Dieter Pfirsch and Darío Correa-Restrepo

New method of deriving local energy and momentum-conserving Maxwell-collisionless drift-kinetic and gyrokinetic theories: basic theory

New method of deriving local energy and momentum-conserving Maxwell-collisionless drift-kinetic and gyrokinetic theories: basic theory

Dieter Pfirsch and Darío Correa-Restrepo

Max-Planck-Institut für Plasmaphysik, EURATOM Association, D-85748 Garching, Germany

(dcr@ipp.mpg.de)

Abstract

In this paper we describe a relatively simple and transparent method of obtaining collisionless drift-kinetic, gyrokinetic and more general theories, including local charge, energy and momentum conservation laws. An important feature of the new formalism is, contrary to present-day theories, the exact gauge invariance, thus avoiding certain inconsistencies. The present paper starts with the introduction and proof of the correctness of a Lagrangian for combined Maxwell-kinetic theories in general coordinates as concerns the particle motion. The kinetic part of it is formulated in Eulerian form by means of the equations of motion in the form of Hamilton-Jacobi's equation, used only as a tool, and Dirac's constraint theory. Charge and current densities automatically distinguish between "particle-like" (guiding-centre), polarization and magnetization contributions. This formalism is applied to averaging coordinates derived by a method similar to Kruskal's. Certain properties of the averaging coordinates, according to the basic requirements imposed on them, can be used to obtain a gyro-angle independent Lagrangian, from which one can obtain a Lagrangian for the combined Maxwell-kinetic theories in a reduced phase space that is applicable to situations in which one is not interested in the dependence on some kind of gyroangle describing the gyromotion, whose treatment can, however, easily be added. The basic perturbation theory, which aims at obtaining averaging phase-space coordinates, is done solely within the Kruskal formalism in which only the electric and magnetic fields appear, but not the corresponding potentials. This formalism provides, in particular, information about the allowed amplitudes of fluctuations. The results are later used to obtain approximate expressions for the Lagrange functions in the drift-kinetic and the gyrokinetic ordering. For the defi-

nition of certain approximations to the exact Lagrangian the central principle is the exact gauge invariance. The terms of the zeroth and first orders needed for such an approximation are given. For the drift-kinetic ordering Littlejohn's Lagrangian is readily rederived and hence the drift-kinetic theory as obtained and investigated by the present authors in some previous work. Conservation laws and their mathematical structures corresponding to the gyroangle-independent Lagrangian are obtained in a following paper. There, in particular, a detailed description of how to obtain variations of gyroangle averaged quantities in the reduced phase space is given, and it is explained how these variations are used in a modified form of Noether's theory which observes exact gauge invariance.

Contents

1 Introduction

In this paper we describe a relatively simple and transparent method of obtaining collisionless drift-kinetic, gyrokinetic and,in principle, more general theories together with local conservation laws by applying a Noether-like formalism. In particular, it presents a rather difficult subject from a point of view different from that of the standard theories and offers an alternative to the usual systematic treatments based on pseudo-canonical theory and Lie-transform theory [1–6]. The very specific feature of our method is the Eulerian - as opposed to Lagrangian - formulation of the Lagrangians for the kinetic equations which are the appropriate base for deriving not only total but also local conservation laws, first introduced by Pfirsch [7], combined with exact gauge invariance. The considered Lagrangians L_K are correspondingly based on a Eulerian instead of a Lagrangian description of the particle motion. Eulerian Lagrangians for the linearized gyrokinetic theory were already obtained by Brizard [8,9] employing the methods of Pfirsch [7] and Pfirsch and Morrison [10,11], and for the nonlinear gyrokinetic theory by means of a variational principle with constrained variations [12], which are performed according to a rule which is equivalent to the principle of dynamical accessibility, as described e.g. in [11,13–15]. Contrary to present-day theories is the observance of exact gauge invariance, thus avoiding certain inconsistencies which can lead to incorrect expressions for the energy density and the energy flux density, and to an erroneous interpretation of the contribution of certain terms as resulting from the polarization current.

Together with the Lagrangian for Maxwell's equations,

$$L_{\mathrm{M}} = \int d^3x \, \frac{1}{8\pi} \left[\mathbf{E}^2 \left(\mathbf{x}, t \right) - \mathbf{B}^2 \left(\mathbf{x}, t \right) \right], \tag{1}$$

the Lagrangians L_{K} form a total Lagrangian,

$$L_{\mathrm{tot}} \left(t \right) = L_{\mathrm{M}} + L_{\mathrm{K}}, \tag{2}$$

from which one obtains, via Hamilton's principle, the kinetic and inhomogeneous Maxwell equations. The latter appear directly in the form

$$\frac{c}{4\pi} \boldsymbol{\nabla} \times \mathbf{B} - \frac{1}{4\pi} \frac{\partial \mathbf{E}}{\partial t} = \mathbf{j}_{\mathrm{particle-like}} + c \boldsymbol{\nabla} \times \mathbf{M} + \frac{\partial \mathbf{P}}{\partial t}, \tag{3}$$

$$\frac{1}{4\pi} \boldsymbol{\nabla} \cdot \mathbf{E} = \rho_{\mathrm{particle-like}} - \boldsymbol{\nabla} \cdot \mathbf{P}, \tag{4}$$

where "particle-like" refers to entities which can and will be chosen as the guiding centres or the gyrocentres, and $\mathbf{M}$ is the magnetization and $\mathbf{P}$ is the polarization connected with these "quasi-particles" . We consider it an important feature of the new method that it automatically yields this kind of representation, unlike the standard methods.

In previous publications of the authors [7,10,11,16] the drift-kinetic and Vlasov theories were already treated in part by the same method, where for the drift-kinetic theory Littlejohn's Lagrangian [3] had to be used. This Lagrangian is now also obtained, and in a very simple way, too, as shown in Sec. 5.4.1.
The method of obtaining the stated Lagrangians for the kinetic equations consists in a series of rather simple steps:

1. Obtain for each particle species ν a transformation $\mathbf{x}(\boldsymbol{\zeta}, t)$, $\mathbf{p}(\boldsymbol{\zeta}, t)$, $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_6)$, from the phase-space variables $\mathbf{x}$, $\mathbf{p}$ to "quasi-particle" variables $(\zeta_1, \ldots, \zeta_4) = (z_1, \ldots, z_4)$, an adiabatic invariant $\zeta_5 = J$ and a gyroangle-like variable $\zeta_6 = \theta$ by applying Kruskal's asymptotic theory of systems with all solutions nearly periodic [17].

2. Use these transformations for transforming the original phase-space Lagrangians $L_\nu(\mathbf{x}, \mathbf{p}, t)$ for each particle species. The transformed Lagrangians are of a non-standard type, which means that Hamilton's equations with a Hamiltonian obtained in the usual way from such a Lagrangian are not identical with the original equations of motion.

3. Obtain, by applying Dirac's "constraint theory" [18, 19], Hamiltonians $H_{\mathrm{D}\nu}$ in an extended phase space which lead, together with certain "constraints", to the correct equations of motion.

4. Use these Hamiltonians in Hamilton-Jacobi equations for the Hamilton-Jacobi functions $S_\nu(\boldsymbol{\zeta}, \boldsymbol{\alpha}, t)$. These equations are the Eulerian forms of the equations of motion.

   The Hamilton-Jacobi equations, however, only play the role of a tool; the Hamilton-Jacobi functions S_ν will not appear in the final expressions for the kinetic equations and the various macroscopic densities, i.e the densities in configuration space. This holds for the nonlinear theory which we present here. For linearized theories it will be advantageous to use a modified version of the Hamilton-Jacobi equations which yields generating functions for the transformation of the perturbed orbits to the unperturbed ones [8, 9, 11]. In this case the first-order generating functions $S_{\nu 1}$ play a central role, particularly in the second-order energy.

5. Obtain the primary form of the Lagrangians for the kinetic equations as

$$\sum_\nu \int d^6\zeta \, d^6\alpha \, f_{\mathrm{p}\nu} \left(\frac{\partial S_\nu}{\partial t} + H_{\mathrm{D}\nu}(\boldsymbol{\zeta}, \frac{\partial S_\nu}{\partial \boldsymbol{\zeta}}, t) \right), \tag{5}$$

   where $f_{\mathrm{p}\nu} = f_{\mathrm{p}\nu}(\boldsymbol{\zeta}, \boldsymbol{\alpha}, t)$ are the primary forms of the distribution functions; the final form of the distribution functions will be denoted by f_ν. The quantities to be varied in Hamilton's principle are $f_{\mathrm{p}\nu}$, S_ν, and the vector and scalar potentials $\mathbf{A}$ and Φ, on which the $H_{\mathrm{D}\nu}$'s depend somehow.

6

6. Prove that this Lagrangian is equivalent to the Vlasov theory

7. Eliminate the S_ν's and $\boldsymbol{\alpha}$ in all final expressions such as the kinetic equations and charge and current densities.

Kruskal's theory (point (a) above), and also a certain modification of it which has to be used here, is based on a formal infinite power series expansion in a certain small parameter ϵ. In the present context it is concerned with the single-particle motion in given electric and magnetic fields. For the drift-kinetic theory ϵ is given by the ratio of the gyroperiod and the time scale for these fields and it is required that the ratio of the gyroradius r_{gyro} to the length scale L for these fields also be of the order of this ϵ. For the gyrokinetic theory this requirement concerns only the "background" fields. There may be in addition fluctuations around these fields with length scales of the order of the gyroradius, but the amplitudes of these fluctuations in some units must be of the order ϵ (details are given in Sec. 3). This restriction, which is an element of all systematic present-day treatments, already arises here within the simple Kruskal formalism in a straightforward way.

Kruskal's method is neither Hamiltonian nor pseudo-Hamiltonian. It also applies to non-Hamiltonian systems. It yields a transformation to new dependent variables implying a reduction of the number of equations. The omitted variable is essentially an angle variable ϕ. Correspondingly, one obtains for Hamiltonian systems an adiabatic invariant. The method is straightforward to any order in ϵ. There is, in particular, in contrast to present-day theories, no need to solve any differential equations.

The basic modification of Kruskal's theory needed here is that the old variables are obtained as functions of the new ones, instead of the other way around as in Kruskal's original paper. Such a modification was first introduced by Larsson within the framework of guiding-centre theory [20].

Another modification is needed for the gyrokinetic scaling. These modifications, however, do not change the characteristics of Kruskal's theory.

Since Kruskal's theory is a complete theory for the particle motion, this theory could be used directly for solving Vlasov's equation and obtaining the charge and current densities for Maxwell's equations. This would, however, have the disadvantage that, when the expansion is only done to a finite order in ϵ, this kind of theory in general lacks, within the approximation made, conservation of energy, momentum and angular momentum. This deficit is avoided in the present-day theories. In our paper it is avoided by incorporating Kruskal's theory for any finite approximation in ϵ in a Eulerian form of a Lagrangian for the combined Maxwell-kinetic theories (points (b) to (e)). This method also allows one to keep background and fluctuating fields together, which is needed when performing variations, although the fluctuating fields must be small in order to allow an expansion according to Kruskal's formalism. Since this Lagrangian is of finite order in ϵ only as concerns the Kruskal theory, its variation with respect to, in particular, the scalar and vector potentials Φ

7

and **A** yields expressions for the charge and current densities which guarantee the various conservation laws for any finite approximation in ϵ. As already shown in [7], these expressions contain particle-like (guiding-centre) as well as polarization and magnetization contributions corresponding to the rotational motion in these particle-like entities [see (3) and (4)].

It is emphasized that the finite approximation in ϵ concerns only the Kruskal part, which implies only the electric and magnetic fields **E** and **B** but not Φ and **A**. In this way, the somewhat doubtful attribution of certain orders in ϵ to Φ and **A**, as done in some present-day theories, is avoided. Such an attribution could be meaningful only for specific gauges. However, instead of doing this, gauge invariance is used as a keynote. A particular result of not attributing A and Φ a certain order is that in the ϵ^{-1} order *of the Kruskal formalism* together with the requirement of gauge invariance, only the gauge-invariant Lagrangian $L = (e/c)\dot{\mathbf{x}} \cdot \mathbf{A}(\mathbf{x}, t) - e\Phi(\mathbf{x}, t)$ is left, leading to a motion parallel to **B** with superimposed $\mathbf{E} \times \mathbf{B}$ drift, according to $\mathbf{E} + (\mathbf{v} \times \mathbf{B})/c = 0$ for the drift-kinetic case, or $(\mathbf{v} \times \mathbf{B})/c = 0$ for the gyrokinetic case. For the purposes of this paper, only the very simple lowest-order contribution and some readily calculated first-order quantities from Kruskal's theory are needed. We would like to mention here also that, contrary to the usual treatments $[2, 3, 5, 12, 21]$, our description of the guiding-centre and gyrocentre dynamics does not require the inversion of the Lagrange matrix in order to obtain the Poisson brackets. We treat this subject in detail in [22]. Also different from the usual theories is our method of obtaining a gyroangle-independent Lagrangian for guiding centres or gyrocentres. Our procedure is based, from the beginning, on Kruskal's method of constructing averaging coordinates whose time evolution is independent of the gyroangle to any order in the expansion parameter ϵ. The particle Lagrangian, expressed in the averaging coordinates, is exact and thus describes the exact particle motion. Although the Euler equations corresponding to this Lagrangian are independent of the gyroangle, the Lagrangian itself usually is not. The dependence of this Lagrangian on the gyroangle is eliminated here by a convenient choice of a gauge function in the phase space of the averaging coordinates, *and not by a change of phase-space coordinates*. The result is an *exact, manifestly gyroangle-independent* Lagrangian. In contrast to this is the usual method introduced by Brizard, e.g. in [5, 12], of obtaining gyrocentre Lagrangians. This method is based on the techniques used in guiding-centre theory by Littlejohn, who preferred to work directly on the particle phase-space Lagrangian; in [21], and also in a detailed treatment concerning a different theory [23], the elimination of the gyroangle dependence from a certain Lagrangian is achieved by performing, order by order in the parameter of smallness, an appropriate transformation of coordinates in phase space, also allowing for a gauge transformation in that space.

As it appears convenient for the particular aspect of the problem being investigated, the description of the particle motion is based on three alternative pictures using either the usual canonical Hamiltonian for particles in an elec-

tromagnetic field, viz.

$$H\left(\mathbf{x}, \mathbf{p}, t\right) = \frac{1}{2m} \left[\mathbf{p} - \frac{e}{c} \mathbf{A}\left(\mathbf{x}, t\right)\right]^2 + e\,\Phi\left(\mathbf{x}, t\right), \tag{6}$$

or Newton's equations of motion as given by (66) below, or the Hamilton-Jacobi equations as done in Sec. 2.4.

The paper is organized as follows: Section 2 concerns the above points (a) to (g) for general transformations $\mathbf{x}(\boldsymbol{\zeta}, t)$, $\mathbf{p}(\boldsymbol{\zeta}, t)$, i.e. not necessarily for a transformation to averaging coordinates. Section 3 presents a reformulation of the equations of motion suitable for application of Kruskal's theory and discussion of the conditions to be satisfied as concerns orders in ϵ. Section 4 presents Kruskal's modified theory. In Sec. 5 we derive a manifestly gyrophase independent Lagrangian corresponding to Kruskal's modified theory for any order in ϵ. In Sec. 6 we obtain the Eulerian form of a Lagrangian for the combined Maxwell-kinetic theories with averaging coordinates based on Kruskal's theory and formulate the whole Maxwell-kinetic theory. The summary is given in Sec. 7.

In Appendix A, direct proof is given that the quantity J of Kruskal's theory is indeed a constant of the motion. In Appendix B, the Lagrange brackets needed for the derivation of a manifestly gyrophase independent phase-space Lagrangian are calculated. In Appendix C, an alternative derivation to that given in the main text of a Liouvillian volume element is presented. Finally, in Appendix D, a concise derivation of a Liouvillian volume element for general kinetic theories is given.

2   Lagrangians for kinetic equations in general coordinates

In the following the index ν for the particle species is suppressed.

2.1   Phase-space Lagrangian

The starting point is the phase-space Lagrangian for each particle species,

$$L\left(\mathbf{x}, \mathbf{p}; \dot{\mathbf{x}}; t\right) = \dot{\mathbf{x}} \cdot \mathbf{p} - H\left(\mathbf{x}, \mathbf{p}, t\right), \tag{7}$$

with the Hamiltonian H given by (6). The quantities to be independently varied in the corresponding action integral are $\mathbf{x}$ and $\mathbf{p}$.

