Analysis of dichroism in the electromagnetic response of superconductors

K. Capelle and E. K. U. Gross
Institut für Theoretische Physik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

B. L. Györffy
H. H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom
(Received 13 May 1997)

The absorption of polarized light in superconductors is studied within the framework of the Bogolubov–de Gennes approach to inhomogeneous superconductors in magnetic fields. Several mechanisms which give rise to a polarization-dependent absorption (i.e., dichroism) in superconductors are analyzed in detail. The relation to the absorption of unpolarized light in superconductors and to the absorption of polarized light in normal conductors is investigated and several effects, not known from either of these cases, are found. These effects arise from the interplay of broken chiral symmetry, which produces dichroism, with the superconducting coherence. One potential source for dichroism, namely spin-orbit coupling, is investigated numerically for a simple model superconductor. [S0163-1829(98)01122-9]

I. INTRODUCTION

In this paper we present a systematic approach to the polarization-dependent absorption of light in superconductors. The phenomenon that left-handed circularly polarized light and right-handed circularly polarized light are absorbed differently by many substances is commonly known as dichroism. Dichroism in the presence of an external magnetic field is usually referred to as the Faraday effect or the Kerr effect, depending on the geometry of the experiment. Dichroism in the absence of external magnetic fields is often referred to as spontaneous dichroism. It is found in magnetically ordered systems, such as iron, but also in systems which break inversion symmetry, such as sugar. In the present paper we use the term dichroism to refer to all situations in which the absorption of light depends on its polarization.

In normal (i.e., not superconducting) and magnetically ordered metals these effects have been the subject of intense study for many years. (See Refs. 1 and 2 for recent reviews.) It is by now unanimously accepted that dichroism in these systems arises mainly from the simultaneous presence of spin-orbit coupling and the spin magnetization. Modern calculations of the optical response to polarized light are therefore usually performed in a relativistic framework.[1,2]

However, similar calculations for superconductors do not exist yet. It is the purpose of the present work to provide a basis for such calculations and to present some first results.

The motivation for this investigation arises partly from the fact that several recent experiments report the observation of dichroic phenomena in superconductors,[3–8] and partly from the fundamental interest in the role of the spin-orbit coupling in superconductors.[9–11] Spin-orbit coupling, being a relativistic effect of second order in v/c, becomes more important in systems with heavy elements (atomic number $Z \gtrsim 40$) in the lattice.[12] Many interesting superconductors, e.g., the heavy-fermion compounds and the high-temperature superconductors, do indeed contain very heavy elements, such as mercury ($Z=80$), uranium ($Z=92$), bismuth ($Z=83$), lanthanum ($Z=57$), platinum ($Z=78$), etc.

In view of the fact that a qualitative and quantitative understanding of dichroism in normal and magnetically ordered metals necessarily requires a relativistic theory, it is of obvious relevance for the present investigations to employ a relativistic theory of superconductivity. Such a theory has recently been constructed.[9–11] From this theory the full form of the spin-orbit operator in superconductors is known to contain not only gradients of the lattice potential (as is the case for the conventional spin-orbit operator), but also gradients of the pair potential of the superconductor. The latter type of spin-orbit coupling is referred to as the anomalous spin-orbit coupling (ASOC).

In the present theory of dichroism we take both types of spin-orbit coupling into account. Several other potential sources for dichroism in superconductors are also considered in detail. To this end we use a perturbative approach, based on the Bogolubov–de Gennes equations in the presence of a magnetic field. The various sources for dichroism are included via first-order stationary perturbation theory for the quasiparticle wave function. The resulting perturbed single-particle states are then, in a second step, used as unperturbed states between which the transitions caused by the polarized light take place. The absorption of light is treated by a generalization of the standard golden rule of first-order time-dependent perturbation theory. First results from this investigation were already presented in a recent paper.[13]

The present paper is organized as follows: Sec. II A contains an outline of the perturbative approach which we employ for the calculations. The unperturbed system is described by the spin-Bogolubov–de Gennes equations, which are a generalization of the conventional Bogolubov–de Gennes equations, designed to treat spin-dependent phenomena. Since perturbation theory has up to now been developed only for the conventional Bogolubov–de Gennes equations, we present the corresponding generalizations for the spin-Bogolubov–de Gennes equations in some detail. The resulting expressions can be used for any type of perturbative calculation for the spin-Bogolubov–de Gennes equations, not

only for investigations of dichroism. Next, we explicitly specify the perturbations to be included in our approach to dichroism. These include the above-mentioned spin-orbit terms, the effect of orbital currents and, if necessary, order-parameter inhomogeneities.

In Sec. II B we employ the perturbative expressions to derive a formula for the power absorption in superconductors as a function of the polarization of the light. In Sec. II C several distinct mechanisms for dichroism in superconductors are identified. A detailed analysis of the physics behind the various mechanisms is performed and the circumstances under which they can produce dichroism in superconductors are discussed.

Section III contains model calculations for one of these mechanisms, namely the conventional spin-orbit coupling. Using simple approximations for the relevant matrix elements and the density of states, which are discussed in Secs. III A and III C, respectively, we draw further analytical conclusions about the physics behind this mechanism in Sec. III B, and evaluate the formulas numerically in Sec. III D. The numerical results are analyzed as functions of temperature, frequency, and magnetic-field strength.

We emphasize that these calculations are not meant as quantitative predictions for experiments, but rather as model calculations illustrating the analytical results of the previous sections and exhibiting surprising *qualitative* features of dichroism in superconductors. The paper ends with a brief discussion of some recent experiments in light of our theory in Sec. IV and a summary in Sec. V.

II. PERTURBATIVE APPROACH TO DICHROISM IN SUPERCONDUCTORS

A. Perturbation theory

1. The spin-Bogolubov–de Gennes equations

The proper microscopic description of inhomogeneous superconductors is provided by the Bogolubov–de Gennes equations (BdGE)

$$\begin{pmatrix} h_0 & \Delta(\mathbf{r}) \\ \Delta^\dagger(\mathbf{r}) & -h_0^* \end{pmatrix} \begin{pmatrix} u_n(\mathbf{r}) \\ v_n(\mathbf{r}) \end{pmatrix} = E_n \begin{pmatrix} u_n(\mathbf{r}) \\ v_n(\mathbf{r}) \end{pmatrix}, \tag{1}$$

where

$$h_0 = \frac{p^2}{2m} + v(\mathbf{r}) - \mu \tag{2}$$

is the normal-state single-particle Hamiltonian, $\Delta(\mathbf{r})$ is the pair potential, and the $u_n(\mathbf{r})$ and $v_n(\mathbf{r})$ are particle and hole amplitudes. The $\mathbf{r}$ dependence of the pair potential $\Delta(\mathbf{r})$ describes the center-of-mass motion of the Cooper pair. The internal degrees-of-freedom of the pair would be described by the dependence of the pair potential on the relative coordinate of the two electrons. This would require the use of the nonlocal version of the Bogolubov–de Gennes equations.[14,15] In the present paper we limit our attention to the local version of the Bogolubov–de Gennes equation, as specified above. This means that we can not adequately treat the effects of the internal degrees of freedom, such as the difference between s-wave and d-wave superconductors. The local Bogolubov–de Gennes equations have been the starting point for many microscopic investigations of superconductors (Refs. 14–25 among others). However, for the present purpose their form is too restrictive because we need to incorporate the spin degrees of freedom of the quasiparticles and spin-orbit coupling terms. There exists a generalization of the BdGE in which these can be included properly, the spin-Bogolubov–de Gennes equations (SBdGE). They read[14,16,26,27]

$$\begin{pmatrix} h_\uparrow & 0 & 0 & \Delta \\ 0 & h_\downarrow & -\Delta & 0 \\ 0 & -\Delta^* & -h_\uparrow^* & 0 \\ \Delta^* & 0 & 0 & -h_\downarrow^* \end{pmatrix} \begin{pmatrix} u_{\uparrow n\sigma}(\mathbf{r}) \\ u_{\downarrow n\sigma}(\mathbf{r}) \\ v_{\uparrow n\sigma}(\mathbf{r}) \\ v_{\downarrow n\sigma}(\mathbf{r}) \end{pmatrix} = E_{n\sigma} \begin{pmatrix} u_{\uparrow n\sigma}(\mathbf{r}) \\ u_{\downarrow n\sigma}(\mathbf{r}) \\ v_{\uparrow n\sigma}(\mathbf{r}) \\ v_{\downarrow n\sigma}(\mathbf{r}) \end{pmatrix}, \tag{3}$$

where

$$h_\tau = h_0 + \tau \mu_B B. \tag{4}$$

$\tau = +1$ for $h_\uparrow$ and $\tau = -1$ for $h_\downarrow$. We have already included the Zeeman coupling of the spins to the external magnetic field B (which is assumed to be spatially constant and to point along the z direction), but not yet the coupling of the orbital degrees of freedom to the vector potential and spin-orbit coupling. The latter two effects are included as perturbations, in Sec. II A 3.

We take these equations to describe the unperturbed superconductor. They can be obtained from a spin-dependent Bogolubov-Valatin transformation.[14,16,26,27] Alternatively, they are found as the nonrelativistic limit of the relativistic Bogolubov–de Gennes equations.[9–11] They relate to the conventional BdGE in exactly the same way as the Pauli equation relates to the Schrödinger equation.

We now summarize a number of properties of the SBdGE which are essential for the following considerations. To every eigenvector

$$\begin{pmatrix} u_{\uparrow n\sigma}(\mathbf{r}) \\ u_{\downarrow n\sigma}(\mathbf{r}) \\ v_{\uparrow n\sigma}(\mathbf{r}) \\ v_{\downarrow n\sigma}(\mathbf{r}) \end{pmatrix} \tag{5}$$

of the SBdGE with eigenvalue $E_{n\sigma} > 0$ belongs a second eigenvector

$$\begin{pmatrix} u_{\uparrow \overline{n\sigma}}(\mathbf{r}) \\ u_{\downarrow \overline{n\sigma}}(\mathbf{r}) \\ v_{\uparrow \overline{n\sigma}}(\mathbf{r}) \\ v_{\downarrow \overline{n\sigma}}(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} v_{\uparrow n\sigma}(\mathbf{r}) \\ v_{\downarrow n\sigma}(\mathbf{r}) \\ u_{\uparrow n\sigma}(\mathbf{r}) \\ u_{\downarrow n\sigma}(\mathbf{r}) \end{pmatrix}^* \tag{6}$$

with eigenvalue $E_{\overline{n\sigma}} = -E_{n\sigma}$. The negative energy solutions describe bound states of the Bogolubov quasiparticles (bogolons), or, equivalently, electrons condensed in Cooper pairs. Below we consider pair-breaking processes as the dominant source for absorption at low temperatures. These processes can simply be described as transitions from a negative energy state of the form (6) to a positive energy state of the form (5).

In the absence of all spin-dependent interactions the SBdGE rigorously reduce to the BdGE. The relation between the respective eigenfunctions is simply $u^0_{\tau n\sigma}(\mathbf{r})=\delta_{\sigma\tau}u_n(\mathbf{r})$ and $v^0_{\tau n\sigma}(\mathbf{r})=\sigma\delta_{\tau\bar{\sigma}}v_n(\mathbf{r})$, where σ and τ are ± 1. The inclusion of the Zeeman term in Eq. (3) does not change these relations, as long as B is spatially constant. If the pair field Δ is spatially constant as well, the spatial dependence of the eigenfunctions of the SBdGE is governed by the normal-state Hamiltonian h_0. The eigenfunctions are thus proportional to normal-state eigenfunctions ϕ_n, which are defined through $h_0\phi_n=\epsilon_n\phi_n$. The constants of proportionality are given by the BCS amplitudes[14] $u_n:=\sqrt{1/2(1+\epsilon_n/E_n)}$ and $v_n:=\sqrt{1/2(1-\epsilon_n/E_n)}$. The full form of the eigenfunctions of the SBdGE under these circumstances thus is

$$u^0_{\tau n\sigma}(\mathbf{r})=u_n\delta_{\sigma\tau}\phi_n(\mathbf{r})=\sqrt{\frac{1}{2}\left(1+\frac{\epsilon_n}{E_n}\right)}\,\delta_{\sigma\tau}\phi_n(\mathbf{r}), \quad (7)$$

$$v^0_{\tau n\sigma}(\mathbf{r})=\sigma v_n\delta_{\tau\bar{\sigma}}\phi_n(\mathbf{r})=\sigma\sqrt{\frac{1}{2}\left(1-\frac{\epsilon_n}{E_n}\right)}\,\delta_{\tau\bar{\sigma}}\phi_n(\mathbf{r}), \quad (8)$$

and the corresponding eigenvalue is given by

$$E_{n\sigma}=E_n+\sigma\mu_B B=\sqrt{\epsilon_n^2+\Delta^2}+\sigma\mu_B B. \quad (9)$$

If the pair potential is not spatially constant, the spatial dependence of the SBdGE eigenfunctions is not determined by the normal-state Hamiltonian anymore. In this case the particle and hole amplitudes are not proportional to $\phi_n(\mathbf{r})$, but have to be determined by solving the full SBdGE. Equations (3)–(9) define the unperturbed superconductor. In the next step we develop perturbation theory for such superconductors.

