HYPERFINE STRUCTURE

HYPERFINE structure in paramagnetic resonance lines was discovered by Penrose (1949). It is due to the coupling of the electrons with the electric and magnetic moments of the nucleus. In this chapter we shall derive the form of the interactions responsible for these couplings. We begin with the electrostatic interactions which, although less important in practice, are conceptually simpler.

17.1. Electrostatic hyperfine interactions

Atomic energy levels are usually calculated under the simplifying assumption of a point nucleus with a charge Ze. If the finite nuclear dimensions are taken into account, the changes in the atomic energy levels are exceedingly small and, for nuclear spins different from zero, are best observed through the removal of the degeneracy of the atomic levels associated with the various orientations of the nuclear spin. The Hamiltonian responsible for the splittings between those levels can be determined as follows, using the correspondence principle as a starting point. If we describe the nucleus and the electron cloud as two classical charge distributions $\rho_n(r_n)$ and $\rho_e(r_e)$, their mutual electrostatic energy is

$$W_E = \iint \frac{\rho_e(r_e)\rho_n(r_n)\, dr_e\, dr_n}{|\mathbf{r}_n - \mathbf{r}_e|}, \tag{17.1}$$

which can be expanded through the classical formula

$$\frac{1}{|\mathbf{r}_n - \mathbf{r}_e|} = \sum_{k=0}^{\infty} \sum_{q=-k}^{k} \frac{r_<^k}{r_>^{k+1}} C_k^q(\theta_n, \varphi_n) C_k^{q*}(\theta_e, \varphi_e), \tag{17.2}$$

where the functions $C_k^q(\theta, \varphi)$ are normalized spherical harmonics: $C_k^q(\theta, \varphi) = \sqrt{\{4\pi/(2k+1)\}} Y_k^q(\theta, \varphi)$ and where the symbols $r_<$ and $r_>$ mean that the larger of the two numbers r_e and r_n is in the denominator, and the smaller in the numerator. If the small penetration of the electron inside the nucleus is neglected we may assume $r_e > r_n$ and write

$$W_E = \sum_{k,q} A_k^q B_k^{q*}, \tag{17.3}$$

where

$$\begin{cases} A_k^q = \int \rho_{\mathrm{n}}(r_{\mathrm{n}}) r_{\mathrm{n}}^k C_k^q(\theta_{\mathrm{n}}, \varphi_{\mathrm{n}}) \, \mathrm{d}r_{\mathrm{n}}, \\ B_k^q = \int \rho_{\mathrm{e}}(r_{\mathrm{e}}) r_{\mathrm{e}}^{-(k+1)} C_k^q(\theta_{\mathrm{e}}, \varphi_{\mathrm{e}}) \, \mathrm{d}r_{\mathrm{e}}. \end{cases} \quad (17.4)$$

If the state of the nucleus is described by a wave-function $\Psi_{\mathrm{n}}(R_1, \dots, R_A)$ of the coordinates of its A nucleons, the nuclear charge density can be written as the expectation value of the operator density of charge at the point r_{n}:

$$\rho_{\mathrm{n}}(r_{\mathrm{n}}) = (\Psi_{\mathrm{n}}| \sum_{i=1}^{A} e_i \, \delta(r_{\mathrm{n}} - R_i) |\Psi_{\mathrm{n}}), \quad (17.5)$$

where $e_i = e$ for a proton and zero for a neutron. From (17.3) and (17.4) A_k^q can be written as an expectation value $A_k^q = \langle \mathscr{A}_k^q \rangle$ where the nuclear operator $\mathscr{A}_k^q$ is defined by

$$\mathscr{A}_k^q = \sum_i e_i R_i^k C_k^q(\Theta_i, \Phi_i), \quad (17.6)$$

R_i, Θ_i, Φ_i being the polar coordinates of the A nucleons. Similarly, B_k^q is the expectation value of the electron operator $\mathscr{B}_k^q$:

$$\mathscr{B}_k^q = -e \sum_{i=1}^{N} r_i^{-(k+1)} C_k^q(\theta_i, \varphi_i), \quad (17.7)$$

where r_i, θ_i, φ_i are the coordinates of the electrons and the minus sign occurs because of the negative charge $(-e)$ of the electron. The energy of electrostatic interaction between the electrons and the nucleus is then the expectation value of a Hamiltonian

$$\mathscr{H}_E = \sum_{k,q} \mathscr{A}_k^q \mathscr{B}_k^{q*}. \quad (17.8)$$

From the definitions (17.6) and (17.7) it is clear that the operators $\mathscr{A}_k^q$ and $\mathscr{B}_k^q$ are tensor operators of order k. The tensor operator $\mathscr{A}_k$ with $2k+1$ components $\mathscr{A}_k^q$ is called the multipole moment of order k of the nucleus. Although for the calculation of the splittings of the atomic energy levels, only the expectation value $A_k^q = \langle \mathscr{A}_k^q \rangle$ of the nuclear multipole operator, taken over the wave-function of a nuclear eigenstate, is required, this operator does in fact have off-diagonal matrix elements between various nuclear states of different energy. The study of these matrix elements, which are responsible for γ-ray transitions between those states, is outside the scope of this book.

For the diagonal matrix elements $A_k^q = \langle \mathscr{A}_k^q \rangle$ odd values of k are forbidden if we assume, as seems well established experimentally, that

stationary nuclear states have well-defined parities. In particular, nuclei should have no permanent electric dipole moments ($k = 1$) in agreement with experimental evidence. Invariance with respect to time reversal (cf. the discussion in § 15.10) also forbids the existence of nuclear electric dipole moments. Off-diagonal matrix elements of the electric dipole nuclear moment between nuclear states of different parity may of course exist. Further information and limitations on the values of the matrix elements of the nuclear multipole operators result from their tensor character and are based on the Wigner–Eckart theorem. It follows from this theorem that for a nucleus of spin I, $A_k^q = \langle \mathscr{A}_k^q \rangle$ will only be different from zero if $k \leqslant 2I$. Thus nuclei of spin $I \geqslant 1$ may have quadrupole moments, nuclei with $I \geqslant 2$ may have moments of order 4, etc, ... The term $k = 0$ of the electrostatic interaction between electron and nucleus clearly corresponds to the coupling with a point charge Ze. Since the nuclear radius R is much smaller than the electronic radius a, the various terms of W_E as given by eqns (17.3), (17.4) decrease rapidly, roughly as $(R/a)^k$. This explains why there is little experimental evidence of electrostatic interactions with $k > 2$.

Only quadrupole interactions will be considered from now on. The components of the nuclear quadrupole moment operator can be written in the form

$$\mathscr{A}_2^0 = \tfrac{1}{2} \sum_i e_i (3z_i^2 - r_i^2),$$

$$\mathscr{A}_2^{\pm 1} = \mp \sqrt{(\tfrac{3}{2})} \sum_i e_i z_i (x_i \pm \mathrm{i} y_i), \qquad (17.9)$$

$$\mathscr{A}_2^{\pm 2} = \sqrt{(\tfrac{3}{8})} \sum_i e_i (x_i \pm \mathrm{i} y_i)^2.$$

According to the Wigner–Eckart theorem the $\mathscr{A}_2^q$ have, for a nuclear spin I, within the manifold of the $(2I+1)$ substates $I_z = m$, the same matrix elements as the Hermitian tensor operator formed from the components of the vector I,

$$Q_2^0 = \langle I \| \alpha \| I \rangle (\tfrac{1}{2}) \{ 3I_z^2 - I(I+1) \},$$

$$Q_2^{\pm 1} = \mp \langle I \| \alpha \| I \rangle \sqrt{(\tfrac{3}{2})} \, \tfrac{1}{2} (I_z I_\pm + I_\pm I_z),$$

$$Q_2^{\pm 2} = \langle I \| \alpha \| I \rangle \sqrt{(\tfrac{3}{8})} \, I_\pm^2.$$

The constant $\langle I \| \alpha \| I \rangle$ is determined, for instance, by the condition that Q_2^0 and $\mathscr{A}_2^0$ have the same expectation value in the substate $I_z = I$ denoted by $|II\rangle$. The usual convention is to represent by the

symbol eQ the quantity

$$eQ = \langle II| \sum_{i=1}^{A} e_i(3z_i^2 - r_i^2) |II\rangle.$$

Then writing

$$eQ = 2\langle II| Q_2^0 |II\rangle = 2\langle II| \mathscr{A}_2^0 |II\rangle$$
$$= \langle I\| \alpha \|I\rangle\langle II| 3I_z^2 - I(I+1) |II\rangle,$$

we find that $\langle I\| \alpha\| I\rangle = eQ/\{I(2I-1)\}$, whence

$$Q_2^0 = \frac{eQ}{I(2I-1)} \tfrac{1}{2}\{3I_z^2 - I(I+1)\},$$

$$Q_2^{\pm 1} = \mp\frac{eQ}{I(2I-1)}\sqrt{(\tfrac{3}{2})}\,\tfrac{1}{2}(I_z I_\pm + I_\pm I_z), \qquad (17.10)$$

$$Q_2^{\pm 2} = \frac{eQ}{I(2I-1)}\sqrt{(\tfrac{3}{8})}\,(I_\pm)^2.$$

The components $\mathscr{B}_2^q$ of the electronic tensor (17.7) can be rewritten in the same way as the $\mathscr{A}_2^q$ in eqn (17.9), giving

$$\mathscr{B}_2^0 = -\tfrac{1}{2}\sum_{i=1}^{N}\frac{e_i(3z_i^2 - r_i^2)}{r_i^5},$$

$$\mathscr{B}_2^{\pm 1} = \pm\sqrt{(\tfrac{3}{2})}\sum_{i=1}^{N}\frac{e_i z_i(x_i \pm iy_i)}{r_i^5}, \qquad (17.11)$$

$$\mathscr{B}_2^{\pm 2} = -\sqrt{(\tfrac{3}{8})}\sum_{i=1}^{N}\frac{e_i(x_i \pm iy_i)^2}{r_i^5},$$

where the sum is over all electrons. The electron operators can be replaced by equivalent operators (as described in the last chapter), involving the orbital angular momentum vector of each electron. This is then summed over all electrons, giving

$$\mathscr{B}_2^0 = -e \sum_i \langle l\| \alpha \|l\rangle\langle r_q^{-3}\rangle_i \tfrac{1}{2}\{3l_z^2 - l(l+1)\},$$

$$\mathscr{B}_2^{\pm 1} = \pm e \sum_i \langle l\| \alpha \|l\rangle\langle r_q^{-3}\rangle_i \sqrt{(\tfrac{3}{2})}\,\tfrac{1}{2}(l_z l_\pm + l_\pm l_z), \qquad (17.12)$$

$$\mathscr{B}_2^{\pm 2} = -e \sum_i \langle l\| \alpha \|l\rangle\langle r_q^{-3}\rangle_i \sqrt{(\tfrac{3}{8})}\, l_\pm^2.$$

For a term in LS- coupling the sums can be replaced by operators involving the components of the total orbital angular momentum just as in the case of the crystal field. This is straightforward if the value of $\langle r_q^{-3}\rangle_i$ is the same for every electron involved but even if not all the values are the same it can be shown that in the non-relativistic limit

only one constant is needed in the quadrupole interaction (Sandars and Beck 1965). Hence we can write

$$\mathscr{B}_2^0 \equiv -e\langle r_q^{-3}\rangle\langle L\| \alpha \|L\rangle\tfrac{1}{2}\{3L_z^2-L(L+1)\},$$
$$\mathscr{B}_2^{\pm 1} \equiv \pm e\langle r_q^{-3}\rangle\langle L\| \alpha \|L\rangle\sqrt{(\tfrac{3}{2})}\ \tfrac{1}{2}(L_zL_\pm+L_\pm L_z), \qquad (17.13)$$
$$\mathscr{B}_2^{\pm 2} \equiv -e\langle r_q^{-3}\rangle\langle L\| \alpha \|L\rangle\sqrt{(\tfrac{3}{8})}\ (L_\pm)^2,$$

where $\langle r_q^{-3}\rangle$ is the mean inverse third power of the electron distance from the nucleus, averaged over the electronic wave-functions.

