CHAPTER 8

ELECTROMAGNETIC PROPERTIES OF SUPERCONDUCTORS

Among the many surprising properties of superconductors, their response to externally applied electric and magnetic fields is the most striking. In 1911, while measuring the electrical resistivity of metals at liquid helium temperature, Kamerlingh Onnes[7] discovered that certain metals passed into a radically new phase in which the voltage drop across the metallic specimen vanished even though a finite current was flowing through it. He characterized the new state as one of infinite electrical conductivity or "superconductivity." Equally striking was Meissner and Ochenfeld's[6] discovery that under ideal conditions a superconductor is a perfect diamagnet, that is, the magnetic field strength B vanishes within the bulk of the superconductor.

8-1 LONDON RIGIDITY

In Chapter 1 we gave a qualitative discussion of the origin of these unique properties. Here we take a more formal approach

and show how these effects follow from microscopic considerations. While the mathematical aspects of the discussion become involved at points, the underlying physics which accounts for the phenomena was clearly stated by London[5] in 1935. He suggested that the wave function Ψ_s of the "superfluid" electrons is "rigid" or "stiff" with respect to perturbations due to the presence of a weak magnetic field. Then, as in the problem of diamagnetism of atoms, the vector potential leads to a finite-current density

$$\langle \mathbf{j} \rangle = \langle -ne\mathbf{v} \rangle = -\frac{ne}{m}\left\langle \mathbf{p} + \frac{e\mathbf{A}}{c} \right\rangle = -\frac{ne^2}{mc}\mathbf{A} \tag{8-1}$$

since $\langle \mathbf{p} \rangle = 0$, owing to the rigidity of Ψ_s. This induced current then gives rise to a magnetic field which screens out the external field and leads to perfect diamagnetism in a large system.

The microscopic implications of London's interpretation of the Meissner effect can be seen in the following manner. Since Maxwell's equations ensure that the magnetic field is necessarily a transverse field ($\nabla \cdot \mathbf{B} = 0$), the magnetic perturbation H' only affects the transverse excitations of the system. If Ψ_s is to be essentially unaffected by this (weak) perturbation, the sum of the squares of the first-order perturbation series amplitudes

$$\sum_\alpha |a_\alpha|^2 = \sum_\alpha \left| \frac{\langle \Psi_\alpha | H' | \Psi_s \rangle}{E_\alpha - E_s} \right|^2 \tag{8-2}$$

must be vanishingly small. Here Ψ_α is the state with transverse excitation α present. Presumably, this anomalously small value of the sum will occur if the matrix elements $\langle \Psi_\alpha | H' | \Psi_s \rangle$ tend to zero while the excitation energies $E_\alpha - E_s$ remain finite. Clearly, this will not be the situation for magnetic fields which vary rapidly in space. Such fields create excitations involving electronic states far from the Fermi surface. These states are presumably unaffected by superconducting correlations and therefore lead to finite matrix elements as in the normal state. Fortunately, in establishing the Meissner effect, one only requires the response of the system to magnetic fields which vary slowly in space. In this limit only electronic states near the Fermi surface enter, and there is no reason to expect the matrix elements not to

vanish in this case. Therefore, London's interpretation of the Meissner effect leads one to suspect that (1) the matrix elements for creating transverse excitations from the superfluid by a magnetic field tend to zero for fields which vary slowly in space; (2) there is an energy gap in the transverse excitation spectrum of the superfluid. As we shall see below, these conditions are satisfied by the BCS theory. It is possible, however, that the matrix elements and the energy denominators *both* vanish in the long wavelength limit, in such a way as to give a finite sum. If this sum does not exactly cancel the diamagnetic current in this limit, a Meissner effect is still obtained, as in superconductors with $l \neq 0$ pairing[170] and "gapless" superconductors.[172] The essential difference between the normal and superconducting states in metals is that the paramagnetic and diamagnetic currents do not exactly cancel in the long wavelength limit in the latter.

On the basis of these arguments one might wonder why an insulator, which has as an energy gap for creating transverse electronic excitations, is not also a perfect diamagnet. The point here is that the energy gap arises from the one-body crystal potential in this case, rather than the effective electron-electron interaction. One can derive a sum rule which shows that for insulators the wave function shifts just enough to make $\langle \mathbf{p} \rangle$ cancel the diamagnetic term $e\mathbf{A}/c$ in (8-1) except for a weak diamagnetism, as in normal metals.

If one can explain the Meissner effect, the "infinite conductivity" observed by Onnes can also be explained. This follows because one can show the currents flowing in his configuration (i.e., a superconducting section in an otherwise normal circuit) were diamagnetic in the superconducting section.[1] That is, the currents in the superconductor were due to electrons described by a wave function which was essentially the same as in the absence of the current. The finite current then arose self-consistently from the magnetic field which was generated by the current itself.

In addition to the Meissner effect, one must understand the stability of persistent currents in a multiply connected body,

for example, a superconducting ring. While the Meissner effect plays a role in the details of the phenomenon, the currents are not primarily diamagnetic in origin in this case. On the contrary, the wave function for the current-carrying state differs greatly from that in the absence of currents, in contrast with the situation described above. However, the effect is again due to a "rigidity" of the wave function with respect to all fluctuations which occur with a finite thermodynamic probability. That is, essentially all fluctuations lead to states of higher free energy and therefore regress without leading to decay of the current.

8-2 WEAK-FIELD RESPONSE

In the beginning of Chapter 2 we argued that transverse electromagnetic fields need not be included directly in calculating the detailed pairing interactions which bring about superconductivity. Their effect can be taken into account in terms of a space- and time-dependent average field which is calculated self-consistently from the external field and the currents flowing in the material. While the externally applied magnetic field generally represents a large perturbation on the system, the induced field arising from the supercurrents cancels the external field over most of the material, as we know from the Meissner effect. Therefore, the net field acts only very near the surface and can often be treated as a weak perturbation on the system as a whole. Thus, we shall formally treat the total transverse electromagnetic field as an externally applied field and solve for self-consistency as a separate problem.

As we have seen above, the Coulomb potential plays an essential role in the pairing theory. It cannot be treated by the self-consistent field scheme we use for the transverse field and therefore we include the total Coulomb interaction in the zero-order Hamiltonian.

We begin by considering a simply connected bulk superconductor of unit volume in the presence of a weak externally applied electromagnetic field described by the vector and scalar

potentials $\mathbf{A}(\mathbf{r}, t)$ and $\varphi(\mathbf{r}, t)$, respectively. As usual, we use periodic boundary conditions. For convenience, we write

$$A_\mu(x) = \begin{cases} A_i(x) & (\mu = i = 1, 2, 3) \\ c\varphi(x) & (\mu = 0) \end{cases} \tag{8-3}$$

where $x \equiv (\mathbf{r}, t)$. To first order in A_μ, the coupling of the electrons to the electromagnetic field is

$$\begin{aligned} H^p &= -\frac{1}{c}\int \sum_\mu j_\mu{}^p(x) A_\mu(x)\, d^3r \\ &= -\frac{1}{c}\int [\mathbf{j}^p(x)\cdot\mathbf{A}(x) - \rho_e(x)c\varphi(x)]\, d^3r \end{aligned} \tag{8-4}$$

where we use the metric $(1, 1, 1, -1)$ in μ-sums for $\mu = 1, 2, 3$, and 0, respectively. We call H^p the paramagnetic coupling. The paramagnetic four-current is defined by

$$j_\mu{}^p(x) = \begin{cases} j_i{}^p(x) \equiv -\dfrac{e}{2mi}\displaystyle\sum_s \{\psi_s{}^+(x)\nabla_i\psi_s(x) - [\nabla_i\psi_s{}^+(x)]\psi_s(x)\} \\ \qquad\qquad (\mu = i = 1, 2, 3) \\ \rho_e(x) = -e\displaystyle\sum_s \psi_s{}^+(x)\psi_s(x) = -e\rho(x) \qquad (\mu = 0) \end{cases} \tag{8-5}$$

the (1, 2, 3) components giving the electronic current density operator in the absence of A and the last component being the electronic charge density operator. The physical current density $j_\mu(x)$ in the presence of A is the sum

$$j_\mu(x) = j_\mu{}^p(x) + j_\mu{}^d(x) \tag{8-6}$$

where the diamagnetic current density j^d is given by

$$j_\mu{}^d(x) = \begin{cases} \dfrac{e}{mc}\rho_e(x)A_i(x) & (\mu = i = 1, 2, 3) \\ 0 & (\mu = 0) \end{cases} \tag{8-7}$$

The full coupling of the electrons to the perturbing electromagnetic field is then

$$H' = H^p + H^d$$

where the diamagnetic coupling is defined by

$$H^d = -\frac{e}{2mc^2}\int \rho_e(x)\sum_{i=1}^{3} A_i{}^2(x)\, d^3r \tag{8-8}$$

Therefore, the total system Hamiltonian is

$$\mathscr{H}' = H + H'$$

If we work in an interaction representation where H' is taken to be the perturbation, and assume that $A_\mu \to 0$ as $t \to -\infty$, the ground state of the system in the presence of A evolves in time according to

$$|\Phi(t)\rangle = T\exp\left[-i\int_{-\infty}^{t} H'(t')\, dt'\right]|0\rangle \equiv U(t, -\infty)|0\rangle \tag{8-9}$$

Here $|0\rangle$ is the ground state of H and all quantities are expressed in the interaction representation. Therefore, the expectation value of the current density in the state $|\Phi(t)\rangle$ is given by

$$J_\mu(x) = \langle\Phi(t)|j_\mu(\mathbf{r}, t)|\Phi(t)\rangle = \langle 0|U^+(t, -\infty)j_\mu(\mathbf{r}, t)U(t, -\infty)|0\rangle \tag{8-10}$$

Since we are interested in the terms of J_μ which are first order in A_μ we have

$$\begin{aligned} J_\mu(x) = \frac{e}{mc}\langle 0|\rho_e(x)|0\rangle A_\mu(x)[1 - \delta_{\mu,0}] \\ - i\langle 0|\left[j_\mu{}^p(\mathbf{r}, t), \int_{-\infty}^{t} H'(t')\, dt'\right]|0\rangle \end{aligned} \tag{8-11}$$

The zeroth-order terms in J_μ vanish except for the average electronic-charge density $\langle j_0(x)\rangle$, which does not interest us here. By using the expressions (8-4) and (8-7) we find that the linear response J_μ and the externally applied potential A_μ are related by a nonlocal kernel $K_{\mu\nu}$:

$$J_\mu(x) = -\frac{c}{4\pi}\sum_\nu \int K_{\mu\nu}(\mathbf{r}, t; \mathbf{r}', t')A_\nu(\mathbf{r}', t')\, d^3r'\, dt' \tag{8-12}$$

where the spatial integral runs over the unit volume and the time integral extends from $-\infty$ to ∞. The electromagnetic response kernel $K_{\mu\nu}$ is given by

$$\begin{aligned} K_{\mu\nu}(x; x') = -\frac{4\pi i}{c^2}\langle 0|[j_\mu{}^p(x), j_\nu{}^p(x')]|0\rangle\theta(t - t') \\ - \frac{4\pi c}{mc^2}\langle 0|\rho_e(x)|0\rangle\, \delta^4(x - x')\, \delta_{\mu\nu}[1 - \delta_{\nu,0}] \end{aligned} \tag{8-13a}$$

where the theta function is defined by

$$\theta(t - t') = \begin{cases} 1 & (t > t') \\ 0 & (t < t') \end{cases} \tag{8-13b}$$

If the system is translationally invariant, $K_{\mu\nu}$ depends only on the difference $x - x' = (\mathbf{r} - \mathbf{r}', t - t')$. In this case it is convenient to work with the spatial Fourier transform of $K_{\mu\nu}$ defined by

$$\begin{aligned} K_{\mu\nu}(\mathbf{q}, t - t') &= \int K_{\mu\nu}(x; x') e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\, d^3r\, d^3r' \\ &= -\frac{4\pi i}{c^2} \langle 0|[j_\mu{}^p(\mathbf{q}, t), j_\nu{}^p(-\mathbf{q}, t')]|0\rangle \theta(t - t') \\ &\quad + \frac{4\pi n e^2}{mc^2} \delta(t - t')\, \delta_{\mu\nu}(1 - \delta_{\nu,0}) \end{aligned} \tag{8-14}$$

where n is the number of electrons per unit volume. Since the diamagnetic (second) term in (8-14) is known explicitly, we concentrate on the paramagnetic (first) term in this expression, and define

$$R_{\mu\nu}(\mathbf{q}, \tau) = -i\langle 0|j_\mu{}^p(\mathbf{q}, \tau),\, j_\nu{}^p(-\mathbf{q}, 0)]|0\rangle \theta(\tau) \tag{8-15}$$

If the ground-state wave function were "rigid" with respect to *all* perturbations (rather than only those which lead to transverse excitations) $R_{\mu\nu}$ would be identically zero and (8-12) would reduce to London's equation

$$J_i(x) = -\frac{ne^2}{mc} A_i(x) \qquad (\mu = i = 1, 2, 3) \tag{8-16}$$

This relation is clearly not gauge-invariant since the predicted current depends upon the choice of gauge. In London's equation, only the transverse part of $\mathbf{A}$ is to be used[1] and therefore $\mathbf{J}$ is properly gauge-invariant. Since the longitudinal part of $\mathbf{A}$ couples to longitudinal excitations, the wave function is not "rigid" with respect to this type of perturbation and the paramagnetic term does not vanish in this case. In fact, if $\mathbf{A}$ is purely a gauge potential, the paramagnetic and diamagnetic terms exactly cancel as required by gauge invariance. In carrying out

an approximate evaluation of $K_{\mu\nu}$, one may be able to accurately treat only excitations which enter the transverse response of the system. In this case the longitudinal part of the paramagnetic and diamagnetic currents will not cancel in general and the resultant current will not be manifestly gauge invariant. Nevertheless, if one recognizes the difficulty and uses only that part of $K_{\mu\nu}$ which is accurately calculated, correct physical predictions would be obtained for transverse fields. This is exactly the situation we shall meet when $K_{\mu\nu}$ is evaluated within the pairing (BCS) approximation. The inclusion of longitudinal collective modes or superfluid flow then restores gauge invariance by correcting the longitudinal part of the paramagnetic term.

