

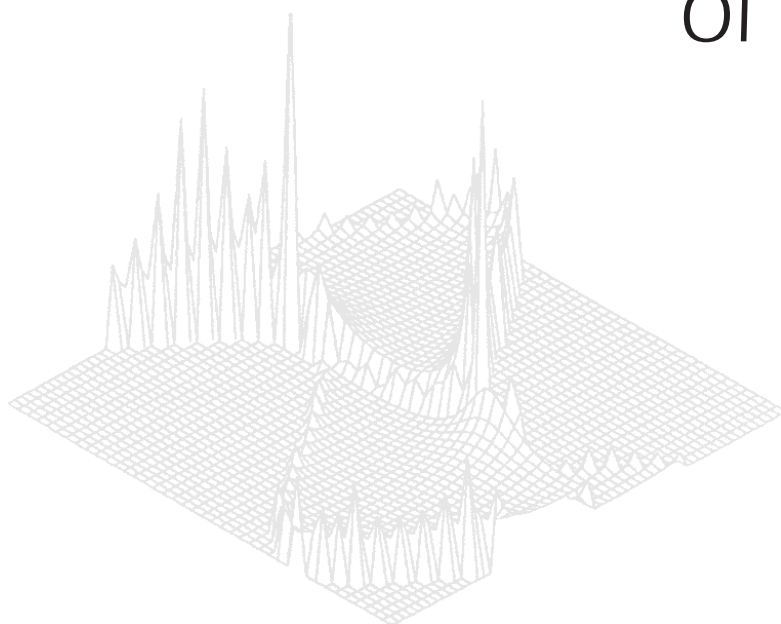
---

CSIRO PUBLISHING

---

# Australian Journal of Physics

Volume 53, 2000  
© CSIRO 2000



A journal for the publication of  
original research in all branches of physics

**[www.publish.csiro.au/journals/ajp](http://www.publish.csiro.au/journals/ajp)**

All enquiries and manuscripts should be directed to

*Australian Journal of Physics*

**CSIRO PUBLISHING**

PO Box 1139 (150 Oxford St)

Collingwood

Vic. 3066

Australia

Telephone: 61 3 9662 7626

Facsimile: 61 3 9662 7611

Email: [peter.robertson@publish.csiro.au](mailto:peter.robertson@publish.csiro.au)



Published by **CSIRO PUBLISHING**  
for CSIRO and  
the Australian Academy of Science



## General Gauge Independence of Diamagnetism plus Paramagnetism\*

A. M. Stewart

Department of Applied Mathematics,  
Research School of Physical Sciences and Engineering,  
Australian National University,  
Canberra, ACT 0200, Australia.  
email: andrew.stewart@anu.edu.au

### Abstract

With the use of the generalised multipolar gauge, a completely general proof is obtained for a result asserted by J. H. Van Vleck in 1932 but never fully proved by him. Specifically it is demonstrated that the matrix elements of the magnetic moment operator are independent of the origin of the vector potential for electromagnetic fields that are non-uniform in space and non-constant in time. A general survey of the principles of the quantum mechanics of solid state and molecular magnetism is also given to set the problem in its context. Based on the multipolar gauge, a simple conversion of the interaction Hamiltonian from the form of  $\mathbf{A} \cdot \mathbf{p}$  to  $\mathbf{E} \cdot \mathbf{r}$  plus higher order terms is demonstrated.

### 1. Introduction

The electric and magnetic Faraday–Maxwell fields  $\mathbf{E}$  and  $\mathbf{B}$  may be expressed in terms of their scalar and vector potentials  $\phi$  and  $\mathbf{A}$  by the relations

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \text{and} \quad \mathbf{E} = -\nabla\phi - \partial\mathbf{A}/\partial t, \quad (1)$$

where all the fields are a function of spatial position  $\mathbf{r}$  and time  $t$  and  $\nabla$  is the gradient operator with respect to  $\mathbf{r}$ , a bold symbol indicating a vector in three-space. If the potentials are changed to

$$\mathbf{A}_\chi = \mathbf{A} + \nabla\chi \quad \text{and} \quad \phi_\chi = \phi - \partial\chi/\partial t, \quad (2)$$

where  $\chi$  is an arbitrary scalar function of position of  $\mathbf{r}$  and  $t$ , then it is simple to see that the fields  $\mathbf{B}$  and  $\mathbf{E}$  are unchanged because in the equation for  $\mathbf{B}$  the curl of a gradient is identically zero and in the expression for  $\mathbf{E}$  two terms  $(\partial/\partial t)\nabla\chi$  appear with opposite signs and cancel each other. The fields  $\mathbf{E}$  and  $\mathbf{B}$  can be measured, for example, through the Lorentz force,

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (3)$$

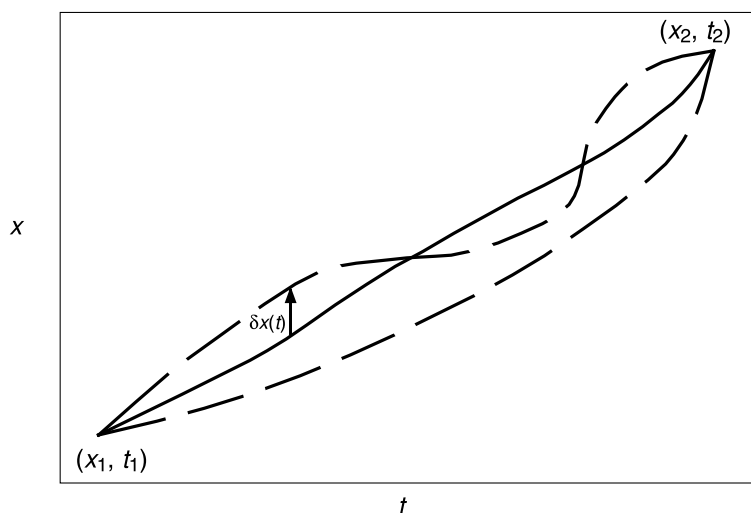
\* Refereed paper based on a contribution to the Ninth Gordon Godfrey Workshop on Condensed Matter in Zero, One and Two Dimensions held at the University of New South Wales, Sydney, in November 1999.

that they exert on a charge  $e$ , but the potentials cannot be measured. It is a generally accepted principle of physics (Doughty 1980; Stewart 1996*b*; Stewart 1997*b*) that any measurable quantity must be unaffected by the gauge transformation of equation (2).

J. H. Van Vleck claimed in 1932 that the magnetic susceptibility of an atom or molecule was independent of the origin of the vector potential  $\mathbf{A}$  used to describe the magnetic field (Van Vleck 1932). He never produced a complete proof of this contention. The main purpose of this paper is to provide a proof of it that applies to the most general situation in which the atomic or molecular system is subjected to fields that vary both in space and time. Another purpose is to give a general introduction to the principles of the quantum mechanics of solid state magnetism in order to demonstrate how the issue arose, set it in context and illustrate some uses of gauge arguments in condensed matter physics. In Section 2 of the paper we discuss classical Lagrangian mechanics, in Section 3 the transition to quantum mechanics and in Section 4 the formulation of the problem posed by Van Vleck. In Section 5 the generalised multipolar gauge is introduced, in Section 6 it is used to transform the interaction Hamiltonian from the  $\mathbf{A} \cdot \mathbf{p}$  to the  $\mathbf{E} \cdot \mathbf{r}$  forms and in Section 7, by showing that the operator for the orbital angular momentum is gauge invariant, Van Vleck's problem is solved in the most general way possible. In Section 8 a discussion is given of the role of gauge in the magnetism of molecules.

## 2. Classical Lagrangian Mechanics

Both classical and quantum dynamics are based on Hamilton's principle of Least Action (Doughty 1980; Feynman *et al.* 1964). To illustrate this we begin with a system which consists of a particle moving in one dimension between points  $(x_1, t_1)$  and  $(x_2, t_2)$ , see Fig. 1. We might imagine a stone falling under gravity. At any intervening time  $t$  the particle is at position  $x$  and has velocity  $v = dx/dt$ . The problem is to find which path the particle actually follows between the two end points (indicated by the continuous line in



**Fig. 1.** Diagram showing possible paths of a particle in the  $x$ - $t$  plane between the given points  $(x_1, t_1)$  and  $(x_2, t_2)$ . The path that is assumed to be that of Least Action is indicated by the continuous line, paths of greater action given by the dotted lines differing from the former by  $\delta x(t)$ .

