Fields Institute Communications
Volume **46**, 2005

A Short Introduction to General Gyrokinetic Theory

H. Qin
Princeton Plasma Physics Laboratory
Princeton University
Princeton, NJ 08543, USA
HongQin@princeton.edu

Abstract. Interesting plasmas in the laboratory and space are magnetized. General gyrokinetic theory is about a symmetry, gyro-symmetry, in the Vlasov-Maxwell system for magnetized plasmas. The most general gyrokinetic theory can be geometrically formulated. First, the coordinate-free, geometric Vlasov-Maxwell equations are developed in the 7D phase space, which is defined as a fiber bundle over the spacetime. The Poincaré-Cartan-Einstein 1-form pullbacked onto the 7D phase space determines particles' worldlines in the phase space, and realizes the momentum integrals in kinetic theory as fiber integrals. The infinite small generator of the gyro-symmetry is then asymptotically constructed as the base for the gyrophase coordinate of the gyrocenter coordinate system. This is accomplished by applying the Lie coordinate perturbation method to the Poincaré-Cartan-Einstein 1-form, which also generates the most relaxed condition under which the gyro-symmetry still exists. General gyrokinetic Vlasov-Maxwell equations are then developed as the Vlasov-Maxwell equations in the gyrocenter coordinate system, rather than a set of new equations. Since the general gyrokinetic system developed is geometrically the same as the Vlasov-Maxwell equations, all the coordinate independent properties of the Vlasov-Maxwell equations, such as energy conservation, momentum conservation, and Liouville volume conservation, are automatically carried over to the general gyrokinetic system. The pullback transformation associated with the coordinate transformation is shown to be an indispensable part of the general gyrokinetic Vlasov-Maxwell equations. Without this vital element, a number of prominent physics features, such as the presence of the compressional Alfvén wave and a proper description of the gyrokinetic equilibrium, cannot be readily recovered. Three examples of applications of the general gyrokinetic theory developed in the areas of plasma equilibrium and plasma waves are given. Interesting topics, such as gyro-center gauge and gyro-gauge, are discussed as well.

2000 *Mathematics Subject Classification.* Primary 82C05, 70H12; Secondary 53Z05.
This research was supported by the U.S. Department of Energy under contract AC02-76CH03073, and by Lawrence Livermore National Laboratory's LDRD Project 04-SI-03, Kinetic Simulation of Boundary Plasma Turbulent Transport.

1 Introduction

General gyrokinetic theory is about a symmetry, gyro-symmetry, in the Vlasov-Maxwell system for magnetized plasmas. In addition to its theoretical importance and elegance, gyro-symmetry can be employed as an effective numerical algorithm for modern large scale computer simulations for magnetized plasmas. Historically, gyrokinetic theory has been developed in various formats in different context [2,4,8–10,12,15,18,22,24,25,27,28,30–32,35–37,41,43,45–51,53,54,56,59,60]. However, gyrokinetic theory can be put into a form much more general and geometric than those found in literature. Here, we will geometrically develop such a general gyrokinetic theory, and leave the computational side of the story [5,11,13, 14,19,21,23,33,34,42,57] to Ref. [58].

2 Geometric Vlasov-Maxwell equations

Since we are looking for the gyro-symmetry of the Vlasov-Maxwell equations, it is necessary to first develop a geometric point of view for the Vlasov-Maxwell equations. Because it turns out that the geometry of the Vlasov-Maxwell equations is best manifested in the spacetime of special relativity, we will start from the relativistic Vlasov-Maxwell equations. The phase space where the Vlasov-Maxwell equations reside is a 7-dimensional manifold

$$P = \{(x,p) \mid x \in M,\ p \in T_x^* M,\ g^{-1}(p,p) = -m^2c^2\}\ , \tag{2.1}$$

where M is the 4-dimensional spacetime, T^*M is the 8-dimensional cotangent bundle of M, and g^{-1} is the inverse of the metric tensor of M defined by

$$(g^{-1})^{\alpha\beta} g_{\beta\gamma} = \delta^\alpha_\gamma\ . \tag{2.2}$$

The phase space is a fiber bundle over spacetime M (see Fig. 1),

$$\pi :\ P \longrightarrow M\ . \tag{2.3}$$

The worldlines of particles on P are determined by the Poincaré-Cartan-Einstein 1-form constructed as follows. First, take the only two geometric objects related to the dynamics of charged particles, the momentum 1-form p and the potential 1-form A on M, then perform the only nontrivial operation, i.e., addition with the right units, to let particles interact with fields,

$$\hat{\gamma} = \frac{e}{c} A + p\ . \tag{2.4}$$

$\hat{\gamma}$ is what we call Poincaré-Cartan-Einstein 1-form on the spacetime M. In a Cartesian inertial coordinate system x^μ $(\mu = 0,1,2,3)$,

$$x^0 = ct \text{ and } A_0 = -\phi\ . \tag{2.5}$$

The Poincaré-Cartan-Einstein 1-form on the phase space P is obtained by pulling back $\hat{\gamma}$,

$$\gamma = \pi^* \hat{\gamma}\ . \tag{2.6}$$

Particles' dynamics is determined by Hamilton's equation

$$i_\tau d\gamma = 0\ , \tag{2.7}$$

where τ is a vector field, whose integrals are particle's worldlines on P.

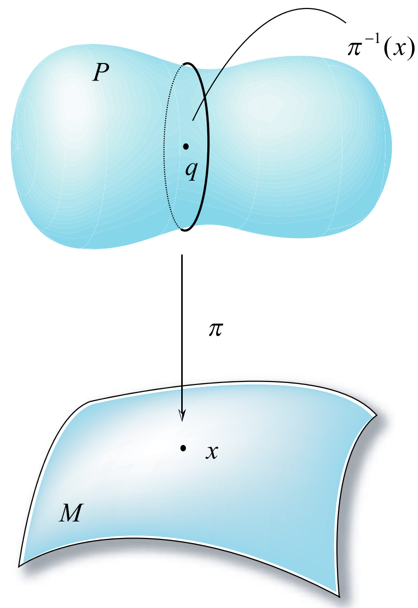

Figure 1 Phase space and fiber integral.

Very elegantly, the Poincaré-Cartan-Einstein 1-form γ also gives the necessary "volume form" needed for the fundamental "velocity integrals" in kinetic theory. Define the Liouville 6-form ω on the 7D phase space P as

$$\omega = -\frac{1}{3!m^3} d\gamma \wedge d\gamma \wedge d\gamma \, . \tag{2.8}$$

We take the viewpoint that the "velocity integrals" in kinetic theory are geometrically fiber integrals [26] defined as follow. For $x \in M$, and $q \in \pi^{-1}(x) \subset P$(see Fig. 1), consider the form

$$\omega_x(q)(u_1, u_2, u_3)[v_1, v_2, v_3] \equiv \omega(q)(u_1, u_2, u_3, \tilde{v}_1, \tilde{v}_2, \tilde{v}_3) \ , \tag{2.9}$$

where

$$u_i \in T_q[\pi^{-1}(x)], \ v_i \in T_x M, \ T_q\pi(\tilde{v}_i) = v_i, \ \tilde{v}_i \in T_q P \ , \ (i = 1, 2, 3).$$

Actually, $\tilde{v}_i$ is not unique because in general $T_q\pi$ is not injective. However, $\omega_x(q)$ is well defined because according to the submersion theorem,

$$Ker(T_q\pi) = T_q[\pi^{-1}(x)] \ . \tag{2.10}$$

Therefore, $\omega_x(q)$ is a 3-form on $\pi^{-1}(x)$, valued in 3-forms on M. The 3-form flux on M corresponding to a distribution function $f : P \longrightarrow R$ is the result of integration of $f\omega_x$ over the fiber $\pi^{-1}(x)$ at x,

$$j(x) = \int_{\pi^{-1}(x)} f\omega_x \ . \tag{2.11}$$

The fact that $j(x)$ is the conventional 3-form flux can be verified by expressing ω in a coordinate system composed of inertial coordinates x^μ $(\mu = 0, 1, 2, 3)$ for M and three corresponding coordinate p_i with $i = 1, 2,$ and 3 for T_xM. In this coordinate system we have the following expressions in the phase space P,

$$p_0 = -\sqrt{m^2c^2+p^2} \ , \tag{2.12}$$

$$d\gamma = \frac{e}{c}A_{i,j}dx^j \wedge dx^i + dp_i \wedge dx^i - e\phi_{,j}dx^j \wedge dt - c\frac{\partial}{\partial p_i}\sqrt{m^2c^2+p^2}dp_i \wedge dt \ , \tag{2.13}$$

$$\begin{aligned}\omega = dp_1 \wedge dp_2 \wedge dp_3 \wedge \Bigg(& dx^1 \wedge dx^2 \wedge dx^3 - \frac{p_1}{m\gamma_r}dt \wedge dx^2 \wedge dx^3 \\ & - \frac{p_2}{m\gamma_r}dx^1 \wedge dt \wedge dx^3 - \frac{p_3}{m\gamma_r}dx^1 \wedge dx^2 \wedge dt\Bigg) \ ,\end{aligned} \tag{2.14}$$

where

$$\gamma_r = \sqrt{1+\frac{p^2}{m^2c^2}} \ . \tag{2.15}$$

The Maxwell equations are

$$d * dA = 4\pi e \int_{\pi^{-1}(x)} f\omega_x \ , \tag{2.16}$$

where $*\alpha$ is the Hodge-dual of α on spacetime M. Overall, the Vlasov-Maxwell equations on the 7D phase space P can be geometrically written as

$$df(v) = 0, \ \ i_v d\gamma = 0 \ , \ \text{and } d * dA = 4\pi e \int_{\pi^{-1}(x)} f\omega_x \ . \tag{2.17}$$

3 Noether's theorem, symmetries, Kruskal ring, and Lie coordinate Perturbation

Noether's theorem links symmetries and invariants. Here, we cast the theorem in the form of forms. Define a symmetry vector field η (infinite small generator) of γ to be a vector field that satisfies

$$L_\eta \gamma = ds \tag{3.1}$$

for some $s : P \longrightarrow R$, where L_η is the Lie derivative. η generates a 1-parameter symmetry group for γ. Using Cartan's magic formula, we have

$$d(\gamma \cdot \eta) + i_\eta d\gamma = ds \ . \tag{3.2}$$

For the vector field τ of a worldline,

$$d(\gamma \cdot \eta) \cdot \tau = ds \cdot \tau \ , \tag{3.3}$$

which implies that $\gamma \cdot \eta - s$ is an invariant.