In our discussion thus far, we have always dealt with time-ordered products of operators rather than retarded commutators of operators as appear in (8-15). It is the former that we can more readily handle by the Green's function scheme. Fortunately, $R_{\mu\nu}$ can be expressed in terms of a time-ordered product of current densities if one works with the time Fourier transforms of these quantities. To see this we note that $R_{\mu\nu}(q, q_0)$ defined by

$$R_{\mu\nu}(\mathbf{q}, \tau) = \int_{-\infty}^{\infty} R_{\mu\nu}(\mathbf{q}, q_0) e^{-iq_0\tau} \frac{dq_0}{2\pi} \tag{8-17}$$

can be expressed in the spectral form

$$R_{\mu\nu}(\mathbf{q}, q_0) = \int_{-\infty}^{\infty} \frac{C_{\mu\nu}(\mathbf{q}, \omega)\, d\omega}{q_0 - \omega + i\delta} \tag{8-18}$$

The spectral weight function $C_{\mu\nu}(\mathbf{q}, \omega)$ is given by

$$\begin{aligned} C_{\mu\nu}(\mathbf{q}, \omega) = &\sum_n \langle 0|j_\mu{}^p(\mathbf{q})|n\rangle\langle n|j_\nu{}^p(-\mathbf{q})|0\rangle\, \delta(E_n - E_0 - \omega) \\ &- \sum_n \langle 0|j_\nu{}^p(-\mathbf{q})|n\rangle\langle n|j_\mu{}^p(\mathbf{q})|0\rangle\, \delta(E_n - E_0 + \omega) \end{aligned} \tag{8-19}$$

where

$$H|n\rangle = E_n|n\rangle \tag{8-20}$$

This spectral representation can be checked by inserting the complete set of intermediate states $|n\rangle$ between the operators in (8-15) and comparing this result with the expression given by combining (8-17) to (8-19).

Consider the corresponding time-ordered product expression

$$P_{\mu\nu}(\mathbf{q}, \tau) = -i\langle 0|T\{j_\mu{}^p(\mathbf{q}, \tau)j_\nu{}^p(-\mathbf{q}, 0)\}|0\rangle \tag{8-21}$$

Its time Fourier transform, defined by

$$P_{\mu\nu}(\mathbf{q}, \tau) = \int_{-\infty}^{\infty} P_{\mu\nu}(\mathbf{q}, q_0)e^{-iq_0\tau}\frac{dq_0}{2\pi} \tag{8-22}$$

is given by the spectral representation

$$P_{\mu\nu}(\mathbf{q}, q_0) = \int_{-\infty}^{\infty} \frac{C_{\mu\nu}(\mathbf{q}, \omega)\, d\omega}{q_0 - \omega + i\,\delta\omega} \tag{8-23}$$

as one can check by direct calculation. By comparing the spectral forms (8-18) and (8-23) we see that in the case that $C_{\mu\nu}(\mathbf{q}, \omega)$ is real, the real parts of $P_{\mu\nu}$ and $R_{\mu\nu}$ are identical while the imaginary parts differ by a minus sign for $q_0 < 0$; thus

$$\text{Re } P_{\mu\nu}(\mathbf{q}, q_0) = \text{Re } R_{\mu\nu}(\mathbf{q}, q_0) \tag{8-24a}$$

$$\text{Im } P_{\mu\nu}(\mathbf{q}, q_0) = \text{sgn } q_0 \text{ Im } R_{\mu\nu}(\mathbf{q}, q_0) \tag{8-24b}$$

More generally, the discontinuity of $P_{\mu\nu}(\mathbf{q}, q_0)$ across the cut determines $C_{\mu\nu}$ from which $R_{\mu\nu}$ can be obtained with the use of (8.18). Therefore, $R_{\mu\nu}$ is known once $P_{\mu\nu}$ is determined. (Since the expression for $K_{\mu\nu}$ involves only the system in the absence of A, the operators j_μ and $j_\mu{}^p$ are identical in this case and we shall often suppress the script p in the operator $j_\mu{}^p$.)

Summarizing the results obtained thus far, we find the response of the system to a weak externally applied potential $A_\mu(q) = [A(q), c\varphi(q)]$ is given by

$$\begin{aligned} J_\mu(q) &= -\frac{c}{4\pi}\sum_\nu K_{\mu\nu}(q)A_\nu(q) \\ &= -\frac{c}{4\pi}\left[\sum_{i=1}^{3} K_{\mu i}(q)A_i(q) - K_{\mu 0}(q)A_0(q)\right] \end{aligned} \tag{8-25}$$

where $q \equiv (\mathbf{q}, q_0)$. The kernel $K_{\mu\nu}$ is given by combining the expressions (8-14) and (8-15),

$$K_{\mu\nu}(q) = \frac{4\pi}{c^2}R_{\mu\nu}(q) + \frac{1}{\lambda_L{}^2}\delta_{\mu,\nu}[1 - \delta_{\nu,0}] \tag{8-26}$$

the two terms giving rise to the paramagnetic and diamagnetic currents, respectively. The quantity $\lambda_L{}^2 = mc^2/4\pi ne^2$ is the

square of the London penetration depth. The paramagnetic kernel $R_{\mu\nu}$ is given in terms of the time-ordered quantity

$$P_{\mu\nu}(q) = \int_{-\infty}^{\infty} (-i)\langle 0|T\{j_\mu(q, \tau)j_\nu(-q, 0)\}|0\rangle e^{iq_0\tau}\, d\tau \tag{8-27}$$

by the relations (8-24).

8-3 THE MEISSNER–OCHSENFELD EFFECT

As Schafroth has shown,[14] the Meissner effect requires that the transverse part of the kernel $K_{\mu\nu}$ remain finite in the long wavelength limit ($\mathbf{q} \to 0$) for zero frequency ($q_0 = 0$). Now gauge invariance and charge conservation require that

$$\sum_{j=1}^{3} K_{ij}q_j = 0 \qquad \text{(gauge invariance)} \tag{8-28a}$$

and

$$\sum_i q_i K_{ij} = 0 \qquad \text{(charge conservation)} \tag{8-28b}$$

for $q_0 = 0$. When these relations are combined with the rotational invariance of the ground state $|0\rangle$, it follows that K_{ij} is of the form

$$K_{ij}(\mathbf{q}, 0) = \left[\delta_{ij} - \frac{q_i q_j}{\mathbf{q}^2}\right]K(\mathbf{q}^2) \tag{8-29}$$

The Meissner effect then requires that

$$K(\mathbf{q}^2) > 0 \qquad \text{as} \qquad \mathbf{q}^2 \to 0 \tag{8-30}$$

since the factor $[\delta_{ij} - q_iq_j/\mathbf{q}^2]$ guarantees that K_{ij} is purely transverse in this case.

The original BCS calculation of J_i was carried out in the transverse gauge, that is, $\mathbf{q} \cdot \mathbf{A}(\mathbf{q}) = 0$. In this gauge only the transverse part of K_{ij} is calculated, and one does not try to ensure that the longitudinal part of K_{ij} vanishes, as required by (8-29). It is, however, instructive to calculate the entire kernel within the BCS approximation so that we can understand the role of collective modes or superfluid flow in giving the correct result for

the longitudinal part of K. To calculate the paramagnetic portion of K we require

$$P_{ij}(\mathbf{q}, \tau) = -i\langle 0|T\{j_i(\mathbf{q}, \tau)j_j(-\mathbf{q}, 0)\}|0\rangle \tag{8-31}$$

The current-density operator $\mathbf{j}(\mathbf{q})$ is given by the Fourier transform of the expression (8-5) and one finds

$$\mathbf{j}(\mathbf{q}) = -\frac{e}{m}\sum_{\mathbf{k}, s}\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)c_{\mathbf{k}s}{}^{+}c_{\mathbf{k}+\mathbf{q}, s} \tag{8-32}$$

Therefore P_{ij} becomes

$$P_{ij}(\mathbf{q}, \tau) = -\frac{ie^2}{m^2}\sum_{\mathbf{k}, \mathbf{k}', s, s'}\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)_i\left(\mathbf{k}' + \frac{\mathbf{q}}{2}\right)_j \times \langle 0|T\{c_{\mathbf{k}s}{}^{+}(\tau)c_{\mathbf{k}+\mathbf{q}, s}(\tau)c_{\mathbf{k}'+\mathbf{q}, s'}{}^{+}(0)c_{\mathbf{k}'s'}(0)\}|0\rangle \tag{8-33}$$

We could evaluate this expression within the pairing approximation by replacing the exact ground state by the BCS ground state (2-33)

$$|\psi_0\rangle = \prod_{\mathbf{k}}(u_{\mathbf{k}} + v_{\mathbf{k}}b_{\mathbf{k}}{}^{+})|0\rangle \tag{8-34}$$

and expressing the c's in terms of the quasi-particle operators by the Bogoliubov–Valatin transformation (2-56). The time dependence is then approximated by that of free quasi-particles and the vacuum expectation value is evaluated in the standard way.

An equivalent procedure, which is more easily generalized beyond the pairing scheme, is to express (8-31) in terms of the Nambu field Ψ_{μ}.[114] One then makes a Hartree factorization of the expectation value. In the Nambu notation, $\mathbf{j}(\mathbf{q})$ takes the form

$$\mathbf{j}(\mathbf{q}) = -\frac{e}{m}\sum_{\mathbf{k}}\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)(\Psi_{\mathbf{k}}{}^{+}\mathbf{1}\Psi_{\mathbf{k}+\mathbf{q}}) \tag{8-35}$$

and P_{ij} becomes

$$P_{ij}(\mathbf{q}, \tau) = -\frac{ie^2}{m^2}\sum_{\mathbf{k}, \mathbf{k}'}\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)_i\left(\mathbf{k}' + \frac{\mathbf{q}}{2}\right)_j \times \langle 0|T\{\Psi_{\mathbf{k}}{}^{+}(\tau)\mathbf{1}\Psi_{\mathbf{k}+\mathbf{q}}(\tau)\Psi_{\mathbf{k}'+\mathbf{q}}{}^{+}(0)\mathbf{1}\Psi_{\mathbf{k}'}(0)\}|0\rangle \tag{8-36}$$

Within the Hartree factorization, the expectation value becomes

$$-\mathrm{Tr}\,[\langle 0|T\{\Psi_{\mathbf{k}+\mathbf{q}}(\tau)\Psi_{\mathbf{k}+\mathbf{q}}{}^{+}(0)\}|0\rangle\langle 0|T\{\Psi_{\mathbf{k}}(-\tau)\Psi_{\mathbf{k}}{}^{+}(0)\}|0\rangle]\,\delta_{\mathbf{k}, \mathbf{k}'} = \mathrm{Tr}\,[\mathbf{G}(\mathbf{k} + \mathbf{q}, \tau)\mathbf{G}(\mathbf{k}, -\tau)]\,\delta_{\mathbf{k}, \mathbf{k}'} \tag{8-37}$$

To make connection with Gor'kov's formulation,[121] we note that in carrying out the trace in (8-37), terms of the form $G_{11}G_{11}'$ and $G_{22}G_{22}'$ correspond to GG' in Gor'kov's notation, while $G_{12}G_{21}'$ and $G_{21}G_{12}'$ correspond to products of his F-functions.

Within this Hartree-like approximation, the time Fourier transform of P_{ij} is given by

$$P_{ij}(q) = -\frac{ie^2}{m^2}\int\frac{d^4k}{(2\pi)^4}\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)_i\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)_j \mathrm{Tr}\,[\mathbf{G}(k + q)\mathbf{G}(k)] \tag{8-38}$$

where $q \equiv (\mathbf{q}, q_0)$. If the pairing potential is nonretarded, we saw in Chapter 7 that within the pairing approximation $\mathbf{G}(k)$ is given by

$$\mathbf{G}(k) = \frac{k_0\mathbf{1} + \epsilon_k\tau_3 + \Delta_k\tau_1}{k_0{}^2 - E_k{}^2 + i\,\delta} \qquad [E_k = (\epsilon_k{}^2 + \Delta_k{}^2)^{1/2}] \tag{8-39}$$

[see (7-41)]. Since we are interested in the static Meissner effect, we set $q_0 = 0$ and P_{ij} reduces to

$$P_{ij}(\mathbf{q}, 0) = -2\left(\frac{e}{m}\right)^2\int\frac{d^3k}{(2\pi)^3}\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)_i\left(\mathbf{k} + \frac{\mathbf{q}}{2}\right)_j L(\mathbf{k}, \mathbf{q}) \tag{8-40a}$$

where the function $L(\mathbf{k}, \mathbf{q})$ is defined by

$$L(\mathbf{k}, \mathbf{q}) = \frac{i}{2}\int_{-\infty}^{\infty}\frac{dk_0}{2\pi}\mathrm{Tr}\,[\mathbf{G}(k + q)\mathbf{G}(k)] = i\int_{-\infty}^{\infty}\frac{dk_0}{2\pi}\frac{(k_0{}^2 + \epsilon_k\epsilon_{k+q} + \Delta_k\Delta_{k+q})}{(k_0{}^2 - E_k{}^2 + i\,\delta)(k_0{}^2 - E_{k+q}{}^2 + i\,\delta)} \tag{8-40b}$$

In the reduction, we have used the relations

$$\begin{aligned}\tau_i{}^2 &= \mathbf{1}\\ \mathrm{Tr}\,\mathbf{1} &= 2\\ \mathrm{Tr}\,\tau_i &= 0 = \mathrm{Tr}\,\tau_i\tau_j \qquad (i \neq j)\end{aligned} \tag{8-40c}$$

The integral is performed by closing the contour in the upper (or lower) half-plane and one finds for $L(\mathbf{k}, \mathbf{q})$ the real quantity

$$L(\mathbf{k}, \mathbf{q}) = \frac{1}{2}\left(1 - \frac{\epsilon_k \epsilon_{k+q} + \Delta_k \Delta_{k+q}}{E_k E_{k+q}}\right) \frac{1}{E_k + E_{k+q}} = \frac{p^2(\mathbf{k}, \mathbf{k} + \mathbf{q})}{E_k + E_{k+q}} \tag{8-41}$$

where $p(\mathbf{k}, \mathbf{k} + \mathbf{q})$ is the coherence factor we met in Chapter 3. This is the result of BCS.[8, 9]

To establish the Meissner effect within the pairing approximation, we note that $L(\mathbf{k}, \mathbf{q}) \to 0$ as $\mathbf{q} \to 0$, owing to the coherence factor $p^2(\mathbf{k}, \mathbf{q})$ vanishing in this limit and the energy denominator remaining finite ($E_k + E_{k+q} \geq 2\Delta_0$). Therefore,

$$\lim_{\mathbf{q}\to 0} P_{ij}(\mathbf{q}, 0) = \lim_{\mathbf{q}\to 0} R_{ij}(\mathbf{q}, 0) = 0 \tag{8-42}$$

and the electromagnetic kernel reduces to the London kernel

$$\lim_{\mathbf{q}\to 0} K_{ij}(\mathbf{q}, 0) = \frac{1}{\lambda_L^2}\delta_{ij} \qquad (i, j = 1, 2, 3) \tag{8-43}$$

The transverse part of this expression is

$$\lim_{q\to 0} K_{ij}(\mathbf{q}, 0) = \frac{\delta_{ij} - q_i q_j}{\mathbf{q}^2}\frac{1}{\lambda_L^2} \tag{8-44}$$

By comparing this result with the general form (8-29) we see that

$$\lim_{\mathbf{q}^2\to 0} K(\mathbf{q}^2) = \frac{1}{\lambda_L^2} > 0 \tag{8-45}$$

which establishes the Meissner effect at zero temperature. Unfortunately, the longitudinal part of K does not vanish in this approximation, but is given by $(q_i q_j/q^2)(1/\lambda_L^2)$; however, this unphysical longitudinal response will be eliminated below.