Fig. 1) out of all the possible paths that it might follow (indicated by the dashed lines). The action  $S$  is defined to be the integral

$$S = \int_{t_1}^{t_2} L(x, v) dt, \quad (4)$$

where  $L(x, v)$  is a quantity called the Lagrangian which describes the complete physics of the motion. It is a function of the coordinate and velocity of the particle, which themselves are functions of time, and will be chosen to reproduce the known equations of motion. Hamilton's principle of Least Action states that the path that will be followed is that for which the action is an extremum or, in the cases dealt with in this paper, a minimum:

$$\delta S = \int_{t_1}^{t_2} \delta L(x, v) dt = 0, \quad (5)$$

where  $\delta L$  is the change in  $L$  from its minimum value occasioned by a change  $\delta x$  of the path from the path which gives the minimum action. Expressing the variation of  $L$  as  $\delta L = (\partial L / \partial x) \delta x + (\partial L / \partial v) \delta v$ , where  $\delta x$  is the deviation of  $x$  from its extremum value and noting that  $\delta v = (d/dt) \delta x$ , we get

$$\delta S = \int_{t_1}^{t_2} \{ (\partial L / \partial x) \delta x + (\partial L / \partial v) (d/dt) \delta x \} dt = 0. \quad (6)$$

An integration of the second term by parts with respect to  $t$  is now made to give

$$\delta S = [(\partial L / \partial v) \delta x]_{t_1}^{t_2} + \int_{t_1}^{t_2} \{ (\partial L / \partial x) - (d/dt) (\partial L / \partial v) \} dt \delta x = 0. \quad (7)$$

The first term vanishes because  $\delta x$  is by definition zero at the two limits, so for the action to be an extremum it is necessary for the integrand to be zero at every time. This gives the Euler-Lagrange equation

$$\partial L / \partial x = (d/dt) (\partial L / \partial v) = dp/dt, \quad (8)$$

where the quantity  $p = \partial L / \partial v$  is called the canonical momentum conjugate to the coordinate  $x$ . The Hamiltonian  $H$  is defined to be  $H(x, p) = pv - L(x, v)$ , assuming that in  $L$   $v$  is able to be expressed in terms of  $p$ . The equations of motion for  $H(p, x)$  are then  $\partial H / \partial x = -\partial L / \partial x = -dp/dt$  and  $\partial H / \partial p = v = dx/dt$ . It also follows from

$$dH(x, p)/dt = (\partial H / \partial x)(dx/dt) + (\partial H / \partial p)(dp/dt) + \partial H / \partial t, \quad (9)$$

that

$$dH/dt = -\partial L / \partial t, \quad (10)$$

as the first two terms in equation (9) cancel. Accordingly, if the Lagrangian does not explicitly depend on time then the Hamiltonian is a constant of the motion. It is given the name energy.

As a simple example of the above consider a particle moving in a potential  $V(x, t)$  having a Lagrangian  $L(x, v, t) = mv^2/2 - V(x, t)$ , in this particular case the non-relativistic kinetic energy minus the potential energy. The canonical momentum is  $p = \partial L / \partial v = mv$  and the Euler–Lagrange equation is  $dp/dt = -\partial V / \partial x$ , which is Newton’s second law. The Hamiltonian is

$$H = pv - L = p^2/2m + V(x, t),$$

the standard form for the total energy. For three dimensions the Lagrangian is generalised to  $L(\mathbf{r}, \mathbf{v}, t) = m\mathbf{v}^2/2 - V(\mathbf{r}, t)$ , the canonical momentum conjugate to each coordinate  $x^i$  is  $p^i = \partial L / \partial v^i = mv^i$  or  $\mathbf{p} = m\mathbf{v}$  and the Euler–Lagrange equation is  $d\mathbf{p}/dt = -\nabla V$ , Newton’s law in vector form. The Hamiltonian is now

$$H(\mathbf{r}, \mathbf{p}, t) = \mathbf{p} \cdot \mathbf{v} - L = \mathbf{p}^2/2m + V(\mathbf{r}, t).$$

We note that if a total time derivative, say  $dQ/dt$ , is added to the Lagrangian, the action gains the extra terms  $Q(t_1) - Q(t_2)$ . These are constants and therefore do not alter the equations of motion (8). The advantages of the Lagrangian–Hamiltonian formulation of dynamics are that it is readily adaptable to non-Cartesian coordinates, to dynamically more complicated systems, to systems of many particles and, as will be discussed in the next section, to the passage to quantum mechanics.

Next we consider the rather more involved system of a particle of charge  $e$  moving in externally applied classical electric and magnetic fields. We take the Lagrangian to be

$$L = m\mathbf{v}^2/2 + e(\mathbf{v} \cdot \mathbf{A} - \phi). \quad (11)$$

The scalar potential is as it was before, and the vector potential, as will be seen, appears in a form that will reproduce the correct equation of motion. The canonical momentum, using the defining relation  $p^i = \partial L / \partial v^i$ , is  $\mathbf{p} = m\mathbf{v} + e\mathbf{A}$  and the Hamiltonian is consequently

$$H = (\mathbf{p} - e\mathbf{A})^2/2m + e\phi = m\mathbf{v}^2/2 + e\phi.$$

Why is the  $L$  that we chose the correct Lagrangian? The Euler–Lagrange equation that it gives rise to is

$$m(d\mathbf{v}/dt) = -e(d\mathbf{A}/dt) - e\nabla\phi + e\nabla(\mathbf{v} \cdot \mathbf{A}).$$

The last term comes to  $e\{\mathbf{v} \times \mathbf{B} + (\mathbf{v} \cdot \nabla)\mathbf{A}\}$  because  $\nabla$  acting on  $\mathbf{v}$  gives zero. We next use the definition of the total derivative  $d\mathbf{A}/dt = \partial\mathbf{A}/\partial t + (\mathbf{v} \cdot \nabla)\mathbf{A}$ , where the second term is the convective derivative which describes how the  $\mathbf{A}$  field felt by the particle varies because its velocity moves the particle into a new position where  $\mathbf{A}$  is different. The Euler–Lagrange equation then becomes, as required, the Lorentz force equation  $m(d\mathbf{v}/dt) = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ . If we carry out the gauge transformation (2) then the Lagrangian acquires the extra term  $e(d\chi/dt)$ , again using the notion of the convective derivative. This is a total time derivative so the Lorentz force equation is unchanged.

### 3. Quantum Mechanics

Semi-classical wave mechanics, which describes the quantum motion of a particle in a prescribed classical electromagnetic field, is obtained from classical mechanics by making two assumptions.

1. The Hamiltonian is taken to be equivalent to the operation  $i\hbar(\partial/\partial t)$  acting upon a wave-function  $\Psi(\mathbf{r}, t)$  which is the probability density amplitude of finding the particle at position  $\mathbf{r}$  at time  $t$  so

$$H(\mathbf{r}, \mathbf{p}, t) \Psi(\mathbf{r}, t) = i\hbar(\partial/\partial t)\Psi(\mathbf{r}, t). \quad (12)$$

2. The canonical momentum  $\mathbf{p}$  of the particle is found from Dirac's commutation relations  $[r_i, p_j] = i\hbar\delta_{ij}$  and  $[p_i, p_j] = 0$ . The general solution of these equations is  $\mathbf{p} = -i\hbar\nabla + \nabla F(r_i)$ , where  $F$  is an arbitrary function of the  $r_i$ . By making the unitary transformation  $\mathbf{p} \rightarrow U\mathbf{p}U^{-1}$ , where  $U = \exp(iF/\hbar)$ ,  $F$  is absorbed and  $\mathbf{p}$  is reduced to its standard form  $\mathbf{p} = -i\hbar\nabla$  (Dirac 1947). This leads to the Schrödinger equation for a charged particle in a prescribed classical electromagnetic field

$$\{(-i\hbar\nabla - e\mathbf{A})^2/2m + e\phi\} \Psi(\mathbf{r}, t) = i\hbar(\partial/\partial t)\Psi(\mathbf{r}, t). \quad (13)$$

Next we chose an  $\mathbf{A}$  such as  $\mathbf{A} = \mathbf{r} \times \mathbf{B}^0/2$ , where  $\mathbf{B}^0$  is a constant vector, that will describe a uniform and time independent magnetic field and choose a scalar potential that is independent of time. It is simple to show that for this vector potential  $\nabla \times \mathbf{A} = \mathbf{B}^0$  and  $\nabla \cdot \mathbf{A} = 0$ . This is called the symmetric uniform gauge, but other equivalent gauges (a particular set of potentials) that differ by a gauge transformation (2) and have  $\nabla \cdot \mathbf{A} = 0$  are equally allowable. Next we multiply out the square in the Hamiltonian on the left-hand side of (13), express the wavefunction as a stationary state  $\Psi(\mathbf{r}, t) = \psi(\mathbf{r})\exp(-iEt/\hbar)$  and separate the space and time variables to get

$$\{-\hbar^2\nabla^2/2m + i\hbar(\mathbf{r} \times \nabla) \cdot \mathbf{B}^0 e/2m + e^2(\mathbf{B}^0)^2(x^2 + y^2)/8m + e\phi\} \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (14)$$

On the left-hand side of this equation the first term comes from  $\mathbf{p}^2$  in the expansion of the square in the first term of equation (13) and the last term is the potential energy. The third term comes from the square of  $\mathbf{A}$ , the second from the cross product of the  $\nabla$  and  $\mathbf{A}$  terms. A term involving  $(\mathbf{p} \cdot \mathbf{A} - \mathbf{A} \cdot \mathbf{p}) = -i\hbar(\nabla \cdot \mathbf{A})$  vanishes for this particular gauge. The particular form of the  $(\mathbf{B}^0)^2$  term assumes that  $\mathbf{B}^0$  is in the  $z$  direction (see Section 6). A term  $-\mathbf{s} \cdot \mathbf{B}^0 e\hbar/m$  involving spin  $s$  may be put in *ad hoc* but is omitted for the time being as spin does not involve gauge in any essential way. The term  $\mathbf{r} \times \mathbf{p} = -i\hbar(\mathbf{r} \times \nabla)$  corresponds to the canonical orbital angular momentum about the origin.

