

Magnetic anisotropy in density-functional theory

H. J. F. Jansen

Department of Physics, Oregon State University, Corvallis, Oregon 97331

(Received 18 April 1988)

Magnetic anisotropy is discussed in terms of relativistic density-functional theory. Spin-orbit coupling in the kinetic energy operator of the Dirac equation is responsible for a nonzero value of the orbital angular momentum in cubic systems. Additional spin-orbit coupling terms originate from the Breit interaction between the electrons. The Hartree approximation to the Breit energy yields a magnetic dipole-dipole interaction, which is the cause of the magnetic shape anisotropy. All these terms have to be included in a discussion of the magnetic anisotropy of cubic systems.

I. INTRODUCTION

Research to explain the origin of the magnetic anisotropy has a long history.¹ Early models are based on an effective quadrupole-quadrupole coupling of the spin fields and an excellent review is given by Van Vleck.² His paper is a good source for references to work performed before 1937. The starting point of the discussions is the relativistic two-electron Hamiltonian in the Pauli approximation, including terms connected to the Breit interaction.³ As a result one derives a coupling between the spin and orbital angular momenta, which in its turn gives rise to effective dipole-dipole and quadrupole-quadrupole forces between local spins. Van Vleck's opinion about the origin of the magnetic anisotropy is stated in his paper:² "We maintain that the coupling between orbital angular momentum vectors of different atoms is a vital ingredient, rather than the interaction of a single vector with the crystalline field."

The first calculations within the itinerant-electron model have been performed by Brooks⁴ in 1940. Here one obtains the orbital angular momentum from the single-particle states. A nonzero value of the orbital angular momentum is possible even in a cubic crystal because of the spin-orbit coupling terms in the kinetic energy operator in the Dirac equation. In subsequent papers⁵ the calculations are more and more refined, and finally the discussion centers around the importance of degenerate states along symmetry lines in the Brillouin zone.⁶ In all of these papers, however, the effect of the Breit interaction is completely ignored.

Recent fully relativistic calculations within local density-functional theory have attempted to give a final answer for the magnetic anisotropy in iron and nickel⁷ and give numbers of the right order of magnitude, but with the wrong sign for iron. There are serious problems, however, with the numerical integrations in \mathbf{k} space and it is well known⁸ that obtaining fully converged results is very difficult. Additionally, there are still some theoretical inconsistencies. In the present paper we give a formal basis for a discussion of the magnetic anisotropy within the framework of density-functional theory. In later papers we will address the relative importance, for different materials, of the terms which could determine the size of the anisotropy.

II. RELATIVISTIC DENSITY-FUNCTIONAL THEORY

Most presentations of density-functional theory are based on a nonrelativistic description of the electrons. The relativistic formulation, however, is a straightforward application of quantum electron dynamics,^{9,10} and has been presented before.¹¹ In this section we will simply state the main results in order to define a frame of reference for our subsequent discussions. We follow the approach of Lieb.¹² The basis of our description of the relativistic many-electron problem is the Hamiltonian density (in second quantization):

$$\mathcal{H}(\mathbf{r}) = \psi^\dagger(\mathbf{r}) \left[\frac{\hbar c}{i} \boldsymbol{\alpha} \nabla + (\beta - 1) mc^2 - e\phi(\mathbf{r}) + e\boldsymbol{\alpha} \mathbf{A}(\mathbf{r}) \right] \times \psi(\mathbf{r}) + \mathcal{H}_{\text{rad}}(\mathbf{r}). \quad (1)$$

We use the notation of Ref. 9 and we define e to be positive. The energy is shifted by subtracting the mass energy of the excess number of electrons in our system. $\mathcal{H}_{\text{rad}}(\mathbf{r})$ denotes the Hamiltonian density of the quantized electromagnetic field and $\phi(\mathbf{r})$ and $\mathbf{A}(\mathbf{r})$ are the scalar and vector potential due to these fields. We assume that these fields are time independent, in order to avoid all problems when energy is not conserved. Since our aim is to obtain the ground-state energy, this restriction is unimportant.

The electromagnetic field in Eq. (1) consists of two parts. First we have the fields generated by the electrons and second, there are the external fields due to the nuclei, external magnets, external electric fields, etc. At this point we make the assumption that the external fields can be described classically, and that quantization of the electromagnetic field is only important for the internal fields related to the electron-electron interaction. The energy expectation value for a state $|a, b\rangle$, where a pertains to the electrons and b to the photons, is then given by

$$E[a, b] = \langle a, b | \int d^3r \mathcal{H}(\mathbf{r}) | a, b \rangle - e \int d^3r \phi_{\text{ext}}(\mathbf{r}) \langle a, b | : \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) : | a, b \rangle + e \int d^3r \mathbf{A}_{\text{ext}}(\mathbf{r}) \langle a, b | : \psi^\dagger(\mathbf{r}) \boldsymbol{\alpha} \psi(\mathbf{r}) : | a, b \rangle. \quad (2)$$

Hence the external fields couple only to the charge and current densities, which are independent of the state of the photons,

$$\rho_a(\mathbf{r}) = -e \langle a, b | : \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) : | a, b \rangle, \quad (3a)$$

$$\mathbf{j}_a(\mathbf{r}) = -ce \langle a, b | : \psi^\dagger(\mathbf{r}) \boldsymbol{\alpha} \psi(\mathbf{r}) : | a, b \rangle. \quad (3b)$$

The first term of Eq. (2) contains all the information pertaining to the electronic system itself. As stated by Lieb,¹² there is a natural separation in two steps of the minimization of the energy. First we consider all states $| a, b \rangle$ which give rise to the charge and current densities $\rho_a(\mathbf{r})$ and $\mathbf{j}_a(\mathbf{r})$. We define the functional F by

$$F[\rho, \mathbf{j}] = \min \langle a, b | \int d^3r \mathcal{H}(\mathbf{r}) | a, b \rangle, \quad (4)$$

where the minimum is only over states $| a, b \rangle$ consistent with (3). The fact that quantum electrodynamics contains positrons as well as electrons does not have to worry us. The normal ordering in Eq. (1) takes care of the nonphysical effects of the negative energy states appearing in first quantization.^{9,10} Of course, when the nuclear potential is strong enough ($Z > 137$), spontaneous creation of electron-positron pairs is possible. For densities corresponding to these cases, states with a nonzero expectation value of positrons will play an important role in the minimum of Eq. (4). This will lead to strong effects on the exchange-correlation potentials. These corrections have been described before,¹¹ but are not well understood. One can only hope, although intuition would certainly confirm this, that they are not important in applications to 3d transition metals.

