Nonlinear relativistic gyrokinetic Vlasov-Maxwell equations

Alain J. Brizard
Lawrence Berkeley National Laboratory and Physics Department, University of California, Berkeley, California 94720

Anthony A. Chan
Department of Space Physics and Astronomy, Rice University, Houston, Texas 77005-1892

(Received 12 July 1999; accepted 20 August 1999)

A set of self-consistent nonlinear gyrokinetic equations is derived for relativistic charged particles in a general nonuniform magnetized plasma. Full electromagnetic-field fluctuations are considered with spatial and temporal scales given by the low-frequency gyrokinetic ordering. Self-consistency is obtained by combining the nonlinear relativistic gyrokinetic Vlasov equation with the low-frequency Maxwell equations in which charge densities and current densities are expressed in terms of moments of the gyrokinetic Vlasov distribution. For these self-consistent gyrokinetic equations, a low-frequency energy conservation law is also derived. © *1999 American Institute of Physics.* [S1070-664X(99)01012-5]

I. INTRODUCTION

Relativistic particles, particularly electrons, are found in many plasmas, including fusion plasmas,[1–4] space plasmas[5] and astrophysical plasmas.[6] In addition to the general question of how these particles are transported and accelerated to relativistic energies, the particles are of interest because they can cause significant damage. In fusion plasmas, for example, relativistic runaway electrons can cause serious damage to the containment vessel.[3,4] In space plasmas, on the other hand, relativistic electrons trapped in the Earth's radiation belts are hazardous to spacecraft hardware[7] and they are a radiation hazard to astronauts.[8]

In high-temperature tokamak plasmas, multi-MeV runaway electrons are typically generated during the start-up phase or following plasma disruptions.[3] After a disruption, a large fraction of the stored magnetic energy in the tokamak plasma can be transferred to the runaway-electron population. The primary mechanism for the generation of relativistic runaway electrons involves an accelerating electric field strong enough to overcome the frictional drag produced by Coulomb collisions.[4,9,10] The confinement properties of relativistic test-particles in tokamak plasmas have been investigated in Refs. 11 and 12; these studies reveal that the guiding-center approximation is appropriate for relativistic runaway-electron dynamics in axisymmetric and non-axisymmetric tokamak plasmas.

In space plasmas there is observational evidence that relativistic electrons may be produced during magnetic storms by a drift-resonant exchange of energy with hydromagnetic waves in the Earth's magnetosphere.[13,14] The drift-resonant interaction conserves the first adiabatic invariant and transports particles to regions of higher magnetic field magnitude, thereby increasing their energy.[15,16] The time-scale separation between the short gyroperiod and the long drift-period motivates the asymptotic removal of the gyromotion time scale from the *exact* dynamical description. The resulting *reduced* dynamical description is expressed in terms of guiding-center and gyrocenter dynamics wherein the equations of motion are independent of the fast gyromotion degree of freedom.

In previous work on radiation-belt transport, various Fokker-Planck transport equations for the average phase-space density as a function of the three adiabatic invariants are used.[5,17] These radiation-belt transport equations are essentially quasilinear transport equations in the space of the three adiabatic invariants. It is important to note that these equations are not derived from first principles; rather, the form of the equation is assumed and the transport coefficients for diffusion and drag are given semiempirical forms and then adjusted by attempting to reproduce measured particle fluxes.

Recently, nonrelativistic quasilinear transport equations have been derived to describe the interaction of nonrelativistic particles with magnetospheric hydromagnetic waves.[18] In these equations it is shown that the wave-particle interactions result in a drag-like term which can be comparable to the diffusion term. We expect that there will be a similar drag-like contribution in the relativistic quasilinear transport equation; in the past this contribution has been ignored because the drag term has been assumed to arise only from Coulomb scattering. Chen's derivation[18] began with the nonrelativistic nonlinear gyrokinetic Vlasov equation derived by Brizard.[19,20]

In this paper a nonlinear relativistic gyrokinetic Vlasov equation is derived, partly as a starting point for a later derivation of relativistic quasilinear transport equations for space physics applications, and partly because it is interesting in its own right as a relativistic generalization of the earlier work of Brizard.[20] Other applications might include the gyrokinetic particle simulation of relativistic runaway electron transport in tokamak plasmas. The nonlinear relativistic gyrokinetic equations derived here also extend the earlier work by Littlejohn[21] and by Tsai, Van Dam and Chen,[22] who derived linear relativistic gyrokinetic equations.

Finite gyroradius effects are usually small for electrons, but they are not necessarily negligible, especially for highly relativistic electrons such as those found in fusion and space plasmas.[23] For generality we retain the finite gyroradius effects in this paper; in a future paper we will obtain the drift-kinetic equations from the appropriate limit of the finite gyroradius equations.

The remainder of the paper is organized as follows. In Sec. II, relativistic Vlasov-Maxwell equations are presented as a foundation for work done in future sections. In Sec. III, an outline of the derivation of relativistic Hamiltonian guiding-center theory based on the phase-space Lagrangian Lie-transform perturbation method is presented. Sections IV and V, where relativistic Hamiltonian gyrocenter theory and nonlinear relativistic gyrokinetic Vlasov-Maxwell equations, respectively, are derived, contain the new results in nonlinear relativistic gyrokinetic theory. Lastly, Sec. VI presents a summary of our work and discusses its applications.

II. RELATIVISTIC VLASOV-MAXWELL EQUATIONS

In this section, we present two Hamiltonian formulations of the relativistic Vlasov-Maxwell equations. Each formulation is defined in terms of a Hamiltonian function H and a Poisson bracket $\{,\}$ derived from a phase-space Lagrangian. The first formulation is based on a covariant description of relativistic charged particle dynamics in an arbitrary electromagnetic field in terms of the phase-space coordinates $(x^0=ct,\mathbf{x};\, p^0=E/c,\mathbf{p})$, where $(x^0,\mathbf{x})$ denotes the space-time position of a particle and $(p^0,\mathbf{p})$ denotes its momentum-energy coordinates. The covariant formulation treats space and time on an equal footing. The second formulation, on the other hand, treats time and space separately and makes use of the extended phase-space coordinates $(\mathbf{x},\mathbf{p};t,w)$, where the energy coordinate w is canonically conjugate to time t.

For each Hamiltonian formulation, a relativistic Vlasov equation $\{f,H\}=0$ for the Vlasov distribution function f is written in terms of the Hamiltonian function H and the Poisson bracket $\{,\}$. Next, self-consistent relativistic Maxwell equations are presented in which the charge and current densities are expressed in terms of moments of the relativistic Vlasov distribution function f. Lastly, an exact energy-momentum conservation law for the self-consistent relativistic Vlasov-Maxwell equations is presented.

A. Relativistic charged particle dynamics

The relativistic motion of a particle of rest-mass m and charge q is described in eight-dimensional phase space in terms of the space-time coordinates $x^\alpha=(x^0=ct,\mathbf{x})$ and the four-momentum $p^\alpha\equiv mu^\alpha=(p^0,\mathbf{p})$, where the four-velocity is

$$u^\alpha\equiv\frac{dx^\alpha}{d\tau}\equiv(u^0=\gamma c,\,\mathbf{u}=\gamma\mathbf{v}), \tag{1}$$

with $\gamma\equiv(1-|\mathbf{v}|^2/c^2)^{-1/2}$ the relativistic factor and $d/d\tau=\gamma\, d/dt$ the derivative with respect to proper time.

The equation of motion for the four-momentum p^α is

$$\frac{dp^\alpha}{d\tau}=\frac{q}{c}F^{\alpha\beta}u_\beta, \tag{2}$$

where summation over repeated indices is assumed and

$$F^{\alpha\beta}\equiv\partial^\alpha A^\beta-\partial^\beta A^\alpha \tag{3}$$

is the Faraday tensor.[24] Here, the space-time contravariant derivative is $\partial^\alpha\equiv g^{\alpha\beta}\partial_\beta=(-\partial/\partial x^0,\nabla)$, where $g^{\alpha\beta}=\mathrm{diag}\,(-1,+1,+1,+1)$ is the space-like metric tensor, and

$$A^\alpha=(\Phi,\mathbf{A}) \tag{4}$$

is the four-dimensional electromagnetic potential.

B. Noncanonical Hamiltonian structures

1. Covariant noncanonical Hamiltonian structure

We now show that the equations (1) and (2) possess a covariant (cov) Hamiltonian structure, i.e., they can be written as $d\mathcal{Z}^a/d\tau=\{\mathcal{Z}^a,\mathcal{H}\}_{\mathrm{cov}}$, where $\mathcal{Z}^a=(x^0,\mathbf{x};p^0,\mathbf{p})$ are eight-dimensional noncanonical phase-space coordinates, $\mathcal{H}$ is the covariant relativistic particle Hamiltonian and $\{,\}_{\mathrm{cov}}$ is the covariant relativistic noncanonical Poisson bracket on eight-dimensional phase space.