2. Perturbation theory for the spin-Bogolubov–de Gennes equations

Perturbation theory for the conventional BdGE has been developed by de Gennes[14] and, in a slightly different formulation, by Kümmel and co-workers.[28,29] While being equivalent to the former approach on the exact level, the latter approach has the advantage that in every order of approximation the formulas are of the same structure as in conventional perturbation theory for the Schrödinger equation. In the following, we generalize the latter approach to the SBdGE, where we allow for perturbations of the normal and of the pair potential and discuss both stationary and time-dependent perturbations. Since the derivation of the formulas is very similar to that of perturbation theory for the Schrödinger equation and to that for the conventional BdGE, we only present the final results and point out the main differences to the conventional case.

The wave functions in the presence of a stationary perturbation δH are, to first order in δH, given by

$$u_{\tau n\sigma}(\mathbf{r})=u^0_{\tau n\sigma}(\mathbf{r})+\sum_{\substack{(m\alpha)\\ \neq(n\sigma)}}\left[\frac{\delta H^{m\alpha,n\sigma}}{E_{n\sigma}-E_{m\alpha}}u^0_{\tau m\alpha}(\mathbf{r})+\frac{\delta H^{\overline{m\alpha},n\sigma}}{E_{n\sigma}-E_{\overline{m\alpha}}}u^0_{\tau\overline{m\alpha}}(\mathbf{r})\right], \quad (10)$$

$$v_{\tau n\sigma}(\mathbf{r})=u^0_{\tau n\sigma}(\mathbf{r})+\sum_{\substack{(m\alpha)\\ \neq(n\sigma)}}\left[\frac{\delta H^{m\alpha,n\sigma}}{E_{n\sigma}-E_{m\alpha}}v^0_{\tau m\alpha}(\mathbf{r})+\frac{\delta H^{\overline{m\alpha},n\sigma}}{E_{n\sigma}-E_{\overline{m\alpha}}}v^0_{\tau\overline{m\alpha}}(\mathbf{r})\right] \quad (11)$$

with corresponding equations for $u_{\tau\overline{n\sigma}}$ and $v_{\tau\overline{n\sigma}}$. δH can be any perturbing 4×4 matrix and may thus contain perturbations of the lattice potential and of the pair potential. All labels $(m\alpha)$ refer to positive energy states of the form (5). The barred labels, such as $(\overline{m\alpha})$, refer to negative energy states of the form (6). For notational simplicity we have suppressed an upper index 0 on the energies, although they are, of course, unperturbed energies. The change of the energies is, to first order, given by the usual result $\delta E_{n\sigma}=\delta H^{n\sigma,n\sigma}$.

A time-dependent perturbation gives rise to transitions from one state of the system to another. We take the perturbation to be

$$\delta\mathcal{H}(t,\mathbf{r})=\begin{pmatrix}\delta h(t,\mathbf{r})\mathbb{1} & 0\\ 0 & -\delta h(t,\mathbf{r})^*\mathbb{1}\end{pmatrix}, \quad (12)$$

where we condensed a 4×4 matrix in 2×2 form ($\mathbb{1}$ is the 2×2 unit matrix) and assumed that there is no time-dependent perturbing pair potential (such a potential would appear off the diagonal in Eq. (12) and could, if necessary, be included without difficulties). $\delta h(t,\mathbf{r})$ is of the general form

$$\delta h(t,\mathbf{r})=\delta h(\mathbf{r})e^{i\omega t}+\delta h(\mathbf{r})^\dagger e^{-i\omega t}. \quad (13)$$

The transition probability per unit time for a transition from the single-particle state $(n\sigma)$ to $(n'\sigma')$ is then

$$w^E_{n\sigma\to n'\sigma'}=\frac{2\pi}{\hbar}|\delta\tilde{\mathcal{H}}^{n\sigma,n'\sigma'}|^2\delta(E_{n'\sigma'}-E_{n\sigma}+\hbar\omega), \quad (14)$$

$$w^A_{n\sigma\to n'\sigma'}=\frac{2\pi}{\hbar}|\delta\tilde{\mathcal{H}}^{\dagger n\sigma,n'\sigma'}|^2\delta(E_{n'\sigma'}-E_{n\sigma}-\hbar\omega), \quad (15)$$

where the first expression describes emission and the second absorption. The labels $n\sigma,n'\sigma'$ refer to arbitrary solutions of the SBdGE. There is one important difference to the conventional golden rule: the matrix elements in Eqs. (14) and (15) are not simply those of the perturbation (12) and its Hermitian conjugate, but those of

$$\delta\tilde{\mathcal{H}}=\begin{pmatrix}\delta h(\mathbf{r})\mathbb{1} & 0\\ 0 & -\delta h(\mathbf{r})^T\mathbb{1}\end{pmatrix} \quad (16)$$

and

$$\delta\tilde{\mathcal{H}}^{\dagger}=\begin{pmatrix}\delta h(\mathbf{r})^{\dagger}\mathbb{1} & 0\\ 0 & -\delta h(\mathbf{r})^{*}\mathbb{1}\end{pmatrix}, \tag{17}$$

which differ in the location of the conjugate operators. The reason for this is that the coefficient of $e^{i\omega t}$ which gives rise to emission, and that of $e^{-i\omega t}$ which produces absorption, are mixed already in Eq. (12).

The above formulas can of course be used for any kind of perturbation of the SBdGE and should therefore be useful not only for investigations of dichroism, but for a large variety of calculations.

3. Explicit form of the perturbations

In specifying the SBdGE eigenfunctions according to Eqs. (7) and (8), we assumed the pair potential to be spatially constant. While this is a good approximation for superconductors with a large coherence length, it is inadequate for those with a short coherence length and for superconducting heterostructures. In order to take this into account we write the full pair potential as

$$\Delta(\mathbf{r})=\overline{\Delta}+\tilde{\Delta}(\mathbf{r}), \tag{18}$$

where $\overline{\Delta}$ is a suitable average of $\Delta(\mathbf{r})$ (e.g., taken over a unit cell) and $\tilde{\Delta}(\mathbf{r})$ is the local deviation from that average. In the SBdGE framework it is included by replacing $\Delta(\mathbf{r})$ by $\overline{\Delta}$ in the unperturbed SBdGE and adding the term

$$\delta H^{(0)}=\begin{pmatrix}0 & i\hat{\sigma}_y\tilde{\Delta}\\ (i\hat{\sigma}_y)^{\dagger}\tilde{\Delta}^{*} & 0\end{pmatrix}. \tag{19}$$

To simplify the notation we have employed the Pauli matrix $\hat{\sigma}_y$ to write the 4×4 equation (19) as a 2×2 equation. For small $\tilde{\Delta}(\mathbf{r})$, $\delta H^{(0)}$ can be treated as a small perturbation. [The case of large $\tilde{\Delta}(\mathbf{r})$ is discussed below Eq. (22).]

In Sec. II A 1 we also assumed that the magnetic field B is spatially constant and acts only on the electron spins. We do not consider the case of an inhomogeneous external magnetic field. However, even for constant fields, there is a coupling of the orbital degrees of freedom of the electrons to the vector potential. To first order in the vector potential this coupling is described by the term

$$\delta H^{(1)}=-\frac{q}{mc}\begin{pmatrix}(\mathbf{A}\cdot\hat{\mathbf{p}})\mathbb{1} & 0\\ 0 & -(\mathbf{A}\cdot\hat{\mathbf{p}})^{*}\mathbb{1}\end{pmatrix}, \tag{20}$$

where $B=(\nabla\times\mathbf{A})_z$ and $\hat{\mathbf{p}}$ is the momentum operator. We choose $\mathbf{A}$ in the Coulomb gauge, so that $\nabla\cdot\mathbf{A}=0$. $\delta H^{(1)}$ is considered a small perturbation as well. For realistic vector potentials this is certainly justified. In the case of the conventional BdGE this is the standard way to treat vector potentials, for instance in deriving the Meissner effect.[14]

Finally, we also consider the spin-orbit coupling as a small perturbation. The relativistic theory of superconductivity[9–11] predicts that there are two distinct types of spin-orbit coupling (SOC) in superconductors. One, the conventional SOC, appears on the diagonal of the SBdGE and contains gradients of the lattice potential $v(\mathbf{r})$. The other, the anomalous SOC (ASOC), appears off the diagonal and contains gradients of the pair potential with respect to the center-of-mass and relative coordinates (C-ASOC and R-ASOC). In the following calculations only the C-ASOC term is included. This amounts to neglecting the internal degrees of freedom of the two electrons in the Cooper pair and retaining only the spin-orbit coupling due to the center of mass motion. The perturbation corresponding to the SOC and C-ASOC terms is given by

$$\delta H^{(2)}=\frac{\hbar}{4m^2c^2}\times\begin{pmatrix}(\boldsymbol{\sigma}\cdot\nabla v(\mathbf{r})\times\hat{\mathbf{p}})\mathbb{1} & (\boldsymbol{\sigma}\cdot\nabla\Delta(\mathbf{r})\times\hat{\mathbf{p}})(i\hat{\sigma}_y)\\ [(\boldsymbol{\sigma}\cdot\nabla\Delta(\mathbf{r})\times\hat{\mathbf{p}})(i\hat{\sigma}_y)]^{\dagger} & -(\boldsymbol{\sigma}\cdot\nabla v(\mathbf{r})\times\hat{\mathbf{p}})^{*}\mathbb{1}\end{pmatrix}. \tag{21}$$

Further terms could be included in $\delta H^{(2)}$, e.g., the conventional and anomalous Darwin terms,[9,11] and the $(q^2/2mc^2)\,n\mathbf{A}^2$ term which goes along with Eq. (20) to second order of (v/c). Both of these are known not to produce dichroism in the normal state and are therefore not included here among the relevant perturbations. We return to this point in the discussion of existence criteria in Sec. II C 2.

The full (stationary) perturbation to be considered in this paper is the sum of all three terms

$$\delta H=\delta H^{(0)}+\delta H^{(1)}+\delta H^{(2)}. \tag{22}$$

The upper index, k, of each term $\delta H^{(k)}$ refers to the order in $1/c$ of the respective term. The relative magnitude of the three terms depends on the particular system under study and does not necessarily correlate with this order.

If the pair potential inhomogeneity $\tilde{\Delta}(\mathbf{r})$ is small enough, the ASOC term, which also contains $\tilde{\Delta}(\mathbf{r})$, can be dropped. If it is large, $\tilde{\Delta}(\mathbf{r})$ must be included already in the unperturbed Hamiltonian. The latter situation arises, e.g., for superconducting heterostructures and in the vortex phase of a type-II superconductor. The pair potential $\Delta(\mathbf{r})$ near the vortices is strongly inhomogeneous and the deviations from its average $\overline{\Delta}$ are so large that $\tilde{\Delta}(\mathbf{r})$ cannot be considered as a small perturbation. With the full pair potential in the unperturbed Hamiltonian, the spatial dependence of its eigenfunctions is, of course, not determined by normal-state eigenfunctions anymore. As a consequence, the particle and hole amplitudes in Eqs. (7) and (8) must be found by solving the SBdGE. The ASOC term then is a first-order perturbation, just as the conventional SOC. Indeed, the vortex lattice is among the situations in which this term is expected to be most important.

In order to facilitate the treatment of strongly and weakly inhomogeneous situations, we will keep both $\tilde{\Delta}(\mathbf{r})$ and the ASOC term in the expression for the perturbation. Depending on the particular system under study, one of the two terms can be dropped in the final result, Eq. (34).