### 4. The Problem posed by Van Vleck

Equation (14) is the time independent Schrödinger equation for a charge in a uniform magnetic field. Van Vleck (1932) used it to calculate the magnetic susceptibilities of magnetic ions in solids. The magnetic moment  $\mathbf{m}$  is obtained from the thermodynamic relation  $\mathbf{m} = -\partial E/\partial \mathbf{B}$  and the magnetic susceptibility  $\chi_M$  from  $\chi_M = -\partial^2 E/\partial \mathbf{B}^2$ . The susceptibility therefore has two contributions arising from treating the two terms in  $\mathbf{B}$  in the Hamiltonian of equation (14) by perturbation theory. The first, the diamagnetic term, comes from treat-

ing the  $\mathbf{B}^2$  term by first order perturbation theory, the second, the paramagnetic, from treating the term linear in  $\mathbf{B}$  by second order perturbation theory. The resulting theory has been very successful in describing the magnetic properties of ions in solids.

If we take the origin of  $\mathbf{A}$  to be  $\mathbf{R}$  and not the origin we get a new vector potential with  $\mathbf{A} = (\mathbf{r} - \mathbf{R}) \times \mathbf{B}^0/2$ . This amounts to making a gauge transformation since the new potential may be expressed as  $\mathbf{A} = \mathbf{r} \times \mathbf{B}^0/2 - \nabla\{\mathbf{r} \cdot (\mathbf{R} \times \mathbf{B}^0)/2\}$ , the gauge function being  $-\mathbf{r} \cdot (\mathbf{R} \times \mathbf{B}^0)/2$ . This new gauge clearly gives the same magnetic field  $\mathbf{B}^0$ . The question that Van Vleck asked was whether the magnetic susceptibility depended upon the origin of the vector potential. He claimed that when quantum conditions and commutation rules were used the susceptibility was found to be independent of the gauge origin, but he never provided a full derivation of this. Some proofs of Van Vleck's claim have been offered since. It was claimed by Griffith (1961) that in the Coulomb gauge ( $\nabla \cdot \mathbf{A} = 0$ ) and with zero scalar potential, the energy to second order in perturbation theory did not depend on the gauge function adopted for any general vector potential. The restriction to zero scalar potential severely limited the usefulness of this work. The same result was derived by Friar and Fallieros (1981) for a non-zero scalar potential with the symmetric uniform gauge using perturbation theory and they showed that the total susceptibility was gauge invariant for spherical systems when the orbital magnetic moment operator was taken to be  $\mathbf{r} \times (\mathbf{p} - e\mathbf{A})e/2m$ . They noted that the discussion was more involved for off-diagonal matrix elements of non-spherical systems. They also used the requirement of gauge invariance to obtain sum rules from perturbation theory. The derivations above applied only to the linear response (the susceptibility) and relied on perturbation theory or had other limitations. Recently a proof has been given that applies to the magnetic moment itself (i.e. the full non-linear response to a magnetic field) and involves essentially operators alone and does not rely on perturbation theory (Stewart 1996a, 1997b).

However, all the proofs mentioned above used the symmetric uniform gauge with vector potential  $\mathbf{A} = \mathbf{B}^0 \times \mathbf{r}/2$ , where  $\mathbf{B}^0$  is a vector that does not depend on the space  $\mathbf{r}$  and time  $t$  coordinates. The scalar potential  $\phi$  is independent of time. From the relations that give the fields in terms of the potentials it follows that in this gauge  $\mathbf{B}(\mathbf{r}, t) = \mathbf{B}^0$  and  $\mathbf{E}(\mathbf{r}, t) = -\nabla\phi(\mathbf{r})$ , both fields being independent of time and  $\mathbf{B}$  being uniform in space.

The main purpose of this paper is to show how the proof of Van Vleck's contention may be extended to the most general situation in which the atomic or molecular system is subjected to fields that vary both in space and time and moreover provide a proof that applies to the matrix elements as well as to the expectation value (Stewart 1999). In the next section of the paper the generalised multipolar gauge is introduced. This gauge describes electric and magnetic fields that vary arbitrarily in space and time. It is shown that in this gauge, as in the uniform gauge, a shift of the origin of the potentials corresponds to a gauge transformation. It is then demonstrated how this leads to a fully general proof of Van Vleck's contention.

## 5. Generalised Multipolar Gauge

The well known electromagnetic retarded potentials (Panofsky and Philips 1955) are expressed in terms of the electric sources, the charges and currents at a retarded time. The rather less well known multipolar gauge expresses the potentials in terms of the fields  $\mathbf{E}$  and  $\mathbf{B}$  at the same time but at non-local positions. This gauge is obtained by making the gauge transformation

$$\mathbf{A}(\mathbf{r}, t) = \mathbf{A}'(\mathbf{r}, t) + \nabla\Lambda(\mathbf{r}, t) \quad \text{and} \quad \phi(\mathbf{r}, t) = \phi'(\mathbf{r}, t) - \partial\Lambda/\partial t \quad (15)$$

from any gauge  $A'$  and  $\phi'$  to the multipolar gauge  $A$  and  $\phi$  using the gauge function

$$\Lambda(\mathbf{r}, t) = \Lambda(\mathbf{R}, t) - \int_{\mathbf{R}, \text{SL}}^{\mathbf{r}} A'(\mathbf{q}, t) \cdot d\mathbf{q} . \quad (16)$$

The first term  $\Lambda(\mathbf{R}, t)$  on the right is not a function of  $\mathbf{r}$ . It is chosen to satisfy the condition  $\partial\Lambda(\mathbf{R}, t)/\partial t = \phi'(\mathbf{R}, t)$  that will be needed later. The second term is the line integral in coordinate space of  $A'$  from  $\mathbf{R}$  to  $\mathbf{r}$ , the variable of integration being the vector  $\mathbf{q}$ . This term in general depends on the path of integration because if it did not the integrand would have to be a gradient, and since the curl of a gradient is zero it would correspond to a zero, and hence trivial, magnetic field. The path of integration therefore needs to be defined and for our purpose it is defined to be the straight line AB in Fig. 2 connecting the points  $\mathbf{R}$  and  $\mathbf{r}$ ; this is indicated by the letters SL on the integral. Hence the integration may be parametrised by a variable  $u$ , where  $0 \leq u \leq 1$  and

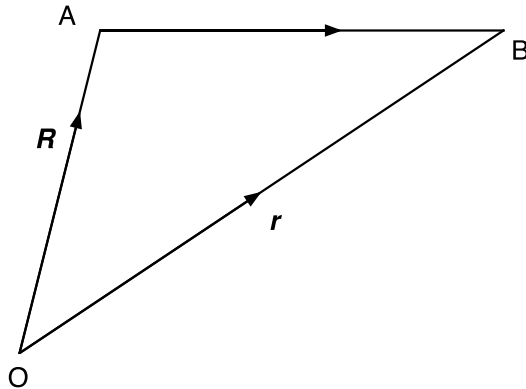
$$\mathbf{q} = u\mathbf{r} + (1-u)\mathbf{R} , \quad (17)$$

with  $u$  being zero at  $\mathbf{q} = \mathbf{R}$  and unity at  $\mathbf{q} = \mathbf{r}$ . A number of expressions relating the derivatives of a scalar function of  $\mathbf{q}$  with respect to  $\mathbf{q}$ ,  $u$ ,  $\mathbf{r}$  and  $\mathbf{R}$  are obtained in the Appendix to this paper. From (17) the path increment  $d\mathbf{q}$  is  $(\mathbf{r} - \mathbf{R})du$ , leading to

$$\Lambda(\mathbf{r}, t) = \Lambda(\mathbf{R}, t) - \int_0^1 du (\mathbf{r} - \mathbf{R}) \cdot A'(\mathbf{q}(u), t) . \quad (18)$$

This gauge function is obtained by requiring  $A$  to be perpendicular to the line  $(\mathbf{r} - \mathbf{R})$  at every point on that line which implies that

$$\int_0^1 du (\mathbf{r} - \mathbf{R}) \cdot A(\mathbf{q}) = 0. \quad (19)$$



**Fig. 2.** Vectors  $\mathbf{R}$  and  $\mathbf{r}$  in three dimensional space and the triangle OBA defined by them. Here O is the origin of the coordinate system,  $\mathbf{r}$  is the point at which the electromagnetic fields  $\mathbf{E}(\mathbf{r})$  and  $\mathbf{B}(\mathbf{r})$  are measured and  $\mathbf{R}$  is a vector on which the fields do not depend but on which the potentials do depend.