For an atom or ion with LS- coupling the form of the Hamiltonian for the electric quadrupole interaction is thus

$$\mathscr{H} = -\frac{e^2 Q\langle r_q^{-3}\rangle}{I(2I-1)}\langle L\| \alpha \|L\rangle\times$$

$$\times\left[\begin{array}{l}\tfrac{1}{4}\{3L_z^2-L(L+1)\}\{3I_z^2-I(I+1)\}+\\[4pt] +\tfrac{3}{8}\{(L_zL_++L_+L_z)(I_zI_-+I_-I_z)+(L_zL_-+L_-L_z)(I_zI_++I_+I_z)\}+\\[4pt] +\tfrac{3}{8}(L_+^2I_-^2+L_-^2I_+^2).\end{array}\right]$$

$$(17.14)$$

For the ground terms of the iron group the values of $\langle L\| \alpha \|L\rangle$ are listed in Table 19. The angular momentum operator expression in square brackets in (17.14) is in a form that is generally the most useful in calculations for weakly-bonded ions of the $3d$ group; it is equivalent to the more usual form

$$\tfrac{1}{2}\{3(\mathbf{L}.\mathbf{I})^2+\tfrac{3}{2}(\mathbf{L}.\mathbf{I})-L(L+1)I(I+1)\} \qquad (17.15)$$

and in the notation of Abragam and Pryce (1951a,b,c)

$$-\langle L\| \alpha \|L\rangle = +\eta_L. \qquad (17.16)$$

When L and S couple to a resultant J, similar formulae for the quadrupole interaction can be obtained in which L is replaced everywhere by J and $\langle L\| \alpha \|L\rangle$ by $\langle J\| \alpha \|J\rangle$. This representation is useful when considering ions of the rare-earth group, and is similar to that used in atomic beam measurements with a constant B which is given by the relation

$$B = -J(2J-1)e^2Q\langle r_q^{-3}\rangle\langle J\| \alpha \|J\rangle. \qquad (17.17)$$

We can of course retain a nomenclature similar to that of Abragam and Pryce by introducing a symbol η_J such that (cf. eqn (17.16))

$$-\langle J\| \alpha \|J\rangle = +\eta_J. \qquad (17.18)$$

For the state $J = L+S$ which is the ground state of the second half of

the rare-earth group we have the simple relationship

$$\frac{\langle J \| \alpha \| J \rangle}{\langle L \| \alpha \| L \rangle} = \frac{\eta_J}{\eta_L} = \frac{\langle JJ| \, 3L_z^2 - L(L+1) \, |JJ\rangle}{\langle JJ| \, 3J_z^2 - J(J+1) \, |JJ\rangle} = \frac{L(2L-1)}{J(2J-1)}. \quad (17.19)$$

For any other value of J we find from eqn (16.30)

$$\eta_J = \eta_L (-1)^{L+S+J} \left(\frac{2J+1}{2L+1} \right)^{\frac{1}{2}} \begin{Bmatrix} J & 2 & J \\ L & S & L \end{Bmatrix}. \quad (17.20)$$

We shall defer until later the discussion of the influence of the nuclear electric quadrupole coupling on the paramagnetic resonance spectrum in bulk matter. However we notice that electron operators such as (17.11) are time-even and will therefore have the same expectation value for the two time-conjugate states of a Kramers doublet and will behave as constants as far as paramagnetic resonance is concerned. In cubic symmetry we meet groups of substates whose degeneracy is greater than twofold, and since for cubic symmetry a symmetrical Cartesian form of the quadrupole operator is often convenient, we make here some remarks about spatial symmetry of the operator in connection with the transformation of (17.14) to a symmetrical Cartesian form.

The electronic and nuclear tensor components (eqns (17.11) and (17.9)) contain two types of terms, one of which transforms like Γ_3, that is like $(3z^2 - r^2)$, $(x^2 - y^2)$, and the other like Γ_5, that is like xy, yz, zx, which are of course just like the orbital wave-functions for a d-electron with $l = 2$. Since the nuclear electric quadrupole operator as a whole must remain unchanged under rotations of *both* electric and nuclear variables belonging to the cubic group it follows that nuclear components of Γ_5 symmetry can only be multiplied by electronic components of Γ_5 symmetry, and nuclear Γ_3 components by electronic Γ_3 components.

It can be seen that these distinctions are preserved in the following transformations of the nuclear electric quadrupole operator components to and from simple Cartesian forms:

$$(m/6) \begin{bmatrix} \{3L_x^2 - L(L+1)\}\{3I_x^2 - I(I+1)\} + \\ +\{3L_y^2 - L(L+1)\}\{3I_y^2 - I(I+1)\} + \\ +\{3L_z^2 - L(L+1)\}\{3I_z^2 - I(I+1)\} \end{bmatrix}$$

$$= (m/4)\{3L_z^2 - L(L+1)\}\{3I_z^2 - I(I+1)\} + $$
$$+ (3m/4)(L_x^2 - L_y^2)(I_x^2 - I_y^2) \quad (17.21)$$

$$= (m/4)\{3L_z^2 - L(L+1)\}\{3I_z^2 - I(I+1)\} + $$
$$+ (3m/16)(L_+^2 + L_-^2)(I_+^2 + I_-^2) \quad (17.22)$$

and

$$(3n/4)\begin{bmatrix} (L_xL_y+L_yL_x)(I_xI_y+I_yI_x)+ \\ +(L_yL_z+L_zL_y)(I_yI_z+I_zI_y)+ \\ +(L_zL_x+L_xL_z)(I_zI_x+I_xI_z) \end{bmatrix}$$
$$= (3n/8)\begin{bmatrix} (L_+L_z+L_zL_+)(I_-I_z+I_zI_-)+ \\ +(L_-L_z+L_zL_-)(I_+I_z+I_zI_+) \end{bmatrix}$$
$$-(3n/16)(L_+^2-L_-^2)(I_+^2-I_-^2). \qquad (17.23)$$

These terms contain products $L_+^2I_+^2$, $L_-^2I_-^2$ which are absent in (17.14). However, if we take $m = n = 1$ and sum (17.22) and (17.23) such products vanish and other products occur in just the ratios appropriate to (17.14). Thus we have the symmetrical Cartesian form, equivalent to (17.14),

$$-\frac{e^2Q\langle r_q^{-3}\rangle}{I(2I-1)}\langle L\|\; \alpha\; \|L\rangle \begin{bmatrix} \frac{1}{6}\{3L_x^2-L(L+1)\}\{3I_x^2-I(I+1)\}+ \\ +\frac{1}{6}\{3L_y^2-L(L+1)\}\{3I_y^2-I(I+1)\}+ \\ +\frac{1}{6}\{3L_z^2-L(L+1)\}\{3I_z^2-I(I+1)\}+ \\ +\frac{3}{4}(L_xL_y+L_yL_x)(I_xI_y+I_yI_x)+ \\ +\frac{3}{4}(L_yL_z+L_zL_y)(I_yI_z+I_zI_y)+ \\ +\frac{3}{4}(L_zL_x+L_xL_z)(I_zI_x+I_xI_z) \end{bmatrix}.$$
$$(17.24)$$

This is appropriate to a free atom or ion with full rotational symmetry (which imposes a relation between the coefficients m and n), but with cubic symmetry m and n are not necessarily equal (see § 18.4) in an equivalent operator where the components of L are replaced by those of an effective spin $\tilde{S}$.

In electron paramagnetic resonance experiments we are faced with the difficulty, which exists also for free atoms, of deciding what value to use for the quantity $\langle r_q^{-3}\rangle$. The assumption often made is that it has the same value as the parameter $\langle r^{-3}\rangle$ that appears in the magnetic hyperfine structure and that it can be determined from the latter if the value of the nuclear magnetic moment is known independently. However, distortion of the charge cloud of the inner closed shells of electrons produces an appreciable change in the electric field gradient at the nucleus due to the magnetic electrons, in contrast to the analogous effect ('diamagnetic shielding') in the magnetic case, which is quite small and generally negligible compared to other effects that produce uncertainties in the value of $\langle r^{-3}\rangle$, particularly in solids. We

may formally represent the effect of this distortion by writing $\langle r_q^{-3} \rangle = (1 - R_q) \langle r^{-3} \rangle$, where the contributions to R_q from the various closed shells may be either positive (shielding) or negative (anti-shielding).

Hitherto we have been concerned primarily with the 'internal' electric field gradient due to the asymmetric charge distribution of the electrons 'on' the magnetic ion. There may also be an 'external' electric field gradient arising from the ligand ions and the lattice generally. If the paramagnetic ion is at a site of local cubic symmetry in the lattice, the 'external' electric field gradient vanishes; otherwise, it is related to the terms of second degree V_2^q in the crystal potential, eqn (16.1). The relation is not however a simple one as there are shielding effects, both for the magnetic electrons and for the nucleus, which may be quite different in size from one another as well as from the effect discussed in the previous paragraph. For the nucleus, the external electric field gradient is multiplied by a factor $(1 - \gamma_\infty)$, where γ_∞ is estimated to be large and negative, with values ranging in some cases up to -100. We cannot simply use this factor to multiply the values of V_2^q obtained empirically by fitting the electronic spectra, since any shielding effects for the magnetic electrons are already included in the latter. However, it is clear that a large value of $|\gamma_\infty|$ will greatly enhance the effect of the external electric field gradient at the nucleus, making it comparable with the internal electric field gradient in some cases, and almost certainly the larger of the two for ions in S-states such as Mn^{2+}, $3d^5$ on non-cubic lattice sites. The general question of shielding and anti-shielding, based on the well-known work of Sternheimer, is discussed by Watson and Freeman (1967a), where further references can be found, and also briefly in § 17.7.

17.2. Magnetic hyperfine interactions

A logical procedure would be to develop the theory of the magnetic interactions between the electron and the nucleus along the same lines as for the electrostatic interactions, that is, to assign electric-current densities to the electrons and to the nucleus (rather than charge densities as in the previous section) and to calculate their interactions according to the laws of classical electromagnetism. One would thus define magnetic multipole operators for the nucleus which, like the electric ones, would be tensor operators of integral order k.

If one recalls the opposite parity properties of the electric field (a polar vector) and of the magnetic field (an axial vector) it is understandable that even, rather than odd, values of k are forbidden for

permanent magnetic multipoles, on the assumption of a well-defined parity for the nuclear energy states. The first non-vanishing nuclear magnetic multipole is thus a magnetic dipole, the next a magnetic octopole, etc. Here again the fact that the magnetic field is a time-odd vector leads to the same conclusion because of invariance through time reversal.