In the present study, we will only consider the non-relativistic case in an inertial coordinate system for M with $x^0 = ct$. In addition, we chose three corresponding coordinate p_i $(i = 1, 2, 3)$ as the fiber coordinates for P at x with

$$p_0 = -\sqrt{m^2c^2+p^2} = -mc - \frac{1}{2}\frac{p^2}{mc} + O\left[(\frac{p}{mc})^4\right] \ . \tag{3.4}$$

We normalize γ by m, A by mc/e, and ϕ by m/e such that

$$\gamma = (\boldsymbol{A}+\boldsymbol{v}) \cdot d\boldsymbol{x} - \left[\frac{v^2}{2}+\phi\right]dt \ , \tag{3.5}$$

$$\boldsymbol{v} \equiv \boldsymbol{p}/m \ . \tag{3.6}$$

Here, the bold mathematical symbols $\boldsymbol{A}$, $\boldsymbol{v}$, and $\boldsymbol{p}$ represent the $i=1,2,3$ components of the 1-forms A, v, and p, $\boldsymbol{dx}$ represents dx^i $(i=1,2,3)$, and $(\boldsymbol{A}+\boldsymbol{v})\cdot\boldsymbol{dx}$ is just a shorthand notation for $\sum_{i=1,2,3}(A_i+v_i)dx^i$. The normalizations for γ, ϕ, and A will be used thereafter, unless it is explicitly stated otherwise.

The symmetry for γ that we are interested is an approximate one. It is an exact symmetry when the electromagnetic fields are constant in spacetime. To demonstrate the basic concept, we first consider the case of constant magnetic field without electrical field. Because of its simplicity, there are several symmetries admitted by γ. The gyro-symmetry is the symmetry given by the infinite small generator (vector field)

$$\eta = v_x\left(\frac{1}{B}\frac{\partial}{\partial x}+\frac{\partial}{\partial v_y}\right)+v_y\left(\frac{1}{B}\frac{\partial}{\partial y}-\frac{\partial}{\partial v_x}\right)\ . \tag{3.7}$$

Applying Noether's theorem, we can verify that the corresponding invariant is the magnetic moment

$$\mu=\frac{v_x^2+v_y^2}{2B}\ . \tag{3.8}$$

The gyro-symmetry η has a rather complicated expression in the Cartesian coordinates (x,y,v_x,v_y). A new coordinate will be constructed such that η is a coordinate base

$$\eta=\frac{\partial}{\partial\theta}\ , \tag{3.9}$$

where θ is the gyrophase coordinate. Eq. (3.7) indicates that the gyro-symmetry η is neither a rotation in the momentum space, nor a rotation in the configuration space or its prolongated version in the phase space. Therefore, θ is not a momentum coordinate or a configuration coordinate. It is a phase space coordinate that depends on particles' momentum as well as their spacetime positions. We will call the orbit of η in phase space Kruskal ring, and points on which Kruskal ring mates [31], which are illustrated in Fig. 2. Shown in Fig. 3 is direct laboratory observation of charged particle gyro-motion in magnetic field [1]. It is the projection of the Kruskal ring in the configuration space.

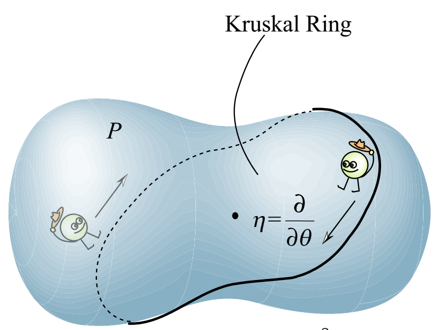

Figure 2 The orbit of the gyro-symmetry $\eta=\frac{\partial}{\partial\theta}$ is Kruskal ring. Points on the ring are ring mates [31].

When the fields are not constant in spacetime, the gyro-symmetry η in Eq. (3.7) is broken. We therefore seek an asymptotic symmetry when the spacetime inhomogeneity is weak. Finding the most relaxed conditions of spacetime inhomogeneity under which an asymptotic gyro-symmetry still exists is our goal as well. The strategy to achieve our objectives has two steps. (i) Construct a non-fibered,

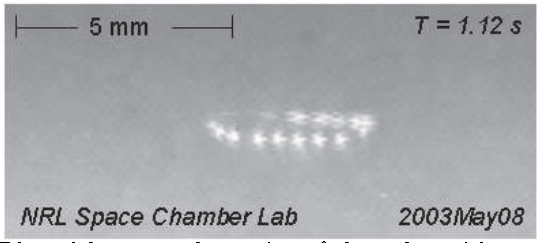

Figure 3 Direct laboratory observation of charged particle gyro-motion in magnetic field. [1]. (Reprint permitted by AIP and Dr. Amatucci.)

non-canonical phase space coordinate system $\bar{Z} = (\bar{\boldsymbol{X}}, \bar{u}, \bar{w}, \bar{\theta})$ such that γ can be expanded into an asymptotic series

$$\gamma = \bar{\gamma}_0 + \bar{\gamma}_1 + \bar{\gamma}_2 + \ \dots \ , \tag{3.10}$$

where $\bar{\gamma}_1 \sim \varepsilon\bar{\gamma}_0$, $\bar{\gamma}_2 \sim \varepsilon\bar{\gamma}_1$, and $\varepsilon \ll 1$. $\bar{Z}$ is the called the zeroth order gyrocenter coordinate. In addition, $\bar{\gamma}_0$ admits the gyro-symmetry $\eta = \partial/\partial\bar{\theta}$, but $\bar{\gamma}_1$ does not necessarily; (ii) Introduce a coordinate perturbation transformation such that in the new coordinates $Z = (\boldsymbol{X}, u, w, \theta)$, γ_1 admits the gyro-symmetry $\eta = \partial/\partial\theta$. In fact, we will seek a stronger symmetry condition

$$\partial\gamma/\partial\theta = 0 \ ,$$

which is sufficient for $\eta = \partial/\partial\theta$ to satisfy Eq. (3.1). Z is the called the first order gyrocenter coordinate. The small parameter ε measures the weakness of spacetime inhomogeneity of the fields. The coordinate perturbation transformation procedure indicates that the most relaxed conditions for the existence of an asymptotic gyro-symmetry is

$$\boldsymbol{E} \equiv \boldsymbol{E}^s + \boldsymbol{E}^l, \ \boldsymbol{B} \equiv \boldsymbol{B}^s + \boldsymbol{B}^l, \tag{3.11}$$

$$\boldsymbol{E}^l \sim \frac{v \times \boldsymbol{B}^l}{c} \ , \ \boldsymbol{E}^s \sim \varepsilon \ \frac{v \times \boldsymbol{B}^l}{c} \ , \ B^s \sim \varepsilon \ B^l, \tag{3.12}$$

$$\left(|\boldsymbol{\rho}| \frac{\nabla E^l}{E^l}, \frac{1}{\Omega E^l}\frac{\partial E^l}{\partial t}\right) \sim \left(|\boldsymbol{\rho}| \frac{\nabla B^l}{B^l}, \frac{1}{\Omega B^l}\frac{\partial B^l}{\partial t}\right) \sim \varepsilon \ , \tag{3.13}$$

$$\left(\ |\boldsymbol{\rho}| \frac{\nabla E^s}{E^s}, \frac{1}{\Omega E^s}\frac{\partial E^s}{\partial t}\right) \sim \left(\ |\boldsymbol{\rho}| \frac{\nabla B^s}{B^s}, \frac{1}{\Omega B^s}\frac{\partial B^s}{\partial t}\right) \sim 1 \ , \tag{3.14}$$

where the fields were split into two parts. $(\boldsymbol{E}^l, \boldsymbol{B}^l)$ are the large amplitude parts with long spacetime scale length comparable to the spacetime gyroradius $\rho = (\boldsymbol{\rho}, 1/\Omega)$, and $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ are the small amplitude parts with spacetime scale length smaller than the spacetime gyroradius.

The coordinate perturbation method we adopt belongs to the class of perturbation techniques generally referred as Lie perturbation method [3, 6, 7, 37]. A coordinate transformation for the 7D phase space P can be locally represented by a map between two subsets of the R^7 space, $T : z \longmapsto Z = T(z)$. As illustrated in Fig. 4, for the same point p in phase space, there could be more than one coordinate systems (patches). The correspondence between two different coordinate systems for the same point in phase space is the coordinate transformation. In the present study, we assume a coordinate transformation can be represented by a single map almost everywhere. The subset of phase space which can not be covered by the single map has zero measure and does not contribute to the fiber integrals.

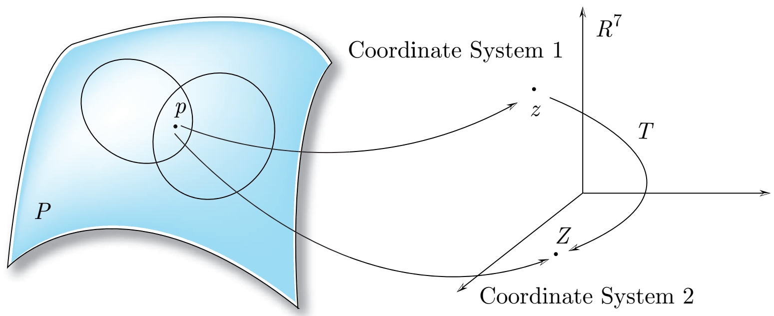

Figure 4 Coordinate transformation as a map in R^7.

To see how γ is transformed by T, let $Z = z + G_1(z)$ and $G_1(z) \sim \varepsilon$,

$$\begin{aligned}\Gamma(Z) &= \gamma(z) = \gamma\left[Z - G_1(z)\right] = \gamma\left[Z - G_1(Z) + O(\varepsilon^2)\right] \\ &= \gamma(Z) - L_{G_1(Z)}\gamma(Z) + O(\varepsilon^2) \\ &= \gamma(Z) - i_{G_1(Z)}d\gamma(Z) - d\left[\gamma \cdot G_1(Z)\right] + O(\varepsilon^2)\ .\end{aligned} \tag{3.15}$$

If γ is an asymptotic series as in Eq. (3.10),

$$\Gamma(Z) = \Gamma_0(Z) + \Gamma_1(Z) + O(\varepsilon^2)\ , \tag{3.16}$$

$$\Gamma_0(Z) = \gamma_0(Z)\ , \tag{3.17}$$

$$\Gamma_1(Z) = \gamma_1(Z) - i_{G_1(Z)}d\gamma_0(Z) - d\left[\gamma_0 \cdot G_1(Z)\right]\ . \tag{3.18}$$

Similar procedure can be straightforwardly carried out to second order. Let $Z = z + G_1(z) + G_2(z)$,

$$\Gamma_2(Z) = \gamma_2(Z) - L_{G_1(Z)}\gamma_1(Z) + \left(\frac{1}{2}L^2_{G_1(Z)} - L_{G_2(Z)}\right)\gamma_0(Z)\ . \tag{3.19}$$

