SPECTROSCOPY OF DISCRETE ENERGY LEVELS IN ULTRASMALL METALLIC GRAINS

Jan VON DELFT[a]**, D.C. RALPH**[b]

[a]*Institut für Theoretische, Festkörperphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*
[b]*Laboratory of Atomic and Solid States Physics, Cornell University, Ithaca, NY 14853, USA*

Spectroscopy of discrete energy levels in ultrasmall metallic grains

Jan von Delft[a],*, D.C. Ralph[b]

[a]*Institut für Theoretische Festkörperphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*
[b]*Laboratory of Atomic and Solid States Physics, Cornell University, Ithaca, NY 14853, USA*

Contents

* Corresponding author. Present address: Physikalisches Institut, Universität Bonn, D-53115 Bonn, Germany.
E-mail addresses: vondelft@th.physik.uni-bonn.de (J. von Delft), ralph@msc.cornell.edu (D.C. Ralph).

Abstract

We review recent experimental and theoretical work on ultrasmall metallic grains, i.e. grains sufficiently small that the conduction electron energy spectrum becomes discrete. The discrete excitation spectrum of an individual grain can be measured by the technique of single-electron tunneling spectroscopy: the spectrum is extracted from the current–voltage characteristics of a single-electron transistor containing the grain as central island. We review experiments studying the influence on the discrete spectrum of superconductivity, nonequilibrium excitations, spin–orbit scattering and ferromagnetism. We also review the theoretical descriptions of these phenomena in ultrasmall grains, which require modifications or extensions of the standard bulk theories to include the effects of level discreteness. © 2001 Elsevier Science B.V. All rights reserved.

1. Introduction

One of the most fundamental features of quantum mechanics is the fact that the energy spectrum of a system of particles confined to a small region is *discrete* or quantized, with the typical spacing between energy levels increasing with decreasing system size. In atomic and nuclear physics, spectroscopic techniques for measuring and analyzing such discrete spectra have for decades been a major source of detailed information on the forces between the particles and the correlations which they experience. In condensed matter physics, however, it has been much more difficult to spectroscopically study the discrete spectrum of an individual sample, since system sizes were typically so large that discrete eigenenergies could not be resolved on the energy scale set by the temperature.

This changed in the course of the last 15 years due to advances in microfabrication techniques, which made it possible to study individual systems of mesoscopic or nanoscopic dimensions, whose characteristic length scales range from a few μm down to a few nm. In the early 1990s, semiconductor devices were used to fabricate the first "quantum dots", i.e. droplets of charge confined to a two-dimensional region so small (radius of order 50 nm) that discrete levels in the conduction electron spectrum [1,2] could be resolved at dilution refrigerator temperatures in the 10–100 mK range. The technique by which this was done is called *single-electron-tunneling spectroscopy*: the dot is connected to two leads via electrostatically defined tunnel barriers to form a so-called single-electron transistor (SET), the current–voltage characteristics of which are measured and analyzed. Under certain conditions the conductance shows well-defined resonances which can be associated with tunneling through discrete eigenstates of the dot. Since quantum dots exhibited various features familiar from atomic physics, such as energy shells that feature magic numbers and are filled according to Hund's rules, etc., they are often aptly referred to as "artificial atoms" (see e.g. the collection of reviews on various aspects of mesoscopic systems may be found in [3], in particular that by Kouwenhoven et al. [4]).

In the mid-1990s, a similar advance was achieved with *metals* [5–10], when Ralph, Black and Tinkham (RBT) succeeded to perform single-electron-tunneling spectroscopy on *individual ultrasmall metallic grains* (of radii $r \lesssim 5$ nm and mean level spacings $d \gtrsim 0.1$ meV): by attaching such a grain via oxide tunnel barriers to two leads they constructed a single-electron transistor having the grain as central island, and showed that a well-resolved, discrete excitation spectrum could indeed be extracted from the conductance. This opened up a new frontier in the study of electron correlations in metals, since the ability to resolve discrete energy levels allows the nature of electron correlations to be studied in unprecedented detail. During the last several years, single-electron-tunneling spectroscopy of ultrasmall metallic grains has been used to probe superconducting pairing correlations in Al grains [6,10], nonequilibrium excitations [10–13] and spin–orbit interactions [5,7,14–16] in normal grains, and ferromagnetic correlations in Co grains [17,18]. A brief overview of some of the early experiments is given in Ref. [9] by Ralph et al., a more thorough one in the thesis of Black [8].

Although tunnel-spectroscopic studies of metallic grains are similar in spirit to those of semiconductor quantum dots, there are a number of important differences:

(1) Metals have much *higher densities of states* than semiconductors (because the latter have smaller electron densities and effective masses), hence metals require much smaller sample sizes ($\lesssim 10$ nm) before discrete levels become resolvable.

(2) Consequently, metallic grains have much *larger charging energies* than quantum dots, which is an advantage when fluctuations in electron number are to be minimized. On the other hand, this also implies that for quantum dots the electron number can be varied over a much larger range than for metallic grains, which is useful for analyses of statistical properties.

(3) For metallic grains, the ability to study a range of materials throughout the periodic table, including samples doped with impurities (which tend to trap or deplete electrons in quantum dots) and alloys, allows some control over the strength and type of electron interactions to be studied. In particular, one may study *superconductivity* and itinerant *ferromagnetism*.