Although the existence of magnetic octopoles may be established by atomic-beam methods, they have never been observed by means of magnetic resonance in bulk matter. Furthermore, the description of the magnetic properties of a nucleus as those of a system of currents is more complicated and at the same time, in our present state of knowledge, much less satisfying than the description of its electrostatic properties as those of a system of charges. We shall therefore be content to describe the magnetic properties of the nucleus as those of a magnetic dipole $\mu_I = \gamma_n \hbar \mathbf{I}$ (see however § 17.7). The reason why the magnetic dipole is collinear with the spin vector $\mathbf{I}$ is again that, within the manifold of the substates of a given nuclear state of spin $\mathbf{I}$, all tensor operators of given k (vectors in the present case) have the same matrix elements. Magnetic fields of impossibly high values, of the order of 10^{16} G or more, would have to be applied to the nucleus before its magnetic energy $-\mu_I \cdot \mathbf{H}$ became comparable to the interval between two different nuclear energy states, invalidating the approximation $\mu_I = \gamma_n \hbar \mathbf{I}$.

The interaction of the nuclear dipole μ_I with the electronic shell is small even compared with atomic-energy splittings (let alone the nuclear ones) and may be computed by a perturbation method.

The behaviour of an electron in a magnetic field $\mathbf{H}$ is obtained by replacing the momentum $\mathbf{p}$ by $\mathbf{p} + (e/c)\mathbf{A}$ in its Hamiltonian, where $\mathbf{A}$ is the magnetic vector potential defined by

$$\operatorname{div} \mathbf{A} = 0, \qquad \operatorname{curl} \mathbf{A} = \mathbf{H}.$$

According to classical electromagnetic theory a magnetic dipole μ produces a magnetic field, at a point removed from it by a vector $\mathbf{r}$, deriving from a vector potential

$$\mathbf{A} = \frac{\mu \wedge \mathbf{r}}{r^3} = \operatorname{curl}\left(\frac{\mu}{r}\right). \tag{17.25}$$

Near the dipole the vector potential $\mathbf{A}$ has a singularity of order r^{-2} and $\mathbf{H} = \operatorname{curl} \mathbf{A}$ a singularity of order r^{-3}, so that some care must be exercised in the calculation of its interaction with an electron. In the

non-relativistic Pauli description of the electron the Hamiltonian in the presence of $\mathbf{A}$ is

$$\mathscr{H} = \frac{1}{2m}\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right)^2 + g_s\beta(\mathbf{s} \cdot \operatorname{curl} \mathbf{A}), \qquad (17.26)$$

where β is the Bohr magneton and $\mathbf{s}$ the electron spin. In a first-order perturbation calculation the only terms of (17.26) to be retained are those linear in $\mathbf{A}$:

$$\mathscr{H}_1 = \frac{e}{2mc}(\mathbf{p} \cdot \mathbf{A}+\mathbf{A} \cdot \mathbf{p})+g_s\beta(\mathbf{s} \cdot \operatorname{curl} \mathbf{A}). \qquad (17.26a)$$

This can be written, using (17.25), as

$$\mathscr{H}_1 = 2\beta\frac{(\mathbf{l} \cdot \boldsymbol{\mu})}{r^3}+g_s\beta\left\{\mathbf{s} \cdot \operatorname{curl} \operatorname{curl}\left(\frac{\boldsymbol{\mu}}{r}\right)\right\}, \qquad (17.27)$$

where $\hbar\mathbf{l} = \mathbf{r} \wedge \mathbf{p}$ is the orbital momentum of the electron.

The spin-dependent part of (17.27) gives

$$\mathscr{H}_1^s = g_s\beta\mathbf{s} \cdot \left\{\boldsymbol{\nabla} \wedge \left(\boldsymbol{\nabla} \wedge \frac{\boldsymbol{\mu}}{r}\right)\right\} = g_s\beta\{(\mathbf{s} \cdot \boldsymbol{\nabla})(\boldsymbol{\mu} \cdot \boldsymbol{\nabla})-(\mathbf{s} \cdot \boldsymbol{\mu})\nabla^2\}\frac{1}{r}, \quad (17.28)$$

which for reasons to appear presently we rewrite as

$$\mathscr{H}_1^s = g_s\beta\{(\mathbf{s} \cdot \boldsymbol{\nabla})(\boldsymbol{\mu} \cdot \boldsymbol{\nabla})-\tfrac{1}{3}(\mathbf{s} \cdot \boldsymbol{\mu})\nabla^2\}\left(\frac{1}{r}\right)-\frac{2\beta}{3}g_s(\mathbf{s} \cdot \boldsymbol{\mu})\nabla^2\left(\frac{1}{r}\right). \quad (17.29)$$

The magnetic interaction of the nuclear moment with the electron spin $W_m^s = (\Psi_e| \mathscr{H}_1^s |\Psi_e)$ is obtained by multiplying (17.29) by the electronic density $\rho = \Psi_e^*\Psi_e$ and integrating over the electron coordinates. For $r \neq 0$, $\mathscr{H}_1^s$, as given by (17.29), is a regular function where the first term is equal to $g_s\beta\{3(\mathbf{s} \cdot \mathbf{r})(\boldsymbol{\mu} \cdot \mathbf{r})/r^5-(\mathbf{s} \cdot \boldsymbol{\mu})/r^3\}$, which is the usual dipole-dipole interaction and the second term vanishes because of Laplace's equation. When r tends toward zero we may remark that the first term $\mathscr{H}_1^{s'}$ of (17.29) behaves as a spherical harmonic of order 2 under a rotation of the coordinate system. Hence if Ψ_e is expanded in a sum of spherical harmonics, $\Psi_e = \sum_l a_l\Psi_l$, the only non-vanishing contributions to $(\Psi_e| \mathscr{H}_1^{s'} |\Psi_e)$ will come from terms $(\Psi_l| \mathscr{H}_1^{s'} |\Psi_{l'})$ such that $l+l' \geqslant 2$. It is well known that a wave-function Ψ_l is of order r^l near the origin so that in the matrix element

$$(\Psi_l| \mathscr{H}_1^{s'} |\Psi_{l'}) = \int \Psi_l^* \mathscr{H}_1^s\Psi_{l'}r^2 \, dr \, d\Omega$$

the integrand varies as $r^{(l+l'+2-3)}$ and the corresponding integral always remains finite since $l+l' \geqslant 2$. According to the theory of the Coulomb

potential the second term of (17.29) is equal to $(8g_s/3)\pi\beta(\mathbf{s} \cdot \boldsymbol{\mu}) \, \delta(\mathbf{r})$ and by integration gives

$$\frac{8g_s}{3}\pi\beta(\mathbf{s} \cdot \boldsymbol{\mu}) \, |\Psi_e(0)|^2,$$

which is finite for s electrons and zero for the others. The Hamiltonian for the magnetic interaction of the electron with the nucleus can then be written without ambiguity as (writing $g_s = 2$)

$$\mathscr{H}_1 = 2\beta\gamma_n\hbar\mathbf{I} \cdot \left\{\frac{1}{r^3} - \frac{\mathbf{s}}{r^3} + 3\frac{\mathbf{r}(\mathbf{s} \cdot \mathbf{r})}{r^5} + \frac{8}{3}\pi\mathbf{s} \, \delta(\mathbf{r})\right\}. \tag{17.30}$$

If several electrons surround the nucleus, the interaction Hamiltonian is the sum of the contributions of the individual electrons. Although the expression (17.30) has been derived for the purpose of calculating its expectation value $(\Psi_e| \, \mathscr{H}_1 \, |\Psi_e)$, it is clear that it also gives unambiguous results for off-diagonal matrix elements $(\Psi_e| \, \mathscr{H}_1 \, |\Phi_e)$ between, say, the ground state and an excited state of the electronic system. Use will be made of this to calculate some effects of $\mathscr{H}_1$, using second-order perturbation theory.

We may write the expression in curly brackets in (17.30) simply as a vector $\mathbf{N}$ so that (17.30) becomes

$$\mathscr{H}_1 = 2\beta\gamma_n\hbar(\mathbf{N} \cdot \mathbf{I}). \tag{17.31}$$

This is equivalent to an interaction

$$-\gamma_n\hbar(\mathbf{H}_e \cdot \mathbf{I})$$

where

$$\mathbf{H}_e = -2\beta\langle r^{-3}\rangle\mathbf{N} \tag{17.32}$$

is the electronic magnetic field at the nucleus.

17.3. Alternative derivation of the magnetic hyperfine interaction

There is a weak point in our derivation of (17.30), namely, the statement that the first part of (17.29) $g_s\beta\{(\mathbf{s} \cdot \boldsymbol{\nabla})(\boldsymbol{\mu} \cdot \boldsymbol{\nabla}) - \frac{1}{3}(\mathbf{s} \cdot \boldsymbol{\mu})\nabla^2\}(1/r)$ gives a vanishing result when the second term $-(2g_s\beta/3)(\mathbf{s} \cdot \boldsymbol{\mu})\nabla^2(1/r)$ does not; that is, for an s-electron. This is based on an angular integration, performed before a radial integration. If we were to interchange these operations we would find an infinite result instead. An alternative derivation based on the relativistic equation of Dirac does not have this drawback.

In a magnetic field Dirac's equation can be written

$$\{c\boldsymbol{\alpha} \cdot (\mathbf{p}+e\mathbf{A}/c)+\beta mc^2-eV\}\Psi = (W+mc^2)\Psi. \qquad (17.33)$$

In this equation $\boldsymbol{\alpha}$ stands for the three Dirac 4 by 4 matrices α_1, α_2, α_3; β is the fourth Dirac matrix (not to be confused with the Bohr magneton); W is the kinetic energy of the electron and Ψ the four-component spinor which is the electron wave-function in Dirac theory.

The spinor Ψ can be written as $\begin{pmatrix}\varphi\\\chi\end{pmatrix}$ where φ and χ are 2 two-component spinors describing what is known as the large and small components of the Dirac wave-function. If we neglect $e\boldsymbol{\alpha} \cdot \mathbf{A}$, which is a small perturbation, it is known from the structure of the matrices α_i and β that (17.33) can be rewritten:

$$\begin{aligned}c(\boldsymbol{\sigma} \cdot \mathbf{p})\chi &= (W+eV)\varphi,\\c(\boldsymbol{\sigma} \cdot \mathbf{p})\varphi &= (W+2mc^2+eV)\chi,\end{aligned} \qquad (17.34)$$

where $\boldsymbol{\sigma}$ is the set of the three Pauli matrices. Neglecting $W+eV$ in comparison with $2mc^2$ we get

$$\chi \cong \frac{1}{2mc}(\boldsymbol{\sigma} \cdot \mathbf{p})\varphi. \qquad (17.35)$$

We may treat the coupling $e(\boldsymbol{\alpha} \cdot \mathbf{A})$ with the magnetic field as a perturbation and proceed to calculate its expectation value:

$$\langle\Psi| e(\boldsymbol{\alpha} \cdot \mathbf{A}) |\Psi\rangle = \langle\varphi| e(\boldsymbol{\sigma} \cdot \mathbf{A}) |\chi\rangle+\langle\chi| e(\boldsymbol{\sigma} \cdot \mathbf{A}) |\varphi\rangle$$

$$\approx \frac{e}{2mc}\langle\varphi| (\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{p})+(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{A}) |\varphi\rangle. \qquad (17.36)$$

Using the commutation relations of the components of $\boldsymbol{\sigma}$ we find that (17.36) becomes

$$\frac{e}{2mc}\langle\varphi| (\mathbf{p} \cdot \mathbf{A})+(\mathbf{A} \cdot \mathbf{p})+i\boldsymbol{\sigma} \cdot (\mathbf{A} \wedge \mathbf{p}+\mathbf{p} \wedge \mathbf{A}) |\varphi\rangle \qquad (17.37)$$

and from the relation $\mathbf{p} = (\hbar/i) \nabla$ we get

$$i(\mathbf{A} \wedge \mathbf{p}+\mathbf{p} \wedge \mathbf{A}) = \hbar \,\mathbf{curl} \cdot \mathbf{A}$$

whence

$$\langle\Psi| e(\boldsymbol{\alpha} \cdot \mathbf{A}) |\Psi\rangle = \langle\varphi| \frac{e}{2mc}\{(\mathbf{p} \cdot \mathbf{A})+(\mathbf{A} \cdot \mathbf{p})\}+\beta\boldsymbol{\sigma} \cdot \mathbf{curl}\,\mathbf{A} |\varphi\rangle, \qquad (17.38)$$

which was precisely our starting point (eqn (17.26a)) for the non-relativistic derivation above, provided that we take $g_s = 2$.

