Relativistic framework for microscopic theories of superconductivity. I. The Dirac equation for superconductors

K. Capelle and E. K. U. Gross
Institut für Theoretische Physik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany
(Received 14 October 1997)

We present a unified treatment of relativistic effects in superconductors. The relativistically correct (Dirac-type) single-particle Hamiltonian describing the quasiparticle spectrum of superconductors is deduced from symmetry considerations and the requirement of the correct nonrelativistic limit. We provide a complete list of all order parameters consistent with the requirement of Lorentz covariance. This list contains the relativistic generalizations of the BCS and the triplet order parameters, among others. Furthermore, we present a symmetry classification of the order parameters according to their behavior under the Lorentz group, generalizing previous treatments that were based on the Galilei group. The considerations in this paper are based only on the concepts of pairing and Lorentz covariance. They can therefore be applied to all situations in which pairing takes place. This includes BCS-type superconductors, as well as the heavy-fermion compounds, high-temperature superconductors, pairing of neutrons and protons in neutron stars, and superfluid helium 3.
[S0163-1829(99)11305-5]

I. INTRODUCTION

This is the first in a series of two papers devoted to an investigation of the effects of relativity in superconductors. It has been noted before by many authors that relativistic effects can have a profound influence on superconductivity. Spin-orbit coupling, a relativistic effect of second order in v/c, is known to influence the symmetry of the order parameter,[1–4] the spin-susceptibility and the Knight shift,[5,6] magnetic impurities in superconductors,[7] Josephson currents,[1,8–10] the value of the upper critical field,[11,12] H_{c2}, and the magnetooptical response of superconductors,[13–15] among other quantities. Relativistic corrections to the Cooper pair mass have been evaluated theoretically and measured experimentally.[16–19] The self-consistent screening of the currents which gives rise to the Meissner effect is due to the current-current interaction,[20,21] which is of relativistic origin and of second order in v/c. Plasma frequency anomalies which were observed in some high-temperature superconductors have been suggested to be due to current-current interactions as well.[22] In the anyonic theory of superconductivity it was recently argued that one needs to start from a relativistic Lagrangian in order to obtain a complete description of the Meissner effect.[23] In studying the electrodynamics of vortices in high-temperature superconductors it was found necessary to employ a relativistically covariant wave equation in order to explain the experimental data.[24] The relativistically covariant theory presented in the present paper provides a unified framework within which such effects can be interpreted and analyzed.

Moreover, whenever there are elements with atomic number $Z \gtrsim 40$ in the lattice,[25] the band structure has to be calculated using relativistic methods.[26–28] 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 ($Z=83$), thallium ($Z=81$), platinum ($Z=78$), etc. In Sec. V of the present paper an equation is derived which allows one to perform such relativistic band-structure calculations for superconductors.

In spite of the fact that relativity thus obviously is relevant for a large number of interesting effects in superconductors, a unified and covariant relativistic approach to superconductivity has not been worked out, until very recently. In a previous paper[29] we presented a first step towards such a relativistic theory of superconductivity. That theory led to a relativistic generalization of the Bogolubov–de Gennes equations of superconductivity. By performing a weakly relativistic expansion up to second order in v/c, where v is a typical velocity of the particles involved, we found the usual relativistic corrections (spin-orbit coupling, mass-velocity correction, and Darwin term) in a form appropriate for superconductors. Furthermore, a number of new relativistic corrections of the *same* order in v/c emerged, which exist in superconductors only. These new terms could be identified as the superconducting counterparts of the spin-orbit and the Darwin term, with the pair potential taking the place of the lattice potential. The appearance of such terms can be traced back to the complex interplay between relativistic symmetry breaking and superconducting coherence. The theory has meanwhile been shown to lead to potentially observable effects on, e.g., the energy spectrum of a superconductor[29] and on the magneto-optical response of superconductors.[13–15]

The present paper provides a more general approach to relativistic effects in superconductors. It is organized as follows: after this introduction we show, in Sec. II, how the Dirac equation is generalized to describe superconductors by introducing the order parameters as 4×4 matrices into the Dirac Hamiltonian.

In Secs. III and IV we present a detailed analysis of the resulting Hamiltonian. We discuss Lorentz covariance of the formalism and investigate the behavior of the order parameters under Lorentz transformations. This leads to a classification of all possible order parameters with respect to their behavior under the operations of the Lorentz group. Next we

show that these order parameters can be interpreted in terms of the symmetries of the underlying Cooper pair states. This gives rise to a relativistic generalization of the concepts of singlet and triplet superconductivity and allows one to specify what kind of physics is described by the various order parameters.

In Sec. V we discuss the diagonalization of the Hamiltonian. This leads to a set of differential equations which generalize the Bogolubov–de Gennes equations of the nonrelativistic theory.

While the present paper thus deals with the Dirac equation for superconductors, the second paper of this series[30] will treat the Pauli equation for superconductors. In that paper we take up the topic of weakly relativistic corrections to the conventional theory of superconductivity and discuss some observable consequences of the new terms.

Unfortunately, the respective terminologies of relativity and superconductivity are rather different and there is little overlap in the literature on these apparently distinct fields of physics. To aid the nonspecialist, we have therefore included two sections in which we briefly review some pertinent aspects of the microscopic theory of inhomogeneous superconductors (Sec. II A) and of relativistic covariance (Sec. III A).

It should be stressed from the outset that we do not attempt to formulate a fully relativistic *interacting* field theory of superconductivity: We quantize only the electron degrees of freedom and treat the external fields as classical fields. Furthermore, the effects of relativity are considered only on the single-particle level, i.e., on the level of the Bogolubov–de Gennes equations. A relativistic treatment of the interaction is beyond the scope of the present paper, although we offer some remarks concerning this topic in Sec. V C.

II. GENERAL FRAMEWORK FOR A RELATIVISTIC THEORY OF SUPERCONDUCTIVITY

A. A basis for the nonrelativistic order parameter

In the nonrelativistic case it is well known that the order parameter (OP) structure is determined by very general symmetry considerations, as follows:[31–34] Cooper pairing takes place between single-particle states which are, for any given state $|n\sigma\rangle$, constructed by application of the time-reversal and parity operators. If n stands for a complete set of quantum numbers needed to label a normal-state eigenfunction, we thus have the states

$$|n\sigma\rangle, \quad |\hat{T}n\sigma\rangle, \quad |\hat{P}n\sigma\rangle, \quad |\hat{P}\hat{T}n\sigma\rangle \tag{1}$$

available for pairing, where $\hat{T}$ and $\hat{P}$ are the time-reversal and parity operators, respectively. The possible pairing states of the superconductor can then be characterized in terms of these states as[31–34]

$$|S\rangle = |n\sigma,\hat{T}n\sigma\rangle - |\hat{P}\hat{T}n\sigma,\hat{P}n\sigma\rangle, \tag{2}$$

$$|T+\rangle = |n\sigma,\hat{P}n\sigma\rangle, \tag{3}$$

$$|T-\rangle = |\hat{P}\hat{T}n\sigma,\hat{T}n\sigma\rangle, \tag{4}$$

$$|T0\rangle = |n\sigma,\hat{T}n\sigma\rangle + |\hat{P}\hat{T}n\sigma,\hat{P}n\sigma\rangle. \tag{5}$$

The physical significance of this construction is easily seen by considering a homogeneous electron gas, where spin and momentum are good quantum numbers. If $|n\sigma\rangle = |\mathbf{k}\uparrow\rangle$, then $|\hat{T}n\sigma\rangle = |-\mathbf{k}\downarrow\rangle$, $|\hat{P}n\sigma\rangle = |-\mathbf{k}\uparrow\rangle$, and $|\hat{P}\hat{T}n\sigma\rangle = |\mathbf{k}\downarrow\rangle$. Taking into account that each of the two-particle states is a Slater determinant, the configuration-space representation of $|S\rangle$ takes the form

$$\langle \mathbf{r},\mathbf{r}'|S\rangle = [\varphi_k(\mathbf{r})\varphi_{-k}(\mathbf{r}') + \varphi_{-k}(\mathbf{r})\varphi_k(\mathbf{r}')](\chi_\uparrow \otimes \chi_\downarrow - \chi_\downarrow \otimes \chi_\uparrow), \tag{6}$$

where the $\varphi_k(\mathbf{r})$ are normal-state single-particle wave functions and the χ_σ are the usual spin functions. Obviously, the spatial part of $|S\rangle$ is an even function under exchange of the two particles, while the spin part is odd. This state describes the conventional singlet Cooper pair.[35,36] Representing the Pauli spinors as $\chi_\uparrow = (1,0)^T$ and $\chi_\downarrow = (0,1)^T$, this can be written as

$$\langle \mathbf{r},\mathbf{r}'|S\rangle = [\varphi_k(\mathbf{r})\varphi_{-k}(\mathbf{r}') + \varphi_{-k}(\mathbf{r})\varphi_k(\mathbf{r}')]\hat{m}_1 \tag{7}$$

with

$$\hat{m}_1 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{8}$$

The other three states are even functions of the spin variables and odd functions of the spatial coordinates. They describe triplet Cooper pairs, as found in superfluid helium 3 (Ref. 37) and, possibly, organic superconductors[38] and heavy-fermion compounds.[39] The spin parts of the triplet states are represented by

$$|T+\rangle \rightarrow \hat{m}_2 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{9}$$

$$|T-\rangle \rightarrow \hat{m}_3 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{10}$$

$$|T0\rangle \rightarrow \hat{m}_4 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{11}$$

These matrices form the basis for many approaches to superconductivity. In particular, the Balian-Werthamer parametrization for the order parameter of triplet superconductors is simply a linear combination of the matrices $|T+\rangle$, $|T-\rangle$, and $|T0\rangle$.[40,41] The fact that the singlet OP is composed of time-reversed single-particle states is the basis for the theory of impurities in BCS superconductors.[42,43] Furthermore, the above defined states (2)–(5) were taken as a starting point for investigations of the order-parameter symmetry in unconventional superconductors, e.g., in Refs. 31–34.

Multiplying these matrices from the left and from the right with spinors whose entries are functions of $\mathbf{r}$, yields a real-space representation of the OP which is commonly used in the Bogolubov–de Gennes equations[43] and in the density-functional theory of superconductivity.[21,41,44–49] Since this will be of considerable importance below, we demonstrate it explicitly for the singlet case. The conventional Bogolubov–de Gennes Hamiltonian for a singlet superconductor is[43]

$$\hat{H}_{\text{non-rel}}=\int d^3r\sum_{\sigma}\ \psi_\sigma^\dagger(\mathbf{r})\left[\frac{\mathbf{p}^2}{2m}+v(\mathbf{r})-\mu\right]\psi_\sigma(\mathbf{r})$$
$$-\int d^3r\,d^3r'[\Delta_{\uparrow\downarrow}^*(\mathbf{r},\mathbf{r}')\hat{\chi}_{\uparrow\downarrow}(\mathbf{r},\mathbf{r}')+\text{H.c.}], \tag{12}$$

where the ψ_σ are second quantized field operators, $v(\mathbf{r})$ is the lattice potential, $\Delta_{\uparrow\downarrow}(\mathbf{r},\mathbf{r}')$ is the pair potential, and the expectation value of $\hat{\chi}_{\uparrow\downarrow}(\mathbf{r},\mathbf{r}')\coloneqq\psi_\uparrow(\mathbf{r})\psi_\downarrow(\mathbf{r}')$ is the (nonlocal) singlet OP.[43] In Sec. V C the role of this Hamiltonian in density-functional theory for superconductors is further discussed.

