Energy Conserving Non-relativistic Guiding Center Mechanics and the Galilean Principle of Relativity

H. K. Wimmel
Max-Planck-Institut für Plasmaphysik, EURATOM Association, Garching

Z. Naturforsch. **38 a,** 841 – 847 (1983); received April 29, 1983

Arguments and representative examples are given that suggest that exact energy conservation and Galilei invariance are incompatible in non-relativistic guiding-center mechanics. Provided that this is true in general it also follows that exact energy conservation and Lorentz invariance are incompatible in relativistic guiding-center mechanics. It would furthermore follow that every guiding-center mechanics with exact energy conservation is a non-unique theory owing to the principle of relativity. The paper also presents a Galilei invariant guiding-center mechanics that does not conserve energy.

1. Introduction

The principle of relativity states that *the laws of physics* are the same in all inertial reference frames (IF) when expressed by the physical quantities defined relative to the IF considered. Hence, those *fundamental equations of physics* that only contain *observable* quantities and are to be *unique* must be *form-invariant* (henceforth simply called "invariant") with respect to an appropriate group of transformations that relate the quantities of one IF with those of another (e.g. Lorentz group, Galilei group). *The principle of relativity* holds independently of the existence of such a transformation group that would leave the equations under consideration form-invariant. Hence, if physical equations are given that are not so invariant they represent a *non-unique* theory with infinitely many branches of equal validity (see Section 2). Actually, this very situation prevails in much of "non-relativistic" plasma theory (see Section 2). This is usually not made very apparent because a change of the IF is seldom considered and rarely necessary. However, in the context of this paper this point has its explicit importance.

Recently, traditional non-relativistic guiding-center mechanics [1] has been greatly improved by the establishment of modified theories that are, in a specific sense, "consistent" [2, 3, 8]. This term is used to express that these theories possess exact energy theorems (in time-independent fields) and exact Liouville theorems. However, these theories, as well as traditional guiding-center mechanics, are *not Galilei invariant*. Even though these "consistent" theories can be derived from the *particle Hamiltonian* [2] or the *particle Lagrangian* [8] Galilei invariance is lost in the course of the approximations used. It has been suggested [3] that "structural reasons" will in fact prevent any of the "*consistent*" versions of non-relativistic guiding-center mechanics from being Galilei invariant. This paper is to corroborate this conjecture by representative examples and physical arguments (Sects. 3 and 4). Specifically, Sect. 3 presents a representative, Galilei invariant guiding-center mechanics that does not conserve energy in general time-independent fields. The general incompatibility of energy conservation and Galilei invariance in guiding-center mechanics is justified and analysed in Section 4. It may be added that the *relativistic* guiding-center mechanics by Morozov and Solov'ev [7] conserves energy in time-independent fields, but lacks Lorentz invariance, as we would expect. The same is true for the consistent relativistic guiding center theory of [9] that conserves energy and phase space volume at the same time.

2. Non-invariant Physical Theories

Different theories result when the same set of dynamic equations is combined with various transformation groups that are to effect transition to the other IFs. If the equations are *non-invariant* with respect to one such transformation group then the principle of relativity will provide that the resulting theory is a *non-unique* one. Such a non-uniqueness cannot always be avoided or amended. In fact, many well-established theories exist in physics that have this non-uniqueness property (see the examples

0340-4811 / 83 / 0800-0841 $ 01.3 0/0. – Please order a reprint rather than making your own copy.

below). In order for such a theory to be physically relevant its "dispersion by non-uniqueness" ought to be sufficiently small in the IFs used. Of course, this non-uniqueness dispersion is related to the approximation error committed when the non-unique theory is derived from a more exact and more fundamental theory that *is* unique and invariant. These and other points are best clarified by a simple example taken from particle dynamics.

The example taken is *non-relativistic mechanics* (NRM) of particles, visualized as an approximation to relativistic mechanics (RM). The usual equations of NRM approximately hold, *for a given mechanical system*, in a restricted class of IFs in which the particle speeds are small compared with c. The equations of NRM are, of course, *not* Lorentz invariant. One may, for a moment, imagine that their Galiei invariance was absent or unknown. One would then argue that an approximation applied to the dynamical equations (of RM) does *not by itself* alter the transformation properties of the physical observables. Hence, one would *extend* the non-relativistic equations of motion to *all other* IFs by applying the *Lorentz transformation group* to all occurring quantities. The resulting new equations would explicitly depend on the transformation velocity V. On the other hand, the principle of relativity provides that, e.g. the equations of NRM *in their usual form* (expressed in quantities appropriate to the IF considered) must be applicable in every IF of the restricted class mentioned. This creates *a 3 parametric family of branches of the theory* that possess equal validity.