2.2   Transformation to new variables $\zeta_1, \ldots, \zeta_6; t$

With

$$\boldsymbol{\zeta} = \left(\zeta_1, \ldots, \zeta_6\right), \ \mathbf{x} = \mathbf{x}\left(\boldsymbol{\zeta}, t\right), \ \mathbf{p} = \mathbf{p}\left(\boldsymbol{\zeta}, t\right), \tag{8}$$

9

where these functions can depend somehow on the electric and magnetic fields, one obtains

$$L\left(\boldsymbol{\zeta}, \dot{\boldsymbol{\zeta}}, t\right) = \dot{\boldsymbol{\zeta}} \cdot \boldsymbol{\gamma} - H_{\mathrm{p}}, \tag{9}$$

with

$$\boldsymbol{\gamma}\left(\boldsymbol{\zeta}, t\right) = \frac{\partial \mathbf{x}}{\partial \boldsymbol{\zeta}} \cdot \mathbf{p}, \tag{10}$$

and

$$H_{\mathrm{p}}\left(\boldsymbol{\zeta}, t\right) = H - \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial t}. \tag{11}$$

The quantities to be varied independently in Hamilton's principle are the ζ_i's, while $\delta\dot{\zeta}_i$ follows via partial integration over t in the action integral, which yields the usual form of the Euler-Lagrange equations. With

$$\frac{\partial L}{\partial \boldsymbol{\zeta}} = \frac{\partial \boldsymbol{\gamma}}{\partial \boldsymbol{\zeta}} \cdot \dot{\boldsymbol{\zeta}} - \frac{\partial H_{\mathrm{p}}}{\partial \boldsymbol{\zeta}} \tag{12}$$

and

$$\frac{\partial L}{\partial \dot{\boldsymbol{\zeta}}} = \boldsymbol{\gamma} \quad , \quad \frac{d\boldsymbol{\gamma}}{dt} = \frac{\partial \boldsymbol{\gamma}}{\partial t} + \dot{\boldsymbol{\zeta}} \cdot \frac{\partial \boldsymbol{\gamma}}{\partial \boldsymbol{\zeta}}, \tag{13}$$

the Euler-Lagrange equations become

$$\frac{\partial \boldsymbol{\gamma}}{\partial \boldsymbol{\zeta}} \cdot \dot{\boldsymbol{\zeta}} - \dot{\boldsymbol{\zeta}} \cdot \frac{\partial \boldsymbol{\gamma}}{\partial \boldsymbol{\zeta}} - \frac{\partial \boldsymbol{\gamma}}{\partial t} - \frac{\partial H_{\mathrm{p}}}{\partial \boldsymbol{\zeta}} = 0, \tag{14}$$

or, equivalently,

$$\sum_{i=1}^{6} \dot{\zeta}_i \left[\frac{\partial \gamma_i}{\partial \zeta_k} - \frac{\partial \gamma_k}{\partial \zeta_i} \right] - \frac{\partial \gamma_k}{\partial t} - \frac{\partial H_{\mathrm{p}}}{\partial \zeta_k} = 0, \tag{15}$$

with solutions

$$\dot{\zeta}_i = V_i\left(\zeta_1, \ldots, \zeta_6; t\right). \tag{16}$$

The momenta canonically conjugated to the ζ_i's are

$$P_i = \frac{\partial L}{\partial \dot{\zeta}_i} = \gamma_i\left(\zeta_1, \ldots, \zeta_6, t\right). \tag{17}$$

The equations for the P_i's have the meaning of "constraints" in the sense that the P_i's are completely determined by the ζ_i's, and have no degrees of freedom of their own. Therefore, the $\dot{\zeta}_i$'s cannot be expressed by the P_i's as usually possible [24]. Lagrangians of the above type belong to the class of non-standard Lagrangians, as Dirac named them.

2.3 Dirac's Hamiltonian

The usual Hamiltonian method applied to the Lagrangian (9) to derive a Hamiltonian from a Lagrangian yields the "primary" Hamiltonian

$$\sum_{i=1}^{6} \dot{\zeta}_i P_i - L\left(\boldsymbol{\zeta}, \dot{\boldsymbol{\zeta}}, t\right) = H_{\mathrm{p}}\left(\boldsymbol{\zeta}, t\right). \tag{18}$$

As mentioned in the Introduction, Hamilton's equations with this Hamiltonian are not identical with the equations of motion, (15). Following Dirac one can, however, define a Hamiltonian H_{D} in an extended phase space which, however, possesses a manifold of solutions twice as large as that following from the original Lagrangian, but reducing again to the original manifold when imposing the constraints. For the present case Dirac's Hamiltonian is

$$H_{\mathrm{D}}\left(\zeta_i, P_i, t\right) = H_{\mathrm{p}}\left(\boldsymbol{\zeta}, t\right) + \sum_{i=0}^{6} V_i \left(P_i - \gamma_i\right). \tag{19}$$

with V_i as defined in (15) and (16). The corresponding Hamilton equations are

$$\dot{\zeta}_k = \frac{\partial H_{\mathrm{D}}}{\partial P_k} = V_k, \tag{20}$$

these being the same as the original equations of motion, (16). But there are the following additional equations:

$$
\begin{aligned}
\dot{P}_k &= -\frac{\partial H_{\mathrm{D}}}{\partial \zeta_k} \\
&= -\frac{\partial H_{\mathrm{p}}}{\partial \zeta_k} - \sum_{i=0}^{6} \frac{\partial V_i}{\partial \zeta_k} \left(P_i - \gamma_i\right) + \sum_{i=0}^{6} V_i \frac{\partial \gamma_i}{\partial \zeta_k}.
\end{aligned} \tag{21}
$$

These equations are now proved to be solved by the constraints (17):
With

$$
\begin{aligned}
\dot{\gamma}_k &= \frac{\partial \gamma_k}{\partial t} + \sum_{i=0}^{6} V_i \frac{\partial \gamma_k}{\partial \zeta_i} \\
&= -\frac{\partial H_{\mathrm{p}}}{\partial \zeta_k} + \sum_{i=0}^{6} V_i \frac{\partial \gamma_i}{\partial \zeta_k},
\end{aligned} \tag{22}
$$

where (15) and (16) have been used, one obtains

$$\dot{P}_k - \dot{\gamma}_k = -\sum_{i=0}^{6} \frac{\partial V_i}{\partial \zeta_k} \left(P_i - \gamma_i\right). \tag{23}$$

These equations have as special solution the constraints $P_k = \gamma_k$. Hence, Dirac's Hamiltonian allows one to obtain the correct equations of motion including the constraints.
There is, however, an important point to be taken into account later. The combinations $(P_k - \gamma_k)$ are not in general constants of motion, and $(P_k - \gamma_k) = 0$ are only special values.

2.4 The Lagrangian for a kinetic equation

As mentioned in the Introduction, this Lagrangian will be based on the Hamilton-Jacobi equation with H_{D}:

$$\frac{\partial S}{\partial t} + H_{\mathrm{D}}\left(\frac{\partial S}{\partial \zeta_i}, \zeta_i, t\right) = 0, \quad S = S(\zeta_i, \alpha_i, t), \quad P_i = \frac{\partial S}{\partial \zeta_i}. \tag{24}$$

The variables α_i are constants of integration for a so-called complete solution. The meaning of the Hamilton-Jacobi equation is briefly recalled. It is the equation for a generating function for a canonical transformation from the old variables ζ_i and momenta P_i to the new variables $\beta_i = \partial S/\partial \alpha_i$ and momenta α_i, and a new Hamiltonian which vanishes identically. The latter has the consequence that α_i , β_i are constants of motion. "Complete" solution means that α_i, β_i form a complete set of constants of motion.

It is now claimed that the Lagrangian $L_{\mathrm{K}}(t)$ for a kinetic equation is given by

$$\begin{aligned} L_{\mathrm{K}} \;=\; &- \sum_{\text{particle species}} \int d^6\zeta \, d^6\alpha \, f_{\mathrm{p}}\left(\zeta_i, \alpha_i, t\right) \left[\frac{\partial S}{\partial t}\left(\zeta_i, \alpha_i, t\right) + H_{\mathrm{p}}\left(\zeta_i, t\right)\right. \\ &+ \left.\sum_{j=1}^{6} V_j\left(\zeta_i, t\right)\left(\frac{\partial S}{\partial \zeta_j}\left(\zeta_i, \alpha_i, t\right) - \gamma_j\left(\zeta_i, t\right)\right)\right]. \end{aligned} \tag{25}$$

Proof of the correctness of this Lagrangian will evolve from the Euler-Lagrange equations for the combined Maxwell-kinetic theory.

With all variations assumed to vanish on the boundaries of the total integration space, including time, variation of the action integral $\mathcal{A} = \int_{t_1}^{t_2} L_{\mathrm{tot}} dt$ corresponding to the Lagrangian (2) (with relation(1)) yields by variation of f_{p}, , S, , Φ and $\mathbf{A}$, respectively,

$$f_{\mathrm{p}} \quad : \qquad \frac{\partial S}{\partial t}\left(\zeta_i, \alpha_i, t\right) + H_{\mathrm{p}}\left(\zeta_i, t\right)$$
$$+ \sum_{j=1}^{6} V_j\left(\zeta_i, t\right)\left[\frac{\partial S}{\partial \zeta_j}\left(\zeta_i, \alpha_i, t\right) - \gamma_j\left(\zeta_i, t\right)\right] = 0, \tag{26}$$

$$S \quad : \qquad \frac{\partial f_{\mathrm{p}}}{\partial t} + \sum_{j=1}^{6} \frac{\partial}{\partial \zeta_j}\left(V_j f_{\mathrm{p}}\right) = 0, \tag{27}$$

$$\Phi \quad : \qquad \frac{1}{4\pi}\boldsymbol{\nabla}\cdot\mathbf{E} = -\sum_{\text{particle species}} \frac{\delta \mathcal{A}_{\mathrm{K}}}{\delta \Phi\left(\mathbf{x}, t\right)}, \tag{28}$$

$$\mathbf{A} \quad : \qquad \frac{1}{4\pi}\left[\boldsymbol{\nabla}\times\mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}\right] = \sum_{\text{particle species}} \frac{\delta \mathcal{A}_{\mathrm{K}}}{\delta \mathbf{A}\left(\mathbf{x}, t\right)}, \tag{29}$$

where $\mathcal{A}_{\mathrm{K}} = \int_{t_1}^{t_2} L_{\mathrm{K}} dt$. Equation (26) is the Hamilton-Jacobi equation for the particle motion. Equation (27) is a continuity equation along the particle orbits (note that the α_i's are constants of motion). It will later become the kinetic

12

equation. Equations (28) and (29) represent the charge and current densities, respectively. They will later become, in particular, the usual equations for these quantities in Vlasov's theory.

A first step in the direction of usual equations is to replace the function $f_{\mathrm{p}}\left(\zeta_i, \alpha_i, t\right)$ by a function which, like the distribution function, is a constant of motion. This is accomplished by expressing the number of particles $dN_{\zeta,\alpha}$ in a comoving volume element $d^6\zeta d^6\alpha$, such that $dN_{\zeta,\alpha}$ is a constant of motion, by

$$dN_{\zeta,\alpha} = f_{\mathrm{p}}\left(\zeta_i, \alpha_i, t\right) d^6\zeta d^6\alpha = w\left(\zeta_i, \alpha_i, t\right) \bar{f}\left(\zeta_i, \alpha_i, t\right) d^6\zeta d^6\alpha, \qquad (30)$$

where

$$w\left(\zeta_i, \alpha_i, t\right) = \left\| \frac{\partial^2 S}{\partial \alpha_i \partial \zeta_k} \right\| \qquad (31)$$

is the Van Vleck determinant. This function has the property of solving the continuity equation, (27) [25, 26]. Hence, the function $\bar{f}\left(\zeta_i, \alpha_i, t\right)$ is a constant of motion and therefore a solution of the equation

$$\frac{\partial \bar{f}}{\partial t} + \sum_{i=1}^{6} V_i \frac{\partial \bar{f}}{\partial \zeta_i} = 0. \qquad (32)$$

The next step is to eliminate the new momenta α_i by again introducing the P_i's with the help of $P_i = \partial S/\partial \zeta_i$. This means that in the integrals over the α_i's one has to replace

$$d^6\alpha \to \left\| \frac{\partial^2 S}{\partial \alpha_i \partial \zeta_k} \right\|^{-1} d^6P = w^{-1} d^6P, \qquad (33)$$

and, as a consequence of this,

$$dN_{\zeta,\alpha} = f_{\mathrm{p}}\left(\zeta_i, \alpha_i, t\right) d^6\zeta d^6\alpha \to$$

$$dN_{\zeta,P} = \bar{f}\left(\zeta_i, \alpha_i(\zeta_i, P_i, t), t\right) d^6\zeta d^6P = \hat{\bar{f}}\left(\zeta_i, P_i, t\right) d^6\zeta d^6P, \qquad (34)$$

where $dN_{\zeta,P}$ is the number of particles in a comoving volume element $d^6\zeta d^6P$ which is itself a constant of motion since the ζ's and P's form pairs of canonical variables.

The last step is to incorporate the constraints, (17). It turns out that the following representation with an additional function $\lambda\left(\zeta_i, t\right)$ included is appropriate:

$$dN_{\zeta,P} = \hat{\bar{f}}\left(\zeta_i, P_i, t\right) d^6\zeta d^6P = \lambda\left(\zeta_i, t\right) \prod_i \delta\left(P_i - \gamma_i\right) f\left(\zeta_1, \ldots, \zeta_6, t\right) d^6\zeta d^6P. $$

$$(35)$$

The function $\lambda\left(\zeta_i, t\right)$ is later to be chosen such that $f\left(\zeta_i, t\right)$ is also a constant of motion in order to become the distribution function. Since $\prod_i \delta\left(P_i - \gamma_i\right)$ is not a constant of motion (as mentioned in Sec. 2.3, $P_i = \gamma_i$ are only special solutions of (23)), λ cannot be a constant.

Since $\hat{\bar{f}}\left(\zeta_i, P_i, t\right) d^6\zeta d^6 P$ is a constant of motion, the factor $\hat{\bar{f}}\left(\zeta_i, P_i, t\right) d^6\zeta$ of $d^6 P$ is a density in P-space satisfying a continuity equation according to the equation of motion for the P_i's. This density has a total "explicit" time dependence through the explicit occurrence of t and the t-dependence of the ζ_i's and $d^6\zeta$. The continuity equation for this density is:

$$\left.\frac{\partial}{\partial t}\right|_{\text{tot}} \left(\hat{\bar{f}} d^6\zeta\right) + \frac{\partial}{\partial P_i}\left(\frac{dP_i}{dt} \hat{\bar{f}} d^6\zeta\right) = 0, \tag{36}$$

with

$$\left.\frac{\partial}{\partial t}\right|_{\text{tot}} = \frac{\partial}{\partial t} + \dot{\boldsymbol{\zeta}} \cdot \frac{\partial}{\partial \boldsymbol{\zeta}} + \frac{d}{dt}\left(d^6\zeta\right)\frac{\partial}{\partial d^6\zeta}. \tag{37}$$

Therefore

$$\frac{d}{dt}\int_{\text{whole P}-\text{space}} d^6 P \left[\hat{\bar{f}} d^6\zeta\right] = 0. \tag{38}$$

This means that the quantity

$$dN_\zeta \equiv \int_{\text{whole P}-\text{space}} d^6 P \left[\hat{\bar{f}} d^6\zeta\right] = \lambda\left(\zeta_i, t\right) f\left(\zeta_1, \ldots, \zeta_6, t\right) d^6\zeta \tag{39}$$

is a constant of motion. dN_ζ is the total number of particles in the comoving volume element $d^6\zeta$. With f required to be a constant of motion it follows that λ must be chosen such that

$$\lambda\left(\zeta_i, t\right) d^6\zeta = \text{constant of the motion}. \tag{40}$$

As shown in Appendix D, such a λ is given by

$$\lambda = \sqrt{\|\omega_{ik}\|}, \tag{41}$$

with

$$\omega_{ik} = \frac{\partial \gamma_i}{\partial \zeta_k} - \frac{\partial \gamma_k}{\partial \zeta_i}. \tag{42}$$

In Sec. 6.2.2 a very simple derivation of λ within the framework of the averaging variables will be given.

The equation for the function $f\left(\zeta_1, \ldots, \zeta_6, t\right)$ is the kinetic equation

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{6} V_i \frac{\partial f}{\partial \zeta_i} = 0. \tag{43}$$

Another way to obtain this equation is to make the ansatz

$$\bar{f}\left(\zeta_i, \alpha_i\left(\zeta_j, P_j, t\right), t\right) = f\left(\zeta_i, t\right)\lambda\left(\zeta_i, t\right)\prod_i \delta\left(P_i - \gamma_i\right) \tag{44}$$

to solve (32). Integration of $d\bar{f}/dt = 0$ with respect to $d^6 P$, taking into account the properties of the delta function and (23), then yields a continuity equation for $f\left(\zeta_i, t\right)\lambda\left(\zeta_i, t\right)$, and thus (43) for $f\left(\zeta_i, t\right)$.

14

The simplest example is the one which leads back to $\mathbf{x}, \mathbf{p}$. This is obtained with

$$\mathbf{x} = (\zeta_1, \zeta_2, \zeta_3) \ , \ \ \mathbf{p} = (\zeta_4, \zeta_5, \zeta_6) \ , \tag{45}$$

which yields

$$\boldsymbol{\gamma} = (\zeta_4, \zeta_5, \zeta_6, 0, 0, 0) \ , \ \ \lambda = \sqrt{||\omega_{ik}||} = 1. \tag{46}$$

This way all relations found above become identical with the usual ones in the Vlasov theory. This only shows, however, that the Lagrangian (25) is correct for the identity transformation. That this Lagrangian is correct for any transformation will follow from a covariance property of this Lagrangian which will be proved in the following subsection.

2.5 Covariance property of the Lagrangian for the kinetic equations

Let

$$\zeta_i = \zeta_i(\xi_k, t) \tag{47}$$

be a transformation from one choice of coordinates ζ_i to another choice of coordinates ξ_i. This transformation results, in particular, in

$$\mathbf{x} = \mathbf{x}(\xi_1, \ldots, \xi_6, t) \ , \ \ \mathbf{p} = \mathbf{p}(\xi_1, \ldots, \xi_6, t). \tag{48}$$

There are then two ways of obtaining the new Lagrangian for the kinetic equations.

