7 Hole Superconductivity

Jorge E. Hirsch
Department of Physics
University of California San Diego, La Jolla, CA, USA

(Lecture Notes of the Autumn School on Correlated Electrons 2024, Correlations and Phase Transitions, Chpt. 7, edited by Eva Pavarini and Erik Koch)

Contents

1 Introduction: Overview of superconductivity

That some metals, called "supraconductors" in the old days [1], make a transition to a state with zero electrical resistance below a critical temperature T_c was discovered experimentally by Kammerlingh Onnes in 1911 [2]. That if a magnetic field is present in the interior of a superconductor for $T > T_c$, it gets expelled when the temperature is lowered to $T < T_c$ was discovered experimentally by Meissner and Ochsenfeld in 1933 [3] and is called the Meissner effect. The Meissner effect was a great surprise: before 1933 it was expected that superconductors would $exclude$ magnetic fields but not that they would $expel$ magnetic fields. This follows from Faraday's law, and was known as "Lippmann's theorem" [4] back in the days [5]: if a magnetic field is applied to a zero resistance material, the material will react by generating a surface current that does not let the field penetrate, thus excluding the magnetic field from its interior. However, Faraday's law / Lippmann's theorem would predict that if a material with finite resistance has a magnetic field in its interior, when it is cooled into the superconducting state with zero resistance no current would flow and the magnetic field would remain in the interior. even if the external source of magnetic field is removed. That is not what superconductors do: metals going superconducting spontaneously generate a surface current that $expels$ magnetic fields from their interior [3]. This appears to violate Faraday's law.

The London equation proposed in 1935 by the London brothers [1, 6] provided a phenomenological description of the magnetic behavior of superconductors, but did not explain how superconductors manage to violate Faraday's law. Neither did the BCS theory, proposed in 1957 by Bardeen, Cooper and Schrieffer [7], based on the electron-phonon interaction. BCS theory provided a microscopic description of superconductors that describes many of their properties accurately, it is generally believed to apply to materials called "conventional superconductors", that include all superconducting elements and many compounds. There are around 30 different classes of superconducting materials [8], only about a third of them are generally agreed to be "conventional superconductors". For the remaining two thirds, there is no generally accepted theory. The field is wide open for further progress.

1.1 The known knowns: London equation, Cooper pairs, BCS theory

The known knowns are what ***we know we know*** about superconductivity.

It took 22 years, from the discovery of zero resistance by Kammerlingh Onnes in 1911, to experimentally discover the Meissner effect in 1933. The London brothers embodied this experimental fact in the London equation [1, 6]

$$\vec{\nabla} \times \vec{J} = -\frac{ne^2}{m_e c}\vec{B} \tag{1}$$

where $\vec{J}$ is the current density, e and m_e are the electron's charge and mass, n is the carrier density and $\vec{B}$ is an applied magnetic field. Eq. (1), together with Ampere's law

$$\vec{\nabla} \times \vec{B} = (4\pi/c)\vec{J} \tag{2}$$

lead to

$$\nabla^2 \vec{B} = \frac{4\pi n e^2}{m_e c^2} \vec{B} \equiv \frac{1}{\lambda_L^2} \vec{B} \tag{3}$$

which says that an external magnetic field decays exponentially to zero in going from the surface to the interior of the superconductor, over a length given by the London penetration depth λ_L. The London Eq. (1) was $derived$ by the Londons for the situation where a magnetic field is applied to a material that is already superconducting, i.e. describing $exclusion$ of an applied magnetic field. They $postulated$ without derivation that it also applies to situations where a normal metal with a magnetic field in its interior is cooled into the superconducting state. If that postulate is valid, the material will $expel$ the magnetic field to reach the state described by Eq. (3), which is what observed experimentally [3]. But no theoretical proof that this should happen was provided by the London brothers nor anybody else.

In 1956, Leon Cooper pointed out [9] that if electrons in a Fermi gas interact through a small net $attractive$ interaction resulting from the electron-phonon interaction, they would form a bound pair with binding energy Δ, and suggested that a system of such bound pairs may Bose-condense into a superconducting state. In 1957 Bardeen, Cooper and Schrieffer formulated a theory of superconductivity [7] describing the many-body state of such Cooper pairs. They showed that the ground state wavefunction of the "reduced Hamiltonian"

$$H_{red} = \sum_{k\sigma} (\epsilon_k - \mu) c^\dagger_{k\sigma} c_{k\sigma} + \sum_{kk'} V_{kk'} c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-k'\downarrow} c_{k'\uparrow} \tag{4}$$

with μ the chemical potential (that determines the band filling) is of the form

$$|\Psi_{BCS}> = \prod_k (u_k^2 + v_k^2 c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}) |0> . \tag{5}$$

where $c^\dagger_{k\sigma}$ creates an electron of spin σ in the single particle state of crystal momentum k, and u_k and v_k are complex amplitudes determined by minimization of the energy. BCS showed that if the net interaction between electrons is attractive, the state Eq. (5) has a lower energy than the normal metal Fermi sea. At finite temperatures, the system has quasiparticle excitations with minimum energy $\Delta(T)$, the superconducting energy gap, which is a decreasing function of temperature. At a critical temperature T_c, $\Delta(T)$ reaches zero and the system transitions into the normal state. They furthermore showed that below T_c the system has many properties that resemble what is experimentally found for superconductors.

BCS formulated their theory under the assumption that the attractive interaction between electrons resulted from the electron-phonon interaction. Under that assumption, a simplified expression for the critical temperature is

$$T_c = \hbar \omega_D e^{-1/\lambda} \tag{6}$$

where λ is the dimensionless electron-phonon coupling and ω_D is the Debye temperature, inversely proportional to the square root of the ionic mass M. Hence Eq. (6) predicts the isotope effect, namely that

$$\frac{\partial ln(T_c)}{\partial M} = -\alpha \tag{7}$$

with $\alpha = 1/2$.

Finally, we should include among the *"known knowns"* the understanding intuited by London [10], implicit in BCS theory, anticipated by Ginzburg and Landau [11], and demonstrated by the physical effects predicted by Brian Josephson [12], that the superconducting condensate can be represented by a macroscopic quantum wavefunction $\psi(\vec{r}) = |\psi(\vec{r})|e^{i\theta(\vec{r})}$ with a unique phase $\theta(\vec{r})$ common to all electrons in the superfluid, and an amplitude $|\psi(\vec{r})|$ whose square gives the density of superfluid electrons.

All of the above, we all agree we know. Hence they are ***known knowns*** . Good references for the above are the books by Tinkham [13] and by de Gennes [14].

1.2 The known unknowns: Mechanisms of unconventional superconductivity

The known unknowns are what ***we know that we don't know*** about superconductivity.

We know (meaning everybody agrees) that for materials that conduct electricity at ambient pressure, hence can potentially be superconductors, the electron-phonon interaction is not strong enough to overcome the Coulomb repulsion between electrons and give rise to superconductivity at temperatures above liquid nitrogen temperature, 77K [15]. Some cuprate superconductors [16], a class of materials discovered in 1986, superconduct up to much higher temperatures, up to 140K. Therefore, we know that there has to be at least one other mechanism that gives rise to superconductivity that is not the electron-phonon interaction. Superconducting materials not driven by the electron-phonon interaction are called "unconventional superconductors".

For a variety of reasons, many classes of materials, even if they have critical temperatures much lower than 77K, are believed to be "unconventional superconductors", as surveyed in Ref. [8]. There are a large number of unconventional theories of superconductivity proposed to describe the cuprates and other unconventional superconductors (see introduction in Ref. [17] for many references), but there is no general agreement on which (if any) of the theories is correct for any materials. These theories are generally proposed to apply to one or more classes of unconventional superconductors but not to all.

Since it is known that there are superconducting materials not described by BCS electron-phonon theory, and there is no general agreement on which mechanisms give rise to superconductivity in the so-called unconventional superconductors, the mechanism(s) that give rise to unconventional superconductivity are ***known unknowns***.

1.3 The unknown unknowns: Does BCS theory explain the Meissner effect?

The unknown unknowns are what ***we don't know that we don't know*** about superconductivity.

It is universally believed that BCS theory explains the Meissner effect. I disagree, I believe this has not been carefully considered. BCS showed that the BCS state with magnetic field ex-

cluded has lower free energy than the normal state with magnetic field in the interior. However, BCS theory has not explained the $process$ by which a normal metal becoming superconducting $expels$ the magnetic field to reach the BCS state. I have argued [18] that BCS superconductors (meaning superconductors described by BCS theory) do not have the physical elements necessary to give rise to a Meissner effect, and that as a consequence, if cooled from above T_c to below T_c in the presence of a magnetic field, they would not make a transition to the superconducting state that excludes the magnetic field but would instead remain in a metastable normal state with the magnetic field remaining in the interior, contrary to what is seen experimentally. This implies, since real superconductors do exhibit a Meissner effect, that real superconductors cannot be described by BCS theory. Since the Meissner effect is not generally considered to be an unexplained phenomenon, I call this an ***unknown unknown***.

Moreover, I believe there is a single theory to describe all superconductors, i.e. the so-called conventional and unconventional superconductors, that also explains the Meissner effect. A survey of the theory in its present state is given in my book [19]. A substantial part of the theory was developed in collaboration with Frank Marsiglio, papers are listed in Ref. [20].

1.4 The unknown knowns: Are there *any* electron-phonon superconductors?

An unknown known is something ***we think we know but in reality we don't know***.

There is essentially universal agreement that the electron-phonon interaction gives rise to superconductivity in conventional materials, including the hydrides under high pressure [21]. Hence, that electron-phonon superconductors exist in nature. For the hydrides, transition temperatures are claimed to approach room temperature, in drastic violation of what was expected [15]. This is argued to come about due to the light mass of hydrogen and the strong electron-phonon coupling in such materials [21], as originally predicted by Ashcroft [22].

There is however no rigorous proof that the electron-phonon interaction gives rise to superconductivity in any material. The direct Coulomb repulsion between electrons is a first order effect, generally much stronger than the second-order frequency-dependent electron-electron interaction mediated by phonons that can be attractive under certain conditions. In calculations that claim to predict superconductivity driven by the electron-phonon interaction, the effect of the Coulomb repulsion is generally lumped into a phenomenological parameter μ^* [23], the "Coulomb pseudopotential", assumed to be of order 0.1, which allows for superconductivity driven by the attractive electron-phonon interaction. However there is no reliable way to calculate μ^* [24]. At the time when no other mechanisms for attractive interactions were known, i.e in the 1960's, this might have been a tenable scenario, it was adopted then and is firmly believed to be valid to this date. I believe this is wrong, that in fact the electron-phonon interaction is irrelevant to superconductivity in all materials. Since everybody believes the opposite, that there are electron-phonon superconductors in nature, I say that this is an ***unknown known***.

2 Correlated electron models for superconductivity

So-called Hubbard models are widely used to describe correlated electrons in solids. Various incarnations of these models have been used to describe essential aspects of superconductivity within a variety of different theories.

2.1 Attractive and repulsive Hubbard models

The simplest Hubbard model is given by the tight binding Hamiltonian

$$H = -\sum_{i,j,\sigma} t_{ij}(c^\dagger_{i\sigma}c_{j\sigma} + h.c.) + U\sum_i n_{i\uparrow}n_{i\downarrow}. \tag{8}$$

describing electrons in a single orbital at each site in a lattice of N sites, with on-site interactions only. In momentum space the Hamiltonian is

$$H = \sum_{k\sigma} \epsilon_k c^\dagger_{k\sigma}c_{k\sigma} + \frac{U}{N}\sum_{kk'q} c^\dagger_{k+q\uparrow}c^\dagger_{k'-q\downarrow}c_{k'\downarrow}c_{k\uparrow} \tag{9}$$

For an attractive interaction $U < 0$, the ground state of its "reduced Hamiltonian", of the form Eq. (4), is of the form Eq. (5). This follows from the fact that the BCS gap equation [13]

$$\Delta_k = -\frac{1}{N}\sum_{k'} V_{kk'}\Delta_{k'}\frac{1-2f(E_{k'})}{2E_{k'}} \tag{10}$$

with f the Fermi distribution function, $V_{kk'} = U$, $\Delta_k = \Delta$, and

$$E_k = \sqrt{(\epsilon_k - \mu)^2 + \Delta^2} \tag{11}$$

has a solution $\Delta \neq 0$ at sufficiently low temperatures for any $U < 0$.

For a repulsive interaction $U > 0$, the gap equation (10) has no solution. However, it has been argued [25] that more elaborate treatments of the repulsive Hubbard model Eq. (8) do give rise to a superconducting state induced by spin fluctuations in that model as well as the related t-J model [26], with the gap function having d-wave symmetry

$$\Delta_k = \Delta_0(cos(k_x a) - cos(k_y a)). \tag{12}$$

There is however no rigurous proof that I know of that the repulsive Hubbard model has a superconducting ground state. The t-J model does have a superconducting ground state, and it is argued that the model results from the repulsive Hubbard model in the limit of strong coupling, $U >> t$ [26]. However I have shown [27] that to the same order in t/U the Hamiltonian resulting from the Hubbard model has, in addition to the t-J terms, three-site terms that exactly cancel the attractive interaction resulting from the two-site terms.

There are no good reasons why the $attractive$ Hubbard model should describe the essential physics of interacting electrons in real materials. And whether the $repulsive$ Hubbard model exhibits superconductivity for any value of $U > 0$ is certainly not established [28, 29]. Thus, the question whether the Hubbard model Eq. (8) has any relevance to the superconductivity of real materials is not a settled question, despite the enormous amounts of research efforts that has been devoted to that hypothesis [30, 31] during the last 40 years [32].

2.2 Electron-phonon attraction versus electron-electron attraction

The interaction between electrons and phonons (lattice vibrations) is typically described by the Hamiltonian

$$H_{e-ph} = \frac{1}{\sqrt{N}} \sum_{k,q,\sigma} g_{kq}(b_q + b^\dagger_{-q}) c^\dagger_{k+q\sigma} c_{k\sigma} \tag{13}$$

where the operators b_q and $b^\dagger_{-q}$ destroy and create phonons with wavevector q and $-q$ respectively, and frequency $\omega_q = \omega_{-q}$. Eliminating the phonons in second order perturbation theory leads to the effective electron-electron interaction [33]

$$H^{ph}_{e-e} = \frac{1}{N} \sum_{k,q,\sigma\sigma'} |g_{kq}|^2 \frac{2\hbar\omega_q}{(\epsilon_{k+q} - \epsilon_k)^2 - (\hbar\omega_q)^2} c^\dagger_{k+q\sigma} c^\dagger_{k'-q\sigma'} c_{k'\sigma'} c_{k\sigma} \tag{14}$$

which is attractive for $|\epsilon_{k+q} - \epsilon_k| < \hbar\omega_q$. It is argued that this frequency-dependent ("retarded") attractive interaction between electrons near the Fermi surface can overcome the much larger instantaneous repulsive interaction between electrons, given in its simplest form by the Hubbard repulsion Eq. (9), through what is called the "Coulomb pseudopotential" effect [23, 24] that we will not go into here.

The electron-electron attraction mediated by phonons Eq. (14), identified in the 1950's [33], arises from *second-order* processes. In the 1960's, it was pointed out that another source of electron-electron attraction could be second order 'excitonic' processes, where the excitations are electronic rather than phononic giving rise to interactions of the form Eq. (14) where the phonon energy $\hbar\omega_q$ is replaced by an electronic excitation. However those proposals did not gain much traction.

It turns out however that a *first order* attractive interaction between electrons exists, originating in the Coulomb interaction between electrons in the presence of the periodic ionic lattice potential [34]. In a tight binding formulation, the Hamiltonian containing diagonal as well as off-diagonal matrix elements resulting from the Coulomb interaction is a 'generalized Hubbard model' [35, 36] given in real space by

$$H_{gen} = -\sum_{i,j,\sigma} t_{ij}(c^\dagger_{i\sigma} c_{j\sigma} + h.c.) + \sum_{ijkl} (ij/kl) c^\dagger_{i\sigma} c^\dagger_{j\sigma'} c_{l\sigma'} c_{k\sigma} \tag{15}$$

where

$$(ij/kl) = \int d^3r d^3r' \varphi^*_i(r) \varphi^*_j(r') \frac{e^2}{|r - r'|} \varphi_l(r') \varphi_k(r) \tag{16}$$

where φ_i is the atomic orbital associated with site i, $U = (ii/ii) > 0$ is the on-site repulsion, and the repulsion between electrons on different sites i, j is $V_{ij} = (ij/ij) > 0$. Here we want to focus on the two-center off-diagonal matrix element of the Coulomb interaction (ii/ij) given by

$$(ii/ij) \equiv \Delta t_{ij} = \int d^3r d^3r' \varphi^*_i(r) \varphi_j(r) \frac{e^2}{|r - r'|} |\varphi^*_i(r')|^2. \tag{17}$$

The sign of the interaction Eq. (17) depends on the orbitals involved, in particular it is positive for s-orbitals and negative for p-orbitals oriented along the i-j direction. The important point

however is that Eq. (17) is *of the same sign* as the single electron hopping matrix element t_{ij} arising from the electron-ion interaction, given approximately by [37]

$$t_{ij} = -\int d^3r \varphi_i^*(r)(-\frac{Ze^2}{|r-R_i|})\varphi_j(r) \tag{18}$$

where $Z|e|$ is the ionic charge, that we can rewrite as

$$t_{ij} = -\int d^3r d^3r' \varphi_i^*(r)\varphi_j(r)(-\frac{Ze^2}{|r-r'|})|\chi_i(r')|^2 \tag{19}$$

with $\chi_i(r')$ the 'ionic wave function' such that $|\chi_i(r')|^2 = \delta(r'-R_i)$, to make its close relationship with Eq. (17) apparent [34]. Thus, it is reasonable to assume that $\Delta t_{ij} = \alpha t_{ij}$ with α a $positive$ constant. The Hamiltonian that results then including this interaction and the Hubbard on-site repulsion U is

$$H = -\sum_{i,j,\sigma}(t_{ij} - \Delta t_{ij}(n_{i,-\sigma} + n_{j,-\sigma}))(c_{i\sigma}^\dagger c_{j\sigma} + h.c.) + U\sum_i n_{i\uparrow}n_{i\downarrow}. \tag{20}$$

The interaction Δt_{ij} is called "correlated hopping". In momentum space the Hamiltonian is

$$H = \sum_{k\sigma}\epsilon_k c_{k\sigma}^\dagger c_{k\sigma} + \frac{1}{N}\sum_{kk'q}(U - \alpha(\epsilon_k + \epsilon_{k+q} + \epsilon_{k'} + \epsilon_{k'-q}))c_{k+q\uparrow}^\dagger c_{k'-q\downarrow}^\dagger c_{k'\downarrow}c_{k\uparrow} \tag{21}$$

with $\epsilon_k = \sum_j t_{0j}e^{ikR_j}$, where $\sum_k \epsilon_k = 0$ since we are defining $t_{ii} = 0$. It can be seen from Eq. (21) that this interaction $increases$ the Hubbard repulsion near the bottom of the band, where $\epsilon_k < 0$, and $decreases$ it near the top of the band where $\epsilon_k > 0$. It is the only interaction of the form Eq. (16) involving two centers that breaks electron-hole symmetry.