To derive the covariant relativistic Poisson bracket $\{,\}_{\mathrm{cov}}$, we proceed along the same lines as with the nonrelativistic case (see Ref. 25 for details). We begin with the covariant relativistic phase-space Lagrangian,

$$\Gamma_{\mathrm{cov}}=\left(p_\alpha+\frac{q}{c}A_\alpha\right)dx^\alpha\equiv\Gamma_a d\mathcal{Z}^a. \tag{5}$$

From the fundamental Lagrange brackets $[\mathcal{Z}^a,\mathcal{Z}^b]\equiv\Omega_{ab}$, where[26]

$$\Omega_{ab}\equiv\frac{\partial\Gamma_b}{\partial\mathcal{Z}^a}-\frac{\partial\Gamma_a}{\partial\mathcal{Z}^b}, \tag{6a}$$

we define the Lagrange (8×8) matrix with components Ω_{ab},

$$\begin{pmatrix} (q/c)\,F_{\alpha\beta} & -g_{\alpha\beta} \\ g_{\alpha\beta} & 0 \end{pmatrix}, \tag{6b}$$

where each component in (6b) is a 4×4 matrix. The inverse of the Lagrange matrix (6b) defines the Poisson matrix,

$$\begin{pmatrix} 0 & g^{\alpha\beta} \\ -g^{\alpha\beta} & (q/c)\,F^{\alpha\beta} \end{pmatrix}, \tag{7}$$

whose elements in turn define the fundamental Poisson brackets $\{\mathcal{Z}^a,\mathcal{Z}^b\}$,

$$\{x^\alpha,x^\beta\}_{\mathrm{cov}}\equiv0,\quad \{x^\alpha,p^\beta\}_{\mathrm{cov}}\equiv g^{\alpha\beta},\ \text{and}\ \{p^\alpha,p^\beta\}_{\mathrm{cov}}\equiv\frac{q}{c}F^{\alpha\beta}.$$

The Poisson bracket $\{\mathcal{A},\mathcal{B}\}_{\mathrm{cov}}$ of two functions $\mathcal{A}$ and $\mathcal{B}$ on eight-dimensional phase space is thus constructed from these fundamental Poisson brackets to yield

$$\{\mathcal{A},\mathcal{B}\}_{\mathrm{cov}}\equiv g^{\alpha\beta}\left(\frac{\partial\mathcal{A}}{\partial x^\alpha}\frac{\partial\mathcal{B}}{\partial p^\beta}-\frac{\partial\mathcal{A}}{\partial p^\beta}\frac{\partial\mathcal{B}}{\partial x^\alpha}\right)+\frac{q}{c}F^{\alpha\beta}\frac{\partial\mathcal{A}}{\partial p^\alpha}\frac{\partial\mathcal{B}}{\partial p^\beta}. \tag{8}$$

We note that the phase-space coordinates $\mathcal{Z}^a$ are noncanonical because of the second term in the Poisson bracket (8). Furthermore, the Poisson bracket (8) is guaranteed to satisfy the Jacobi identity as well as other Poisson-bracket properties since it was derived from a phase-space Lagrangian.

We now determine the relativistic Hamiltonian function $\mathcal{H}$ by requiring that $d\mathcal{Z}^a/d\tau = \{\mathcal{Z}^a, \mathcal{H}\}_{\rm cov}$. Substituting $\mathcal{A} = x^\alpha$ or p^α and $\mathcal{B} = \mathcal{H}$ in (8), we find

$$\{x^\alpha, \mathcal{H}\}_{\rm cov} = g^{\alpha\beta} \frac{\partial \mathcal{H}}{\partial p^\beta}$$

and

$$\{p^\alpha, \mathcal{H}\}_{\rm cov} = -\, g^{\alpha\beta} \frac{\partial \mathcal{H}}{\partial x^\beta} + \frac{q}{c} F^{\alpha\beta} \frac{\partial \mathcal{H}}{\partial p^\beta}.$$

We recover (1) and (2) from these equations if the Hamiltonian function $\mathcal{H}$ is

$$\mathcal{H} \equiv \frac{1}{2m} g_{\alpha\beta} p^\alpha p^\beta = \frac{m}{2} g_{\alpha\beta} u^\alpha u^\beta. \tag{9}$$

The covariant dynamical Eqs. (1) and (2) can thus be written in Hamiltonian form,

$$\left. \begin{array}{l} dx^\alpha/d\tau = \{x^\alpha, \mathcal{H}\}_{\rm cov} = p^\alpha/m \equiv u^\alpha \\ dp^\alpha/d\tau = \{p^\alpha, \mathcal{H}\}_{\rm cov} = (q/mc)\, F^{\alpha\beta}\, p_\beta \equiv (q/c)\, F^{\alpha\beta}\, u_\beta \end{array} \right\}. \tag{10}$$

Note that covariant relativistic charged-particle motion takes place on the surface $\mathcal{H} \equiv -mc^2/2$, where the right side is obtained by substituting (1) into (9). Hence the Hamiltonian $\mathcal{H}$ is a Lorentz scalar.[24] More importantly, however, the Hamiltonian function (9) has no classical analog (i.e., the limit $c \rightarrow \infty$ is not defined). This property of the covariant Hamiltonian (9) forces us to seek an alternative Hamiltonian formulation for relativistic charged-particle motion with a well-defined classical limit.

2. Extended phase-space Hamiltonian structure

The relativistic Hamilton Eqs. (1) and (2) can alternately be written in terms of the reference-frame time t as $dz^a/dt \equiv \{z^a, H\}$, where $z^a \equiv (\mathbf{x}, \mathbf{p}; t, w)$ are extended phase-space coordinates and the Hamiltonian structure is given in terms of the extended relativistic Hamiltonian (for positive-energy particles),

$$H \equiv \gamma mc^2 + q\,\Phi - w, \tag{11}$$

where $\gamma \equiv \sqrt{1 + |\mathbf{p}/mc|^2}$ and, since the relativistic charged-particle motion now takes place on the surface $H = 0$, $w = \gamma mc^2 + q\,\Phi$ represents the total energy of a charged particle (including its rest energy). In contrast to the covariant Hamiltonian (9), the Hamiltonian (11) has a well-defined classical limit and is an energy-like quantity. The phase-space Lagrangian, on the other hand, is

$$\Gamma \equiv \left(\mathbf{p} + \frac{q}{c} \mathbf{A} \right) \cdot d\mathbf{x} - w\, dt, \tag{12a}$$

and its associated noncanonical Poisson bracket is derived by the same procedure used in going from (5) to (8) above; for two arbitrary functions F and G on extended phase space, it is given as

$$\{F, G\} \equiv \left(\frac{\partial F}{\partial w} \frac{\partial G}{\partial t} - \frac{\partial F}{\partial t} \frac{\partial G}{\partial w} \right) + \left(\nabla F \cdot \frac{\partial G}{\partial \mathbf{p}} - \frac{\partial F}{\partial \mathbf{p}} \cdot \nabla G \right) + \frac{q}{c} \left[\mathbf{B} \cdot \frac{\partial F}{\partial \mathbf{p}} \times \frac{\partial G}{\partial \mathbf{p}} + \frac{\partial \mathbf{A}}{\partial t} \cdot \left(\frac{\partial F}{\partial \mathbf{p}} \frac{\partial G}{\partial w} - \frac{\partial F}{\partial w} \frac{\partial G}{\partial \mathbf{p}} \right) \right]. \tag{12b}$$

Using (11) and (12b), the relativistic noncanonical Hamilton equations for charged particle motion are now

$$\frac{dt}{dt} \equiv -\frac{\partial H}{\partial w} = 1, \tag{13a}$$

$$\frac{dw}{dt} \equiv \frac{\partial H}{\partial t} - \frac{q}{c} \frac{\partial \mathbf{A}}{\partial t} \cdot \frac{\partial H}{\partial \mathbf{p}} = q \left(\frac{\partial \Phi}{\partial t} - \frac{\mathbf{v}}{c} \cdot \frac{\partial \mathbf{A}}{\partial t} \right), \tag{13b}$$

$$\frac{d\mathbf{x}}{dt} \equiv \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p}}{m\gamma} = \mathbf{v}, \tag{13c}$$

$$\frac{d\mathbf{p}}{dt} \equiv -\nabla H + \frac{q}{c} \frac{\partial \mathbf{A}}{\partial t} \frac{\partial H}{\partial w} + \frac{q}{c} \frac{\partial H}{\partial \mathbf{p}} \times \mathbf{B} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right). \tag{13d}$$

This set of Hamilton equations presents an alternative description to the Hamilton equations (10) derived within the covariant formulation. In the present formulation, we note from (13a) that the time coordinate t can be identified with the Hamiltonian orbit parameter.

C. Relativistic Vlasov equations

The relativistic Vlasov equation describes the fact that the Vlasov distribution function is conserved along a relativistic Hamiltonian orbit in phase space. First, using the covariant formalism discussed above, this statement is mathematically expressed as

$$\frac{d\mathcal{F}}{d\tau} \equiv \{\mathcal{F}, \mathcal{H}\}_{\rm cov} = 0, \tag{14}$$

where $\mathcal{F}(x,p)$ is the covariant Vlasov distribution function on eight-dimensional phase space. It can be explicitly written as

$$0 = u^\alpha \frac{\partial \mathcal{F}}{\partial x^\alpha} + \frac{q}{c} \frac{\partial \mathcal{F}}{\partial p^\alpha} F^{\alpha\beta} u_\beta, \tag{15a}$$

or

$$0 = \frac{\partial}{\partial \mathcal{Z}^a} \left(\frac{d\mathcal{Z}^a}{d\tau} \mathcal{F} \right), \tag{15b}$$

since Hamilton's equations are incompressible: $\partial(d\mathcal{Z}^a/d\tau)/\partial \mathcal{Z}^a \equiv 0$.

