Covariant electromagnetic projection operators and a covariant description of charged particle guiding centre motion



1978 J. Phys. A: Math. Gen. 11 1069

(http://iopscience.iop.org/0305-4470/11/6/010)

J. Phys. A: Math. Gen., Vol. 11, No. 6, 1978. Printed in Great Britain

# Covariant electromagnetic projection operators and a covariant description of charged particle guiding centre motion

D M Fradkin†

Department of Physics, University of Edinburgh, Edinburgh EH9 3JZ, Scotland, UK

Received 31 October 1977, in final form 23 January 1978

**Abstract.** For non-null electromagnetic fields, a pair of covariant field projection operators is obtained which separate four-space into mutually orthogonal two-flats spanned by the field eigenvectors associated, respectively, with real or imaginary eigenvalues. In appropriate limits, these projections generalise the notion of components perpendicular to and parallel with a particular $\boldsymbol{E}$ or $\boldsymbol{B}$ direction. These operators are used in a relativistic covariant analysis of the motion of a charged particle subject to a constant electromagnetic field. The covariant projections of the motion are uncoupled and separate orbit equations for each projection are presented. This analysis is related to a generalisation of the guiding centre concept. A covariant guiding centre treatment is developed and applied to the situation where slowly changing field inhomogeneities exist.

## 1. Introduction

When all its eigenvalues are distinct, the electromagnetic field tensor provides, via its eigenvectors, a field of basis vectors in terms of which an arbitrary four-vector may be decomposed. Furthermore, the subspaces spanned by those eigenvectors which are associated with either real or imaginary eigenvalues provide a unique decomposition of the four-dimensional continuum into orthogonal two-flats, a decomposition which is meaningful even if the eigenvalues are doubly degenerate. In this paper such a decomposition is investigated in terms of the relativistically covariant projection operators that, among other things, generate quantities that generalise the notion of projections perpendicular to and parallel with a particular electromagnetic three-vector, be it $\boldsymbol{E}$ or $\boldsymbol{B}$ (in the limit that one of these is small). The utility of these covariant operators is then indicated with reference to the analyses of the motion of a charged particle in an electromagnetic field, and the use of these operators is related to a relativistic guiding centre decomposition of such motion.

The organisation of this paper is as follows. First, in § 2.1, we review some of the basic algebraic properties of $F$ such as its eigenvalues and various product relations involving powers of $F$ and its dual. In § 2.2, we construct by means of projection techniques the eigenvectors of $F$ and display equivalent eigenvector sets. The same type of technique is employed in § 2.3 to construct covariant electromagnetic field projection operators. The nature of these operators is discussed, and various limiting

† On sabbatical leave from the Physics Department, Wayne State University, Detroit, Michigan 48202, USA.

0305-4770/78/0006-1069$03.00  

cases of their projection are given. Our covariant projection operators are shown to be a generalisation of the covariant operators devised for the case $\boldsymbol{E}\,.\,\boldsymbol{B}=0$ by Derfler (1976). (Similar operators for this special case had earlier been referred to by Mangeney and Signore (1974).)

The covariant projections are utilised in § 3.1 to determine covariant expressions for the motion of a charged particle under the influence of constant electromagnetic fields. It turns out that the covariant projections decouple the equations of motion into two parts. One part describes oscillatory proper time behaviour and the other describes exponential proper time behaviour. For each projection of the motion there is an associated orbit equation, as well as an orbit equation for the motion as a whole. Section 3.2 is devoted to relating the two preceding projections of the motion to a covariant generalisation of the guiding centre of motion.

In § 4.1, the covariant approximate guiding centre motion of a charged particle under the influence of inhomogeneous fields is discussed. This approximation involves a time average for slowly changing quantities over a period of the presumed more rapid oscillatory motion. Covariant expressions for such a time-averaged guiding centre (which in a non-relativistic context has been termed by Jancel and Kahan (1966) as a 'glide centre' or 'drift centre') are obtained, and a covariant derivation of the various drift velocities is given in § 4.2. The paper concludes with a simple derivation of the invariant guiding centre momentum–energy as well as an expression which relates to a covariant generalisation of the usual non-relativistic adiabatic electromagnetic moment.

The concept of the guiding centre has been used extensively in plasma physics and astrophysics. Although perhaps the genesis of this concept lies with Hipparchus and Ptolemy, the idea that the instantaneous motion of a charged particle could be separated into an oscillatory motion superimposed upon the motion of the centre of the oscillation (called the guiding centre) is generally attributed to Alfvén (1950) who made such a decomposition in discussing the non-relativistic motion of a charged particle generally spiralling around and accelerating along a $\boldsymbol{B}$ field direction and drifting normally to that direction in the presence of small field inhomogeneities. Later, others (for example, Hellwig 1955, Spitzer 1956, Northrop and Teller 1960, Northrop 1961, Kruskal 1965, pp 67–90, and Schmidt 1966) have discussed and amplified to some degree these non-relativistic arguments, and have examined in some detail the effects of specific inhomogeneities. The derivations yielding the various terms in the guiding centre drift and equation of motion tended to be rather involved, partially due to the non-covariant nature of the treatment. Some define the guiding centre position $\boldsymbol{r}_{\rm c}$ by following Alfvén's definition $\boldsymbol{r}_{\rm c}=\boldsymbol{r}+c(eB^2)^{-1}\boldsymbol{p}\times\boldsymbol{B}$, while others directly sought expressions for the guiding centre by a suitable time average of the basic dynamical relations over a single period associated with the cyclotron frequency.

In addition to a guiding centre description, Alfvén and Fälthammar (1963) introduced the idea of centre of gyration movement which they obtained by considering a (non-relativistic) transformation to a coordinate system in which no force exists on the particle except that of a magnetic field $\boldsymbol{B}$. Of course, this type of description cannot apply to an arbitrary electromagnetic field since $\boldsymbol{E}\,.\,\boldsymbol{B}$ is a Lorentz invariant and hence one cannot find a reference frame (speaking from the standpoint of a relativistic generalisation) in which only a magnetic field exists unless it should happen that $\boldsymbol{E}$ and $\boldsymbol{B}$ are crossed fields or that to some desired approximation, $E/B \ll 1$. For the non-relativistic treatment, Kruskal (1959, 1965, pp 91–102) has analysed the first-order inhomogeneity corrections in terms of an asymptotic approximation to the exact motion (controlled by the parameter $m/e$), and Bernstein (1971) has devised a two-time-scales scheme to develop an iteration procedure to obtain higher-order terms.

The literature referring to a relativistic treatment of guiding centre motion is much more sparse. A number of years ago, Hellwig (1955) did sketch a covariant relativistic approach based on a perturbative variation of constants technique applied to the oscillatory motion and he identified the modulus of the imaginary eigenvalue of $(e/mc)F$ as the relevant frequency. His prescription for the motion of the guiding centre (which is called the '*ersatzteilchen*' by him), based on a coupled set of equations involving a number of parameters, does not appear to lend itself reaily to an explicit description of the motion in terms of the initial conditions which must be applied to the dynamical equations. Later, in their discussion of charged particle motion in the earth's field, Northrop and Teller (1960) quote expressions (without derivation), valid in the small-$\boldsymbol{E}$ regime, for the guiding centre motion of relativistic particles, accelerating along $\boldsymbol{B}$ and drifting perpendicular to $\boldsymbol{B}$. Then, a lengthy treatment of the relativistic situation involving inhomogeneous fields was given by Vandervoort (1960). He expanded the equations of motion via a Taylor's expansion about guiding centre variables in order to obtain differential equations for the gyration variables, and techniques of solution to successive orders of approximation for a number of relevant variables were indicated. The guiding centre motion and the question of the non-constancy of the adiabatic invariant was also considered by him. Again, it is somewhat awkward to extract results in terms of field quantities and initial conditions. More recently, Mangeney and Signore (1974) adopt an averaged Lagrangian approach to obtain drift expressions for a relativistic particle in a strong magnetic field for the special case $\boldsymbol{E}\,.\,\boldsymbol{B}=0$.

