

Home Search Collections Journals About Contact us My IOPscience

General proof of gauge independence of the sum of the dia- and paramagnetic magnetic moments

This article has been downloaded from IOPscience. Please scroll down to see the full text article. 1999 J. Phys. A: Math. Gen. 32 6091 (http://iopscience.iop.org/0305-4470/32/33/307) View the table of contents for this issue, or go to the journal homepage for more

Download details: IP Address: 171.66.16.105 The article was downloaded on 02/06/2010 at 07:39

Please note that terms and conditions apply.

General proof of gauge independence of the sum of the diaand paramagnetic magnetic moments

A M Stewart

Department of Applied Mathematics, Research School of Physical Sciences and Engineering, The Australian National University, Canberra ACT 0200, Australia

E-mail: andrew.stewart@anu.edu.au

Received 6 April 1999, in final form 9 June 1999

Abstract. By making use of the generalized multipolar gauge it is proved that the matrix elements of the magnetic moment operator are independent of gauge origin for electromagnetic fields that are non-uniform in space and non-constant in time.

1. Introduction

In 1932 Van Vleck asserted, but did not prove fully, that the sum of the orbital paramagnetic and diamagnetic susceptibilities of an atom or molecule was independent of the origin of the vector potential used to describe the magnetic field. This contention, in its most general form to be discussed, might be called the fundamental theorem of magnetism since if it were not true the theory of magnetism, and much else, would be untenable.

Van Vleck took the diamagnetic susceptibility to be proportional to $\langle r^2 \rangle$ and the paramagnetic term to $\langle r \mathbf{x} p \rangle$ and stated that when quantum conditions and commutation rules were used the susceptibility was found to be independent of gauge origin, but he did not provide a complete derivation of this. Of the proofs of Van Vleck's contention that have been offered since, Griffith (1961) claimed that in the Coulomb gauge 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. Friar and Fallieros (1981) derived the same result for the uniform gauge (see later) using perturbation theory and showed that the total susceptibility was gauge invariant for spherical systems when the magnetic moment operator was taken to be $r \ge (p - eA)e/2m$. 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. Geersten (1989) showed that the total susceptibility was independent of gauge origin using the theory of the polarization propagator but that this could be violated in numerical calculations if a restricted set of basis states was used. All these derivations only applied to the linear response (the susceptibility) and relied on perturbation theory or had other limitations. Recently (Stewart 1996a, 1997) a proof was given that applies to the magnetic moment itself (i.e. the full nonlinear response to a magnetic field) and essentially involves operators alone and is independent of any particular set of basis states.

However, all the proofs mentioned above used the symmetric uniform gauge with vector potential $A = B_0 \times r/2$, where B_0 is a vector that does not depend on the space r and time

6092 A M Stewart

t coordinates. The scalar potential ϕ is independent of time. From the relations that give the fields in terms of the potentials

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

where ∇ is the gradient operator with respect to r, it follows that in this gauge $B(r, t) = B_0$ and $E(r, t) = -\nabla \phi(r)$, both fields being independent of time and B being uniform in space.

The purpose of this paper is to extend the proof of Van Vleck's contention 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. In section 2 it is shown that in the generalized multipole gauge, as in the uniform gauge, a shift of the origin of the potentials corresponds to a gauge transformation. In section 3 it is shown how this leads to the general proof of Van Vleck's contention and in the appendix an algebraic result needed for the analysis of the generalized multipolar gauge is obtained.

2. Generalized multipolar gauge

The generalized multipolar gauge is obtained by making the gauge transformation

$$A(\mathbf{r},t) = A'(\mathbf{r},t) + \nabla \Lambda(\mathbf{r},t)$$
 and $\phi(\mathbf{r},t) = \phi'(\mathbf{r},t) - \partial \Lambda/\partial t$ (2)

from any gauge A' and ϕ' to the new gauge A and ϕ using the gauge function

$$\Lambda(\mathbf{r},t) = \Lambda(\mathbf{R},t) - \int_0^1 \mathrm{d}u \,(\mathbf{r}-\mathbf{R}) \cdot \mathbf{A}'(\mathbf{q}(u),t) \tag{3}$$