The last step in finding the ground-state energy, charge density, and current density of the interacting electron system is a minimization of the energy with respect to the charge and current density. One defines¹² the kinetic energy of a noninteracting electron system by

$$K[\rho, \mathbf{j}] = \min \langle a, b | \int d^3r : \psi^\dagger(\mathbf{r}) \left[\frac{\hbar c}{i} \boldsymbol{\alpha} \nabla + (\beta - 1) mc^2 \right] \times \psi(\mathbf{r}) : | a, b \rangle \quad (5)$$

and writes $F = K + G$. The minimization is again only over states consistent with (3). Hence the functional of $\rho(\mathbf{r})$ and $\mathbf{j}(\mathbf{r})$ one has to minimize is given by

$$E[\rho, \mathbf{j}] = K[\rho, \mathbf{j}] + G[\rho, \mathbf{j}] + \int d^3r \rho(\mathbf{r}) \phi_{\text{ext}}(\mathbf{r}) - \frac{1}{c} \int d^3r \mathbf{j}(\mathbf{r}) \mathbf{A}_{\text{ext}}(\mathbf{r}). \quad (6)$$

III. MINIMIZATION PROCEDURE

Before one can discuss a minimization procedure for (6), one has to be able to evaluate expression (5). At this point we assume that we have a satisfactory approximation for the functional G and focus on the noninteracting kinetic energy. Suppose we want to evaluate K at ρ_0, \mathbf{j}_0 . In a small range around this point G can be expanded as

$$\tilde{G}[\rho, \mathbf{j}] = G[\rho_0, \mathbf{j}_0] + \int d^3r Z(\mathbf{r}) [\rho(\mathbf{r}) - \rho_0(\mathbf{r})] - \frac{1}{c} \int d^3r \mathbf{Y}(\mathbf{r}) [\mathbf{j}(\mathbf{r}) - \mathbf{j}_0(\mathbf{r})] \quad (7)$$

and therefore we first try to solve for the minimum of

$$\tilde{E}[\rho, \mathbf{j}] = K[\rho, \mathbf{j}] + \tilde{G}[\rho, \mathbf{j}] + \int d^3r \rho(\mathbf{r}) \phi_{\text{ext}}(\mathbf{r}) - \frac{1}{c} \int d^3r \mathbf{j}(\mathbf{r}) \mathbf{A}_{\text{ext}}(\mathbf{r}). \quad (8)$$

But this is a noninteracting problem because all terms apart from K are linear in ρ and \mathbf{j} . Hence minimizing (8) is equivalent to minimizing

$$\langle a | \int d^3r : \psi^\dagger(\mathbf{r}) \left[\frac{\hbar c}{i} \boldsymbol{\alpha} \nabla + (\beta - 1) mc^2 - e[Z(\mathbf{r}) + \phi_{\text{ext}}(\mathbf{r})] + e\boldsymbol{\alpha}[\mathbf{Y}(\mathbf{r}) + \mathbf{A}_{\text{ext}}(\mathbf{r})] \right] \psi(\mathbf{r}) : | a \rangle + \tilde{G}[\rho_0, \mathbf{j}_0]. \quad (9)$$

The minimum is given when the wave function corresponds to a Slater determinant, where the single-particle wave functions obey the standard Kohn-Sham equations

$$\left[\frac{\hbar c}{i} \boldsymbol{\alpha} \nabla + (\beta - 1) mc^2 - e[Z(\mathbf{r}) + \phi_{\text{ext}}(\mathbf{r})] + e\boldsymbol{\alpha}[\mathbf{Y}(\mathbf{r}) + \mathbf{A}_{\text{ext}}(\mathbf{r})] \right] \phi_n(\mathbf{r}) = \epsilon_n \phi_n(\mathbf{r}). \quad (10)$$

The values of the density and current which minimize (10) are found by summing over the lowest energy states

$$\rho_m(\mathbf{r}) = -e \sum_{\text{occ}} \phi_n^\dagger(\mathbf{r}) \phi_n(\mathbf{r}), \quad (11a)$$

$$\mathbf{j}_m(\mathbf{r}) = -ce \sum_{\text{occ}} \phi_n^\dagger(\mathbf{r}) \boldsymbol{\alpha} \phi_n(\mathbf{r}). \quad (11b)$$

It is directly clear that because of the definition (5) of the

kinetic energy one has to use the ground state of the noninteracting problem and that one is not allowed to leave holes below the noninteracting Fermi level. The latter would change the definition of K and hence of G (and thus the exchange-correlation potential). The value of the noninteracting kinetic energy at (11) follows from

$$K[\rho_m, \mathbf{j}_m] = \sum_{\text{occ}} \epsilon_n - \int d^3r \rho_m(\mathbf{r}) [Z(\mathbf{r}) + \phi_{\text{ext}}(\mathbf{r})] + \frac{1}{c} \int d^3r \mathbf{j}_m(\mathbf{r}) [\mathbf{Y}(\mathbf{r}) + \mathbf{A}_{\text{ext}}(\mathbf{r})]. \quad (12)$$

In this approach it is impossible to find the kinetic energy at the starting densities ρ_0 and \mathbf{j}_0 . Rather, one obtains new densities ρ_m and \mathbf{j}_m with corresponding kinetic energy and hence total energy from

$$E[\rho_m, \mathbf{j}_m] = K[\rho_m, \mathbf{j}_m] + G[\rho_m, \mathbf{j}_m] + \int d^3r \rho_m(\mathbf{r}) \phi_{\text{ext}}(\mathbf{r}) - \frac{1}{c} \int d^3r \mathbf{j}_m(\mathbf{r}) \mathbf{A}_{\text{ext}}(\mathbf{r}). \quad (13)$$

At this point we have made a choice for the effective potentials in (9). A different approach is to find an effective potential which will also yield the initial densities as a solution of the Kohn-Sham equations.¹³ This approach certainly has its merits, but will not be discussed here.