The terms in Eq. (12) referring to superconductivity can be rewritten as

$$\int d^3r\,d^3r'\Delta_{\uparrow\downarrow}^*(\mathbf{r},\mathbf{r}')\hat{\chi}_{\uparrow\downarrow}(\mathbf{r},\mathbf{r}')$$
$$=\frac{1}{2}\int d^3r\,d^3r'\Delta_{\uparrow\downarrow}^*(\mathbf{r},\mathbf{r}')$$
$$\times\begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix}^T\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix} \tag{13}$$

and similarly for its Hermitian conjugate. We have made use here of the fact that the pair potential for a singlet state is even under interchange of (r) and $\mathbf{r}'$. This is guaranteed because the spin part of the singlet state is an odd function of the spins, and the full pair potential, having the symmetries of a fermionic pair wave function, must be odd under particle exchange.

It is seen that the matrix (8) appears naturally in Eq. (13). In exactly the same way the matrices (9)–(11) appear in the generalization of the Bogolubov–de Gennes equations to triplet superconductors, as demonstrated explicitly in Ref. 41. The corresponding pair potentials are odd functions of the spatial coordinates.

In the most general case, all these OP are present simultaneously. Then all the order parameters enter the Hamiltonian, multiplied by the appropriate pair fields and the superconducting term in the Hamiltonian becomes

$$\left(\frac{1}{2}\right)\int d^3r\int d^3r'\begin{pmatrix}\psi_\uparrow(\mathbf{r})\\ \psi_\downarrow(\mathbf{r})\end{pmatrix}^T\sum_j\ \Delta_j(\mathbf{r},\mathbf{r}')\hat{m}_j\begin{pmatrix}\psi_\uparrow(\mathbf{r}')\\ \psi_\downarrow(\mathbf{r}')\end{pmatrix}, \tag{14}$$

where $\hat{m}_j$ are the four matrices defined in Eqs. (8)–(11). Since $\hat{m}_2$ and $\hat{m}_3$ have only one entry, the factor 1/2 is present only for $j=1,4$. [We can replace it by $\frac{1}{2}(1+\delta_{j2}+\delta_{j3})$ to have a closed expression for all cases.]

It is by no means necessary to use this particular set of matrices. Any linear combination of them [respectively, of the states (2)–(5)] can be used just as well. Writing the general OP as

$$\begin{pmatrix}\psi_\uparrow(\mathbf{r})\\ \psi_\downarrow(\mathbf{r})\end{pmatrix}^T\hat{m}(\mathbf{r},\mathbf{r}')\begin{pmatrix}\psi_\uparrow(\mathbf{r}')\\ \psi_\downarrow(\mathbf{r}')\end{pmatrix} \tag{15}$$

with an arbitrary 2×2 matrix $\hat{m}(\mathbf{r},\mathbf{r}')$, we can expand $\hat{m}(\mathbf{r},\mathbf{r}')$ in any complete basis in spin space instead of the $\hat{m}_j$, such as, for example, the set of the three Pauli matrices and the 2×2 unit matrix.

B. A basis for the relativistic order parameter

We are now in a position to formulate the relativistic generalization of the Bogolubov–de Gennes Hamiltonian (12) or, equivalently, the superconducting generalization of the Dirac Hamiltonian. Proceeding in the spirit of a single-particle theory it is obvious, that the first (i.e., the Schrödinger) part of Eq. (12) is to be replaced by the corresponding Dirac Hamiltonian. In analogy to Eq. (15) the second part can always be expressed in terms of a general and as yet undetermined 4×4 matrix $\hat{M}(\mathbf{r},\mathbf{r}')$, i.e.,

$$\hat{H}=\int d^3r\ \overline{\Psi}(\mathbf{r})[c\,\hat{\boldsymbol{\gamma}}\cdot\mathbf{p}+mc^2+q\,\hat{\gamma}^\mu A_\mu]\Psi(\mathbf{r})$$
$$-\frac{1}{2}\int d^3r\,d^3r'[\Psi^T(\mathbf{r})\hat{M}(\mathbf{r},\mathbf{r}')\Psi(\mathbf{r}')+\text{H.c.}], \tag{16}$$

where $\Psi(\mathbf{r})$ is a four-component field operator (Dirac-spinor operator), $\overline{\Psi}(\mathbf{r})\coloneqq\Psi^\dagger(\mathbf{r})\hat{\gamma}^0$ and A_μ is the electromagnetic four-potential[50,51]

$$A_\mu=\left(\frac{1}{q}(v(\mathbf{r})-\mu),-\mathbf{A}\right). \tag{17}$$

q is the charge of the particles involved. In writing the above, we made use of the summation convention, i.e., a summation over *greek* indices which appear once as an upper and once as a lower index is implied. We now expand the matrix $\hat{M}(\mathbf{r},\mathbf{r}')$ in a basis set of 16 linearly independent 4×4 matrices $\hat{M}_j$ with expansion coefficients $\Delta_j^*(\mathbf{r},\mathbf{r}')$:

$$\Psi^T(\mathbf{r})\hat{M}(\mathbf{r},\mathbf{r}')\Psi(\mathbf{r}')=\Psi^T(\mathbf{r})\left[\sum_j\ \Delta_j^*(\mathbf{r},\mathbf{r}')\hat{M}_j\right]\Psi(\mathbf{r}'). \tag{18}$$

Just as in the nonrelativistic case, in principle, any complete set of 4×4 matrices can be chosen as a basis for this expansion. Below we employ one particular such set, namely that of the $\hat{\eta}$ matrices defined by Eqs. (19)–(34). This set will be shown to ensure relativistic covariance of the formalism (cf. Sec. III) and to permit an interpretation of the physics of the paired states analogous to Eqs. (2)–(5) (cf. Sec. IV). For clarity we display below only those matrix entries which are different from zero:

$$\hat{\eta}\coloneqq\begin{pmatrix} & & 1 & \\ -1 & & & \\ & & & 1\\ & & -1 & \end{pmatrix}, \tag{19}$$

$$\hat{\eta}_V^0 := \begin{pmatrix} & & 1 & \\ -1 & & & \\ & & & -1 \\ & 1 & & \end{pmatrix}, \tag{20}$$

$$\hat{\eta}_V^1 := \begin{pmatrix} & & 1 & \\ & & & -1 \\ -1 & & & \\ & 1 & & \end{pmatrix}, \tag{21}$$

$$\hat{\eta}_V^2 := \begin{pmatrix} & & i & \\ & & & i \\ -i & & & \\ & -i & & \end{pmatrix}, \tag{22}$$

$$\hat{\eta}_V^3 := \begin{pmatrix} & & & -1 \\ & & -1 & \\ & 1 & & \\ 1 & & & \end{pmatrix}, \tag{23}$$

$$\hat{\eta}^5 := \begin{pmatrix} & & & 1 \\ & & -1 & \\ & 1 & & \\ -1 & & & \end{pmatrix} = i\hat{\eta}\hat{\eta}_V^0\hat{\eta}_V^1\hat{\eta}_V^2\hat{\eta}_V^3, \tag{24}$$

$$\hat{\eta}_A^0 := \begin{pmatrix} & & & 1 \\ & & -1 & \\ & -1 & & \\ 1 & & & \end{pmatrix} = \hat{\eta}\hat{\eta}^5\hat{\eta}_V^0, \tag{25}$$

$$\hat{\eta}_A^1 := \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} = \hat{\eta}\hat{\eta}^5\hat{\eta}_V^1, \tag{26}$$

$$\hat{\eta}_A^2 := \begin{pmatrix} i & & & \\ & i & & \\ & & -i & \\ & & & -i \end{pmatrix} = \hat{\eta}\hat{\eta}^5\hat{\eta}_V^2, \tag{27}$$

$$\hat{\eta}_A^3 := \begin{pmatrix} & -1 & & \\ -1 & & & \\ & & & 1 \\ & & 1 & \end{pmatrix} = \hat{\eta}\hat{\eta}^5\hat{\eta}_V^3, \tag{28}$$

$$\hat{\eta}_T^{01} := \begin{pmatrix} & -1 & & \\ -1 & & & \\ & & & 1 \\ & & 1 & \end{pmatrix} = \hat{\eta}\hat{\eta}_V^0\hat{\eta}_V^1, \tag{29}$$

$$\hat{\eta}_T^{02} := -\begin{pmatrix} & & i & \\ & & & i \\ i & & & \\ & i & & \end{pmatrix} = \hat{\eta}\hat{\eta}_V^0\hat{\eta}_V^2, \tag{30}$$

$$\hat{\eta}_T^{03} := \begin{pmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{pmatrix} = \hat{\eta}\hat{\eta}_V^0\hat{\eta}_V^3, \tag{31}$$

$$\hat{\eta}_T^{12} := \begin{pmatrix} & i & & \\ i & & & \\ & & & i \\ & & i & \end{pmatrix} = \hat{\eta}\hat{\eta}_V^1\hat{\eta}_V^2, \tag{32}$$

$$\hat{\eta}_T^{13} := -\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} = \hat{\eta}\hat{\eta}_V^1\hat{\eta}_V^3, \tag{33}$$

$$\hat{\eta}_T^{23} := \begin{pmatrix} i & & & \\ & -i & & \\ & & i & \\ & & & -i \end{pmatrix} = \hat{\eta}\hat{\eta}_V^2\hat{\eta}_V^3. \tag{34}$$

These 16 matrices are linearly independent and hence constitute a complete set in which every 4×4 matrix can be expanded.[52] The labeling of the $\hat{\eta}$ matrices refers to the transformation behavior of the corresponding OP under Lorentz transformations, as will be explained in Sec. III B and in Table I.

Clearly, the nonrelativistic limit of Eq. (16) is the Bogolubov–de Gennes Hamiltonian for singlet *and* triplet OP (Refs. 43 and 41) (the spin degrees of freedom are naturally contained in the relativistic framework), while the nonsuperconducting limit (all $\Delta_j \to 0$) is the conventional Dirac Hamiltonian.

We can now also see another reason, why it is useful to introduce the matrix notation for the order parameter: The corresponding terms in the Hamiltonian are all of the form

$$spinor\cdot(4\times4\ \ matrix)\cdot spinor\cdot field, \tag{35}$$

just as are the terms of the Dirac Hamiltonian. In particular, the pairing fields Δ_j and the electromagnetic field A_μ enter the Hamiltonian on the same footing. The four-current density is defined, as usual, by $j^\mu = c\overline{\Psi}\hat{\gamma}^\mu\Psi$, while the order

TABLE I. This table is a complete list of all order-parameter matrices which are consistent with the requirement of Lorentz covariance. The first column denotes the transformation behavior of the OP. The second column contains the familiar bilinear covariants of the conventional Dirac equation. The third column contains the corresponding anomalous bilinear covariants. Every individual component of each entry in this table is a 4×4 matrix. The labels SE, SO, TE, and TO, assigned to each component of the anomalous bilinear covariants refer to the spin character (S=singlet, T=triplet) and the behavior under space inversion (E=even, O=odd) of the corresponding pairs. These labels are explained in detail in Secs. IV B and IV C.