Of course, the alternative procedure of employing the *Galilei transformation group* exists. The resulting theory is invariant and unique, as is well known. Yet, application of the Galilei transformations is limited to $V \ll c$, while the above Lorentz transformations admitted all values $V < c$ for the transformation velocity. Thus, the choice between the two ways of proceeding may sometimes be a matter of taste. *Examples of the first kind* are offered by several "non-relativistic" plasma theories that consist of some form of NRM together with the *full set* of *unabridged* Maxwell's equations (see, e.g. ref. [4]). The *combined* set of equations is neither Lorentz invariant nor Galilei invariant. Hence, they form a non-unique theory in the sense defined above. Other non-unique theories are given by the guiding-center theories mentioned in Section 1.

3. A Galilei Invariant Guiding-center Mechanics Lacking Energy Conservation

Be $\{\boldsymbol{x}, \mu, v_\parallel\}$ the coordinates of 5-dimension guiding-center (G.C.) phase space, with $\boldsymbol{x}$ the G.C. position in ordinary space, μ the magnetic moment, and $v_\parallel$ the component of G.C. velocity parallel to $\boldsymbol{B}$ [3]. We use the Galilei transformation, with the transformation velocity $\boldsymbol{V}$, in its usual form, viz.

$$\boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{V}t\,, \quad t' = t\,, \tag{3.1}$$

$$\boldsymbol{v}' = \boldsymbol{v} - \boldsymbol{V}\,, \quad \boldsymbol{v} \equiv \dot{\boldsymbol{x}}\,, \tag{3.2}$$

$$\dot{\boldsymbol{v}}' = \dot{\boldsymbol{v}}\,, \tag{3.3}$$

$$\mu' = \mu\,, \tag{3.4}$$

$$\boldsymbol{B}' = \boldsymbol{B}\,, \tag{3.5}$$

$$\boldsymbol{E}' = \boldsymbol{E} + \frac{1}{c}\boldsymbol{V} \times \boldsymbol{B}\,, \tag{3.6}$$

with $\boldsymbol{E}$ and $\boldsymbol{B}$ being the electromagnetic fields. The Lorentz force, viz.

$$\boldsymbol{K}_\mathrm{L} \equiv e\boldsymbol{E} + \frac{e}{c}\boldsymbol{v} \times \boldsymbol{B}\,, \tag{3.7}$$

is seen to be Galilei invariant. From $v_\parallel \equiv \boldsymbol{v} \cdot \hat{\boldsymbol{b}}$, $\hat{\boldsymbol{b}} \equiv \boldsymbol{B}/B$, it follows that

$$\dot{v}'_\parallel = \dot{v}_\parallel - \boldsymbol{V} \cdot \mathrm{d}\hat{\boldsymbol{b}}/\mathrm{d}t\,. \tag{3.8}$$

The dot in (3.3) and (3.8) indicates the ordinary time derivative, which is equivalent to the *total* time derivative in phase space, viz.

$$\frac{\mathrm{d}}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \boldsymbol{v} \cdot \nabla + \dot{v}_\parallel \frac{\partial}{\partial v_\parallel} + \dot{\mu}\frac{\partial}{\partial \mu}\,, \tag{3.9}$$

with $\dot{\mu} = 0$ henceforth. The ∇ operation is, of course, performed with μ and $v_\parallel$ kept constant. It follows that

$$\begin{aligned}\frac{\partial}{\partial t'} &= \frac{\partial}{\partial t} + \boldsymbol{V} \cdot \nabla + \left(\boldsymbol{V} \cdot \frac{\partial \hat{\boldsymbol{b}}}{\partial t'}\right)\frac{\partial}{\partial v_\parallel} \\ &= \frac{\partial}{\partial t} + \boldsymbol{V} \cdot \nabla + \left[\left(\frac{\partial}{\partial t} + \boldsymbol{V} \cdot \nabla\right)\hat{\boldsymbol{b}}\right] \cdot \boldsymbol{V}\frac{\partial}{\partial v_\parallel}\,,\end{aligned} \tag{3.10}$$

$$\nabla' = \nabla + \nabla\hat{\boldsymbol{b}} \cdot \boldsymbol{V}\frac{\partial}{\partial v_\parallel}\,, \tag{3.11}$$

$$\partial/\partial v'_\parallel = \partial/\partial v_\parallel\,, \tag{3.12}$$

$$\mathrm{d}/\mathrm{d}t' = \mathrm{d}/\mathrm{d}t\,. \tag{3.13}$$

In (3.11) the vector notation of [5] is used, i.e. ∇ operates on $\hat{\boldsymbol{b}}$ only. Considering the fields it turns out that the homogeneous Maxwell equations, viz.