The previous discussion shows clearly that the practical complications encountered in density-functional theory are related to the fact that one does not have an analytical form of the noninteracting kinetic energy. Very often one finds the expression (or its nonrelativistic equivalent)

$$\sum_n \phi_n^\dagger(\mathbf{r}) \left[\frac{\hbar c}{i} \boldsymbol{\alpha} \nabla + (\beta - 1) m c^2 \right] \phi_n(\mathbf{r}), \quad (14)$$

where the single-particle states sum up to the densities ρ and \mathbf{j} . The expression above, however, is not a functional of ρ and \mathbf{j} , since it is possible to make changes in the single-particle states, preserving ρ and \mathbf{j} , but changing the value of (14) in an arbitrary way.¹⁴

In order to find the minimum of (13), one has to find new guesses for the densities ρ_0 and \mathbf{j}_0 . These are used as input to evaluate the corresponding densities ρ_m and \mathbf{j}_m . Finally, one obtains the kinetic energy and the total energy through (13). If all changes in the initial densities produce a larger value of the total energy, the minimum of the total energy is found and the densities ρ_m and \mathbf{j}_m correspond to the true densities ρ_{min} and \mathbf{j}_{min} of the many-body problem.

When we use the true minimum densities ρ_{min} and \mathbf{j}_{min} as starting guesses, the difference between (8) and (6) is in lowest order:

$$\begin{aligned} & \frac{1}{2} \int d^3r \int d^3r' \left[[\rho(\mathbf{r}) - \rho_{\text{min}}(\mathbf{r})] \frac{\partial^2 G}{\partial \rho^2} [\rho(\mathbf{r}') - \rho_{\text{min}}(\mathbf{r}')] \right. \\ & \quad + [\mathbf{j}(\mathbf{r}) - \mathbf{j}_{\text{min}}(\mathbf{r})] \frac{\partial^2 G}{\partial \mathbf{j}^2} \\ & \quad \times [\mathbf{j}(\mathbf{r}') - \mathbf{j}_{\text{min}}(\mathbf{r}')] + [\rho(\mathbf{r}) - \rho_{\text{min}}(\mathbf{r})] \\ & \quad \left. \times \frac{\partial^2 G}{\partial \rho \partial \mathbf{j}} [\mathbf{j}(\mathbf{r}') - \mathbf{j}_{\text{min}}(\mathbf{r}')] \right]. \quad (15) \end{aligned}$$

Therefore, (8) will also have an extremum at ρ_{min} and \mathbf{j}_{min} because (6) is minimal. Since any noninteracting Hamiltonian is unbounded towards plus infinity (one can simply take states with higher and higher kinetic energies), the only extremum for (8) is the ground state and, by necessity, this is a minimum. Hence we find $\rho_m = \rho_{\text{min}}$ and $\mathbf{j}_m = \mathbf{j}_{\text{min}}$, indicating that these solutions are self-consistent. Therefore, if one is able to define functional derivatives through (7) and (15), a self-consistent solution of the Kohn-Sham equations, corresponding to the interactive ground state, does always exist. A self-consistent solution of the Kohn-Sham equations, however, exists for any extremum (maximum or minimum) of (6) and hence self-consistency by itself is not a sufficient criterium for a global minimum of the energy (6). The general conclusion is that, assuming the functional derivatives (7) and (15) do exist, a necessary and sufficient condition for ρ, \mathbf{j} to be an extremum of the energy (6) is that ρ, \mathbf{j} are self-consistent solutions of the Kohn-Sham equations (10).

A next question is whether expansion (7) is always possible. Certainly all local density functionals have a functional derivative, but it is unclear if the true functional G also has this property. If one does not require that the functions Z and Y in (7) are functional derivatives, but some close approximations, then (15) does not hold anymore, and the minimum of (6) does not correspond to a self-consistent solution of the Kohn-Sham equations. Also, it follows immediately from (6) that for all densities, for which the functional derivatives of F exist, one can find external potentials in such a way that these densities are the ground-state densities of the interacting electron system with these potentials. One simply takes the external scalar and vector potentials to be the negative of these functional derivatives. Similarly, all densities for which the functional derivatives of K exist can be represented as the ground state of a noninteractive problem. It is an open question what happens when the functional derivatives of K or G do not exist, and whether that will introduce some special physical effects.