1. Direct use of the latter functions in the general definition of the Lagrangian for the kinetic equations.

2. Transformation of the integrand in the old Lagrangian for the kinetic equations.

If both ways lead, up to a total time derivative, to the same Lagrangian, then direct use of the identity transformation and the transformation inside the integral on the right-hand side of (25) leads to the same results in Hamilton's principle. Since the integral is independent of the choice of the integration variables, this "covariance" proves the general correctness of the Lagrangian (25).
The result of the first way for the contribution of particle species ν is simply

$$L_\nu = \int d^6\xi \, d^6\alpha \, \hat{f}_\mathrm{p}(\xi_i, \alpha_i, t) \left[\frac{\partial \hat{S}(\xi_i, \alpha_i, t)}{\partial t} \right.$$
$$\left. + \hat{H}_p(\xi_i, t) + \sum_i \hat{V}_i \left(\frac{\partial \hat{S}}{\partial \xi_i} - \hat{\gamma}_i \right) \right], \tag{49}$$

with

$$\hat{H}_p = \frac{m}{2}\mathbf{v}^2 + e\Phi - \left.\frac{\partial\mathbf{x}}{\partial t}\right|_{\xi_i} \cdot \left(m\mathbf{v} + \frac{e}{c}\mathbf{A}\right),\tag{50}$$

$$\hat{\gamma}_i(\xi_1,\ldots,\xi_6) = \left.\frac{\partial\mathbf{x}}{\partial\xi_i}\right|_t \cdot \left(m\mathbf{v} + \frac{e}{c}\mathbf{A}\right),\tag{51}$$

$$\sum_i \hat{V}_i\left(\frac{\partial\hat{\gamma}_i}{\partial\xi_k} - \frac{\partial\hat{\gamma}_k}{\partial\xi_i}\right) - \frac{\partial\hat{\gamma}_k}{\partial t} - \frac{\partial\hat{H}_p}{\partial\xi_k} = 0.\tag{52}$$

For the second way the following transformation relations are needed:

$$\left.\frac{\partial\mathbf{x}}{\partial t}\right|_{\xi} = \left.\frac{\partial\mathbf{x}}{\partial t}\right|_{\zeta} + \sum_i \left.\frac{\partial\zeta_i}{\partial t}\right|_{\xi} \frac{\partial\mathbf{x}}{\partial\zeta_i}.\tag{53}$$

Hence it follows that

$$\begin{aligned}
\left.\frac{\partial\mathbf{x}}{\partial t}\right|_{\zeta} \cdot \mathbf{p} &= \left.\frac{\partial\mathbf{x}}{\partial t}\right|_{\xi} \cdot \mathbf{p} - \sum_i \left.\frac{\partial\zeta_i}{\partial t}\right|_{\xi} \frac{\partial\mathbf{x}}{\partial\zeta_i} \cdot \mathbf{p} \\
&= \left.\frac{\partial\mathbf{x}}{\partial t}\right|_{\xi} \cdot \mathbf{p} - \sum_i \left.\frac{\partial\zeta_i}{\partial t}\right|_{\xi} \gamma_i.
\end{aligned}\tag{54}$$

This yields

$$H_p = \hat{H}_p + \sum_i \left.\frac{\partial\zeta_i}{\partial t}\right|_{\xi} \gamma_i.\tag{55}$$

Furthermore, one obtains

$$\frac{d\zeta_i}{dt} = \sum_k \frac{\partial\zeta_i}{\partial\xi_k} \frac{d\xi_k}{dt} + \left.\frac{\partial\zeta_i}{\partial t}\right|_{\xi},\tag{56}$$

and, therefore

$$V_i = \sum_k \frac{\partial\zeta_i}{\partial\xi_k} \hat{V}_k + \left.\frac{\partial\zeta_i}{\partial t}\right|_{\xi}.\tag{57}$$

With

$$S = S(\zeta_i(\xi_k,t),\alpha_i,t) = \hat{S}(\xi_k,\alpha_i,t)\tag{58}$$

one has

$$\frac{dS}{dt} = \frac{d\hat{S}}{dt}\tag{59}$$

or

$$\frac{\partial S}{\partial t} + \sum_i V_i \frac{\partial S}{\partial\zeta_i} = \frac{\partial\hat{S}}{\partial t} + \sum_i \hat{V}_i \frac{\partial\hat{S}}{\partial\xi_i}.\tag{60}$$

Furthermore, one gets

$$\begin{aligned}
\hat{\gamma}_i &= \left.\frac{\partial\mathbf{x}}{\partial\xi_i}\right|_t \cdot \mathbf{p} = \sum_l \frac{\partial\zeta_l}{\partial\xi_i}\frac{\partial\mathbf{x}}{\partial\zeta_l} \cdot \mathbf{p} \\
&= \sum_l \frac{\partial\zeta_l}{\partial\xi_i}\gamma_l.
\end{aligned}\tag{61}$$

16

This yields

$$V_i \gamma_i = \left(\sum_k \frac{\partial \zeta_i}{\partial \xi_k} \hat{V}_k + \left. \frac{\partial \zeta_i}{\partial t} \right|_\xi \right) \gamma_i, \tag{62}$$

$$\sum_i V_i \gamma_i = \sum_k \hat{V}_k \hat{\gamma}_k + \sum_i \left. \frac{\partial \zeta_i}{\partial t} \right|_\xi \gamma_i. \tag{63}$$

The combination of all these relations yields

$$\frac{\partial S(\zeta_i, \alpha_i, t)}{\partial t} + H_p(\zeta_i, t) + \sum_i V_i \left(\frac{\partial S}{\partial \zeta_i} - \gamma_i \right) =$$

$$\frac{\partial \hat{S}(\xi_i, \alpha_i, t)}{\partial t} + \hat{H}_p(\xi_i, t) + \sum_i \hat{V}_i \left(\frac{\partial \hat{S}}{\partial \xi_i} - \hat{\gamma}_i \right). \tag{64}$$

It then remains to mention that, corresponding to the transformation of the volume element, one obtains

$$\hat{f}_\mathrm{p} = \frac{\partial(\zeta_1, \ldots, \zeta_6)}{\partial(\xi_1, \ldots, \xi_6)} f_\mathrm{p}. \tag{65}$$

These last two relations together prove the covariance property of the Lagrangian for the kinetic equations and therefore also its correctness.

3 Reformulation of the equations of motion suitable for Kruskal's formalism

The equation of motion

$$\frac{d\mathbf{v}}{dt} = \frac{e}{m} \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right) \tag{66}$$

is reformulated in a form suitable for applying Kruskal's method. This is done by representing the particle velocity as

$$\mathbf{v} = u_\parallel \mathbf{b} + u_\perp \mathbf{n}_1 + \mathbf{v}_E \tag{67}$$

and introducing the five-component vector $\mathbf{y}$ and an angle variable ϑ given by

$$\mathbf{y} = (y_1, y_2, y_3, y_4, y_5) = (x, y, z, u_\parallel, u_\perp) = (\mathbf{x}, u_\parallel, u_\perp), \ \vartheta = \frac{\theta}{2\pi}. \tag{68}$$

Here, the dot means d/dt, $\mathbf{b} = \mathbf{B}/B$, and $\mathbf{v}_E = (c/B^2)\mathbf{E} \times \mathbf{B}$. $\mathbf{n}_1 = \mathbf{n}_1(\mathbf{x}, \theta, t)$ is a unit vector perpendicular to the magnetic field at point $\mathbf{x}$ and time t, characterized by the angle θ as defined by (70) below. This description makes use of two orthonormal basis sets. The first one, $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, is related to the magnetic field and thus depends only on $\mathbf{x}$ and t. The second one, $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{e}_3)$,

17

is a particle-related basis set with $\mathbf{n}_1$ and $\mathbf{n}_2$ depending not only on $\mathbf{x}$ and t, but also on the instantaneous orientation of the particle velocity in terms of the angle θ. Given in right-hand-rule order, these two sets are defined by the relations

$$\mathbf{e}_1 = \frac{\mathbf{b} \cdot \boldsymbol{\nabla} \mathbf{b}}{|\mathbf{b} \cdot \boldsymbol{\nabla} \mathbf{b}|} \quad , \quad \mathbf{e}_2 = \mathbf{b} \times \mathbf{e}_1 \quad , \quad \mathbf{e}_3 = \mathbf{b}, \tag{69}$$

$$\mathbf{n}_1 \left(\mathbf{x}, \theta, t\right) = -\sin\theta \, \mathbf{e}_1 - \cos\theta \, \mathbf{e}_2 \quad , \quad \mathbf{n}_2 \left(\mathbf{x}, \theta, t\right) = \cos\theta \, \mathbf{e}_1 - \sin\theta \, \mathbf{e}_2. \tag{70}$$

Substituting (67) in (66) yields for the LHS of this equation

$$\dot{\mathbf{v}} = \dot{u}_{\|}\mathbf{b} + u_{\|}\dot{\mathbf{b}} + \dot{u}_{\perp}\mathbf{n}_1 - \dot{\theta}u_{\perp}\mathbf{n}_2 - u_{\perp}(\dot{\mathbf{e}}_1 \sin\theta + \dot{\mathbf{e}}_2 \cos\theta) + \dot{\mathbf{v}}_{\mathbf{E}}, \tag{71}$$

while the right-hand side is given by

$$\frac{e}{m}\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right) = \frac{e}{m}\left(\mathbf{b}\cdot\mathbf{E}\right)\mathbf{b} - \frac{eB}{mc}u_{\perp}\mathbf{n}_2. \tag{72}$$

Scalar multiplication of (71) by $\mathbf{b}$, $\mathbf{n}_1$ and $\mathbf{n}_2$ yields the equations for $\dot{u}_{\|}$, $\dot{u}_{\perp}$ and $\dot{\theta}$, respectively. With the operator $\mathcal{D}/\mathcal{D}t$ defined by

$$\frac{\mathcal{D}}{\mathcal{D}t} = \frac{d}{dt} - u_{\perp}\mathbf{n}_1 \cdot \boldsymbol{\nabla} \tag{73}$$

in order to separate the explicit θ-dependence in d/dt, one obtains

$$\dot{u}_{\|} = \frac{e}{m}\mathbf{b}\cdot\mathbf{E} + [u_{\perp}\mathbf{n}_1 + \mathbf{v}_E]\cdot\left[\frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + u_{\perp}\mathbf{n}_1\cdot\boldsymbol{\nabla}\mathbf{b}\right], \tag{74}$$

$$\dot{u}_{\perp} = -\mathbf{n}_1\cdot\left[u_{\|}\frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + \frac{\mathcal{D}\mathbf{v}_E}{\mathcal{D}t}\right] - u_{\perp}\mathbf{n}_1\cdot\left[\mathbf{n}_1\cdot\left(u_{\|}\boldsymbol{\nabla}\mathbf{b} + \boldsymbol{\nabla}\mathbf{v}_E\right)\right], \tag{75}$$

$$\begin{aligned}
\dot{\vartheta} \;\; := \;\; & \frac{\dot{\theta}}{2\pi} \\
= \;\; & \frac{1}{2\pi}\left\{\frac{eB}{mc} + \frac{\mathbf{n}_2}{u_{\perp}}\cdot\left[u_{\|}\frac{d\mathbf{b}}{dt} + \frac{d\mathbf{v}_E}{dt}\right] + \mathbf{e}_2\cdot\frac{d\mathbf{e}_1}{dt}\right\} \\
= \;\; & \frac{1}{2\pi}\left\{\frac{eB}{mc} + \frac{\mathbf{n}_2}{u_{\perp}}\cdot\left[u_{\|}\frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + \frac{\mathcal{D}\mathbf{v}_E}{\mathcal{D}t}\right] + \mathbf{n}_2\cdot\left[\mathbf{n}_1\cdot\left(u_{\|}\boldsymbol{\nabla}\mathbf{b} + \boldsymbol{\nabla}\mathbf{v}_E\right)\right]\right. \\
& \left. + \mathbf{e}_2\cdot\frac{\partial\mathbf{e}_1}{\partial t} + \left(u_{\|}\mathbf{b} + \mathbf{v}_E + u_{\perp}\mathbf{n}_1\right)\cdot\left[\boldsymbol{\nabla}\mathbf{e}_1\cdot\mathbf{e}_2\right]\right\}.
\end{aligned} \tag{76}$$

3.1 The Kruskal form of the equations of motion and related requirements

The Kruskal form of the equations of motion is obtained by changing the time scale,

$$t \rightarrow s = \frac{t}{\epsilon}. \tag{77}$$

The basic assumption of the drift-kinetic and gyrokinetic theories is that the fast time dependence is described by the gyrofrequency alone. One can therefore write

$$\frac{eB}{mc} = \Omega_g = \frac{1}{\epsilon}\Omega_s. \tag{78}$$

This yields for (67) and (74)-(76)

$$\frac{d\mathbf{x}}{ds} = \epsilon \left(u_{\parallel}\mathbf{b} + u_{\perp}\mathbf{n}_1 + \mathbf{v}_E \right), \tag{79}$$

$$\frac{du_{\parallel}}{ds} = \epsilon \left[\frac{e}{m}\mathbf{b}\cdot\mathbf{E} + [u_{\perp}\mathbf{n}_1 + \mathbf{v}_E]\cdot\left[\frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + u_{\perp}\mathbf{n}_1\cdot\boldsymbol{\nabla}\mathbf{b} \right] \right], \tag{80}$$

$$\frac{du_{\perp}}{ds} = -\epsilon \left[\mathbf{n}_1\cdot\left[u_{\parallel}\frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + \frac{\mathcal{D}\mathbf{v}_E}{\mathcal{D}t} \right] + u_{\perp}\mathbf{n}_1\cdot\left[\mathbf{n}_1\cdot\left(u_{\parallel}\boldsymbol{\nabla}\mathbf{b} + \boldsymbol{\nabla}\mathbf{v}_E \right) \right] \right], \tag{81}$$

$$\begin{aligned}
\frac{d\vartheta}{ds} = & \frac{1}{2\pi}\Omega_s + \epsilon\frac{1}{2\pi}\left\{ \frac{\mathbf{n}_2}{u_{\perp}}\cdot\left[u_{\parallel}\frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + \frac{\mathcal{D}\mathbf{v}_E}{\mathcal{D}t} \right] + \mathbf{n}_2\cdot\left[\mathbf{n}_1\cdot\left(u_{\parallel}\boldsymbol{\nabla}\mathbf{b} + \boldsymbol{\nabla}\mathbf{v}_E \right) \right] \right. \\
& \left. + \mathbf{e}_2\cdot\frac{\partial\mathbf{e}_1}{\partial t} + \left(u_{\parallel}\mathbf{b} + \mathbf{v}_E + u_{\perp}\mathbf{n}_1 \right)\cdot\left[\boldsymbol{\nabla}\mathbf{e}_1\cdot\mathbf{e}_2 \right] \right\}. \tag{82}
\end{aligned}$$

With the definition of the five-component vector $\mathbf{y}$ in (68), these equations are of the form

$$\frac{d\mathbf{y}}{ds} = \epsilon\,\mathbf{g}\left(\mathbf{y}, \vartheta \right) \qquad , \qquad \frac{d\vartheta}{ds} = \psi\left(\mathbf{y}, \vartheta \right), \tag{83}$$

with

$$\mathbf{g}\left(\mathbf{y}, \vartheta + 1 \right) = \mathbf{g}(\mathbf{y}, \vartheta) \quad , \quad \psi\left(\mathbf{y}, \vartheta + 1 \right) = \psi(\mathbf{y}, \vartheta). \tag{84}$$

Equations (83) and (84) are of the same type as the starting equations in Kruskal's method if they satisfy the conditions

$$\mathbf{g}(\mathbf{y}, \vartheta) = \mathcal{O}(\epsilon^0) \quad , \quad \psi(\mathbf{y}, \vartheta) = \mathcal{O}(\epsilon^0). \tag{85}$$

For the inverse Kruskal method there is the additional requirement that

$$\psi(\mathbf{y}, \vartheta) = \psi_0(\mathbf{y}) + \mathcal{O}\left(\epsilon^1 \right). \tag{86}$$

Weaker conditions, which would be of interest for the gyrokinetic theory are not possible, as follows from (120) and (121) below. These latter relations require that the ϕ-average of $\mathbf{g}_0$ possess only a parallel spatial component, which is the case if $\mathbf{v}_E$ is of the order $\epsilon u_{\perp}$, $\epsilon u_{\parallel}$. This way it is ensured that, like $\partial/\partial t$, also $\mathbf{k}_{\perp}\cdot\mathbf{v}_E = \mathcal{O}(\epsilon)\Omega_g$, with $k_{\perp} \sim 1/r_{\text{gyro}}$. (In the drift-kinetic case this is guaranteed with $\mathbf{v}_E = \mathcal{O}\left(\epsilon^0 \right)$ because of $k_{\perp} \sim 1/L$.) This and the above conditions are satisfied if with

$$u_{\perp} = \mathcal{O}\left(\epsilon^0 \right) \qquad , \qquad u_{\parallel} = \mathcal{O}\left(\epsilon^0 \right) \tag{87}$$

it holds that
$$\mathbf{b} \cdot \mathbf{E} = \mathcal{O}\left(\epsilon^0\right) \quad , \quad \boldsymbol{\nabla} \mathbf{v}_E = \mathcal{O}\left(\epsilon^0\right), \tag{88}$$

$$\frac{\partial}{\partial t} = \Omega_s \mathcal{O}\left(\epsilon^0\right) \quad , \quad \mathbf{v}_E = \mathcal{O}\left(\epsilon^0\right) \text{ (for the drift-kinetic case)}$$
$$\mathbf{v}_E = \mathcal{O}\left(\epsilon\right) \text{ (for the gyrokinetic case)}, \tag{89}$$

$$\boldsymbol{\nabla} \mathbf{b} = \mathcal{O}\left(\epsilon^0\right) . \tag{90}$$

In the drift-kinetic case, one has

$$\boldsymbol{\nabla} \sim \frac{1}{L} \text{ with } L \sim \frac{r_{\text{gyro}}}{\epsilon}, \tag{91}$$

and the conditions (87)-(90) are therefore satisfied if

$$\mathbf{b} \cdot \mathbf{E} = \frac{v}{c} B \, \mathcal{O}(\epsilon) \quad , \quad \mathbf{b} \times \mathbf{E} = \frac{v}{c} B \, \mathcal{O}(\epsilon^0). \tag{92}$$

In the gyrokinetic case, one also considers fluctuations with large perpendicular derivatives,

$$\mathbf{b} \times \boldsymbol{\nabla} \sim \frac{1}{r_{\text{gyro}}} \text{ for fluctuations.} \tag{93}$$

Equations (87)-(90) then require that the fluctuations of $\mathbf{v}_E$ and $\mathbf{b}$, given by $\delta \mathbf{v}_E$ and $\delta \mathbf{b}$, satisfy the condition

$$\delta \mathbf{v}_E = \mathcal{O}(\epsilon) \quad , \quad \delta \mathbf{b} = \mathcal{O}(\epsilon). \tag{94}$$

This means, together with the characterization of $\mathbf{v}_E$ in (89), that

$$\mathbf{b} \times \mathbf{E}_{\text{fluctuations}} = \frac{v}{c} B \, \mathcal{O}\left(\epsilon\right) \quad , \quad \mathbf{B}_{\text{fluctuations}} = B \, \mathcal{O}\left(\epsilon\right). \tag{95}$$

In Sec. 4 it will be shown that the recursion procedure of the inverse Kruskal method does not introduce any further conditions. This is due to the fact that the perpendicular derivatives which appear in the recursion procedure occur only together with a factor ϵ^n, with $n \geq 1$.