Transformation behavior under Lorentz transformations	Conventional bilinear covariants of the Dirac equation $\bar{\Psi}(\mathbf{r})\ldots\Psi(\mathbf{r})$	Anomalous bilinear covariants of Eq. (43) $\Psi^T(\mathbf{r})\ldots\Psi(\mathbf{r}')$
Scalar	$\mathbb{1}$	$\hat{\eta}\quad(SE)$
Four vector	$\hat{\gamma}^\mu=\begin{pmatrix}\hat{\gamma}^0\\ \hat{\gamma}^1\\ \hat{\gamma}^2\\ \hat{\gamma}^3\end{pmatrix}$	$\hat{\eta}_V^\mu=\begin{pmatrix}\hat{\eta}_V^0\\ \hat{\eta}_V^1\\ \hat{\eta}_V^2\\ \hat{\eta}_V^3\end{pmatrix}\quad\begin{pmatrix}SE\\ SO\\ SO\\ SO\end{pmatrix}$
Pseudoscalar	$\hat{\gamma}^5=i\hat{\gamma}^0\hat{\gamma}^1\hat{\gamma}^2\hat{\gamma}^3$	$\hat{\eta}^5=i\hat{\eta}\hat{\eta}_V^0\hat{\eta}_V^1\hat{\eta}_V^2\hat{\eta}_V^3\quad(SO)$
Axial four vector	$\hat{\gamma}^5\hat{\gamma}^\mu=\begin{pmatrix}\hat{\gamma}^5\hat{\gamma}^0\\ \hat{\gamma}^5\hat{\gamma}^1\\ \hat{\gamma}^5\hat{\gamma}^2\\ \hat{\gamma}^5\hat{\gamma}^3\end{pmatrix}$	$\hat{\eta}_A^\mu=\hat{\eta}\hat{\eta}^5\hat{\eta}_V^\mu\quad\begin{pmatrix}TO\\ TE\\ TE\\ TE\end{pmatrix}$
Antisymmetric tensor	$-i\hat{\sigma}^{\mu\nu}=$	$\hat{\eta}_T^{\mu\nu}=$
	$\begin{pmatrix}0 & \hat{\gamma}^0\hat{\gamma}^1 & \hat{\gamma}^0\hat{\gamma}^2 & \hat{\gamma}^0\hat{\gamma}^3\\ -\hat{\gamma}^0\hat{\gamma}^1 & 0 & \hat{\gamma}^1\hat{\gamma}^2 & \hat{\gamma}^1\hat{\gamma}^3\\ -\hat{\gamma}^0\hat{\gamma}^2 & -\hat{\gamma}^1\hat{\gamma}^2 & 0 & \hat{\gamma}^2\hat{\gamma}^3\\ -\hat{\gamma}^0\hat{\gamma}^3 & -\hat{\gamma}^1\hat{\gamma}^3 & -\hat{\gamma}^2\hat{\gamma}^3 & 0\end{pmatrix}$	$\hat{\eta}\begin{pmatrix}0 & \hat{\eta}_V^0\hat{\eta}_V^1 & \hat{\eta}_V^0\hat{\eta}_V^2 & \hat{\eta}_V^0\hat{\eta}_V^3\\ -\hat{\eta}_V^0\hat{\eta}_V^1 & 0 & \hat{\eta}_V^1\hat{\eta}_V^2 & \hat{\eta}_V^1\hat{\eta}_V^3\\ -\hat{\eta}_V^0\hat{\eta}_V^2 & -\hat{\eta}_V^1\hat{\eta}_V^2 & 0 & \hat{\eta}_V^2\hat{\eta}_V^3\\ -\hat{\eta}_V^0\hat{\eta}_V^3 & -\hat{\eta}_V^1\hat{\eta}_V^3 & -\hat{\eta}_V^2\hat{\eta}_V^3 & 0\end{pmatrix}$
		$\begin{pmatrix}0 & TO & TO & TO\\ TO & 0 & TE & TE\\ TO & TE & 0 & TE\\ O & TE & TE & 0\end{pmatrix}$

parameters are analogously given by $\hat{\chi}_j=\Psi^T\hat{M}_j\Psi$. The four-current density and the order parameters are thus also treated in a parallel way. This is particularly appealing from the point of view of the density-functional theory of superconductivity,[21,41,44–49] where the Bogolubov–de Gennes equations become the Kohn-Sham equations for superconductors and the order parameters enter the formalism as ''anomalous densities,'' in addition to the normal density $n(\mathbf{r})$ and, if necessary, the current density $\mathbf{j}$. We call Eq. (16) the generalized Dirac–Bogolubov–de Gennes Hamiltonian because it contains the Dirac–Bogolubov–de Gennes Hamiltonian previously suggested[29] as a special case.

III. SYMMETRY ANALYSIS WITH RESPECT TO THE LORENTZ GROUP

A. Bilinear covariants of the conventional Dirac equation

Every relativistic theory needs to be covariant, i.e., the equations must have the same form in every inertial system. We, therefore, have to verify that our equations are form invariant under Lorentz transformations, i.e., invariant modulo a Lorentz transformation of the arguments. This is clearly the case for the conventional Dirac theory,[50,51,53] i.e., without the terms containing the Δ_j. To determine whether the additional terms in Eq. (16) are covariant, we need to

explicitly study their transformation behavior under Lorentz transformations.

In principle, a discussion of Lorentz invariance should be based on the Lagrangian density and not the Hamiltonian because the former is invariant itself, while the latter is not.[50] The Hamiltonian formulation, though, is much more common and convenient in solid-state physics. Since the Lagrangian density $\mathcal{L}$ is related to the Hamiltonian $H = \int d^3r\, \mathcal{H}$ by

$$\mathcal{H} = \sum_i \dot{\psi}_i \frac{\partial \mathcal{L}}{\partial \dot{\psi}_i} - \mathcal{L}, \tag{36}$$

all terms in $\mathcal{L}$ which do not depend on $\dot{\psi}_i$ [such as the terms describing superconductivity in Eq. (16)] appear in $\mathcal{H}$ and $\mathcal{L}$ with opposite sign, but the same transformation behavior. We can thus carry on with the Hamiltonian formulation, if we keep in mind that the actual invariant quantity is the Lagrangian density.

The prescription how to perform a Lorentz transformation for Dirac spinors can be found in many textbooks.[50,51,53] If Ψ is the original spinor, then the spinor in a moving inertial system is related to it via

$$\Psi'(x') = \hat{S}\Psi(x), \tag{37}$$

where $x = (ct, \mathbf{r})$ and $x' = (ct', \mathbf{r}')$ are the space-time coordinates in the original and the new inertial system, respectively. For a transformation to an inertial system moving with uniform velocity v_k along the k axis $(k = x,y,z)$, the transformation matrix is given by

$$\hat{S} = \left[\cosh\left(\frac{\omega}{2}\right)\mathbb{1} - \sinh\left(\frac{\omega}{2}\right)\begin{pmatrix} 0 & \hat{\sigma}_k \\ \hat{\sigma}_k & 0 \end{pmatrix} \right], \tag{38}$$

where

$$\tanh\left(\frac{\omega}{2}\right) = \frac{v_k}{c} \tag{39}$$

and $\hat{\sigma}_k$ are the usual 2×2 Pauli matrices. This form applies to finite, homogeneous (not involving translations), orthochronous (not involving time reflections) and pure (not involving spatial rotations and reflections) Lorentz transformations (i.e., ''Lorentz boosts'').

One can now form all possible combinations of the Dirac-$\hat{\gamma}$ matrices and investigate the transformation behavior of the quantities[50,51,53]

$$\overline{\Psi}'\hat{\gamma}\Psi' \equiv \Psi'^{\dagger}\hat{\gamma}^0\hat{\gamma}\Psi' \equiv (S\Psi)^{\dagger}\hat{\gamma}^0\hat{\gamma}(S\Psi) \tag{40}$$

under Lorentz transformations, where $\hat{\gamma}$ stands for any of the linearly independent Dirac matrices. Since there are 16 linearly independent 4×4 matrices, one can find 16 such quantities which can be classified according to their transformation behavior. Explicitly, one finds that one can form a scalar, a pseudoscalar, a (polar) four vector, an axial four vector and an antisymmetric tensor of rank two. Each of these transforms under Lorentz transformations and space inversion according to an irreducible representation of the Lorentz group[54] (see Sec. III C for more details). These five entities are normally called the *bilinear covariants* of the Dirac equation. They are listed in the second column of Table I. Their construction and a detailed analysis of their properties is found in many standard textbooks on relativistic quantum field theory.[50,51,53] Together they have 16 components. Since we had only 16 matrices to start with, we have exhausted all possible combinations. For the Dirac equation there are exactly these five and no other bilinear covariants.

Their importance stems from the fact that the Hamiltonian (or the Lagrangian) needs to be expressed in terms of these bilinear covariants in order to be manifestly covariant itself. Furthermore, since there is just a small number of different bilinear covariants, one can classify all interactions with respect to the bilinear covariants which one needs for a proper description. It turns out, that almost every interaction can be described in terms of just one bilinear covariant. The most prominent exception is the weak interaction of high-energy physics. A description of this interaction requires the presence of a (polar) four vector *and* an axial four vector in the Lagrangian. Since under a spatial reflection the former picks up a relative sign with respect to the latter, the weak interaction violates parity invariance.[51,53] We have discussed these facts at length because they will turn out to be highly relevant for the superconducting case as well.

B. Symmetry classification of all possible order parameters

We first consider the transformation behavior of the OP formed with $\hat{\eta}$. Given Eq. (38), it is a matter of simple matrix multiplication to verify that

$$\Psi'^{T}\hat{\eta}\Psi' = (\hat{S}\Psi)^{T}\hat{\eta}(\hat{S}\Psi) \equiv \Psi^{T}\hat{\eta}\Psi, \tag{41}$$

i.e., $\Psi^T\hat{\eta}\Psi$ is form invariant under Lorentz transformations.

$\hat{\eta}$ can be expressed in terms of the $\hat{\gamma}$ matrices as $\hat{\gamma}^1\hat{\gamma}^3$. The bilinear covariant which is formed with these $\hat{\gamma}$ matrices, namely $\overline{\Psi}\hat{\gamma}^1\hat{\gamma}^3\Psi$, is, on the other hand, *not* invariant under Lorentz transformations. As displayed in the second column of Table I, it transforms as the component $\hat{\sigma}^{13}$ of the antisymmetric tensor $\hat{\sigma}^{\mu\nu}$.

A similar situation is found for all OP: their transformation behavior cannot simply be obtained from their representation in terms of the $\hat{\gamma}$ matrices in Eqs. (19)–(34) and a comparison with the corresponding representation of the conventional bilinear covariants in terms of the same matrices. The reason for this is, that the OP contain Ψ^T instead of $\overline{\Psi} \equiv \Psi^{\dagger}\hat{\gamma}^0$. Therefore, it is the transposed transformation matrix $\hat{S}^T$ which enters and not, as usual, the Hermitian conjugate $\hat{S}^{\dagger}$. Since $\hat{S}$ is, in general, a complex matrix, this difference leads, together with the additional $\hat{\gamma}^0$, to a different transformation behavior, as compared to the conventional case. This phenomenon is known also from the nonrelativistic triplet OP, where it leads to using the Balian-Werthamer matrices instead of the Pauli matrices in order to obtain a quantity which transforms as a three vector under spatial rotations.[37,40,55]

We can, of course, investigate the transformation behavior under space reflections in the same way. Here the transformation matrix is simply $\hat{\gamma}^0$ and it is readily verified that

$$\Psi'^{T}\hat{\eta}\Psi' = (\hat{\gamma}^0\Psi)^T\hat{\eta}(\hat{\gamma}^0\Psi) \equiv \Psi^T\hat{\eta}\Psi \tag{42}$$

is invariant under this operation as well. We conclude that $\Psi^T\hat{\eta}\Psi$ is a Lorentz scalar, which does not change sign under spatial reflections (the latter property excludes a pseudoscalar, which would change sign). We do not need to consider spatial rotations and translations because they do not influence the classification of the bilinear covariants.[50,51,53] These operations do not, in general, leave the underlying lattice of the superconductor invariant. The symmetry group of the lattice can be used to further classify the OP.[1,33,56–58] The considerations of the present section can be viewed as an extension of these works from using the Galilei group to the Lorentz group.