4 Gyrocenter coordinates

To construct the zeroth order gyrocenter coordinate $\bar{Z} = (\bar{\boldsymbol{X}}, \bar{u}, \bar{w}, \bar{\theta})$, we first define two vector fields on M (or more rigorously sections of a vector bundle over the spacetime M)

$$\boldsymbol{D}(\boldsymbol{y}) \equiv \frac{\boldsymbol{E}^l(\boldsymbol{y}) \times \boldsymbol{B}^l(\boldsymbol{y})}{\left[B^l(\boldsymbol{y})\right]^2}\ ,\ \boldsymbol{b}(\boldsymbol{y}) \equiv \frac{\boldsymbol{B}^l(\boldsymbol{y})}{B^l(\boldsymbol{y})}\ , \tag{4.1}$$

where $\boldsymbol{y} \in M$. In addition, we define the following vector fields which also depend on $\boldsymbol{v_x}$, the velocity at another spacetime position $\boldsymbol{x} \in M$,

$$u(\boldsymbol{y}, \boldsymbol{v_x})\boldsymbol{b}(\boldsymbol{y}) \equiv [\boldsymbol{v_x}(\boldsymbol{y}) - \boldsymbol{D}(\boldsymbol{y})] \cdot\ \boldsymbol{b}(\boldsymbol{y})\ \boldsymbol{b}(\boldsymbol{y})\ , \tag{4.2}$$

$$w(\boldsymbol{y}, \boldsymbol{v_x})\boldsymbol{c}(\boldsymbol{y}, \boldsymbol{v_x}) \equiv [\boldsymbol{v_x}(\boldsymbol{y}) - \boldsymbol{D}(\boldsymbol{y})] \times \boldsymbol{b}(\boldsymbol{y}) \times \boldsymbol{b}(\boldsymbol{y})\ , \tag{4.3}$$

$$\boldsymbol{c}(\boldsymbol{y}, \boldsymbol{v_x}) \cdot \boldsymbol{c}(\boldsymbol{y}, \boldsymbol{v_x}) = 1\ , \tag{4.4}$$

$$\boldsymbol{a}(\boldsymbol{y}, \boldsymbol{v_x}) \equiv \boldsymbol{b}(\boldsymbol{y}) \times \boldsymbol{c}(\boldsymbol{y}, \boldsymbol{v_x})\ , \tag{4.5}$$

$$\boldsymbol{\rho}(\boldsymbol{y}, \boldsymbol{v_x}) \equiv \frac{\boldsymbol{b}(\boldsymbol{y}) \times [\boldsymbol{v_x}(\boldsymbol{y}) - \boldsymbol{D}(\boldsymbol{y})]}{B^l(\boldsymbol{y})}, \tag{4.6}$$

where $\boldsymbol{v}_{\boldsymbol{x}}(\boldsymbol{y})$ is the velocity at $\boldsymbol{x}$ parallel transported to $\boldsymbol{y}$, and all the fields can depend on t. In the flat spacetime considered here, $\boldsymbol{v}_{\boldsymbol{x}}(\boldsymbol{y}) = \boldsymbol{v}_{\boldsymbol{x}}$. The parallel transported velocity $\boldsymbol{v}_{\boldsymbol{x}}(\boldsymbol{y})$ has the following partition

$$\boldsymbol{v}_{\boldsymbol{x}}(\boldsymbol{y}) \equiv \boldsymbol{D}(\boldsymbol{y}) + u(\boldsymbol{y}, \boldsymbol{v}_{\boldsymbol{x}})\boldsymbol{b}(\boldsymbol{y}) + w(\boldsymbol{y}, \boldsymbol{v}_{\boldsymbol{x}})\boldsymbol{c}(\boldsymbol{y}, \boldsymbol{v}_{\boldsymbol{x}}) \ . \tag{4.7}$$

The zeroth order gyrocenter coordinate transformation

$$g_0 : z = (\boldsymbol{x}, \boldsymbol{v}, t) \mapsto \bar{Z} = (\bar{\boldsymbol{X}}, \bar{u}, \bar{w}, \bar{\theta}, t) \tag{4.8}$$

is defined by

$$\boldsymbol{x} \equiv \bar{\boldsymbol{X}} + \boldsymbol{\rho}(\bar{\boldsymbol{X}}, \boldsymbol{v}) \ , \tag{4.9}$$

$$\bar{u} \equiv u(\bar{\boldsymbol{X}}, \boldsymbol{v}) \ , \tag{4.10}$$

$$\bar{w} \equiv w(\bar{\boldsymbol{X}}, \boldsymbol{v}) \ , \tag{4.11}$$

$$\sin\bar{\theta} \equiv -\boldsymbol{c}(\bar{\boldsymbol{X}}) \cdot \boldsymbol{e}_1(\bar{\boldsymbol{X}}) \ , \tag{4.12}$$

$$t \equiv t \ , \tag{4.13}$$

where $\boldsymbol{e}_1(\bar{\boldsymbol{X}})$ is an arbitrary unit vector field in the perpendicular direction, and it can depend on t as well. Consequently,

$$\boldsymbol{v} = \boldsymbol{D}(\bar{\boldsymbol{X}}) + \bar{u}\boldsymbol{b}(\bar{\boldsymbol{X}}) + \bar{w}\boldsymbol{c}(\bar{\boldsymbol{X}}) \ . \tag{4.14}$$

Substituting Eqs. (4.9)–(4.14) into Eq. (3.5), and expanding terms using the ordering Eqs. (3.11)-(3.14), we have

$$\gamma = \bar{\gamma}_0 + \bar{\gamma}_1 + \ O(\varepsilon^2) \ , \tag{4.15}$$

$$\begin{aligned}\bar{\gamma}_0 = {} & \left[A^l(\bar{\boldsymbol{X}}, t) + \bar{u}\boldsymbol{b}(\bar{\boldsymbol{X}}, t) + \boldsymbol{D}(\bar{\boldsymbol{X}}, t)\right] \cdot d\bar{\boldsymbol{X}} + \frac{\bar{w}^2}{2B^l(\bar{\boldsymbol{X}}, t)} d\bar{\theta} \\ & - \left[\frac{\bar{u}^2 + \bar{w}^2 + D(\bar{\boldsymbol{X}}, t)^2}{2} + \phi^l(\bar{\boldsymbol{X}}, t)\right] dt \ ,\end{aligned} \tag{4.16}$$

$$\begin{aligned}\bar{\gamma}_1 = {} & \left[\frac{\bar{w}}{B^l}\boldsymbol{\nabla}\boldsymbol{a} \cdot \left(\bar{u}\boldsymbol{b} + \frac{\bar{w}\boldsymbol{c}}{2}\right) + \frac{1}{2}\boldsymbol{\rho} \cdot \boldsymbol{\nabla}\boldsymbol{B}^l \times \boldsymbol{\rho} - \frac{\bar{w}}{B^l}\boldsymbol{\nabla}\boldsymbol{D} \cdot \boldsymbol{a} + \boldsymbol{A}^s(\bar{\boldsymbol{X}} + \boldsymbol{\rho})\right] \cdot d\bar{\boldsymbol{X}} \\ & + \left[-\frac{\bar{w}^3}{2B^{l3}}\boldsymbol{a} \cdot \boldsymbol{\nabla}\boldsymbol{B}^l \cdot \boldsymbol{b} + \frac{\bar{w}^2}{B^l}\boldsymbol{A}^s(\bar{\boldsymbol{X}} + \boldsymbol{\rho}) \cdot \boldsymbol{c}\right] d\bar{\theta} + \left[\frac{1}{B^l}\boldsymbol{A}^s(\bar{\boldsymbol{X}} + \boldsymbol{\rho}) \cdot \boldsymbol{a}\right] d\bar{w} \\ & - \left[\phi^s(\bar{\boldsymbol{X}} + \boldsymbol{\rho}) + \boldsymbol{\rho} \cdot \frac{\partial \boldsymbol{D}}{\partial t} - \frac{1}{2}\boldsymbol{\rho} \cdot \boldsymbol{\nabla}\boldsymbol{E}^l \cdot \boldsymbol{\rho} - \left(\bar{u}\boldsymbol{b} + \frac{\bar{w}\boldsymbol{c}}{2}\right) \cdot \frac{\bar{w}}{B^l}\frac{\partial \boldsymbol{a}}{\partial t}\right] dt \ .\end{aligned} \tag{4.17}$$

Here, every field is evaluated at $\bar{Z}$ and can depend on t, and exact terms of the form $d\alpha$ for some $\alpha : P \to R$ have been discarded because their insignificance in Hamilton's equation (2.7). Computation needed in deriving the above equations is indeed involving. It can be easily verified that $\partial\bar{\gamma}_0/\partial\bar{\theta} = 0$, but $\partial\bar{\gamma}_1/\partial\bar{\theta} \neq 0$. As discussed before, we now introduce a coordinate perturbation to the zeroth order gyrocenter coordinates $\bar{Z}$,

$$Z = g_1(\bar{Z}) = \bar{Z} + G_1(\bar{Z}) \ , \tag{4.18}$$

such that $\partial\gamma_1/\partial\theta = 0$ in the first order gyrocenter coordinates $Z = (\boldsymbol{X}, u, w, \theta)$. Considering the fact that an arbitrary exact term of the form $d\alpha$ can be added to γ_1, we have

$$\gamma_1(Z) = \bar{\gamma}_1(Z) - i_{G_1(Z)}d\gamma_0(Z) + dS_1(Z) \ , \tag{4.19}$$

which, with $G_t = 0$, expands into

$$
\begin{aligned}
\gamma_1(Z) = & \Big[\boldsymbol{G}_{1\boldsymbol{X}} \times \boldsymbol{B}^l - G_{1u}\boldsymbol{b} + \boldsymbol{\nabla} S_1 + \frac{w}{B^l}\boldsymbol{\nabla a}\cdot\Big(u\boldsymbol{b} + \frac{w\boldsymbol{c}}{2}\Big) + \frac{1}{2}\boldsymbol{\rho}\cdot\boldsymbol{\nabla B}^l\times\boldsymbol{\rho} \\
& -\frac{w}{B^l}\boldsymbol{\nabla D}\cdot\boldsymbol{a} + \boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})\Big]\cdot d\boldsymbol{X} + \Big[\boldsymbol{G}_{1\boldsymbol{X}}\cdot\boldsymbol{b} + \frac{\partial S_1}{\partial u}\Big]du + \Big[\frac{w}{B^l}G_{1\theta} + \frac{\partial S_1}{\partial w} \\
& +\frac{1}{B^l}\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})\cdot\boldsymbol{a}\Big]dw + \Big[-\frac{w}{B^l}G_{1w} + \frac{\partial S_1}{\partial\theta} - \frac{w^3}{2B^{l3}}\boldsymbol{a}\cdot\boldsymbol{\nabla B}^l\cdot\boldsymbol{b} \\
& +\frac{w}{B^l}\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})\cdot\boldsymbol{c}\Big]d\theta + \Big[-\boldsymbol{E}^l\cdot\boldsymbol{G}_{1\boldsymbol{X}} + uG_{1u} + wG_{1w} + \frac{\partial S_1}{\partial t} - \phi^s(\boldsymbol{X}+\boldsymbol{\rho}) \\
& -\boldsymbol{\rho}\cdot\frac{\partial\boldsymbol{D}}{\partial t} + \frac{1}{2}\boldsymbol{\rho}\cdot\boldsymbol{\nabla E}^l\cdot\boldsymbol{\rho} + \Big(u\boldsymbol{b} + \frac{w\boldsymbol{c}}{2}\Big)\cdot\frac{w}{B^l}\frac{\partial\boldsymbol{a}}{\partial t}\Big]dt\ . \qquad (4.20)
\end{aligned}
$$