The operation $\langle\varphi| \quad |\varphi\rangle$ corresponds to an integration over orbital and spin variables. In the Pauli approximation $|\varphi\rangle$ is the product of the Schrödinger orbital wave-function Ψ_e times a spin function (or a sum of such products). Leaving out the integration over spin variables, the second part of (17.37) is a spin-operator

$$i\frac{\beta}{\hbar}\boldsymbol{\sigma} \, . \, \langle\Psi_e|\, \mathbf{A} \wedge \mathbf{p} + \mathbf{p} \wedge \mathbf{A}\, |\Psi_e\rangle, \tag{17.39}$$

which, since $\mathbf{p}$ is a Hermitian operator, can be rewritten in terms of the density $\rho = \varphi_e^*\varphi_e$ as

$$\beta\boldsymbol{\sigma} \, . \, \int(\nabla\rho \wedge \mathbf{A})\, \mathrm{d}\tau = \beta\int\mathrm{curl}\,\{(\rho\boldsymbol{\sigma}) \, . \, \mathbf{A}\}\, \mathrm{d}\tau. \tag{17.40}$$

If one assigns to the electron a spin current density

$$\mathbf{j}_s = \beta\,\mathrm{curl}\,(\rho\boldsymbol{\sigma}), \tag{17.41}$$

we can write (17.40) in the form

$$\int(\mathbf{j}_s \, . \, \mathbf{A})\, \mathrm{d}\tau. \tag{17.42}$$

The expression (17.40) is equivalent to our starting point (17.38) but has the advantage that if we replace $\mathbf{A}$ by $\mathrm{curl}\,(\boldsymbol{\mu}/r)$ the integral converges absolutely; the singularity of the integrand is of order $1/r^2$ rather than $1/r^3$ and for an s-electron it leads unambiguously to the last term of (17.30).

17.4. Equivalent operators for the magnetic hyperfine structure

The dipole-dipole coupling $\{3(\mathbf{I} \, . \, \mathbf{r})(\mathbf{s} \, . \, \mathbf{r}) - r^2(\mathbf{I} \, . \, \mathbf{s})\}/r^5$ is a product of two second-rank tensors and can be rewritten as

$$\sum_{j,k}\{I_j s_k - \tfrac{1}{3}\delta_{jk}(\mathbf{I} \, . \, \mathbf{s})\}(3x_j x_k - r^2\delta_{jk})r^{-5}. \tag{17.43}$$

By a process similar to that used in § 17.1 in deriving the nuclear electric quadrupole interaction, this is found for a single electron l to be equivalent to

$$\langle r^{-3}\rangle\frac{2}{(2l+3)(2l-1)}\{l(l+1)(\mathbf{I} \, . \, \mathbf{s}) - \tfrac{3}{2}(\mathbf{l} \, . \, \mathbf{I})(\mathbf{l} \, . \, \mathbf{s}) - \tfrac{3}{2}(\mathbf{l} \, . \, \mathbf{s})(\mathbf{l} \, . \, \mathbf{I})\}. \tag{17.44}$$

Similarly for a term (L, S) obeying Hund's rule we obtain

$$\sum_i\frac{3(\mathbf{I} \, . \, \mathbf{r}_i)(\mathbf{s}_i \, . \, \mathbf{r}_i) - r_i^2(\mathbf{I} \, . \, \mathbf{s}_i)}{r_i^5}$$
$$= \xi\langle r^{-3}\rangle\{L(L+1)(\mathbf{I} \, . \, \mathbf{S}) - \tfrac{3}{2}(\mathbf{L} \, . \, \mathbf{I})(\mathbf{L} \, . \, \mathbf{S}) - \tfrac{3}{2}(\mathbf{L} \, . \, \mathbf{S})(\mathbf{L} \, . \, \mathbf{I})\}, \tag{17.45}$$

in which

$$\xi = \frac{2l+1-4S}{S(2l-1)(2l+3)(2L-1)} = \pm\frac{\eta_L}{2S} = \mp\frac{1}{2S}\langle L\| \alpha \|L\rangle, \quad (17.46)$$

where η_L is given by (17.16) and $\langle L\| \alpha \|L\rangle$ by (16.25).

Since the operator (17.43) is time-odd with respect to electron variables, ξ has the same sign for two complementary configurations of x and $2(2l+1)-x$ electrons, in contrast to η_L or $\langle L\| \alpha \|L\rangle$ which were introduced in dealing with interactions that are time-even with respect to the electron variables.

Within a ground term (L, S) of electrons for which $l \neq 0$ the magnetic hyperfine coupling can then be written

$$2\beta\gamma_n\hbar\langle r^{-3}\rangle\{(\mathbf{L} \cdot \mathbf{I})+\xi L(L+1)(\mathbf{I} \cdot \mathbf{S})-\tfrac{3}{2}\xi(\mathbf{L} \cdot \mathbf{I})(\mathbf{L} \cdot \mathbf{S})-\tfrac{3}{2}\xi(\mathbf{L} \cdot \mathbf{S})(\mathbf{L} \cdot \mathbf{I})\}.$$
$$(17.47)$$

We now construct an equivalent operator for the hfs interaction within the manifold (J, L, S). We define the vector $\mathbf{N}$ by

$$\mathbf{N} = \sum_i \left\{\mathbf{1}_i-\mathbf{s}_i+3\frac{(\mathbf{r}_i \cdot \mathbf{s}_i)}{r_i^2}\mathbf{r}_i\right\}, \quad (17.48)$$

which is equivalent to (17.31) except that we have assumed the term $\delta(r)$ in (17.30) to be absent. Inside the manifold (L, S) this is equivalent to, using (17.47),

$$\langle LS\| \mathbf{N} \|LS\rangle = \mathbf{L}+\xi L(L+1)\mathbf{S}-\tfrac{3}{2}\xi\{\mathbf{L}(\mathbf{L} \cdot \mathbf{S})+(\mathbf{L} \cdot \mathbf{S})\mathbf{L}\}. \quad (17.49)$$

Within the manifold (J, L, S) we can replace $\mathbf{N}$ by a vector collinear with $\mathbf{J}$

$$\frac{(\mathbf{N} \cdot \mathbf{J})\mathbf{J}}{J(J+1)} = \langle J\| N \|J\rangle\mathbf{J}, \quad (17.50)$$

where

$$\langle J\| N \|J\rangle = \frac{1}{J(J+1)}\{(\mathbf{L} \cdot \mathbf{J})+\xi L(L+1)(\mathbf{S} \cdot \mathbf{J})-3\xi(\mathbf{L} \cdot \mathbf{J})(\mathbf{L} \cdot \mathbf{S})\},$$
$$(17.51)$$

in which

$$(\mathbf{L} \cdot \mathbf{J}) = \tfrac{1}{2}\{J(J+1)+L(L+1)-S(S+1)\},$$
$$(\mathbf{S} \cdot \mathbf{J}) = \tfrac{1}{2}\{J(J+1)+S(S+1)-L(L+1)\}, \quad (17.52)$$
$$(\mathbf{L} \cdot \mathbf{S}) = \tfrac{1}{2}\{J(J+1)-L(L+1)-S(S+1)\}.$$

Here it may be useful to note that

$$(\mathbf{L} \cdot \mathbf{J})/J(J+1) = 2 - \langle J \| \Lambda \| J \rangle,$$
$$(\mathbf{S} \cdot \mathbf{J})/J(J+1) = \langle J \| \Lambda \| J \rangle - 1,$$
(17.53)

and that the form of the magnetic hyperfine interaction is now simply

$$2\beta\gamma_n\hbar\langle r^{-3}\rangle\langle J \| N \| J \rangle(\mathbf{J} \cdot \mathbf{I}).$$
(17.54)

The values of $\langle J \| N \| J \rangle$ are tabulated for the rare-earth group in Table 20.

For a single electron with $j = l \pm \frac{1}{2}$ a direct derivation leads to a much simpler result for (17.51) since

$$(\mathbf{j} \cdot \mathbf{N}) = \mathbf{j} \cdot \left\{1 - \mathbf{s} + 3\frac{(\mathbf{r} \cdot \mathbf{s})}{r^2}\mathbf{r}\right\} = \mathbf{l} \cdot \left\{1 - \mathbf{s} + 3\frac{(\mathbf{r} \cdot \mathbf{s})}{r^2}\mathbf{r}\right\} + \mathbf{s} \cdot \left\{1 - \mathbf{s} + 3\frac{(\mathbf{r} \cdot \mathbf{s})}{r^2}\mathbf{r}\right\}$$

$$= l(l+1) - s(s+1) + 3\frac{(\mathbf{r} \cdot \mathbf{s})^2}{r^2} = l(l+1),$$

where we have used the fact that $(\mathbf{r} \cdot \mathbf{l})$ is proportional to $\mathbf{r} \cdot (\mathbf{r} \wedge \mathbf{p}) = 0$. Hence

$$\langle j \| N \| j \rangle = \frac{l(l+1)}{j(j+1)}.$$
(17.55)

To find the off-diagonal elements $\langle J'M' | \mathbf{N} | JM \rangle$ of the vector $\mathbf{N}$ between two manifolds J', J belonging to the same (L, S) multiplet we make use of the fact that, from eqn (17.49), the matrix elements of $\mathbf{N}$ are a linear combination of those of $\mathbf{L}$ and $\mathbf{S}$. It then follows from (16.38) that the off-diagonal elements of $\mathbf{N}$ are proportional to those of $\mathbf{L}+2\mathbf{S}$, and vanish except for $|J'-J| = 1$. Introducing a symbol $\langle J+1 \| N \| J \rangle$, we can therefore write

$$\langle J+1, M' | \mathbf{N} | J, M \rangle = \frac{\langle J+1 \| N \| J \rangle}{\langle J+1 \| \Lambda \| J \rangle}\langle J+1, M' | \mathbf{L}+2\mathbf{S} | J, M \rangle$$
(17.56)

and from (17.49) and (16.38) we have

$$\frac{\langle J+1 \| N \| J \rangle}{\langle J+1 \| \Lambda \| J \rangle}$$
$$= -1 + \xi L(L+1) + \tfrac{3}{2}\xi\{\langle J | (\mathbf{L} \cdot \mathbf{S}) | J \rangle + \langle J+1 | (\mathbf{L} \cdot \mathbf{S}) | J+1 \rangle\},$$
(17.57)

where ξ (which is equal to $-\nu$ in the notation of Elliott and Stevens (1953a)) is given by eqn (17.46) and $\langle J | (\mathbf{L} \cdot \mathbf{S}) | J \rangle$ by eqn (17.52).