(4) For metallic grains, the tunnel barriers to the leads are insulting oxide layers and hence insensitive to applied voltages, whereas for quantum dots they are electrostatically defined and hence *tunable*. Tunability is usually an advantage, but not always: for example, *nonequilibrium effects* are easier to study quantitatively for a metallic grain than a quantum dot, because for the latter a large source–drain voltage lowers the tunnel barrier in poorly controlled ways.

(5) *Spin effects* are easily probed in metallic grains by applying a magnetic field and studying the Zeeman-splitting of time-reversed pairs of states, which for nm-scale metallic grains is a much stronger effect than that of the applied field on orbital properties. In contrast, in semiconductors the latter are dominant over spin effects.

(6) For the same reason, *spin–orbit* effects are more easily studied in metallic grains than in quantum dots.

A particularly interesting feature of experiments on ultrasmall metallic grains is that they probe the ways in which finite-size effects modify a system's characteristic correlations relative to their bulk properties, both due to mesoscopic fluctuations and via level discreteness. Such modifications arise when the new energy scale characterizing the spectrum's discreteness, namely the single-particle mean level spacing $d = 1/\mathcal{N}(\varepsilon_F) \sim 1/\text{Vol}$ (where $\mathcal{N}(\varepsilon)$ is the density of states per spin species), becomes comparable to the energy scale characterizing the correlations in bulk systems (such as the energy gap in superconductors or the Kondo temperature in magnetic alloys). Such quantum finite-size effects in metals had attracted considerable attention in the past, but could hitherto only be studied in ensemble-averaged quantities (for reviews, see [19,20]). Spectroscopic studies of discrete spectra of *individual* grains yield significant new information. Let us briefly mention the most important examples, organized according to the section in which they will be discussed in detail later.

RBT were able to determine the *number parity* (even or odd) of a given grain (Section 3), by studying the evolution of the discrete spectrum in an applied magnetic field: For an odd grain, the ground state energy was observed to Zeeman-split in a magnetic field H, as expected for a spin-$\frac{1}{2}$ Kramers doublet; moreover, for pure Al grains, the Landé g factor extracted from the size of the splitting was close to the expected value of $g^{\text{pure}} = 2$. In contrast, for an even grain the ground state is a nondegenerate spin singlet, and accordingly no ground state splitting was observed.

Parity effects were also observed in RBT's experiments on largish ($r \gtrsim 5\,\text{nm}$) Al grains [6,10,21]: an even grain had a distinct spectroscopic gap ($\gg d$) but an odd grain did not, which is clear evidence for the presence of *superconducting pairing correlations* in these grains (Section 4). The spectroscopic gap for even grains was driven to zero by an applied magnetic field, hence the paramagnetic breakdown of pairing correlations could be studied in detail. The corresponding theory was worked out by Braun, von Delft, Ralph and Tinkham [22,23].

In RBT's smallest grains ($r \lesssim 3$ nm), however, no such distinct spectroscopic gap could be discerned. This observation revived an old but fundamental question (Section 5): *What is the lower size limit for the existence of superconductivity in small grains?*, and thereby stimulated a large number of theoretical investigations [21–44]. Anderson [45] had addressed this question already in 1959, arguing that "superconductivity would no longer be possible" when the mean level spacing d becomes larger than the bulk gap, to be denoted by $\tilde{\Delta}$, because then $\tilde{\Delta}$ looses its special significance as gap in an otherwise continuous spectrum. RBT's new experiments stimulated a number of theoretical attempts to quantitatively describe the *crossover* from the bulk limit $d \ll \tilde{\Delta}$, where superconductivity is well developed, to the fluctuation-dominated regime of $d \gg \tilde{\Delta}$, where pairing correlations survive only in the form of weak fluctuations. Describing this crossover constituted a conceptual challenge, since the standard grand-canonical mean-field BCS treatment of pairing correlations [21–28] breaks down for $d \gtrsim \tilde{\Delta}$. This challenge elicited a series of increasingly sophisticated canonical treatments of pairing correlations [33–41], based on a simple reduced BCS-Hamiltonian for discrete energy levels, which showed that the crossover is completely smooth, but, interestingly, depends on the parity of the number of electrons on the grain, as pointed out by von Delft et al. [21]. Very recently, the main conclusions of these works were confirmed [39] using an exact solution of the reduced BCS model, discovered by Richardson in the context of nuclear physics in the 1960s [46–54]. (The existence of this solution came as a surprise – in the form of a polite letter from its inventor – to those involved with ultrasmall grains, since hitherto it had apparently completely escaped the attention of the condensed-matter community.)

An interesting finite-size effect was also revealed in *nonequilibrium* grains (Section 6), for which BRT observed the excitation spectrum to consist of *clusters of resonances*, with the spacing between clusters comparable to free-electron estimates of the single-particle mean level spacing d, but the spacings between subresonances of the same cluster much smaller than d. A theory for this effect was developed by Agam et al. [11,13], who showed that clusters of resonances can be caused by nonequilibrium excitations on the grain, provided that mesoscopic fluctuations of the matrix elements of the electron–electron interaction are sufficiently strong. Such fluctuations are neglected in the so-called "orthodox model" that is commonly used to describe single-electron transistors, but are expected, according to the general theory of disordered interacting electron systems (e.g. [13]), to become significant for sufficiently small grains, and can be described using random matrix theory. The experimental observation of such clusters thus constitutes a beautiful and direct illustration of the importance of mesoscopic fluctuations in ultrasmall grains.