$$\partial \boldsymbol{B}/\partial t = - c \nabla \times \boldsymbol{E}\,, \tag{3.14}$$

$$\nabla \cdot \boldsymbol{B} = 0\,, \tag{3.14a}$$

are Galilei invariant while the inhomogeneous Maxwell equations, which are not needed in a G.C. mechanics – as opposed to a self consistent G.C. kinetic theory – are *not* Galilei invariant (see Appendix A). The Galilei transforms of the fields, e.g. (3.5) and (3.6), deviate from the Lorentz transforms by terms of the relative magnitude $\{O(V^2/c^2) + O(VE/cB)\}$, i.e. the relative deviation is not simply $O(V^2/c^2)$. This is not specific to the field transforms; the coordinate transforms (including time) exhibit a similar behavior. Of the two *relativistic* invariants only $\boldsymbol{E} \cdot \boldsymbol{B}$ is Galilei invariant, whereas $(B^2 - E^2)$ is *not*.

In order to construct a Galilei invariant guiding-center mechanics we start by putting

$$\boldsymbol{v} = v_\parallel \hat{\boldsymbol{b}} + \boldsymbol{v}_\mathrm{D}, \tag{3.15}$$

whence

$$\dot{\boldsymbol{v}} = \dot{v}_\parallel \hat{\boldsymbol{b}} + v_\parallel (\mathrm{d}\hat{\boldsymbol{b}}/\mathrm{d}t) + (\mathrm{d}\boldsymbol{v}_\mathrm{D}/\mathrm{d}t)\,, \tag{3.16}$$

with $\boldsymbol{v}_\mathrm{D}$ a drift velocity to be determined. In what follows, drift scaling in $\varepsilon \equiv r_\mathrm{g}/L$ is used (see Appendix A of [3]), where r_g is the gyro-radius and L a characteristic macroscopic length. It follows that $\boldsymbol{v}_\mathrm{D} = O(\varepsilon)$. In order for $\boldsymbol{v}$ to be Galilei covariant and for $\boldsymbol{v}_\mathrm{D}$ to agree to order ε with conventional theory [1, 3] we put

$$\boldsymbol{v}_\mathrm{D} \equiv \boldsymbol{v}_\mathrm{E} + \boldsymbol{v}_{\nabla\mathrm{B}} + \boldsymbol{v}_{\varkappa 1}\,, \tag{3.17}$$

where

$$\boldsymbol{v}_\mathrm{E} \equiv \left(\frac{c}{B}\right) \boldsymbol{E} \times \hat{\boldsymbol{b}}\,, \tag{3.18}$$

$$\boldsymbol{v}_{\nabla\mathrm{B}} \equiv \frac{c\mu}{eB} \hat{\boldsymbol{b}} \times \nabla B\,, \tag{3.19}$$

$$\boldsymbol{v}_{\varkappa 1} \equiv \frac{1}{\Omega} \hat{\boldsymbol{b}} \times \frac{\mathrm{d}_1}{\mathrm{d}t} \boldsymbol{v}_1 = \boldsymbol{v}_\varkappa + O(\varepsilon^2)\,, \tag{3.20}$$

with $\Omega \equiv eB/mc$, and the definitions

$$\boldsymbol{v}_1 \equiv v_\parallel \hat{\boldsymbol{b}} + \boldsymbol{v}_\mathrm{E} + \boldsymbol{v}_{\nabla\mathrm{B}}\,, \tag{3.21}$$

$$\frac{\mathrm{d}_1}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \boldsymbol{v}_1 \cdot \nabla + \dot{v}_\parallel \frac{\partial}{\partial v_\parallel}\,, \tag{3.22}$$

and $\boldsymbol{v}_\varkappa$ the usual curvature drift [1, 3]. It is seen that $\boldsymbol{v}_\mathrm{D}$ and $\boldsymbol{v}_\mathrm{E}$ satisfy (3.2), while $\boldsymbol{v}_{\nabla\mathrm{B}}$ and $\boldsymbol{v}_{\varkappa 1}$ transform like *velocity increments*, i.e. they are Galilei invariant. The resulting $\boldsymbol{v}$, as defined by (3.15) to (3.22), can also be written in the form

$$\boldsymbol{v} = \boldsymbol{v}_1 + \frac{1}{\Omega} \hat{\boldsymbol{b}} \times \frac{\mathrm{d}_1}{\mathrm{d}t} \boldsymbol{v}_1\,. \tag{3.15a}$$