IV. BREIT INTERACTION

The definition of the universal functional F contains radiation effects because the photon Hamiltonian and photon states are included. It is well known that this changes the effective electron-electron interaction and that one has to include the so-called Breit^{3,15} modification

$$W = \frac{e^2}{2 |\mathbf{r} - \mathbf{r}'|} : \psi_\kappa^\dagger(\mathbf{r}) \psi_\lambda^\dagger(\mathbf{r}') \left[\delta_{\lambda\mu} \delta_{\kappa\nu} - \frac{1}{2} \boldsymbol{\alpha}_{\lambda\mu} \boldsymbol{\alpha}_{\kappa\nu} - \frac{1}{2} \frac{[\boldsymbol{\alpha}_{\lambda\mu}(\mathbf{r} - \mathbf{r}')][\boldsymbol{\alpha}_{\kappa\nu}(\mathbf{r} - \mathbf{r}')] }{(\mathbf{r} - \mathbf{r}')^2} \right] \psi_\mu(\mathbf{r}') \psi_\nu(\mathbf{r}) :. \quad (16)$$

Like the charge-charge interaction the other terms should also be taken into account in a Hartree approximation. There will be additional current-current exchange and correlation effects, which modify the effective

exchange and correlation potential.¹¹ In this paper, however, we will not deal with these modifications and only include the Breit interaction in the Hartree approximation. The effective scalar potential in Eq. (10) has the fa-

miliar form

$$\phi_{\text{eff}}(\mathbf{r}) = \phi_{\text{ext}}(\mathbf{r}) + \phi_{\text{xc}}(\mathbf{r}) + \int d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \quad (17)$$

while the vector potential becomes

$$\begin{aligned} \mathbf{A}_{\text{eff}}(\mathbf{r}) = & \mathbf{A}_{\text{ext}}(\mathbf{r}) + \mathbf{A}_{\text{xc}}(\mathbf{r}) \\ & + \frac{1}{2c} \int d^3r' \frac{1}{|\mathbf{r}-\mathbf{r}'|} \\ & \times \left[\mathbf{j}(\mathbf{r}') + \frac{(\mathbf{r}-\mathbf{r}')[\mathbf{j}(\mathbf{r}')(\mathbf{r}-\mathbf{r}')] }{(\mathbf{r}-\mathbf{r}')^2} \right]. \quad (18) \end{aligned}$$

We can always add the divergence of a scalar field to the vector field, and hence the Breit term in \mathbf{A}_{eff} is equivalent to

$$\frac{1}{c} \int d^3r' \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \quad (19)$$

In other words, this last form yields the same results for physical observable quantities as the previous form. This modification is not possible in the total interaction energy, because there one has to consider the current-current correlation function, which in general does not factorize as a product of two currents. We are now able to reformulate the effective vector potential as an effective magnetic field by

$$\mathbf{B}_{\text{eff}}(\mathbf{r}) = \mathbf{B}_{\text{ext}}(\mathbf{r}) + \mathbf{B}_{\text{xc}}(\mathbf{r}) + \frac{1}{c} \int d^3r' \frac{\mathbf{j}(\mathbf{r}') \times (\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}. \quad (20)$$

V. FORMULATION WITH MAGNETIZATION DENSITY

Because the divergence of the current is zero, it is possible to write this current as the curl of a vector field \mathbf{M} , through

$$\mathbf{j}(\mathbf{r}) = c \nabla \times \mathbf{M}(\mathbf{r}), \quad (21)$$

where \mathbf{M} is defined up to the divergence of some field $\Xi(\mathbf{r})$. This is a standard ambiguity which has no physical consequences. $\mathbf{M}(\mathbf{r})$ is the magnetization density due to the current $\mathbf{j}(\mathbf{r})$. We are now able to rewrite the last term of Eq. (2) as

$$- \int d^3r \mathbf{B}_{\text{ext}}(\mathbf{r}) \mathbf{M}(\mathbf{r}) \quad (22)$$

when we assume that the accompanying surface integral vanishes. This is certainly the case when the external magnetic field vanishes at infinity, as in normal experiments. One has to keep this point in mind, though, when using a constant external field in the calculations. In that case $\mathbf{M}(\mathbf{r})$ should be defined to be bounded, which is only possible for a finite sample.

As defined in Eq. (3), the current $\mathbf{j}(\mathbf{r})$ has no clear physical interpretation. But when we know that the field satisfies the Dirac equation, it is possible to use the Gordon decomposition to rewrite the current in a different way. The Gordon decomposition can be performed in two slightly different ways (see Appendix A), but the final results are equivalent. We approximate the current by

$$-c\mu_B \left[\frac{1}{i} : \psi^\dagger(\mathbf{r})(\nabla - \bar{\nabla})\psi(\mathbf{r}) : + \nabla \times : \psi^\dagger(\mathbf{r})\Sigma\psi(\mathbf{r}) : \right] \quad (23)$$

in which the first term expresses the charge flow of the electrons and the second term relates to the spins of the electrons. One can compare this to the total angular momentum density of the Dirac field:

$$\mathbf{g}(\mathbf{r}) = : \psi^\dagger(\mathbf{r}) \left[\mathbf{r} \times \frac{\hbar}{i} \nabla + \frac{\hbar}{2} \Sigma \right] \psi(\mathbf{r}) :. \quad (24)$$

Note⁹ that we multiply by ψ^\dagger on the left and not by $\bar{\psi}$. The angular momentum is a pseudovector and not part of a tensor, as it would be with $\bar{\psi}$ on the left-hand side of the expression. Further, here too one clearly identifies the first part of (19) as the orbital momentum and the second as the spin.

The magnetic moment density associated with the orbital current corresponds exactly to the first term of Eq. (19) and therefore we see that the vector field \mathbf{M} is the sum of a term related to the angular momentum of the Dirac field, plus terms depending on higher-order multipole moments and relativistic corrections. Approximating \mathbf{M} by (see Appendix B)

$$-\mu_B : \psi^\dagger(\mathbf{r}) \left[\frac{\mathbf{r} \times \nabla}{i} + \Sigma \right] \psi(\mathbf{r}) : \quad (25)$$

is therefore very reasonable, because this only neglects three small effects. First, one ignores the coupling of the effective field to the quadrupole and higher moments of the electronic current.¹⁶ This is certainly justified for the external field, but it introduces errors in the energy related with the exchange-correlation field. Second, the term with the electromagnetic field \mathbf{A} is not included. This term couples the charge density to the square of the electromagnetic field, and hence describes processes where two photons are created (destroyed) simultaneously or where a photon is created (destroyed) when the electron scatters in the external field. The latter process should not be confused with Compton scattering, where the scattering and creation are virtually independent events (e.g., the intermediate state is not real, because it violates energy conservation). The last set of terms which are excluded describe relativistic corrections, of which the order of magnitude is smaller by the square of the fine-structure constant.