The covariant relativistic Vlasov equation (14) has one additional degree of freedom compared to its non-relativistic version. This is due to the fact that the energy coordinate p^0 is not really independent of the other coordinates because the

Hamiltonian (9) is constrained to be equal to $-mc^2/2$ (i.e., $p_\alpha p^\alpha = -m^2c^2$). Hence the covariant Vlasov distribution function $\mathcal{F}(x,p)$ must vanish for $p^0 \neq \gamma mc$ (for positive-energy particles). This thus implies that the covariant Vlasov distribution $\mathcal{F}(x,p)$ should be written as[27]

$$\mathcal{F}(x,p) \equiv c\,\delta\left(\mathcal{H} + \frac{mc^2}{2}\right) f(\mathbf{x},\mathbf{p},t) = \delta(p^0 - \gamma mc)\,\frac{f(\mathbf{x},\mathbf{p},t)}{\gamma}, \tag{16}$$

where $f(\mathbf{x},\mathbf{p},t)$ is the Vlasov distribution on the six-dimensional phase space with coordinates $(\mathbf{x},\mathbf{p})$ and $\gamma \equiv (1 + |\mathbf{p}/mc|^2)^{1/2}$ is the relativistic factor expressed in terms of the relativistic kinetic momentum $\mathbf{p}$; note that $f(\mathbf{x},\mathbf{p},t)$ is explicitly independent of the energy coordinate w. When (16) is substituted into (14), one obtains

$$0 = c\,\delta\left(\mathcal{H} + \frac{mc^2}{2}\right)\{f,\mathcal{H}\}_{\rm cov} \equiv \delta(p^0 - \gamma mc) \times \left[\frac{\partial f}{\partial t} + \frac{\mathbf{u}}{\gamma}\cdot\nabla f + q\left(\mathbf{E} + \frac{\mathbf{u}}{\gamma c}\times\mathbf{B}\right)\cdot\frac{\partial f}{\partial \mathbf{p}}\right]. \tag{17a}$$

After performing a p^0-integration on (17a), we obtain the relativistic Vlasov equation on $6+1$ phase space,

$$0 = \frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f + q\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right)\cdot\frac{\partial f}{\partial \mathbf{p}}, \tag{17b}$$

which looks exactly like the nonrelativistic Vlasov equation except that $\mathbf{p}$ is the relativistic kinetic momentum $\mathbf{p} = m\gamma\mathbf{v}$.[27] The relativistic Vlasov equation (17b) can also be written as

$$\frac{df}{dt} \equiv \{f,H\} = 0, \tag{18}$$

where the relativistic Hamiltonian is given by (11) and the Poisson bracket is given by (12b). Note that this new Hamiltonian formulation has separated the components of the electromagnetic four-potential $(\Phi,\mathbf{A})$: the electrostatic potential Φ now appears in the Hamiltonian (11) while the magnetic vector potential $\mathbf{A}$ and its derivatives remain in the phase-space Lagrangian (12a) and the Poisson bracket (12b).

D. Self-consistent relativistic Vlasov-Maxwell equations

To obtain a self-consistent relativistic Vlasov-Maxwell theory one combines the covariant relativistic Vlasov equation, (14) or (18), with the Maxwell equations

$$\partial_\beta F^{\alpha\beta}(r) = 4\pi\sum q\int d^8\mathcal{Z}\,\delta^4(x-r)\,\frac{p^\alpha}{mc}\mathcal{F}(x,p) = 4\pi\sum q\int d^6z\,\delta^3(\mathbf{x}-\mathbf{r})\,\frac{p^\alpha}{mc\gamma}f(\mathbf{x},\mathbf{p};t), \tag{19}$$

where the sum is over particle species and $r^\alpha = (ct,\mathbf{r})$ denotes the space-time location where the charge and current densities are evaluated. The second expression on the right side of (19) is obtained by substituting (16) into the first expression and performing the p^0-integration.

The relativistic Vlasov-Maxwell equations (14) or (18) and (19) satisfy an energy-momentum conservation law

$$\frac{\partial}{\partial x^\alpha}(T_M^{\alpha\beta} + T_V^{\alpha\beta}) = 0, \tag{20a}$$

where the energy-momentum stress tensor $T^{\alpha\beta} \equiv T_M^{\alpha\beta} + T_V^{\alpha\beta}$ is divided into the Maxwell (field) part,

$$T_M^{\alpha\beta} \equiv \frac{1}{16\pi}(F_{\mu\nu}F^{\nu\mu})g^{\alpha\beta} - \frac{1}{4\pi}F^{\alpha\mu}g_{\mu\nu}F^{\nu\beta}, \tag{20b}$$

and the Vlasov (particle) part,

$$T_V^{\alpha\beta} \equiv \sum\int d^4p\,mu^\alpha u^\beta\mathcal{F} = \sum\int d^3p\,\frac{mu^\alpha u^\beta}{\gamma}f. \tag{20c}$$

Of particular importance in the development of self-consistent Hamiltonian gyrokinetic particle simulation techniques is the energy conservation law

$$\frac{\partial\mathcal{E}}{\partial t} + \nabla\cdot\mathbf{S} = 0, \tag{21a}$$

where the energy density is

$$\mathcal{E} = \frac{1}{8\pi}(|\mathbf{E}|^2 + |\mathbf{B}|^2) + \sum\int d^3p\,\gamma mc^2 f, \tag{21b}$$

and the energy-density flux is

$$\mathbf{S} = \frac{c}{4\pi}\mathbf{E}\times\mathbf{B} + \sum\int d^3p\,mc^2\,\mathbf{u}f. \tag{21c}$$

We return to this conservation law in Sec. V C, where an approximate energy conservation law is derived for the nonlinear relativistic gyrokinetic Vlasov-Maxwell equations.

In summary, two Hamiltonian formulations for relativistic charged-particle motion are presented in this section: the covariant formulation based on the Hamiltonian (9) and the Poisson bracket (8); and the extended phase-space formulation based on the Hamiltonian (11) and the Poisson bracket (12b). For each formulation, self-consistent relativistic Vlasov-Maxwell equations [(14) or (18) and (19)] are given in addition to the exact energy-momentum conservation law (20a). In the remainder of this paper, we adopt the extended phase-space formulation since its Hamiltonian function has a well-defined classical limit and the separation of the time coordinate from the other phase-space coordinates allows us to transform the latter coordinates without transforming the time coordinate itself.

III. RELATIVISTIC HAMILTONIAN GUIDING-CENTER THEORY

A. Preliminary coordinate transformation

Relativistic charged-particle motion in a strong magnetic field is characterized by three disparate time scales: the fast gyromotion time scale associated with gyration about a single magnetic field line, the intermediate parallel (or bounce) time scale associated with motion along a field line, and the slow drift time scale associated with motion across magnetic-field lines.[28]

In the following analysis (only results are presented here), we work in a preferred reference frame in which the background magnetic field $\mathbf{B}_0 \equiv \nabla \times \mathbf{A}_0$ is time independent ($\partial \mathbf{A}_0/\partial t \equiv 0$) and the background electrostatic field Φ_0 is zero. We note that the choice $\Phi_0 \equiv 0$ is not crucial to the present analysis and that nonrelativistic gyrokinetic theory in the presence of background nonuniform electrostatic fields has been systematically derived elsewhere using phase-space Lagrangian methods.[19,29] In the present paper, any quasistatic electrostatic potential is thus treated as a perturbation. We further note that the covariant formulation of relativistic guiding-center Hamiltonian theory has been developed in Ref. 30 in which the background electromagnetic field is expressed in terms of Φ_0 and $\mathbf{A}_0$.

The analysis begins with the introduction of the following *local* momentum coordinates $(p_{\|0}, \mu_0, \theta_0)$,

$$\mathbf{p} \equiv p_{\|0}\,\hat{\mathbf{b}} - \sqrt{2m\mu_0 B_0}\,(\hat{\mathbf{a}}\cos\theta_0 + \hat{\mathbf{c}}\sin\theta_0), \tag{22}$$

where $p_{\|0} \equiv m\gamma\mathbf{v}\cdot\hat{\mathbf{b}}$ is the component of the relativistic momentum parallel to the local background magnetic field ($\mathbf{B}_0\cdot\hat{\mathbf{b}} \equiv B_0$), $\mu_0 \equiv m\gamma^2|\mathbf{v}_\perp|^2/2B_0$ is the relativistic magnetic moment, while $\theta_0 \equiv \tan^{-1}[(-\mathbf{p}\cdot\hat{\mathbf{c}})/(-\mathbf{p}\cdot\hat{\mathbf{a}})]$ is the local gyrorangle, and $(\hat{\mathbf{a}},\hat{\mathbf{b}},\hat{\mathbf{c}})$ form a right-handed orthogonal unit-vector set. The Jacobian for this preliminary coordinate transformation is mB_0.