In order to avoid problems in doing variations of the electric and magnetic potentials in Hamilton's principle with the Lagrangian for the combined gyrokinetic-Maxwell theory which will be derived in Sec. 6, no splitting into slow and fast spatial dependences will be done. This problem does not appear in the drift-kinetic case.

3.2 Estimate of the allowed fluctuation densities for electrostatic perturbations

If $\mathbf{E}$ is assumed to be mainly of electrostatic origin, the electron density can be approximated by

$$n_{\text{fluctuation},\,e} = n_{0e}\, e^{\frac{e\Phi}{T}}.\tag{96}$$

In the drift-kinetic case one has for all components

$$\mathbf{E} \sim \frac{\Phi}{L}.\tag{97}$$

According to (92), it then follows that

$$\Phi \sim r_{\text{gyro}}\, \frac{v}{c} B,\tag{98}$$

and, therefore, that

$$\frac{e\Phi}{T} \sim \frac{mvc}{eB}\, \frac{v}{c}\, eB\, \frac{1}{T} \sim 1.\tag{99}$$

This means that large density fluctuations are allowed.

In the gyrokinetic case one has

$$\mathbf{b} \cdot \mathbf{E} \sim \frac{\Phi}{L} \quad , \quad \mathbf{b} \times \mathbf{E} \sim \frac{\Phi}{r_{\text{gyro}}}.\tag{100}$$

According to (92) and (95), which give the same limits for all components of $\mathbf{E}$, the second relation, (100), is the critical one and yields

$$\Phi \sim r_{\text{gyro}}\, \frac{r_{\text{gyro}}}{L}\, \frac{v}{c}\, B,\tag{101}$$

and, therefore, it follows that

$$\frac{e\Phi}{T} \sim \frac{r_{\text{gyro}}}{L}.\tag{102}$$

Hence, only very small fluctuation amplitudes are allowed.

4 The inverse Kruskal method

4.1 General outline

Kruskal's method [17] (with slightly different notation here and later, and adapted to explicitly account for a possible slow dependence of the fields on time, i.e. a dependence on $t = \epsilon s$) starts with a set of ordinary differential equations which have the following structure:

$$\mathbf{y}_s = \frac{d\mathbf{y}}{ds} = \epsilon\, \mathbf{g}(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \vartheta_s = \frac{d\vartheta}{ds} = \psi(\mathbf{y}, \vartheta, \epsilon s)\tag{103}$$

with $\mathbf{y} = (y_1, \ldots, y_n)$ and

$$\mathbf{g}(\mathbf{y}, \vartheta + 1, \epsilon s) = \mathbf{g}(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \psi(\mathbf{y}, \vartheta + 1, \epsilon s) = \psi(\mathbf{y}, \vartheta, \epsilon s). \tag{104}$$

The goal of the inverse Kruskal method is to obtain a transformation (cf. also [20])

$$\mathbf{y} = \mathbf{Y}(\hat{\mathbf{z}}, \phi, \epsilon s) \quad , \quad \vartheta = \Theta(\hat{\mathbf{z}}, \phi, \epsilon s) \quad , \quad \hat{\mathbf{z}} = (\hat{z}_1, \ldots, \hat{z}_n), \tag{105}$$

where ϕ is an angle variable with period 1 such that

$$\hat{\mathbf{z}}_s = \epsilon \, \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \quad , \quad \phi_s = \omega(\hat{\mathbf{z}}, \epsilon s). \tag{106}$$

These transformations will be obtained below, analogously to the original Kruskal method, iteratively via a recursion procedure to any approximation in ϵ without solving any differential equations.
For Hamiltonian systems the quantity

$$J(\hat{\mathbf{z}}, t) = \oint_{\hat{\mathbf{z}} = const} \mathbf{p} \cdot d\mathbf{x} = \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial \phi} \, d\phi \tag{107}$$

turns out to be canonically conjugated to ϕ and a constant of motion for the approximation considered. A direct proof of $dJ(\hat{\mathbf{z}}, t)/dt = 0$ is given in Appendix A. This constant of motion is gauge-invariant since it is unchanged by replacing $\mathbf{A}$ with $\mathbf{A} + \boldsymbol{\nabla}\Psi$,

$$\int_0^1 \boldsymbol{\nabla}\Psi(\mathbf{x}, t) \cdot \frac{\partial \mathbf{x}}{\partial \phi} \, d\phi = \int_0^1 \frac{\partial \Psi}{\partial \phi} \, d\phi = \oint d\Psi = 0, \tag{108}$$

and therefore does not explicitly depend on the vector potential, but only on the magnetic field and its derivatives. Note that by Stokes' theorem the contribution of the vector potential $\mathbf{A}$ to J is essentially the magnetic flux inside the loop $\hat{\mathbf{z}} = const$. This is discussed in detail in [27].

4.2   Recursion procedure

As in Kruskal's original theory, one can find a recursion procedure which enables one to obtain the transformation of (105) together with the functions $\mathbf{h}$ and ω in (106) to arbitrary approximation in ϵ without solving any differential equations. This goal is reached in the following way:
From (103) and (105) one obtains

$$\begin{aligned} \mathbf{y}_s &= \hat{\mathbf{z}}_s \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} + \phi_s \frac{\partial \mathbf{Y}}{\partial \phi} + \epsilon \frac{\partial \mathbf{Y}}{\partial(\epsilon s)} \\ &= \epsilon \, \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} + \omega(\hat{\mathbf{z}}, \epsilon s) \frac{\partial \mathbf{Y}}{\partial \phi} + \epsilon \frac{\partial \mathbf{Y}}{\partial(\epsilon s)} = \epsilon \, \mathbf{g}(\mathbf{Y}, \Theta, \epsilon s), \end{aligned} \tag{109}$$

$$\vartheta_s = \hat{\mathbf{z}}_s \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} + \phi_s \frac{\partial \Theta}{\partial \phi} + \epsilon \frac{\partial \Theta}{\partial (\epsilon s)}$$

$$= \epsilon \, \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} + \omega \left(\hat{\mathbf{z}}, \epsilon s \right) \frac{\partial \Theta}{\partial \phi} + \epsilon \frac{\partial \Theta}{\partial (\epsilon s)} = \psi \left(\mathbf{Y}, \Theta, \epsilon s \right). \quad (110)$$

Formal integration of (109) and (110) over ϕ yields

$$\mathbf{Y} = \mathbf{Y} \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right)$$
$$+ \frac{\epsilon}{\omega(\hat{\mathbf{z}}, \epsilon s)} \int_0^\phi \left[\mathbf{g} \left(\mathbf{Y}, \Theta, \epsilon s \right) - \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} - \frac{\partial \mathbf{Y}}{\partial (\epsilon s)} \right] d\tilde{\phi}, \quad (111)$$

$$\Theta = \Theta \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right)$$
$$+ \frac{1}{\omega(\hat{\mathbf{z}}, \epsilon s)} \int_0^\phi \left[\psi \left(\mathbf{Y}, \Theta, \epsilon s \right) - \epsilon \, \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} - \epsilon \frac{\partial \Theta}{\partial (\epsilon s)} \right] d\tilde{\phi}. \quad (112)$$

The transformation functions $\mathbf{Y}$ and Θ must satisfy the periodicity conditions

$$\mathbf{Y}(\hat{\mathbf{z}}, \phi + 1, \epsilon s) = \mathbf{Y}(\hat{\mathbf{z}}, \phi, \epsilon s),$$
$$\Theta(\hat{\mathbf{z}}, \phi + 1, \epsilon s) = \Theta(\hat{\mathbf{z}}, \phi, \epsilon s) + 1. \quad (113)$$

They need to be imposed only for $\phi = 0$ as starting values. With (111) and (112), these conditions become

$$\int_0^1 \left[\mathbf{g} \left(\mathbf{Y}, \Theta, \epsilon s \right) - \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} - \frac{\partial \mathbf{Y}}{\partial (\epsilon s)} \right] d\phi = 0, \quad (114)$$

$$\frac{1}{\omega(\hat{\mathbf{z}}, \epsilon s)} \int_0^1 \left[\psi \left(\mathbf{Y}, \Theta, \epsilon s \right) - \epsilon \, \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} - \epsilon \frac{\partial \Theta}{\partial (\epsilon s)} \right] d\phi = 1. \quad (115)$$

The functions $\mathbf{Y} \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right)$ and $\Theta \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right)$ may be prescribed arbitrarily. The choice $\mathbf{Y} \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right) = \hat{\mathbf{z}}$, $\Theta \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right) = 0$, which is analogous to that in the original Kruskal method, though perfectly valid, does not yield the usual guiding centre position. Instead we choose

$$\mathbf{Y} \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right) = \hat{\mathbf{z}} + \epsilon \mathbf{Y}_c \left(\hat{\mathbf{z}}, \epsilon s \right), \quad (116)$$

$$\Theta \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right) = \epsilon \Theta_c \left(\hat{\mathbf{z}}, \epsilon s \right), \quad (117)$$

where $\mathbf{Y}_c \left(\hat{\mathbf{z}}, \epsilon s \right)$ and $\Theta_c \left(\hat{\mathbf{z}}, \epsilon s \right)$ are power series in ϵ:

$$\epsilon \mathbf{Y}_c \left(\hat{\mathbf{z}}, \epsilon s \right) = \epsilon \mathbf{Y}_{c1} \left(\hat{\mathbf{z}}, \epsilon s \right) + \epsilon^2 \mathbf{Y}_{c2} \left(\hat{\mathbf{z}}, \epsilon s \right) + \cdots, \quad (118)$$

$$\epsilon \Theta_c \left(\hat{\mathbf{z}}, \epsilon s \right) = \epsilon \Theta_{c1} \left(\hat{\mathbf{z}}, \epsilon s \right) + \epsilon^2 \Theta_{c2} \left(\hat{\mathbf{z}}, \epsilon s \right) + \cdots. \quad (119)$$

This choice guarantees that the variables $\hat{\mathbf{z}}$ and ϕ and the particle variables $\mathbf{y}$ and θ coincide in the limit $\epsilon \to 0$, leaving, however, each term in the series for $\mathbf{Y}_c \left(\hat{\mathbf{z}}, \epsilon s \right)$ and $\Theta_c \left(\hat{\mathbf{z}}, \epsilon s \right)$ still arbitrary, these being chosen later as appears convenient. A subject related to the freedom to prescribe the functions $\mathbf{Y} \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right)$ and $\Theta \left(\hat{\mathbf{z}}, \phi = 0, \epsilon s \right)$ is discussed in [28].

Equations (111) and (112) together with (114) and (115) determine the functions $\mathbf{Y}(\hat{\mathbf{z}}, \phi, \epsilon s)$, $\Theta(\hat{\mathbf{z}}, \phi, \epsilon s)$, $\mathbf{h}(\hat{\mathbf{z}}, \epsilon s)$ and $\omega(\hat{\mathbf{z}}, \epsilon s)$.

4.3 Solution by iteration

4.3.1 *Remarks on the gyrokinetic theory*

Because of the fast spatial variations allowed in the gyrokinetic theory, spatial derivatives can change the order in ϵ. In (111), (112), (114) and (115) such derivatives occur only in the form

$$\mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial}{\partial \hat{\mathbf{z}}}. \tag{120}$$

As will be shown below, $\mathbf{h}$ is given to zeroth order in ϵ by

$$\mathbf{h}_0 = \int_0^1 \mathbf{g}_0 \, d\phi \equiv \; < \mathbf{g}_0 > . \tag{121}$$

According to (79), (83), (89) and (94), $< \mathbf{g}_0 >$ possesses only a parallel spatial component. The derivatives (120) therefore cannot lower the order in ϵ. This means that an iterative solution procedure is possible. This was the reason for introducing the condition $\mathbf{v}_E = \mathcal{O}\left(\epsilon\right)$ for the gyrokinetic theory. There is, of course, no problem in the case of drift-kinetic theory.

4.3.2 *Notations*

The following notations will be used:

$$\begin{aligned}
\mathbf{Y}^{(n)} &= \mathbf{Y}, \;\; \text{obtained in the n-th iteration,} \\
< f > &= \int_0^1 f \, d\phi \;\; \text{for any function } f.
\end{aligned} \tag{122}$$

The upper indices introduced here must not be mixed up with the lower indices indicating the order in ϵ.

4.3.3 *Structure of the iteration procedure*

The iteration is done in the following way: (111), (112), (114) and (115) will be solved in the order (111), (115), (112) and (114) using in each step the results obtained in the preceding one.

The iteration procedure will make use of the following properties of the functions $\mathbf{g}$ and ψ, which, according to (79)-(82) and (83), have the form

$$\mathbf{g} = \mathbf{g}_0(\mathbf{y}, \vartheta, \epsilon s) + \epsilon \mathbf{g}_1(\mathbf{y}, \vartheta, \epsilon s) \; , \;\; \psi = \psi_0(\mathbf{y}, \epsilon s) + \epsilon \psi_1(\mathbf{y}, \vartheta, \epsilon s) \; ,$$
$$\psi_0(\mathbf{y}, \epsilon s) = \frac{\Omega_s(\mathbf{x}, \epsilon s)}{2\pi}. \tag{123}$$

4.3.4 *Orders in ϵ needed*

The orders in ϵ needed are ϵ^0 and ϵ^1. From the first order of $\mathbf{Y}$, only $\mathbf{x}$ is necessary in order to calculate J, and also for the spatial arguments in the gyrokinetic case and for the factor in front of the scalar and vector potentials in order to guarantee gauge invariance.

4.3.5 *Zeroth order*

The zeroth-order quantities follow from (111), (112), (114) and (115) with $\epsilon = 0$:

From (111):

$$\mathbf{Y}^{(0)} = \hat{\mathbf{z}}. \tag{124}$$

From (115):

$$\omega^{(0)}(\hat{\mathbf{z}}, \epsilon s) = \psi_0(\hat{\mathbf{z}}, \epsilon s) = \frac{1}{2\pi}\Omega_s(\hat{\mathbf{z}}, \epsilon s). \tag{125}$$

From (112):

$$\Theta^{(0)} = \phi. \tag{126}$$

From (114):

$$\mathbf{h}^{(0)} = < \mathbf{g}_0(\hat{\mathbf{z}}, \phi, \epsilon s) > . \tag{127}$$

In the later applications only the first three components $\mathbf{h}^{(0)}$ are of interest. These are

$$\left(h_1^{(0)}, h_2^{(0)}, h_3^{(0)}\right) = u_\| \mathbf{b} + \mathbf{v}_E, \tag{128}$$

where the right-hand side is understood to be given in terms of the phase-space coordinates $\hat{\mathbf{z}}$ and ϕ. Note that in the gyrokinetic theory $\mathbf{v}_E$ does not appear because of $\mathbf{v}_E = \mathcal{O}(\epsilon)$.