An expression for $\dot{v}_\parallel$ is also needed. In order for it to satisfy (3.8) and agree to leading order in ε with traditional theory [1, 3] we put

$$\dot{v}_\parallel = \hat{\boldsymbol{b}} \cdot \left(\frac{e}{m} \boldsymbol{E} - \frac{\mu}{m} \nabla B\right) + \boldsymbol{v} \cdot \frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}\,. \tag{3.23}$$

Finally, a Galilei covariant expression for the kinetic energy is

$$W_\mathrm{k} \equiv \frac{m}{2} v^2 + \mu B\,, \tag{3.24}$$

whence

$$\dot{W}_\mathrm{k} \equiv m \boldsymbol{v} \cdot \dot{\boldsymbol{v}} + \mu (\mathrm{d}B/\mathrm{d}t)\,. \tag{3.25}$$

In order to obtain the desired energy equation, or "power balance equation" [3], in terms of the fields and derivatives the r.h.s. of (3.25) must be transformed with the aid of

$$m \boldsymbol{v} \cdot \dot{\boldsymbol{v}} = m v_\parallel \dot{v}_\parallel + \frac{m}{2} \frac{\mathrm{d}}{\mathrm{d}t} (v_\mathrm{D}^2) \tag{3.26}$$

and (3.23) and (3.7). After that, by specializing to time-independent fields, we will investigate whether an energy theorem exists.

A straight-forward calculation brings (3.25) to the form

$$\begin{aligned} \dot{W}_\mathrm{K} = {} & e\boldsymbol{E} \cdot \boldsymbol{v} + \mu \frac{\partial B}{\partial t} + \frac{m}{2} \frac{\mathrm{d}}{\mathrm{d}t} (v_\mathrm{D}^2) \\ & + m v_\parallel \boldsymbol{v} \cdot \mathrm{d}\hat{\boldsymbol{b}}/\mathrm{d}t \\ & - \boldsymbol{v}_{\varkappa 1} \cdot (e\boldsymbol{E} - \mu \nabla B)\,. \end{aligned} \tag{3.27}$$

An alternative form of the last term of the r.h.s. of (3.27) is

$$- \boldsymbol{v}_\mathrm{D} \cdot \frac{\mathrm{d}_1}{\mathrm{d}t} \boldsymbol{v}_1\,. \tag{3.27a}$$

The calculation uses the identity

$$\boldsymbol{v}_0 \cdot (e\boldsymbol{E} - \mu \nabla B) = 0\,, \tag{3.28}$$

with the definition

$$\boldsymbol{v}_0 \equiv \boldsymbol{v}_\mathrm{E} + \boldsymbol{v}_{\nabla\mathrm{B}}\,. \tag{3.29}$$

In time-independent fields ($\partial/\partial t = 0$) (3.27) reduces to

$$\dot{W}_K = \frac{d}{dt}\left(-e\Phi + \frac{m}{2}v_D^2\right) + m v_\parallel \boldsymbol{v} \cdot \nabla \hat{\boldsymbol{b}} \cdot \boldsymbol{v} + \boldsymbol{v}_{\times 1} \cdot \nabla (e\Phi + \mu B), \tag{3.30}$$

where the vector notation of [5] is used, i.e. the ∇ operator in the first term of the second line only operates on $\hat{\boldsymbol{b}}$. The quantity Φ is the electric potential, i.e. $\boldsymbol{E} = -\nabla\Phi$.

In *special* time-independent field configurations, e.g. with $\nabla\hat{\boldsymbol{b}} \equiv 0$, $\boldsymbol{v}_{\times 1} \equiv 0$, $\boldsymbol{v}_D \equiv \boldsymbol{v}_0$, an energy theorem results, viz.

$$W \equiv W_K + e\Phi - \frac{m}{2}v_0^2 = \text{const} \tag{3.31}$$

along G.C. orbits. An energy theorem does *not* exist for *general* time-independent fields. The reason is the following. The second line of (3.30) cannot, for general time-independent fields, be transformed into a total time derivative of a function $\Psi(\boldsymbol{x}, \mu, v_\parallel)$, i.e. independent of t. To be sure, transformation into a total time derivative of a function $C(t, \boldsymbol{x}, \mu, v_\parallel)$ would be formally possible be employing the formal solutions of the equations of motion. But in order to construct a *first integral* of the form

$$W \equiv W_k + V(\boldsymbol{x}, \mu, v_\parallel) = \text{const} \tag{3.31 a}$$

along G.C. orbits the functions W_k, V, and Ψ may not explicitly depend on t (except for t dependences that would cancel in W, to yield $\partial W/\partial t = 0$). Two independent, informal non-existence proofs for the energy integral are carried through in Appendix B. It is seen that this result does not depend on the exact definition employed for W_k [Eq. (3.24)], as long as a Galilei covariant definition is chosen.