The gauge function of equation (18) is obtained by using (15) in (19) and noting that  $\nabla_{\mathbf{q}}\Lambda(\mathbf{q}) = \nabla\Lambda(\mathbf{q})/u$ , where  $\nabla_{\mathbf{q}}$  is the gradient with respect to  $\mathbf{q}$  and  $\nabla$  is the gradient with respect to  $\mathbf{r}$ . The relation  $(\mathbf{r} - \mathbf{R}) \cdot \nabla\Lambda(\mathbf{q}) = u(\partial\Lambda/\partial u)$  derived in the Appendix is then used to give (18).

We obtain the potentials from (18) using the methods of Kobe (1982) but generalised to the case of non-zero  $\mathbf{R}$ . First we obtain  $\mathbf{A}(\mathbf{r})$  by taking the gradient with respect to  $\mathbf{r}$  of the gauge function in equation (18). Henceforth we omit the time argument  $t$  as it is the same in all the potentials and fields; this is an instantaneous gauge. There are four terms arising from  $\nabla\Lambda$ . One of them vanishes from  $\nabla \times \mathbf{r} = 0$ . Two others, using the result  $(\mathbf{r} - \mathbf{R}) \cdot \nabla\Lambda(\mathbf{q}) = u(\partial\Lambda/\partial u)$  come to  $-\partial/\partial u(u\mathbf{A}')$  whose integral cancels  $\mathbf{A}'$  in equation (15). The remaining term, using the relation  $\nabla \times \mathbf{A}(\mathbf{q}) = u\mathbf{B}(\mathbf{q})$ , comes to

$$\mathbf{A}(\mathbf{r}, \mathbf{R}) = -(\mathbf{r} - \mathbf{R}) \times \int_0^1 u \, du \, \mathbf{B}(\mathbf{q}), \quad (20)$$

where the argument  $\mathbf{R}$  is explicitly exhibited. The potential  $\mathbf{A}(\mathbf{r}, \mathbf{R})$  has the property that at every point  $\mathbf{r}$  it is perpendicular to the vector  $(\mathbf{r} - \mathbf{R})$ .

The scalar potential is obtained from equation (15) by noting that  $\partial\mathbf{A}'(\mathbf{q})/\partial t = -\mathbf{E}(\mathbf{q}) - \nabla\phi'(\mathbf{q})/u$ . Using the result in the Appendix, the second term gives rise to  $-\partial/\partial u\{\phi'(\mathbf{q})\}$  whose integral cancels  $\phi'$ , giving, with the condition  $\partial\Lambda(\mathbf{R}, t)/\partial t = \phi'(\mathbf{R}, t)$ , the scalar potential

$$\phi(\mathbf{r}, \mathbf{R}) = -(\mathbf{r} - \mathbf{R}) \cdot \int_0^1 du \, \mathbf{E}(\mathbf{q}), \quad (21)$$

so all the components of the potential are zero at  $\mathbf{r} = \mathbf{R}$ .

Equations (20) and (21) specify the generalised multipolar gauge. Its validity is confirmed by verifying that operating on the potentials (20) and (21) with equations (1) recovers the fields  $\mathbf{B}(\mathbf{r}, t)$  and  $\mathbf{E}(\mathbf{r}, t)$ . The curl of equation (20) has four terms. One is zero from  $\nabla \cdot \mathbf{B} = 0$ . The other three, using the identities  $\{\mathbf{B}(\mathbf{r}) \cdot \nabla\}\mathbf{r} = \mathbf{B}(\mathbf{r})$  and  $\nabla \cdot \mathbf{r} = 3$ , come to  $\partial/\partial u\{u^2\mathbf{B}(\mathbf{q})\}$  whose integral over  $u$  gives  $\mathbf{B}(\mathbf{r})$  as required. The electric field  $\mathbf{E}(\mathbf{r})$  is obtained from equations (1), (20) and (21) by noting that  $\partial\mathbf{B}(\mathbf{q})/\partial t = -\nabla \times \mathbf{E}(\mathbf{q})/u$ . The remaining terms of the integrand come to  $\partial/\partial u\{u\mathbf{E}(\mathbf{q})\}$  whose integral is  $\mathbf{E}(\mathbf{r})$ . Neither  $\mathbf{E}(\mathbf{r})$  nor  $\mathbf{B}(\mathbf{r})$  involve  $\mathbf{R}$ , so its presence in  $\mathbf{A}$  and  $\phi$  indicates that a gauge freedom associated with  $\mathbf{R}$  exists in these potentials. If the magnetic field happens to be uniform in space then the integral over  $u$  may be performed trivially to give the symmetric uniform gauge with  $\mathbf{A} = \mathbf{B} \times (\mathbf{r} - \mathbf{R})/2$ . In the symmetric uniform gauge  $\mathbf{R}$  corresponds to a change of the origin of the potentials and is associated with a gauge transformation with gauge function  $-\mathbf{r} \cdot (\mathbf{B}_0 \times \mathbf{R})/2$ , and  $\mathbf{R}$  will be found to play a similar role in the generalised multipolar gauge. If the electric field is uniform in space then the scalar potential becomes  $\phi = -(\mathbf{r} - \mathbf{R}) \cdot \mathbf{E}$ .

The four-vector version of the multipolar gauge, otherwise known as the Poincaré gauge, the line gauge or the point gauge, was introduced by Valatin (1954). Its three-vector version, which we use here, was suggested later by Woolley (1973, 1974) and discussed by Kobe (1982). Skagerstam (1983) has examined the relation between the three- and four-vector versions. Brittin *et al.* (1982) gave an interesting application of this gauge to calculate the voltage induced in the closed circuit formed by the edges of a Mobius strip in a changing uniform magnetic field. A prototype of the multipolar gauge had been proposed

earlier by Goppert-Mayer (1931) who used  $A'(q(u), t) = A'(\mathbf{R}, t)$  in (18) to give a gauge function  $-(\mathbf{r} - \mathbf{R}) \cdot \mathbf{A}'(\mathbf{R}, t)$ . The multipolar gauge with  $\mathbf{R} = \mathbf{0}$  we call the simple *multipolar* gauge. An explanation of the descriptor *multipolar* is given in the next section.

Can we express the gauge freedom implied by the presence of  $\mathbf{R}$  in the vector potential in the form of a gauge transformation between  $\mathbf{A}(\mathbf{r}, \mathbf{R})$  and  $\mathbf{A}(\mathbf{r}, \mathbf{0})$  and if so what is the gauge function for the transformation? To answer this question we use equations (15) and (16) to obtain the difference

$$A(\mathbf{r}, \mathbf{R}) - A(\mathbf{r}, \mathbf{0}) = \nabla \left[ \int_{\mathbf{O}}^{\mathbf{B}} + \int_{\mathbf{B}}^{\mathbf{A}} \right] A'(q) \cdot d\mathbf{q}, \quad (22)$$

noting that the  $A'$  terms cancel, that the gradients with respect to  $\mathbf{r}$  of  $A(\mathbf{R}, t)$  and  $A(\mathbf{0}, t)$  are zero and reversing some limits of integration. But by Stokes' law the magnetic flux  $\Phi(\mathbf{r}, \mathbf{R})$  that is linked by the triangle OBA of Fig. 2 formed by the vectors  $\mathbf{r}$  and  $\mathbf{R}$  is given by the line integral of the vector potential around the sides of the triangle

$$\Phi(\mathbf{r}, \mathbf{R}) = \left[ \int_{\mathbf{O}}^{\mathbf{B}} + \int_{\mathbf{B}}^{\mathbf{A}} + \int_{\mathbf{A}}^{\mathbf{O}} \right] A'(q) \cdot d\mathbf{q}, \quad (23)$$

and so

$$A(\mathbf{r}, \mathbf{R}) - A(\mathbf{r}, \mathbf{0}) = \nabla \left[ \Phi + \int_{\mathbf{O}}^{\mathbf{A}} A'(q) \cdot d\mathbf{q} \right]. \quad (24)$$

Since the line integral on the right-hand side of (24) does not depend on  $\mathbf{r}$  (the limits of the integral are  $\mathbf{0}$  and  $\mathbf{R}$ ) its gradient with respect to  $\mathbf{r}$  is zero and we conclude that the gauge function for the transformation is equal to the magnetic flux linked with triangle OAB. This result can be obtained by other methods (Stewart 1999). It is consistent with the uniform gauge where the corresponding gauge function is  $-\mathbf{r} \cdot (\mathbf{B}^0 \times \mathbf{R})/2$  or  $\mathbf{B}^0(\mathbf{r} \times \mathbf{R})/2$ , the scalar product of the uniform magnetic field and the area of the triangle. We have therefore shown that changing the origin  $\mathbf{R}$  of the potentials in the generalised multipolar gauge, just as in the uniform gauge, amounts to making a gauge transformation, a result that will be used in a later section of this paper. It is to be noted that the arguments given above apply just as well to time dependent fields and potentials as to time independent ones. It is also worth noting that although the defining equations (20) and (21) for the multipolar gauge appear to be gauge invariant because they involve only the fields  $\mathbf{E}$  and  $\mathbf{B}$ , full gauge arbitrariness is nevertheless still present because the gradient of any gauge function can be added to (20) and its time derivative to (21) without changing the fields.