The reduced Hamiltonian for Eq. (21) is

$$H_{red} = \sum_{k\sigma}(\epsilon_k - \mu)c_{k\sigma}^\dagger c_{k\sigma} + \frac{1}{N}\sum_{kk'} V_{kk'}c_{k\uparrow}^\dagger c_{-k\downarrow}^\dagger c_{-k'\downarrow}c_{k'\uparrow} \tag{22a}$$

$$V_{kk'} = U - 2\alpha(\epsilon_k + \epsilon_{k'}) \tag{22b}$$

with $-(D/2) \le \epsilon_k \le (D/2)$, with D the bandwidth. For a system with only nearest neighbor hopping with $t_{ij} = t$, $\Delta t_{ij} = \Delta t$, and z nearest neighbors to each atom, the bandwidth is $D = 2zt$ and

$$U - 4z\Delta t < V_{kk'} < U + 4z\Delta t \tag{23}$$

so the reduction of the on-site repulsion U increases with the coordination number of the lattice. Even for parameter values such that $V_{kk'}$ is always repulsive, i.e. $U > 4z\Delta t$, the BCS gap equation Eq. (10) will have solutions [38]. This happens because Δ_k changes sign and becomes negative far from the Fermi energy when the interaction is most repulsive. We have called this a "spatial pseudopotential effect" [38], since it is analogous to what occurs for the electron-phonon interaction due to its frequency dependence [23].

We can also write the Hamiltonian Eq. (20), assuming we are dealing with a bipartite lattice, in terms of hole operators, by performing the transformation

$$c_{i\sigma}^{\dagger} \rightarrow (-1)^i c_{i\sigma} \tag{24}$$

and it becomes

$$H = -\sum_{i,j,\sigma}(\bar{t}_{ij} + \Delta t_{ij}(n_{i,-\sigma} + n_{j,-\sigma}))(c_{i\sigma}^{\dagger}c_{j\sigma} + h.c.) + U\sum_i n_{i\uparrow}n_{i\downarrow}. \tag{25}$$

with $\bar{t}_{ij} = t_{ij} - 2\Delta t_{ij}$. It will be more useful in this form for further developments.

There is a caveat however associated with the derivation given here. In the computation of the Coulomb matrix elements Eq. (16) we have assumed that the orbitals are atomic orbitals, which are not orthogonal at neighboring sites. Instead, the Hamiltonian Eq. (15) implicitly assumes that the fermion operators create electrons in orthogonal orbitals. If we use orthogonal orbitals instead to compute the matrix elements of the Coulomb interaction, it is found that the off-diagonal element given by Eq. (17) is nearly zero [39, 40]. However, we will show in the next section that there are other physical reasons for why the Hamiltonian Eq. (20) with appreciable values of Δt_{ij} is relevant.

Thus, both the electron-phonon Hamiltonian Eq. (13) and the purely electronic Hamiltonian without electron-phonon interaction Eq. (21) can give rise to superconductivity in model systems. Whether or not both or one of them or neither of them gives rise to superconductivity in real materials is a question for which there is no universally agreed answer.

2.3 Electron-hole asymmetry, correlated hopping and dynamic Hubbard models

After having been a devoted fan of the Hubbard model in the early stages of my physics career [41], I came to the conclusion 35 years ago that the Hubbard model Eq. (8) has a fundamental flaw: it is *electron-hole symmetric*. By that I mean, the properties of a system with n $electrons$ per site are identical to those of a system with n $holes$ per site, or equivalently (2-n) electrons per site. Around the same time, I came to the conclusion that electron-hole asymmetry is the key to superconductivity [42]. That moment marked my definitive divorce from the Hubbard model Eq. (8). Years later I attempted reconciliation with electron-hole $asymmetric$ versions of the Hubbard model, namely $dynamic$ Hubbard models [43], discussed later in this paper.

The fundamental electron-hole $asymmetry$ of condensed matter systems follows from the basic fact that the mass of the electron is 2000 smaller than the mass of the proton. It manifests itself for example in the fact that the mean inner potential of solids is necessarily positive [44, 45]. It renders the repulsive Hubbard model Eq. (8) irrelevant for the description of real systems [46]. Let's see why. Eq. (8) contains two energy scales, the hopping parameter t and the on-site repulsion U. But it excludes a third energy scale ϵ that is always in-between t and U, namely the spacing between atomic energy levels. When a second electron comes to occupy the orbital

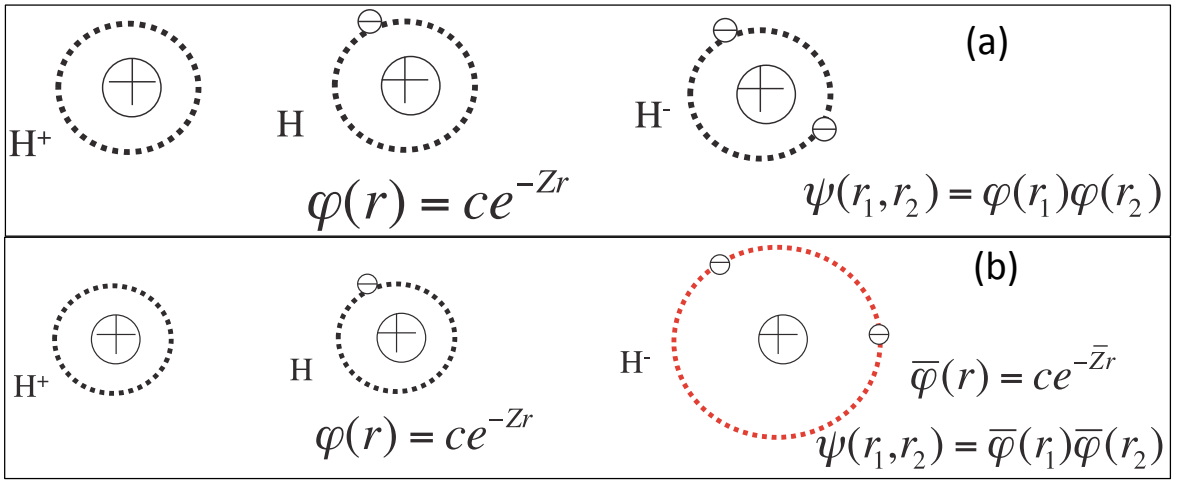

Fig. 1: *The top panel shows the atomic physics assumed in most models of interacting electrons such as the Hubbard model: the atomic orbital does not change with electron occupancy. The bottom panel shows the real physics: the atomic orbital expands when it is doubly occupied.*

already occupied by the first electron, the first electron doesn't sit idle to pay the very large Coulomb repulsion that would result from the second electron invading its turf. Instead, it (as well as the second electron) will expand their wavefunction, partially occupying higher energy single-particle states, thus reducing their Coulomb repulsion. This is shown schematically in Fig. 1. For electrons in the ground state of hydrogen-like ions of charge Z, the radius of the expanded orbital, or equivalently the value of $\bar{Z}$ is, within the Hartree approximation,

$$\bar{Z} = Z - 5/16. \tag{26}$$

This reduces the Coulomb repulsion between electrons, from $U = 17eV \times Z = (5/4)\epsilon_0 Z$, with $\epsilon_0 = 13.6eV$, to

$$\bar{U} = \frac{5}{4}\bar{Z} = U - \frac{25}{64}\epsilon_0 = U - 5.31eV \tag{27}$$

and reduces the kinetic energy of each electron, because of the orbital expansion, from $K = \epsilon_0 Z^2$ to $K = \epsilon_0 \bar{Z}^2$, while it increases the potential energy of each electron, because they are further away from the nucleus, from $-2\epsilon_0 Z^2$ to $-2\epsilon_0 Z\bar{Z}$, for a net energy reduction of

$$E(\bar{Z}) - E(Z) = -\frac{25}{128}\epsilon_0 = -2.66eV \tag{28}$$

so that the effective repulsion between electrons, defined as $U_{eff} = E(2) + E(0) - 2E(1)$, with $E(n)$ the energy of the ion with n electrons, is [43]

$$U_{eff} = U - \frac{25}{128}\epsilon_0 = U - 2.66eV. \tag{29}$$

From experimental values of ionization energies of hydrogen-like ions we find that

$$U_{eff} \sim U - 4.1eV, \tag{30}$$

the reduction is greater than Eq. (29) because of radial and angular correlations for electrons in the doubly occupied orbital that we didn't take into account .

The Hubbard model ignores the fundamental fact that the orbitals are modified upon double occupancy. In the Hubbard model, both the singly and the doubly occupied atomic orbitals are assumed to be single Slater determinants with the same single particle wavefunctions:

$$| \uparrow >= c_{\uparrow}^{\dagger}|0 > \tag{31a}$$

$$| \uparrow\downarrow >= c_{\uparrow}^{\dagger}c_{\downarrow}^{\dagger}|0 > \tag{31b}$$

so that

$$< 0|c_{\uparrow}| \uparrow >=<\downarrow |c_{\uparrow}| \uparrow\downarrow >= 1 \tag{31c}$$

which embodies the fundamental electron-hole $symmetry$ of the Hubbard Hamiltonian. However, this is qualitatively incorrect because the doubly occupied state is never a single Slater determinant but rather a linear combination of Slater determinants involving higher single electron states:

$$| \uparrow\downarrow >= \sum_{m,n} A_{mn} c_{m\uparrow}^{\dagger} c_{n\downarrow}^{\dagger}|0 > \tag{32a}$$

$$\sum_{m,n} |A_{mn}|^2 = 1 \tag{32b}$$

where the sum runs over a complete set of atomic orbitals with the lowest single particle orbital denoted by $m = 0$, i.e. $c_{0\sigma} = c_{\sigma}$, as well as over continuum states [47]. Eq. (32b) implies of course that $A_{mn} < 1$ for any m,n. Hence we have

$$c_{\uparrow}| \uparrow\downarrow >= \sum_{n} A_{0n} c_{n\downarrow}^{\dagger}|0 >= A_{00}| \downarrow > + \sum_{n\neq 0} A_{0n} c_{n\downarrow}^{\dagger}|0 > \tag{33}$$

and

$$1 =< 0|c_{\uparrow}| \uparrow >\neq<\downarrow |c_{\uparrow}| \uparrow\downarrow >= A_{00} < 1. \tag{34}$$

In other words, creating an electron into an empty orbital (or destroying an electron in the singly occupied orbital) is qualitatively different from creating a hole in the doubly occupied orbital (or creating an electron in the single occupied orbital).

For an electronic energy band that is close to empty, when electrons hop between sites transitions occur mostly between empty and singly occupied orbitals, so no other states are involved and the spectral function $A(k, \omega)$ will be a δ function with quasiparticle weight $z = 1$, $A(k, w) = \delta(\omega - \epsilon_k)$. Instead, for a band that is close to full, when a hole hops from a site to a neighboring site with no hole (i.e. doubly occupied), the final state can involve any of the atomic excited states at the two sites. The spectral function will have a quasiparticle part with quasiparticle weight $z < 1$, representing ground state to ground state transitions, and a broad incoherent part, as shown schematically in Fig. 2. We can describe this physics by writing the electron creation operator at site i as [48, 49]

$$c_{i\sigma}^{\dagger} = [1 + (S - 1)\tilde{n}_{i,-\sigma}]\tilde{c}_{i\sigma}^{\dagger} \tag{35}$$

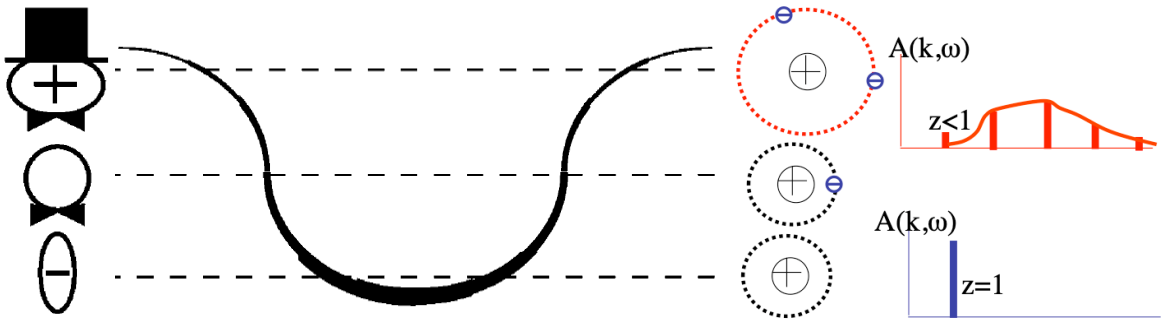

Fig. 2: *Electronic energy band ϵ_k versus k. Near the top of the band, most orbitals are doubly occupied and hence expanded. As the Fermi level moves up in the band, quasiparticles at the Fermi energy become increasingly dressed by the atomic electron-electron interaction, and increasingly heavier, and turn from electrons to holes. On the right we show schematically the spectral function evolving from a δ-function near the bottom of the band with quasiparticle weight z=1 to one with small quasiparticle weight $z < 1$ near the top of the band, with the rest of the spectral weight spread out in incoherent processes. The thickness of the curve giving the ϵ_k versus k relation indicates schematically the magnitude of the quasiparticle weight, which can be vanishingly small near the top of the band when the effective ionic charge is small.*

where $\tilde{c}^\dagger_{i\sigma}$ creates a quasielectron at site i, and $S =< \bar{\varphi}(r)|\varphi(r) >$ is the overlap matrix element between expanded and unexpanded orbitals shown in Fig. 1. The quasiparticle spectral weight as a function of band filling $0 < n_e < 2$ is

$$z(n_e) = [1 + (S-1)\frac{n_e}{2}]^2 \tag{36}$$

and goes from $z = 1$ for an almost empty band to $z = S^2$ for an almost full band. Alternatively, in terms of hole operators (denoted by the same symbols to avoid proliferation of symbols)

$$c^\dagger_{i\sigma} = [S + (1-S)\tilde{n}_{i,-\sigma}]\tilde{c}^\dagger_{i\sigma} \tag{37}$$

and the quasiparticle weight as function of hole concentration $n_h = 2 - n_e$ is

$$z(n_h) = [S + (1-S)\frac{n_h}{2}]^2 \tag{38}$$

The kinetic energy part of the Hamiltonian is

$$H_{kin} = -\sum_{ij,\sigma} t^\sigma_{ij}(c^\dagger_{i\sigma}c_{j\sigma} + h.c.) \tag{39}$$

and replacing the electron operators by their expression in terms of quasiparticle operators Eq. (35) we obtain

$$H_{kin} = -\sum_{ij,\sigma} t^\sigma_{ij}(\tilde{c}^\dagger_{i\sigma}\tilde{c}_{j\sigma} + h.c.) \tag{40}$$

with

$$t^\sigma_{ij} = t_{ij}[1 + (S-1)(\tilde{n}_{i,-\sigma} + \tilde{n}_{j,-\sigma}) + (S-1)^2\tilde{n}_{i,-\sigma}\tilde{n}_{j,-\sigma}]. \tag{41}$$

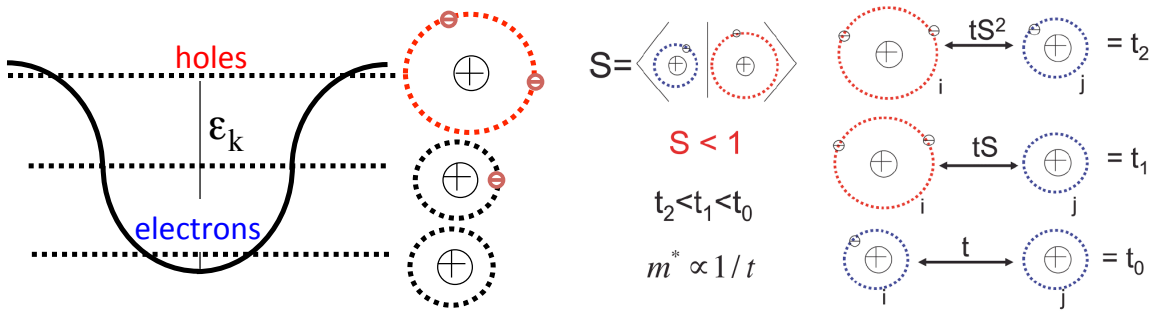

Fig. 3: *Hopping amplitudes for electrons at the Fermi energy as function of band occupation. As the number of electrons increases, hopping amplitudes are suppressed due to the modulation of the hopping by the overlap matrix element of expanded and unexpanded orbitals, S.*

Alternatively, in terms of hole operators, we have the same Eq. (40), with the hopping amplitudes given by

$$t_{ij}^{\sigma} = t_{ij}S^2[1 + (\frac{1}{S} - 1)(\tilde{n}_{i,-\sigma} + \tilde{n}_{j,-\sigma}) + (\frac{1}{S} - 1)^2\tilde{n}_{i,-\sigma}\tilde{n}_{j,-\sigma}]. \tag{42}$$

Fig. 3 shows the hopping amplitudes resulting from these equations.

The hopping amplitudes Eqs. (41) or (42) give rise to four-fermion and six-fermion terms in the Hamiltonian. In the presence of on-site repulsion, the six-fermion term will be irrelevant for the form Eq. (41) (Eq. (42)) when the Fermi level is close to the bottom (top) of the band. For the latter case, the Hamiltonian for hole operators is then

$$H = -\sum_{<ij>,\sigma} (\bar{t}_{ij} + \Delta t_{ij}(n_{i,-\sigma} + n_{j,-\sigma}))(c_{i\sigma}^{\dagger}c_{j\sigma} + h.c.) + U\sum_i n_{i\uparrow}n_{i\downarrow} \tag{43}$$

of the same form as Eq. (25), with $\bar{t}_{ij} = t_{ij}S^2$ and $\Delta t_{ij} = t_{ij}S(1 - S)$.