## 2. Algebraic properties of the electromagnetic field tensor

### *2.1. Eigenvalues and product relations*

We shall adopt Minkowski coordinates (Latin indices ranging from 1 to 3, Greek indices from 1 to 4) so that the electromagnetic field tensor $F$ has elements

$$F_{\mu\nu}=-F_{\nu\mu}, \qquad \text{with } F_{ij}=\epsilon_{ijk}B_k, F_{j4}=-\mathrm{i}E_j, \tag{2.1}$$

and its dual $F^{\mathrm{D}}$ is defined by

$$F^{\mathrm{D}}_{\mu\nu}=\tfrac{1}{2}\mathrm{i}\epsilon_{\mu\nu\rho\pi}F_{\rho\pi}. \tag{2.2}$$

As is well known, the eigenvalues $\eta$ of the field tensor consist of a set of two pairs, one real and the other pure imaginary, $\{\eta\}=\{+b, -b, +\mathrm{i}a, -\mathrm{i}a\}$, where

$$\begin{aligned} a &= (2)^{-1/2}\{[(B^2-E^2)^2+4(\boldsymbol{E}\,.\,\boldsymbol{B})^2]^{1/2}+(B^2-E^2)\}^{1/2} \\ &= \tfrac{1}{2}\{[(F_{\mu\pi}F_{\pi\mu})^2+(F_{\mu\pi}F^{\mathrm{D}}_{\pi\mu})^2]^{1/2}-(F_{\mu\pi}F_{\pi\mu})\}^{1/2}, \end{aligned} \tag{2.3}$$

and

$$\begin{aligned} b &= (2)^{-1/2}\{[(B^2-E^2)+4(\boldsymbol{E}\,.\,\boldsymbol{B})^2]^{1/2}-(B^2-E^2)\}^{1/2} \\ &= \tfrac{1}{2}\{[F_{\mu\pi}F_{\pi\mu})^2+(F_{\mu\pi}F^{\mathrm{D}}_{\pi\mu})^2]^{1/2}+(F_{\mu\pi}F_{\pi\mu})\}^{1/2}. \end{aligned} \tag{2.4}$$

The parameter $a$ is zero only when $\boldsymbol{E}.\boldsymbol{B}=0$ and $B^2-E^2\leqslant 0$, whereas $b$ is zero only when $\boldsymbol{E}.\boldsymbol{B}=0$ and $B^2-E^2\geqslant 0$. A null electromagnetic field is characterised by $a=b=0$. The ratio $\epsilon=b/a$ indicates the relative sizes of the magnetic and electric fields. If $\epsilon=1$, then $B^2=E^2$ but specification of this ratio does not in itself determine the relative size of $|\boldsymbol{E}.\boldsymbol{B}|$. If $\epsilon\ll 1$, then $B^2$ is larger than $E^2$ and $|\boldsymbol{E}.\boldsymbol{B}|\ll|B^2-E^2|$. On the other hand, if $\epsilon\gg 1$, then $B^2$ is less than $E^2$ and again $|\boldsymbol{E}.\boldsymbol{B}|\ll|B^2-E^2|$. The eigenvalue $a$ (cf § 3.1) is directly proportional to the generalised cyclotron frequency appropriate to a charged particle oscillating in a constant, but otherwise arbitrary, electromagnetic field.

In terms of the parameters $a$ and $b$, the product relations (cf Synge 1956) satisfied by the skew-symmetric field tensors are

$$FF=(b^2-a^2)I+F^{\mathrm{D}}F^{\mathrm{D}}, \tag{2.5}$$

$$FF^{\mathrm{D}}=F^{\mathrm{D}}F=-sabI, \tag{2.6}$$

so that

$$FFF=F^3=(b^2-a^2)F-sabF^{\mathrm{D}}. \tag{2.7}$$

Here $s$ is a sign parameter defined by

$$s\equiv(\boldsymbol{E}.\boldsymbol{B})/|(\boldsymbol{E}.\boldsymbol{B})| \tag{2.8}$$

and $I$ is the unit $4\times 4$ matrix with elements $\delta_{\mu\nu}$.

Power of $F$ higher than the cubic may be reduced by invoking the Hamilton–Cayley theorem which states that each square matrix satisfies its own characteristic equation. Consequently,

$$F^4+(a^2-b^2)F^2-(ab)^2I=0. \tag{2.9}$$

Hence, any power of $F$ may be reduced to a linear combination of $I$, $F$, $F^2$, and $F^3$, or to a linear combination of $I$, $F$, $F^{\mathrm{D}}$, and $F^2$. In particular, it follows that

$$F^{2n}=(a^2+b^2)^{-1}\{[b^{2n}+(-1)^{n+1}a^{2n}]F^2+(ab)^2[b^{2(n-1)}+(-1)^n a^{2(n-1)}]I\}, \tag{2.10}$$

from which one may establish that

$$\begin{aligned}\mathrm{e}^{tF}=(a^2+b^2)^{-1}\{&(a^2I+F^2)[(\cosh bt)I+b^{-1}(\sinh bt)F]\\&+(b^2I-F^2)[(\cos at)I+a^{-1}(\sin at)F]\}.\end{aligned} \tag{2.11}$$

The special cases for $a$ and/or $b$ equal to zero may be recovered by taking the appropriate limits of this relation.

For future convenience, we will display the matrix elements of $F^2$ in terms of $\boldsymbol{E}$ and $\boldsymbol{B}$:

$$\begin{aligned}&(F^2)_{ij}=B_iB_j+E_iE_j-\delta_{ij}B^2\\&(F^2)_{j4}=(F^2)_{4j}=-\mathrm{i}(\boldsymbol{E}\times\boldsymbol{B})_j\\&(F^2)_{44}=E^2.\end{aligned} \tag{2.12}$$

For the special case $a=b=0$, the vectors $\boldsymbol{E}$, $\boldsymbol{B}$ and the unit vector $\hat{\boldsymbol{n}}=\boldsymbol{E}\times\boldsymbol{B}/E^2$ represent an orthogonal triad, so then the matrix elements of $F^2$ simply reduce to

$$\begin{aligned}&(F^2)_{ij}=-E^2\hat{n}_i\hat{n}_j\\&(F^2)_{j4}=(F^2)_{4j}=-\mathrm{i}E^2\hat{n}_j\\&(F^2)_{44}=E^2.\end{aligned} \tag{2.13}$$

### 2.2. *Field eigenvectors*

We discuss first the situation in which all the eigenvalues of $F$ are distinct. (The case of degeneracy is treated in the concluding paragraph of this section.) For a particular eigenvalue $\eta_\beta$, where $\beta$ is a label indicating a specific member of the eigenvalue set { }, the remaining numbers of the set are $-\eta_\beta$, $-\mathrm{i}ab\eta_\beta^{-1}$, and $\mathrm{i}ab\eta_\beta^{-1}$. Thus, the associated eigenvector $\psi^\beta$ satsifying

$$(F_{\mu\nu}-\eta_\beta\delta_{\mu\nu})\psi^\beta_\nu=0 \qquad (\textit{no}\text{ sum on }\beta) \tag{2.14}$$

may be constructed by means of a projective technique using the operator

$$(F+\eta_\beta I)(F-\mathrm{i}ab\eta_\beta^{-1}I)(F+\mathrm{i}ab\eta_\beta^{-1}I)$$

acting on an arbitrary four-vector $\Phi$. Indeed, invoking the Hamilton–Cayley theorem,

$$(F_{\mu\nu}-\eta_\beta\delta_{\mu\nu})[(F+\eta_\beta I)(F-\mathrm{i}ab\eta_\beta^{-1})(F+\mathrm{i}ab\eta_\beta^{-1}I)]_{\nu\pi}=0, \tag{2.15}$$

we note that the expression in square brackets itself may be taken to be proportional to $\psi^\beta_\nu$.

One may define the eigenvector (for fixed $\pi$) as

$$\psi^\beta_{\nu\pi}=\mathrm{i}\eta_\beta^{-1}\{(F+\eta_\beta I)[F^2+(ab/\eta_\beta)^2I]\}_{\nu\pi}. \tag{2.16}$$

The assignment of different variables to the second index ($\pi$) produces four-vectors that are proportional since eigenvectors associated with the same eigenvalue are generated. This fact is reflected in the quadratic relation

$$\psi^\beta_{\nu\pi}\psi^\beta_{\mu\alpha}-\psi^\beta_{\nu\alpha}\psi^\beta_{\mu\pi}=0 \qquad (\textit{no}\text{ sum on }\beta). \tag{2.17}$$

Thus, using the result

$$\psi^\beta_{\nu\nu}=2\mathrm{i}(2\eta_\beta^2+a^2-b^2), \tag{2.18}$$

one obtains

$$\psi^\beta_{\nu\pi}\psi^\beta_{\pi\alpha}=2\mathrm{i}(2\eta_\beta^2+a^2-b^2)\psi^\beta_{\nu\alpha} \qquad (\textit{no}\text{ sum on }\beta). \tag{2.19}$$

The quadratic relation involving summation of indices in the same position is quite different, yielding the result

$$\psi^\beta_{\pi\nu}\psi^\beta_{\pi\alpha}=0=\psi^\beta_{\nu\pi}\psi^\beta_{\alpha\pi} \qquad (\textit{no}\text{ sum on }\beta). \tag{2.20}$$

Hence, $\psi^\beta_{\nu\alpha}$ for fixed $\alpha$ is a null four-vector.