One point is very important. According to (8) the effective exchange-correlation potential couples to the total current and hence the effective exchange-correlation magnetic field couples to both the spin and orbital angular momentum. This is clearly not a physical situation, since this would give rise to large orbital contributions to the magnetic moments, even in 3d transition metals. It is easy to show, however, that the standard approximation to the exchange interaction only couples the spin of the electrons to the effective magnetic field.¹⁷ Hence the orbital moment only couples to the external magnetic field, the Hartree part of the additional terms in the Breit interaction, and some parts of the correlation potential. This coupling is therefore not dominant in determining

the general magnetic properties of solids, but they will show up in small effects like the magnetic anisotropy. Since we have used different approximations for the spin and orbital parts of the current to arrive at (21), the fact that these terms couple to different fields can be expected. At this point we therefore assume that the orbital contri-

bution only couples to the first and the third term of (20), and these are exactly the cases investigated in Appendix B.

After making the approximation (25), the set of single-particle equations (10) is transformed to

$$\left[\frac{\hbar c}{i} \boldsymbol{\alpha} \nabla + (\beta - 1) m c^2 - e \phi_{\text{eff}}(\mathbf{r}) + \mu_B (\mathbf{B}_{\text{eff}}(\mathbf{r}) - \mathbf{B}_{\text{xc}}(\mathbf{r})) \frac{\mathbf{r} \times \nabla}{i} + \mu_B \mathbf{B}_{\text{eff}}(\mathbf{r}) \boldsymbol{\Sigma} \right] \phi_n(\mathbf{r}) = \epsilon_n \phi_n(\mathbf{r}). \quad (26)$$

The term related to the orbital angular momentum is not periodic, however, leading to an effect on the energy which is a function of the shape of the sample. This also introduces some errors, and (26) should only be regarded as a first approximation to the inherently nonlocal effects of the diamagnetic terms. Fortunately, for the localized core electrons the integral of the current over a unit cell is zero, and in that case a shift of the origin has no physical consequences.

VI. MAGNETIC SHAPE ANISOTROPY

The Hartree energy part of the Breit interaction is written in the following way:

$$\frac{1}{2} \int d^3r d^3r' \frac{1}{|\mathbf{r} - \mathbf{r}'|^3} \left[\mathbf{M}(\mathbf{r}) \mathbf{M}(\mathbf{r}') - 3 \frac{[(\mathbf{r} - \mathbf{r}') \mathbf{M}(\mathbf{r})][(\mathbf{r} - \mathbf{r}') \mathbf{M}(\mathbf{r}')] }{|\mathbf{r} - \mathbf{r}'|^2} \right] - \frac{4\pi}{3} \int d^3r \mathbf{M}(\mathbf{r}) \mathbf{M}(\mathbf{r}). \quad (27)$$

This contribution to the total energy of the Breit interaction favors magnetic ordering. The magnitude of $\mathbf{M}(\mathbf{r})$ is determined by μ_B , as is seen from (25). The value of the Bohr magneton μ_B in atomic units is 5.2×10^{-3} and hence the order of magnitude of the Breit energy is 25 μRy or a few Kelvin only. It was already pointed out long time ago that this interaction therefore cannot explain ferromagnetism in materials like iron, and that the exchange energy is dominant.

As is well known, any nonspherical sample will have a preferred direction of the magnetization simply due to the shape of the sample, and one has to correct for this. The standard correction assumes that the magnetization is uniform throughout the crystal and the corresponding dipole-dipole interaction is considered to be the origin of the shape anisotropy. According to Eq. (27) this shape anisotropy is directly related to the Breit interaction. When we define the average magnetic moment by

$$\mathbf{m}_0 \int d^3r = \int d^3r \mathbf{M}(\mathbf{r}) \quad (28)$$

we find that the energy term giving the shape anisotropy is

$$\frac{1}{2} \int d^3r d^3r' \frac{1}{|\mathbf{r} - \mathbf{r}'|^3} \left[\mathbf{m}_0 \mathbf{m}_0 - 3 \frac{[(\mathbf{r} - \mathbf{r}') \mathbf{m}_0]^2}{|\mathbf{r} - \mathbf{r}'|^2} \right] - \frac{4\pi}{3} \mathbf{m}_0^2 \int d^3r. \quad (29)$$

We have included the contact term in this definition, although obviously this term is isotropic. This will affect the total magnetization, but that can be neglected since the exchange energy is much larger than this contact energy.

The intrinsic anisotropy is defined by subtracting (29) from (27). It is easy to show that terms containing a factor $\mathbf{M}(\mathbf{r})$ and a factor \mathbf{m}_0 give a zero contribution, because one essentially integrates a spherical harmonic with $l=2, m=0$ over all angles. The intrinsic part of the Breit interaction then is identical to (27) with $\mathbf{M}(\mathbf{r})$ replaced by $\mathbf{M}(\mathbf{r}) - \mathbf{m}_0$. This integral can be defined for an infinite periodic system, since the dependence on $|\mathbf{r} - \mathbf{r}'|$ has now changed to an inverse fifth power for large separations. For a bulk solid we therefore define the Breit energy per unit cell as