Values of ξ and $\langle J+1 \| N \| J \rangle$, as well as of $\langle J+1 \| \Lambda \| J \rangle$ (eqn (16.41)), are listed for the rare-earth group in Table 20. The off-diagonal elements of $\mathbf{L} + 2\mathbf{S}$ are given by eqns (16.39), (16.40), whence we have

$$\langle J+1, J_z| N_z |J, J_z \rangle = \langle J+1 \| N \| J \rangle \{(J+1)^2 - J_z^2\}^{\frac{1}{2}}, \tag{17.58}$$

$$\langle J+1, J_z \pm 1| N_x |J, J_z \rangle = \mp \langle J+1 \| N \| J \rangle \tfrac{1}{2} \{(J \pm J_z + 1)(J \pm J_z + 2)\}^{\frac{1}{2}}. \tag{17.59}$$

17.5. The effect of s-electrons: configuration interaction

We have already commented on the fact that, to calculate the coupling between an atom (or an ion) and a homogeneous applied field, a detailed knowledge of the atomic wave-function is not required (§ 11.3); in particular, for a term (L, S) it is not necessary to know in what configuration this term originated. The position is quite different for the coupling with the very inhomogeneous magnetic field produced by a nuclear moment; a formula such as (17.45) rests on the explicit assumption that the term (L, S) is built from a specific configuration of i electrons each with angular momentum l.

In our calculations of the effects of the crystal field we have assumed that the configuration is a good quantum number and on the basis of this assumption we would have expected eqn (17.47) to give reasonable agreement with experiment. In fact in the important case of the iron group the agreement is miserable. The explanation resides in the fact that s-electrons, especially those of low principal quantum number n, have hyperfine couplings very much stronger than those of other electrons with $l \neq 0$. It follows that if a term (L, S) belonging to an ion of the transition group, whose ground configuration has no unpaired s-electrons, is 'contaminated' even slightly with excited configurations containing unpaired s-electrons (for brevity s-configurations), there may be an appreciable change in the hyperfine coupling although all other properties of the term remain practically unaffected. If there is an admixture from s-configurations the contribution of the so-called contact term, the last term of (17.30), will not be zero; within the manifold of the term (L, S), this contribution can be represented quite generally as $A_s \mathbf{I} \cdot \mathbf{S}$, the value of the constant A_s depending on the nature and amount of s-configuration present. It is customary (though perhaps unfortunate) to define a dimensionless constant κ through the relation

$$A_s = -2\gamma_n \beta \hbar \langle r^{-3} \rangle \kappa. \tag{17.60}$$

To sum up, within a term (L, S) the total hyperfine structure, electric and magnetic, can be written

$$W_n = 2\beta\gamma_n\hbar\langle r^{-3}\rangle\{(\mathbf{L} \cdot \mathbf{I}) + \{\xi L(L+1) - \kappa\}(\mathbf{S} \cdot \mathbf{I})\} -$$
$$- \tfrac{3}{2}\xi\{(\mathbf{L} \cdot \mathbf{S})(\mathbf{L} \cdot \mathbf{I}) + (\mathbf{L} \cdot \mathbf{I})(\mathbf{L} \cdot \mathbf{S})\} -$$
$$- \langle L\| \alpha \|L\rangle \frac{e^2Q}{2I(2I-1)} \langle r^{-3}\rangle\{3(\mathbf{L} \cdot \mathbf{I})^2 + \tfrac{3}{2}(\mathbf{L} \cdot \mathbf{I}) - L(L+1)I(I+1)\}.$$

$$(17.61)$$

We examine below and in § 17.6 methods by which theoretical estimates can be made of the magnitude of the constant κ.

The original idea that configuration interaction is responsible for anomalies in atomic magnetic hyperfine structure is due to Fermi and Segré (1933a, b). Among the examples they studied, the most relevant for our purpose, as will appear shortly, is the $6s^26p$, ^2P ground term of atomic thallium perturbed by the excited ^2P term $6s6p7s$. Fermi and Segré used a perturbation method to estimate the small admixture α from the excited term into the ground manifold; the unperturbed one-electron wave-functions in each configuration were eigenstates of an approximate central-field Hamiltonian, the matrix elements of the Coulomb repulsion $V = \sum_{i<k} e^2/r_{ik}$ between the two terms were computed numerically, and the energy separation Δ_{10} between the terms was taken from spectroscopic data.

Let Ψ_0 and Ψ_1 be atomic states, belonging respectively to the ground and to the excited configuration, and linked by the electronic repulsion V, which in first order changes Ψ_0 into $\Psi = \Psi_0 + \alpha\Psi_1$, where

$$\alpha \approx -\frac{\langle\Psi_0| V |\Psi_1\rangle}{\Delta_{10}}. \qquad (17.62)$$

Since the coupling Hamiltonian V is invariant through spatial rotation and does not contain the spin variables it is clear that the states Ψ_0 and Ψ_1 must have the same quantum numbers L, S, M_L, M_S.

For our purpose the essential feature of the excited configuration $6s6p7s$ is that it is obtained from the ground configuration $6s^26p$ by promotion of one electron from an s-orbit φ_{6s} to another s-orbit φ_{7s}. Therefore the contact part $\mathscr{H}_c$ of the hyperfine Hamiltonian does have a non-vanishing matrix element between Ψ_0 and Ψ_1, and according to the rules given in § 11.6 this is proportional to a one-electron matrix element $(\varphi_{6s}| \delta(\mathbf{r}) |\varphi_{7s})$, that is, to the product $\varphi_6(0) \cdot \varphi_7(0)$. The expectation value $\langle\Psi| \mathscr{H}_c |\Psi\rangle$ of the contact hyperfine interaction in

the state Ψ' given by (17.62) contains therefore a term proportional to $2\alpha\varphi_{6s}(0)\varphi_{7s}(0)$. This linear dependence on α of $\langle\Psi|\,\mathscr{H}_c\,|\Psi\rangle$ has two far-reaching consequences: first, as α is small, the correction to the hyperfine structure is very much larger than the second-order term $\alpha^2\langle\Psi_1|\,\mathscr{H}_c\,|\Psi_1\rangle$, which is proportional to $\alpha^2\varphi_{7s}^2(0)$; secondly, this correction can have either sign, again in contrast to the second-order term.

This feature proves to be of paramount importance in the explanation, that now follows, of anomalous hyperfine structure in the paramagnetic resonance spectrum of ions belonging to the iron group (Abragam 1950; Abragam and Pryce 1951a, b, c; Abragam, Horowitz, and Pryce 1955). The conventional configuration assigned to the ions of the iron group in their ground state is

$$\mathscr{C}_0 = 1s^2 2s^2 2p^6 3s^2 3p^6 3d^x. \tag{17.63}$$

In the ground term, constructed from this configuration according to Hund's rule, the total spin S is either $x/2$ or $5-(x/2)$ depending on whether $x < 5$ or $x > 5$. This configuration does not allow for a term such as $A_s\mathbf{I}\,.\,\mathbf{S} = -2\gamma_n\beta\hbar\langle r^{-3}\rangle\kappa\mathbf{I}\,.\,\mathbf{S}$ (eqn (17.60)) in the hyperfine Hamiltonian (17.61); however such a term is found to be necessary in order to fit the experimental data in the iron group (see Chapter 7). The sign of the coefficient κ is found to be positive, that is such as to make the electronic hyperfine field $\mathbf{H}_e$ at the nucleus parallel to the total ionic spin $\mathbf{S}$. In contrast the field $\mathbf{H}_e$ produced at the nucleus by a single s-electron is antiparallel to its spin.

It is convenient to introduce a quantity χ characteristic of the density of unpaired spins at the nucleus:

$$\chi = \frac{4\pi}{S}\Big\langle\sum_k \delta(\mathbf{r}_k)s_{kz}\Big\rangle_{S_z=S}. \tag{17.64}$$

The expectation value in (17.64) is to be taken over the state $S_z = S$ of the ground term. According to (17.30) and (17.60) χ is related to κ by

$$\kappa\langle r^{-3}\rangle = -(\tfrac{2}{3})\chi \tag{17.65}$$

so that χ must be negative if κ is positive.

For each ion of the iron group χ is found to depend little on the surroundings of the ion, which makes it plausible that it is a feature of the free ion. χ is also found experimentally to vary little (by no more than 25 per cent) inside the iron group, the mean value being of the order of -3 in atomic units, a fact for which there is no simple physical

explanation. The proportionality coefficient between χ, expressed in atomic units, and the electronic field H_e at the nucleus, expressed in gauss, is given by $H_e = -2S \times 4\cdot21 \times 10^4\chi$.

In contrast with the hyperfine structure the other features of the spectrum such as the g-factor and fine structure agree well with the ground configuration (17.63). It is therefore reasonable to look for small admixtures into this configuration that contain unpaired s-electrons. In order for such admixtures to be able to modify appreciably the hyperfine structure, while leaving practically unaffected the other features of the spectrum, and also to produce hyperfine fields $\mathbf{H_e}$ of the right sign, they must differ from the ground configuration through promotion of one electron from an s-orbit into another s-orbit, as explained in connection with the work of Fermi and Segré. Let us denote by $\mathscr{C}_3$ one such configuration, in which a $3s$ electron has been promoted,

$$\mathscr{C}_3 = 1s^2 2s^2 2p^6 3s Rs 3p^6 3d^x = 3s Rs(\mathscr{C}_0 - 3s^2). \qquad (17.66)$$

In (17.66) Rs means an s-orbital that is orthogonal to all the s-orbitals of $\mathscr{C}_0$, $1s$, $2s$, $3s$ (it is automatically orthogonal to p- and d-orbitals), and at this point our treatment departs from that of Fermi and Segré. In the latter, the excited s-orbital $7s$ is a physical orbital of the thallium atom; states belonging to the excited configuration $6s6p7s$ have actually been observed and their energies have been measured by the methods of optical spectroscopy. In a configuration such as $\mathscr{C}_3$, given by (17.66), it is not clear what meaning one should attach to the s-orbit Rs. In order to elucidate this point let us return to the ground configuration $\mathscr{C}_0$ and recall how the orbits $1s$ to $3d$ of which it is composed are obtained in the variational Hartree–Fock method. As explained in § 11.4 we build from these orbits, each of which has the usual form

$$\varphi_{n,l,m_l,m_s} = P_{nl}(r) Y_l^{m_l}(\theta, \varphi)\chi(m_s), \qquad (17.67)$$

a many-electron wave-function Ψ_0 which has the required symmetry and quantum numbers L, S, M_L, M_S. Terms obeying Hund's rule have a simpler structure than the others: the magnetic substate of such a term with $M_L = L$, $M_S = S$ (sometimes called the 'stretched' state) can be written as a single Slater determinant. Then we write down the expectation value in the state Ψ_0 of an electronic Hamiltonian $\mathscr{H}$ which is the sum of the kinetic energies of the electrons and of their potential energy due to their attraction by the nucleus and to their mutual repulsion (smaller magnetic interactions are neglected at this stage). The expression $\langle\Psi_0|\mathscr{H}|\Psi_0\rangle$ is expanded into a sum of one-electron and

two-electron matrix elements by the rules of § 11.6, and minimized with respect to the trial radial functions $P_{nl}(r)$ which are considered as unknown. The generalization whereby the configuration $\mathscr{C}_3$ of (17.63) is introduced and the orbit Rs as defined is now straightforward: as a trial many-electron wave-function we use the linear combination

$$\Psi = \cos \alpha\, \Psi_0 + \sin \alpha\, \Psi_3, \tag{17.68}$$

where Ψ_3 is a wave-function belonging to $\mathscr{C}_3$ and having the same quantum numbers L, S, M_L, M_S as Ψ_0. The latter, being the 'stretched' substate of a Hund's term, is a single Slater determinant. (We have tacitly assumed that the configuration $\mathscr{C}_3$ contains only one state Ψ_3 with the right quantum numbers; we shall return to this point later.) We formulate again the expression $\langle\Psi|\,\mathscr{H}\,|\Psi\rangle$ and minimize it with respect to the functions $P_{nl}(r)$ and also with respect to the new trial function Rs, which appears in the configuration $\mathscr{C}_3$, and with respect to the normalizing variable α. Since the set of trial functions Ψ is wider than that of the functions Ψ_0, which it contains, it must be better than or at least as good as the latter for the calculation of every physical parameter. The parameter of interest to us is the hyperfine structure and this is naturally the reason why we have augmented the set of the trial functions by adding to it components to which the hyperfine structure is most sensitive.