Eigenvalue degeneracy can occur only if a particular $\eta_\beta$ is zero, which in turn occurs only when $|\boldsymbol{E}\,.\,\boldsymbol{B}|=ab=0$, i.e., whenever $a$ and/or $b$ equals zero. Equation (2.6) shows that $F$ has no inverse then. In any of these singular $F$ situations, the tensor $F$ itself does not have a complete set of eigenvectors in terms of which an arbitrary four-vector can be expanded. Nevertheless, the eigenvectors it does have may be constructed by a procedure similar to that previously used. For example, for the doubly degenerate case $b=0$, the operators $(F^2+a^2I)$ and $F(F\pm\mathrm{i}aI)$ operating on an arbitrary four-vector $\Phi$ produce eigenvectors associated, respectively, with the $b=0$ and $\pm\mathrm{i}a$ eigenvalues. For the quadruply degenerate case ($a=b=0$), $F^2\Phi$ is the eigenvector associated with zero eigenvalue.

### 2.3. *Electromagnetic field projection operators*

By the reasoning of the last section, the operator $(F-bI)(F+bI)$ when acting on an arbitrary four-vector produces a four-vector which is composed of eigenvectors of $F$ with associated eigenvalues of $\pm ia$, in other words, an eigenvector of $F^2$ with eigenvalue $-a^2$. Similarly, $(F-iaI)(F+iaI)$ produces an eigenvector of $F^2$ with eigenvalue $b^2$. Multiplying these operators by appropriate scalar factors so that idempotency results (cf equations (2.23) and (2.24)), one may thus define the following electromagnetic field projection operators:

$$\mathcal{O}^{(a)} = -(a^2+b^2)^{-1}(F^2-b^2I), \tag{2.21}$$

and

$$\mathcal{O}^{(b)} = (a^2+b^2)^{-1}(F^2+a^2I). \tag{2.22}$$

These operators are each symmetric. Although they remain well defined if either $a$ or $b$ equals zero (the doubly degenerate case in which $\boldsymbol{E}\cdot\boldsymbol{B}=0$), they are not defined for the case of a null electromagnetic field (both $a$ and $b$ equal to zero). In the rest of this paper we shall exclude the null field from consideration and only deal with situations for which these operators are well defined.

By its construction, it is apparent that $\mathcal{O}^{(a)}$ is an operator which projects into the two-flat subspace spanned by the eigenvectors of $F$ having eigenvalues $\pm ia$. An equivalent description of this two-flat is that it is the subspace which is orthogonal to the two-flat spanned by the eigenvectors of $F$ having eigenvalues $\pm b$. (As has been emphasised by Synge (1956, p 61), four-dimensional geometry permits the existence of mutually orthogonal two-flats.) Similarly $\mathcal{O}^{(b)}$ projects into the two-flat spanned by eigenvectors of $F$ having eigenvalues $\pm b$ (which in turn is orthogonal to the two-flat characterised by the $\pm ia$ eigenvalues).

Using equation (2.9), it is readily shown that these operators satisfy the projection algebra

$$\mathcal{O}^{(a)}\mathcal{O}^{(a)} = \mathcal{O}^{(a)}, \tag{2.23}$$

$$\mathcal{O}^{(b)}\mathcal{O}^{(b)} = \mathcal{O}^{(b)}, \tag{2.24}$$

$$\mathcal{O}^{(a)}+\mathcal{O}^{(b)} = I, \tag{2.25}$$

and

$$\mathcal{O}^{(a)}\mathcal{O}^{(b)} = \mathcal{O}^{(b)}\mathcal{O}^{(a)} = 0. \tag{2.26}$$

Furthermore, the relationship of these operators to the eigenvectors of $F^2$ may be expressed by the equations

$$F^2\mathcal{O}^{(a)} = \mathcal{O}^{(a)}F^2 = -a^2\mathcal{O}^{(a)} \tag{2.27}$$

$$F^2\mathcal{O}^{(b)} = \mathcal{O}^{(b)}F^2 = b^2\mathcal{O}^{(b)}. \tag{2.28}$$

Analogously, from equation (2.5) it follows that

$$(F^{\mathrm{D}})^2\mathcal{O}^{(a)} = \mathcal{O}^{(a)}(F^{\mathrm{D}})^2 = -b^2\mathcal{O}^{(a)} \tag{2.29}$$

$$(F^{\mathrm{D}})^2\mathcal{O}^{(b)} = \mathcal{O}^{(b)}(F^{\mathrm{D}})^2 = a^2\mathcal{O}^{(b)}. \tag{2.30}$$

A further relation satisfied by these projection operators is

$$a^{-2}F\mathcal{O}^{(a)} - b^{-2}F\mathcal{O}^{(b)} = (sab)^{-1}F^{\mathrm{D}} \tag{2.31}$$

from which it follows that

$$sbF\mathcal{O}^{(a)} = aF^{\mathrm{D}}\mathcal{O}^{(a)}, \qquad saF\mathcal{O}^{(b)} = -bF^{\mathrm{D}}\mathcal{O}^{(b)}. \tag{2.32}$$

For an arbitrary four-vector $g_\mu = (\boldsymbol{g}, g_4 = \mathrm{i}g_0)$, the two-flat components produced by these projection operators may be written out in terms of the electromagnetic field three-vectors $\boldsymbol{E}$ and $\boldsymbol{B}$. The results are

$$(\mathcal{O}^{(a)}g)_i = -(a^2+b^2)^{-1}[(\boldsymbol{B}\cdot\boldsymbol{g})B_i + (\boldsymbol{E}\cdot\boldsymbol{g})E_i - (b^2+B^2)g_i + g_0(\boldsymbol{E}\times\boldsymbol{B})_i], \tag{2.34}$$

$$(\mathcal{O}^{(a)}g)_4 = \mathrm{i}(a^2+b^2)^{-1}[(\boldsymbol{E}\times\boldsymbol{B}\cdot\boldsymbol{g}) + (b^2-E^2)g_0], \tag{2.35}$$

and

$$(\mathcal{O}^{(b)}g)_i = (a^2+b^2)^{-1}[(\boldsymbol{B}\cdot\boldsymbol{g})B_i + (\boldsymbol{E}\cdot\boldsymbol{g})E_i + (a^2-B^2)g_i + g_0(\boldsymbol{E}\times\boldsymbol{B})_i], \tag{2.36}$$

$$(\mathcal{O}^{(b)}g)_4 = \mathrm{i}(a^2+b^2)^{-1}[-(\boldsymbol{E}\times\boldsymbol{B}\cdot\boldsymbol{g}) + (a^2+E^2)g_0]. \tag{2.37}$$

The effect of the projection operators becomes especially simple in two limiting cases to be described below.

In the limit in which $\boldsymbol{E}\to 0$, then $b^2\to 0$, and $a^2\to B^2$. One then obtains

$$(\mathcal{O}^{(a)}g)_i \to B^{-2}[\boldsymbol{B}\times(\boldsymbol{g}\times\boldsymbol{B})]_i, \qquad (\mathcal{O}^{(a)}g)_4\to 0, \tag{2.38}$$

and

$$(\mathcal{O}^{(b)}g)_i \to B^{-2}(\boldsymbol{B}\cdot\boldsymbol{g})B_i, \qquad (\mathcal{O}^{(b)}g)_4\to g_4. \tag{2.39}$$

Thus, in this limit, $\mathcal{O}^{(a)}$ projects out the component of $\boldsymbol{g}$ perpendicular to $\boldsymbol{B}$, and $\mathcal{O}^{(b)}$ projects out the component of $\boldsymbol{g}$ parallel to $\boldsymbol{B}$.