In Eq. (4.20), every field is evaluated at Z and can depend on t. Extensive calculations are needed to solve for G_1 and S_1 from the requirement that $\partial\gamma_1/\partial\theta = 0$. We listed the results without giving the details of the derivation,

$$
\begin{aligned}
\boldsymbol{G}_{1\boldsymbol{X}} &= -\frac{\partial S_1}{\partial u} + \frac{w^2}{2B^{l3}}\boldsymbol{aa}\cdot\boldsymbol{\nabla B}^l + \frac{wu}{B^{l2}}(\boldsymbol{\nabla a}\cdot\boldsymbol{b})\times\boldsymbol{b} - \frac{w}{B^{l2}}(\boldsymbol{\nabla D}\cdot\boldsymbol{a})\times\boldsymbol{b} \\
&\quad + \frac{\boldsymbol{\nabla} S_1 + \boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})}{B^l}\times\boldsymbol{b} \qquad (4.21) \\
G_{1u} &= \frac{w^2}{2B^{l2}}\boldsymbol{a}\cdot\boldsymbol{\nabla B}^l\cdot\boldsymbol{c} + \frac{wu}{B^l}\boldsymbol{b}\cdot\boldsymbol{\nabla a}\cdot\boldsymbol{b} - \frac{w}{B^l}\boldsymbol{b}\cdot\boldsymbol{\nabla D}\cdot\boldsymbol{a} \\
&\quad - \boldsymbol{b}\cdot[\boldsymbol{\nabla} S_1 + \boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})]\ , \qquad (4.22) \\
G_{1w} &= \frac{B^l}{w}\frac{\partial S_1}{\partial\theta} - \frac{w^2}{2B^{l2}}\boldsymbol{a}\cdot\boldsymbol{\nabla B}^l\cdot\boldsymbol{b} + \boldsymbol{c}\cdot\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})\ , \qquad (4.23) \\
G_{1\theta} &= -\frac{B^l}{w}\frac{\partial S_1}{\partial w} + \frac{1}{w}\boldsymbol{a}\cdot\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})\ . \qquad (4.24)
\end{aligned}
$$

The determining equation for S_1 is

$$
\begin{aligned}
&\frac{\partial S_1}{\partial t} + \Big(\frac{\boldsymbol{E}^l_\perp\times\boldsymbol{b}}{B^l} + u\boldsymbol{b}\Big)\cdot\boldsymbol{\nabla} S_1 + E^l_\parallel\frac{\partial S_1}{\partial u} + B^l\frac{\partial S_1}{\partial\theta} = \boldsymbol{E}^l_\perp\cdot\Big[\frac{w^2}{2B^{l3}}\widetilde{\boldsymbol{aa}}\cdot\boldsymbol{\nabla} B^l \\
&\quad +\frac{wu}{B^{l2}}(\boldsymbol{\nabla a}\cdot\boldsymbol{b})\times\boldsymbol{b} - \frac{w}{B^{l2}}(\boldsymbol{\nabla D}\cdot\boldsymbol{a})\times\boldsymbol{b}\Big] - \frac{w^2u}{2B^{l2}}\boldsymbol{\nabla B}^l : \widetilde{\boldsymbol{ca}} \\
&\quad -\frac{wu^2}{B^l}\boldsymbol{b}\cdot\boldsymbol{\nabla a}\cdot\boldsymbol{b} + \frac{wu}{B^l}\boldsymbol{b}\cdot\boldsymbol{\nabla D}\cdot\boldsymbol{a} + \frac{w^3}{2B^{l2}}\boldsymbol{a}\cdot\boldsymbol{\nabla B}^l\cdot\boldsymbol{b} \\
&\quad +\frac{w}{B^l}\boldsymbol{a}\cdot\frac{\partial\boldsymbol{D}}{\partial t} + \widetilde{\psi^s} - \frac{w^2}{2B^{l2}}\boldsymbol{\nabla E}^l : \widetilde{\boldsymbol{aa}} + \frac{uw}{B^l}\boldsymbol{a}\cdot\frac{\partial\boldsymbol{b}}{\partial t}\ . \qquad (4.25)
\end{aligned}
$$

The G_1 and S_1 in Eqs. (4.21)-(4.25) remove the θ-dependence in γ_1, i.e.,

$$\gamma(Z) = \gamma_0(Z) + \gamma_1(Z) \ , \tag{4.26}$$

$$\begin{aligned}\gamma_0 &= \left[A^l(\boldsymbol{X},t) + u\boldsymbol{b}(\boldsymbol{X},t) + \boldsymbol{D}(\boldsymbol{X},t)\right]\cdot d\boldsymbol{X} + \frac{w^2}{2B^l(\boldsymbol{X},t)}d\theta \\ &- \left[\frac{u^2+w^2+D(\boldsymbol{X},t)^2}{2} + \phi^l(\boldsymbol{X},t)\right] dt \ ,\end{aligned} \tag{4.27}$$

$$\gamma_1(Z) = -\frac{w^2}{2B^l}\boldsymbol{R}\cdot d\boldsymbol{X} - H_1 dt \ , \tag{4.28}$$

$$\begin{aligned}H_1 &= \boldsymbol{E}^l_\perp \cdot \frac{w^2}{2B^{l3}}\boldsymbol{\nabla}\boldsymbol{B}^l + \frac{w^2 u}{4B^l}\boldsymbol{b}\cdot\boldsymbol{\nabla}\times\boldsymbol{b} + \langle\psi^s\rangle \\ &- \frac{w^2}{4B^{l2}}\left(\nabla\cdot\boldsymbol{E}^l - \boldsymbol{b}\boldsymbol{b}:\boldsymbol{\nabla}\boldsymbol{E}^l\right) - \frac{w^2}{2B^l}R_0 \ ,\end{aligned} \tag{4.29}$$

$$\boldsymbol{R} \equiv \boldsymbol{\nabla}\boldsymbol{c}\cdot\boldsymbol{a} \ , \ R_0 \equiv -\frac{\partial\boldsymbol{c}}{\partial t}\cdot\boldsymbol{a} \ , \tag{4.30}$$

$$\psi^s \equiv \phi^s(\boldsymbol{X}+\boldsymbol{\rho}) - \frac{\boldsymbol{E}^l_\perp\times\boldsymbol{b}}{B^l}\cdot\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho}) - w\boldsymbol{c}\cdot\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho}) \ , \tag{4.31}$$

$$\langle\alpha\rangle \equiv \frac{1}{2\pi}\int_0^{2\pi}\alpha d\theta \ , \ \widetilde{\alpha} \equiv \alpha - \langle\alpha\rangle \ . \tag{4.32}$$

Even though Eqs. (4.21)-(4.32) are displayed without derivation, it may be necessary to demonstrate the basic procedures of the derivation. For this purpose, we will outline here the derivation of the $\boldsymbol{X}_\perp$ and w components of G_1 in $\gamma(z)$. Let

$$\begin{aligned}\boldsymbol{\gamma}_{1\boldsymbol{X}}(Z) = \Bigg[&\boldsymbol{G}_{1\boldsymbol{X}}\times\boldsymbol{B}^l - G_{1u}\boldsymbol{b} + \boldsymbol{\nabla}S_1 + \frac{w}{B^l}\boldsymbol{\nabla}\boldsymbol{a}\cdot\left(u\boldsymbol{b}+\frac{w\boldsymbol{c}}{2}\right) + \frac{1}{2}\boldsymbol{\rho}\cdot\boldsymbol{\nabla}\boldsymbol{B}^l\times\boldsymbol{\rho} \\ &-\frac{w}{B^l}\boldsymbol{\nabla}\boldsymbol{D}\cdot\boldsymbol{a} + \boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho})\Bigg] \ .\end{aligned} \tag{4.33}$$

We look at the following partition of $\gamma_{1\boldsymbol{X}}(Z)\cdot d\boldsymbol{X}$,

$$\gamma_{1\boldsymbol{X}}(Z)\cdot d\boldsymbol{X} = \boldsymbol{b}\cdot\gamma_{1\boldsymbol{X}}(Z)\boldsymbol{b}\cdot d\boldsymbol{X} + \gamma_{1\boldsymbol{X}}(Z)\times\boldsymbol{b}\times\boldsymbol{b}\cdot d\boldsymbol{X} \ . \tag{4.34}$$

For the first term in the right hand side of Eq. (4.34)

$$\begin{aligned}\boldsymbol{b}\cdot\gamma_{1\boldsymbol{X}}(Z) &= -G_{1u} + \boldsymbol{b}\cdot\boldsymbol{\nabla}S_1 + \boldsymbol{b}\cdot\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho}) - \frac{w}{B^l}\boldsymbol{b}\cdot\boldsymbol{\nabla}\boldsymbol{D}\cdot\boldsymbol{a} \\ &+ \frac{1}{2}\left(\boldsymbol{\rho}\cdot\boldsymbol{\nabla}\boldsymbol{B}^l\times\boldsymbol{\rho}\right)\cdot\boldsymbol{b} + \frac{w}{B^l}\boldsymbol{\nabla}\boldsymbol{a}\cdot\left(u\boldsymbol{b}+\frac{w\boldsymbol{c}}{2}\right)\cdot\boldsymbol{b} \ .\end{aligned} \tag{4.35}$$

Choosing G_{1u} to be the form displayed in Eq. (4.22), we are left with the following expression

$$\boldsymbol{b}\cdot\gamma_{1\boldsymbol{X}}(Z)\boldsymbol{b}\cdot d\boldsymbol{X} = \left(-\frac{w^2}{2B^l}\boldsymbol{R}\cdot\boldsymbol{b}\right)\boldsymbol{b}\cdot d\boldsymbol{X} \ . \tag{4.36}$$

Similarly, for the second term in the right hand side of Eq. (4.34)

$$\begin{aligned}\gamma_{1\boldsymbol{X}}(Z)\times\boldsymbol{b} &= -\boldsymbol{G}_{1\boldsymbol{X}\perp}B^l + \boldsymbol{b}\times\boldsymbol{\nabla}S_1 - \boldsymbol{b}\times\boldsymbol{A}^s(\boldsymbol{X}+\boldsymbol{\rho}) + \frac{w}{B^l}\boldsymbol{b}\times\boldsymbol{\nabla}\boldsymbol{D}\cdot\boldsymbol{a} \\ &+ \frac{1}{2}\boldsymbol{\rho}\left(\boldsymbol{\rho}\cdot\boldsymbol{\nabla}\boldsymbol{B}^l\right) + \frac{w}{B^l}\boldsymbol{\nabla}\boldsymbol{a}\cdot\left(u\boldsymbol{b}+\frac{w\boldsymbol{c}}{2}\right)\times\boldsymbol{b} \ .\end{aligned} \tag{4.37}$$

Choosing $\boldsymbol{G}_{1\boldsymbol{X}\perp}$ to be the perpendicular component part of the result displayed in Eq. (4.21), we are left with

$$\gamma_{1\boldsymbol{X}}(Z) \times \boldsymbol{b} \times \boldsymbol{b} \cdot d\boldsymbol{X} = \left(-\frac{w^2}{2B^l}\boldsymbol{R}_\perp\right) \cdot d\boldsymbol{X} \,. \tag{4.38}$$

Combining Eqs. (4.36) and (4.38), we obtain the first term on the right hand side of Eq. (4.28). The rest of the derivation for Eqs. (4.21)-(4.32) can be carried out in similar procedures.