For small S, $\bar{t}_{ij}$ will be very small, giving rise to a narrow energy band, of width D that grows as the number of holes in the band increases. For a hypercubic lattice with nearest neighbor hopping only $\bar{t}_{ij} = \bar{t}$ the bandwidth is

$$D(n_h) = 2z\bar{t}(1 + n_h\Delta t) \tag{44}$$

with z the number of nearest neighbors to a site.

The physics discussed above is not present in the conventional Hubbard model. For those that cannot renounce the credo that the Hubbard model describes the essential physics of electron correlation in solids I offer"dynamic Hubbard models" [43] that embody the essential ubiquitous presence of electron-hole asymmetry. The physics is schematically shown in Fig. 4. We introduce a fictitious local boson displacement coordinate q_i for atom i that modulates the Hubbard U:

$$U(q_i) = U + \alpha q_i \tag{45}$$

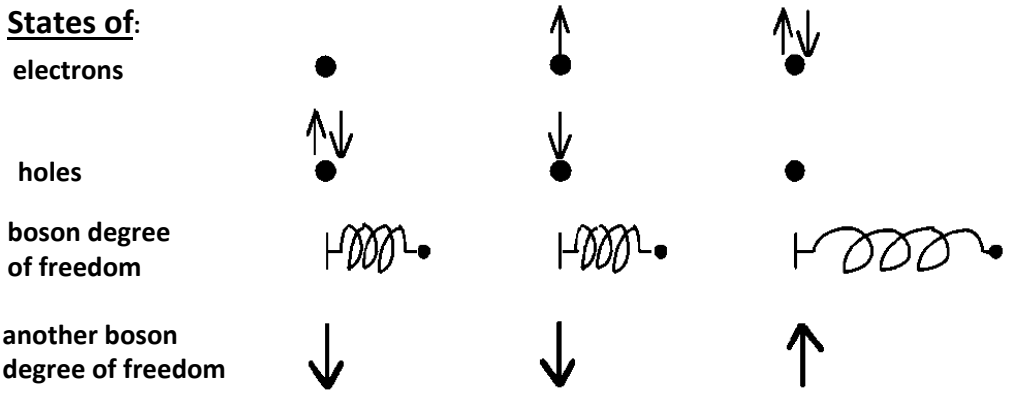

Fig. 4: *Dynamic Hubbard models, or electron-hole asymmetric polarons: a boson degree of freedom is associated with each site. The first electron at the site causes no change or a small change in the ground state of this degree of freedom, and the second electron causes a large change. For holes, the situation is reversed. Two examples of the boson degree of freedom are shown, an oscillator and a spin 1/2.*

that will relax when double occupancy occurs. As the simplest model we describe the boson dynamics by a harmonic oscillator of frequency $\omega_0 = \sqrt{K/M}$:

$$H_i = \frac{p_i^2}{2M} + \frac{1}{2}Kq_i^2 + (U + \alpha q_i)n_{i\uparrow}n_{i\downarrow} \tag{46}$$

The equilibrium position of the boson is $q_i = 0$ for the site empty or singly occupied and $q_i = -\alpha/K$ for the site doubly occupied. This embodies the physics of orbital expansion of the doubly occupied sites discussed above, and increasingly "dresses" the quasiparticles as the Fermi level goes up in the band. In terms of boson creation and annihilation operators $a_i^\dagger, a_i$ the site Hamiltonian and the effective Coulomb repulsion between electrons are

$$H_i = \omega_0 a_i^\dagger a_i + [U + g\omega_0(a_i^\dagger + a_i)]n_{i\uparrow}n_{i\downarrow} \tag{47a}$$

$$U_{eff} = U - \frac{\alpha^2}{2K} = U - \omega_0 g^2 \tag{47b}$$

with $g = \alpha/(2K\omega_0)^{1/2}$. The boson degree of freedom describes the electronic excitation of an electron when a second electron is added to the orbital. Hence the frequency ω_0 is related to the excitation energies of the atom, and we expect

$$\omega_0 = cZ^2 \tag{48}$$

where c is a constant of order eV, since the excitation energies in an atom scale with the square of the nuclear charge. From Eqs. (29), (47b) and (48) we conclude that

$$g^2 = \frac{c'}{Z^2}. \tag{49}$$

For a lattice system with hopping amplitude t_{ij} the Hamiltonian is then

$$H = -\sum_{ij,\sigma} t_{ij}(c_{i\sigma}^\dagger c_{j\sigma} + h.c.) + \sum_i [U + g\omega_0(a_i^\dagger + a_i))]n_{i\uparrow}n_{i\downarrow} + \sum_i \omega_0 a_i^\dagger a_i \tag{50}$$

Treating the four-fermion term in mean field, the electron-boson part of the Hamiltonian Eq. (50) is

$$H_{el-b} = g(n)\omega_0(a_i^\dagger + a_i)(n_{i\uparrow} + n_{i\downarrow}) \tag{51a}$$

$$g(n) = \frac{n}{2}g \tag{51b}$$

that is, an ordinary electron-boson coupling with a coupling constant that increases with band filling. Hence as in the usual electron-phonon interaction it will give rise to an effective mass enhancement and a quasiparticle weight reduction which increases as the band filling increases. Performing a generalized Lang-Firsov transformation on the fermion and boson operators [48] we obtain

$$c_{i\sigma} = e^{g(a_i^\dagger - a_i)\tilde{n}_{i,-\sigma}}\tilde{c}_{i\sigma} \equiv X_{i\sigma}\tilde{c}_{i\sigma} \tag{52a}$$

$$a_i = \tilde{a}_i - g\tilde{n}_{i\uparrow}\tilde{n}_{i\downarrow} \tag{52b}$$

and the Hamiltonian Eq. (50) becomes

$$H = -\sum_{ij,\sigma} t_{ij}(X_{i\sigma}^\dagger X_{j\sigma}\tilde{c}_{i\sigma}^\dagger\tilde{c}_{j\sigma} + h.c.) + \sum_i U_{eff}\tilde{n}_{i\uparrow}\tilde{n}_{i\downarrow} + \sum_i \omega_0\tilde{a}_i^\dagger\tilde{a}_i \tag{53}$$

with U_{eff} given by Eq. (47b). The ground state expectation value of the $X_{i\sigma}$ operator is

$$< X_{i\sigma} >_0 = e^{-(g^2/2)\tilde{n}_{i,-\sigma}} = 1 + (S-1)\tilde{n}_{i,-\sigma} \tag{54a}$$

$$S = e^{-g^2/2} \tag{54b}$$

The part of the fermion operator Eq. (52a) associated with ground state to ground state transitions of the boson field is the coherent part of the operator, the quasiparticle. We have then

$$c_{i\sigma} = |0><0|[1 + (S-1)\tilde{n}_{i,-\sigma}]\tilde{c}_{i\sigma} + c_{i\sigma}^{incoh} \tag{55}$$

where the coherent part was given in Eq. (35). $|0>$ denotes the ground state of the auxiliary boson. The incoherent part of the operator

$$c_{i\sigma}^{incoh} = [\tilde{n}_{i,-\sigma}\sum_{(l,l')\neq(0,0)} |l><l|e^{g(a_i^\dagger - a_i)}|l'><l'| + \sum_{l\neq 0}|l><l|]\tilde{c}_{i\sigma} \tag{56}$$

describes processes where the boson field makes transitions to and from excited states $|l>, l \neq 0$, which only take place if $\tilde{n}_{i,-\sigma} = 1$, that is if the orbital is occupied by another electron of opposite spin.

Replacing the ground state expectation values of the $X_{i\sigma}$ operators in the Hamiltonian Eq. (53), gives rise to the hopping amplitudes discussed earlier, Eq. (41). This will be accurate for $\omega >> t_{ij}$. The quasiparticle weight in this model is, from Eq. (55)

$$z(n) = [1 + \frac{n}{2}(S-1)]^2 \tag{57}$$

as was already given in Eq. (36), it decreases monotonically with electronic band filling n_e, $0 \leq n_e \leq 2$, so that quasiparticles become increasingly dressed as the band filling increases.

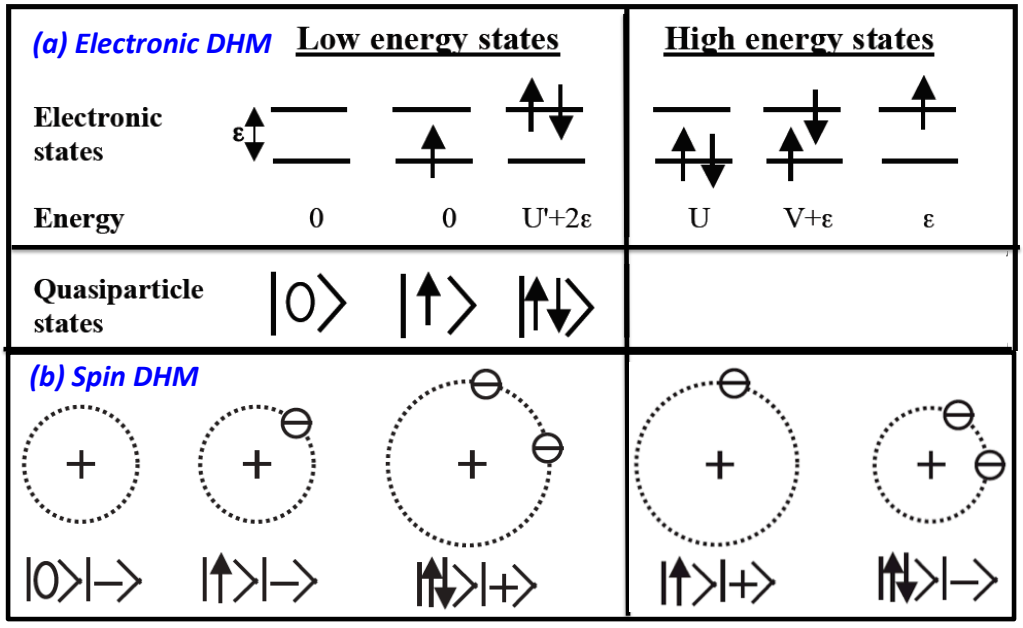

Fig. 5: *(a) Site states of an electronic dynamic Hubbard model (DHM) with two orbitals per site. (b) Site states of a dynamic Hubbard model with an auxiliary spin degree of freedom denoted by states* $|+>, |->$*. For both cases, as indicated in the center panel, the left three states are lowest in energy and are the quasiparticle states in the low energy effective Hamiltonian with the correlated hopping term Eq. (43).*

The factor S is the overlap matrix element of the oscillator ground states with and without site double occupancy [48], and S^2 gives the quasiparticle weight for a hole in the filled band ($n = 2$ in Eq. (57)). According to Eqs. (49) and (54b), as the ionic charge Z decreases S decreases rapidly, implying that hole quasiparticles become increasingly incoherent.

We can estimate S from first principles for a hydrogen-like ion. In the Hartree approximation, S will be given by the overlap matrix element of the electron wave function in the presence and in the absence of another electron in the orbital:

$$S = | < \varphi_{1s} | \bar{\varphi}_{1s} > | = \frac{(1 - \frac{5}{16Z})^{3/2}}{(1 - \frac{5}{32Z})^3} \tag{58}$$

with $\bar{\varphi}_{1s}$ the 1s orbital Eq. (1b) with Z replaced by $\bar{Z} = Z - 5/16$, as appropriate for the Hartree wavefunction [43]. Better estimates can be obtain with other more accurate approximations to the two-electron wavefunction [50].

Other forms of dynamic Hubbard models have also been proposed and studied, where the auxiliary boson degree of freedom is a spin 1/2 instead of an oscillator, as well as one with purely electronic degrees of freedom involving two orbitals per site [50, 51]. The low and high energy states in these models are shown in Fig. 5. The essential physics is always the same and leads to the correlated hopping Hamiltonian Eq. (43) in the limit where the boson excitation energies are large. Dynamic Hubbard models have been studied numerically using exact diagonalization [51–53], quantum Monte Carlo [52, 54] and dynamical mean field theory [55]. The

effect of finite boson frequency was studied analytically in Ref. [56] using Eliashberg theory. Both from exact diagonalization for small clusters [52] and analytically [56] it was found that finite frequencies enlarges the parameter regime that gives rise to pairing and superconductivity relative to that in the antiadiabatic limit Eq. (66).

2.4 Superconductivity in dynamic Hubbard models

The dynamic Hubbard models discussed in the previous section describe increased dressing of the quasiparticles when the Fermi level goes up in the band. When the Fermi level is close to the top of the band, and carriers pair, it means that locally the band filling decreases, hence carriers partially undress. This provides a mechanism for superconductivity [57,58], driven by lowering of kinetic energy [59]. The undressing gives rise to experimental signatures, in particular transfer of spectral weight from high to low frequencies in both the single and two-particle spectral function [49,53,60], that can be detected experimentally in photoemission [61] and optical [62] experiments respectively. This physics leads to an apparent violation of the conductivity sum rule, that was predicted theoretically in 1992 [63,60] and first observed experimentally 10 years later [62]. The fact that the superconductivity mechanism is tied to electron-hole asymmetry also gives rise to experimental signatures, in particular a tunneling asymmetry of universal sign, predicted theoretically in 1989 [64] and first observed experimentally around 1995 and thereafter [65].

We focus on the low energy physics that results from the correlated hopping Hamiltonian Eq. (21), with the addition of off-site Coulomb repulsion $V_{ij}n_in_j$. In momentum space the Hamiltonian is, in hole representation

$$H = \sum_{k\sigma}(\epsilon_k - \mu)c^\dagger_{k\sigma}c_{k\sigma} + \frac{1}{N}\sum_{kk'q}(V(q) + \alpha(\epsilon_k + \epsilon_{k+q} + \epsilon_{k'} + \epsilon_{k'-q}))c^\dagger_{k+q\uparrow}c^\dagger_{k'-q\downarrow}c_{k'\downarrow}c_{k\uparrow} \quad (59a)$$

$$V(q) = \sum_j e^{iqR_j}V_{0j} \quad (59b)$$

with $V_{00} = U$, and $\alpha = \Delta t_{ij}/\bar{t}_{ij}$. Assuming only nearest neighbor hopping and only on-site and nearest-neighbor repulsion, we write the pair interaction in the BCS reduced Hamiltonian Eq. (4) as

$$V_{kk'} = V(\epsilon_k, \epsilon_{k'}) = U + \frac{K}{D/2}(\epsilon_k + \epsilon_{k'}) + \frac{W}{D/2}\epsilon_k\epsilon_{k'} \quad (59c)$$

with the bandwidth $D = 2z\bar{t}$, with z the number of nearest neighbors to a site, $K = 2z\Delta t$ and $W = zV$. We have left out some terms in $V(k-k')$ that are odd under $k \to -k$ or $k' \to -k'$ that drop out in the subsequent development. Note that everything depends on kinetic energy rather than momentum, hence the resulting gap function will obey $\Delta_k = \Delta(\epsilon_k)$, and in particular will be constant over the Fermi surface. The usual BCS gap equation is

$$\Delta_k = -\frac{1}{N}\sum_{k'} V(\epsilon_k, \epsilon_{k'})\Delta_{k'}\frac{tanh(\frac{\beta E_{k'}}{2.})}{2E_{k'}} \quad (60)$$

with

$$E_k = \sqrt{(\epsilon_k - \mu)^2 + \Delta_k^2}. \tag{61}$$

From the form of $V(\epsilon_k, \epsilon_{k'})$ it follows that $\Delta_k = \Delta(\epsilon_k)$ is a linear function of ϵ_k, which we parametrize as

$$\Delta(\epsilon_k) = \Delta_m(-\frac{\epsilon_k}{D/2} + c) \tag{62}$$

and replacement of Eq. (62) in Eq. (60) yields the following two equations [57, 58] :

$$1 = K(I_1 + cI_0) - W(I_2 + cI_1) \tag{63a}$$

$$c = K(I_2 + cI_1) - U(I_1 + cI_0) \tag{63b}$$

with

$$I_\ell = \frac{1}{N}\sum_k (-\frac{\epsilon_k}{D/2})^\ell \frac{tanh(\frac{\beta E_{k'}}{2.})}{2E_{k'}} \tag{64}$$

These equations are solved numerically for Δ_m and c as function of temperature and band filling determined by μ. To obtain the critical temperature, a single equation needs to be solved, obtained by combining Eqs. (63a) and (63b)

$$1 = 2KI_1 - WI_2 - UI_0 + (K^2 - WU)(I_0 I_2 - I_1^2) \tag{65}$$

with $E_k = |\epsilon_k - \mu|$ in the formulas for I_ℓ. Eq. (65) will have a solution, i.e. give rise to superconductivity, when the parameters in the Hamiltonian Eq. (59) satisfy the condition [36]

$$k > \sqrt{(1+u)(1+w)} - 1 \tag{66}$$

with $u = gU, w = gW, k = gK$, with g the density of states at the Fermi energy.

Fig. 6 (a) shows the typical behavior for critical temperature versus hole concentration resulting from this model for a set of parameters appropriate to describe cuprate superconductors. In Fig. 6 (b) we show the results for the parameter Δt versus interatomic distance for various values of the effective ionic charge Z, obtained from an approximate first-principles calculation for a diatomic molecule [66]. It can be seen that Δt is larger for negatively charged ions ($Z < 2$) in close proximity.

The parameters for Fig. 6 (a) are $U = 5eV$, $\bar{t} = 0.03eV$, $\Delta t = 0.1875eV$. The bare bandwidth is $3.24eV$ when the Fermi level is near the bottom of the band, but when it is near the top it is narrowed to $0.24eV$ since $\bar{t} = t - 2\Delta t$. Superconductivity only occurs near the top of the band, with the characteristic dome-type structure seen in the cuprates as well as in other materials like in the transition metal series [67]. The figure also shows that the effective mass decreases as the Fermi level moves down in the band, and the superconducting coherence length increases. There is a cross-over between strong and weak coupling regimes as the hole concentration increases, as seen in the cuprate superconductors.