In the same way as for the $\hat{\eta}$-OP we can investigate the transformation behavior of all the other 15 matrices. To determine their transformation behavior we perform Lorentz transformations along all three spatial axes. This allows us to distinguish scalars, vectors, and tensors. To further distinguish between proper scalars and pseudoscalars and between (polar) four vectors and axial (pseudo) four vectors, we also need to investigate the behavior under spatial reflections. The covariant quantities one finds in this way will be termed *anomalous* bilinear covariants.[59] We summarize the results of this investigation in Table I, where we display the conventional and the anomalous bilinear covariants, classified according to their transformation behavior.

The 16 components of the five distinct covariant quantities of the Dirac equation can all be expressed in a well-known way[50,53] in terms of just five matrices, namely $\hat{\gamma}^0,\hat{\gamma}^1,\hat{\gamma}^2,\hat{\gamma}^3$, which form the (polar) four vector, and the unit matrix $\mathbb{1}$, which is the scalar. In just the same way, the 16 components of the five distinct anomalous covariants can all be expressed in terms of five matrices, namely $\hat{\eta}_V^0,\hat{\eta}_V^1,\hat{\eta}_V^2,\hat{\eta}_V^3$, which form the (polar) four vector, and the $\hat{\eta}$ matrix, which is the scalar. There is thus a far reaching correspondence between the conventional bilinear covariants of the Dirac equation for normal electrons and the anomalous bilinear covariants (order parameters) of the Dirac–Bogolubov–de Gennes equation for superconductors.

In view of this correspondence it is not surprising that there exists a transformation which, when applied to one of the 16 components of the conventional covariants, yields the corresponding anomalous covariant. If every entry in the second column of Table I is multiplied with $\hat{\gamma}^1\hat{\gamma}^3$, there results (up to factors ± 1) the corresponding entry in the third column. In the case of the pseudoscalar $\hat{\gamma}^5$, for instance, we find $\hat{\gamma}^1\hat{\gamma}^3\hat{\gamma}^5=\hat{\eta}^5$, while the tensor component $\hat{\gamma}^0\hat{\gamma}^2$ becomes $\hat{\gamma}^1\hat{\gamma}^3\hat{\gamma}^0\hat{\gamma}^2=-\hat{\eta}\hat{\eta}_V^0\hat{\eta}_V^2$. This rule transforms a conventional bilinear covariant in the corresponding anomalous covariant, which has the same transformation behavior under Lorentz boosts and spatial reflections. It summarizes in a concise way the lengthy calculations necessary to determine the matrix representations of the 16 $\hat{\eta}$-OP by performing the Lorentz transformations explicitly.

This rule, extracted from Table I, can also be arrived at by means of a generalization of the nonelativistic Balian-Werthamer prescription for the construction of matrices for the singlet and triplet OP's: Observing that ordinarily a scalar is expressed in terms of the unit matrix, while a vector is expressed in terms of the three Pauli matrices, one finds matrices yielding scalar and vector OP simply by multiplication with the nonrelativistic time-reversal matrix $i\hat{\sigma}_y$. The resulting matrices $i\hat{\sigma}_y$ and $i\hat{\sigma}_y\hat{\sigma}_x, i\hat{\sigma}_y\hat{\sigma}_y, i\hat{\sigma}_y\hat{\sigma}_z$, respectively, are those employed in the Balian-Werthamer parametrization for the nonrelativistic OP. In the same spirit all matrices describing relativistic OP can be obtained from the conventional Dirac matrices by multiplication with the relativistic time-reversal matrix $\hat{\gamma}^1\hat{\gamma}^3$. Although this procedure is highly plausible, its general validity can only be verified by performing the Lorentz transformations as described above.

Just as in the nonsuperconducting case, the importance of the anomalous bilinear covariants stems from the fact that a Lagrangian expressed in terms of them is manifestly covariant and from the possibility to classify systems with respect to the appropriate covariants. If for any given superconductor more then one of the five covariants appear in the Hamiltonian, a very general symmetry, such as parity, would be broken.

In the spirit of a mean-field (Hartree-type) approximation, where one has a linear relation between the particle-particle interaction and the effective single-particle potential, we can classify the interaction leading to superconductivity with respect to the symmetry of the corresponding effective pair potential. If, for instance, for a certain superconductor only the scalar OP is present in the Hamiltonian, the interaction leading to superconductivity must be such, that the pair fields multiplying the other bilinear OP's vanish identically. As we will show below, this example corresponds to the conventional BCS-OP. In this way we can classify the interactions as ''scalar interaction,'' ''vector interaction,'' etc.

We can now formulate the following theorem: *No matter what the OP symmetry and the detailed nature of the pairing interaction leading to superconductivity, there are only five OP with a total of 16 components, consistent with the principle of general covariance. Each of these is a bilinear form of the field operators Ψ^T and Ψ transforming according to an irreducible representation of the Lorentz group.* This theorem, derived here by explicit construction of a complete set of matrices, is rederived in Sec. III C by purely group-theoretical considerations. The resulting anomalous bilinear covariants are listed in Table I. Using these results we can now write the Hamiltonian (16) with Eq. (18) as

$$\begin{aligned}\hat{H} = &\int d^3r\,\overline{\Psi}(\mathbf{r})[c\hat{\boldsymbol{\gamma}}\cdot\mathbf{p}+mc^2+q\hat{\gamma}^\mu A_\mu]\Psi(\mathbf{r})\\ &-\frac{1}{2}\int d^3r\,d^3r'\{\Psi^T(\mathbf{r})[\hat{\eta}\Delta^*(\mathbf{r},\mathbf{r}')+\hat{\eta}^5\Delta_P^*(\mathbf{r},\mathbf{r}')\\ &+\hat{\eta}_V^\mu\Delta_{V\mu}^*(\mathbf{r},\mathbf{r}')+\hat{\eta}_A^\mu\Delta_{A\mu}^*(\mathbf{r},\mathbf{r}')\\ &+\hat{\eta}_T^{\mu\nu}\Delta_{T,\mu\nu}^*(\mathbf{r},\mathbf{r}')]\Psi(\mathbf{r}')+\text{H.c.}\}.\end{aligned} \tag{43}$$

Now every contribution to the superconducting part of the Hamiltonian is expressed in terms of covariant quantities: the scalar is formed with $\hat{\eta}$ and Δ, the pseudoscalar with $\hat{\eta}^5$ and Δ_P, the four vector with $\hat{\eta}_V^\mu$ and $\Delta_{V\mu}$, the axial vector with $\hat{\eta}_A^\mu$ and $\Delta_{A\mu}$ and the tensor, finally, with $\hat{\eta}_T^{\mu\nu}$ and $\Delta_{T,\mu\nu}$.

Note the formal correspondence between the four-vector product $\hat{\gamma}^\mu A_\mu$, describing the coupling of the four current, $\overline{\Psi}\hat{\gamma}^\mu\Psi$, to the four potential, A_μ, and the product $\hat{\eta}_V^\mu\Delta_{V\mu}$, describing the coupling of the four vector OP, $\Psi^T\hat{\eta}_V^\mu\Psi$, to the four potential, $\Delta_{V\mu}$.

Equation (43) is one of the central equations of the present work. It constitutes the desired relativistic Hamiltonian for superconductors and contains all possible anomalous bilinear covariants.

C. Irreducible representations of the Lorentz group

The method used above to determine the relativistic OP is based on explicit expressions for each OP in terms of a 4×4 matrix. The main conclusion of the previous section can, however, also be found without any recourse to explicit matrices, simply by exploiting the properties of the irreducible representations (IR's) of the Lorentz group. To this end we first recapitulate a number of properties of such IR's which will be used below.[60–62]

The IR's of the Lorentz group, denoted τ, are characterized by two labels, p and q, which can take positive integer and half-integer values only. (Infinite dimensional representations characterized by continuous labels exist as well, but are not relevant for the present considerations.) The dimension of the IR τ_{pq} is $(2p+1)(2q+1)$. A four vector, for instance, transforms according to the (four-dimensional) representation

$$\tau_{1/2\ 1/2}=\tau_{1/2\ 0}\otimes\tau_{0\ 1/2}, \tag{44}$$

while a four-component Dirac spinor transforms according to the direct sum representation $\tau_{1/2\ 0}\oplus\tau_{0\ 1/2}$.

Product representations of such IR's are, in general, reducible, and can be written as a direct sum of irreducible representations. This is greatly facilitated by noting that the IR's of the Lorentz group satisfy

$$\tau_{pq}\otimes\tau_{nm}=\bigoplus_{\substack{|p-n|\le k\le p+n\\ |q-m|\le l\le q+m}}\tau_{kl}. \tag{45}$$

Obviously, this is just the familiar rule for the coupling of two angular momenta (i.e., for forming the product of two IR's of the rotation group), applied, however, to each of the two labels of the IR's individually. Equation (44) is a special case of this rule. Another example is a general tensor of rank two. Being defined as the direct product of two vectors, it transforms according to

$$\tau_{1/2\ 1/2}\otimes\tau_{1/2\ 1/2}=\tau_{00}\oplus\tau_{11}\oplus\tau_{10}\oplus\tau_{01}. \tag{46}$$

Here τ_{00} denotes the identical representation. Symmetric tensors transform according to the ten-dimensional direct sum representation $\tau_{00}\oplus\tau_{11}$ (note that a symmetric tensor of rank two, defined in a four-dimensional vector space, has ten independent entries). Antisymmetric tensors transform according to the six-dimensional IR's $\tau_{10}\oplus\tau_{01}$ (an antisymmetric tensor in a four-dimensional vector space has six independent entries).