A particle's worldline is given by a vector field τ on phase space P which satisfies

$$i_\tau d\gamma = 0 \ . \tag{4.39}$$

The conventional gyrocenter motion equation can be obtained through

$$\frac{d\boldsymbol{X}}{dt} = \frac{\tau_{\boldsymbol{X}}}{\tau_t} \ , \ \frac{du}{dt} = \frac{\tau_u}{\tau_t} \ , \ \frac{dw}{dt} = \frac{\tau_w}{\tau_t} \ , \ \frac{d\theta}{dt} = \frac{\tau_\theta}{\tau_t} \ . \tag{4.40}$$

After some calculation, we obtain the following explicit expressions up to order ε for gyrocenter dynamics,

$$\frac{d\boldsymbol{X}}{dt} = \frac{\boldsymbol{B}^\dagger}{\boldsymbol{b}\cdot\boldsymbol{B}^\dagger}(u + \frac{\mu}{2}\boldsymbol{b}\cdot\boldsymbol{\nabla}\times\boldsymbol{b}) - \frac{\boldsymbol{b}\times\boldsymbol{E}^\dagger}{\boldsymbol{b}\cdot\boldsymbol{B}^\dagger} \ , \tag{4.41}$$

$$\frac{du}{dt} = \frac{\boldsymbol{B}^\dagger\cdot\boldsymbol{E}^\dagger}{\boldsymbol{B}^\dagger\cdot\boldsymbol{b}} \ , \tag{4.42}$$

$$\begin{aligned}\frac{d\theta}{dt} = B^l + \boldsymbol{R}\cdot\frac{d\boldsymbol{X}}{dt} - R_0 + \Bigg\{&\frac{\boldsymbol{E}^l\cdot\nabla B^l}{B^{l2}} + \frac{u}{2}b\cdot\nabla\times b \\ &+\frac{\partial}{\partial\mu}\langle\psi^s\rangle - \frac{1}{2B^l}\left[\nabla\cdot\boldsymbol{E}^l - \boldsymbol{bb}:\boldsymbol{\nabla}\boldsymbol{E}^l\right]\Bigg\} \ ,\end{aligned} \tag{4.43}$$

$$\frac{d\mu}{dt} = 0 \ , \ \mu \equiv \frac{w^2}{2B^l} \ , \tag{4.44}$$

$$\boldsymbol{B}^\dagger \equiv \boldsymbol{\nabla}\times\left[\boldsymbol{A}^l(\boldsymbol{X},t) + u\boldsymbol{b}(\boldsymbol{X},t) + \boldsymbol{D}(\boldsymbol{X},t)\right] \ , \tag{4.45}$$

$$\boldsymbol{E}^\dagger \equiv \boldsymbol{E}^l - \boldsymbol{\nabla}\left[\mu B^l + \frac{D^2}{2} + \langle\psi^s\rangle\right] - u\frac{\partial\boldsymbol{b}}{\partial t} - \frac{\partial\boldsymbol{D}}{\partial t} \ . \tag{4.46}$$

The modified fields $\boldsymbol{B}^\dagger$ and $\boldsymbol{E}^\dagger$ can be viewed as those generated by a modified potential $A^\dagger = (\phi^\dagger, \boldsymbol{A}^\dagger)$,

$$\phi^\dagger(\boldsymbol{X},t) \equiv \phi^l(\boldsymbol{X},t) + \mu B^l(\boldsymbol{X},t) + \frac{D(\boldsymbol{X},t)^2}{2} + \langle\psi^s(\boldsymbol{X},t)\rangle \ , \tag{4.47}$$

$$\boldsymbol{A}^\dagger(\boldsymbol{X},t) \equiv \boldsymbol{A}^l(\boldsymbol{X},t) + u\boldsymbol{b}(\boldsymbol{X},t) + \boldsymbol{D}(\boldsymbol{X},t) \ , \tag{4.48}$$

$$\boldsymbol{B}^\dagger = \boldsymbol{\nabla}\times\boldsymbol{A}^\dagger, \ \boldsymbol{E}^\dagger = -\boldsymbol{\nabla}\phi^\dagger - \frac{\partial\boldsymbol{A}^\dagger}{\partial t} \ . \tag{4.49}$$

In Eq. (4.44), the conserved magnetic momentum μ is constructed asymptotically when the spacetime inhomogeneities are weak. Recently, the concept of adiabatic invariant has been extended to cases with strong spatial inhomogeneities for magnetic field [20, 61].

5 Gyrocenter gauge and gyro-gauge

An important fact is that the requirement $\partial\gamma_1/\partial\theta = 0$ does not uniquely determine the coordinate perturbation G and the gauge function S, and therefore the first order gyrocenter coordinates. There are freedoms in defining the zeroth order gyrocenter coordinates as well. For example, in Ref. [38], the following definition of the zeroth order gyrocenter coordinates are used

$$\boldsymbol{x} \equiv \bar{\boldsymbol{X}} + \boldsymbol{\rho}(\bar{\boldsymbol{X}}, \boldsymbol{v}) \ , \tag{5.1}$$

$$\bar{u} \equiv u(\boldsymbol{x}, \boldsymbol{v}) \ , \tag{5.2}$$

$$\bar{w} \equiv w(\boldsymbol{x}, \boldsymbol{v}) \ , \tag{5.3}$$

$$\sin\bar{\theta} \equiv -\boldsymbol{c}(\boldsymbol{x}) \cdot \boldsymbol{e}_1(\boldsymbol{x}) \ , \tag{5.4}$$

$$t \equiv t \ . \tag{5.5}$$

This choice results in more terms in the expression for $\bar{\gamma}_1$. We will call the freedoms in selecting the gyrocenter coordinates gyro-center gauges. In Eq. (4.30), $\boldsymbol{R}$ and R_0 are θ-independent, even though $\boldsymbol{a}$ and $\boldsymbol{c}$ are θ-dependent. Let $R = (R_0, \boldsymbol{R})$, $X = (t, \boldsymbol{X})$, and $\nabla = (-\partial/\partial t, \boldsymbol{\nabla})$. The γ in Eq. (4.26) is invariant under the following group of transformation

$$R \longrightarrow R' + \nabla\delta(X) \ , \tag{5.6}$$

$$\theta \longrightarrow \theta' + \delta(X) \ . \tag{5.7}$$

Apparently, this is a gauge group associated how the gyrophase θ is measured or how Kruskal ring mates are labeled. Naturally, an appropriate name for this gauge would be gyro-gauge. The $\boldsymbol{R}$ components of gyro-gauge group were first rigorously derived in Ref. [39]. Without R, γ will not be invariant under the gyro-gauge group transformation.

6 Pullback transformation

Even though the γ in Eq. (4.26) is gyro-gauge invariant, it does not need to be. Different gyro-center gauges can be chosen such that γ is not gyro-gauge invariant. The gyrocenter coordinate system constructed is just a useful coordinate system for physics, but not the physics itself. It can depend on the gauges (freedoms) we choose, as long as it is useful. Gyrocenter coordinate system and the gyrokinetic equation are not the total of physics under investigation. What is gauge invariant is the system of gyrokinetic equation and the gyrokinetic Maxwell equations. The key element which makes this gyrokinetic system gauge invariant is the pullback transformation associated with the gyrocenter coordinate system. Without this vital element, a number of prominent physics features, such as the presence of the compressional Alfvén wave and a proper description of the gyrokinetic equilibrium, cannot be readily recovered.

Kinetic theory deals with particle distribution function f, which is a function defined on the phase space P, $f : P \to R$. As discussed in Sec. 2, the familiar density and momentum velocity integrals needed for Maxwell's equations are the fiber integrals $j(x) = \int_{\pi^{-1}(x)} f\omega_x$ at x, which returns a 3-form flux. A coordinate system (x, v) for P is fibered if x are the coordinates for the base, i.e., the spacetime M. In gyrokinetic theory, however, the useful gyrocenter coordinate system is non-fibered because $\boldsymbol{X}$ are not coordinates for spacetime. The gyrocenter transformation $g : z \longmapsto Z$ is a non-fibered coordinate transformation. No matter which coordinate system is used, non-fibered or fibered, the moment integrals are still defined on the fiber $\pi^{-1}(x)$ at each x, and $j(x)$ should be invariant under general non-fibered coordinate transformations. For the new non-fibered coordinate system Z to be useful, it is necessary to know the construction of $j(x)$ in it. To be specific, the current scenario is that the distribution function f is known in the transformed non-fibered coordinate system Z as $F(Z)$. Given $F(Z)$, we need to pull back the distribution function $F(Z)$ into $f(z)$,

$$j(x) = \int_{\pi^{-1}(x)} g^* \left[F(Z)\right] \omega_x \,, \tag{6.1}$$

$$g^* \left[F(Z)\right] = F\left(g(z)\right) = f(z) \,. \tag{6.2}$$

Considering the asymptotic nature of the construction of the gyrocenter transformation g,

$$g = g_1 g_0 \;,\; g_0 : z \longmapsto \bar{Z} \;,\; g_1 : \bar{Z} \longmapsto Z \;, \tag{6.3}$$

we write

$$\begin{aligned} f(z) &= g^* F(Z) = g_0^* \; g_1^* \; F(Z) = g_0^* F\left[g_1(\bar{Z})\right] \\ &= g_0^* \left[F(\bar{Z}) + G \cdot \nabla F(\bar{Z}) + O(\varepsilon^2)\right] \\ &= F\left[g_0(z)\right] + G\left[g_0(z)\right] \cdot \nabla F\left[g_0(z)\right] + O(\varepsilon^2) \;. \end{aligned} \tag{6.4}$$

7 General gyrokinetic Vlasov-Maxwell equations

After constructing the gyrocenter coordinates and the corresponding pullback transformation, we are ready to cast the coordinate independent (geometric) Vlasov-Maxwell equations (2.17) in the gyrocenter coordinates to obtain the general gyrokinetic Vlasov-Maxwell equations. The gyrokinetic Vlasov equation is simply the Vlasov equation $df(\tau) = 0$ in the gyrocenter coordinates Z, which is explicitly

$$\frac{dZ_j}{dt}\frac{\partial F}{\partial Z_j} = 0 \;,\; (0 \le j \le 6) \;. \tag{7.1}$$

Because

$$\frac{\partial}{\partial \theta}\left(\frac{dZ}{dt}\right) = 0 \;, \tag{7.2}$$

the gyrokinetic equation can be easily split into two parts

$$F = \langle F \rangle + \widetilde{F} \;, \tag{7.3}$$

$$\frac{\partial \langle F \rangle}{\partial t} + \frac{d\boldsymbol{X}}{dt} \cdot \boldsymbol{\nabla}_{\boldsymbol{X}} \langle F \rangle + \frac{du}{dt}\frac{\partial \langle F \rangle}{\partial u} = 0 \;, \tag{7.4}$$

$$\frac{\partial \widetilde{F}}{\partial t} + \frac{d\boldsymbol{X}}{dt} \cdot \boldsymbol{\nabla}_{\boldsymbol{X}} \widetilde{F} + \frac{du}{dt}\frac{\partial \widetilde{F}}{\partial u} + \frac{d\theta}{dt}\frac{\partial \widetilde{F}}{\partial \theta} = 0 \;, \tag{7.5}$$

where $d\boldsymbol{X}/dt$, du/dt, and $d\theta/dt$ are given by Eqs. (4.41)-(4.43). It is necessary to complete the kinetic equations for F with Maxwell's equation. With the pullback transformation (6.4), the gyrokinetic Maxwell's equation can be written as

$$d * dA = 4\pi \int_{\pi^{-1}(x)} \left[(F \circ g_0) + (G \circ g_0) \cdot \nabla \left(F \circ g_0\right)\right] \omega_x \;. \tag{7.6}$$

We emphasize that Eq. (7.6) is not a new equation which contains different physics than the original Maxwell's equation with moment integral. The more appropriate name for this equation should be "Maxwell's equation with moment integral (fiber integral) in the gyrocenter coordinates".