$$\frac{1}{2} \int_{\Omega} d^3r \int_{\Omega} d^3r' \sum_{\mathbf{R}} \frac{1}{|\mathbf{r} - \mathbf{r}' + \mathbf{R}|^3} \left[[\mathbf{M}(\mathbf{r}) - \mathbf{m}_0][\mathbf{M}(\mathbf{r}') - \mathbf{m}_0] - 3 \frac{\{(\mathbf{r} - \mathbf{r}' + \mathbf{R})[\mathbf{M}(\mathbf{r}) - \mathbf{m}_0]\} \{(\mathbf{r} - \mathbf{r}' + \mathbf{R})[\mathbf{M}(\mathbf{r}') - \mathbf{m}_0]\}}{|\mathbf{r} - \mathbf{r}' + \mathbf{R}|^2} \right] - \frac{4\pi}{3} \int_{\Omega} d^3r [\mathbf{M}(\mathbf{r}) - \mathbf{m}_0]^2, \quad (30)$$

where the integrations are performed over the unit cell Ω only, and the summation is over all Bravais lattice vectors \mathbf{R} . The average magnetization is defined by

$$\mathbf{m}_0\Omega = \int_{\Omega} d^3r \mathbf{M}(\mathbf{r}). \quad (31)$$

By expanding terms for large values of \mathbf{R} , we see that the term with R^{-3} has a prefactor zero, and that the series converges like a sum over R^{-5} .

VII. CRYSTALLINE ANISOTROPY

After having identified the term in the electron-electron interaction which is responsible for the shape anisotropy of a material, the next step is to find the origin of the crystalline anisotropy. As is well known, the orbital angular momentum is quenched in $3d$ transition metals, which means that the magnetic moment is essentially spin derived. Even in that case, however, the Breit interaction (30) couples the magnetic moment to the crystalline lattice. Bozorth and McKeehan¹⁸ showed that a nonspherical spin distribution around the nuclei can give rise to the observed anisotropies, depending on the shape of the spin distribution. But Van Vleck already argued² that this contribution would be too small by 3 orders of magnitude. His reasoning is based on the observation that spin-orbit coupling is weak and that the asymmetry in the electronic charge cloud does change very much as a function of the direction of the spins. Spin-orbit interaction is necessary, since the direct spin-spin part of the Breit interaction only yields terms proportional to \cos^2 , which have to average out in a cubic crystal.

A nonzero value of the orbital angular momentum density hence is essential for the existence of crystalline magnetic anisotropy. The only reason that the orbital angular momentum is nonzero is the spin-orbit coupling due to the kinetic energy operator of the Dirac equation. Fully relativistic calculations are therefore needed. These kinetic terms, however, only couple the spin of a given electron to its own orbit and, again, Van Vleck has shown² that the corresponding contribution to the energy is too small to explain the observed values of the anisotropy constants. He stresses that the coupling between orbital angular momentum vectors of different atoms is important. Later papers¹⁹ also focus on the role of spin to other orbit coupling. Both these terms are included in the Hartree part of the Breit interaction and hence in the effective potential for the Kohn-Sham equations (26). Like for the Coulomb interaction, there will be corrections due to exchange and correlation, but at this point we will make the *ad hoc* assumption that these corrections are smaller than the main effect represented in the Hartree part.

When we assume that the direction of the magnetization is the same everywhere in the unit cell

$$\mathbf{M}(\mathbf{r}) - \mathbf{m}_0 = \hat{\mathbf{m}}_0 f(\mathbf{r}) \quad (32)$$

Eq. (30) simplifies to

$$- \int_{\Omega} d^3r \int_{\Omega} d^3r' \sum_{\mathbf{R}} \frac{f(\mathbf{r})f(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'+\mathbf{R}|^3} P_2(\cos(\hat{\mathbf{m}}_0, \mathbf{r}-\mathbf{r}'+\mathbf{R})) - \frac{4\pi}{3} \int_{\Omega} d^3r [f(\mathbf{r})]^2. \quad (33)$$

If $f(\mathbf{r})$ has cubic symmetry, the first (anisotropic) term is zero, because it contains a summation of $P_2(\cos(\theta))$ over all cubic rotations. But due to the relativistic terms in the kinetic energy $f(\mathbf{r})$ will not have cubic symmetry, although the deviation is small. According to van Vleck, the sum over \mathbf{R} in (33) will give a contribution which is large enough to explain the magnitude of the magnetic anisotropy.

VIII. CONCLUSIONS

In this paper we have shown that it is possible to discuss the origins of the magnetic anisotropy in terms of relativistic density-functional theory. We have identified two terms leading to a spin-orbit coupling needed for the existence of the magnetic anisotropy. The magnetic shape anisotropy results from a dipole-dipole interaction due to the Breit modification of the relativistic two-electron energy. Future calculations will have to show whether (30) is an important contribution to the crystalline anisotropy.

ACKNOWLEDGMENTS

The research published in this paper was supported by the Office of Naval Research under Contract No. N00014-86-K-0294. I would like to thank Dr. M. F. H. Schuurmans, Dr. G. H. O. Daalderop, and Dr. P. J. Kelly for many stimulating and helpful discussions during my stay at the Philips Research Laboratory in Eindhoven. I am also very grateful to Dr. K. B. Hathaway and Dr. A. J. Freeman for many supportive discussions.

APPENDIX A: THE GORDON DECOMPOSITION

The Gordon decomposition yields an expression for the expectation value of the current for a solution of the time-independent form of the Dirac equation

$$\left\{ \frac{\hbar c}{i} \boldsymbol{\alpha} \nabla + (\beta - 1) m c^2 - e\phi(\mathbf{r}) + e\boldsymbol{\alpha} \mathbf{A}(\mathbf{r}) \right\} \psi_n(\mathbf{r}) = \epsilon_n \psi_n(\mathbf{r}). \quad (A1)$$

The conjugate equation is

$$\psi_n^\dagger(\mathbf{r}) \left\{ -\frac{\hbar c}{i} \boldsymbol{\alpha} \bar{\nabla} + (\beta - 1) m c^2 - e\phi(\mathbf{r}) + e\boldsymbol{\alpha} \mathbf{A}(\mathbf{r}) \right\} = \epsilon_n \psi_n^\dagger(\mathbf{r}). \quad (A2)$$