The time-dependent perturbation is given by the interaction of the polarized light with the quasiparticles. It is of the same form as the stationary perturbation $\delta H^{(1)}$, but with a time-dependent vector potential $\mathbf{A}_L(t,\mathbf{r})$. The electromagnetic radiation is specified by its electric-field amplitude E_0,

the frequency ω, the wave vector $\mathbf{q}$, and the polarization vector $\boldsymbol{\epsilon}$, leading to the vector potential

$$\mathbf{A}_L(t,\mathbf{r}) = \frac{ic}{2}\frac{E_0}{\omega}\boldsymbol{\epsilon} e^{i(\omega t - \mathbf{qr})} + \text{c.c.} \tag{23}$$

In the Coulomb gauge ($\nabla \cdot \mathbf{A}_L = 0$), $\delta h(\mathbf{r})$, as needed in Eqs. (13), (16), and (17), is given by

$$\delta h(\mathbf{r}) = -\frac{e}{2mi}\frac{E_0}{\omega} e^{-i\mathbf{qr}} \boldsymbol{\epsilon}\hat{\mathbf{p}}, \tag{24}$$

where $-e$ is the charge on the electron. We assume perpendicular incidence of the light, so that $\mathbf{q}$ is parallel to the static magnetic field $\mathbf{B}$ and perpendicular to the sample surface, while $\boldsymbol{\epsilon}$ is a unit vector in the plane of the surface. This geometry corresponds to very common experimental setups.

B. Absorption of polarized light in superconductors

The quantity we use to describe the interaction of light with a superconductor is the power absorption P. It can be calculated within the above perturbation theoretical framework by evaluating

$$P = \sum_{fi} (E_f - E_i) f(E_i)[1 - f(E_f)] w^A_{i\to f}, \tag{25}$$

where i and f denote the initial and the final single-particle states, respectively, $w^A_{i\to f}$ is given by Eq. (15) and $f(E)$ stands for the Fermi function. The power emission can be calculated in a similar way from $w^E_{i\to f}$. In the following we will consider only power absorption. The emitted (scattered) power, being diluted over the entire solid angle, is usually not measured experimentally. Furthermore, as long as we consider the ground state or situations in which only a few excited single-particle states are occupied, absorption is by far the dominant process.

At sufficiently low temperatures almost all electrons are condensed in Cooper pairs. The most important mechanism for absorption under these circumstances is pair breaking. Scattering from unpaired electrons (i.e., from thermally excited quasiparticles) can safely be neglected at low temperatures. The inclusion of single-particle scattering normally produces only some additional absorption below the absorption edge.[30] Therefore we will limit ourselves to considering only pair breaking (i.e., bogolon creation) as the mechanism for absorption. It should be stressed that neglecting emission processes and scattering from broken pairs is physically justified, but in no way necessary for the further development.

Within the linear-response regime, the absorption of light is generally described by the conductivity tensor $\hat{\sigma}$. Expressions for the conductivity tensor in superconductors are known in a number of different approximations.[30–33]

For any system the elements of the conductivity tensor can be related in a simple fashion to the power absorption for various polarization directions.[34] For systems with cubic symmetry one finds within the dipole approximation, $\mathbf{q}=0$,

$$\mathrm{Im}[\sigma_{xy}(\omega)] = \frac{1}{VE_0^2}[P(\omega,\boldsymbol{\epsilon}_L) - P(\omega,\boldsymbol{\epsilon}_R)] \tag{26}$$

and

$$\mathrm{Re}[\sigma_{xx}(\omega)] = \frac{1}{VE_0^2}[P(\omega,\boldsymbol{\epsilon}_L) + P(\omega,\boldsymbol{\epsilon}_R)], \tag{27}$$

where V is the sample volume, $P(\omega,\boldsymbol{\epsilon})$ is the power absorption for polarization $\boldsymbol{\epsilon}$ and frequency ω, and E_0 is the electric-field strength of the light. In the frequency range relevant for pair breaking, the dipole approximation is generally valid and will be used exclusively. For light with left-handed polarization (LHP) we have

$$\boldsymbol{\epsilon}_L = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \tag{28}$$

while light with right-handed polarization (RHP) has a polarization vector

$$\boldsymbol{\epsilon}_R = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}. \tag{29}$$

Equations (26) and (27) hold regardless of whether the system is superconducting, magnetic, or in the normal state, because the microscopic properties of the system do not enter the derivation.[34] Similar formulas can also be derived for noncubic crystals[2,35] or beyond the dipole approximation.[2,36,37]

Equation (26) shows that the imaginary part of the off-diagonal elements of the conductivity tensor is nonzero only if light with left- and right-handed polarization is absorbed in a different way. These elements thus provide a direct measure for dichroism which can be used to relate experimental results to theoretical calculations. For the case of magnetic and normal metals they are routinely used for this purpose. (See the review Ref. 2 and the proceedings volume Ref. 1 for further discussion and applications.)

By substituting Eqs. (24) and (17) in Eq. (15) and the result in Eq. (25), we arrive at

$$P(\boldsymbol{\epsilon}) = \frac{\pi e^2 E_0^2}{2m^2\omega}\sum_{NN'} f(-E_N) f(-E_{N'}) |(u_N|\boldsymbol{\epsilon}^*\mathbf{p}|v^*_{N'}) + (v_N|\boldsymbol{\epsilon}^*\mathbf{p}|u^*_{N'})|^2 \delta(E_N + E_{N'} - \hbar\omega). \tag{30}$$

Here we employed the particle-hole convention, i.e., all energies are positive, the sums are limited to positive energy single-particle states and the explicit form of the negative energy states, as given by Eq. (6) was used. In writing Eq. (30) we used a short-hand notation for the matrix elements. Written out in complete detail, the first matrix element is given by

$$(u_N|\boldsymbol{\epsilon}^*\mathbf{p}|v^*_{N'}) = \int d^3r \begin{pmatrix} u_{\uparrow n\sigma}(\mathbf{r}) \\ u_{\downarrow n\sigma}(\mathbf{r}) \end{pmatrix}^T \frac{\hbar}{i}\boldsymbol{\epsilon}^*\cdot\nabla_{\mathbf{r}} \begin{pmatrix} v_{\uparrow n'\sigma'}(\mathbf{r}) \\ v_{\downarrow n'\sigma'}(\mathbf{r}) \end{pmatrix}^* \tag{31}$$

and the second one reads

$$(v_N|\boldsymbol{\epsilon}^*\mathbf{p}|u^*_{N'})=\int d^3r\begin{pmatrix}v_{\uparrow n\sigma}(\mathbf{r})\\ v_{\downarrow n\sigma}(\mathbf{r})\end{pmatrix}^T\frac{\hbar}{i}\boldsymbol{\epsilon}^*\cdot\nabla_{\mathbf{r}}\begin{pmatrix}u_{\uparrow n'\sigma'}(\mathbf{r})\\ u_{\downarrow n'\sigma'}(\mathbf{r})\end{pmatrix}^*. \tag{32}$$

If we now substitute the unperturbed solutions to the SBdGE, $u^0_{\tau n\sigma}$ and $v^0_{\tau n\sigma}$, as given by Eqs. (7) and (8), in Eqs. (31) and (32), we can calculate the power absorption in the superconductor as a function of polarization. Evaluating the result once for $\boldsymbol{\epsilon}$ given by Eq. (28) and once for $\boldsymbol{\epsilon}$ given by Eq. (29), then leads to a direct measure for dichroism, and, according to Eq. (26), an expression for the off-diagonal elements of the imaginary part of the conductivity tensor. Proceeding in this way it is immediately found from Eq. (30) that

$$\Delta P\propto\sum_{nn'}[p_y^{n\bar{n}'}(p_x^{n\bar{n}'})^*-p_x^{n\bar{n}'}(p_y^{n\bar{n}'})^*]\equiv 0, \tag{33}$$

i.e., there is no dichroism. The p_i^{kl} in Eq. (33) are normal-state matrix elements of Cartesian components of the momentum operator. A barred index stands for the complex conjugate normal-state wave function $\phi_{\bar{n}}(\mathbf{r}):=\phi_n(\mathbf{r})^*=\phi_n(-\mathbf{r})$. The latter equality holds if the original normal-state Hamiltonian, h_0, does not break time-reversal and inversion symmetry.[38,39] As a consequence all momentum matrix elements are real, even if the individual wave functions are not, and ΔP vanishes identically. This result, which holds independently of the explicit form of the normal-state wave functions, was of course to be expected. Indeed, it is well known from normal and magnetically ordered metals that to find dichroism one needs to include mechanisms which break chiral symmetry, so that the system becomes susceptible to the difference between left- and right-handed polarization.[1,2,34]

In the next step we include all the stationary perturbations which were discussed above in order to find out if, and under which circumstances, these produce dichroism in superconductors. To this end we substitute Eqs. (19), (20), and (21) in Eq. (22). Equation (22) is then used, together with Eqs. (7) and (8), in Eqs. (10) and (11) to determine the form of the perturbed wave functions. These wave functions are, in a next step, used as *unperturbed* single-particle states with respect to the time-dependent perturbation in Eqs. (30)–(32). Multiplying out all terms to first order in δH we find

$$\begin{aligned}\Delta P:&=P(\boldsymbol{\epsilon}_L)-P(\boldsymbol{\epsilon}_R)\\ &=\frac{\pi e^2E_0^2}{m^2\omega}\sum_{nn'\sigma}p(n,n')f(-E_{n\sigma})f(-E_{n'\bar{\sigma}})\\ &\quad\times\delta(E_{n\sigma}+E_{n'\bar{\sigma}}-\hbar\omega)\\ &\quad\times\sum_m\mathrm{Re}\left[\frac{p(m,n')C^m_{nn'}}{E_n-E_m}T^+_{m\sigma}+\frac{l(m,n')C^{\bar{m}}_{nn'}}{E_n+E_m}T^-_{m\sigma}\right],\end{aligned} \tag{34}$$

where Re denotes the real part. All sums are restricted to single-particle states with positive energies (particle-hole convention). The first term under the sum on m represents the contribution of the broken pairs (for this term the case $E_m=E_n$ is excluded from the sum), while the second represents that of the condensed quasiparticles. The matrix elements of the stationary perturbations are contained in the quantities

$$T^+_{m\sigma}=u_nu_mh^{nm}_{\sigma\sigma}+u_nv_md^{n\bar{m}}_{\sigma\bar{\sigma}}+v_nu_md^{*\bar{m}n}_{\sigma\bar{\sigma}}-v_nv_mh^{*\bar{n}\bar{m}}_{\sigma\sigma} \tag{35}$$

and

$$T^-_{m\sigma}=u_nv_mh^{n\bar{m}}_{\sigma\sigma}-u_nu_md^{n\bar{m}}_{\sigma\bar{\sigma}}+v_nv_md^{*\bar{m}n}_{\sigma\bar{\sigma}}+v_nu_mh^{*\bar{n}\bar{m}}_{\sigma\bar{\sigma}}, \tag{36}$$

where

$$h^{nm}_{\sigma\sigma}=\frac{\sigma\hbar^2}{4im^2c^2}\langle n|[\nabla v\times\nabla]_z|m\rangle+\frac{iq\hbar}{mc}\langle n|\mathbf{A}\cdot\nabla|m\rangle, \tag{37}$$

$$h^{*nm}_{\sigma\sigma}=-h^{nm}_{\sigma\sigma}, \tag{38}$$

$$d^{n\bar{m}}_{\sigma\bar{\sigma}}=\frac{\sigma\hbar^2}{4im^2c^2}\langle n|[\nabla\Delta\times\nabla]_z|m\rangle+\langle n|\tilde{\Delta}|m\rangle, \tag{39}$$

$$d^{*\bar{m}n}_{\sigma\bar{\sigma}}=-\frac{\sigma\hbar^2}{4im^2c^2}\langle n|[\nabla\Delta^*\times\nabla]_z|m\rangle+\langle n|\tilde{\Delta}^*|m\rangle. \tag{40}$$

A is the vector potential of the static magnetic field **B**, and should not be confused with that of the light wave, $\mathbf{A}_L$. [Normally, $\mathbf{A}_L(t,\mathbf{r})\ll\mathbf{A}(\mathbf{r})$.] The matrix elements of the time-dependent perturbation enter through

$$C^m_{nn'}=2i[p_y^{n\bar{n}'}p_x^{m\bar{n}'}-p_x^{n\bar{n}'}p_y^{m\bar{n}'}], \tag{41}$$

which generalizes the term found in Eq. (33). The combination of momentum matrix elements in Eq. (41) is typical of dichroism and also appears in many approaches to dichroism in the normal state and in magnetically ordered materials.[2,34,35,40]

$$p(n,n')=u_nv_{n'}-v_nu_{n'} \tag{42}$$

and

$$l(n,n')=u_nu_{n'}+v_nv_{n'} \tag{43}$$

are coherence factors, with the BCS amplitudes u_n and v_n as defined in Eqs. (7) and (8). These are the coherence factors normally found in treatments of optical absorption in superconductors.[30,41] The additional factors u_nv_n, etc., in Eqs. (35) and (36) result from the effect of the coherence on the stationary perturbations.

