UNIVERSITY OF CALIFORNIA

Lawrence Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

THE ADIABATIC MOTION OF CHARGED PARTICLES IN ELECTROMAGNETIC FIELDS

Theodore G. Northrop

January 1961

TABLE OF CONTENTS

THE ADIABATIC MOTION OF CHARGED PARTICLES IN ELECTROMAGNETIC FIELDS

Theodore G. Northrop

Lawrence Radiation Laboratory
University of California
Berkeley, California

I. INTRODUCTION

The trajectory of a charged particle in an electromagnetic field is in general very complicated and must be obtained by a numerical integration of the equation of motion[1] $d\vec{p}/dt = \frac{e}{c}\, d\vec{r}/dt \times \vec{B}(\vec{r}, t) + e\vec{E}(\vec{r}, t) + m\vec{g}(\vec{r}, t)$, where $\vec{r}$ is the position of the particle, $\vec{p}$ is the relativistic momentum $m_0 \gamma \dot{\vec{r}}$, $\vec{B}$ and $\vec{E}$ are the magnetic and electric fields, m_0 is the rest mass, γ is m/m_0, and $\vec{g}$ is the total non-electromagnetic force per unit mass. In special cases, analytic solutions can be obtained, such as in the trivial case of the uniform static magnetic field, where the particle gyrates in a helix about $\vec{B}$. Usually such solutions are possible only because of a symmetry.

In a general field, a Taylor expansion of $\vec{r}(t)$ about the initial conditions would be practical only for short times. Suppose the particle is in an approximately uniform magnetic field (one which varies slowly in space and time -- slowly compared to the gyration radius and period); the particle motion is approximately helical. A Taylor expansion would be practical only over a fraction of a gyration period; a different approximation is needed if one wishes to follow the particle over many gyration periods. The gyration and motion parallel to the field line should be introduced explicitly into the expansion and deviations from strict helical motion treated as the perturbation. This is the so called "guiding center" or "adiabatic" approximation. The terminology "guiding center" arises because in a slowly varying field the particle moves approximately in a circle whose center drifts slowly across the lines of force and moves rapidly along the

The approximation would be of no use where the fields vary at a frequency comparable with the gyration frequency. In spite of the limitation to slowly varying fields, there are many applications of the guiding center approximation, especially to plasma physics and to particle motion in the terrestrial (Van Allen) radiation belts.

Not only the equations governing the guiding center motion, but also any approximate constants ("adiabatic invariants") of the particle motion will be useful in applications. The guiding center motion and the adiabatic invariants will be treated fully in this review.

It should be pointed out that the approximate (actually asymptotic) solution and adiabatic invariants of the charged particle motion in an electromagnetic field is a special case of a more general theory of asymptotic solutions and adiabatic invariants of differential equations of a certain type. The more general theory will be reviewed to some extent in Section IX.

II THE GUIDING CENTER MOTION

In this section we will first give an informal derivation of the guiding center motion. Although it is informal, it makes good physical sense and of course gives the correct results. There are, however, points where objections can be raised. These objections cannot be met without the more formal mathematical proofs, which will be outlined in Section III. The author believes that the derivation presented here is a good one for classroom use. It may leave the exacting reader wondering if something has been overlooked.

The derivation does not proceed by considering separately special field arrangements in which one guiding center drift or another appears alone, and then superposing the various drifts. This latter method has been frequently used[(2),(3)], but is lengthy. Instead, all the drifts will be obtained deductively at once by starting with a field of general geometry and doing a small amount of vector algebra. The method is similar to Hellwig's[(11)] but is carried out for ~~a more general situation~~ a more general situation. It will be assumed that the particle has non-relativistic energy and the non-relativistic equation of motion used as a starting point. After understanding the geometric reason for each drift, it is easy to write the modified equations for a particle with relativistic energy, at least for the case where $\vec{E}$ is small.

A. The Equation of Motion in Dimensionless Form

As a preliminary, it is worthwhile writing in dimensionless form the non-relativistic equation of motion

$$\frac{mc}{e}\left[\ddot{\vec{r}} - \vec{g}(\vec{r}, t)\right] = \dot{\vec{r}} \times \vec{B}(\vec{r}, t) + c\vec{E}(\vec{r}, t). \qquad (1)$$

Let

$$\vec{\mathcal{B}} = \vec{B}(\vec{r}, t)/B_0(t), \quad \vec{\mathcal{E}} = \frac{c}{v_0 B_0}\vec{E}(\vec{r}, t), \quad \mathcal{T} = \frac{v_0\, t}{L}, \quad \vec{\mathcal{R}} = \vec{r}/L,$$

$$\vec{\mathcal{G}} = \frac{L}{v_0^{\,2}}\vec{g},$$

where $\vec{v}_0$ is the initial velocity, $B_0(t)$ is the magnetic field at a typical point at time t, and L is a characteristic dimension or distance over which the fields change. Then the equation of motion becomes

$$\frac{mc\, v_0}{eB_0 L}\left[\frac{d^2\vec{\mathcal{R}}}{d\mathcal{T}^2} - \vec{\mathcal{G}}(\vec{\mathcal{R}},\mathcal{T})\right] = \frac{d\vec{\mathcal{R}}}{d\mathcal{T}} \times \vec{\mathcal{B}}(\vec{\mathcal{R}},\mathcal{T}) + \vec{\mathcal{E}}(\vec{\mathcal{R}},\mathcal{T}), \qquad (2)$$

Greek "nu"

with the initial conditions that at $\mathcal{T} = 0$, $\vec{\mathcal{R}} = \vec{r}_0/L$ and $\frac{d\vec{\mathcal{R}}}{d\mathcal{T}} = \hat{\nu}_0$, where $\vec{r}_0$ is the initial position of the particle and $\vec{v}_0$ equals $v_0\hat{\nu}_0$. Equation (2) is formally identical with (1) by the substitutions

$$\frac{mc\, v_0}{eB_0 L} \longrightarrow \frac{mc}{e},\ \vec{\mathcal{R}} \longrightarrow \vec{r},\ \mathcal{T} \longrightarrow t,\ \vec{\mathcal{G}} \longrightarrow \vec{g},\ \vec{\mathcal{B}} \longrightarrow \vec{B},\ \text{and}\ \vec{\mathcal{E}} \longrightarrow c\vec{E}.$$

Thus any solution of (2) gives a solution of (1) by these substitutions. Now $\frac{mcv_0}{eB_0L}$ is the ratio of the radius of gyration to the characteristic distance over which fields change and therefore is the quantity which one expects must be made smaller in order that the adiabatic approximation become more valid. But because equation (1) is formally identical with (2), we can work with (1) and just use the dimensional quantity $m/e \equiv \epsilon$ as the smallness parameter. It obviously is impossible to make the m/e of say an electron smaller in a series of experiments in order to make its behavior more adiabatic. But it really is only $\frac{mcv_0}{eB_0L}$ that must be made smaller, and this is possible by changing v_0, B_0 or L. The advantage of using m/e instead of $\frac{mcv_0}{eB_0L}$ is that one does not have to work with dimensionless equations, and results are obtained directly in dimensional form. A more complete discussion of the scaling of the equation of motion has been given in reference (5).

B. Derivation of the Guiding Center Equation of Motion

i.e. "rho"

To derive the equation of motion of the guiding center, let $\vec{r} = \vec{R} + \vec{\rho}$, where $\vec{r}$ is the instantaneous position of the particle, $\vec{R}$ is the position of the guiding center, and $\vec{\rho}$ is a vector from the guiding center to the particle (Fig. 1). The vector $\vec{\rho}$ can be given a precise definition by the equation $\vec{\rho} = -(mc/eB^2)\,\vec{B} \times (\vec{v} - c\vec{E}\times\vec{B}/B^2)$, where $\vec{E}$ and $\vec{B}$ are evaluated at $\vec{r}$. This combined with $\vec{r} = \vec{R} + \vec{\rho}$ gives a precise definition of $\vec{R}$. To lowest order in ϵ the fields can of course be evaluated at either $\vec{r}$ or $\vec{R}$, the difference being of order ϵ^2 in the equation for $\vec{\rho}$.

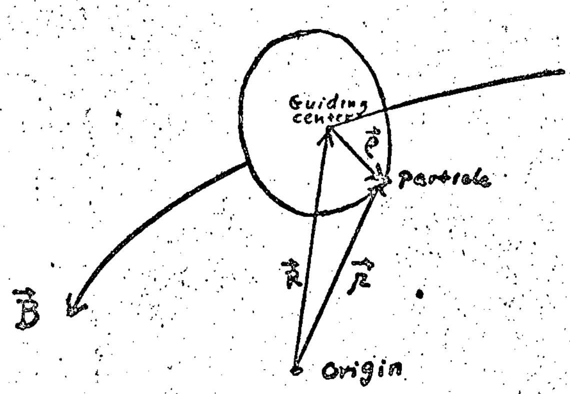


Fig. 1 The charged particle and its guiding center.

Now substitute $\vec{r} = \vec{R} + \vec{\rho}$ into Eq. (1). Since the radius of gyration is proportional to ϵ, terms containing ρ^2 can be neglected compared to those in ρ.

The result of substituting $\vec{r} = \vec{R} + \vec{\rho}$ into Eq. (1) and expanding the fields in a Taylor series about $\vec{R}$ is

$$\ddot{\vec{R}} + \ddot{\vec{\rho}} = \vec{g} + (e/m)\left\{\vec{E}(\vec{R}) + \vec{\rho}\cdot\nabla\vec{E}(\vec{R}) + (1/c)(\dot{\vec{R}}+\dot{\vec{\rho}})\times\left[\vec{B}(\vec{R}) + \vec{\rho}\cdot\nabla\vec{B}(\vec{R})\right]\right\} + \mathcal{O}(\epsilon), \tag{3}$$

where $\mathcal{O}(\epsilon)$ means terms of order ϵ. The term $(\dot{\vec{\rho}}/c)\times\vec{\rho}\cdot\nabla\vec{B}(\vec{R})$ must be retained in Eq. (3); as will become apparent shortly, this term is not of order ϵ^2 but is of order ϵ. Now define three orthogonal unit vectors: let $\hat{e}_1$ equal $\vec{B}/B$, let $\hat{e}_2$ be a unit vector directed towards the center of curvature of the line of force, and let $\hat{e}_3$ be $\hat{e}_1 \times \hat{e}_2$, a unit vector along the binormal. In order to correspond to the picture of the particle moving about a circle of radius ρ, let $\vec{\rho} = \rho(\hat{e}_2 \sin\theta + \hat{e}_3\cos\theta)$, where $\theta = \int\omega\,dt$, ω being the gyro frequency $eB(\vec{R})/mc$. Then $\dot{\vec{\rho}} = \omega\rho(\hat{e}_2\cos\theta - \hat{e}_3\sin\theta) + (\rho\hat{e}_2)^{\circ}\sin\theta + (\rho\hat{e}_3)^{\circ}\cos\theta$. The first term contains $\omega\rho$ and is of zero order in ϵ, since $\omega\sim 1/\epsilon$ and $\rho\sim\epsilon$. The second and third terms contain ρ or $\dot{\rho}$ and are of order ϵ. The reason for retaining $\dot{\vec{\rho}}\times(\vec{\rho}\cdot\nabla)\vec{B}$ in Eq. (3) is now formally apparent, since it is of order ϵ, whereas a term such as $(\vec{\rho}\cdot\nabla)^2\vec{E}$ in the Taylor expansion is of order ϵ^2. A second differentiation gives

$$\ddot{\vec{\rho}} = -\left[\omega^2\rho\,(\hat{e}_2\sin\theta + \hat{e}_3\cos\theta)\right] + \dot{\omega}\rho\left[\hat{e}_2\cos\theta - \hat{e}_3\sin\theta\right]$$
$$+ 2\omega\left[(\rho\hat{e}_2)^{\circ}\cos\theta - (\rho\hat{e}_3)^{\circ}\sin\theta\right] + \left[(\rho\hat{e}_2)^{\circ\circ}\sin\theta + (\rho\hat{e}_3)^{\circ\circ}\cos\theta\right],$$

the four terms being of order $1/\epsilon$, 1, 1, and ϵ, respectively. These expressions for $\vec{\rho}$, $\dot{\vec{\rho}}$, and $\ddot{\vec{\rho}}$ are now substituted into Eq. (3) and the resulting equation time-averaged over a gyration period, by taking $\int_0^{2\pi}(\ldots)d\theta$ and considering coefficients, such as $(\rho\hat{e}_2)^{\circ}$, to be constants. Then $\langle\vec{\rho}\rangle = \langle\dot{\vec{\rho}}\rangle = \langle\ddot{\vec{\rho}}\rangle = 0$, where the angular brackets denote the average. The result of time-averaging Eq. (3) is

$$\ddot{\vec{R}} = \vec{g}(\vec{R}) + \frac{e}{m}\left[\vec{E}(\vec{R}) + \frac{\dot{\vec{R}}}{c} \times \vec{B}(\vec{R})\right] + \frac{e}{mc}\frac{\rho^2\omega}{2}\left[\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} - \hat{e}_3 \times (\hat{e}_2 \cdot \nabla)\vec{B}\right] + \theta(\epsilon) \quad (4)$$

since

$$\langle \dot{\vec{\rho}} \times (\vec{\rho} \cdot \nabla)\vec{B} \rangle = (\rho^2\omega/2)\left[\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} - \hat{e}_3 \times (\hat{e}_2 \cdot \nabla)\vec{B}\right] \quad (5)$$

The coefficient $\rho^2\omega/2$ is an approximate invariant of the motion and is Mc/e, where M is the well-known magnetic moment. That M is an adiabatic invariant of the particle motion has been demonstrated in reference (6) and in numerous other places. The adiabatic invariants will be treated in Section III. The invariance of M is not used in deriving the guiding center equations of motion and therefore does not have to be assumed now.

The right hand side of Eq. (4) can be simplified as follows:

$$\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} = (\hat{e}_3 \times \hat{e}_1)\times(\hat{e}_3 \cdot \nabla)\vec{B} = \hat{e}_1\left[\hat{e}_3 \cdot (\hat{e}_3 \cdot \nabla)\vec{B}\right] - \hat{e}_3\left[\hat{e}_1 \cdot (\hat{e}_3 \cdot \nabla)\vec{B}\right] \quad (6)$$

Now

$$\hat{e}_1 \cdot (\hat{e}_3 \cdot \nabla)\vec{B} = \hat{e}_1 \cdot (e_3 \cdot \nabla)(\hat{e}_1 B) = (B/2)(e_3 \cdot \nabla)\hat{e}_1^{\,2} + \hat{e}_3 \cdot \nabla B = \hat{e}_3 \cdot \nabla B, \quad (7)$$

since $e_1^{\,2} = 1$. Therefore Eq. (6) becomes

$$\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} = \hat{e}_1\left[\hat{e}_3 \cdot (\hat{e}_3 \cdot \nabla)\vec{B}\right] - \hat{e}_3(\hat{e}_3 \cdot \nabla)B. \quad (8)$$

Similarly

$$\hat{e}_3 \times (\hat{e}_2 \cdot \nabla)\vec{B} = -\hat{e}_1\left[\hat{e}_2 \cdot (\hat{e}_2 \cdot \nabla)\vec{B}\right] + \hat{e}_2(\hat{e}_2 \cdot \nabla)B. \quad (9)$$

The fact that $\nabla \cdot \vec{B} = 0$ must now be used. The operator ∇ can be expressed as $\hat{e}_1(\hat{e}_1 \cdot \nabla) + \hat{e}_2(\hat{e}_2 \cdot \nabla) + \hat{e}_3(\hat{e}_3 \cdot \nabla)$, so that

$$\nabla \cdot \vec{B} = \hat{e}_1 \cdot (\hat{e}_1 \cdot \nabla)\vec{B} + \hat{e}_2 \cdot (\hat{e}_2 \cdot \nabla)\vec{B} + \hat{e}_3 \cdot (\hat{e}_3 \cdot \nabla)\vec{B} = 0. \quad (10)$$

But $\hat{e}_1 \cdot (\hat{e}_1 \cdot \nabla)\vec{B} = \hat{e}_1 \cdot \partial\vec{B}/\partial s = \partial B/\partial s$, where s is distance along the line of force. Therefore by subtracting (9) from (8) and using $\nabla \cdot \vec{B} = 0$, one obtains

$$\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} - \hat{e}_3 \times (\hat{e}_2 \cdot \nabla)\vec{B} = -\hat{e}_1(\partial B/\partial s) - \hat{e}_2(\hat{e}_2 \cdot \nabla)B - \hat{e}_3(\hat{e}_3 \cdot \nabla)B = -\nabla B \quad (11)$$

The time average of Eq. (3) then is

$$\boxed{\ddot{\vec{R}} = \vec{g}(\vec{R}) + (e/m)\left[\vec{E}(\vec{R}) + (\dot{\vec{R}}/c) \times \vec{B}(\vec{R})\right] - (M/m)\nabla B(\vec{R}) + \mathcal{O}(\epsilon),} \quad (12)$$

with an initial velocity $\dot{\vec{R}}(0)$ equal to $\left[\hat{e}_1\hat{e}_1 \cdot \vec{v} + (c\vec{E} \times \vec{B}/B^2)\right]_{t=0} + \mathcal{O}(\epsilon)$.

Equation (12) is the basic differential equation for the guiding center motion.(4) It is the same as the equation of motion of a particle in a magnetic field $\vec{B}$ and an equivalent electric field $\vec{E} - (M/e)\nabla B$, and therefore for numerical integration is more complicated than was the original equation of motion (1). If a numerical solution of (12) were performed, it would be found that the guiding center $\vec{R}$ travels in roughly a helix about the field line, just as the particle does. However, it can be shown that the radius of this helix is one order of ϵ smaller than the radius of gyration of the particle, as would be expected for the guiding center. This small amplitude oscillation of the guiding center is to be ignored, since it is of order ϵ^2 and of no significance in a first-order theory. Furthermore in the preceding analysis other terms of this order have been neglected. Equation (12) will next be solved by iteration to obtain the equations for the guiding center motion parallel and perpendicular to $\vec{B}$.

C. The Drift Velocity and the Longitudinal Equation of Motion

The differential equation for the guiding center motion can be separated into components parallel and perpendicular to $\vec{B}$. Crossing Eq. (12) on the right with $\hat{e}_1(\vec{R})$ gives the perpendicular component of the vector equation as

$$\dot{\vec{R}} - \hat{e}_1 \cdot \dot{\vec{R}} = \dot{\vec{R}}_\perp = \frac{c\vec{E}\times\hat{e}_1}{B} + \frac{Mc}{e}\,\frac{\hat{e}_1\times\nabla B}{B} + \frac{mc}{e}\,\frac{(\vec{g}-\ddot{\vec{R}})\times\hat{e}_1}{B} + \mathcal{O}(\epsilon^2) \qquad (13)$$

where $\dot{\vec{R}}_\perp$ is the component of $\dot{\vec{R}}$ perpendicular to $\hat{e}_1(\vec{R})$. It is called the drift velocity. The first term is the usual "$\vec{E} \times \vec{B}$" drift. The second term is the "gradient B" drift, and the third is the "acceleration drift". By dotting Eq. (12) with $\hat{e}_1(\vec{R})$ one obtains the scalar parallel component as

$$\frac{m}{e}\ddot{\vec{R}} \cdot \hat{e}_1 = \frac{m}{e}\vec{g} \cdot \hat{e}_1 + \vec{E} \cdot \hat{e}_1 - \frac{M}{e}\,\frac{\partial B}{\partial s} + \mathcal{O}(\epsilon^2) \qquad (14)$$

In Eq. (13) the guiding center acceleration $\ddot{\vec{R}}$ is needed to calculate the drift velocity; but because the term in which it occurs already contains ϵ as a coefficient, $\ddot{\vec{R}}$ is needed only to zero order in ϵ. It is assumed that $\ddot{\vec{R}}$ is not of negative order, such as $1/\epsilon$. If it were of negative order, the fields would change by a large amount in a gyration period when ϵ is small, and the guiding center picture would not be valid.

The acceleration $\ddot{\vec{R}} = d\dot{\vec{R}}/dt = (d/dt)(\dot{\vec{R}}_\perp + \hat{e}_1\,\dot{\vec{R}} \cdot \hat{e}_1)$, and $d\dot{\vec{R}}_\perp/dt$ can be obtained to zero order in ϵ from Eq. (13) as $d\dot{\vec{R}}_\perp/dt = (d/dt)(c\vec{E}\times\hat{e}_1/B) + \mathcal{O}(\epsilon)$. Only the first term in the drift velocity (13) is needed, since the third term is $\sim\epsilon$ and the second term contains $M/e = m(\rho\omega)^2/2eB \sim \epsilon$. If the perpendicular electric field happens to be of order ϵ instead of zero order, the retention of $c\vec{E}\times\hat{e}_1/B \equiv \vec{u}_E$ would be unnecessary in the calculation of $\ddot{\vec{R}}$. The acceleration ~~then~~ is

$$\ddot{\vec{R}} = \frac{d}{dt}\dot{\vec{R}} = \frac{d}{dt}(v_{\parallel}\hat{e}_1 + \vec{u}_E) + \mathcal{O}(\epsilon) \tag{15}$$

$$\ddot{\vec{R}} - \mathcal{O}(\epsilon) = \hat{e}_1\frac{dv_{\parallel}}{dt} + v_{\parallel}\frac{d\hat{e}_1}{dt} + \frac{d\vec{u}_E}{dt} \tag{16}$$

$$= \hat{e}_1\frac{dv_{\parallel}}{dt} + v_{\parallel}\left[\frac{\partial\hat{e}_1}{\partial t} + (\hat{e}_1 v_{\parallel} + \vec{u}_E)\cdot\nabla\hat{e}_1\right] + \left[\frac{\partial\vec{u}_E}{\partial t} + (\hat{e}_1 v_{\parallel} + \vec{u}_E)\cdot\nabla\vec{u}_E\right]$$

$$= \hat{e}_1\frac{dv_{\parallel}}{dt} + v_{\parallel}\frac{\partial e_1}{\partial t} + v_{\parallel}^2\frac{\partial\hat{e}_1}{\partial s} + v_{\parallel}\vec{u}_E\cdot\nabla\hat{e}_1 + \frac{\partial\vec{u}_E}{\partial t} + v_{\parallel}\frac{\partial\vec{u}_E}{\partial s} + \vec{u}_E\cdot\nabla\vec{u}_E$$

where $v_{\parallel}$ means $\dot{\vec{R}}\cdot\hat{e}_1(\vec{R})$. Other "parallel" velocities can be defined, such as $\vec{v}\cdot\hat{e}_1(\vec{r})$ or $\vec{v}\cdot\hat{e}_1(\vec{R})$, but here $v_{\parallel}$ always stands for $\dot{\vec{R}}\cdot\hat{e}_1(\vec{R})$. The first term is the tangential acceleration, the third is the centripetal acceleration, the second occurs in nonstatic fields, where the direction of the line of force changes with time, while the last four terms occur in the presence of a zero-order electric field. It should be stated that the presence of a "zero-order electric field" means that in a series of experiments with constant $\vec{v}_0$, $\vec{r}_0$ and in which m/e is successively reduced, the electric field is held constant. The electric field is of order ϵ if it is reduced in proportion to m/e. Whether the $\partial\hat{e}_1/\partial t$ term need be retained or not depends on how the time in which fields vary is to be scaled in the series of experiments. If the time scale is held constant, then $\partial/\partial t$ is of zero order and $\partial\hat{e}_1/\partial t$ contributes a first-order drift. If the time scale is increased in proportion to $1/\epsilon$, then $\partial/\partial t$ is of order ϵ and $\partial\hat{e}_1/\partial t$ is not needed.[7] Of course, because m/e is fixed, it is really B_0, v_0, or L that must be varied in the series of experiments. Appropriate modifications of the actual fields and time scale must be made to keep the dimensionless fields $\vec{\mathcal{B}}$, $\vec{\mathcal{E}}$, and $\vec{\mathcal{G}}$ unchanged at a given $\vec{\mathcal{R}}$ and $\mathcal{T}$. Details are given in Ref. (5).

With expression (16) for $\ddot{\vec{R}}$, Eq. (13) for the drift becomes

$$\boxed{\begin{aligned}\dot{\vec{R}}_\perp &= \frac{\hat{e}_1}{B} \times \left\{ - c\vec{E} + \frac{Mc}{e} \nabla B + \frac{mc}{e}\left(-\vec{g} + v_\parallel \frac{d\hat{e}_1}{dt} + \frac{d\vec{u}_E}{dt}\right)\right\} \\ &= \frac{\hat{e}_1}{B} \times \left\{ - c\vec{E} + \frac{Mc}{e}\nabla B + \frac{mc}{e}\left[-\vec{g} + v_\parallel \frac{\partial \hat{e}_1}{\partial t} + v_\parallel{}^2 \frac{\partial \hat{e}_1}{\partial s} + v_\parallel \vec{u}_E \cdot \nabla \hat{e}_1 \right.\right. \\ &\qquad \left.\left. + \frac{\partial}{\partial t} \vec{u}_E + v_\parallel \frac{\partial}{\partial s} \vec{u}_E + \vec{u}_E \cdot \nabla \vec{u}_E\right]\right\} + \mathcal{O}(\epsilon^2)\end{aligned}} \tag{17}$$

where $\vec{u}_E = c\vec{E}\times\hat{e}_1/B$.

The longitudinal Eq. (14) shows $E_\parallel = \vec{E} \cdot \hat{e}_1$ must be of order ϵ if $\ddot{\vec{R}}$ is to be of non-negative order. Thus in contrast to $\vec{E}_\perp$, which may be of zero order, $E_\parallel$ must be of order ϵ. If this were not so, the parallel acceleration would be $\sim 1/\epsilon$.

Equation (14) can be put in a form more useful for obtaining an energy integral by rewriting $\ddot{\vec{R}} \cdot \hat{e}_1$ as

$$\ddot{\vec{R}} \cdot \hat{e}_1 = (d/dt)(\dot{\vec{R}} \cdot \hat{e}_1) - \dot{\vec{R}} \cdot \dot{\hat{e}}_1 = (dv_\parallel/dt) - \dot{\vec{R}} \cdot \dot{\hat{e}}_1 \tag{18}$$

and noting that

$$\begin{aligned}\dot{\vec{R}} \cdot \dot{\hat{e}}_1 &= (\hat{e}_1 v_\parallel + \vec{u}_E + \mathcal{O}(\epsilon)) \cdot \dot{\hat{e}}_1 = \vec{u}_E \cdot \dot{\hat{e}}_1 + \mathcal{O}(\epsilon) \\ &= \vec{u}_E \cdot \left[(\partial \hat{e}_1/\partial t) + (\hat{e}_1 v_\parallel + \vec{u}_E) \cdot \nabla \hat{e}_1\right] + \mathcal{O}(\epsilon)\end{aligned} \tag{19}$$

In Eq's. (18) and (19) we consider only the contribution of the zero order motion to d/dt. This is all that is required, since $\vec{R} \cdot \hat{e}_1$ has ϵ for a coefficient in Eq. (14). The longitudinal Eq. (14) then becomes

$$\boxed{\begin{aligned}\frac{m}{e}\frac{dv_{\parallel}}{dt} &= \frac{m}{e} g_{\parallel} + E_{\parallel} - \frac{M}{e}\frac{\partial B}{\partial s} + \frac{m}{e}\vec{u}_E \cdot \frac{d\hat{e}_1}{dt} + \theta(\epsilon^2) \\ &= \frac{m}{e} g_{\parallel} + E_{\parallel} - \frac{M}{e}\frac{\partial B}{\partial s} + \frac{m}{e}\vec{u}_E \cdot \left(\frac{\partial \hat{e}_1}{\partial t} + v_{\parallel}\frac{\partial \hat{e}_1}{\partial s} + \vec{u}_E \cdot \nabla \hat{e}_1\right) + \theta(\epsilon^2)\end{aligned}}$$

(20)

Equations (17) and (20) are equivalent to the original differential Eq. (12).