Inserting the local momentum coordinates into (12a), the unperturbed phase-space Lagrangian is

$$\Gamma_0 \equiv \left(\frac{q}{c}\,\mathbf{A}_0 + p_{\|0}\,\hat{\mathbf{b}} + m\omega_B\frac{\partial\boldsymbol{\rho}_0}{\partial\theta_0}\right)\cdot d\mathbf{x} - w\,dt, \tag{23}$$

where $\omega_B \equiv qB_0/mc$ is the signed (rest-mass) gyrofrequency and the lowest-order relativistic gyroradius vector $\boldsymbol{\rho}_0$ is

$$\boldsymbol{\rho}_0 \equiv \frac{\hat{\mathbf{b}}\times\mathbf{u}}{\omega_B} = \frac{1}{\omega_B}\sqrt{\frac{2\mu_0 B_0}{m}}\,(\hat{\mathbf{c}}\cos\theta_0 - \hat{\mathbf{a}}\sin\theta_0), \tag{24}$$

while the unperturbed Hamiltonian function is

$$H_0 = mc^2\sqrt{1 + (2\mu_0 B_0/mc^2) + (p_{\|0}/mc)^2} - w. \tag{25}$$

After this preparatory transformation, we now proceed with the asymptotic elimination of the fast gyromotion time scale from the relativistic Hamiltonian system represented by (23) and (25).

B. Relativistic guiding-center Hamiltonian dynamics

Using the phase-space Lagrangian Lie-perturbation method on the unperturbed phase-space Lagrangian (23) and the unperturbed Hamiltonian (25), we obtain (to zeroth-order in magnetic-field nonuniformity) the unperturbed relativistic guiding-center (gc) phase-space Lagrangian,

$$\Gamma_{0gc} = \left(\frac{q}{\epsilon c}\,\mathbf{A}_0 + p_\|\,\hat{\mathbf{b}}\right)\cdot d\mathbf{R} + \epsilon\,(mc/q)\,\mu\,d\theta - W\,dt, \tag{26}$$

and the unperturbed relativistic guiding-center Hamiltonian function

$$\begin{aligned} H_{0gc} &= \gamma mc^2 - W \\ &\equiv mc^2\sqrt{1 + (2\mu B_0/mc^2) + (p_\|/mc)^2} - W. \end{aligned} \tag{27}$$

In (26), $\mathbf{R}$ denotes the guiding-center position, $(p_\|, \mu, \theta)$ are the guiding-center momentum coordinates, and (W,t) are the guiding-center energy coordinate and time coordinate (we henceforth omit the subscript 0 to denote the background magnetic field). In addition, the dimensionless parameter $\epsilon \ll 1$ denotes terms of order of the ratio of the gyroradius $|\boldsymbol{\rho}|$ over the magnetic-field nonuniformity length scale L_B (see Ref. 28 for details).

The relation between the guiding-center phase-space coordinates $Z^a = (t, W, \mathbf{R}, p_\|, \mu, \theta)$ and the local phase-space coordinates $z^a = (t, w, \mathbf{x}, p_{\|0}, \mu_0, \theta_0)$ is given to lowest order in magnetic-field nonuniformity as

$$\left.\begin{aligned} w &= W \\ \mathbf{x} &= \mathbf{R} + \epsilon\,\boldsymbol{\rho} \\ p_{\|0} &= p_\| \\ \mu_0 &= \mu \\ \theta_0 &= \theta \end{aligned}\right\}, \tag{28}$$

where $\boldsymbol{\rho}(\mu,\theta) \equiv \boldsymbol{\rho}_0(\mu_0,\theta_0)$ denotes the gyroradius vector expressed in guiding-center coordinates.

The Poisson bracket corresponding to the relativistic guiding-center phase-space Lagrangian (26) is defined in terms of two arbitrary functions F and G on the guiding-center eight-dimensional phase space as

$$\begin{aligned} \{F,G\}_{gc} \equiv {} & \left(\frac{\partial F}{\partial W}\frac{\partial G}{\partial t} - \frac{\partial F}{\partial t}\frac{\partial G}{\partial W}\right) + \frac{\mathbf{B}^*}{B_\|^*}\cdot\left(\nabla F\,\frac{\partial G}{\partial p_\|} - \frac{\partial F}{\partial p_\|}\nabla G\right) \\ & - \frac{\mathbf{B}}{m\omega_B B_\|^*}\cdot\nabla F\times\nabla G + \frac{q}{mc}\left(\frac{\partial F}{\partial\theta}\frac{\partial G}{\partial\mu} - \frac{\partial F}{\partial\mu}\frac{\partial G}{\partial\theta}\right), \end{aligned} \tag{29a}$$

where the small parameter ϵ was omitted for simplicity and

$$\left.\begin{aligned} \mathbf{B}^* &\equiv \mathbf{B} + (cp_\|/q)\,\nabla\times\hat{\mathbf{b}} \\ B_\|^* &\equiv \hat{\mathbf{b}}\cdot\mathbf{B}^* = B[\,1 + (p_\|/m\omega_B)\,\hat{\mathbf{b}}\cdot\nabla\times\hat{\mathbf{b}}\,] \end{aligned}\right\}. \tag{29b}$$

We note that the Jacobian for the guiding-center transformation is $B_\|^*/B$ and thus the Jacobian for the total transformation $(t,w,\mathbf{x},\mathbf{p}) \rightarrow (t,W,\mathbf{R},p_\|,\mu,\theta)$ is $mB_\|^*$.

Using the relativistic guiding-center Hamiltonian (27) and the relativistic guiding-center Poisson bracket [(29a) and (29b)], we obtain the relativistic guiding-center equations of motion: $dZ^a/dt = \{Z^a, H_{0gc}\}_{gc}$. The relativistic guiding-center equations for the canonically conjugate coordinates (t,W) are

$$\left.\begin{aligned} dt/dt &= -\,\partial H_{0gc}/\partial W = 1 \\ dW/dt &= \partial H_{0gc}/\partial t = 0 \end{aligned}\right\}. \tag{30a}$$

The relativistic guiding-center velocity

$$\frac{d\mathbf{R}}{dt}=\frac{\partial H_{0gc}}{\partial p_\parallel}\frac{\mathbf{B}^*}{B_\parallel^*}+\frac{c\hat{\mathbf{b}}}{qB_\parallel^*}\times\nabla H_{0gc}$$

$$\equiv\frac{p_\parallel}{\gamma m}\frac{\mathbf{B}^*}{B_\parallel^*}+\frac{c\hat{\mathbf{b}}}{q\gamma B_\parallel^*}\times\mu\nabla B \tag{30b}$$

includes parallel motion ($\hat{\mathbf{b}}\cdot d\mathbf{R}/dt\equiv p_\parallel/\gamma m$) as well as the gradient-B and curvature drifts (which characterize drift motion). The relativistic guiding-center force equation for the parallel momentum

$$\frac{dp_\parallel}{dt}=-\frac{\mathbf{B}^*}{B_\parallel^*}\cdot\nabla H_{0gc}\equiv-\frac{\mu\,\mathbf{B}^*}{\gamma B_\parallel^*}\cdot\nabla B, \tag{30c}$$

shows the mirror force associated with magnetic-field nonuniformity along the magnetic field lines (which characterizes the bounce motion). We point out that (30b) and (30c) agree with the standard relativistic guiding-center equations found in Northrop[28] if we neglect phase-space volume-preserving terms (such as the difference between $B_\parallel^*$ and B); see (30e) below. Lastly, the relativistic guiding-center equations for the canonically conjugate coordinates (μ,θ) are

$$\left.\begin{array}{l} d\mu/dt=-(q/mc)\,\partial H_{0gc}/\partial\theta=0 \\ d\theta/dt=(q/mc)\,\partial H_{0gc}/\partial\mu\equiv\omega_B\,\gamma^{-1}\end{array}\right\}. \tag{30d}$$

These equations indicate that, through the asymptotic elimination of the fast gyromotion time scale, an adiabatic invariant, the relativistic magnetic moment μ, has been constructed. Hence, within relativistic guiding-center Hamiltonian theory, the relativistic magnetic moment μ is a constant of the motion. We also note that the guiding-center Hamilton equations [(30a)–(30d)] satisfy the identity

$$\frac{\partial}{\partial Z^a}\left(B_\parallel^*\frac{dZ^a}{dt}\right)\equiv 0, \tag{30e}$$

i.e., the relativistic guiding-center Hamiltonian flow is incompressible; note that the appearance of the Jacobian $B_\parallel^*$ in (30e) ensures the property of phase-space volume conservation.

IV. RELATIVISTIC HAMILTONIAN GYROCENTER THEORY

When small-amplitude electromagnetic-field fluctuations $\Phi_1(\mathbf{x},t)$ and $\mathbf{A}_1(\mathbf{x},t)$ are introduced into the relativistic guiding-center Hamiltonian theory, the guiding-center phase-space Lagrangian (26) and the guiding-center Hamiltonian (27) are perturbed,

$$\left.\begin{array}{l} \Gamma_{gc}=\Gamma_{0gc}+\epsilon_\delta\Gamma_{1gc} \\ H_{gc}=H_{0gc}+\epsilon_\delta H_{1gc}\end{array}\right\}, \tag{31}$$

where ϵ_δ is an ordering parameter associated with the amplitude of the electromagnetic field perturbations.