The above derivation emphasizes the role of the energy gap in bringing about the Meissner effect. Aside from scale factors, the quantity $L(\mathbf{k}, \mathbf{q})$, given by (8-41), is the square of magnetic perturbation matrix element taken between the ground state and a transverse excited state, divided by the excitation energy for

the two quasi-particles which are excited. In the superconducting state, the matrix element vanishes as $\mathbf{q}^2 \to 0$ [i.e., $p^2(\mathbf{k}, \mathbf{q}) \to 0$] and the energy denominator remains finite, in accordance with the discussion in the beginning of this chapter. Therefore, only the diamagnetic term in $K_{\mu\nu}$ survives in this long wavelength limit. In the normal metal $L(\mathbf{k}, \mathbf{q})$ goes over to the conventional result of second-order perturbation theory:

$$L_N(\mathbf{k}, \mathbf{q}) = \frac{1}{|\epsilon_{k+q} - \epsilon_k|} \tag{8-46}$$

if the $\mathbf{k}$ and $\mathbf{k} + \mathbf{q}$ are on opposite sides of the Fermi surface, and zero otherwise, as required by the Pauli principle. As we argued above, a finite value of $P_{ij}(\mathbf{q}, 0)$ arises in the normal metal as $\mathbf{q}^2 \to 0$ despite the fact that most of the matrix elements vanish, because of the vanishingly small-energy denominator in this case. If one calculates the magnitude of P_{ij} one finds that it almost exactly cancels the diamagnetic term, leaving the weak Landau diamagnetism of the normal state.

In "gapless" superconductors[172] both the matrix elements and energy denominators vanish, but the density of states near the Fermi surface is small enough to ensure that P_{ij} does not cancel the diamagnetic term as $\mathbf{q} \to 0$.

To extend this calculation to finite temperature, we use the prescription discussed in Chapter 7 to convert the zero temperature form (8-38) to one involving the discrete frequency sums. The only change is that $L(\mathbf{k}, \mathbf{q})$ becomes

$$L(\mathbf{k}, \mathbf{q}) = -\frac{1}{2\beta}\sum_{n=-\infty}^{\infty} \mathrm{Tr}\,[\mathbf{G}(\mathbf{k} + \mathbf{q}, i\omega_n)\mathbf{G}(\mathbf{k}, i\omega_n)] \tag{8-47}$$

where $\omega_n = (2n + 1)\pi/\beta$. As before we convert the sum to an integral using (7-107) and obtain

$$L(\mathbf{k}, \mathbf{q}) = -\frac{i}{2}\int_c \frac{d\omega}{2\pi}\,\mathrm{Tr}\,[\mathbf{G}(\mathbf{k} + \mathbf{q}, \omega)\mathbf{G}(\mathbf{k}, \omega)]f(\omega) \tag{8-48}$$

where the integral encircles the entire imaginary axis in a counter-clockwise sense and $f(\omega) \equiv [e^{\beta\omega} + 1]^{-1}$ is the Fermi function. By expanding the contour to infinity we pick up residues at the four poles $\pm E_k \equiv \pm E$ and $\pm E_{k+q} \equiv \pm E'$, as shown in Figure 8-1. These residues lead to the real expression

$$L(\mathbf{k}, \mathbf{q}) = \frac{E^2 + \epsilon\epsilon' + \Delta\Delta'}{2E(E^2 - E'^2)}[1 - 2f(E)] + \frac{E'^2 + \epsilon\epsilon' + \Delta\Delta'}{2E'(E'^2 - E^2)}[1 - 2f(E')] \quad (8\text{-}49)$$

which can be written as

$$L(\mathbf{k}, \mathbf{q}) = \frac{p^2(\mathbf{k}, \mathbf{q})}{E_\mathbf{k} + E_{\mathbf{k}+\mathbf{q}}}[1 - f(E_\mathbf{k}) - f(E_{\mathbf{k}+\mathbf{q}})] + \frac{l^2(\mathbf{k}, \mathbf{q})}{E_\mathbf{k} - E_{\mathbf{k}+\mathbf{q}}}[f(E_{\mathbf{k}+\mathbf{q}}) - f(E_\mathbf{k})] \quad (8\text{-}50)$$

The coherence factors p^2 and l^2 are given by

$$\frac{1}{2}\left(1 \mp \frac{\epsilon_\mathbf{k}\epsilon_{\mathbf{k}+\mathbf{q}} + \Delta_\mathbf{k}\Delta_{\mathbf{k}+\mathbf{q}}}{E_\mathbf{k}E_{\mathbf{k}+\mathbf{q}}}\right) \quad (8\text{-}51)$$

the upper and lower signs applying to p and l, respectively. Since we are interested in establishing the Meissner effect at finite temperature we consider the limit of (8-50) as $\mathbf{q} \to 0$. The first

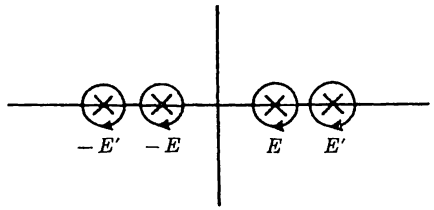

FIGURE 8-1 Poles contributing to the static electromagnetic kernel.

term gives the contribution of the "superfluid" electrons (i.e., the superfluid component in the two-fluid model), and vanishes in this limit as it did at zero temperature. The second term gives the contribution of the thermally excited quasi-particles (i.e., the "normal" fluid component) and does not vanish as $q \to 0$ since the denominator vanishes in this limit. The essential physical difference between the two terms is that the superfluid term involves creation of two quasi-particles, with the minimum excitation energy being $2\Delta_0$. On the other hand, the normal fluid term involves scattering of quasi-particles *already present* and the excitation energy in this case can be arbitrarily small, as in the normal metal. Therefore $L(\mathbf{k}, 0)$ becomes

$$\lim_{\mathbf{q}\to 0} L(\mathbf{k}, \mathbf{q}) = -\frac{\partial f(E_k)}{\partial E_k} = \frac{\beta e^{\beta E_k}}{(e^{\beta E_k} + 1)^2} = \beta f_k(1 - f_k) \quad (8\text{-}52)$$

and from (8-40a) we find

$$\lim_{\mathbf{q}\to 0} P_{ij}(\mathbf{q}, 0) = -2\left(\frac{e}{m}\right)^2 \int \frac{d^3k}{(2\pi)^3} k_i k_j \beta f_k(1 - f_k) \quad (8\text{-}53)$$

It is convenient to define the effective density $\rho_s(T)$ of superfluid electrons at temperature T by

$$\frac{\rho_s(T)}{\rho_s(0)} = 1 - \frac{2\beta E_F}{{k_F}^5} \int_0^\infty k^4 \frac{e^{\beta E_k}\, dk}{(e^{\beta E_k} + 1)^2} \quad (8\text{-}54a)$$

where the Fermi energy is given by $E_F = {k_F}^2/2m$. The relation

$$\rho_s(0) = n = \frac{{k_F}^3}{3\pi^2} \quad (8\text{-}54b)$$

states that all the valence electrons act as superfluid electrons at $T = 0$. On combining (8-53) and (8-54) we find the simple form

$$\lim_{\mathbf{q}\to 0} P_{ij}(\mathbf{q}, 0) = -\frac{ne^2}{m}\left[1 - \frac{\rho_s(T)}{\rho_s(0)}\right]\delta_{ij} \quad (8\text{-}55)$$

By using this result in the expression (8-24) for K_{ij}, one finds within the pairing approximation that

$$\lim_{\mathbf{q}\to 0} K_{ij}(\mathbf{q}, 0) = \frac{1}{\lambda_L^2(0)}\left[\frac{\rho_s(T)}{\rho_s(0)}\right]\delta_{ij} \tag{8-56}$$

Thus, as long as $\rho_s(T)$ is nonzero, $K(q^2)$ [defined by (8-29)], is nonzero as $\mathbf{q}^2 \to 0$ and the Meissner effect is obtained. A plot of $\rho_s(T)/\rho_s(0)$ is shown in Figure 8-2. As $T \to T_c$, the density of superfluid electrons vanishes and one goes over to the normal state with its weak Landau diamagnetism.[92]

In summary, we find within the pairing approximation the following phenomena:

1. The Meissner effect is obtained for all $T \leqslant T_c$.

2. The transverse part of the electromagnetic response kernel K goes to the London form in the long wavelength limit (8-56).

3. By recognizing that only the superfluid electrons give a finite contribution to the transverse part of K as $\mathbf{q} \to 0$, we obtain an expression for the density of superfluid electrons $\rho_s(T)$ as a function of T (8-54).

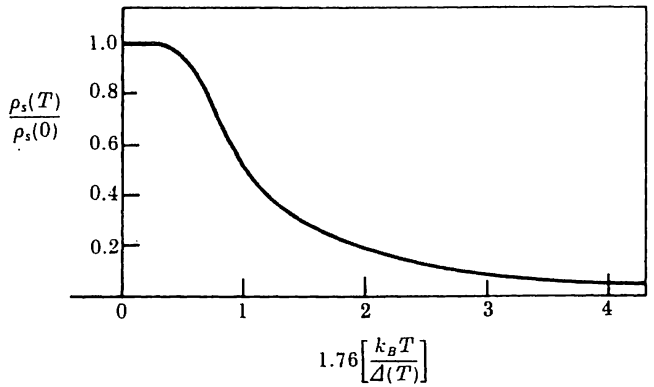

FIGURE 8-2 The superfluid density as a function of temperature. At $T = 0$ all the electrons are in the superfluid, while at $T \geqslant T_c$ all electrons are in the normal fluid.

4. The kernel K is not manifestly gauge invariant within this approximation.

8-4 ELECTROMAGNETIC PROPERTIES FOR FINITE $\mathbf{q}$ AND ω

Although it is gratifying to see the Meissner effect emerge within the pairing approximation, one would like to know the kernel $K_{ij}(\mathbf{q}, \omega, T)$ for general values of $\mathbf{q}$, ω, and T, as well as the effect of impurities on this function. This problem has been worked out by Mattis and Bardeen.[69a] Rather than rederiving their results by the Green's function formalism,[69b,c] we simply state their conclusions. For many purposes it is more convenient to express the kernel in coordinate space rather than in $\mathbf{q}$-space. If one works in the transverse gauge

$$\nabla \cdot \mathbf{A}(\mathbf{r}, \omega) = 0 \tag{8-57}$$

they find

$$\mathbf{J}(\mathbf{r}, \omega) = -\alpha \int d^3r' \frac{\mathbf{R}[\mathbf{R} \cdot \mathbf{A}(\mathbf{r}')]}{R^4} I(\omega, R, T) e^{-R/l} \tag{8-58}$$

Here $\mathbf{R} \equiv \mathbf{r} - \mathbf{r}'$ and the constant α is given by

$$\alpha = \frac{e^2 N(0) v_F}{2\pi^2 \hbar c} \tag{8-59}$$

v_F being the Fermi velocity (we do not set $\hbar = 1$ in this section). The form (8-58) is arranged to resemble the forms given by Pippard[33] and by Chambers[34] for the Meissner effect and the anomalous skin effect, respectively. The factor $e^{-R/l}$ accounts for impurity scattering effects, where l is the electronic mean free path in the normal state. The all important function $I(\omega, R, T)$ is given by

$$I(\omega, R, T) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{L(\omega, \epsilon, \epsilon') - \frac{[f(\epsilon) - f(\epsilon')]}{\epsilon' - \epsilon}\right\} \times \cos\left[\frac{R(\epsilon - \epsilon')}{\hbar v_F}\right] d\epsilon\, d\epsilon' \tag{8-60}$$

where as usual f is the Fermi function. The function $L(\omega, \epsilon, \epsilon')$ is the generalization of the function $L(\mathbf{k}, \mathbf{q})$ of the previous section, and is given by

$$\begin{aligned} L(\omega, \epsilon, \epsilon') = \frac{1}{2} p^2(\epsilon, \epsilon') \left[\frac{1}{E + E' + \hbar\omega - i\delta} + \frac{1}{E + E' - \hbar\omega + i\delta} \right] \\ \times [1 - f(E) - f(E')] \\ + \frac{1}{2} l^2(\epsilon, \epsilon') \left[\frac{1}{E - E' + \hbar\omega - i\delta} \right. \\ \left. + \frac{1}{E - E' - \hbar\omega + i\delta} \right] \\ \times [f(E') - f(E)] \end{aligned} \tag{8-61}$$

The coherence factors are defined by

$$\begin{aligned} p^2(\epsilon, \epsilon') &= \frac{1}{2} \left(1 - \frac{\epsilon\epsilon' + \Delta\Delta'}{EE'} \right) \\ l^2(\epsilon, \epsilon') &= \frac{1}{2} \left(1 + \frac{\epsilon\epsilon' + \Delta\Delta'}{EE'} \right) \end{aligned} \tag{8-62}$$

and $E = (\epsilon^2 + \Delta^2)^{1/2}$. In several limiting cases $I(\omega, R, T)$ takes a simple form:

1. $\hbar\omega \gg \Delta_0$: In this limit, which includes the normal metal as a special case, I becomes

$$I(\omega, R, T) = i\pi\hbar\omega e^{iR\omega/v_F} \tag{8-63}$$

and (8-58) reduces to Chambers' expression for the anomalous skin effect. This allows the coefficient α to be evaluated in terms of the surface impedance of the normal metal in the extreme anomalous limit.