The gyrophase dependent $\widetilde{F}$ can be decoupled from the system. Letting $\widetilde{F} = 0$, Eqs. (7.4) and (7.6) form a close system for $\langle F \rangle$ and A. We note that $\widetilde{F} = 0$ does not imply that $\widetilde{f} = 0$. f becomes gyrophase dependent through the pullback transformation (6.4) and G. Indeed, $\widetilde{f}$ and G contain significant amount of important physics, which will be demonstrated in the next two sections.

The spirit of the general gyrokinetic theory is to decouple the gyro-phase dynamics from the rest of particle dynamics by finding the gyro-symmetry, which is fundamentally different from the conventional gyrokinetic concept of "averaging out" the "fast gyro-motion". This objective is accomplished by asymptotically constructing a good coordinate system, which is of course a nontrivial task [16, 17, 40] (see Fig. 5). Indeed, it is almost impossible without the Lie coordinate perturbation method enabled by the geometric nature of the phase space dynamics. We developed the gyrokinetic Vlasov-Maxwell equations not as a new set of equations, but rather as the Vlasov-Maxwell equations in the gyrocenter coordinates. Since the general gyrokinetic system developed is geometrically the same as the Vlasov-Maxwell equations, all the coordinate independent properties of the Vlasov-Maxwell equations, such as energy conservation, momentum conservation and Liouville volume conservation, are automatically carried over to the general gyrokinetic system.

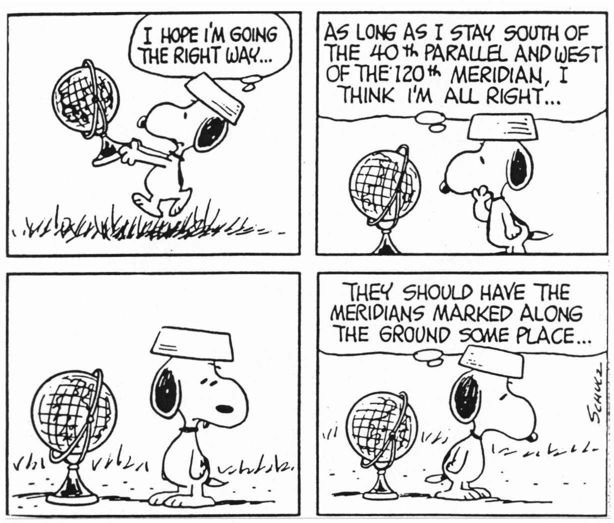

Figure 5 Quest of useful coordinates [40]. (Peanuts by Charles Schulz. Reprint permitted by UFS, Inc.)

8 Application: Spitzer paradox

Now, we turn to the applications of the gyrokinetic theory developed. The first application is related to how to describe plasma equilibrium using the gyrokinetic theory. Spitzer first noticed the obvious differences between the currents described by the fluid equations and the guiding center motion [53, 54]. There are two aspects of these obvious differences in an equilibrium plasma without flow and electric field. First, the perpendicular current given by the fluid model is the diamagnetic current $\boldsymbol{b} \times \nabla p / B$, which is not in the guiding center drift motion. On the other hand, the curvature drift and the gradient drift for the guiding center motion are not found in the fluid results. This puzzle, first posed and discussed by Spitzer, is what we call the Spitzer paradox. To resolve it, we must explain, qualitatively as well as quantitatively, how the diamagnetic current is microscopically generated, and what happens to the macroscopic counterparts of the curvature drift and the gradient drift. Here, we will only discuss the first part of the puzzle — how the diamagnetic current is generated microscopically. A detailed study of the puzzle and other relevant topics can be found in Ref. [50].

Spitzer gave the well-known physical picture, which is illustrated in Fig. 6. The basic setup is an equilibrium plasma with a constant magnetic field and a pressure (density) gradient in the perpendicular direction. From the fluid equation $\boldsymbol{b} \times B = \nabla p$, we know that the perpendicular current is $\boldsymbol{b} \times \nabla p/B$. However, if we look at the microscopic picture, for each gyrocenter, the drift motion does not produce any gyrocenter current or flow because the magnetic field is constant in spacetime. Spitzer pointed out that there are more particles on the left than on the right; thus macroscopically gyromotion generates current and flow at each spatial location. The key to resolve the paradox is the realization that the flow

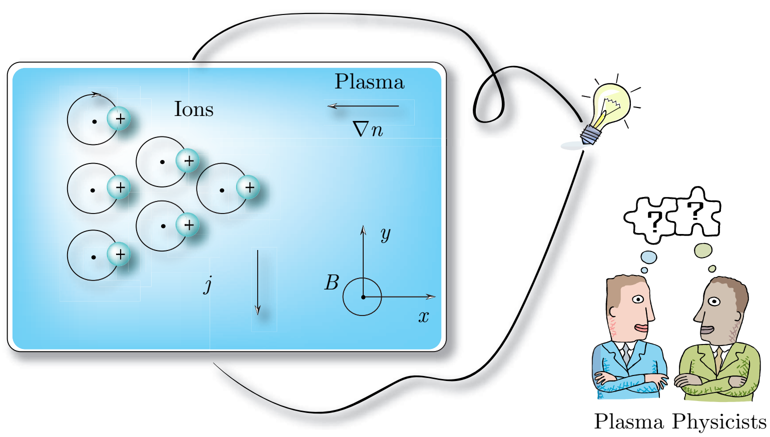

Figure 6 Spitzer Paradox. In memory of Lyman Spitzer Jr. (1914-1997) [52].

of particles is fundamentally different from that of gyrocenters. The difference is rigorously described by the pullback transformation discussed in Sec. 6. Because $\boldsymbol{B}$ is constant, $G_1 = 0$. Using Eqs. (6.1) and (6.4), the $dx \wedge dt \wedge dz$ component of j is

$$\begin{aligned} j_y &= \int v_y g_0^* \left[F(\bar{Z})\right] d\boldsymbol{v}^3 = \int v_y F(\boldsymbol{x} + \boldsymbol{\rho}, \boldsymbol{v}) d\boldsymbol{v}^3 \\ &= \int v_y \left[F(\boldsymbol{x}) + \boldsymbol{\rho} \cdot \boldsymbol{\nabla} F(\boldsymbol{x}) + O(\varepsilon^2)\right] d\boldsymbol{v}^3 \\ &= \int v_y^2 \frac{\partial F}{\partial x} \frac{1}{B} d\boldsymbol{v}^3 = \left(b \times \frac{\nabla p}{B}\right)_y . \end{aligned} \tag{8.1}$$

The physics captured in Eq. (8.1) is clear. Even though the gyrocenter flow is zero, particle flow can be generated by the pullback transformation g_0^* associated with the zeroth order gyrocenter transformation g_0. The Spitzer paradox highlights the "seeming conflict" between the theory of gyromotion and the fluid equations, two most fundamental concepts in plasma physics, and emphasizes the important physics content in the pullback transformation.

9 Application: Bernstein wave and compressional Alfvén wave

As examples of applications of the gyrokinetic theory developed to plasma waves, we derive the dispersion relations for the Bernstein wave and the compressional Alfvén wave in this section. A detailed derivation of the complete dispersion relation for plasma waves with arbitrary wavelength and frequency using the gyrokinetic theory can be found in Ref. [46].

For the Bernstein wave, we consider an electrostatic wave propagating perpendicularly in a homogeneous magnetized plasma. Let $\boldsymbol{B}^l = B\boldsymbol{e}_z = \Omega\boldsymbol{e}_z$, $\boldsymbol{E}^l = 0$, $\boldsymbol{A}^s = 0$, $\boldsymbol{k} = k\boldsymbol{e}_x$, and

$$\phi^s \sim \phi \exp(ikx - i\omega t) \ . \tag{9.1}$$

Linearizing the gyrokinetic equation for $F = F_0 + F_1$, we have

$$\frac{dF_1}{dt} = \frac{\partial F_1}{\partial t} + u\boldsymbol{b}\cdot\boldsymbol{\nabla}F = -\boldsymbol{b}\cdot\boldsymbol{\nabla}\langle\phi\rangle\frac{\partial F_0}{\partial u} \ . \tag{9.2}$$

Assuming the equilibrium distribution function F_0 to be Maxwellian

$$F_0 = \frac{n_0}{(2\pi T/m)^{3/2}}\exp\left(\frac{-v^2}{2T/m}\right) \ , \tag{9.3}$$

the solution for the linear gyrokinetic equation is degenerate because $k_\| \equiv \boldsymbol{b}\cdot\boldsymbol{k} = 0$,

$$F_1 = -\frac{1}{T/m}F_0\frac{-k_\| u}{\omega - k_\| u}\langle\phi\rangle = 0 \ . \tag{9.4}$$

The only physics content is found in the pull-back of the perturbed density, which requires expressing the gauge function S_1 in terms of the perturbed fields. The equation for S_1 is

$$\Omega\frac{\partial S_1}{\partial\theta} + \frac{\partial S_1}{\partial t} = \widetilde{\phi}(\boldsymbol{X}+\boldsymbol{\rho}) = [\ \mathrm{e}^{\boldsymbol{\rho}\cdot\nabla} \ - J_0(\frac{\boldsymbol{\rho}\cdot\boldsymbol{\nabla}}{i})]\phi(\boldsymbol{X}) \ . \tag{9.5}$$

Using the identity

$$\exp(\lambda\cos\theta) = \sum_{n=-\infty}^{\infty} I_n(\lambda)\exp(in\theta) \ , \tag{9.6}$$

we can easily solve Eq. (9.5) for S_1,

$$S_1 = \frac{1}{\Omega i\bar{\omega}}J_0\phi + \frac{1}{\Omega}\sum_{n=-\infty}^{\infty}\frac{I_n(i\rho k)}{i(n-\bar{\omega})}\ \mathrm{e}^{in\theta}\ \phi \ . \tag{9.7}$$

where $\bar{\omega} = \omega/\Omega$. Since $F_1 = 0$, the density response (i.e., the $dx\wedge dt\wedge dz$ component of the 3-form flux in spacetime) comes only from S_1 in the pull-back transformation.

