

D. Pfirsch, D. Correa-Restrepo

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

D. Pfirsch, D. Correa-Restrepo

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

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

Dieter Pfirsch and Darío Correa-Restrepo

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

Abstract

This is the first of two papers: I general formalism, II further development of the theory with particular emphasis on the technique for performing variations when the arguments of the quantities to be varied are themselves functions of these quantities, and application to special kinetic theories, in which we describe a relatively simple and transparent method of obtaining collisionless drift-kinetic, gyrokinetic and more general theories, including local conservation laws.

The present paper starts with the introduction and proof of the correctness of a Lagrangian for the combined Maxwell-kinetic theory 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 as a tool and Dirac's constraint theory. Charge and current densities automatically distinguish between "particle-like" (guiding-center), 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 can be used to obtain a Lagrangian for the combined Maxwell-kinetic theory 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.

The perturbation theory with respect to small gyroperiods, and in the case of drift-kinetic ordering also to small gyroradii, is done solely within the framework of the rather simple part of Kruskal's method explicitly derived in this paper. 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 zeroth and first orders needed 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. The gyrokinetic ordering will be treated in the second paper.

Contents

I	INTRODUCTION	5
II	LAGRANGIANS FOR KINETIC EQUATIONS IN GENERAL COORDINATES	9
A	Phase-space Lagrangian	9
B	Transformation to new variables $\zeta_1, \dots, \zeta_6; t$	9
C	Dirac's Hamiltonian	10
D	The Lagrangian for a kinetic equation	11
E	Covariance property of the Lagrangian for the kinetic equations	14
III	REFORMULATION OF THE EQUATIONS OF MOTION SUITABLE FOR KRUSKAL'S FORMALISM	16
A	The Kruskal form of the equations of motion and related requirements	18
B	Estimate of the allowed fluctuation densities for electrostatic perturbations	20
IV	THE INVERSE KRUSKAL METHOD	21
A	General outline	21
B	Recursion procedure	22
C	Solution by iteration	23
1	<i>Remarks on the gyrokinetic theory</i>	23
2	<i>Notations</i>	24
3	<i>Structure of the iteration procedure</i>	24
4	<i>Orders in ϵ needed</i>	24
5	<i>Zeroth order</i>	24
6	<i>First order</i>	25
V	REDUCED PHASE-SPACE LAGRANGIAN	26
A	Convenient choice of a gauge function in the phase-space Lagrangian	26
B	Reduced phase-space Lagrangian	28
C	Vector representation of the reduced Lagrangian and the Euler-Lagrange equations	29
1	<i>The Littlejohn Lagrangian</i>	31
VI	MAXWELL-KINETIC THEORY IN THE REDUCED PHASE SPACE	32
A	Quantities entering the Lagrangian for the kinetic theory	32
1	<i>The canonical momenta in the reduced phase space</i>	32
2	<i>Dirac's constrained Hamiltonian</i>	32

B	Maxwell-kinetic theory	33
1	<i>The Lagrangian</i>	33
2	<i>Liouville's theorem</i>	33
3	<i>The kinetic equation</i>	34
4	<i>Sketch of a procedure for obtaining the inhomogeneous Maxwell equations and conservation laws</i>	34
VII SUMMARY		35
Appendix A	DIRECT PROOF OF $\frac{d}{dt}J(\hat{z}, t) = 0$	36
Appendix B	PROPERTIES OF SOME LAGRANGE BRACKETS	37
Appendix C	ALTERNATIVE DERIVATION OF A LIOUVILLIAN VOLUME ELEMENT	40
Appendix D	LIOUVILLE'S THEOREM	41

I. INTRODUCTION

This is the first of two papers: I general formalism, II application to special kinetic theories - in which we describe a relatively simple and transparent method of obtaining collisionless drift-kinetic, gyrokinetic and more general theories together with local conservation laws by applying a Noether-like formalism. In particular, they present a rather difficult subject from a point of view different to 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. These Lagrangians L_K are correspondingly based on a Eulerian instead of the more usual Lagrangian description of the particle motion. Eulerian Lagrangians for the linearized gyrokinetic theory were already obtained by Brizard [7, 8] employing the methods of Pfirsch [9] and Pfirsch and Morrison [10, 11].

Together with the Lagrangian for Maxwell's equations,

$$L_M = \int d^3x \frac{1}{8\pi} [\mathbf{E}^2(\mathbf{x}, t) - \mathbf{B}^2(\mathbf{x}, t)] , \quad (1)$$

the Lagrangians L_K form a total Lagrangian,

$$L_{\text{tot}}(t) = L_M + L_K , \quad (2)$$

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

$$\frac{c}{4\pi} \nabla \times \mathbf{B} - \frac{1}{4\pi} \frac{\partial \mathbf{E}}{\partial t} = \mathbf{j}_{\text{particle-like}} + c \nabla \times \mathbf{M} + \frac{\partial \mathbf{P}}{\partial t} , \quad (3)$$

$$\frac{c}{4\pi} \nabla \cdot \mathbf{E} = \rho_{\text{particle-like}} - \nabla \cdot \mathbf{P} , \quad (4)$$

where "particle-like" refers to entities which can and will be chosen as the guiding centers or the gyrocenters, 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 [9-12] 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. V.C.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 $\mathbf{x}(\zeta, t)$, $\mathbf{p}(\zeta, t)$, $\zeta = (\zeta_1, \dots, \zeta_6)$, from the phase-space variables \mathbf{x} , \mathbf{p} to "quasi-particle" variables $(\zeta_1, \dots, \zeta_4) = (z_1, \dots, z_4)$, 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 [13].
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" [14, 15], Hamiltonians $H_{D\nu}$ 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 $S_\nu(\zeta, \alpha, t)$. 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 [7, 8, 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

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

where $f_{p\nu} = f_{p\nu}(\zeta, \alpha, t)$ 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 $f_{p\nu}$, S_ν , and the vector and scalar potentials \mathbf{A} and Φ , on which the $H_{D\nu}$'s depend somehow.

6. Prove that this Lagrangian is equivalent to the Vlasov theory
7. Eliminate the S_ν 's and α in all final expressions such as the kinetic equations and charge and current densities.

Kruskal's theory (point 1 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. III). 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-center theory [16].

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 2 to 5). 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 and \mathbf{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 [9], these expressions contain particle-like (guiding-center) as well as polarization and magnetization contributions corresponding to the rotational motion in these particle-like entities [see Eqs. (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 \mathbf{E} and \mathbf{B} but not Φ and \mathbf{A} . In this way, the somewhat doubtful attribution of certain orders in ϵ to Φ and \mathbf{A} , as done in some present-day theories, is avoided. Such an attribution could be meaningful only for specific gauges. However, in order to obtain approximate Lagrangians, the following procedure is applied: since A must scale at least as B , one can formally treat it as being of the order ϵ^{-1} , although it might contain components of order ϵ^{-n} with $n > 1$ or $n \leq 0$. Similarly, one can formally treat Φ as being of order ϵ^{-1} , so that these scalings yield to leading order a motion parallel to \mathbf{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 $(\mathbf{v} \times \mathbf{B}_{\text{background}})/c = 0$ 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.

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 electromagnetic field, *viz.*

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

or Newton's equations of motion as given by Eq. (61) below, or the Hamilton-Jacobi equations as done in Sec. II.D.

The paper is organized as follows: Section II concerns the above points 1 to 7 for general transformations $\mathbf{x}(\zeta, t)$, $\mathbf{p}(\zeta, t)$, *i.e.* not necessarily for a transformation to averaging coordinates. Section III 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 IV presents Kruskal's modified theory. In Sec. V we derive a manifestly gyrophase independent Lagrangian corresponding to Kruskal's modified theory for any order in ϵ . In Sec. VI 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. Section VII presents a summary of the first paper.

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.

II. LAGRANGIANS FOR KINETIC EQUATIONS IN GENERAL COORDINATES

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

A. Phase-space Lagrangian

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

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

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

B. Transformation to new variables $\zeta_1, \dots, \zeta_6; t$

With

$$\zeta = (\zeta_1, \dots, \zeta_6), \quad \mathbf{x} = \mathbf{x}(\zeta, t), \quad \mathbf{p} = \mathbf{p}(\zeta, t), \quad (8)$$

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

$$L(\zeta, \dot{\zeta}, t) = \dot{\zeta} \cdot \gamma - H_p, \quad (9)$$

with

$$\gamma(\zeta, t) = \frac{\partial \mathbf{x}}{\partial \zeta} \cdot \mathbf{p}, \quad (10)$$

and

$$H_p(\zeta, t) = H - \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial t}. \quad (11)$$

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

$$\frac{\partial L}{\partial \zeta} = \frac{\partial \gamma}{\partial \zeta} \cdot \dot{\zeta} - \frac{\partial H_p}{\partial \zeta} \quad (12)$$

and

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

the Euler-Lagrange equations become

$$\frac{\partial \gamma}{\partial \zeta} \cdot \dot{\zeta} - \dot{\zeta} \cdot \frac{\partial \gamma}{\partial \zeta} - \frac{\partial \gamma}{\partial t} - \frac{\partial H_p}{\partial \zeta} = 0, \quad (14)$$

or, equivalently,

$$\sum_{i=1}^6 \dot{\zeta}_i \left[\frac{\partial \gamma_i}{\partial \zeta_k} - \frac{\partial \gamma_k}{\partial \zeta_i} \right] - \frac{\partial \gamma_k}{\partial t} - \frac{\partial H_p}{\partial \zeta_k} = 0 , \quad (15)$$

with solutions

$$\dot{\zeta}_i = V_i(\zeta_1, \dots, \zeta_6; t) . \quad (16)$$

The momenta canonically conjugated to the ζ_i 's are

$$P_i = \frac{\partial L}{\partial \dot{\zeta}_i} = \gamma_i(\zeta_1, \dots, \zeta_6, t) . \quad (17)$$

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. Lagrangians of the above type belong to the class of non-standard Lagrangians, as Dirac named them.

C. Dirac's Hamiltonian

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

$$\sum_{i=1}^6 \dot{\zeta}_i P_i - L(\zeta, \dot{\zeta}, t) = H_p(\zeta, t) . \quad (18)$$

As mentioned in the Introduction, Hamilton's equations with this Hamiltonian are not identical with the equations of motion, Eqs (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

$$H_D(\zeta_i, P_i, t) = H_p(\zeta, t) + \sum_{i=0}^6 V_i(P_i - \gamma_i) . \quad (19)$$

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

$$\dot{\zeta}_k = \frac{\partial H_D}{\partial P_k} = V_k , \quad (20)$$

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

$$\begin{aligned} \dot{P}_k &= -\frac{\partial H_D}{\partial \zeta_k} \\ &= -\frac{\partial H_p}{\partial \zeta_k} - \sum_{i=0}^6 \frac{\partial V_i}{\partial \zeta_k} (P_i - \gamma_i) + \sum_{i=0}^6 V_i \frac{\partial \gamma_i}{\partial \zeta_k} . \end{aligned} \quad (21)$$

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

$$\begin{aligned}\dot{\gamma}_k &= \frac{\partial \gamma_k}{\partial t} + \sum_{i=0}^6 V_i \frac{\partial \gamma_k}{\partial \zeta_i} \\ &= -\frac{\partial H_p}{\partial \zeta_k} + \sum_{i=0}^6 V_i \frac{\partial \gamma_i}{\partial \zeta_k},\end{aligned}\quad (22)$$

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

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

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.