At this point we multiply (A1) by α_k and (A2) by α_k from the right and add the two equations together. This results in

$$\frac{\hbar c}{i} \psi^\dagger(\mathbf{r}) [\alpha_k \boldsymbol{\alpha} \nabla - \boldsymbol{\alpha} \bar{\nabla} \alpha_k] \psi(\mathbf{r}) - 2[mc^2 + e\phi(\mathbf{r})] \psi^\dagger(\mathbf{r}) \alpha_k \psi(\mathbf{r}) + e \psi^\dagger(\mathbf{r}) [\alpha_k \boldsymbol{\alpha} \mathbf{A}(\mathbf{r}) + \boldsymbol{\alpha} \mathbf{A}(\mathbf{r}) \alpha_k] \psi(\mathbf{r}) = 2\varepsilon_n \psi^\dagger(\mathbf{r}) \alpha_k \psi(\mathbf{r}). \quad (\text{A3})$$

Alternatively, one can multiply (A1) by $\beta \alpha_k$ and (A2) by $\beta \alpha_k$ from the right, subtract, and get

$$\frac{\hbar c}{i} \psi^\dagger(\mathbf{r}) \beta [\alpha_k \boldsymbol{\alpha} \nabla - \boldsymbol{\alpha} \bar{\nabla} \alpha_k] \psi(\mathbf{r}) + 2mc^2 \psi^\dagger(\mathbf{r}) \beta \alpha_k \beta \psi(\mathbf{r}) + e \psi^\dagger(\mathbf{r}) \beta [\alpha_k \boldsymbol{\alpha} \mathbf{A}(\mathbf{r}) + \boldsymbol{\alpha} \mathbf{A}(\mathbf{r}) \alpha_k] \psi(\mathbf{r}) = 0. \quad (\text{A4})$$

Using some standard identities,²⁰ (A3) and (A4) can be rewritten in the form

$$\psi^\dagger(\mathbf{r}) \alpha_k \psi(\mathbf{r}) = \frac{\left[\frac{\hbar}{2mc} \left[\frac{1}{i} \psi^\dagger(\mathbf{r}) (\nabla_k - \bar{\nabla}_k) \psi(\mathbf{r}) + [\nabla \times \psi^\dagger(\mathbf{r}) \boldsymbol{\Sigma} \psi(\mathbf{r})]_k \right] + \frac{e}{mc^2} \psi^\dagger(\mathbf{r}) A_k(\mathbf{r}) \psi(\mathbf{r}) \right]}{\left[1 + \frac{\varepsilon_n + e\phi(\mathbf{r})}{mc^2} \right]}, \quad (\text{A5})$$

$$\psi^\dagger(\mathbf{r}) \alpha_k \psi(\mathbf{r}) = \frac{\hbar}{2mc} \left[\frac{1}{i} \psi^\dagger(\mathbf{r}) \beta (\nabla - \bar{\nabla}) \psi(\mathbf{r}) + \nabla \times \psi^\dagger(\mathbf{r}) \beta \boldsymbol{\Sigma} \psi(\mathbf{r}) \right] + \frac{e}{mc^2} \psi^\dagger(\mathbf{r}) \beta A_k(\mathbf{r}) \psi(\mathbf{r}), \quad (\text{A6})$$

and Eq. (A5) immediately shows how the current is related to the particle current and the spin, with an additional term proportional to the expectation value of the external field. The denominator is equal to one plus terms of order $(1/c)^2$. The second form (A6) has incorporated this denominator by introducing factors of β , thereby losing the direct physical meaning of these terms.

APPENDIX B: ORBITAL ANGULAR MOMENTUM

Ignoring terms of order $(1/c)^2$ in the expression (A5) for the current, and ignoring the term with the field \mathbf{A} , we can write the current in the following form:

$$\begin{aligned} \mathbf{j}(\mathbf{r}) &= -ec \langle a | : \frac{\hbar}{2mc} \\ &\times \left[\frac{1}{i} \psi^\dagger(\mathbf{r}) (\nabla - \bar{\nabla}) \psi(\mathbf{r}) + \nabla \times \psi^\dagger(\mathbf{r}) \boldsymbol{\Sigma} \psi(\mathbf{r}) \right] : | a \rangle \end{aligned} \quad (\text{B1})$$

from which it is immediately clear that

$$\mathbf{M}_{\text{spin}}(\mathbf{r}) = -\mu_B \langle a | : \psi^\dagger(\mathbf{r}) \boldsymbol{\Sigma} \psi(\mathbf{r}) : | a \rangle. \quad (\text{B2})$$

The orbital part of the current introduces a term

$$\frac{\mu_B}{i} \int d^3r \mathbf{A}_{\text{eff}}(\mathbf{r}) \psi^\dagger(\mathbf{r}) \{ \nabla - \bar{\nabla} \} \psi(\mathbf{r}) \quad (\text{B3})$$

in the expression for the total energy. Because $\nabla \cdot \mathbf{A} = 0$, one can rewrite this term as

$$\frac{2\mu_B}{i} \int d^2r \mathbf{A}_{\text{eff}}(\mathbf{r}) \psi^\dagger(\mathbf{r}) \nabla \psi(\mathbf{r}). \quad (\text{B4})$$

If the effective field \mathbf{B} is constant, which normally is a good approximation to the external field, one defines the field \mathbf{A} as

$$\mathbf{A}(\mathbf{r}) = -\frac{1}{2} (\mathbf{r} - \mathbf{r}_0) \times \mathbf{B} \quad (\text{B5})$$

and the integral in the total energy then leads to an effective orbital magnetization of the form

$$\mathbf{M}_{\text{orb}}(\mathbf{r}) = -\frac{\mu_B}{i} (\mathbf{r} - \mathbf{r}_0) \times \psi^\dagger(\mathbf{r}) \nabla \psi(\mathbf{r}). \quad (\text{B6})$$