From Eqs. (62) and (63), we find that the quasiparticle excitation energy in the superconducting state is given by

$$E_k = \sqrt{a^2(\epsilon_k - \mu - \nu)^2 + \Delta_0^2} \tag{67a}$$

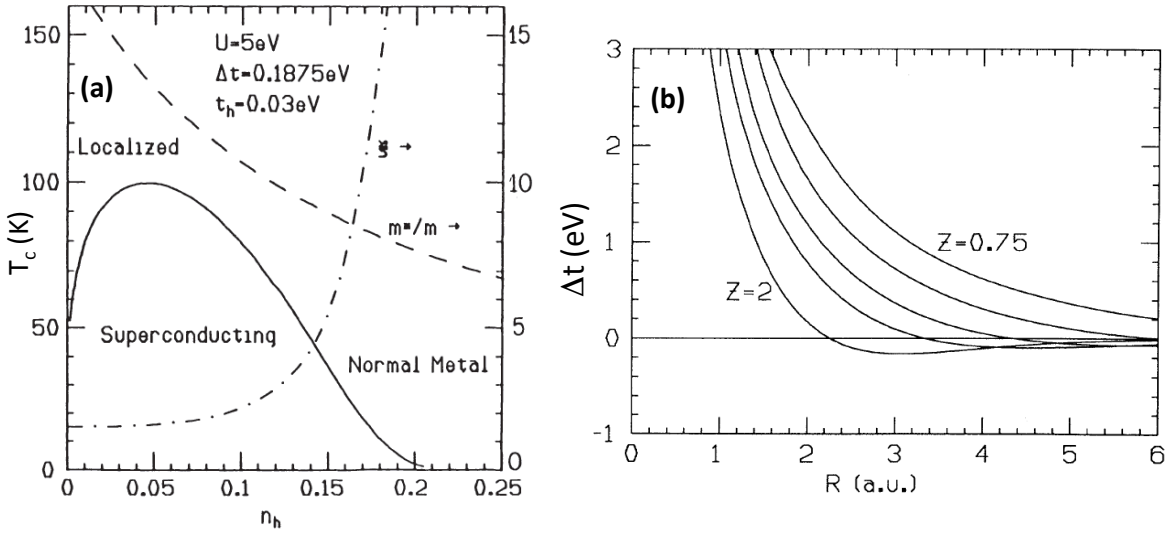

Fig. 6: *The left panel shows the typical behavior of critical temperature versus hole concentration in this model for one set of parameters appropriate for high T_c cuprates; the behavior of the coherence length versus hole concentration in units of lattice spacing and of the ratio of effective mass to band mass is shown on the right-hand scale ($m^*/m = .(\bar{t}+2\Delta t)/(\bar{t}+n_h\Delta t)$). The right panel shows the behavior of the parameter Δt as function of interatomic distance and various values of the effective ionic charge Z (Z=0.75, 1, 1.25, 1.5, 2).*

$$a = [1 + (\frac{\Delta_m}{D/2})^2]^{1/2} \tag{67b}$$

$$\Delta_0 = \frac{\Delta(\mu)}{a} \tag{67c}$$

$$\nu = \frac{1}{a}\frac{\Delta_m}{D/2}\Delta_0 \tag{67d}$$

This is shown in Fig. 7 (a). It can be seen that the minimum in the E_k versus ϵ_k relation is not at the chemical potential, as in usual BCS, because of the energy dependence of the gap. Instead, it is shifted to $\mu + \nu$. Thus, the quasiparticle excitations are not charge-neutral as in usual BCS, they are positively charged. The behavior of the asymmetry parameter ν versus temperature is shown in Fig. 7 (b). Both the gap slope $\Delta_m/(D/2)$ and the gap Δ_0 vanish at T_c as the square root of $(T_c - T)$, hence ν goes linearly to zero at T_c.

The parameter ν gives rise to the asymmetry of universal sign in S-I-N tunneling mentioned earlier. It also gives rise to $positive$ thermoelectric power of tunnel junction: under open circuit conditions, the thermoelectric voltage is predicted to be [68]

$$V_t = \frac{\nu}{e}\frac{T_s - T_n}{T_n} \tag{68}$$

with T_s and T_n the temperatures in the normal and superconducting sides of the junction respectively. Thus, this effect provides a direct measure of the fundamental asymmetry parameter ν, or equivalently of the gap function slope $\Delta_m/(D/2)$, given an independent estimate of the gap Δ_0. The parameter ν is expected to be of order meV for cuprate superconductors and μeV for

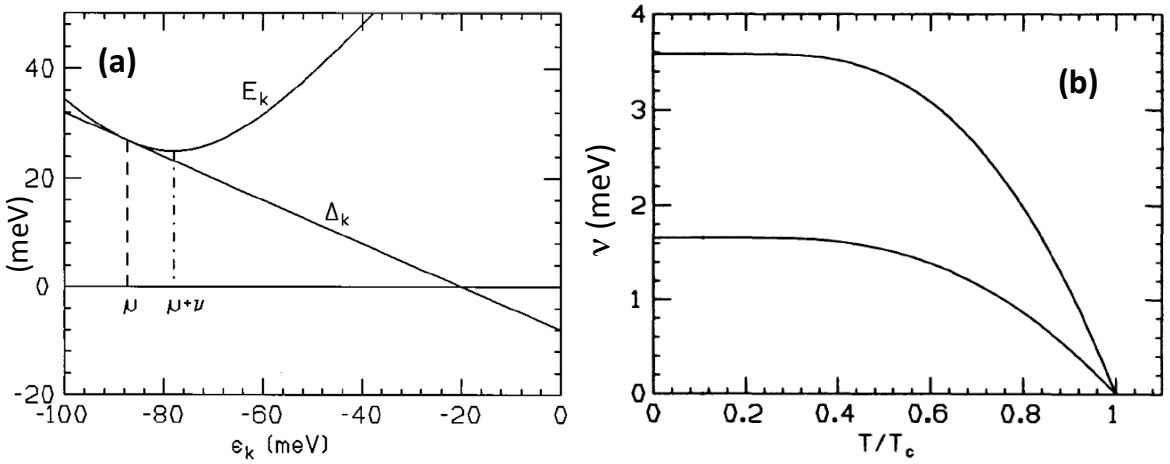

Fig. 7: *(a) shows the gap function and the quasiparticle energy as function of kinetic energy in a hole representation. Note that the gap function changes sign as ϵ_k increases. The minimum in the quasiparticle energy is shifted from the chemical potential μ by the asymmetry parameter ν. (b) shows the temperature dependence of ν for two sets of parameter values. Lower curve: $U = 5eV, K = 3.61eV, W = 2.24eV, D_h = 0.24eV$. Upper curve: $U = 5eV, K = 3.78eV, W = 2.60eV, D_h = 0$. $K = 2z\Delta t, W = zV, z = 4$. z is the number of nearest neighbors to a site.*

conventional superconductors. For quasiparticle tunneling between two superconductors A and B the thermoelectric voltage under open circuit conditions is

$$V_{AB} = \frac{\nu_A + \nu_B}{e} \frac{T_B - T_A}{(T_A + T_B)/2} \tag{69}$$

neglecting a small correction of order $\nu_A \nu_B / \Delta_{0A} \Delta_{0B}$. These predictions have not been experimentally tested.

3 Charge expulsion and alternative London electrodynamics

In the models discussed in the previous section, charge asymmetry plays an essential role. Superconductivity only occurs when the band is almost full, i.e. a lot of $negative$ electrons are present. The essential physics at the atomic level is that the doubly occupied orbital expands, hence negative charge moves outward, driven by both lowering of Coulomb repulsion and lowering of quantum kinetic energy. The pairing interaction Δt is larger when the atoms involved are $negatively$ charged anions, i.e. the effective nuclear charge Z is small (Fig. 6b), in which case the atomic orbital expansion is larger. Having fewer electrons in the vicinity of a given electron allows it to hop with larger hopping amplitude, hence its kinetic energy decreases. All of this suggests that systems governed by this physics will have a tendency to $expel$ electrons from their interior [69]. That is indeed what they do, and it has fundamental consequences.

3.1 Charge expulsion in dynamic Hubbard models

As was shown in Fig. 7 (a), the quasiparticle excitation energy E_k is not symmetric around the chemical potential due to the energy dependence of the gap. The BCS coherence factors are given by the usual form

$$u_k^2 = \frac{1}{2}(1 + \frac{\epsilon_k - \mu}{E_k}) = \frac{1}{2}(1 + \frac{\epsilon_k - \mu - \nu}{E_k}) + \frac{\nu}{2E_k} \tag{70a}$$

$$v_k^2 = \frac{1}{2}(1 - \frac{\epsilon_k - \mu}{E_k}) = \frac{1}{2}(1 - \frac{\epsilon_k - \mu - \nu}{E_k}) - \frac{\nu}{2E_k} \tag{70b}$$

and as a consequence quasiparticles are positively charged on average. The net quasiparticle charge per site is given by

$$Q^* = \frac{2}{N}\sum_k (u_k^2 - v_k^2) f(E_k) = 2\nu \frac{1}{N}\sum_k \frac{1}{E_k}. \tag{71}$$

As a consequence of quasiparticles being positively charged, the condensate will acquire an extra negative charge.

Hence the superconductor is characterized by having two different 'chemical potentials'. The chemical potential μ corresponds to the condensate, and $\mu' = \mu + \nu$ to the quasiparticle excitations. In a hole representation, $\mu' > \mu$, in an electron representation $\mu' < \mu$. The negatively charged condensate, by virtue of being a superfluid as well as because of the effective mass reduction that occurs due to pairing and undressing, is highly mobile, in contrast to the quasiparticles which experience normal scattering and have the higher effective mass characteristic of the normal state dressed carriers. As a consequence, one expects that the negative condensate will have a tendency to move out of the bulk of the superconductor, so as to tend to equate the chemical potentials μ and μ' in the bulk. Because of overall charge neutrality, the negative charge will accumulate near the surface of the superconductor.

An estimate of the maximum amount of charge that will be expelled from the bulk of the superconductor is given by the ratio of the difference in chemical potentials to the bandwidth D:

$$n_{max} = \frac{2(\mu' - \mu)}{D} = \frac{2\nu}{D} \tag{72}$$

carriers per site, so it is very small. However, the tendency to charge expulsion will be counteracted by Coulomb charging energy.

That this physics takes place is confirmed by numerical analysis of the underlying Hamiltonian [70]. We consider the Hamiltonian in the hole representation

$$H = -\sum_{ij\sigma} t_{ij}^{\sigma}[c_{i\sigma}^{\dagger} c_{j\sigma} + h.c.] + U\sum_i n_{i\uparrow} n_{i\downarrow} \tag{73a}$$

$$t_{ij}^{\sigma} = t_h + \Delta t(n_{i,-\sigma} + n_{j,-\sigma}) + \Delta t_2 n_{i,-\sigma} n_{j,-\sigma} \tag{73b}$$

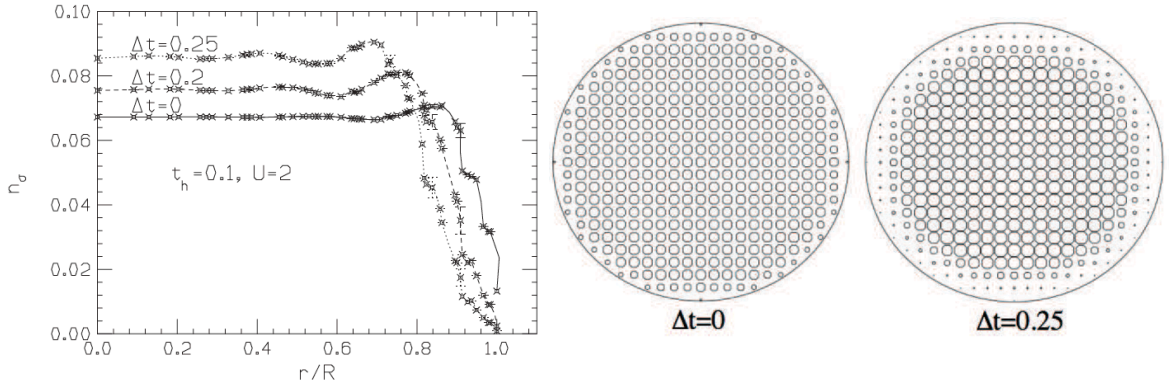

Fig. 8: *Left panel: hole site occupation per spin n_σ for a cylinder of radius R=11 as a function of r/R, with r the distance to the center, for a cubic lattice of side length 1. There are 377 sites in a cross-sectional area ($\pi R^2 = 380.1$). The average occupation for both spins is $n = 0.126$ holes per site. Parameters in the Hamiltonian are shown on the left panel. $\Delta t_2 = 0$. On the right panel we show the results for two cases from the left panel, representing the hole occupation at the site with circles of diameter proportional to it. Note that for finite Δt the hole occupation increases in the interior and is depleted near the surface, leading to charge inhomogeneity with excess negative charge near the surface and excess positive charge in the interior, relative to a neutralizing background of charge density n.*

with $t_h = tS^2$ $\Delta t = tS(1-S)$, $\Delta t_2 = t(1-S)^2 = (\Delta t)^2/t_h$. The fact that the hopping amplitudes Eq. (73b) increase with hole occupation suggests that the system will have a tendency to *expel electrons* from its interior to the surface, because the coordination of sites in the interior is larger than of sites at the surface. This is indeed what we find numerically. We assume a cylindrical geometry of radius R and infinite length in the z direction, and decouple the interaction terms within a simple mean field approximation assuming $< n_{i\sigma} >= n_i/2$ with n_i the hole occupation at site i, yielding the mean field Hamiltonian

$$\begin{aligned} H_{mf} = & - \sum_{<ij>,\sigma} [t_h + \Delta t n_i + \Delta t_2 \frac{n_i^2}{4}][c_{i\sigma}^\dagger c_{j\sigma} + h.c.] + \frac{U}{4}\sum_i n_i^2 \\ & - \sum_{<ij>} n_i[\Delta t + \frac{n_j}{2}\Delta t_2]\sum_\sigma < c_{i\sigma}^\dagger c_{j\sigma} > \end{aligned} \tag{74}$$

Note that the local average bond occupation modifies the local chemical potential. Assuming a band filling of n holes per site, we diagonalize the Hamiltonian Eq. (74) on a finite lattice with initial values $n_i = n$ and fill the lowest energy levels until the occupation n is achieved. From the Slater determinant of that state we obtain new values of n_i for each site and for the local bond occupation, and iterate this procedure until self-consistently is achieved. We then examine the resulting occupation of the sites as function of the distance r to the center of the cylinder. Figure 8 shows a typical example of the behavior found. Here we assumed $\Delta t_2 = 0$, corresponding to the simpler Hubbard model with correlated hopping and no six-fermion operator term. Even for $\Delta t = 0$ the hole occupation is somewhat larger in the interior than near the surface. When the interaction Δt is turned on, the hole occupation increases in the interior and decreases near the surface. This indicates that the system expels electrons from the interior to the surface.

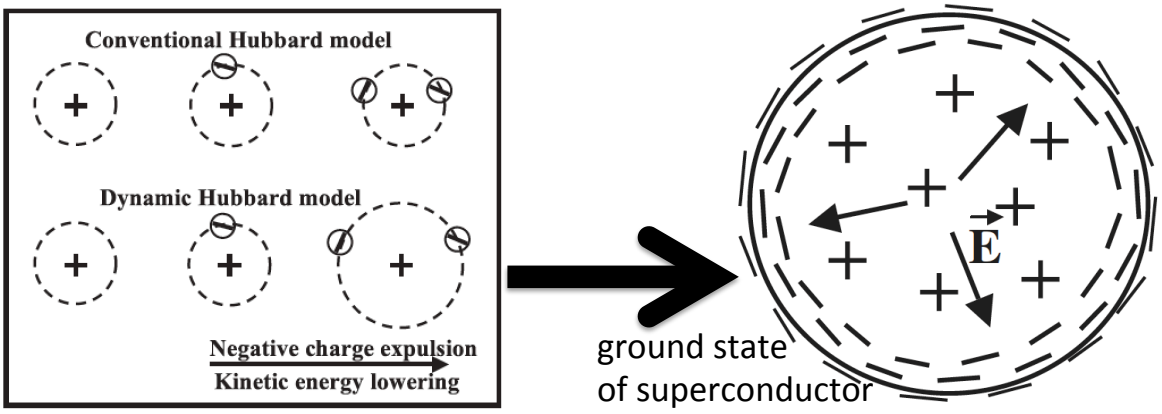

Fig. 9: *The dynamic Hubbard model describes negative charge expulsion and kinetic energy lowering at the atomic level, due to orbital expansion (left panel). In a lattice system it leads to expulsion of negative charge from the interior to the surface. In the superconducting state, a macroscopic charge inhomogeneity develops as shown on the right panel, with an electric field $\vec{E}(\vec{r})$ in the interior that points towards the surface.*

The effect becomes more pronounced when Δt is increased or t_h is decreased. Finite Δt_2 enhances the effect. For larger values of the parameters the system develops a tendency to phase separation, where holes condense in the interior and the outer region of the cylinder has no holes. The instability condition for phase separation can be found analytically.

Of course in a real material this tendency to charge segregation will be countered by longer range Coulomb repulsion. But it is clear that this physics will cause a tendency to develop local charge inhomogeneity: because kinetic energy dominates the physics of the dynamic Hubbard model, the system will develop charge inhomogeneity at a cost in potential energy if it can thereby lower its kinetic energy more, unlike systems where the dominant physics is potential-energy driven like the conventional Hubbard model. High T_c cuprates, for which Δt should be large, show a strong tendency to charge inhomogeneity. .

3.2 Electric fields in superconductors and alternative London charge electrodynamics

The physics discussed in the previous sections leads to the prediction that when a system goes superconducting it will expel electrons from the interior to the surface, resulting in macroscopic charge inhomogeneity, as shown on the right panel of Fig. 9. In the normal state this cannot occur, the tendency to charge expulsion is countered by the Coulomb repulsion and no electric field can exist in the interior of a normal metal. That minimizes the potential energy. However the superconducting state is a macroscopic quantum state, where the sum of potential and quantum kinetic energy need to be minimized. Just like in the microscopic atom the charge distribution is inhogeneous, with the negative charge more "spread out" than the positive charge to lower its quantum kinetic energy, the same will be true in the superconductor, which is in some sense a "giant atom". This is also suggested by the fact that the superfluid condensate is

described by a macroscopic quantum wavefunction $\psi(\vec{r}) = |\psi(\vec{r})|e^{i\theta(\vec{r})}$ [11, 12], just like the single electron in the hydrogen atom. This results in the existence of an electric field in the interior of superconductors, just like in the hydrogen atom, as shown on the right panel of Fig. 9.