where

$$q = ur + (1 - u)R \tag{4}$$

with

$$\partial \Lambda(\mathbf{R}, t) / \partial t = \phi'(\mathbf{R}, t)$$
 (5)

the integral being carried out along the line AB in figure 1. This gauge function is obtained by requiring A to be perpendicular to the line (r - R) at every point on that line which implies

$$\int_0^1 \mathrm{d}u \,(\boldsymbol{r} - \boldsymbol{R}) \cdot \boldsymbol{A}(\boldsymbol{q}) = 0. \tag{6}$$

Equation (3) is obtained from equation (6) noting that $\nabla_q \Lambda(q) = \nabla \Lambda(q)/u$, where ∇_q is the gradient with respect to q and ∇ is the gradient with respect to r.

The four-vector version of the multipolar gauge was introduced by Valatin (1954) and its three-vector version, which we use here, was suggested later by Woolley (1973, 1974). Skagerstam (1983) has discussed the relation between the three- and four-vector versions. An analysis of the three-vector version was given by Kobe (1982) for the $\mathbf{R} = \mathbf{0}$ form only. We sketch the derivation of the three-vector version for non-zero \mathbf{R} as it is needed for what follows.

Following the methods of Kobe (1982) we first obtain A(r) by taking the gradient of the gauge function in equation (3). Henceforth we omit the time argument *t*, as it is the same in all the potentials and fields. There are four terms arising from $\nabla \Lambda$. One of them vanishes from $\nabla \mathbf{x} \mathbf{r} = \mathbf{0}$. Two others, using the result $(r - R) \cdot \nabla f(q) = u(\partial f/\partial u)$ derived in the appendix, come to $-\partial/\partial u(uA')$ whose integral cancels A' in equation (2). The remaining term, using the relation $\nabla \mathbf{x} \mathbf{A}(q) = uB(q)$, comes to

$$A(\mathbf{r}, \mathbf{R}) = -(\mathbf{r} - \mathbf{R}) \times \int_0^1 u \, \mathrm{d}u \, B(\mathbf{q}).$$
⁽⁷⁾

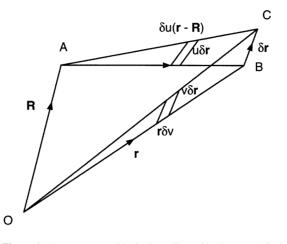


Figure 1. Vectors R, r and δr in three-dimensional space. *O* is the coordinate origin, δr is not in general coplanar with r and R. When r is changed by δr the area of triangle *OAB* is increased by triangle *ABC* and decreased by triangle *OBC*. The dimensions of the area elements are shown on the diagram. The position of the area element in triangle *ABC* is q, given by equation (4). The area element in triangle *OBC* is at vr where $1 \ge v \ge 0$.

A(r, R) has the property that at every point r it is perpendicular to the vector (r - R).

The scalar potential is obtained from equation (2) by noting that $\partial A'(q)/\partial t = -E(q) - \nabla \phi'(q)/u$. Using the result in the appendix, the second term gives rise to $-\partial/\partial u \{\phi'(q)\}$ whose integral cancels ϕ' , giving, with equation (5), the scalar potential

$$\phi(\mathbf{r}, \mathbf{R}) = -(\mathbf{r} - \mathbf{R}) \cdot \int_0^1 \mathrm{d}u \, \mathbf{E}(\mathbf{q}) \tag{8}$$

so all the components of the potential are zero at r = R.