## 6. $A.p$ and $E.r$

The multipolar gauge was given this name (Kobe 1982) because if the fields are expanded about the point  $\mathbf{R}$  using the relation  $\mathbf{E}(\mathbf{R} + \mathbf{y}) = \exp(\mathbf{y} \cdot \nabla) \mathbf{E}(\mathbf{R})$ , where the spatial derivatives of the vector field  $\mathbf{E}$  are evaluated at the point  $\mathbf{R}$ , then the integrals over  $u$  may be

performed trivially and the potentials may be expressed in a multipole expansion, expanding as far as the quadrupole terms, as (Stewart 1999)

$$\phi = -(\mathbf{r} - \mathbf{R}) \cdot \mathbf{E}(\mathbf{R}) - \sum_{i,j} (x^i - X^i)(x^j - X^j)(\partial E^i / \partial x^j) / 2 + \dots \quad (25)$$

$$A^k(\mathbf{r}) = -(\mathbf{r} - \mathbf{R}) \times \mathbf{B}(\mathbf{R}) \cdot \mathbf{p} / 2 - \sum_{i,j,l} \varepsilon_{ijk} (x^i - X^i)(x^j - X^j)(\partial B^l / \partial x^k) / 3 + \dots \quad (26)$$

where  $\varepsilon_{ijk}$  is the antisymmetric unit tensor and the derivatives are evaluated at  $\mathbf{R}$ . Although the simple and generalised multipolar gauges do not in general satisfy the Coulomb gauge condition  $\nabla \cdot \mathbf{A} = 0$ , their multipole expansions up to but not including the magnetic quadrupole term do, which makes them convenient for use in many atomic spectroscopy calculations. By substituting (25) and (26) into equation (13) the Hamiltonian in (12) becomes

$$H = -\hbar^2 \nabla^2 / 2m - e(\mathbf{r} - \mathbf{R}) \cdot \mathbf{E}(\mathbf{R}) - e \sum_{i,j} (x^i - X^i)(x^j - X^j)(\partial E^i / \partial x^j) / 2 + H_M, \quad (27)$$

where  $H_M$  is the part of the Hamiltonian containing  $\mathbf{B}$ . This contains two terms; the cross terms between  $\mathbf{A}$  and  $\mathbf{p}$  is  $-(\mathbf{r} - \mathbf{R}) \times \mathbf{B}(\mathbf{R}) \cdot \mathbf{p} / 2m$  and the diamagnetic  $A^2$  term is  $\{(\mathbf{r} - \mathbf{R}) \times \mathbf{B}(\mathbf{R})\}^2 / 8m$  giving

$$H_M = -\mu_B \mathbf{l} \cdot \mathbf{B}(\mathbf{R}) + [(\mathbf{r} - \mathbf{R})^2 \mathbf{B}(\mathbf{R})^2 - \{(\mathbf{r} - \mathbf{R}) \cdot \mathbf{B}(\mathbf{R})\}^2] e^2 / 8m + \dots \quad (28)$$

where the canonical orbital angular momentum about  $\mathbf{R}$  is  $\mathbf{l} = (\mathbf{r} - \mathbf{R}) \times \mathbf{p} / \hbar$ . The first term is the scalar product of the magnetic field at  $\mathbf{R}$  with the canonical orbital angular momentum about  $\mathbf{R}$ . The square bracket in the diamagnetic term may be written in a manifestly positive definite form as

$$\{(x - X)B^y - (y - Y)B^x\}^2 + \{(x - X)B^z - (z - Z)B^x\}^2 + \{(y - Y)B^z - (z - Z)B^y\}^2.$$

When  $\mathbf{B}$  is in the  $z$  direction this becomes the familiar  $\{(x - X)^2 + (y - Y)^2\} B^z^2$ . The magnetic quadrupole and higher order terms are neglected. If higher order terms are needed in (27) then they can be obtained by expanding (24) and (25) further, noting that the  $\nabla \cdot \mathbf{A}$  contribution must be included for terms for which this is not zero.

In a second case we consider the system to be, as in a typical spectroscopy experiment, subjected to a static atomic potential and a static and uniform magnetic field  $\mathbf{B}^0$  together with perturbing time dependent but spatially uniform electric and magnetic fields  $\mathbf{E}^1$  and  $\mathbf{B}^1$ , the electric quadrupole and higher components of the driving fields being neglected. In this situation the potentials are the sum of two terms  $(A^0, \phi^0)$  and  $(A^1, \phi^1)$ , where  $A^0 = -(\mathbf{r} - \mathbf{R}) \times \mathbf{B}^0 / 2$ ,  $\phi^0$  is the atomic potential,  $A^1(t) = -(\mathbf{r} - \mathbf{R}) \times \mathbf{B}^1(t) / 2$  and  $\phi^1 = -(\mathbf{r} - \mathbf{R}) \cdot \mathbf{E}^1(t)$ , the terms with superscript zero being independent of time, the terms with superscript one depending on time. The Hamiltonian is

$$H = \{\mathbf{p} - e(\mathbf{A}^0 + \mathbf{A}^1)\}^2 / 2m + e(\phi^0 + \phi^1) - 2\mu_B (\mathbf{B}^0 + \mathbf{B}^1) \cdot \mathbf{s}, \quad (29)$$

where an interaction of the spin with the magnetic fields has been introduced *ad hoc* and  $\mu_B = e\hbar/2m$  is the Bohr magneton. By multiplying out the square, the Hamiltonian may be

partitioned into two terms, a time independent one  $H^0$  and a time dependent one  $V(t)$  with  $H = H^0 + V(t)$ , where

$$H^0 = (\mathbf{p} - e\mathbf{A}^0)^2/2m + e\phi^0 - 2\mu_B \mathbf{B}^0 \cdot \mathbf{s} \dots \text{plus any other static terms} \quad (30)$$

$$V(t) = e\phi^1 - \mathbf{A}^1 \cdot (\mathbf{p} - e\mathbf{A}^0)/m - 2\mu_B \mathbf{B}^1 \cdot \mathbf{s} + (\mathbf{A}^1)^2 e^2/2m, \quad (31)$$

which, by substituting the fields for the potentials, may be expressed as

$$V(t) = -e(\mathbf{r} - \mathbf{R}) \cdot \mathbf{E}^1 - \mu_B (\mathbf{l} + 2\mathbf{s}) \cdot \mathbf{B}^1 + (\mathbf{r} \times \mathbf{B}^1) \cdot (\mathbf{r} \times \mathbf{B}^0) e^2/4m + (\mathbf{r} \times \mathbf{B}^1)^2 e^2/8m. \quad (32)$$

For systems with many electrons the electron–electron interaction is added to  $H^0$  which is summed over electrons to give eigenstates which form a set of stationary states between which transitions may be induced by the time dependent perturbation  $V(t)$ , again summed over electrons. The first term of equation (32) is the electric dipole interaction, the second the magnetic dipole interaction, typically two orders of magnitude smaller except in cases where electric dipole transitions are forbidden, the third is a cross term between  $\mathbf{B}^0$  and  $\mathbf{B}^1$  and the last a diamagnetic term. If  $(\mathbf{l} + 2\mathbf{s})$  is taken to be 1 Bohr magneton,  $\mathbf{B}^0$  to be 10 Tesla and  $r = 0.035$  nm, the radius of the 4f shell of a Rare Earth atom, the ratio of the magnitude of the cross term to that of the magnetic dipole term is  $10^{-5}$ , so in these circumstances the cross term is negligible. If  $\mathbf{B}^0$  were large enough to make the cross term significant it is likely that a fully relativistic treatment would be needed. The diamagnetic term is smaller than the cross term by a factor  $B^0/B^1$ . The formalism above may be extended to include an arbitrary gauge function but it is found that transition amplitudes are essentially independent of gauge (Stewart 1997a).