We can rewrite Eq. (34) in a more compact form, by noting that, due to the ansatz (7), (8), the amplitudes u_n and v_n are solutions of the BdGE (not the SBdGE). They have the property that if

$$\begin{pmatrix}u_n(\mathbf{r})\\ v_n(\mathbf{r})\end{pmatrix}=\begin{pmatrix}u_n\\ v_n\end{pmatrix}\phi_n(\mathbf{r}) \tag{44}$$

is a solution with eigenvalue E_n, then

$$\begin{pmatrix}u_{\bar{n}}(\mathbf{r})\\ v_{\bar{n}}(\mathbf{r})\end{pmatrix}=\begin{pmatrix}-v_n\\ u_n\end{pmatrix}\phi_n(\mathbf{r})^*=\begin{pmatrix}-v_n\\ u_n\end{pmatrix}\phi_{\bar{n}}(\mathbf{r}) \tag{45}$$

is a solution with eigenvalue $-E_n$. [This is a well-known property of the BdGE, which is the counterpart to Eq. (6) for the SBdGE.]

Hence the second term under the sum on m in Eq. (34) can be written in the same form as the first term, with every index m replaced by $\bar{m}$. It represents the contribution of the negative energy states. We can therefore allow Σ_m to range over *all* m, not just over those with positive energies and find, after slight rearrangements,

$$\Delta P = \frac{\pi e^2 E_0^2}{m^2 \omega} \sum_{nn'\sigma} \sum_{m} \frac{p(n,n')p(m,n')}{E_n - E_m} \mathrm{Re}[C^m_{nn'} T^+_{m\sigma}] \times f(-E_{n\sigma}) f(-E_{n'\bar{\sigma}}) \delta(E_{n\sigma} + E_{n'\bar{\sigma}} - \hbar\omega), \quad (46)$$

where the sums on n and n' are still restricted to positive energies. In order to obtain this convenient and compact version of Eq. (34), the particle-hole convention is used only for E_n and $E_{n'}$, but not for E_m.

The arguments of the δ and Fermi functions, in Eqs. (34) or (46) imply that two quasiparticles, one with spin up, the other with spin down, are created by the absorption process. This expresses the fact that absorption of light does not induce spin flips, so that the original spin direction of the particles in the Cooper pair is not affected by breaking the pair. (This is in contrast to the paramagnetic pair breaking discussed below.)

C. Analytical results

Before explicitly evaluating the above formulas numerically for a simple model we first want to demonstrate that a number of general conclusions follow already from their structure.

1. Mechanisms for dichroism in superconductors

On the basis of Eqs. (34) or (46) we can identify several distinct mechanisms for dichroism in superconductors.

Mechanism 1 is the conventional spin-orbit coupling (SOC) produced by gradients of the lattice potential $\nabla v(\mathbf{r})$. The relevant matrix element for this mechanism is

$$\frac{\sigma\hbar^2}{4im^2c^2}\langle n|[\nabla v \times \nabla]_z|m\rangle, \quad (47)$$

which contributes to $h^{nm}_{\sigma\sigma}$ and $h^{*nm}_{\sigma\sigma}$ in Eqs. (37) and (38).

This mechanism is easy to interpret physically: the magnetic field breaks time-reversal invariance and leads to a finite spin magnetization. Since the initial single-particle states are occupied in pairs with total spin zero, the main contribution to this magnetization arises from the final states. The spin-orbit coupling converts the broken orientational symmetry in spin space into a broken chiral symmetry in real space. The polarization of the light is sensitive to the chiral symmetry of the system, hence dichroism arises.

This mechanism is known from normal and magnetically ordered metals to give rise to the Faraday and Kerr effects and to x-ray magnetic dichroism, respectively.[1,2] We thus find that the same mechanism is operative in superconductors as well, strongly modified, though, by the presence of the energy gap and the coherence factors. We present numerical results for this mechanism in Sec. III D.

From the above physical explanation it is seen that the presence of an (external or internal) magnetic field is a necessary condition for the mechanism to work. This is easily verified from the equations. In the absence of any type of magnetic field, the energies in Eq. (34) do not depend on the spin. The only spin dependence then comes from the factor σ in front of the matrix elements (47). The sum over the spins then reduces to

$$\sum_{\sigma=\pm 1} \sigma \equiv 0. \quad (48)$$

Hence, to have dichroism from the spin-orbit coupling, the presence of magnetic fields and thus of broken time-reversal invariance is mandatory. This is confirmed by many investigations on normal and magnetic materials.[1,2]

Mechanism 2 is the anomalous spin-orbit coupling (ASOC) produced by gradients of the pair potential $\nabla\Delta(\mathbf{r})$. The relevant matrix elements for this mechanism are

$$\frac{\sigma\hbar^2}{4im^2c^2}\langle n|[\nabla\Delta \times \nabla]_z|m\rangle \quad (49)$$

and the corresponding term containing Δ^* appearing in Eqs. (39) and (40). The ASOC term produces dichroism for the same reason as the SOC term and also requires the presence of magnetic fields. The temperature behavior of this term is very different from that of the SOC term because the pair potential itself is strongly temperature dependent. Moreover, the coherence factors for this mechanism, in Eqs. (35) and (36), are different from those for the SOC mechanism.

Since both the SOC and the ASOC mechanism depend on gradients of the respective potentials, they do not operate in homogeneous systems. On the other hand, the more inhomogeneous the system, the more important they become. The ASOC term, for instance, being produced by gradients of the pair potential, becomes large in the vortex state of a type-II superconductor and for superconductors with a short coherence length.

Mechanism 3 is provided by orbital currents, which flow in the presence of the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$. The relevant matrix element is

$$\frac{iq\hbar}{mc}\langle n|\mathbf{A}\cdot\nabla|m\rangle. \quad (50)$$

The physical interpretation is that these currents, circulating in the plane perpendicular to the magnetic field, give the material a definite handedness and hence result in dichroism. This mechanism is known also from the normal state, where it is usually much smaller than the SOC mechanism and therefore neglected in most calculations. In superconductors the screening currents which give rise to the Meissner effect are stronger than typical screening currents in normal metals. Current-induced dichroism may thus be stronger in superconductors than in normal conductors. Indeed, this type of dichroism in superconductors was already observed experimentally. We discuss these experiments in Sec. IV.

Since all of the above mechanisms require the presence of a magnetic field, we conclude that dichroism in the Meissner

phase is a surface effect because in the bulk the magnetic field does not coexist with superconductivity. The region in which coexistence is found depends on the penetration depth. The larger the penetration depth, the larger the region of the sample in which dichroism is produced. In the vortex phase the field penetrates into the bulk of the superconductor and coexists with the order parameter near the vortices. Every vortex is thus a source for dichroism due to each of the three above mechanisms.

Mechanism 4 is caused by inhomogeneities of the superconducting order parameter itself. The relevant matrix elements are

$$\langle n|\tilde{\Delta}|m\rangle \quad \text{and} \quad \langle n|\tilde{\Delta}^*|m\rangle. \tag{51}$$

To see how the order parameter can lead to a finite result for ΔP we first look at the quantity $C^m_{nn'}$, as defined by Eq. (41). As mentioned above, if the original normal-state Hamiltonian does not break time-reversal and inversion symmetry all momentum matrix elements are real. Under these circumstances $C^m_{nn'}$ is purely imaginary. Since in Eq. (34) the real part of the expression under the sum on m is taken, $C^m_{nn'}$ must be multiplied by a quantity with a finite imaginary part in order to lead to a finite ΔP. Since the matrix elements for the SOC, ASOC, and orbital mechanisms explicitly contain a factor i, this condition is normally satisfied in these cases. The matrix elements of the pair potential are easily seen to be purely real (and thus not to produce dichroism) if

$$\tilde{\Delta}^*(\mathbf{r}) = \tilde{\Delta}(-\mathbf{r}). \tag{52}$$

The condition (52) is violated if $\tilde{\Delta}(\mathbf{r})$ is real and breaks inversion symmetry (mechanism 4a), or if $\tilde{\Delta}(\mathbf{r})$ is symmetric under inversion, but complex, in which case it breaks time-reversal symmetry (mechanism 4b).

Order parameters which break inversion symmetry, while the underlying normal-state Hamiltonian does not, are examples of ''unconventional order parameters'' in the sense that the superconducting phase has different spatial symmetries from the normal phase.[42] The fact that broken inversion symmetry can give rise to dichroism is well known, e.g., in chemistry, where molecules without a center of inversion are called ''optically active'' since their optical properties depend on the polarization of the light. In principle, inversion symmetry is also broken at every surface. However, as long as the light penetrates sufficiently deeply into the material to sample the bulk properties, this contribution is expected to be very small. This expectation is corroborated by detailed investigations of dichroism in the normal state, where this type of surface-induced dichroism was not found to play an appreciable role in explaining experimental observations.

A complex order parameter is found in the presence of external magnetic fields, where the gradient of its phase corresponds to supercurrents.[14] This aspect of mechanism 4b is thus closely related to mechanism 3. Complex order parameters are also possible in the absence of magnetic fields, if the superconducting phase itself breaks time-reversal symmetry. This is the case, for example, for the ''$d+is$'' type of order parameter discussed in connection with the heavy fermions, the high-temperature superconductors, and anionic superconductivity.[42–44]

Dichroism produced by order parameters which break inversion- or time-reversal symmetry was discussed previously by several authors. In Ref. 45, a phenomenological analysis of this mechanism for the high-temperature superconductors is performed. Our analysis provides possible microscopic mechanisms underlying the phenomenological treatment of Ref. 45. In Ref. 46, the breaking of these symmetries is discussed in the context of collective modes of the order parameter (an effect we do not discuss in the present paper). In Ref. 47, the effect of broken chiral symmetry in the normal state of superconductors is analyzed. This effect follows in our framework if for the wave functions $\phi_n(\mathbf{r})$ one uses the eigenstates of a normal-state Hamiltonian with broken chiral symmetry. In this case the momentum matrix elements are complex, and dichroism is found from Eq. (33) already in zeroth order, i.e., even in the absence of the above four mechanisms.

Mechanism 5 is therefore provided by a normal state which already displays dichroism. In the present paper we focus on the superconducting state and do not discuss further this type of normal-state-induced dichroism.

In conclusion, we have identified five distinct mechanisms for dichroism in the superconducting phase. Two of these mechanisms, (1) and (3), are already known from the normal state, strongly modified, though, by the presence of the superconducting order parameter. Mechanisms (2), (4a), and (4b) on the other hand, exist only in superconductors. One mechanism of each type, namely (1) and (2), is of relativistic origin, arising from the interplay between relativistic symmetry breaking and superconducting coherence. Mechanism (5) operates only in superconductors in which the normal state already breaks time-reversal or inversion symmetry.

2. Existence criteria

We can extract from the general formula a number of ''existence criteria'' which determine under which circumstances one can expect dichroism at all. From the above discussion of the mechanisms we already have

Existence criterium 1: There is no dichroism due to spin-orbit coupling in the absence of magnetic fields.

Existence criterium 2: There is no dichroism due to BCS-type order parameters, i.e., order parameters which are real and spatially constant.

Existence criterium 3: Any perturbation whose contribution to Eqs. (37)–(40) is purely real does not give rise to dichroism.