In the other limit in which $\boldsymbol{B}\to 0$, then $a^2\to 0$, and $b^2\to E^2$. One then obtains

$$(\mathcal{O}^{(a)}g)_i \to E^{-2}[\boldsymbol{E}\times(\boldsymbol{g}\times\boldsymbol{E})], \qquad (\mathcal{O}^{(a)}g)_4\to 0, \tag{2.40}$$

and

$$(\mathcal{O}^{(b)}g)_i \to E^{-2}(\boldsymbol{E}\cdot\boldsymbol{g})E_i, \qquad (\mathcal{O}^{(b)}g)_4\to g_4. \tag{2.41}$$

So, in this limit, $\mathcal{O}^{(a)}$ projects out the component of $\boldsymbol{g}$ perpendicular to $\boldsymbol{E}$, and $\mathcal{O}^{(b)}$ projects out the component of $\boldsymbol{g}$ parallel to $\boldsymbol{E}$.

Thus it is a remarkable fact that if only one of the three-vectors $\boldsymbol{E}$, $\boldsymbol{B}$ is non-zero, the projection operators $\mathcal{O}^{(a)}$ and $\mathcal{O}^{(b)}$ respectively project out the three-vectors perpendicular to and parallel to the non-zero field. Hence, with respect to these limiting situations, one may characterise $\mathcal{O}^{(a)}$ as the perpendicular projection operator, and $\mathcal{O}^{(b)}$ as the parallel projection operator.

In general, as a consequence of equation (2.26), it follows that four-vectors projected out by the two different projection operators are orthogonal, i.e. that

$$g_\mu^{(a)}h_\mu^{(b)} = 0, \tag{2.42}$$

where we have adapted the notation

$$g^{(a)} \equiv \mathcal{O}^{(a)}g, \qquad h^{(b)} \equiv \mathcal{O}^{(b)}h, \tag{2.43}$$

in which $g$ and $h$ represent four-vectors.

For the special case $a = 0$, $b \neq 0$, the covariant projection operators become

$$\mathcal{O}^{(a)} \xrightarrow[a\to 0]{} -(F^{\mathrm{D}})^2/b^2, \qquad \mathcal{O}^{(b)} \xrightarrow[a\to 0]{} F^2/b^2,$$

while for the special case $b = 0$, $a \neq 0$, these operators become

$$\mathcal{O}^{(a)} \xrightarrow[b\to 0]{} -F^2/a^2, \qquad \mathcal{O}^{(b)} \xrightarrow[b\to 0]{} (F^{\mathrm{D}})^2/a^2.$$

Recently, for applications in which $\boldsymbol{E}.\boldsymbol{B} = ab = 0$, Derfler (1976) has introduced operators, which in our notation may be written as $(b^2 - a^2)^{-1}F^2$ and $-(b^2 - a^2)(F^{\mathrm{D}})^2$, in order to generalise the concept of perpendicular and parallel projections. It is readily seen that our set of covariant projection operators reduces to his set in the appropriate limit.

## 3. Dynamics of a charged particle in constant fields

### 3.1. Dynamics of the projected components

For a particle of charge $e$, rest mass $m$, moving in an electromagnetic field described by the tensor $F$ and acted upon by an additional (non-electromagnetic) four-vector force field $g$, the covariant equation of motion (neglecting radiation reaction) for the four-vector momentum $p$ is

$$\frac{\mathrm{d}}{\mathrm{d}\tau} p = \frac{e}{mc} Fp + g. \tag{3.1}$$

Here $p = m\ \mathrm{d}x/\mathrm{d}\tau$ (in component form $p_\mu = m\ \mathrm{d}x_\mu/\mathrm{d}\tau$), where $\tau$ is the proper time.

In this section it shall be assumed that all elements of $F$ and $g$ have no spatial or temporal dependence. This restriction on $g$ requires that it be zero since the relativistic condition $p_\mu g_\mu = 0$ forces $g$ to have a non-constant behaviour for $p$ being non-constant. Thus, the constant field equation that we are dealing with is,

$$\frac{\mathrm{d}}{\mathrm{d}\tau} p = \frac{e}{mc} Fp. \tag{3.2}$$

Now, this equation can be directly integrated to give $p = \{\exp[(e\tau/mc)F]\}p(0)$, where the exponential operator may be simplified using equation (2.11). Then, the resulting equation can be integrated again to give the particle's four-vector position as a function of proper time.

Alternatively, equation (3.2) may be solved by employing the projection operators $\mathcal{O}^{(a)}$ and $\mathcal{O}^{(b)}$ of the previous section. Decomposing $p = p^{(a)} + p^{(b)}$, and recognising that the projection operators commute with $\mathrm{d}/\mathrm{d}\tau$ and $F$, one sees that

$$\frac{\mathrm{d}}{\mathrm{d}\tau} p^{(a)} = \frac{e}{mc} Fp^{(a)}, \qquad \frac{\mathrm{d}}{\mathrm{d}\tau} p^{(b)} = \frac{e}{mc} Fp^{(b)}. \tag{3.3}$$

Thus the two projections of the motion are completely uncoupled from each other. Furthermore, from equations (2.27) and (2.28), it follows that

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} p^{(a)} = \left(\frac{e}{mc}\right)^2 F^2 p^{(a)} = -\left(\frac{ea}{mc}\right)^2 p^{(a)} \tag{3.4}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} p^{(b)} = \left(\frac{e}{mc}\right)^2 F^2 p^{(b)} = \left(\frac{eb}{mc}\right)^2 p^{(b)}. \tag{3.5}$$

Thus the '$a$ projection' behaves like a harmonic oscillator with frequency

$$\omega_a = ea/mc, \tag{3.6}$$

while the '$b$ projection' exhibits an exponential proper time behaviour characterised by an inverse time constant

$$\lambda_b = eb/mc. \tag{3.7}$$

Thus from equations (3.4) through (3.7) it follows that

$$p^{(a)}(\tau) = (\cos\omega_a\tau)p^{(a)}(0) + a^{-1}(\sin\omega_a\tau)Fp^{(a)}(0), \tag{3.8}$$

$$p^{(b)}(\tau) = (\cosh\lambda_b\tau)p^{(b)}(0) + b^{-1}(\sinh\lambda_b\tau)Fp^{(b)}(0). \tag{3.9}$$

Using the relations

$$\begin{aligned}(a^2+b^2)F\mathcal{O}^{(a)} &= a(aF+sbF^{\mathrm{D}}),\\ (a^2+b^2)F\mathcal{O}^{(b)} &= b(bF-saF^{\mathrm{D}}),\end{aligned} \tag{3.10}$$

the preceding relations become

$$(a^2+b^2)p^{(a)}(\tau) = \{(\sin\omega_a\tau)[(aF+sbF^{\mathrm{D}})p(0)] - (\cos\omega_a\tau)[(F^2-b^2I)p(0)]\} \tag{3.11}$$

and

$$(a^2+b^2)p^{(b)}(\tau) = \{(\sinh\lambda_b\tau)[(bF-saF^{\mathrm{D}})p(0)] + (\cosh\lambda_b\tau)[(F^2+a^2I)p(0)]\}. \tag{3.12}$$

Note that if $\boldsymbol{E}\,.\,\boldsymbol{B}=0$ but $E^2 \neq B^2$ so that $a$ and $b$ are not zero simultaneously, the limits obtained for both $p^{(a)}(\tau)$ and $p^{(b)}(\tau)$ are well behaved if either $a\to 0$ or $b\to 0$. However, if $a$ and $b$ are both zero, then neither $p^{(a)}(\tau)$ nor $p^{(b)}(\tau)$ exist in this simultaneous limit, but the combination $p(\tau)=p^{(a)}(\tau)+p^{(b)}(\tau)$ is well behaved, yielding

$$p(\tau) \xrightarrow[\substack{a\to 0\\ b\to 0}]{} p(0) + (e/mc)\tau Fp(0) + \tfrac{1}{2}(e/mc)^2\tau^2F^2p(0). \tag{3.13}$$

In the limit $b/a \ll 1$ and also with $|\boldsymbol{p}/p_4| \ll |E/B|$, the quantity $p^{(b)}(0)/m$ reduces to the usually cited (cf Schmidt 1966, p 9) non-relativistic $(\boldsymbol{E}\times\boldsymbol{B})/B^2$ drift velocity attributable to the constant electric field. (*Note.* By making the replacement $\boldsymbol{E}\to c\boldsymbol{G}/e$, the non-relativistic equation of motion $\mathrm{d}\boldsymbol{p}/\mathrm{d}t = (e/c)(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B})$ changes into that describing a charged particle in a constant $\boldsymbol{B}$ field in the presence of a constant gravitational force $\boldsymbol{G}$. Thus, such a replacement in the drift velocity gives the corresponding expression for this situation (cf Schmidt 1966, p 10).)