4.3.6 *First order*

From (111) and (116):

$$\begin{aligned}
\mathbf{Y}^{(1)} &= \hat{\mathbf{z}} + \epsilon \mathbf{Y}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{\omega_0(\hat{\mathbf{z}}, \epsilon s)} \int_0^\phi \left[\mathbf{g}_0(\hat{\mathbf{z}}, \tilde{\phi}, \epsilon s) - < \mathbf{g}_0 > \right] d\tilde{\phi} \\
&= \hat{\mathbf{z}} + \epsilon \mathbf{Y}_1.
\end{aligned} \tag{129}$$

In the later applications, again only the spatial components $\mathbf{x}^{(1)}$ are of interest. With

$$\mathbf{R} := (\hat{z}_1, \hat{z}_2, \hat{z}_3) \quad , \quad V_\perp := \hat{z}_5 \quad , \tag{130}$$

they are

$$\begin{aligned}
\mathbf{x}^{(1)} &= \mathbf{R} + \epsilon\, \mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{\omega_0(\mathbf{R}, \epsilon s)} \int_0^\phi \mathbf{n}_1\left(\mathbf{R}, \tilde{\phi}, \epsilon s\right) V_\perp\, d\tilde{\phi} \\
&= \mathbf{R} + \epsilon\, \mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{2\pi\omega_0(\mathbf{R}, \epsilon s)} V_\perp(\mathbf{n}_2(\mathbf{R}, \phi, \epsilon s) - \mathbf{e}_1(\mathbf{R}, \epsilon s)) \\
&= \mathbf{R} + \epsilon\, \mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{\Omega_s(\mathbf{R}, \epsilon s)} V_\perp(\mathbf{n}_2(\mathbf{R}, \phi, \epsilon s) - \mathbf{e}_1(\mathbf{R}, \epsilon s)).
\end{aligned} \tag{131}$$

The arbitrary function $\mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s)$ can be chosen such as to obtain the usual guiding centre position to first order in ϵ:

$$\mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) = \frac{V_\perp}{\Omega_s(\mathbf{R}, \epsilon s)} \mathbf{e}_1(\mathbf{R}, \epsilon s). \tag{132}$$

This yields

$$\mathbf{x}^{(1)} = \mathbf{R} + \epsilon \frac{V_\perp}{\Omega_s(\mathbf{R}, \epsilon s)} \mathbf{n}_2 (\mathbf{R}, \phi, \epsilon s) \tag{133}$$

A particular result is obtained by substituting this relation in (107):

$$J^{(1)} = \epsilon \frac{2\pi}{\Omega_s(\mathbf{R}, \epsilon s)} \frac{m V_\perp^2}{2}. \tag{134}$$

Therefore,

$$\mathbf{x}^{(1)} = \mathbf{R} + \boldsymbol{\rho}^{(1)} \quad , \quad \boldsymbol{\rho}^{(1)} = \epsilon \, \rho_1 \mathbf{n}_2 (\mathbf{R}, \phi, \epsilon s) = \epsilon \sqrt{\frac{2}{m} \frac{1}{2\pi\Omega_s} \frac{J^{(1)}}{\epsilon}} \mathbf{n}_2 (\mathbf{R}, \phi, \epsilon s) . \tag{135}$$

5 Reduced phase-space Lagrangian

5.1 Convenient choice of a gauge function in the phase-space Lagrangian

In order to obtain a Lagrangian in a lower-dimensional space in terms of J-based averaging variables (with J taken as one of the coordinates in phase space), we start, as in Sec. 2.1, with a phase-space particle Lagrangian in the form

$$L(\mathbf{p}, \mathbf{x}; \dot{\mathbf{x}}; t) = \mathbf{p} \cdot \dot{\mathbf{x}} - H(\mathbf{x}, \mathbf{p}, t) + \frac{d}{dt} F(\mathbf{x}, \mathbf{p}, t), \tag{136}$$

where, as compared with Sec. 2.1, a total time derivative of an arbitrary gauge function $F(\mathbf{x}, \mathbf{p}, t)$ in phase space is added. By expressing the position vector $\mathbf{x}$ and the canonical momentum $\mathbf{p}$ in terms of J-based averaging variables $z_1 = \hat{z}_1, z_2 = \hat{z}_2, z_3 = \hat{z}_3, z_4 = \hat{z}_4, J$ and ϕ,

$$\mathbf{x} = \mathfrak{x} (z_1, \dots, z_4, J, \phi, t) \quad , \quad \mathbf{p} = \mathfrak{p} (z_1, \dots, z_4, J, \phi, t) , \tag{137}$$

where $\mathfrak{x}$ and $\mathfrak{p}$ are periodic functions of ϕ with period 1, one obtains the Lagrangian

$$L \left(z_1, \dots, z_4, J, \phi; \dot{z}_1, \dots, \dot{z}_4, \dot{J}, \dot{\phi}; t \right) =$$
$$\sum_{i=1}^{4} \dot{z}_i \left[\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} + \frac{\partial F}{\partial z_i} \right] + \dot{\phi} \left[\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} + \frac{\partial F}{\partial \phi} \right]$$
$$+ \dot{J} \left[\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial J} + \frac{\partial F}{\partial J} \right] - \left[H - \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t} - \frac{\partial F}{\partial t} \right] . \tag{138}$$

Here, the Fraktur letters $\mathfrak{x}$ and $\mathfrak{p}$ are used to stress the functional dependence on the J-based averaging coordinates. The arbitrary gauge function F can be

chosen such as to make the coefficient of $\dot{\phi}$ in (138) independent of ϕ and equal to J,

$$\frac{\partial F}{\partial \phi} := J - \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \phi}. \tag{139}$$

Integration of this equation yields

$$
\begin{aligned}
F\left(z_1, \ldots, z_4, J, \phi; t\right) &= \langle F \rangle \left(z_1, \ldots, z_4, J; t\right) + F^{(\phi)}\left(z_1, \ldots, z_4, J, \phi; t\right) \\
&\quad - \left\langle F^{(\phi)}\left(z_1, \ldots, z_4, J, \phi; t\right) \right\rangle,
\end{aligned}
\tag{140}
$$

where $\langle F \rangle \left(z_1, \ldots, z_4, J; t\right)$ is an arbitrary function, and

$$F^{(\phi)}\left(z_1, \ldots, z_4, J, \phi; t\right) = J\phi - \int_0^\phi \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \hat{\phi}} \, d\hat{\phi}. \tag{141}$$

The function $F^{(\phi)}$ is periodic in ϕ.
With the identity

$$f = \langle f \rangle + \int_0^\phi \frac{\partial f}{\partial \hat{\phi}} \, d\hat{\phi} - \left\langle \int_0^\phi \frac{\partial f}{\partial \hat{\phi}} \, d\hat{\phi} \right\rangle \tag{142}$$

for any function f taken into account, one obtains

$$
\begin{aligned}
\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} + \frac{\partial F}{\partial z_i} &= \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} \right\rangle + \frac{\partial \langle F \rangle}{\partial z_i} \\
&\quad + \int_0^\phi \left[\frac{\partial \mathfrak{p}}{\partial \hat{\phi}} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} - \frac{\partial \mathfrak{p}}{\partial z_i} \cdot \frac{\partial \mathfrak{x}}{\partial \hat{\phi}} \right] d\hat{\phi} \\
&\quad - \left\langle \int_0^\phi \left[\frac{\partial \mathfrak{p}}{\partial \hat{\phi}} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} - \frac{\partial \mathfrak{p}}{\partial z_i} \cdot \frac{\partial \mathfrak{x}}{\partial \hat{\phi}} \right] d\hat{\phi} \right\rangle,
\end{aligned}
\tag{143}
$$

$$
\begin{aligned}
\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial J} + \frac{\partial F}{\partial J} &= \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial J} \right\rangle + \frac{\partial \langle F \rangle}{\partial J} \\
&\quad + \int_0^\phi \left[1 + \frac{\partial \mathfrak{p}}{\partial \hat{\phi}} \cdot \frac{\partial \mathfrak{x}}{\partial J} - \frac{\partial \mathfrak{p}}{\partial J} \cdot \frac{\partial \mathfrak{x}}{\partial \hat{\phi}} \right] d\hat{\phi} \\
&\quad - \left\langle \int_0^\phi \left[1 + \frac{\partial \mathfrak{p}}{\partial \hat{\phi}} \cdot \frac{\partial \mathfrak{x}}{\partial J} - \frac{\partial \mathfrak{p}}{\partial J} \cdot \frac{\partial \mathfrak{x}}{\partial \hat{\phi}} \right] d\hat{\phi} \right\rangle,
\end{aligned}
\tag{144}
$$

$$
\begin{aligned}
H - \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t} - \frac{\partial F}{\partial t} &= \langle H \rangle - \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t} \right\rangle - \frac{\partial \langle F \rangle}{\partial t} \\
&\quad - \int_0^\phi \left[-\frac{\partial H}{\partial \hat{\phi}} + \frac{\partial \mathfrak{p}}{\partial \hat{\phi}} \cdot \frac{\partial \mathfrak{x}}{\partial t} - \frac{\partial \mathfrak{p}}{\partial t} \cdot \frac{\partial \mathfrak{x}}{\partial \hat{\phi}} \right] d\hat{\phi} \\
&\quad + \left\langle \int_0^\phi \left[-\frac{\partial H}{\partial \hat{\phi}} + \frac{\partial \mathfrak{p}}{\partial \hat{\phi}} \cdot \frac{\partial \mathfrak{x}}{\partial t} - \frac{\partial \mathfrak{p}}{\partial t} \cdot \frac{\partial \mathfrak{x}}{\partial \hat{\phi}} \right] d\hat{\phi} \right\rangle.
\end{aligned}
\tag{145}
$$

As shown in Appendix B, the Lagrange brackets which appear in these equations satisfy the following relations

$$\frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} - \frac{\partial \mathfrak{p}}{\partial z_i} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} = 0, \tag{146}$$

$$\frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{x}}{\partial J} - \frac{\partial \mathfrak{p}}{\partial J} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} = -1, \tag{147}$$

$$\frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{x}}{\partial t} - \frac{\partial \mathfrak{p}}{\partial t} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} = \frac{\partial H}{\partial \phi} . \tag{148}$$

By inserting (143)-(148) in (138), one obtains

$$L\left(z_1, \ldots, z_4, J; \dot{z}_1, \ldots, \dot{z}_4, \dot{J}, \dot{\phi}; t\right) =$$
$$\sum_{i=1}^{4} \dot{z}_i \left[\left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial z_i}\right\rangle + \frac{\partial \langle F \rangle}{\partial z_i}\right] + J\dot{\phi}$$
$$+\dot{J}\left[\left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial J}\right\rangle + \frac{\partial \langle F \rangle}{\partial J}\right]$$
$$-\langle H \rangle + \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t}\right\rangle + \frac{\partial \langle F \rangle}{\partial t}. \tag{149}$$

This Lagrangian obviously does not depend on ϕ.

5.2   Gauge invariance of the Lagrangian

Under a gauge transformation,

$$\begin{aligned}
\mathbf{A}\left(\mathbf{x}, t\right) &\rightarrow \mathbf{A}\left(\mathbf{x}, t\right) + \boldsymbol{\nabla}\Psi\left(\mathbf{x}, t\right) \\
\Phi\left(\mathbf{x}, t\right) &\rightarrow \Phi\left(\mathbf{x}, t\right) - (1/c)\left(\partial\Psi\left(\mathbf{x}, t\right)/\partial t\right) \\
\langle F \rangle\left(z_1, \ldots, z_4, J; t\right) &\rightarrow \langle F \rangle - (e/c)\left\langle \Psi\left(\mathfrak{x}\left(z_1, \ldots, z_4, J, \phi\,; t\right), t\right)\right\rangle, \tag{150}
\end{aligned}$$

the ϕ-independent Lagrangian L is left unchanged. Explicitly, the changes of the different terms are:

$$\left\langle \frac{\partial \mathfrak{x}}{\partial z_i} \cdot \mathfrak{p}\right\rangle \rightarrow \left\langle \frac{\partial \mathfrak{x}}{\partial z_i} \cdot \mathfrak{p}\right\rangle + \frac{e}{c}\left\langle \frac{\partial \mathfrak{x}}{\partial z_i} \cdot \frac{\partial\Psi\left(\mathfrak{x}, t\right)}{\partial \mathfrak{x}} \cdot\right\rangle, \tag{151}$$

$$\left\langle \frac{\partial \mathfrak{x}}{\partial J} \cdot \mathfrak{p}\right\rangle \rightarrow \left\langle \frac{\partial \mathfrak{x}}{\partial J} \cdot \mathfrak{p}\right\rangle + \frac{e}{c}\left\langle \frac{\partial \mathfrak{x}}{\partial J} \cdot \frac{\partial\Psi\left(\mathfrak{x}, t\right)}{\partial \mathfrak{x}} \cdot\right\rangle, \tag{152}$$

$$-\langle H \rangle + \left\langle \frac{\partial \mathfrak{x}}{\partial t} \cdot \mathfrak{p}\right\rangle \rightarrow -\langle H \rangle + \frac{e}{c}\left\langle \frac{\partial\Psi\left(\mathfrak{x}, t\right)}{\partial t}\right\rangle + \frac{e}{c}\left\langle \frac{\partial \mathfrak{x}}{\partial t} \cdot \frac{\partial\Psi\left(\mathfrak{x}, t\right)}{\partial \mathfrak{x}}\right\rangle. \tag{153}$$

The Ψ-dependent terms in these relations add in the Lagrangian to $(e/c)\, d\langle\Psi\rangle/dt$ and thus compensate the $\langle F \rangle$-dependent change in L.

5.3 Reduced phase-space Lagrangian

The simplest expression for the Lagrangian L is obtained by choosing the free function $\langle F \rangle$ as

$$\langle F \rangle = 0. \tag{154}$$

With this choice of $\langle F \rangle$ the Lagrangian becomes

$$L\left(z_1, \ldots, z_4, J; \dot{z}_1, \ldots \dot{z}_4, \dot{J}, \dot{\phi}; t\right) =$$

$$\sum_{i=1}^{4} \dot{z}_i \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} \right\rangle + J\dot{\phi}$$

$$+ \dot{J} \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial J} \right\rangle$$

$$- \langle H \rangle + \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t} \right\rangle . \tag{155}$$

This agrees with Larsson's results [20] obtained in a somewhat different way within the framework of guiding-centre theory. Of course, it would be perfectly valid to choose any arbitrary $\langle F \rangle \neq 0$ in (5.14). This would change the form of the different terms in the Lagrangian leaving, however, the equations of motion unchanged. In particular, one could obtain different contributions for the symplectic and the Hamiltonian parts of the Lagrangian leading, within the framework of certain approximations, to the two (still gauge-invariant) different expressions derived by Brizard in [5], Eqs. (6), (7) and (10), (11), respectively.

For the applications envisaged one is not interested in the dynamics of ϕ. It then suffices to describe the reduced dynamics involving only $z_1, \ldots, z_4$, with the constant of motion J having the character of a parameter. This is possible since L does not depend explicitly on ϕ. Such a description can be made in terms of a reduced Lagrangian L_r which is obtained from the Lagrangian L, (155), by omitting the term $J\dot{\phi}$ and the terms proportional to $\dot{J}$. It must, however, be stressed that it is easy to keep the terms which describe the fast gyromotion [22] and to obtain the Euler-Lagrange equation for $\dot{\phi}$ and $\dot{J}$, the last one being just $\dot{J} = 0$. Here we choose not to do this, since such terms shall not be needed.

With the total time derivative in the reduced phase space given by

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \sum_{j=1}^{4} \dot{z}_j \frac{\partial}{\partial z_j}, \tag{156}$$

the Euler-Lagrange equations for the z_i obtained from the Lagrangian L_r,

$$L_\mathrm{r}\left(z_1, \ldots, z_4; J; \dot{z}_1, \ldots, \dot{z}_4; t\right) :=$$

$$\sum_{i=1}^{4} \dot{z}_i \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} \right\rangle$$

$$-\langle H \rangle + \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t} \right\rangle , \tag{157}$$

are the same as those obtained from (155).

There are no approximations in the Lagrangian L_{r} from which $\dot{z}_1, \ldots, \dot{z}_4$ are derived. In particular, there is no approximation concerning ϕ. Any results obtained from this Lagrangian are therefore exact.

For the Dirac Hamiltonian the solutions $\dot{z}_i \, (z_1, \cdots, z_4, J, t)$ of the Euler-Lagrange equations corresponding to (157) are needed. These will be obtained in the next section by using an appropriate vector notation.

