivotal matrix M_∞ is indeed invertible when restricted to the phase-space, since the Lagrange 2-form ω_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 ∂_{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 M_∞, the induction relation for high orders $n \geqslant n_c$ just relies on a matrix inverse M_∞^{-1}. Then, the basic picture of formula (A1) becomes efficient: equations (13) are solved at arbitrary order through a formula completely analogous to (A1), 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.

2. 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 M_∞. At the end $n = n_b$, the invertible high-order pivotal matrix M_∞ 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 M_∞. 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 (A1), the pivotal matrix must be extended to include the gauge function S_n, and some coefficients of the inverted matrix M_∞^{-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 M_∞ involves only $\overline{\varpi}_{-1}$ and $\overline{\varpi}_0$. Then, the equation (13) or (A6) writes

$$\overline{\Gamma}_n = \mathbf{G}_{n+1}\cdot\overline{\varpi}_{-1} + \mathbf{G}_n\cdot\overline{\varpi}_0 + dS_n + \mathsf{R}_n\,, \tag{A7}$$

where R_n indicates all terms of (13), 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

$$\mathbf{g}_n := \big(\mathbf{G}_{n+1}, \mathbf{G}_n, S_n\big)\,. \tag{A8}$$

The pivotal matrix M_n is then extended to act on all $\mathbf{g}_n$ (including the gauge function) in (A7) and is defined by

$$\mathsf{M}_n\cdot\mathbf{g}_n := \mathbf{G}_{n+1}\cdot\overline{\varpi}_{-1} + \mathbf{G}_n\cdot\overline{\varpi}_0 + dS_n\,.$$

As announced, some of its coefficients (the ones acting on the component S_n) are differential operators. With these conventions, equation (A7) 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 ∞:

$$\mathbf{g} = \big((\mathbf{g}_n)_a; (\mathbf{g}_n)_\infty\big) \,,$$

with $(\mathbf{g}_n)_a$ fully identified and all terms of $(\mathbf{g}_n)_\infty$ fully unknown. The Lagrangian writes

$$\overline{\Gamma}_n = (\mathsf{M}_n)_a \cdot (\mathbf{g}_n)_a + (\mathsf{M}_n)_\infty \cdot (\mathbf{g}_n)_\infty + \mathsf{R}_n \,, \tag{A9}$$

with obvious definitions for the linear operators $(\mathsf{M}_n)_a$ and $(\mathsf{M}_n)_\infty$. By assumption, the quantities 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

$$(\mathbf{g}_n)_\infty = \mathsf{M}_\infty^{-1} \cdot \left[\overline{\Gamma}_n - (\mathsf{M}_n)_a \cdot (\mathbf{g}_n)_a - \mathsf{R}_n\right] \,. \tag{A10}$$

It is explicit and makes the basic picture (A1) apply to all orders $n \geqslant n_c$. Some coefficients of M_∞^{-1} are integral operators, since in the inverse matrix M_∞ 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 ω_{-1} and will be determined at the next order; this is well illustrated by (61) and (65)-(66). 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 R_n (see formula (63)), 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 (61) and (65)-(66).

The final algorithm to be iterated for the n-th-order term is trivial: in formula (A10), replace the lowest orders terms by their expression, already known, then compute the Lie derivatives involved in the term R_n, and last apply the matrix product with M_∞^{-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 (A1) and (A10) 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 [12].

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 (17), which means to put to zero the gyro-fluctuating part of the Lagrangian; and the averaged part is imposed by the other requirements (18) and (20), 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 equation (A7): $\mathtt{avg}(\Gamma_n^\mu)$ remains as a parameter in the right-hand side of (A10).

Furthermore, the integrability conditions on S_n cannot be fully satisfied, one of the optional requirements (18) 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 (A8) and (A9), by including in this vector only the unknowns that are involved at that order (see formula (61)), and by distinguishing between the unknowns that will be identified and the ones that will remain parameters (see formulae (65)-(66)).

References

[1] T. G. Northrop, "The adiabatic motion of charged particles", Wiley, New York (1963).

[2] J. R. Cary and A. J. Brizard, Rev. Mod. Phys. **81**, 693 (2009).

[3] A. J. Brizard and T. S. Hahm, Rev. Mod. Phys. **79**, 421 (2007).

[4] R. G. Littlejohn, J. Math. Phys. **20**, 2445 (1979).

[5] R. G. Littlejohn, Phys. Fluids **24**, 1730 (1981).

[6] R. G. Littlejohn, J. Plasm. Phys. **29**, 111 (1983).

[7] R. G. Littlejohn, Phys. Rev. A **38**, 6034 (1988).

[8] L. E. Sugiyama, Phys. Plasmas **15**, 092112 (2008).

[9] J. A. Krommes, Phys. Plasmas **16**, 084701 (2009).

[10] L. E. Sugiyama, Phys. Plasmas **16**, 084702 (2009).

[11] J. W. Burby and H. Qin, Phys. Plasmas **19**, 052106 (2012).

[12] L. de Guillebon and M. Vittot, Phys. Plasmas **20**, 082505 (2013).

[13] J. R. Cary, Phys. Reports **79**, 129 (1981).

[14] R. G. Littlejohn, J. Math. Phys. **23**, 742 (1982).

[15] 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.

[16] M. Kruskal, J. Math. Phys. **3**, 806 (1962).

[17] L. de Guillebon and M. Vittot, "On an intrinsic formulation of the guiding-center anholonomy and gyro-gauge arbitrariness", submitted, arXiv:1304.4563 (2013).

[18] "Equivalent representations of higher-order Hamiltonian guiding-center theory", A. J. Brizard and N. Tronko, arXiv:1205.5772 (2012).

[19] J. W. Burby, J. Squire, and H. Qin, Phys. Plasmas **20**, 072105 (2013).

[20] F. I. Parra and P. J. Catto, Phys. Plasmas **17**, 056106 (2010).

[21] F. I. Parra and I. Calvo, Plasma Phys. Control. Fusion **53**, 045001 (2011).

[22] J. Liu and H. Qin, Phys. Plasmas **18**, 072505 (2011).

[23] A. J. Brizard and L. de Guillebon, Phys. Plasm. **19**, 094701 (2012).

[24] A. J. Brizard, private communication (2011).

[25] J. Berkowitz and C. S. Gardner, Institute of mathematical sciences, Magneto-fluid dynamics division, NYO-7975, New-York University (1957).

[26] E. Frenod and M. Lutz, "On the Geometrical Gyro-Kinetic Theory", arXiv:1306.5639 (2013).

[27] N. N. Bogoliubov, J. A. Mitropolski, and A. M. Samoilenko, *Methods of accelerated convergence in non-linear mechanics*, Springer-Verlag, Berlin, 1976.