D. The Lagrangian for a kinetic equation

As mentioned in the Introduction, this Lagrangian will be based on the Hamilton-Jacobi equation with H_D :

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

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 $L_K(t)$ for a kinetic equation is given by

$$\begin{aligned}L_K &= - \sum_{\text{particle species}} \int d^6 \zeta d^6 \alpha f_p(\zeta_i, \alpha_i, t) \left[\frac{\partial S}{\partial t}(\zeta_i, \alpha_i, t) + H_p(\zeta_i, t) \right. \\ &\quad \left. + \sum_{j=1}^6 V_j(\zeta_i, t) \left(\frac{\partial S}{\partial \zeta_j}(\zeta_i, \alpha_i, t) - \gamma_j(\zeta_i, t) \right) \right].\end{aligned}\quad (25)$$

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

Variation of the total action integral corresponding to the Lagrangian (2) (with relation(1)) yields by variation of f_p , S , Φ and \mathbf{A} , respectively,

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

$$S : \quad \frac{\partial f_p}{\partial t} + \sum_{j=1}^6 \frac{\partial}{\partial \zeta_j} (V_j f_p) = 0, \quad (27)$$

$$\Phi : \quad \frac{1}{4\pi} \nabla \cdot \mathbf{E} = - \sum_{\text{particle species}} \frac{\delta \mathcal{A}_K}{\delta \Phi (\mathbf{x}, t)}, \quad (28)$$

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

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 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 $f_p (\zeta_i, \alpha_i, t)$ by a function which, like the distribution function, is a constant of motion. This is accomplished by

$$f_p (\zeta_i, \alpha_i, t) = w (\zeta_i, \alpha_i, t) \bar{f} (\zeta_i, \alpha_i, t), \quad (30)$$

where

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

is the Van Vleck determinant. This function has the property of solving the continuity equation, Eq. (27) [18, 19]. Hence, the function $\bar{f} (\zeta_i, \alpha_i, t)$ is a solution of the equation

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

which means that $\bar{f} (\zeta_i, \alpha_i, t)$ is a constant of motion.

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

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

and, as a consequence of this,

$$f_p(\zeta_i, \alpha_i, t) d^6 \alpha \rightarrow \bar{f}(\zeta_i, P_i, t) d^6 P . \quad (34)$$

The last step is to incorporate the constraints, Eq. (17). They require that

$$\bar{f}(\zeta_i, P_i, t) = \lambda(\zeta_i, P_i, t) \prod_i \delta(P_i - \gamma_i) f(\zeta_1, \dots, \zeta_6, t) . \quad (35)$$

The function $\lambda(\zeta_i, P_i, t)$ is to be chosen such that $f(\zeta_i, t)$ is also a constant of motion in order to become the distribution function. Since $\prod_i \delta(P_i - \gamma_i)$ is not a constant of motion (as mentioned in Sec. II. C, $P_i = \gamma_i$ are only special solutions of Eq. (23)), λ cannot be a constant. Since \bar{f} and the canonical volume element in extended phase space $d^6 \zeta d^6 P$ are constants of motion, and $\int \prod_i \delta(P_i - \gamma_i) d^6 P = 1$, it follows that λ must be chosen such that

$$\lambda(\zeta_i, P_i, t) d^6 \zeta = \text{constant of the motion} . \quad (36)$$

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

$$\lambda = \sqrt{|\omega_{ik}|} , \quad (37)$$

with

$$\omega_{ik} = \frac{\partial \gamma_i}{\partial \zeta_k} - \frac{\partial \gamma_k}{\partial \zeta_i} . \quad (38)$$

In Sec. VI.B.2 a very simple derivation of λ within the framework of the averaging variables will be given.

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

$$\frac{\partial f}{\partial t} + \sum_{i=1}^6 V_i \frac{\partial f}{\partial \zeta_i} = 0 . \quad (39)$$

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

$$\mathbf{x} = (\zeta_1, \zeta_2, \zeta_3) , \quad \mathbf{p} = (\zeta_4, \zeta_5, \zeta_6) , \quad (40)$$

which yields

$$\gamma = (\zeta_4, \zeta_5, \zeta_6, 0, 0, 0) , \quad \lambda = \sqrt{|\omega_{ik}|} = 1 . \quad (41)$$

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.

E. Covariance property of the Lagrangian for the kinetic equations

Let

$$\zeta_i = \zeta_i(\xi_k, t) \quad (42)$$

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

$$\mathbf{x} = \mathbf{x}(\xi_1, \dots, \xi_6, t), \quad \mathbf{p} = \mathbf{p}(\xi_1, \dots, \xi_6, t). \quad (43)$$

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 RHS of Eq. (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

$$L_\nu = \int d^6\xi d^6\alpha \hat{f}_p(\xi_i, \alpha_i, t) \left[\frac{\partial \hat{S}(\xi_i, \alpha_i, t)}{\partial t} + \hat{H}_p(\xi_i, t) + \sum_i \hat{V}_i \left(\frac{\partial \hat{S}}{\partial \xi_i} - \hat{\gamma}_i \right) \right], \quad (44)$$

with

$$\hat{H}_p = \frac{m}{2} \mathbf{v}^2 + e\Phi - \left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\xi_i} \cdot \left(m\mathbf{v} + \frac{e}{c} \mathbf{A} \right), \quad (45)$$

$$\hat{\gamma}_i(\xi_1, \dots, \xi_6) = \left. \frac{\partial \mathbf{x}}{\partial \xi_i} \right|_t \cdot \left(m\mathbf{v} + \frac{e}{c} \mathbf{A} \right), \quad (46)$$

$$\sum_i \hat{V}_i \left(\frac{\partial \hat{\gamma}_i}{\partial \xi_k} - \frac{\partial \hat{\gamma}_k}{\partial \xi_i} \right) - \frac{\partial \hat{\gamma}_k}{\partial t} - \frac{\partial \hat{H}_p}{\partial \xi_k} = 0. \quad (47)$$

For the second way the following transformation relations are needed:

$$\left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\xi} = \left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\zeta} + \sum_i \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi} \frac{\partial \mathbf{x}}{\partial \zeta_i}. \quad (48)$$

Hence it follows that

$$\begin{aligned}\left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\zeta} \cdot \mathbf{p} &= \left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\xi} \cdot \mathbf{p} - \sum_i \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi} \frac{\partial \mathbf{x}}{\partial \zeta_i} \cdot \mathbf{p} \\ &= \left. \frac{\partial \mathbf{x}}{\partial t} \right|_{\xi} \cdot \mathbf{p} - \sum_i \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi} \gamma_i.\end{aligned}\quad (49)$$

This yields

$$H_p = \hat{H}_p + \sum_i \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi} \gamma_i. \quad (50)$$

Furthermore, one obtains

$$\frac{d\zeta_i}{dt} = \sum_k \frac{\partial \zeta_i}{\partial \xi_k} \frac{d\xi_k}{dt} + \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi}, \quad (51)$$

and, therefore

$$V_i = \sum_k \frac{\partial \zeta_i}{\partial \xi_k} \hat{V}_k + \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi}. \quad (52)$$

With

$$S = S(\zeta_i(\xi_k, t), \alpha_i, t) = \hat{S}(\xi_k, \alpha_i, t) \quad (53)$$

one has

$$\frac{dS}{dt} = \frac{d\hat{S}}{dt} \quad (54)$$

or

$$\frac{\partial S}{\partial t} + \sum_i V_i \frac{\partial S}{\partial \zeta_i} = \frac{\partial \hat{S}}{\partial t} + \sum_i \hat{V}_i \frac{\partial \hat{S}}{\partial \xi_i}. \quad (55)$$

Furthermore, one gets

$$\begin{aligned}\hat{\gamma}_i &= \left. \frac{\partial \mathbf{x}}{\partial \xi_i} \right|_t \cdot \mathbf{p} = \sum_l \frac{\partial \zeta_l}{\partial \xi_i} \frac{\partial \mathbf{x}}{\partial \zeta_l} \cdot \mathbf{p} \\ &= \sum_l \frac{\partial \zeta_l}{\partial \xi_i} \gamma_l.\end{aligned}\quad (56)$$

This yields

$$V_i \gamma_i = \left(\sum_k \frac{\partial \zeta_i}{\partial \xi_k} \hat{V}_k + \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi} \right) \gamma_i, \quad (57)$$

$$\sum_i V_i \gamma_i = \sum_k \hat{V}_k \hat{\gamma}_k + \sum_i \left. \frac{\partial \zeta_i}{\partial t} \right|_{\xi} \gamma_i. \quad (58)$$

The combination of all these relations yields

$$\begin{aligned} \frac{\partial S(\zeta_i, \alpha_i, t)}{\partial t} + H_p(\zeta_i, t) + \sum_i V_i \left(\frac{\partial S}{\partial \zeta_i} - \gamma_i \right) = \\ \frac{\partial \hat{S}(\xi_i, \alpha_i, t)}{\partial t} + \hat{H}_p(\xi_i, t) + \sum_i \hat{V}_i \left(\frac{\partial \hat{S}}{\partial \xi_i} - \hat{\gamma}_i \right). \end{aligned} \quad (59)$$

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

$$\hat{f}_p = \frac{\partial(\zeta_1, \dots, \zeta_6)}{\partial(\xi_1, \dots, \xi_6)} f_p. \quad (60)$$

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

III. REFORMULATION OF THE EQUATIONS OF MOTION SUITABLE FOR KRUSKAL'S FORMALISM

The equation of motion

$$\frac{d\mathbf{v}}{dt} = \frac{e}{m} \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right) \quad (61)$$

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

$$\mathbf{v} = v_{\parallel} \mathbf{b} + v_{\perp} \mathbf{n}_1 + \mathbf{v}_E \quad (62)$$

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

$$\mathbf{y} = (y_1, y_2, y_3, y_4, y_5) = (x, y, z, v_{\parallel}, v_{\perp}) = (\mathbf{x}, v_{\parallel}, v_{\perp}), \quad \vartheta = \frac{\theta}{2\pi}. \quad (63)$$