A further example of such mesoscopic fluctuations was published by Salinas et al. [14], who experimentally studied the effect of the *spin–orbit interaction* (Section 7) in Al grains doped with Au, and by Davidović and Tinkham [15,16], who studied Au grains. These authors observed effective g^{eff} factors significantly smaller than $g^{\text{pure}} = 2$, as expected, since the spin–orbit interaction mixes states with opposite spin. Interestingly, the measured g^{eff} factors were also found to vary from one Kramers doublet to the next, which is a clear signature for mesoscopic fluctuations. Very recently, a theory for the statistics of the fluctuations of g^{eff} was worked out by Matveev et al. [55] and by Brouwer et al. [56].

During the last year, Guéron et al. [18] published the first set of detailed experimental results on *ferromagnetic Co grains* (Section 8). Among the novel features found for these are a strong asymmetry between the tunneling probabilities for spin-up or spin-down electrons, a larger-than-expected density of low-lying excitations, and hysteretic behavior of the excitation spectrum as

function of an applied magnetic field. The properties of the hysteresis loop reflect the changes in the direction of the grain's magnetic moment as the applied field is ramped. The latter features hold great promise for being a potential tool with which to study the dynamics of magnetization reversal in individual nm-scale ferromagnets. No detailed theory for these experiments exists at present.

An ultrasmall grain containing a single magnetic impurity can, in principle, be used to study the effect of level discreteness on the *Kondo effect* (Section 9). A theory for such a "Kondo box" was worked out by Thimm et al. [57]. They found that when the mean level spacing in the grain becomes larger than the Kondo temperature, the Kondo resonance is strongly affected in a way that depends on the parity of the number of electrons on the grain, and that should be detectable by tunnel-spectroscopic measurements.

The present review summarizes the above developments. The guiding principle in the choice of topics was to focus on the experimental studies of discrete spectra in ultrasmall metallic grains that have been carried out since 1995 using single-electron-tunneling spectroscopy, and on those theoretical developments that were directly inspired by them. The notation to be used throughout is introduced in Section 2, which describes how tunneling spectroscopy works in practice and reviews the well-known orthodox theory for single-electron transistors (also reviewed in, e.g., Refs. [58–60]), emphasizing those features that have special relevance for ultrasmall grains. The contents of Sections 3–9, which for the most part can be read independently of each other, were already outlined above. The main conclusions of each section are summarized concisely in Section 10, which also gives an outlook towards directions for further work.

For each subject, an attempt was made to give a detailed account of the arguments and calculations that have a direct bearing on understanding experimental data, and to qualitatively explain the theoretical ideas required for their interpretation. Theoretical developments beyond those of direct relevance to experiment are usually either summarized or only briefly mentioned, but seldom reproduced in detail. The fact that more than a third of the review is devoted to superconductivity in ultrasmall grains is a reflection of the number of papers that have appeared on this subject since 1995.

2. Single-electron-tunneling spectroscopy

The experiments in which Ralph, Black and Tinkham (RBT) succeeded to resolve discrete energy levels of an ultrasmall metallic grain used the technique of single-electron-tunneling spectroscopy. In Section 2.1 we explain the idea behind this technique and illustrate its capabilities by showing some representative data sets (without delving into the interesting physics contained in this data, which will be done in later sections). In Section 2.2 we estimate theoretically under which conditions discrete states should be resolvable, in Section 2.3 develop a theoretical description for ultrasmall single-electron transistors (SETs), and in Section 2.4 discuss experimental details concerning fabrication and measurement techniques.

2.1. Ultrasmall single-electron transistor

In the first generation of experiments of 1995 [5–7], a grain made from Al (a superconducting material) was connected to two metal leads via high-resistance tunnel junctions, with capacitances
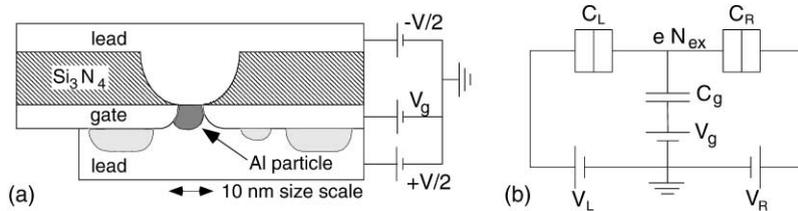
Fig. 1. (a) Schematic cross section of the ultrasmall SETs studied by RBT in [10], and (b) the corresponding circuit diagram.

C_L and C_R, say. In the next generation of 1997 [10], the grain was also coupled capacitatively to a gate, with capacitance C_g. The resulting device, schematically depicted in Fig. 1(a), has the structure of a SET, with the grain as central island. The circuit diagram for an SET is shown in Fig. 1(b). Applying a bias voltage V between the two leads causes a tunnel current I to flow between the leads through the grain, via incoherent sequential tunneling through the tunnel junctions. The current can be influenced by changing the gate voltage V_g (hence the name "transistor"), which tunes the electrostatic potential on the grain and thereby also its average number of electrons N. (For devices without a gate these two quantities cannot be tuned and instead have some sample-dependent, fixed value. For such devices, set $C_g = 0$ in all formulas below.)