It is seen that the transformation to the multipolar gauge has the effect of converting the leading term in the Hamiltonian of equations (13) and (14) that describes the interaction between the charge and the fields from the minimal coupling form  $\mathbf{A} \cdot \mathbf{p}$  to the electric dipole coupling form  $\mathbf{E} \cdot \mathbf{r}$  plus higher terms. The latter is often easier to use in calculations (Cohen-Tannoudji *et al.* 1989). The conversion from minimal coupling to electric dipole interaction may also be accomplished by means of a unitary transformation (Craig and Thirunamachandran 1998; Power and Zienau 1959) in the scheme in which both fields and particles are quantised. The approach of the present paper is semi-classical in that only the particle motion is quantised and the electromagnetic field is treated classically (Stewart 1997a). Although such an approach cannot account for effects associated with the quantisation of the electromagnetic field, such as the Lamb shift and long range interactions between atoms (Babiker and Loudon 1983), it does not contribute any error to the results of the calculations contained in this paper.

## 7. Gauge Invariance of the Magnetic Moment

The classical expression for the orbital magnetic moment  $\mathbf{m}$  about the point  $\mathbf{R}'$  of a particle of charge  $e$  at position  $\mathbf{r}$  moving with velocity  $\mathbf{v}$  is  $\mathbf{m} = (\mathbf{r} - \mathbf{R}') \times e\mathbf{v}/2$ . Because the physical system considered is of finite extent, such as an atom or molecule, the drift velocity, which is the expectation value of  $\mathbf{v}$ , is zero and consequently the expectation value of  $\mathbf{m}$  is independent of  $\mathbf{R}'$ . We do not consider systems in which there is a net drift current. The quantum mechanical operator identified with the particle velocity is  $\mathbf{v} = d\mathbf{r}/dt = [\mathbf{r}, H]/i\hbar$ , where  $H$  is the Hamiltonian. By commuting  $\mathbf{r}$  with the non-relativistic Hamiltonian  $H = (\mathbf{p} - e\mathbf{A})^2/2m + e\phi$  the velocity operator  $\mathbf{v}$  is given by  $m\mathbf{v} = \mathbf{p} - e\mathbf{A}$ . If

other terms involving  $\mathbf{p}$  are present in the Hamiltonian the commutator of them with  $\mathbf{r}$  will contribute further to the velocity operator. For example, the spin-orbit interaction  $\mathbf{s} \times \mathbf{E} \cdot (\mathbf{p} - e\mathbf{A})e\hbar/4m^2c^2$  (Frohlich and Studer 1993) will add a term  $\mathbf{s} \times \mathbf{E}e\hbar/4m^2c^2$  to  $\mathbf{v}$  and so add a term  $[\mathbf{s}\{(\mathbf{r} - \mathbf{R}') \cdot \mathbf{E}\} - \mathbf{E}\{(\mathbf{r} - \mathbf{R}') \cdot \mathbf{s}\}]e^2\hbar/8m^2c^2$  that is manifestly gauge invariant to the orbital magnetic moment. If the Dirac Hamiltonian is used then the velocity is  $\mathbf{v} = c\boldsymbol{\alpha}$ , where  $\boldsymbol{\alpha}$  is the Dirac matrix, and gauge is not involved at all. The operator for the orbital moment in the non-relativistic case therefore contains two terms  $\mathbf{m} = \mathbf{m}^p + \mathbf{m}^d$ , where the paramagnetic moment is  $\mathbf{m}^p = (\mathbf{r} - \mathbf{R}') \times \mathbf{p}/2m$  and the diamagnetic moment is  $\mathbf{m}^d = -(\mathbf{r} - \mathbf{R}') \times \mathbf{A}e^2/2m$ .

We now make a gauge transformation to a new gauge described by a gauge function  $\chi(\mathbf{r}, t)$ . The electromagnetic potentials are transformed according to equation (2) but the wavefunction  $\Psi$  of the particle is transformed according to

$$\Psi_0(\mathbf{r}, t) \rightarrow \Psi_\chi(\mathbf{r}, t) = \Psi_0(\mathbf{r}, t) \exp \{ie\chi(\mathbf{r}, t)/\hbar\}. \quad (33)$$

By doing this the Schrödinger equation

$$\{(-i\hbar\nabla - e\mathbf{A})^2/2m + e\phi\}\Psi(\mathbf{r}, t) = i\hbar(\partial/\partial t)\Psi(\mathbf{r}, t) \quad (34)$$

remains invariant in form under the gauge transformation (Stewart 1997a, 1997b).

The result of this transformation is that the operator  $\mathbf{m}^d$  becomes  $\mathbf{m}_\chi^d = -(\mathbf{r} - \mathbf{R}') \times (\mathbf{A} + \nabla\chi)e^2/2m$  but  $\mathbf{m}_\chi^p = \mathbf{m}^p$  remains unchanged since in the Schrödinger representation in which we work the operator  $\mathbf{p} = -i\hbar\nabla$  is independent of gauge. However, the wavefunction changes according to equation (33). When the effect of the operator  $\mathbf{p} = -i\hbar\nabla$  acting on the transformed wavefunction is allowed for, the matrix elements between particle states  $\Psi'$  and  $\Psi$  of  $\mathbf{m}^p$  and  $\mathbf{m}^d$  in the new gauge are found to be

$$\begin{aligned} \langle \Psi_\chi | \mathbf{m}_\chi^d | \Psi_\chi \rangle &= -(e^2/2m) \langle \Psi_0 | (\mathbf{r} - \mathbf{R}') \times \mathbf{A}(\mathbf{r}, \mathbf{0}) | \Psi_0 \rangle \\ &\quad - (e^2/2m) \langle \Psi_0 | (\mathbf{r} - \mathbf{R}') \times \nabla\chi | \Psi_0 \rangle \end{aligned} \quad (35a)$$

$$\begin{aligned} \langle \Psi_\chi | \mathbf{m}_\chi^p | \Psi_\chi \rangle &= (e/2m) \langle \Psi_0 | (\mathbf{r} - \mathbf{R}') \times \mathbf{p} | \Psi_0 \rangle \\ &\quad + (e^2/2m) \langle \Psi_0 | (\mathbf{r} - \mathbf{R}') \times \nabla\chi | \Psi_0 \rangle. \end{aligned} \quad (35b)$$

It is seen that under any gauge transformation the matrix elements of the paramagnetic and diamagnetic moments are changed by equal and opposite amounts. The sum of the two is independent of gauge. Since, as shown in Section 5, a change of origin of the vector potential of the generalised multipolar gauge is equivalent to making a gauge transformation we obtain the result that the matrix elements of the total orbital moment are independent of the origin of the vector potential even when the fields are time dependent and non-uniform. The same is true for the expectation value thereby proving Van Vleck's contention in the most general sense.

This achieves the aim of the paper, but it is possible to write down a further simple and useful expression for the diamagnetic moment. A preferred coordinate system may be obtained by choosing the origin of coordinates  $\mathbf{r}$  to be such that  $\langle \Psi_0 | \mathbf{A}(\mathbf{r}, \mathbf{0}) | \Psi_0 \rangle = \mathbf{0}$  or

$$\int \Psi_0^*(\mathbf{r}) \mathbf{A}(\mathbf{r}, \mathbf{0}) \Psi_0 d\mathbf{r} = \mathbf{0}. \quad (36)$$

For the uniform gauge this results in the centre of charge  $\langle \Psi_0 | \mathbf{r} | \Psi_0 \rangle$  being at the origin of coordinates (Stewart 1996a), but this is not necessarily the case when the fields are non-uniform. Because the drift velocity is zero it follows that in the gauge  $\mathbf{A}(\mathbf{r}, \mathbf{0})$  with this coordinate system  $\langle \Psi_0 | \mathbf{p} | \Psi_0 \rangle$  is zero too. Accordingly

$$\begin{aligned}\langle \Psi_0 | \mathbf{m}_0^p | \Psi_0 \rangle &= (e/2m) \langle \Psi_0 | \mathbf{r} \times \mathbf{p} | \Psi_0 \rangle, \\ \langle \Psi_0 | \mathbf{m}_0^d | \Psi_0 \rangle &= - (e^2/2m) \langle \Psi_0 | \mathbf{r} \times \mathbf{A}(\mathbf{r}, \mathbf{0}) | \Psi_0 \rangle,\end{aligned}$$

so in this particular coordinate system the paramagnetic and diamagnetic terms individually are independent of  $\mathbf{R}'$  the origin of the orbital angular momentum. For this case the diamagnetic moment in a non-uniform field may be expressed as an integral over the field

$$\langle \Psi_0 | \mathbf{m}_0^d | \Psi_0 \rangle = \frac{e^2}{2m} \langle \Psi_0 | \mathbf{r} \times \left\{ \mathbf{r} \times \int_0^1 u \, du \, \mathbf{B}(u\mathbf{r}) \right\} | \Psi_0 \rangle \quad (37)$$