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 Eq. (65) 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, $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{e}_3)$, 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

$$\mathbf{e}_1 = \frac{\mathbf{b} \cdot \nabla \mathbf{b}}{|\mathbf{b} \cdot \nabla \mathbf{b}|} \quad , \quad \mathbf{e}_2 = \mathbf{b} \times \mathbf{e}_1 \quad , \quad \mathbf{e}_3 = \mathbf{b} \quad , \quad (64)$$

$$\mathbf{n}_1(\mathbf{x}, \theta, t) = -\sin \theta \mathbf{e}_1 - \cos \theta \mathbf{e}_2 \quad , \quad \mathbf{n}_2(\mathbf{x}, \theta, t) = \cos \theta \mathbf{e}_1 - \sin \theta \mathbf{e}_2 \quad . \quad (65)$$

Substituting Eq. (62) in Eq. (61) yields for the LHS of this equation

$$\dot{\mathbf{v}} = \dot{v}_{\parallel} \mathbf{b} + v_{\parallel} \dot{\mathbf{b}} + \dot{v}_{\perp} \mathbf{n}_1 - \dot{\theta} v_{\perp} \mathbf{n}_2 - v_{\perp} (\dot{\mathbf{e}}_1 \sin \theta + \dot{\mathbf{e}}_2 \cos \theta) + \dot{\mathbf{v}}_{\mathbf{E}} \quad , \quad (66)$$

while the RHS is given by

$$\frac{e}{m} \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right) = \frac{e}{m} (\mathbf{b} \cdot \mathbf{E}) \mathbf{b} - \frac{eB}{mc} v_{\perp} \mathbf{n}_2 \quad . \quad (67)$$

Scalar multiplication of Eq. (66) by \mathbf{b} , \mathbf{n}_1 and \mathbf{n}_2 yields the equations for \dot{v}_{\parallel} , \dot{v}_{\perp} and $\dot{\theta}$, respectively. With the operator $\mathcal{D}/\mathcal{D}t$ defined by

$$\frac{\mathcal{D}}{\mathcal{D}t} = \frac{d}{dt} - v_{\perp} \mathbf{n}_1 \cdot \nabla \quad (68)$$

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

$$\dot{v}_{\parallel} = \frac{e}{m} \mathbf{b} \cdot \mathbf{E} + [v_{\perp} \mathbf{n}_1 + \mathbf{v}_{\mathbf{E}}] \cdot \left[\frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + v_{\perp} \mathbf{n}_1 \cdot \nabla \mathbf{b} \right] \quad , \quad (69)$$

$$\dot{v}_{\perp} = -\mathbf{n}_1 \cdot \left[v_{\parallel} \frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + \frac{\mathcal{D}\mathbf{v}_{\mathbf{E}}}{\mathcal{D}t} \right] - v_{\perp} \mathbf{n}_1 \cdot \left[\mathbf{n}_1 \cdot (v_{\parallel} \nabla \mathbf{b} + \nabla \mathbf{v}_{\mathbf{E}}) \right] \quad , \quad (70)$$

$$\begin{aligned} \dot{\vartheta} &:= \frac{\dot{\theta}}{2\pi} \\ &= \frac{1}{2\pi} \left\{ \frac{eB}{mc} + \frac{\mathbf{n}_2}{v_{\perp}} \cdot \left[v_{\parallel} \frac{d\mathbf{b}}{dt} + \frac{d\mathbf{v}_{\mathbf{E}}}{dt} \right] + \mathbf{e}_2 \cdot \frac{d\mathbf{e}_1}{dt} \right\} \\ &= \frac{1}{2\pi} \left\{ \frac{eB}{mc} + \frac{\mathbf{n}_2}{v_{\perp}} \cdot \left[v_{\parallel} \frac{\mathcal{D}\mathbf{b}}{\mathcal{D}t} + \frac{\mathcal{D}\mathbf{v}_{\mathbf{E}}}{\mathcal{D}t} \right] + \mathbf{n}_2 \cdot \left[\mathbf{n}_1 \cdot (v_{\parallel} \nabla \mathbf{b} + \nabla \mathbf{v}_{\mathbf{E}}) \right] \right. \\ &\quad \left. + \mathbf{e}_2 \cdot \frac{\partial \mathbf{e}_1}{\partial t} + (v_{\parallel} \mathbf{b} + \mathbf{v}_{\mathbf{E}} + v_{\perp} \mathbf{n}_1) \cdot [\nabla \mathbf{e}_1 \cdot \mathbf{e}_2] \right\} \quad . \quad (71) \end{aligned}$$

A. 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,

$$t \rightarrow s = \frac{t}{\epsilon} . \quad (72)$$

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

$$\frac{eB}{mc} = \Omega_g = \frac{1}{\epsilon} \Omega_s . \quad (73)$$

This yields for Eqs. (62) and (69)-(71)

$$\frac{d\mathbf{x}}{ds} = \epsilon (v_{\parallel} \mathbf{b} + v_{\perp} \mathbf{n}_1 + \mathbf{v}_E) , \quad (74)$$

$$\frac{dv_{\parallel}}{ds} = \epsilon \left[\frac{e}{m} \mathbf{b} \cdot \mathbf{E} + [v_{\perp} \mathbf{n}_1 + \mathbf{v}_E] \cdot \left[\frac{D\mathbf{b}}{Dt} + v_{\perp} \mathbf{n}_1 \cdot \nabla \mathbf{b} \right] \right] , \quad (75)$$

$$\frac{dv_{\perp}}{ds} = -\epsilon \left[\mathbf{n}_1 \cdot \left[v_{\parallel} \frac{D\mathbf{b}}{Dt} + \frac{D\mathbf{v}_E}{Dt} \right] + v_{\perp} \mathbf{n}_1 \cdot \left[\mathbf{n}_1 \cdot (v_{\parallel} \nabla \mathbf{b} + \nabla \mathbf{v}_E) \right] \right] , \quad (76)$$

$$\begin{aligned} \frac{d\vartheta}{ds} = & \frac{1}{2\pi} \Omega_s + \epsilon \frac{1}{2\pi} \left\{ \frac{\mathbf{n}_2}{v_{\perp}} \cdot \left[v_{\parallel} \frac{D\mathbf{b}}{Dt} + \frac{D\mathbf{v}_E}{Dt} \right] + \mathbf{n}_2 \cdot \left[\mathbf{n}_1 \cdot (v_{\parallel} \nabla \mathbf{b} + \nabla \mathbf{v}_E) \right] \right. \\ & \left. + \mathbf{e}_2 \cdot \frac{\partial \mathbf{e}_1}{\partial t} + (v_{\parallel} \mathbf{b} + \mathbf{v}_E + v_{\perp} \mathbf{n}_1) \cdot [\nabla \mathbf{e}_1 \cdot \mathbf{e}_2] \right\} . \end{aligned} \quad (77)$$

With the definition of the five-component vector \mathbf{y} in Eq. (63), these equations are of the form

$$\frac{d\mathbf{y}}{ds} = \epsilon \mathbf{g}(\mathbf{y}, \vartheta) \quad , \quad \frac{d\vartheta}{ds} = \psi(\mathbf{y}, \vartheta) , \quad (78)$$

with

$$\mathbf{g}(\mathbf{y}, \vartheta + 1) = \mathbf{g}(\mathbf{y}, \vartheta) \quad , \quad \psi(\mathbf{y}, \vartheta + 1) = \psi(\mathbf{y}, \vartheta) . \quad (79)$$

Equations (78) and (79) are of the same type as the starting equations in Kruskal's method if they satisfy the conditions

$$\mathbf{g}(\mathbf{y}, \vartheta) = \mathcal{O}(\epsilon^0) \quad , \quad \psi(\mathbf{y}, \vartheta) = \mathcal{O}(\epsilon^0) . \quad (80)$$

For the inverse Kruskal method there is the additional requirement that

$$\psi(\mathbf{y}, \vartheta) = \psi_0(\mathbf{y}) + \mathcal{O}(\epsilon^1) . \quad (81)$$

Weaker conditions, which would be of interest for the gyrokinetic theory are not possible, as follows from Eqs. (115) and (116) 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 ϵv_\perp , ϵv_\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 $\mathbf{v}_E = \mathcal{O}(\epsilon^0)$ because of $k_\perp \sim 1/L$.) This and the above conditions are satisfied if with

$$v_\perp = \mathcal{O}(\epsilon^0) \quad , \quad v_\parallel = \mathcal{O}(\epsilon^0) \quad (82)$$

it holds that

$$\mathbf{b} \cdot \mathbf{E} = \mathcal{O}(\epsilon^0) \quad , \quad \nabla \mathbf{v}_E = \mathcal{O}(\epsilon^0) \quad , \quad (83)$$

$$\frac{\partial}{\partial t} = \Omega_s \mathcal{O}(\epsilon^0) \quad , \quad \mathbf{v}_E = \mathcal{O}(\epsilon^0) \quad (\text{drift-kinetic case}) \quad , \quad \mathbf{v}_E = \mathcal{O}(\epsilon) \quad (\text{gyrokinetic case}) \quad , \quad (84)$$

$$\nabla \mathbf{b} = \mathcal{O}(\epsilon^0) \quad . \quad (85)$$

In the drift-kinetic case, one has

$$\nabla \sim \frac{1}{L} \quad \text{with} \quad L \sim \frac{r_{\text{gyro}}}{\epsilon} \quad , \quad (86)$$

and the conditions (82)-(85) are therefore satisfied if

$$\mathbf{b} \cdot \mathbf{E} = \frac{v}{c} B \mathcal{O}(\epsilon) \quad , \quad \mathbf{b} \times \mathbf{E} = \frac{v}{c} B \mathcal{O}(\epsilon^0) \quad . \quad (87)$$

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

$$\mathbf{b} \times \nabla \sim \frac{1}{r_{\text{gyro}}} \quad \text{for fluctuations} \quad . \quad (88)$$

Equations (82)-(85) 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

$$\delta \mathbf{v}_E = \mathcal{O}(\epsilon) \quad , \quad \delta \mathbf{b} = \mathcal{O}(\epsilon) \quad . \quad (89)$$

This means, together with the characterization of \mathbf{v}_E in Eq. (84), that

$$\mathbf{b} \times \mathbf{E}_{\text{fluctuations}} = \frac{v}{c} B \mathcal{O}(\epsilon) \quad , \quad \mathbf{B}_{\text{fluctuations}} = B \mathcal{O}(\epsilon) \quad . \quad (90)$$

In Sec. IV 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. VI, no splitting into slow and fast spatial dependences will be done. This implies that expansions with respect to terms $\mathcal{O}(\epsilon)$ occurring in the arguments characterizing slow spatial dependences can only be done after the stated variations. This problem does not appear in the drift-kinetic case.