The first-order guiding-center phase-space Lagrangian is[20]

$$\Gamma_{1gc}=\frac{q}{c}\mathbf{A}_1(\mathbf{R}+\boldsymbol{\rho},t)\cdot(d\mathbf{R}+d\boldsymbol{\rho})\equiv\frac{q}{c}\mathbf{A}_{1gc}\cdot(d\mathbf{R}+d\boldsymbol{\rho}), \tag{32a}$$

and the first-order guiding-center Hamiltonian is

$$H_{1gc}=q\,\Phi_1(\mathbf{R}+\boldsymbol{\rho},t)\equiv q\,\Phi_{1gc}, \tag{32b}$$

where the perturbation fields Φ_1 and $\mathbf{A}_1$ are evaluated at the particle position $\mathbf{x}\equiv\mathbf{R}+\boldsymbol{\rho}$ at time t and we define $\Phi_{1gc}\equiv\Phi_1(\mathbf{R}+\boldsymbol{\rho},t)$ and $\mathbf{A}_{1gc}\equiv\mathbf{A}_1(\mathbf{R}+\boldsymbol{\rho},t)$. The subscript gc is used here to denote the fact that the fields $\Phi_1(R_\parallel,\mathbf{R}_\perp+\boldsymbol{\rho},t)$ and $\mathbf{A}_1(R_\parallel,\mathbf{R}_\perp+\boldsymbol{\rho},t)$ have acquired gyroangle-dependence through the gyroradius vector $\boldsymbol{\rho}$.

The perturbations in [(32a) and (32b)] reintroduce gyroangle dependence into the guiding-center Hamiltonian dynamics (31). To remove this fast time-scale dependence, we again use the phase-space Lagrangian Lie-perturbation method. The outcome of this analysis yields the *gyrocenter* Hamiltonian dynamics, i.e., perturbed guiding-center Hamiltonian dynamics from which the fast gyroangle dependence has been asymptotically eliminated. In what follows, we shall treat the nonuniformity in the background field only to lowest order, i.e., $d\boldsymbol{\rho}\equiv(\partial\boldsymbol{\rho}/\partial\mu)d\mu+(\partial\boldsymbol{\rho}/\partial\theta)d\theta$ in (32a).

A. Low-frequency gyrokinetic ordering

The electromagnetic-field perturbations in [(32a) and (32b)] are assumed to satisfy the following space-time scale ordering,[31]

$$\left.\begin{array}{l} \omega/\omega_c=\mathcal{O}(\epsilon_\delta) \\ k_\parallel\,\rho=\mathcal{O}(\epsilon_\delta) \\ |\mathbf{k}_\perp|\,\rho=\mathcal{O}(1)\end{array}\right\}, \tag{33}$$

where ρ and $\omega_c\equiv\omega_B/\gamma$ are a typical gyroradius and gyrofrequency for a particle of mass m and charge q, respectively, while $(\omega,k_\parallel,\mathbf{k}_\perp)$ are the characteristic frequency, parallel and perpendicular wave numbers of the electromagnetic-field fluctuations of interest. This is the usual low-frequency gyrokinetic ordering[31] in which the characteristic fluctuation time scales are long compared to the gyroperiod $2\pi/\omega_c$, the characteristic wavelengths parallel to the magnetic field $2\pi/k_\parallel$ are long compared to the typical gyroradius ρ, while the characteristic wavelengths perpendicular to the magnetic field $2\pi/|\mathbf{k}_\perp|$ are comparable to ρ.

The space-time scale ordering used for the background field, on the other hand, considers a characteristic time scale ordered at $\mathcal{O}(\epsilon_\delta^3)$ compared to the gyroperiod $2\pi/\omega_c$ and a characteristic spatial scale ordered at $\epsilon\equiv\mathcal{O}(\epsilon_\delta)$ compared to the typical gyroradius ρ. Since the present nonlinear analysis retains only terms up to second order in ϵ_δ in the Hamiltonian, the background fields are thus considered as time-independent nonuniform fields with weak spatial gradients ($\rho\,|\nabla\ln B|\ll 1$).

B. Phase-space Lagrangian Lie-perturbation method

To remove the gyroangle dependence reintroduced by the electromagnetic field perturbations in (31), we use the phase-space Lagrangian Lie-transform perturbation method[20,32] wherein a new (gyrocenter) phase-space Lagrangian Γ_{gy} and a new (gyrocenter) Hamiltonian H_{gy} are constructed. These new gyrocenter expressions are given as asymptotic expansions in powers of ϵ_δ,

$$\left.\begin{array}{l}\Gamma_{gy}=\Gamma_{0gy}+\epsilon_\delta\Gamma_{1gy}+\epsilon_\delta^2\Gamma_{2gy}+\cdots\\ H_{gy}=H_{0gy}+\epsilon_\delta H_{1gy}+\epsilon_\delta^2 H_{2gy}+\cdots\end{array}\right\}, \tag{34}$$

where the lowest-order terms $\Gamma_{0gy}\equiv\Gamma_{0gc}$ and $H_{0gy}\equiv H_{0gc}$ are given by the unperturbed guiding-center expressions (26) and (27), respectively. The first- and second-order terms for the phase-space Lagrangian are[20,32]

$$\left.\begin{array}{l}\Gamma_{1gy}=\Gamma_{1gc}-i_1\cdot\Omega_{0gc}+dS_1\\ \Gamma_{2gy}=-\,i_2\cdot\Omega_{0gc}-i_1\cdot\Omega_{1gc}+i_1\cdot d(i_1\cdot\Omega_{0gc})/2+dS_2\end{array}\right\}, \tag{35a}$$

and the first- and second-order terms for the Hamiltonian are

$$\left.\begin{array}{l}H_{1gy}=H_{1gc}-g_1\cdot dH_{0gc}\\ H_{2gy}=-\,g_2\cdot dH_{0gc}-g_1\cdot dH_{1gc}+g_1\cdot d(g_1\cdot dH_{0gc})/2\end{array}\right\}, \tag{35b}$$

where g_n and S_n are the nth-order generating vector and the phase-space gauge function, respectively. In (35a), where the elements of the two-form Ω_{gc} are defined as in (6a), we have $i_n\cdot\Omega_{gc}=g_n^a\,(\Omega_{gc})_{ab}\,dZ^b$, while in (35b) we have $g_n\cdot dH_{gc}=g_n^a\,\partial H_{gc}/\partial Z^a$. We note that, within the phase-space Lagrangian formalism, the Hamilton equations are independent of the choice of phase-space gauge functions S_n.

C. Gyrocenter phase-space Lagrangian

We choose the generating vectors g_n in [(35a) and (35b)] so that the gyrocenter phase-space Lagrangian (31) retains the guiding-center form given by (26)

$$\Gamma_{gy}\equiv\left(\frac{q}{c}\mathbf{A}+\overline{p}_\|\,\hat{\mathbf{b}}\right)\cdot d\overline{\mathbf{R}}+(mc/q)\,\overline{\mu}\,d\overline{\theta}-\overline{W}\,d\overline{t}, \tag{36}$$

i.e., $(\Gamma_{gy})_n\equiv 0$ for $n\geqslant 1$ and the gyrocenter phase-space transformation is canonical. Here $\overline{Z}^a=(\overline{t},\overline{W},\overline{\mathbf{R}},\overline{p}_\|,\overline{\mu},\overline{\theta})$ are the new gyrocenter phase-space coordinates [see (45a)–(45f) below].

Solving (35a) for $\Gamma_{1gy}\equiv 0$ and $\Gamma_{2gy}\equiv 0$ (by inverting the Lagrange Ω-matrices) yields the following expressions for the first- and second-order generating vectors,

$$g_1^a=\{S_1,\overline{Z}^a\}_{gc}+\frac{q}{c}\mathbf{A}_{1gc}\cdot\{\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}},\overline{Z}^a\}_{gc} \tag{37a}$$

and

$$g_2^a=\{S_2,\overline{Z}^a\}_{gc}+\frac{q}{2c}\mathbf{G}_1\times\mathbf{B}_{1gc}\cdot\{\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}},\overline{Z}^a\}_{gc}, \tag{37b}$$

where $\mathbf{G}_1=g_1\cdot d(\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}})\equiv\{S_1,\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}}\}_{gc}$ and $\mathbf{B}_{1gc}\equiv\overline{\nabla}\times\mathbf{A}_{1gc}$ is the perturbed magnetic field expressed in guiding-center phase-space coordinates. At this point, the gauge functions S_1 and S_2 are still arbitrary; they will be chosen next by demanding that the gyrocenter Hamiltonian H_{gy} be independent of the gyrocenter gyroangle $\overline{\theta}$.