The physics of SETs had been clarified in the early 1990s [3] through extensive studies of lithographically defined SETs of *mesoscopic* size, i.e. with *micron*-scale central islands. The fundamentally new aspect of RBT's work was that their SETs, made by a novel fabrication technique (described in Section 2.4), were *nanoscopic* in size: they had ultrasmall grains with radii between 15 and 2 nm as central islands, which were thus several orders of magnitude smaller in volume than in previous experiments. This had two important consequences:

(i) *The grain's charging energy $E_C \equiv e^2/2C$ was much larger than for mesoscopic SETs*, ranging roughly between 5 and 50 meV (where $C \equiv C_L + C_R + C_g$). E_C is the scale that determines the energy cost for changing N by one. Since for ultrasmall grains it far exceeds all other typical energy scales of the SET, such as those set by the bias voltage ($V \lesssim 1$ mV), the temperature ($T \lesssim 4.2$ K) and the bulk superconducting gap for Al ($\Delta_{\text{bulk}} = 0.18$ meV), fluctuations in electron number are strongly suppressed.

(ii) *Discrete eigenstates of the conduction electron energy spectrum became resolvable* – their mean level spacing d ranged from 0.02 to 0.3 meV, which is in order-of-magnitude agreement with the free-electron estimate of $d = 2\pi^2\hbar^2/(mk_F \text{Vol})$ for the single-particle level spacing. Such d-values are much larger than $k_B T$ for the lowest temperatures attained (around $T \simeq 30$ mK), but on the order of Δ_{bulk}. However, the number of conduction electrons for grains of this size is still rather large (between 10^4 and 10^5).

Since the two scales E_C and d differ by at least an order of magnitude, they manifest themselves in two distinct and easily separable ways in the low-temperature I–V curves of the devices, shown in Fig. 2 for a series of different V_g values:

(i) When V is varied on a large scale of tens of mV (Fig. 2), the I–V curves have a typical "Coulomb-staircase form" characteristic of SETs: zero current at low $|V|$ (the "Coulomb
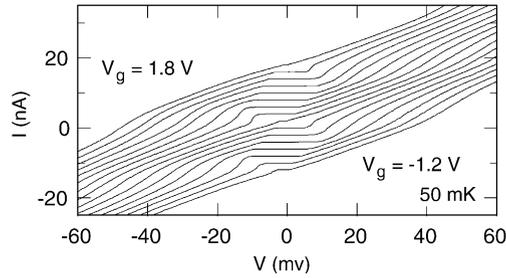
Fig. 2. Current–voltage curves for an ultrasmall SET [10] at 50 mK, artificially offset on the vertical axis for a set of equally spaced values of V_g between -1.2 and 1.8 V. The I–V curves display Coulomb-staircase structure on a bias voltage scale of tens of mV. By fitting these to standard SET theory [62], the SET's basic parameters can be determined: $R_L = 3.5$ MΩ, $R_R = 0.2$ MΩ, $C_L = 3.5$ aF, $C_R = 9.4$ aF, $C_g = 0.09$ aF, $E_C = 46$ meV. The grain radius and mean level spacing are estimated as $r \simeq 4.5$ nm and $d \simeq 0.45$ meV, using assumptions stated in Section 2.4.2.
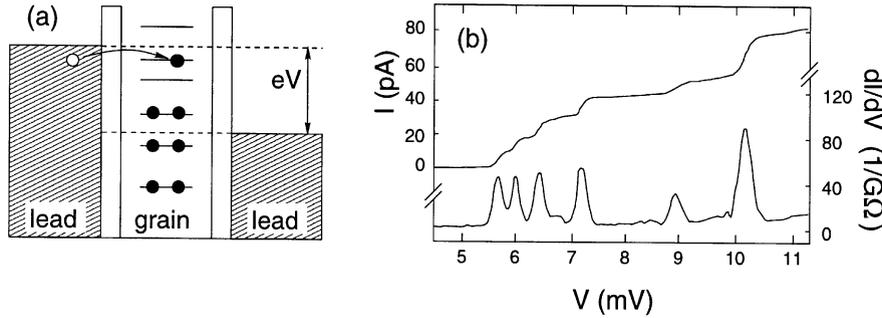
Fig. 3. (a) Cartoon of an dI/dV-measurement: The number of available transport channels through discrete states (here three) is determined by the bias voltage V. (b) The current and differential conductance as functions of bias voltage V for one of RBT's ultrasmall grains; beyond the Coulomb-blockade threshold at around 5.5 mV, the current displays fine steps and the conductance fine peaks, on a voltage scale of a few mV, reflecting the grain's discrete eigenspectrum.

blockade" regime), sloping steps equally spaced in V, and step thresholds sensitive to V_g. This proves that the tunnel current flows only through *one* grain. The maximal width of the flat step of zero current around $|eV| = 0$ is governed, in order of magnitude, by E_C and typically varies between 5 and 50 mV. As V_g is varied, the I–V curves periodically repeat, with a period e/C_g.

(ii) When V is varied on the much smaller scale of a few mV around the threshold of the Coulomb blockade regime and the temperature is sufficiently low ($T \ll d$), the I–V curves have a step-like substructure, shown in Fig. 3(b) [see also Fig. 5(b)]. As first pointed out by Averin and Korotkov [61], such small steps in the I–V curve are expected to occur whenever the voltage drop across one of the tunnel junctions equals the threshold energy at which the rate for tunneling across that junction into or out of one of the grain's *discrete energy eigenstates* becomes nonzero, since this opens up another channel for carrying current across that junction. Correspondingly, the differential conductance (dI/dV) curves contain a series of fine peaks, see Fig. 3(b). Under certain conditions described in Section 2.3.4(a), the distances between these peaks directly reflect the energy differences between eigenenergies of same-N eigenstates of the

grain. *Such conductance curves thus directly yield the grain's fixed-N excitation spectrum*, or more precisely, the set of energy differences

$$\delta\mathscr{E}_{\alpha\alpha'}^{N} = \mathscr{E}_{\alpha}^{N} - \mathscr{E}_{\alpha'}^{N} \tag{1}$$

for those eigenstates $|\alpha\rangle_N$ of the N-electron grain that are accessible final states for a tunneling process that removes or adds an electron to the grain if its initial ground state has $(N + 1)$ or $(N - 1)$ electrons, respectively.