Summarizing this section, we have presented a Galilei invariant G.C. mechanics that does *not*, however, conserve energy in general time-independent fields. The analysis of Sect. 4 makes it probable that Galilei invariance and energy conservation are in fact incompatible in G.C. mechanics.

4. Incompatibility of Energy Conservation and Galilei Invariance in Guiding-center Mechanics

This section exemplifies the conjecture that energy conservation and Galilei invariance are *incompatible* in non-relativistic guiding-center mechanics, except for specialized field configurations. We define the kinetic energy W_k and the energy equation (i.e. for $\dot{W}_k$) such that the two expressions agree to leading order in ε with other non-relativistic guiding center theories [1, 2, 3, 8], drift ordering [3] presupposed, and such that energy is conserved in time-independent fields. The requirement of Galilei invariance then leads to a set of *generalized* G.C. equations of motion that are in conflict with the use of a G.C. drift velocity for general field configurations. This discrepancy is further discussed and interpreted as incompatibility between energy conservation and Galilei invariance in the case of G.C. mechanics.

We use again the Galilei-covariant definition

$$W_k \equiv \tfrac{1}{2} m v^2 + \mu B \tag{4.1}$$

whence (remember $\dot{\mu} = 0$):

$$\dot{W}_k \equiv m\boldsymbol{v} \cdot \frac{d\boldsymbol{v}}{dt} + \mu\frac{dB}{dt}. \tag{4.2}$$

We require the energy equation to read (see Appendix B of [3]).

$$\dot{W}_k = e\boldsymbol{E} \cdot \boldsymbol{v} + \mu(\partial B/\partial t). \tag{4.3}$$

In *time-independent fields* this yields exact conservation of energy, viz.

$$W_k + e\Phi = \text{const}, \tag{4.4}$$

with $\nabla\Phi = -\boldsymbol{E}$. Comparison of (4.2) and (4.3) yields (for arbitrary fields):

$$\boldsymbol{v} \cdot (e\boldsymbol{E} - \mu\nabla B - m(d\boldsymbol{v}/dt)) = 0. \tag{4.5}$$

Comparing the Galilei transforms of (4.2) and (4.3) furnishes the conditions

$$\frac{d\boldsymbol{v}}{dt} = \frac{e}{m}\boldsymbol{E} - \frac{\mu}{m}\nabla B + \boldsymbol{v} \times \boldsymbol{\Omega}, \tag{4.6}$$

which are nothing else than *the G.C. equations of motion required* by the above definitions and Galilei invariance. Of course, $\boldsymbol{\Omega} = (e/mc)\boldsymbol{B}$. These equations have an unusual, generalized form because they contain the full acceleration vector $d\boldsymbol{v}/dt$. Equation (4.6) implies (4.5). Up to now the above theory is Galilei invariant and, at the same time, conserves energy in time-independent fields. It may be derived from a Lagrangian, viz.

$$L \equiv \frac{e}{c}\boldsymbol{A} \cdot \dot{\boldsymbol{x}} - e\Phi - \mu B - \frac{m}{2}\dot{\boldsymbol{x}}^2,$$

which is a modified particle Lagrangian that depends on the independent variables t, $\boldsymbol{x}$, $\dot{\boldsymbol{x}}$, while μ is merely a parameter. However, it is *not* an ordinary G.C. mechanics because $\boldsymbol{v}$ does not have the form

$$\boldsymbol{v} = v_{\parallel}\hat{\boldsymbol{b}} + \boldsymbol{v}_{\mathrm{D}}(t, \boldsymbol{x}, \mu, v_{\parallel}), \tag{4.7}$$

with the *drift velocity* $\boldsymbol{v}_{\mathrm{D}}$ a *given function* of its arguments. Rather $\boldsymbol{v}$ must be determined by integrating (4.6) and introducing initial conditions. The gyro-motion of the guiding center is not yet eliminated here.

Before discussing the approximation of $\boldsymbol{v}_{\perp}$ by a drift velocity $\boldsymbol{v}_{\mathrm{D}}$ let us further consider (4.6). An equation for $\dot{v}_{\parallel}$ can be derived from it, viz.

$$\dot{v}_{\parallel} = \frac{e}{m}E_{\parallel} - \frac{\mu}{m}\frac{\partial B}{\partial s} + \boldsymbol{v}\cdot\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}, \tag{4.8}$$

which is *formally* identical to (3.23). The perpendicular components of (4.6) can be rewritten to read

$$\boldsymbol{v}_{\perp} = \boldsymbol{v}_0 + \frac{1}{\Omega}\hat{\boldsymbol{b}} \times \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}, \tag{4.9}$$

with $\boldsymbol{v}_0$ defined by (3.29). This is equivalent with the following equation for $\boldsymbol{v}$:

$$\boldsymbol{v} = \boldsymbol{v}_1 + \tau\boldsymbol{v}, \tag{4.10}$$

with $\boldsymbol{v}_1$ given by (3.21) and τ defined by

$$\tau \equiv \frac{1}{\Omega}\hat{\boldsymbol{b}} \times \frac{\mathrm{d}}{\mathrm{d}t}. \tag{4.11}$$