B. 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

$$n_{\text{fluctuation}, e} = n_{0e} e^{\frac{e\Phi}{T}}. \quad (91)$$

In the drift-kinetic case one has for all components

$$\mathbf{E} \sim \frac{\Phi}{L}. \quad (92)$$

According to Eq. (87), it then follows that

$$\Phi \sim r_{\text{gyro}} \frac{v}{c} B, \quad (93)$$

and, therefore, that

$$\frac{e\Phi}{T} \sim \frac{mvc}{eB} \frac{v}{c} eB \frac{1}{T} \sim 1. \quad (94)$$

This means that large density fluctuations are allowed.

In the gyrokinetic case one has

$$\mathbf{b} \cdot \mathbf{E} \sim \frac{\Phi}{L}, \quad \mathbf{b} \times \mathbf{E} \sim \frac{\Phi}{r_{\text{gyro}}}. \quad (95)$$

According to Eq. (87) and (90), which give the same limits for all components of \mathbf{E} , the second relation, Eq. (95), is the critical one and yields

$$\Phi \sim r_{\text{gyro}} \frac{r_{\text{gyro}}}{L} \frac{v}{c} B, \quad (96)$$

and, therefore, it follows that

$$\frac{e\Phi}{T} \sim \frac{r_{\text{gyro}}}{L}. \quad (97)$$

Hence, only very small fluctuation amplitudes are allowed.

IV. THE INVERSE KRUSKAL METHOD

A. General outline

Kruskal's method [13] (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:

$$\mathbf{y}_s = \frac{d\mathbf{y}}{ds} = \epsilon \mathbf{g}(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \vartheta_s = \frac{d\vartheta}{ds} = \psi(\mathbf{y}, \vartheta, \epsilon s) \quad (98)$$

with

$$\mathbf{g}(\mathbf{y}, \vartheta + 1, \epsilon s) = \mathbf{g}(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \psi(\mathbf{y}, \vartheta + 1, \epsilon s) = \psi(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \mathbf{y} = (y_1, \dots, y_n) . \quad (99)$$

The goal of the inverse Kruskal method is to obtain a transformation (cf. also Ref.[16])

$$\mathbf{y} = \mathbf{Y}(\hat{\mathbf{z}}, \phi, \epsilon s) \quad , \quad \vartheta = \Theta(\hat{\mathbf{z}}, \phi, \epsilon s) \quad , \quad \hat{\mathbf{z}} = (\hat{z}_1, \dots, \hat{z}_n) \quad , \quad (100)$$

where ϕ is an angle variable with period 1 such that

$$\hat{\mathbf{z}}_s = \epsilon \mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \quad , \quad \phi_s = \omega(\hat{\mathbf{z}}, \epsilon s) . \quad (101)$$

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

$$J(\hat{\mathbf{z}}, t) = \oint_{\hat{\mathbf{z}}=const} \mathbf{p} \cdot d\mathbf{x} = \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial \phi} d\phi \quad (102)$$

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 $\mathbf{A} + \nabla\Psi$,

$$\int_0^1 \nabla\Psi(\mathbf{x}, t) \cdot \frac{\partial \mathbf{x}}{\partial \phi} d\phi = \int_0^1 \frac{\partial\Psi}{\partial \phi} d\phi = \oint d\Psi = 0 \quad , \quad (103)$$

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 Ref. [20].

B. Recursion procedure

As in Kruskal's original theory, one can find a recursion procedure which enables one to obtain the transformation of Eqs. (100) together with the functions h and ω in Eqs. (101) to arbitrary approximation in ϵ without solving any differential equations. This goal is reached in the following way:

From Eqs. (98) and (100) one obtains

$$\begin{aligned} y_s &= \hat{z}_s \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} + \phi_s \frac{\partial \mathbf{Y}}{\partial \phi} + \epsilon \frac{\partial \mathbf{Y}}{\partial (\epsilon s)} \\ &= \epsilon h(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} + \omega(\hat{\mathbf{z}}, \epsilon s) \frac{\partial \mathbf{Y}}{\partial \phi} + \epsilon \frac{\partial \mathbf{Y}}{\partial (\epsilon s)} = \epsilon \mathbf{g}(\mathbf{Y}, \Theta, \epsilon s) , \end{aligned} \quad (104)$$

$$\begin{aligned} \vartheta_s &= \hat{z}_s \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} + \phi_s \frac{\partial \Theta}{\partial \phi} + \epsilon \frac{\partial \Theta}{\partial (\epsilon s)} \\ &= \epsilon h(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} + \omega(\hat{\mathbf{z}}, \epsilon s) \frac{\partial \Theta}{\partial \phi} + \epsilon \frac{\partial \Theta}{\partial (\epsilon s)} = \psi(\mathbf{Y}, \Theta, \epsilon s) . \end{aligned} \quad (105)$$

Formal integration of Eqs. (104) and (105) over ϕ yields

$$\begin{aligned} \mathbf{Y} &= \mathbf{Y}(\hat{\mathbf{z}}, \phi = 0, \epsilon s) \\ &\quad + \frac{\epsilon}{\omega(\hat{\mathbf{z}}, \epsilon s)} \int_0^\phi \left[\mathbf{g}(\mathbf{Y}, \Theta, \epsilon s) - h(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} - \frac{\partial \mathbf{Y}}{\partial (\epsilon s)} \right] d\tilde{\phi} , \end{aligned} \quad (106)$$

$$\begin{aligned} \Theta &= \Theta(\hat{\mathbf{z}}, \phi = 0, \epsilon s) \\ &\quad + \frac{1}{\omega(\hat{\mathbf{z}}, \epsilon s)} \int_0^\phi \left[\psi(\mathbf{Y}, \Theta, \epsilon s) - \epsilon h(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} - \epsilon \frac{\partial \Theta}{\partial (\epsilon s)} \right] d\tilde{\phi} . \end{aligned} \quad (107)$$

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

$$\begin{aligned} \mathbf{Y}(\hat{\mathbf{z}}, \phi + 1, \epsilon s) &= \mathbf{Y}(\hat{\mathbf{z}}, \phi, \epsilon s) , \\ \Theta(\hat{\mathbf{z}}, \phi + 1, \epsilon s) &= \Theta(\hat{\mathbf{z}}, \phi, \epsilon s) + 1 . \end{aligned} \quad (108)$$

They need to be imposed only for $\phi = 0$ as starting values. With Eqs. (106) and (107), these conditions become

$$\int_0^1 \left[\mathbf{g}(\mathbf{Y}, \Theta, \epsilon s) - h(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \mathbf{Y}}{\partial \hat{\mathbf{z}}} - \frac{\partial \mathbf{Y}}{\partial (\epsilon s)} \right] d\phi = 0 , \quad (109)$$

$$\frac{1}{\omega(\hat{\mathbf{z}}, \epsilon s)} \int_0^1 \left[\psi(\mathbf{Y}, \Theta, \epsilon s) - \epsilon h(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial \Theta}{\partial \hat{\mathbf{z}}} - \epsilon \frac{\partial \Theta}{\partial (\epsilon s)} \right] d\phi = 1 . \quad (110)$$

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

$$\mathbf{Y}(\hat{\mathbf{z}}, \phi = 0, \epsilon s) = \hat{\mathbf{z}} + \epsilon \mathbf{Y}_c(\hat{\mathbf{z}}, \epsilon s) , \quad (111)$$

$$\Theta(\hat{\mathbf{z}}, \phi = 0, \epsilon s) = \epsilon \Theta_c(\hat{\mathbf{z}}, \epsilon s) , \quad (112)$$

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

$$\epsilon \mathbf{Y}_c(\hat{\mathbf{z}}, \epsilon s) = \epsilon \mathbf{Y}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \epsilon^2 \mathbf{Y}_{c2}(\hat{\mathbf{z}}, \epsilon s) + \dots , \quad (113)$$

$$\epsilon \Theta_c(\hat{\mathbf{z}}, \epsilon s) = \epsilon \Theta_{c1}(\hat{\mathbf{z}}, \epsilon s) + \epsilon^2 \Theta_{c2}(\hat{\mathbf{z}}, \epsilon s) + \dots . \quad (114)$$

This choice guarantees that the variables $\hat{\mathbf{z}}$ and ϕ and the particle variables \mathbf{y} and θ coincide in the limit $\epsilon \rightarrow 0$, leaving, however, each term in the series for $\mathbf{Y}_c(\hat{\mathbf{z}}, \epsilon s)$ and $\Theta_c(\hat{\mathbf{z}}, \epsilon s)$ still arbitrary, these being chosen later as appears convenient. A subject related to the freedom to prescribe the functions $\mathbf{Y}(\hat{\mathbf{z}}, \phi = 0, \epsilon s)$ and $\Theta(\hat{\mathbf{z}}, \phi = 0, \epsilon s)$ is discussed in Ref. [21].

Equations (106) and (107) together with (109) and (112) 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 $\omega(\hat{\mathbf{z}}, \epsilon s)$.

C. Solution by iteration

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 Eqs. (106), (107), (109) and (110) such derivatives occur only in the form

$$\mathbf{h}(\hat{\mathbf{z}}, \epsilon s) \cdot \frac{\partial}{\partial \hat{\mathbf{z}}} . \quad (115)$$

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

$$\mathbf{h}_0 = \int_0^1 \mathbf{g}_0 d\phi \equiv \langle \mathbf{g}_0 \rangle . \quad (116)$$

According to Eqs. (74), (78), (84) and (89), $\langle \mathbf{g}_0 \rangle$ possesses only a parallel spatial component. The derivatives (115) therefore cannot lower the order in ϵ . This means that an iterative solution procedure is possible. This was the reason for introducing the condition $\mathbf{v}_E = \mathcal{O}(\epsilon)$ for the gyrokinetic theory. There is, of course, no problem in the case of drift-kinetic theory.

2. Notations

The following notations will be used:

$$\begin{aligned} \mathbf{Y}^{(n)} &= \mathbf{Y}, \text{ obtained in the } n\text{-th iteration,} \\ \langle f \rangle &= \int_0^1 f d\phi \text{ for any function } f. \end{aligned} \quad (117)$$

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

3. Structure of the iteration procedure

The iteration is done in the following way: Eqs. (106), (107), (109) and (110) will be solved in the order (106), (110), (107) and (109) 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 Eqs. (74)-(77) and (78), have the form

$$\mathbf{g} = \mathbf{g}_0(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \psi = \psi_0(\mathbf{y}, \epsilon s) + \epsilon \psi_1(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \psi_0(\mathbf{y}, \epsilon s) = \frac{\Omega_s(\mathbf{x}, \epsilon s)}{2\pi}. \quad (118)$$

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.