Let us now introduce a true curvilinear coordinate system (α, β, s) such that $\alpha(\vec{r}, t)$ and $\beta(\vec{r}, t)$ are two parameters specifying a line of force and therefore constant on it; s is distance along the line as previously. For a divergence-free field such as $\vec{B}$, α and β can be chosen so that the vector potential $\vec{A}$ is $\alpha \nabla \beta$ and $\vec{B} = \nabla\alpha \times \nabla\beta$. To prove that this is possible, let $c_1(\vec{r}, t)$ and $c_2(\vec{r}, t)$ be two parameters which are constant on a line of force, but such that $\vec{A}$ is not $c_1 \nabla c_2$. The lines of force are at time t the intersections of the two families of surfaces given by $c_1(\vec{r}, t)$ equals constant and $c_2(\vec{r}, t)$ equals constant. Now $\hat{e}_1 = \dfrac{\nabla c_1 \times \nabla c_2}{|\nabla c_1 \times \nabla c_2|}$, and since $\nabla \cdot \vec{B}$ vanishes and equals $\nabla \cdot (\hat{e}_1 B)$,

we have $\nabla \cdot \left(\dfrac{\nabla c_1 \times \nabla c_2}{|\nabla c_1 \times \nabla c_2|} B\right) = (\nabla c_1 \times \nabla c_2) \cdot \nabla \dfrac{B}{|\nabla c_1 \times \nabla c_2|} = 0.$

This says that $\dfrac{B}{|\nabla c_1 \times \nabla c_2|}$ is constant along a line of force, since $\nabla c_1 \times \nabla c_2 = |\nabla c_1 \times \nabla c_2|\, \hat{e}_1$ and $\hat{e}_1 \cdot \nabla = \dfrac{\partial}{\partial s}$. Thus a general c_1, c_2 coordinate system has the property that $\dfrac{B}{|\nabla c_1 \times \nabla c_2|}$ is constant along a given line, but varies from line to line. Starting with a c_1, c_2 system for naming the lines of force, an α, β system can be obtained as follows: since $B/|\nabla c_1 \times \nabla c_2|$ is independent of s, let it equal $\xi(c_1, c_2)$. Then $\vec{B}$ equals $\xi \nabla c_1 \times \nabla c_2$. Let $\beta = c_2$ and $\alpha = \alpha(c_1, c_2)$, where the functional form is to be determined. Because $\vec{B}$ is to equal $\nabla\alpha \times \nabla\beta$ it must be true that $\xi \nabla c_1 \times \nabla c_2 = \nabla\alpha \times \nabla\beta$, or that

$$\xi \nabla c_1 \times \nabla c_2 = \left(\frac{\partial \alpha}{\partial c_1} \nabla c_1 + \frac{\partial \alpha}{\partial c_2} \nabla c_2\right) \times \nabla c_2, \text{ or that } \xi = \frac{\partial \alpha}{\partial c_1}, \text{ or}$$

finally that $\alpha = \int_{c_1} \xi(c_1, c_2) dc_1$. It is therefore possible to go from a general c_1, c_2 system to an α, β system. Because α is the integral of ξ with respect to c_1, it is clear that the α, β system for a given field is not unique; an arbitrary function of c_2 can be added to $\alpha(c_1, c_2)$. In an α, β system $\frac{|\nabla\alpha \times \nabla\beta|}{B}$ is constant not only on a line of force but everywhere, being equal to unity. This fact often simplifies the algebra, especially in connection with the adiabatic invariants.

In equation (20) the parallel electric field is given by

$$E_{\parallel} = \hat{e}_1 \cdot \left(-\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \nabla \phi\right) = -\frac{\partial}{\partial s} (\psi + \phi), \tag{21}$$

where $\phi(r, t)$ is the scalar potential for $\vec{E}$, and ψ is $\frac{\alpha}{c} \frac{\partial \beta}{\partial t}$. Multiplication of Eq. (20) by $v_{\parallel}$ and substitution of (21) for $E_{\parallel}$ gives (the $\vec{g}$ fields will be omitted from here on, but could be retained if desired)

$$\frac{d}{dt}\left(\frac{m}{2e} v_{\parallel}^2\right) = \frac{m}{e} v_{\parallel} \vec{u}_E \cdot \frac{d\hat{e}_1}{dt} - v_{\parallel} \frac{\partial}{\partial s}\left(\frac{MB}{e} + \psi + \phi\right) + \mathcal{O}(\epsilon^2). \tag{22}$$

Furthermore

$$\frac{d}{dt}\left(\frac{MB}{e}\right) = \frac{\partial}{\partial t}\left(\frac{MB}{e}\right) + v_{\parallel} \frac{\partial}{\partial s}\left(\frac{MB}{e}\right) + \vec{u}_E \cdot \nabla\left(\frac{MB}{e}\right) + \mathcal{O}(\epsilon^2) \tag{23}$$

Equation (22) can then be written as

$$\frac{d}{dt}\left(\frac{m}{2e}\ v_{\parallel}^{\ 2} + \frac{MB}{e}\right) = -v_{\parallel}\frac{\partial}{\partial s}(\psi + \phi) + \frac{\partial}{\partial t}\left(\frac{MB}{e}\right) +$$

$$\vec{u}_E \cdot \left[\nabla\left(\frac{MB}{e}\right) + \frac{m}{e} v_{\parallel} \frac{d\hat{e}_1}{dt}\right] + \mathcal{O}(\epsilon^2) . \qquad (24)$$

The rate of change of kinetic energy of the particle, averaged over a gyration period, can now be calculated. The average kinetic energy is $\frac{mv_{\parallel}^{\ 2}}{2} + MB + \frac{mu_E^{\ 2}}{2}$, since MB is the energy of rotation about the guiding center, and $\frac{mv_{\parallel}^{\ 2}}{2} + \frac{mu_E^{\ 2}}{2}$ is the kinetic energy of the guiding center motion, since $\dot{\vec{R}}$ equals $\hat{e}_1 v_{\parallel} + \vec{u}_E + \mathcal{O}(\epsilon)$. Or to give a more formal proof, start with $\vec{v} = \dot{\vec{R}} + \dot{\vec{\rho}} = \dot{\vec{R}} + \rho\omega(\hat{e}_2 \cos\omega t - \hat{e}_3 \sin\omega t) + \mathcal{O}(\epsilon)$.

"rho"

Then $\frac{m\overline{v^2}}{2e} = \frac{m}{2e}\dot{\vec{R}}^2 + \frac{m\rho^2\omega^2}{2e} = \frac{m}{2e}\left(u_E^{\ 2} + v_{\parallel}^{\ 2}\right) + \frac{MB}{e} + \mathcal{O}(\epsilon^2)$.

From Eq. (24)

$$\frac{1}{e}\ \frac{d}{dt}\left(\frac{m}{2} v_{\parallel}^{\ 2} + MB + \frac{m}{2} u_E^{\ 2}\right) =$$

$$= v_{\parallel}\frac{\partial}{\partial s}(\psi + \phi) + \frac{M}{e}\frac{\partial B}{\partial t} + \vec{u}_E \cdot \frac{M}{e}\nabla B + \frac{m}{e}\vec{u}_E \cdot v_{\parallel}\frac{d\hat{e}_1}{dt}$$

$$+ \frac{m}{2e}\ \frac{d}{dt}\ u_E^{\ 2} + \mathcal{O}(\epsilon^2). \qquad (25)$$

Now in Eq. (25)

$$\vec{u}_E \cdot \frac{M}{e}\nabla B + \frac{m}{e}\,\vec{u}_E \cdot v_\parallel \frac{d\hat{e}_1}{dt} + \frac{m}{2e}\,\frac{d}{dt}\,u_E^{\,2} = \frac{c\vec{E}\times\hat{e}_1}{B} \cdot \left(\frac{M}{e}\nabla B + \frac{m}{e} v_\parallel \frac{d\hat{e}_1}{dt} + \frac{m}{e}\frac{d\vec{u}_E}{dt} \right)$$

$$= \vec{E} \cdot \frac{\hat{e}_1}{B} \times \left(\frac{Mc}{e}\nabla B + \frac{mc}{e} v_\parallel \frac{d\hat{e}_1}{dt} + \frac{mc}{e}\frac{d\vec{u}_E}{dt} \right),$$

$$= \vec{E} \cdot \left[\frac{\hat{e}_1}{B} \times \left(- c\vec{E} + \frac{Mc}{e}\nabla B + \frac{mc}{e} v_\parallel \frac{d\hat{e}_1}{dt} + \frac{mc}{e}\frac{d\vec{u}_E}{dt} \right)\right]$$

$$= \vec{E} \cdot \dot{\vec{R}}_\perp \,. \qquad (26)$$

Finally, therefore

$$\frac{1}{e}\frac{d}{dt}\,(\text{kinetic energy}) = - v_\parallel \frac{\partial}{\partial s}(\psi + \phi) + \dot{\vec{R}}_\perp \cdot \vec{E} + \frac{M}{e}\frac{\partial B}{\partial t} + \mathcal{O}(\epsilon^2) \qquad (27)$$

$$= (\hat{e}_1 v_\parallel + \dot{\vec{R}}_\perp) \cdot \vec{E} + \frac{M}{e}\frac{\partial B}{\partial t} + \mathcal{O}(\epsilon^2)$$

or

$$\boxed{\frac{1}{e}\frac{d}{dt}\,(\text{kinetic energy}) = \dot{\vec{R}} \cdot \vec{E}\,(\vec{R}, t) + \frac{M}{e}\,\frac{\partial B(\vec{R}, t)}{\partial t} + \mathcal{O}(\epsilon^2)\,.} \qquad (28)$$

This is not a surprising result and probably could be written down directly. The term $e\dot{\vec{R}} \cdot \vec{E}$ is the rate of increase of energy due to work done by the electric field on the guiding center, while $M\frac{\partial B}{\partial t}$ is the induction effect of a time dependent field and is due to the curl of $\vec{E}$ acting about the circle of gyration. One might wonder why the second term does not have the total time derivative $\frac{dB}{dt} = \frac{\partial B}{\partial t} + v_\parallel \frac{\partial B}{\partial s} + \vec{u}_E \cdot \nabla B$. The reason is that a magnetic field gradient (as represented by $v_\parallel \frac{\partial B}{\partial s} + \vec{u}_E \cdot \nabla B$) does not change the total kinetic energy, but merely interchanges energy between the perpendicular and parallel components.

We now have the three fundamental equations (17), (20), and (28) for the guiding center motion and rate of change of energy.

If $\frac{\partial}{\partial t}$ and $\vec{E}_\perp$ (i.e., $\vec{u}_E$, ψ, and ϕ) are $\mathcal{O}(\epsilon)$ instead of $\mathcal{O}(1)$, Eq. (27) can be integrated. In that case

$$\frac{d}{dt}(\psi + \phi) = v_{\parallel} \frac{\partial(\psi + \phi)}{\partial s} + \mathcal{O}(\epsilon^2)$$

and Eq. (27) can be written as

$$\frac{d}{dt}\left(\frac{m v_{\parallel}^2}{2e} + \frac{MB}{e} + \phi + \psi \right) = 0 + \mathcal{O}(\epsilon^2), \qquad (29)$$

Since $\vec{R}_\perp \cdot \vec{E} = \mathcal{O}(\epsilon^2)$. ~~since $\vec{R}_\perp \cdot \vec{E} = \mathcal{O}(\epsilon^2)$~~. Thus $\frac{m v_{\parallel}^2}{2} + MB + e(\psi + \phi)$ is a constant of the zero order motion along a line of force. In a static field, $\psi = 0$, and (29) is just conservation of energy.

Although it does not appear generally possible to integrate Eq. (27) when $\vec{E}_\perp$ is $\mathcal{O}(1)$, there is one geometry, with $\frac{\partial}{\partial t} = 0$, for which it can be done. This is the case of a static magnetic field with rotational symmetry (such as a mirror machine), and a static $\vec{E}$, where $\vec{E}_\perp$ has no azimuthal component and $E_{\parallel} = 0$ (Fig 2). The potential ϕ is thus a constant on a line of force. Such a mirror machine has been named Ixion, and is discussed by Longmire, Nagle and Ribe.[8] The zero-order drift $\vec{u}_E$ is in the azimuthal direction; the component parallel to $\vec{B}$ of the resulting radial centrifugal force $m u_E^2/r$ has the desirable property of making it more difficult for the particle to escape at the ends.

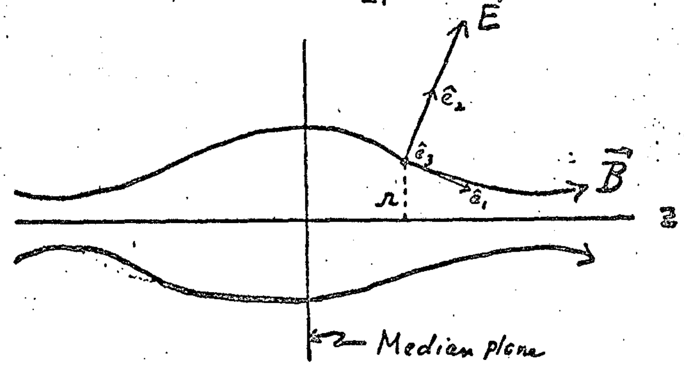


Fig. 2 Mirror machine with large electric field perpendicular to $\vec{B}$.

The effect is just that which would be observed if a bead were placed on a smooth wire bent in the shape of the line of force, and the wire then rotated about the z axis. This analogy will become apparent in the following analysis.

Under the specified restriction on the $\vec{E}$ and $\vec{B}$ fields, all terms on the right side of Eq. (27) vanish except $\dot{\vec{R}}_\perp \cdot \vec{E}$, which reduces to $\frac{mv_\parallel}{e} \vec{u}_E \cdot (\vec{u}_E \cdot \nabla) \hat{e}_1$ in this special case, and equals $\frac{mv_\parallel}{e} (cE/B)^2 \hat{e}_3 \cdot (\hat{e}_3 \cdot \nabla) \hat{e}_1$. Because $\hat{e}_3 \cdot \hat{e}_1 = 0$, the factor $\hat{e}_3 \cdot (\hat{e}_3 \cdot \nabla) \hat{e}_1$ equals $- \hat{e}_1 \cdot (\hat{e}_3 \cdot \nabla) \hat{e}_3$. But $(\hat{e}_3 \cdot \nabla) \hat{e}_3 = -\hat{e}_r/r$, where $\hat{e}_r$ is a unit vector in the radial direction. Therefore $- \hat{e}_1 \cdot (\hat{e}_3 \cdot \nabla)\hat{e}_3 = \hat{e}_1 \cdot \hat{e}_r/r$.

In order to integrate $(cE/B)^2 (\hat{e}_1 \cdot \hat{e}_r/r)$ over the zero-order motion on a flux surface (defined as the surface formed by revolving a line of force about z), the variation of cE/B and $\hat{e}_1 \cdot \hat{e}_r/r$ with longitudinal position must be known. The following is a proof that cE/rB is independent of position on a flux surface. Let $\Psi(r, z)$ be the stream function [9] for the magnetic field; the stream function has the property that Ψ = constant is the equation of a line of force and that $B_z = (1/r)\partial\Psi/\partial r$ and $B_r = -(1/r)\partial\Psi/\partial z$. Since $\vec{E}$ is perpendicular to $\vec{B}$, flux surfaces are also equipotentials and ϕ is therefore a function of Ψ. The components of electric field are

$E_r = -\partial\phi/\partial r = -(d\phi/d\Psi)\partial\Psi/\partial r$ and $E_z = -(d\phi/d\Psi)\partial\Psi/\partial z$. Thus $E = \left[(\partial\Psi/\partial r)^2 + (\partial\Psi/\partial z)^2\right]^{1/2} d\phi/d\Psi = rB(d\phi/d\Psi)$ and $cE/rB = c\,d\phi/d\Psi$, which is constant on a flux surface. The quantity cE/rB is the angular velocity of the $\vec{u}_E$ drift about z and will be denoted by Ω. Therefore $\frac{mv_{\parallel}}{e}\,\vec{u}_E\cdot(\vec{u}_E\cdot\nabla)\hat{e}_1$ is $(m/e)\,v_{\parallel}\Omega^2 r\hat{e}_1\cdot\hat{e}_r$, which equals $(m/e)(d/dt)(\Omega^2 r^2/2)$ because

$$\frac{1}{2}\frac{d}{dt}(\Omega^2 r^2) = \frac{\Omega^2}{2}\frac{dr^2}{dt} = \Omega^2 r\frac{dr}{dt} = \Omega^2 r\,(\hat{e}_1 v_{\parallel} + \vec{u}_E)\cdot\nabla r + \mathcal{O}(\epsilon)$$

$$= \Omega^2 v_{\parallel} r\,\frac{\partial r}{\partial s} = \Omega^2 v_{\parallel} r\,(\hat{e}_1\cdot\hat{e}_r) + \mathcal{O}(\epsilon). \qquad (30)$$

The integral of Eq. (27) is then

$$mv_{\parallel}^2/2 + MB - m\Omega^2 r^2/2 \text{ equals a constant of the zero-order motion on the flux surface.} \qquad (31)$$

If the subscript c denotes quantities at the median plane of Fig. 2 and e at the mirror (i.e., at the location of maximum magnetic field on the flux surface), Eq. (31) becomes

$$v_{\parallel e}^2 = v_{\parallel c}^2 + (2MB_c/m)(1 - B_e/B_c) - \Omega^2 r_c^2\,(1 - r_e^2/r_c^2). \qquad (32)$$

Therefore $v_{\parallel e}^2 \leq 0$ - i.e., the particle is contained, if

$$v_{\parallel c}^2 \leq v_{\perp c}^2\left[(B_e/B_c) - 1\right] + u_{Ec}^2\,(1 - r_e^2/r_c^2). \qquad (33)$$

If M is set equal to zero in Eq. (32) the change in parallel kinetic energy between the median plane and the mirror is $(m/2)\Omega^2(r_c^2 - r_e^2)$, which is just the work done against the centrifugal force. Thus when

$M = 0$, the problem is that of the bead sliding on the wire described previously.

Terms containing $\vec{u}_E$ in the drift Eq. (17) give a small (order ϵ) motion in or normal to a flux surface, the zero-order velocity being $\dot{\vec{R}} = \hat{e}_1 v_{\parallel} + \vec{u}_E$ in the surface. When crossed with $\hat{e}_1/B$, the third term in the square brackets is in the azimuthal direction. If $\vec{E}$ is outward as in Fig. 2, the fourth term is

$$-v_{\parallel} \frac{cE}{B} \hat{e}_3 \cdot \nabla \hat{e}_1 = -v_{\parallel}\Omega \frac{\partial \hat{e}_1}{\partial \theta} = -v_{\parallel}\Omega \frac{\partial}{\partial \theta} (\hat{e}_r\hat{e}_r \cdot \hat{e}_1 + \hat{e}_z\hat{e}_z \cdot \hat{e}_1) = -v_{\parallel}\Omega \hat{e}_3(\hat{e}_r \cdot \hat{e}_1),$$

where θ is the azimuthal angle in cylindrical coordinates and $\hat{e}_z$ is a unit vector in the z direction. When crossed with $\hat{e}_1$, this fourth term gives a drift normal to the flux surface. The sixth term in brackets is $v_{\parallel} (\partial \vec{u}_E/\partial s) = - v_{\parallel} (\partial/\partial s)(\Omega r\hat{e}_3) = -v_{\parallel}\Omega \hat{e}_3(\partial r/\partial s) = -v_{\parallel}\Omega \hat{e}_3(\hat{e}_r \cdot \hat{e}_1)$, hence is the same in this geometry as the fourth term. The last term in the square brackets is $\Omega r\hat{e}_3 \cdot \nabla (\Omega r\hat{e}_3) = \Omega^2 r^2 (\hat{e}_3 \cdot \nabla)\hat{e}_3 = -\Omega^2 r\hat{e}_r$. When crossed with $\hat{e}_1$ this last term gives another order ϵ drift in the surface, in addition to the ∇B and line curvature drifts.

Because of the two order ϵ drift terms perpendicular to the flux surface, there is an order ϵ change in ϕ (and therefore of kinetic energy) as the particle traverses the surface. This change in ϕ can be calculated directly from the product of the drift velocity normal to the surface and the electric field

$$\frac{d\phi}{dt} = 2 \frac{\hat{e}_1}{B} \times \left[-\frac{mc}{e} v_{\parallel}\Omega \hat{e}_3(\hat{e}_r \cdot \hat{e}_1)\right] \cdot (-\vec{E})$$

$$= \frac{2mc}{eB} v_{\parallel}\Omega \frac{\partial r}{\partial s} \hat{e}_1 \cdot (\hat{e}_3 \times \vec{E}) \qquad (34)$$

$$= -\frac{2mc}{eB} v_{\parallel}\Omega \frac{\partial r}{\partial s} E = -\frac{2m}{e} v_{\parallel}\Omega^2 r \frac{\partial r}{\partial s}$$

On integrating

$$\Delta(e\phi) = - m\Omega^2\Delta(r^2) = - m\Delta(u_E^2) \quad . \qquad (35)$$

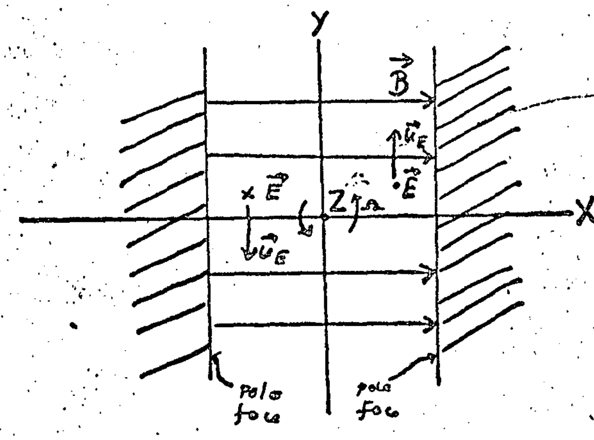
The change in $e\phi$ caused by the first-order drift off the surface equals twice the change in $(m/2)u_E^2$ as the particle moves in zero order on the surface. This result can also be obtained by energy conservation. The total average energy associated with the perpendicular motion is $MB + mu_E^2/2$. Therefore $(mv_{\parallel}^2/2) + MB + (mu_E^2/2) + e\phi$ is a constant of the zero plus first-order motion. But from Eq. (31) $(m/2)v_{\parallel}^2 + MB - (m/2)u_E^2$ is a constant of the zero-order motion. By subtraction $\Delta(e\phi) = - 2\Delta(mu_E^2/2)$.

This drift normal to the flux surface is not cumulative, since the sign of $v_{\parallel}$ reverses when the particle reflects near the mirror.

D. The Geometric Interpretation of the Drift Equation and the Longitudinal Equation of Motion

To get some physical insight into equations (17) and (20), it is instructive to look at the geometric reason for the occurrence of some of the terms. All the drift terms in Eq. (17) arise because of a variation at the gyration frequency of the curvature of the particle trajectory. This variation results in a cycloid like motion at right angles to $\vec{B}$. The reason for the variation at the gyrofrequency is different for each drift. The familiar drifts $c\vec{E} \times \hat{e}_1/B$, $\frac{mc}{eB}\vec{g} \times \hat{e}_1$, and $\frac{Mc}{eB}\hat{e}_1 \times \nabla B$ have often been illustrated in the literature(2) and will not be diagrammed here. For example, the ∇B drift occurs because B varies during a gyration period, and therefore the radius of curvature does also. The remaining drift terms in Eq. (17) come from the $\ddot{\vec{R}} \times \hat{e}_1$ term in Eq. (13), and are usually described as the result of a d'Alembertian force due to the guiding center acceleration. However, such a description does not help one's geometric understanding.

Consider, for example, a physical situation in which the $\frac{\partial \hat{e}_1}{\partial t}$ drift appears. Let a magnet with large parallel pole faces as shown in Fig. 3 be rotated about the Z-axis to give a $\partial \hat{e}_1/\partial t = \Omega \hat{y}$, where $\hat{y}$ is a unit vector along the Y axis and Ω is the magnet's angular velocity ($\Omega \ll \omega$). Because there is a $\partial \vec{B}/\partial t$ there will in general be an $\vec{E}$,



note to draftsman: $\vec{U}_E$ is the vector in the y-direction not the horizontal vector

Fig. 3 Rotating magnet gives a $\partial \hat{e}_1/\partial t$ drift in the Z direction

and therefore there will also occur the drift $\vec{u}_E = -c\, \hat{e}_1 \times \vec{E}/B$. The two terms in Eq. (17) that are proportional to $\hat{e}_1 \times \partial \vec{u}_E/\partial t$ and to $\hat{e}_1 \times \vec{u}_E \cdot \nabla \vec{u}_E$ are not obviously zero, although this will turn out to be the case. From $c\, \nabla \times \vec{E} = -\partial \vec{B}/\partial t$ one finds $\partial E_z/\partial x = \Omega B/c$ at zero time, if we take E_x and E_y zero. Then $E_z = \Omega Bx/c + E_z(x{=}0)$. Now $E_z(x{=}0)$ also equals E_z everywhere for $\Omega = 0$. Let us assume there is no $\vec{E}$ in the absence of rotation, so that $E_z(x{=}0)$ is zero. The $\vec{u}_E$ drift is

$$\vec{u}_E = c\vec{E} \times \hat{e}_1/B = x\Omega \hat{y} \text{ at zero time.} \tag{36}$$

Since $\vec{u}_E$ is independent of y, $\vec{u}_E \cdot \nabla \vec{u}_E = 0$. Also $\partial \vec{u}_E/\partial t$ is parallel to $\hat{e}_1$, so that $\hat{e}_1 \times \partial \vec{u}_E/\partial t = 0$. Thus both of these drift terms vanish leaving

$$\dot{\vec{R}}_\perp = c\vec{E} \times (\hat{e}_1/B) + v_{\parallel} (mc/eB)\, \hat{e}_1 \times \partial\hat{e}_1/\partial t = \hat{y}\Omega x + \hat{z}v_{\parallel}\Omega/\omega \, . \qquad (37)$$

The $\partial\hat{e}_1/\partial t$ drift is perpendicular to the page and of magnitude $v_{\parallel}\Omega/\omega$.

The parallel equation of motion (20) becomes

$$dv_{\parallel}/dt = \vec{u}_E \cdot \partial\hat{e}_1/\partial t = \Omega^2 x. \qquad (38)$$

This is just the centrifugal acceleration at a distance x from the axis of rotation.

The geometric reason that the curvature of the trajectory varies in the presence of $\partial\hat{e}_1/\partial t$ is that the perpendicular velocity $|\vec{v} \times \hat{e}_1|$ varies as $\hat{e}_1$ changes direction. The drift velocity can be derived (except possibly for a numerical factor) by holding $\hat{e}_1$ fixed for half a gyration period and then changing its direction for the next half period, etc. A view along the X axis of Fig. 3 will appear as in Fig. 4.

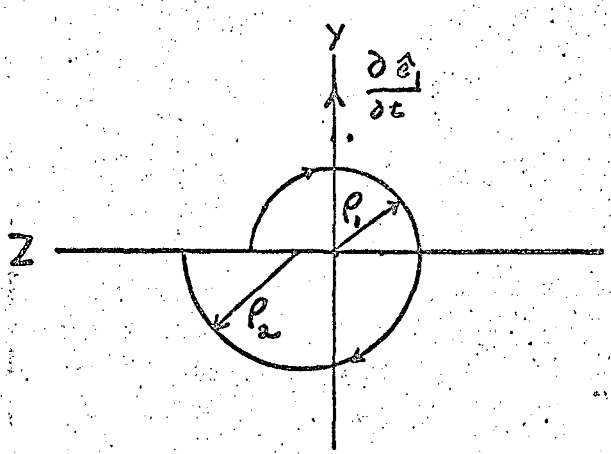


Fig. 4 Geometric explanation of the $\partial\hat{e}_1/\partial t$ drift.

Let δ be the angle between $\vec{v}$ and $\hat{e}_1$, so that $v_\perp$ is $v \sin\delta$ and $v_{\parallel}$ is $v\cos\delta$. At the end of the first half period ($y > 0$) let $\hat{e}_1$ change by $\Delta\hat{e}_1$ in the y direction. For the second half period ($y < 0$) $v_\perp$ will be changed by $\Delta v_\perp = v \cos\delta\, \Delta\delta = v_{\parallel}\Delta\delta$. The drift velocity equals the difference in the diameters of the two semicircles divided by the gyration period, or $\omega(\rho_2 - \rho_1)/\pi$. Since ρ equals $v_\perp/\omega$, $\Delta\rho$ is $\Delta v_\perp/\omega$ or $v_\parallel \Delta\delta/\omega$. And $\Delta\delta = \Omega\pi/\omega$. Thus the drift velocity equals $v_\parallel \Omega/\omega$, which in this case happens to be correct even to numerical factors.

Similar geometric derivations can be given of the other drifts containing $v_\parallel$ and $\vec{u}_E$ in Eq. (17).

In Eq. (20) the parallel acceleration of the guiding center comes from several sources. It is obvious that the parallel components of $\vec{g}$ and $\vec{E}$ will accelerate the particle along the line of force. The third term $-\frac{M}{e}\frac{\partial B}{\partial s}$ is the well-known mirror effect. It can be understood by considering the special case of a particle gyrating about the axis of symmetry of a mirror-type field, (Fig. 5). The force on the particle is $\frac{e}{c}\vec{v}\times\vec{B}(\vec{r}) = \frac{e}{c}(\hat{e}_1 v_\parallel + \vec{v}_\perp)\times(\vec{B}_r + \vec{B}_z) = \frac{e}{c}\vec{v}_\perp\times(\vec{B}_r + \vec{B}_z)$. The $\frac{e}{c}\vec{v}_\perp \times \vec{B}_z$ term is a radial force and gives the centripetal acceleration. The

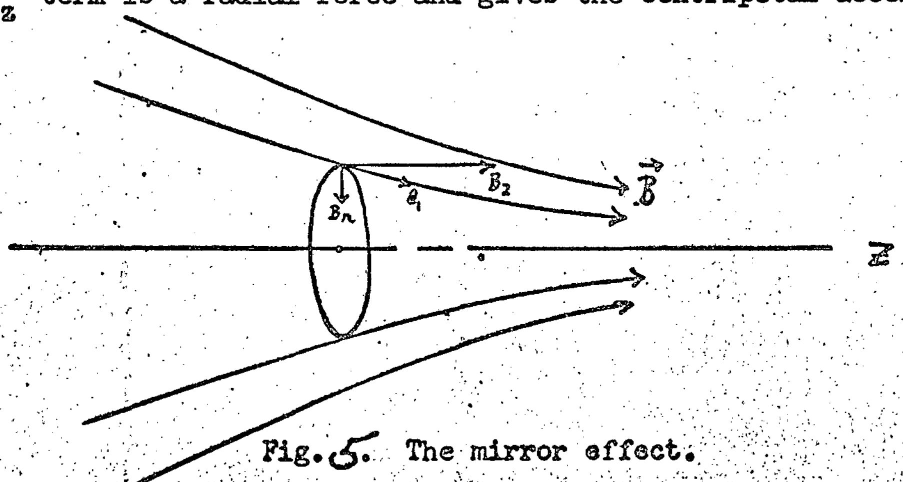


Fig. 5. The mirror effect.

term is a radial force and gives the centripetal acceleration. The

$\frac{e}{c}\vec{v}_\perp \times \vec{B}_r$ term is directed to the left and equals $-M\frac{\partial B}{\partial s}$, since the B_r at a gyration radius from the z-axis is related to $\frac{\partial B}{\partial s}$ by $\nabla \cdot \vec{B} = 0$.