When the magnetic field  $\mathbf{B}^0$  is uniform there is much simplification. If the quantities  $\mathbf{A}(\mathbf{r}, \mathbf{0}) = \mathbf{B}^0 \times \mathbf{r}/2$ , the gradient of the gauge function  $\nabla\chi = \mathbf{R} \times \mathbf{B}^0/2$  and  $\mathbf{R}_i = \langle \mathbf{r} \rangle$  the centre of charge of orbital  $i$  responsible for the magnetic moment are substituted into equations (35) we obtain

$$\langle \mathbf{m}_\chi^d \rangle = - (e^2/4m) \langle (\mathbf{r} - \mathbf{R}_i) \times \{ \mathbf{B}^0 \times (\mathbf{r} - \mathbf{R}_i) \} \rangle - (e^2/4m) (\mathbf{R}_i - \mathbf{R}') \times \{ \mathbf{B}^0 \times (\mathbf{R}_i - \mathbf{R}') \} \quad (38a)$$

$$\langle \mathbf{m}_\chi^p \rangle = (e/2m) \langle (\mathbf{r} - \mathbf{R}_i) \times \mathbf{p} \rangle + (e^2/4m) (\mathbf{R}_i - \mathbf{R}') \times \{ \mathbf{B}^0 \times (\mathbf{R}_i - \mathbf{R}') \}. \quad (38b)$$

If either  $\mathbf{R}$  or  $\mathbf{R}'$  are at  $\mathbf{R}_i$  the above equations reduce to the simpler forms

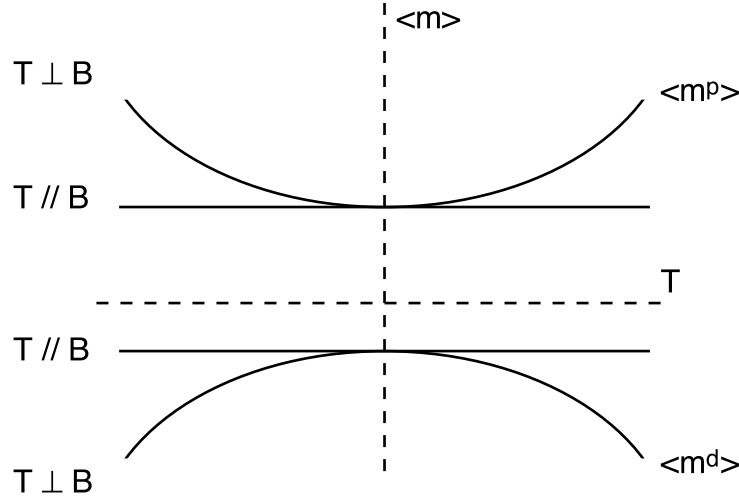
$$\begin{aligned}\langle \mathbf{m}_\chi^p \rangle &= \langle \mathbf{m}_0^p \rangle = (e/2m) \langle (\mathbf{r} - \mathbf{R}_i) \times \mathbf{p} \rangle, \\ \langle \mathbf{m}_\chi^d \rangle &= \langle \mathbf{m}_0^d \rangle = - (e^2/4m) \langle (\mathbf{r} - \mathbf{R}_i) \times \{ \mathbf{B}^0 \times (\mathbf{r} - \mathbf{R}_i) \} \rangle.\end{aligned}$$

This shows that the centre of charge is the natural origin for the paramagnetic and diamagnetic orbital moments and gives an answer to the question raised some time ago by Hameka (1965) as to the nature of the natural origin. If the point about which the angular momentum is taken is at the origin of the vector potential so  $\mathbf{R}' = \mathbf{R}$ , then with  $\mathbf{T} = \mathbf{R}_i - \mathbf{R}$  the expectation values of the moments become

$$\langle \mathbf{m}_\chi^p \rangle = \langle \mathbf{m}_0^p \rangle + (e^2/4m) \mathbf{T} \times (\mathbf{B}^0 \times \mathbf{T}) \quad (39a)$$

$$\langle \mathbf{m}_\chi^d \rangle = \langle \mathbf{m}_0^d \rangle - (e^2/4m) \mathbf{T} \times (\mathbf{B}^0 \times \mathbf{T}) \quad (39b)$$

To illustrate the meaning of the above two equations we consider a system consisting of a single orbital with the centre of charge at the origin so that  $\mathbf{R}_i = \langle \mathbf{r} \rangle = 0$ . When both the angular momentum and the vector potential are referred to the origin as well ( $\mathbf{R} = \mathbf{0}$  and  $\mathbf{R}' = \mathbf{0}$ ) the paramagnetic and diamagnetic moments are  $\langle \mathbf{m}_0^p \rangle$  and  $\langle \mathbf{m}_0^d \rangle$ . If the orbital is then physically displaced from the origin in a direction perpendicular to the field by a distance  $\mathbf{T}$  while  $\mathbf{R}$  and  $\mathbf{R}'$  remain unchanged (so  $\mathbf{T} = \mathbf{R}_i$ ) then, according to equations (39), the paramagnetic moment is increased by an amount  $(e^2/4m) \mathbf{T}^2 \mathbf{B}^0$  and the diamagnetic moment decreased by a similar amount, as shown in Fig. 3. If the displacement is parallel to the field there is no change. It is seen that when the centre of charge of the orbital is at



**Fig. 3.** Expectation values of the paramagnetic  $\langle m^p \rangle$  and diamagnetic moments  $\langle m^d \rangle$ , referred to the origin, of an atomic orbital in a static and uniform magnetic field as the orbital is displaced a distance  $T$  from the origin. The vertical dashed axis gives the magnetic moment, the horizontal axis the magnitude of the displacement  $T$ . The origin is where the horizontal and vertical dashed lines cross. If  $T$  is perpendicular to the field the paramagnetic and diamagnetic moments respectively increase and decrease quadratically with  $T$  according to equations (39). If  $T$  is parallel to the field the vector cross product in equations (39) is zero and so no change occurs. For the situation shown in the figure the paramagnetic moment is greater in magnitude than the diamagnetic moment giving a system that is overall paramagnetic.

the origin the paramagnetic and diamagnetic moments have their minimum and maximum values respectively.

## 8. Magnetism of Molecules

As a final illustration of the use of gauge arguments in condensed matter physics, we discuss the magnetism of molecules. The usual method of describing molecular structure is to build the molecule up from molecular orbitals that are spread over the whole molecule and are made up from linear combinations of atomic orbitals (LCAO) localised at the atomic sites. One electron occupies each molecular orbital

$$\Psi(\mathbf{r}) = \sum_{a,i} a_{a,i} \phi'_a(\mathbf{r} - \mathbf{R}_i) \quad (40)$$

The quantity  $\Psi(\mathbf{r})$  is the molecular orbital and  $\phi'_a(\mathbf{r} - \mathbf{R}_i)$  is an atomic orbital, whose nature is to be determined, localised at site  $i$  with quantum numbers  $a$ . The complex coefficients  $a_{a,i}$  are varied to make the energy of the orbital

$$E = \int \Psi^*(\mathbf{r}) H \Psi(\mathbf{r}) d\mathbf{r} / \int \Psi^*(\mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r} \quad (41)$$

a minimum. The molecule is then constructed by putting electrons into the molecular orbitals of lowest energy and calculating the exchange and correlation energy. The magnetic susceptibility is given by the second derivative of the total energy with respect to magnetic field.

The Hamiltonian for the molecular orbital is

$$H = \{p - eA(\mathbf{r}, \mathbf{R}_A)\}^2/2m + \sum_k V_k(\mathbf{r} - \mathbf{R}_k), \quad (42)$$

where  $\mathbf{R}_A$  is the gauge origin of the vector potential that describes the magnetic field and  $V_k$  is the atomic scalar potential localised at site  $k$ . The gauge origin must be the same for every atomic orbital as each molecular orbital consists of only one electronic state. We need to isolate the terms in equation (42) that involve site  $i$ . To do this we use equation (24) to shift the origin of the vector potential from  $\mathbf{R}_A$  to  $\mathbf{R}_i$ . This gives  $A(\mathbf{r}, \mathbf{R}_A) = A(\mathbf{r}, \mathbf{R}_i) + \nabla\chi_i$  where, for a uniform field  $\mathbf{B}^0$ ,  $\chi_i = \mathbf{r} \cdot (\mathbf{R}_A - \mathbf{R}_i) \times \mathbf{B}^0/2$ . Accordingly we have

$$H = \{p - eA(\mathbf{r}, \mathbf{R}_i) - e\nabla\chi_i\}^2/2m + V_i(\mathbf{r} - \mathbf{R}_i) + \sum_{k \neq i} V_k(\mathbf{r} - \mathbf{R}_k), \quad (43)$$

where the first two terms involve only site  $i$ .