We now write out the wave-function Ψ_3 which has the same quantum numbers as Ψ_0. Since $\mathscr{C}_3$ results from $\mathscr{C}_0$ through promotion of an electron from one s-orbit to another it is clear that Ψ_3 will have the same M_L and L as Ψ_0, provided each electron has the same m_l in Ψ_3 and Ψ_0. Let us write the configuration $\mathscr{C}_0$ as P^2Q and the stretched state Ψ_0 as

$$\Psi_0 = P^+P^-Q^S, \tag{17.69}$$

where P is the orbital $3s$, Q stands for $3d^x$, and Q^S is the Slater determinant formed from the $3d^x$ electrons with all their spins up, thus adding to a total spin $S_z = S$; the orbitals belonging to the other closed shells of $\mathscr{C}_0$ are omitted in this shorthand notation. Ψ_0 itself is a Slater determinant so that

$$P^+P^-Q^S = -P^-P^+Q^S = (-1)^x P^+Q^S P^-. \tag{17.70}$$

We can similarly define

$$Q^{S-1} = \frac{1}{\sqrt{(2S)}}(S_x - iS_y)Q^S \tag{17.71}$$

as the state of spin S with $M_S = S-1$.

The configuration $\mathscr{C}_3$ can be written as PRQ and two states with total spin S and $M_S = S$ can be formed from it: Ψ_{3a} where the spins of the electrons Ps and Rs couple to zero and which is obviously of the form

$$\Psi_{3a} = \frac{1}{\sqrt{2}}(P^+R^- - P^-R^+)Q^S \qquad (17.72)$$

and the state where the spins of P and R couple to a spin $S' = 1$ which in turn couples to the spin S of Q to give a spin S again. This state can be written

$$\Psi_{3b} = \frac{\sqrt{(2S)}}{2\sqrt{(S+1)}}(P^+R^-Q^S + P^-R^+Q^S - \frac{2}{\sqrt{(2S)}}P^+R^+Q^{S-1}), \quad (17.73)$$

which is justified by the fact that, from (17.72), (17.73),

$$\langle \Psi_{3b} \mid \Psi_{3a} \rangle = 0; \qquad S_z\Psi_{3b} = S\Psi_{3b}; \qquad S_+\Psi_{3b} = 0. \quad (17.74)$$

We can now write the expression to be minimized as

$$\langle \Psi | \mathscr{H} | \Psi \rangle = \cos^2\alpha\langle \Psi_0 | \mathscr{H} | \Psi_0 \rangle +$$
$$+ \sin 2\alpha\langle \Psi_0 | \mathscr{H} | \Psi_3 \rangle + \sin^2\alpha\langle \Psi_3 | \mathscr{H} | \Psi_3 \rangle, \quad (17.75)$$

where Ψ_0 and $\Psi_3 = \Psi_{3b}$ are given in shorthand notation by (17.69) and (17.73). (We shall show shortly that the state Ψ_{3a} given by (17.72) is not coupled to Ψ_0.) Minimizing (17.75) with respect to α, we find

$$\tan 2\alpha = -2\frac{\langle \Psi_0 | \mathscr{H} | \Psi_3 \rangle}{\langle \Psi_3 | \mathscr{H} | \Psi_3 \rangle - \langle \Psi_0 | \mathscr{H} | \Psi_0 \rangle}. \quad (17.76)$$

If we use our assumption that the admixture from the excited configuration is small, (17.76) becomes, in a good approximation,

$$\alpha = -\frac{\langle \Psi_0 | \mathscr{H} | \Psi_3 \rangle}{\langle \Psi_3 | \mathscr{H} | \Psi_3 \rangle - \langle \Psi_0 | \mathscr{H} | \Psi_0 \rangle}, \quad (17.77)$$

which resembles the perturbation formula (17.62). It remains to vary (17.75) with respect to all the radial functions P_{nl} of $\mathscr{C}_0$ and with respect to Rs. Since the admixture α is very small, it is a good approximation to take for the P_{nl} the unperturbed functions of the configuration $\mathscr{C}_0$, that is the functions that minimized $\langle \Psi_0 | \mathscr{H} | \Psi_0 \rangle$. This enables us to prove the statement that the state Ψ_{3a} given by (17.72) is not coupled to Ψ_0 in first order in α; using (17.69) and (17.72) we can rewrite $\Psi_0 + \alpha\Psi_{3a}$, where α is an arbitrary small quantity, as

$$\Psi = \Psi_0 + \alpha\Psi_{3a} = \left(P^+ + \frac{\alpha}{\sqrt{2}}R^+, \ P^- + \frac{\alpha}{\sqrt{2}}R^-\right)Q^S + O(\alpha^2). \quad (17.78)$$

This equation shows that as far as terms $O(\alpha^2)$, Ψ results from Ψ_0 by giving a first-order increase $\alpha R/\sqrt{2}$ to P. Since, however, the functions P have been determined precisely by the condition that they should make $\langle \Psi_0| \mathscr{H} |\Psi_0 \rangle$ a minimum, it follows that with Ψ given by (17.78)

$$\langle \Psi| \mathscr{H} |\Psi \rangle = \langle \Psi_0| \mathscr{H} |\Psi_0 \rangle + O(\alpha^2). \tag{17.79}$$

Since we also have from (17.78)

$$\langle \Psi| \mathscr{H} |\Psi \rangle = \langle \Psi_0| \mathscr{H} |\Psi_0 \rangle + 2\alpha \langle \Psi_0| \mathscr{H} |\Psi_{3a} \rangle + O(\alpha^2), \tag{17.80}$$

we must have

$$\langle \Psi_0| \mathscr{H} |\Psi_{3a} \rangle = 0 \tag{17.81}$$

and, according to (17.77), the value of α that minimizes $\langle \Psi| \mathscr{H} |\Psi \rangle$ and gives the amplitude of Ψ_{3a} in Ψ vanishes. If we now minimize (17.75) with respect to Rs we get for this function an integro-differential equation that must be integrated numerically. We shall not write out its expression which is lengthy (see Abragam, Horowitz, and Pryce (1955) where slightly different notations are used). Once this is done and Rs is known we can calculate χ using eqns (17.64), (17.68), and (17.73) and we find, neglecting terms of order $O(\alpha^2)$,

$$\chi = \frac{4\pi}{S}\langle \Psi| \sum_k \delta(\mathbf{r}_k)s_{zk} |\Psi \rangle \approx -2\alpha \frac{4\pi}{S}\langle \Psi_0| \sum_k \delta(\mathbf{r}_k)s_{zk} |\Psi_{3b} \rangle$$

$$= 4\pi\alpha \Big/ \sqrt{\left\{\frac{2}{S(S+1)}\right\}} \varphi_{3s}(0)\varphi_{Rs}(0), \tag{17.82}$$

or if we write

$$\varphi_{3s} = \frac{P_{3s}(r)}{r}\frac{1}{\sqrt{(4\pi)}}, \; \varphi_{Rs} = \frac{Rs(r)}{r}\frac{1}{\sqrt{(4\pi)}}, \tag{17.83}$$

$$\chi = -\alpha \Big/ \sqrt{\left\{\frac{2}{S(S+1)}\right\}} P'_{1s}(0)R's(0). \tag{17.84}$$

It was hoped originally (Abragam and Pryce 1951a) that the bulk of the anomalous hyperfine structure in the iron group was due to the configuration $\mathscr{C}_3$ and a numerical calculation of χ as given by (17.84) was attempted for Mn^{2+} (Abragam, Horowitz, and Pryce 1955) leading to a value of $\chi = -0.3 \cdot 10^{-3}$ in atomic units (instead of the experimental value $\chi \approx -3$). This dismal failure is due to three facts.

(i) The unperturbed wave-functions P_{nl} for the configuration $\mathscr{C}_0$ of Mn^{2+} were not available at the time of the calculation and those of the cuprous ion Cu^+ for the configuration $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10}$ were used

instead. It turns out that in the equation satisfied by Rs there occur overlap integrals of the type $\langle 3s, 3d| \ 1/r_{12} \ |3d, Rs\rangle$ with large positive and negative parts which nearly cancel each other, making the integrals very sensitive to small changes in the wave-function. In consequence, large errors may result through starting from incorrect unperturbed wave-functions.

(ii) The assumption that the configuration $\mathscr{C}_3$ is the only one to contribute appreciably to the anomalous hyperfine structure is unwarranted, as will appear shortly. Instead of the set of functions Ψ given by (17.63) a wider set should be used in the expression $\langle\Psi| \ \mathscr{H} \ |\Psi\rangle$ to be minimized, namely,

$$\Psi = \nu_0\Psi_0 + \sum_{i=1}^{3} \nu_i\Psi_i. \tag{17.85}$$

In this equation Ψ_i belongs to a configuration $\mathscr{C}_i$ where an s-electron from the shell i is promoted. Such a calculation was not attempted through lack of adequate computing facilities.

(iii) Last but not least, it is not impossible that a numerical error was made in the calculation by those in charge of the numerical computation (W. Marshall, private communication). All further calculations of anomalous hyperfine structure were made using the so-called core polarization method to be described now.

17.6. The effect of s-electrons: core polarization

The idea of magnetic polarization of an atomic or ionic core of closed shells by an unfilled external shell with a total spin S goes back to the pioneer work of Sternheimer (1951). This idea is physically much easier to grasp than the configuration interaction method outlined in the previous section: let us consider a paramagnetic ion, say Mn^{2+}, in the state $S_z = S = +\frac{5}{2}$ where the spins of its five $3d$ electrons are all 'up'. The electrostatic repulsion between an s-electron of an inner shell, with spin up, say $1s^+$, and these $3d$ electrons will not be the same as that for an electron in the state $1s^-$ (for brevity, in the following, spin up or $+$ will mean spin parallel to the total spin of the ion). This is a consequence of the exclusion principle that prevents two electrons with parallel spins from occupying the same position in space. Mathematically this principle is built into the formalism by using wave-functions that are antisymmetric with respect to interchange of both spin and orbital coordinates of any two electrons. For wave-functions approximated by Slater determinants this results in the vanishing of the exchange matrix

elements of the electrostatic repulsion between electrons of opposite spins, such as $\langle 3s^-(1), 3d^+(2)| \, e^2/r_{12} \, |3d^+(1), 3s^-(2)\rangle$ (eqn (11.25)). It is therefore not unreasonable to think that for the electron with spin up, the density at the origin, $|\varphi_{1s}^+(0)|^2$, will be different from $|\varphi_{1s}^-(0)|^2$. Can one predict the sign of the difference? Hand-waving arguments can be produced for either sign. One could argue that the inner $1s^+$ electron with spin up is kept apart from the outer $3d^+$ electrons by the exclusion principle and thus has a greater probability to be pushed near the nucleus. One would then have

$$|\varphi_{1s}^+(0)|^2 - |\varphi_{1s}^-(0)|^2 > 0 \qquad (17.86)$$

for the spin density at the origin. On the other hand, one can also argue that for electrons with parallel spins the fact of being kept apart by the exclusion principle reduces the expectation value of their electro-static repulsion by so much that on the whole they are less repelled by each other than by electrons with antiparallel spins. The sign of (17.86) would then be reversed. Whatever the conclusion reached for the sign of the spin density at the origin for innermost s-electrons it would be reversed for s-electrons which are farther removed from the nucleus than the magnetic $3d$ electrons, such as the two $4s$ electrons of a free atom of the iron group.