5. Zeroth order

The zeroth-order quantities follow from Eqs. (106), (107), (109) and (110) with $\epsilon = 0$:

From Eq. (106):

$$\mathbf{Y}^{(0)} = \hat{\mathbf{z}}. \quad (119)$$

From Eq. (110):

$$\omega^{(0)}(\hat{\mathbf{z}}, \epsilon s) = \psi_0(\hat{\mathbf{z}}, \epsilon s) = \frac{1}{2\pi} \Omega_s(\hat{\mathbf{z}}, \epsilon s). \quad (120)$$

From Eq. (107):

$$\Theta^{(0)} = \phi. \quad (121)$$

From Eq. (109):

$$\mathbf{h}^{(0)} = \langle \mathbf{g}_0(\hat{\mathbf{z}}, \phi, \epsilon s) \rangle. \quad (122)$$

Corrigenda

Page 24: the correct reading of Eq.(118) is:

$$\begin{aligned} \mathbf{g} &= \mathbf{g}_0(\mathbf{y}, \vartheta, \epsilon s) + \epsilon \mathbf{g}_1(\mathbf{y}, \vartheta, \epsilon s) \quad , \quad \psi = \psi_0(\mathbf{y}, \epsilon s) + \epsilon \psi_1(\mathbf{y}, \vartheta, \epsilon s) \quad , \\ \psi_0(\mathbf{y}, \epsilon s) &= \frac{\Omega_s(\mathbf{x}, \epsilon s)}{2\pi} . \end{aligned} \quad (118)$$

Page 29: the last but one sentence of Sec. V B should read:

For the Dirac Hamiltonian the solutions $\dot{z}_i(z_1, \dots, z_4, J, t)$ of the Euler-Lagrange equations corresponding to Eq. (147) are needed.

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

$$(h_1^{(0)}, h_2^{(0)}, h_3^{(0)}) = v_{\parallel} \mathbf{b} + \mathbf{v}_E, \quad (123)$$

where the RHS 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 $\mathbf{v}_E = \mathcal{O}(\epsilon)$.

6. First order

From Eqs. (106) and (111):

$$\begin{aligned} \mathbf{Y}^{(1)} &= \hat{\mathbf{z}} + \epsilon \mathbf{Y}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{\omega_0(\hat{\mathbf{z}}, \epsilon s)} \int_0^\phi [\mathbf{g}_0(\hat{\mathbf{z}}, \phi, \epsilon s) - \langle \mathbf{g}_0 \rangle] d\tilde{\phi} \\ &= \hat{\mathbf{z}} + \epsilon \mathbf{Y}_1. \end{aligned} \quad (124)$$

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

$$\mathbf{R} := (\hat{z}_1, \hat{z}_2, \hat{z}_3), \quad V_{\perp} := \hat{z}_5, \quad (125)$$

they are

$$\begin{aligned} \mathbf{x}^{(1)} &= \mathbf{R} + \epsilon \mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{\omega_0(\mathbf{R}, \epsilon s)} \int_0^\phi n_1(\mathbf{R}, \Phi, \epsilon s) V_{\perp} d\tilde{\phi} \\ &= \mathbf{R} + \epsilon \mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{2\pi\omega_0(\mathbf{R}, \epsilon s)} V_{\perp} (n_2(\mathbf{R}, \Phi, \epsilon s) - \mathbf{e}_1(\mathbf{R}, \epsilon s)) \\ &= \mathbf{R} + \epsilon \mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) + \frac{\epsilon}{\Omega_s(\mathbf{R}, \epsilon s)} V_{\perp} (n_2(\mathbf{R}, \Phi, \epsilon s) - \mathbf{e}_1(\mathbf{R}, \epsilon s)). \end{aligned} \quad (126)$$

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

$$\mathbf{x}_{c1}(\hat{\mathbf{z}}, \epsilon s) = \frac{V_{\perp}}{\Omega_s(\mathbf{R}, \epsilon s)} \mathbf{e}_1(\mathbf{R}, \epsilon s). \quad (127)$$

This yields

$$\mathbf{x}^{(1)} = \mathbf{R} + \epsilon \frac{V_{\perp}}{\Omega_s(\mathbf{R}, \epsilon s)} n_2(\mathbf{R}, \phi, \epsilon s) \quad (128)$$

A particular result is obtained by substituting this relation in Eq. (102):

$$J^{(1)} = \epsilon \frac{2\pi}{\Omega_s(\mathbf{R}, \epsilon s)} \frac{mV_{\perp}^2}{2}. \quad (129)$$

V. REDUCED PHASE-SPACE LAGRANGIAN

A. 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. II.A, with a phase-space particle Lagrangian in the form

$$L(\mathbf{p}, \mathbf{x}; \dot{\mathbf{x}}; t) = \mathbf{p} \cdot \dot{\mathbf{x}} - H(\mathbf{x}, \mathbf{p}, t) + \frac{d}{dt} F(\mathbf{x}, \mathbf{p}, t), \quad (130)$$

where, as compared with Sec. II.A, 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 ,

$$\mathbf{x} = \mathfrak{x}(z_1, \dots, z_4, J, \phi, t) \quad , \quad \mathbf{p} = \mathfrak{p}(z_1, \dots, z_4, J, \phi, t) \quad , \quad (131)$$

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

$$\begin{aligned} L(z_1, \dots, z_4, J, \phi; \dot{z}_1, \dots, \dot{z}_4, \dot{J}, \dot{\phi}; t) = \\ \sum_{i=1}^4 \dot{z}_i \left[\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial z_i} + \frac{\partial F}{\partial z_i} \right] + \dot{\phi} \left[\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} + \frac{\partial F}{\partial \phi} \right] \\ + \dot{J} \left[\mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial J} + \frac{\partial F}{\partial J} \right] - \left[H - \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial t} - \frac{\partial F}{\partial t} \right]. \end{aligned} \quad (132)$$

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 Eq. (132) independent of ϕ and equal to J ,

$$\frac{\partial F}{\partial \phi} := J - \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \phi}. \quad (133)$$

Integration of this equation yields

$$\begin{aligned} F(z_1, \dots, z_4, J, \phi; t) = \langle F \rangle(z_1, \dots, z_4, J; t) + F^{(\phi)}(z_1, \dots, z_4, J, \phi; t) \\ - \langle F^{(\phi)}(z_1, \dots, z_4, J, \phi; t) \rangle, \end{aligned} \quad (134)$$

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

$$F^{(\phi)}(z_1, \dots, z_4, J, \phi; t) = J\phi - \int_0^\phi \mathfrak{p} \cdot \frac{\partial \mathfrak{x}}{\partial \phi} d\tilde{\phi}. \quad (135)$$

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

With the identity

$$f = \langle f \rangle + \int_0^\phi \frac{\partial f}{\partial \tilde{\phi}} d\tilde{\phi} - \left\langle \int_0^\phi \frac{\partial f}{\partial \tilde{\phi}} d\tilde{\phi} \right\rangle \quad (136)$$

for any function f taken into account, one obtains

$$\begin{aligned} \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial z_i} + \frac{\partial F}{\partial z_i} &= \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial z_i} \right\rangle + \frac{\partial \langle F \rangle}{\partial z_i} \\ &+ \int_0^\phi \left[\frac{\partial \mathbf{p}}{\partial \tilde{\phi}} \cdot \frac{\partial \mathbf{r}}{\partial z_i} - \frac{\partial \mathbf{p}}{\partial z_i} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} \right] d\tilde{\phi} \\ &- \left\langle \int_0^\phi \left[\frac{\partial \mathbf{p}}{\partial \tilde{\phi}} \cdot \frac{\partial \mathbf{r}}{\partial z_i} - \frac{\partial \mathbf{p}}{\partial z_i} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} \right] d\tilde{\phi} \right\rangle, \end{aligned} \quad (137)$$

$$\begin{aligned} \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial J} + \frac{\partial F}{\partial J} &= \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial J} \right\rangle + \frac{\partial \langle F \rangle}{\partial J} \\ &+ \int_0^\phi \left[1 + \frac{\partial \mathbf{p}}{\partial \tilde{\phi}} \cdot \frac{\partial \mathbf{r}}{\partial J} - \frac{\partial \mathbf{p}}{\partial J} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} \right] d\tilde{\phi} \\ &- \left\langle \int_0^\phi \left[1 + \frac{\partial \mathbf{p}}{\partial \tilde{\phi}} \cdot \frac{\partial \mathbf{r}}{\partial J} - \frac{\partial \mathbf{p}}{\partial J} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} \right] d\tilde{\phi} \right\rangle, \end{aligned} \quad (138)$$

$$\begin{aligned} H - \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial t} - \frac{\partial F}{\partial t} &= \langle H \rangle - \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial t} \right\rangle - \frac{\partial \langle F \rangle}{\partial t} \\ &- \int_0^\phi \left[-\frac{\partial H}{\partial \tilde{\phi}} + \frac{\partial \mathbf{p}}{\partial \tilde{\phi}} \cdot \frac{\partial \mathbf{r}}{\partial t} - \frac{\partial \mathbf{p}}{\partial t} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} \right] d\tilde{\phi} \\ &+ \left\langle \int_0^\phi \left[-\frac{\partial H}{\partial \tilde{\phi}} + \frac{\partial \mathbf{p}}{\partial \tilde{\phi}} \cdot \frac{\partial \mathbf{r}}{\partial t} - \frac{\partial \mathbf{p}}{\partial t} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} \right] d\tilde{\phi} \right\rangle. \end{aligned} \quad (139)$$

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

$$\frac{\partial \mathbf{p}}{\partial \phi} \cdot \frac{\partial \mathbf{r}}{\partial z_i} - \frac{\partial \mathbf{p}}{\partial z_i} \cdot \frac{\partial \mathbf{r}}{\partial \phi} = 0, \quad (140)$$

$$\frac{\partial \mathbf{p}}{\partial \phi} \cdot \frac{\partial \mathbf{r}}{\partial J} - \frac{\partial \mathbf{p}}{\partial J} \cdot \frac{\partial \mathbf{r}}{\partial \phi} = -1, \quad (141)$$

$$\frac{\partial \mathbf{p}}{\partial \phi} \cdot \frac{\partial \mathbf{r}}{\partial t} - \frac{\partial \mathbf{p}}{\partial t} \cdot \frac{\partial \mathbf{r}}{\partial \phi} = \frac{\partial H}{\partial \phi}. \quad (142)$$

By inserting Eqs. (137)-(142) in Eq. (132), one obtains

$$\begin{aligned}
L(z_1, \dots, z_4, J; \dot{z}_1, \dots, \dot{z}_4, \dot{J}, \dot{\phi}; t) = & \\
\sum_{i=1}^4 \dot{z}_i \left[\left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial z_i} \right\rangle + \frac{\partial \langle F \rangle}{\partial z_i} \right] + J \dot{\phi} & \\
+ \dot{J} \left[\left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial J} \right\rangle + \frac{\partial \langle F \rangle}{\partial J} \right] & \\
- \langle H \rangle + \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial t} \right\rangle + \frac{\partial \langle F \rangle}{\partial t} . & \quad (143)
\end{aligned}$$

This Lagrangian obviously does not depend on ϕ . The simplest expression for L is obtained by choosing the free function $\langle F \rangle$ as

$$\langle F \rangle = 0 . \quad (144)$$

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

$$\begin{aligned}
L(z_1, \dots, z_4, J; \dot{z}_1, \dots, \dot{z}_4, \dot{J}, \dot{\phi}; t) = & \\
\sum_{i=1}^4 \dot{z}_i \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial z_i} \right\rangle + J \dot{\phi} & \\
+ \dot{J} \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial J} \right\rangle & \\
- \langle H \rangle + \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial t} \right\rangle . & \quad (145)
\end{aligned}$$

This agrees with Larsson's results [16] obtained in a somewhat different way within the framework of guiding-center theory.