The magnetic-field dependence of the fixed-N excitation spectrum can be obtained by simply tracing the motion of the conductance peak positions as a magnetic field is turned on (at fixed V_g). Examples are given in Fig. 8 in Section 3.1, Fig. 10 in Section 4.1, Fig. 31 in Section 7.1 and Fig. 34 in Section 8.1. If a gate is present, two further very interesting options exist: firstly, by tuning V_g by an amount large enough ($\simeq E_C/e$) to change N by one unit, the *influence on the spectrum of the parity* of the number of electrons on the grain can be studied. Secondly, by tuning V_g such that the Coulomb blockade regime is large or small, so that the V-threshold at which current begins to flow is large or small, *nonequilibrium effects can be maximized or minimized*, respectively, depending on whether one chooses to study them or not.

2.2. Conditions under which discrete states are resolvable

Having shown in Fig. 3 an example of the data by which RBT demonstrated that discrete energy levels of an ultrasmall metallic grain can be resolved, let us take a step back and recall the theoretical arguments, mainly due to Averin and Korotkov [61], for why this should be possible. This requires comparing various characteristic energy scales of a 3D grain and their dependence on its size, such as the mean level spacing d, the charging energy E_C, the Thouless energy E_{Thouless}, and the amount of level broadening due to inelastic relaxation processes (with rate Γ^{inel}) and tunneling (with rate Γ_{tun}).

The mean level spacing $d = 1/\mathscr{N}(\varepsilon_F)$ can be estimated using the free-electron expression for the density of states at the Fermi surface of a 3D grain:

$$d = \frac{2\pi^2\hbar^2}{mk_F\text{Vol}} = \frac{1.50\,\text{eV nm}^2}{k_F\text{Vol}} \ , \tag{2}$$

where, for example, $k_F = 17.5\,\text{nm}^{-1}$ for Al and $k_F = 12.1\,\text{nm}^{-1}$ for Au. For the sake of order-of-magnitude estimates below, we shall use $k_F = 10\,\text{nm}^{-1}$ and crudely assume the grain to have linear dimension r, volume r^3 and contact area r^2 (via tunnel junctions) with each lead.

Next, we consider the charging energy $E_C = e^2/2C$. The capacitances C_L and C_R to the left and right leads are each of order $C_\square r^2$, where $C_\square$ is the capacitance per unit area of the tunnel junctions, typically of order $C_\square \simeq 0.05\text{aF/nm}^2$ [63]; if a gate is present, C_g is typically an order of magnitude smaller. We therefore have

$$E_C \simeq \frac{e^2}{4C_\square r^2} \simeq \frac{0.8\,\text{eV nm}^2}{r^2} \ . \tag{3}$$

Although the mean level spacing increases faster (r^{-3}) with decreasing size than the charging energy (r^{-2}), they become comparable only for grains of almost atomic size ($r \simeq 2$ Å). For nm-scale

grains (say $r = 15$–2 nm), d is still significantly smaller (0.04–20 meV) than E_C (4–200 meV). Moreover, the total number of conduction electrons $N \propto \varepsilon_F/d$ is still very large ($\simeq 300\,000$–500, for $\varepsilon_F \simeq 10$ eV). As long as r is substantially larger than the Thomas–Fermi screening length $\lambda_{TF} \simeq 1/k_F$, nm-scale grains can still be regarded as metallic.

Since $d \ll E_C$ for nm-scale SETs, the *large-scale* features of their current–voltage characteristics (e.g. Fig. 2) can be correctly described [61] using continuous spectra for both island and leads and adopting the standard "orthodox" theory [58] for Coulomb-blockade phenomena in metallic SETs. For additional *fine-structure* due to a discrete spectrum to be resolvable, three more conditions must be met:

(i) To avoid thermal smearing, $T \ll d$ is required.

(ii) Γ_α^r, the tunneling rate out of a given discrete state $|\alpha\rangle$ on the grain into lead r ($=$ L, R), must be small enough that the tunneling-induced level widths $\hbar\Gamma_\alpha^r$ do not cause neighboring levels to overlap, i.e. $\hbar\Gamma_\alpha^r \ll d$; but this condition is equivalent to one which must hold anyway for Coulomb-blockade effects to occur [59,60], namely $G_r \ll 2e^2/h$, where $G_r = e^2 \mathcal{N}(\varepsilon_F)\Gamma_{tun}^r$ is, roughly, the total conductance across junction r, and Γ_{tun}^r the average of the Γ_α^r's near ε_F.