D. Nonlinear gyrocenter Hamiltonian

The first- and second-order phase-space gauge functions S_1 and S_2 in [(37a) and (37b)] are chosen by demanding that the first- and second-order gyrocenter Hamiltonians H_{1gy} and H_{2gy} be gyroangle-independent, respectively. From (32b), (35b) and (37a), we find the expression for the first-order gyrocenter Hamiltonian

$$\begin{aligned}H_{1gy}&=-\{S_1,H_{0gy}\}_{gc}+q\left(\Phi_{1gc}-\frac{\overline{\mathbf{u}}}{\gamma c}\cdot\mathbf{A}_{1gc}\right)\\&\equiv-\{S_1,H_{0gy}\}_{gc}+K_{1gc}.\end{aligned} \tag{38}$$

The phase-space gauge functions S_n are chosen to have the following properties (for all n):

$$\langle S_n\rangle\equiv 0\quad\text{and}\quad\partial S_n/\partial\overline{W}\equiv 0, \tag{39a}$$

where $\langle\rangle$ denotes averaging with respect to $\overline{\theta}$ (with the other gyrocenter phase-space coordinates held constant). The first choice in (39a) simplifies the analysis while the second choice ensures that the time coordinate remains unaffected by the gyrocenter extended phase-space transformation. Since we want H_{1gy} to be gyroangle-independent, noting that the gyroangle-averaging operation commutes with the guiding-center Poisson bracket, it can be shown that the first-order gyrocenter Hamiltonian can be written as

$$H_{1gy}\equiv\langle K_{1gc}\rangle. \tag{39b}$$

The phase-space gauge function S_1 is determined from the equation

$$\{S_1,H_{0gy}\}_{gc}=K_{1gc}-\langle K_{1gc}\rangle\equiv\tilde{K}_{1gc}, \tag{40a}$$

which is obtained by subtracting (39b) from (38). This equation can also be written as $\tilde{K}_{1gc}\equiv\gamma^{-1}L_\tau S_1$, where the operator $L_\tau\equiv\overline{\gamma}\,\partial/\partial t+(\overline{p}_\|/m)\,\hat{\mathbf{b}}\cdot\overline{\nabla}+\omega_B\,\partial/\partial\overline{\theta}$ denotes the total proper-time derivative along unperturbed particle orbits. To lowest order in the low-frequency gyrokinetic ordering (33), the operator L_τ becomes $\omega_B\,\partial/\partial\overline{\theta}$, and the solution for S_1 is written explicitly as

$$S_1=\overline{\gamma}L_\tau^{-1}\tilde{K}_{1gc}\equiv(\overline{\gamma}/\omega_B)\int\tilde{K}_{1gc}\,d\overline{\theta}. \tag{40b}$$

We easily verify that (39a) is satisfied for $n=1$. We note that for high-frequency gyrokinetics,[22] the inverse operator L_τ^{-1} involves an integration along unperturbed particle orbits.

Next, from (35b) and [(37a)–(37b)], we find the expression for the second-order gyrocenter Hamiltonian

$$\begin{aligned}H_{2gy}=&\frac{q^2}{2c^2}\left\langle\mathbf{A}_{1gc}\cdot\left\{\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}},\frac{\overline{\mathbf{u}}}{\gamma}\right\}_{gc}\cdot\mathbf{A}_{1gc}\right\rangle\\&-\frac{1}{2}\langle\{\overline{\gamma}L_\tau^{-1}\tilde{K}_{1gc},\tilde{K}_{1gc}\}_{gc}\rangle,\end{aligned} \tag{41}$$

where the second term is evaluated only to lowest order in the low-frequency gyrokinetic ordering (33). This expression shows how ponderomotive effects enter into the low-frequency gyrokinetic formalism (the derivation of a second-order ponderomotive Hamiltonian for arbitrary frequencies is discussed in Ref. 33). The two terms in (41) contain the following relativistic corrections:

$$\left\{ \overline{\mathbf{R}} + \overline{\boldsymbol{\rho}}, \frac{\overline{\mathbf{u}}}{\overline{\gamma}} \right\}_{gc} = \frac{1}{m\overline{\gamma}} \left(\mathbf{I} - \frac{\overline{\mathbf{u}\mathbf{u}}}{\overline{\gamma}^2 c^2} \right) \tag{42a}$$

and

$$\langle \{ \overline{\gamma} L_\tau^{-1} \tilde{K}_{1gc}, \tilde{K}_{1gc} \}_{gc} \rangle = \overline{\gamma} \langle \{ L_\tau^{-1} \tilde{K}_{1gc}, \tilde{K}_{1gc} \}_{gc} \rangle - \frac{1}{\overline{\gamma} m c^2} \langle (\tilde{K}_{1gc})^2 \rangle, \tag{42b}$$

where the first terms in [(42a) and (42b)] correspond to relativistic generalizations of the classical terms ($c \to \infty$) while the second terms are relativistic corrections.

The nonlinear gyrocenter Hamiltonian function $H_{gy} = H_{0gy} + \epsilon_\delta H_{1gy} + \cdots$ can now be written as

$$H_{gy} \equiv \overline{\gamma} m c^2 + \Psi_{gy} - \overline{W}, \tag{43a}$$

where the nonlinear perturbation gyrocenter Hamiltonian is

$$\Psi_{gy} \equiv \epsilon_\delta \langle K_{1gc} \rangle - \frac{\epsilon_\delta^2}{2} \langle \{ \overline{\gamma} L_\tau^{-1} \tilde{K}_{1gc}, \tilde{K}_{1gc} \}_{gc} \rangle + \epsilon_\delta^2 \frac{q^2}{2c^2} \left\langle \mathbf{A}_{1gc} \cdot \left\{ \overline{\mathbf{R}} + \overline{\boldsymbol{\rho}}, \frac{\overline{\mathbf{u}}}{\overline{\gamma}} \right\}_{gc} \cdot \mathbf{A}_{1gc} \right\rangle. \tag{43b}$$

We note that all perturbation effects appear exclusively in the gyrocenter Hamiltonian (43a), while the gyrocenter Poisson bracket is identical to the unperturbed guiding-center Poisson bracket (29a). In the classical limit, we recover the previous nonrelativistic nonlinear gyrokinetic results[20] from (43b).

E. Gyrocenter phase-space coordinates

The new gyrocenter phase-space coordinates $\overline{Z}^a = (\overline{t}, \overline{W}, \overline{\mathbf{R}}, \overline{p}_\parallel, \overline{\mu}, \overline{\theta})$ are related to the old guiding-center phase-space coordinates $Z^a = (t, W, \mathbf{R}, p_\parallel, \mu, \theta)$ by the asymptotic expansion in powers of ϵ_δ,

$$Z^a = \overline{Z}^a - \epsilon_\delta g_1^a - \epsilon_\delta^2 \left(g_2^a - \tfrac{1}{2} g_1 \cdot d g_1^a \right) + \mathcal{O}(\epsilon_\delta^3), \tag{44}$$

where g_1^a and g_2^a are defined in (37a) and (37b). When (44) is written explicitly using [(37a) and (37b)], we find first that the time coordinate is unaffected by the gyrocenter extended phase-space transformation

$$t = \overline{t} \ (\text{to all orders in } \epsilon_\delta), \tag{45a}$$

because of the choice (39a) for the phase-space gauge functions S_n. The remaining expressions are

$$W = \overline{W} + \mathcal{O}(\epsilon_\delta^2), \tag{45b}$$

$$\mathbf{R} = \overline{\mathbf{R}} + \epsilon_\delta \frac{c\hat{\mathbf{b}}}{qB_\parallel^*} \times \left(\overline{\nabla} S_1 + \frac{q}{c} \mathbf{A}_{1gc} \right) + \epsilon_\delta \frac{\partial S_1}{\partial \overline{p}_\parallel} \hat{\mathbf{b}} + \mathcal{O}(\epsilon_\delta^2), \tag{45c}$$

$$p_\parallel = \overline{p}_\parallel - \epsilon_\delta \frac{q}{c} \mathbf{A}_{1gc} \cdot \hat{\mathbf{b}} + \mathcal{O}(\epsilon_\delta^2), \tag{45d}$$

$$\mu = \overline{\mu} - \epsilon_\delta \frac{\omega_B}{B} \left(\frac{\partial S_1}{\partial \overline{\theta}} + \frac{q}{c} \mathbf{A}_{1gc} \cdot \frac{\partial \overline{\boldsymbol{\rho}}}{\partial \overline{\theta}} \right) + \mathcal{O}(\epsilon_\delta^2), \tag{45e}$$

$$\theta = \overline{\theta} + \epsilon_\delta \frac{\omega_B}{B} \left(\frac{\partial S_1}{\partial \overline{\mu}} + \frac{q}{c} \mathbf{A}_{1gc} \cdot \frac{\partial \overline{\boldsymbol{\rho}}}{\partial \overline{\mu}} \right) + \mathcal{O}(\epsilon_\delta^2), \tag{45f}$$

where the low-frequency gyrokinetic ordering (33) was used. We note from (45c) that the gyrocenter position $\overline{\mathbf{R}}$ is shifted from the guiding-center position $\mathbf{R}$ as a result of the electromagnetic fluctuations and that this shift has both gyroangle-dependent and gyroangle-independent parts. Furthermore, the gyrocenter parallel momentum $\overline{p}_\parallel$ can be interpreted from (45d) as a canonical momentum, while (45e) shows that the gyrocenter relativistic magnetic moment $\overline{\mu}$ is constructed as an asymptotic expansion in powers of ϵ_δ which makes it into an adiabatic invariant for gyrocenter Hamiltonian dynamics [see (46d) below].