B. Reduced phase-space Lagrangian

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, \dots, 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 , Eq. (145), by omitting the term $J\dot{\phi}$ and the terms proportional to \dot{J} . With the total time derivative in the reduced phase space given by

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \sum_{j=1}^4 \dot{z}_j \frac{\partial}{\partial z_j} , \quad (146)$$

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

$$L_r(z_1, \dots, z_4; J; \dot{z}_1, \dots, \dot{z}_4; t) := \sum_{i=1}^4 \dot{z}_i \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial z_i} \right\rangle - \langle H \rangle + \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial t} \right\rangle, \quad (147)$$

are the same as those obtained from Eq. (145).

There are no approximations in the Lagrangian L_r from which $\dot{z}_1, \dots, \dot{z}_4$ 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 $\dot{z}_i(z_1, \dots, z_4, J, t)$ of Eq. (147) are needed. These will be obtained in the next section by using an appropriate vector notation.

C. 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,

$$\mathbf{R} = (R_1, R_2, R_3) \quad , \quad R_i := z_i \quad , \quad i = 1, 2, 3 \quad . \quad (148)$$

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

$$U := z_4 \quad . \quad (149)$$

The gradient with respect to \mathbf{R} will be denoted by both $\partial/\partial\mathbf{R}$ and $\nabla_{\mathbf{R}}$, as it appears convenient. The reduced Lagrangian L_r , Eq. (147), can then be written as

$$L_r(\mathbf{R}, U; J; \dot{\mathbf{R}}, \dot{U}; t) = \frac{e}{c} \left[\dot{\mathbf{R}} \cdot \hat{\mathbf{A}}(\mathbf{R}, U; J; t) + \dot{U} \hat{A}_U \right] - e \hat{\Phi}(\mathbf{R}, U; J; t) \quad , \quad (150)$$

where

$$\hat{\mathbf{A}}(\mathbf{R}, U; J; t) := \frac{c}{e} \left\langle \frac{\partial \mathbf{r}}{\partial \mathbf{R}} \cdot \mathbf{p} \right\rangle \quad , \quad (151)$$

$$\hat{A}_U(\mathbf{R}, U; J; t) := \frac{c}{e} \left\langle \frac{\partial \mathbf{r}}{\partial U} \cdot \mathbf{p} \right\rangle \quad , \quad (152)$$

$$e\hat{\Phi}(\mathbf{R}, U; J; t) := \langle H \rangle - \left\langle \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial t} \right\rangle . \quad (153)$$

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

$$-\frac{1}{c} \frac{d}{dt} \hat{\mathbf{A}} + \frac{1}{c} \left(\frac{\partial \hat{\mathbf{A}}}{\partial \mathbf{R}} \right) \cdot \dot{\mathbf{R}} + \frac{1}{c} \left(\frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right) \dot{U} - \frac{\partial \hat{\Phi}}{\partial \mathbf{R}} = 0 . \quad (154)$$

The Euler-Lagrange equation obtained by varying U is

$$-\frac{1}{c} \frac{d}{dt} \hat{A}_U + \frac{1}{c} \left(\frac{\partial \hat{\mathbf{A}}}{\partial U} \right) \cdot \dot{\mathbf{R}} + \frac{1}{c} \left(\frac{\partial \hat{A}_U}{\partial U} \right) \dot{U} - \frac{\partial \hat{\Phi}}{\partial U} = 0 . \quad (155)$$

With the total time derivative given by

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \dot{\mathbf{R}} \cdot \frac{\partial}{\partial \mathbf{R}} + \dot{U} \frac{\partial}{\partial U} , \quad (156)$$

and

$$\hat{\mathbf{B}} := \nabla_{\mathbf{R}} \times \hat{\mathbf{A}} , \quad (157)$$

$$\hat{\mathbf{E}} := -\frac{1}{c} \frac{\partial \hat{\mathbf{A}}}{\partial t} - \nabla_{\mathbf{R}} \hat{\Phi} , \quad (158)$$

the two Euler-Lagrange equations become, respectively,

$$\frac{1}{c} \dot{\mathbf{R}} \times \hat{\mathbf{B}} + \hat{\mathbf{E}} - \dot{U} \frac{1}{c} \left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right) = 0 \quad (159)$$

and

$$\frac{1}{c} \dot{\mathbf{R}} \cdot \left[\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right] - \frac{\partial \hat{\Phi}}{\partial U} - \frac{1}{c} \frac{\partial \hat{A}_U}{\partial t} = 0 . \quad (160)$$

Crossing Eq. (159) with $\partial \hat{\mathbf{A}}/\partial U$ and $\partial \hat{A}_U/\partial \mathbf{R}$ and subtracting the second from the first equation resulting yields

$$\hat{\mathbf{E}} \times \left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right) + \hat{\mathbf{B}} \left(\frac{\partial \hat{\Phi}}{\partial U} + \frac{1}{c} \frac{\partial \hat{A}_U}{\partial t} \right) - \frac{1}{c} \dot{\mathbf{R}} \left[\hat{\mathbf{B}} \cdot \left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right) \right] = 0 . \quad (161)$$

With the definition

$$\mathcal{B} := \hat{\mathbf{B}} \cdot \frac{e}{mc} \left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right) \quad (162)$$

one obtains from this relation

$$\dot{\mathbf{R}} = \mathbf{V}_{\mathbf{R}}(\mathbf{R}, U; J; t) , \quad (163)$$

with

$$\mathbf{V}_{\mathbf{R}}(\mathbf{R}, U; J; t) = \frac{e}{m\mathcal{B}} \left[\hat{\mathbf{B}} \left(\frac{\partial \hat{\Phi}}{\partial U} + \frac{1}{c} \frac{\partial \hat{A}_U}{\partial t} \right) + \hat{\mathbf{E}} \times \left(\frac{\partial \hat{\mathbf{A}}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right) \right]. \quad (164)$$

The scalar product of Eq. (159) with $\hat{\mathbf{B}}$ yields

$$\dot{U} = V_U(\mathbf{R}, U; J; t), \quad (165)$$

where

$$V_U(\mathbf{R}, U; J; t) = \frac{e}{m\mathcal{B}} \hat{\mathbf{B}} \cdot \hat{\mathbf{E}}, \quad (166)$$

which, using Eq. (164), can also be written in an alternative form which is sometimes useful,

$$V_U = \frac{\mathbf{V}_{\mathbf{R}} \cdot \hat{\mathbf{E}}}{\frac{\partial \hat{\Phi}}{\partial U} + \frac{1}{c} \frac{\partial \hat{A}_U}{\partial t}}. \quad (167)$$

Thus, 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 $\dot{z}_j(z_1, \dots, z_4; J; t)$, $j = 1, \dots, 4$, obtained from the Lagrangian L_r , Eq. (147), is more involved.

1. The Littlejohn Lagrangian

Littlejohn's Lagrangian, basically in the form of Ref. [3], Eq. (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 Sec. IV. C. 5 and Sec. IV. C. 6. One has

$$\begin{aligned} \mathbf{p}_{-1} &= \frac{e}{c} \mathbf{A}(\mathbf{R}, t), \quad \Phi_{-1} = \Phi(\mathbf{R}, t), \\ \mathbf{x}^{(0)} &= \mathbf{R}, \quad v_{\parallel}^{(0)} = U, \quad v_{\perp}^{(0)} = \sqrt{\frac{2 \Omega_s J^{(1)}}{m 2\pi \epsilon}}, \quad \theta^{(0)} = \phi, \quad \mathbf{v}_E^{(0)} = c \frac{\mathbf{E}(\mathbf{R}, t) \times \mathbf{B}(\mathbf{R}, t)}{B^2}, \\ \mathbf{x}_1 &= \sqrt{\frac{2 \Omega_s J^{(1)}}{m 2\pi \epsilon}} \mathbf{n}_2(\mathbf{R}, \phi, t), \\ \mathbf{p}^{(0)} &= m \left[U \mathbf{b}(\mathbf{R}, t) + \sqrt{\frac{2 \Omega_s J^{(1)}}{m 2\pi \epsilon}} \mathbf{n}_1(\mathbf{R}, \phi, t) + \mathbf{v}_E^{(0)} \right] + \frac{e}{c} \mathbf{A}(\mathbf{R}, t) \\ &+ \frac{e}{c} \mathbf{x}_1 \cdot \nabla_{\mathbf{R}} \mathbf{A}(\mathbf{R}, t), \\ H^{(0)} &= \frac{m}{2} \left[U \mathbf{b}(\mathbf{R}, t) + \sqrt{\frac{2 \Omega_s J^{(1)}}{m 2\pi \epsilon}} \mathbf{n}_1(\mathbf{R}, \phi, t) + \mathbf{v}_E^{(0)} \right]^2 + e\Phi(\mathbf{R}, t) \\ &+ e\mathbf{x}_1 \cdot \nabla_{\mathbf{R}} \Phi(\mathbf{R}, t). \end{aligned} \quad (168)$$

With these relations, one obtains:

$$\begin{aligned}
\hat{\mathbf{A}}^{(0)} &= \frac{c}{e} \langle \mathbf{p}^{(0)} \rangle \\
&= \frac{cm}{e} [U \mathbf{b}(\mathbf{R}, t) + \mathbf{v}_E^{(0)}] + \mathbf{A}(\mathbf{R}, t) , \\
\hat{A}_U^{(0)} &= 0 , \\
e\hat{\Phi}^{(0)} &= -\langle H^{(0)} \rangle \\
&= -\frac{m}{2} [U^2 + (\mathbf{v}_E^{(0)})^2] - \frac{1}{2\pi} \frac{J^{(1)}}{\epsilon} \Omega_s - e\Phi(\mathbf{R}, t) . \tag{169}
\end{aligned}$$