The task of enumerating all possible relativistic OP is now reduced to finding all distinct IR's according to which the anomalous bilinear forms of Dirac spinors transform. Explicitly, the bilinear form $\Psi^T\hat{\eta}_i\Psi$ transforms, for any $\hat{\eta}_i$, according to the direct product representation of two spinor representations, i.e.,

$$[\tau_{1/2\ 0}\oplus\tau_{0\ 1/2}]\otimes[\tau_{1/2\ 0}\oplus\tau_{0\ 1/2}]. \tag{47}$$

The resulting product representation is immediately seen to be reducible. Using Eq. (45) one finds

$$[\tau_{1/2\ 0}\oplus\tau_{0\ 1/2}]\otimes[\tau_{1/2\ 0}\oplus\tau_{0\ 1/2}]=\tau_{00}\oplus\tau_{00}\oplus\tau_{1/2\ 1/2}\oplus\tau_{1/2\ 1/2}\oplus\tau_{10}\oplus\tau_{01}. \tag{48}$$

We find that there are two Lorentz scalars (twice the identical representation), two four vectors (twice $\tau_{1/21/2}$) and one antisymmetric tensor of rank two contained in the product representation. An analysis of the proper Lorentz group alone does not suffice to distinguish between scalars and pseudoscalars, or vectors and axial vectors, respectively. We have therefore obtained two scalars and two vectors. Parity then serves to further classify these quantities according to their behavior under space inversion. It should be noted that exactly the same analysis goes through for the conventional bilinear covariants $\overline{\Psi}\hat{\Gamma}\Psi$, where $\hat{\Gamma}$ is any of the 16 conventional Dirac matrices. As these quantities are formed with the adjoint spinor they transform according to

$$[\tau^*_{1/2\ 0}\oplus\tau^*_{1/2\ 0}\otimes[\tau_{1/2\ 0}\oplus\tau_{0\ 1/2}], \tag{49}$$

where τ^*_{pq} is the complex conjugate representation of τ_{pq}. For the Lorentz group an IR is not generally equivalent to its conjugate IR. It follows immediately that anomalous and conventional bilinear covariants describe distinct physical objects. The transformation behavior of these objects, however, depends only on the labels p and q and is therefore the same for the conventional as for the anomalous bilinear covariants. The resulting classification of the conventional bilinear covariants into scalars, vectors, and an antisymmetric tensor of rank two is, of course, not new, but found in many standard textbooks by way of explicit manipulations of $\hat{\gamma}$ matrices.[50,51,53]

As far as the anomalous bilinear covariants are concerned, Eq. (48) constitutes a rederivation of the above theorem by purely group-theoretical considerations, without having to write down explicit matrices or Dirac-type equations.

IV. SYMMETRY ANALYSIS WITH RESPECT TO DISCRETE SYMMETRIES

A. Discrete symmetries of the Dirac equation

We now generalize the considerations of Sec. II A to the relativistic case, i.e., we construct a symmetry adapted basis set of two-particle states into which any OP can be expanded. In a relativistic situation one has *three* symmetry operations, instead of two, namely time-reversal $\hat{T}$, parity $\hat{P}$ and charge conjugation $\hat{C}$.[50,51,53] Every solution, $|n\rangle$, of the Dirac equation yields four linearly independent states, according to

$$|n\rangle,\quad |\hat{T}n\rangle,\quad |\hat{P}n\rangle,\quad |\hat{P}\hat{T}n\rangle. \tag{50}$$

The corresponding charge-conjugate states are

$$|\hat{C}n\rangle, \quad |\hat{C}\hat{T}n\rangle, \quad |\hat{C}\hat{P}n\rangle, \quad |\hat{C}\hat{P}\hat{T}n\rangle. \tag{51}$$

As will turn out below, a complete set of OP's are only obtained if both types of states, Eqs. (50) and (51), are included.[63] The BCS prescription for pairing requires that the single-particle states are paired to yield a two-particle state with zero center-of-mass momentum. The Pauli principle requires that these two-particle states be antisymmetric under particle exchange. We now construct the appropriate symmetry adapted basis functions, which are the counterparts to Eqs. (2)–(5) in the nonrelativistic case. To this end, we first consider homogeneous systems, in which the momentum is a good quantum number, and introduce four-component spinors of the form

$$\langle \mathbf{r}|ki\rangle = \phi_k(\mathbf{r})\chi_i, \tag{52}$$

where

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \chi_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \chi_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \chi_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{53}$$

We will refer to the space defined by the label i as i space. It is the formal generalization of the nonrelativistic spin space. The three symmetry operators act on $\phi_k\chi_i$ according to[50,51] $\hat{T}[\phi_k\chi_1] = \phi_k^*\chi_2$, $\hat{P}[\phi_k\chi_1] = \phi_{-k}\chi_1$, $\hat{C}[\phi_k\chi_1] = \phi_k^*\chi_4$, etc. Using this notation, the relativistic counterpart of the state $|S\rangle$ is, for homogeneous systems, given by

$$|1\rangle = |ki, \hat{T}ki\rangle - |\hat{P}\hat{T}ki, \hat{P}ki\rangle + |\hat{C}\hat{P}\hat{T}ki, \hat{C}\hat{P}ki\rangle - |\hat{C}ki, \hat{C}\hat{T}ki\rangle. \tag{54}$$

It is readily verified that the configuration-space representation of this state is

$$\langle \mathbf{r}\mathbf{r}'|1\rangle = [\phi_k(\mathbf{r})\phi_{-k}(\mathbf{r}') + \phi_{-k}(\mathbf{r})\phi_k(\mathbf{r}')] \times (\chi_1\otimes\chi_2 - \chi_2\otimes\chi_1 + \chi_3\otimes\chi_4 - \chi_4\otimes\chi_3). \tag{55}$$

$\langle \mathbf{r}\mathbf{r}'|1\rangle$ is seen to be even in the spatial coordinates and odd in the i space coordinates. In this sense it describes the relativistic counterpart of a singlet state. Performing the tensor products of the χ_i yields the matrix

$$\begin{pmatrix} & 1 & & \\ -1 & & & \\ & & & 1 \\ & & -1 & \end{pmatrix}, \tag{56}$$

which is just $\hat{\eta}$, as defined in Eq. (19).

In inhomogeneous systems k is not a good quantum number any more. In complete analogy to the nonrelativistic case one therefore defines the general basis state as

$$|1\rangle = |ni, \hat{T}ni\rangle - |\hat{P}\hat{T}ni, \hat{P}ni\rangle + |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{P}ni\rangle - |\hat{C}ni, \hat{C}\hat{T}ni\rangle, \tag{57}$$

where n stands for a complete set of quantum numbers necessary to label a solution of the Dirac equation. Evidently, Eq. (57) contains Eqs. (54) and (2) as special cases. In the same way in which $|1\rangle$ is mapped onto the matrix $\hat{\eta}$, the matrices $\hat{\eta}_V^0$ to $\hat{\eta}_T^{23}$ correspond, in this order, (up to global factors $\pm 1, \pm i$) to the basis states

$$|2\rangle = |ni, \hat{T}ni\rangle - |\hat{P}\hat{T}ni, \hat{P}ni\rangle - |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{P}ni\rangle + |\hat{C}ni, \hat{C}\hat{T}ni\rangle, \tag{58}$$

$$|3\rangle = |ni, \hat{C}\hat{P}\hat{T}ni\rangle - |\hat{C}\hat{T}ni, \hat{P}ni\rangle - |\hat{T}ni, \hat{C}\hat{P}ni\rangle + |\hat{C}ni, \hat{P}\hat{T}ni\rangle, \tag{59}$$

$$|4\rangle = |ni, \hat{C}\hat{P}\hat{T}ni\rangle - |\hat{C}\hat{T}ni, \hat{P}ni\rangle + |\hat{T}ni, \hat{C}\hat{P}ni\rangle - |\hat{C}ni, \hat{P}\hat{T}ni\rangle, \tag{60}$$

$$|5\rangle = |ni, \hat{C}ni\rangle - |\hat{C}\hat{P}ni, \hat{P}ni\rangle + |\hat{T}ni, \hat{C}\hat{T}ni\rangle - |\hat{C}\hat{P}\hat{T}ni, \hat{P}\hat{T}ni\rangle, \tag{61}$$

$$|6\rangle = |ni, \hat{C}ni\rangle - |\hat{C}\hat{P}ni, \hat{P}ni\rangle - |\hat{T}ni, \hat{C}\hat{T}ni\rangle + |\hat{C}\hat{P}\hat{T}ni, \hat{P}\hat{T}ni\rangle, \tag{62}$$

$$|7\rangle = |ni, \hat{C}ni\rangle + |\hat{C}\hat{P}ni, \hat{P}ni\rangle + |\hat{T}ni, \hat{C}\hat{T}ni\rangle + |\hat{C}\hat{P}\hat{T}ni, \hat{P}\hat{T}ni\rangle, \tag{63}$$

$$|8\rangle = |ni, \hat{P}ni\rangle - |\hat{P}\hat{T}ni, \hat{T}ni\rangle + |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{T}ni\rangle - |\hat{C}ni, \hat{C}\hat{P}ni\rangle, \tag{64}$$

$$|9\rangle = |ni, \hat{P}ni\rangle + |\hat{P}\hat{T}ni, \hat{T}ni\rangle + |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{T}ni\rangle + |\hat{C}ni, \hat{C}\hat{P}ni\rangle, \tag{65}$$

$$|10\rangle = |ni, \hat{T}ni\rangle + |\hat{P}\hat{T}ni, \hat{P}ni\rangle + |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{P}ni\rangle + |\hat{C}ni, \hat{C}\hat{T}ni\rangle, \tag{66}$$

$$|11\rangle = |ni, \hat{C}\hat{P}\hat{T}ni\rangle + |\hat{C}\hat{T}ni, \hat{P}ni\rangle + |\hat{T}ni, \hat{C}\hat{P}ni\rangle + |\hat{C}ni, \hat{P}\hat{T}ni\rangle, \tag{67}$$

$$|12\rangle = |ni, \hat{C}\hat{P}\hat{T}ni\rangle + |\hat{C}\hat{T}ni, \hat{P}ni\rangle - |\hat{T}ni, \hat{C}\hat{P}ni\rangle - |\hat{C}ni, \hat{P}\hat{T}ni\rangle, \tag{68}$$

$$|13\rangle = |ni, \hat{C}ni\rangle + |\hat{C}\hat{P}ni, \hat{P}ni\rangle - |\hat{T}ni, \hat{C}\hat{T}ni\rangle - |\hat{C}\hat{P}\hat{T}ni, \hat{P}\hat{T}ni\rangle, \tag{69}$$

$$|14\rangle = |ni, \hat{T}ni\rangle + |\hat{P}\hat{T}ni, \hat{P}ni\rangle - |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{P}ni\rangle - |\hat{C}ni, \hat{C}\hat{T}ni\rangle, \tag{70}$$

$$|15\rangle = |ni, \hat{P}ni\rangle + |\hat{P}\hat{T}ni, \hat{T}ni\rangle - |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{T}ni\rangle - |\hat{C}ni, \hat{C}\hat{P}ni\rangle, \tag{71}$$

$$|16\rangle = |ni, \hat{P}ni\rangle - |\hat{P}\hat{T}ni, \hat{T}ni\rangle - |\hat{C}\hat{P}\hat{T}ni, \hat{C}\hat{T}ni\rangle + |\hat{C}ni, \hat{C}\hat{P}ni\rangle, \tag{72}$$

respectively. The states $|1\rangle$–$|16\rangle$ are independent of the choice for the label i in the sense that $i=2,3,4$ lead, up to global phase factors, to the same 16 matrices as $i=1$. These matrices are determined only up to within an overall unitary transformation. The same applies, evidently, to the basis states (57)–(72). The present choice is simply a matter of convenience.

This connection between the discrete symmetries of the Dirac equation and the behavior of the relativistic OP under Lorentz transformations[64] can now be used to discuss the physical meaning of the various terms in Eq. (43).