Now, integrating equations (3.11) and (3.12) we obtain the two different projections of the four-vector position. The results are:

$$\begin{aligned} m(a^2+b^2)x^{(a)}(\tau) &\\ = \{&2\omega_a^{-1}(\sin\omega_a\tau/2)^2[(aF+sbF^{\mathrm{D}})p(0)]\\ &-\omega_a^{-1}(\sin\omega_a\tau)[(F^2-b^2I)p(0)] - m(F^2-b^2I)x(0)\}, \end{aligned} \tag{3.14}$$

$$\begin{aligned} m(a^2+b^2)x^{(b)}(\tau) &\\ = \{&2\lambda_b^{-1}(\sinh\lambda_b\tau/2)^2[(bF-saF^{\mathrm{D}})p(0)]\\ &+\lambda_b^{-1}(\sinh\lambda_b\tau)[(F^2+a^2I)p(0)] + m(F^2+a^2I)x(0)\}. \end{aligned} \tag{3.15}$$

In principle one may solve for $\tau$ in terms of the time $t$ via $x_4 = \mathrm{i}ct = x_4^{(a)}(\tau)+x_4^{(b)}(\tau)$, and then express all functional relationships in terms of the time $t$.

For the '$a$' projection, the oscillatory functions of $\tau$ may be eliminated between equations (3.11) and (3.14) to yield an orbit equation for the '$a$' projected motion:

$$ma\omega_a(x^{(a)}(\tau)-x^{(a)}(0))=-F(p^{(a)}(\tau)-p^{(a)}(0)). \tag{3.16}$$

Similarly, for the '$b$' projection, the hyperbolic functions of $\tau$ may be eliminated between equations (3.12) and (3.15) to yield an orbit equation for the '$b$' projected motion:

$$mb\lambda_b(x^{(b)}(\tau)-x^{(b)}(0))=F(p^{(b)}(\tau)-p^{(b)}(0)). \tag{3.17}$$

Finally, using equation (2.31), the two preceding orbit equations for the separate projected motions may be combined to give the total motion orbit relation

$$(e/c)(x(\tau)-x(0))=-(\boldsymbol{E}\,.\,\boldsymbol{B})^{-1}F^{\mathrm{D}}(p(\tau)-p(0)), \tag{3.18}$$

or, multiplying by $F$,

$$(e/c)F(x(\tau)-x(0))=p(\tau)-p(0). \tag{3.19}$$

(*Note.* The right-hand side of equation (3.18) is well defined as $(\boldsymbol{E}\,.\,\boldsymbol{B})\to 0$, since in this limit, $(p(\tau)-p(0))\to F$ operating on a well behaved four-vector. Thus, since by equation (2.6), $F^{\mathrm{D}}F=-(\boldsymbol{E}\,.\,\boldsymbol{B})I$, the limit of equation (3.18) becomes just the aforementioned well behaved four-vector.)

### 3.2. *Guiding centre motion*

In the usual treatment of the non-relativistic motion of a charged particle undergoing oscillatory motion while being subjected to the influence of constant electromagnetic fields as well as, perhaps, an additional constant non-electromagnetic force, it has become customary to decompose the motion into two parts, one part referring to the 'guiding centre' about which the oscillation takes place, and the other part referring to the oscillatory motion itself. In this section we will relate the covariant generalisations of these concepts to the '$a$' and '$b$' projections of the motion described in the preceding section.

Letting $R$ represent the four-vector position of the guiding centre and $r$ represent the four-vector displacement of the actual position from the guiding centre, one may write

$$x(\tau)=R(\tau)+r(\tau). \tag{3.20}$$

We may take the defining characteristic of $r(\tau)$ to be that it time averages to zero over a cycle of the oscillatory motion. In terms of our previous decomposition $x(\tau)=x^{(a)}(\tau)+x^{(b)}(\tau)$, we see by reference to the explicit solutions, equations (3.14) and (3.15), or more generally from the equations of motion (3.4) and (3.5), that $r(\tau)$ is just that part of $x^{a}(\tau)$ that oscillates with period $1/\omega_a$. (Note, for oscillatory motion to occur, $a$ (and hence $\omega_a$) is non-zero.) Thus, we may extract from equation (3.1) the oscillatory component, and relate it to the $x^{(a)}$ projection by

$$r(\tau)=x^{(a)}(\tau)-q^{(a)}=\{(m\omega_a)^{-1}[(\sin\omega_a\tau)p^{(a)}(0)-(\cos\omega_a\tau)Fp^{(a)}(0)]\} \tag{3.21}$$

where $q^{(a)}$ is the initial constant given by

$$q^{(a)}=[x^{(a)}(0)+(ma\omega_a)^{-1}Fp^{(a)}(0)]. \tag{3.22}$$

Note that $r(\tau)$ depends only on the initial condition $p^{(a)}(0)$ and is completely independent of the initial condition $x(0)$.

The guiding centre solution itself is

$$R(\tau)=x^{(b)}(\tau)+q^{(a)}. \tag{3.23}$$

Designating a derivative with respect to proper time by a dot over the quantity involved, we find that the oscillatory and guiding centre momenta are, respectively

$$m\dot{r}(\tau)=p^{(a)}(\tau)=(\cos\omega_a\tau)p^{(a)}(0)+a^{-1}(\sin\omega_a\tau)Fp^{(a)}(0), \tag{3.24}$$

and

$$m\dot{R}(\tau)=p^{(b)}(\tau)=(\cosh\lambda_b\tau)p^{(b)}(0)+b^{-1}(\sinh\lambda_b\tau)Fp^{(b)}(0). \tag{3.25}$$

The quantity $\dot{R}(0)$ is the 'drift' four-velocity of the guiding centre.

As an alternative to using the actual particle's proper time $\tau$, one might employ (Hellwig 1955) the guiding centre proper time $\tau_b$ defined as the proper time which would elapse for a ficticious particle on the guiding centre world-line. However it is easy to show that

$$c^2(\mathrm{d}\tau_b/\mathrm{d}\tau)^2\equiv-\dot{R}_\mu\dot{R}_\mu=c^2+\dot{r}_\mu\dot{r}_\mu=-m^{-2}p_\mu^{(b)}(0)p_\mu^{(b)}(0)$$

so, for the situation of constant fields under discussion, rate relations involving $\tau_b$ would only amount to a uniform scale change in proper time measurement.

Both the oscillatory and the guiding centre momenta satisfy the same form of equations of motion:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(m\dot{r}(\tau))=\frac{e}{mc}F(m\dot{r}(\tau)),\qquad\text{and}\qquad\frac{\mathrm{d}}{\mathrm{d}\tau}(m\dot{R}(\tau))=\frac{e}{mc}F(m\dot{R}(\tau)), \tag{3.26}$$

the difference being in the respective initial conditions. We note that although the rate of change of the guiding centre momentum $(\mathrm{d}/\mathrm{d}\tau)(m\dot{R}(\tau))$ does not have an '$a$' projection, the initial conditions for $\dot{R}(\tau)$ are such that there is a constant '$a$' projection contribution to the guiding centre momentum $\dot{R}(\tau)$ itself.