5.4   Vector representation of the reduced Lagrangian and the Euler-Lagrange equations

For the further development of the theory it will be convenient to express the Lagrangian L_{r} and the corresponding Euler-Lagrange equations in vector notation. For this purpose, it is useful to introduce a vector $\mathbf{R}$ with three Cartesian components,

$$\mathbf{R} = (R_1, R_2, R_3) \quad , \quad R_i := z_i \quad , \quad i = 1, 2, 3 \quad . \tag{158}$$

The fourth coordinate in the reduced phase space will be designated by U,

$$U := z_4. \tag{159}$$

The gradient with respect to $\mathbf{R}$ will be denoted by both $\partial/\partial \mathbf{R}$ and $\boldsymbol{\nabla}_{\mathbf{R}}$, as it appears convenient. The reduced Lagrangian L_{r}, (157), can then be written as

$$L_{\mathrm{r}} \left(\mathbf{R}, U; J; \dot{\mathbf{R}}, \dot{U}; t \right) =$$
$$\frac{e}{c} \left[\dot{\mathbf{R}} \cdot \hat{\mathbf{A}} \left(\mathbf{R}, U; J; t \right) + \dot{U} \hat{A}_{\mathrm{U}} \right] - e \hat{\Phi} \left(\mathbf{R}, U; J; t \right) , \tag{160}$$

where

$$\hat{\mathbf{A}} \left(\mathbf{R}, U; J; t \right) := \frac{c}{e} \left\langle \frac{\partial \mathfrak{x}}{\partial \mathbf{R}} \cdot \mathfrak{p} \right\rangle , \tag{161}$$

$$\hat{A}_{\mathrm{U}} \left(\mathbf{R}, U; J; t \right) := \frac{c}{e} \left\langle \frac{\partial \mathfrak{x}}{\partial U} \cdot \mathfrak{p} \right\rangle , \tag{162}$$

$$e \hat{\Phi} \left(\mathbf{R}, U; J; t \right) := \langle H \rangle - \left\langle \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t} \right\rangle . \tag{163}$$

The Euler-Lagrange equation obtained by varying $\mathbf{R}$ is

$$-\frac{1}{c} \frac{d}{dt} \hat{\mathbf{A}} + \frac{1}{c} \left(\frac{\partial \hat{\mathbf{A}}}{\partial \mathbf{R}} \right) \cdot \dot{\mathbf{R}} + \frac{1}{c} \left(\frac{\partial \hat{A}_{\mathrm{U}}}{\partial \mathbf{R}} \right) \dot{U} - \frac{\partial \hat{\Phi}}{\partial \mathbf{R}} = 0. \tag{164}$$

The Euler-Lagrange equation obtained by varying U is

$$-\frac{1}{c}\frac{d}{dt}\hat{A}_U + \frac{1}{c}\left(\frac{\partial \hat{\mathbf{A}}}{\partial U}\right)\cdot\dot{\mathbf{R}} + \frac{1}{c}\left(\frac{\partial \hat{A}_U}{\partial U}\right)\dot{U} - \frac{\partial \hat{\Phi}}{\partial U} = 0\ . \tag{165}$$

Since $\dot{J} = 0$, the total time derivative for ϕ-independent functions is given by

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \dot{\mathbf{R}}\cdot\frac{\partial}{\partial \mathbf{R}} + \dot{U}\frac{\partial}{\partial U}. \tag{166}$$

With the fields $\hat{\mathbf{B}}$, $\hat{\mathbf{E}}$ and $\mathbf{b}_U$ defined by

$$\hat{\mathbf{B}} := \boldsymbol{\nabla}_{\mathbf{R}}\times\hat{\mathbf{A}}, \tag{167}$$

$$\hat{\mathbf{E}} := -\frac{1}{c}\frac{\partial \hat{\mathbf{A}}}{\partial t} - \boldsymbol{\nabla}_{\mathbf{R}}\hat{\Phi}, \tag{168}$$

$$\mathbf{b}_U := \frac{e}{mc}\left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}}\right), \tag{169}$$

(164) and (165) become, respectively,

$$\frac{1}{c}\dot{\mathbf{R}}\times\hat{\mathbf{B}} + \hat{\mathbf{E}} - \frac{m}{e}\dot{U}\,\mathbf{b}_U = 0 \tag{170}$$

and

$$\frac{m}{e}\dot{\mathbf{R}}\cdot\mathbf{b}_U - \frac{\partial \hat{\Phi}}{\partial U} - \frac{1}{c}\frac{\partial \hat{A}_U}{\partial t} = 0. \tag{171}$$

After crossing (170) with $\mathbf{b}_U$ and inserting $\dot{\mathbf{R}}\cdot\mathbf{b}_U$ from (171), one can solve for $\dot{\mathbf{R}}$. With the definition

$$\mathcal{B}\left(\mathbf{R}, U, J; t\right) := \hat{\mathbf{B}}\cdot\mathbf{b}_U = \hat{\mathbf{B}}\cdot\frac{e}{mc}\left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}}\right) \tag{172}$$

one obtains the velocity $\dot{\mathbf{R}}$,

$$\dot{\mathbf{R}} = \mathbf{V}_{\mathbf{R}}\left(\mathbf{R}, U, J; t\right), \tag{173}$$

with

$$\mathbf{V}_{\mathbf{R}}\left(\mathbf{R}, U, J; t\right) = \frac{e}{m\mathcal{B}}\left(\frac{\partial \hat{\Phi}}{\partial U} + \frac{1}{c}\frac{\partial \hat{A}_U}{\partial t}\right)\hat{\mathbf{B}} + \frac{c}{\mathcal{B}}\hat{\mathbf{E}}\times\mathbf{b}_U. \tag{174}$$

The scalar product of (170) with $\hat{\mathbf{B}}$ yields

$$\dot{U} = V_U\left(\mathbf{R}, U, J; t\right), \tag{175}$$

where

$$V_U\left(\mathbf{R}, U, J; t\right) = \frac{e}{m\mathcal{B}}\hat{\mathbf{B}}\cdot\hat{\mathbf{E}}. \tag{176}$$

which, using (174), can also be written in the alternative form

$$V_U = \frac{\mathbf{V_R} \cdot \hat{\mathbf{E}}}{\dfrac{\partial \hat{\Phi}}{\partial U} + \dfrac{1}{c} \dfrac{\partial \hat{A}_{\mathrm{U}}}{\partial t}}.$$ (177)

By replacing $\hat{\mathbf{E}}$ according to (168), (177) yields the energy-related equation,

$$\frac{d\left(e\,\hat{\Phi}\right)}{dt} = -\mathbf{V_R} \cdot \frac{e}{c} \frac{\partial \hat{\mathbf{A}}}{\partial t} - \frac{e}{c} V_U \frac{\partial \hat{A}_{\mathrm{U}}}{\partial t} + \left.\frac{\partial\left(e\,\hat{\Phi}\right)}{\partial t}\right|_{\mathbf{R},U,J}.$$ (178)

Thus, $\hat{\Phi}$ is conserved when $\hat{\mathbf{A}}$, $\hat{A}_{\mathrm{U}}$ and $\hat{\Phi}$ do not depend explicitly on time and $e\,\hat{\Phi} = \langle H \rangle$.

It is seen that one advantage of using the vector representation is that it is relatively easy to solve for the time derivatives $\dot{\mathbf{R}}$ and $\dot{U}$ as functions of $\mathbf{R}, U; J; t$, while the (equivalent) solution of the equations for $\dot{z}_j\left(z_1, \ldots, z_4; J; t\right)$, $\quad j = 1, \ldots, 4$, obtained from the Lagrangian L_{r}, (157), is more involved.

5.4.1 *The Littlejohn Lagrangian*

Littlejohn's Lagrangian, basically in the form of [3], (20) but with the electric drift being taken into account, can readily be derived from the vector representation for the reduced Lagrangian obtained in the preceding section and from the results of Secs. 4.3.5 and 4.3.6. In the following, again, an upper index (n) indicates that a quantity has been obtained in the n-th iteration and thus includes all orders below and equal to (n). These upper indices must not be mixed up with the lower indices indicating the order in ϵ. One has

$$\mathbf{p}_{-1} = \frac{e}{c} \mathbf{A}\left(\mathbf{R}, t\right), \quad \Phi_{-1} = \Phi\left(\mathbf{R}, t\right),$$

$$\mathbf{x}^{(0)} = \mathbf{R}, \quad u_\parallel^{(0)} = U, \quad u_\perp^{(0)} = \sqrt{\frac{2}{m} \frac{\Omega_s}{2\pi} \frac{J^{(1)}}{\epsilon}}, \quad \theta^{(0)} = \phi,$$

$$\mathbf{v}_E^{(0)} = c\frac{\mathbf{E}\left(\mathbf{R}, t\right) \times \mathbf{B}\left(\mathbf{R}, t\right)}{B^2},$$

$$\mathbf{x}_1 = \sqrt{\frac{2}{m} \frac{1}{2\pi\Omega_s} \frac{J^{(1)}}{\epsilon}}\,\mathbf{n}_2\left(\mathbf{R}, \phi, t\right),$$

$$\mathbf{p}^{(0)} = m\left[U\,\mathbf{b}\left(\mathbf{R}, t\right) + \sqrt{\frac{2}{m} \frac{\Omega_s}{2\pi} \frac{J^{(1)}}{\epsilon}}\,\mathbf{n}_1\left(\mathbf{R}, \phi, t\right) + \mathbf{v}_E^{(0)}\right] + \frac{e}{c}\mathbf{A}\left(\mathbf{R}, t\right)$$

$$+ \frac{e}{c}\mathbf{x}_1 \cdot \boldsymbol{\nabla}_{\mathbf{R}}\mathbf{A}\left(\mathbf{R}, t\right),$$

$$H^{(0)} = \frac{m}{2}\left[U\,\mathbf{b}\left(\mathbf{R}, t\right) + \sqrt{\frac{2}{m} \frac{\Omega_s}{2\pi} \frac{J^{(1)}}{\epsilon}}\,\mathbf{n}_1\left(\mathbf{R}, \phi, t\right) + \mathbf{v}_E^{(0)}\right]^2 + e\,\Phi\left(\mathbf{R}, t\right)$$

$$+ e\,\mathbf{x}_1 \cdot \boldsymbol{\nabla}_{\mathbf{R}}\Phi\left(\mathbf{R}, t\right).$$ (179)

With these relations, one obtains:

$$
\begin{aligned}
\hat{\mathbf{A}}^{(0)} &= \frac{c}{e}\left\langle \mathbf{p}^{(0)} \right\rangle = \frac{c\,m}{e}\left[U\,\mathbf{b}\left(\mathbf{R},t\right) + \mathbf{v}_E^{(0)} \right] + \mathbf{A}\left(\mathbf{R},t\right), \\
\hat{A}_{\mathrm{U}}^{(0)} &= 0, \\
e\,\hat{\Phi}^{(0)} &= \left\langle H^{(0)} \right\rangle = \frac{m}{2}\left[U^2 + \left(\mathbf{v}_E^{(0)}\right)^2 \right] + \frac{1}{2\pi}\frac{J^{(1)}}{\epsilon}\Omega_s + e\,\Phi\left(\mathbf{R},t\right). \quad (180)
\end{aligned}
$$

With the replacement

$$
J^{(1)} = \frac{2\pi}{e} mc\,\mu^{(1)},
$$

this yields the Lagrangian

$$
\begin{aligned}
L_{\mathrm{Littlejohn}} &= \dot{\mathbf{R}} \cdot \left[m\left[U\,\mathbf{b}\left(\mathbf{R},t\right) + \mathbf{v}_E^{(0)} \right] + \frac{e}{c}\mathbf{A}\left(\mathbf{R},t\right) \right] \\
&\quad - \frac{m}{2}\left[U^2 + \left(\mathbf{v}_E^{(0)}\right)^2 \right] - \mu^{(1)}B - e\Phi\left(\mathbf{R},t\right). \quad (181)
\end{aligned}
$$

Gyrocentre particle Lagrangians similar to those derived by Brizard [5,12] can be obtained, using appropriate approximations, from our gyroangle-independent Lagrangian by a procedure similar to that used to obtain Littlejohn's guiding-centre Lagrangian. However, particularly as concerns the gyrocentre case, but also for the guiding-centre approximation, we prefer not make any further specific approximations at this stage. Leaving the expressions as generally valid as possible has two important advantages. First, our treatment of gauge-invariant Maxwell-kinetic theories remains relatively clear and tractable and, second, the final results concerning the kinetic equations, the conservations laws and the energy-momentum tensor are obtained in a form that comprises several different theories which can then be obtained by introducing in those expressions the desired approximations, as will be seen in a following paper [29].

6 Maxwell-kinetic theory in the reduced phase space

In this section the formalism of Sec. 2 is specialized to the reduced Lagrangian in the form of (160). The situation with this Lagrangian is the following:

6.1 Quantities entering the Lagrangian for the kinetic theory

6.1.1 *The canonical momenta in the reduced phase space*

These momenta are

$$
\mathbf{P}_{\mathbf{R}} = \frac{\partial L_{\mathrm{r}}}{\partial \dot{\mathbf{R}}} = \frac{e}{c}\hat{\mathbf{A}}, \quad (182)
$$

$$
P_U = \frac{\partial L_{\mathrm{r}}}{\partial \dot{U}} = \frac{e}{c}\hat{A}_{\mathrm{U}}. \quad (183)
$$

6.1.2 *Dirac's constrained Hamiltonian*

Therefore, the primary Hamiltonian is

$$
\begin{aligned}
H_p &= \dot{\mathbf{R}} \cdot \mathbf{P_R} + \dot{U} P_U - L_{\mathrm{r}} \\
&= e\hat{\Phi},
\end{aligned}
\tag{184}
$$

and Dirac's Hamiltonian becomes

$$
H_{\mathrm{D}}\left(\mathbf{R}, U; \mathbf{P}, P_U; J; t\right) := e\hat{\Phi} + \mathbf{V_R} \cdot \left(\mathbf{P_R} - \frac{e}{c}\hat{\mathbf{A}}\right) + V_U \left(P_U - \frac{e}{c}\hat{A}_{\mathrm{U}}\right), \tag{185}
$$

with $\hat{\mathbf{A}}\left(\mathbf{R}, U; J; t\right)$, $\hat{A}_{\mathrm{U}}\left(\mathbf{R}, U; J; t\right)$ and $\hat{\Phi}\left(\mathbf{R}, U; J; t\right)$ from (161), (162) and (163), respectively, and $\mathbf{V_R}\left(\mathbf{R}, U; J; t\right)$ and $V_U\left(\mathbf{R}, U; J; t\right)$ given by (174) and (176).

6.2 Maxwell-kinetic theory

6.2.1 *The Lagrangian*

The total Lagrangian for the Maxwell-kinetic theory according to (1), (2) and (25) is now

$$
L_{\mathrm{tot}}\left(t\right) = L_{\mathrm{M}} + L_{\mathrm{K}}, \tag{186}
$$

$$
L_{\mathrm{M}} = \int d^3x \, \frac{1}{8\pi} \left[\mathbf{E}^2\left(\mathbf{x}, t\right) - \mathbf{B}^2\left(\mathbf{x}, t\right)\right], \tag{187}
$$

$$
\begin{aligned}
L_{\mathrm{K}} &= - \sum_{\text{particle species}} \int d^3R \, dU \, d^4\alpha \, dJ \, f_{\mathrm{p}}\left(\mathbf{R}, U; \alpha_i; J; t\right) \left[\frac{\partial S}{\partial t}\left(\mathbf{R}, U; \alpha_i; J; t\right) + e\hat{\Phi} \right. \\
&\quad \left. + \mathbf{V_R} \cdot \left(\frac{\partial S}{\partial \mathbf{R}} - \frac{e}{c}\hat{\mathbf{A}}\right) + V_U \left(\frac{\partial S}{\partial U} - \frac{e}{c}\hat{A}_{\mathrm{U}}\right)\right],
\end{aligned}
\tag{188}
$$

with $\hat{\mathbf{A}}$, $\hat{A}_{\mathrm{U}}$ and $\hat{\Phi}$ given by (161), (162) and (163) and $\mathbf{V_R}$ and V_U given by (174) and (176), respectively, and with the integration over J representing the summation over the parameter J (cf. Sec. 5.3).

6.2.2 *Liouville's theorem*

The quantity λ, introduced in (40) for general coordinates ζ_i, for obtaining a volume element $d\tau$ in the $\mathbf{R}$, U, J phase space which is conserved along the orbits in $\mathbf{R}$, U, J phase space can easily be derived in the following way. From (167), (168) and (173)-(176) one finds

$$
\frac{\partial \mathcal{B}}{\partial t} + \boldsymbol{\nabla}_{\mathbf{R}} \cdot \left(\mathcal{B}\mathbf{V_R}\right) + \frac{\partial}{\partial U}\left(\mathcal{B}V_U\right) = 0. \tag{189}
$$

The quantity $\mathcal{B}$ thus satisfies a continuity equation in the reduced phase space. Therefore, the phase-space volume element $d\tau_{\mathbf{R},U}$ defined as

$$d\tau_{\mathbf{R},U} := |\mathcal{B}(R_1, \ldots, R_3, U; J; t)| \, dR_1 dR_2 dR_3 dU \tag{190}$$

is Liouvillian, i.e.

$$\frac{d}{dt}(d\tau_{\mathbf{R},U}) = 0, \tag{191}$$

where the time derivative along orbits in phase space is given by (166). Since J is only a parameter which is constant along the orbits, the same is also valid for the volume element $d\tau$ in extended phase space which is obtained by adding the parameter J,

$$d\tau := |\mathcal{B}(R_1, \ldots, R_3, U; J; t)| \, dR_1 dR_2 dR_3 dU \, dJ. \tag{192}$$

Hence, it holds that

$$\lambda = |\mathcal{B}(R_1, \ldots, R_3, U; J; t)|. \tag{193}$$

An alternative derivation of λ is given in Appendix C.

6.2.3 *The kinetic equation*

The kinetic equation for the distribution function f is obtained from the Lagrangian (188) by variation of S, in a way similar to that in Sec. 2.4, (43). One obtains

$$\frac{\partial f}{\partial t} + \mathbf{V}_{\mathbf{R}} \cdot \frac{\partial f}{\partial \mathbf{R}} + V_U \frac{\partial f}{\partial U} = 0, \tag{194}$$

with $\mathbf{V}_{\mathbf{R}}$ and V_U given by (174) and (176), respectively.

6.2.4 *Sketch of a procedure for obtaining the inhomogeneous Maxwell equations and conservation laws*

Variation of the total action integral

$$\mathcal{A} = \int_{t_1}^{t_2} L_{\text{tot}} dt \tag{195}$$

with respect to the potentials $\Phi(\mathbf{x}, t)$ and $\mathbf{A}(\mathbf{x}, t)$ yields the expressions for the charge and current densities, respectively. By means of the Noether formalism one can also obtain the energy-momentum tensor density. In these expressions one has to replace

$$f_{\mathrm{p}} \, d^4\alpha \;\; \rightarrow \;\; \mathcal{B} \, \delta\left(\mathbf{P}_{\mathbf{R}} - \frac{e}{c}\hat{\mathbf{A}}\right) \delta\left(P_U - \frac{e}{c}\hat{A}_{\mathrm{U}}\right) \, f(\mathbf{R}, U; J; t) \, d^3 P_{\mathbf{R}} \, dP_U,$$

$$\frac{\partial S}{\partial \mathbf{R}} \;\; \rightarrow \;\; \mathbf{P}_{\mathbf{R}},$$

$$\frac{\partial S}{\partial U} \;\; \rightarrow \;\; P_U, \tag{196}$$

where $\mathcal{B}$ is given by (172).