B. Order parameters involving charge conjugation

The charge-conjugation operation $\hat{C}$ relates positive and negative-energy solutions of the Dirac equation to each other. The nature of the corresponding pairs can be understood by expressing the order parameters in terms of creation and annihilation operators of single-particle states rather than in terms of field operators. The field operators Ψ have an expansion of the form:[53]

$$\Psi = \sum_p \sum_{s=1}^{4} f_{s,p} \hat{b}_p^{(s)}, \tag{73}$$

where the $\hat{b}_p^{(s)}$ annihilate positive energy electrons for $s=1,2$ and negative energy electrons for $s=3,4$. The coefficient $f_{s,p}$ can be expressed in terms of solutions of the corresponding Dirac equation.[53] The order parameters thus contain operators for two positive-energy electrons (e.g., $\hat{b}_p^{(1)}\hat{b}_p^{(2)}$), two negative-energy electrons (e.g., $\hat{b}_p^{(3)}\hat{b}_p^{(3)}$) or a positive and a negative-energy electron (e.g., $\hat{b}_p^{(1)}\hat{b}_p^{(3)}$).[65] It follows that there are no electron-positron pairs contained in our theory, as these would require terms like $\hat{b}_p^{(1)}\ (\hat{b}_p^{(3)})^\dagger$.

The appearance of pair states formed with one $\hat{C}$ operation, such as $|ni, \hat{C}ni\rangle$, in the basis set thus reflects the possibility of pairs consisting of a positive and a negative-energy electron. Such pairs require the pairing interaction to bridge the gap of $2mc^2$. As this energy is far bigger than the typical energies of pairing mechanisms discussed in connection with ordinary superconductors, such pairs are unlikely to be realized in solid-state situations. In particular, in the nonrelativistic limit the negative-energy states are completely decoupled from the positive-energy states, so that the pairs formed with one $\hat{C}$ operation do not contribute at all in this limit.

On the other hand, it is *necessary* to include these pairs in the formal development of the theory, because without them one does not recover the complete set of 16 4×4 matrices. Furthermore it should be noted that such pairs are not forbidden by either relativity or symmetry. Whether they are actually realized in nature depends, therefore, only on the existence of a suitable interaction.

In the nonrelativistic limit we only have the states with zero or two $\hat{C}$ operations, i.e., pairs between two positive or two negative-energy states at our disposal to form Cooper pairs. Considering these states it is important to note that the theory is exactly symmetric with respect to the operation of charge conjugation. To every single-particle state in Eq. (50) corresponds one in Eq. (51) which differs from it just by application of $\hat{C}$. Furthermore, as we did not specify the sign of the external potential, A_μ, we have in fact no way to distinguish superconductivity involving two electrons, moving in a lattice of ordinary atoms, from that involving two positrons in a lattice composed of anti atoms. It is therefore a definite (and rather plausible) prediction of the theory that in an ''antiworld'' superconductivity takes place between two positrons instead of two electrons. Since both electrons and positrons are solutions of a Dirac equation there are relativistic corrections to both types of pairs.

Already to second order in v/c, at the weakly relativistic level, these corrections lead to important effects, such as spin-orbit coupling, the Darwin term, etc. From the above considerations it appears highly plausible that there are such corrections to the order parameters as well, which indeed turns out to be the case (see the second paper in this series).

In Table I each matrix is assigned two labels to indicate the nature of the underlying basis states. The *second* of these labels refers to the behavior of the pairs under space inversion. The matrices corresponding to the states formed with zero or two $\hat{C}$ operations are all even under space inversion. These matrices are labeled *E*. Those corresponding to pairs formed with one application of $\hat{C}$ turn out to be odd under inversion and are labeled *O*.

C. Order parameters involving triplet pairing

In the nonrelativistic case, where spin is a good quantum number, all pair states can be classified as either singlet or triplet states. This classification is reflected in the form of the pair states (2)–(5). In the relativistic case, singlet and triplet states lose their meaning because spin is not a good quantum number any more. However, as long as a center of inversion is present, a classification according to the behavior under particle exchange is still possible.[34,1] Order parameters which are even in i space and odd in $\mathbf{r}$ space are the generalization of triplet OP. Those OP which are odd in i space and even in $\mathbf{r}$ space generalize singlet OP.

The symmetry in i space can be immediately read off from the matrix representations of the various OP in Eqs. (19)–(34). All matrices satisfying $\hat{\eta}^T = \hat{\eta}$ are even in i space and lead to triplet OP's in the nonrelativistic limit. These matrices correspond to the states 7–16. In Table I these matrices carry a *T* (for triplet pair) as a first label. Matrices satisfying $\hat{\eta}^T = -\hat{\eta}$, on the other hand, give rise to singlet OP in the nonrelativistic limit. These matrices correspond to the basis states 1–6. In Table I they carry an *S* (for a singlet pair) as a first label.

The matrices labeled *TE* in Table I (corresponding to the states 8,9,10,14,15,16) describe triplet pairs between two positive or two negative-energy states. These are the relativ-

istic generalizations of the Balian-Werthamer matrices $i\hat{\sigma}_y\boldsymbol{\sigma}$, to which they rigorously reduce (up to phase factors ± 1, $\pm i$ which are a consequence of our definition of the $\hat{\eta}$ matrices) in the nonrelativistic limit.

D. Order parameters of the generalized BCS-type

It is perfectly possible to formulate a relativistic theory of superconductivity on the basis of the complete Hamiltonian (43). One would carry on with 16 different order parameters and the associated pair fields, which transform as the components of the various bilinear covariants.

Obviously, this Hamiltonian leads to very general equations of motion whose solution will be rather involved. In most, if not all, realistic situations the interaction leading to superconductivity will pick one of the various coupling channels (i.e., bilinear covariants) and favor one (or a few) of the great variety of order parameters. In the following sections we will thus not continue with the full equation (43).

We assume instead that the nonelativistic limit of our equations must describe a singlet superconductor with pairing between two positive-energy states. OP's with such a nonrelativistic limit will in the following be termed ''order parameters of the generalized BCS type.''

For an analysis of triplet superconductors and more exotic pairs one needs to go back to Eq. (43) and select the relevant terms from Eqs. (19)–(34), just as we shall do below for the BCS case.

By inspection of Table I we find that only two out of the 16 possible OP are of the generalized BCS type. One of them is the Lorentz scalar, formed using the matrix $\hat{\eta}$. The other is the zeroth component of the (polar) four vector, which is formed with the matrix $\hat{\eta}_V^0$.

If we took both of these two OP's into account, the resulting expression for the Lagrangian would not be covariant, because $\hat{\eta}_V^0$ is only a single component of a bilinear covariant. This on its own does not disqualify the $\hat{\eta}_V^0$-OP as a valid OP, but it has the immediate consequence that to form a covariant quantity we need to combine it with the OP $\hat{\eta}_V^1$, $\hat{\eta}_V^2$, and $\hat{\eta}_V^3$ which were already excluded above, because they correspond to pairs between a positive and a negative-energy electron. Therefore we will focus mainly on the scalar OP in the remainder of this paper.

For later reference we now explicitly write out the Hamiltonian containing the generalized BCS-type OP

$$\hat{\chi}(\mathbf{r},\mathbf{r}') = \Psi^T(\mathbf{r})\,\hat{\eta}\Psi(\mathbf{r}'), \tag{74}$$

namely

$$\hat{H} = \int d^3r\,\overline{\Psi}(\mathbf{r})[c\,\hat{\boldsymbol{\gamma}}\cdot\mathbf{p} + mc^2 + q\,\hat{\gamma}^\mu A_\mu]\Psi(\mathbf{r})$$
$$-\frac{1}{2}\int d^3r\,d^3r'[\Psi^T(\mathbf{r})\underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}}_{=\hat{\eta}}\Psi(\mathbf{r}')\Delta^*(\mathbf{r},\mathbf{r}') + H.c.]. \tag{75}$$

This is the BCS version of (43) or, equivalently, the relativistic version of (12) and (13). The other OP of the generalized BCS-type, $\hat{\eta}_V^0$, leads to the term

$$-\frac{1}{2}\int d^3r\,d^3r'[\Psi^T(\mathbf{r})\underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{=\hat{\eta}_V^0}\Psi(\mathbf{r}')\Delta_{V0}^*(\mathbf{r},\mathbf{r}') + H.c.] \tag{76}$$

If (76) is included in Eq. (75), the resulting Hamiltonian contains all terms which in the nonrelativistic limit reduce to BCS-type OP's, but the Lagrangian it corresponds to is not covariant.

The formulation of the relativistic theory of superconductivity given in Ref. 29 is a special case of the present formulation, in which only the center-of-mass degrees of freedom of the Cooper pair are considered and only the scalar OP, Eq. (74), formed with $\hat{\eta}$, is included. The nonrelativistic limit of the approach of the present section has been shown in Ref. 29 to lead to the standard Bogolubov–de Gennes theory of inhomogeneous superconductors[43] with a local OP.

The essence of this reduction can already be seen from the structure of the $\hat{\eta}$-OP. Using the anticommutation relations of the four components of the Ψ, ψ_i, and the symmetry of the pair field $\Delta(\mathbf{r},\mathbf{r}')$ under particle exchange, we can rewrite

$$\frac{1}{2}\int d^3r\,d^3r'\,\Psi^T(\mathbf{r})\,\hat{\eta}\Psi(\mathbf{r}')\Delta^*(\mathbf{r},\mathbf{r}') \tag{77}$$

as

$$\int d^3r\,d^3r'[\psi_1(\mathbf{r})\psi_2(\mathbf{r}') + \psi_3(\mathbf{r})\psi_4(\mathbf{r}')]\Delta^*(\mathbf{r},\mathbf{r}'). \tag{78}$$

Here we have essentially performed the steps in reverse, which led from Eq. (12) to Eq. (13) in the nonrelativistic case. We now see the physical significance of the matrix entries in $\hat{\eta}$. The first product, $\psi_1(\mathbf{r})\psi_2(\mathbf{r}')$, is just the relativistic counterpart to the familiar $\psi_\uparrow(\mathbf{r})\psi_\downarrow(\mathbf{r}')$, to which it rigorously reduces in the nonrelativistic limit.

The second product, $\psi_3(\mathbf{r})\psi_4(\mathbf{r}')$, is the analogous term for the lower components of the Dirac spinor. Since in a relativistic theory upper and lower components must be treated on the same footing, the appearance of these terms is highly plausible. In the nonrelativistic limit the lower components are by a factor of v/c smaller than the upper components. Therefore, the second product, being of order $(v/c)^2$, does not contribute in this limit.

From these considerations it follows that it is not necessary to appeal to the discrete symmetries of the Dirac equation in order to identify the relativistic generalization of the BCS OP, as we have done in the beginning of this subsection. It is sufficient to look at the upper left corner of the various $\hat{\eta}$ matrices, which must be of the form

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{79}$$

in order to reduce to the BCS OP $\psi_\uparrow(\mathbf{r})\psi_\downarrow(\mathbf{r}')$ in the nonrelativistic limit. Obviously, both ways to identify the relativistic generalization of the BCS OP lead to the same result.