In order to settle this question, and more generally to calculate χ, eqn (17.64) is rewritten as

$$\chi = \frac{4\pi}{2S} \sum_i \{|\varphi_{is}^+(0)|^2 - |\varphi_{is}^-(0)|^2\}. \qquad (17.87)$$

For the $3d$ group, i goes from 1 to 3 for ions and from 1 to 4 for atoms, and we again use a variational method. In contrast to the configuration interaction approach, the wave-function Ψ for the stretched state of the ground term ($S_z = S$, $L_z = L$, L and S maximum) is now written as a single Slater determinant constructed solely from orbitals belonging to the unperturbed configuration $\mathscr{C}_0$ given by (17.63). The difference from the conventional function Ψ_0 is in the choice of the trial one-electron functions which, in the conventional Hartree–Fock method, are given by (17.67). The trial wave-functions $P_{nl}(r)$ of (17.67) are assumed to be independent of the orientation of the spin, a restriction made for the sake of simplicity. On the other hand, we have explained that the radial functions of electrons with spin up and down are expected to be different because of their different coupling with the magnetic d-electrons. This is why it is reasonable to use trial one-electron functions

where $P_{nl}^{+}(r)$ is not required to be equal to $P_{nl}^{-}(r)$. This method is called the unrestricted Hartree–Fock (UHF) method. The new ionic wave-functions Ψ form a wider set than the functions Ψ_{0} of the conventional Hartree–Fock method, sometimes called restricted Hartree–Fock method (RHF) because there the restriction is imposed that the orbitals must be independent of the orientation of the spin. The functions Ψ are therefore at least as good as the functions Ψ_{0}; the foregoing argument suggests that they can be expected to be considerably better for the description of the spin-density near the nucleus, that is of the anomalous magnetic hyperfine structure. We shall not describe the procedure used for writing down and solving the equations obeyed by the spin-polarized orbitals in the UHF method (any more than we did for the conventional RHF method). Some details and an abundant bibliography can be found in the review articles of Freeman and Watson (1965), Watson and Freeman (1967a).

The calculated values of χ are in surprisingly good agreement with experiment considering the uncertainties involved in the calculation. First, the spin densities, that is the differences between the electronic densities $|\varphi_{ns}^{+}(0)|^{2}$ and $|\varphi_{ns}^{-}(0)|^{2}$ are, for each shell, very small fractions of the electronic densities themselves. Secondly, the contributions of different shells have opposite signs and cancel each other to a large extent. The calculation shows that the electronic densities behave as if electrons with parallel spins did attract each other, the second of our hand-waving conclusions. This is why, for the inner shells $1s$, $2s$ where the electrons with spin up are attracted away from the nucleus by the magnetic electrons $3d$, the spin density is negative, i.e. opposite to the orientation of the ionic spin. On the other hand, the spin density is positive for $3s$ electrons and, in free atoms, for $4s$ electrons also. Table 17.1 (Freeman and Watson 1965; Watson and Freeman 1967a)

TABLE 17.1

Ion	$Mn^{2+}(3d^5)$	$Fe^{2+}(3d^6)$	$Ni^{2+}(3d^8)$	$Mn^0(3d^54s^2)$
χ (atomic units)	−3·34	−3·29	−3·94	−0·54
$1s$ shell contribution to χ	−0·16	−0·21	−0·27	−0·03
$2s$ shell contribution to χ	−6·73	−7·80	−9·62	−6·63
$3s$ shell contribution to χ	+3·55	+4·72	+5·95	+3·23
$4s$ shell contribution to χ				+2·89

illustrates these statements. The ratio

$$\rho_n = \frac{|\varphi_{ns}^{+}|^{2}-|\varphi_{ns}^{-}|^{2}}{|\varphi_{ns}^{+}|^{2}+|\varphi_{ns}^{-}|^{2}}$$

is always a very small number; for example, for Mn^{2+}: $\rho_1 \approx 6 \times 10^{-6}$, $\rho_2 \approx 3 \times 10^{-3}$, $\rho_3 \approx 1 \cdot 1 \times 10^{-2}$.

The foregoing discussion was centered on the hyperfine structure in the iron group. For the $4f$ group of the rare earths where the orbital momentum is largely unquenched the effects of core polarization are only a small correction and cannot be separated from orbital hyperfine structure except for ions with the half-closed shell configuration $4f^7$ (see § 5.4 for a discussion).

Comparison between core polarization and configuration interaction

The explanations provided for the origin of the anomalous magnetic hyperfine interaction by these two approaches appear to be qualitatively different. It is true that quantitatively both are variational methods that go beyond the conventional Hartree–Fock calculations by enlarging the set of the many-electron trial functions Ψ from which are chosen the solutions that minimize $\langle \Psi | \mathscr{H} | \Psi \rangle$. However the ways in which these sets are enlarged in either method appear to differ considerably. In the configuration-interaction method the one-electron orbitals, from which are constructed the many-electron Slater determinants, are the conventional functions (17.67) with a radial part independent of the orientation of the individual electronic spins, but more than one configuration $\mathscr{C}$ is required to represent adequately the ionic functions Ψ. In contrast, a single configuration is used in the core-polarization method but the one-electron orbitals from which it is built are spin-dependent as explained earlier. We now show that this distinction is more apparent than real and in the process we shall point out the main theoretical weakness of the core-polarization method, however attractive its other features. Let us write again the ionic wave-function Ψ,

$$\Psi = \Psi_0 + \alpha \Psi_{3b} + O(\alpha^2), \qquad (17.88)$$

used in the configuration-interaction method, with Ψ_0 given by (17.69) and Ψ_{3b} by (17.73). Equation (17.88) can be rewritten

$$\Psi = \left(P^+ - \frac{\alpha\sqrt{(2S)}}{2\sqrt{(S+1)}} R^+, \ P^- + \frac{\alpha\sqrt{(2S)}}{2\sqrt{(S+1)}} R^- \right) Q^S +$$
$$+ O(\alpha^2) - \frac{\alpha}{\sqrt{(S+1)}} P^+ R^+ Q^{S-1}. \quad (17.89)$$

If we disregard the last term of (17.89), which is there to make Ψ an eigenstate of S, we see that the wave-function Ψ is nothing but a spin-polarized UHF function where the two radial trial functions with spin

up and spin down used in the core-polarization method are, respectively,

$$P-\frac{\alpha\sqrt{(2S)}}{2\sqrt{(S+1)}}R \text{ and } P+\frac{\alpha\sqrt{(2S)}}{2\sqrt{(S+1)}}R. \qquad (17.90)$$

It remains to explain the last term of (17.89). Its absence in the core polarization method is the main weakness of the method, in that the Slater determinants it uses are no longer eigenstates of S. In principle this is a serious matter, since the true ionic state *is* an eigenstate of S, and the errors committed through disregard of this fact are very difficult to evaluate theoretically (see, for a discussion and references, Freeman and Watson 1965; Watson and Freeman 1967a; Moser 1967). The justification for the neglect of a proper symmetrization of the wave-function Ψ is twofold: first, beyond the lightest atoms the calculations become unmanageable; secondly, even in its imperfect form the agreement of the UHF method with experiment is satisfactory in most cases.

To conclude this discussion one can say that the origin of the anomalous magnetic hyperfine structure in the iron group is well understood qualitatively. It is due to the existence of a finite spin-density at the nucleus for which the method of core polarization provides both a satisfactory physical model and calculated values in reasonable agreement with experiment.

17.7. Finer effects in the theory of hyperfine structure

Since the discovery of Endor (see Chapter 4) the accuracy in the measurements of the various parameters of hyperfine structure has increased considerably. Many small effects that hitherto could be observed only by the very precise method of atomic beams have become accessible to paramagnetic resonance measurements. Some of these effects are described below.

Hyperfine structure anomaly

This effect (not to be confused with the anomalous hyperfine structure discussed in the last section) is related to the finite size of the atomic nucleus. When an s-electron penetrates inside the nucleus the magnetic properties of the latter are no longer those of a point-like magnetic moment. Instead the electron probes the detailed distribution of the magnetization inside the nucleus which, like the electronic magnetization of an atom, has a part due to the magnetic spin moments of the individual nucleons and a part related to their orbital motion.

For two different isotopes the shapes of these distributions may be quite different and the ratio of their contributions to the magnetic hyperfine structure has no reason to be equal to the ratio of the nuclear magnetic moments. Another effect related to the former, but not identical with it, is the fact that although two isotopes have the same charge Ze, it may be distributed differently inside the two nuclei. This modifies the electronic wave-function inside the nucleus, which is another reason why the ratio of the hyperfine structures may differ from that of the nuclear moments.

Let $A_S\mathbf{I} . \mathbf{S}$ and $A'_S\mathbf{I}' . \mathbf{S}$ be the scalar parts of the hyperfine interactions for two isotopes with spins I and I' and magnetic moments $\mu_I = \gamma_n\hbar\mathbf{I}$, $\mu'_I = \gamma'_n\hbar\mathbf{I}'$. The hyperfine anomaly can be defined as the parameter Δ in the following relation

$$\frac{A_S}{A'_S} = \frac{\gamma_n}{\gamma'_n}(1+\Delta). \tag{17.91}$$

The ratios A_S/A'_S and γ_n/γ'_n can be obtained in principle from an analysis of Endor measurements (or for γ_n/γ'_n from NMR measurements performed on diamagnetic compounds) and Δ, extracted from (17.91), can be compared with theoretical values derived from nuclear models, thus providing a test of their validity. For light nuclei Δ is very small, of the order of a few parts in a million, but for heavy nuclei where, because of the large Z, the electron spends an appreciable part of its time inside the nucleus, Δ can exceed 1 per cent. In the analysis of Endor measurements leading to Δ one should beware of spurious effects (see, for example, § 4.8) which can simulate a hyperfine anomaly. These are the second-order magnetic hyperfine structure and the pseudo-nuclear Zeeman effects to be described in § 18.1; also, very light nuclei have a zero-point motion that modifies the value of the wave-function at the nucleus, making the ratio A_S/A'_S different from γ_n/γ'_n even though the electronic penetration inside the nucleus is negligible. The best example of this situation is provided by the nuclei of ^{6}Li and ^{7}Li in silicon where an apparent Δ of the order of 1 per cent was observed (Feher 1959).