Equations (7) and (8) constitute the generalized multipolar gauge. Next we verify that operating on the potentials (7) and (8) with equations (1) recovers the fields B(r, t) and E(r, t). The curl of equation (7) has four terms. One is zero from $\nabla \cdot B = 0$. The other three come to $\partial/\partial u \{u^2 B(q)\}$ whose integral is B(r) as required. The electric field E(r) is obtained from equations (1), (7) and (8) by noting that $\partial B(q)/\partial t = -\nabla \mathbf{x} E(q)/u$. The remaining terms of the integrand come to $\partial/\partial u \{u E(q)\}$ whose integral is E(r). Neither E(r) nor B(r) involve R, so its presence in A and ϕ indicates that a gauge freedom associated with R exists in these potentials. If the B field happens to be uniform in space then the integral over u may be performed trivially and the symmetric uniform gauge is obtained with $A = B_0 \mathbf{x} (r - R)/2$. In the symmetric uniform gauge R corresponds to a change of the origin of the potentials and is associated with a gauge transformation with gauge function $-r \cdot (B_0 \mathbf{x} R)/2$, and R will be found to play a similar role in the generalized multipolar gauge.

The multipolar gauge was given this name by 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 are evaluated at the point \mathbf{R} , then the integrals over u may be performed trivially and the potentials may be expressed, expanding as far as the quadrupole terms, as

$$\phi = -(\mathbf{r} - \mathbf{R}) \cdot \mathbf{E}(\mathbf{R}) - \Sigma_{ij} (x^i - X^i) (x^j - X^j) (\partial E^i / \partial x^j) / 2 + \cdots$$
(9)

$$A^{k}(\mathbf{r}) = -(\mathbf{r} - \mathbf{R}) \mathbf{x} B(\mathbf{R})|^{k}/2 - \sum_{ijl} \varepsilon_{ijk} (x^{i} - X^{i})(x^{1} - X^{1})(\partial B^{j}/\partial x^{1})/3 + \cdots$$
(10)

where ε_{ijk} is the antisymmetric unit tensor and the derivatives are evaluated at R.

Although the simple (with R = 0) and generalized multipolar gauges do not in general satisfy the Coulomb gauge condition $\nabla \cdot A = 0$, their multipole expansions up to but not

6094 A M Stewart

including the magnetic quadrupole term do, which makes them convenient for use in many atomic spectroscopy calculations.

Can the gauge freedom implied by R be expressed in the form of a gauge transformation

$$A(r, R) = A(r, 0) + \nabla \chi(r, R)$$
(11)

where χ is a gauge function and if so what is it? To answer this question we consider the magnetic flux $\Phi(r, R)$ that is linked with the triangle *OAB* of figure 1 formed by the vectors r and R. Let r vary by a small vector δr ; this latter vector is not necessarily coplanar with R and r. As a result the flux linked with the triangle changes by an amount $\delta \Phi$. This quantity is equal to the flux linked with the triangle *ABC* minus that linked with the triangle *OBC*.

First we calculate the flux linked with *ABC*. The vector element of the area at point q on line *AB* is $u\delta r \mathbf{x} (r - R)\delta u$ and the flux $\delta \Phi_1$ linking this area is $B(q) \cdot \delta r \mathbf{x} (r - R)u\delta u$ or $(r - R) \mathbf{x} B(q) \cdot \delta r u\delta u$. The integral of this over u is seen from equation (7) to be the scalar product of the vector potential of equation (7) with δr , namely $\delta \Phi_1 = A(r, R) \cdot \delta r$. From this must be subtracted the flux $\delta \Phi_2$ linking the triangle *OBC*. The area element here is $-r\mathbf{x}\delta r v\delta v$ which leads to a flux $\delta \Phi_2 = A(r, \mathbf{0}) \cdot \delta r$ so $\delta \Phi = \delta \Phi_1 + \delta \Phi_2 = \delta r \cdot \{A(r, R) - A(r, \mathbf{0})\}$. Accordingly

$$A(r, R) - A(r, 0) = \nabla \Phi(r, R)$$
(12)

so $\chi(\mathbf{r}, \mathbf{R}) = \Phi(\mathbf{r}, \mathbf{R})$ and we find that the gauge function for this transformation is equal to the magnetic flux linked with triangle *OAB*. This is consistent with the uniform gauge where the corresponding gauge function is $-\mathbf{r} \cdot (\mathbf{B}_0 \times \mathbf{R})/2$ or $\mathbf{B}_0 \cdot (\mathbf{r} \times \mathbf{R})/2$, the scalar product of the uniform \mathbf{B} field and the area. We have therefore shown that changing the origin \mathbf{R} of the potentials in the generalized multipolar gauge, just as in the uniform gauge, amounts to making a gauge transformation. It is to be noted that the arguments given above apply as much to time-dependent fields and potentials as to time-independent ones.