In equation (41) we have to evaluate  $H$  acting on  $\phi'_a(\mathbf{r} - \mathbf{R}_i)$ . We make the substitution

$$\begin{aligned} \phi'_a(\mathbf{r} - \mathbf{R}_i) &= \phi_a(\mathbf{r} - \mathbf{R}_i) \exp(i e \chi_i / \hbar) \\ &= \phi_a(\mathbf{r} - \mathbf{R}_i) \exp\{i e \mathbf{r} \cdot (\mathbf{R}_A - \mathbf{R}_i) \times \mathbf{B}^0 / 2\hbar\} \end{aligned} \quad (44)$$

which defines London's (1937) gauge invariant atomic orbitals  $\phi_a(\mathbf{r} - \mathbf{R}_i)$ . By doing this we obtain

$$H\phi'_a(\mathbf{r} - \mathbf{R}_i) = \exp(i e \chi_i / \hbar) [\{p - eA(\mathbf{r}, \mathbf{R}_i)\}^2/2m + V_i(\mathbf{r} - \mathbf{R}_i) + \sum_{k \neq i} V_k(\mathbf{r} - \mathbf{R}_k)] \phi_a(\mathbf{r} - \mathbf{R}_i), \quad (45)$$

or

$$H\phi'_a(\mathbf{r} - \mathbf{R}_i) = \exp(i e \chi_i / \hbar) [\varepsilon_{a,i} + \sum_{k \neq i} V_k(\mathbf{r} - \mathbf{R}_k)] \phi_a(\mathbf{r} - \mathbf{R}_i), \quad (46)$$

where  $\varepsilon_{a,i}$  is the single particle energy of the orbital  $\phi_a(\mathbf{r} - \mathbf{R}_i)$  in the presence of the magnetic field given by

$$[\{p - eA(\mathbf{r}, \mathbf{R}_i)\}^2/2m + V_i(\mathbf{r} - \mathbf{R}_i)] \phi_a(\mathbf{r} - \mathbf{R}_i) = \varepsilon_{a,i} \phi_a(\mathbf{r} - \mathbf{R}_i).$$

The orbital  $\phi_a(\mathbf{r} - \mathbf{R}_i)$  is defined by this equation and is the eigenfunction of the Hamiltonian for site  $i$  with  $\varepsilon_{a,i}$  as its eigenvalue. It is to be noted that when the atomic orbital is other than an s-state orbital the energy  $\varepsilon_{a,i}$  will have a Zeeman component linearly proportional to the magnitude of the magnetic field  $\mathbf{B}^0$ .

The numerator of equation (41) is now expressed as

$$\begin{aligned} \sum_{b,j} \sum_{a,l} a_{b,j}^* a_{a,i} \int d\mathbf{r} \phi_b^*(\mathbf{r} - \mathbf{R}_j) \exp\{i e \mathbf{r} \cdot (\mathbf{R}_j - \mathbf{R}_i) \times \mathbf{B}^0 / 2\hbar\} \\ \times [\delta_{a,b} \delta_{i,j} \varepsilon_{a,i} + \sum_{k \neq i} V_k(\mathbf{r} - \mathbf{R}_k)] \phi_a(\mathbf{r} - \mathbf{R}_i) \end{aligned} \quad (47)$$

and it is seen that the gauge origin  $\mathbf{R}_A$  for the molecule has disappeared making the result independent of gauge. The same is true for the denominator

$$\sum_{b,j} \sum_{a,l} a_{b,j}^* a_{a,i} \int d\mathbf{r} \phi_b^*(\mathbf{r} - \mathbf{R}_j) \exp\{i e \mathbf{r} \cdot (\mathbf{R}_j - \mathbf{R}_i) \times \mathbf{B}^0 / 2\hbar\} \phi_a(\mathbf{r} - \mathbf{R}_i). \quad (48)$$



It appears that the equations have the same structure as those of a non-magnetic variational calculation, except that the orbitals are replaced by London's gauge invariant orbitals  $\phi_a(\mathbf{r} - \mathbf{R}_i)\exp(-i\mathbf{e}\mathbf{r}\cdot\mathbf{R}_i \times \mathbf{B}^0/2\hbar)$ , where the  $\phi_a$  are the atomic eigenfunctions in the presence of the field and the eigenvalues  $\varepsilon_{a,i}$  may in general depend on field. Further approximations needed to compute the integrals in the above two equations in a practical calculation were discussed by Pople (1962).

## References

- Babiker, M., and Loudon, R. (1983). *Proc. R. Soc. London A* **385**, 439.
- Brittin, W. E., Smythe, W. R., and Wyss, W. (1982). *Am. J. Phys.* **50**, 693.
- Cohen-Tannoudji, C., Dupont-Roc, J., and Gilbert, G. (1989). 'Photons and Atoms' (Wiley: New York).
- Craig, D. P., and Thirunamachandran, T. (1998). 'Molecular Quantum Mechanics' (Dover: New York).
- Dirac, P. A. M. (1947). 'Quantum Mechanics' (Clarendon Press: Oxford).
- Doughty, N. A. (1980). 'Lagrangian Interaction' (Addison-Wesley: Sydney).
- Feynman, R. P., Leighton, R. B., and Sands, M. (1964). 'The Feynman Lectures on Physics' (Addison-Wesley: Reading).
- Friar, J. L., and Fallieros, S. (1981). *Am. J. Phys.* **49**, 847.
- Frohlich, J., and Studer, U. M. (1993). *Rev. Mod. Phys.* **65**, 733.
- Goppert-Mayer, M. (1931). *Ann. der Phys. (Leipzig)* **9**, 273.
- Griffith, J. S. (1961). 'Theory of Transition Metal Ions' (Cambridge Univ. Press).
- Hameka, H. F. (1965). 'Advanced Quantum Chemistry' (Addison-Wesley: Reading, Mass.).
- Kobe, D. H. (1982). *Am. J. Phys.* **50**, 128.
- London, F. (1937). *J. Phys. Radium (Paris)* **8**, 387.
- Panofsky, W. K. H., and Phillips, M. (1955). 'Classical Electricity and Magnetism' (Addison-Wesley: Cambridge, Mass.).
- Pople, J. A. (1962). *J. Chem. Phys.* **37**, 53.
- Power, E. A., and Zienau, S. (1959). *Phil. Trans. R. Soc. London* **251**, 427.
- Skagerstam, B. K. (1983). *Am. J. Phys.* **52**, 1148.
- Stewart, A. M. (1996a). *Aust. J. Phys.* **49**, 683.
- Stewart, A. M. (1996b). *J. Phys. A* **29**, 1411.
- Stewart, A. M. (1997a). *Aust. J. Phys.* **50**, 869.
- Stewart, A. M. (1997b). *Aust. J. Phys.* **50**, 1061.
- Stewart, A. M. (1999). *J. Phys. A* **32**, 6091.
- Valatin, J. G. (1954). *Proc. R. Soc. London A* **222**, 93.
- Van Vleck, J. H. (1932). 'The Theory of Electric and Magnetic Susceptibilities' (Oxford Univ. Press).
- Woolley, R. G. (1973). *J. Phys. B* **6**, L97.
- Woolley, R. G. (1974). *J. Phys. B* **7**, 488.

## Appendix

Here we derive the relation  $(\mathbf{r} - \mathbf{R})\cdot\nabla h(\mathbf{q}) = u(\partial h/\partial u)$  and other useful relations between the derivatives of scalar functions  $h(\mathbf{q})$  of the vector  $\mathbf{q}$ , where  $\nabla$  is the gradient operator with respect to  $\mathbf{r}$  and  $\mathbf{q}$  is given by equation (17). The relations apply also to the individual components of a vector.

Noting from equation (17) that

$$\partial q^i / \partial r^j = u \delta_{ij}, \quad \partial q^i / \partial R^j = -u \delta_{ij}, \quad \partial q^i / \partial u = (r^i - R^i),$$

we find that

$$\partial h / \partial r^i = u(\partial h / \partial q^i), \quad \partial h / \partial R^i = -u(\partial h / \partial q^i), \quad \partial h / \partial u = \Sigma_i (r^i - R^i)(\partial h / \partial q^i).$$

Accordingly we get

$$\nabla h(\mathbf{q}) = u \nabla_{\mathbf{q}} h(\mathbf{q}), \quad \nabla_{\mathbf{R}} h(\mathbf{q}) = -u \nabla_{\mathbf{q}} h(\mathbf{q}), \quad \partial h / \partial u = (\mathbf{r} - \mathbf{R}) \cdot \nabla_{\mathbf{q}} h(\mathbf{q}),$$

where the subscript shows the vector with respect to which the gradient is taken when it is not the field point  $\mathbf{r}$ . The result at the beginning of this Appendix follows from multiplying both sides of the last relation by  $u$  and using the relation between  $\nabla$  and  $\nabla_{\mathbf{q}}$ .

Manuscript received 17 December 1999, accepted 15 March 2000