Separate orbit equations also exist for the oscillatory and guiding centre motions. From equation (3.21) and (3.24) there follows the simplest orbit relations

$$a\omega_a r(\tau)=-F\dot{r}(\tau),\qquad\qquad \dot{r}(\tau)=(e/mc)Fr(\tau). \tag{3.27}$$

Since $F$ is antisymmetric, it immediately follows that the oscillatory four-position and four-momentum are orthogonal, i.e.

$$r_\mu(\tau)\dot{r}_\mu(\tau)=0. \tag{3.28}$$

One may also obtain the guiding centre orbit equation

$$b\lambda_b(R(\tau)-R(0))=F(\dot{R}(\tau)-\dot{R}(0)) \tag{3.29}$$

which may be written in the equivalent form

$$(b\lambda_bR(\tau)-F\dot{R}(\tau))=sb(ma)^{-1}F^{\mathrm{D}}p(0)+b\lambda_bx(0). \tag{3.30}$$

The current associated with the oscillatory motion of the charged particle (which in ordinary three-space is motion in an elliptical path) gives rise to an electromagnetic moment. One may define the antisymmetric generalised electromagnetic moment tensor $A$ associated with the oscillating component of the motion by its matrix elements

$$A_{\mu\nu}=\tfrac{1}{2}(e/c)(r_\mu\dot{r}_\nu-r_\nu\dot{r}_\mu). \tag{3.31}$$

It is easy to show that, as a consequence of equations (3.26) and (3.27), the electromagnetic moment tensor $A$ is a constant, i.e. that

$$\mathrm{d}A/\mathrm{d}\tau = 0. \tag{3.32}$$

On the other hand, the symmetric tensor $S$ with elements

$$S_{\mu\nu} = \tfrac{1}{2}(e/c)(r_\mu \dot{r}_\nu + r_\nu \dot{r}_\mu) \tag{3.33}$$

satisfies

$$\mathrm{d}^2S/\mathrm{d}\tau^2 = -(2\omega_a)^2 S, \tag{3.34}$$

so it has twice the frequency of $r$. Consequently,

$$\langle S \rangle = 0$$

where $\langle\ \rangle$ signifies a time average over the gyration period of the oscillatory motion.

Use of the orbit equation (equation (3.27)) leads to the following relation between the oscillatoty Lorentz invariant $\dot{r}_\mu \dot{r}_\mu$ and the electromagnetic moment tensor:

$$\dot{r}_\mu \dot{r}_\mu = m^{-1} F_{\mu\nu} A_{\nu\mu}. \tag{3.36}$$

For the special case $\boldsymbol{E} \to 0$, the antisymmetric tensor $A$ becomes

$$\tfrac{1}{2} e_{ijk} A_{jk} \to -\tfrac{1}{2} |\boldsymbol{p} \times \boldsymbol{B}|^2 B_i / m B^4 \equiv -\mu_m B_i / B \tag{3.37}$$

$$A_{i4} \to 0.$$

Here $\mu_m$ signifies the magnitude of the magnetic moment. In this limit, $\dot{r}_4 \to 0$, so equation (3.36) reduces to the oscillatory kinetic energy relation

$$\tfrac{1}{2} m \dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} \to \mu_m B. \tag{3.38}$$

## 4. Guiding centre approximation for inhomogeneous fields

### *4.1. Equations of motion*

If the electromagnetic fields $F$ that a charged particle passes through are inhomogeneous, it may still be useful to analyse the dynamics in terms of a guiding centre decomposition provided that the field changes are small during a gyration period. Taking the gyration period to be approximately $1/\omega_a$, the criterion may be stated as

$$\left| \frac{\mathrm{d}}{\mathrm{d}\tau} (\text{field quantity}) \right| \ll \omega_a \, (\text{field quantity}). \tag{4.1}$$

An alternative way of stating this condition is that during a period of the oscillatory motion, any explicitly field dependent quantity may be represented by its proper time average over that period.

To obtain the equations of motion of the guiding centre, we once again start from equation (3.1)

$$\frac{\mathrm{d}p}{\mathrm{d}\tau} = \frac{e}{mc} Fp + g \tag{3.1$'$}$$

where we now allow a non-electromagnetic force $g$ which is assumed to vary slowly with respect to a gyration period. We use the decomposition

$$p = m(\dot{R} + \dot{r}) \tag{4.2}$$

where $R$ refers to the position of the guiding centre and $r$ refers to the displacment of the particle from the guiding centre. This displacement $r$ is characterised by $\langle r \rangle = 0$ where $\langle\ \rangle$ refers to a time average over a gyration period. Now in equation (3.1′), we expand $F$ and $g$, which may be functions of $x = R + r$, about $R$ so that to first order

$$\begin{aligned} F_{\mu\nu} &= F_{\mu\nu}^{(R)} + \left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right)^{(R)} r_\lambda \\ g_\mu &= g_\mu^{(R)} + \left(\frac{\partial}{\partial x_\lambda} g_\mu\right)^{(R)} r_\lambda. \end{aligned} \tag{4.3}$$

Here the superscript notation $(R)$ signifies 'at the position of the guiding centre'. Substituting equations (4.2) and (4.3) into (3.1), we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\tau}(\dot{R}_\mu + \dot{r}_\mu) = \Bigg[ & \frac{e}{mc} F_{\mu\nu}^{(R)} \dot{R}_\nu + \frac{1}{m} g_\mu^{(R)} + \frac{e}{mc}\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right)^{(R)} r_\lambda \dot{r}_\nu \\ & + \frac{e}{mc} F_{\mu\nu}^{(R)} \dot{r}_\nu + \frac{1}{m}\left(\frac{\partial}{\partial x_\lambda} g_\mu\right)^{(R)} r_\lambda + \frac{e}{mc}\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right)^{(R)} r_\lambda \dot{R}_\nu \Bigg]. \end{aligned} \tag{4.4}$$

With the understanding that all explicitly field dependent quantities are to be evaluated at $R$, we will now suppress the superscript notation $(R)$. Also, by the following manipulations, we can write the third term of equation (4.4) in an alternative way. Interchanging $\lambda$ and $\nu$, one may write

$$\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right) r_\lambda \dot{r}_\nu = \frac{1}{2}\left[\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right) r_\lambda \dot{r}_\nu + \left(\frac{\partial}{\partial x_\nu} F_{\mu\lambda}\right) r_\nu \dot{r}_\lambda\right] \tag{4.5}$$

which upon substitution for $(\partial/\partial x_\nu) F_{\mu\lambda}$ via Maxwell's equation

$$\frac{\partial}{\partial x_\nu} F_{\mu\lambda} + \frac{\partial}{\partial x_\lambda} F_{\nu\mu} + \frac{\partial}{\partial x_\mu} F_{\lambda\nu} = 0 \tag{4.6}$$

yields

$$\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right) r_\lambda \dot{r}_\nu = \frac{1}{2}\left[\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right) r_\lambda \dot{r}_\nu - \left(\frac{\partial}{\partial x_\lambda} F_{\nu\mu}\right) r_\nu \dot{r}_\lambda\right] - \frac{1}{4}\left[\left(\frac{\partial}{\partial x_\mu} F_{\lambda\nu}\right) r_\nu \dot{r}_\lambda + \left(\frac{\partial}{\partial x_\mu} F_{\nu\lambda}\right) r_\lambda \dot{r}_\nu\right].$$

Thus, using the antisymmetry of $F$ and the definitions (equations (3.31) and (3.33)) of the antisymmetric electromagnetic moment tensor $A$ and the symmetric tensor $S$, we obtain

$$\frac{e}{mc}\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right) r_\lambda \dot{r}_\nu = \frac{1}{m}\left(\frac{\partial}{\partial x_\lambda} F_{\mu\nu}\right) S_{\lambda\nu} + \frac{1}{2m}\left(\frac{\partial}{\partial x_\mu} F_{\nu\lambda}\right) A_{\nu\lambda}. \tag{4.7}$$

Thus, equation (4.4) becomes

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \dot{R}_\mu = \frac{e}{mc} F_{\mu\nu} \dot{R}_\nu + \frac{1}{m} g_\mu + \frac{1}{2m}\left(\frac{\partial}{\partial x_\mu} F_{\nu\lambda}\right) A_{\nu\lambda} + H_\mu \tag{4.8}$$

where

$$H_\mu = -\frac{\mathrm{d}}{\mathrm{d}\tau}\dot{r}_\mu + \frac{e}{mc}F_{\mu\nu}\dot{r}_\nu + \frac{1}{m}\left(\frac{\partial}{\partial x_\lambda}g_\mu\right)r_\lambda + \frac{1}{m}\left(\frac{\partial}{\partial x_\lambda}F_{\mu\nu}\right)S_{\lambda\nu}. \tag{4.9}$$

In the expression for $H_\mu$, the quantity $S_{\lambda\nu}$ may be evaluated in zero order since $(\partial/\partial x_\lambda)F_{\mu\nu}$ is already a first-order correction in the inhomogeneity. Thus, through first order, the time average of the $H_\mu$ over a gyration period vanishes, i.e.

$$\langle H_\mu\rangle = 0. \tag{4.10}$$

Additionally, if the time constant $\lambda_b$ is considerably less than the frequency $\omega_a$ (i.e. $b<a$) so that no appreciable change in the guiding centre motion occurs over an oscillation cycle, then $\langle R_\mu\rangle$ may be taken to be $R_\mu$ itself. At any event, redefining $R$ as $\langle R_\mu\rangle$ and referring to this quantity as the guiding centre, we see that the guiding centre motion is governed by the equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\dot{R}_\mu = \frac{e}{mc}F_{\mu\nu}\dot{R}_\nu + \frac{1}{m}g_\mu + \frac{1}{2m}\left(\frac{\partial}{\partial x_\mu}F_{\nu\lambda}\right)A_{\nu\lambda}. \tag{4.11}$$

(In a somewhat different form, this equation has been derived earlier by Hellwig (1955) and Vandervoort (1960).)