The conventional London equations do not allow for electric fields inside superconductors. However a simple modification of them does [1, 71, 72]. The London equation is derived as shown in Eq. (75):

$$\frac{\partial \vec{J}}{\partial t} = \frac{n_s e^2}{m_e}\vec{E} \rightarrow \vec{\nabla} \times \vec{J} = -\frac{n_s e^2}{m_e c}\vec{B} = -\frac{c}{4\pi\lambda_L^2}\vec{B}, \frac{1}{\lambda_L^2} = \frac{4\pi n_s e^2}{m_e c^2} \rightarrow \nabla^2 \vec{B} = \frac{1}{\lambda_L^2}\vec{B} \quad (75)$$

The Eq. on the left describes the colisionless response of a conducting fluid of density n_s to an applied electric field $\vec{E}$, i.e. free acceleration of superfluid carriers of charge e and mass m_e, giving rise to the supercurrent $\vec{J} = n_s e\vec{v}$, with $\vec{v}$ the carrier velocity. Upon application of the curl on both sides, using Faraday's law and integrating over the time derivatives, the second equation results, called the London equation. From applying the curl to both sides of Ampere's law $\vec{\nabla} \times \vec{B} = (4\pi/c)\vec{J}$ and replacing $\vec{\nabla} \times \vec{J}$ in Eq. (75), the right-hand side expression results, that predicts that the magnetic field decays exponentially over a distance λ_L in going from the surface towards the interior of the superconductor.

We note however that the London equation can be written as

$$\vec{\nabla} \times \vec{J} = -\frac{c}{4\pi\lambda_L^2}\vec{B} \rightarrow \vec{J} = -\frac{c}{4\pi\lambda_L^2}\vec{A} \quad (76)$$

with $\vec{A}$ the magnetic vector potential, $\vec{\nabla} \times \vec{A} = \vec{B}$. Faraday's law, upon integration, leads to

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \rightarrow \vec{E} = -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \quad (77)$$

where ϕ is the electric potential. Taking the time derivative of the right hand side of Eq. (76) and using Eq. (77) leads to

$$\frac{\partial \vec{J}}{\partial t} = \frac{n_s e^2}{m_e}(\vec{E} + \vec{\nabla}\phi) \quad (78)$$

which, unlike the left-hand-side of Eq. (75), allows for the presence of an electric field that derives from a potential that will not give rise to an infinite current . Note that the left-hand-side of Eq. (75) is derived from Newton's equation by replacing the total time derivative by the partial time derivative, which is not correct.

Note that the right-hand-side of Eq. (76) relates the electric current density, a physical quantity, to the magnetic vector potential, that is gauge-dependent. Assuming different gauges for $\vec{A}$ in Eq. (76) leads to different physics. The London brothers assumed that $\vec{\nabla} \cdot \vec{A} = 0$, the "London gauge", which has as a consequence that no electric fields can exist in the interior of superconductors. But that was just an unproven assumption. Instead, we will assume that $\vec{A}$ obeys the Lorentz gauge, as was also done in the first London paper [1]

$$\vec{\nabla} \cdot \vec{A} = -\frac{1}{c}\frac{\partial \phi}{\partial t}. \quad (79)$$

Upon taking the divergence of both sides of the right-hand-side of Eq. (76), and using the continuity equation $\vec{\nabla}\cdot\vec{J} = -\partial\rho/\partial t$ with ρ the charge density and the gauge condition Eq. (79) we obtain

$$\vec{J} = -\frac{c}{4\pi\lambda_L^2}\vec{A} \rightarrow \frac{\partial\rho}{\partial t} = -\frac{1}{4\pi\lambda_L^2}\frac{\partial\phi}{\partial t} \tag{80}$$

and integrating with respect to time to

$$\phi(\vec{r},t) - \phi_0(\vec{r}) = -4\pi\lambda_L^2(\rho(\vec{r},t) - \rho_0(\vec{r})) \tag{81}$$

where $\phi_0(\vec{r})$ and $\rho_0(\vec{r})$ are constants of integration. A possible choice would be $\phi_0 = \rho_0 = 0$ [1]. Instead, motivated by the physics discussed in the previous sections, we choose $\rho(\vec{r}) = \rho_0 > 0$, that is, a uniform positive charge density in the interior of the superconductor. This then implies that the electrostatic potential $\phi(\vec{r},t)$ equals $\phi_0(\vec{r})$ when the charge density in the interior of the superconductor is constant, uniform, and equal to ρ_0. From Maxwell's equations we deduce that $\phi_0(\vec{r})$ is given by

$$\phi_0(\vec{r}) = \int_V d^3r' \frac{\rho_0}{|\vec{r}-\vec{r'}|} \tag{82}$$

where the integral is over the volume of the superconducting body. ρ_0 is a function of the material, the temperature, and the volume and shape of the superconducting body. Before discussing its value, we discuss some consequences of these equations.

In the absence of time dependence, the electrostatic equations can be solved analytically for simple geometries (sphere, cylinder, plane) and numerically for other geometries. For example, for a sphere of radius R we obtain for the charge density and electric field

$$\rho(r) = \rho_0[1 - \frac{1}{3}\frac{R^3}{\lambda_L^2 r}\frac{sinh(r/\lambda_L)}{f(R/\lambda_L)}]; \vec{E}(r) = \frac{4}{3}\pi\rho_0\vec{r}[1 - \frac{R^3}{r^3}\frac{f(r/\lambda_L)}{f(R/\lambda_L)}] \tag{83}$$

with $f(x) = xcoshx - sinhx$. Within a layer of thickness λ_L from the surface there is excess negative charge density $\rho_- = -R/(3\lambda_L)\rho_0$. The electric field grows linearly with distance from the center of the sphere, peaks at distance λ_L from the surface, with peak value $E_m = -4\pi\lambda_L\rho_-$, decays to zero at $r = R$, and is of course zero for $r > R$. For a long cylinder of radius R, the peak value of the electric field E_m at distance λ_L from the surface is given by the same expression in terms of ρ_-, and $\rho_- = -(R/(2\lambda_L)\rho_0$. The expressions for the charge density and electric field as function of r involve Bessel functions of imaginary argument.

For more general geometries, the electric field will "leak out" from the interior and be non-zero outside the superconducting sample. In particular, ellipsoidal samples give rise to quadrupolar electric fields in the exterior [73, 74]. Figure 10 shows examples of field lines and position dependence of charge density and electric field in the interior of an ellipsoidal sample. Note that the electric field lines outside the samples go out from regions of lower surface curvature and go into regions of higher surface curvature. This is easy to understand qualitatively. As we will discuss later, there are spin currents flowing near the surface of the superconducting body. In the regions of high curvature (low curvature) electrons slow down (speed up), just like racing cars would, so their kinetic energy decreases (increases) and consequently their potential energy

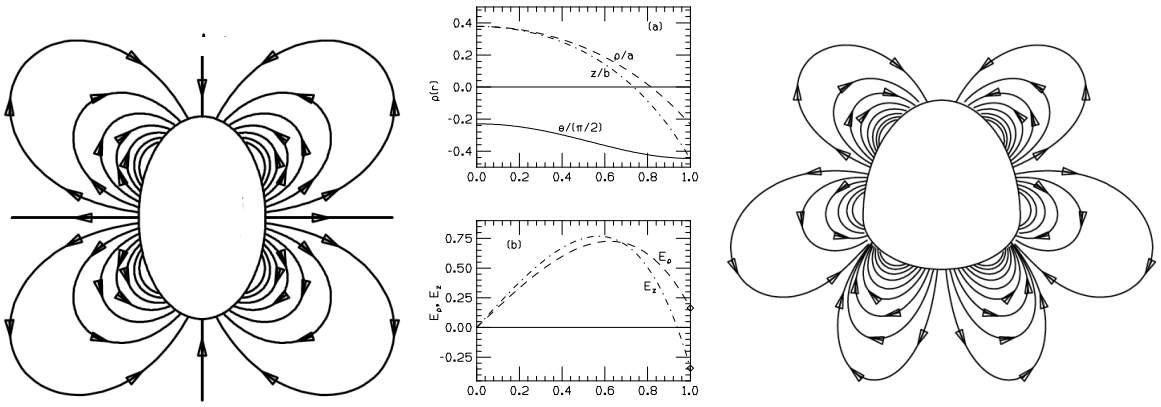

Fig. 10: *Results from numerical solution of the electrostatic equations. Left panel: electric field lines in the exterior of a sample of ellipsoidal shape, of dimensions* $a = 1, b = 1.5$. *London penetration depth is* $\lambda_L = 0.5$ *and* $\rho_0 = 0.1$. *Middle panel top: Charge density in the interior of the superconductor along the horizontal axis plotted versus* ρ/a *(curve labeled* ρ/a *and along the vertical axis plotted versus* z/b *(curve labeled* z/b*) and on the boundary plotted versus* $\theta/(\pi/2)$*, with* $\theta = tan^{-1}(z/b)/(\rho/a)$*; note that the negative charge density near the surface is larger in magnitude along the z direction. Middle panel bottom: Electric fields in the interior along the* ρ/a *and* z/b *(dot-dashed) directions. Note that the electric field along the z direction changes sign near the surface, and that the electric fields are finite at the surface. Right panel: Electric field lines in the exterior of a sample of egg-like shape*

increases (decreases), to keep the total energy constant. Higher (lower) potential energy for the electron means lower (higher) electric potential, and electric field lines go from high potential to low potential.

From Eq. (81) and Maxwell's equations, we deduce that electric and magnetic fields, charges and currents in superconductors, obey the following equations:

$$\nabla^2 \vec{B} = \frac{1}{\lambda_L^2}\vec{B} + \frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} \tag{84a}$$

$$\nabla^2 (\vec{E} - \vec{E}_0) = \frac{1}{\lambda_L^2}(\vec{E} - \vec{E}_0) + \frac{1}{c^2}\frac{\partial^2 (\vec{E} - \vec{E}_0)}{\partial t^2} \tag{84b}$$

$$\nabla^2 \vec{J} = \frac{1}{\lambda_L^2}\vec{J} + \frac{1}{c^2}\frac{\partial^2 \vec{J}}{\partial t^2} \tag{84c}$$

$$\nabla^2 (\rho - \rho_0) = \frac{1}{\lambda_L^2}(\rho - \rho_0) + \frac{1}{c^2}\frac{\partial^2 (\rho - \rho_0)}{\partial t^2} \tag{84d}$$

so that all quantities obey exactly the same equation.

The simplicity of eqs. (84) derives from the fact that the theory is relativistically covariant. We define the current 4-vector and the four-vector potential in the usual way

$$J = (\vec{J}(\vec{r}, t), ic\rho(\vec{r}, t)) \quad ; \quad A = (\vec{A}(\vec{r}, t), i\phi(\vec{r}, t)) \tag{85}$$

The continuity equation sets the four-dimensional divergence of the four-vector J equal to zero, where the fourth derivative is $\partial/\partial(ict)$, and the Lorenz gauge condition sets the divergence of the four-vector A to zero

$$DivJ = 0 \quad ; \quad DivA = 0. \tag{86}$$

Furthermore we define the four-vectors associated with the positive uniform charge density ρ_0 and its associated current $\vec{J_0}$, denoted by J_0, and the associated four-vector potential A_0. In the frame of reference where the superconducting body is at rest the spatial part of these four-vectors is zero, hence

$$J_0 = (0, ic\rho_0) \quad ; \quad A_0 = (0, i\phi_0(\vec{r})) \tag{87}$$

in that reference frame. In any inertial reference frame, A_0 and J_0, as well the four-vectors J and A obey

$$\Box^2 A_0 = -\frac{4\pi}{c} J_0 \quad ; \quad \Box^2 A = -\frac{4\pi}{c} J \tag{88}$$

with the d'Alembertian operator

$$\Box^2 = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}.. \tag{89}$$

Our fundamental equation is then either of the following relations between four-vectors

$$\Box^2(A - A_0) = \frac{1}{\lambda_L^2}(A - A_0) \quad ; \quad J - J_0 = -\frac{c}{4\pi\lambda_L^2}(A - A_0) \tag{90}$$

valid in any inertial reference frame. In the frame of reference at rest with respect to the superconducting body, J_0 and A_0 have only time-like components, in another reference frame they will also have space-like components. Eq. (84) for the fields, current and charge density can be written in covariant form as

$$\Box^2(J - J_0) = \frac{1}{\lambda_L^2}(J - J_0) \quad ; \quad \Box^2(F - F_0) = \frac{1}{\lambda_L^2}(F - F_0) \tag{91}$$

where F is the usual electromagnetic field tensor and F_0 is the field tensor with entries $\vec{E}_0$ and 0 for $\vec{E}$ and $\vec{B}$ respectively when expressed in the reference frame at rest with respect to the ions.

An important consequence of these equations is that they predict that externally applied electrostatic fields should be screened over a distance λ_L, the London penetration depth [75, 76], rather than over the much shorter Thomas-Fermi screening length, as the conventional theory predicts. This should be so at zero temperature. At finite temperatures, the effective screening length decreases since excited quasiparticles screen with the much shorter Thomas Fermi screening length.

3.3 Spin electrodynamics and the Spin-Meissner effect

The canonical momentum of an electron with superfluid velocity $\vec{v}_s$ is $\vec{p} = m_e\vec{v}_s + \frac{e}{c}\vec{A}$, with $\vec{A}$ the magnetic vector potential. In the BCS ground state the expectation value $<\vec{p}>= 0$, hence the superfluid velocity is given by $\vec{v}_s = -\frac{e}{m_e c}\vec{A} = -\frac{e\lambda_L}{m_e c}\vec{B}\times\hat{n}$. The second equality applies

to a cylindrical geometry, where $\hat{n}$ is the outward pointing normal of the lateral surface of the cylinder and $\vec{B}$ is the magnetic field along the axis of the cylinder.

Consider an electron that moves radially outward from the axis of a cylinder in the presence of a magnetic field $\vec{B}$ parallel to the cylinder. The equation of motion is

$$m_e \frac{d\vec{v}}{dt} = \frac{e}{c}\vec{v} \times \vec{B} + \vec{F}_r \tag{92}$$

where the first term is the magnetic Lorentz force and the second term is a radial force arising from "quantum pressure" that drives the electron outward. From it we infer

$$\vec{r} \times \frac{d\vec{v}}{dt} = \frac{e}{m_e c}\vec{r} \times (\vec{v} \times \vec{B}) \tag{93}$$

where $\vec{r}$ is in the plane perpendicular to the axis of the cylinder. Hence $\vec{r} \cdot \vec{B} = 0$ and $\vec{r} \times (\vec{v} \times \vec{B}) = -(\vec{r} \cdot \vec{v})\vec{B}$, and

$$\frac{d}{dt}(\vec{r} \times \vec{v}) = -\frac{e}{m_e c}(\vec{r} \cdot \vec{v})\vec{B} = -\frac{e}{2m_e c}(\frac{d}{dt}r^2)\vec{B} \tag{94}$$

so that $\vec{r} \times \vec{v} = -(e/2m_e c)r^2\vec{B}$, and the acquired azimuthal velocity in moving out a distance r is

$$v_\phi = -\frac{e}{2m_e c}rB \tag{95}$$

Thus, to acquire the azimuthal speed needed for the Meissner current, $v_s = -e\lambda_L/(m_e c)B$, requires the action of the Lorentz force *over a radially outgoing motion to radius* $r = 2\lambda_L$. This is shown schematically in Fig. 11 left panel.

Consider next a magnetic moment $\vec{\mu}$ along the z direction that moves radially outward with velocity $\vec{v}$. It is equivalent to an electric dipole moment $\vec{p} = \frac{\vec{v}}{c} \times \vec{\mu}$. The radial electric field of the cylinder that results from the positive charge that compensates the superfluid negative charge density en_s is $\vec{E} = 2\pi\rho\vec{r} = 2\pi|e|n_s\vec{r}$. The electric dipole experiences a torque

$$\vec{\tau} = \vec{p} \times \vec{E} = (\frac{\vec{v}}{c} \times \vec{\mu}) \times \vec{E} = -2\pi|e|n_s\vec{r} \times (\frac{\vec{v}}{c} \times \vec{\mu}) \tag{96}$$

which causes a change in its angular momentum

$$\frac{d\vec{L}}{dt} = m_e \frac{d}{dt}(\vec{r} \times \vec{v}) = \vec{\tau} \tag{97}$$

hence

$$\vec{r} \times \frac{d\vec{v}}{dt} = \frac{2\pi e n_s}{m_e}\vec{r} \times (\frac{\vec{v}}{c} \times \vec{\mu}). \tag{98}$$

Eq. (98) is identical to Eq. (93) if we define the 'effective' magnetic field

$$\vec{B}_\sigma = 2\pi n_s \vec{\mu} \tag{99}$$

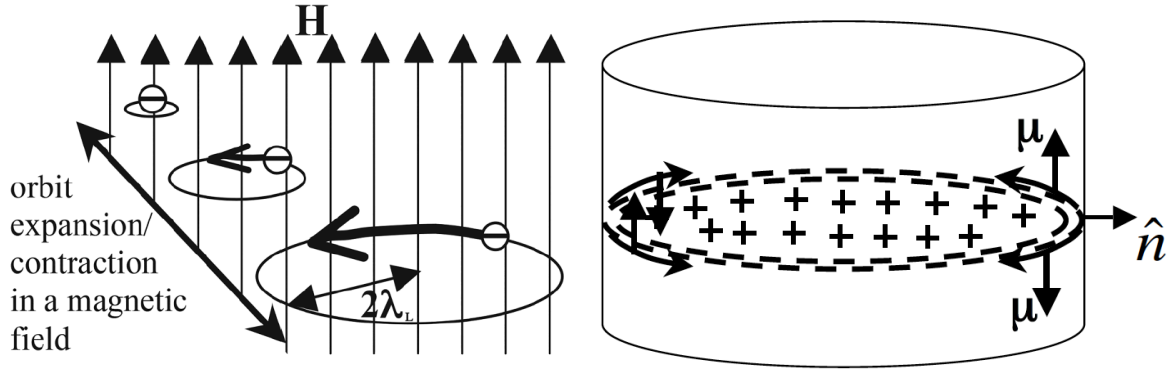

Fig. 11: *Left panel: when electrons forming a Cooper pair enlarge their orbits from a microscopic radius to radius $2\lambda_L$, they acquire the azimuthal velocity required for the Meissner current. The interior currents cancel out, and a charge current circulating in the surface layer of thickness λ_L results, the Meissner current. Right panel: electrons in expanding orbits acquire also azimuthal velocity through the spin-orbit interaction of their magnetic moment $\mu_z = -2\mu_B S_z$, with μ_B the Bohr magneton, leading to a spin current circulating in the surface layer of thickness λ_L, that adds to the charge current (if any), or is a pure spin current in the absence of applied magnetic field.*

and hence leads to the azimuthal velocity as derived earlier, $v_s = -e\lambda_L/(m_e c)B$, with B_σ replacing B and $|\vec{\mu}| = \mu_B$ the intrinsic magnetic moment of the electron

$$v_\phi = -\frac{\pi e n_s}{m_e c} r\mu_B \quad ; \quad v_\phi = \frac{\pi n_s e^2 \hbar r}{2m_e^2 c^2} = \frac{\hbar r}{8m_e\lambda_L^2} \tag{100}$$

with $\mu_B = |e|\hbar/2m_e c$ the Bohr magneton. The two electrons in a Cooper pair have opposite spin and orbit in opposite directions. The orbital angular momentum of each electron is $l = m_e r v_\phi = \hbar r^2/(8\lambda_L^2)$.