The last term in Eq. (20) is $\frac{m}{e}\vec{\dot{R}} \cdot \dot{\hat{e}}_1$ and obviously arises because the parallel velocity $v_{\parallel}$ is altered not only by an acceleration $\ddot{\vec{R}}$ of the guiding center, but also by a change in direction of the $\vec{B}$ field without a change in $\dot{\vec{R}}$. A specific example of this term was seen in the illustration of the drift due to $\frac{\partial \hat{e}_1}{\partial t}$.

E. The Relativistic Case

The drifts and longitudinal equation of motion for a particle of relativistic energy can now be deduced, at least for the case where $\vec{E}_\perp$ and $\frac{\partial}{\partial t}$ are of $\mathcal{O}(\epsilon)$. To begin with, let us consider the case where $\vec{E} = 0$; the relativistic equation of motion is

$$\frac{m_0}{\sqrt{1-\rho^2}}\frac{d\vec{v}}{dt} = \frac{e\vec{v}}{c} \times \vec{B}, \qquad (39)$$

since ρ^2 is constant in the absence of $\vec{E}$. Thus the trajectory is correctly given in detail by the non-relativistic equation for a particle of mass $m_0\gamma$. Therefore the guiding center drifts are given by Eq. (17) with $\vec{E}$ set equal to zero and m replaced by $m_0\gamma$:

$$\dot{\vec{R}}_\perp = \frac{\hat{e}_1}{B} \times \left[\frac{m_0\gamma v_\perp^2 c}{2eB}\nabla B + \frac{m_0\gamma c}{e} v_\parallel^2 \frac{\partial \hat{e}_1}{\partial s}\right], \qquad (40)$$

or in terms of momenta

$$\dot{\vec{R}}_{\perp} = \frac{\hat{e}_1}{B} \times \left[\frac{M_r c}{\gamma e} \nabla B + \frac{c}{\gamma e} \frac{p_{\parallel}^{\;2}}{m_0} \frac{\partial \hat{e}_1}{\partial s} \right], \tag{41}$$

where M_r is $\frac{p_{\perp}^{\;2}}{2m_0 B}$, the relativistic magnetic moment.[10] To understand physically the relativistic effect, consider the ∇B drift term in Eq. (40). It is γ times as large as the non-relativistic expression for the same $v_{\perp}$. Relativistically the particle has γ times the mass, hence γ times the radius of gyration, and therefore samples γ times as much of the B field inhomogeneity as non-relativistically. Since the ∇B drift comes from the small difference in radii of curvature on opposite sides of the "circle" of gyration, the larger the radius, the larger the difference, and therefore the larger the drift velocity. The fact that the period of gyration is increased by γ is just compensated by the fact that the difference $\Delta\rho$ in radius of curvature on the two sides of the orbit for a given ΔB between the two sides is also γ times larger. These statements will be clarified by an order of magnitude calculation similar to that in Fig. (4) for the $\frac{\partial \hat{e}_1}{\partial t}$ drift. The radius of gyration is

$$\rho = \frac{m_0 \gamma c v_{\perp}}{eB} . \tag{42}$$

The difference in radii on the two sides is

$$\Delta\rho = - \frac{m_0 \gamma c v_{\perp}}{eB^2} \Delta B = - \frac{m_0 \gamma c v_{\perp}}{eB^2} \rho \nabla B . \tag{43}$$

$$\text{Drift velocity due to } \nabla B = \frac{|\Delta \rho|}{\text{Period}} \cong \omega |\Delta \rho| = \frac{m_0 \gamma c v_\perp}{eB^2}\, \omega \rho \nabla B$$

$$= v_\perp \frac{\rho \nabla B}{B} \tag{44}$$

and $\Delta B = \rho \nabla B$ is γ times larger relativistically than non-relativistically.

A similar argument holds for the term in Eq. (40) containing $\frac{\partial \hat{e}_1}{\partial s}$. Non-relativistically, $v_\parallel{}^2 \frac{\partial \hat{e}_1}{\partial s}$ is $v_\parallel \frac{d\hat{e}_1}{dt}$, and the explanation of the drift is similar to that given in Sect. II D for the $\frac{\partial \hat{e}_1}{\partial t}$ drift. Relativistically, the increase in mass multiplies the gyration period by γ. The difference in radii of curvature between one half period and the next is (from Eq. 42)

$$\Delta \rho = \frac{m_0 \gamma c}{eB} \Delta v_\perp \ . \tag{45}$$

Drift velocity due to $\frac{\partial \hat{e}_1}{\partial s}$ is

$$\frac{|\Delta \rho|}{\text{Period}} \cong \omega |\Delta \rho| = \Delta v_\perp$$

$$= \frac{v_\parallel}{\omega} \frac{\partial \hat{e}_1}{\partial s} , \tag{46}$$

and $\frac{1}{\omega}$ is γ times larger relativistically than non-relativistically.

The parallel equation of motion (20) in the absence of $\vec{E}$ becomes

$$m_0 \gamma \frac{dv_\parallel}{dt} = - \frac{m_0 \gamma v_\perp{}^2}{2B} \frac{\partial B}{\partial s} \tag{47}$$

or

$$\frac{dp_\parallel}{dt} = - \frac{M_r}{\gamma} \frac{\partial B}{\partial s} \tag{48}$$

From Eq. (47), the parallel force is γ times larger than non-relativistically. This can be understood from Fig. 5 and the non-relativistic explanation of the mirror effect. Relativistically, the radius of gyration is larger by γ, and therefore B_r is larger by γ at the position of the particle because of the convergence of the field lines.

If $\vec{E}$ is different from zero and if the fields are non-static, the drift and parallel equations cannot logically be obtained from the non-relativistic ones, since γ is no longer constant. One might surmise that the drift equation (41) would be modified by the addition of the term $\frac{c\vec{E}\times\hat{e}_1}{B}$, and that the longitudinal equation (48) would have the term $eE_{\parallel}$ added. This surmise is correct for the case that $\vec{E}_{\perp}$ is $\mathcal{O}(\epsilon)$, and the relativistic equations become

Eq(49) →
$$\dot{\vec{R}}_{\perp} = \frac{\hat{e}_1}{B} \times \left[-c\vec{E} + \frac{M_r c}{\gamma e} \nabla B + \frac{c}{\gamma e} \frac{p_{\parallel}^2}{m_0} \frac{\partial \hat{e}_1}{\partial s} \right] \tag{49}$$

Eq(50) →
$$\frac{dp_{\parallel}}{dt} = eE_{\parallel} - \frac{M_r}{\gamma} \frac{\partial B}{\partial s} \tag{50}$$

Eq.(51) →
$$\frac{p_{\perp}^2}{2m_0 B} = M_r = \text{constant, where } p_{\perp} = \text{perpendicular relativistic momentum} \tag{51}$$

Relativistic;) $E_{\parallel}$, $\vec{E}_{\perp}$ and $\frac{\partial}{\partial t} \sim \epsilon$

Note to editor: the label 50 with the three boxed equations (49), (50), (51)

These equations were given (without proof) in reference (20).

The more general case where $\vec{E}_{\perp}$ is $\mathcal{O}(1)$ has been studied by Vandervoort.(11) The method is necessarily covariant and is not restricted to cases where $E_{\parallel}$ is of $\mathcal{O}(\epsilon)$. The analysis will not be reviewed in detail here, but only the method indicated and the resulting guiding center equations of motion given (for the case where $E_{\parallel}$ is $\mathcal{O}(\epsilon)$). The starting point is the relativistic four-dimensional equation of motion (12)

$$\frac{d^2x_i}{d\tau^2} = F_{ik}\,\frac{dx_k}{d\tau} \,, \qquad (52)$$

where

$$F_{ik} = \frac{e}{m_0c}\begin{pmatrix} 0 & B_z & -B_y & -iE_x \\ -B_z & 0 & B_x & -iE_y \\ B_y & -B_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{pmatrix} \qquad (53)$$

x_i is the four-vector (x, y, z, ict) and τ is the proper time. The first three components of equation (52) are the relativistic vector equation $\frac{d\vec{p}}{dt} = \frac{e}{c}\frac{d\vec{r}}{dt} \times \vec{B} + e\vec{E}$, and the fourth component is the rate of change of energy $\frac{d}{dt}\left(\frac{m_0c^2}{(1-\beta^2)^{1/2}}\right) = e\vec{v}\cdot\vec{E}$, where $\vec{v} = (\dot{x}, \dot{y}, \dot{z})$. If F_{ik} is independent of x_i (i.e., fields independent of position and time), the solution of Eq. (52) exhibits a gyration in four-dimensional space at the frequency

$$\omega_\tau = \frac{e}{m_0c}\left\{\frac{B^2-E^2}{2} + \frac{1}{2}\left[(B^2-E^2)^2 + 4(E\cdot B)^2\right]^{1/2}\right\}^{1/2} \,, \qquad (54)$$

where ω_τ is the angular frequency in radians per unit proper time. The actual gyration frequency is $\omega = \omega_\tau \frac{d\tau}{dt} = \frac{\omega_\tau}{\gamma}$ radians per unit real time. When $\vec{E} = 0$, the frequency ω reduces to the usual relativistic frequency $\frac{eB}{m_0\gamma c}$.

The guiding center equations of motion are the equations of motion with the gyration at frequency ω averaged out. There are three equations, corresponding to the first three components of Eq. (52), which give the actual guiding center motion in three-space. And corresponding to the fourth component of Eq. (52) there is an equation for the average (over a gyration period) rate of increase of the total particle energy. Written in the present notation, these equations are (for $E_\parallel$ of $\mathcal{O}(\epsilon)$)

$$\dot{\vec{R}}_\perp = \frac{\hat{e}_1}{B(1-E_\perp^2/B^2)} \times \Bigg\{ -\left(1-\frac{E_\perp^2}{B^2}\right)c\vec{E} + \frac{M_r c}{\gamma e}\nabla\left[B\left(1-\frac{E_\perp^2}{B^2}\right)^{1/2}\right] + \frac{m_0 c\gamma}{e}\left(v_\parallel \frac{d\hat{e}_1}{dt} + \frac{d\vec{u}_E}{dt}\right) + \frac{v_\parallel E_\parallel}{c}\vec{u}_E + \frac{M_r}{\gamma e}\frac{\vec{u}_E}{c}\frac{\partial}{\partial t}\left[B\left(1-\frac{E_\perp^2}{B^2}\right)^{1/2}\right]\Bigg\} + \mathcal{O}(\epsilon^2) \qquad (55)$$

$$\frac{m_0 d(\gamma v_\parallel)}{dt} = \frac{dp_\parallel}{dt} = m_0\gamma\vec{u}_E \cdot \frac{d\hat{e}_1}{dt} + eE_\parallel - \frac{M_r}{\gamma}\frac{\partial}{\partial s}\left[B\left(1-\frac{E_\perp^2}{B^2}\right)^{1/2}\right] \qquad (56)$$

Relativistic: $E_\parallel = \mathcal{O}(\epsilon)$, $\vec{E}_\perp$ and $\frac{\partial}{\partial t} = \mathcal{O}(1)$.

$$\frac{p_\perp^{*2}}{2m_0 B^*} = M_r = \text{constant} \qquad (57)$$

$$\frac{d}{dt}(m_0 c^2\gamma) = e(\dot{\vec{R}}_\perp + \hat{e}_1 v_\parallel)\cdot\vec{E} + \frac{M_r}{\gamma}\frac{\partial}{\partial t}\left[B\left(1-\frac{E_\perp^2}{B^2}\right)^{1/2}\right] \qquad (58)$$

Here $p_\perp^*$ is the perpendicular momentum the particle has when observed from the frame of reference moving at $\vec{u}_E$, and B^* is the magnetic field observed in that frame. It is given by $B^* = B(1 - E_\perp^2/B^2)^{1/2} + \mathcal{O}(\epsilon)$. M_r is actually proportional to the flux through the circle of gyration. as observed in the frame of reference moving at $\vec{u}_E$. When $\vec{E}_\perp$ is $\mathcal{O}(\epsilon)$, $p_\perp^{*2}/2m_0B^*$ equals $p_\perp^2/2m_0B$ to lowest order in ϵ, and M_r is as defined previously for that case.

In Equations (55) - (58), γ oscillates at the gyrofrequency. However, this oscillation can be averaged out to give $\gamma_{avg} = \gamma^*(1-E_\perp^2/B^2)^{-1/2}$.

Equations (55) and (56) are the guiding center equations of motion in three-space, while (58) is the average rate of energy increase.

Because of the denominator $1 - E_{\perp}^2/B^2$, it is apparent that $E_\perp$ must be less than B for the equations to be valid.

Equations (55), (56), and (57) are respectively the generalizations of the relativistic equations (49), (50), and (51) to the case where $\vec{E}_\perp$ is $\mathcal{O}(1)$ instead of $\mathcal{O}(\epsilon)$. If $\vec{E}_\perp$ is $\mathcal{O}(\epsilon)$, then so are $\vec{u}_E$ and $E_\perp/B$, and equations (55) and (56) reduce to (49) and (50) upon dropping terms of order ϵ^2.

Equations (55), (56), and (58) are respectively the relativistic forms of (17), (20), and (28). A comparison of (55) and (17) shows that relativistic effects not only modify the existing terms in (17), but also introduce two new drift terms in the direction of $\hat{e}_1 \times \vec{u}_E$ — i.e., in the direction of $\vec{E}_\perp$. It is possible to prove directly that these two new drift terms are of order v^2/c^2 smaller than the others, and therefore are indeed purely relativistic effects. To show this, the relativistic equation of motion $\frac{d\vec{p}}{dt} = \frac{e}{c} \frac{d\vec{r}}{dt} \times \vec{B} + e\vec{E}$ must be written in dimensionless form. The scaling is similar to that for the non-relativistic equation. Let $\vec{\mathcal{B}} = \vec{B}(\vec{r}, t)/B_0(t)$, $\vec{\mathcal{E}} = \frac{m_0 c}{p_0 B_0} \vec{E}(\vec{r}, t)$, $\mathcal{T} = \frac{p_0 t}{m_0 L}$, and $\vec{\mathcal{R}} = \vec{r}/L$, where p_0 is the initial momentum $\frac{m_0 v_0}{(1-\beta_0^2)^{1/2}}$ and the other symbols are the same as non-relativistically. In terms of these dimensionless quantities, the equation of motion becomes

$$\frac{p_0 c}{eB_0 L} \frac{d}{d\mathcal{T}} \left\{ \frac{d\vec{\mathcal{R}}}{d\mathcal{T}} \left[1 - \left(\frac{p_0}{m_0 c} \frac{d\vec{\mathcal{R}}}{d\mathcal{T}} \right)^2 \right]^{-1/2} \right\} = \frac{d\vec{\mathcal{R}}}{d\mathcal{T}} \times \vec{\mathcal{B}} + \vec{\mathcal{E}}, \qquad (59)$$

with the initial conditions that at $\mathcal{T} = 0$, $\vec{\mathcal{R}} = \vec{r}_0/L$, $\frac{d\vec{\mathcal{R}}}{d\mathcal{T}} = \hat{v}_0 \left[1 + \left(\frac{p_0}{m_0 c} \right)^2 \right]^{-1/2} = \hat{v}_0 \times$ rest energy/initial total energy. The problem now contains the two dimensionless parameters $p_0 c/eB_0 L$ and $p_0/m_0 c$, whereas the non-relativistic problem contained just the first one. The new parameter $p_0/m_0 c$ is $m_0 \gamma_0 v_0/c = v_0/c + \mathcal{O}(v_0^3/c^3)$. An order of magnitude comparison of say

$$\frac{M_r}{e}\ \frac{\vec{u}_E}{c}\ \frac{\partial}{\partial t}\left[B(1-\frac{E_\perp^2}{B^2})^{1/2}\right] \quad \text{with the term} \quad \frac{M_r c}{e}\ \nabla\left[B(1-\frac{E_\perp^2}{B^2})^{1/2}\right]$$

gives

$$\frac{M_r}{e}\ \frac{u_E}{c}\ \frac{\partial}{\partial t}\left[B(1-E_\perp^2/B^2)^{1/2}\right] \Big/ \frac{M_r c}{\gamma e}\ \nabla\left[B(1-E_\perp^2/B^2)^{1/2}\right]$$

$$\sim \frac{u_E}{c^2 t} \sim \frac{E}{cBt}\ . \tag{60}$$

But E/cBt equals $(P_0/m_0 c)^2 \frac{\mathcal{E}}{\mathcal{B}\mathcal{T}} = \frac{v_0^2}{c^2}\ (\frac{\mathcal{E}}{\mathcal{B}\mathcal{T}})$ and therefore is of order v^2/c^2. Similarly $(v_\parallel E_\parallel/c)\vec{u}_E$ is v^2/c^2 smaller than the preceding drift terms in Eq. (55).

It may seem strange that drifts in the direction of $\vec{E}_\perp$ occur. The origin of the drift which is proportional to $v_\parallel E_\parallel$ is easily understood. Because of the parallel electric field, the magnetic field is not in the same direction when viewed from the frame of reference moving at $\vec{u}_E$ as when viewed from the laboratory frame. Consequently a velocity which is parallel to $\vec{B}^*$ (asterisk refers to $\vec{u}_E$ frame) will have components both parallel and perpendicular to $\vec{B}$ when observed from the laboratory frame. Suppose that in the $\vec{u}_E$ frame there is a uniform static field $\vec{B}^*$ and an electric field $\vec{E}^*$ parallel to it, as shown in Fig. 6a. The guiding center velocity $\dot{\vec{R}}^*$ will consist of $\vec{v}_\parallel^*$ only. When a Lorentz transformation along X^* is made to the laboratory frame, the fields and guiding center velocity appear as in Fig. 6b. The $\vec{B}$ and $\vec{E}$ vectors lie in the Y-Z plane, but $\vec{B}$ does not lie along Z and $\vec{E}$ is not parallel to Y. The angle between $\vec{B}$ and the Z axis is proportional to $E_\parallel$. As shown, $\dot{\vec{R}}$ can be resolved into two components, one of which is $\vec{u}_E$ and the other is the drift $(v_\parallel E_\parallel \vec{E}_\perp)/(B^2(1-E_\perp^2/B^2))$ in the direction of $\vec{E}_\perp$.

The origin of the last drift term in Eq. (55), which is also in the direction of $\vec{E}_\perp$, is less obvious. One can start with a magnetic field having straight lines of force and a gradient of B at right angles to $\vec{B}$. There is a drift velocity due to ∇B. If one now makes a Lorentz transformation to a frame (denoted by *) moving at right angles to the drift and to $\vec{E}$, there will appear an $\vec{E}_\perp^*$ and a $\frac{\partial B^*}{\partial t^*}$. The ∇B drift velocity, when transformed to this new frame, would equal the sum of several drifts as calculated for the new frame from Eq. (55), including the second drift term in the direction of $\vec{E}_\perp^*$. Making the transformation does not really make the origin of that drift more transparent geometrically and therefore will not be done here.

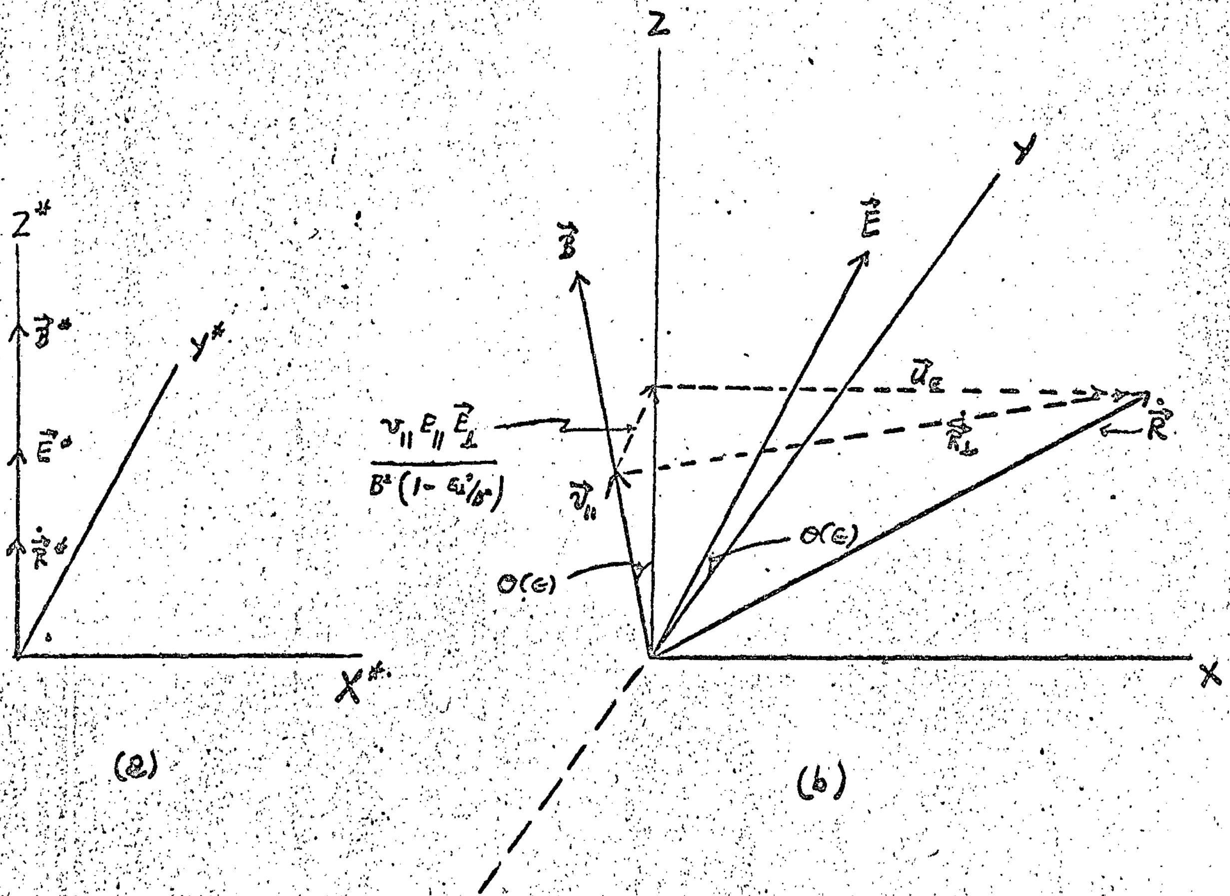


Fig. 6. The explanation of the drift proportional to $v_\parallel E_\parallel \vec{E}_\perp$

Finally, a comparison of Eq. (56) with (20) and Eq. (58) with (28) shows no new terms, only a modification of existing terms.

III. A MORE FORMAL DERIVATION OF THE NON-RELATIVISTIC GUIDING CENTER EQUATION

The derivation of the guiding center equations of motion for non-relativistic particles presented in Sect. II requires rigorous justification. The work of Kruskal(6) and of Berkowitz and Gardner(13) provides the justification of Eq. (4). Kruskal derives equations for the $\vec{R}_n$ apppearing in a series of the form

$$\vec{r} = \sum_{-\infty}^{\infty} \epsilon^{|n|} \vec{R}_n(t) \exp\left(\frac{in}{\epsilon c} \int B(R_0)\, dt\right) \qquad (61)$$

by equating coefficients of equal powers of $\exp\left(\frac{i}{\epsilon c}\int B\, dt\right)$ after substituting the series into Eq. (1). The fields are expanded in Taylor series about $\vec{R}_0$. Each $\vec{R}_n$ is itself a power series in ϵ, so that $\vec{R}_n(t) = \vec{R}_{n0}(t) + \epsilon \vec{R}_{n1}(t) + \ldots$. ($\vec{R}_{-n}$ must equal the complex conjugate of $\vec{R}_n$, in order that $\vec{r}$ be real. It is not immediately obvious that equating the coefficient of each $e^{in\theta}$ (where $\theta = \int \omega(\vec{R}_0)\, dt$) to zero is justified; the $\vec{R}_n$ are functions of time, so that the series is not simply a Fourier series. However, Berkowitz and Gardner prove that the series obtained by this process is actually an asymptotic expansion of $\vec{r}$ for small ϵ. Their proof is necessarily a formal mathematical one and will not be repeated here. But it is worthwhile to discuss what is usually meant by an asymptotic expansion and what it is they proved. The usual definition of an asymptotic expansion, to be found in Whittaker and Watson(14) for example, can be stated as follows: given the power series $S(\epsilon) = A_0 + A_1\epsilon + A_2\epsilon^2 + \ldots$. It is called the asymptotic expansion in ϵ of a function $f(\epsilon)$ if $\lim_{\epsilon\to 0} \left|\frac{f(\epsilon) - S_n(\epsilon)}{\epsilon^n}\right|$ is zero, where $S_n(\epsilon)$ is the sum of $n+1$ terms (i.e., including $A_n\epsilon^n$) of the series. Or by the definition of a limit, for every number Q there is a number $\epsilon_0(Q)$ such that

$\left|\dfrac{f(\epsilon) - S_n(\epsilon)}{\epsilon^n}\right|$ is $< Q$ for $\epsilon < \epsilon_0$. (We take ϵ as positive only). If this is true, one can show that given a value of ϵ, say ϵ_1, a number $A(\epsilon_1)$ can be found(15) such that $|f(\epsilon) - S_n(\epsilon)| < A(\epsilon_1)\,\epsilon^{n+1}$ for $0 < \epsilon < \epsilon_1$. This is the precise meaning of the statement that $|f(\epsilon) - S_n(\epsilon)|$ is $\mathcal{O}(\epsilon^{n+1})$. Now the series in Eq. (61) is not a simple power series in ϵ. Each term is itself an infinite series in ϵ, multiplied by $\exp\left(\frac{in}{\epsilon c}\int B\,dt\right)$; and this exponential is not expansible in a power series in ϵ. Nevertheless, Berkowitz and Gardner prove that if $\vec{\Sigma}_n$ is the sum of terms between and including(16) $\vec{R}_n$ and $\vec{R}_{-n}$, then there is an $A(\epsilon_1)$ for which $|\vec{r}(\epsilon, t) - \vec{\Sigma}_n| < A(\epsilon_1)\,\epsilon^{n-1}$ if $\epsilon < \epsilon_1$. Therefore $|\vec{r} - \vec{\Sigma}_n| = \mathcal{O}(\epsilon^{n-1})$ and in this sense $\vec{\Sigma}_n$ is the asymptotic series for $\vec{r}$. It is also true that $\lim_{\epsilon\to 0} \dfrac{|\vec{r} - \vec{\Sigma}_n|}{\epsilon^{n-2}} = 0$. With n equal to 2, it follows that $\vec{r} - \vec{\Sigma}_2 = \mathcal{O}(\epsilon)$. But $\vec{\Sigma}_2 = \vec{R}_0 + \epsilon(\vec{R}_1 e^{i\theta} + \vec{R}_{-1} e^{-i\theta}) + \epsilon^2 (\vec{R}_2 e^{2i\theta} + \vec{R}_{-2} e^{-2i\theta})$, so that $\vec{\Sigma}_2 - \vec{R}_0 = \mathcal{O}(\epsilon)$. By subtraction, $(\vec{r} - \vec{\Sigma}_2) + (\vec{\Sigma}_2 - \vec{R}_0) = \vec{r} - \vec{R}_0 = \mathcal{O}(\epsilon)$. This means that the difference between the actual particle position and $\vec{R}_0$ is of first order in the radius of gyration, and therefore $\vec{R}_0$ is a suitable definition(17) of the guiding center position.

Following Kruskal, we will now derive equations (some algebraic, some differential) for the $\vec{R}_n$. It will be found that the equation for $\vec{R}_0$, the guiding center position, is equation (4). Substituting Eq. (61) into Eq. (1) and collecting coefficients of $e^{in\theta}$ gives for

n = 0

$$\epsilon\ddot{\vec{R}}_0 = \vec{E}(\vec{R}_0, t) + \dot{\vec{R}}_0 \times \vec{B}(\vec{R}_0, t) + i\epsilon B(\vec{R}_0, t)\Big[\vec{R}_1 \times (\vec{R}_{-1}\cdot\nabla)\vec{B}(\vec{R}_0, t) - \vec{R}_{-1} \times (\vec{R}_1\cdot\nabla)\vec{B}(\vec{R}_0, t)\Big] + \mathcal{O}(\epsilon^2) \qquad (62)$$

<u>n = 1</u>

$$B^2\vec{R}_1 + iB\vec{R}_1 \times \vec{B} = -\epsilon\left[\vec{R}_1 \cdot \nabla\vec{E} + \dot{\vec{R}}_1 \times \vec{B} + \dot{\vec{R}}_0 \times (\vec{R}_1 \cdot \nabla)\vec{B} - i\dot{B}\vec{R}_1 - 2iB\dot{\vec{R}}_1\right] + O(\epsilon^2) \tag{63}$$

<u>n = 2</u>

$$-4B^2\vec{R}_2 - 2iB\vec{R}_2 \times \vec{B} = iB\vec{R}_1 \times (\vec{R}_1 \cdot \nabla)\vec{B} + O(\epsilon) \tag{64}$$

<u>n ≥ 2</u>

$$-n^2B^2\vec{R}_n - inB\,\vec{R}_n \times \vec{B} = \vec{G}_n, \tag{65}$$

where $\vec{G}_n$ is a function of the $\vec{R}$'s.