The third criterium applies in particular to the normal and anomalous Darwin terms, containing $\nabla^2 v(\mathbf{r})$ and $\nabla^2\Delta(\mathbf{r})$. If the potentials $v(\mathbf{r})$ and $\Delta(\mathbf{r})$ are real and inversion symmetric, these terms lead only to real contributions under $\mathrm{Re}[i\ldots]$ in Eq. (34) and therefore do not produce dichroism. If $v(\mathbf{r})$ and $\Delta(\mathbf{r})$ break inversion symmetry or are complex, dichroism already arises from the zero-order wave functions [in the case of $v(\mathbf{r})$] or from the matrix elements of $\tilde{\Delta}$ in Eqs. (37)–(40). In this case, the Darwin terms are small additional corrections which do not break any further symmetries. Hence, the Darwin terms themselves are not a source of dichroism. A similar argument applies to the $n\mathbf{A}^2$ coupling of the external vector potential to the density. These conclusions had been anticipated, in discussing the relevant perturbations below Eq. (21).

A number of further criteria can be found from the various ingredients of Eq. (34). Since the minimum value the unperturbed energies can take is $\overline{\Delta}$, it follows from the arguments of the δ function that there is an absorption edge at $\hbar\omega=2\overline{\Delta}$. This edge is a consequence of the fact that we limited ourselves to the consideration of pair-breaking processes. For sufficiently low temperatures, where pair breaking is the only absorption process, we thus have

Existence criterium 4: At low temperatures there is no dichroism for frequencies below the absorption edge.

The SOC and ASOC mechanisms require the presence of gradients of the lattice- and pair potentials. However, not every gradient produces dichroism. In both cases we have a term of the form

$$[(\nabla f(\mathbf{r}))\times\nabla]_z=\frac{\partial f(\mathbf{r})}{\partial x}\frac{\partial}{\partial y}-\frac{\partial f(\mathbf{r})}{\partial y}\frac{\partial}{\partial x}, \tag{53}$$

where $f(\mathbf{r})$ is $v(\mathbf{r})$ or $\Delta(\mathbf{r})$, respectively. Obviously, the gradients in the z direction, i.e., the direction of the magnetic field and the incident light, do not enter. This leads to

Existence criterium 5: In the polar geometry ($\mathbf{B}\|\mathbf{q}$) dichroism arises entirely from the lateral (in-plane) inhomogeneity, not from the perpendicular inhomogeneity.

This observation is particularly relevant for surface geometries. As long as light incidence and magnetic field are perpendicular to the surface, the surface gradients themselves do not produce dichroism due to mechanisms 1 or 2 (SOC and ASOC), only the much smaller lateral gradients do. Mechanism 4b (broken inversion symmetry), on the other hand, is operative at surfaces, as inversion symmetry is necessarily broken. However, as mentioned above, this mechanism is expected to yield only a very small contribution, as long as inversion symmetry is not broken in the bulk of the material as well.

Finally, we note that all the energies in Eq. (34) are unperturbed energies. The stationary perturbations enter only through their effect on the wave functions, as determined by Eqs. (7) and (8). The reason for this is that in calculating ΔP to first order in the perturbations, the coefficient of the terms containing the energy shift is, of course, the difference of the *zero-order* matrix elements. This difference, specified in Eq. (33), is zero. Thus these contributions vanish identically. In second order in the stationary perturbations the energy shifts reappear. This gives us

Existence criterium 6: For sufficiently small perturbations (such that a first-order treatment is justified) perturbations leaving the zero-order wave functions unchanged do not give rise to dichroism.

This existence criterium also provides a different point of view of the above discussed fact that the Zeeman ($\mathbf{m}\cdot\mathbf{B}$) coupling of a constant magnetic field to the electron spins does not produce dichroism without the simultaneous presence of SOC, while the $\mathbf{j}\cdot\mathbf{A}$ coupling does: Namely, unlike the coupling to the orbital currents, the Zeeman term alone does not change the form of the wave functions.

3. The normal-state limit

In the normal state, the pair potential $\Delta(\mathbf{r})$ vanishes identically. Therefore the matrix elements of the ASOC term and of $\tilde{\Delta}$ are zero, while $C^m_{nn'}$ is unaffected. The transition we consider is from an occupied state n' (which in the normal state lies below the Fermi surface) to an unoccupied state n (lying above the Fermi surface in the normal state). In the particle-hole convention the energies of states below the Fermi surface are interpreted as hole energies and taken to be positive. It is only the quantum number n which distinguishes between occupied and unoccupied states.

Using this convention, the $\Delta\to 0$ limit of the coherence factors can be evaluated straightforwardly. Explicitly we have

$$\begin{aligned} p(n,n')&=u_n v_{n'}-v_n u_{n'}\\ &=\sqrt{\frac{1}{2}\left(1+\frac{\epsilon_n}{E_n}\right)}\sqrt{\frac{1}{2}\left(1+\frac{\epsilon_{n'}}{E_{n'}}\right)}\\ &\quad-\sqrt{\frac{1}{2}\left(1-\frac{\epsilon_n}{E_n}\right)}\sqrt{\frac{1}{2}\left(1-\frac{\epsilon_{n'}}{E_{n'}}\right)}\\ &\to 1. \end{aligned} \tag{54}$$

Similarly we find

$$p(m,n')\to\begin{cases}1 \text{ if } |m\rangle \text{ is above } \epsilon_F\\ 0 \text{ if } |m\rangle \text{ is below } \epsilon_F\end{cases}, \tag{55}$$

$$l(m,n')\to\begin{cases}0 \text{ if } |m\rangle \text{ is above } \epsilon_F\\ 1 \text{ if } |m\rangle \text{ is below } \epsilon_F\end{cases}. \tag{56}$$

As $\Delta\to 0$, Eq. (34) therefore becomes

$$\begin{aligned}\Delta P^{(N)}=&\frac{\pi e^2E_0^2}{m^2\omega}\sum_{nn'\sigma} f(-\epsilon_{n\sigma})f(-\epsilon_{n'\bar{\sigma}})\delta(\epsilon_{n\sigma}+\epsilon_{n'\bar{\sigma}}-\hbar\omega)\\ &\times\sum_m \mathrm{Re}\left[C^m_{nn'}\frac{h^{nm}_{\sigma\sigma}}{\epsilon_n-\epsilon_m}+C^{\bar{m}}_{nn'}\frac{h^{n\bar{m}}_{\sigma\sigma}}{\epsilon_n+\epsilon_m}\right].\end{aligned} \tag{57}$$

Similar to the superconducting case, the second term in the sum on m represents the contribution of the single-particle states below the Fermi surface, while the first term represents that of states above it. This formula describes dichroism due to the creation of particle-hole excitations in the normal metal. It is readily verified that the same result is obtained in a normal-state calculation: if the unperturbed Hamiltonian is taken to be Eq. (4), simple first-order perturbation theory leads to Eq. (57). The $\Delta\to 0$ limit of our formula thus correctly reproduces the corresponding normal-state result. Results for nonsuperconducting materials, which are of the type of Eq. (57), were derived previously by several authors.[34,40,48,49]

Many of the above-mentioned conclusions about mechanisms and existence criteria are valid also in the normal state (where they are mostly well known). We just mention a few: (i) There is no effect to zero order in the stationary perturbations, even in the presence of polarized light and Zeeman splitting. (ii) The constraints on the geometry of the experiment are the same. (iii) SOC alone, without Zeeman splitting of the energies, does not lead to dichroism. (iv) The argument that the Darwin term and the $\mathbf{A}^2$ term do not, on their

own, produce dichroism, remains valid. (v) The above-mentioned result that the effect arises to first order only from the change of the wave functions due to the perturbation and not from the change in the energies still holds. (In higher-order perturbation theory, of course, the energy shifts reappear.) Conclusions (i)–(iv) are well known in the theory of dichroism in normal and magnetically ordered metals, but are usually arrived at using more complicated methods. The dominance of the change of the wave function over that of the energies [conclusion (v)] was noted by several authors, who arrived at this or very similar conclusions in a variety of ways.[40,48,50–52]

III. MODEL CALCULATIONS

A. A model for the SOC mechanism

After discussing general consequences of the results we now turn to their approximate numerical evaluation. A full evaluation of Eq. (34) requires self-consistent solutions of the SBdGE in the presence of magnetic fields and remains a project for the future. Here we only look for a simple model which illustrates the main physical aspects of the theory. To this end we specialize to mechanism 1.

By imposing a number of physical conditions on the superconductor under study we can make sure that the contributions of the remaining mechanisms are small. Mechanisms (2) and (4) do not contribute if the pair potential is real and spatially constant, i.e., for order parameters of the type considered in the original BCS model. Moreover, in materials with heavy elements (high atomic number Z) in the lattice, the SOC (which increases roughly as Z^4) is strongly enhanced, as compared to low Z materials.

For superconductors with constant pair potential and heavy elements in the lattice, in sufficiently weak external fields, we can therefore limit ourselves, for the purpose of a model calculation, to the SOC mechanism (1). We thus neglect the matrix elements for the other mechanisms and keep only

$$h^{nm}_{\sigma\sigma}=\sigma M_{nm}[v] \tag{58}$$

with

$$M_{nm}[v]:=\frac{1}{4im^2c^2}\langle n|[\nabla v\times\nabla]_z|m\rangle \tag{59}$$

in the equations. (In this section we set $\hbar\equiv 1$.) The various coherence factors can all be expressed in terms of the function

$$p(E,E'):=\sqrt{\frac{1}{2}\left(1+\frac{\epsilon}{E}\right)}\sqrt{\frac{1}{2}\left(1+\frac{\epsilon'}{E'}\right)} -\sqrt{\frac{1}{2}\left(1-\frac{\epsilon}{E}\right)}\sqrt{\frac{1}{2}\left(1-\frac{\epsilon'}{E'}\right)}, \tag{60}$$

which in the respective energy intervals reduces to the appropriate coherence factors. ϵ is defined through

$$\epsilon:=\sqrt{E^2-\Delta^2}. \tag{61}$$

Finally, we define

$$F(E,E',H):=f(-E_\uparrow)f(-E'_\downarrow)-f(-E_\downarrow)f(-E'_\uparrow). \tag{62}$$

By introducing the density of states, $N(E)$, in the superconducting state (SDOS), we can convert the sums in Eq. (46) in integrals. One of the integrals can immediately be performed due to the δ function. Utilizing the above defined quantities, Eq. (46) is written as

$$\begin{aligned}\Delta P=&\Theta(\omega-2\Delta)\langle CM\rangle_F\frac{\pi e^2E_0^2}{m^2\omega}\\ &\times\int_\Delta^{\omega-\Delta}dE_1\left[\int_{-\infty}^{-\Delta}dE_2+\int_\Delta^{\infty}dE_2\right]\frac{F(E_1,\omega-E_1,H)}{E_1-E_2}\\ &\times p(E_1,\omega-E_1)p(E_1,E_2)p(E_2,\omega-E_1)\\ &\times N(E_1)N(E_2)N(\omega-E_1),\end{aligned} \tag{63}$$

where the E_1 integration arises from the sum on n and the E_2 integration from the sum on m. The signature of superconductivity in this equation is the appearance of the coherence functions and the SDOS. The matrix elements M_{nm} and $C^m_{nn'}$, on the other hand, are *normal-state* matrix elements. We have approximated these by their average $\langle CM\rangle_F$ over the Fermi surface. Since superconductivity happens essentially within a few meV around the Fermi surface, this is a very reasonable approximation. The above procedure has the advantage that the remaining two integrals can be evaluated numerically, once a model for the SDOS is chosen. By forming the ratio of the superconducting result to its normal-state limit,

$$\frac{\Delta P^S}{\Delta P^N}:=\frac{\Delta P}{\lim_{\Delta\to 0}\Delta P}, \tag{64}$$

the average matrix elements cancel and we obtain a direct measure for the interplay between the superconducting coherence and spin-orbit coupling. The same strategy is also applied in a large number of similar model calculations for other properties of superconductors, e.g., for the absorption of sound,[30,41] the nuclear-spin relaxation rate,[53] the thermal conductivity,[54] the spin susceptibility,[55] the absorption of unpolarized light,[31] etc.