(iii) For an excited state on the grain with energy ε above the ground state to be resolvable, its inelastic relaxation rate $\Gamma_\varepsilon^{inel}$ must be small enough that the corresponding line width is less than the level spacing, $\hbar\Gamma_\varepsilon^{inel} \lesssim d$. The question when this is satisfied depends (neglecting phonons) on the grain's so-called "dimensionless conductance" g_{dim}, defined for finite systems by [64]

$$g_{dim} \equiv E_{Thouless}/d \,, \tag{4}$$

$$E_{Thouless} \simeq \begin{cases} \hbar D_{diff}/r^2 = (0.25 k_F l_{tr}/r^2)\,\text{meV nm}^2 & \text{(diffusive)}\,, \\ \hbar v_F/(a2r) = (38 k_F/ar)\,\text{meV nm}^2 & \text{(ballistic)}\,. \end{cases} \tag{5}$$

Here the Thouless energy $E_{Thouless}$ is the inverse time for an electron near ε_F to travel once across the system, $D_{diff} = v_F l_{tr}/3$ is the diffusion constant, l_{tr} the transport mean free path, and a a geometrical constant. Eqs. (5) are order-of-magnitude estimates[1] for diffusive or ballistic 3D grains, for which $g_{dim} \propto r$ or r^2, respectively. It was shown in Refs. [66,64] that for grains so small[2] that $g_{dim} \lesssim 1$, the spectrum of low-lying excitations consists of δ-function like peaks with $\hbar\Gamma_\varepsilon^{inel} \ll d$ that can definitely be resolved [64], whereas for larger grains with $g \gg 1$, the condition $\hbar\Gamma_\varepsilon^{inel} \lesssim d$ can be

---

[1] The expressions for g_{dim} and $E_{Thouless}$ given by various authors differ by factors of order unity. (For g_{dim} we followed [64], and for $E_{Thouless}$ [55] for the diffusive and [65,15] for the ballistic case.) Note that the ballistic definition for $E_{Thouless}$ is similar to the diffusive one, except that the grain size determines the effective mean-free path ($l_{tr} \to 3r/2a$). The proper choice of a depends on the precise geometry and the amount of disorder present [13]; in practice, it is somewhat of a "fudge factor". For a hemispherical grain, Davidović and Tinkham [15] used $a = 3$ (the number of dimensions); they argue that if the electronic motion in the grain is ballistic and surface scattering diffusive, an electron's mean-free path is approximately $2r$, but to diffuse through the entire particle volume, it should scatter from the surface about three times [15]. For a pancake-shaped grain with thickness z, radius r and diffusive dynamics, Agam et al. [11] argued that $a \propto (r/z)^2$.

[2] For a ballistic, hemispherical Al grain, for example, Eqs. (2)–(5) yield $g_{dim} = (16/a)(r/\text{nm})^2$, which, for $a = 3$, gives $g_{dim} \le 1$ if $r \le 0.4$ nm. An experimental estimate [11] of g_{dim} for two ballistic Al grains with Vol ≈ 40 and 100 nm^3 yielded values of $g_{dim} \approx 5$ in both cases, as discussed in Section 6.1.3. If these grains are assumed to be hemispherical, this would imply very large a factors of ≈ 24 and 42; this might imply that an assumption of pancake-shaped grains would be more appropriate, for which a does contain a large factor $(r/z)^2$, cf. footnote 1 and, in Section 6.1.3, footnote 16.

met as long as $\varepsilon \lesssim E_{\text{Thouless}}$. For grains with $d \ll E_{\text{Thouless}} \ll E_C$, the fine structure due to level discreteness will thus be smeared out for $\varepsilon \approx E_C$, but should be clearly resolvable for $\varepsilon \lesssim$ several units of d. (For a more detailed discussion of $\Gamma_\varepsilon^{\text{inel}}$, including the effect of phonons, see Section 6.3.)

For future reference, we remark here that the parameter g_{dim} also controls the importance or not of corrections to the orthodox model, which can be shown [67–71,13] to be small as $1/g_{\text{dim}}$ (see Section 6.1.3).

Finally, note that the amount of nonequilibrium effects on the grain is controlled by the ratio of Γ_{tun}^r to $\Gamma_\varepsilon^{\text{inel}}$. If $\Gamma_{\text{tun}}^r \ll \Gamma_\varepsilon^{\text{inel}}$, nonequilibrium effects are negligible, since an electron tunneling into an excited state of the grain has plenty of time to relax before the next tunneling event. Conversely, nonequilibrium effects should become very important for $\Gamma_{\text{tun}}^r \gg \Gamma_\varepsilon^{\text{inel}}$. Indeed, they have been observed directly [10–13] and are discussed in Section 6.

2.3. Theoretical description of an ultrasmall SET

In this section we set up a general formalism for describing transport through ultrasmall metallic grains, in order to show explicitly how level discreteness causes steps in the tunneling current and to clarify precisely what information can be extracted from the latter. Since ultrasmall grains may, due to their discrete states, be regarded as "metallic quantum dots", we can carry over much of the formalism developed in the literature for quantum dots; indeed, in spirit and notation we shall very closely follow a review by Schoeller entitled "Transport Theory of Interacting Quantum Dots" [72]. Although this material may be familiar to many readers, it is needed here to establish the notation to be used throughout this review.

2.3.1. Charging energy and Coulomb blockade
We start by deriving the charging energy for the grain, using the so-called "orthodox" or "Coulomb-blockade" model [58], which assumes the electrostatic potential to be *homogeneous* on the grain. This assumption was argued to be reasonable as long as the Thomas–Fermi screening length $\lambda_{\text{TF}} \simeq 1/k_F$ is very much shorter than the grain's linear dimensions; a more rigorous condition is $g_{\text{dim}} \gg 1$ [67–71,13]. Nevertheless, deviations from the predictions of the orthodox model can be expected in principle, and indeed were observed in some of RBT's smallest grains with $g_{\text{dim}} \simeq 5$ (see Section 6).