Note (The velocity of light c will be taken = 1. It can be re-introduced by dividing $\vec{B}$ and B by c.) In all four equations the fields are evaluated at $\vec{R}_0$ and t. In Eq. (65) the lowest order $\vec{G}_n$, denoted by $\vec{G}_{no}$, will always contain only $\vec{R}_{p0}$ (where $1 \leq p \leq n-1$). For example, from Eq. (64), $\vec{G}_{20} = iB(\vec{R}_0,t)\,\vec{R}_{10} \times (\vec{R}_{10} \cdot \nabla)\vec{B}(\vec{R}_0, t)$. In general $\vec{G}_{n1}$ will contain $\vec{R}_n$. It is thus possible to solve algebraically for $\vec{R}_{n0}$ in terms of the $\vec{R}_{10}, \vec{R}_{20}, \ldots \vec{R}_{n-1,0}$. To solve Eq. (65) for $\vec{R}_{n0}$, take its scalar product with $\hat{e}_1 \equiv \vec{B}/B$ to get

$$\vec{R}_{n0} \cdot \hat{e}_1 = -\frac{\vec{G}_{n0} \cdot \hat{e}_1}{B^2\,n^2} \tag{66}$$

and its cross product with $\hat{e}_1$

$$n^2\vec{R}_{n0} \times \hat{e}_1 + in\,(\hat{e}_1\,\vec{R}_{n0} \cdot \hat{e}_1 - \vec{R}_{n0}) = -\frac{\vec{G}_{n0} \times \hat{e}_1}{B^2} \tag{67}$$

Substitute $\vec{R}_{n0} \cdot \hat{e}_1 = -\dfrac{\vec{G}_{no} \cdot \hat{e}_1}{n^2\,B^2}$ from (66) and $\vec{R}_{n0} \times \hat{e}_1 = \frac{i}{n}\left(n^2\vec{R}_{n0} + \dfrac{\vec{G}_{n0}}{B^2}\right)$

from (65) into Eq. (67); the resulting equation can be solved for $\vec{R}_{n0}$:

$$\vec{R}_{n0} = \frac{-n^2\,\vec{G}_{n0} + (\hat{e}_1 \cdot \vec{G}_{n0})\,\hat{e}_1 + in\,\vec{G}_{n0} \times \hat{e}_1}{B^2\,n^2\,(n^2 - 1)}\,, \tag{68}$$

where the fields and $\hat{e}_1$ are evaluated at $(\vec{R}_0, t)$ as usual.

In order to find $\vec{R}_{n1}$, equation (65) is written to next order as

$$-in^2B\,\vec{R}_{n1} - inB\,\vec{R}_{n1} \times \vec{B} = \vec{G}_{n1}, \tag{69}$$

where $\vec{G}_{n1}$ will contain $\vec{R}_p$ $(p \leq n-1)$ and $\vec{R}_{n0}$. Thus there exists a recursion scheme for $\vec{R}_2$ and higher, provided $\vec{R}_0$ and $\vec{R}_1$ are known so that the recursion scheme can be initiated. ¶ Equations (62) and (63) determine $\vec{R}_0$ and $\vec{R}_1$. The zero order equation (63) is

Note Paragraph

$$\vec{R}_{10} + i\vec{R}_{10} \times \hat{e}_1 = 0\,. \tag{70}$$

Dotting with $\hat{e}_1(\vec{R}_0)$ gives $\vec{R}_{10} \cdot \hat{e}_1 = 0$, so that $\vec{R}_{10}$ is a vector perpendicular to the magnetic field. Therefore let $\vec{R}_{10}$ equal $(a + ib)\hat{e}_2 + (c + id)\hat{e}_3$, where $\hat{e}_2(\vec{R}_0)$ and $\hat{e}_3(\vec{R}_0)$ are perpendicular to each other and to $\hat{e}_1(\vec{R}_0)$ as previously, and a and b are real. Substitution into Eq. (70) gives $c = -b$ and $a = d$. As a result $\vec{R}_{10}$ is of the form

$$\vec{R}_{10} = (a + ib)(\hat{e}_2 + i\hat{e}_3). \tag{71}$$

Equation (71) contains all the information present in Eq. (70).

If Eq. (71) is now substituted into the square brackets in Eq. (62), the result is

$$\begin{aligned} \epsilon\,\ddot{\vec{R}}_0 &= \vec{E} + \dot{\vec{R}}_0 \times B + 2\epsilon B(a^2 + b^2)[\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} - \hat{e}_3 \times (\hat{e}_2 \cdot \nabla)\vec{B}] + O(\epsilon^2) \\ &= \vec{E} + \dot{\vec{R}}_0 \times B + \epsilon B\vec{R}_{10} \cdot \vec{R}_{10}^{*}\,[\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} - \hat{e}_3 \times (\hat{e}_2 \cdot \nabla)\vec{B}] + O(\epsilon^2)\,. \end{aligned} \tag{72}$$

The coefficient $B\vec{R}_{10} \cdot \vec{R}_{10}^{*}$ will now be identified as the magnetic moment by differentiating Eq. (61) with respect to time. Since $\vec{R}_n = \vec{R}_{-n}^{*}$, differentiation yields.

$$\vec{v} = \dot{\vec{r}} = \dot{\vec{R}}_0 + iB\,(\vec{R}_{10}\,e^{i\theta} - \vec{R}_{10}^{*}\,e^{-i\theta}) + \mathcal{O}(\epsilon) \qquad (73)$$

$$(\vec{v} - \dot{\vec{R}}_0)_{\perp} = iB(\vec{R}_{10}\,e^{i\theta} - \vec{R}_{10}^{*}\,e^{-i\theta}) + \mathcal{O}(\epsilon), \qquad (74)$$

since $\vec{R}_{10}$ is perpendicular to $\vec{B}$ by Eq. (71). Squaring (74) gives

$$(v - \dot{\vec{R}}_0)_{\perp}^2 = 2B^2\,\vec{R}_{10} \cdot \vec{R}_{10}^{*} + \mathcal{O}(\epsilon) \qquad (75)$$

because $\vec{R}_{10} \cdot \vec{R}_{10} = 0$. However, $(\vec{v} - \dot{\vec{R}}_0)_{\perp}$ is the perpendicular or gyration velocity in the frame of reference moving at the guiding center velocity, and this is $\rho\omega$. Therefore $\epsilon B\vec{R}_{10} \cdot \vec{R}_{10}^{*} = \dfrac{\epsilon\rho^2\omega^2}{2B} = \dfrac{\rho^2\omega}{2c} = \dfrac{M}{e}$,

and Eq. (62) becomes

$$\epsilon\ddot{\vec{R}}_0 = \vec{E}(\vec{R}_0, t) + \frac{\dot{\vec{R}}_0}{c} \times \vec{B}(\vec{R}_0, t) + \frac{\rho^2\omega}{2c}\left[\hat{e}_2 \times (\hat{e}_3 \cdot \nabla)\vec{B} - \hat{e}_3 \times (\hat{e}_2 \cdot \nabla)\vec{B}\right] + \mathcal{O}(\epsilon^2), \qquad (76)$$

which is the same as equation (4) for $\dot{\vec{R}}$. We have thus justified the averaging process used to get Eq. (4) in Sect. II.

In the next Section, equation (63) will be used to next higher order to show that $\dfrac{d}{dt}\,(B\vec{R}_{10} \cdot \vec{R}_{10}^{*}) = 0 + \mathcal{O}(\epsilon)$, and this will be a proof of the adiabatic invariance of the magnetic moment in the general case where $\vec{E}_{\perp}$ and $\dfrac{\partial}{\partial t}$ are of $\mathcal{O}(1)$.

IV The Adiabatic Invariants of the Motion

Since the solution of the equation of motion has been obtained as an asymptotic series in ϵ in the previous section, it is reasonable that any approximate constants (adiabatic invariants) of the particle motion should be obtainable as asymptotic series in ϵ. In analytical dynamics exact constants of the motion are usually obtained by the canonical formulation: if the Hamiltonian is independent of a given coordinate, the conjugate momentum is an invariant. By analogy, we expect that an expansion of the Hamiltonian in an asymptotic series in ϵ should reveal the adiabatic invariants, provided that at each step of the expansion, variables can be found which makes the Hamiltonian independent of one of the coordinates. (The last term H_n in the expansion of H as $H_0 + \epsilon H_1 + \ldots \epsilon^n H_n$ may contain the coordinate). The systematic procedure for finding the proper variables has been given for the non-relativistic case by Gardner.(18)

It is similarly true that to find exact invariants the Hamiltonian must be expressed in terms of the proper variables. As an example, if the Hamiltonian for a charged particle in a magnetic field having azimuthal symmetry is written in rectangular coordinates, it is not at all evident that the canonical momentum $P_\theta = mr^2\dot{\theta} + \frac{e}{c} r A_\theta$ is an exact invariant of the motion. It only becomes apparent when H is written in cylindrical coordinates.

There is an even more general theory of asymptotic solutions and adiabatic invariants of coupled first order differential equations of a certain type. This is due to Kruskal.(19) The equation of motion of a charged particle is a special case. This more general theory will be discussed in some detail later in this Section.

Although the adiabatic invariants are asymptotic series of the form: constant $= A_0 + \epsilon A_1 + \epsilon^2 A_2 + \ldots$, it is customary to speak of the lowest order invariant A_0 as "the" adiabatic invariant. For the charged particle there are as many as three such series, one for the magnetic moment, one for the longitudinal invariant, and one for the "flux" invariant. These three series will be designated by $M + \epsilon M' + \ldots$, $J + \epsilon J' + \ldots$, $\Phi + \epsilon \Phi' + \ldots$, respectively. Proofs of the invariance of the lowest orders M, J, and Φ will be given below.

The number of adiabatic invariants is less than or equal to the number of degrees of freedom of the system. The charged particle, which has three degrees of freedom, may or may not have M, J, and Φ, depending on the field geometry. The number of adiabatic invariants is determined by the number of periodicities. To illustrate, suppose that B is nowhere large enough to reflect the particle. The particle motion is nearly periodic because of the gyration about the line of force, but there is no semblance of periodicity in the motion along the line of force. There would be only one adiabatic invariant series, the one for the magnetic moment, even though there are three degrees of freedom. If now the field is such that a particle is always trapped and oscillating between two mirrors, there will be a second or longitudinal invariant J. Finally, if the drift from line to line as the particle oscillates between mirrors with constant M and J carries the particle repeatedly around a closed surface, there is a third periodicity associated with the motion and there will exist a third adiabatic invariant Φ. The motion of charged particles which comprise the Van Allen radiation possesses all three ~~periodicities~~ and invariants.[20] The periodicities are: the gyration about the geomagnetic field lines, the north-south oscillation, and the precession about the earth.

A. The Magnetic Moment

The invariance of the magnetic moment will now be proven via Eq. (63) and Maxwell's equations. It will be shown that $B\vec{R}_{10} \cdot \vec{R}_{10}^{*}$, which is proportional to the magnetic moment, is independent of time to lowest order in ϵ. By way of clarification it should be emphasized that the magnetic moment is not always $\frac{"mv_\perp^2"}{2B}$, where $v_\perp$ is the perpendicular velocity in the laboratory frame of reference. If there is an electric field $\vec{E}_\perp$, then the magnetic moment is $\frac{m(\vec{v}-\dot{\vec{R}}_0)_\perp^2}{2B}$, where $(\vec{v}-\dot{\vec{R}}_0)_\perp$ equals $(\vec{v}-\vec{u}_E)_\perp$. In short, the "perpendicular velocity" must be that observed in the frame of reference moving at $\vec{u}_E$. This can be verified from Eq. (75). It does not matter whether the component of $\vec{v}-\vec{u}_E$ perpendicular to $\vec{B}(\vec{r})$ or to $\vec{B}(\vec{R}_0)$ is used, the difference being of $\mathcal{O}(\epsilon)$ and therefore appearing in the $\epsilon M'$ term of the magnetic moment series.

Superscript "$\frac{mv_\perp^2}{2B}$"

Equation (63) is of the form

$$\vec{R}_1 + i\vec{R}_1 \times \hat{e}_1 = -\frac{\epsilon\vec{F}}{B^2} \, . \tag{77}$$

If this is first dotted and then crossed with $\hat{e}_1$ in an attempt to solve for $\vec{R}_1$ in a fashion similar to the solution for $\vec{R}_{n0}$ (Eq. 68), the result is (by setting $n = 1$ in Eq. (68)).

$$-\vec{F} + (\hat{e}_1 \cdot \vec{F})\,\hat{e}_1 + i\vec{F} \times \hat{e}_1 = 0$$

or

$$\vec{F}_\perp - i\vec{F}_\perp \times \hat{e}_1 = 0 \tag{78}$$

This is a condition on $\vec{F}_{\perp}$, which to $\mathcal{O}(1)$ is a condition on $\vec{F}_{0\perp}$, where

$$\vec{F}_0 = \vec{R}_{10} \cdot \nabla \vec{E}(\vec{R}_0, t) + \dot{\vec{R}}_{10} \times \vec{B}(\vec{R}_0, t) + \dot{\vec{R}}_0 \times \left[\vec{R}_{10} \cdot \nabla \vec{B}(\vec{R}_0, t) - i\dot{B}(\vec{R}_0, t)\vec{R}_{10} - 2iB(\vec{R}_0, t)\dot{\vec{R}}_{10}\right] \tag{79}$$

Equations (78) and (79) constitute a differential equation for $\vec{R}_{10}$. If the terms with $\dot{\vec{R}}_{10}$ are separated on the left hand side, the differential equation is

$$(\dot{\vec{R}}_{10})_{\perp} - i\,(\dot{\vec{R}}_{10})_{\perp} \times \hat{e}_1 = \frac{1}{B}\,(\vec{L}_{\perp} - i\vec{L}_{\perp} \times \hat{e}_1), \tag{80}$$

where $\vec{L}_{\perp} = \left[-i\vec{R}_{10} \cdot \nabla \vec{E} - i\dot{\vec{R}}_{00} \times (\vec{R}_{10} \cdot \nabla)\vec{B} - \dot{B}\,\vec{R}_{10}\right]_{\perp}$ and $\dot{\vec{R}}_{00}$ is the zero order motion $v_{\parallel}\hat{e}_1 + \vec{u}_E$. One might conclude from Eq. (80) that $(\dot{\vec{R}}_{10})_{\perp}$ must equal $\frac{\vec{L}_{\perp}}{B}$. Such is not the case, however; any complex vector quantity of the form $\vec{W}_{\perp} - i\vec{W}_{\perp} \times \hat{e}_1$ can be written as $(g + ih)(\hat{e}_2 + i\hat{e}_3)$, where $\vec{W} = g\hat{e}_2 + h\hat{e}_3$, and g and h are complex. By collecting all terms on one side of Eq. (80), one finds that $(\dot{\vec{R}}_{10})_{\perp} - \frac{\vec{L}_{\perp}}{B}$ also is of this form and is not necessarily zero. Equation (80) must be used as it stands to prove that $\frac{d}{dt}(B\,\vec{R}_{10} \cdot \vec{R}_{10}^{\,*})$ equals zero.

$$\frac{d}{dt}\,(B\vec{R}_{10} \cdot \vec{R}_{10}^{\,*}) = B(\dot{\vec{R}}_{10} \cdot \vec{R}_{10}^{\,*} + \vec{R}_{10} \cdot \dot{\vec{R}}_{10}^{\,*} + \frac{\dot{B}}{B}\vec{R}_{10} \cdot \vec{R}_{10}^{\,*}). \tag{81}$$

Equation (80) contains only $(\dot{\vec{R}}_{10})_{\perp}$; but this is all that is required in Eq. (81) because of the fact that $\vec{R}_{10}^{\,*}$ and $\vec{R}_{10}$ have no parallel components. Now Eq. (71) gives, when differentiated with respect to time

$$\dot{\vec{R}}_{10} = (\dot{a} + i\dot{b})(\hat{e}_2 + i\hat{e}_3) + (a + ib)(\dot{\hat{e}}_2 + i\dot{\hat{e}}_3) \qquad (82)$$

Substitution of this into Eq. (80) gives (via either the $\hat{e}_2$ or $\hat{e}_3$ component)

$$\dot{a} + i\dot{b} = -i(a + ib)(\hat{e}_2 \cdot \dot{\hat{e}}_3) + \frac{g + ih}{2B}, \qquad (83)$$

where $(g + ih)(\hat{e}_2 + i\hat{e}_3)$ now stands for $\vec{L}_\perp - i\vec{L}_\perp \times \hat{e}_1$. From equations (82) and (83) it follows that

$$\dot{\vec{R}}_{10} = \left[- i\,(a + ib)(\hat{e}_2 \cdot \dot{\hat{e}}_3) + \frac{g + ih}{2B}\right](\hat{e}_2 + i\hat{e}_3) + (a + ib)(\dot{\hat{e}}_2 + i\dot{\hat{e}}_3)$$

$$= -i\,\hat{e}_2 \cdot \dot{\hat{e}}_3\,\vec{R}_{10} + \frac{g + ih}{2B(a + ib)}\,\vec{R}_{10} + (a + ib)(\dot{\hat{e}}_2 + i\dot{\hat{e}}_3) \qquad (84)$$

$$\dot{\vec{R}}_{10} \cdot \vec{R}_{10}^{\,*} = \left[- i\hat{e}_2 \cdot \dot{\hat{e}}_3 + \frac{g + ih}{2B(a + ib)}\right]\vec{R}_{10} \cdot \vec{R}_{10}^{\,*} + i(a^2 + b^2)(\dot{\hat{e}}_2 \cdot \hat{e}_3 - \dot{\hat{e}}_3 \cdot \hat{e}_2) \qquad (85)$$

The sum of Eq. (85) with its complex conjugate is

$$2\,\mathcal{R}e\,(\dot{\vec{R}}_{10} \cdot \vec{R}_{10}^{\,*}) - \frac{\vec{R}_{10} \cdot \vec{R}_{10}^{\,*}}{B}\,\mathcal{R}e\left(\frac{g + ih}{a + ib}\right) = 0 \quad . \qquad (86)$$

In order to prove that the right hand side of Eq. (81) vanishes, it remains only to show that $\mathcal{R}e\left(\frac{g + ih}{a + ib}\right)$ equals $-\frac{dB}{dt}$. Now $g + ih$ is defined by $\vec{L}_\perp - i\vec{L}_\perp \times \hat{e}_1 = (g + ih)(\hat{e}_2 + i\hat{e}_3)$. This can be solved for $g + ih$ by dotting with $\hat{e}_2$ or $\hat{e}_3$

$$g + ih = \vec{L}_\perp \cdot (\hat{e}_2 - i\hat{e}_3). \qquad (87)$$

The explicit expression for $\vec{L}_\perp$ given immediately following Eq. (80) must now be used; $\vec{L}_\perp$ contains $\vec{R}_{10}$, which is to be replaced by $(a + ib)(\hat{e}_2 + i\hat{e}_3)$.

$$\frac{g+ih}{a+ib} = \frac{\hat{L}_\perp \cdot (\hat{e}_2 - i\hat{e}_3)}{a+ib} = \left[-i\,(\hat{e}_2 + i\hat{e}_3)\cdot\nabla\vec{E} - i\dot{\vec{R}}_{00} \times (\hat{e}_2 + i\hat{e}_3)\cdot\nabla\vec{B} - \dot{B}(\hat{e}_2 + i\hat{e}_3)\right] \cdot (\hat{e}_2 - i\hat{e}_3) \qquad (88)$$

The first term on the right hand side is

$$-(\hat{e}_3 + i\hat{e}_2)\cdot\left[(\hat{e}_2 + i\hat{e}_3)\cdot\nabla\right]\vec{E} = \hat{e}_2\cdot(\hat{e}_3\cdot\nabla)\vec{E} - \hat{e}_3\cdot(\hat{e}_2\cdot\nabla)\vec{E} + i\,(\ldots)$$

$$= -\hat{e}_1\cdot\left[\hat{e}_3 \times (\hat{e}_3\cdot\nabla)\vec{E} + \hat{e}_2 \times (\hat{e}_2\cdot\nabla)\vec{E}\right] + i\,(\ldots)$$

$$= -\hat{e}_1\cdot\nabla\times\vec{E} + i\,(\ldots). \qquad (89)$$

The imaginary part does not have to be evaluated because only the real part of (89) is required. The second term on the right side of Eq. (88) is

$$\dot{\vec{R}}_{00}\cdot\left[(\hat{e}_3 - i\hat{e}_2)\cdot\nabla\right]\vec{B} \times (\hat{e}_2 - i\hat{e}_3) = \dot{\vec{R}}_{00}\cdot\left\{(i\hat{e}_3 - \hat{e}_2)\times\left[(\hat{e}_3 - i\hat{e}_2)\cdot\nabla\right](B\hat{e}_1)\right\}$$

$$= \dot{\vec{R}}_{00}\cdot B\left[\hat{e}_3 \times (\hat{e}_2\cdot\nabla)\,\hat{e}_1 - \hat{e}_2 \times (\hat{e}_3\cdot\nabla)\,\hat{e}_1\right]$$

$$+ \dot{\vec{R}}_{00}\cdot\left[\hat{e}_3\,(\hat{e}_3\cdot\nabla)B + \hat{e}_2\,(\hat{e}_2\cdot\nabla)B\right] + i\,(\ldots)$$

$$= \dot{\vec{R}}_{00}\cdot B\hat{e}_1\left[-\hat{e}_2\cdot(\hat{e}_2\cdot\nabla)\hat{e}_1 - \hat{e}_3\cdot(\hat{e}_3\cdot\nabla)\hat{e}_1\right] + \dot{\vec{R}}_{00}\cdot\nabla_\perp B + i(\ldots)$$

$$= \dot{\vec{R}}_{00}\cdot\left[-\vec{B}\,\nabla\cdot\hat{e}_1 + \nabla_\perp B\right] + i\,(\ldots). \qquad (90)$$

Because

$$0 = \nabla\cdot\vec{B} = \nabla\cdot(\hat{e}_1 B) = \hat{e}_1\cdot\nabla B + B\,\nabla\cdot\hat{e}_1, \qquad (91)$$

it follows that $-\vec{B}\,\nabla\cdot\hat{e}_1 = \hat{e}_1\,\hat{e}_1\cdot\nabla B$ and the real part of the right side of Eq. (90) is simply $\dot{\vec{R}}_{00}\cdot\nabla B$.

The last term on the right side of (88) is

$$-\mathring{B}\,(\hat{e}_2 + i\hat{e}_3)\cdot(\hat{e}_2 - i\hat{e}_3) = -2\mathring{B} \tag{92}$$

and

$$\mathcal{R}e\left(\frac{g+ih}{a+ib}\right) = -\hat{e}_1\cdot\nabla\times\vec{E} + \dot{\vec{R}}_{00}\cdot\nabla B - 2\mathring{B}\,. \tag{93}$$

By Maxwell's equation, $\nabla\times\vec{E} = -\dfrac{\partial\vec{B}}{\partial t}$ and $\hat{e}_1\cdot\nabla\times\vec{E} = -\dfrac{\partial B}{\partial t}$, since $\hat{e}_1\cdot\dfrac{\partial\hat{e}_1}{\partial t} = 0.$

$$\mathcal{R}e\left(\frac{g+ih}{a+ib}\right) = \frac{\partial B}{\partial t} + \dot{\vec{R}}_{00}\cdot\nabla B - 2\mathring{B} = -\dot{B}, \tag{94}$$

since $\dot{B} = \dfrac{\partial B}{\partial t} + \dot{\vec{R}}_{00}\cdot\nabla B + \mathcal{O}(\epsilon)$. Thus it is proven that

$\dfrac{d}{dt}\left(B\vec{R}_{10}\cdot\vec{R}_{10}^{\,*}\right) = 0 + \mathcal{O}(\epsilon)$ or that $B\vec{R}_{10}\cdot\vec{R}_{10}^{\,*}$ equals a constant $+\ \mathcal{O}(\epsilon)$.

B. The Longitudinal Adiabatic Invariant

The next adiabatic invariant to be studied is the longitudinal invariant

$$J = \oint p_{\parallel}\, ds, \qquad (95)$$

where $p_{\parallel}$ is the guiding center momentum parallel to the line of force, and the integral is taken over a complete oscillation from one mirror point to the other and back again. As stated previously, the longitudinal motion must be periodic for J to exist. The procedure thus far has been to start with the equation of motion (1) of a charged particle and to average over the gyration. The resulting guiding center equations (19) and (20) are new equations of motion. The next step is to average over the longitudinal oscillation and obtain a third set of equations of motion governing the average drift from line to line, and then to show that this average drift conserves J.

As the guiding center moves along a line of force in accord with Eq. (20), it drifts at right angles to the line at a rate given by Eq. (19). See Fig. 7.

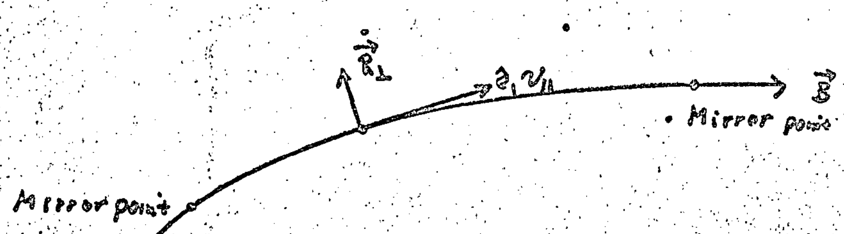


Fig. 7. Guiding center oscillates along a line of force and drifts slowly at right angles to it.

If the drift at right angles is slow compared to the longitudinal motion (i.e., if $\vec{E}_\perp$ and $\frac{\partial}{\partial t}$ are of $\mathcal{O}(\epsilon)$), one can calculate the average drift rate at right angles to the line during a longitudinal oscillation as if the guiding center did not deviate from the line of force, the error being of order ϵ^2 since the drift rate $\dot{\vec{R}}_\perp$ is $\mathcal{O}(\epsilon)$. If $\vec{E}_\perp$ is $\mathcal{O}(1)$, $\dot{\vec{R}}_\perp$ contains the $\mathcal{O}(1)$ term $\vec{u}_E$ and the guiding center moves a long ways from a given line of force in one oscillation. It is then no longer possible to ignore the deviation from the line of force; the guiding center does not remain even approximately on the line and it will be found that $\oint^{l.c.} p_\parallel ds$ is not conserved.

A proof has been given in reference (20) that the average drift from line to line conserves J; the proof is for relativistic particles. Rather than repeating tthat proof in the present review, a somewhat different one will be given which is more algebraic and less geometric. Also it will be done for non-relativistic energies, but the modifications of this proof to relativistic energies are slight since the relativistic expressions (49), (50), and (51) are known.

To formulate the problem explicitly, the α, β, s curvilinear coordinate system introduced previously will be used. Let $K = \frac{mv_\parallel^2}{2} + MB + e(\phi + \psi)$, where $\psi = \frac{\alpha}{c}\frac{\partial\beta}{\partial t}$. By Eq. (29), K is a constant of the longitudinal motion. Then J is given by[21]

$$J(\alpha, \beta, K, t) = \oint \left\{ 2m\left[K - e(\psi + \phi) - MB\right]\right\}^{1/2} ds \qquad (96)$$

where ψ, ϕ, and B are all functions of (α, β, s, t). The rate of change of J is

$$\frac{dJ}{dt} = \frac{\partial J}{\partial \alpha}\frac{d\alpha}{dt} + \frac{\partial J}{\partial \beta}\frac{d\beta}{dt} + \frac{\partial J}{\partial K}\frac{dK}{dt} + \frac{\partial J}{\partial t}, \tag{97}$$

where

$$\frac{\partial J}{\partial \alpha} = -me \oint \frac{ds}{\left\{2m\left[K - e(\psi+\phi) - MB\right]\right\}^{1/2}} \cdot \frac{\partial}{\partial \alpha}\left(\psi+\phi+\frac{MB}{e}\right)$$

$$= -e \oint \frac{ds}{v_{\parallel}} \frac{\partial}{\partial \alpha}\left(\psi+\phi+\frac{M}{e}B\right) \tag{98}$$

$$\frac{\partial J}{\partial \beta} = -me \oint \frac{ds}{\left\{2m\left[K - e(\psi+\phi) - MB\right]\right\}^{1/2}} \frac{\partial}{\partial \beta}\left(\psi+\phi+\frac{MB}{e}\right)$$

$$= -e \oint \frac{ds}{v_{\parallel}} \frac{\partial}{\partial \beta}\left(\psi+\phi+\frac{MB}{e}\right) \tag{99}$$

$$\frac{\partial J}{\partial K} = m \oint \frac{ds}{\left\{2m\left[K - e(\psi+\phi) - MB\right]\right\}^{1/2}} = \oint \frac{ds}{v_{\parallel}} = T, \tag{100}$$

where T is the period of the longitudinal oscillation.

$$\frac{\partial J}{\partial t} = -me \oint \frac{ds}{\left\{2m\left[K - e(\psi+\phi) - MB\right]\right\}^{1/2}} \frac{\partial}{\partial t}\left(\psi+\phi+\frac{MB}{e}\right)$$

$$= -e \oint \frac{ds}{v_{\parallel}} \frac{\partial}{\partial t}\left(\psi+\phi+\frac{MB}{e}\right). \tag{101}$$

In equations (98) - (101), ψ, ϕ, and B are to be considered as functions of (α, β, s, t) and not of the guiding center position $\vec{R}$ and of t. $\vec{R}$ is itself a function $\vec{R}(\alpha, \beta, s, t)$. It must be remembered that

$$\frac{\partial \psi(\alpha,\beta,s,t)}{\partial t} = \frac{\partial \psi(\vec{R},t)}{\partial t} + \frac{\partial \vec{R}(\alpha,\beta,s,t)}{\partial t} \cdot \nabla \psi(\vec{R}, t).$$

(102)

$$\frac{\partial \psi(\alpha, \beta, s, t)}{\partial \alpha} = \frac{\partial \vec{R}}{\partial \alpha}(\alpha, \beta, s, t) \cdot \nabla \psi(\vec{R}, t), \text{ etc. for } \beta \text{ and } s.$$

Since $\alpha(\vec{R}, t)$ and $\beta(\vec{R}, t)$ are constants on a line of force, their time derivatives contain $\dot{\vec{R}}_{\perp}$ and not $v_{\parallel}$.