This program will be carried out in a following paper for the drift-kinetic and the gyrokinetic theory.

7 Summary

The paper starts with the introduction and proof of the correctness of a Lagrangian for combined Maxwell-kinetic theories in general coordinates as concerns the particle motion. The kinetic part of it is formulated Eulerian with the help of the equations of motion in form of Hamilton-Jacobi's equation as a tool and Dirac's constraint theory. In all final expressions such as charge and current densities as well as the energy-momentum tensor density and the kinetic equations the Hamilton-Jacobi function is eliminated. There is correspondingly no need to solve the Hamilton-Jacobi equation. Charge and current densities distinguish automatically between "particle-like" (gyration centre), polarization and magnetization contributions. This formalism is applied to averaging coordinates derived by Kruskal's method. With the help of certain properties of the averaging coordinates according to the basic requirements imposed on them, it was possible to obtain a Lagrangian for the combined Maxwell-kinetic theory in a reduced phase space which is applicable to situations in which one is not interested in the dependence on some kind of gyroangle describing the gyromotion, whose treatment, however, can easily be added [22]. There is no approximation with this Lagrangian. This reduced phase-space Lagrangian is obtained from a form of the exact particle Lagrangian which is manifestly independent of the gyrophase ϕ.

The basic perturbation theory, which aims at obtaining averaging phase-space coordinates, is done solely within the framework of the inverse Kruskal formalism given explicitly in this paper. The results are later used to obtain approximate expressions for the Lagrange functions. For the definition of certain approximations to the exact Lagrangian the vector potential is formally treated as scaling with the magnetic field B times the scale length of the background plasma, and the electric potential Φ is formally treated as scaling with B, but there is no need to really attribute certain orders to the potentials by a corresponding choice of gauge. The terms of the needed zeroth and first orders are given. For the drift-kinetic ordering Littlejohn's Lagrangian is rederived immediately and with this the drift-kinetic theory as obtained and investigated by the present authors in some previous work. Gyrocentre particle Lagrangians similar to those derived by Brizard [5, 12] can be obtained, using appropriate approximations, from our gyroangle-independent Lagrangian by a procedure similar to that used to obtain Littlejohn's guiding-centre Lagrangian. However, particularly as concerns the gyrocentre case, but also for the guiding-centre approximation, we prefer not make any further specific approximations at this stage. Leaving the expressions as generally valid as possible has two important advantages. First, our treatment of gauge-invariant Maxwell-kinetic theories remains relatively clear and tractable and, second, the final results concerning the kinetic equations, the conservations laws and the energy-momentum tensor are obtained in a form that comprises several different theories which can then be obtained by introducing in those expres-

sions the desired approximations, as will be seen in a following paper [29]. That paper will apply the above exact Lagrangian to the combined Maxwell-kinetic theories, with the gyroradius approximated by its first-order Kruskal expression. In that paper one will obtain by means of a modified Noether formalism which is gauge invariant [30] all the local conservation laws in a form which exhibits in a very transparent way their structures. In the gyrokinetic theory one will encounter complications compared with the linearized theory for the above Lagrangian and with the drift-kinetic ordering which result from the occurrence of first-order terms in the arguments of the potentials. This will lead to certain averages of these potentials over the gyromotion similar as in other present-day theories. The most important point concerns the energy-momentum tensor. Although the gauge-invariant Noether formalism automatically yields the fully symmetric energy-momentum tensor for the relativistic Maxwell-Vlasov theory and the energy-momentum tensor symmetric in the spatial indices for the non-relativistic Maxwell-Vlasov theory, as shown in [30], and also for an approximate theory such as the drift-kinetic theory, as shown in [10], this is, in general, not the case in other approximate theories. To these theories belongs the gyrokinetic theory. The nonsymmetric terms in the energy-momentum tensor are of higher order and can therefore disappear when going to higher order approximations. However, it will be shown in the already mentioned following paper that even within the framework of the gyrokinetic theory the nonsymmetric terms can easily be eliminated without having to consider higher-order approximations. The reason for this is that the gauge-invariant Noether formalism yields expressions for the momentum and angular momentum conservation laws from which one gets a clear indication how to easily derive the symmetric energy-momentum tensor. This tensor describes the correct energy and momentum densities and their corresponding flux densities. The final expressions for these densities have a physically clear and transparent structure. In particular, they contain spin-like contributions everywhere, except in the energy density. All the results concerning the gyrokinetic energy-momentum tensor are new and cannot be found elsewhere.

Appendix A. Direct proof of $\frac{d}{dt} J(\hat{\mathbf{z}}, t) = 0$

$J(\hat{z}_1, \ldots, \hat{z}_5; t)$ is defined as

$$J(\hat{\mathbf{z}}, t) = \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial \phi} \, d\phi. \tag{A1}$$

Its time derivative is

$$\begin{aligned}
\frac{d}{dt} J(\hat{\mathbf{z}}, t) &= \dot{\hat{z}}_i \frac{\partial}{\partial \hat{z}_i} \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial \phi} \, d\phi + \frac{\partial}{\partial t} \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial \phi} \, d\phi \\
&= h_i \frac{\partial}{\partial \hat{z}_i} \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial \phi} \, d\phi + \frac{\partial}{\partial t} \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial \phi} \, d\phi
\end{aligned}$$

$$= \int_0^1 \left[\left(h_i \frac{\partial \mathbf{p}}{\partial \hat{z}_i} + \frac{\partial \mathbf{p}}{\partial t} \right) \cdot \frac{\partial \mathbf{x}}{\partial \phi} + \mathbf{p} \cdot \left(h_i \frac{\partial}{\partial \hat{z}_i} + \frac{\partial}{\partial t} \right) \frac{\partial \mathbf{x}}{\partial \phi} \right] d\phi$$

$$= \int_0^1 \left[\left(h_i \frac{\partial \mathbf{p}}{\partial \hat{z}_i} + \frac{\partial \mathbf{p}}{\partial t} \right) \cdot \frac{\partial \mathbf{x}}{\partial \phi} - \frac{\partial \mathbf{p}}{\partial \phi} \cdot \left(h_i \frac{\partial \mathbf{x}}{\partial \hat{z}_i} + \frac{\partial \mathbf{x}}{\partial t} \right) \right] d\phi. \quad \text{(A2)}$$

This relation can also be written in the form

$$\frac{d}{dt} J(\hat{\mathbf{z}}, t) = \int_0^1 \left[\left(\frac{\partial \mathbf{p}}{\partial t} + h_i \frac{\partial \mathbf{p}}{\partial \hat{z}_i} + \frac{\omega(\hat{\mathbf{z}})}{\epsilon} \frac{\partial \mathbf{p}}{\partial \phi} \right) \cdot \frac{\partial \mathbf{x}}{\partial \phi} \right.$$
$$\left. - \frac{\partial \mathbf{p}}{\partial \phi} \cdot \left(\frac{\partial \mathbf{x}}{\partial t} + h_i \frac{\partial \mathbf{x}}{\partial \hat{z}_i} + \frac{\omega(\hat{\mathbf{z}})}{\epsilon} \frac{\partial \mathbf{x}}{\partial \phi} \right) \right] d\phi. \quad \text{(A3)}$$

The additional terms obviously cancel each other. The function $\omega(\hat{\mathbf{z}})/\epsilon$ is the time derivative of ϕ. The expressions in parentheses are therefore the total time derivatives of $\mathbf{p}$ and $\mathbf{x}$. They can be expressed by Hamilton's equations of motion as

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{x}} \;, \quad \frac{d\mathbf{x}}{dt} = \frac{\partial H}{\partial \mathbf{p}} \;.$$

Equation (A3) can therefore be written as

$$\frac{d}{dt} J(\hat{\mathbf{z}}, t) = \int_0^1 \left[-\frac{\partial H}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \phi} - \frac{\partial \mathbf{p}}{\partial \phi} \cdot \frac{\partial H}{\partial \mathbf{p}} \right] d\phi = \int_0^1 \left[-\frac{\partial H}{\partial \phi} \right] d\phi = 0. \quad \text{(A4)}$$

Note that it was essential that $\mathbf{h}$, (106), should not depend on ϕ, and that it was not required that H be independent of ϕ, which it in general is not.

Appendix B. Properties of some Lagrange brackets

Equations (146)-(148) are derived in this appendix.
In the phase-space coordinates $z_1, \ldots, z_4, J, \phi$ the constant of the motion J, (107), can be written as

$$J = \oint_{\{z_1,\ldots,z_4,J,t\}=const} \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} \, d\phi = \int_0^1 \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} \, d\phi. \quad \text{(B1)}$$

Partial derivation of J with respect to z_i, J and t yields, respectively,

$$0 = \int_0^1 \left[\frac{\partial \mathfrak{p}}{\partial z_i} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} - \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} \right] d\phi := \int_0^1 l_i d\phi, \quad \text{(B2)}$$

$$1 = \int_0^1 \left[\frac{\partial \mathfrak{p}}{\partial J} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} - \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{x}}{\partial J} \right] d\phi := \int_0^1 l_J d\phi \quad \text{(B3)}$$

and

$$0 = \int_0^1 \left[\frac{\partial \mathfrak{p}}{\partial t} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} - \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{x}}{\partial t} \right] d\phi := \int_0^1 l_t \, d\phi, \tag{B4}$$

where the convenient notation l_i, l_J, l_t has been introduced for the Lagrange brackets concerned. Thus, $l_i, l_J - 1$ and l_t have vanishing mean values. It will now be shown that also their oscillating parts vanish. The calculations are made with basically the same methods used to derive the transformation equations (109)-(112) in Sec. 4, but employing the phase-space coordinates $z_1, \ldots, z_4, J, \phi$ instead of $\hat{z}_1, \ldots, \hat{z}_5, \phi$.

Some relations needed for the calculations can be derived from the canonical Hamiltonian H,

$$H\left(\mathbf{x}, \mathbf{p}, t = \epsilon s\right) = H\left(\mathfrak{x}\left(z_1, \ldots, z_4, J, \phi; t = \epsilon s\right), \mathfrak{p}\left(z_1, \ldots, z_4, J, \phi; t = \epsilon s\right); t = \epsilon s\right), \tag{B5}$$

from which one obtains as derivatives with respect to the independent variables $z_1, \ldots, z_4, J, \phi$ and t the relations

$$\frac{\partial H}{\partial z_i} = \left. \frac{\partial H}{\partial \mathfrak{x}} \right|_{\mathfrak{p},t} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} + \left. \frac{\partial H}{\partial \mathfrak{p}} \right|_{\mathfrak{x},t} \cdot \frac{\partial \mathfrak{p}}{\partial z_i}$$

$$= -\frac{d\mathfrak{p}}{dt} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} + \frac{d\mathfrak{x}}{dt} \cdot \frac{\partial \mathfrak{p}}{\partial z_i}, \tag{B6}$$

$$\frac{\partial H}{\partial J} = -\frac{d\mathfrak{p}}{dt} \cdot \frac{\partial \mathfrak{x}}{\partial J} + \frac{d\mathfrak{x}}{dt} \cdot \frac{\partial \mathfrak{p}}{\partial J}, \tag{B7}$$

$$\frac{\partial H}{\partial \phi} = -\frac{d\mathfrak{p}}{dt} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} + \frac{d\mathfrak{x}}{dt} \cdot \frac{\partial \mathfrak{p}}{\partial \phi} \tag{B8}$$

and

$$\left. \frac{\partial H}{\partial t} \right|_{z_i,J,\phi} = -\frac{d\mathfrak{p}}{dt} \cdot \frac{\partial \mathfrak{x}}{\partial t} + \frac{d\mathfrak{x}}{dt} \cdot \frac{\partial \mathfrak{p}}{\partial t} + \frac{dH}{dt}, \tag{B9}$$

where $\left. \frac{\partial H}{\partial t} \right|_{\mathfrak{x},\mathfrak{p}} = \frac{dH}{dt}$ has been used in the last equation. With (B6) and $\frac{\partial}{\partial \phi}\left(\frac{\partial H}{\partial z_i}\right) = \frac{\partial}{\partial z_i}\left(\frac{\partial H}{\partial \phi}\right)$ (and similarly for J and t) taken into account, differentiation of (B6), (B7) and (B9) with respect to ϕ yields

$$-\frac{\partial \mathfrak{x}}{\partial z_i} \cdot \frac{\partial}{\partial \phi}\left[\frac{d\mathfrak{p}}{dt}\right] + \frac{\partial \mathfrak{p}}{\partial z_i} \cdot \frac{\partial}{\partial \phi}\left[\frac{d\mathfrak{x}}{dt}\right] = -\frac{\partial \mathfrak{x}}{\partial \phi} \cdot \frac{\partial}{\partial z_i}\left[\frac{d\mathfrak{p}}{dt}\right] + \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial}{\partial z_i}\left[\frac{d\mathfrak{x}}{dt}\right], \tag{B10}$$

$$-\frac{\partial \mathfrak{x}}{\partial J} \cdot \frac{\partial}{\partial \phi}\left[\frac{d\mathfrak{p}}{dt}\right] + \frac{\partial \mathfrak{p}}{\partial J} \cdot \frac{\partial}{\partial \phi}\left[\frac{d\mathfrak{x}}{dt}\right] = -\frac{\partial \mathfrak{x}}{\partial \phi} \cdot \frac{\partial}{\partial J}\left[\frac{d\mathfrak{p}}{dt}\right] + \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial}{\partial J}\left[\frac{d\mathfrak{x}}{dt}\right] \tag{B11}$$

and

$$-\frac{\partial \mathfrak{x}}{\partial t} \cdot \frac{\partial}{\partial \phi}\left[\frac{d\mathfrak{p}}{dt}\right] + \frac{\partial \mathfrak{p}}{\partial t} \cdot \frac{\partial}{\partial \phi}\left[\frac{d\mathfrak{x}}{dt}\right] + \frac{\partial}{\partial \phi}\left[\frac{dH}{dt}\right]$$

$$= -\frac{\partial \mathfrak{x}}{\partial \phi} \cdot \frac{\partial}{\partial t}\left[\frac{d\mathfrak{p}}{dt}\right] + \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial}{\partial t}\left[\frac{d\mathfrak{x}}{dt}\right]. \tag{B12}$$

In the phase-space coordinates $z_1, \ldots, z_4, J, \phi$, the total time derivative is given by

$$\frac{d}{dt} = \sum_{j=1}^{4} h_j \left(z_1, \ldots, z_4, J, t = \epsilon s\right) \frac{\partial}{\partial z_j} + \frac{\omega}{\epsilon} \left(z_1, \ldots, z_4, J, t = \epsilon s\right) \frac{\partial}{\partial \phi} + \frac{\partial}{\partial t} \quad \text{(B13)}$$

(cf. (106). It is easily seen that

$$\frac{\partial}{\partial \phi} \frac{d}{dt} = \frac{d}{dt} \frac{\partial}{\partial \phi}, \tag{B14}$$

$$\frac{\partial}{\partial z_i} \frac{d}{dt} = \frac{d}{dt} \frac{\partial}{\partial z_i} + \sum_{j=1}^{4} \frac{\partial h_j}{\partial z_i} \frac{\partial}{\partial z_j} + \frac{1}{\epsilon} \frac{\partial \omega}{\partial z_i} \frac{\partial}{\partial \phi}, \tag{B15}$$

and correspondingly for $\partial/\partial J$ and $\partial/\partial t$. When this is taken into account, (B10)-(B12) yield, respectively,

$$\frac{d\mathbf{l}_{(\mathbf{z})}}{dt} = -\frac{\partial \mathbf{h}_{(\mathbf{z})}}{\partial \mathbf{z}} \cdot \mathbf{l}_{(\mathbf{z})}, \text{with } \mathbf{l}_{(\mathbf{z})} = \left(l_1, l_2, l_3, l_4\right), \ \mathbf{h}_{(\mathbf{z})} = \left(h_1, h_2, h_3, h_4\right), \quad \text{(B16)}$$

$$\frac{dl_J}{dt} = -\frac{\partial \mathbf{h}_{(\mathbf{z})}}{\partial J} \cdot \mathbf{l}_{(\mathbf{z})}, \tag{B17}$$

$$\frac{d}{dt} \left[l_t + \frac{\partial H}{\partial \phi}\right] = -\frac{\partial \mathbf{h}_{(\mathbf{z})}}{\partial t} \cdot \mathbf{l}_{(\mathbf{z})}. \tag{B18}$$

Since the mean values of $\mathbf{l}_{(\mathbf{z})}$ and l_t vanish, and the mean value of l_J is 1 according to (B2)-(B4), $\mathbf{l}_{(\mathbf{z})}$, l_J and l_t can be replaced by their oscillatory parts $\tilde{\mathbf{l}}_{(\mathbf{z})}$, $\tilde{l}_J$ and $\tilde{l}_t$ in (B16)-(B18). By substituting d/dt from (B13) in (B16) and solving for $\partial \tilde{\mathbf{l}}_{(\mathbf{z})}/\partial \phi$, one obtains

$$\frac{\partial \tilde{\mathbf{l}}_{(\mathbf{z})}}{\partial \phi} = -\frac{\epsilon}{\omega} \left[\frac{\partial \mathbf{h}_{(\mathbf{z})}}{\partial \mathbf{z}} \cdot \tilde{\mathbf{l}}_{(\mathbf{z})} + \mathbf{h}_{(\mathbf{z})} \cdot \frac{\partial \tilde{\mathbf{l}}_{(\mathbf{z})}}{\partial \mathbf{z}} + \frac{\partial \tilde{\mathbf{l}}_{(\mathbf{z})}}{\partial t}\right]. \tag{B19}$$