This yields the Lagrangian

$$\begin{aligned}
L_{\text{Littlejohn}} &= \dot{\mathbf{R}} \cdot \left[m [U \mathbf{b}(\mathbf{R}, t) + \mathbf{v}_E^{(0)}] + \frac{e}{c} \mathbf{A}(\mathbf{R}, t) \right] \\
&\quad - \frac{m}{2} [U^2 + (\mathbf{v}_E^{(0)})^2] - \frac{1}{2\pi} \frac{J^{(1)}}{\epsilon} \Omega_s - e\Phi(\mathbf{R}, t) . \tag{170}
\end{aligned}$$

VI. MAXWELL-KINETIC THEORY IN THE REDUCED PHASE SPACE

In this section the formalism of Sec. II is specialized to the reduced Lagrangian in the form of Eq. (150). The situation with this Lagrangian is the following:

A. Quantities entering the Lagrangian for the kinetic theory

1. The canonical momenta in the reduced phase space

These momenta are

$$\mathbf{P}_R = \frac{\partial L_r}{\partial \dot{\mathbf{R}}} = \frac{e}{c} \hat{\mathbf{A}} , \tag{171}$$

$$P_U = \frac{\partial L_r}{\partial \dot{U}} = \frac{e}{c} \hat{A}_U . \tag{172}$$

2. Dirac's constrained Hamiltonian

Therefore, the primary Hamiltonian is

$$\begin{aligned}
H_p &= \dot{\mathbf{R}} \cdot \mathbf{P}_R + \dot{U} P_U - L_r \\
&= e\hat{\Phi} , \tag{173}
\end{aligned}$$

and Dirac's Hamiltonian becomes

$$H_D(\mathbf{R}, U; \mathbf{P}, P_U; J; t) := e\hat{\Phi} + \mathbf{V}_R \cdot \left(\mathbf{P}_R - \frac{e}{c}\hat{\mathbf{A}} \right) + V_U \left(P_U - \frac{e}{c}\hat{A}_U \right), \quad (174)$$

with $\hat{\mathbf{A}}(\mathbf{R}, U; J; t)$, $\hat{A}_U(\mathbf{R}, U; J; t)$ and $\hat{\Phi}(\mathbf{R}, U; J; t)$ from Eqs. (151), (152) and (153), respectively, and $\mathbf{V}_R(\mathbf{R}, U; J; t)$ and $V_U(\mathbf{R}, U; J; t)$ given by Eqs. (164) and (166).

B. Maxwell-kinetic theory

1. The Lagrangian

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

$$L_{\text{tot}}(t) = L_M + L_K, \quad (175)$$

$$L_M = \int d^3x \frac{1}{8\pi} [\mathbf{E}^2(\mathbf{x}, t) - \mathbf{B}^2(\mathbf{x}, t)], \quad (176)$$

$$L_K = - \sum_{\text{particle species}} \int d^3R dU d^4\alpha dJ f_p(\mathbf{R}, U; \alpha_i; J; t) \left[\frac{\partial S}{\partial t}(\mathbf{R}, U; \alpha_i; J; t) + e\hat{\Phi} + \mathbf{V}_R \cdot \left(\frac{\partial S}{\partial \mathbf{R}} - \frac{e}{c}\hat{\mathbf{A}} \right) + V_U \left(\frac{\partial S}{\partial U} - \frac{e}{c}\hat{A}_U \right) \right], \quad (177)$$

with $\hat{\mathbf{A}}$, \hat{A}_U and $\hat{\Phi}$ given by Eqs. (151), (152) and (153) and \mathbf{V}_R and V_U given by Eqs. (164) and (166), respectively, and with the integration over J representing the summation over the parameter J (cf. Sec. V.B).

2. Liouville's theorem

The quantity λ , introduced in Eq. (36) 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 Eqs. (157), (158) and (163)-(166) one finds

$$\frac{\partial \mathcal{B}}{\partial t} + \nabla_{\mathbf{R}} \cdot (\mathcal{B}\mathbf{V}_R) + \frac{\partial}{\partial U} (\mathcal{B}V_U) = 0. \quad (178)$$

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

$$d\tau_{\mathbf{R}, U} := |\mathcal{B}(R_1, \dots, R_3, U; J; t)| dR_1 dR_2 dR_3 dU \quad (179)$$

is Liouillian, *i.e.*

$$\frac{d}{dt}(d\tau_{\mathbf{R},U}) = 0, \quad (180)$$

where the time derivative along orbits in this extended phase space is given by Eq. (156). 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 ,

$$d\tau := |\mathcal{B}(R_1, \dots, R_3, U; J; t)| dR_1 dR_2 dR_3 dU dJ. \quad (181)$$

Hence, it holds that

$$\lambda = |\mathcal{B}(R_1, \dots, R_3, U; J; t)|. \quad (182)$$

3. The kinetic equation

The kinetic equation for the distribution function f , which here takes the form

$$\frac{\partial f}{\partial t} + \frac{e}{m\mathcal{B}} \left[\frac{\partial \hat{\Phi}}{\partial U} \hat{\mathbf{B}} - \frac{\partial \hat{\mathbf{A}}}{\partial U} \times \hat{\mathbf{E}} \right] \cdot \frac{\partial f}{\partial \mathbf{R}} + \frac{e}{m\mathcal{B}} \hat{\mathbf{B}} \cdot \hat{\mathbf{E}} \frac{\partial f}{\partial U} = 0, \quad (183)$$

is obtained from the Lagrangian (177) by variation of S , in a way similar to that in Sec. II.D.

4. Sketch of a procedure for obtaining the inhomogeneous Maxwell equations and conservation laws

Variation of the total action integral

$$\mathcal{A} = \int_{t_1}^{t_2} L_{\text{tot}} dt \quad (184)$$

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

$$\begin{aligned} f_p d^4\alpha &\rightarrow \mathcal{B} \delta \left(\mathbf{P}_{\mathbf{R}} - \frac{e}{c} \hat{\mathbf{A}} \right) \delta \left(P_U - \frac{e}{c} \hat{A}_U \right) f(\mathbf{R}, U; J; t) d^3\mathbf{P}_{\mathbf{R}} dP_U, \\ \frac{\partial S}{\partial \mathbf{R}} &\rightarrow \mathbf{P}_{\mathbf{R}}, \\ \frac{\partial S}{\partial U} &\rightarrow P_U, \end{aligned} \quad (185)$$

where \mathcal{B} is given by Eq. (162).

This program will be carried out in paper II with special application to the gyrokinetic theory.

VII. SUMMARY

The paper starts with the introduction and proof of the correctness of a Lagrangian for the combined Maxwell-kinetic theory 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" (guiding center), 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 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. 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 perturbation theory with respect to small gyroperiods, and in the case of drift-kinetic ordering also to small gyroradii, is done solely within the framework of the rather simple part of the inverse Kruskal formalism derived explicitly in this paper. 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.

The forthcoming paper will apply the above method to the gyrokinetic ordering. There one will encounter complications compared with the linearized theory 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.

APPENDIX A: DIRECT PROOF OF $\frac{d}{dt}J(\hat{\mathbf{z}}, t) = 0$

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

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

Its time derivative is

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

This relation can also be written in the form

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

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

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

Note that it was essential that h , Eq. (101), 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 (140)-(142) are derived in this appendix.

In the phase-space coordinates z_1, \dots, z_4, J, ϕ the constant of the motion J , Eq. (102), can be written as

$$J = \oint_{\{z_1, \dots, z_4, J, t\} = \text{const}} \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} d\tilde{\phi} = \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} d\tilde{\phi} . \quad (\text{B1})$$

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

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

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

and

$$0 = \int_0^1 \left[\frac{\partial \mathbf{p}}{\partial t} \cdot \frac{\partial \mathbf{r}}{\partial \tilde{\phi}} - \frac{\partial \mathbf{p}}{\partial \tilde{\phi}} \cdot \frac{\partial \mathbf{r}}{\partial t} \right] d\tilde{\phi} := \int_0^1 l_t d\tilde{\phi} , \quad (\text{B4})$$

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 (104)-(107) in Sec. IV, but employing the phase-space coordinates z_1, \dots, z_4, J, ϕ instead of $\hat{z}_1, \dots, \hat{z}_5, \phi$.

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

$$H(\mathbf{x}, \mathbf{p}, t = \epsilon s) = H(\mathbf{r}(z_1, \dots, z_4, J, \phi; t = \epsilon s), \mathbf{p}(z_1, \dots, z_4, J, \phi; t = \epsilon s); t = \epsilon s) , \quad (\text{B5})$$

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

$$\begin{aligned} \frac{\partial H}{\partial z_i} &= \left. \frac{\partial H}{\partial \mathbf{r}} \right|_{\mathbf{p}, t} \cdot \frac{\partial \mathbf{r}}{\partial z_i} + \left. \frac{\partial H}{\partial \mathbf{p}} \right|_{\mathbf{r}, t} \cdot \frac{\partial \mathbf{p}}{\partial z_i} \\ &= -\frac{d\mathbf{p}}{dt} \cdot \frac{\partial \mathbf{r}}{\partial z_i} + \frac{d\mathbf{r}}{dt} \cdot \frac{\partial \mathbf{p}}{\partial z_i} , \end{aligned} \quad (\text{B6})$$

$$\frac{\partial H}{\partial J} = -\frac{d\mathbf{p}}{dt} \cdot \frac{\partial \mathbf{r}}{\partial J} + \frac{d\mathbf{r}}{dt} \cdot \frac{\partial \mathbf{p}}{\partial J} , \quad (\text{B7})$$

$$\frac{\partial H}{\partial \phi} = -\frac{d\mathbf{p}}{dt} \cdot \frac{\partial \mathbf{r}}{\partial \phi} + \frac{d\mathbf{r}}{dt} \cdot \frac{\partial \mathbf{p}}{\partial \phi} \quad (\text{B8})$$

and

$$\left. \frac{\partial H}{\partial t} \right|_{z_i, J, \phi} = -\frac{d\mathbf{p}}{dt} \cdot \frac{\partial \mathbf{r}}{\partial t} + \frac{d\mathbf{r}}{dt} \cdot \frac{\partial \mathbf{p}}{\partial t} + \frac{dH}{dt}, \quad (\text{B9})$$

where $\left. \frac{\partial H}{\partial t} \right|_{\mathbf{r}, \mathbf{p}} = \frac{dH}{dt}$ has been used in the last equation. With $\frac{\partial}{\partial \phi} \left(\frac{\partial H}{\partial z_i} \right) = \frac{\partial}{\partial z_i} \left(\frac{\partial H}{\partial \phi} \right)$ (and similarly for J and t) taken into account, these equations yield