$\frac{\partial \alpha(\vec{R}, t)}{\partial t}$

$$\frac{d\alpha}{dt} = \frac{\partial \alpha}{\partial t}(\vec{R}, t) + \dot{\vec{R}}_{\perp} \cdot \nabla \alpha(\vec{R}, t) + \mathcal{O}(\epsilon^2) \qquad (103)$$

$$\frac{d\beta}{dt} = \frac{\partial \beta}{\partial t} + \dot{\vec{R}}_{\perp} \cdot \nabla \beta + \mathcal{O}(\epsilon^2) \qquad (104)$$

The expression for $\dot{\vec{R}}_{\perp}$ (with $\vec{u}_E$ and $\frac{\partial}{\partial t} \sim \epsilon$) must now be substituted into equations (103) and (104). The procedure will be carried out in detail only for $\frac{d\alpha}{dt}$. It is the same for $d\beta/dt$ (except for a sign.).

$$\frac{d\alpha}{dt} = \frac{\partial \alpha}{\partial t} + \frac{\hat{e}_1}{B} \times \left\{ -c\vec{E} + \frac{Mc}{e}\nabla B + \frac{mc}{e} v_{\parallel}^{2} \frac{\partial \hat{e}_1}{\partial s} \right\} \cdot \nabla \alpha + \mathcal{O}(\epsilon^2) \qquad (105),$$

Now $\alpha = \alpha[\vec{R}(\alpha, \beta, s, t), t]$, $\beta = \beta[\vec{R}(\alpha, \beta, s, t), t]$ and $s = s[\vec{R}(\alpha, \beta, s, t), t]$. By implicit differentiation of α with respect to α, β, s, and t, and of β and of s with respect to the same four variables, one obtains the following equations:

$$1 = \nabla \alpha(\vec{R}, t) \cdot \frac{\partial \vec{R}}{\partial \alpha}(\alpha, \beta, s, t)$$

$$0 = \nabla \alpha \cdot \frac{\partial \vec{R}}{\partial \beta}$$

$$0 = \nabla \alpha \cdot \frac{\partial \vec{R}}{\partial s}$$

$$0 = \frac{\partial \alpha}{\partial t} + \nabla \alpha \cdot \frac{\partial \vec{R}}{\partial t}$$

leave a space →

$$0 = \nabla \beta(\vec{R}, t) \cdot \frac{\partial \vec{R}}{\partial \alpha}(\alpha, \beta, s, t)$$

$$1 = \nabla \beta \cdot \frac{\partial \vec{R}}{\partial \beta} \qquad (106)$$

$$0 = \nabla \beta \cdot \frac{\partial \vec{R}}{\partial s}$$

$$0 = \frac{\partial \beta}{\partial t} + \nabla \beta \cdot \frac{\partial \vec{R}}{\partial t}$$

leave a space →

$$0 = \nabla s(\vec{R}, t) \cdot \frac{\partial \vec{R}}{\partial \alpha}(\alpha, \beta, s, t)$$

$$0 = \nabla s \cdot \frac{\partial \vec{R}}{\partial \beta}$$

$$1 = \nabla s \cdot \frac{\partial \vec{R}}{\partial s}$$

$$0 = \frac{\partial s}{\partial t} + \nabla s \cdot \frac{\partial \vec{R}}{\partial t}.$$

In Eq. (105) $\nabla\alpha$ can therefore be replaced by $\frac{\partial\vec{R}}{\partial\beta} \times \vec{B}$,
since $\vec{B}(\vec{R}) = \nabla\alpha \times \nabla\beta$ and

beta

$$\frac{\partial\vec{R}}{\partial\beta} \times \vec{B} = \frac{\partial\vec{R}}{\partial\beta} \times (\nabla\alpha \times \nabla\beta) = -\left(\nabla\alpha \cdot \frac{\partial\vec{R}}{\partial\beta}\right)\nabla\beta + \left(\nabla\beta \cdot \frac{\partial\vec{R}}{\partial\beta}\right)\nabla\alpha = \nabla\alpha \qquad (107)$$

by equation (106). Also in Eq. (105)

$$\vec{E} = -\frac{1}{c}\frac{\partial\vec{A}}{\partial t} - \nabla\phi = -\frac{1}{c}\frac{\partial}{\partial t}(\alpha\nabla\beta) - \nabla\phi = -\nabla\psi + \frac{1}{c}\left(\frac{\partial\beta}{\partial t}\nabla\alpha - \frac{\partial\alpha}{\partial t}\nabla\beta\right) - \nabla\phi \qquad (108)$$

With these substitutions Eq. (105) becomes, after interchanging the dot and cross, and expanding the triple vector product $\left(\frac{\partial\vec{R}}{\partial\beta} \times \hat{e}_1\right) \times \hat{e}_1$:

$$\frac{d\alpha}{dt} = \frac{\partial\alpha}{\partial t} + \left(\hat{e}_1\hat{e}_1 \cdot \frac{\partial\vec{R}}{\partial\beta} - \frac{\partial\vec{R}}{\partial\beta}\right) \cdot \left(c\nabla\psi + c\nabla\phi - \frac{\partial\beta}{\partial t}\nabla\alpha + \frac{\partial\alpha}{\partial t}\nabla\beta + \frac{Mc}{e}\nabla B + \frac{mc}{e}v_\parallel^2\frac{\partial\hat{e}_1}{\partial s}\right) + O(\epsilon^2)$$

$$= \frac{\partial\alpha}{\partial t} + c\,\hat{e}_1 \cdot \frac{\partial\vec{R}}{\partial\beta}\frac{\partial}{\partial s}\left(\psi + \phi + \frac{MB}{e}\right) + \left(\frac{\partial\vec{R}}{\partial\beta} \cdot \nabla\alpha\right)\frac{\partial\beta}{\partial t} - \left(\frac{\partial\vec{R}}{\partial\beta} \cdot \nabla\beta\right)\frac{\partial\alpha}{\partial t}$$

$$- c\frac{\partial\vec{R}}{\partial\beta} \cdot \nabla\left(\psi + \phi + \frac{MB}{e}\right) - \frac{mc}{e}v_\parallel^2\frac{\partial\vec{R}}{\partial\beta} \cdot \frac{\partial\hat{e}_1}{\partial s} + O(\epsilon^2) \qquad (109)$$

By Eq. (106), and because $\frac{\partial\vec{R}}{\partial\beta}(\alpha,\beta,s,t) \cdot \nabla = \frac{\partial}{\partial\beta}$, equation (109) becomes

$$\frac{d\alpha}{dt} = c\,\hat{e}_1 \cdot \frac{\partial\vec{R}}{\partial\beta}(\alpha,\beta,s,t)\frac{\partial}{\partial s}\left(\psi + \phi + \frac{MB}{e}\right) - c\frac{\partial}{\partial\beta}\left(\psi + \phi + \frac{MB}{e}\right) - \frac{mc}{e}v_\parallel^2\frac{\partial\vec{R}}{\partial\beta} \cdot \frac{\partial\hat{e}_1(\alpha,\beta,s,t)}{\partial s} + O(\epsilon^2) \qquad (110)$$

where $\psi + \phi + \frac{MB}{e}$ is to be considered a function of (α, β, s, t).
By Eq. (22) $\frac{\partial}{\partial s}\left(\psi + \phi + \frac{MB}{e}\right)$ equals $-\frac{m}{e}\frac{dv_\parallel}{dt} + O(\epsilon^2)$. In the last term of Eq. (110)

$$\frac{\partial \vec{R}}{\partial \beta} \cdot \frac{\partial \hat{e}_1}{\partial s} = \frac{\partial}{\partial s}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right) - \hat{e}_1 \cdot \frac{\partial}{\partial s}\frac{\partial \vec{R}(\alpha,\beta,s,t)}{\partial \beta}$$

$$= \frac{\partial}{\partial s}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right) - \hat{e}_1 \cdot \frac{\partial}{\partial \beta}\frac{\partial \vec{R}}{\partial s} = \frac{\partial}{\partial s}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right) - \hat{e}_1 \cdot \frac{\partial \hat{e}_1}{\partial \beta}$$

$$= \frac{\partial}{\partial s}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right), \qquad (111)$$

since $\hat{e}_1 = \partial \vec{R}/\partial s$. The last term of Eq. (110) then is

$$\frac{mc}{e} v_{\parallel}^2 \frac{\partial \vec{R}}{\partial \beta} \cdot \frac{\partial \hat{e}_1}{\partial s} = \frac{mc}{e} v_{\parallel}^2 \frac{\partial}{\partial s}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right) = \frac{mcv_{\parallel}}{e}\frac{d}{dt}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right) + O(\epsilon^2) \qquad (112)$$

The first and last terms on the right side of (110) combine to

$-\frac{mc}{e}\frac{d}{dt}\left(v_{\parallel}\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right) + O(\epsilon^2)$ and

$$\frac{d\alpha}{dt} = -c\frac{\partial}{\partial \beta}\left(\psi + \phi + \frac{M}{e}B\right) - \frac{mc}{e}\frac{d}{dt}\left(v_{\parallel}\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \beta}\right) + O(\epsilon^2) \qquad (113)$$

Similarly

$$\frac{d\beta}{dt} = c\frac{\partial}{\partial \alpha}\left(\psi + \phi + \frac{M}{e}B\right) + \frac{mc}{e}\frac{d}{dt}\left(v_{\parallel}\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial \alpha}\right) + O(\epsilon^2) \qquad (114)$$

An analogous procedure is needed for dK/dt in Eq. (97). The quantity K was defined as the particle kinetic energy plus $e(\psi + \phi)$. From Eq. (29), it is true that $\frac{1}{e}\frac{dK}{dt} = 0 + O(\epsilon^2)$ when $\vec{E}_{\perp}$ and $\frac{\partial}{\partial t}$ are $O(\epsilon)$ as here. However, for present purposes it is necessary to know the $O(\epsilon^2)$ term of $\frac{1}{e}\frac{dK}{dt}$. The time rate of change of $\psi + \phi$ is correct through order $O(\epsilon^2)$ is

$$\frac{d(\psi+\phi)}{dt} = \frac{\partial}{\partial t}\left[\psi(\vec{R},t) + \phi(\vec{R},t)\right] + \dot{\vec{R}} \cdot \nabla\left[\psi(\vec{R},t) + \phi(\vec{R},t)\right] + O(\epsilon^3) \qquad (115)$$

If $\vec{E}_\perp$ and $\partial/\partial t$ are $\mathcal{O}(\epsilon)$, both ψ and ϕ are themselves $\mathcal{O}(\epsilon)$ and their partial time derivatives are $\mathcal{O}(\epsilon^2)$. From Eq. (28) the rate of change of kinetic energy is the sum of the drift velocity in the direction of $\vec{E}$, plus the induction effect $\frac{M}{e}\frac{\partial B}{\partial t}$, plus terms of $\mathcal{O}(\epsilon^2)$, at least when $\vec{E}$ and $\frac{\partial}{\partial t}$ are $\mathcal{O}(1)$. It is also true that if $\vec{E}_\perp$ and $\frac{\partial}{\partial t}$ are of $\mathcal{O}(\epsilon)$,

$$\frac{d(\text{Kinetic energy})}{e\,dt} = \dot{\vec{R}}\cdot\vec{E} + \frac{M}{e}\frac{\partial B}{\partial t}(\vec{R},t) + \mathcal{O}(\epsilon^3) \tag{116}$$

This can be verified by starting with Eq. (61) for $\vec{r}$ and calculating $\frac{d(\text{Kinetic energy})}{dt} = e\vec{v}\cdot\vec{E}(\vec{r}, t)$. Upon expanding $\vec{E}(\vec{r}, t)$ about $\vec{E}(\vec{R}_0, t)$ and time averaging over a gyration (over θ, that is), equation (116) results. The Maxwell equation $\nabla\times\vec{E} = -\partial\vec{B}/\partial t$ must also be used.

Addition of (115) and (116), with $\vec{E}$ from Eq. (108), yields

$$\frac{\dot{K}}{e} = \frac{\partial}{\partial t}\left(\psi+\phi+\frac{MB}{e}\right) + \dot{\vec{R}}\cdot\left[\vec{E}+\nabla(\psi+\phi)\right] + \mathcal{O}(\epsilon^3)$$

$$= \frac{\partial}{\partial t}\left(\psi+\phi+\frac{MB}{e}\right) + \dot{\vec{R}}_\perp\cdot\frac{1}{c}\left(\frac{\partial\beta}{\partial t}\nabla\alpha - \frac{\partial\alpha}{\partial t}\nabla\beta\right) + \mathcal{O}(\epsilon^3), \tag{117}$$

where ψ, ϕ, B, α, β are all to be considered functions of $\vec{R}$ and t. The vector $\frac{\partial\beta}{\partial t}\nabla\alpha - \frac{\partial\alpha}{\partial t}\nabla\beta$ appears frequently and will be denoted by $\vec{w}$. The drift $\dot{\vec{R}}_\perp$ is now replaced by $\frac{\vec{e}_1}{B}\times\left[-c\vec{E} + \frac{Mc}{e}\nabla B + \frac{mc}{e}v_\parallel^2\frac{\partial\vec{e}_1}{\partial s}\right]$, where $\vec{E}$ is $-\nabla(\psi+\phi)+\frac{\vec{w}}{c}$.

$$\frac{\dot{\vec R}_\perp \cdot \vec\omega}{c} = \frac{\hat e_1}{B}\times\left[c\nabla(\psi+\phi) - \vec\omega + \frac{Mc}{e}\nabla B + \frac{mc}{e} v_\parallel^2 \frac{\partial \hat e_1}{\partial s}\right]\cdot\frac{\vec\omega}{c}$$

$$= \left[\nabla\left(\psi+\phi+\frac{M}{e}B\right) + \frac{m}{e} v_\parallel^2 \frac{\partial\hat e_1}{\partial s}\right]\cdot\frac{\vec\omega\times\hat e_1}{B} \qquad (118)$$

$$\frac{\vec\omega\times\hat e_1}{B} = \frac{\partial\beta}{\partial t}\,\frac{\nabla\alpha\times\hat e_1}{B} - \frac{\partial\alpha}{\partial t}\,\frac{\nabla\beta\times\hat e_1}{B}$$

$$= \frac{\partial\beta}{\partial t}\left(\frac{\partial\vec R}{\partial\beta}\times\hat e_1\right)\times\hat e_1 - \frac{\partial\alpha}{\partial t}\left(\frac{\partial \vec R}{\partial\alpha}\times\hat e_1\right)\times\hat e_1 = -\left(\frac{\partial\vec R}{\partial\beta}\frac{\partial\beta}{\partial t}+\frac{\partial\vec R}{\partial\alpha}\frac{\partial\alpha}{\partial t}\right)_\perp \qquad (119)$$

where $\nabla\alpha$ has been replaced by $\frac{\partial\vec R}{\partial\beta}(\alpha, \beta, s, t) \times \vec B$ from Eq. (107), and similarly for $\nabla\beta$. Now $\vec R = \vec R(\alpha, \beta, s, t)$ $= \vec R\left[\alpha(\vec r, t), \beta(\vec r, t), s(r, t), t\right]$. By implicit differentiation with respect to time,

$$0 = \frac{\partial\vec R}{\partial\alpha}\frac{\partial\alpha}{\partial t} + \frac{\partial\vec R}{\partial\beta}\frac{\partial\beta}{\partial t} + \frac{\partial\vec R}{\partial s}\frac{\partial s}{\partial t} + \frac{\partial\vec R}{\partial t}. \qquad (120)$$

From (119) and (120), and the fact that $\partial\vec R/\partial s = \hat e_1$, it follows that

$$\frac{\vec\omega\times\hat e_1}{B} = \left[\hat e_1\frac{\partial s}{\partial t}(\vec r,t) + \frac{\partial\vec R}{\partial t}(\alpha,\beta, s,t)\right]_\perp = \left(\frac{\partial\vec R}{\partial t}\right)_\perp = \frac{\partial\vec R}{\partial t} - \hat e_1\,\hat e_1\cdot\frac{\partial\vec R}{\partial t}, \qquad (121)$$

and

$$\frac{\dot{\vec{R}}_\perp \cdot \vec{\omega}}{e} = \left[\nabla\left(\psi + \phi + \frac{MB}{e}\right) + \frac{m}{e} v_\parallel^2 \frac{\partial \hat{e}_1}{\partial s}\right] \cdot \left(\frac{\partial \vec{R}}{\partial t} - \hat{e}_1 \hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial t}\right)$$

$$= \nabla\left(\psi + \phi + \frac{MB}{e}\right) \cdot \frac{\partial \vec{R}}{\partial t}(\alpha, \beta, s, t) - \frac{\partial}{\partial s}\left(\psi + \phi + \frac{MB}{e}\right) \hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial t}$$

$$+ \frac{m}{e} v_\parallel^2 \frac{\partial \vec{R}}{\partial t}(\alpha, \beta, s, t) \cdot \frac{\partial \hat{e}_1}{\partial s} \qquad (122)$$

The last term in Eq. (122) is $\overbrace{(m/e) v_\parallel^2 \, \partial/\partial s(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial t})}^{\frac{m v_\parallel^2}{e} \frac{\partial(\hat{e}_1 \cdot \partial \vec{R}/\partial t)}{\partial s}}$, since $\hat{e}_1 \cdot \frac{\partial}{\partial s} \frac{\partial \vec{R}}{\partial t}$ equals $\hat{e}_1 \cdot \frac{\partial}{\partial t} \frac{\partial \vec{R}}{\partial s}$, which is $\hat{e}_1 \cdot \frac{\partial \hat{e}_1}{\partial t}$ and vanishes. In addition

$$\frac{m}{e} v_\parallel^2 \frac{\partial}{\partial s}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial t}\right) = \frac{m}{e} v_\parallel \frac{d}{dt}\left(\hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial t}\right) + O(\epsilon^2) \qquad (123)$$

Replacing $-\frac{\partial}{\partial s}\left(\psi + \phi + \frac{MB}{e}\right)$ by $-\frac{m}{e} \frac{dv_\parallel}{dt} + O(\epsilon^2)$ Eq.(72) finally gives

$$\frac{\dot{\vec{R}}_\perp \cdot \vec{\omega}}{e} = \nabla\left(\psi + \phi + \frac{MB}{e}\right) \cdot \frac{\partial \vec{R}}{\partial t} + \frac{m}{e} \frac{d}{dt}\left(v_\parallel \hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial t}\right) + O(\epsilon^2) \qquad (124)$$

and

$$\frac{1}{e}\frac{dK}{dt} = \frac{\partial}{\partial t}\left[\psi(\vec{R},t) + \phi(\vec{R},t) + \frac{MB(\vec{R},t)}{e}\right] + \nabla\left(\psi + \phi + \frac{MB}{e}\right) \cdot \frac{\partial \vec{R}}{\partial t} + \frac{m}{e}\frac{d}{dt}\left(v_\parallel \hat{e}_1 \cdot \frac{\partial \vec{R}}{\partial t}\right) + O(\epsilon^3) \qquad (125)$$

As indicated, ψ, ϕ, and B are functions of $\vec{R}$ and t, while $\vec{R} = \vec{R}(\alpha, \beta, s, t)$. If now ψ, ϕ, and B are regarded as functions of α, β, s, t as in Eq. (101), then by Eq. (102)

$$\frac{1}{e}\frac{dK}{dt} = \frac{\partial}{\partial t}\left[\Psi(\alpha,\beta,s,t) + \phi + \frac{MB}{e}\right] + \frac{m}{e}\frac{d}{dt}\left(v_{\parallel}\,\hat{e}_1\cdot\frac{\partial \vec{R}}{\partial t}\right) + O(\epsilon^2) \tag{126}$$

Finally we are prepared to evaluate $\frac{dJ}{dt}$. Comparison of $\frac{d\alpha}{dt}$ in Eq. (113) with $\frac{\partial J}{\partial \beta}$ from Eq. (99) reveals a similarity between $\frac{d\alpha}{dt}$ and the integrand of $\frac{\partial J}{\partial \beta}$. In fact

$$\frac{e}{c}\oint \frac{ds}{v_{\parallel}}\frac{d\alpha}{dt} = -e\oint \frac{ds}{v_{\parallel}}\frac{\partial}{\partial \beta}\left(\Psi + \phi + \frac{MB}{e}\right) - m\oint \frac{ds}{v_{\parallel}}\frac{d}{dt}\left(v_{\parallel}\,\hat{e}_1\cdot\frac{\partial \vec{R}}{\partial \beta}\right) \tag{127}$$

But $\frac{ds}{v_{\parallel}}$ equals dt, the time for the guiding center to traverse ds, so that the last integral in (127) is the net change in $v_{\parallel}\,\hat{e}_1\cdot\frac{\partial \vec{R}}{\partial \beta}$ over a period of the longitudinal oscillation. This change is zero. As a special case if one takes the period from one mirror reflection to the next reflection at the same end of the line of force, $v_{\parallel}$ vanishes both times and therefore the integral vanishes. Equation (127) can be written as

$$\frac{e}{c}T\langle\dot{\alpha}\rangle = \frac{\partial J}{\partial \beta}(\alpha,\beta,K,t) \,. \tag{128}$$

where $\langle\dot{\alpha}\rangle$ is the average rate of change of α over a longitudinal oscillation, the rate of change of α being caused by the time-dependent fields and the guiding center drift. The period of the longitudinal oscillation is T. Similarly

$$\frac{e}{c}T\langle\dot{\beta}\rangle = -\frac{\partial J}{\partial \alpha} \tag{129}$$

$$T\langle\dot{K}\rangle = -\frac{\partial J}{\partial t} \tag{130}$$

and by Eq. (100)

$$1 = \frac{1}{T}\frac{\partial J}{\partial K} \tag{131}$$

Equations (128) - (131) may be regarded as a third set of equations of motion; they govern the average motion from one line of force to another.[22]

Equation (97) for $\frac{dJ}{dt}$ when the guiding center is at some point (α, β, s) can, by virtue of these equations of motion, be written as

$$\left(\frac{dJ}{dt}\right)_s = \frac{eT}{c}\left[\langle\dot{\alpha}\rangle\dot{\beta}(s) - \langle\dot{\beta}\rangle\dot{\alpha}(s)\right] + T\left[\dot{K}(s) - \langle\dot{K}\rangle\right]$$

$$= \oint\frac{ds'}{v_{\parallel}'}\left\{\frac{e}{c}\left[\dot{\beta}(s)\dot{\alpha}(s') - \dot{\alpha}(s)\dot{\beta}(s')\right] + \dot{K}(s) - \dot{K}(s')\right\}, \tag{132}$$

where $v_{\parallel}'$ is the parallel velocity the guiding center has at s'. There is no reason in general that $\frac{dJ}{dt}$ should vanish. However,

$$\oint\frac{ds}{v_{\parallel}}\frac{dJ}{dt} = \oint\oint\frac{ds}{v_{\parallel}}\frac{ds'}{v_{\parallel}'}\left\{\dot{K}(s) - \dot{K}(s') + \frac{e}{c}\left[\dot{\beta}(s)\dot{\alpha}(s') - \dot{\alpha}(s)\dot{\beta}(s')\right]\right\} = 0 \tag{133}$$

The double integral vanishes because of the antisymmetry of the integrand in s and s'. Equation (133) means that although the instantaneous rate of change of J is not zero, the change averaged over a complete oscillation is zero.[23]

The above proof of the invariance of J has been carried out non-relativistically. The relativistic modifications are[20]

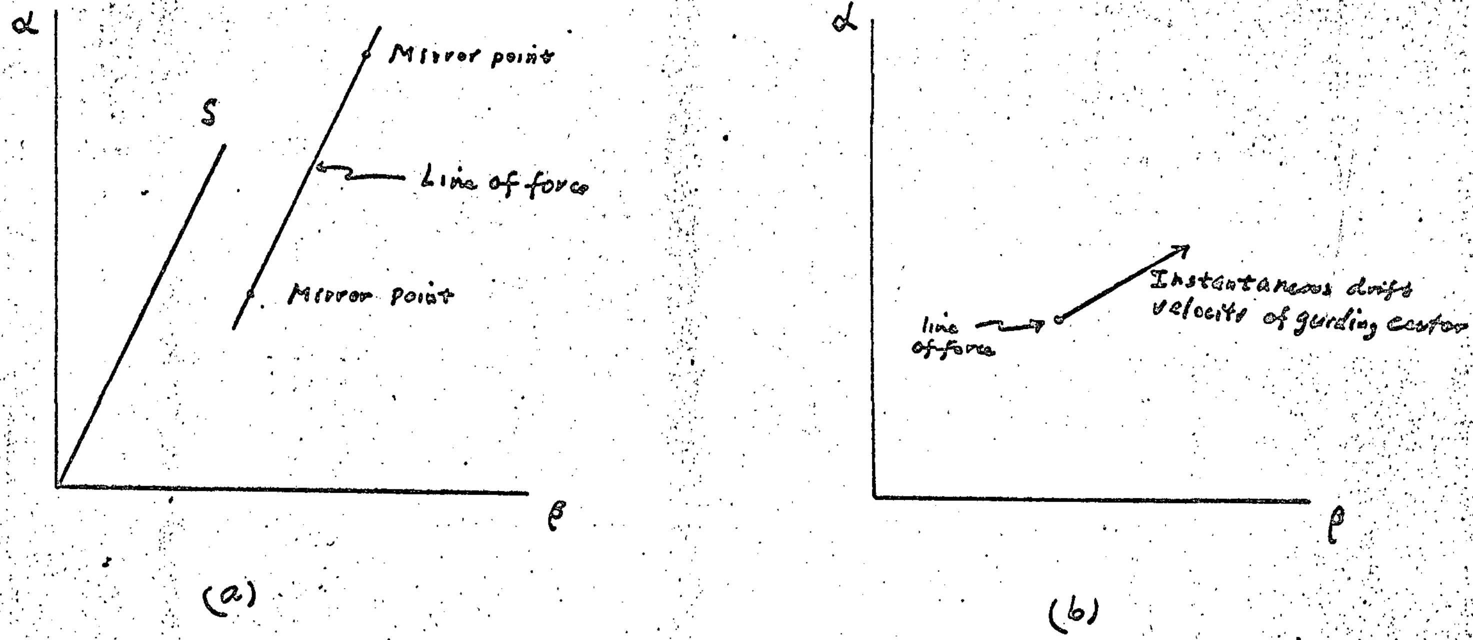
$$K = \left[p^2 c^2 + m_0^2 c^4\right]^{\frac{1}{2}} + e(\psi + \phi) \tag{134}$$

$$\frac{1}{e}\frac{dK}{dt} = \dot{\vec{R}}\cdot\vec{E} + \frac{M_r}{\gamma e}\frac{\partial B}{\partial t} + \dot{\vec{R}}\cdot\nabla(\psi+\phi) + \frac{\partial}{\partial t}(\psi+\phi)$$

$$= \frac{\dot{\vec{R}}_\perp\cdot\vec{\omega}}{c} + \frac{M_r}{\gamma e}\frac{\partial B}{\partial t} + \frac{\partial}{\partial t}(\psi+\phi) \tag{135}$$

and the relativistic guiding center equations (49) - (51). With these, a proof similar to the preceding one can be carried out, or the proof in reference (20) can be used. In either case exactly the same equations of motion (128) - (131) result.

A convenient space in which to illustrate the guiding center motion is a Cartesian (α, β, s) space as shown in Fig (8a). In this space a line of force appears as a straight line parallel to s. At any instant of time the guiding center is drifting in the α, β plane with velocity components $\dot{\alpha}$ and $\dot{\beta}$ as given in (113) and (114) and illustrated in Fig. (8b). As the guiding center moves along s, the direction



A line of force and the drift velocity in a Cartesian α, β, s, space.