2. $\omega = 0$: In this low-frequency limit (8-58) reduces to a form closely related to Pippard's equation. It is conventional to introduce the function $J(R, T)$ (not to be confused with the current density $\mathbf{J}$) by the relation

$$I(0, R, T) = \left[\frac{\rho_s(T)}{\rho_s(0)} \right] \frac{\pi\hbar v_F}{\xi_0} J(R, T) \tag{8-64}$$

where Pippard's coherence length ξ_0 is defined in terms of microscopic parameters by

$$\xi_0 = \frac{\hbar v_F}{\pi \Delta_0} \tag{8-65}$$

In this limit (8-58) becomes

$$\mathbf{J}(\mathbf{r}) = \frac{-3}{4\pi c \Lambda(T) \xi_0} \int \frac{\mathbf{R}[\mathbf{R} \cdot \mathbf{A}(\mathbf{r}')]}{R^4} J(R, T) e^{-R/l} \, d^3r' \tag{8-66}$$

where $\Lambda(T) \equiv m/\rho_s(T)e^2$ is London's parameter.[1] This expression agrees with Pippard's equation except for the factor e^{-R/ξ_0} being replaced by $J(R, T)$ in (8-66). The definition (8-64) ensures that the $J(R, T)$ and e^{-R/ξ_0} have the same integral

$$\int_0^\infty J(R, T) \, dR = \xi_0 = \int_0^\infty e^{-R/\xi_0} \, dR \tag{8-67}$$

for all $T \leqslant T_c$. One finds that not only are the integrals of the two functions the same, but also the functions themselves resemble each other over the entire range of R and T. For example $J(R, 0)$ is within 5 per cent of e^{-R/ξ_0} for all R and $J(0, 0) = 1$, $J(0, T_c) = 1.33$.

3. $q\xi_0 \ll 1$, $\omega = 0$: In this long wavelength, zero-frequency limit we have already seen that one has the London expression

$$\mathbf{J}_s(\mathbf{r}) = -\frac{\rho_s(T)e^2}{mc} \mathbf{A}(\mathbf{r}) = -\frac{1}{c\Lambda(T)} \mathbf{A}(\mathbf{r}) \tag{8-68}$$

for the pure superconductor, in agreement with (8-66) as can be seen by taking $\mathbf{A}$ outside of the integral in this limit. For a short mean free path $l \ll \xi_0$, one obtains an extra factor $J(0, T)l/\xi_0 \simeq l/\xi_0$, which shows that the London penetration depth increases with impurity concentration.

4. $R/\xi_0 \ll 1$ ($q\xi_0 \gg 1$): If the field is well localized on space, for example, by a skin depth $\lambda \ll \xi_0$ or by the geometry of a thin film where $d \ll \xi_0$, one can evaluate $I(\omega, R, T)$ at $R = 0$ and take it outside of the integral in (8-58). Since the remaining integral is the same as in the normal state of the metal in this limit (i.e., the factor $e^{i\omega R/v_F} \simeq 1$), we can normalize the current

to that in the normal state and express the ratio in terms of the complex surface conductivities σ in the two states. Thus,

$$\frac{\sigma_s}{\sigma_n} = \frac{\sigma_1 + i\sigma_2}{\sigma_n} = \frac{I(\omega, 0, T)}{i\pi\hbar\omega} \tag{8-69}$$

The expression for σ_1/σ_n was given in Chapter 3 while the expression for σ_2 is

$$\frac{\sigma_2}{\sigma_n} = \frac{1}{\hbar\omega}\int_{\Delta-\hbar\omega,\,-\Delta}^{\Delta} \frac{[1 - 2f(E+\hbar\omega)][E^2 + \hbar\omega E + \Delta^2]\,dE}{\{[\Delta^2 - E^2][(E+\hbar\omega)^2 - \Delta^2]\}^{1/2}} \tag{8-70}$$

the lower limit being the larger of the two quantities $\Delta - \hbar\omega$ and $-\Delta$. At zero temperature, the ratio σ_2/σ_n is

$$\frac{\sigma_2}{\sigma_n} = \frac{1}{2}\left(1 + \frac{2\Delta}{\hbar\omega}\right)E(k') - \frac{1}{2}\left(1 - \frac{2\Delta}{\hbar\omega}\right)K(k') \tag{8-71}$$

while the absorptive part is given by

$$\frac{\sigma_1}{\sigma_n} = \left(1 + \frac{2\Delta}{\hbar\omega}\right)E(k) - \frac{4\Delta}{\hbar\omega}K(k) \tag{8-72}$$

In these expressions E and K are the complete elliptic integrals and

$$k' = (1 - k^2)^{1/2} \qquad \text{where} \qquad k = \left|\frac{2\Delta - \hbar\omega}{2\Delta + \hbar\omega}\right| \tag{8-73}$$

The functions σ_1 and σ_2 have been calculated by Tinkham[129] for $T = 0$. For $T \neq 0$ numerical calculations are necessary to determine the surface conductivity; however, a simple low-frequency limit is

$$\frac{\sigma_2}{\sigma_n} = \frac{\pi\Delta}{\hbar\omega}\tanh\frac{\Delta}{2k_BT} \tag{8-74}$$

Calculations for a wide range of frequency and temperature have been carried out by Miller.[130a]

In general one finds remarkably good quantitative agreement between these predictions of the pairing theory and experiment.

As mentioned in Ch. 3, a precursor absorption had been observed below the gap edge $\omega/2\Delta \sim 0.85$. While $\ell \neq 0$ collective modes give absorption in this region, the absorption is too weak to account for these experimental results, which were later found to be spurious.

8-5 GAUGE INVARIANCE

While the simple pairing approximation gives an accurate account of the response of the system to transverse electromagnetic fields, it does not in general give the correct response to longitudinal fields. In particular, we saw in a previous section that it predicts unphysical longitudinal currents which depend on the choice of gauge of the electromagnetic potentials. The physical origin of this difficulty was first recognized by Bardeen,[131] who pointed out that a (longitudinal) gauge potential couples primarily to the collective density fluctuation mode of the electron system (i.e., the plasmons of the charged electron gas). He argued that if one generalizes the pairing scheme to include this mode in a consistent way, a gauge-invariant theory would be obtained. While a number of authors have contributed to the detailed resolution of this problem, the pioneering work of Anderson[47] followed by that of Rickayzen[132] gave the essentials of a generalized pairing scheme which includes these effects. In essence, their approach is to extend the random phase approximation to include pairing correlations.

It is well known that a gauge-invariant response is a consequence of local charge conservation in the system. By local charge conservation we mean that the electronic current and charge density operators satisfy the continuity equation at each point in space and time,

$$\nabla \cdot \mathbf{j}(\mathbf{r}, t) + \frac{\partial \rho_e(\mathbf{r}, t)}{\partial t} = 0 \tag{8-75}$$

In Fourier transform variables this becomes

$$\mathbf{q} \cdot \mathbf{j}(\mathbf{q}, q_0) - q_0\rho_e(\mathbf{q}, q_0) = 0 \tag{8-76a}$$

With the definitions (8-5) and (8-6) plus the metric (1, 1, 1, −1) used previously, the continuity equation for the four-current becomes

$$\sum_{\mu=0}^{3} q_\mu j_\mu(q) = 0 \tag{8-76b}$$

where $q = (\mathbf{q}, q_0)$ as usual. From these relations, it follows that the expected current

$$J_\mu(\mathbf{r}, t) = \langle j_\mu(\mathbf{r}, t) \rangle \tag{8-77}$$

satisfies the continuity equation

$$\nabla \cdot \mathbf{J}(\mathbf{r}, t) + \frac{\partial J_0(\mathbf{r}, t)}{\partial t} = 0 \tag{8-78a}$$

or in Fourier transform space,

$$\sum_{\mu=0}^{3} q_\mu J_\mu(q) = 0 \tag{8-78b}$$

If we concentrate on the linear response of the system to the potential A_μ, we define [see (8-25)]

$$J_\mu(q) = -\frac{c}{4\pi} \sum_{\nu=0}^{3} K_{\mu\nu}(q) A_\nu(q) \tag{8-79}$$

It follows from the continuity equation (8-78b) that the response kernel $K_{\mu\nu}(q)$ must satisfy the equation

$$\sum_{\mu=0}^{3} q_\mu K_{\mu\nu}(q) = 0 \tag{8-80}$$

[For $q_0 = 0$, this condition reduces to the condition (8-28b), used in discussing the static Meissner effect.]

Turning now to the restrictions imposed on K by gauge invariance, we note that the most general gauge transformation is of the form

$$\begin{aligned} \mathbf{A}(\mathbf{r}, t) &\Rightarrow \mathbf{A}(\mathbf{r}, t) + \nabla \Lambda(\mathbf{r}, t) \\ c\varphi(\mathbf{r}, t) &\Rightarrow c\varphi(\mathbf{r}, t) - \frac{\partial \Lambda(\mathbf{r}, t)}{\partial t} \end{aligned} \tag{8-81a}$$

The observable fields

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) &= -\frac{1}{c} \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} - \nabla \varphi(\mathbf{r}, t) \\ \mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}(\mathbf{r}, t) \end{aligned} \tag{8-81b}$$

are invariant under this transformation. In Fourier-transform space, the gauge transformation becomes

$$A_\mu(q) \Rightarrow A_\mu(q) + i q_\mu \Lambda(q) \tag{8-81c}$$

If the observable current is to be unaffected by the gauge transformation, we must require that K satisfies

$$\sum_{\nu=0}^{3} K_{\mu\nu}(q) q_\nu = 0 \tag{8-82}$$

To show the equivalence between the restrictions of local charge conservation (8-80) and of gauge invariance (8-82), we note that $K_{\mu\nu}(q)$ satisfies the symmetry relations

$$\operatorname{Re} K_{\mu\nu}(q) = \operatorname{Re} K_{\nu\mu}(-q) \tag{8-83a}$$

$$\operatorname{Im} K_{\mu\nu}(q) = -\operatorname{Im} K_{\nu\mu}(-q) \tag{8-83b}$$

as a consequence of the definition (8-26) and the spectral representation (8-18) for the retarded commutator $R_{\mu\nu}$. Therefore by changing dummy indices, the real part of the charge-conservation restriction (8-80) can be written as

$$\sum_\nu q_\nu \operatorname{Re} K_{\nu\mu}(q) = \sum_\nu \operatorname{Re} K_{\mu\nu}(-q) q_\nu = 0 \tag{8-84a}$$

or sending $q_\mu \to -q_\mu$

$$\sum_\nu \operatorname{Re} K_{\mu\nu}(q) q_\nu = 0 \tag{8-84b}$$

which is the real part of the gauge-invariance restriction (8-82). In a similar way one finds for imaginary part

$$\sum_\nu q_\nu \operatorname{Im} K_{\nu\mu}(q) = -\sum_\nu \operatorname{Im} K_{\mu\nu}(-q) q_\nu = 0 \tag{8-84c}$$

or

$$\sum_\nu \operatorname{Im} K_{\mu\nu}(q) q_\nu = 0 \tag{8-84d}$$

which agrees with the imaginary part of the gauge-invariance restriction (8-82). Therefore, in an exact calculation, gauge invariance would follow as a consequence of local charge conservation.

Unfortunately, approximate calculations of $K_{\mu\nu}$ may not maintain local charge conservation and therefore not lead to gauge-invariant results. This difficulty, however, is *not* peculiar to the superconducting state. It is a commonly held view that the gauge-invariance problem of the simple pairing approximation is due to the use of wave functions which do not describe a system with a fixed number of particles. That this is not the source of difficulty is seen by realizing that the matrix elements entering the kernel K involve the operator j_μ, which only connects states $|\alpha, N\rangle$ with the *same* number of particles. Therefore, if the states $|\alpha\rangle$ used in calculating these matrix elements are an ensemble average of states $|\alpha, N\rangle$, each with a fixed number of particles, as in the BCS approach, one simply obtains an ensemble average of matrix elements, each of which is evaluated between two states with the same number of particles. Since these fixed N matrix elements are slowly varying functions of N, the ensemble average does not affect the over-all result.

The actual source of error is that the quasi-particle excitations are not treated accurately enough in the simple pairing scheme to ensure local charge conservation under all conditions. Exactly the same situation exists in the normal state if one does not work within a "conserving" approximation, as Baym and Kadanoff[133] put it, even if one uses states which explicitly describe N-particles. In treating the motion of an electron in the medium, one must include the "backflow" of other electrons around the electron in question,[132, 134] as Feynman and Cohen[135] stressed in their work on excitations in superfluid helium. This backflow has a dipolar form at large distance. As one can show, the backflow cancels itself out if the quasi-particles are excited by a transverse field.[9] Therefore, in calculating the response of the system to transverse fields, the backflow currents play no role and one can obtain correct results by an approximation which neglects these complicating effects, as we did in the previous sections.

On the other hand, for the longitudinal response of the system, the backflow around a quasi-particle has a coupling to the external potential which in the long wavelength limit is equal in magnitude but opposite in sign to that given by the bare quasi-particle. Thus the dressed quasi-particle (i.e., the bare quasi-particle plus its associated backflow cloud) is very weakly coupled to a slowly varying longitudinal potential, as Pines and the author have discussed.[134] There is however an extra longitudinal mode of the system which occurs once backflow is properly taken into account and this is the collective density fluctuation mode.[47, 52] Physically, one can think of the collective mode simply as a compressional wave in the superfluid. From this physical picture it is reasonable that the current and particle densities associated with this mode will satisfy the continuity equation. Since in the long wavelength limit only the density fluctuation mode is coupled to a longitudinal potential, it is reasonable that a gauge-invariant response will be obtained once these effects are included.

Kadanoff and Ambegaokar have shown that in the long wavelength limit the collective mode can be described as a state in which the phase of the energy-gap parameter varies periodically in space and time, while the magnitude of the gap parameter remains fixed. Since the phase of the gap parameter gives the mean local center-of-mass momentum of the superfluid pairs, a periodically varying phase is exactly what one would expect if the superfluid momentum density varies periodically.

There are now a number of formalisms for including the backflow and the collective mode. One of the simplest ways to handle the problem is to make use of a "generalized Ward's identity," which is the Green's function analog of the continuity equation. By making approximations which are consistent with this identity one can ensure local charge conservation and therefore gauge invariance. This approach was first discussed by Nambu,[114] and we follow his line of argument below.