The Hartree part of the Breit energy has the form

$$-\frac{1}{2c^2} \int d^3r d^3r' \frac{\mathbf{j}(\mathbf{r}) \mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (\text{B7})$$

When we integrate \mathbf{r}' over a unit cell Ω at the origin, the contributions of the integral to a unit cell at a large distance \mathbf{r} can be approximated by writing¹⁶

$$\mathbf{A}(\mathbf{r}) = \frac{1}{c} \int_{\Omega} d^3r' \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \simeq \frac{\mathbf{m} \times (\mathbf{r} - \mathbf{r}'_0)}{|\mathbf{r} - \mathbf{r}'_0|^3}, \quad (\text{B8})$$

where \mathbf{r}'_0 is the center of the unit cell Ω at the origin and

$$\mathbf{m} = \frac{1}{2c} \int_{\Omega} d^3r' (\mathbf{r}' - \mathbf{r}) \times \mathbf{j}(\mathbf{r}'). \quad (\text{B9})$$

The unit cell Ω has to be chosen such that the integral of the current over Ω is 0. Therefore we can only incorporate the effects of the diamagnetic currents of core shells and ignore all macroscopic currents which are induced by the external field. Again, at this point that is a reasonable approximation. Next, we approximate (B7) by

$$\begin{aligned} &-\frac{1}{2c} \int d^3r \mathbf{j}(\mathbf{r}) \frac{\mathbf{m} \times (\mathbf{r} - \mathbf{r}'_0)}{|\mathbf{r} - \mathbf{r}'_0|^3} \\ &= -\frac{1}{2c} \int d^3r \mathbf{m} \frac{(\mathbf{r} - \mathbf{r}'_0) \times \mathbf{j}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'_0|^3}. \end{aligned} \quad (\text{B10})$$

Because of (B4) one can write the orbital part of the current as

$$-\frac{2c\mu_B}{i} \psi^\dagger(\mathbf{r}) \nabla \psi(\mathbf{r}) \quad (\text{B11})$$

and hence (B10) becomes

$$-\int d^3r \mathbf{m} \mathbf{M}_{\text{orb}}(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'_0|^3}. \quad (\text{B12})$$

Inserting the definition (B9) of \mathbf{m} into this equation then yields approximately

$$-\frac{1}{2c} \int d^3r d^3r' \frac{(\mathbf{r}'-\mathbf{r}) \times \mathbf{j}(\mathbf{r}') \mathbf{M}_{\text{orb}}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|^3}. \quad (\text{B13})$$

Introducing the standard formula for $\mathbf{B}(\mathbf{r})$ (Ref. 16) then

transforms this last equation into

$$-\frac{1}{2} \int d^3r \mathbf{B}(\mathbf{r}) \mathbf{M}_{\text{orb}}(\mathbf{r}) \quad (\text{B14})$$

showing that (B6) is also a good approximation in this case. Obviously, we have neglected here all higher multipoles of the current field, and these will certainly play a role for contributions of cells which are close together. But as a first approximation (B6) seems an adequate definition for coupling the angular momentum to the external and Breit fields.

¹G. S. Mahanji, Philos. Trans. R. Soc. London **228**, 63 (1929).

²J. H. Van Vleck, Phys. Rev. **52**, 1178 (1937).

³H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems* (Plenum, New York, 1977).

⁴H. Brooks, Phys. Rev. **58**, 909 (1940).

⁵G. C. Fletcher, Proc. Phys. Soc. **67A**, 505 (1954); N. Mori, J. Phys. Soc. Jpn. **27**, 307 (1969).

⁶E. I. Kondorskii and E. Straube, Zh. Eksp. Teor. Fiz. **63**, 356 (1972) [Sov. Phys.—JETP **36**, 188 (1973)]; N. Mori, Y. Fukuda, and T. Ukai, J. Phys. Soc. Jpn. **37**, 1263 (1974).

⁷H. Eckardt, L. Fritsche, and J. Noffke, J. Phys. F **17**, 943 (1987).

⁸H. J. F. Jansen, B. Gutekunst, G. H. O. Daalderop, P. J. Kelly, and M. F. H. Schuurmans, Bull. Am. Phys. Soc. **33**, 271 (1988); J. G. Gay and R. Richter, J. Appl. Phys. **61**, 3362 (1987).

⁹F. Mandl and G. Shaw, *Quantum Field Theory* (Wiley, New York, 1984).

¹⁰S. Gasiorowicz, *Elementary Particle Physics* (Wiley, New York, 1966).

¹¹A. K. Rajagopal and J. Callaway, Phys. Rev. B **7**, 1912 (1973); A. K. Rajagopal, in *Density Functional Methods in Physics*,

edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985), p. 159, and references therein; A. H. MacDonald, J. Phys. C **16**, 3869 (1983); H. Eschrig, G. Seifert, and P. Ziesche, Solid State Commun. **56**, 777 (1985).

¹²E. H. Lieb, in *Density Functional Methods in Physics*, Ref. 11, p. 31.

¹³J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976); M. R. Norman and D. D. Koelling, Phys. Rev. B **30**, 5530 (1984).

¹⁴T. L. Gilbert, Phys. Rev. B **12**, 2111 (1975).

¹⁵A. O. Barut and B. W. Xu, Ann. Phys. (NY) **148**, 135 (1983); G. Breit, Phys. Rev. **34**, 553 (1929); **39**, 616 (1932).

¹⁶J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).

¹⁷J. Callaway, *Quantum Theory of the Solid State* (Academic, New York, 1974), p. 331.

¹⁸R. M. Bozorth and L. W. McKeehan, Phys. Rev. **51**, 216 (1937).

¹⁹M. Blume, A. J. Freeman, and R. E. Watson, Phys. Rev. **134**, A320 (1964).

²⁰P. K. Misra and J. Callaway, Phys. Rev. B **36**, 3731 (1987).