Fig. 8

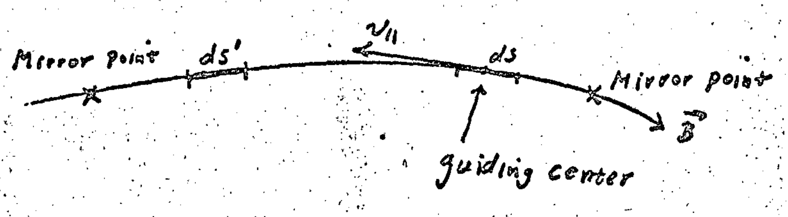
and magnitude of the drift vector in Fig. (8b) change, since $\dot{\alpha}$ and $\dot{\beta}$ are functions of s. The guiding center therefore does not always drift towards the same adjacent line of force during its rapid motion along s. In a special case it could however always drift towards the same adjacent line. Consider for example a static field, so that $\dot{K} = 0$. If the drift is towards the same adjacent line at all s, then $\dot{\alpha}/\dot{\beta}$ must be constant and the constant consequently is $\langle\dot{\alpha}\rangle/\langle\dot{\beta}\rangle$ Under these circumstances $\frac{dJ}{dt}$ is instantaneously zero, by Eq. (132). Such would be the case in an azimuthally symmetric mirror machine, where the drift is always in the azimuthal direction.

The equations (128) and (129) for $\langle\dot{\alpha}\rangle$ and $\langle\dot{\beta}\rangle$ may appear to be canonical, with J as the Hamiltonian. But they are not, for the reason that T is also a function of (α, β, K, t) If $J = J(\alpha, \beta, K, t)$ is solved as $K = K(\alpha, \beta, J, t)$, then by implicit differentiation $\frac{\partial \mathcal{J}}{\partial \beta} = -(\partial K/\partial \beta)/(\partial K/\partial \mathcal{J})$, etc., and

$$\begin{aligned} \langle\dot{\alpha}\rangle &= -\frac{c}{e}\frac{\partial K(\alpha,\beta,\mathcal{J},t)}{\partial \beta} \\ \langle\dot{\beta}\rangle &= \frac{c}{e}\frac{\partial K}{\partial \alpha} \\ \langle\dot{K}\rangle &= \frac{\partial K}{\partial t} \\ 1 &= T\frac{\partial K}{\partial \mathcal{J}} \end{aligned} \tag{136}$$

These are canonical in form, with the result that a Liouville theorem exists in $(\alpha, \beta, \mathcal{J})$ space. See reference (20).

The antisymmetry of the integrand in Eq. (133) has an interesting physical meaning. By equation (132) the contribution of ds' to the rate of change of J when the guiding center is at s can be (but does not have to be, as will become apparent below) considered to be the integrand. $\frac{ds'}{v_{\parallel}'}\{\ldots s, s'\ldots\}$. The contribution of ds' to the change in J while the guiding center traverses ds is $dt = \frac{ds}{v_{\parallel}}$ times this rate, or $\frac{ds}{v_{\parallel}}\frac{ds'}{v_{\parallel}'}\{\ldots s, s'\ldots\}$. See Fig 9. At a later time, when the $\sqrt{9}$



Cancellation of the effects of ds and ds' on J.

Fig. 9

guiding center has actually arrived at ds', the contribution of ds to the change in J while the guiding center traverses ds' is $\frac{ds'}{v_{\parallel}'}\frac{ds}{v_{\parallel}}\left\{\ldots s', s\ldots\right\}$, which is just the negative of the contribution of ds' to the change in J when the guiding center traverses ds. This cancellation is an interpretation of the antisymmetry of the integrand in Eq. (133) and holds for all pairs of arc elements ds and ds'. Such a cancellation is somewhat remarkable, especially if one recalls that the guiding center does not even drift towards the same adjacent line of force at s and s'. The remarkability disappears to some extent if one realizes that Eq. (132) can also be written as (for simplicity let $\vec{E} = \frac{\partial}{\partial t} = 0$)

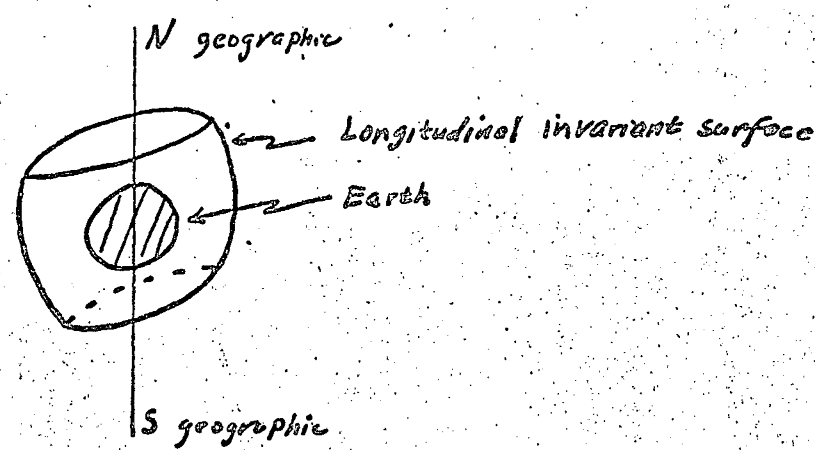
$$\left(\frac{dJ}{dt}\right)_s = e\oint \frac{ds'}{v_{\parallel}'}\left[\frac{\partial}{\partial \alpha'}\left(\frac{MB'}{e}\right)\dot{\alpha}(s) - \frac{\partial}{\partial \beta'}\left(\frac{MB'}{e}\right)\dot{\beta}(s)\right], \qquad (137)$$

where the primes on B, α, and β mean evaluated at s'. It can be written this way because the difference between $\dot{\beta}(s')$ and $\frac{\partial}{\partial \alpha'}(MB')$ is $m\frac{d}{dt}\left(v_{\parallel}\vec{e}_1\cdot\frac{\partial \vec{R}}{\partial \alpha}\right)'$ which integrates to zero when the s' integration is performed. In truth, anything can be added to $\dot{\alpha}(s')$ and $\dot{\beta}(s')$ in (132) which integrates to zero over s'. The integrand in (137) is no longer antisymmetric in s and s', and the nice physical interpretation is no longer present. The interpretation of the antisymmetry is therefore not unique. It is nevertheless true that

$$\iint \frac{ds}{v_{\parallel}} \frac{ds'}{v'_{\parallel}} \left[\frac{\partial}{\partial \alpha'} \left(\frac{MB'}{e} \right) \dot{\alpha}(s) - \frac{\partial}{\partial \beta'} \left(\frac{MB'}{e} \right) \dot{\beta}(s) \right] \text{ is zero.}$$

C. The Third or Flux Adiabatic Invariant Φ

As a particle oscillates between mirror points it drifts across lines of force on which J is constant. These lines form a surface in space. It may happen that these surfaces are closed. Such is the case in laboratory type mirror machines and probably in the geomagnetic field. A "longitudinal invariant surface" is illustrated in Fig. 10 for the earth's field.

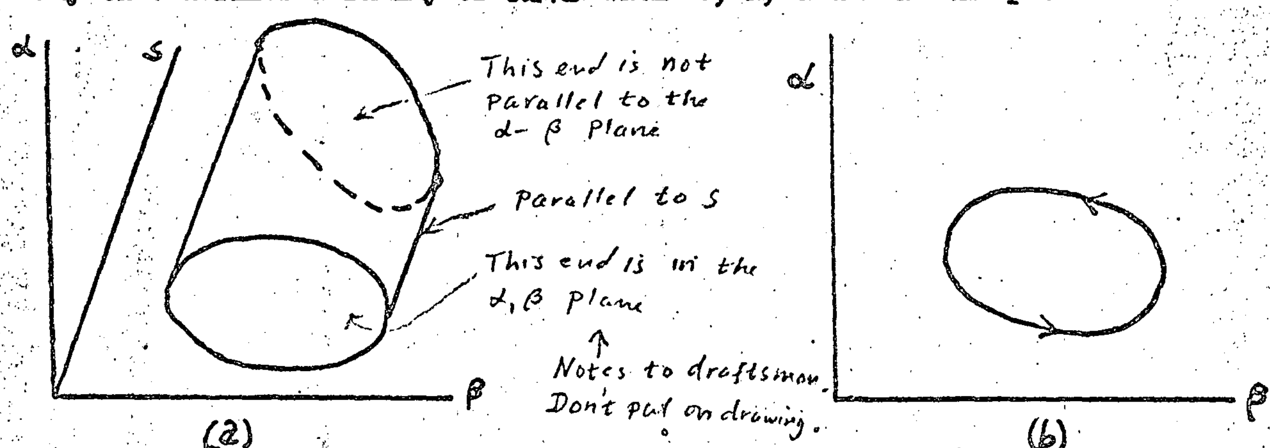


A longitudinal invariant surface in the geomagnetic field

Fig. 10

In the Cartesian α, β, s space of Fig. 8a longitudinal invariant surfaces are curved in one direction only, the elements of the surfaces being straight lines parallel to s. If the invariant surfaces are closed, they are represented by cylinders in α, β, s space as shown in Fig. 11a. The elements of a cylinder are not all of equal length because the distance between reflection points is not a constant of the guiding center motion. The intersection of a cylinder with the α, β plane appears as a closed curve in Fig. 11b.

The intersections of the cylinders with the α, β plane form a three parameter family of curves. The longitudinal invariant J is a function $J(\alpha,\beta,K,M,t)$, as can be seen from the defining integral in (96). The M dependence has been suppressed heretofore, but will now be exhibited explicitly. The equation $J = J(\alpha,\beta,K,M,t)$ can be solved as $\alpha = \alpha(\beta,J,M,K,t)$, which at any time defines a family of curves with J, M, and K as the parameters.



A longitudinal invariant surface in α, β, s space

Fig. 11

Suppose that at a given instant the time-dependence of the fields were turned off; K as well as J and M would be constant and the guiding center would precess about the corresponding (J, M, K) surface. The third adiabatic invariant Φ is the flux of $\vec{B}$ through this surface. That Φ is a constant is a trivial statement if the fields are static, for the guiding center repeatedly precesses around the same surface and the surface does not change with time. An analogous trivial statement would be that the magnetic moment is constant in a uniform static magnetic field. If, on the contrary, the fields are time dependent, the time dependence being slow compared to the precession time once around the surface, the third invariant applies and yields non-trivial information. Since K is no longer constant, the guiding center gradually moves from one J, M, K surface to another as it rapidly precesses, and is at all times on a surface with the same J, M, and Φ. Thus although one constant of the motion K has been lost, another one, Φ, replaces it.

The proof of the invariance of Φ is fortunately much simpler than that for M or J. The flux through the invariant surface is

$$\Phi(J,M,K,t) = \oint \vec{A}\cdot d\vec{l} = \oint \alpha \nabla\beta \cdot d\vec{l} = \oint \alpha(\beta, J, M, K, t)\, d\beta \tag{138}$$

where the contour is any closed curve lying on the surface (in real space) and going once around it. However $\oint \alpha\, d\beta$ is the area inside the closed curve in Fig. 11b on which the guiding center is located.

$$\frac{d\Phi(K,t)}{dt} = \frac{\partial \Phi(K,t)}{\partial K}\langle \dot{K} \rangle + \frac{\partial \Phi(K,t)}{\partial t} \tag{139}$$

The J and M dependence of Φ will not be carried explicitly since they are constants. The average, $\langle \dot{K} \rangle$, has been used in place of $\dot{K}$ because we are not interested in fluctuations of Φ over a longitudinal oscillation. Now

$$\frac{\partial \Phi}{\partial K} = \oint \frac{\partial \alpha(\beta,K,t)}{\partial K}\, d\beta \tag{140}$$

By implicit differentiation of $K = K\left[\alpha(\beta, K, t), \beta, t\right]$ we obtain

$$1 = \frac{\partial K}{\partial \alpha}(\alpha, \beta, t)\frac{\partial \alpha}{\partial K}(\beta, K, t) \tag{141}$$

$$0 = \frac{\partial K}{\partial \alpha}\frac{\partial \alpha}{\partial t} + \frac{\partial K}{\partial t} \tag{142}$$

So,

$$\frac{\partial \Phi}{\partial K} = \oint \frac{d\beta}{\partial K/\partial \alpha} = \frac{c}{e}\oint \frac{d\beta}{\langle \dot{\beta} \rangle}, \tag{143}$$

the last equality being via Eq. (136). But $\oint \frac{d\beta}{\langle \dot{\beta} \rangle}$ equals $\oint dt$, which is the time T_p for the guiding center to precess once around the surface. Therefore $\frac{\partial \Phi}{\partial K} = \frac{c}{e} T_p$, and this the analog of Eq. (131). The last term in (139) is

$$\frac{\partial \Phi}{\partial t} = \oint \frac{\partial \alpha}{\partial t}(\beta, K, t)\, d\beta = -\oint \frac{\partial K/\partial t}{\partial K/\partial \alpha}\, d\beta. \tag{144}$$

Replacing $\frac{\partial K}{\partial \alpha}$ by $\langle \dot{\beta} \rangle$ again and $\frac{\partial K}{\partial t}$ by $\langle \dot{K} \rangle$ from Eq. (136), we have

$$\frac{\partial \Phi}{\partial t} = -\frac{c}{e}\oint \frac{\langle \dot{K} \rangle\, d\beta}{\langle \dot{\beta} \rangle} = -\frac{c}{e}\oint \langle \dot{K} \rangle\, dt = -\frac{c T_p}{e}\langle\langle \dot{K} \rangle\rangle \tag{145}$$

where $\langle\langle \dot{K} \rangle\rangle$ is the average of $\langle \dot{K} \rangle$ over a precession period. Equation (145) is the analog of Eq. (130). The instantaneous rate of change of Φ is therefore

$$\frac{d\Phi}{dt} = \frac{c T_p}{e}\left[\langle \dot{K} \rangle - \langle\langle \dot{K} \rangle\rangle\right] \tag{146}$$

As with $\frac{d\mathcal{J}}{dt}$, we again find that $\frac{d\Phi}{dt} \neq 0$, but that $\left\langle \frac{d\Phi}{dt} \right\rangle$, the average rate of change over a precession period, does vanish, again by the antisymmetry of the integrand in

$$\left(\frac{d\Phi}{dt}\right)_{\sigma} = \frac{c}{e} \oint \frac{d\sigma'}{v'_{\alpha\beta}} \left[\langle \dot{K} \rangle_{\sigma} - \langle \dot{K} \rangle_{\sigma'} \right], \tag{147}$$

sigma

where $d\sigma'$ is the element of arc length and $v'_{\alpha\beta}$ is the velocity at σ' about the closed curve in Fig. 11b.

The new and final set of equations of motion is

$$\langle\langle \dot{K} \rangle\rangle = \frac{-e}{cT_p} \frac{\partial \Phi}{\partial t}(\mathcal{J}, M, K, t) \tag{148}$$

$$1 = \frac{e}{cT_p} \frac{\partial \Phi}{\partial K} \tag{149}$$

l.c. p

The averaging processes used to establish the three adiabatic invariants and sets of equations of motion are summarized in the following diagram.

Equation of motion of a particle (1)

$\xrightarrow{\text{Average over the gyration}}$

Guiding center equation of motion (17) and (20), and the adiabatic invariant M

$\downarrow$ Average over the longitudinal oscillation

Average guiding center equations of motion (136), and the adiabatic invariant J

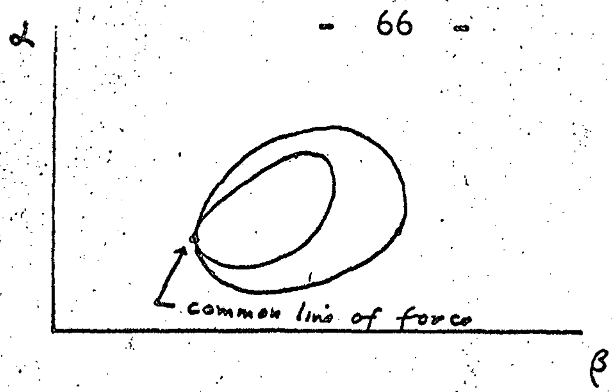
$\xrightarrow{\text{Average over the precessional motion}}$

Equations of motion (148) and (149), and the adiabatic invariant Φ

The averaging process used to obtain the three adiabatic invariants and equations of motion.

Fig. 12

A few facts about the family of longitudinal invariant surfaces should be made clear now. In the first place, at any instant of time they are not simply nested but intersect in a very complex fashion; this is to be expected because there are three parameters, not one. For example, two particles with the same K which are oscillating along the same line of force, but which have different mirror points, have different M and consequently by Eq. (96) different J. There is no reason for these two particles to traverse the same line of force anywhere else. Fig. 13 illustrates how their two invariant surfaces might appear in the α, β plane. (Of course



Longitudinal invariant surfaces are not simply nested and may intersect.

Fig. 13

they may intersect elsewhere; there is nothing to prohibit it in general.) An infinite number of surfaces therefore intersect along any line of force; if particles are injected on one line of force with a distribution of mirror points, they spread into a layer of finite thickness elsewhere.

Secondly, in the α, β plane a surface denoted by fixed values of J, M, K will move in time as indicated schematically in Fig. 14, by curves I and II. It is clear that the surface must move, because α equals $\alpha(\beta, J, M, K, t)$ and therefore the value of α for a given β varies with t at fixed J, M, K. The fact that a constant J, M, K surface moves with time does not mean that the lines of force move in α, β, s space. In fact, they can be considered fixed, even though they move in real space. The velocity of a magnetic line of force is not a physically observable quantity, and therefore must be defined. The usual definition is the "flux preserving" one: (24) let an arbitrary closed loop be drawn in space and let the velocity $\vec{u}$ of each arc element of the loop be such that the flux of $\vec{B}$ through the loop is constant in time. The velocity of the line of force at each point is defined as $\vec{u}$. It is easy to show that $\vec{u}$ must satisfy $\nabla \times (\vec{E} + \frac{\vec{u}}{c} \times \vec{B}) = 0$. The velocity $\frac{\vec{W} \times \hat{e}_1}{B}$ is such a velocity, where $\vec{W} = \frac{\partial\beta}{\partial t}\nabla\alpha - \frac{\partial\alpha}{\partial t}\nabla\beta$, as defined previously.

$$\vec{E} + \frac{\vec{W} \times \hat{e}_1}{\cancel{c}B} \times \vec{B} = \vec{E} - \frac{\vec{W}}{c} . \tag{150}$$

By Eq. (108), $\vec{E}$ equals $-\nabla(\psi + \phi) + \frac{\vec{W}}{c}$, so

$$\vec{E} + \frac{\vec{W} \times \hat{e}_1}{\cancel{c}B} \times \vec{B} = -\nabla(\psi + \phi) \tag{151}$$

and indeed the curl vanishes. It has thus been established that $\frac{\vec{W} \times \hat{e}_1}{B}$ can be taken as the velocity of a line of force. It will now be proven that if an observer moves at this velocity, the total rate of change of α (and β) he observes is zero. In other words the (α, β) label on a line of force is not changed by the motion of the line and all lines are fixed in α, β, s space consequently. The rate of change of α under the time dependence and the velocity of the line of force is

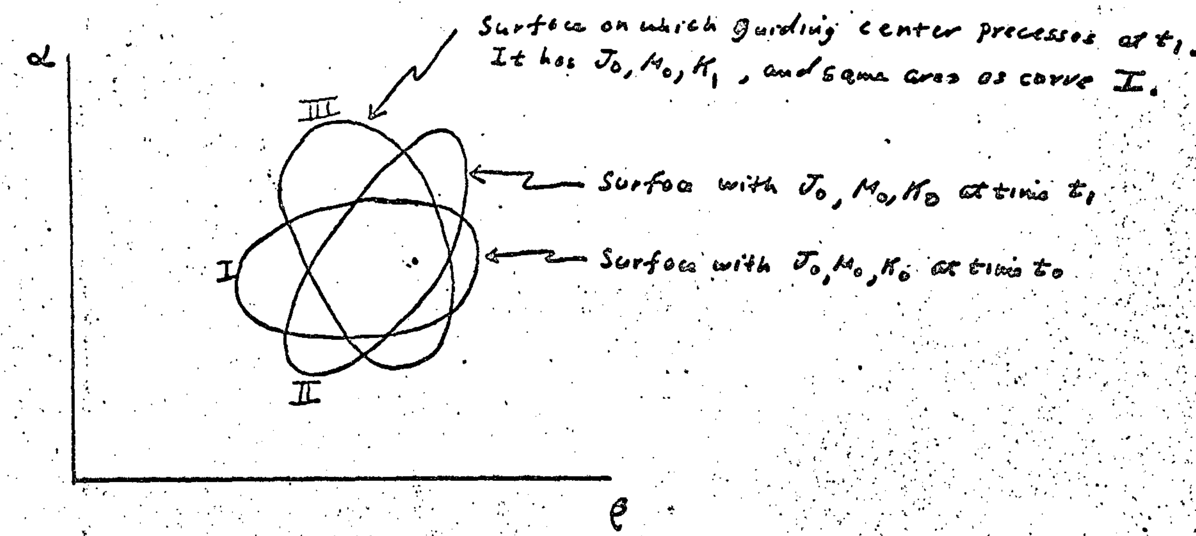
$$\frac{\partial\alpha}{\partial t} + \frac{\vec{w} \times \hat{e}_1}{B} \cdot \nabla\alpha = \frac{\partial\alpha}{\partial t} + \left(\frac{\partial\beta}{\partial t}\nabla\alpha - \frac{\partial\alpha}{\partial t}\nabla\beta\right) \times \frac{\hat{e}_1}{B} \cdot \nabla\alpha$$

$$= \frac{\partial\alpha}{\partial t} - \frac{\partial\alpha}{\partial t}\,\frac{\nabla\beta \times \hat{e}_1 \cdot \nabla\alpha}{B}$$

$$= 0 , \tag{152}$$

because $B = \hat{e}_1 \cdot \vec{B} = \hat{e}_1 \cdot (\nabla\alpha \times \nabla\beta)$. It is therefore clear that although lines of force do not move in (α, β, s) space, the locus of lines which form an invariant surface with constant J, M, K does. In Fig.14 curves I and II are the projections on the α, β plane of two surfaces having the same J, M, and K at different times t_0 and t_1.

A third point is: the actual surface in α, β, s space on which the guiding center precesses at time t_1 is neither curve I or II, but some third curve represented schematically by III. This is necessarily so because the particle's K has changed and is no longer K_0. Curve III has the same enclosed area as curve I because the flux invariant Φ is this area. It is true that if the guiding center were on curve I at time t_1, it would have the same Φ as at t_0, but its J would have to be different.



Longitudinal invariant surfaces at different times.

Fig. 14

D. Adiabatic Invariants to Higher Order

It is apparent from the foregoing extensive analysis that to proceed to higher orders in the three adiabatic invariant series by use of these direct methods would be laborious. Indeed, even to guess what $\epsilon M'$, $\epsilon J'$, and $\epsilon \Phi'$ are would tax one's imagination. Proofs of the type presented for M, J, and Φ are valuable in producing a physical picture of what is happening, but must be abandoned in favor of a systematic, canonical method like that of Gardner[18] or Kruskal[19] to go to the next higher order. Probably nothing is lost in the way of physical understanding either, since effects which are second order in the gyration radius are difficult to visualize anyway.

In Gardner's procedure, to obtain each new term in one of the adiabatic invariant series a canonical transformation must be made from the variables used in the previous order. A prescription is available for obtaining the generating function of each successive transformation. At any order, all the preceding canonical transformations must be inverted to express the adiabatic invariant series thus far in terms of the original variables (velocity and position of the particle). In practice this may be very laborious, but at least a deductive method is available and no guessing of the higher order terms is required. To the author's knowledge, $\epsilon M'$ is the only higher term in any of the three series that has been worked out. It is in reference (6) and is in a slightly different (and more useful, it will turn out) form than one might expect. The series correct through $\mathcal{O}(\epsilon)$ is given as (for the case where $\hat{E} = 0$)

$$\text{const.} = \frac{v_\perp^{\,2}}{B(\vec{r})} - \frac{\epsilon}{B^3}\left\{ (v^2\hat{e}_1 + \vec{v}\,\vec{v}\cdot\hat{e}_1)\cdot\left[(\vec{v}\times\hat{e}_1)\cdot\nabla\right]\vec{B} \right.$$

$$\left. + (\hat{e}_1\cdot\vec{v})(\nabla\times\vec{B})\cdot\left(\frac{v_\perp^{\,2}}{2}\hat{e}_1 + 2\vec{v}_\perp\,\hat{e}_1\cdot\vec{v}\right)\right\} + \mathcal{O}(\epsilon^2) \tag{153}$$

where $\vec{v}_\perp$ is the instantaneous particle velocity perpendicular to $\hat{e}_1(\vec{r})$, not $\hat{e}_1(\vec{R}_0)$. Also note that in the first term B is evaluated at $\vec{r}$ and not $\vec{R}_0$. In the $\mathcal{O}(\epsilon)$ term, it does not matter whether the field and $\hat{e}_1$ are at $\vec{r}$ or $\vec{R}_0$, nor does it matter whether $\vec{v}_\perp$ is perpendicular to $\hat{e}_1(\vec{r})$ or $\hat{e}_1(\vec{R}_0)$, since the difference is $\mathcal{O}(\epsilon^2)$. The form (153) of the series is useful for comparison with a numerical integration of the equations of motion because the result of such a computation would most likely be the particle velocity and position as functions of time. The fields at the particle position would therefore already be present in the code, whereas the fields at $\vec{R}_0$ would require an auxiliary computation.

The first term of Eq. (153) can be converted to velocities perpendicular to $\hat{e}_1(\vec{R}_0)$ and to fields at the guiding center $\vec{R}_0$ as follows:

$$\begin{aligned}
\vec{v}_{\perp(\vec{r})} &= \vec{v} - \hat{e}_1(\vec{r})\,\hat{e}_1(\vec{r})\cdot\vec{v} = \vec{v} - \hat{e}_1(\vec{r})v_{\parallel}(r)\\
&= \vec{v} - \left[\hat{e}_1(\vec{R}_0) + \vec{\rho}\cdot\nabla\,\hat{e}_1(\vec{R}_0)\right]\left[\hat{e}_1(\vec{R}_0) + \vec{\rho}\cdot\nabla\hat{e}_1(\vec{R}_0)\right]\cdot\vec{v}\\
&= \vec{v}_\perp(\vec{R}_0) - v_{\parallel}(\vec{R}_0)\,\vec{\rho}\cdot\nabla\hat{e}_1 - \hat{e}_1\,\vec{v}\cdot(\vec{\rho}\cdot\nabla)\hat{e}_1 + \mathcal{O}(\epsilon^2)
\end{aligned} \tag{154}$$

$$v_{\perp(\vec{r})}^2 = v_\perp^{\,2}(\vec{R}_0) - 2v_{\parallel}(\vec{R}_0)\,\vec{v}\cdot(\vec{\rho}\cdot\nabla)\hat{e}_1 + \mathcal{O}(\epsilon^2) \tag{155}$$

where the $\vec{r}$ or $\vec{R}_0$ following the $\perp$ or $\parallel$ subscript signifies the direction to which the velocity component is perpendicular or parallel. The vector $\vec{\rho}$ is $(\epsilon\vec{R}_{10}\,e^{i\theta} + c.c.) = \dfrac{\hat{e}_1\times\vec{v}_\perp}{\omega} + \mathcal{O}(\epsilon^2)$. In addition $\dfrac{1}{B(\vec{r})}$ must be transformed:

$$\frac{1}{B(\vec{r})} = \frac{1}{B(\vec{R}_0)} - \frac{\vec{\rho} \cdot \nabla B}{B^2} + \mathcal{O}(\epsilon^2). \qquad (156)$$

With the substitutions (155) and (156) and some vector algebra, Eq. (153) becomes:

$$\text{constant} = \frac{v_\perp^2(\vec{R}_0)}{B(\vec{R}_0)} - \epsilon \frac{v_\parallel}{B^2} \vec{v}_\perp \cdot \left[(\hat{e}_1 \times \vec{v}_\perp) \cdot \nabla \hat{e}_1\right] + \epsilon \frac{2v_\parallel^2}{B^3} (\hat{e}_1 \times \vec{v}_\perp) \cdot \nabla B$$

$$- \frac{\epsilon}{B^3} v_\parallel (\nabla \times \vec{B}) \cdot \left(\frac{v_\perp^2 \hat{e}_1}{2} + 2\vec{v}_\perp v_\parallel \right) + \mathcal{O}(\epsilon^2). \qquad (157)$$

It may be observed that if $v_\parallel$ is zero at all times, then $v_\perp^2(\vec{R}_0)/B(\vec{R}_0)$ is constant to one higher order in ϵ.