The operator τ contains $\boldsymbol{v}$ [see (3.9)]; hence (4.10) is a non-linear partial D. Eq. for $\boldsymbol{v}$. The guiding-center velocity $\boldsymbol{v}$ of (3.15a) does not usually satisfy (4.10).

The only known method for solving (4.6), (4.10), or (4.11) by means of a drift velocity $\boldsymbol{v}_{\mathrm{D}}$ [cf. (4.7)] is by expanding the equations in terms of the small parameter $\varepsilon \equiv r_g/L$ and *truncating the resulting series*. Galilei invariance will so be lost because (4.6), which is the condition for (4.3) to be Galilei invariant, will then only approximately be satisfied, except for specialized field configurations for which the higher-order terms in ε vanish identically.

The manner in which violation of Galilei invariance comes about in the above analysis makes it probable that this defect will exist for any "consistent" version of G.C. mechanics. Specifically, Galilei invariance of any energy equation that would guarantee energy conservation in time-independent fields hinges on exactly satisfying *the resulting equations of motion*; for these represent the very conditions of Galilei invariance of the energy equation. However, any such *resulting* equations of motion can be expected to contradict the use of a G.C. drift velocity when *general* field configurations are admitted.

Provided that energy conservation and uniqueness by Galilei invariance are generally incompatible in G.C. theories, which of the two symmetries should be preferred? It appears to us that *conservation of energy* is more important for the following reasons. When the "dispersion by non-uniqueness" is not larger than the approximation error incurred in deriving a theory, then a non-unique theory is no worse than a unique one. Furthermore, a preferred IF is defined by boundary conditions in many problems, and a change of IFs is often not required. On the other hand, energy is of practical importance *as a first integral* (in time-independent fields) when considering orbits as well as when doing kinetic theory (see [3]). We therefore recommend the use of "*consistent*" *guiding-center theories*, as given in [2, 3], even though they lack Galilei invariance and, hence, uniqueness.

5. Conclusion

We have given arguments why energy conservation and Galilei invariance can be expected to be generally incompatible in non-relativistic guiding-center mechanics. It turns out that Galilei invariance of the G.C. energy equation requires validity of a set of generalized G.C. equations of motion that are generally in conflict with the use of a G.C. drift velocity (Section 4). Energy conserving types of G.C. mechanics are therefore non-unique theories in the sense explained in Sects. 1 and 2. In addition, a Galilei invariant, energy non-conserving version of G.C. mechanics has been given in Section 3. From the above it may be inferred that energy conservation and *Lorentz invariance* will be equally incompatible in any relativistic G.C. mechanics. The "consistent" types of G.C. mechanics, as given in [2, 3, 8, 9], are preferable over any Galilei invariant (or Lorentz invariant, respectively) versions in practice as is discussed in Section 4. Independently of that, a *self-consistent, Galilei invariant, kinetic G.C. theory does not exist* because the inhomogeneous Maxwell equations are not Galilei invariant (see Appendix A).

Acknowledgement

The author thanks D. Correa Restrepo for pointing out [7] and G. O. Spies for referring him to [6].

Appendix A. Lacking Galilei Invariance of the Inhomogeneous Maxwell Equations

We assume that charge density ϱ and electric current density $\boldsymbol{j}$ are created by nonrelativistic fluids of *charged particles*. Hence these quantities Galilei transform thus:

$$\varrho' = \varrho\,, \tag{A.1}$$

$$\boldsymbol{j}' = \boldsymbol{j} - \varrho \boldsymbol{V}\,, \tag{A.2}$$

where $\boldsymbol{V}$ is again the transformation velocity. The *equation of continuity* is Galilei invariant, viz.

$$(\partial\varrho/\partial t + \nabla\cdot\boldsymbol{j})' = \partial\varrho/\partial t + \nabla\cdot\boldsymbol{j}\,. \tag{A.3}$$

According to (3.5) and (3.6) the inhomogeneous Maxwell equations Galilei transform in the following way. The Poisson equation transforms as

$$(\nabla\cdot\boldsymbol{E} - 4\pi\varrho)' - (\nabla\cdot\boldsymbol{E} - 4\pi\varrho) = -\frac{1}{c}\boldsymbol{V}\cdot(\nabla\times\boldsymbol{B})\,. \tag{A.4}$$