F. Nonlinear relativistic gyrocenter Hamiltonian dynamics

The nonlinear relativistic gyrocenter Hamilton equations $d\overline{Z}^a/dt = \{ \overline{Z}^a, H_{gy} \}_{gc}$ are

$$\frac{d\overline{W}}{dt} = \frac{\partial \Psi_{gy}}{\partial \overline{t}}, \tag{46a}$$

$$\frac{d\overline{\mathbf{R}}}{dt} = \left(\frac{\overline{p}_\parallel}{\overline{\gamma} m} + \frac{\partial \Psi_{gy}}{\partial \overline{p}_\parallel} \right) \frac{\mathbf{B}^*}{B_\parallel^*} + \frac{c\hat{\mathbf{b}}}{qB_\parallel^*} \times \left(\frac{\overline{\mu}}{\overline{\gamma}} \overline{\nabla} B + \overline{\nabla} \Psi_{gy} \right), \tag{46b}$$

$$\frac{d\overline{p}_\parallel}{dt} = -\frac{\mathbf{B}^*}{B_\parallel^*} \cdot \left(\frac{\overline{\mu}}{\overline{\gamma}} \overline{\nabla} B + \overline{\nabla} \Psi_{gy} \right), \tag{46c}$$

$$\frac{d\overline{\mu}}{dt} \equiv 0. \tag{46d}$$

In (46b), the term $\partial \Psi_{gy} / \partial \overline{p}_\parallel$ is associated with perturbed parallel motion and perturbed curvature drift, while the term $\overline{\nabla} \Psi_{gy}$ is associated with a perpendicular perturbed $E \times B$ flow. In (46c), the term $\mathbf{B}^* \cdot \overline{\nabla} \Psi_{gy}$ includes the effects of a parallel electric field, although the usual induction term is absent from the right side of (46c). This is due to the fact that $\overline{p}_\parallel$, as defined in (45d), is the parallel component of the gyrocenter canonical momentum; this important feature provides useful computational advantages in gyrokinetic particle simulations.[20,34] We also note that the nonlinear relativistic gyrocenter Hamilton equations [(46a)–(46c)] satisfy the incompressibility condition

$$\frac{\partial}{\partial \overline{Z}^a} \left(B_\parallel^* \frac{d\overline{Z}^a}{dt} \right) = 0. \tag{47}$$

Note that since $\overline{p}_\parallel$ is the canonical parallel momentum [see (45c)], the perturbed parallel induction term $\hat{\mathbf{b}}\cdot\partial\mathbf{A}_{1gc}/\partial t$ does not appear in the gyrocenter Hamilton equations [(46a)–(46c)].

V. NONLINEAR RELATIVISTIC GYROKINETIC VLASOV-MAXWELL EQUATIONS

A. Nonlinear relativistic gyrokinetic Vlasov equation

The nonlinear relativistic gyrokinetic Vlasov equation in gyrocenter phase space is simply

$$\{f_{gy},H_{gy}\}_{gc}=0, \tag{48a}$$

or using (36) and (29a),

$$0=\frac{\partial f_{gy}}{\partial t}+\left(\frac{\mathbf{B}^*}{B_\parallel^*}\frac{\partial H_{gy}}{\partial\overline{p}_\parallel}+\frac{c\hat{\mathbf{b}}}{qB_\parallel^*}\times\overline{\nabla}H_{gy}\right)\cdot\overline{\nabla}f_{gy}$$
$$-\frac{\mathbf{B}^*}{B_\parallel^*}\cdot\overline{\nabla}H_{gy}\frac{\partial f_{gy}}{\partial\overline{p}_\parallel}. \tag{48b}$$

By defining the new gyrocenter Hamiltonian

$$h_{gy}\equiv H_{gy}+\overline{W}=\overline{\gamma}mc^2+\Psi_{gy}, \tag{49a}$$

the nonlinear relativistic gyrokinetic Vlasov equation can also be written as

$$\frac{\partial f_{gy}}{\partial t}+\{f_{gy},h_{gy}\}_{gc}=0. \tag{49b}$$

From (48b) we recover the linear gyrokinetic Vlasov equation previously derived by Littlejohn,[21] who used the Hamiltonian Lie-perturbation method, and by Tsai, Van Dam and Chen,[22] who used the standard method of gyroangle-averaging the relativistic Vlasov equation directly.

The relationship between the gyrocenter Vlasov distribution function f_{gy}, the guiding-center Vlasov distribution function f_{gc}, and the particle Vlasov distribution function f is discussed in terms of two operators:[20] the guiding-center operator $\mathcal{T}_{gc}$ and the gyrocenter operator $\mathcal{T}_{gy}$. To lowest order in magnetic-field nonuniformity, the guiding-center operator is $\mathcal{T}_{gc}\equiv\exp(-\epsilon\boldsymbol{\rho}\cdot\nabla)$ and the relation between the guiding-center Vlasov distribution function f_{gc} and the particle Vlasov distribution function f is expressed in terms of the scalar-invariance property: $f_{gc}(\mathbf{R},p_\parallel,\mu,\theta,t)\equiv\mathcal{T}_{gc}^{-1}f(\mathbf{R},p_\parallel,\mu,\theta,t)=f(\mathbf{R}+\boldsymbol{\rho},p_\parallel,\mu,\theta,t)$.

For the gyrocenter transformation, the scalar-invariance property yields $f_{gy}\equiv\mathcal{T}_{gy}^{-1}f_{gc}$, where

$$\mathcal{T}_{gy}\equiv\exp(\epsilon_\delta g_1\cdot d+\epsilon_\delta^2 g_2\cdot d+\cdots), \tag{50a}$$

with g_1 and g_2 given in [(37a) and(37b)]. When (50a) is expanded up to second order in ϵ_δ, we find

$$f_{gy}\equiv\mathcal{T}_{gy}^{-1}f_{gc}=f_{gc}-\epsilon_\delta g_1\cdot df_{gc}$$
$$-\epsilon_\delta^2[g_2\cdot df_{gc}-\tfrac{1}{2}g_1\cdot d(g_1\cdot df_{gc})]+\mathcal{O}(\epsilon_\delta^3). \tag{50b}$$

We note that, by construction, the gyrocenter Vlasov distribution function f_{gy} is independent of the gyrocenter gyroangle $\overline{\theta}$ to all orders in ϵ_δ.

B. Low-frequency gyrokinetic Maxwell equations

The low-frequency gyrokinetic Maxwell equations are

$$-\frac{\epsilon_\delta}{4\pi}\nabla_\perp^2A_1^\alpha(\mathbf{r},t)=\sum q\int d^6\overline{Z}\,\delta^3(\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}}-\mathbf{r})$$
$$\times\left(\frac{\overline{u}^\alpha}{\overline{\gamma}c}\right)\mathcal{T}_{gy}f_{gy}(\overline{\mathbf{R}},\overline{p}_\parallel,\overline{\mu},\overline{\theta};t), \tag{51}$$

where $A_1^\alpha\equiv(\Phi_1,\mathbf{A}_1)$, $\overline{u}^\alpha\equiv(\overline{\gamma}c,\overline{\mathbf{u}})$, and we use the gauge condition $\nabla_\perp\cdot\mathbf{A}_1=0$.[35] Also in (51), we find $d^6\overline{Z}\equiv mB_\parallel^*\,d^3\overline{R}d\overline{p}_\parallel d\overline{\mu}d\overline{\theta}$, the term $\delta^3(\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}}-\mathbf{r})$ relates the particle source point $\mathbf{r}$ at which the fields are evaluated to the gyrocenter position $\overline{\mathbf{R}}$, and (ignoring terms of order ϵ_δ^2),

$$\mathcal{T}_{gy}f_{gy}\equiv f_{gy}+\epsilon_\delta\{S_1,f_{gy}\}_{gc}+\epsilon_\delta\frac{q\mathbf{A}_{1gc}}{c}\cdot\{\overline{\mathbf{R}}+\overline{\boldsymbol{\rho}},f_{gy}\}_{gc}. \tag{52}$$

Here terms of order ϵ_δ^2 are necessarily omitted in (52) to ensure energy conservation.[20,36,37]

C. Nonlinear gyrokinetic energy conservation law

The nonlinear relativistic gyrokinetic Vlasov-Maxwell equations (48b) and (51) possess the following energy conservation law:

$$\frac{d}{dt}(E_M+E_{GV})\equiv\mathcal{O}(\epsilon_\delta^3), \tag{53}$$

i.e., energy is conserved to the order considered in this work. Here, the field energy is

$$E_M\equiv\int\frac{d^3x}{8\pi}(|\epsilon_\delta\nabla_\perp\Phi_1|^2+|\mathbf{B}+\epsilon_\delta\nabla_\perp\times\mathbf{A}_1|^2), \tag{54}$$

and the gyrokinetic particle energy is

$$E_{GV}\equiv\sum\int d^6\overline{Z}\langle\mathcal{T}_{gy}^{-1}(\overline{\gamma}mc^2)\rangle f_{gy}, \tag{55}$$

where

$$\langle\mathcal{T}_{gy}^{-1}(\overline{\gamma}mc^2)\rangle\equiv h_{gy}-\epsilon_\delta q\,\langle\Phi_{1gc}\rangle+\epsilon_\delta^2 q\,\langle\{S_1,\Phi_{1gc}\}_{gc}\rangle. \tag{56}$$

We note that the energy conservation law (53) is consistent with the time-scale ordering of the background fields discussed in Sec. IV A.