Consider the SET shown in Fig. 1(b), and let $E_{\text{pot}}(N_{\text{ex}})$ denote the electrostatic work required to add N_{ex} excess electrons with a total charge of $Q_{\text{ex}} = e N_{\text{ex}}$ (with $e < 0$) to a grain with initial random off-set charge Q_0, while the time-independent voltages V_L, V_R and V_g of the left and right leads and the gate electrode, respectively, are held fixed. Within the orthodox model, $E_{\text{pot}}(N_{\text{ex}}) = \int_{Q_0}^{Q_{\text{ex}} + Q_0} dQ \, V(Q)$, where $V(Q)$ is the electrostatic potential of the grain for given grain charge Q. It is determined by $C_r[V_r - V(Q)] = Q_r$, where Q_r is the screening charge on capacitor $r(= L,R,g)$ and $Q_L + Q_R + Q_g = -Q$. (In the absence of a gate one has $Q_g = 0$ and hence should set $C_g = 0$.) Using the definitions $C \equiv C_L + C_R + C_g$ and $q_D \equiv \sum_{r = L,R,g} C_r V_r$, we obtain $V(Q) = (q_D + Q)/C$ and thus

$$E_{\text{pot}}(N_{\text{ex}}) = Q_{\text{ex}}(q_D + Q_0)/C + Q_{\text{ex}}^2/2C = e V_D N_{\text{ex}} + E_C N_{\text{ex}}^2 \ . \tag{6}$$

Since the first term on the r.h.s. is linear in the number of excess electrons, $V_D \equiv (q_D + Q_0)/C$ can be viewed as the electrostatic potential on the grain (the subscript D, for "dot", is used to conform to

Schoeller's notation [72]). For gated devices, Q_0 may be absorbed into q_D by shifting V_g by an amount $-Q_0/C_g$, but for gateless ones, it must be treated as a fit parameter. The second term on the r.h.s. of Eq. (6) represents the Coulomb interaction energy of the N_{ex} excess electrons due to their mutual repulsion, and its scale is set by the charging energy $E_C \equiv e^2/2C$.

E_{pot} is often written as $E_C(N_{ex} - n_D)^2 - E_C n_D^2$, where $n_D \equiv -eV_D/2E_C$, since this makes it clear that for a given value of n_D, the system will adjust N_{ex} to be the integer closest to n_D, in order to minimize E_{pot}. The particle number on the grain can thus be controlled in discrete units by varying n_D via the gate voltage V_g. Since the electrostatic energy difference

$$\delta E_{pot}^\pm(N_{ex}) \equiv E_{pot}(N_{ex} \pm 1) - E_{pot}(N_{ex}) \tag{7}$$

between grains with $N_{ex} \pm 1$ or N_{ex} excess electrons vanishes when n_D is tuned to lie half-way between $N_{ex} \pm 1$ and N_{ex}, half-integer values of n_D are called "degeneracy points". Transport is possible through a grain whose ground state has N_{ex} electrons only if, roughly speaking, $\min[\delta E_{pot}^\pm(N_{ex})] \lesssim \max[k_B T, |eV|]$, i.e. only if the "Coulomb barrier" presented by Eq. (7) can be overcome by the temperature or bias voltage. If both of these are small ($k_B T, |eV| \ll E_C$), a complete suppression of transport through the grain, the so-called "Coulomb blockade", occurs far away from the degeneracy points; in particular, the grain's low-temperature, linear-response conductance shows "Coulomb oscillations" as function of V_g, i.e. a series of peaks, with a uniform n_D-spacing of 1, i.e. V_g-spacing of e/C_g. In between degeneracy points, low-temperature transport is possible only with a large bias voltage, which will, in general, lead to nonequilibrium effects.

2.3.2. General Hamiltonian

To describe transport through the grain, we shall adopt a Hamiltonian $\hat{H} = \hat{H}_L + \hat{H}_R + \hat{H}_D + \hat{H}_{tun}$ of the following rather general form:

$$\hat{H}_r = \sum_{k\sigma} (\varepsilon_{kr} + eV_r)c_{k\sigma r}^\dagger c_{k\sigma r} \quad (r = L, R), \tag{8}$$

$$\hat{H}_D = \sum_\alpha (eV_D N_{ex,\alpha} + \mathcal{E}_\alpha)|\alpha\rangle\langle\alpha|, \tag{9}$$

$$\hat{H}_{tun} = \sum_{r=L,R} \sum_{kl\sigma} T_{kl\sigma}^r c_{k\sigma r}^\dagger c_{l\sigma D} + (h.c.), \tag{10}$$

$$= \sum_{r=L,R} \sum_{k\sigma,\alpha\alpha'} T_{k\sigma,\alpha\alpha'}^r c_{k\sigma r}^\dagger |\alpha\rangle\langle\alpha'| + (h.c.), \tag{11}$$

$$T_{k\sigma,\alpha\alpha'}^r = \sum_l T_{kl\sigma}^r \langle\alpha|c_{l\sigma D}|\alpha'\rangle. \tag{12}$$