We consider the time-ordered quantity $\Lambda_\mu(x, y, z)$ defined in terms of the Nambu field Ψ by

$$\Lambda_\mu(x, y, z) = \langle 0|T\{j_\mu(z)\Psi(x)\Psi^+(y)\}|0\rangle \qquad (8\text{-}85)$$

The four-current density j_μ is defined by (8-6). It is clear that the paramagnetic kernel $R_{\mu\nu}$ can be calculated from Λ_μ by taking the appropriate gradients and traces of (8-85), as we shall see below. We define the vertex function $\mathbf{\Gamma}_\mu(x', y', z)$ by the integral relation

$$\mathbf{\Lambda}(x, y, z) = e \int \mathbf{G}(x, x')\mathbf{\Gamma}_\mu(x', y', z)\mathbf{G}(y', y)\, d^4x'\, d^4y' \tag{8-86}$$

where

$$\mathbf{G}(x, x') = -i\langle 0|T\{\Psi(x)\Psi^+(x')\}|0\rangle \tag{8-87}$$

is Nambu's one-particle Green's function. We assume the system to be translationally invariant so that we can write

$$\mathbf{\Gamma}_\mu(x', y', z) = \int \mathbf{\Gamma}_\mu(p + q, p)e^{i[p(x'-y')+q(x'-z)]}\frac{d^4p\, d^4q}{(2\pi)^8} \tag{8-88}$$

The generalized Ward's identity for the superconductor is then

$$\begin{aligned}\sum_\mu q_\mu\mathbf{\Gamma}_\mu(p + q, p) &= \sum_{i=1}^{3} q_i\mathbf{\Gamma}_i(p + q, p) - q_0\mathbf{\Gamma}_0(p + q, p) \\ &= \tau_3\mathbf{G}^{-1}(p) - \mathbf{G}^{-1}(p + q)\tau_3\end{aligned} \tag{8-89}$$

To prove this identity, we take the four-divergence of Λ_μ with respect to $z \equiv (z, z_0 = t_z)$:

$$\begin{aligned}\sum_{i=1}^{3}\frac{\partial\mathbf{\Lambda}_i}{\partial z_i} + \frac{\partial\mathbf{\Lambda}_0}{\partial z_0} &= \langle 0|T\left\{\left[\sum_{i=1}^{3}\frac{\partial j_i(z)}{\partial z_i} + \frac{\partial j_0(z)}{\partial z_0}\right]\Psi(x)\Psi^+(y)\right\}|0\rangle \\ &+ \langle 0|T\{[j_0(z), \Psi(x)]\Psi^+(y)\}|0\rangle\,\delta(z_0 - x_0) \\ &+ \langle 0|T\{\Psi(x)[j_0(z), \Psi^+(y)]\}|0\rangle\,\delta(z_0 - y_0)\end{aligned} \tag{8-90}$$

The last two terms on the right-hand side arise from differentiating the time dependence due to the time-ordering symbol T. Now the first term on the right-hand side of this expression vanishes by virtue of the continuity equation (8-75). If we use the equal time anticommutation relations of the Ψ's (7-21), the commutators in (8-90) can be reduced to

$$[j_0(z), \Psi(x)]\,\delta(z_0 - x_0) = e\tau_3\Psi(z)\,\delta^4(z - x) \tag{8-91a}$$

and

$$[j_0(z), \Psi^+(y)]\,\delta(z_0 - y_0) = -e\Psi^+(y)\tau_3\,\delta^4(z - y) \tag{8-91b}$$

By inserting these expressions into (8-90) and using the definition (8-86), and (8-87) we find

$$\begin{aligned}&i\mathbf{G}(x - z)\tau_3\,\delta^4(z - y) - i\tau_3\mathbf{G}(z - y)\,\delta^4(z - x) \\ &\quad = -\int \mathbf{G}(x - x')\left[\sum_{i=1}^{3}\frac{\partial\mathbf{\Gamma}_i}{\partial z_i} + \frac{\partial\mathbf{\Gamma}_0}{\partial z_0}\right]\mathbf{G}(y' - y)\, d^4x'\, d^4y'\end{aligned} \tag{8-92}$$

Going over to Fourier transform variables we find (8-92) reduces to

$$\mathbf{G}(p + q)\tau_3 - \tau_3\mathbf{G}(p) = \mathbf{G}(p + q)\sum_{\mu=0}^{3} q_\mu\mathbf{\Gamma}_\mu(p + q, p)\mathbf{G}(p) \tag{8-93a}$$

or finally by operating with $\mathbf{G}^{-1}(p + q)$ on the left and with $\mathbf{G}^{-1}(p)$ on the right we obtain the generalized Ward's identity

$$\sum_{\mu=0}^{3} q_\mu\mathbf{\Gamma}_\mu(p + q, p) = \tau_3\mathbf{G}^{-1}(p) - \mathbf{G}^{-1}(p + q)\tau_3 \tag{8-93b}$$

as stated.

What is the physical significance of $\mathbf{\Gamma}_\mu$ and why is this identity of interest? The significance of $\mathbf{\Gamma}_\mu$ can be best understood by noticing that the four-current density operator $j_\mu{}^p(\mathbf{q})$ can be written in the Nambu notation as

$$j_\mu{}^p(\mathbf{q}) = \begin{cases}\displaystyle\sum_{\mathbf{p}} \Psi_{\mathbf{p}}{}^+\left[-e\left(\mathbf{p} + \frac{\mathbf{q}}{2}\right)_i\mathbf{1}\right]\Psi_{\mathbf{p}+\mathbf{q}} & (\mu = i = 1, 2, 3) \\ \displaystyle\sum_{\mathbf{p}} \Psi_{\mathbf{p}}{}^+\left[-e\tau_3\right]\Psi_{\mathbf{p}+\mathbf{q}} & (\mu = 0)\end{cases} \tag{8-94}$$

If we define the "free" vertex function $\gamma_\mu(\mathbf{p} + \mathbf{q}, \mathbf{p})$ as

$$\gamma_\mu(\mathbf{p} + \mathbf{q}, \mathbf{p}) = \begin{cases}\dfrac{1}{m}\left(\mathbf{p} + \dfrac{\mathbf{q}}{2}\right)_i\mathbf{1} & (\mu = i = 1, 2, 3) \\ \tau_3 & (\mu = 0)\end{cases} \tag{8-95}$$

then $j_\mu{}^p(\mathbf{q})$ can be written as

$$j_\mu{}^p(\mathbf{q}) = -e\sum_{\mathbf{p}} \Psi_{\mathbf{p}}{}^+\gamma_\mu(\mathbf{p} + \mathbf{q}, \mathbf{p})\Psi_{\mathbf{p}+\mathbf{q}} \tag{8-96}$$

The vertex function $\mathbf{\Gamma}_\mu(p+q,p)$ is then a "*dressed*" version of the *free* vertex $\gamma_\mu(p+q,p)$. To make this plausible, we apply the generalized Ward's identity to a system of noninteracting electrons. Since $G^{-1}(p)$ becomes

$$\mathbf{G}_0^{-1}(p) = p_0\mathbf{1} - \epsilon_p\tau_3 \tag{8-97}$$

in this limit, where $\epsilon_p = (p^2/2m) - \mu$, (8-93b) reduces to

$$\sum_\mu q_\mu\mathbf{\Gamma}_\mu(p+q,p) = (\epsilon_{p+q} - \epsilon_p)\mathbf{1} - q_0\tau_3 \tag{8-98}$$

However, this relation is identically satisfied if $\mathbf{\Gamma}_\mu$ is the free vertex γ_μ. Therefore, we can think of the dressed electrons as interacting with the electromagnetic field through the dressed vertex $(-e\mathbf{\Gamma}_\mu)$.

As to why the generalized Ward's identity is of interest, we shall now see that the paramagnetic kernel can be simply expressed in terms of $\mathbf{G}$ and $\mathbf{\Gamma}_\mu$. Furthermore, if we approximate $\mathbf{G}$ and $\mathbf{\Gamma}$ in a way which maintains Ward's identity, the full electromagnetic kernel $K_{\mu\nu}$ will be manifestly gauge-invariant. From the definitions of the time-ordered kernel $P_{\mu\nu}$ (8-21) and the vertex function $\mathbf{\Gamma}_\mu$ (8-85) and (8-86), it is straightforward to show that $P_{\mu\nu}(q)$ is given by

$$P_{\mu\nu}(q) = -ie^2\int \mathrm{Tr}\,[\gamma_\mu(p,\,p+q)\mathbf{G}(p+q)\mathbf{\Gamma}_\nu(p+q,p)\mathbf{G}(p)]\frac{d^4p}{(2\pi)^4} \tag{8-99}$$

If we consider only the components of $P_{\mu\nu}$ with μ and $\nu \neq 0$, and approximate the vertex function $\mathbf{\Gamma}_\nu$ by the bare vertex γ_ν, we retrieve the expression (8-38) given by the pairing approximation. The lack of gauge invariance within this approximation is a consequence of calculating $P_{\mu\nu}$ with dressed G's but bare vertex functions, thereby violating the generalized Ward's identity.

It is convenient to represent the relation (8-99) in graphical form. In Figure 8-3 we show $P_{\mu\nu}(q)$ represented in terms of the dressed electron lines and the vertex parts. While $P_{\mu\nu}$ appears

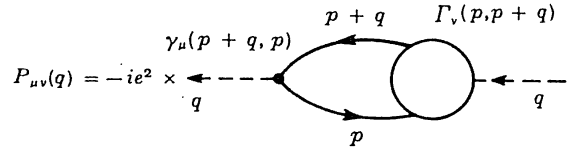

FIGURE 8-3 The polarizability kernel $P_{\mu\nu}$ represented in terms of the bare and dressed vertices γ_μ and Γ_ν, respectively.

to be unsymmetrical in μ and ν, the expression (8-99) is equally valid if $\gamma_\mu \to \mathbf{\Gamma}_\mu$ and $\mathbf{\Gamma}_\nu \to \gamma_\nu$.

We now prove that $K_{\mu\nu}$ is gauge-invariant if the generalized Ward's identity is satisfied. This follows by writing

$$\begin{aligned}\sum_\nu P_{\mu\nu}(q)q_\nu &= -ie^2\int \mathrm{Tr}\left[\gamma_\mu(p,\,p+q)\mathbf{G}(p+q)\sum_\nu q_\nu\mathbf{\Gamma}_\nu(p+q,p)\right.\\ &\qquad\left.\times\,\mathbf{G}(p)\right]\frac{d^4p}{(2\pi)^4}\\ &= ie^2\int \mathrm{Tr}\,[\gamma_\mu(p,\,p+q)\mathbf{G}(p+q)\{\mathbf{G}^{-1}(p+q)\tau_3\\ &\qquad - \tau_3\mathbf{G}^{-1}(p)\}\mathbf{G}(p)]\frac{d^4p}{(2\pi)^4}\\ &= ie^2\int \mathrm{Tr}\,[\gamma_\mu(p,\,p+q)\{\tau_3\mathbf{G}(p) - \mathbf{G}(p+q)\tau_3\}]\\ &\qquad\times\frac{d^4p}{(2\pi)^4}\end{aligned} \tag{8-100}$$

In the second equality we have used Ward's identity (8-93b). Since τ_3 commutes with γ_μ, the last equality in (8-100) reduces to

$$\sum_\mu P_{\mu\nu}(q)q_\nu = ie^2\int \mathrm{Tr}\,[\{\gamma_\mu(p+q,p) - \gamma_\mu(p,p-q)\}\times\tau_3\mathbf{G}(p)]\frac{d^4p}{(2\pi)^4} \tag{8-101}$$

when we use cyclic invariance of the trace. From the definition of γ_μ (8-95) we see that

$$\gamma_\mu(p + q, p) - \gamma_\mu(p, p - q) = \frac{q_\mu}{m}(1 - \delta_{\mu,0}) \tag{8-102}$$

Therefore, we obtain the simple expression

$$\sum_\nu P_{\mu\nu}(q)q_\nu = -\frac{ne^2}{m} q_\mu(1 - \delta_{\mu,0}) \tag{8-103}$$

where we have used the relation (7-35) giving the number of electrons per unit volume in terms of $\mathbf{G}$. Since the right-hand side of this equation is real, (8-24) allows us to replace $P_{\mu\nu}$ in (8-103) by the physically relevant kernel $R_{\mu\nu}$ and we finally obtain

$$\sum_\nu R_{\mu\nu}(q)q_\nu = -\frac{ne^2}{m} q_\mu[1 - \delta_{\mu,0}] \tag{8-104}$$

To check that $K_{\mu\nu}$ is manifestly gauge-invariant due to this result, we use (8-26) to write the gauge-invariance condition as

$$\sum_\nu K_{\mu\nu}(q)q_\nu = 0 = \frac{4\pi}{c^2}\sum_\nu R_{\mu\nu}(q)q_\nu + \frac{1}{\lambda_L^2} q_\mu[1 - \delta_{\mu,0}] \tag{8-105}$$

Since $1/\lambda_L^2 = 4\pi ne^2/mc^2$, we see from (8-104) that the gauge-invariance condition is satisfied identically,

$$\sum_\nu K_{\mu\nu}(q)q_\nu = \left[-\frac{4\pi ne^2}{mc^2} + \frac{1}{\lambda_L^2}\right] q_\mu[1 - \delta_{\mu,0}] = 0 \tag{8-106}$$

In the next section we shall discuss the generalization of the pairing scheme which is required in order that the generalized Ward's identity is to be satisfied.

8-6 THE VERTEX FUNCTION AND COLLECTIVE MODES

In seeking a gauge-invariant generalization of the pairing scheme, Nambu used a well-known prescription of quantum-field theory for constructing approximations which satisfy the generalized Ward's identity (GWI) (8-89).[137] If $\mathbf{G}$ is described

FIGURE 8-4 The pairing approximation sums all no-line-crossing graphs contributing to Σ.

by a certain set of perturbation series graphs, the corresponding vertex function $\boldsymbol{\Gamma}_\mu$ (which satisfies the GWI) is given by the sum of all graphs in which the free vertex γ_μ is inserted in each bare electron line in this set. The pairing approximation for $\mathbf{G}$,

$$\mathbf{G}^{-1}(p) = p_0\mathbf{1} - \epsilon_p\tau_3 - \boldsymbol{\Sigma}(p) \tag{8-107}$$

$$\boldsymbol{\Sigma}(p) = i\int \tau_3\mathbf{G}(p')\tau_3\gamma\!\!\!/(p - p')\frac{d^4p'}{(2\pi)^4} \tag{8-108}$$

can be formally thought of as the sum of all graphs in which no two interaction lines cross, as shown in Figure 8-4. If the vertex γ_μ is inserted at all possible places in these graphs, the resultant series is summed by the ladder graph approximation for $\boldsymbol{\Gamma}_\mu$, as shown in Figure 8-5. Therefore $\boldsymbol{\Gamma}_\mu$ satisfies the linear integral equation

$$\boldsymbol{\Gamma}_\mu(p + q, p) = \gamma_\mu(p + q, p) + i\int \tau_3\mathbf{G}(k + q)\boldsymbol{\Gamma}_\mu(k + q, k)\mathbf{G}(k)\tau_3\gamma\!\!\!/(p - k)\frac{d^4k}{(2\pi)^4} \tag{8-109}$$

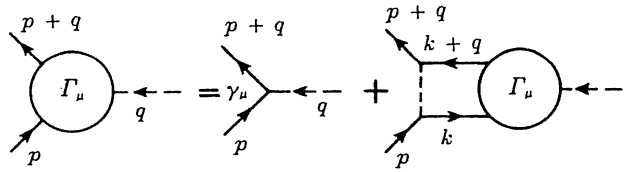

FIGURE 8-5 An equation for the vertex function $\boldsymbol{\Gamma}_\mu$ which leads to manifestly gauge-invariant results for the electromagnetic kernel within the pairing scheme.