V. THE RELATIVISTIC BOGOLUBOV–DE GENNES EQUATIONS

A. Relativistic Bogolubov-Valatin transformation

Equation (75) defines the Dirac–Bogolubov–de Gennes Hamiltonian for the generalized BCS-OP $\hat{\eta}$. This Hamiltonian can be diagonalized by a unitary canonical transformation from the field operators $\Psi(\mathbf{r})$ to new operators a_k. (Usually these new operators are labeled γ_k in the literature on superconductivity. We use a_k in order to avoid confusion with the Dirac matrices $\hat{\gamma}$.) In the nonrelativistic case, the transformation which diagonalizes Eq. (12) is given by the Bogolubov-Valatin transformation[43,66–68]

$$\psi_\uparrow(\mathbf{r}) = \sum_k \left(u_k(\mathbf{r})a_{k\uparrow} - v_k(\mathbf{r})^* a^\dagger_{k\downarrow}\right), \tag{80}$$

$$\psi_\downarrow(\mathbf{r}) = \sum_k \left(u_k(\mathbf{r})a_{k\downarrow} + v_k(\mathbf{r})^* a^\dagger_{k\uparrow}\right), \tag{81}$$

where the coefficients $u_k(\mathbf{r})$ and $v_k(\mathbf{r})$ are determined from the requirement that the transformed Hamiltonian be diagonal. Obviously, the spin of the quasiparticles enters the transformation only in a fixed combination. To treat magnetic impurities, spin-orbit coupling, triplet pairing, etc., this transformation needs to be replaced by the more general form[1,37,41,43,45,69]

$$\psi_\tau(\mathbf{r}) = \sum_{\sigma k} \left[u_{\tau\sigma k}(\mathbf{r})a_{\sigma k} + v^*_{\tau\sigma k}(\mathbf{r})a^\dagger_{\sigma k}\right], \tag{82}$$

where the spin degrees of freedom are involved in the transformation as well. In the relativistic case, the spinlike quantum numbers σ and τ have to be replaced by component labels of the Dirac spinors. The relativistic generalization of Eq. (82) is thus

$$\psi_i(\mathbf{r}) = \sum_{jk} \left[u_{ijk}(\mathbf{r})a_{jk} + v^*_{ijk}(\mathbf{r})a^\dagger_{jk}\right]. \tag{83}$$

There are several conditions the transformation (83) has to satisfy.[49,70] First of all it needs to be unitary (i.e., preserve the normalization of the quasiparticle wave functions) and canonical (i.e., preserve the anticommutation relations of the field operators).[70] Unitarity requires that

$$\int d^3r \sum_i \left[v_{ijk}(\mathbf{r})v^*_{ij'k'}(\mathbf{r}) + u_{ijk}(\mathbf{r})u^*_{ij'k'}(\mathbf{r})\right] = \delta_{kk'}\delta_{jj'} \tag{84}$$

and

$$\int d^3r \sum_i \left[v_{ijk}(\mathbf{r})u_{ij'k'}(\mathbf{r}) + u_{ijk}(\mathbf{r})v_{ij'k'}(\mathbf{r})\right] = 0, \tag{85}$$

while the conditions

$$\sum_{kj} \left[u_{ijk}(\mathbf{r})v^*_{i'jk}(\mathbf{r}') + v^*_{ijk}(\mathbf{r})u_{i'jk}(\mathbf{r}')\right] = 0 \tag{86}$$

and

$$\sum_{kj} \left[u^*_{ijk}(\mathbf{r})u_{i'jk}(\mathbf{r}') + v_{ijk}(\mathbf{r})v^*_{i'jk}(\mathbf{r}')\right] = \delta_{ii'}\delta(\mathbf{r}-\mathbf{r}') \tag{87}$$

ensure that the transformation is canonical. As in the nonrelativistic case, it turns out that the same relations are also obtained by demanding that the solutions of the resulting single-particle equations be complete and orthonormal.[45,49] Explicitly we find that completeness of the solutions follows, if the transformation is canonical, while orthonormality follows, if it is unitary.

B. Dirac–Bogolubov–de Gennes equations

Further conditions on the coefficients $u_{ijk}(\mathbf{r})$ and $v_{ijk}(\mathbf{r})$ follow from demanding that the Hamiltonian (75) be diagonal in the new creation and annihilation operators a_{jk}:

$$H = \sum_{jk} E_{jk}a^\dagger_{jk}a_{jk} + E_0, \tag{88}$$

where E_0 is the ground-state energy and the a_{jk} create and annihilate elementary excitations (Bogolons) with energy E_{jk}. In the same way as in the nonrelativistic case,[43] one finds from Eqs. (83) and (88) that the $u_{ijk}(\mathbf{r})$ and $v_{ijk}(\mathbf{r})$ which diagonalize Eq. (75) satisfy a set of coupled integrodifferential equations of the Bogolubov–de Gennes type. These equations are most conveniently written in a matrix notation as

$$\begin{pmatrix} \hat{h} & \mathcal{D} \\ -\mathcal{D}^* & -\hat{h}^* \end{pmatrix} \begin{pmatrix} u_{jk}(\mathbf{r}) \\ v_{jk}(\mathbf{r}) \end{pmatrix} = E_{jk} \begin{pmatrix} u_{jk}(\mathbf{r}) \\ v_{jk}(\mathbf{r}) \end{pmatrix}. \tag{89}$$

Here h is the kernel of the Dirac Hamiltonian

$$\hat{h} = \hat{\gamma}^0[c\,\hat{\boldsymbol{\gamma}}\cdot\mathbf{p} + mc^2(1-\hat{\gamma}^0) + q\,\hat{\gamma}^\mu A_\mu]. \tag{90}$$

The term $mc^2(1-\hat{\gamma}^0)$ arises from subtracting mc^2 from the energy eigenvalues of Eq. (75) before diagonalization, i.e., measuring the energies relative to the rest energy. $\mathcal{D}$ is an integral operator that contains the pair potential as kernel

$$\mathcal{D}=\int d^3r' \ldots \Delta(\mathbf{r},\mathbf{r}')\,\hat{\eta}. \tag{91}$$

For the case of a local pair potential it reduces to the multiplicative operator $\Delta(\mathbf{R})\,\hat{\eta}$, where $\mathbf{R}$ is the center-of-mass coordinate of the Cooper pairs. Each entry in the matrix in Eq. (89) is thus a 4×4 matrix. Accordingly, the four-component spinors $u_{jk}(\mathbf{r})$ and $v_{jk}(\mathbf{r})$ are given by

$$u_{jk}(\mathbf{r})=\begin{pmatrix} u_{1jk} \\ u_{2jk} \\ u_{3jk} \\ u_{4jk} \end{pmatrix} \qquad v_{jk}(\mathbf{r})=\begin{pmatrix} v_{1jk} \\ v_{2jk} \\ v_{3jk} \\ v_{4jk} \end{pmatrix}. \tag{92}$$

Equation (89) with Eqs. (90)–(92) constitutes the relativistic generalization of the Bogolubov–de Gennes equation. It will in the following be referred to as the Dirac–Bogolubov–de Gennes equation. We can immediately verify that a number of important special cases is contained correctly in Eq. (89): (i) The *nonrelativistic limit* is obtained if we neglect the lower two components of the spinors $u_{jk}(\mathbf{r})$ and $v_{jk}(\mathbf{r})$, which are small in the weakly relativistic limit and zero in the nonrelativistic case. The 8×8 equation (89) then reduces to a 4×4 equation, which is identical to the nonrelativistic 4×4 spin-Bogolubov–de Gennes equation.[43,45,69] (ii) In the *nonsuperconducting limit*, $\mathcal{D}\equiv0$, we obtain the conventional Dirac Hamiltonians for electrons and holes. (iii) In the *local limit* the integral operator $\mathcal{D}$ becomes a multiplicative operator $\Delta(\mathbf{r})$ and we obtain the local version of the Dirac–Bogolubov–de Gennes equation, derived in Ref. 29.

Equations of a similar algebraic form as Eq. (89) were previously proposed in the context of nuclear physics and of Hartree-Fock-Bogolubov theory by Kucharek and Ring[71] and by Zimdahl.[72] The present derivation within the framework of superconductivity and density-functional theory, the detailed symmetry analysis of the order parameter, and the investigation of the weakly relativistic limit and its consequences, presented in the present and the following paper,[30] however, are not contained in these older works.

C. Density-functional aspects

The nonrelativistic Hamiltonian (12) is the starting point for many microscopic investigations of superconductivity, of which we can just discuss a few.[43,69,73–79]

In the nonrelativistic case the potentials $v(\mathbf{r})$ and $\Delta(\mathbf{r},\mathbf{r}')$ appearing in Eq. (12) are either used as parameters in order to simulate, e.g., superconducting multilayers and heterostructures[74,75] or determined microscopically in a self-consistent fashion. The first approach can be used in the relativistic case as well. If, for example, one of the two materials at the interface contains heavy atoms, a relativistic description is called for.

The self-consistent numerical calculations are often done in a mean-field framework[43,76–79,69] or, more recently, using the apparatus of the density-functional theory (DFT).[21,41,44–49] In the DFT for superconductors the interaction leading to superconductivity and the Coulomb interaction are formally eliminated in favor of suitably choosen effective pair, $\Delta(\mathbf{r},\mathbf{r}')$, and lattice, $v(\mathbf{r})$, potentials. The potentials $v(\mathbf{r})$ and $\Delta(\mathbf{r},\mathbf{r}')$ are determined self-consistently as functionals of the density and the order parameter, by solving the Kohn-Sham Bogolubov–de Gennes equations. These effectice potentials can thus be viewed as a convenient way to deal with the interactions at hand.[21,41,44–49]

That such relativistic calculations can become necessary for realistic superconductors containing heavy elements is exemplified by the results of Singh and co-workers,[28] who performed (conventional) relativistic band-structure calculations for $Ba(Sn,Sb)O_3$ and concluded that the absence of superconductivity in these materials is due to relativistic effects on the band structure. The Dirac–Bogolubov–de Gennes equations derived above provide the opportunity to improve such calculations by treating the effects of relativity and superconductivity on the same footing.

A proper relativistic DFT for superconductors has not been formulated as yet. The main reason for this is the problem of variational stability of the relativistic electron gas which, from a purists point of view, is only partly solved, even for normal (i.e., nonsuperonducting) systems.[25] It is not the intention of the present paper to tackle this question for the superconducting case. It is, however, a definite consequence of the present paper that the Kohn-Sham equations of any conceivable relativistic DFT for superconductors *must have the algebraic form of Eq. (89)*. In lieu of a microscopic prescription how to determine the effective potentials A_μ and Δ in this ''Kohn–Sham–Dirac–Bogolubov–de Gennes'' equation, we suggest that they be treated either as adjustable parameters to model realistic materials, as, e.g., in Refs. 74 and 75, or by parametrizing them in terms of the underlying orbitals of the system under study as, e.g., in Refs. 47 and 48.

VI. SUMMARY AND OUTLOOK

The main results of this work are summarized in Table I and Eq. (43). In Table I we classified all possible order parameters consistent with the requirement of relativistic covariance, according to their transformation behavior under Lorentz transformations. This table therefore generalizes previous symmetry classifications of the order parameter from the Galilei group to the Lorentz group. The table contains the relativistic generalizations of the standard BCS order parameter, as well as that of the triplet (Balian-Werthamer) order parameters. It also predicts that there are several other types of order parameters which have not yet been considered in the literature on superconductivity.

Equation (43) contains all these order parameters in a manifestly covariant fashion. It expresses the theorem that all order parameters have to transform like bilinear covariants of the Dirac equation in order to ensure Lorentz invariance.

Many of these results are derived from more than one viewpoint. In particular, the classification of all relativistic OP's as scalar, four vector, etc., expressed in the above theorem and Table I, follows from either one of the following methods: (i) A decomposition of bilinear forms of Dirac spinors according to irreducible representations of the Lor-

entz group. This method is completely general and algebraically extremely simple. However, it does not yield explicit expressions for the OP. (ii) The construction of explicit matrices having the indicated transformation properties. Such matrices in turn can be found in at least three ways: (iia) By making educated guesses as to the form of the matrix, followed by verification through explicit Lorentz transformations. This method is laborious but explicit. (iib) By generalizing the Balian-Werthamer construction (i.e., multiplication of the Pauli matrices with the nonrelativistic time-reversal matrix) to the relativistic domain (i.e., multiplication of Dirac matrices with the relativistic time-reversal matrix). (iic) By forming linear combinations of pair states constructed using the discrete symmetries of the Dirac equation, instead of the Lorentz group.