3. Gauge invariance of the magnetic moment

The classical expression for the orbital magnetic moment m about the point \mathbf{R}' of a particle of charge e at position r is $m = (r - \mathbf{R}') \times ve/2$. Because we deal with a physical system of finite extent, such as an atom or molecule the drift velocity, the expectation value of v, is zero and consequently the expectation value of m is independent of \mathbf{R}' . The quantum mechanical operator identified with the particle velocity is $v = dr/dt = [r, H]/i\hbar$ where H is the Hamiltonian. By commuting r with the non-relativistic Hamiltonian $H = (p - eA)^2/2m + e\phi$ the velocity operator v is given by mv = p - eA. If other terms involving p are present in the Hamiltonian the commutator of them with r will contribute further to the velocity operator. For example, the spin-orbit interaction $s \times E \cdot (p - eA)e\hbar/4m^2c$ (Frohlich and Studer 1993) will add a term $s \times Ee\hbar/4m^2c$ to v and so add a manifestly gauge invariant term $[s\{(r - R') \cdot E\} - E\{(r - R') \cdot s\}]e^2\hbar/8m^2c$ to the magnetic moment. If the Dirac Hamiltonian is used then $v = c\alpha$, where α is the Dirac operator and gauge is not involved at all. The operator for the orbital moment in the non-relativistic case is therefore composed of two terms $m = m^p + m^d$, where the paramagnetic moment is $m^p = (r - \mathbf{R}') \times pe/2m$ and the diamagnetic moment is $m^d = -(r - \mathbf{R}') \times Ae^2/2m$.

Now let us 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 Ψ of the particle is transformed according to

$$\Psi_0(\mathbf{r},t) \to \Psi_{\chi}(\mathbf{r},t) = \Psi_0(\mathbf{r},t) \exp\{ie\chi(\mathbf{r},t)/\hbar\}.$$
(13)

By doing this the Schrödinger equation

$$[(\boldsymbol{p} - \boldsymbol{e}\boldsymbol{A})^2/2\boldsymbol{m} + \boldsymbol{e}\boldsymbol{\phi}]\Psi(\boldsymbol{r}, t) = \mathrm{i}\hbar(\partial/\partial t)\Psi(\boldsymbol{r}, t)$$
(14)

remains invariant in form under the gauge transformation (Stewart 1996b, 1997).

Under this transformation the operator m^d becomes $m_{\chi}^d = -(r - R') \mathbf{x} (A + \nabla \chi) e^2/4m$ and $m_{\chi}^p = m^p$ since in the Schrödinger representation the operator $p = -i\hbar\nabla$ is independent of gauge. However, the wavefunction changes according to equation (13). When the effect of the operator $p = -i\hbar\nabla$ acting on the transformed wavefunction is allowed for, the matrix elements between particle states Ψ' and Ψ of m^p and m^d in the new gauge are found to be

$$\langle \Psi'_{\chi} | \boldsymbol{m}^{d}_{\chi} | \Psi_{\chi} \rangle = -(e^{2}/2m) \langle \Psi'_{0} | (\boldsymbol{r} - \boldsymbol{R}') \times \boldsymbol{A}(\boldsymbol{r}, \boldsymbol{0}) | \Psi_{0} \rangle - (e^{2}/2m) \langle \Psi'_{0} | (\boldsymbol{r} - \boldsymbol{R}') \times \nabla \chi | \Psi_{0} \rangle$$
(15a)

$$\langle \Psi_{\chi}' | \boldsymbol{m}_{\chi}^{p} | \Psi_{\chi} \rangle = (e/2m) \langle \Psi_{0}' | (\boldsymbol{r} - \boldsymbol{R}') \mathbf{x} \boldsymbol{p} | \Psi_{0} \rangle + (e^{2}/2m) \langle \Psi_{0}' | (\boldsymbol{r} - \boldsymbol{R}') \mathbf{x} \nabla \chi | \Psi_{0} \rangle.$$
(15b)

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 2, a change of origin of the vector potential of the generalized multipolar gauge is equivalent to making a gauge transformation it is thereby proved 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 its fullest generality.