Hence, the Poisson equation is generally not Galilei invariant. Similarly, Ampère's law shows non-invariance, viz.

$$\begin{aligned}(\partial\boldsymbol{E}/\partial t - c\nabla\times\boldsymbol{B} + 4\pi\boldsymbol{j})' - (\partial\boldsymbol{E}/\partial t - c\nabla\times\boldsymbol{B} + 4\pi\boldsymbol{j})\\ = \boldsymbol{V}\cdot\{\nabla\boldsymbol{E} - \mathbf{1}(\nabla\cdot\boldsymbol{E}) - (1/c)\nabla\boldsymbol{B}\times\boldsymbol{V}\}\\ - \boldsymbol{V}\times(\nabla\times\boldsymbol{E})\,,\end{aligned} \tag{A.5}$$

where $\mathbf{1}$ is the unit dyad, and the vector notation of [5] is used again. In particular the operator ∇ only operates on $\boldsymbol{B}$ in the expression $\nabla\boldsymbol{B}\times\boldsymbol{V}$. This completes the proof of lacking Galilei invariance of the inhomogeneous Maxwell equations.

Appendix B. Non-Integrability of Equation (3.30)

The first line of (3.30) is a total time derivative (in time-independent fields). The second line is given by

$$\dot{W}_{k2} \equiv m v_\parallel\, \boldsymbol{v}\cdot\nabla\hat{\boldsymbol{b}}\cdot\boldsymbol{v} + \boldsymbol{v}_{\times 1}\cdot\nabla(e\Phi + \mu B)\,, \tag{B.1}$$

again for time-independent fields. Here the vector notation of [5] is used (see Sect. 3 above), the ∇ operator operates on $\boldsymbol{x}$ with μ and $v_\parallel$ kept constant, and $\boldsymbol{v}_{\times 1}$ is given by (3.20). In order to show that $\dot{W}_{k2}$ cannot be expressed, for general time-independent fields, as a total time derivative in the way explained in Sect. 3 ("non-integrability") one first notes that it would suffice to show this impossibility for one special time-independent field configuration and one particular point $\{\boldsymbol{x}, \mu, v_\parallel\}$ in phase space. If $\dot{W}_{k2}$ was "integrable" it would have the form

$$\dot{W}_{k2} = \boldsymbol{v}\cdot\nabla\Psi + \dot{v}_\parallel(\partial\Psi/\partial v_\parallel)\,, \tag{B.2}$$

with $\dot{v}_\parallel$ substituted by the r.h.s. of (3.23), and $\Psi(\boldsymbol{x}, \mu, v_\parallel)$ being a potential function defined in phase space. It can be demonstrated in several ways that (B.1) cannot assume the form of (B.2) in the general, time-independent case.

To give a first, informal proof, consider a field with $\nabla\hat{\boldsymbol{b}} = 0$, $\boldsymbol{v}_{\times 1} \neq 0$ at $\{\boldsymbol{x}_0, \mu, v_\parallel\}$, i.e. with $\boldsymbol{x} = \boldsymbol{x}_0$ fixed, but μ and $v_\parallel$ arbitrary. Then the first term on the r.h.s. of (B.1) vanishes and it suffices to consider the second term. This term does not contain $E_\parallel$, hence the term $\dot{v}_\parallel\,\partial\Psi/\partial v_\parallel$ in (B.2) must not contribute, and $\partial\Psi/\partial v_\parallel = 0$ at $\{\boldsymbol{x}_0, \mu, v_\parallel\}$ is necessary. On the other hand, it is not possible to express the second term as $\boldsymbol{v}\cdot\nabla\Psi$. Firstly, it has the form $\boldsymbol{v}_{\times 1}\cdot\nabla\Psi$ and, generally, one has $\boldsymbol{v}_{\times 1} \neq \boldsymbol{v}$ even though $\nabla\hat{\boldsymbol{b}} = 0$ at $x = x_0$. Further inspection shows that no other possibility exists to transform the second term to the form $\boldsymbol{v}\cdot\nabla\Psi$.