B. Qualitative analysis of the SOC mechanism

The simple form of Eq. (63) allows us to draw some further conclusions about the nature of SOC-induced dichroism in superconductors, before proceeding to its numerical evaluation. Close to $T=0$ we can replace the pair potential $\Delta(T)$ by its zero-temperature value $\Delta(0)$, and the Fermi functions in $F(E_1,\omega-E_1,H)$ by step functions, according to $f(-E_{n\sigma})=f(-E_n-\sigma\mu_BH)\to\Theta(E_n+\sigma\mu_BH)$. These step functions impose restrictions on the integration limits of the E_1 integral, which can symbolically be written as

$$\int_{\Delta}^{\omega-\Delta} dE_1 F(E_1,\omega-E_1,H)\cdots \rightarrow \int_{\Delta}^{\mathrm{Min}[\omega-\Delta,\omega-\mu_B H]} dE_1\cdots - \int_{\mathrm{Max}[\Delta,\mu_B H]}^{\omega-\Delta} dE_1\cdots. \tag{65}$$

In the (usual) case that $\mu_B H<\Delta$ we immediately see that the two integrals on the right-hand side of Eq. (65) cancel each other exactly. If $\mu_B H$ exceeds Δ there is no such cancellation. There is thus no SOC-induced dichroism in superconductors at $T=0$, unless the magnetic field is of the order of magnitude of the energy gap. Physically, this can be understood as follows: at $T=0$ all electrons are condensed in pairs with net spin zero. The system is thus not spin polarized, even for finite magnetic fields. In this case the symmetry in spin space is not broken and hence SOC cannot induce broken chiral symmetry. Hence, no dichroism results. A more general viewpoint is that time-reversal invariance must be broken to have dichroism from the SOC mechanism. Since, by construction, the ground state of a superconductor at $T=0$ consists of pairs of mutually time-conjugate states, it is invariant under time reversal, even in the presence of finite fields, so that no dichroism can arise. If, on the other hand, the magnetic field is strong enough to break Cooper pairs paramagnetically, there exist unpaired electrons even at $T=0$. In this case one finds a finite spin polarization (i.e., a ground-state breaking time reversal) at $T=0$ and dichroism results. Thus, for $T=0$ and $\mu_B H<\Delta$ there is no dichroism, while for $T=0$ and $\mu_B H\gtrsim\Delta$ there is a finite amount of dichroism, in accordance with the above analytical findings. This is a direct manifestation of paramagnetic pair breaking, a phenomenon which is hard to observe experimentally by other means. Of course, the details of the phase transition induced by paramagnetic pair breaking cannot be described with our simple model for the unperturbed superconductor and our approximate treatment of the matrix elements. In particular, real superconductors display paramagnetic limiting already at $H=\Delta/\sqrt{2}$ and not at $H=\Delta$.[30] The factor of $\sqrt{2}$ follows from a comparison of total ground-state energies and is not contained in our single-particle model. The mere fact, however, that there is no spin-orbit-induced dichroism at $T=0$ until the field is of the order of magnitude of the gap, is a very definite consequence of our results.

Interestingly, if we perform the corresponding analysis for the orbital current mechanism (3), we find that the two integrals in Eq. (65) are not subtracted, but added, hence leading to a finite result even for $T=0$ and $\mu_B H\ll\Delta$. Physically, this can be interpreted by noting that even the smallest magnetic field gives rise to orbital currents in the superconductor, regardless of temperature. *The difference between the orbital and the spin response of a superconductor thus has a decisive influence on the low-temperature behavior of dichroism in superconductors*.

We now return to the case of finite temperature and investigate the behavior of ΔP as a function of the (external or internal) magnetic field. The field enters only through $F(E,E',H)$, which can be expanded, for sufficiently small H, about $H=0$. One finds

$$\Delta P\propto H[f'(-E)f(E-\omega)-f'(E-\omega)f(-E)]+O(H^2), \tag{66}$$

where $f'(E)$ stands for the derivative of the Fermi function. Hence, for small fields, ΔP, and the imaginary part of the off-diagonal elements of the conductivity tensor for superconductors, are linear functions of H. The absence of the zero-order term means that the SOC mechanism alone will not give rise to dichroism without magnetic fields. This is the same conclusion we already arrived at in Sec. II C 1 by following a somewhat different line of reasoning.

C. A BCS superconductor in the Meissner phase

Now we turn to the numerical evaluation of Eq. (63). For this purpose we have to assume a particular form for the SDOS. The simplest choice, the BCS form

$$N_{\mathrm{BCS}}(E)=N_N\frac{E}{\sqrt{E^2-\Delta^2}}, \tag{67}$$

where N_N is the DOS of the normal metal, leads to unphysical singularities at $E=\Delta$. In every real superconductor these singularities are smoothed by a large number of processes, such as gap anisotropy, residual interactions with phonons, strong-coupling effects, inelastic scattering from impurities, etc.[30,53,56,57] Many parametrizations of the SDOS which take these effects into account have been suggested in the literature.[41,53,56–58]

In the following, we model the SDOS in a way similar to that of Hebel and Slichter in their seminal work on NMR in superconductors.[53,57,58] The SDOS is written as a weighted average over the BCS-DOS,

$$N(E):=\mathrm{Re}\int_{-\infty}^{+\infty} dE' N_{\mathrm{BCS}}(E')w(E,E'). \tag{68}$$

The weighting function $w(E,E')$ can take a variety of forms, but in the case of NMR experiments on BCS superconductors the simple square form

$$w(E,E')=\frac{1}{2\delta}\Theta(E'-(E-\delta))\Theta((E+\delta)-E'), \tag{69}$$

which makes w constant for E' between $E-\delta$ and $E+\delta$, and zero everywhere else, already leads to quantitative agreement with experiments.[53,56–58] The broadening δ is typically chosen to be $\delta\approx 0.1\Delta(T)$. Equations (68) and (69) convert the singularity into a peak. In the case of NMR this peak, in conjunction with the coherence factors, gives rise to the well-known Hebel-Slichter peak in the nuclear-spin relaxation rate.[30,53,57,58] The form, Eqs. (68) and (69), of the SDOS can be physically justified in terms of the anisotropy of the gap,[53,56–58] but for the present purpose it can also be regarded as just a simple phenomenological model. Apart from this simple form we also considered several more sophisticated expressions for the SDOS, but it turned out that its detailed form does not qualitatively (and quantitatively only up to within 5–10 %) affect our results, as long as the singularity is smoothed out in whatever fashion. We have therefore chosen to work with Eqs. (68) and (69) for numerical simplicity.

In order to investigate dichroism as a function of temperature, we also need to specify the temperature dependence of the pair potential. For BCS superconductors this is found by interpolating between two analytically known limiting cases. Close to $T=0$ we use the formula derived in Ref. 59 in terms of Bessel functions, while close to $T=T_c$ we employ the parametrization of Ref. 56 which very close to $T=T_c$ reduces to the standard BCS square root form. (More details can be found in these references.)

In the actual calculation we need to assign physically reasonable values to the parameters entering Eq. (63). For the zero-temperature pair potential we choose a value of $\Delta(T=0)=1$ meV. Assuming a BCS-type relation between the pair potential at $T=0$ and the critical temperature this leads to $T_c=6.6$ K, which is close to that of Pb. The normal-state DOS (NDOS) is assumed to be constant between -0.3 and $+0.3$ eV and zero everywhere else. This corresponds to a square-shaped NDOS centered around the Fermi energy. The absolute value of the NDOS does not enter because it cancels when forming the ratio (64). Of course, all these numbers can be modified easily, without changing any of our conclusions in an essential way.

The assumptions made and the numerical values chosen, specify the superconductor under consideration to be a BCS-type s-wave superconductor in the Meissner phase. Thus the dichroism we discuss in the following is induced in the region where the magnetic field penetrates. The more this superconductor is type II (i.e., the larger the ratio of penetration depth to coherence length), the larger is the region of the sample where the magnetic field and the order parameter coexist and dichroism results. We therefore assume the model superconductor to be a strong type-II superconductor. [Note that this assumption does *not* underly the general theory leading to Eq. (34)]. Since under these circumstances the penetration depth is much larger than both the coherence length and the lattice constant, the light probes the bulk of the superconductor. At this point we therefore neglect any surface-specific sources for dichroism and use typical bulk values for the energy gap, the SDOS, etc.

In concluding this section we stress that our calculations are meant only to illustrate the physics of the SOC mechanism and not to provide quantitative predictions for real superconductors. For the latter purpose, a self-consistent solution of the SBdGE is required.

D. Numerical results

The integrations in Eq. (63) are performed numerically. In the following subsections we display the results of these calculations as a function of temperature, magnetic field, and frequency.

For each of these we first present the result for a normal conductor. Although the corresponding calculations can be done for a normal conductor in a self-consistent and fully relativistic manner,[1,2] we perform them using the same approximations as discussed above for the superconductor, in order to facilitate the comparison with the results for the superconductor. Strictly speaking, the calculation for a normal conductor loses its meaning below T_c. However, as compared to the superconductor, the response of the normal conductor does not vary significantly with temperature. In any case, it constitutes a convenient normalization. We display the result in arbitrary units, because the prefactors are determined by the absolute values of the matrix elements M_{nm} and $C^m_{nn'}$ and the NDOS at the Fermi surface, none of which can be calculated within the simple model of the present section.

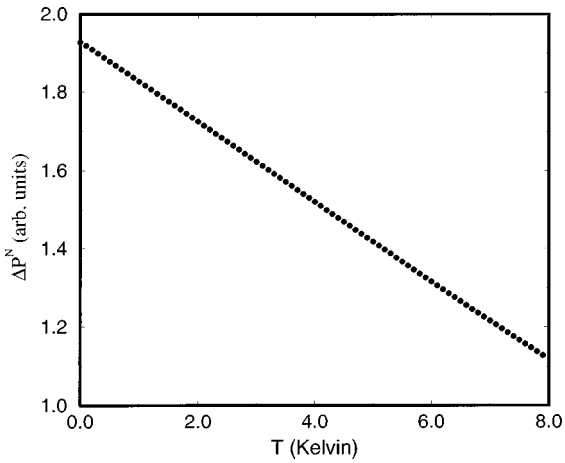

FIG. 1. Dichroism in the normal state vs temperature T. This and all other figures display only the contribution of SOC-induced dichroism. The other mechanisms are excluded from the calculation, as discussed in the main text. The numerical values of the parameters specifying the system are given in the main text.

We also show the ratio of the superconducting results to those for the normal conductor. Here the matrix elements cancel and no ambiguity is left. We then have a direct illustration of the change in the response of the metal to polarized light due to the presence of the superconducting coherence. This strategy, of course, is standard, and was used successfully in many similar calculations [cf. the references below Eq. (64)].

1. Dependence on temperature

Figure 1 shows that dichroism in the normal conductor increases with decreasing temperature. Figure 2 shows the corresponding calculation for the superconductor, divided by the values for the normal conductor. The calculations were

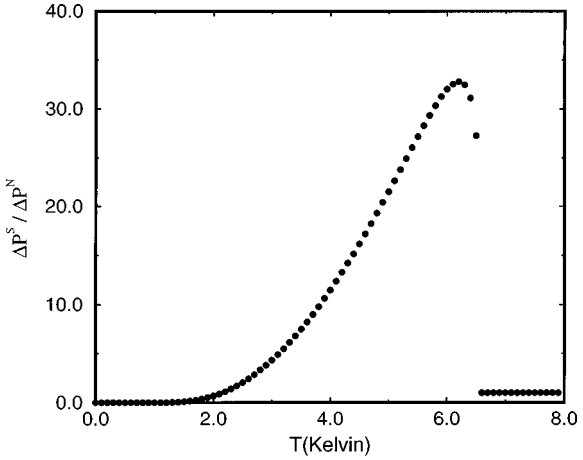

FIG. 2. Dichroism ratio vs T at small magnetic fields. A strong coherence peak is seen close to T_c, while near $T=0$ the curve approaches 0.

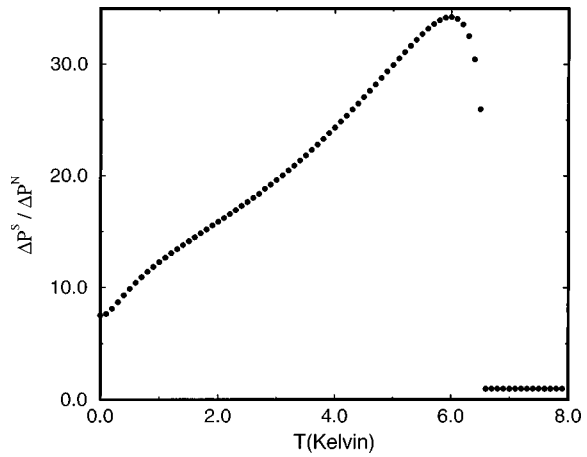

FIG. 3. Dichroism ratio vs T at large magnetic fields. Paramagnetic pair breaking leads to a finite value at $T=0$.