$$-\frac{\partial \mathbf{r}}{\partial z_i} \cdot \frac{\partial}{\partial \phi} \left[\frac{d\mathbf{p}}{dt} \right] + \frac{\partial \mathbf{p}}{\partial z_i} \cdot \frac{\partial}{\partial \phi} \left[\frac{d\mathbf{r}}{dt} \right] = -\frac{\partial \mathbf{r}}{\partial \phi} \cdot \frac{\partial}{\partial z_i} \left[\frac{d\mathbf{p}}{dt} \right] + \frac{\partial \mathbf{p}}{\partial \phi} \cdot \frac{\partial}{\partial z_i} \left[\frac{d\mathbf{r}}{dt} \right], \quad (\text{B10})$$

$$-\frac{\partial \mathbf{r}}{\partial J} \cdot \frac{\partial}{\partial \phi} \left[\frac{d\mathbf{p}}{dt} \right] + \frac{\partial \mathbf{p}}{\partial J} \cdot \frac{\partial}{\partial \phi} \left[\frac{d\mathbf{r}}{dt} \right] = -\frac{\partial \mathbf{r}}{\partial \phi} \cdot \frac{\partial}{\partial J} \left[\frac{d\mathbf{p}}{dt} \right] + \frac{\partial \mathbf{p}}{\partial \phi} \cdot \frac{\partial}{\partial J} \left[\frac{d\mathbf{r}}{dt} \right] \quad (\text{B11})$$

and

$$-\frac{\partial \mathbf{r}}{\partial t} \cdot \frac{\partial}{\partial \phi} \left[\frac{d\mathbf{p}}{dt} \right] + \frac{\partial \mathbf{p}}{\partial t} \cdot \frac{\partial}{\partial \phi} \left[\frac{d\mathbf{r}}{dt} \right] + \frac{\partial}{\partial \phi} \left[\frac{dH}{dt} \right] = -\frac{\partial \mathbf{r}}{\partial \phi} \cdot \frac{\partial}{\partial t} \left[\frac{d\mathbf{p}}{dt} \right] + \frac{\partial \mathbf{p}}{\partial \phi} \cdot \frac{\partial}{\partial t} \left[\frac{d\mathbf{r}}{dt} \right]. \quad (\text{B12})$$

In the phase-space coordinates z_1, \dots, z_4, J, ϕ , the total time derivative is given by

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

(cf. Eqs. (101)). It is easily seen that

$$\frac{\partial}{\partial \phi} \frac{d}{dt} = \frac{d}{dt} \frac{\partial}{\partial \phi}, \quad (\text{B14})$$

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

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

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

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

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

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

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

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

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

Therefore,

$$\frac{\partial\mathbf{p}}{\partial z_i} \cdot \frac{\partial\mathbf{r}}{\partial\phi} - \frac{\partial\mathbf{p}}{\partial\phi} \cdot \frac{\partial\mathbf{r}}{\partial z_i} = 0, \quad i = 1, \dots, 4, \quad (\text{B21})$$

$$\frac{\partial\mathbf{p}}{\partial J} \cdot \frac{\partial\mathbf{r}}{\partial\phi} - \frac{\partial\mathbf{p}}{\partial\phi} \cdot \frac{\partial\mathbf{r}}{\partial J} = 1, \quad (\text{B22})$$

and

$$\frac{\partial\mathbf{p}}{\partial t} \cdot \frac{\partial\mathbf{r}}{\partial\phi} - \frac{\partial\mathbf{p}}{\partial\phi} \cdot \frac{\partial\mathbf{r}}{\partial t} + \frac{\partial H}{\partial\phi} = 0. \quad (\text{B23})$$

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, \dots, z_4, J, ϕ . Then, from

$$q_\lambda = q_\lambda(\mathbf{r}(q_\mu, t), \mathbf{p}(q_\mu, t), t), \quad (\text{B24})$$

one obtains

$$\frac{\partial q_\lambda}{\partial q_\mu} = \delta_{\lambda\mu} = \frac{\partial\mathbf{r}}{\partial q_\mu} \cdot \frac{\partial q_\lambda}{\partial\mathbf{r}} + \frac{\partial\mathbf{p}}{\partial q_\mu} \cdot \frac{\partial q_\lambda}{\partial\mathbf{p}}, \quad (\text{B25})$$

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

$$\delta_{\lambda\mu} = \sum_{\nu=1}^6 [q_\lambda, q_\nu] \left[\frac{\partial\mathbf{r}}{\partial q_\mu} \cdot \frac{\partial\mathbf{p}}{\partial q_\nu} - \frac{\partial\mathbf{p}}{\partial q_\mu} \cdot \frac{\partial\mathbf{r}}{\partial q_\nu} \right]. \quad (\text{B26})$$

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

$$[z_i, J] = 0 \quad i = 1, \dots, 4 \quad (\text{B27})$$

and

$$[\phi, J] = 1. \quad (\text{B28})$$

APPENDIX C: ALTERNATIVE DERIVATION OF A LIOUVILLIAN VOLUME ELEMENT

In Sec. VI . B . 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-center theories.

The Lagrangian equations derived from L_r , Eq. (147), can be written as

$$\sum_{i=1}^4 \dot{z}_i \omega_{ik} = \frac{\partial H}{\partial z_k} + \left\langle \frac{\partial \mathbf{p}}{\partial t} \cdot \frac{\partial \mathbf{r}}{\partial z_k} - \frac{\partial \mathbf{p}}{\partial z_k} \cdot \frac{\partial \mathbf{r}}{\partial t} \right\rangle, \quad i = 1, \dots, 4, \quad (\text{C1})$$

with ω_{ik} the antisymmetric matrix given by

$$\omega_{ik} := \left\langle \frac{\partial \mathbf{r}}{\partial z_i} \cdot \frac{\partial \mathbf{p}}{\partial z_k} - \frac{\partial \mathbf{r}}{\partial z_k} \cdot \frac{\partial \mathbf{p}}{\partial z_i} \right\rangle. \quad (\text{C2})$$

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

$$d\tau_\lambda := \lambda(z_1, \dots, z_4; J; t) dz_1 dz_2 dz_3 dz_4, \quad (\text{C3})$$

with

$$\lambda := |\det(\omega_{ik})|^{1/2}. \quad (\text{C4})$$

A proof of this theorem is given in Appendix D. Here, an expression for λ will be derived, and it will be shown that it coincides with \mathcal{B} , Eq. (162).

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

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

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 Eq. (C1) with Eq. (159):

$$\omega_{k4} = \frac{e}{c} \hat{\mathbf{x}}_k \cdot \left[\frac{\partial \hat{A}}{\partial U} - \frac{\partial \hat{A}_U}{\partial \mathbf{R}} \right], \quad k = 1, \dots, 3, \quad (\text{C6})$$

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

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

Since $\hat{\mathbf{B}} \cdot (\partial \hat{\mathbf{A}} / \partial U - \partial \hat{A}_U / \partial \mathbf{R}) = (mc/e)\mathcal{B}$ according Eq. (162), one obtains

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

The Liouvillian volume element given by Eq. (C3) is thus the same (within a constant factor) as that given by Eq. (179).

APPENDIX D: LIOUVILLE'S THEOREM

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

$$\sum_{i=1}^n V_i \omega_{ik} = \frac{\partial \gamma_k}{\partial t} + \frac{\partial H_p}{\partial \zeta_k}, \quad (\text{D1})$$

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

$$\omega_{ik} = \frac{\partial \gamma_i}{\partial \zeta_k} - \frac{\partial \gamma_k}{\partial \zeta_i}. \quad (\text{D2})$$

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

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

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

$$\Delta = \sum_{k=1}^n \omega_{ik} C_{ik}, \quad (\text{D4})$$

where i can be any of the numbers $1, \dots, n$. 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,

$$\frac{\partial \Delta}{\partial \omega_{ik}} = C_{ik}. \quad (\text{D5})$$

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

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

By using the relation

$$(C_{ik}) = \Delta (J_{ki}) , \quad (D7)$$

where the matrix (J_{ik}) is the inverse of the matrix (ω_{ik}) , *i.e.*

$$\sum_l \omega_{il} J_{lk} = \delta_{ik} , \quad (D8)$$

one obtains

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

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

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

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

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

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

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

can be written as

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

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

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

By inserting this in Eq. (D9), one obtains

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

The quantity

$$\lambda \equiv |\Delta|^{1/2} \quad (D16)$$

then satisfies the continuity equation

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

Thus, the volume element

$$d\tau \equiv \lambda d\zeta_1 \dots d\zeta_n \quad (\text{D18})$$

is Liouvillian, *i.e.*

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

*

References

- [1] R. G. Littlejohn, *J. Math. Phys.* **20**, 2445 (1979).
- [2] R. G. Littlejohn, *Phys. Fluids* **24**, 1730 (1981).
- [3] R. G. Littlejohn, *J. Plasma Phys.* **29**, 111 (1983).
- [4] W. Weyssow and R. Balescu, *J. Plasma Phys.* **35**, 449 (1986).
- [5] A. Brizard, *J. Plasma Phys.* **41**, 541 (1989).
- [6] H. Sugama, *Phys. Plasmas* **7**, 466 (2000).
- [7] A. Brizard, *Phys. Plasmas* **1**, 2460 (1994).
- [8] A. Brizard, *Phys. Plasmas* **1**, 2473 (1994).
- [9] D. Pfirsch, *Z. Naturforsch.* **39a**, 1 (1984).
- [10] D. Pfirsch and P. J. Morrison, *Phys. Rev. A* **32**, 1714 (1985).
- [11] D. Pfirsch and P. J. Morrison, *Phys. Fluids B* **3**, 271 (1991).
- [12] D. Correa-Restrepo, D. Pfirsch, and H. K. Wimmel, *Physica A* **136**, 453 (1986).
- [13] M. Kruskal, *J. Math. Phys.* **3**, 806 (1962).
- [14] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York, 1964), Lecture No.1.
- [15] K. Sundermeyer, *Constrained Dynamics* (Springer Verlag, Berlin, 1982), Chapters I and II.
- [16] J. Larsson, *J. Math. Phys.* **27**, 495 (1986).
- [17] E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (John Wiley & Sons, New York, 1974), pp. 11 and 79.
- [18] D. Pfirsch, *Arch. Rational Mech. Anal.* **29**, 323 (1968).
- [19] J. H. Van Vleck, *Proc. Nat. Acad. Sci.* **14**, 178 (1928).

- [20] M. Kruskal, in *Plasma Physics* (IAEA, Vienna, 1965), p. 100.
- [21] R. Balescu, *Transport Processes in Plasmas* (North-Holland, Amsterdam, 1988), Vol. I, p. 54.