There is a third form in which the series may be written. The particle velocity may be eliminated from (157) and the $\vec{R}_n$ of the series (61) used instead. Differentiation of the asymptotic series (61) for $\vec{r}$ gives $\vec{v}$ and therefore $\vec{v}_\perp$. In Eq. (73) $\vec{v}$ was obtained correct through zero order in ϵ. Now, however, $\vec{v}_\perp$ is needed correct through $\mathcal{O}(\epsilon)$, because the first term $\frac{v_\perp^2(\vec{R}_0)}{B(\vec{R}_0)}$ will then yield more $\mathcal{O}(\epsilon)$ terms. It should not be overlooked in differentiating the series for $\vec{r}$ that $\epsilon^2 \vec{R}_{20} e^{2i\theta}$ terms ~~in the $\vec{r}$ series~~ contribute $\mathcal{O}(\epsilon)$ terms to $\vec{v}$, as a result of differentiating the exponential. When $\vec{v}$ is substituted into (157), the result is ~~(for the special case where $\nabla \times \vec{B} = 0$)~~

"one one" parallel sign

$$\text{constant} = \cancel{2B\vec{R}_{10} \cdot \vec{R}_{10}^* + 2\epsilon \left[\left(B\vec{R}_{11} + \frac{3}{2} i v_\parallel \vec{R}_{10} \cdot \nabla \hat{e}_1 + \frac{i\dot{B}}{B} \vec{R}_{10} \right) \cdot \vec{R}_{10}^* + c.c. \right] + \mathcal{O}(\epsilon^2)}, \qquad (158)$$

$$= 2B\left(\vec{R}_{10} + \epsilon \vec{R}_{11}\right) \cdot \left(\vec{R}_{10}^* + \epsilon \vec{R}_{11}^*\right)\left(1 + \frac{\epsilon v_\parallel \hat{e}_1 \cdot \nabla \times \vec{B}}{B^2}\right) + \mathcal{O}(\epsilon^2)$$

where $\vec{R}_{11}$ is the second coefficient in $\vec{R}_1 = \vec{R}_{10} + \epsilon \vec{R}_{11} + \ldots$. This is certainly the least useful form for comparison with numerical computation. Because $\vec{v}$ contains $e^{in\theta}$ it may seem surprising that there are no exponentials of this type left in (158). The reason is that the adiabatic invariant series is an integral is an integral over θ and therefore θ cannot appear. This will become apparent in the discussion in Section E of adiabatic invariants of systems of coupled differential equations.

The statement is often made that the magnetic moment $\frac{mv_\perp^2}{2B}$ is "constant to all orders". The meaning of this statement is that if a particle goes from one region in space and time where $\vec{E}$ and $\vec{B}$ are constant to another such region via time and space dependent fields, that the change in $v_\perp^2/B$ between the initial and final states vanishes faster than any power of ϵ, even though the change at intermediate times may not. This conclusion follows from the fact that $\mathcal{O}(\epsilon)$ and higher terms in the magnetic moment series vanish in uniform fields. That the $\mathcal{O}(\epsilon)$ term vanishes when $\vec{B}$ is constant can be seen in (153) or (157). Higher terms always contain field gradients and vanish in uniform fields. If the magnetic moment series were convergent instead of asymptotic, the change in $v_\perp^2/B$ would be rigorously zero, but because of the asymptotic property, all we know is that the change goes to zero faster than any power of ϵ. It is frequently suggested that the change in $v_\perp^2/B$ is proportional to $\exp(-\text{constant}/\epsilon)$, which does indeed vanish faster than any power of ϵ, but the change certainly could be some other function of ϵ that has no power series expansion in ϵ.

E. The Adiabatic Invariants of Singly Periodic Systems

A more general theory of asymptotic solutions and adiabatic invariants has been given by Kruskal[19] for coupled differential equations of a certain type. Let

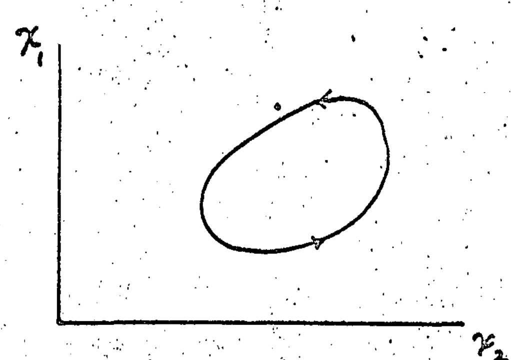
$$\frac{dx_1}{ds} = f_1 \, (x_1, \, x_2, \, \ldots \, x_N, \; \epsilon)$$

$$\frac{dx_N}{ds} = f_N \, (x_1 \ldots x_N, \; \epsilon) \tag{159}$$

be a set of coupled first order differential equations in which the independent variable s does not appear explicitly on the right hand side; ϵ is a parameter. The set of equations may be written vectorically as

$$\frac{d\vec{x}}{ds} = \vec{f}(\, \vec{x} \, , \; \epsilon \,), \tag{160}$$

where $\vec{x}$ is the vector $(x_1, \, x_2 \cdots x_N)$. Distinguish between the independent variable s used here and distance along the line of force; we will use the notation of reference (19) insofar as possible. Let the system (160) have the property that solutions of $\frac{d\vec{x}}{ds} = f(\vec{x}, \, 0)$ are simple closed curves in $\vec{x}$ space, as illustrated in Fig. 15 for two dimensions. All the components of $\vec{x}$ are periodic with the same fundamental frequency. It is also assumed that $\vec{f}$ possesses



The unperturbed solution is a closed curve in $\vec{x}$ space.

Fig. 15

a power series expansion in ϵ. Under these conditions there exists a transformation $\vec{x} = \vec{x} \, (\vec{z}, \, \theta, \, \epsilon)$, where $\vec{x}$ is periodic in θ, such that the transformed equations (160) assume the form

$$\frac{d\vec{z}}{ds} = \epsilon\vec{h}\,(\vec{z},\ \epsilon)$$

$$\frac{d\theta}{ds} = \omega(\vec{z},\ \epsilon)\ . \tag{161}$$

The vector $\vec{z}$ has only N-1 components. It is not immediately obvious that there even exists such a transformation, which makes $\vec{h}$ and ω independent of θ. The first part of reference (19) is devoted to proving this and to developing the step by step recursion method which gives the transformation $\vec{x}\,(\vec{z},\ \theta,\ \epsilon)$ and the functions $\vec{h}$ and ω as series in ϵ. According to Kruskal, it is possible to prove that if the solution $\vec{z}(s,\ \epsilon)$, $\theta(s,\ \epsilon)$ of (161) is substituted into $\vec{x} = \vec{x}\,(\vec{z},\ \theta,\ \epsilon)$, the result is an asymptotic (in ϵ) approximation to the exact solution of (159).

The equation of motion of a charged particle may be written in the form (159). If $\vec{x}$ were the six-component vector $(\vec{v},\ \vec{r})$, the equation of motion would not be of the required form because the independent variable t would appear or the right hand sides when the fields are time dependent. But if time is treated as a seventh dependent variable via the substitution $t = \epsilon s$ (where $\epsilon = m/e$), seven equations of motion of the required form result. They are

$$\frac{d\vec{v}}{ds} = \vec{E}\,(\vec{r},\ t) + \frac{\vec{v}}{c} \times \vec{B}(\vec{r},\ t)$$

$$\frac{d\vec{r}}{ds} = \epsilon\vec{v}$$

$$\frac{dt}{ds} = \epsilon \tag{162}$$

When $\epsilon = 0$, this set reduces to

$$\frac{d\vec{v}}{ds} = \vec{E} + \frac{\vec{v}}{c} \times \vec{B}$$

$$\frac{d\vec{r}}{ds} = 0$$

$$\frac{dt}{ds} = 0 \quad . \tag{163}$$

Because $d\vec{r}/ds$ and dt/ds vanish, $\vec{r}$ is a constant $\equiv \vec{r}_0$, and t is a constant $\equiv t_0$, so $\frac{d\vec{v}}{ds} = \vec{E}(\vec{r}_0, t_0) + \frac{\vec{v}}{c} \times \vec{B}(\vec{r}_0, t_0)$. Because $\vec{E}$ and $\vec{B}$ are constants when their arguments are constant, the zero order motion is that of a particle in a uniform field. In $\vec{v}$ space, it is a circle with center at $\frac{c\vec{E}(\vec{r}_0, t_0) \times \vec{B}(\vec{r}_0, t_0)}{B^2}$, provided $E_{\parallel}$ is assumed of $\mathcal{O}(\epsilon)$ as usual, and therefore the motion in the complete seven-dimensional space is periodic when $\epsilon = 0$. Consequently the theory applies and would give the asymptotic series for $\vec{r}$ in Eq. (61) and its derivative for $\vec{v}$.

Thus far the theory has produced no adiabatic invariants, only asymptotic solutions. If, however, the equations (159) are of canonical form, there exists one or more adiabatic invariants in addition to asymptotic solutions. Let the vector $\vec{x}$ be $(\vec{p}, \vec{q})$ and s be t. The canonical equations are

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}(\vec{p}, \vec{q}, \epsilon)$$

$$\dot{q}_i = \frac{\partial H}{\partial p_i}(\vec{p}, \vec{q}, \epsilon), \tag{164}$$

and these are of the required form, ~~[illegible]~~. If the Hamiltonian is time-dependent, time and energy can be used as conjugate variables and the number of degrees of freedom increased by one. Given the necessary conditions on the periodicity of the zero order solutions, there exist the transformations $\vec{p} = \vec{p}(\vec{z}, \theta, \epsilon)$ and $\vec{q} = \vec{q}(\vec{z}, \theta, \epsilon)$

which are periodic in θ and are such that the transformed canonical equations are

$$\frac{d\vec{z}}{dt} = \epsilon \vec{h} (\vec{z}, \epsilon)$$

$$\frac{d\theta}{dt} = \omega (\vec{z}, \epsilon) . \tag{165}$$

The solutions $\vec{z} = \vec{z} (t, \epsilon)$ and $\theta = \theta (t, \epsilon)$ of (165) can be substituted to give $\vec{p} = \vec{p} \left[\vec{z} (t, \epsilon), \theta(t, \epsilon), \epsilon \right]$ as a series in ϵ, and likewise for $\vec{q}$. These are not exact solutions of the canonical equations, only asymptotic solutions.

The adiabatic invariant is

$$I(\vec{z}, \epsilon) = \int_0^1 d\theta . \vec{p} (\vec{z}, \theta, \epsilon) \cdot \frac{\partial \vec{q}}{\partial \theta} (\vec{z}, \theta, \epsilon) \, d\theta = \oint_{\vec{z} = \text{constant}} \vec{p} \cdot d\vec{q} \tag{166}$$

where to be specific the angle variable θ is assumed to have the period 0 to 1. The invariant I will in fact be obtained as a series in ϵ, since $\vec{p}$ and $\vec{q}$ are themselves series in ϵ. The proof of the adiabatic invariance of I consists of showing that $\frac{dI}{dt}$ is zero, or that $\frac{d}{dt} (I_0 + \epsilon I_1 + \dots \epsilon^n I_n) = 0 + \mathcal{O}(\epsilon^{n+1})$.

$$\frac{dI}{dt}(t, \epsilon) = \frac{dz_j (t, \epsilon)}{dt} \frac{\partial I}{\partial z_j} (\vec{z}, \epsilon)$$

$$= \int_0^1 d\theta \frac{dz_j}{dt} \left(\frac{\partial p_k}{\partial z_j} \frac{\partial q_k}{\partial \theta} + p_k \frac{\partial^2 q_k}{\partial \theta \, \partial z_j} \right), \tag{167}$$

where sums are to be taken over repeated indices. The factor $\frac{dz_j}{dt}$ is not a function of θ and therefore has been put under the integral. The second term in the integrand may be integrated by parts to give

$$\frac{dI}{dt} = \int_0^1 d\theta \, \frac{dz_j}{dt} \left(\frac{\partial p_k}{\partial z_j} \frac{\partial q_k}{\partial \theta} - \frac{\partial q_k}{\partial z_j} \frac{\partial p_k}{\partial \theta} \right), \tag{168}$$

where the extra term $\int_0^1 d\theta \, \frac{\partial}{\partial \theta} \left(p_k \frac{\partial q_k}{\partial z_j} \right)$ vanished because $\vec{p}$ and $\vec{q}$ are periodic in θ. Let us now show that the integrand in (168) is simply $\frac{\partial H}{\partial \theta}$.

$$\frac{\partial H(\vec{p}(\vec{z}, \theta, \epsilon), \vec{q}(\vec{z}, \theta, \epsilon)\ \epsilon)}{\partial \theta} = \frac{\partial H}{\partial p_k} \frac{\partial p_k}{\partial \theta} + \frac{\partial H}{\partial q_k} \frac{\partial q_k}{\partial \theta} = \frac{dq_k}{dt} \frac{\partial p_k}{\partial \theta} - \frac{dp_k}{dt} \frac{\partial q_k}{\partial \theta}$$

$$= \left(\frac{\partial q_k}{\partial z_j} \frac{dz_j}{dt} + \frac{\partial q_k}{\partial \theta} \frac{d\theta}{dt} \right) \frac{\partial p_k}{\partial \theta}$$

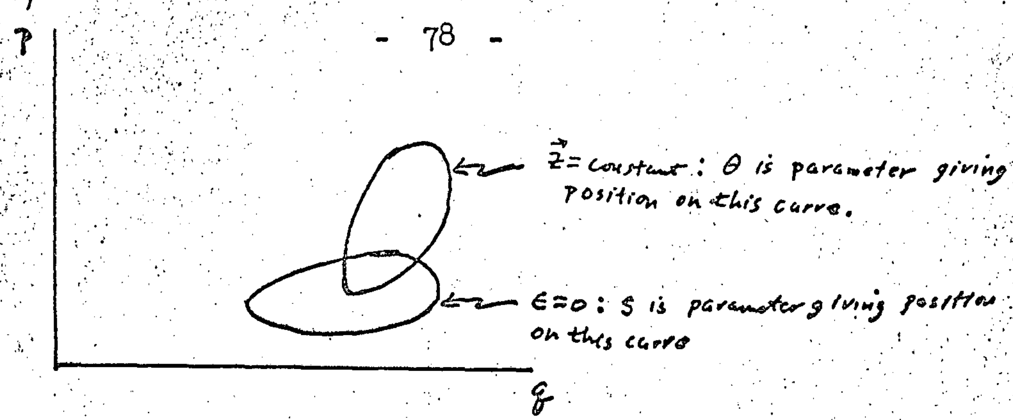
$$- \left(\frac{\partial p_k}{\partial z_j} \frac{dz_j}{dt} + \frac{\partial p_k}{\partial \theta} \frac{d\theta}{dt} \right) \frac{\partial q_k}{\partial \theta}$$

$$= \left(\frac{\partial q_k}{\partial z_j} \frac{\partial p_k}{\partial \theta} - \frac{\partial p_k}{\partial z_j} \frac{\partial q_k}{-\partial \theta} \right) \frac{dz_j}{dt} . \tag{169}$$

Therefore

$$\frac{dI}{dt} = - \int_0^1 d\theta \, \frac{\partial H}{\partial \theta} = 0; \tag{170}$$

the integral vanishes due to the periodicity of $\vec{p}$ and $\vec{q}$ in θ. It should be emphasized that the adiabatic invariant is the integral of $\vec{p} \cdot d\vec{q}$ around a closed curve on which $\vec{z}$ is constant in $\vec{p}, \vec{q}$ phase space, and not about the closed curve representing the zero order (i.e. $\epsilon = 0$) periodic motion. The difference between the two curves is shown schematically in Fig. 16 for a system with one degree of freedom. To lowest order in ϵ it does turn out however that the $\vec{z}$ = constant curve is the same as the unperturbed curve. It is in calculating higher order terms in the series for I that the difference appears.



The unperturbed path in phase space differs from the constant $\vec{z}$ curve.
Fig. 16

The constant $\vec{z}$ curves are known to be closed due to the periodicity of $\vec{x} = \vec{x}\,(\vec{z},\ \theta,\ \epsilon)$ in θ.

In the present situation, where the coordinates all have the same fundamental frequency in the unperturbed state, the adiabatic invariant (166) has turned out to be the sum of action variables rather than any one of them alone. Of course for a system with one degree of freedom there is no distinction. The integral for I can be written

$$I = \iint d\vec{p} \cdot d\vec{q}, \qquad (171)$$

where the double integral is over the area of the $\vec{z}$ = constant curve, and this is one of the Poincaré invariants. It is not surprising that I should be a Poincaré invariant: the value of the adiabatic invariant for a given system should be independent of the canonical variables used in (164) and (166), and indeed the Poincaré integrals are invariant under canonical transformations, whereas each term of $\int_0^1 d\theta\ \vec{p} \cdot \frac{d\vec{q}}{\partial\theta}$ is not.

It is now clear that the number of degrees of freedom is not necessarily the number of adiabatic invariants, but is the number of terms in the adiabatic invariant integral. A system with say two degrees of freedom might have just the one adiabatic invariant.

On the other hand, there may be additional adiabatic invariants. It is possible to prove that the Poisson bracket $[\theta, I]$ equals unity. The proof is a little lengthy to be included here, but it is outlined in reference (19) in sufficient detail. The fact that $[\theta, I] = 1$ means that θ and I can be used as new conjugate variables in a canonical transformation of the form $(p_1 \ldots p_N; q_1 \ldots q_N) \rightarrow (P_1 \ldots P_{N-1}, I; Q_1 \ldots Q_{N-1}, \theta)$. Let $H' (\vec{Q}, \vec{P}, I, \epsilon)$ be the transformed Hamiltonian. As indicated the new Hamiltonian is not a function of θ because $\dot{I} = -\partial H'/\partial \theta$ and $\dot{I}$ is zero. Thus H' will have one less degree of freedom than H, and I will be merely a parameter. If now the new canonical equations

$$\dot{Q}_i = \frac{\partial H'}{\partial P_i} (P_1 \ldots P_{N-1}, Q_1 \ldots Q_{N-1}, \epsilon)$$

$$\dot{P}_i = \frac{\partial H'}{\partial Q_i} \tag{172}$$

have solutions which are periodic in $\vec{P}$, $\vec{Q}$ space when ϵ or some other smallness parameter is zero, the entire process from eqn. (164) on can be repeated and a second adiabatic invariant $\oint \vec{P} \cdot d\vec{Q}$ obtained. For the case of the charged particle this would be the longitudinal invariant $\oint p_{\parallel} ds$. If it so happens that the new canonical equations again have periodic solutions, there will be a third adiabatic invariant, the flux invariant Φ in the case of the charged particle. Since the number of degrees of freedom is reduced by one for each adiabatic invariant, it is now clear why for singly periodic systems the number of adiabatic invariants is at most equal to the number of degrees of freedom, and will be less if at any step the new canonical equations corresponding to (172) fail to have periodic solutions in the unperturbed state. In connection with the charged particle, this would be found the case if the motion along the line of force were not periodic.

"ioto"

By use of the four dimensional relativistic Hamiltonian[25] $H^{(\tau)} = -\frac{1}{2mc^2}\eta_{\iota\kappa}(p^{\iota} - \frac{e}{c}\varphi^{\iota})(p^{\kappa} - \frac{e}{c}\varphi^{\kappa})$ the method of this section would furnish an alternate (to Ref. 11) way to study the adiabatic motion of relativistic particles. To the author's knowledge this has not been done.

The preceding proof that $\frac{dI}{dt}$ is zero would establish I as a rigorous (not adiabatic) invariant of the motion provided the series for $\vec{p}$ and $\vec{q}$ were rigorous solutions of the equations of motion instead of non-convergent asymptotic solutions. It is the asymptotic property of the series for $\vec{p}$ and $\vec{q}$ that makes I an adiabatic invariant. The theorem establishing a rigorous invariant would be: Given (1) two functions $\vec{p}(\vec{z}, \theta)$ and $\vec{q}(\vec{z}, \theta)$, which are periodic in θ (2) that $\theta = \theta(t)$ and $\vec{z} = \vec{z}(t)$ are such functions of time that $\vec{p}(\vec{z}(t), \theta(t))$ and $\vec{q}(\vec{z}(t), \theta(t))$ are solutions of the equations of motion

$$\dot{p} = -\frac{\partial H(\vec{p}, \vec{q})}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$$

derived from a Hamiltonian. Then

$$\int_{\text{period}} \vec{p}(\vec{z}(t), \theta) \cdot \frac{\partial \vec{q}(\vec{z}(t), \theta)}{\partial \theta} d\theta$$

is independent of t. Such a theorem is not of muchuse unless one knows how to find functions $\vec{p}$, $\vec{q}$, $\vec{z}$, θ with the required properties. The theory discussed in the present section supplies the functions $\vec{p}(\vec{z}, \theta)$, $\vec{q}(\vec{z}, \theta)$ and $\vec{z}$ and θ as asymptotic series in ϵ.

V. APPLICATIONS OF ADIABATIC THEORY

A. The Current Density in a Plasma

In electricity and magnetism texts[(26)] it is usually proved that the equivalent current density due to magnetization of a material is given by $c\nabla \times \vec{\mathcal{M}}$, where $\vec{\mathcal{M}}$ is the magnetic moment per unit volume. One expects that under adiabatic conditions the extension to a collisionless plasma would be to add to $c\nabla \times \vec{\mathcal{M}}$ the current of the guiding center motion. This is indeed correct, but is not so easy to prove rigorously for a general $\vec{B}$ field geometry. The author is aware of no published proof and therefore will outline one here, with many of the details omitted. The complete proof is lengthy.

We wish to prove that the current density at $\vec{r}$ is

$$\vec{J}(\vec{r}, t) = Ne \langle \dot{\vec{R}}_0 \rangle + c\nabla \times \vec{\mathcal{M}} \tag{173}$$

l.c.

where

N = number of guiding centers per unit volume at $\vec{r}$ and t.

$\langle \dot{\vec{R}}_0 \rangle$ = average guidng center velocity of particles with guiding centers at $\vec{r}$.

$\vec{\mathcal{M}}$ = total magnetic moment per unit volume of particles with guiding centers at $\vec{r}$.

There will be an equation of this type for each charged component, ions and electrons for example. There exists a reasonably simple demonstration of the perpendicular component, $\vec{j}_\perp = Ne\langle \dot{\vec{R}}_{0\perp} \rangle + c(\nabla \times \vec{\mathcal{M}})_\perp$. The proof starts with the second moment[(27)] of the collisionless Boltzmann (Vlasov) equation. The second moment is

$$nm \frac{d\langle\vec{v}\rangle}{dt} = -\nabla \cdot \overleftrightarrow{P} + ne \frac{\langle\vec{v}\rangle}{c} \times \vec{B} + ne\vec{E}, \tag{174}$$

where $\langle \vec{v} \rangle$ is the average particle velocity, n is the particle density and $\overleftrightarrow{P} = nm \langle (\vec{v} - \langle\vec{v}\rangle)(\vec{v} - \langle\vec{v}\rangle) \rangle$. The current density perpendicular to $\vec{B}$ is $\vec{j}_\perp = ne\langle \vec{v} \rangle_\perp$. Consequently $\vec{j}_\perp$ can be obtained from (174) by crossing it with $\vec{B}$, but $\vec{j}_\parallel$ cannot. The analysis has been given in detail in reference (5).

In order to derive the parallel component of (173), it is necessay to work from the asymptotic series (61) and the collisionless Boltzmann equation itself rather than its second moment. Equation (173) complete is obtained this way and not merely its parallel component. Let $f(\vec{r}, \vec{v}, t)$ be the particle distribution function satisfying the collisionless Boltzmann equation: $\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla f + \frac{e}{m}(\vec{E} + \frac{\vec{v}}{c} \times \vec{B}) \cdot \nabla_{\vec{v}} f = 0$. The current density is

$$\vec{j}(\vec{r}, t) = ne\langle\vec{v}\rangle = e\int \vec{v} f(\vec{r}, \vec{v}, t)\, d^3v. \tag{175}$$

The procedure is first to differentiate the series (61), for $\vec{r}$ with respect to time and obtain $\vec{v}$ correct through terms of $\mathcal{O}(\epsilon)$. Next the "guiding center variables"(28) $\vec{R}_0$ and $\vec{V}$ are introduced, where the components of $\vec{V}$ along the directions of $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ are $V_1 = v_\parallel$, $V_2 = \rho\omega\cos\theta$, and $V_3 = -\rho\omega\sin\theta$. $\vec{V}_\perp$ is the gyration velocity in the frame of reference moving at the guiding center velocity $\dot{\vec{R}}_0$. The particle velocity $\vec{v}$ from the asymptotic series can then be expressed in terms of $\vec{R}_0$ and $\vec{V}$:

$$\vec{v} = (\dot{\vec{R}}_0)_\perp + \hat{e}_1 V_1 + V_2\hat{e}_2 + V_3\hat{e}_3 + \frac{V_3}{\omega}\left(\frac{\dot{B}}{2B}\hat{e}_2 - \dot{\hat{e}}_2\right) - \frac{V_2}{\omega}\left(\frac{\dot{B}}{2B}\hat{e}_3 - \dot{\hat{e}}_3\right)$$

$$+ \frac{i\epsilon^2}{\rho}(\vec{R}_{11} - \vec{R}^*_{11})V_2 + \frac{\epsilon^2}{\rho}(\vec{R}_{11} + \vec{R}^*_{11})V_3 + \frac{2i\epsilon^2}{\omega}(\vec{R}_{20} - \vec{R}^*_{20})(V_2^2 - V_3^2)$$

$$+ \frac{4\epsilon^2}{\omega}(\vec{R}_{20} + \vec{R}^*_{20})V_2 V_3 + \mathcal{O}(\epsilon^2), \tag{176}$$

where all vectors and fields are evaluated at $\vec{R}_0$. It will be noticed that the $\vec{R}_{20}$ term, which was of order ϵ^2 in $\vec{r}$, gives an order ϵ contribution to $\vec{v}$ due to differentiation of $e^{2i\int\omega dt}$ with respect to time. Terms of $\mathcal{O}(\epsilon)$ are needed in $\vec{v}$ if $\int \vec{v} f d^3v$ is to be correct through $\mathcal{O}(\epsilon)$, and this is necessary since $c\nabla \times \vec{\mathcal{M}}$ is of order ϵ.

The next step is to transform from the particle distribution function to the guiding center distribution function F defined by:

$$dn = f(\vec{r}, \vec{v}, t)\, d^3r\, d^3v = F(\vec{R}_0, \vec{V}, t)\, d^3R_0\, d^3V \quad . \tag{177}$$

Thus $F(\vec{R}_0, \vec{V}, t)\, d^3R_0 d^3V$ is the number of guiding centers in d^3R_0 of particles with parallel velocity in dV_1 and gyration velocity in $dV_2 dV_3$. Division by d^3r gives

$$f(\vec{r}, \vec{v}, t)\; d^3v = F(\vec{R}_0, \vec{V}, t)\; \mathcal{J}\left(\frac{\vec{R}_0}{\vec{r}}\right) d^3V, \tag{178}$$

where $\mathcal{J}$ is the Jacobian of $\vec{R}_0$ with respect to $\vec{r}$ at constant $\vec{V}$. Because $\vec{R}_0 = \vec{r} + \dfrac{\vec{V} \times \hat{e}_1}{\omega} + \mathcal{O}(\epsilon^2)$, the Jacobian is

$$\mathcal{J}\left(\frac{\vec{R}_0}{\vec{r}}\right) = 1 + \frac{V_2}{\omega}\;\hat{e}_3 \cdot \left[(\hat{e}_2 \cdot \nabla)\,\hat{e}_2 + (\hat{e}_1 \cdot \nabla)\hat{e}_1 + \frac{\nabla B}{B}\right]$$

$$- \frac{V_3}{\omega}\;\hat{e}_2 \cdot \left[(\hat{e}_3 \cdot \nabla)\hat{e}_3 + (\hat{e}_1 \cdot \nabla)\hat{e}_1 + \frac{\nabla B}{B}\right] + \mathcal{O}(\epsilon^2) \tag{179}$$

Zero order terms in the expression for $\vec{v}$, such as $\hat{e}_1(\vec{R}_0)\, V_2$, must be expanded about $\vec{r}$ because we wish $j(\vec{r})$ in terms of guiding center velocities at $\vec{r}$ and in terms of the magnetic moment of particles with guiding centers at $\vec{r}$, not $\vec{R}_0$. For the same reason $F(\vec{V}, \vec{R}_0)$ must be expanded about $F(\vec{V}, \vec{r})$. In addition $F(\vec{V}, \vec{r})$ can be expanded as $F_0(\vec{V}, \vec{r}) + \epsilon F_1(\vec{V}, \vec{r}) + \ldots$. After these expansions have been made, equations (176), (178), and (179) are substituted into $\int \vec{v}\, f d^3v$. Even after terms of $\mathcal{O}(\epsilon^2)$ are dropped, there still remain many terms in the integrand. A large fraction of these integrate to zero however by virtue of the following facts: (1) F_0 and F_1 are functions of V_2 and V_3 only through the combination $V_2^{\,2} + V_3^{\,2}$; this is not obvious and must be demonstrated by use of the "Boltzmann" equation which F satisfies. (2) R_{11} is a function of V_1, $V_2^{\,2} + V_3^{\,2}$, $\vec{R}_0$, t . (3) $\vec{R}_{20}$ is a function of $V_2^{\,2} + V_3^{\,2}$, $\vec{R}_0$, t. The last two statements again are not obvious but can be verified by examining the equations $\vec{R}_{11}$ and $\vec{R}_{20}$ satisfy.