Remarkably, for $r = 2\lambda_L$, the size of the orbit required to explain the Meissner effect, the orbital angular momentum is $\boxed{l = \hbar/2}$.

The azimuthal velocity has magnitude $v_\phi \equiv v_\sigma^0 = \hbar/(4m_e\lambda_L)$. In the interior, the azimuthal velocities cancel out. Within a layer of thickness λ_L from the surface, they give rise to a spin current, where electrons of opposite spin flow in opposite directions, as shown schematically in Fig. 11 right panel [77, 78].

Figure 12 shows the three key aspects of the physics of superconductors within the theory discussed here. (a) The charge distribution in the superconductor is macroscopically inhomogeneous, with excess negative charge near the surface and excess positive charge in the interior. (b) Superfluid carriers reside in overlapping mesoscopic orbits of radius $2\lambda_L$. (c) A macroscopic spin current flows near the surface of superconductors in the absence of applied external fields. Macroscopic phase coherence results from the fact that the $2\lambda_L$ orbits are strongly overlapping.

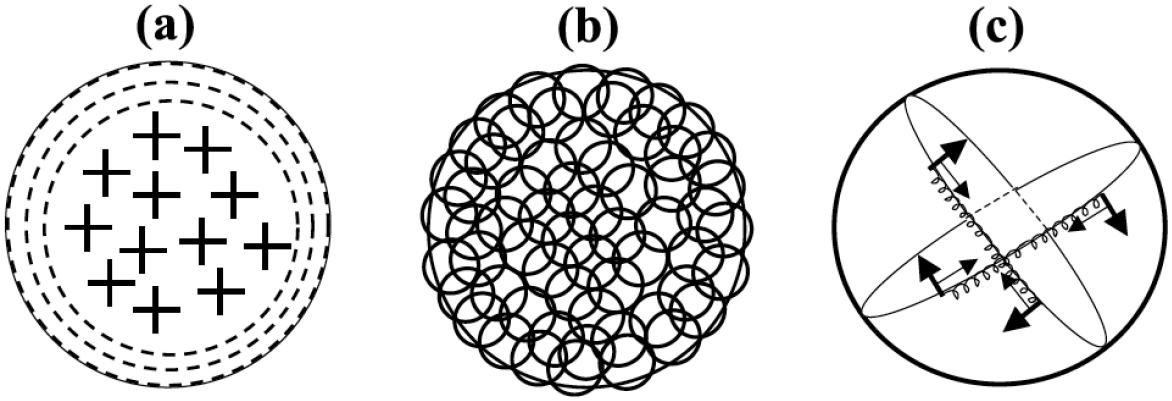

Fig. 12: *Illustration of three key aspects of the physics of superconductors discussed here. (a) : Superconductors expel negative charge from their interior to the region near the surface. (b) : Carriers reside in mesoscopic overlapping orbits of radius $2\lambda_L$. (c) : A spin current flows near the surface of superconductors (the arrow perpendicular to the orbit denotes the direction of the electron magnetic moment).*

The fact that superfluid electrons reside in mesoscopic orbits of radius $2\lambda_L$ can be seen from the equivalence of the following two expressions for the total angular momentum L of n_s electrons per unit volume flowing with velocity v along the lateral surface of a cylinder of radius R and height h carrying the Meissner current within a distance λ_L from the surface:

$$L = [2\pi R\lambda_L h n_s][m_e v R] = [\pi R^2 h n_s][m_e v(2\lambda_L)]. \tag{101}$$

When spin is taken into account, the four-vector current is $J = J_\uparrow + J_\downarrow$ and the electrodynamic equation (88) becomes [78]

$$J_\sigma - J_{\sigma 0} = -\frac{c}{8\pi\lambda_L^2}(A_\sigma - A_{\sigma 0}) \quad ; \quad J_\sigma = (\vec{J}_\sigma, ic\rho_\sigma) \quad ; \quad A_\sigma = (\vec{A}_\sigma, i\phi_\sigma) \tag{102}$$

$\vec{J}_\sigma = e(n_s/2)\vec{v}_\sigma$ is the component of the current of spin σ and ρ_σ is the charge density with spin σ. The spin potentials are given by [78]

$$\vec{A}_\sigma = \lambda_L \vec{\sigma} \times \vec{E}(\vec{r}, t) + \vec{A}(\vec{r}, t) \quad ; \quad \phi_\sigma(\vec{r}, t) = -\lambda_L \vec{\sigma} \cdot \vec{B}(\vec{r}, t) + \phi(\vec{r}, t) \tag{103}$$

Finally, the quantities with subindex 0 are

$$J_{\sigma 0} = (\vec{J}_{\sigma 0}(\vec{r}), ic\rho_{\sigma 0}) \quad ; \quad \vec{J}_{\sigma 0}(\vec{r}) = -\frac{c\rho_0}{2}\vec{\sigma} \times \hat{r} \quad ; \quad \rho_{\sigma 0} = \frac{\rho_0}{2} \tag{104}$$

and

$$A_{\sigma 0} = (\vec{A}_{\sigma 0}(\vec{r}), i\phi_{\sigma 0}(\vec{r})) \quad ; \quad \vec{A}_{\sigma 0}(\vec{r}) = \lambda_L \vec{\sigma} \times \vec{E}_0(\vec{r}) \quad ; \quad \phi_{\sigma 0}(\vec{r}) = \phi_0(\vec{r}). \tag{105}$$

These equations predict the existence of a spontaneous spin current flowing within a London penetration depth of the surface of the superconductor in the absence of applied fields, with carrier densities $(n_s/2)$ and opposite spin electrons flowing in opposite direction with speed $v_\sigma^0 = \hbar/(4m_e\lambda_L)$, and a spontaneous electric field throughout the interior of the superconductor.

Remarkably, the formalism uniquely determines the value of the expelled charge and maximum electric field in the interior [78], as

$$E_m = -\frac{\hbar c}{4e\lambda_L^2} \quad ; \quad \rho_- = -\frac{E_m}{4\pi\lambda_L} \quad ; \quad \rho_- = en_s\frac{v_{\sigma_0}}{c} \quad ; \quad v_\sigma^0 = \frac{\hbar}{4m_e\lambda_L}. \tag{106}$$

Note that E_m is the same as the lower critical magnetic field of a BCS superconductor H_{c1}. In the absence of an applied magnetic field, electrons near the surface move in opposite direction with speed v_σ^0, as shown in Fig. 12. When a magnetic field is applied, electrons of one spin speed up and those of opposite spin slow down, according to $v_\sigma = v_s + \sigma v_\sigma^0$. The total excess negative charge density $\rho_- = \rho_\uparrow + \rho_\downarrow$ doesn't change, but it has different magnitudes for spin up and down, according to

$$\rho_\sigma = \frac{n_s e}{2}(\frac{v_\sigma^0 + \sigma v_s}{c}) \quad ; \quad v_s = -\frac{e\lambda_L}{m_e c}B \quad ; \quad \frac{1}{\lambda_L^2} = \frac{4\pi n_s e^2}{m_e c^2}. \tag{107}$$

These equations imply that for applied magnetic field $H_{c1} = E_m$ one of the components of the spin current stops, at which point the magnetic field penetrates the sample [78].

The value of the interior positive charge density ρ_0 is determined by charge neutrality. For a cylinder and a sphere, it is $\rho_0 = -2\lambda_L/R\rho_-$ and $\rho_0 = -3\lambda_L/R\rho_-$ respectively.

4 How the Meissner effect works

We next will show that the physics discussed in the previous sections leads to a dynamical explanation of the Meissner effect. Contrary to what is generally believed, BCS theory has not provided a dynamical explanation of the Meissner effect. Why is it that it is generally believed that BCS theory explains the Meissner effect?

4.1 BCS theory *does not explain* the Meissner effect

Within BCS theory, the Meissner effect is explained as follows [7]. One considers the linear response of a system *in the BCS state* Eq. (5) to the perturbation created by a magnetic field, as shown in Fig. 13. The perturbing Hamiltonian H_1 is the linear term in the magnetic vector potential $\vec{A}$ that results from the kinetic energy $(\vec{p} - (e/c)\vec{A})^2/2m$, and it causes a change in the BCS ground state $|\Psi_{BCS}>$ to first order in $\vec{A}$:

$$H_1 = \frac{ie\hbar}{2mc}\sum_i(\vec{\nabla}_i \cdot A + \vec{A}\cdot\vec{\nabla}) \quad ; \quad |\Psi> = |\Psi_{BCS}> - \sum_n \frac{<\Psi_n|H_1|\Psi_{BCS}>}{E_n}|\Psi_n> \tag{108}$$

where $|\Psi_n>$ are states obtained from the BCS state $|\Psi_{BCS}>$ by exciting 2 quasiparticles, and E_n is the excitation energy. The expectation value of the current operator $\vec{J}_{op}$ with this wave function gives the electric current $\vec{J}$, and hence the "London Kernel" K [13]. In the long wavelength limit this calculation yields

$$\vec{J} = <\Psi|\vec{J}_{op}|\Psi> = -\frac{c}{4\pi}K\vec{A} \quad ; \quad K = \frac{1}{\lambda_L^2} \tag{109}$$

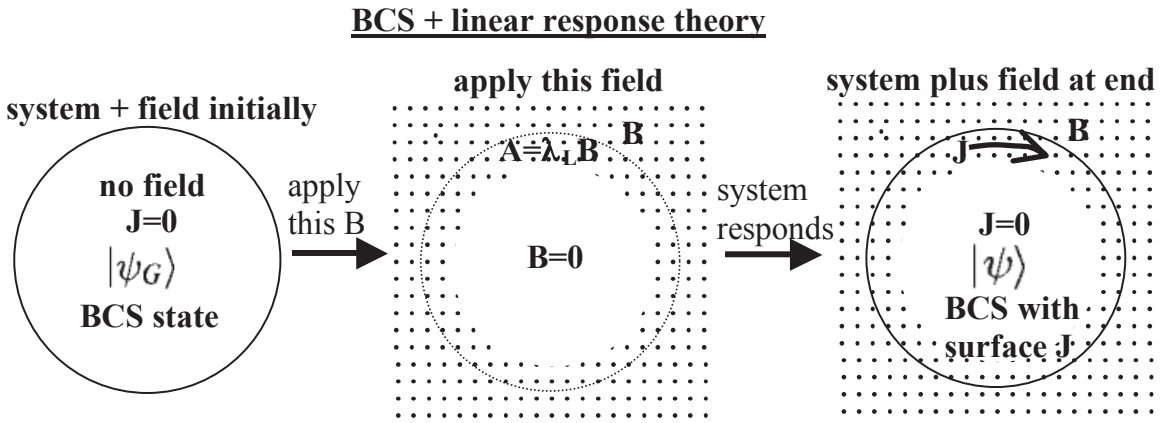

Fig. 13: *The BCS explanation of the Meissner effect. The system (cylinder, top view) is initially in the BCS state (left panel) with no magnetic field. Its linear response to the magnetic field shown in the middle panel (dots) is computed to first order in the magnetic field. The result is the state shown in the right panel, with a surface current J circulating.*

where λ_L is the London penetration depth. Eq. (109) is the (second) London equation Eq. (76). In combination with Ampere's law, Eq. (109) predicts that the magnetic field does not penetrate the superconductor beyond a distance λ_L from the surface, where the current $\vec{J}$ circulates, as shown schematically in Fig. 13 right panel.

However, note that this calculation *uses only the BCS wavefunction in and around the BCS state*, namely the ground state wavefunction $|\Psi_{BCS}>$ and the wavefunctions $|\Psi_n>$ that result from breaking one Cooper pair at a time. The wavefunction of the normal metal never appears. This is not explaining the Meissner effect. The Meissner effect is what is shown in Fig. 14: the $process$ by which a system starting in the normal metallic state expels a magnetic field in the process of becoming a superconductor. It cannot be explained by starting from the assumption that the system is in the final BCS state and gets perturbed by H_1. Explaining this process requires explaining how the interface between normal and superconducting regions moves (center panel in Fig. 14). Because calculations of the sort described in Eqs. (108), (109) contain no information about what is the nature of the initial state when the Meissner effect starts, namely the normal metal, they cannot be a microscopic derivation of the Meissner effect.

During the process of field expulsion, as well as its reverse, the process where a superconductor with a magnetic field excluded turns normal and the field penetrates, a Faraday electric field is generated that opposes the process. This electric field drives current in direction opposite to the current that develops. So it is necessary to explain: (i) How can a Meissner current start to flow in direction opposite to the Faraday electric force resisting magnetic flux change (Lenz's law)? (ii) How is the angular momentum of the developing supercurrent compensated so that momentum conservation is not violated? (iii) When a supercurrent stops, what happens to the angular momentum that the supercurrent had? (iv) How can a supercurrent stop without generation of Joule heat and associated with it an irreversible increase in the entropy of the universe that is known not to occur? None of these questions are addressed in the BCS literature.

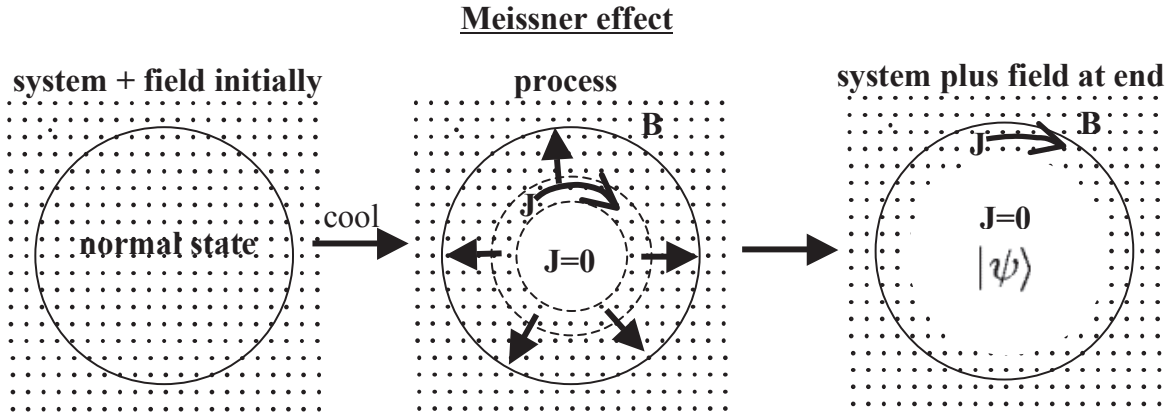

Fig. 14: *What the Meissner effect really is: the process by which a normal metal becomes superconducting in the presence of a magnetic field throughout its interior initially. The simplest route in this process (not the only one) is depicted in the figure. The superconducting region (white region) expands gradually from the center to fill the entire volume, expelling the magnetic field in the process.*

4.2 The Meissner effect *necessitates* charge expulsion

That the London derivation of the London equation does not account for the Meissner effect is clear. To get from the first to the second equality in Eq. (75), a time integration was performed after taking the curl and using Faraday's law. More explicitly,

$$\begin{aligned} \frac{\partial \vec{J}}{\partial t} &= \frac{n_s e^2}{m_e}\vec{E} \rightarrow \frac{\partial}{\partial t}(\vec{\nabla}\times\vec{J}) = -\frac{n_s e^2}{m_e c}\frac{\partial \vec{B}}{\partial t} \rightarrow \\ &\rightarrow \vec{\nabla}\times\vec{J}(\vec{r},t) - \vec{\nabla}\times\vec{J}(\vec{r},0) = -\frac{n_s e^2}{m_e c}(\vec{B}(\vec{r},t) - \vec{B}(\vec{r},0)). \end{aligned} \tag{110}$$

The last equality in Eq. (110) leads to the London equation Eq. (75) if $\vec{B}(\vec{r},0) = \vec{J}(\vec{r},0) = 0$. However, under the initial conditions appropriate to the Meissner effect, namely $\vec{B}(\vec{r},0) = \vec{B}_0, \vec{J}(\vec{r},0) = 0$ it leads instead to the solution $\vec{B}(\vec{r},t) = \vec{B}_0, \vec{J}(\vec{r},t) = 0$. No current is generated, and the magnetic field is not expelled, contrary to what experiment tells us.