$$\begin{aligned} n_1 &= \int g_0^*\left[F_1(\bar{Z}) + G_1\cdot\boldsymbol{\nabla}F_0(\bar{Z}) + O(\varepsilon^2)\right]d\boldsymbol{v}^3 \\ &= \int \mathrm{e}^{-\boldsymbol{\rho}\cdot\nabla}G_1\cdot\boldsymbol{\nabla}F_0(z)d\boldsymbol{v}^3 + O(\varepsilon^2) \\ &= \int \mathrm{e}^{-\boldsymbol{\rho}\cdot\nabla}\frac{\Omega}{w}\frac{\partial S_1}{\partial\theta}\frac{\partial F_0}{\partial w}d\boldsymbol{v}^3 + O(\varepsilon^2) \\ &= \int -\mathrm{e}^{\boldsymbol{\rho}\cdot\nabla}\frac{F_0}{T/m}\sum_{n=-\infty}^{\infty}\frac{nI_n(i\rho k)}{(n-\bar{\omega})}\ \mathrm{e}^{in\theta}\ \phi(\boldsymbol{x})\ d^3\boldsymbol{v} + O(\varepsilon^2) \ . \end{aligned} \tag{9.8}$$

Using the facts that

$$\int_0^{2\pi} \mathrm{e}^{i(m+n)\xi} \, d\xi = \delta_{m,-n} 2\pi \;, \tag{9.9}$$

we have

$$n_1 = \frac{2\pi}{(2\pi T)^{3/2}} \int \frac{-n_0\phi}{T/m} \exp\left(-\frac{v_\parallel^2 + v_\perp^2}{2T/m}\right) \sum_{n=-\infty}^{\infty} \frac{n I_{-n}(-i\rho k) I_n(i\rho k)}{(n-\bar{\omega})} \, v_\perp dv_\parallel dv_\perp \;. \tag{9.10}$$

Carrying out the algebra with the help of some identities related to the Bessel functions, we obtain

$$n_1 = n_0 \frac{\phi}{T/m} \sum_{n=1}^{\infty} \frac{2n^2}{(\frac{\omega}{\Omega})^2 - n^2} \exp\left(-\frac{k^2 T}{\Omega^2 m}\right) I_n\left(\frac{k^2 T}{\Omega^2 m}\right) \;. \tag{9.11}$$

Finally, the Poisson equation (in unnormalized units)

$$-\nabla^2 \phi = \sum_{spec} 4\pi e n_1 \tag{9.12}$$

gives the dispersion relation of the Bernstein wave

$$1 = \sum_{spec} \frac{4\pi n_0 e^2}{T k^2} \sum_{n=1}^{\infty} \frac{2n^2}{(\frac{\omega}{\Omega})^2 - n^2} \exp\left(-\frac{k^2 T}{\Omega^2 m}\right) I_n\left(\frac{k^2 T}{\Omega^2 m}\right) \;. \tag{9.13}$$

For low frequency and long wavelength modes, the leading order n_1 in Eq. (9.11) is

$$n_1 = -n_0 \frac{k^2 \phi}{\Omega^2} \;.$$

Historically, this term has been referred as "the polarization drift term in the Poisson equation". It has played an important role in the development of gyrokinetic particle simulation methods [5,11,13,14,19,21,23,33,34,42,57]. However, its derivation were almost always heuristic. Using the general gyrokinetic theory developed here, this term is rigorously recovered as a special case of the general pullback transformation. In an inhomogeneous equilibrium, it is generalized into [47]

$$n_1 = \boldsymbol{\nabla} \cdot \left(\frac{n_0}{\Omega^2} \boldsymbol{\nabla} \phi \right) \;. \tag{9.14}$$

Let's rewrite the Poisson equation for the current case as,

$$\boldsymbol{\nabla} \cdot (\varepsilon \boldsymbol{E}_\perp) = 0 \;, \tag{9.15}$$

$$\varepsilon = 1 + \sum_{spec} \frac{4\pi n_0 e^2}{\Omega^2 m} \;. \tag{9.16}$$

Here ε can be viewed as the dielectric constant of the plasma in the perpendicular direction. This point of view can be justified by the following alternative derivation of Eq. (9.16). Because

$$\boldsymbol{x} = \boldsymbol{X} + \boldsymbol{\rho} + \boldsymbol{G}_{1\boldsymbol{X}} \;, \tag{9.17}$$

if we treat gyrocenters $\boldsymbol{X}$ as individual particles, then there is a charge separation due to the $\boldsymbol{G}_{1\boldsymbol{X}}$ displacement (see Fig. 7). The induced electric polarization p

(dipole moment) for each gyrocenter is [29]

$$p = \frac{e}{2\pi} \int \boldsymbol{G}_{1\boldsymbol{X}} \, d\theta \ , \tag{9.18}$$

$$\boldsymbol{G}_{1\boldsymbol{X}} = \frac{1}{B} \times \boldsymbol{\nabla} S_1 \ . \tag{9.19}$$

Calculation shows

$$p = \frac{-e}{\Omega^2} \boldsymbol{\nabla}_{\perp} \phi \ . \tag{9.20}$$

Therefore, the electric susceptibility for each species is

$$\chi = \frac{n_0 e^2}{\Omega^2 m} \ , \tag{9.21}$$

which is consistent with Eq. (9.16).

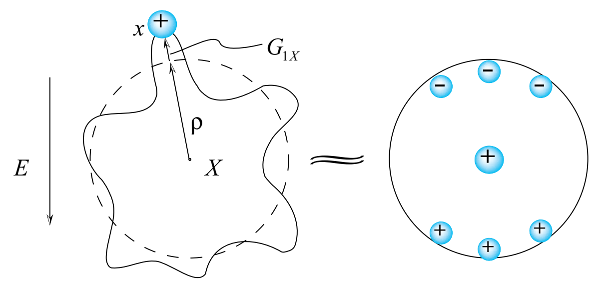

Figure 7 The $\boldsymbol{G}_{1\boldsymbol{X}}$ displacement induces an electric polarization p (dipole moment) for each gyrocenter $\boldsymbol{X}$.

We observe that the second term in Eq. (9.16), or the dielectric constant due to the polarizaiton drift, agrees with the well-known classical result. In this case, it shows that the magnetized plasma described by the gyrokinetic theory physically has the same linear response as that described by the classical theory. We take the viewpoint that the general gyrokinetic theory should not contain different physics that are not described by the Vlasov-Maxwell equations in the regular laboratory phase space coordinates. However, there are gyrocenter coordinates where the Vlasov-Maxwell equations have different forms more suitable for theoretical analysis and numerical simulations. The challenge of the general gyrokinetic theory is to construct such a useful coordinate system and associated pull-back transformation without losing or adding any physics content to the Vlasov-Maxwell equations. Indeed, the dielectric constant in Eq. (9.16) is just a limiting case of the most general classical dielectric constant tensor for magnetized plasmas [55], which has been recovered exactly from the general gyrokinetic theory with the most general pullback transformation [46]. Alternative viewpoint on the dielectric response in gyrokinetic plasma, which has important implications for numerical methods for gyrokinetic systems, was discussed by Krommes [30].

To derive the dispersion relation for the compressional Alfvén wave, we consider an electromagnetic wave propagating perpendicularly with $\boldsymbol{B}^l = B\boldsymbol{e}_z = \Omega \boldsymbol{e}_z$, $\boldsymbol{E}^l = 0$, $\phi^s = 0$, $\boldsymbol{k} = k\boldsymbol{e}_y$, and

$$\boldsymbol{A}^s = A_x \exp\left(iky - i\omega t\right) \boldsymbol{e}_x \ . \tag{9.22}$$

As in the case of the Bernstein wave, $F_1 = 0$ since $k_\| = 0$. Ignoring the finite gyro-radius effect, the equation for S_1 is

$$\Omega \frac{\partial S_1}{\partial \theta} + \frac{\partial S_1}{\partial t} = w \sin\theta A_x \ . \tag{9.23}$$

The solution for S_1 is

$$S_1 = \frac{w}{\Omega} A_x \frac{\cos(\theta) + i\bar{\omega}\sin(\theta)}{(\bar{\omega}+1)(\bar{\omega}-1)} \ . \tag{9.24}$$

The perpendicular components of j are

$$\begin{aligned} j_\perp &= \int -\frac{\partial F_0}{\partial w} G_{1w} \left(v_x dt \wedge dy \wedge dz + v_y dx \wedge dt \wedge dz \right) \\ &= n_0 A_x \left[\frac{-\bar{\omega}^2}{-\bar{\omega}^2+1} dt \wedge dy \wedge dz + \frac{i\bar{\omega}}{(\bar{\omega}+1)(\bar{\omega}-1)} dx \wedge dt \wedge dz \right] \ . \end{aligned} \tag{9.25}$$

The $dt \wedge dy \wedge dz$ component is the polarization drift flow, and the $dx \wedge dt \wedge dz$ is the $\boldsymbol{E} \times \boldsymbol{B}$ flow. When $\bar{\omega} \ll 1$, the $\boldsymbol{E} \times \boldsymbol{B}$ flow from different species cancels out in neutral plasma, and

$$j_\perp = -\sum_{spec} \frac{e n_0}{B^2} \omega^2 A_x dt \wedge dy \wedge dz \ . \tag{9.26}$$

Invoking the Maxwell's equation $d * dA = 4\pi j$, we obtain the dispersion relation for the compressional Alfvén wave

$$\omega^2 = k^2 v_A^2 \ , \tag{9.27}$$

where $v_A^2 \equiv B^2/4\pi n_0$ is the Alfvén velocity in unnormalized units.

10 Further development

Physics is geometry. The geometric point of view for the gyrokinetic theory has been proven to be efficient and productive. The geometry of the gyrokinetic theory is rich. Many interesting topics, such as gyrocenter dynamics as a Hamiltonian system in a cotangent bundle [44], gyrocenter dynamics as an anholonomy of a connection [44], collision operator for the gyrokinetic system, and gyrokinetic concept for magnetized plasmas with strong spacetime inhomogeneities [20, 61] are currently being investigated.

Acknowledgment

I sincerely thank Prof. William M. Tang and Prof. Ronald C. Davidson for their continuous support. In the last ten years, I have learned a lot on this subject from my teachers and colleagues, for which I am grateful. Especially, I would like to thank J. R. Cary, P. J. Channell, L. Chen, Y. Chen, B. I. Cohen, R. H. Cohen, R. C. Davidson, A. M. Dimits, W. D. Dorland, N. J. Fisch, A. Friedman, G. W. Hammet, T. S. Hahm, J. A. Krommes, M. Kruskal, R. M. Kulsrud, G-L. Kuang, W. W. Lee, J-G. Li, Z. Lin, W. M. Nevins, G. Rewoldt, E. Sonnendrucker, H. Sugama, W. M. Tang, X. Tang, Y-X. Wan, T-S. F. Wang, S. J. Wang, and X. Xu for inspirations and fruitful discussions. Finally, I thank Prof. Thierry Passot, Prof. Pierre-Louis Sulem, and Prof. Catherine Sulem for the invitation to give lectures on this subject at the Fields Institute.