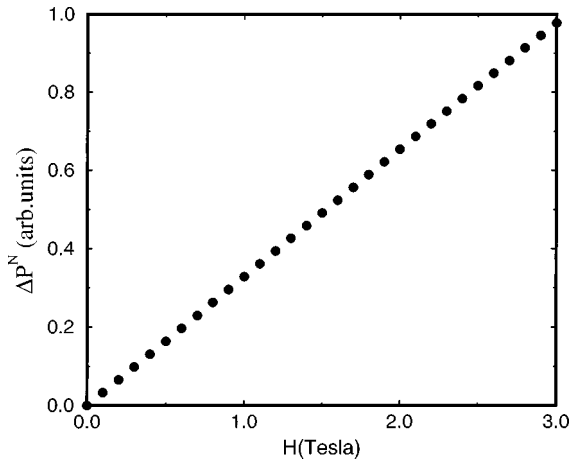

FIG. 4. Dichroism in the normal state vs magnetic field H at finite temperature. ΔP rises almost exactly linearly with H.

done for a frequency of 4.5 meV and a magnetic field of 0.05 T. Figure 3 repeats Fig. 2, but for a magnetic field close to the paramagnetic limit of superconductivity. (This plot corresponds to a superconductor in which the critical field for orbital pair breaking H_{c2} is high enough that the paramagnetic limit is actually reached.)

Since $T_c=6.6$ K, the data points between 6.6 and 8 K are the same in Figs. 1 and 2. The above statement that the response of the normal conductor does not change very much, if compared with the superconducting response, is immediately verified.

Clearly, the dominating feature of Fig. 2 is the huge peak just below T_c. Similar peaks in other observables arise from the peak in the SDOS in conjunction with the coherence factors. These peaks are known as Hebel-Slichter or coherence peaks.[30,41] However, the conventional Hebel-Slichter peak, as seen, e.g., in NMR experiments,[53,57,58,60] typically reaches values in the range (1–5), while the peak in Fig. 2 goes up to almost 33.

Some part of this additional enhancement may be due to the oversimplified approximations for the matrix elements, the NDOS and the SDOS. On the other hand, the approximations for the matrix elements are the same in the superconductor and in the normal conductor and do not lead to peaks in the latter case. Furthermore, different models for the SDOS lead to enhancements of the same order of magnitude.

This suggests that there is a second mechanism at work which is responsible for the additional enhancement of the peak. Numerically, this effect arises from the regions of integration in Eq. (63) which are excluded due to the energy gap in the superconductor, but contribute to the integrals for the normal conductor. Performing a numerical experiment, we can integrate the normal conductor with the integration limits of the superconductor. It turns out that it is mainly the limits of the inner integration (over E_2) which in this case lead to a strong enhancement of the normal-conducting results as well.

This effect is independent of the form of the DOS and thus not the same as the conventional Hebel-Slichter mechanism. It can be interpreted by noting that, apart from producing the peak in the SDOS, the gap also acts as an absorption edge which is present in the superconductor but not in the normal conductor. Experiments and theoretical calculations for magnetically ordered materials show that dichroism is always greatly enhanced near absorption edges.[1,2] The above numerical experiment and the additional enhancement of dichroism in superconductors, seen in Fig. 2, indicate that this is the case for superconductors as well.

We can thus identify *two* independent ways in which dichroism in superconductors is enhanced as compared to the normal conductor: The pileup in the density of states, giving rise to the Hebel-Slichter-type enhancement and the gap itself, producing an absorption edge. Both the original Hebel-Slichter mechanism and the additional ''gap-enhancement'' crucially depend on the existence of an energy gap. For superconductors in which the energy gap is zero on some parts of the Fermi surface, such as in d-wave superconductors, they will be strongly reduced or even not present at all.[61]

The physics at low temperatures is very different from that at higher temperatures. By comparing the low-temperature behavior of Figs. 1, 2, and 3 we can verify the analytical results obtained in Sec. III B. In the normal conductor and the superconductor in the realm of paramagnetic pair breaking there is a finite amount of SOC-induced dichroism at $T=0$, while in the superconductor at lower fields dichroism is quenched at $T=0$.

2. Dependence on the magnetic field

Figure 4 displays the behavior of dichroism in the normal conductor as a function of the magnetic field. In accordance with the discussion in Sec. III B it starts from zero at $H=0$ and rises almost linearly. (For fields about 10 times as large as in the figure a slight deviation from linearity is found.) Figure 5 illustrates that, although both the normal and the superconductor are approximately proportional to H, their ratio is not constant. This is due to the higher-order terms in H. Both figures are for a temperature of 2.5 K and a frequency of 4 meV.

Figure 6 displays the ratio for the case of zero temperature and very large magnetic fields. All data points at $H<17.5$ T are zero, while above 17.5 T suddenly finite values show up. This reflects the paramagnetic limit, as discussed above. Indeed, $H=17.5$ T corresponds to an energy of 1 meV, which is just the value chosen for the pair potential at $T=0$. Super-

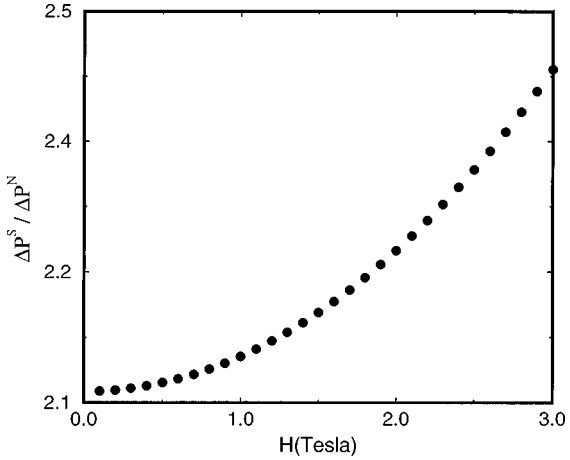

FIG. 5. Dichroism ratio vs H at finite temperature. The superconductor displays a stronger variation with the magnetic field as compared to the normal conductor.

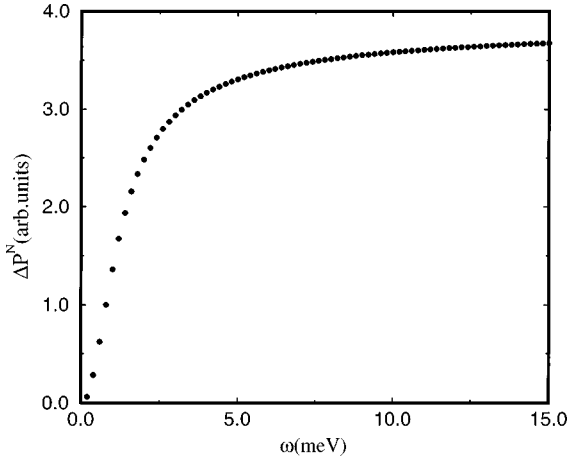

FIG. 7. Dichroism in the normal state vs frequency ω.

conductors where the upper critical field is thought to be influenced by paramagnetic limiting are, e.g.,[62] the heavy-fermion compounds UBe_{13} and $CeCu_2Si_2$, the Chevrel phase material $Gd_{0.2}PbMo_6S_8$, the A-15 superconductor $Nb_3Al_{0.75}Ge_{0.25}$ and thin films of, e.g., Al or Sn.[63,64] While plausible, this is not supported by direct experimental evidence. Clearly, under such circumstances, the observation of dichroism at $H \sim H_{c2}$, as in Fig. 6, could be decisive.

3. Dependence on frequency

Figure 7 shows ΔP in a normal conductor versus frequency. At $\omega = 0$ there is, of course, no dichroism because no transitions can take place. At higher frequencies ΔP approaches an almost constant value. This reflects the featureless NDOS which was used in the calculations (namely its average value at the Fermi surface).

The corresponding plot for the ratio, Fig. 8, displays an absorption edge at $\omega = 2\Delta$. This edge is due to our limitation to pair-breaking processes. At low temperatures very few excited quasiparticles are present and pair breaking is the only available mechanism for absorption. The inclusion of scattering from excitations (i.e., broken pairs) is known from investigations of the absorption of unpolarized light,[30] to give rise to some additional absorption below the edge, which does not significantly affect the part of the curve due to pair breaking.

The shape of the peak directly above the edge reflects the behavior of the perturbation, acting on the system, under time reversal (cf. the discussion in Ref. 30, in particular Fig. 2-9). The curve in Fig. 8 is of a mixed type, which reflects that we have two perturbations acting on the system: the spin-orbit coupling, which is even under time reversal, and the magnetic field, which is odd.

Both plots were done for a temperature of 3 K (well below the strong peak seen in Fig. 1) and a (relatively strong) magnetic field of 0.1 T.

IV. EXPERIMENTAL ASPECTS

A number of experimental results on dichroism in the vortex state of high-temperature superconductors are available.[3,4] Below we offer some speculative remarks on the interpretation of these experiments on the basis of the present theory.

First, we note that the difference LHP-RHP, which can be

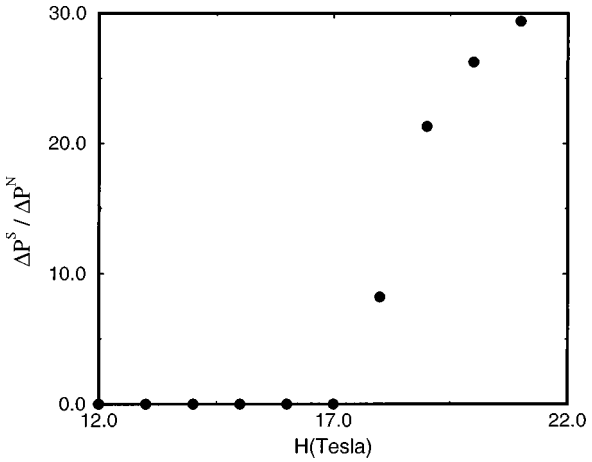

FIG. 6. Dichroism ratio vs H at zero temperature. A finite-spin polarization is not produced until the magnetic field is comparable to the energy gap.

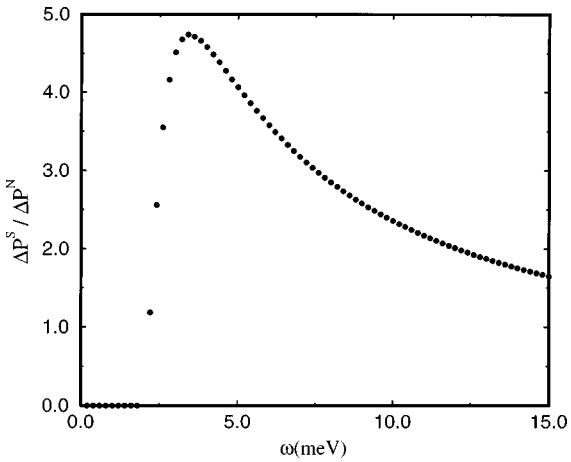

FIG. 8. Dichroism ratio vs ω. The absorption edge at $\omega = 2\Delta$ has a mixed type I-II character, reflecting the behavior of the perturbations under time reversal.

read off from Fig. 3 of Ref. 4, displays a strong enhancement between $T \sim T_c$ and $T \sim T_c/2$, which is consistent with our general prediction of enhanced dichroism below T_c.

This difference was theoretically analyzed on the basis of the cyclotron motion of the orbital currents circulating around the vortex core.[25,65] In our framework this corresponds to mechanism (3). The orbital mechanism can be distinguished experimentally from the SOC mechanism because, as explained in Sec. III B, it leads to a finite value for ΔP and $\mathrm{Im}[\hat{\sigma}_{xy}]$ at $T=0$. By contrast, the SOC-induced dichroism vanishes at $T=0$, as demonstrated analytically in Sec. III B and numerically in Sec. III D 1. The data of Ref. 4 do indeed extrapolate to a finite value at $T=0$. The present analysis thus supports the assumptions made, as to the mechanism responsible for the dichroism, in evaluating the above-mentioned experiments.

There are at least two features of these experiments which suggest that the spin-orbit mechanisms are present as well. First, the observed difference LHP-RHP displays changes of sign[3,4] between $T=T_c$ and $T=0$. Such a behavior finds a natural explanation in the presence of two distinct mechanisms for dichroism which have a different temperature behavior, but are operative for roughly the same values of temperature, magnetic field, and frequency. This interpretation is further supported by the fact that, in order to phenomenologically fit their measured data, the authors of Refs. 3 and 4 had to assume *two* independent Lorentzian oscillators. Evidently, coexistence of the SOC or the ASOC mechanism (1) and (2), with the orbital mechanism (3), provides a possible explanation for these observations.