The iterative solution of this equation yields $\tilde{\mathbf{l}}_{(\mathbf{z})} = 0$ to all orders in ϵ. With this result taken into account, (B17) and (B18) yield in a similar way

$$\tilde{l}_J = 0 \quad , \quad \tilde{l}_t + \frac{\partial H}{\partial \phi} = 0. \tag{B20}$$

Therefore,

$$\frac{\partial \mathfrak{p}}{\partial z_i} \cdot \frac{\partial \mathfrak{r}}{\partial \phi} - \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{r}}{\partial z_i} = 0 \quad , \ i = 1, \ldots, 4 \ , \tag{B21}$$

$$\frac{\partial \mathfrak{p}}{\partial J} \cdot \frac{\partial \mathfrak{r}}{\partial \phi} - \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{r}}{\partial J} = 1, \tag{B22}$$

and

$$\frac{\partial \mathfrak{p}}{\partial t} \cdot \frac{\partial \mathfrak{r}}{\partial \phi} - \frac{\partial \mathfrak{p}}{\partial \phi} \cdot \frac{\partial \mathfrak{r}}{\partial t} + \frac{\partial H}{\partial \phi} = 0. \tag{B23}$$

40

It is easy to calculate the Poisson brackets $[z_i, J]$ and $[\phi, J]$ by making use of these results. Let q_λ be any of the coordinates $z_1, \ldots, z_4, J, \phi$. Then, from

$$q_\lambda = q_\lambda \left(\mathfrak{x} \left(q_\mu, t \right), \mathfrak{p} \left(q_\mu, t \right), t \right), \tag{B24}$$

one obtains

$$\frac{\partial q_\lambda}{\partial q_\mu} = \delta_{\lambda\mu} = \frac{\partial \mathfrak{x}}{\partial q_\mu} \cdot \frac{\partial q_\lambda}{\partial \mathfrak{x}} + \frac{\partial \mathfrak{p}}{\partial q_\mu} \cdot \frac{\partial q_\lambda}{\partial \mathfrak{p}}, \tag{B25}$$

which, with $\partial q_\lambda / \partial \mathfrak{x} = [q_\lambda, \mathfrak{p}] = \sum_{\nu=1}^{6} [q_\lambda, q_\nu] \partial \mathfrak{p} / \partial q_\nu$ and $\partial q_\lambda / \partial \mathfrak{p} = - [q_\lambda, \mathfrak{x}] = - \sum_{\nu=1}^{6} [q_\lambda, q_\nu] \partial \mathfrak{x} / \partial q_\nu$ yields the orthogonality relations between the Poisson and the Lagrange brackets, viz.

$$\delta_{\lambda\mu} = \sum_{\nu=1}^{6} [q_\lambda, q_\nu] \left[\frac{\partial \mathfrak{x}}{\partial q_\mu} \cdot \frac{\partial \mathfrak{p}}{\partial q_\nu} - \frac{\partial \mathfrak{p}}{\partial q_\mu} \cdot \frac{\partial \mathfrak{x}}{\partial q_\nu} \right]. \tag{B26}$$

By taking $q_\mu = \phi$, this equation yields

$$[z_i, J] = 0 \quad , \quad i = 1, \ldots, 4 \tag{B27}$$

and

$$[\phi, J] = 1. \tag{B28}$$

Appendix C. Alternative derivation of a Liouvillian volume element

In Sec. 6.2.2 a Liouvillian volume element was derived in a simple way by making use of the vector representation of the Euler-Lagrange equations. Here, an alternative (though more complicated) derivation will be given which is similar to the usual derivation of a Liouvillian volume element within the framework of guiding-centre theories.

The Lagrangian equations derived from L_r, (157), can be written as

$$\sum_{i=1}^{4} \dot{z}_i \, \omega_{ik} = \frac{\partial \langle H \rangle}{\partial z_k} + \left\langle \frac{\partial \mathfrak{p}}{\partial t} \cdot \frac{\partial \mathfrak{x}}{\partial z_k} - \frac{\partial \mathfrak{p}}{\partial z_k} \cdot \frac{\partial \mathfrak{x}}{\partial t} \right\rangle \quad , \quad i = 1, \ldots, 4 \quad , \tag{C1}$$

with ω_{ik} the antisymmetric matrix given by

$$\omega_{ik} := \left\langle \frac{\partial \mathfrak{x}}{\partial z_i} \cdot \frac{\partial \mathfrak{p}}{\partial z_k} - \frac{\partial \mathfrak{x}}{\partial z_k} \cdot \frac{\partial \mathfrak{p}}{\partial z_i} \right\rangle. \tag{C2}$$

The following phase-space volume element $d\tau_\lambda$ is then Liouvillian, i.e. $d \left(d\tau_\lambda \right) / dt = 0$ along orbits:

$$d\tau_\lambda := \lambda \left(z_1, \ldots, z_4; J; t \right) dz_1 dz_2 dz_3 dz_4, \tag{C3}$$

with

$$\lambda := \left| \det \left(\omega_{ik} \right) \right|^{1/2}. \tag{C4}$$

A proof of this theorem is given in Appendix D. Here, it will be shown that this λ coincides with $\mathcal{B}$, (172).

The matrix elements ω_{ik} satisfy the following conditions: $\omega_{ii} = 0$ for $i = 1, \ldots, 4$; $\omega_{4i} = -\omega_{i4}$ for $i = 1, \ldots, 3$; $\omega_{12} = -\omega_{21}$; $\omega_{13} = -\omega_{31}$ and $\omega_{23} = -\omega_{32}$. Calculation of $\det(\omega_{ik})$ yields

$$\det(\omega_{ik}) = (\omega_{14}\omega_{23} - \omega_{13}\omega_{24} + \omega_{12}\omega_{34})^2. \tag{C5}$$

By identifying z_1, z_2, z_3 and z_4 with R_1, R_2, R_3 and U, respectively, the following relations can be derived by comparing (C1) with (170):

$$\omega_{k4} = \frac{e}{c}\hat{\mathbf{x}}_k \cdot \left[\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_{\mathrm{U}}}{\partial \mathbf{R}}\right] \quad , \quad k = 1, \ldots, 3 \quad , \tag{C6}$$

where the $\hat{\mathbf{x}}_k$, $k = 1, \ldots, 3$ are Cartesian unit vectors. Next, one has

$$\omega_{23} = -\frac{e}{c}\hat{B}_1 \quad , \quad \omega_{13} = \frac{e}{c}\hat{B}_2 \quad , \quad \omega_{12} = -\frac{e}{c}\hat{B}_3. \tag{C7}$$

Since $\hat{\mathbf{B}} \cdot \left(\partial \hat{\mathbf{A}}/\partial U - \partial \hat{A}_{\mathrm{U}}/\partial \mathbf{R}\right) = (mc/e)\mathcal{B}$ according (172), one obtains

$$\begin{aligned}
\det(\omega_{ik}) &= \left[\sum_{i=1}^{3}\left(\frac{e}{c}\right)^2 \hat{\mathbf{x}}_i \cdot \left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_{\mathrm{U}}}{\partial \mathbf{R}}\right)\hat{B}_i\right]^2 \\
&= \left[\frac{e\,m}{c}\mathcal{B}\right]^2. \tag{C8}
\end{aligned}$$

The Liouvillian volume element given by (C3) is thus the same (within a constant factor) as that given by (190).

Appendix D. Liouville's theorem

The equations of motion as given by (15) and (16) can be written as

$$\sum_{i=1}^{n} V_i \omega_{ik} = \frac{\partial \gamma_k}{\partial t} + \frac{\partial H_{\mathrm{p}}}{\partial \zeta_k}, \tag{D1}$$

with the elements of the matrix (ω_{ik}) given by

$$\omega_{ik} = \frac{\partial \gamma_i}{\partial \zeta_k} - \frac{\partial \gamma_k}{\partial \zeta_i}. \tag{D2}$$

In order to derive Liouville's theorem, one needs the time derivative of

$$\Delta \equiv \det(\omega_{ik}). \tag{D3}$$

It is convenient to use the Laplacian development of Δ, viz.

$$\Delta = \sum_{k=1}^{n} \omega_{ik}C_{ik}, \tag{D4}$$

where i can be any of the numbers $1, \ldots, n$. Here, the determinant is expanded as a linear combination of the products of the elements of any row and the corresponding signed minors, viz. the cofactors C_{ik}. By definition, the C_{ik} do not contain the element ω_{ik} and, therefore,

$$\frac{\partial \Delta}{\partial \omega_{ik}} = C_{ik}. \tag{D5}$$

Since Δ is a function of the ω_{ik}'s, its derivative with respect to time is

$$\begin{aligned}
\frac{d\Delta}{dt} &= \sum_{i,k} \dot{\omega}_{ik} \frac{\partial \Delta}{\partial \omega_{ik}} \\
&= \sum_{i,k} \dot{\omega}_{ik} C_{ik}. \tag{D6}
\end{aligned}$$

By using the relation

$$\left(C_{ik} \right) = \Delta \left(J_{ki} \right), \tag{D7}$$

where the matrix $\left(J_{ik} \right)$ is the inverse of the matrix $\left(\omega_{ik} \right)$, i.e.

$$\sum_l \omega_{il} J_{lk} = \delta_{ik}, \tag{D8}$$

one obtains

$$\frac{d\Delta}{dt} = \Delta \sum_{i,k} \dot{\omega}_{ik} J_{ki}. \tag{D9}$$

From the definition of ω_{ik}, (D2), one can derive the relation

$$\frac{\partial \omega_{li}}{\partial \zeta_k} + \frac{\partial \omega_{kl}}{\partial \zeta_i} + \frac{\partial \omega_{ik}}{\partial \zeta_l} = 0. \tag{D10}$$

By using (D1) in the expression for the partial derivative of (D2) with respect to time, one obtains

$$\frac{\partial \omega_{ik}}{\partial t} = \sum_l \left[V_l \left[\frac{\partial \omega_{li}}{\partial \zeta_k} + \frac{\partial \omega_{kl}}{\partial \zeta_i} \right] + \frac{\partial V_l}{\partial \zeta_k} \omega_{li} + \frac{\partial V_l}{\partial \zeta_i} \omega_{kl} \right]. \tag{D11}$$

By making use of the last two equations, the total time derivative of ω_{ik},

$$\frac{d\omega_{ik}}{dt} = \frac{\partial \omega_{ik}}{\partial t} + \sum_l V_l \frac{\partial \omega_{ik}}{\partial \zeta_l} \tag{D12}$$

can be written as

$$\frac{d\omega_{ik}}{dt} = \sum_l \left[\frac{\partial V_l}{\partial \zeta_k} \omega_{li} + \frac{\partial V_l}{\partial \zeta_i} \omega_{kl} \right]. \tag{D13}$$

Multiplication of this equation by J_{ki} and summation over i and k, with $J_{ki} = -J_{ik}$ and $\omega_{kl} = -\omega_{lk}$ and the relations $\sum_i \omega_{li} J_{ik} = \delta_{lk}$ and $\sum_k \omega_{lk} J_{ki} = \delta_{li}$ being taken into account, yields

$$\sum_{i,k} \frac{d\omega_{ik}}{dt} J_{ki} = -2 \sum_l \frac{\partial V_l}{\partial \zeta_l}. \tag{D14}$$

43

By inserting this in (D9), one obtains

$$\frac{d\Delta}{dt} = -2\Delta \sum_l \frac{\partial V_l}{\partial \zeta_l}. \tag{D15}$$

The quantity

$$\lambda \equiv |\Delta|^{1/2} \tag{D16}$$

then satisfies the continuity equation

$$\frac{\partial \lambda}{\partial t} + \sum_l \frac{\partial}{\partial \zeta_l} (\lambda V_l) = 0. \tag{D17}$$

Thus, the volume element

$$d\tau \equiv \lambda \, d\zeta_1 \dots d\zeta_n \tag{D18}$$

is Liouvillian, i.e.

$$\frac{d}{dt}(d\tau) = 0. \tag{D19}$$

References

[1] Littlejohn, R. G., A guiding center Hamiltonian: a new approach. *J. Math. Phys.* **20**, 2445–2458 (1979).

[2] Littlejohn, R. G., Hamiltonian formulation of guiding center motion. *Phys. Fluids* **24**, 1730–1749 (1981).

[3] Littlejohn, R. G., Variational principles of guiding centre motion. *J. Plasma Phys.* **29**, 111–125 (1983).

[4] Weyssow, W. and Balescu, R., Hamiltonian theory of guiding centre motion revisited. *J. Plasma Phys.* **35**, 449–471 (1986).

[5] Brizard, A., Nonlinear gyrokinetic Maxwell-Vlasov equations using magnetic co-ordinates. *J. Plasma Phys.* **41**, 541–559 (1989).

[6] Sugama, H., Gyrokinetic field theory. *Phys. Plasmas* **7**, 466–480 (2000).

[7] Pfirsch, D., New variational formulation of Maxwell-Vlasov and guiding center theories. Local charge and energy conservation laws. *Z. Naturforsch.* **39a**, 1–8 (1984).

[8] Brizard, A., Eulerian action principles for linearized reduced dynamical equations. *Phys. Plasmas* **1**, 2460–2472 (1994).

[9] Brizard, A., Quadratic free energy for the linearized gyrokinetic Vlasov-Maxwell equations. *Phys. Plasmas* **1**, 2473–2479 (1994).

[10] Pfirsch, D. and Morrison, P. J., Local conservation laws for the Maxwell-Vlasov and collisionless kinetic guiding-center theories. *Phys. Rev. A* **32**, 1714–1721 (1985).

[11] Pfirsch, D. and Morrison, P. J., The energy-momentum tensor for the linearized Maxwell-Vlasov and kinetic guiding center theories. *Phys. Fluids* B **3**, 271–283 (1991).

[12] Brizard, A., Variational principle for nonlinear gyrokinetic Maxwell-Vlasov equations. *Phys. Plasmas* **7**, 4816–4822 (2000).

[13] Morrison, P. J. and Pfirsch, D., Free-energy expressions for Vlasov equilibria. *Phys. Rev. A* **40**, 3898–3910 (1989).

[14] Morrison, P. J. and Pfirsch, D., The free energy of Maxwell-Vlasov equilibria. *Phys. Fluids* B **2**, 1105–1113 (1990).

[15] Morrison, P. J. and Pfirsch, D., Dielectric energy versus plasma energy, and Hamiltonian action-angle variables for the Vlasov equation. *Phys. Fluids* B **4**, 3038–3057 (1992).

[16] Correa-Restrepo, D., Pfirsch, D. and Wimmel, H. K., Combined Maxwell and kinetic guiding center theory with polarization drift: regularized variational formulation with local charge and energy conservation. *Physica* A **136**, 453–474 (1986).

[17] Kruskal, M., Asymptotic theory of Hamiltonian and other Systems with all Solutions nearly periodic. *J. Math. Phys.* **3**, 806–828 (1962).

[18] Dirac, P. A. M., *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York, 1964), Lecture No.1.

[19] Sundermeyer, K., *Constrained Dynamics* (Springer Verlag, Berlin, 1982), Chapters I and II.

[20] Larsson, J., Hamiltonian and non-Hamiltonian perturbation theory for nearly periodic motion. *J. Math. Phys.* **27**, 495–501 (1986).

[21] Littlejohn, R. G., Hamiltonian perturbation theory in noncanonical coordinates. *J. Math. Phys.* **23**, 742–747 (1982).

[22] Correa-Restrepo, D. and Pfirsch, D., Poisson brackets for guiding-centre and gyrocentre theories. Accepted for publication, *J. Plasma Phys.*, **70**, *part 5*, 2004.

[23] Cary, J. R. and Littlejohn, R. G., Noncanonical Hamiltonian mechanics and its application to magnetic field line flow. *Ann. Phys.* **151**, 1–34 (1983).

[24] Sudarshan E. C. G. and Mukunda, N. M., *Classical Dynamics: A Modern Perspective* (John Wiley & Sons, New York, 1974), pp. 11 and 79.

[25] Pfirsch, D., De Theoremate conservandi quodam differentiali Theoriae ab Hamilton atque Jacobi doctae. *Arch. Rational Mech. Anal.* **29**, 323–324 (1968).

[26] Van Vleck, J. H., The correspondence principle in the statistical interpretation of quantum mechanics. *Proc. Nat. Acad. Sci.* **14**, 178–188 (1928).

[27] Kruskal, M., Advanced theory of gyrating particles. in *Plasma Physics* (IAEA, Vienna, 1965), p. 100.

[28] Balescu, R., *Transport Processes in Plasmas* (North-Holland, Amsterdam, 1988), Vol. I, p. 54.

[29] Correa-Restrepo, D. and Pfirsch, D. New method of deriving local energy and momentum-conserving Maxwell-collisionless drift-kinetic and gyrokinetic theories: conservation laws and their structures. Accepted for publication, *J. of Plasma Phys.*, **70**, *part 6*, 2004.

[30] Correa-Restrepo, D. and Pfirsch, D. 2004 Noether formalism with gauge-invariant variations. *J. Plasma Phys.*, **70**, 199–213.