The Sternheimer antishielding factor

We have already mentioned at the end of § 17.1 the fact that in contrast to magnetic hyperfine structure, quadrupole interactions in ions do not require the existence of an unfilled shell of d or f electrons; if a diamagnetic ion such as say F$^-$, with only closed shells, is placed

in a non-cubic environment, there will be an electric-field gradient at the nucleus. This gradient can be written as

$$q = q_c(1 - \gamma_\infty), \qquad (17.92)$$

where q_c is the gradient produced by the charges outside the ion and γ_∞ is a coefficient, introduced and calculated by Sternheimer, which represents the effect of the polarization of the closed shells by the external charges. It is large and negative, reaching values of the order of 100 for the heaviest ions. In paramagnetic ions the electrons belonging to the unfilled shells d and f produce at the nucleus electric field gradients proportional to $\langle r^{-3} \rangle$, which are usually larger than those given by (17.92) for diamagnetic ions even when the anti-shielding amplification is included. We must also add to the electric-field gradient at the nucleus the effect of the polarization of the closed shells of the paramagnetic ions by the electrons of the unfilled shell. This problem has also been considered by Sternheimer and the polarization is taken into account by writing

$$\langle r_q^{-3} \rangle = (1 - R_q)\langle r^{-3} \rangle. \qquad (17.93)$$

The remarkable fact is that in contrast to $|\gamma_\infty|$, which is large, $|R_q|$ is a small number of the order $0 \cdot 1 - 0 \cdot 2$.

Without going into details of the heavy calculations of γ_∞ and R_q it is interesting to explain qualitatively the disparity between their magnitudes. Let us consider a spherical electronic shell polarized by a point charge e with polar coordinates R and Θ. In the absence of this charge the ground state $|O\rangle$ of the electronic shell is spherically symmetrical and the expectation value of the quadrupole gradient $\langle V_{zz} \rangle = e\langle O| P_2(\cos \theta)/r^3 |O\rangle$ vanishes. The point charge itself produces at the nucleus a gradient $q_c = e(P_2(\cos \Theta)/R^3)$. Its electrostatic interaction with the closed shell is $e^2/|\mathbf{r} - \mathbf{R}|$, which admixes excited states $|n\rangle$ into the ground state $|O\rangle$ so that the expectation value of $P_2(\cos \theta)/r^3$ has now a finite value that can be written symbolically

$$\frac{e}{\Delta_{on}} \langle O| \frac{P_2(\cos \theta)}{r^3} |n\rangle \langle n| \frac{e^2}{|\mathbf{r} - \mathbf{R}|} |O\rangle. \qquad (17.94)$$

Let us assume first that the charge density of the closed shell vanishes for $r \geqslant R$ (external polarizing charge). A typical term in the expansion of $e^2/|\mathbf{r} - \mathbf{R}|$ will be

$$e^2 P_2(\cos \theta) P_2(\cos \Theta) r^2/R^3, \qquad (17.95)$$

and (17.94) can be rewritten

$$\frac{1}{\Delta_{on}} \langle O| \frac{P_2(\cos\theta)}{r^3} |n\rangle\langle n| \, e^2 r^2 P_2(\cos\theta) \, |O\rangle \left\{ \frac{eP_2(\cos\Theta)}{R^3} \right\} \qquad (17.96)$$

where the expression multiplying $eP_2(\cos\Theta)/R^3$ appears as an anti-shielding factor.

Suppose instead that the charge density of the closed shell vanishes for $r \leqslant R$ (internal polarizing charge). In the expansion of $e^2/|\mathbf{r}-\mathbf{R}|$ we must now replace (17.95) by $e^2 P_2(\cos\theta)P_2(\cos\Theta)R^2/r^3$ and (17.96) is replaced by

$$\frac{1}{\Delta_{on}} \langle O| \frac{P_2(\cos\theta)}{r^3} |n\rangle\langle n| \, e^2 r^2 P_2(\cos\theta) \left(\frac{R}{r}\right)^5 |O\rangle \left\{ \frac{eP_2(\cos\Theta)}{R^3} \right\}. \qquad (17.97)$$

A comparison of (17.96) and (17.97) shows that because of the factor $(R/r)^5$ in (17.97), where $R/r < 1$, the anti-shielding factor is very much larger when the polarizing charge is outside the closed shell than when it is inside. In applying these arguments to the electrons of the unfilled shells d or f which play the role of the polarizing charges, we have the problem that they are neither inside nor outside the closed shells but are spread all over them. However the main contribution to their quadrupole interactions comes from regions near the nucleus, where they play the part of inside charges with respect to the closed shells and thus should have a small anti-shielding factor. The foregoing argument (Watson and Freeman 1967a) is admittedly very crude but perhaps sufficient to give an inkling as to the causes of the smallness of $|R_q|$ relative to $|\gamma_\infty|$.

The different values of $\langle r^{-3} \rangle$

Equation (17.61), which is the operator expression of the hyperfine structure for the manifold of states belonging to a term (L, S), contains (besides the universal constants β and $\hbar$) three types of parameters; first, the purely nuclear constants γ_{n} and Q; secondly, what one might call angular parameters, namely ξ and $\langle L\| \, \alpha \, \|L\rangle$, which are determined by the algebra of vector coupling of angular momenta; and thirdly, the so-called radial constants, $\langle r^{-3} \rangle$ and $\kappa\langle r^{-3} \rangle$, which have to be computed numerically from a detailed knowledge of the electronic wave-functions. We have explained in some detail in §§ 17.5 and 17.6 the origin of the scalar interaction proportional to $\kappa\langle r^{-3} \rangle$. Whatever the language used to express it, core polarization or configuration interaction, this interaction is due to the presence at the nucleus of a finite density of

unpaired s-electrons and it is a large effect, responsible in particular for the bulk of magnetic hyperfine structure in S-state ions. Apart from this 'anomalous' scalar interaction, the remainder of (17.61) or what we might call its 'normal' part, results from the following assumptions: the term (L, S) is constructed from a single configuration $\mathscr{C}$, which contains a single unfilled shell; the one-electron orbitals are of the restricted Hartree–Fock type given by (17.67); the values of L and S are such that only one (L, S) term can be constructed from the configuration $\mathscr{C}$. The last condition is always fulfilled if the term (L, S) obeys Hund's rule; in that case ξ and $\langle L \| \alpha \| L \rangle$ are given by (17.46). A direct consequence of these assumptions is that a single 'radial' parameter $\langle r^{-3} \rangle$ multiplies all the 'normal' terms of (17.61). We have already shown in the preceding section that this assumption is invalid for the quadrupole interaction, because of the polarization of the inner shells by the electrons of the unfilled shell.

On the other hand it can be shown, using rotational invariance arguments, that (17.61) is the most general expression that can be written *a priori* for magnetic dipole and electric quadrupole hyperfine structure within a term (L, S) if the constants ξ and $\langle L \| \alpha \| L$ can be considered as adjustable parameters rather than well-defined algebraic functions such as those given by (17.46). This is a direct consequence of the fact that the hyperfine interaction is a sum of one-electron operators and that individual electrons each have a spin $\frac{1}{2}$ (Sandars and Beck 1965). However the common practice is to retain for ξ and $\langle L \| \alpha \| L$ their calculated values and to assume that for the three terms of (17.61), proportional respectively to $\langle r^{-3} \rangle$, $\xi \langle r^{-3} \rangle$, and $\langle L \| \alpha \| L \rangle \langle r^{-3} \rangle$, it is the constant $\langle r^{-3} \rangle$, which can have three different values, that we write as $\langle r_l^{-3} \rangle$, $\langle r_{sc}^{-3} \rangle$, and $\langle r_q^{-3} \rangle$. We had already done so for the quadrupole interactions in eqn (17.93). We can even extend it to the 'anomalous' scalar term by taking $-\kappa = 1$ and introducing a fourth value $\langle r_s^{-3} \rangle$. We can then rewrite the magnetic hyperfine structure as

$$\frac{W_n^m}{\beta \gamma_n \hbar} = 2 \langle r_l^{-3} \rangle (\mathbf{L} \cdot \mathbf{I}) + g_s \langle r_s^{-3} \rangle (\mathbf{S} \cdot \mathbf{I}) + $$

$$+ g_s \xi \langle r_{sc}^{-3} \rangle \{ L(L+I)(\mathbf{S} \cdot \mathbf{I}) - \tfrac{3}{2}(\mathbf{L} \cdot \mathbf{S})(\mathbf{L} \cdot \mathbf{I}) - \tfrac{3}{2}(\mathbf{L} \cdot \mathbf{I})(\mathbf{L} \cdot \mathbf{S}) \}, \quad (17.98)$$

where we have used for the spin gyromagnetic factor the correct value g_s rather than the approximate value 2.

In contrast to the comparatively large effects connected with the 'anomalous' scalar coupling the difference between $\langle r_l^{-3} \rangle$ and $\langle r_{sc}^{-3} \rangle$ is rather small. Experimentally a difference of 10 per cent has been

observed in oxygen and fluorine atoms (Harvey 1965) and 1 per cent for Sm (Woodgate 1966). There is no evidence so far for such a difference from paramagnetic resonance experiments. The origin of this difference must again be sought in the insufficient accuracy of restricted Hartree–Fock functions. The arguments we used in § 17.6 to convince ourselves that radial functions with spin up and spin down should be slightly different is by no means restricted to s-electrons but applies equally well to all the electrons of the ion whether they belong to closed shells or to the unfilled shell, provided there *is* an unfilled shell. Furthermore there is no reason to restrict the one-electron orbits with different values of the magnetic quantum number m_l to have the same radial part: as long as there is an imbalance between the values of m_l in the unfilled shell, the variational equations to be satisfied by two trial radial functions P_{nlm_l} and $P_{nlm_{l'}}$ will be different, and in particular P_{nlm_l} will differ from $P_{nl,-m_l}$. It is clear that under these conditions closed shells will bring a finite contribution to the orbital and to the dipolar hyperfine couplings and that the corresponding values of $\langle r_l^{-3} \rangle$ and $\langle r_{sc}^{-3} \rangle$ have no reason to coincide. The same will naturally also be true for the contributions of the unfilled shell.

For quadrupole interactions the fact that $P_{nlm_l} \neq P_{nl,-m_l}$ will lead to a finite quadrupole interaction for S-state ions with half-filled shells such as $3d^5$ or $4f^7$ since it is proportional to

$$\sum_{m_l=-l}^{l} \{3m_l^2 - l(l+1)\} \langle r_q^{-3} \rangle_{m_l}. \tag{17.99}$$

Such a small but finite quadrupole interaction has been observed in paramagnetic resonance for the $Eu^{2+}(4f^7)$ ion in cubic symmetry in CaF_2 (Baker and Williams 1962) and also in the free atoms Mn and Eu (Evans, Sandars, and Woodgate 1965).

In the foregoing we had implicitly assumed that the non-relativistic expression (17.30) for the one-electron Hamiltonian was correct. However, as Sandars and Beck (1965) have shown, the magnetic part of the hyperfine interaction can still be represented correctly by the expression (17.98) when relativistic corrections, neglected in the derivation of (17.30), are taken into account. It is therefore very difficult to separate relativistic corrections from those due to the use of unrestricted Hartree–Fock functions (UHF). Nevertheless this can still be achieved under favorable circumstances for the scalar term $g_s \langle r_s^{-3} \rangle (\mathbf{S} \cdot \mathbf{I})$ of (17.98). If this term originates in core polarization it implies an unpaired spin-density at the nucleus and thus makes

possible the existence of a hyperfine structure anomaly observable by comparing the spectra of two isotopes. A scalar term of relativistic origin is not necessarily associated with a finite spin-density at the nucleus and in this case does not lead to a hyperfine structure anomaly. The hyperfine structure measurements on the europium atom ($4f^7$, ^8S) show a small isotropic magnetic hyperfine structure but no hyperfine anomaly for the isotopes (151,153). Apart from contributions arising from the breakdown of LS-coupling, the hyperfine structure is therefore attributed mainly to relativistic effects (Evans, Sandars, and Woodgate 1965) (see § 4.7).