The expression for $\int \vec{v} f d^3v$ now simplifies to

$$\int \vec{v}\, f(\vec{r}, \vec{v}, t)\, d^3v = \int d^3V\, \dot{\vec{R}}_0(\vec{r})\, F(\vec{r}, \vec{V}, t) + \int d^3V\, \hat{e}_1 \times \nabla \left[\frac{V_2^{\,2} + V_3^{\,2}}{2\omega} F(\vec{r}, \vec{V}, t)\right]$$

$$- \int d^3V\, \frac{V_2^{\,2} + V_3^{\,2}}{2\omega} \Big[(\hat{e}_3 \cdot \nabla)\hat{e}_2 - (\hat{e}_2 \cdot \nabla)\hat{e}_3 - \hat{e}_2\hat{e}_3 \cdot (\hat{e}_2 \cdot \nabla)\hat{e}_2 - \hat{e}_2\hat{e}_3 \cdot (\hat{e}_1 \cdot \nabla)\hat{e}_1$$

$$+ \hat{e}_3\hat{e}_2 \cdot (\hat{e}_3 \cdot \nabla)\hat{e}_3 + \hat{e}_3\hat{e}_2 \cdot (\hat{e}_1 \cdot \nabla)\hat{e}_1\Big]\, F(\vec{r}, \vec{V}, t), \qquad (180)$$

Where $\dot{\vec{R}}_0(\vec{r})$ is the velocity of guiding centers at $\vec{r}$, not $\vec{R}_0$. The unit vectors and ω are at $\vec{r}$ too. The long vector expression in the last integral of (180) turns out to be $\nabla \times \hat{e}_1$; this is by no means obvious. So,

$$\int \vec{v} f(\vec{r}, \vec{v}, t)\, d^3v = \int d^3V\, \dot{\vec{R}}_0\, F(\vec{r}, \vec{V}, t) - \nabla \times \int \hat{e}_1(\vec{r})\, \frac{V_2^{\,2} + V_3^{\,2}}{2\omega(\vec{r})}\, F(\vec{r}, \vec{V}, t)\, d^3V. \tag{181}$$

The magnetic moment per unit volume of particles with guiding centers at $\vec{r}$ is

$$\vec{\mathcal{M}} = -\frac{e}{c} \int \hat{e}_1(\vec{r})\, \frac{V_2^{\,2} + V_3^{\,2}}{2\omega(\vec{r})}\, F(\vec{r}, \vec{V}, t)\, d^3V, \tag{182}$$

and therefore

$$\int \vec{v} f\, d^3v = \int d^3V\, \dot{\vec{R}}_0\, F(\vec{r}, \vec{V}, t) + \frac{c}{e} \nabla \times \vec{\mathcal{M}} + \mathcal{O}(\epsilon^2) \tag{183}$$

$$\vec{J}(\vec{r}, t) = e \int \vec{v} f d^3v = Ne \langle \dot{\vec{R}}_0 \rangle + c\nabla \times \vec{\mathcal{M}}, \tag{184}$$

and the theorem is proved.

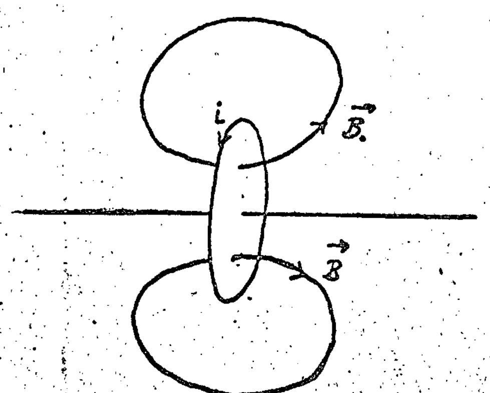
~~The preceding is quite sketchy; the detailed proof will be available in the future.~~

B. Loss-Free Geometries

The ordinary laboratory mirror machine with a magnetic field like that of Fig. (2) has a "loss cone": If at any point the velocity vector makes too small an angle with the field line, the particle will escape through the mirror (by Eq. 29). In the absence of diffusion in velocity space due to collisions with neutral or other charged particles, there would be no way adiabatically for a particle to get into this loss cone if it did not start out in it. But in practice such collisions do occur. The question therefore arises: Are there any static field configurations from which a particle cannot

(39)

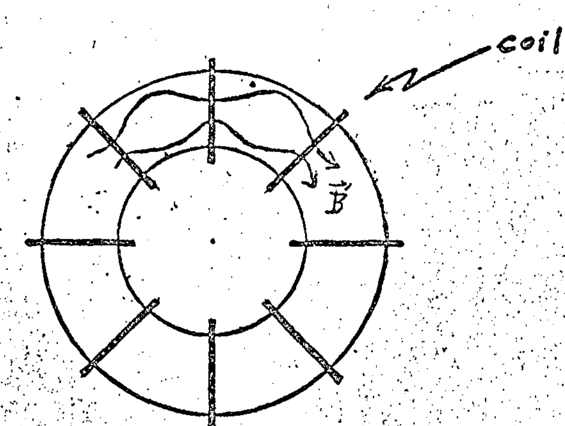
escape so long as it behaves adiabatically? (Non-adiabatic behavior is another possibility that must be considered). Although the adiabatic invariants do not seem to tell how to find such a configuration, they do provide a way to test a proposed one. The criterion is that all the longitudinal invariant surfaces passing through an arbitrary point in the containing volume must at no other place intercept a physical obstacle, such as a vacuum chamber wall or a current carrying wire. Since the surfaces are not simply nested and different velocity vector directions at the same point in space correspond to different surfaces, considerable effort may be needed to calculate where the surfaces lie. One loss-free configuration is that of a circular current loop (Fig. 17). For that matter the field of Fig. 2 is loss free if the external return flux is included in the vacuum system. The guiding center of a particle remains on a line of force and repeatedly transverses it. The problem is to supply the current to the coil and to support it mechanically without introducing wires and other obstacles into the path of the lines of force.



Field of a current loop is loss cone free

Fig. 17

Another geometry which has been extensively studied[30] for loss-cone free properties is the "bumpy torus". It consists of several discrete circular windings arranged around a toroidal vacuum chamber (Fig. 18) much as if several mirror machines were arranged end to end and then bent into a circle. It does indeed have regions where a particle will not be lost regardless of the direction of its velocity vector, and the coils can be supported in such a way that the supports do not interfere with the containment region. However, losses still occur due to a rather novel type diffusion, which is a consequence of there being many invariant surfaces through any one point. If the velocity vector of a particle is changed by scattering at a point, the particle changes invariant surfaces and may change to one that lies nearer an obstacle at some other point in the system. When the particle reaches this second point, it may again scatter onto a third invariant surface which lies still further out, etc. Thus because of the presence of many invariant surfaces at any one point, the particle may still work its way out of the system, but more slowly than by scattering in an ordinary mirror machine. Lauer, Gibson, and Jordan estimate that at



Bumpy torus

Fig. 18

least a factor of 9 increase in containment time over the ordinary mirror machine is possible with the bumpy torus. Of course a loss-free geometry does not eliminate losses by ordinary diffusion across the lines of force. If the loss due to the new ~~type~~ mechanism in the bumpy torus is less than that due to ordinary diffusion across the lines, ~~it~~ the torus is for practical purposes loss-free. The hydromagnetic stability of the bumpy torus is another question.

Loss-free geometries are of particular interest in connection with the change of e/m injection methods which have been proposed[31], [32] for building up a thermonuclear plasma. In these methods particles such as H_2^+ or neutral H are injected into a magnetic field at energies of many kilovolts or even an Mev. Some of these particles undergo a change in e/m by means of collisions with neutral background gas or with previously injected particles, and thereby become trapped if the geometry of the field is proper. The time scale for such systems to build up have been calculated to be as long as several minutes. The final steady state density is limited by the loss rate of particles already trapped. One loss mechanism is scattering into a loss region of velocity space, so that a loss-free geometry would increase the final density.

Another advantage that a loss-free geometry may possess is that of greater plasma stability. Sufficiently anisotropic velocity distributions are known theoretically to be unstable[33] and the instability may have been observed experimentally[34] in an ordinary mirror machine. If a field had no loss region in velocity space, the distribution function could be nearly isotropic.

C. The Geomagnetic Case

The behavior to be expected of adiabatic particles in the earth's field has been studied in some detail.[20] The geomagnetic field approximates that of a ~~loop of wire~~ current loop, but the system is not loss-cone free because of the presence of the earth. Neither is the field exactly azimuthally symmetric about any axis. Nevertheless, so long as the three adiabatic invariants are conserved, a trapped charged particle cannot escape. One interesting theorem in connection with the trapped (Van Allen) radiation surrounding the earth is a result of the canonical form of equations (136): Contours of constant B on an invariant surface are also contours of constant particle density provided there is a steady state and no electric field.[20]

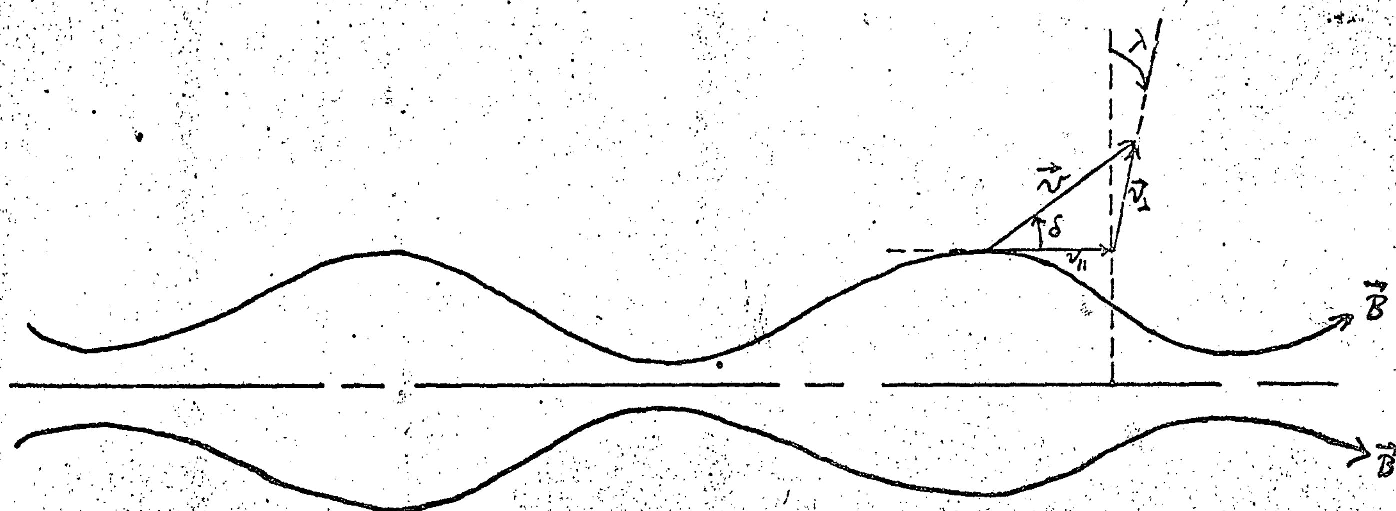
One as yet unsolved problem in connection with the Van Allen radiation is why the density of the inner or proton radiation belt falls off so rapidly with radius. It is observed[35] that the density of protons with energy greater than 75 Mev falls from a maximum at about 10,000 km. from the center of the earth to practically zero at 12,000 km in the equatorial plane. If the theory is to be tenable that the protons are the product of β-decay of cosmic ray neutrons, it is necessary to explain the large decrease in density beyond 10,000 km. The decrease in the source strength is nothing like this rapid. One possible explanation is in terms of a large loss rate due to non-adiabatic effects. If, for example, the magnetic moment decreased during many longitudinal oscillations, the particle would eventually be absorbed by the earth's atmosphere. Such a process might be described as scattering into the loss cone by the magnetic field, rather than by other particles. Possibly the gyration radius of a proton of say 100 Mev energy is so large at 12,000 km that non-adiabatic effects are important even in a static dipole field, as suggested by Singer.[36] It is also possible that rapid spatial variation of the field due to Alfvén waves is responsible, as suggested by Welch and Whitaker.[37]

Non-adiabatic effects in a static mirror geometry will be reviewed in more detail in the next section.

VI. UNRESOLVED PROBLEMS

A. Non-adiabatic Effects in Mirror Machines-Numerical

A principal unresolved question is the extent to which particles obey adiabatic predictions. The question is probably not capable of a general answer, and each situation must be examined independently. It is possible to follow particle trajectories by numerical methods and to compare the results with the adiabatic predictions. This has been done (38) for the laboratory type mirror machine, or more precisely for many mirror machines end to end in a straight line line (fields taken as periodic functions of z, Fig. 19). The somewhat surprising observation was made that as the particle oscillated between mirrors the magnetic moment could vary by large amounts but that the variations were highly self cancelling from one oscillation to the next. To be more specific, the particle was started off at the plane midway between mirrors as shown in Fig. 19. The

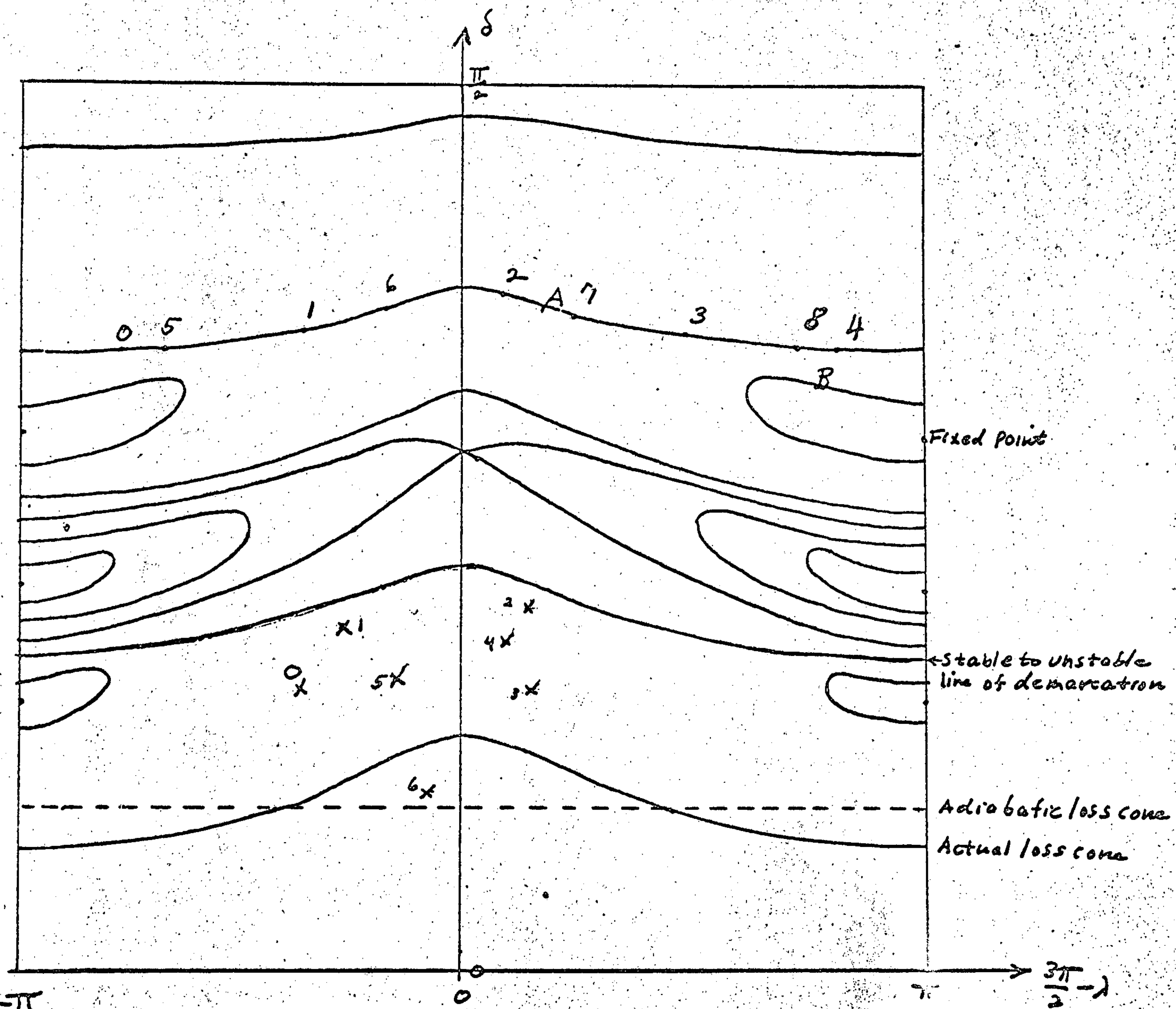


A periodic mirror geometry for which numerical calculations have been compared to the adiabatic predictions.

Fig. 19

angle between $\vec{v}$ and $\vec{B}(\vec{r})$ is δ, while λ is the angle between $\vec{v}_\perp$ and the plane of the page. Each time the particle returned to the midplane, new δ and λ were computed. A cumulative drift of δ would imply a gradual change in the magnetic moment; in an azimuthally symmetric field, conservation of the canonical angular momentum p_θ prohibits a radial drift, and therefore $B(\vec{R}_0)$ cannot change. Thus a change in $v_\perp{}^2/B(\vec{R}_0)$ must come from a change in $v_\perp = v \sin\delta$ and therefore from a change in δ since v is constant. Phase plots of the type shown qualitatively in Fig. (20) were obtained. The numbered points represent successive traversals of the median plane, zero being the initial point. There are two types of particle behavior evident. The first might be termed stable (curves A or B) and the second type is unstable as exemplified by points X. The δ of a stable particle may undergo severe variations, and the larger the gyration radius, the larger the variation. However, the variation is highly self cancelling; the phase points always appear to fall on a well-defined curve. On a large scale plot the points appear to fall on a smooth curve to within less than 10^{-3} degrees in δ, and that much scatter can be attributed to numerical errors. Thus it appears that the stable points are very stable indeed as a result of a "memory" from one median plane traversal to the next.

By contrast the unstable particles show no such memory and form no regular pattern. They usually escape within ten or so mirror reflections. The line of demarcation between the stable and unstable regions also is quite sharp, although just how sudden the transition is has not been carefully studied. It would be interesting to investigate the transition region more minutely. The unstable region, which begins at the loss cone and extends upward in δ to the line of demarcation does not appear to cover the entire range of λ from 0 to 2π. Near $\lambda = \pi/2$ there are stable curves of type B even below the line of demarcation. The fixed points at the center of type B patterns are rigorously fixed by virtue of the symmetry about the median plane. Because



Phase plots for particle in the mirror geometry of Fig. 19.

Fig. 20

the unstable particles escape in a few reflections, the demarcation between the stable and unstable regions is the loss cone for practical purposes. It has been found that the unstable region vanishes when the ratio of gyration radius to the distance between mirrors is less than something like 0.03.

The behavior of particles within the loss cone as they go from one mirror section to the next in this periodic machine has not been studied. It would be worthwhile studying, since these particles are analogous to those in the bumpy torus that are not trapped in one section but go clear around the torus.

A theory of the behavior observed in Fig. 20 has been given by Chirikov.[(39)] According to the theory particles are lost which exhibit a resonance between the fundamental longitudinal oscillation frequency (or one of its harmonics) and the gyration frequency. The theory has been given only in the approximation of straight lines of force. This is equivalent to assuming $\nabla \cdot \vec{B} \neq 0$. The inportance in this theory of line curvature would be worth investigating. It should be pointed out that the fixed points at the center of type B patterns in Fig. 20 possess this resonance and are precisely the ones which cannot be lost because of the symmetry. These are however very special cases and constitute a class of measure zero.

The type A stable curves are predicted qualitatively by the first two terms of the magnetic moment series (Eq. 153). From the fact that the canonical angular momentum p_θ is a constant of the motion, it is possible to write Eq. 153 in the form $f(\delta,\lambda) = \text{constant}$. The field used in the numerical calculations of Fig. 20 is derived (in cylindrical coordinates) from the vector potential:

$$A_\theta = \frac{LB_0}{2\pi}\left[\frac{\rho}{2} + \alpha \cos \zeta \, I_1(\rho)\right], \tag{185}$$

where L is the distance between mirrors, ρ is $\frac{2\pi r}{L}$, ζ is $\frac{2\pi z}{L}$, α is 0.2, and B_0 is the field halfway from the mirror to the median plane. With this field

$$f(\delta, \lambda) = \sin\delta\,(\sin\delta + a\cos^2\delta\,\sin\lambda) = \text{constant}, \tag{186}$$

where

$$a = \frac{4\pi\alpha\, I_1(\rho_0)}{\left[1 - \alpha\, I_0(\rho_0)\right]^2} \quad \frac{mcv}{eB_0L} \tag{187}$$

and p_0 is the solution of

$$\frac{4\pi^2 c p_\theta}{eB_0L^2} = \rho_0\left[\frac{\rho_0}{2} - \alpha I_1(\rho_0)\right]. \tag{188}$$

Higher terms in the magnetic moment series might give more details of the stable orbits. ~~the type B curves for instance~~. But the unstable behavior could not be predicted from the series.

An investigation of particle orbits in other geometries, such as the bumpy torus or a mirror machine without a median plane of symmetry, or without azimuthal symmetry would be of interest.

B. Non Adiabatic Effects-Experimental

The containment time of relativistic positrons in a laboratory mirror machine has been measured by Lauer, Gibson, and Jordan.(40) Exponential decay of the positron density is observed, with time constants as long as many seconds. The observed loss rate is quantitatively attributable to scattering from the background neutral gas, and therefore is not a non-adiabatic effect, even though a particle has made of the order of 10^{10} mirror reflections in the decay time. The accuracy of the self-cancelling effect in the stable orbits A and B of Fig. 20 therefore is quite phenonemal. It would be impossible to test for such long containment by numerical methods because of the computer time required and the numerical errors that would accumulate.

If in the experiment the gyration radius is increased so that the ratio of gyration radius to L is $\gtrsim 0.03$, evidence of the unstable region and enhanced loss cone appears.

Similar containment experiments with the bumpy torus or with a field resembling the earth's would be worthwhile.

C. Higher Order Terms

The first order terms $\epsilon J'$ and $\epsilon \Phi'$ in the adiabatic series for J and Φ have not been calculated. The first correction to the longitudinal invariant would be particularly valuable in studying the motion of the high energy protons in the earth's field.

VII. SUMMARY

The equations of motion of a charged particle in an electromagnetic field have been used to obtain the so-called "guiding center" motion. In a slowly varying field the particle motion is approximately a rapid rotation about a circle whose center (the guiding center) has velocity components along and at right angles to the magnetic field. The equations governing the guiding center motion have been obtained deductively for the general time-dependent field. Several unfamiliar drifts appeared because the time dependence and the perpendicular electric field were assumed, for the sake of generality, to be zero order in the radius of gyration rather than first order. These less familiar drifts have been illustrated via practical examples where they occur.

The modifications of the guiding center equations needed for particles of relativistic energy have been given.

The three adiabatic invariants of the particle motion have been defined and the invariance proven for the general case of time dependent fields. Several applications have been made of the guiding center equations of motion and the adiabatic invariants.

Finally, deviations from the adiabatic predictions have been discussed.

REFERENCES AND FOOTNOTES

1. The radiation reaction force will be omitted. In many practical cases its effect on the particle motion is negligible.

2. H. Alfvén, <u>Cosmical Electrodynamics</u>, (Oxford, Clarendon Press, 1950).

3. J. Linhart, <u>Plasma Physics</u>, (Interscience, New York, 1960).

4. G. Hellwig, Z. Naturforschung <u>10A</u>, 508 (1955).

5. T. Northrop, "The Guiding Center Approximation to Charged Particle Motion", Univ. of Calif., Lawrence Rad. Lab. Report, UCRL-5708-T, March 15, 1960, and Annals of Physics <u>15</u>, 79 (1961)

6. M. Kruskal, "The Gyration of a Charged Particle", Princeton University, Project Matterhorn Report PM-S-33 (NYO-7903), March 1958.

7. For a given piece of experimental equipment it is never in principle wrong to make the more general assumption that $\frac{\partial}{\partial t}$ and $\vec{u}_E$ are of zero order and to retain all the terms. Some of them may turn out to be small.

8. C. Longmire, D. Nagle, F. Ribe, Phys. Rev. <u>114</u>, 1187 (1959).

9. H. Lamb, <u>Hydrodynamics</u>, (Dover Publications, New York 1945) 6th Ed., p. 373.

10. It is simple to demonstrate that M_r is the adiabatic invariant for the special case of a uniform, azimuthally symmetric, time-dependent field with the particle gyrating about the axis of symmetry. It is also true (see Ref. 11) that $P_\perp^2/2m_0B$ is the adiabatic invariant in a general geometry so long as $\vec{E}_\perp$ is $\mathcal{O}(\epsilon)$, and is the flux of $\vec{B}$ enclosed by the circle of gyration.

11. P. Vandervoort, Annals of Physics <u>10</u>, 401 (1960).

12. P. G. Bergmann, <u>Introduction to the Theory of Relativity</u>, (Prentice-Hall, New York, 1942), equation 7.49. Bergmann uses a different metric tensor than in this paper, however.

13. J. Berkowitz and C. S. Gardner, "On the Asymptotic Series Expansion of the Motion of the Charged Particle in Slowly Varying Fields", New York University Report NYO-7975, December, 1957.

14. E. T. Whittaker and G. N. Watson, A Course of Modern Analysis, (Cambridge, 1945), 4th Ed., Ch. VIII.

15. The converse is also true. If there is an $A(\epsilon_1)$ such that $|f(\epsilon) - S_n(\epsilon)| \angle A(\epsilon_1)\, \epsilon^{n+1}$ for $0 < \epsilon < \epsilon_1$, then

$$\lim_{\epsilon \to 0} \frac{f(\epsilon) - S_n(\epsilon)}{\epsilon^n} = 0.$$

16. For example Σ_2 means $\vec{R}_0 + \epsilon(\vec{R}_1 e^{i\theta} + c.c.) + \epsilon^2 (\vec{R}_2 e^{2i\theta} + c.\ c.)$.

17. $\vec{R}$, as defined at the beginning of Section II, and $\vec{R}_0$ do not agree to all orders in ϵ. In fact $\vec{R} - \vec{R}_0 = \mathcal{O}(\epsilon^2)$. However, this is of no significance here, since we are considering only effects which are first order in gyration radius.

18. C. S. Gardner, Phys. Rev. 115, 791 (1959).

19. M. Kruskal, in The Theory of Neutral Ionized [illegible] (John Wiley, New York, 1960).

20. T. G. Northrop and E. Teller, Phys. Rev. 117, 215 (1960)

21. It is best to define J by equation (96), rather than by (95), and to define K as particle kinetic energy + e(ψ + φ).

22. B. B. Kadomtsev, Plasma Physics and the Problem of Controlled Thermonuclear Reactions, Vol. III (Akad. Nauk USSR, 1958), has given a similar proof for the case of static fields and $\vec{E}$ zero.

23. If the orders of ϵ are followed through, it is found that $\frac{1}{m}\left\langle\frac{dJ}{dt}\right\rangle = 0 + \mathcal{O}(\epsilon^2)$ and $\frac{1}{m}\frac{dJ}{dt} = 0 + \mathcal{O}(\epsilon)$.

24. W. Newcomb, "Motion of Magnetic Lines of Force", ~~Princeton University Observatory Technical Report No. 1, April 1955~~. Annals of Physics 3, 347 (1958)

25. Reference 12, equation 7.53

26. R. Becker, Theorie der Elektrizität, Vol II (Edwards Brothers, Ann Arbor), paragraph 21.

27. L. Spitzer, Physics of Fully Ionized Gases (Interscience, New York, 1956) P. 94.

28. These variables have been used by Watson (Phys. Rev. 102, 12 (1956)) to prove the perpendicular component of equation (173) for the special case where the magnetic field is static, torsionless, and curl free.

29. H. Margenau and M. Lewis (Rev. Mod. Phys. 31, 569 (1959)) point out that the opposite of "adiabatic" is "diabatic". However, "non-adiabatic" seems to be firmly entrenched in the literature and therefore will be used here.

30. G. Gibson, W. Jordan, E. J. Lauer, Phys. Rev. Letters 4, 217 (1960). See also A. I. Morozov and L. S. Solov'ev, Zh. Tekhn. Fiz. 30, 261 (1960).

31. G. Gibson, W. A. S. Lamb, E. J. Lauer, Phys. Rev. 114, 937 (1959).

32. A. Simon, Physics of Fluids 1, 495 (1958)

33. (a) W. Newcomb, "Stability of a Plasma in a Magnetic Field," University of California Lawrence Radiation Lab. Report UCRL 6197 (in preparation).

(b) M. N. Rosenbluth, Proceedings of the Second United Nations International Conference on the Peaceful Uses of Atomic Energy, V[illegible], p. 85.

(c) S. Chandrasekhar, A Kaufman, K. Watson, Proc. Roy. Soc. A24[illegible] 435, (1958)

34. W. A. Perkins and R. F. Post, Bull. Am. Phys. Soc. 5, 353 ([illegible])

35. C. Y. Fan, P. Meyer, J. A. Simpson; "Trapped and Cosmic Radiation Measurements from Explorer VI", paper given at the First International Space Science Symposium (Nice, France) Univ. of Chicago Enrico Fermi Institute for Nuclear Studies, Report EFINS - 59-73 (Dec. 1959)

36. S. F. Singer, Phys. Rev. Letters 3, 188 (1959)

37. J. A. Welch, Jr. and W. A. Whitaker, J. Geophys. Research 64, 909 (1959)

38. A. Garren, et al. "Individual Particle Motion and the Effect of Scattering in an Axially Symmetric Magnetic Field", Univ. of Calif. Lawrence Rad. Lab. Report UCRL 8076 (March 1958), or Proceedings of the Second United Nations International Conference on the Peaceful Uses of Atomic Energy, Vol. 31, p.65.

39. B. V. Chirikov, Atomnaya Energiya 6, 630 (1959)

40. G. Gibson, W. C. Jordan, E. J. Lauer, Phys. Rev. Letters 5, 141 (1960)