where the **G**'s are the Nambu functions evaluated within the pairing approximation (8-108). To check that the solution of this vertex equation is in fact consistent with the GWI, we form the quantity

$$\sum_\mu q_\mu \mathbf{\Gamma}_\mu(p+q,p) = \sum_\mu q_\mu \gamma_\mu(p+q,p) + i\int \tau_3 \mathbf{G}(k+q) \sum_\mu q_\mu \mathbf{\Gamma}_\mu(k+q,k) \times \mathbf{G}(k)\tau_3 \mathscr{V}(p-k)\frac{d^4k}{(2\pi)^4} \quad (8\text{-}110)$$

The right-hand side of this equation can be reduced by use of the assumed GWI. By using the relation

$$\sum_\mu q_\mu \mathbf{\Gamma}_\mu(k+q,k) = \tau_3 \mathbf{G}^{-1}(k) - \mathbf{G}^{-1}(k+q)\tau_3 \quad (8\text{-}111)$$

we see that the second term becomes

$$\left[i\int \tau_3 \mathbf{G}(k+q)\tau_3 \mathscr{V}(p-k)\frac{d^4k}{(2\pi)^4}\right]\tau_3 - \tau_3\left[i\int \tau_3 \mathbf{G}(k)\tau_3 \mathscr{V}(p-k)\frac{d^4k}{(2\pi)^4}\right] = \mathbf{\Sigma}(p+q)\tau_3 - \tau_3\mathbf{\Sigma}(p) \quad (8\text{-}112)$$

where we have used the equation determining $\mathbf{\Sigma}(p)$ (8-108). In addition, from (8-95) we see that the free vertex satisfies

$$\sum_\mu q_\mu \gamma_\mu(p+q,p) = (\epsilon_{p+q} - \epsilon_p)\mathbf{1} - q_0\tau_3 \quad (8\text{-}113)$$

On combining these results, (8-93b) reduces to

$$\sum_\mu q_\mu \mathbf{\Gamma}_\mu(p+q,p) = \tau_3[p_0\mathbf{1} - \epsilon_p\tau_3 - \mathbf{\Sigma}(p)] - [(p_0+q_0)\mathbf{1} - \epsilon_{p+q}\tau_3 - \mathbf{\Sigma}(p+q)]\tau_3 = \tau_3\mathbf{G}^{-1}(p) - \mathbf{G}^{-1}(p+q)\tau_3 \quad (8\text{-}114)$$

which is the required GWI. Therefore, if **G** and $\mathbf{\Gamma}_\mu$ are given by solutions of Eqs. (8-107), (8-108), and (8-109), the electromagnetic kernel $K_{\mu\nu}$ determined through (8-99) will be *manifestly* gauge-invariant.

To understand the mechanism by which gauge invariance has been restored in this rather formal scheme, we again look at the

GWI. If we assume that the dressed vertex Γ_μ is a well-behaved function in the limit **q** and q_0 go to zero, the left-hand side of (8-114) vanishes while the right-hand side becomes the finite quantity

$$\tau_3\mathbf{\Sigma}(p) - \mathbf{\Sigma}(p)\tau_3 = 2i\tau_2\phi(p) = 2i\tau_2\Delta_p \quad (8\text{-}115)$$

(the second equality holding for a nonretarded pairing potential). Therefore, $\mathbf{\Gamma}_\mu(p+q,p)$ must be singular for $q=0$. If one thinks of the coupling as going through a set of excited states of the system (i.e., one thinks of Γ_μ as written in a spectral form), one is tempted to argue that the $q=0$ singularity in Γ reflects the existence of a low-lying collective mode whose frequency Ω_q goes to zero in the long-wavelength limit. To check this idea we would like to obtain an explicit solution of the vertex equation (8-109) and see if Γ_μ is actually singular for $q_0 = \Omega_q$ and $\mathbf{q} \neq 0$. As in Chapter 7, this t-matrix-like equation can be solved if $\mathscr{V}(k-p)$ is approximated by a factorizable potential

$$\mathscr{V}(k-p) = \lambda w^*(\mathbf{k})w(\mathbf{p}) \quad (8\text{-}116)$$

In solving (8-109) it is convenient to think of the 2×2 matrices $\mathbf{\Gamma}_\mu$ and γ_μ as being represented by four-component column vectors. Thus we replace the matrix component $\langle l|\mathbf{\Gamma}_\mu|r\rangle$ by the column vector component $(\mathbf{\Gamma}_\mu)_{lr}$ and (8-109) becomes

$$\mathbf{\Gamma}_\mu(p+q,p) = \gamma_\mu(p+q,p) + i\lambda w^*(\mathbf{p})\int \tau_3{}^l\mathbf{G}^l(k+q)\tau_3{}^r\tilde{\mathbf{G}}^r(k) \times \mathbf{\Gamma}_\mu(k+q,k)w(\mathbf{k})\frac{d^4k}{(2\pi)^4} \quad (8\text{-}117)$$

The scripts l and r indicate which part of the double script (lr) the matrices act and $\tilde{\mathbf{G}}$ means the transpose of **G**. Since the last term in this expression is a constant (matrix) multiple of $w^*(\mathbf{p})$, we have

$$\mathbf{\Gamma}_\mu(p+q,p) = \gamma_\mu(p+q,p) + w^*(\mathbf{p})C_q \quad (8\text{-}118)$$

where the constant C_q is defined by

$$C_q = i\lambda\int \tau_3{}^l\mathbf{G}^l(k+q)\tau_3{}^r\tilde{\mathbf{G}}^r(k)\mathbf{\Gamma}_\mu(k+q,k)w(\mathbf{k})\frac{d^4k}{(2\pi)^4} \quad (8\text{-}119)$$

By inserting the relation (8-118) into (8-119) and solving for C_q, one finds the explicit solution

$$\mathbf{\Gamma}_\mu(p + q, p) = \gamma_\mu(p + q, p) + [1 - \lambda\phi(q)]^{-1}\chi_\mu(q)w^*(\mathbf{p}) \quad (8\text{-}120)$$

where $\phi(q)$ is a generalization of the function $\tilde{\Phi}(q)$ [see (5-21)] used in discussing the instability of the normal state. The matrix function $\phi(q)$ is given by

$$\phi(q) = i\int \tau_3{}^l\mathbf{G}^l(k + q)\tau_3{}^r\mathbf{G}^r(k)|w(\mathbf{k})|^2\frac{d^4k}{(2\pi)^4} \quad (8\text{-}121)$$

The matrix $\chi(q)$ is defined by

$$\chi_\mu(q) = i\lambda\int \tau_3{}^l\mathbf{G}^l(k + q)\tau_3{}^r\mathbf{G}^r(k)\gamma_\mu(k + q, k)w(k)\frac{d^4k}{(2\pi)^4} \quad (8\text{-}122)$$

Since χ_μ is regular for $q_0 < 2\Delta_0$, the singularity of $\mathbf{\Gamma}_\mu$ must arise from a singularity of $[1 - \lambda\phi(q)]^{-1}$. For this matrix to be singular, the determinant of its inverse must be zero. Therefore, the dispersion law for the collective mode (or modes) is

$$\det[1 - \lambda\phi(\mathbf{q}, \Omega_\mathbf{q})] = 0 \quad (8\text{-}123)$$

where $\Omega_\mathbf{q}$ is the frequency of the mode in question. If we assume s-state pairing and as in Chapter 7 take $w(\mathbf{k})$ to be

$$w(k) = \begin{cases} 1 & |\epsilon_k| < \omega_c \\ 0 & \text{otherwise} \end{cases} \quad (8\text{-}124)$$

there is a root of (8-123) which for $|\mathbf{q}|\xi_0 \ll 1$, satisfies the dispersion law

$$\Omega_\mathbf{q} = \frac{v_F}{(3)^{1/2}}|\mathbf{q}| \quad (8\text{-}125)$$

This sound-wave mode was first discovered by Bogoliubov.[52] Physically, it corresponds to long wavelength density fluctuations of the electron system as a whole. Since the pairing correlations are not expected to be changed appreciably by slowly varying the electron density in space and time, one might expect such a collective mode on physical grounds. In fact (8-125) follows if one uses the standard hydrodynamic expression for the speed of sound s:

$$s^2 = \frac{dP}{d\rho} \quad (8\text{-}126)$$

where ρ and P are the mass density and pressure of the *free*-electron gas. Therefore within this approximation pairing correlations play no direct role in determining the velocity of this mode, their main function being to remove low-lying single-particle states which would otherwise lead to damping of the wave.

Returning to the solution (8-120) for the vertex function, if the two-body potential is purely s-wave, the transverse part of χ_μ vanishes, as one can easily see on symmetry grounds from (8-122). Therefore vertex corrections *do not affect* the Meissner kernel in this case. If there is a finite attractive d-wave part of the potential, d-state excitons exist and will contribute to the vertex function. Calculations by Rickayzen show that these collective corrections to the Meissner kernel are in general small.

If there is a strongly attractive d-wave potential one should see a precursor for infrared absorption below the gap edge, owing to creation of d-state excitons.[138, 139] Such anomalies were first observed by Ginsberg, Richards, and Tinkham,[71, 72] although as Tsuneto[138] has shown, the predicted absorption is an order of

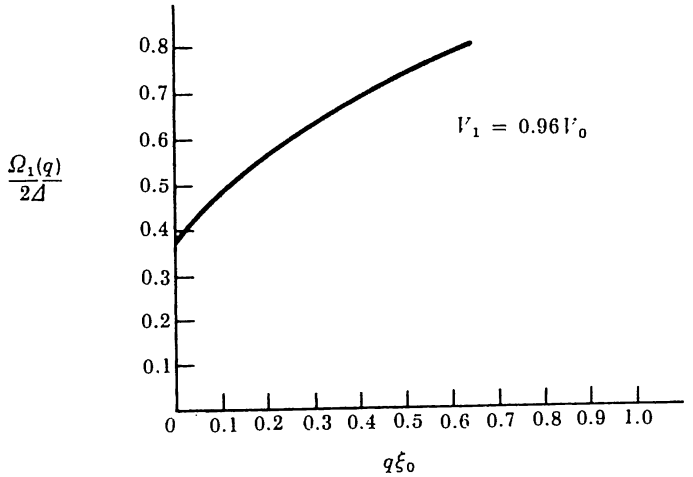

FIGURE 8-6 Energy of a p-state exciton as a function of momentum.

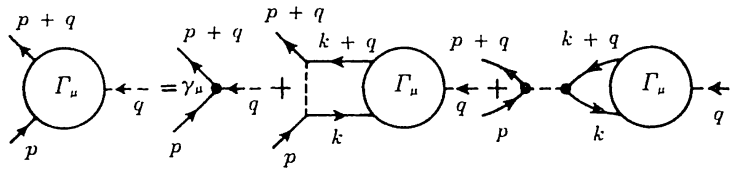

FIGURE 8-7 Vacuum-polarization correction to the equation shown in Figure 8-5. This correction is all important in the longitudinal part of $P_{\mu\nu}$ and leads to plasma oscillations.

magnitude weaker than that observed experimentally. Since the calculations were carried out on a continuum model, momentum was conserved in the absorption process. Owing to the fact that the excitons pass into the single-particle continuum for $q \gtrsim 1/\xi_0$, as shown in Figure 8-6, only the long wavelength components of the penetrating field contribute to exciton creation, and these small q-components are small compared to those for $q \sim 1/\lambda \gg 1/\xi_0$, thereby giving a small absorption. Experimentally, the precursor is essentially unaffected by impurities.[130b] Since the exciton state is destroyed by impurities,[130c] it appears that the precursor is due to another mechanism. [Subsequent experiments have shown the precursor absorption to be spurious].

Thus far we have neglected vacuum-polarization processes in the vertex function $\mathbf{\Gamma}_\mu$. As we saw in Chapter 6, these processes dominate the long wavelength polarizability of the electron gas because of the long-range Coulomb potential and therefore *must* be included in $\mathbf{\Gamma}$. The vertex equation including vacuum polarization processes is illustrated in Figure 8-7. The only change in (8-109) is to add the term

$$-iV_B(q)\int \mathrm{Tr}\,\{\tau_3\mathbf{G}(k+q)\mathbf{\Gamma}_\mu(k+q,k)\mathbf{G}(k)\}\frac{d^4k}{(2\pi)^4} \quad (8\text{-}127)$$

to the right-hand side. Here $V_B(q)$ is the sum of the bare Coulomb and bare longitudinal phonon interactions:

$$V_B(q) = \frac{4\pi e^2}{q^2} + |g_{ql}|^2 D_{0l}(q) \quad (8\text{-}128)$$

To make connection with the RPA for the electron gas, we note that P_{RPA} (4-2) is proportional to P_{00} in (8-99) when we (1) include *only* the polarization term (8-127) in the equation for $\mathbf{\Gamma}_\mu$, (2) replace all G's by G_0's, and (3) set $g_q = 0$. It is easily seen that when (8-127) is included, the solution $\mathbf{\Gamma}_\mu(p+q,p)$ continues to satisfy the GWI and $K_{\mu\nu}$ is still manifestly gauge-invariant. By explicitly solving this improved vertex equation, one can see that the Bogoliubov sound-wave mode continues to exist if $V_B(q)$ approaches a finite value as $\mathbf{q} \to 0$, as Anderson[47] first showed. In the presence of the Coulomb potential, which of course always exists in real metals, the Bogoliubov-Anderson mode is pushed up to high energy and becomes the plasma oscillation of the electron system. Therefore, the $q \equiv 0$ singularity of $\mathbf{\Gamma}_\mu(p+q,p)$ required by (8-115) *does not* imply a low-lying boson mode in physical metals due to the long-range Coulomb interaction between electrons.

8-7 FLUX QUANTIZATION

A qualitatively new effect arises when we investigate the electromagnetic behavior of a multiply connected superconducting system, e.g., a long, hollow cylinder. In this case, flux can be trapped in the hole and persist in the absence of an externally applied field. On the basis of London's "rigidity" concept he concluded that for a cylinder with walls thick compared to the penetration depth λ, flux could be trapped only in multiples of $hc/e = 4 \times 10^{-7}$ gauss cm^2.[1] This value of the flux quantum follows if one assumes that the *only* low-lying current-carrying states of the superfluid are those given by multiplying the superfluid ground state by a single-valued phase factor, as we saw in Chapter 1. We shall see below that there are *two* distinct sets of low-lying states, one being the set considered by London, the other arising from phase factors multiplying a basic state which is not included in London's set. Owing to these two sets of states, the flux quantum in superconductors is actually $hc/2e$, i.e., one-half the London unit. The even multiples are associated with London-type states, while the odd multiples are due to the

other series of states, as Byers and Yang[19] first pointed out. The value $hc/2e$ was observed experimentally by Deaver and Fairbank[20a] and by Doll and Näbauer[20b] prior to Byers and Yang's work.