The initial condition for this equation is found by the following reasoning. Restricting the use of this equation (for the moment) to changes taking place within one cycle, and making the replacement

$$\dot{R} = \dot{\mathcal{R}} + c(esab)^{-1}F^{\mathrm{D}}g$$

the *one-cycle* dynamical equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\dot{\mathcal{R}} = \frac{e}{mc}F\dot{\mathcal{R}} + (\text{higher-order terms})$$

in which $\mathrm{d}g/\mathrm{d}\tau$ is included in the higher-order terms. Comparison with the analogous constant field dynamical equation (equation (3.26)) and reference to its initial condition (cf equations (3.25) and (4.2)) yields

$$\dot{\mathcal{R}}(0) = \mathcal{O}^{(b)}[m^{-1}p(0) - c(esab)^{-1}F^{\mathrm{D}}g]$$

which with the aid of equation (2.32) may be manipulated to give the result

$$\dot{R}(0) = m^{-1}\mathcal{O}^{(b)}p(0) + (ma\omega_a)^{-1}F\mathcal{O}^{(a)}g. \tag{4.12}$$

This initial condition, to be used in conjunction with equation (4.11), is correct to zero order. It is important, however, to retain first-order terms as well as the functional dependence of field dependent quantities in the dynamical equation (4.11), since although the inhomogeneous corrections per gyration cycle are assumed to be small, over a large number of cycles those corrections could constructively add and become comparable to zero-order expressions. In any specific problem, the functional form describing the inhomogeneous fields must be known or estimated before equation (4.11) can be (numerically) integrated.

For the special case, $g\to 0$ and $\boldsymbol{E}\to 0$, one sees with the aid of equation (3.36) that the inhomogeneous contribution in equation (4.11) to the spatial component of the guiding centre equation of motion becomes the usually cited (cf Schmidt 1966, p 14) expression

$$\frac{1}{2m}\left(\frac{\partial}{\partial x_i}F_{\nu\lambda}\right)A_{\nu\lambda}$$

$$\rightarrow\frac{1}{4m}(\epsilon_{jkl}A_{kl})\left(\frac{\partial}{\partial x_i}\epsilon_{jnp}F_{np}\right)$$

$$=-\frac{\mu_m}{mB}B_j\frac{\partial}{\partial x_i}B_j=-\frac{\mu_m}{2mB}\left(\frac{\partial}{\partial x_i}B^2\right)=-\frac{\mu_m}{mB}(\boldsymbol{B}\cdot\nabla)B_i,$$

taking curl $\boldsymbol{B}\rightarrow 0$.

### 4.2. Drift velocities

Multiplying by $-(mc/esab)F^{\mathrm{D}}$, one may invert equation (4.11) and solve for the guiding centre velocity

$$\dot{R}_\mu=\frac{-s}{\omega_a b}F^{\mathrm{D}}_{\mu\nu}(\ddot{R}_\nu-m^{-1}g_\nu)+\frac{s}{2m\omega_a b}F^{\mathrm{D}}_{\mu\nu}A_{\pi\lambda}\left(\frac{\partial}{\partial x_\nu}F_{\pi\lambda}\right). \tag{4.13}$$

From this expression can be obtained those quantities that have been termed 'drift velocities'. For example, by applying the operator $\mathcal{O}^{(a)}$ (which for small $\boldsymbol{E}$ projects out the perpendicular component to the field vector $\boldsymbol{B}$) and employing equation (2.32), we obtain

$$\dot{R}^{(a)}_\mu=-(ma\omega_a)^{-1}F_{\mu\nu}(m\ddot{R}^{(a)}_\nu-g^{(a)}_\nu)+(2ma\omega_a)^{-1}F_{\mu\nu}\mathcal{O}^{(a)}_{\nu\xi}A_{\pi\lambda}\left(\frac{\partial}{\partial x_\xi}F_{\pi\lambda}\right). \tag{4.14}$$

The term involving $(m\ddot{R}^{(a)}_\nu-g^{(a)}_\nu)$ is the 'acceleration drift' (as defined by Northrop 1961) and the term involving $(\partial/\partial x_\xi)F_{\pi\lambda}$ is the 'gradient drift'. When all electromagnetic fields are strictly constant, $g^{(a)}_\nu$ represents the leading contribution since $\ddot{R}^{(a)}_\nu$ is of higher order (cf reasoning for equation (4.12)).

When small inhomogeneities are present, $\ddot{R}^{(a)}_\nu$ is at most first order in small quantities, so equation (4.14) has the structure

$$\dot{R}^{(a)}_\mu=\begin{pmatrix}\text{at most first-order correction}\\ \text{in inhomogeneities}\end{pmatrix}+(ma\omega_a)^{-1}F_{\mu\nu}g^{(a)}_\nu.$$

Since the (proper) time derivative of an inhomogeneity dependent term produces a term which is an order smaller, it follows from the preceding equation that

$$\ddot{R}^{(a)}_\mu\equiv\mathcal{O}^{(a)}_{\mu\nu}\ddot{R}_\nu$$

$$=\frac{\mathrm{d}}{\mathrm{d}\tau}(\mathcal{O}^{(a)}_{\mu\nu}\dot{R}_\nu)-\left(\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{O}^{(a)}_{\mu\nu}\right)\dot{R}_\nu=\frac{\mathrm{d}}{\mathrm{d}\tau}\dot{R}^{(a)}_\mu-\left(\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{O}^{(a)}_{\mu\nu}\right)\dot{R}_\nu$$

$$=\frac{\mathrm{d}}{\mathrm{d}\tau}[(ma\omega_a)^{-1}F_{\mu\nu}g^{(a)}_\nu]-\left(\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{O}^{(a)}_{\mu\nu}\right)\dot{R}_\nu\qquad\text{to first order.}$$

Thus, to first order one can replace $\ddot{R}^{(a)}$ in equation (4.14) by this expression. Non-relativistically $-\mathrm{i}\dot{R}_4\approx mc\gg|\boldsymbol{R}|$, so for $E/B\ll 1$,

$$\left(\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{O}^{(a)}_{j\nu}\right)\dot{R}_\nu\rightarrow mc\frac{\mathrm{d}}{\mathrm{d}\tau}[B^{-2}(\boldsymbol{E}\times\boldsymbol{B})_j].$$

This first-order contribution to the acceleration drift is due to the fact that the '$a$' projection direction is not constant, i.e. the field lines are not straight but are curved. Replacing $\ddot{R}^{(a)}$ in equation (4.14) by this expression, one obtains (through first order)

$$\dot{R}_\mu^{(a)} = -(ma\omega_a)^{-1}F_{\mu\nu}\left[\frac{d}{d\tau}[(a\omega_a)^{-1}F_{\nu\pi}g_\pi^{(a)}] - m\left(\frac{d}{d\tau}\mathcal{O}_{\nu\pi}^{(a)}\right)\dot{R}_\pi\right] + (ma\omega_a)^{-1}F_{\mu\nu}g_\nu^{(a)}$$

$$+ (2ma\omega_a)^{-1}F_{\mu\nu}\mathcal{O}_{\nu\xi}^{(a)}A_{\pi\lambda}\left(\frac{\partial}{\partial x_\xi}F_{\pi\lambda}\right). \tag{4.15}$$

Note that non-relativistically $-i\dot{R}_4 \approx c \gg |R_j|$, hence for $E/B \to 0$, one obtains in this limit

$$\frac{1}{a\omega_a}F_{j\nu}\left(\frac{d}{d\tau}\mathcal{O}_{\nu\pi}^{(a)}\right)\dot{R}_\pi \to \frac{-mc}{eB^2}\left(\boldsymbol{B}\times\frac{d}{d\tau}\left|\frac{\boldsymbol{E}\times\boldsymbol{B}}{B^2}\right|\right)_j.$$

Also, in this limit the gradient drift becomes

$$\frac{1}{2ma\omega_a}F_{j\nu}\mathcal{O}_{\nu\xi}^{(a)}A_{\pi\lambda}\left(\frac{\partial}{\partial x_\xi}F_{\pi\lambda}\right) \to \frac{c\mu_m}{2eB^3}(\boldsymbol{B}\times\nabla B^2)_j.$$

Thus, in the appropriate limit, the more general expressions reduce to those quoted by Northrop (1961) and Schmidt (1966).