To understand what is needed to expel the magnetic field, let us consider more carefully the equation of motion for an electron of charge e and mass m_e in the presence of electric and magnetic fields:

$$\frac{d\mathbf{v}}{dt} = \frac{e}{m_e}\mathbf{E} + \frac{e}{m_e c}\mathbf{v}\times\mathbf{B} \tag{111}$$

The left-hand side of Eq. (111) is the total (convective) time derivative, which is related to the local (partial) time derivative by

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = \frac{\partial \mathbf{v}}{\partial t} + \nabla(\frac{\mathbf{v}^2}{2}) - \mathbf{v}\times(\nabla\times\mathbf{v}) \tag{112}$$

Defining the 'generalized vorticity'

$$\mathbf{w} = \nabla\times\mathbf{v} + \frac{e}{m_e c}\mathbf{B}, \tag{113}$$

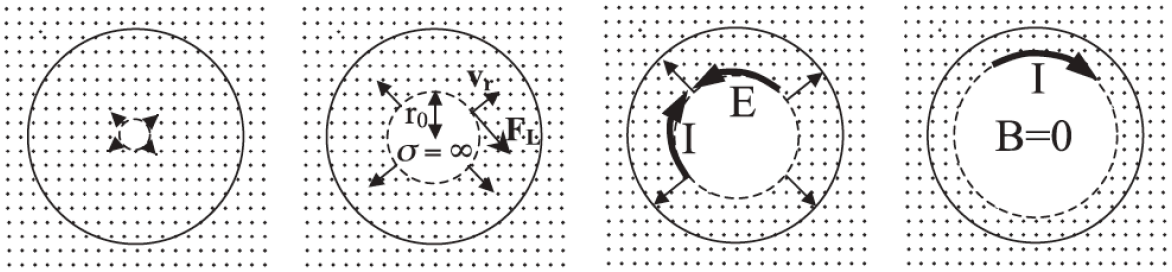

Fig. 15: *Schematic depiction of a perfectly conducting fluid ($\sigma = \infty$) that flows radially outward with radial velocity v_r in a uniform magnetic field perpendicular to the plane of the paper. The carriers at the boundary experience a Lorentz force F_L. Assuming the sign of the charge q is positive for definiteness the Lorentz force $F_L = (q/c)v_r B$ points in the clockwise direction. The resulting electric current I at the boundary flows clockwise (for negative charge carriers the Lorentz force would be in the opposite direction, the current in the same direction), generating a magnetic field opposite to the external field so that no magnetic field lines can penetrate the fluid. During this process, a Faraday electric field E is generated that opposes the current flow.*

taking the curl of Eq. (111) and using Eq. (112) and Faraday's law $\nabla \times \mathbf{E} = -(1/c)\partial\mathbf{B}/\partial t$ leads to the following equation of motion for $\mathbf{w}$:

$$\frac{\partial \mathbf{w}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{w}) \tag{114}$$

Note that $\mathbf{w}$ is essentially the curl of the canonical momentum $\mathbf{p} = m_e\mathbf{v} + (e/c)\mathbf{A}$, with $\mathbf{A}$ the magnetic vector potential. In the Meissner process we have at time $t = 0$: $\mathbf{w}(\mathbf{r}, t = 0) = \frac{e}{m_e c}\mathbf{B}(t = 0) \equiv \mathbf{w}_0$ independent of position $\mathbf{r}$. We set $\nabla \times \mathbf{v} = 0$ because in the normal state there is no net macroscopic charge flow. Hence the canonical momentum $\mathbf{p}$ is nonzero throughout the interior of the superconductor in the initial state. In the superconducting state, the superfluid velocity $\mathbf{v}$ obeys the London equation $\nabla \times \mathbf{v} = -\frac{e}{m_e c}\mathbf{B}$. Therefore, $\mathbf{w}(\mathbf{r}, t = \infty) = 0$ everywhere in the superconducting body. Equivalently, the canonical momentum $\mathbf{p} = 0$ throughout the interior of the (simply connected) superconductor. In a cylindrical geometry, assuming azimuthal symmetry as well as translational symmetry along the cylinder axis (z) direction (infinitely long cylinder) $\mathbf{w}(\mathbf{r}, t) = w(r, t)\hat{z}$ and Eq. (114) takes the form

$$\frac{\partial w}{\partial t} = -\frac{1}{r}\frac{\partial}{\partial r}(rwv_r) \tag{115}$$

with r the radius in cylindrical coordinates. Eq. (115) implies that w *can only change if there is radial flow of charge* ($v_r \neq 0$). Moreover, for w to evolve towards its final value 0 requires $v_r > 0$, i.e. a radial $outflow$ of electrons. This is a particular case of what's called Alfven's theorem [79], that says that in a perfectly conducting fluid magnetic field lines are frozen into the fluid and move with the fluid. It predicts what is shown in Fig. 15: if a perfectly conducting fluid expands from the center in the presence of magnetic field, it will push the magnetic field lines out as it expands, since otherwise Alfven's theorem would be violated.

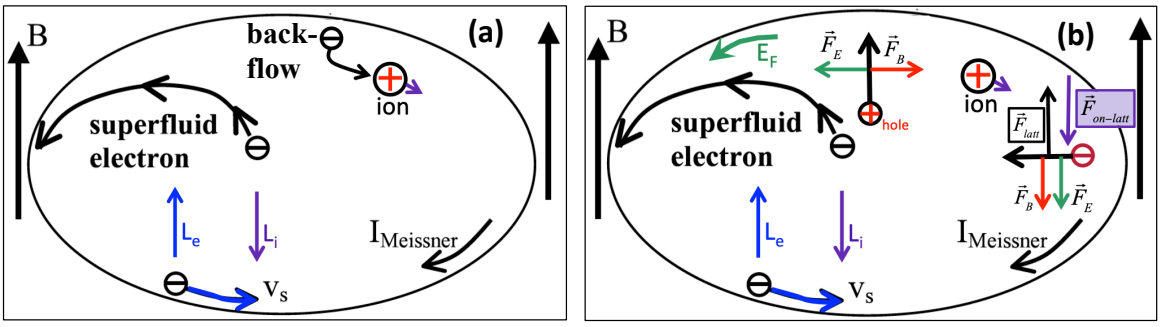

Fig. 16: *Left panel: electrons becoming superconducting move out, and are deflected by the Lorentz force in counterclockwise direction, generating the clockwise Meissner current $I_{Meissner}$. A backflow of normal electrons gets deflected clockwise and transmit their azimuthal momentum to the ions, hence to the body as a whole. The right panel explains how the backflowing electrons transmit their azimuthal momentum to the ions without scattering processes, that would generate Joule heat in contradiction with the observation that the transition is reversible (see text).*

4.3 Why holes are indispensable to understand the Meissner effect

The left panel of Fig. 16 shows qualitatively how the Meissner effect works [18]. Electrons condensing into the superconducting state move radially outward, and in the presence of a magnetic field B acquire counterclockwise azimuthal velocity, giving rise to the Meissner current $I_{Meissner}$ flowing near the surface that generates a magnetic field in opposite direction to the applied field and cancels it in the interior. There is also a backflow of normal electrons to preserve near charge neutrality, that acquire through the Lorentz force an azimuthal velocity in opposite direction. The backflowing electrons do not cancel the Meissner current because they transmit their azimuthal momentum to the ions, i.e. to the body as a whole. The body rotates very slowly in clockwise direction, compensating the counterclockwise motion of superfluid electrons in the Meissner current, so that momentum conservation is maintained.

However, the question arises: how do the backflowing electrons transmit their azimuthal momentum to the body? It cannot happen through collisions with impurities or phonons, because such collisions would generate Joule heat. However it is known that the transition is thermodynamically reversible, this has been carefully tested experimentally [80] and is also implicit in BCS theory, allowing the understanding of the transition as a thermodynamic phase transition, of second order without a magnetic field and first order in the presence of a magnetic field.

The answer to this question requires $holes$, or equivalently antibonding electrons. *The backflowing electrons need to have negative effective mass*. If so, their motion is purely radial, as Fig. 16 (b) shows: they experience magnetic and electric forces in the same direction, clockwise, and the lattice exerts a counterclockwise force F_{latt} that exactly cancels the electromagnetic forces. By Newton's third law, a force on the ions $F_{on-latt} = -F_{latt}$ is exerted by the electrons, transferring their azimuthal momentum to the body without any dissipation. If we prefer to describe the backflowing normal electrons equivalently as outflowing normal holes we can also do that.

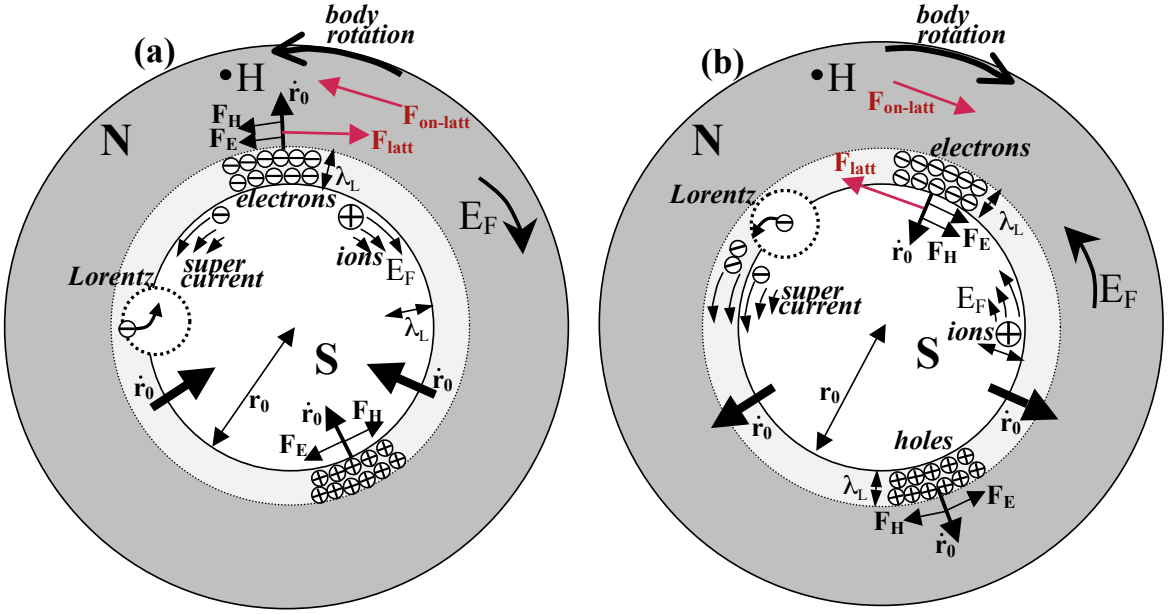

Fig. 17: *Transition from the normal into the superconducting state (left panel) and from the superconducting into the normal state (right panel) in the presence of a magnetic field. The processes are discussed in the text.*

In that case, electric and magnetic forces are exactly cancelled, as Fig. 16 (b) shows.

The electric force discussed above arises from the Faraday electric field that exists during the process of flux expulsion. If the phase boundary is moving at speed v_r, the Faraday electric field at the boundary is $E_F = (v_r/c)B$, and the magnetic force on normal carriers moving radially with the boundary, $F_B = e(v_r/c)B$, is of the same magnitude as the electric force exerted by the Faraday field $F_E = eE_F$, in opposite direction for holes and in the same direction for electrons, as shown in Fig. 16 (b).

Fig. 17 illustrates the processes in more detail. Starting with the left panel, that describes the Meissner effect, the phase boundary is moving outward with speed $\dot{r}_0$. Normal electrons at the boundary expand their orbits to radius $2\lambda_L$, as discussed earlier, and this expansion imparts them the azimuthal speed of carriers in the Meissner current, as was shown in Eqs. (92)-(93). The resulting outward motion of negative charge gives rise to an inflow of normal electrons just outside the boundary, moving inward with speed $\dot{r}_0$. They experience a Lorentz force pointing clockwise, and are also subject to the clockwise electric force resulting from Faraday's electric field pointing counterclockwise that originates in the outward motion of magnetic flux. The forces are balanced by a force exerted by the lattice on the backflowing electron that has negative effective mass, F_{latt}, as discussed earlier, pointing in counterclockwise direction, so that the backflow is radial. In turn the backflowing electron exerts an equal and opposite force on the lattice, $F_{on-latt}$, thus transmitting azimuthal momentum to the body without dissipation, compensating for the counterclockwise momentum acquired by electrons expanding their orbits and joining the Meissner current. Instead of backflowing electrons we can understand the process with outward moving normal holes, as shown on the lower part of the left panel. As the phase

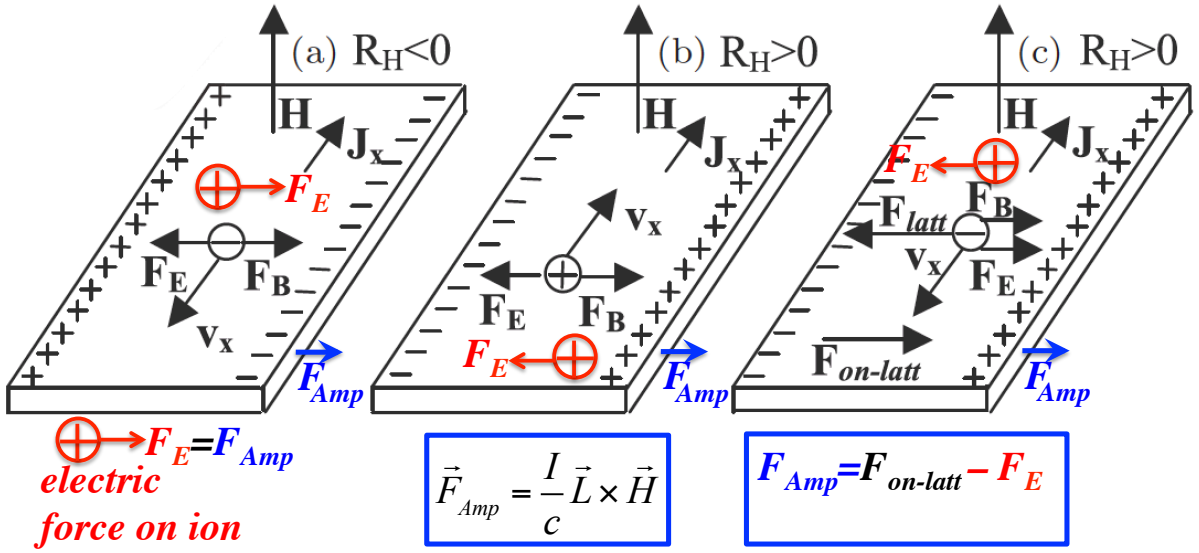

Fig. 18: *Illustration of momentum transfer without energy dissipation in a Hall bar. The Ampere force points to the right independently of whether the material has negative or positive Hall coefficient. For the material with negative Hall coefficient (left panel) the Ampere force results from the electric force resulting from the electric field created by the charge imbalance on the positive ions (in red). For the material with positive Hall coefficient (middle and right panels) the force from the electric field acting on positive ions is opposite to the Ampere force.*

boundary moves further out, the superelectrons at its boundary that acquired azimuthal velocity through orbit expansion get slowed down by the Faraday electric field and stop contributing to the supercurrent when the phase boundary has moved beyond them a distance λ_L. All this is discussed quantitatively in Refs. [81, 82]. The same momentum transfer without energy dissipation explains the origin of the Ampere force on conductors with positive Hall coefficient, as illustrated in Fig. 18.

Switching the sign of all the processes we can understand the right panel of Fig. 17, namely the process by which a superconductor in a magnetic field carrying a Meissner current near its surface, turns normal and the supercurrent stops [83]. As the phase boundary moves in, electrons that are in the superconducting region at distance λ_L from it get accelerated by the Faraday electric field in counterclockwise direction, reaching maximum velocity when the phase boundary reaches them, at which point their orbits shrink and they are stopped by the action of the Lorentz force on the shrinking orbit pushing in clockwise direction. This explains how the supercurrent stops when the system becomes normal, without collisions and hence no dissipation. Momentum conservation results from the compensating backflow of normal electrons of negative effective mass as discussed earlier.

Finally, the same physics that explain how the Meissner effect works explains how magnetic fields are generated in rotating superconductors [84], and how charge flows in a superconducting wire connected to normal metal leads [85]. BCS theory does not have the physical elements necessary to describe any of these processes.

5 Theory of hole superconductivity versus the conventional theory

In the conventional understanding of superconductivity, there is a large set of materials considered to be "conventional superconductors", for which the pairing mechanism is believed to be the electron-phonon interaction. There is another large set of materials considered to be "unconventional superconductors", for which the pairing interaction is believed not to be the electron-phonon interaction. There is no general agreement on how many other pairing mechanisms exist, nor what is their nature, although most physicists believe that magnetic interactions of some kind are responsible for pairing in various unconventional superconductors such as the cuprates. A survey of 32 classes of superconducting materials is given in Ref. [8].

Instead, within the theory of hole superconductivity discussed here, there is a single mechanism of superconductivity for all materials, that originates in the fundamental charge asymmetry of matter, namely the fact that the proton is 2000 times heavier than the electron, leading to electron-hole asymmetry in condensed matter, essential to understand both the pairing mechanism of charge carriers near the top of electronic energy bands, as well as their ability to interchange momentum with the body as a whole without dissipation due to their negative effective mass, which is necessary to understand how electric currents start and stop in superconductors without dissipating Joule heat.

Both points of view could be wrong, but not both can be right. What do we learn from superconducting materials? We discuss this in the next section.

5.1 Superconducting materials: judge and jury of theories of superconductivity

It is difficult to prove theories wrong in condensed matter physics, because they are based on model Hamiltonians whose connection with real materials is difficult to ascertain. It is especially difficult for the case of BCS theory, because it is generally assumed that if a material does not conform to it this does not indicate that the theory is wrong but rather that the material is "wrong", i.e. non-conventional.

Here we have argued that Hamiltonians such as dynamic Hubbard models contain the essential physics necessary to describe superconductivity, and Hamiltonians describing the electron-phonon interaction do not. How do we decide which is right and which is wrong? One way is to consider what real material tell us [17].

The highest T_c unconventional superconductors are cuprates. The highest T_c $proven$ conventional superconductor is magnesium diboride (we exclude the hydrides for reasons discussed later). What do MgB_2 and cuprates have in common?

According to the conventional view, nothing. MgB_2 becomes superconducting because of a strong electron-phonon interaction, and cuprates become superconducting because of mag-

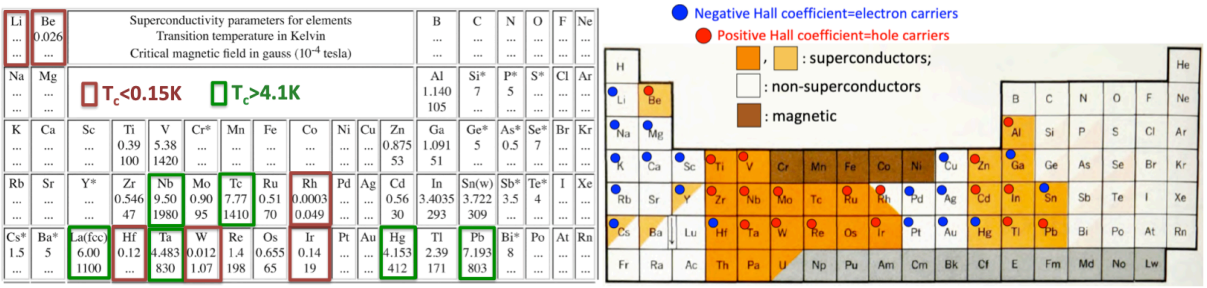

Fig. 19: *Two images of superconductors in the periodic table. The left image highlights the superconducting elements with highest T_c in green and those with lowest T_c in brown. The ionic mass increases as we move down and to the right in the table, and T_c should decrease as the ionic mass increases according to the conventional theory. It does not. The right image shows which elements among superconducting and non-superconducting ones have positive and negative Hall coefficients. It is clear that positive Hall coefficients are predominant in superconducting elements, and negative Hall coefficients are predominant in non-superconducting elements.*

netic fluctuations. MgB_2 does not have magnetic fluctuations, and cuprates don't have strong electron-phonon interaction.

Instead, according to the theory of hole superconductivity, what they have in common is that in both materials holes conduct through negatively charged anions in close proximity: B^- ions in MgB_2, $O^=$ ions in cuprates. That is what makes them high temperature superconductors.