We demonstrate the approximate energy conservation law (53) as follows. After some simple manipulations, the expression for $d(E_M+E_{GV})/dt$ can be written as

$$\sum\int d^6\overline{Z}\bigg[(\partial_tf_{gy})h_{gy}-\frac{\epsilon_\delta^2}{2}f_{gy}(\langle\{\partial_tS_1,\tilde{K}_{1gc}\}_{gc}\rangle$$
$$-\langle\{S_1,\partial_t\tilde{K}_{1gc}\}_{gc}\rangle)\bigg]. \tag{57}$$

The first term on the right vanishes because

$$(\partial_t f_{gy})\, h_{gy} = -\{f_{gy}, h_{gy}\}_{gc} h_{gy} \equiv -\{(f_{gy} h_{gy}), h_{gy}\}_{gc}, \tag{58a}$$

and the guiding-center Poisson bracket has the following property $\int d^6\overline{Z}\{a,b\}_{gc} \equiv 0$ for arbitrary functions a and b. The second and third terms satisfy the following identity:

$$\langle\{\partial_t S_1, \widetilde{K}_{1gc}\}_{gc}\rangle = \langle\{S_1, \partial_t \widetilde{K}_{1gc}\}_{gc}\rangle + \langle\{H_{0gy}, \{S_1, \partial_t S_1\}_{gc}\}_{gc}\rangle, \tag{58b}$$

which is obtained by using (40a) and the Jacobi identity for the gyrocenter Poisson bracket. Using the gyroangle-independence of the guiding-center Poisson bracket (29a) and the low-frequency gyrokinetic ordering (33), the last term in (58b) yields

$$\epsilon_\delta^2 \langle\{H_{gy}, \{S_1, \partial_t S_1\}_{gc}\}_{gc}\rangle = \epsilon_\delta^2 \{H_{gy}, \langle\{S_1, \partial_t S_1\}_{gc}\rangle\}_{gc} \equiv \mathcal{O}(\epsilon_\delta^3),$$

and thus we recover the energy conservation law (53) for the nonlinear relativistic gyrokinetic Vlasov-Maxwell equations. The relativistic gyrokinetic energy conservation law (53) is a simple extension of the classical gyrokinetic conservation law derived in Refs. 20, 36, and 37; note that this result requires the Maxwell equations to retain only first-order terms in ϵ_δ in (51) and (52), while the gyrocenter Hamiltonian (43a) and the gyrokinetic particle energy (55) retain second-order terms.

VI. DISCUSSION

We now summarize our work. In Sec. II, relativistic Vlasov-Maxwell equations were presented as a foundation for the remainder of the paper. In Sec. III, an outline of the derivation of relativistic Hamiltonian guiding-center theory based on the phase-space Lagrangian Lie-transform perturbation method was presented. Relativistic Hamiltonian gyrocenter theory [based on (46a)–(46d)] was constructed in Sec. IV while nonlinear relativistic gyrokinetic Vlasov-Maxwell equations [(48b) and (51)] were derived in Sec. V. These two sections contain new results which extend previous nonrelativistic nonlinear gyrokinetic work[20] and previous linear relativistic gyrokinetic work.[21,22] A nonlinear relativistic gyrokinetic energy conservation law [(53)] was also derived in Sec. V.

In future work we plan to initially obtain nonlinear relativistic drift-kinetic equations by taking the small- $(k_\perp \rho)$ limit of the gyrokinetic equations, and perhaps later consider the finite gyroradius effects. From the nonlinear drift-kinetic equations, following the work of Chen,[18] quasilinear transport equations can be obtained, presumably in a relativistic Fokker-Planck form. Such equations are expected to be very useful for evaluating, for example, the effectiveness of low-frequency waves in the transport and acceleration of radiation-belt electrons.

ACKNOWLEDGMENTS

The authors wish to thank Professor Liu Chen for useful discussions and for his constant support and encouragement. Professor Allan N. Kaufman is gratefully acknowledged for his careful reading of the manuscript and for his suggestions toward improving it. The comments of the anonymous referee are also acknowledged concerning applications of our work to the investigation of relativistic electrons in tokamak plasmas. This work was supported by National Science Foundation Grants ATM-9613800 and ATM-9901127 and partially supported by U.S. Department of Energy Contract No. DE-AC03-76SFOO098.

[1] O. J. Kwon, P. H. Diamond, F. Wagner, G. Fussmann, and the ASDEX and NI Teams, Nucl. Fusion **28**, 1931 (1988).
[2] R. Jaspers, K. H. Finken, G. Mank, F. Hoenen, J. A. Boedo, N. J. Lopes Cardozo, and F. C. Schuller, Nucl. Fusion **33**, 1775 (1993).
[3] A. J. Russo and R. B. Campbell, Nucl. Fusion **33**, 1305 (1993); R. D. Gill, *ibid.* **33**, 1613 (1993).
[4] P. B. Parks, M. N. Rosenbluth, and S. V. Putvinski, Phys. Plasmas **6**, 2523 (1999).
[5] M. Schulz and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belts* (Springer-Verlag, New York, 1974).
[6] *Particle Acceleration in Cosmic Plasmas*, edited by G. P. Zank and T. K. Gaisser, AIP Conf. Proc. No. 23 (American Institute of Physics, NY, 1991) Vol. **264.**
[7] D. N. Baker, S. Kanekal, J. B. Blake, B. Klecker, and G. Rostoker, EOS Trans. Am. Geophys. Union **75**, 401 (1994).
[8] National Space Weather Program Implementation Plan, FCM-P31-1997 (Office of the Federal Coordinator for Meteorological Services and Supporting Research, Silver Spring, Maryland, 1997).
[9] J. W. Connor and R. J. Hastie, Nucl. Fusion **15**, 415 (1995).
[10] J. R. Martin-Solis, J. D. Alvarez, R. Sanchez, and B. Esposito, Phys. Plasmas **5**, 2370 (1998); J. R. Martin-Solis, B. Esposito, R. Sanchez, and J. D. Alvarez, *ibid.* **6**, 238 (1999).
[11] R. Jaspers, N. J. Lopes Cardozo, K. H. Finken, B. C. Shokker, G. Mank, G. Fuchs, and F. C. Schuller, Phys. Rev. Lett. **72**, 4093 (1994); B. Kurzan, K.-H. Steuer, and G. Fussmann, *ibid.* **75**, 4626 (1995).
[12] M. de Rover, N. J. Lopes Cardozo, and A. Montvai, Phys. Plasmas **3**, 4468 (1996); **3**, 4478 (1996).
[13] D. N. Baker, T. I. Pulkinnen, X. Li, *et al.*, Geophys. Res. Lett. **25**, 2975 (1998).
[14] G. Rostoker, S. Skone, and D. N. Baker, Geophys. Res. Lett. **25**, 3701 (1998).
[15] M. K. Hudson, S. R. Elkington, J. G. Lyon, C. C. Goodrich, and T. J. Rosenberg, in *Sun-Earth Plasma Connections* (American Geophysical Union, Washington, DC, 1998), volume 109, pp. 171–182.
[16] A. A. Chan and M. K. Hudson, EOS Trans. Am. Geophys. Union **74**, 501 (1998).
[17] S. Bourdarie, D. Boscher, T. Beutier, J.-A. Sauvaud, and M. Blanc, J. Geophys. Res. **101**, 27171 (1996).
[18] Liu Chen, J. Geophys. Res. **104**, 2421 (1999).
[19] A. J. Brizard, Phys. Plasmas **2**, 459 (1995).
[20] A. J. Brizard, J. Plasma Phys. **41**, 541 (1989).
[21] R. G. Littlejohn, Phys. Fluids **27**, 976 (1984).
[22] S. T. Tsai, J. W. Van Dam, and L. Chen, Plasma Phys. Controlled Fusion **26**, 907 (1984).
[23] The gyroradius of a relativistic electron is approximately given as $\rho = mc^2(\gamma^2-1)^{1/2}/eB \approx 17(\gamma/B)$ m, where B is evaluated in Gauss and γ is the relativistic factor. For example, the gyroradius of a 10 MeV electron ($\gamma \approx 20$) in a 5 T tokamak magnetic field ($B = 5\times10^4$ G) is approximately 0.7 cm. At a few Earth radii ($B\sim$ a few mG), on the other hand, a 5 MeV electron ($\gamma \approx 10$) has a gyroradius in the range of 10 to 20 km.
[24] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1975), Chap. 12.
[25] R. G. Littlejohn, J. Plasma Phys. **29**, 11 (1983).
[26] R. G. Littlejohn, J. Math. Phys. **23**, 742 (1982).

[27]R. L. Liboff, *Kinetic Theory: Classical, Quantum and Relativistic Descriptions* (Prentice Hall, Englewood Cliffs, New Jersey, 1990), Chap. 6.
[28]T. G. Northrop, *The Adiabatic Motion of Charged Particles* (Wiley, New York, 1963).
[29]T. S. Hahm, Phys. Plasmas **3**, 4658 (1996).
[30]B. M. Boghosian, Ph.D. dissertation, University of California at Davis, 1987.
[31]E. A. Frieman and L. Chen, Phys. Fluids **25**, 502 (1982).
[32]J. R. Cary and R. G. Littlejohn, Ann. Phys. (N.Y.) **151**, 1 (1983).
[33]H. Ye and A. N. Kaufman, Phys. Fluids B **4**, 1735 (1992).
[34]T. S. Hahm, W. W. Lee, and A. J. Brizard, Phys. Fluids **31**, 1940 (1988).
[35]Using the low-frequency gyrokinetic ordering (33), we note that the Lorentz gauge $c^{-1}\partial_t\Phi_1+\nabla\cdot\mathbf{A}_1=0$ and the Coulomb gauge $\nabla\cdot\mathbf{A}_1=0$ both give the same low-frequency gyrokinetic gauge condition $\nabla_\perp\cdot\mathbf{A}_1=0$.
[36]D. H. E. Dubin, J. A. Krommes, C. Oberman, and W. W. Lee, Phys. Fluids **26**, 3524 (1983).
[37]A. J. Brizard, Phys. Fluids B **1**, 1381 (1989).