As we have seen in § 3, it is from the '$a$' projection and not the '$b$' projection of the four-vector velocity that the zero-order '$\boldsymbol{E}\times\boldsymbol{B}$ drift velocity' arises. Indeed, to zero order

$$\dot{R}_\mu^{(b)} = \mathcal{O}_{\mu\nu}^{(b)}\dot{R}_\nu = \mathcal{O}_{\mu\nu}^{(b)}(m^{-1}p_\nu - \dot{r}_\nu) \simeq m^{-1}\mathcal{O}_{\mu\nu}^{(b)}p_\nu$$

since to zero order in inhomogeneities, $\dot{r}_\nu$ is an '$a$' projection. Again, from equation (2.36) for $|\boldsymbol{p}/p_4| \ll E/B \ll 1$, $m^{-1}\mathcal{O}_{j\nu}^{(b)}p_\nu \to cB^{-2}(\boldsymbol{E}\times\boldsymbol{B})_j$.

### *4.3. Adiabatic moment invariance and guiding centre energy generalisations*

When field inhomogeneities are present, the orbit equation for the oscillatory motion must be of the form

$$\dot{r} = (e/mc)Fr + \Gamma$$

where $\Gamma$ is a first-order (or higher) correction in the inhomogeneities. Thus, the quantity $r_\mu\dot{r}_\mu$ is of first order (or higher) and hence $d/d\tau(r_\mu\dot{r}_\mu)$ is of second order (or higher). But,

$$d/d\tau(r_\mu\dot{r}_\mu) = \dot{r}_\mu\dot{r}_\mu + r_\mu\ddot{r}_\mu,$$

so substituting for $\ddot{r}_\mu$ from equation (4.8), which is correct through first order, and one-cycle time averaging the resultant expression, one obtains

$$\langle\dot{r}_\mu\dot{r}_\mu\rangle = m^{-1}F_{\mu\nu}\langle A_{\nu\mu}\rangle + (\text{terms of first order}). \tag{4.16}$$

In arriving at this result, we have used the fact that $\langle r_\mu\rangle = \langle\dot{r}_\mu\rangle = \langle S_{\mu\nu}\rangle = \langle r_\mu S_{\nu\lambda}\rangle = \langle r_\mu A_{\nu\lambda}\rangle = 0$ to zero order, and we have assumed a conservative non-electromagnetic force field

$$g_\mu = -\partial V/\partial x_\mu, \tag{4.17}$$

so that $dV/d\tau = -g_\mu p_\mu = 0$.

Differentiating equation (4.16), we obtain

$$\frac{d}{d\tau}\langle \dot{r}_\mu \dot{r}_\mu \rangle = m^{-1}\frac{d}{d\tau}(F_{\mu\nu}\langle A_{\nu\mu}\rangle) + (\text{terms of second order}). \tag{4.18}$$

Another expression involving $d/d\tau\langle \dot{r}_\mu \dot{r}_\mu\rangle$ may be obtained from $m^{-1}\,d/d\tau(p_\mu p_\mu) = m^{-1}p_\mu g_\mu = 0$, which by setting $p_\mu = m(\dot{R}_\mu + \dot{r}_\mu)$, and one-cycle time averaging yields the energy–momentum invariance relation

$$\frac{d}{d\tau}[\tfrac{1}{2}m(\dot{R}_\mu \dot{R}_\mu + \langle \dot{r}_\mu \dot{r}_\mu\rangle)] = 0 + (\text{terms of second order}). \tag{4.19}$$

However, multiplying equation (4.8) by $\dot{R}_\mu$, then one-cycle averaging, one obtains

$$\frac{d}{d\tau}(\tfrac{1}{2}m\dot{R}_\mu \dot{R}_\mu) = \tfrac{1}{2}\langle A_{\nu\lambda}\rangle\left(\frac{d}{d\tau}F_{\nu\lambda}\right) + (\text{terms of second order}). \tag{4.20}$$

On subtracting this equation from the preceding one, it follows that

$$\frac{d}{d\tau}\langle \dot{r}_\mu \dot{r}_\mu\rangle = m^{-1}\langle A_{\lambda\nu}\rangle\left(\frac{d}{d\tau}F_{\nu\lambda}\right) + (\text{terms of second order}). \tag{4.21}$$

Comparison of equation (4.21) with (4.18) leads one to the conclusion that *through* first order in inhomogeneity corrections,

$$F_{\nu\lambda}\frac{d}{d\tau}\langle A_{\lambda\nu}\rangle = 0. \tag{4.22}$$

This relation is a covariant generalisation of the adiabatic invariance of the magnetic moment. Indeed, in the limit $g \to 0$, $\boldsymbol{E} \to 0$, equation (4.22) reduces to

$$B_k\, d(\mu_m B_k/B)/d\tau = B(d\mu_m/d\tau) = 0,$$

from which it follows that $\mu_m$ is an adiabatic invariant. (*Note.* Vandervoort (1960) has discussed a quantity $4\pi\mu$ which he describes as the relativistic analogue of the magnetic flux through the region enclosed by the gyration. In terms of our notation (his equation (194)), $\mu = (2m\omega)^{-1}F_{\nu\lambda}A_{\lambda\nu}$. Although $F_{\nu\lambda}A_{\lambda\nu}$ is not an invariant through first order, $\mu$ is because $\omega$ varies according to $\dot{\omega} = (2\mu m)^{-1}\dot{F}_{\nu\lambda}A_{\lambda\nu}$ (which may be established from Vandervoort's analysis), so it follows (through first order) that $\dot{\mu} = (2m\omega)^{-1}F_{\nu\lambda}\dot{A}_{\lambda\nu} = 0$.)

Using the generalised adiabatic relation (equation (4.22)) in equation (4.20), one obtains a generalised momentum–energy invariance relation involving only guiding centre quantities and the electromagnetic moment tensor,

$$\frac{d}{d\tau}(\tfrac{1}{2}m\dot{R}_\mu \dot{R}_\mu + \tfrac{1}{2}\langle A_{\lambda\nu}\rangle F_{\nu\lambda}) = 0. \tag{4.23}$$

This expression, valid through first order in field inhomogeneity corrections, is equivalent to that given by Vandervoort (1960) (his equation (199)) who also showed that it followed as a consequence of the invariance of the square of the four-velocity. It follows that the relation between the particle's proper time $\tau$ and the guiding centre proper time $\tau_b$ no longer is uniformly proportional since $(d\tau_b/d\tau)^2$ now contains a non-constant first-order contribution.

## Acknowledgment

I would like to thank Professor N Kemmer for hospitality extended at the University of Edinburgh where this work was completed.

## References

Alfvén H 1950 *Cosmical Electrodynamics* (Oxford: Clarendon)
Alfvén H and Fälthammar C 1963 *Cosmical Electrodynamics* 2nd edn (Oxford: Clarendon)
Bernstein I 1971 *Advances in Plasma Physics* vol. 4 (New York: Interscience) pp 311–33
Derfler H 1976 *Bull. Am. Phys. Soc.* **21** 1032
Hellwig G 1955 *Z. Naturf.* a **10** 508–16
Jancel R and Kahan T 1966 *Electrodynamics of Plasmas* (London: Wiley) pp 174–238
Kruskal M 1959 *La Théorie des Gaz Neutres et Ionisés* (Paris: Hermann) pp 277–84
—— 1965 *Plasma Physics* (Vienna: International Atomic Energy Agency) pp 67–102
Mangeney A and Signore M 1974 *Phys. Lett.* **48A** 65–7
Northrop T 1961 *Ann. Phys., NY* **15** 79–101
Northrop T and Teller E 1960 *Phys. Rev.* **117** 215–25
Schmidt G 1966 *Physics of High Temperature Plasmas* (New York: Academic)
Spitzer L 1956 *Physics of Fully Ionized Gases* (New York: Interscience)
Synge J 1956 *Relativity: The Special Theory* (Amsterdam: North-Holland)
Vandervoort P 1960 *Ann. Phys., NY* **10** 401–53