Second, as stated in Ref. 3, the orbital mechanism alone does not fully explain the observations because there are indications for ''other (electronlike) chiral resonances in addition to the simple cyclotron resonance.''[3] Clearly, the spin-orbit mechanisms SOC and ASOC can play the role of these electronlike chiral resonances.

In a different set of experiments[5–7] spontaneous dichroism in high-temperature superconductors, both above and below T_c, was observed to exist at frequencies much higher than those considered in the present paper. In view of this difference in frequency it is unlikely that these effects are produced by the pair-breaking contribution to the SOC, ASOC, or orbital mechanisms discussed above. The fact that dichroism is observed also above T_c points to mechanism 5 (normal-state dichroism which continues to be present below T_c), in which case the energy scale is not set by the energy gap, but by normal-state parameters. More recent experiments[8] indeed suggest that this is the correct explanation.

Although the results obtained in this paper are thus consistent with all available experiments, more detailed experimental and theoretical investigations are called for, in order to identify the various mechanisms analyzed in this paper in an unequivocal way.

The SOC mechanism may be easiest to identify in superconductors which satisfy as many of the following criteria as possible: (i) heavy atoms in the lattice (this favors the conventional SOC, which rises approximately as Z^4), (ii) no zeros of the energy gap (such zeros can eliminate the coherence peak), (iii) short coherence length and large penetration depth (to maximize the region in which the order parameter and the magnetic field are present simultaneously), (iv) weak shielding currents (to minimize the influence of the orbital mechanism). These conditions are approximately satisfied, e.g., by lead.

For the ASOC mechanism to be observed we suggest: (i) light atoms in the lattice (to minimize the effect of the SOC term), (ii) large gradient of the pair potential (i.e., large energy gap and short coherence length), (iii) a large penetration depth, (iv) weak shielding currents, as for the SOC mechanism. It should be stressed that the observation of the ASOC mechanism would be of general significance because it could confirm or reject the form of the relativistic BCS Hamiltonian which has been proposed only recently[9–11] and is not yet experimentally verified.

The pair potential mechanism (4) is unique in that it does not depend on the existence of an (external or internal) magnetic field. The observation of dichroism below T_c in the absence of such fields would be a strong hint at mechanism (4).

V. CONCLUSION AND OUTLOOK

We have presented a perturbative approach to the absorption of polarized light in superconductors. Several distinct mechanisms for dichroism in superconductors were identified and interpreted. These are (1) the conventional spin-orbit coupling, (2) the anomalous spin-orbit coupling, (3) orbital currents, (4a) complex order parameters, (4b) inversion symmetry-breaking order parameters, and (5) a normal state which already displays dichroism.

Using perturbation theory for the spin-Bogolubov–de Gennes equations, we derived a general formula which contains the contributions of all these mechanisms. On the basis of this result several analytical conclusions concerning these mechanisms were drawn. These include an investigation of the circumstances under which the various mechanisms can be operative and an analysis of their behavior as a function of temperature and magnetic field.

The important role played by time-reversal symmetry is pointed out. While a Cooper pair consists of two mutually time conjugate single-particle states and the ground state of a superconductor is thus invariant under time reversal, several mechanisms for dichroism require the breaking of this symmetry.

For more detailed illustrations we have chosen the mechanism based on the conventional spin-orbit coupling because it is known to be the dominant mechanism in the normal state. Numerical calculations for a simple model superconductor confirm the analytical results and provide first hints at experimental signatures of SOC-induced dichroism in superconductors.

The influence of paramagnetic limiting on dichroism at low temperatures and high fields, unconventional order parameters as a source for dichroism, the quenching of spin-orbit-induced dichroism at $T=0$, and an additional enhancement of the coherence peaks are definite predictions of our theory which can be verified experimentally. None of these phenomena show up in the absorption of unpolarized light in superconductors or in dichroism in the normal state.

ACKNOWLEDGMENTS

We want to thank M. Lüders, R. Kümmel, H. Plehn, and H. Ebert for very useful discussions about various aspects of this work. This work has benefited from collaborations within, and has been partially funded by, the HCM network *Ab-initio (from electronic structure) calculation of complex processes in materials (Contract: ERBCHRXCT930369)*, and the program *Relativistic effects in heavy-element chemistry and physics* of the European Science Foundation. Partial financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

[1] *Spin-Orbit Influenced Spectroscopies of Magnetic Solids*, edited by H. Ebert and G. Schütz (Springer, Heidelberg, 1996), and references therein.
[2] H. Ebert, Rep. Prog. Phys. **59**, 1665 (1996).
[3] H.-T. S. Lihn, S. Wu, H. D. Drew, S. Kaplan, Q. Li, and D. B. Fenner, Phys. Rev. Lett. **78**, 3810 (1996).
[4] S. Wu, S. G. Kaplan, H.-T. S. Lihn, H. D. Drew, S. Y. Hou, J. M. Phillips, J. C. Barbour, E. L. Venturini, Q. Li, and D. B. Fenner, Phys. Rev. B **54**, 13 343 (1996).
[5] K. B. Lyons, J. F. Dillon, Jr., E. S. Hellman, E. H. Hartford, and M. McGlashan-Powell, Phys. Rev. B **43**, 11 408 (1991).
[6] K. B. Lyons, J. Kwo, J. F. Dillon, Jr., G. P. Espinosa, M. McGlashan-Powell, A. P. Ramirez, and L. F. Schneemeyer, Phys. Rev. Lett. **64**, 2949 (1990).
[7] H. J. Weber, D. Weitbrecht, D. Brach, A. L. Shelankov, H. Keiter, W. Weber, Th. Wolf, J. Geerk, G. Linker, G. Roth, P. C. Splittgerber-Hünnekes, and G. Güntherodt, Solid State Commun. **76**, 511 (1990).
[8] T. W. Lawrence, A. Szöke, and R. B. Laughlin, Phys. Rev. Lett. **69**, 1439 (1992).
[9] K. Capelle and E. K. U. Gross, Phys. Lett. A **198**, 261 (1995).
[10] K. Capelle and E. K. U. Gross (unpublished).
[11] K. Capelle and E. K. U. Gross (unpublished).
[12] E. Engel, H. Müller, C. Speicher, and R. M. Dreizler, in *Density-Functional Theory*, Vol. 337 of NATO *Advanced Study Institute, Series B: Physics*, edited by E. K. U. Gross and R. M. Dreizler, (Plenum, New York 1995).
[13] K. Capelle, E. K. U. Gross, and B. L. Györffy, Phys. Rev. Lett. **78**, 3753 (1997).
[14] P. G. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966).
[15] L. N. Oliveira, E. K. U. Gross, and W. Kohn, Phys. Rev. Lett. **60**, 2430 (1988).
[16] W. N. Mathews, Jr., Phys. Status Solidi B **90**, 327 (1978).
[17] J. Bardeen, R. Kümmel, A. E. Jacobs, and L. Tewordt, Phys. Rev. **187**, 556 (1969).
[18] R. Kümmel, U. Schüssler, U. Gunsenheimer, and H. Plehn, Physica C **185-189**, 221 (1991).
[19] H. Plehn, O.-J. Wacker, and R. Kümmel, Phys. Rev. B **49**, 12 140 (1994).
[20] U. Schüssler and R. Kümmel, Phys. Rev. B **47**, 2754 (1994).
[21] U. Gunsenheimer, U. Schüssler, and R. Kümmel, Phys. Rev. B **49**, 6111 (1994).
[22] F. Gygi and M. Schlüter, Phys. Rev. Lett. **65**, 1820 (1990).
[23] F. Gygi and M. Schlüter, Phys. Rev. B **43**, 7609 (1991).
[24] M. B. Suvasini, W. M. Temmerman, and B. L. Gyorffy, Phys. Rev. B **48**, 1202 (1993).
[25] B. Janko and J. D. Shore, Phys. Rev. B **46**, 9270 (1992).
[26] K. Capelle and E. K. U. Gross, Int. J. Quantum Chem. **61**, 325 (1997).
[27] E. K. U. Gross, S. Kurth, K. Capelle, and M. Lüders, in *Density-Functional Theory*, Vol. 337 of NATO *Advanced Study Institute, Series B: Physics*, edited by E. K. U. Gross and R. M. Dreizler (Plenum, New York, 1995).
[28] C. V. Nino and R. Kümmel, Phys. Rev. B **29**, 3957 (1984).
[29] H. Plehn and R. Kümmel, Ann. Phys. (N.Y.) **5**, 559 (1996).
[30] M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, New York, 1975).
[31] D. C. Mattis and J. Bardeen, Phys. Rev. **111**, 412 (1958).
[32] J. J. Chang and D. J. Scalapino, Phys. Rev. B **40**, 4299 (1989).
[33] H. Chen, Phys. Rev. Lett. **71**, 2304 (1993).
[34] H. Bennett and E. A. Stern, Phys. Rev. **137**, A448 (1965).
[35] C. S. Wang and J. Callaway, Phys. Rev. B **9**, 4897 (1974).
[36] P. Strange, P. J. Durham, and B. L. Gyorffy, Phys. Rev. Lett. **67**, 3590 (1991).
[37] H. Ebert, P. Strange, and B. L. Gyorffy, J. Appl. Phys. **63**, 3055 (1988).
[38] J. Callaway, *Energy Band Theory* (Academic, New York, 1964).
[39] C. Kittel, *Quantum Theory of Solids* (Wiley, New York, 1987).
[40] Yu. A. Uspenskii and S. V. Khalilov, Sov. Phys. JETP **68**, 588 (1989).
[41] J. R. Schrieffer, *Theory of Superconductivity* (Addison-Wesley, Reading, MA, 1988).
[42] M. Sigrist and K. Ueda, Rev. Mod. Phys. **63**, 239 (1991).
[43] G. Preosti and M. Palumbo, Phys. Rev. B **55**, 8430 (1997).
[44] X. G. Wen and A. Zee, Phys. Rev. Lett. **62**, 2873 (1989).
[45] Q. P. Li and R. Joynt, Phys. Rev. B **44**, 4720 (1991).
[46] S. K. Yip and J. A. Sauls, J. Low Temp. Phys. **86**, 257 (1992).
[47] B. Arfi and L. P. Gorkov, Phys. Rev. B **46**, 9163 (1992).
[48] P. Argyres, Phys. Rev. **97**, 334 (1955).
[49] I. M. Boswarva, Proc. R. Soc. London, Ser. A **269**, 125 (1962).
[50] J. L. Erskine and E. A. Stern, Phys. Rev. B **12**, 5016 (1975).
[51] C. Kittel, Phys. Rev. **83**, 208 (1951), and as quoted in Ref. 48.
[52] R. A. de Groot, F. M. Mueller, P. G. van Engen, and K. H. J. Buschow, J. Appl. Phys. **55**, 2151 (1984).
[53] L. C. Hebel and C. P. Slichter, Phys. Rev. **107**, 901 (1957); **113**, 1504 (1959).
[54] J. Bardeen, G. Rickayzen, and L. Tewordt, Phys. Rev. **119**, 982 (1959).
[55] K. Yosida, Phys. Rev. **110**, 769 (1958).
[56] J. R. Clem, Ann. Phys. (N.Y.) **40**, 268 (1966).
[57] D. E. Maclaughlin, in *Solid State Physics: Advances in Research and Applications*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1976), Vol. 31, p. 1.
[58] Y. Masuda, Phys. Rev. **126**, 1271 (1962).
[59] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover, New York, 1963).
[60] C. H. Pennington and V. A. Stenger, Rev. Mod. Phys. **68**, 855 (1996).
[61] J. R. Schrieffer, Solid State Commun. **92**, 129 (1994), and references therein.

[62] C. P. Poole, H. A. Farach, and R. J. Creswick, *Superconductivity* (Academic, San Diego, 1995).
[63] M. Strongin and O. F. Kammerer, Phys. Rev. Lett. **16**, 456 (1966).
[64] J. E. Crow, M. Strongin, and A. K. Bhatnagar, Phys. Rev. B **9**, 135 (1974).
[65] Y.-D. Zhu, F.-C. Zhang, and H. D. Drew, Phys. Rev. B **47**, 586 (1993).