The conventional theory predicts that light atoms should give rise to high critical temperatures, because the lattice vibration frequencies are high. But there is no evidence from materials for that, as illustrated in Fig. 19. For example, Pb, a very heavy element, has the third highest T_c among the elements. Li, a very light element, has the lowest T_c among the elements. On the other hand, there is a very strong correlation between positive sign of the Hall coefficient and the element being a superconductor, as the right panel of Fig. 19 shows, and as the theory of hole superconductivity predicts. Correlations between superconductivity and a variety of normal state properties of elements are analyzed in Ref. [86]. The same non-correlation with ionic mass and strong correlation with sign of the Hall coefficient is seen in compounds [17, 67, 87].

5.2 Experimental tests and open questions

Of course the ultimate test of theories is experiments. It is generally believed that BCS theory has been proven right by experiments. However, many of the predictions of BCS theory are common to other theories including the theory of hole superconductivity. The role of the electron-phonon interaction in causing superconductivity has not been proven experimentally. The isotope effect is not a proof, since many materials considered to be conventional do not obey the BCS prediction, including several elements and the compound PdH, where the mass of H can be increased by a factor of 2 by substitution with the isotope deuterium, and T_c goes up rather than down [88]. Small wiggles in tunneling characteristics [89], attributed to electron-

phonon coupling and generally believed to prove that superconductivity is caused by it [90], may also result from modulation of the pairing interaction Δt discussed here by phonons [91], implying that pairing would persist even if the ionic mass is infinite, i.e. if the lattice doesn't vibrate. It has been claimed that hydrogen-rich materials at high pressures superconduct at temperatures close to room temperature, proving the importance of the electron phonon interaction, that is predicted to give highest T_c for light ions such as hydrogen [92]. We have analyzed multiple experiments reporting such claims and in every case concluded that the experimental observations are incompatible with superconductivity [93].

There are several predictions of the theory of hole superconductivity that are specific to it, but most have not been tested experimentally to date. Some of the predictions are: (i) tunneling asymmetry of universal sign [64], (ii) positive thermoelectric power of superconductive tunnel junctions [68], (iii) apparent violation of the conductivity sum rule [63], (iv) electric screening length in the superconducting state much larger than in the normal metallic state [75, 76], (v) electric fields in the interior and in the vicinity of superconducting samples [73], (vi) increase in the mean inner potential in the superconducting state [45], (vii) charge imbalance in the absence of applied fields [94], (viii) radial electric fields during the normal-superconductor transition [95], (ix) Alfven-like waves along superconductor-normal phase boundaries [96], (x) absence of superconductivity in any material that doesn't have hole carriers [67].

Why is it important and urgent to decide which theory of superconductivity describes real materials? One important reason is that it would allow to make real progress in the theoretical search for new materials, to guide experimental search and discovery of superconductors that work at room temperature. Room temperature superconductors will change the world. Imagine how different our lives would be today if semiconductors only worked at temperatures below 150K. The current theoretical guidance based on BCS theory, that focuses on light elements [97], has not led to progress.

In conclusion, I would like to stress that the understanding of superconductivity based on the principles discussed in this paper is far from complete. There are many opportunities for further advances through theoretical and experimental research. Furthermore, a full understanding of how quantum mechanics operates on a macroscopic scale in superconductors may well lead to new insights on how it operates on the microscopic scale [98].

Acknowledgments
The author is grateful to Frank Marsiglio for collaboration in substantial parts of this work.

Corresponding author email address: jhirsch@ucsd.edu

References

[1] F. London and H. London, "The electromagnetic equations of the supraconductor", Proc.Roy.Soc. A 149, 71 (1935).

[2] H. Kamerlingh Onnes, Comm. Leiden 1911, Nr I22b, I24C; 1913, Nr 133a, I33C.

[3] W. Meissner and R. Ochsenfeld, "Ein neuer Effekt bei Eintritt der Supraleitfähigkeit", Naturwissenschaften **21**, 787 (1933).

[4] G. Lippmann, "Sur les proprietes des circuits electriques denues de resistance," Academie des Sciences, Comptes Rendus, **168**, 73 (1919).

[5] T. Sauer, "Einstein and the Early Theory of Superconductivity, 1919-1922", Archive for History of Exact Sciences, vol 61, p. 159 (2007).

[6] F. London and H. London, "Supraleitung und diamagnetismus", Physica 2, 341 (1935).

[7] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, "Theory of Superconductivity", Phys. Rev. 108, 1175 (1957).

[8] Physica C Special Issue, "Superconducting Materials: Conventional, Unconventional and Undetermined. Dedicated to Theodore H. Geballe on the year of his 95th birthday", ed. by J.E. Hirsch, M.B. Maple, F. Marsiglio, Vol. 514, 1-444 (2015).

[9] L.N. Cooper, "Bound Electron Pairs in a Degenerate Fermi Gas", Phys. Rev. 104, 1189 (1956).

[10] F. London, "Superfluids", Vol. I, Dover, New York, 1961.

[11] V. L. Ginzburg and L. D. Landau, Zh. Eksp. Teor. Fiz. 20, 1064 (1950).

[12] B. D. Josephson, "Possible new effects in superconductive tunnelling", Physics Letters 1, 251 (1962).

[13] M. Tinkham, "Introduction to superconductivity", Second Edition, McGraw Hill, New York, 1996.

[14] P. G. de Gennes, *Superconductivity of Metals and Alloys* (W.A. Benjamin, Inc. New York, 1966).

[15] M. L. Cohen and P. W. Anderson, "Comments on the Maximum Superconducting Transition Temperature ", AIP Conf. Proc. 4, 17 (1972).

[16] C.W. Chu, L.Z. Deng and B. Lv, "Hole-doped cuprate high temperature superconductors", Physica C: 514, 290 (2015).

[17] J. E. Hirsch, "Superconducting materials: Judge and jury of BCS-electron-phonon theory", Appl. Phys. Lett. 121, 080501 (2022).

[18] J. E. Hirsch, "On the dynamics of the Meissner effect", Physica Scripta **91**, 035801 (2016).

[19] J. E. Hirsch, "Superconductivity begins with H", World Scientific, Singapore, 2020.

[20] See https://jorge.physics.ucsd.edu/hole.html for a list of references.

[21] W. E. Pickett, "Room temperature superconductivity: The roles of theory and materials design", Rev. Mod. Phys. 95, 021001 (2023).

[22] N. W. Ashcroft, "Hydrogen Dominant Metallic Alloys: High Temperature Superconductors?", Phys. Rev. Lett. 92, 187002 (2004).

[23] P. Morel and P. W. Anderson, "Calculation of the Superconducting State Parameters with Retarded Electron-Phonon Interaction", Phys. Rev. 125, 1263 (1962).

[24] J. S. Bauer, J. E. Han and O. Gunnarsson, "The theory of electron-phonon superconductivity: does retardation really lead to a small Coulomb pseudopotential?", J. Phys. Condens. Matter 24, 492202 (2012).

[25] D. J. Scalapino, "Superconductivity and Spin Fluctuations", Journal of Low Temperature Physics 117, 179 (1999) and references therein.

[26] E. Dagotto and J. Riera, "Superconductivity in the two-dimensional t-J model", Phys. Rev. B 46, 12084(R) (1992).

[27] J. E. Hirsch, "Singlet pairs, covalent bonds, superexchange, and superconductivity", Physics Letters A 136, 163 (1989).

[28] J. E. Hirsch, E. Loh, D. J. Scalapino and S. Tang, "Antiferromagnetism and superconductivity: Can a Hubbard U do it all by itself?", Physica C 153-155, 549 (1988).

[29] M. Qin, Chia-Min Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S. R. White, and S. Zhang, "Absence of Superconductivity in the Pure Two-Dimensional Hubbard Model", Phys. Rev. X 10, 031016 (2020).

[30] J. E. Hirsch, "Attractive Interaction and Pairing in Fermion Systems with Strong On-Site Repulsion", Phys. Rev. Lett. 54, 1317 (1985).

[31] D. J. Scalapino, E. Loh, Jr., and J. E. Hirsch, "d-wave pairing near a spin-density-wave instability", Phys. Rev. B 34, 8190(R) 1986).

[32] "The Hubbard model at half a century", Nature Phys 9, 523 (2013).

[33] J. Bardeen and D. Pines, "Electron-Phonon Interaction in Metals", Phys. Rev. 99, 1140 (1955).

[34] J. E. Hirsch, "Coulomb attraction between Bloch electrons", Physics Letters A 138, 83 (1989).

[35] S. Kivelson, W.-P. Su, J. R. Schrieffer, and A. J. Heeger, "Missing bond-charge repulsion in the extended Hubbard model: Effects in polyacetylene", Phys. Rev. Lett. 58, 1899 (1987).

[36] J.E. Hirsch, "Bond-charge repulsion and hole superconductivity", Physica C 158, 326 (1989).

[37] N. W. Ashcroft and N. D. Mermin, "Solid State Physics", Cengage, Boston, USA, 1976, Eq. (10.18).

[38] J. E. Hirsch and F. Marsiglio, "Superconducting state in an oxygen hole metal", Phys. Rev. B 39, 11515 (1989).

[39] A. Painelli and A. Girlando, "Interacting electrons in 1D: Applicability of Hubbard models", Synthetic Metals 27, A15 (1988).

[40] J. E. Hirsch, "Empirical estimate of a Coulomb matrix element of relevance to superconductivity", Chemical Physics Letters 171, 161 (1990).

[41] J. E. Hirsch et al, in A. Montorsi, "The Hubbard Model: A Reprint Volume", World Scientific, SIngapore, 1992, p. 123-126, 206-213, 217-233, 274-282.

[42] J. E. Hirsch, "Electron-Hole Asymmetry: The Key to Superconductivity", in "High-Temperature Superconductivity", ed. by J. Ashkenazi et al, , p. 295 (1991).

[43] J.E. Hirsch, "Dynamic Hubbard Model", Phys.Rev. Lett. **87**, 206402 (2001).

[44] H. Bethe, "Theorie der Beugung von Elektronen an Kristallen", Ann. Physik 392, 55 (1928).

[45] J. E. Hirsch, "Superconductivity, diamagnetism, and the mean inner potential of solids", Annalen der Physik **526**, 63 (2013).

[46] J. E. Hirsch, "Inapplicability of the Hubbard model for the description of real strongly correlated electrons", Physica B 199-200, 366 (1994).

[47] J. Hutchinson, M. Baker and F. Marsiglio, "The spectral decomposition of the helium atom two-electron configuration in terms of hydrogenic orbitals", Europ. J. of Phys. 34, 111 (2013).

[48] J.E. Hirsch, "Superconductivity from undressing", Phys. Rev. B 62, 14487 (2000).

[49] J.E. Hirsch, "Superconductivity from undressing. II. Single-particle Green's function and photoemission in cuprates", Phys. Rev. B 62, 14998 (2000).

[50] J.E. Hirsch, "Why holes are not like electrons: A microscopic analysis of the differences between holes and electrons in condensed matter", Phys.Rev.B **65**, 184502 (2002).

[51] J.E. Hirsch, "Electronic dynamic Hubbard model: Exact diagonalization study", Phys. Rev. B 67, 035103 (2003).

[52] J. E. Hirsch, "Quantum Monte Carlo and exact diagonalization study of a dynamic Hubbard model", Phys. Rev. B 65, 214510 (2002).

[53] J. E. Hirsch, "Quasiparticle undressing in a dynamic Hubbard model: Exact diagonalization study", Phys. Rev. B 66, 064507 (2002).

[54] K. Bouadim et al, "Sign change of the extended s-wave pairing vertex in the dynamic Hubbard model: A quantum Monte Carlo study", Phys. Rev. B77, 014516 (2008).

[55] G.H. Bach, J. E. Hirsch and F. Marsiglio, "Two-site dynamical mean field theory for the dynamic Hubbard model", Phys. Rev. B 82, 155122 (2010).

[56] F. Marsiglio, R. Teshima and J.E. Hirsch, "Dynamic Hubbard Model: Effect of Finite Boson Frequency", Phys. Rev. B 68, 224507 (2003).

[57] J. E. Hirsch and F. Marsiglio, "Superconducting state in an oxygen hole metal", Phys. Rev. B 39, 11515 (1989).

[58] F. Marsiglio and J.E. Hirsch, "Hole superconductivity and the high-T_c oxides", Phys. Rev. B 41, 6435 (1990).

[59] J. E. Hirsch, 'Kinetic energy driven superconductivity, the origin of the Meissner effect, and the reductionist frontier', Int. J. Mod. Phys. B **225**, 1173 (2011).

[60] J. E. Hirsch, "Superconductors that change color when they become superconducting", Physica C 201, 347 (1992).

[61] H. Ding et al, 'Coherent Quasiparticle Weight and Its Connection to High- T_c Superconductivity from Angle-Resolved Photoemission', Phys. Rev. Lett. 87, 227001 (2001).

[62] H. J. A. Molegraaf, C. Presura, D. van der Marel, P. H. Kes, and M. Li, 'Superconductivity-Induced Transfer of In-Plane Spectral Weight in $Bi_2Sr_2CaCu_2O_{8+\delta}$', Science **295**, 2239 (2002).

[63] J. E. Hirsch, 'Apparent violation of the conductivity sum rule in certain superconductors', Physica C **199**, 305 (1992).

[64] F. Marsiglio and J.E. Hirsch, Physica C **159**, 157 (1989).

[65] Ch. Renner and O. Fischer, "Vacuum tunneling spectroscopy and asymmetric density of states of $Bi_2Sr_2CaCu_2O_{8+\delta}$", Phys. Rev. B 51, 9208 (1995).

[66] J. E. Hirsch, "Electron and hole hopping amplitudes in a diatomic molecule", Phys.Rev.B 48, 3327 (1993).

[67] J.E. Hirsch, "Materials and mechanisms of hole superconductivity", Physica C 472, 78 (2012) and references therein.

[68] J. E. Hirsch, "Thermoelectric power of superconductive tunnel junctions", Phys. Rev. Lett. 72, 558 (1994).

[69] J. E. Hirsch, "Consequences of charge imbalance in superconductors within the theory of hole superconductivity", Phys. Lett. A 281, 44 (2001).

[70] J. E. Hirsch, "Charge expulsion, charge inhomogeneity, and phase separation in dynamic Hubbard models", Phys. Rev. B 87, 184506 (2013).

[71] J.E. Hirsch, "Charge expulsion and electric field in superconductors", Phys. Rev. B 68, 184502 (2003).

[72] J. E. Hirsch, "Electrodynamics of superconductors", Phys. Rev. B 69, 214515 (2004).

[73] J. E. Hirsch, "Predicted Electric Field near Small Superconducting Ellipsoids", Phys. Rev. Lett. 92, 016402 (2004).

[74] J. E. Hirsch, "Correcting 100 Years of Misunderstanding: Electric Fields in Superconductors, Hole Superconductivity, and the Meissner Effect", Jour. Supercond. Novel Mag. **25**, 1357 (2012).

[75] J. E. Hirsch, "Experimental consequences of predicted charge rigidity of superconductors", Physica C: 478, A42 (2012).

[76] J. E. Hirsch, "Proposed experimental test of an alternative electrodynamic theory of superconductors", Elsevier Physica C 508, 21 (2015).

[77] J.E. Hirsch, "Spin Meissner effect in superconductors and the origin of the Meissner effect", Europhys. Lett. 81, 67003 (2008).

[78] J.E. Hirsch, 'Electrodynamics of spin currents in superconductors', Ann. Phys. (Berlin) **17**, 380 (2008).

[79] H. Alfven, "Existence of Electromagnetic-Hydrodynamic Waves", Nature 150, 405 (1942); Arkiv foer Matematik, Astronomi och Fysik, 39: 2 (1943).

[80] P.H. van Laer and W.H. Keesom, "On the reversibility of the transition process between the superconductive and the normal state", Physica **5**, 993 (1938).

[81] J.E. Hirsch, "Momentum of superconducting electrons and the explanation of the Meissner effect", Phys. Rev. B **95**, 014503 (2017).

[82] J. E. Hirsch, "On the reversibitity of the Meissner effect and the angular momentum puzzle", Annals of Physics **373**, 230 (2016).

[83] J. E. Hirsch, "The disappearing momentum of the supercurrent in the superconductor to normal phase transformation", Europhys. Lett. **114**, 57001 (2016).

[84] J. E. Hirsch, "Defying Inertia: How Rotating Superconductors Generate Magnetic Fields", Ann. Phys. (Berlin) 531, 1900212 (2019).

[85] J. E. Hirsch, "Magnetic flux expulsion in a superconducting wire", Phys. Lett. A 413, 127592 (2021).

[86] J.E. Hirsch, 'Correlations between normal-state properties and superconductivity', Phys. Rev. B 55, 9007 (1997).

[87] J. E. Hirsch, 'Superconducting materials: the w*hole* story' (Dedicated to Ted Geballe on his 100th birthday), J. Supercond. Nov. Mag. (2019).

[88] T Skoskiewicz, A W Szafranski, W Bujnowski and B Baranowski, "Isotope effect in the superconducting palladium-hydrogen-deuterium system", J. Phys. C: Solid State Phys. 7, 2670 (1974).

[89] I. Giaever, H. R. Hart Jr., and K. Megerle, "Tunneling into Superconductors at Temperatures below 1^oK, Phys. Rev. 126, 941 (1962).

[90] D. J. Scalapino, J. R. Schrieffer, and J. W. Wilkins, "Strong-Coupling Superconductivity. I", Phys. Rev. 148, 263 (1966).

[91] F. Marsiglio and J. E. Hirsch, unpublished.

[92] W. Pickett and M. I. Eremets, "The quest for room-temperature superconductivity in hydrides", Physics Today 72, 52 (2019).

[93] J. E. Hirsch, "Are hydrides under high pressure high temperature superconductors?", National Science Review, nwad174 (2023) and references therein.

[94] J.E. Hirsch, "Thermoelectric effect in superconductive tunnel junctions", Phys. Rev. B 58, 8727 (1998).

[95] J. E. Hirsch, Physica C 544, 54 (2018).

[96] J.E. Hirsch, 'Alfven-like waves along normal-superconductor phase boundaries', Physica C 564, 42 (2019).

[97] L. Boeri and . B. Bachelet, "Viewpoint: the road to room-temperature conventional superconductivity", J. Phys.: Condens. Matter 31, 234002 (2019).

[98] J.E. Hirsch, Phys. Lett. A **374**, 3777 (2010); Mod. Phys. Lett. B **24**, 2201 (2010).