Perhaps the simplest way of understanding flux quantization is by considering a long, hollow cylinder of inner and outer radii a and b, respectively. Suppose the cylinder is initially in the normal state in the presence of a magnetic field and the temperature is lowered so that the cylinder becomes superconducting. Owing to the Meissner effect, the magnetic field will be expelled from the material and in general there will be a finite magnetic flux Φ trapped in the hole. If we assume the wall thickness $b - a$ is much greater than the penetration depth, the magnetic field in the penetration layer will be a small perturbation on the system as a whole and cannot affect our results. If we use cylindrical coordinates (r, θ, z), the vector potential is given by

$$\oint \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}' = \int \mathbf{B} \cdot d\mathbf{S} = \Phi(r) \tag{8-129}$$

where the line integral is taken around a circle of radius r and $\Phi(r)$ is the flux enclosed by the path. Since $\Phi(r)$ goes to a constant Φ (the total trapped flux) for $r - a \gg \lambda$, we write

$$A_\theta(r) = \frac{\Phi}{2\pi r} + \frac{\Phi(r) - \Phi}{2\pi r} \equiv A_\theta^{(0)}(r) + A_\theta^{(1)}(r) \tag{8-130}$$

and only include $A_\theta^{(0)}(r)$ in zero order, $A^{(1)}$ being treated as a perturbation. We first consider the single-particle states defined in the presence of $A^{(0)}$ and then pair up these states to form the superconducting phase. The azimuthal part of the single-particle eigenfunctions satisfy

$$\frac{\left[p_\theta + \dfrac{eA_\theta^{(0)}}{c}\right]^2}{2m_e} \psi_M(\theta) = \frac{\hbar^2}{2m_e r^2} (M + \varphi)^2 \psi_M(\theta) \tag{8-131a}$$

where

$$\varphi \equiv \frac{e\Phi}{hc} \tag{8-131b}$$

is the flux measured in units of London's flux quantum and

$$\psi_M(\theta) = e^{iM\theta} \tag{8-131c}$$

In order that Ψ_M be single-valued, M must be an integer. The dynamics is simplified if we assume the thickness of the cylinder $b - a$ is small compared to the radius a. The angular kinetic energy is then $\hbar^2(M + \varphi)^2/2m_e a^2$, i.e., a parabolic function of M, centered about $M = -\varphi$. If we are to obtain a low-energy state of the system we must pair single-particle states which are (a) degenerate with each other and (b) coupled to other paired states by the two-body potential. Condition (a) means that the paired states M and $\bar{M}$ must satisfy

$$|M + \varphi| = |\bar{M} + \varphi| \tag{8-132}$$

This condition can be satisfied if $M = \bar{M}$ but conservation of angular momentum forbids the two-body potential from connecting states paired in this manner. The other choice is $M + \varphi = -(\bar{M} + \varphi)$, that is, $M \equiv m - \varphi$ and $\bar{M} = -m - \varphi$ are paired, so that the pairing is symmetric in M-space about the value $M = -\varphi$. Since M and $\bar{M}$ are required to be integers, it follows that m and φ are both integers or both half-odd integers. Therefore, we conclude from this result and (8-131b) that one obtains a large pairing energy and therefore a low-energy state of the system only if the trapped flux is given by

$$\Phi = n\left(\frac{hc}{2e}\right) \tag{8-133}$$

where n is an integer. In Figure 8-8a, b, c, and d we illustrate the pairing for $n = 0$, 1, 2, and 3, respectively. Geometrically, one simply pairs symmetrically about the value $M = -n/2$. Notice that the states being paired for $n = 0$ and $n = 2$ (Figure 8-8a and c) differ only by a shift of all the angular momentum quantum numbers by the fixed amount -1. Thus, the system wave functions for the $n = 0$ and $n = 2$ states are related by the phase factor

$$\psi_2(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N) = e^{-i\sum_j \theta_j} \psi_0(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N) \tag{8-134}$$

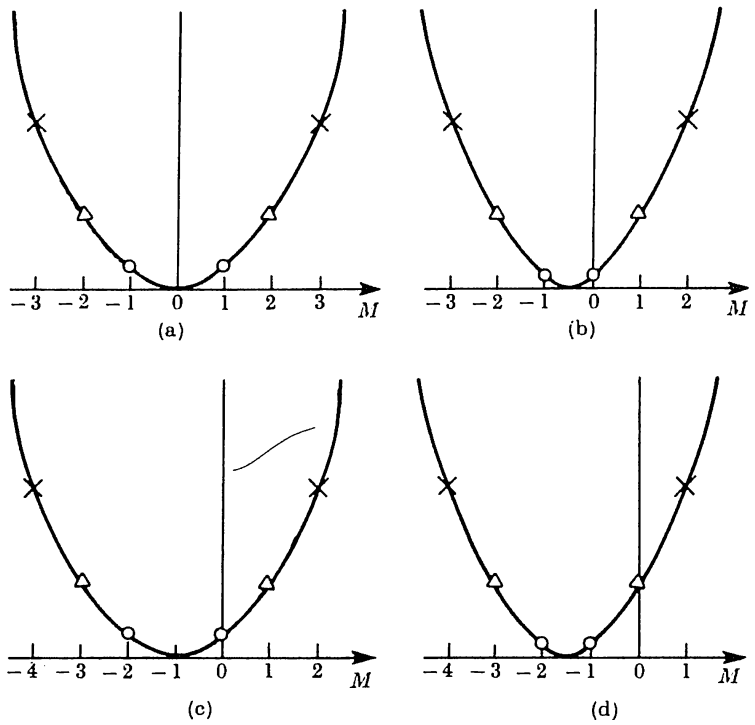

FIGURE 8-8 Pairings of azimuthal quantum numbers for flux quantum numbers $n = 0, 1, 2, 3$ are shown in (a), (b), (c), and (d), respectively [see Eq. (8-133)].

in agreement with London's argument. The states for $n = 1$ and $n = 3$ are related in the same manner:

$$\psi_3(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N) = e^{-i\sum_j \theta_j} \psi_1(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N) \qquad (8\text{-}135)$$

The essential point is that the even n- and odd n-states are *not* related by a phase factor. For example, the $n = 0$ and $n = 1$ states differ by sliding *only* the mates to the right of $M = 0$ left one notch in forming the $n = 1$ state from $n = 0$. This cannot be done by a phase factor, yet the states differ little in

energy, contrary to London's assumption. It is this "extra degree of freedom" which was missing in London's argument and leads to the flux quantum being $hc/2e$ rather than hc/e.

The quantization of flux leads naturally to the stability of persistent supercurrents. In order for the supercurrent in the cylinder to "decay," the system must make transitions between macroscopically distinct states with different flux quantum number n. Thus the current cannot slowly "dribble" away, but it must decrease in macroscopic steps. The probability for such a macroscopic thermodynamic fluctuation is presumably vanishingly small since typical fluctuations involve the excitation of a few particles at a time rather than the entire system as a whole. This point of view has been emphasized by Bohr and Mottelson,[141] and while it is not an airtight theory, it is a compelling argument in the absence of a more complete calculation.

8-8 THE KNIGHT SHIFT

On the basis of our earlier discussion it would appear that the electronic spin susceptibility should vanish in a superconductor as $T \to 0$ if one uses s-state pairing. In this case the Pauli principle ensures that the spins are paired in the singlet state so that a finite spin magnetization can result only if the spin Zeeman energy $2\mu_B H$ is greater than $2\Delta_0$, the minimum energy to break up a pair. A means of checking this prediction is the Knight shift[142] (the change in nuclear magnetic resonance frequency due to the coupling of the nuclear spins with the polarization of the electrons). If only the electronic *spin* polarization (as opposed to orbital effects) is important in the shift, the Knight shift gives a measure of the electronic spin susceptibility. Reif[143] found that the shift in superconducting mercury, when extrapolated to $0°K$ was about two-thirds of the value in the normal state, contrary to the simple pairing theory. Androes and Knight[144] found similar results in tin, while the shift in vanadium is nearly the same in the N- and S-phases.

There have been several attempts to explain this discrepancy, none of which is widely accepted at present as providing the essential mechanism. It may well be that a fraction of the observed shift comes from several of the following mechanisms.

Parallel Spin Pairing

If one uses p-state (or any odd l-state) pairing, the antisymmetry of the wave-function forces the spins to be paired in triplet states. Fisher[59a] studied the model in which one pairs states with equal z-components of spin (i.e., up paired with up, down with down). The reduced Hamiltonian is then the sum of two noninteracting parts. In this case there is no energy gap for creating spin polarization since a down-spin pair can be changed into an up-spin pair with no change of energy. Unfortunately, this type of pairing leads to an anisotropic energy gap which vanishes in certain directions and a nonexponential electronic specific heat at low temperature. A more general triplet pairing has been treated by Balian and Werthamer[59c], who include all three components of the triplet state. In a pure material they obtain an isotropic energy gap; however, a small amount of disorder or impurity (which was certainly present in the above experiments) destroys their state as well as that of Fisher.

Spin-Orbit Coupling at the Surface

Since the Meissner effect screens out the magnetic field within a distance $\lambda \sim 5 \times 10^{-6}$ cm of the surface, the Knight shift experiments are carried out on specimens whose dimensions are small compared to λ in order to eliminate line broadening due to the field inhomogeniety. The smallness of the particles led Ferrell[145] to suggest that the spin-orbit coupling near the surface might be sufficiently strong to mix the one-electron spin states during scattering and thereby lead to a finite spin susceptibility. A semiquantitative theory of this effect was worked out by Ferrell and by Anderson,[146] however, we shall not discuss the calculations here. Unfortunately, as mentioned above, recent measurements on vanadium and aluminum,[147] which are light metals and should

have appreciably weaker spin-orbit coupling than mercury and tin, give essentially the *same* Knight shift in the normal and superconducting states. This rules out the spin-orbit mechanism, at least in this case.

Collective Magnetization

In the original BCS paper[8] it was suggested that there might be low-lying collective spin wave states in a superconductor which give rise to the observed Knight shift. Bardasis and Schrieffer[139] found that by retaining the p-wave part of the two-body potential, spin wave states can exist in the energy gap. However, to obtain a nonzero long wavelength susceptibility the spectrum must go down to zero energy in this limit. These authors found that two situations can exist. If the p-wave part of the two-body potential is weaker than the s-wave part, the spin waves possess a finite energy as their momentum goes to zero. If the p-wave potential is stronger than the s-wave potential, the spin wave states are unstable and the ground state is then formed by p-state pairing. One is then led back to the difficulties of the first proposal.

Modified Antiparallel Spin Pairing

Soon after the BCS theory was proposed, Heine and Pippard[148] suggested that one might relax the strict BCS rule of pairing a given state $\mathbf{k}$ with only one other state $\bar{\mathbf{k}}$. They argued that if $\mathbf{k}$ is paired with a *group* of states centered about $\bar{\mathbf{k}}$, a finite spin susceptibility might result. Their argument was based on an assumed form of the two-particle density matrix which is not consistent with that given by the most general form of the pairing theory. Since it has not been possible to construct a wavefunction which gives their density matrix, it appears that their basic assumption cannot be justified and we are forced to reject this point of view.

Subsequent to Heine and Pippard's proposal, Schrieffer[149] argued that a finite spin susceptibility would be obtained if the pairing condition were modified in the magnetized state. In

particular, he suggested pairing states $\mathbf{k}\uparrow$ and $-\mathbf{k}'\downarrow$, which are degenerate when the spin Zeeman energy is included, rather than the BCS prescription of pairing degenerate orbital states $\mathbf{k}\uparrow$ and $-\mathbf{k}\downarrow$. This scheme then leads to a net spin magnetization. In a pure unbounded specimen the modified pairing gives essentially no pairing energy since momentum conservation forbids a pair with center-of-mass momentum $\mathbf{k} - \mathbf{k}'$ from scattering into a pair state with center-of-mass momentum $\tilde{\mathbf{k}} - \tilde{\mathbf{k}}'$, owing to these momenta being different in general. Under experimental conditions, impurity and surface scattering is sufficiently strong to spread the single-particle states "$\mathbf{k}$" in momentum space by an amount large compared to the "center-of-mass" momentum $\hbar|\mathbf{k} - \mathbf{k}'| \sim 2\mu_B H/v_F$. Therefore, the actual single-particle eigenstates (including these one-particle scattering effects) can be formed into pairs in the above manner and coupled by the two-body potential. The pairing still occurs between *two definite* single-particle states. It is reasonable to assume that the reduction in the pairing energy will be a smoothly varying function of the net spin magnetic moment and therefore a finite spin susceptibility will result. This scheme is very different in point of view from that proposed by Heine and Pippard, who argue that a *group* of single-particle states are strongly correlated in occupancy (even in a pure, unbounded specimen).

Recently Cooper[150] has reinvestigated Schrieffer's idea by introducing a phenomenological two-body potential whose matrix elements are taken to be a slowly varying function of the center-of-mass momentum of *each* pair. This leads to a pairing energy which varies slowly as a function of the spin magnetization and gives a finite spin paramagnetism. Cooper stressed the added possibility of momentum nonconservation being due to a nontranslationally invariant two-body potential.

Orbital Paramagnetism

Clogston, Gossard, Jaccarino, and Yafet[151] have argued convincingly that in vanadium essentially *all* of the Knight shift is due to Kubo–Obata temperature-independent orbital paramagnetism.[152] Since this orbital paramagnetism should be the same in the N- and S-states, the lack of change in the Knight shift in vanadium is presumably explained without modifying the pairing theory. It appears unlikely that this mechanism can account for the observed shift in all cases.

8-9 THE GINSBURG–LANDAU–GOR'KOV THEORY

Thus far we have concentrated on the response of a superconductor to weak electromagnetic fields. There are many important problems, e.g., N-S phase boundary, the intermediate and mixed states, etc., in which the magnetic field enters in a nonperturbative manner. These problems typically involve the energy-gap parameter Δ varying as a function of position in the materials. As we discussed in Chapter 1, the phenomenological theory of Ginsburg and Landau[36] (proposed in 1950) gives in many instances a good account of these strong field situations. An important advance in the microscopic theory was made by Gor'kov,[37] who showed how the GL equations follow from the pairing theory when T is near T_c and the magnetic field varies slowly in space over a coherence length. Gor'kov found that the GL effective wave function $\Psi(\mathbf{r})$ is proportional to the local value of the gap parameter $\Delta(\mathbf{r})$ and the effective charge e^* of the GL theory is equal to $2e$, the charge of a pair of electrons. It is interesting to note that these results were guessed prior to Gor'kov's work, the identification of $\Psi(\mathbf{r})$ and $\Delta(\mathbf{r})$ being suggested by Bardeen[8] and the effective charge $e^* = 2e$ being suggested by Ginsburg[153] prior to the BCS theory in order to fit the GL theory with experiment.