The latter construction allows for a physical interpretation of the 16 order parameters, leading to the distinction between between triplet and singlet pairs and to identifying pairs involving positive and negative-energy solutions of the Dirac equation.

We identified one of the 16 order parameters as the relativistic version of the BCS order parameter. The Hamiltonian containing this order parameter was diagonalized. The resulting single-particle equations can be regarded as the Dirac equivalent of the Bogolubov–de Gennes equations.

All our considerations in this paper are based on the concepts of pairing and Lorentz invariance. They can therefore be applied whenever pairing takes place, not merely in the case of proper superconductors. Other situations to which our results apply are superfluid helium 3,[37,41] nuclear matter,[80] and the pairing of neutrons and protons in neutron stars.[81,82]

In order to predict observable consequences of the relativistic terms it is advisable to proceed to the weakly relativistic limit. This will be the subject of the second paper in this series, in which a number of reduction techniques are applied to the Dirac–Bogolubov–de Gennes equations. These methods allow one to recover the familiar nonrelativistic equations in zeroth order and to derive relativistic corrections in higher order of v/c. Explicit forms for these corrections will be derived and it will be pointed out in which situations they are relevant for realistic superconductors.

ACKNOWLEDGMENTS

Many helpful discussions with M. Lüders, S. Kurth, T. Kreibich, M. Marques, J. Annett, and B. L. Gyorffy are gratefully acknowleged. 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 No. 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 acknowleged.

[1] M. Sigrist and K. Ueda, Rev. Mod. Phys. **63**, 239 (1991).
[2] J. Appel and P. Hertel, Phys. Rev. B **35**, 155 (1987).
[3] G. E. Volovik and L. P. Gorkov, JETP Lett. **39**, 674 (1984).
[4] J. Annett, Adv. Phys. **39**, 83 (1990).
[5] J. Appel, Phys. Rev. **139**, A1536 (1965).
[6] P. W. Anderson, Phys. Rev. Lett. **3**, 325 (1959).
[7] P. Fulde and K. Maki, Phys. Rev. **141**, 275 (1966).
[8] V. B. Geshkenbein and A. I. Larkin, JETP Lett. **43**, 395 (1986).
[9] A. J. Millis, D. Rainer, and J. A. Sauls, Phys. Rev. **38**, 4504 (1988).
[10] G. S. Lee, Physica C **292**, 171 (1998).
[11] N. R. Werthamer *et al.*, Phys. Rev. **147**, 296 (1966).
[12] K. Maki, Phys. Rev. **148**, 362 (1966).
[13] K. Capelle, E. K. U. Gross, and B. L. Györffy, Phys. Rev. Lett. **78**, 3753 (1997).
[14] K. Capelle, E. K. U. Gross, and B. L. Györffy, Phys. Rev. B **58**, 473 (1998).
[15] K. Capelle, E. K. U. Gross, and B. L. Györffy (unpublished).
[16] B. Cabrera, H. Gutfreund, and W. A. Little, Phys. Rev. B **25**, 6644 (1982).
[17] B. Cabrera and M. E. Peskin, Phys. Rev. B **39**, 6425 (1989).
[18] S. B. Felch *et al.*, Phys. Rev. B **31**, 7006 (1985).
[19] B. Cabrera *et al.*, Phys. Rev. B **42**, 7885 (1990).
[20] T. Holstein, R. E. Norton, and P. Pincus, Phys. Rev. B **8**, 2649 (1973).
[21] W. Kohn, E. K. U. Gross, and L. N. Oliveira, J. Phys. (France) **50**, 2601 (1989).
[22] M. Weger (private communication).
[23] M. Eliashvili and G. Tsitsishvili, Phys. Rev. B **57**, 2713 (1998).
[24] T. Ichiguchi, Phys. Rev. B **57**, 638 (1998).
[25] E. Engel *et al.*, 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).
[26] D. D. Koelling and A. H. MacDonald, in *Relativistic Effects in Atoms, Molecules and Solids*, Vol. 87 of *NATO ASI Series B: Physics*, edited by G. L. Malli (Plenum, New York, 1983).
[27] B. L. Gyorffy *et al.*, in *The Effects of Relativity in Atoms, Molecules and the Solid State*, edited by S. Wilson, I. P. Grant, B. L. Gyorffy (Plenum, New York, 1991).
[28] D. J. Singh, D. A. Papaconstantopoulos, J. P. Julien, and F. Cyrot-Lackman, Phys. Rev. B **44**, 9519 (1991).
[29] K. Capelle and E. K. U. Gross, Phys. Lett. A **218**, 261 (1995).
[30] K. Capelle and E. K. U. Gross, following paper, Phys. Rev. B **59**, 7155 (1999).
[31] K. Ueda and T. M. Rice, Phys. Rev. B **31**, 7114 (1985).
[32] M. R. Norman, Phys. Rev. Lett. **72**, 2077 (1994).
[33] P. Fulde, J. Keller, and G. Zwicknagl, in *Solid State Phyics: Advances in Research and Applications*, edited by H. Ehrenreich and D. Turnbull (Academic Press, San Diego, 1988), Vol. 41, p. 1.
[34] P. W. Anderson, Phys. Rev. B **30**, 4000 (1984).
[35] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
[36] J. R. Schrieffer, *Theory of Superconductivity* (Addison-Wesley, Reading, MA, 1988).
[37] D. Vollhardt and P. Woelfle, *The Superfluid Phases of Helium 3* (Taylor&Francis, London, 1990).
[38] T. Ishiguro and K. Yamagij, *Organic Superconductors*, Springer

Series in Solid State Sciences, Vol. 88 (Springer-Verlag, Berlin, 1990).
[39] H. Tou *et al.*, Phys. Rev. Lett. **77**, 1374 (1996).
[40] R. Balian and N. R. Werthamer, Phys. Rev. **131**, 1553 (1963).
[41] K. Capelle and E. K. U. Gross, Int. J. Quantum Chem. **61**, 325 (1997).
[42] P. W. Anderson, J. Phys. Chem. Solids **11**, 25 (1959).
[43] P. G. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966).
[44] L. N. Oliveira, E. K. U. Gross, and W. Kohn, Phys. Rev. Lett. **60**, 2430 (1988).
[45] 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).
[46] O.-J. Wacker *et al.*, Phys. Rev. Lett. **73**, 2915 (1994).
[47] M. B. Suvasini *et al.*, Phys. Rev. B **48**, 1202 (1993).
[48] M. B. Suvasini and B. L. Gyorffy, Physica B **195**, 109 (1992).
[49] E. K. U. Gross and S. Kurth, Int. J. Quantum Chem., Symp. **25**, 289 (1991).
[50] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* and *Relativistic Quantum Fields* (McGraw-Hill, New York, 1964).
[51] C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).
[52] The matrix $\hat{\eta}$ in Eq. (19) is twice the matrix called $\hat{\eta}$ in Ref. 29. The definition used here is more convenient for the present purpose. Defining $\hat{\eta}$, and consequently $\hat{\eta}^0, \ldots, \hat{\eta}^3$, with a factor $\frac{1}{2}$ (as in Ref. 29) would lead to unpleasant factors $\frac{1}{8}$ and $\frac{1}{32}$ in Eqs. (24)–(34) and in Table I.
[53] J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, MA, 1976).
[54] J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, 2nd ed. (Springer-Verlag, Berlin, 1980).
[55] P. W. Anderson, *Basic Notions of Condensed Matter Physics* (Benjamin/Cummings, Menlo Park, California, 1984), p. 311.
[56] G. E. Volovik and L. P. Gorkov, JETP **61**, 843 (1985).
[57] S. Yip and A. Garg, Phys. Rev. B **48**, 3304 (1993).
[58] M. Lüders, S. Kurth, and E. K. U. Gross (unpublished).
[59] The possibility of forming bilinear covariants with Ψ^T instead of $\overline{\Psi} = \Psi^\dagger \hat{\gamma}^0$ is also pointed out in Sec. 3.4 of Ref. 54, where it is stated that these alternative covariants ''*do not seem to have any physical significance in radiation theory.*'' As follows from the above consideration it is *superconductivity* where the anomalous bilinear covariants aquire physical significance as the relativistic generalization of the order parameter.
[60] G. J. Ljubarski, *Anwendungen der Gruppentheorie in der Physik* (Deutscher Verlag der Wissenschaften, Berlin, 1962).
[61] J. S. Lomont, *Applications of Finite Groups* (Academic, New York, 1959).
[62] J. P. Elliot and P. G. Dawber, *Symmetry in Physics* (MacMillan, London, 1979).
[63] Whether the states (51) are interpreted as positrons or negative-energy electrons depends on whether they are occupied or not. For the construction of a basis set, this is immaterial, because the occupation of the states of the basis set is not a property of the basis set itself. The interpretation of the states involving $\hat{C}$ is discussed in detail in Sec. IV B and the second paper in this series.
[64] From the discussion in Sec. II A it follows that the same connection exists also in the nonrelativistic theory of superconductivity: Linear combinations of pair states constructed from the discrete symmetries of the Schrödinger equation ($\hat{T}$ and $\hat{P}$) are represented in spin space by combinations of matrices which reflect the behavior of the OP under spatial rotations (singlet: scalar, triplet: vector).
[65] It follows that there are no electron-positron pairs contained in our theory, as these would require terms like $\hat{b}_p^{(1)}\,(\hat{b}_p^{(3)})^\dagger$. Note that electron-positron pairs would yield neither supercurrents nor a Meissner effect.
[66] N. N. Bogolubov, Nuovo Cimento **7**, 794 (1958).
[67] J. G. Valatin, Nuovo Cimento **7**, 843 (1958).
[68] M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, New York, 1975).
[69] W. N. Mathews Jr., Phys. Status Solidi B **90**, 327 (1978).
[70] J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems* (MIT Press, Cambridge, 1986).
[71] H. Kucharek and P. Ring, Z. Phys. A **339**, 23 (1997).
[72] W. Zimdahl, Czech. J. Phys. Sect. B **29**, 713 (1979).
[73] J. Bardeen *et al.*, Phys. Rev. **187**, 556 (1969).
[74] H. Plehn *et al.*, J. Low Temp. Phys. **83**, 71 (1991).
[75] H. Plehn *et al.*, Phys. Rev. B **49**, 12 140 (1994).
[76] F. Gygi and M. Schlüter, Phys. Rev. Lett. **65**, 1820 (1990).
[77] F. Gygi and M. Schlüter, Phys. Rev. B **43**, 7609 (1991).
[78] J. D. Shore *et al.*, Phys. Rev. Lett. **62**, 3089 (1989).
[79] M. Rasolt and Z. Tesanovic, Rev. Mod. Phys. **64**, 709 (1992).
[80] P. Ring and P. Schuck, *The Nuclear Many-Body Problem* (Springer-Verlag, New York, 1980).
[81] D. Pines and A. Alpar, Nature (London) **316**, 27 (1985).
[82] J. A. Sauls, D. L. Stein, and J. W. Serene, Phys. Rev. D **25**, 967 (1982).