References

[1] W. E. Amatucci, D. N. Walker, G. Gatling, and E. E. Scime, *Direct observation of microparticle gyromotion in a magnetized direct current glow discharge dusty plasma*, Physics of Plasmas **11** (2004), no. 5, 2097–2105.

[2] T. M. Antonsen and B. Lane, *Kinetic-equations for low-frequency instabilities in inhomogeneous plasmas*, Physics of Fluids **23** (1980), no. 6, 1205–1214.

[3] V. I. Arnold, *Mathematical methods of classicalmechanics*, second ed., pp. 240–241, Springer-Verlag, New York, 1989.

[4] A. Brizard, *Nonlinear gyrokinetic Maxwell-Vlasov equations using magnetic coordinates*, Journal of Plasma Physics **41** (1989), no. 6, 541–559.

[5] J. Candy and R. E. Waltz, *Anomalous transport scaling in the DIII-D tokamak matched by supercomputer simulation*, Physical Review Letters **91** (2003), no. 4, 045001.

[6] J. R. Cary, *Lie transform perturbation-theory for Hamiltonian-systems*, Physics Reports-Review Section of Physics Letters **79** (1981), no. 2, 129–159.

[7] J. R. Cary and R. G. Littlejohn, *Noncanonical Hamiltonian-mechanics and its application to magnetic-field line flow*, Annals of Physics **151** (1983), no. 1, 1–34.

[8] P. J. Catto, *Linearized gyro-kinetics*, Plasma Physics and Controlled Fusion **20** (1978), no. 7, 719–722.

[9] P. J. Catto, W. M. Tang, and D. E. Baldwin, *Generalized gyrokinetics*, Plasma Physics and Controlled Fusion **23** (1981), no. 7, 639–650.

[10] L. Chen and S. T. Tsai, *Electrostatic-waves in general magnetic-field configurations*, Physics of Fluids **26** (1983), no. 1, 141–145.

[11] Y. Chen and S. E. Parker, *A delta-f particle method for gyrokinetic simulations with kinetic electrons and electromagnetic perturbations*, Journal of Computational Physics **189** (2003), no. 2, 463–475.

[12] G. F. Chew, M. L. Goldberger, and F. E. Low, *The Boltzmann equation and the one-fluid hydromagnetic equations in the absence of particle collisions*, Proceedings of the Royal Society of London Series A-Mathematical and Physical Sciences **236** (1956), no. 1204, 112–118.

[13] B. I. Cohen, T. A. Brengle, D. B. Conley, and R. P. Freis, *An orbit averaged particle code*, Journal of Computational Physics **38** (1980), no. 1, 45–63.

[14] B. I. Cohen, T. J. Williams, A. M. Dimits, and J. A. Byers, *Gyrokinetic simulations of EXB velocity-shear effects on ion-temperature-gradient modes*, Physics of Fluids B-Plasma Physics **5** (1993), no. 8, 2967–2980.

[15] R. C. Davidson, *Invariant of higher-order Chew-Goldberger-Low theory*, Physics of Fluids **10** (1967), no. 3, 669–670.

[16] R. C. Davidson and H. Qin, *Physics of intense charged particle beams in high energy accelerators*, pp. 301–352, World Scientific, Singapore, 2001.

[17] R. C. Davidson, H. Qin, and P. J. Channell, *Periodically-focused solutions to the nonlinear Vlasov-Maxwell equations for intense charged particle beams*, Physics Letters a **258** (1999), no. 4-6, 297–304.

[18] A. M. Dimits, L. L. Lodestro, and D. H. E. Dubin, *Gyroaveraged equations for both the gyrokinetic and drift-kinetic regimes*, Physics of Fluids B-Plasma Physics **4** (1992), no. 1, 274–277.

[19] A. M. Dimits, T. J. Williams, J. A. Byers, and B. I. Cohen, *Scalings of ion-temperature-gradient-driven anomalous transport in tokamaks*, Physical Review Letters **77** (1996), no. 1, 71–74.

[20] I. Y. Dodin and N. J. Fisch, *Motion of charged particles near magnetic-field discontinuities*, Physical Review E **6401** (2001), no. 1, 016405.

[21] W. Dorland, F. Jenko, M. Kotschenreuther, and B. N. Rogers, *Electron temperature gradient turbulence*, Physical Review Letters **85** (2000), no. 26, 5579–5582.

[22] D. H. E. Dubin, J. A. Krommes, C. Oberman, and W. W. Lee, *Non-linear gyrokinetic equations*, Physics of Fluids **26** (1983), no. 12, 3524–3535.

[23] A. Friedman, A. B. Langdon, and B. I. Cohen, *A direct method for implicit particle-in-cell simulation*, Comments Plasma Phys. Controlled Fusion **6** (1981), 225–236.

[24] E. A. Frieman and L. Chen, *Non-linear gyrokinetic equations for low-frequency electromagnetic-waves in general plasma equilibria*, Physics of Fluids **25** (1982), no. 3, 502–508.

[25] E. A. Frieman, R. C. Davidson, and B. Langdon, *Higher-order corrections to chew-goldberger-low theory*, Physics of Fluids **9** (1966), no. 8, 1475–1482.
[26] W. Greub, S. Halperin, and R. Vanstone, *Connections and curvature and cohomology*, vol. II, pp. 242–243, Academic Press, New York, 1973.
[27] T. S. Hahm, *Nonlinear gyrokinetic equations for tokamak microturbulence*, Physics of Fluids **31** (1988), no. 9, 2670–2673.
[28] T. S. Hahm, W. W. Lee, and A. Brizard, *Nonlinear gyrokinetic theory for finite-beta plasmas*, Physics of Fluids **31** (1988), no. 7, 1940–1948.
[29] J. D. Jackson, *Classical electrodynamics*, pp. 137–146, John Wiley & Sons, New York, 1975.
[30] J. A. Krommes, *Dielectric response and thermal fluctuations in gyrokinetic plasma*, Physics of Fluids B-Plasma Physics **5** (1993), no. 4, 1066–1100.
[31] M. Kruskal, *Asymptotic theory of Hamiltonian and other systems with all solutions nearly periodic*, Journal of Mathematical Physics **3** (1962), no. 4, 806–828.
[32] R.M. Kulsrud, *Adiabatic invairant of the harmonic oscillator*, Physical Review **106** (1957), no. 2, 205–207.
[33] W. W. Lee, *Gyrokinetic approach in particle simulation*, Physics of Fluids **26** (1983), no. 2, 556–562.
[34] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, *Turbulent transport reduction by zonal flows: Massively parallel simulations*, Science **281** (1998), no. 5384, 1835–1837.
[35] R. G. Littlejohn, *Guiding center Hamiltonian - new approach*, Journal of Mathematical Physics **20** (1979), no. 12, 2445–2458.
[36] ———, *Hamiltonian-formulation of guiding center motion*, Physics of Fluids **24** (1981), no. 9, 1730–1749.
[37] ———, *Hamiltonian theory of guiding center bounce motion*, Physica Scripta **T2** (1982), 119–125.
[38] ———, *Variational-principles of guiding center motion*, Journal of Plasma Physics **29** (1983), no. 2, 111–125.
[39] ———, *Phase anholonomy in the classical adiabatic motion of charged-particles*, Physical Review a **38** (1988), no. 12, 6034–6045.
[40] L. Michelotti, *Intermediate classical dynamics with applications to beam physics*, p. 240, John Wiley & Sons, New York, 1995.
[41] T. G. Northrop, *The adiabatic motion of charged particles*, pp. 78–95, John Wiley & Sons, New York, 1963.
[42] S. E. Parker, W. W. Lee, and R. A. Santoro, *Gyrokinetic simulation of ion-temperature-gradient-driven turbulence in 3d toroidal geometry*, Physical Review Letters **71** (1993), no. 13, 2042–2045.
[43] H. Qin, *Gyrokinetic theory and computational methods for electromagnetic perturbations in tokamaks*, Ph.D. thesis, Princeton University, Princeton, NJ 08540, 1998.
[44] ———, *Differential geometry and gyrokinetic theory*, Bulletin of the American Physical Society (New York), vol. 47, American Physical Society, 2002, p. 295.
[45] H. Qin and W. M. Tang, *Pullback transformations in gyrokinetic theory*, Physics of Plasmas **11** (2004), no. 3, 1052–1063.
[46] H. Qin, W. M. Tang, and W. W. Lee, *Gyrocenter-gauge kinetic theory*, Physics of Plasmas **7** (2000), no. 11, 4433–4445.
[47] H. Qin, W. M. Tang, W. W. Lee, and G. Rewoldt, *Gyrokinetic perpendicular dynamics*, Physics of Plasmas **6** (1999), no. 5, 1575–1588.
[48] H. Qin, W. M. Tang, and G. Rewoldt, *Gyrokinetic theory for arbitrary wavelength electromagnetic modes in tokamaks*, Physics of Plasmas **5** (1998), no. 4, 1035–1049.
[49] ———, *Linear gyrokinetic theory for kinetic magnetohydrodynamic eigenmodes in tokamak plasmas*, Physics of Plasmas **6** (1999), no. 6, 2544–2562.
[50] H. Qin, W. M. Tang, G. Rewoldt, and W. W. Lee, *On the gyrokinetic equilibrium*, Physics of Plasmas **7** (2000), no. 3, 991–1000.
[51] P. H. Rutherford and E. A. Frieman, *Drift instabilities in general magnetic field configurations*, Physics of Fluids **11** (1968), no. 3, 569–585.
[52] D. C. Spitzer, 2004, Private communication with Mrs. D. C. Spitzer.
[53] L. Spitzer, *Equations of motion for an ideal plasma*, Astrophysical Journal **116** (1952), no. 2, 299–316.

[54] L. Spitzer, *Physics of fully ionized gases*, second ed., pp. 30–40, Interscience Publications, New York, 1962.

[55] T. H. Stix, *Waves in plasmas*, pp. 247–262, American Institute of Physics, New York, 1992.

[56] H. Sugama, *Gyrokinetic field theory*, Physics of Plasmas **7** (2000), no. 2, 466–480.

[57] R. D. Sydora, V. K. Decyk, and J. M. Dawson, *Fluctuation-induced heat transport results from a large global 3D toroidal particle simulation model*, Plasma Physics and Controlled Fusion **38** (1996), no. 12A, A281–A294.

[58] W. M. Tang, *Introduction to gyrokinetic theory with applications in magnetic confinement research in plasma physics*, Topics in Kinetic Theory (D. Levermore, T. Passot, C. Sulem, and P.L. Sulem, eds.), Fields Institute Communications, American Mathematical Society, 2004, in this volume.

[59] J. B. Taylor and R. J. Hastie, *Stability of general plasma equilibria .I. formal theory*, Plasma Physics **10** (1968), no. 5, 479–494.

[60] S. J. Wang, *Canonical Hamiltonian gyrocenter variables and gauge invariant representation of the gyrokinetic equation*, Physical Review E **64** (2001), no. 5, 056404.

[61] H. Weitzner and C. Chang, *Extensions of adiabatic invariant theory for a charged particle*, Physics of Plasmas **12** (2005), no. 12, 012106.