An alternative informal proof is the following. Be the magnetic field $\boldsymbol{B}$ chosen such that $\hat{\boldsymbol{b}}\cdot\operatorname{curl}\hat{\boldsymbol{b}} = 0$. It follows that orthogonal surfaces exist for the field of $\boldsymbol{B}$ lines [6]. Choose $\boldsymbol{E} = -\nabla\Phi$ and $\mu = \mu_0$ such that

$$\nabla_\perp(e\Phi + \mu_0 B) \equiv 0 \tag{B.3}$$

for $\mu = \mu_0$, all $\boldsymbol{x}$ and $v_\parallel$, but

$$\nabla_\parallel(e\Phi + \mu_0 B) \not\equiv 0\,. \tag{B.4}$$

This means that the function $(e\Phi + \mu_0 B)$ is constant on the orthogonal surfaces mentioned above. Then the following relations hold for $\mu = \mu_0$:

$$v_0 = 0\,,\quad \boldsymbol{v}_1 = v_\parallel\hat{\boldsymbol{b}}\,,\quad \mathrm{d}_1/\mathrm{d}t = v_\parallel\,\partial/\partial s\,,$$

$$\boldsymbol{v}_{\times 1} = \frac{v_\parallel^2}{\Omega}\hat{\boldsymbol{b}}\times\frac{\partial\hat{\boldsymbol{b}}}{\partial s}\,, \tag{B.5}$$

$$\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t} = \boldsymbol{v} \cdot \nabla \hat{\boldsymbol{b}} = v_\parallel \frac{\partial \hat{\boldsymbol{b}}}{\partial s} + \boldsymbol{v}_{\times 1} \cdot \nabla \hat{\boldsymbol{b}}\,. \tag{B.6}$$

For $\mu = \mu_0$ the second term of the r.h.s. of (B.1) vanishes, and

$$\dot{W}_{k2} = m v_\parallel \boldsymbol{v} \cdot \nabla \hat{\boldsymbol{b}} \cdot \boldsymbol{v}\,. \tag{B.7}$$

Again, this expression does *not* contain $E_\parallel$ and, hence, $\partial \Psi / \partial v_\parallel = 0$ for $\mu = \mu_0$ in (B.2). It is easy to see that the r.h.s. of (B.3) does not have the form $\boldsymbol{v} \cdot \nabla \Psi$. If it had this would imply that either

$$m v_\parallel \nabla \hat{\boldsymbol{b}} \cdot \boldsymbol{v} = \nabla \Psi_1 \tag{B.8}$$

or

$$m v_\parallel \boldsymbol{v} \cdot \nabla \hat{\boldsymbol{b}} = \nabla \Psi_2 \tag{B.9}$$

would have to hold. Here the velocity $\boldsymbol{v}$ is explicitly given by

$$\boldsymbol{v} = v_\parallel \hat{\boldsymbol{b}} + \frac{v_\parallel^2}{\Omega} \hat{\boldsymbol{b}} \times \frac{\partial \hat{\boldsymbol{b}}}{\partial s} \tag{B.10}$$

at $\mu = \mu_0$. Taking into account that $\partial \Psi_1 / \partial v_\parallel = \partial \Psi_2 / \partial v_\parallel = 0$ at $\mu = \mu_0$ it is clear that the two sides of (B.8) cannot have the same dependence on $v_\parallel$. The same is true for (B.9). Hence, (B.8) and/or (B.9) cannot be satisfied at $\mu = \mu_0$. To perform this second proof it would, of course, have been sufficient to consider only a neighborhood of $\boldsymbol{x} = \boldsymbol{x}_0$ in position space instead of all $\boldsymbol{x}$. This completes the proof of the proposition that (3.30) does not have the form of a total time-derivative in the way specified in Sect. 3 and, hence, does not yield an energy theorem.

[1] T. G. Northrop, The Adiabatic Motion of Charged Particles, Interscience, New York 1963.
[2] R. G. Littlejohn, Phys. Fluids **24,** 1730 (1981).
[3] H. K. Wimmel, Z. Naturforsch. **37 a,** 985 (1982).
[4] T. H. Stix, The Theory of Plasma Waves, McGraw-Hill, New York 1962.
[5] H. K. Wimmel, Extended Standard Vector Analysis for Plasma Physics, Report IPP 6/210, Max-Planck-Institut für Plasmaphysik, 8046 Garching, 1982. Europ. J. Physics **3,** 223 (1982).
[6] W. A. Newcomb, Ann. Physics **3,** 347 (1958).
[7] A. I. Morozov and L. S. Solov'ev, in Reviews of Plasma Physics (M. A. Leontovich, Ed.) Vol. **2,** 201 (1966).
[8] R. G. Littlejohn, Variational Principles of Guiding Center Motion, Report PPG-611, Center for Plasma Physics and Fusion Engineering, Univ. of California (UCLA), 1982; J. Plasma Phys. **29,** 111 (1983).
[9] H. K. Wimmel, Lagrangian Formulation of a Consistent Relativistic Guiding Center Theory, Report IPP 6/218, Max-Planck-Institut für Plasmaphysik, D-8046 Garching bei München, 1983; Z. Naturforsch. **38 a,** 601 (1983).