Although this achieves the object of the paper some simplification is still possible. A preferred coordinate system may be obtained by choosing the origin of coordinates r to be such that $\langle \Psi_0 | A(r, 0) | \Psi_0 \rangle = 0$ or

$$\int \Psi_0^*(\boldsymbol{r}) \boldsymbol{A}(\boldsymbol{r}, \boldsymbol{0}) \Psi_0 \, \mathrm{d}\boldsymbol{r} = \boldsymbol{0}.$$
(16)

For the uniform gauge this results in the centre of charge $\langle \Psi_0 | \mathbf{r} | \Psi_0 \rangle$ being at the origin (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 $\langle \Psi_0 | \mathbf{m}_0^p | \Psi_0 \rangle = (e/2m) \langle \Psi_0 | \mathbf{r} \mathbf{x} \mathbf{p} | \Psi_0 \rangle$ and $\langle \Psi_0 | \mathbf{m}_0^d | \Psi_0 \rangle = -(e^2/2m) \langle \Psi_0 | \mathbf{r} \mathbf{x} \mathbf{A}(\mathbf{r}, \mathbf{0}) | \Psi_0 \rangle$ so in this coordinate system the paramagnetic and diamagnetic terms individually are independent of \mathbf{R}' the origin of the orbital angular momentum. In this case the diamagnetic moment in a non-uniform field may be expressed explicitly as

$$\langle \Psi_0 | \boldsymbol{m}_0^d | \Psi_0 \rangle = \frac{e^2}{2m} \langle \Psi_0 | \boldsymbol{r} \, \mathbf{x} \left\{ \boldsymbol{r} \, \mathbf{x} \int_0^1 u \, \mathrm{d}u \, \boldsymbol{B}(\boldsymbol{u}\boldsymbol{r}) \right\} | \Psi_0 \rangle. \tag{17}$$

Appendix

We wish to show that $(r - R) \cdot \nabla f(q) = u(\partial f/\partial u)$ where ∇ is the gradient operator with respect to r, q is given by equation (4) and f is any function of the vector q.

Noting from equation (4) that $\partial q^i / \partial x^j = u \delta_{ij}$, $\partial q^i / \partial X^j = -u \delta_{ij}$ and $\partial q^i / \partial u = (x^i - X^i)$ we find that $\partial f / \partial x^i = u (\partial f / \partial q^i)$, $\partial f / \partial X^i = -u (\partial f / \partial q^i)$ and $\partial f / \partial u = \sum_i (x^i - X^i) (\partial f / \partial q^i)$. Accordingly $u \partial f (q^j) / \partial u = \sum_i u (x^i - X^i) (\partial f / \partial q^i)$. Next $(r - R) \cdot \nabla f(q) = \sum_i (x^i - X^i) (\partial f / \partial q^i)$ and the result is proved.

References

Friar J L and Fallieros S 1981 *Am. J. Phys.* **49** 847–9 Frohlich J and Studer U M 1993 *Rev. Mod. Phys.* **65** 733–802 6095

6096 A M Stewart

Geersten J 1989 J. Chem. Phys. 90 4892-4

- Griffith J S 1961 Theory of Transition Metal Ions (Cambridge: Cambridge University Press) pp 434-6
- Kobe D H 1982 Am. J. Phys. 50 128–33
- Skagerstam B K 1983 Am. J. Phys. **51** 1148–9 Stewart A M 1996a Aust. J. Phys. **49** 683–7

- Valatin J G 1954 Proc. R. Soc. A 222 93–108
- Van Vleck J H 1932 The Theory of Electric and Magnetic Susceptibilities (Oxford: Oxford University Press) p 276
- Woolley R G 1973 J. Phys. B: At. Mol. Phys. 6 L97-9