We give a brief summary of Gor'kov's derivation below. To familiarize the reader with Gor'kov's scheme, we use his notation. For simplicity Gor'kov used a nonretarded zero-range attractive potential to describe the pairing interactions. Since this singular potential leads to divergences, it is cut off in momentum space at the appropriate point in the derivation. The vector

potential $\mathbf{A}(\mathbf{r})$ is treated self-consistently as we did above in the weak-field case. The system Hamiltonian is then

$$H = -\sum_s \int \psi_s{}^+(\mathbf{r})\left\{\frac{1}{2m}\left[\nabla - \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu\right\}\psi_s(\mathbf{r})\, d^3r - v\int \psi_\uparrow{}^+(\mathbf{r})\psi_\uparrow(\mathbf{r})\psi_\downarrow{}^+(\mathbf{r})\psi_\downarrow(\mathbf{r})\, d^3r \quad (8\text{-}136)$$

where $e = -|e|$ is the charge of an electron, and we measure single-particle energies relative to the chemical potential μ. The thermodynamic Green's function

$$G(x, x') = -\frac{\mathrm{Tr}\,[e^{-\beta H}T\{\psi_\uparrow(x)\psi_\uparrow{}^+(x')\}]}{\mathrm{Tr}\, e^{-\beta H}} \equiv -\langle T\{\psi_\uparrow(x)\psi_\uparrow{}^+(x')\}\rangle \quad (8\text{-}137a)$$

is defined for pure imaginary time

$$\psi(x) \equiv \psi(r, \tau) = e^{H\tau}\psi(\mathbf{r}, 0)e^{-H\tau} \quad (8\text{-}137b)$$

and satisfies the equation of motion

$$\left\{-\frac{\partial}{\partial\tau} + \frac{1}{2m}\left[\nabla - \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu\right\}G(x, x') + V\langle T\{\psi_\uparrow{}^+(x')\psi_\downarrow{}^+(x)\psi_\downarrow(x)\psi_\uparrow(x)\}\rangle = \delta(x - x') \quad (8\text{-}138)$$

This result follows from (8-136) and (8-137) since

$$\frac{\partial\psi_\uparrow(x)}{\partial\tau} = [H, \psi(x)] = \left\{\frac{1}{2m}\left[\nabla - \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu\right\}\psi_\uparrow(x) + V\psi_\downarrow{}^+(x)\psi_\downarrow(x)\psi_\uparrow(x) \quad (8\text{-}139)$$

In Gor'kov's scheme, the pairing approximation for this zero-range potential corresponds to factorizing the four-point function in (8-138),

$$\langle T\{\psi_\uparrow{}^+(x')\psi_\downarrow{}^+(x)\psi_\downarrow(x)\psi_\uparrow(x)\}\rangle \Rightarrow \langle T\{\psi_\uparrow{}^+(x')\psi_\downarrow{}^+(x)\}\rangle\langle\psi_\downarrow(x)\psi_\uparrow(x)\rangle \quad (8\text{-}140)$$

Therefore the equation for G becomes

$$\left\{-\frac{\partial}{\partial\tau} + \frac{1}{2m}\left[\nabla - \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu\right\}G(x, x') + \Delta(\mathbf{r})F^+(x, x') = \delta(x - x') \quad (8\text{-}141)$$

where the "anomalous" Green's function $F^+(x, x')$ is defined by

$$F^+(x, x') = -\langle T\{\psi_\downarrow{}^+(x)\psi_\uparrow{}^+(x')\}\rangle \quad (8\text{-}142)$$

and the energy-gap parameter $\Delta(\mathbf{r})$ is given by

$$\Delta^*(\mathbf{r}) = V\langle\psi_\downarrow(\mathbf{r})\psi_\uparrow(\mathbf{r})\rangle^* = VF^+(x, x') \quad (8\text{-}143)$$

The functions G and F^+ correspond to Nambu's G_{11} and G_{21}. Since F^+ is an unknown function, it must be determined from its equation of motion. By making a factorization similar to that in (8-140), (except that the four-point function now contains three ψ^+'s and one ψ), one finds

$$\left\{\frac{\partial}{\partial\tau} + \frac{1}{2m}\left[\nabla + \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu\right\}F^+(x, x') - \Delta^*(\mathbf{r})G(x, x') = 0 \quad (8\text{-}144)$$

As we discussed in Chapter 7, the pure imaginary time Green's functions can be expressed in the Fourier series variable $\omega_n = (2n + 1)\pi/\beta$ (n = integer) and one finds the Fourier components $\mathscr{G}_\omega(\mathbf{r}, \mathbf{r}')$ and $\mathscr{F}_\omega(\mathbf{r}, \mathbf{r}')$ satisfy

$$\left\{i\omega_n + \frac{1}{2m}\left[\nabla - \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu\right\}\mathscr{G}_\omega(\mathbf{r}, \mathbf{r}') + \Delta(\mathbf{r})\mathscr{F}_\omega(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')$$

$$\left\{-i\omega_n + \frac{1}{2m}\left[\nabla + \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu\right\}\mathscr{F}_\omega(\mathbf{r}, \mathbf{r}') - \Delta^*(\mathbf{r})\mathscr{G}_\omega(\mathbf{r}, \mathbf{r}') = 0 \quad (8\text{-}145)$$

These equations together with the condition (8-143)

$$\Delta^*(\mathbf{r}) = \frac{V}{\beta}\sum_n \mathscr{F}_\omega{}^+(\mathbf{r}, \mathbf{r}') \quad (8\text{-}146)$$

in principle determine the behavior of the superconductor in the presence of an arbitrarily strong potential A at any temperature $k_BT = 1/\beta$.

The nonlinearity of the coupled equations makes them difficult to handle. Gor'kov restricted his attention to the region T near T_c, where the gap parameter is small so that a perturbation expansion in powers of Δ can be carried out (in the spirit of the GL theory). Furthermore, the penetration depth λ becomes

large compared to Pippard's coherence length ξ_0 for $T \sim T_c$ and therefore **A** will vary slowly over a coherence length. In this limit the linear relation between the current density and the vector potential reduces to London's equation. To carry out the series solution of (8-145) in powers of Δ, Gor'kov wrote these equations in integral form

$$\mathcal{G}_\omega(\mathbf{r}, \mathbf{r}') = \tilde{\mathcal{G}}_\omega(\mathbf{r}, \mathbf{r}') - \int \tilde{\mathcal{G}}_\omega(\mathbf{r}, \mathbf{s})\, \Delta(\mathbf{s}) \mathcal{F}_\omega{}^+(\mathbf{s}, \mathbf{r}')\, d^3s \quad (8\text{-}147a)$$

$$\mathcal{F}_\omega{}^+(\mathbf{r}, \mathbf{r}') = \int \mathcal{G}_\omega(\mathbf{s}, \mathbf{r}')\, \Delta^*(\mathbf{s}) \tilde{\mathcal{G}}_{-\omega}(\mathbf{s}, \mathbf{r})\, d^3s \quad (8\text{-}147b)$$

where $\tilde{\mathcal{G}}$ is the Green's function for an electron in the normal metal in the presence of the magnetic field

$$\left\{ i\omega_n + \frac{1}{2m}\left[\nabla - \frac{ie}{c}\mathbf{A}(\mathbf{r})\right]^2 + \mu \right\} \tilde{\mathcal{G}}_\omega(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \quad (8\text{-}148)$$

If one solves the equations for F^+, accurate to terms of order Δ^4, one finds that the equation (8-146) determining Δ is

$$\begin{aligned} \Delta^*(\mathbf{r}) = {} & \frac{V}{\beta} \sum_n \int \tilde{\mathcal{G}}_\omega(\mathbf{r}, \mathbf{r}') \tilde{\mathcal{G}}_{-\omega}(\mathbf{r}, \mathbf{r}')\, \Delta^*(\mathbf{r}')\, d^3r' \\ & - \frac{V}{\beta} \sum_n \int \tilde{\mathcal{G}}_\omega(\mathbf{s}, \mathbf{r}) \tilde{\mathcal{G}}_{-\omega}(\mathbf{s}, \mathbf{l}) \tilde{\mathcal{G}}_\omega(\mathbf{m}, \mathbf{l}) \tilde{\mathcal{G}}_{-\omega}(\mathbf{m}, \mathbf{r})\, \Delta(\mathbf{s}) \\ & \times \Delta^*(\mathbf{l})\, \Delta^*(\mathbf{m})\, d^3s\, d^3l\, d^3m \end{aligned} \quad (8\text{-}149)$$

The first term on the right-hand side of the form

$$\int K(\mathbf{r}, \mathbf{r}')\, \Delta^*(\mathbf{r}')\, d^3r \quad (8\text{-}150a)$$

where the kernel $K(\mathbf{r}, \mathbf{r}')$

$$K(\mathbf{r}, \mathbf{r}') = \frac{1}{\beta} \sum_n \tilde{\mathcal{G}}_\omega(\mathbf{r}, \mathbf{r}') \tilde{\mathcal{G}}_{-\omega}(\mathbf{r}, \mathbf{r}') \quad (8\text{-}150b)$$

is given by

$$K_0(\mathbf{r} - \mathbf{r}') = K_0(R) = \left[\frac{m}{2\pi R}\right]^2 \frac{1}{\beta} \sinh\left(\frac{2\pi R}{\beta v_F}\right) \quad (8\text{-}150c)$$

and

$$\tilde{\mathcal{G}}_\omega{}^0(\mathbf{r} - \mathbf{r}') = -\frac{m}{2\pi R} \exp\left[ip_F R \operatorname{sgn} \omega_n - \frac{|\omega_n|}{v_F} R\right] \quad (8\text{-}150d)$$

if $\mathbf{A} = 0$. Since **A** is assumed to vary slowly over a coherence length and $\tilde{\mathcal{G}}_\omega(\mathbf{r}, \mathbf{r}')$ decreases exponentially for $|\mathbf{r} - \mathbf{r}'| > v_F/\omega \simeq \xi_0$ ($\omega \simeq \Delta$), a WKB-like approximation can be used to include **A** and one finds

$$K(\mathbf{r}, \mathbf{r}') = K_0(\mathbf{r} - \mathbf{r}') \exp\left[2\frac{ie}{c}(\mathbf{r} - \mathbf{r}') \cdot \mathbf{A}\ (\mathbf{r})\right] \quad (8\text{-}150e)$$

and

$$\tilde{\mathcal{G}}_\omega(\mathbf{r}, \mathbf{r}') = \tilde{\mathcal{G}}_\omega{}^0(\mathbf{r} - \mathbf{r}') \exp\left[\frac{ie}{c}(\mathbf{r} - \mathbf{r}') \cdot \mathbf{A}\ (\mathbf{r})\right] \quad (8\text{-}150f)$$

The singularity of K_0 as $\mathbf{r} \to \mathbf{r}'$ arises from the zero-range two-body potential. If one cuts off the potential $V_{kk'}$ outside the energy range $-\omega_0 \to \omega_0$ (centered about the Fermi surface), one has

$$\begin{aligned} \int K_0(R)\, dR &= N(0) \int_0^{\omega_0} \frac{1}{\epsilon} \tanh\left(\frac{\beta\epsilon}{2}\right) d\epsilon \\ &= N(0)\left[\int_0^{\omega} \frac{1}{\epsilon} \tanh\left(\frac{\beta_c\epsilon}{2}\right) d\epsilon + \int_{\beta_c\omega_0}^{\beta\omega_0} \frac{\tanh x}{x}\, dx\right] \\ &= N(0)\left[\frac{1}{N(0)V} + \ln\left(\frac{T_c}{T}\right)\right] \end{aligned} \quad (8\text{-}151)$$

In the reduction we have used the equation determining $k_B T_c = 1/\beta_c$. By expanding the normal metal Green's functions in powers of the small quantity $(e/c)(\mathbf{r} - \mathbf{r}') \cdot \mathbf{A}(\mathbf{r})$ and assuming that $\Delta(\mathbf{r})$ varies slowly over a coherence length, Gor'kov obtains the equation

$$\begin{aligned} \left\{ \frac{1}{2m}\left[\nabla + i\frac{2e}{c}\mathbf{A}(\mathbf{r})\right]^2 \right. \\ \left. + \frac{1}{\lambda_G}\left[\left(1 - \frac{T}{T_c}\right) - \frac{7\zeta(3)}{8(\pi k_B T_c)^2} |\Delta(\mathbf{r})|^2\right] \right\} \Delta^*(\mathbf{r}) = 0 \end{aligned} \quad (8\text{-}152)$$

where Gor'kov's parameter λ_G is

$$\lambda_G = \frac{7\zeta(3) E_F}{12(\pi k_B T_c)^2} \quad (8\text{-}153)$$

and $\zeta(x)$ is the Riemann zeta function.

By introducing the "wave function"

$$\psi(\mathbf{r}) = \frac{\Delta(\mathbf{r})[7\zeta(3)n]^{1/2}}{4\pi T_c} \tag{8-154}$$

one obtains the Ginsburg–Landau-like equation

$$\left\{\frac{1}{2m}\left[\nabla - \frac{ie^*}{c}\mathbf{A}(\mathbf{r})\right]^2 + \frac{1}{\lambda_G}\left[\left(1 - \frac{T}{T_c}\right) - \frac{2}{N}|\psi(\mathbf{r})|^2\right]\right\}\psi(\mathbf{r}) = 0 \tag{8-155}$$

where $e^* = 2e$. One can also calculate the current density

$$\mathbf{J}(\mathbf{r}) = \left[\frac{ie}{m}(\nabla_{r'} - \nabla_r)G(x, x') - \frac{2e^2}{mc}\mathbf{A}(\mathbf{r})G(x, x')\right]_{t'=t^+, \mathbf{r}=\mathbf{r}'} \tag{8-156}$$

to second order in Δ by the perturbation expansion used above and one finds on using the relation between ψ and Δ:

$$\mathbf{J}(\mathbf{r}) = -\frac{ie^*}{2m}(\psi^*\nabla\psi + \psi\nabla\psi^*) - \frac{e^{*2}}{mc}\mathbf{A}(\mathbf{r})|\psi(\mathbf{r})|^2 \tag{8-157}$$

as in the GL theory.

Recently, the derivation of Gor'kov has been extended to all temperatures by Werthamer[154] and by Tewordt,[155] who continue to assume the system is such that $\mathbf{A}$ and Δ vary slowly over a coherence length. Their equations are somewhat more complicated than the GLG form, as one might expect. Gor'kov has extended his treatment to include finite mean-free-path effects. He finds the equations have the same form as above, except that the "mass" m is increased relative to that of the pure material.