Here $c_{k\sigma r}^\dagger$ creates an electron in lead r, in a single-particle state with spin σ, kinetic energy ε_{kr} (measured relative to the Fermi energy $\varepsilon_{k_F r}$ of that lead) and electrostatic energy eV_r. $|\alpha\rangle$ is a many-body eigenstate of the *isolated* grain in the absence of tunneling, with definite total electron number N_α. The sum on α in Eq. (9) is not restricted to a fixed-N_α Hilbert space, but is over *all* eigenstates in the isolated grain's Fock space. The eigenvalue $eV_D N_{ex,\alpha} + \mathcal{E}_\alpha$ of $|\alpha\rangle$ has been split

into two parts: the first contains, via V_D, all dependences on the external voltages arising from E_{pot} of Eq. (6), while $\mathscr{E}_\alpha$ is the voltage-independent remainder and includes the interaction contribution $(E_C N_{ex,\alpha}^2)$ to E_{pot}. In the absence of many-body correlations, e.g. for a grain made from a normal metal, one could take, as Averin and Korotkov did [61],

$$\hat{H}_D^{normal} = E_{pot}(\hat{N}_{ex}) + \sum_{l\sigma} \varepsilon_{l\sigma D} c_{l\sigma D}^\dagger c_{l\sigma D} \; , \tag{13}$$

$$|\alpha\rangle = |\{n_{l\sigma D}^{(\alpha)}\}\rangle, \quad \mathscr{E}_\alpha = E_C N_{ex,\alpha}^2 + \sum_{l\sigma} \varepsilon_{l\sigma D} n_{l\sigma D}^{(\alpha)} \; . \tag{14}$$

Here $c_{l\sigma D}^\dagger$ is a creation operator for one of the grain's single-particle states with energy $\varepsilon_{l\sigma D}$, and each many-body eigenstate $|\alpha\rangle$, with eigenenergy $eV_D N_{ex,\alpha} + \mathscr{E}_\alpha$, is specified by a set $\{n_{l\sigma D}^{(\alpha)}\}$ of single-particle occupation numbers. However, since such a single-particle description fails, e.g., for superconducting or ferromagnetic grains, we shall treat the $|\alpha\rangle$ as general many-body eigenstates below.

The tunneling term $\hat{H}_{tun}$ has been written in two forms: Eq. (10) describes the tunneling of electrons between *single-particle* states in the leads and grain, with matrix elements $T_{kl\sigma}^r$ for tunneling events involving lead r; and in Eq. (11), $\hat{H}_{tun}$ is represented in the $|\alpha\rangle$ eigenbasis, with corresponding matrix elements $T_{k\sigma,\alpha\alpha'}^r$ given by Eq. (12). When doing perturbation theory in $\hat{H}_{tun}$, it turns out [72] that the $T_{kl\sigma}^r$ matrix elements always occur in the combination

$$\frac{2\pi}{\hbar} \sum_k T_{kl\sigma}^{r*} T_{kl'\sigma}^r \delta(\omega - \varepsilon_{kr}) \simeq \delta_{ll'} \Gamma_{l\sigma}^r \; , \tag{15}$$

which represents the total tunneling rate from lead r across barrier r into the single-particle state $|l\sigma\rangle$ on the grain. On the right-hand side, we have neglected all off-diagonal terms, since for $l \neq l'$ the sum on k would involve matrix elements with randomly varying phases that average to zero, and also the ω dependence, which typically is of order ω/ε_F (since ε_F sets the energy scale for changes in the leads' density of states).

2.3.3. Tunneling current and master equation.

The operator for the current carried by electrons tunneling between lead r and the grain, $\hat{I}_r = -e \, \partial \hat{N}_r / \partial t$ (defined to be positive if N_r increases), is

$$\hat{I}_r = ie[\hat{N}_r, \hat{H}] = ie \sum_{k\sigma,\alpha\alpha'} T_{k\sigma,\alpha\alpha'}^r c_{k\sigma r}^\dagger |\alpha\rangle\langle\alpha'| + (\text{h.c.}) \; . \tag{16}$$

Its steady-state expectation value $I_r = \text{Tr}_{L,R,\alpha}(\hat{\rho}\hat{I}_r)$, in general, requires knowledge of the system's full density matrix $\hat{\rho}$. However, since the Hamiltonian is quadratic in the lead degrees of freedom, the latter can be integrated out, so that the remaining degrees of freedom are described by the reduced density matrix $\hat{P} = \text{Tr}_{L,R}(\hat{\rho})$ for the grain. Its diagonal elements $P_\alpha = \langle\alpha|\hat{P}|\alpha\rangle$ give the probability to find the grain in state $|\alpha\rangle$. They satisfy, and may be found by numerically solving, a normalization condition and master equation of the form

$$\sum_{\alpha'} P_{\alpha'} = 1, \quad 0 = \sum_{\alpha' \neq \alpha} (\Sigma_{\alpha\alpha'} P_{\alpha'} - \Sigma_{\alpha'\alpha} P_\alpha) \quad \text{for each } \alpha \; , \tag{17}$$

where $\Sigma_{\alpha\alpha'}$ (for $\alpha \neq \alpha'$) is the total transition rate from initial state $|\alpha'\rangle$ to final state $|\alpha\rangle$. Furthermore, the current I_r can be written as

$$I_r = e \sum_{\alpha\alpha'} \sum_{p=1}^{\infty} p(\Sigma_{\alpha\alpha'}^{r\,+p} - \Sigma_{\alpha\alpha'}^{r\,-p})P_{\alpha'} \;, \tag{18}$$

where $\Sigma_{\alpha\alpha'}^{r\,+p}(\Sigma_{\alpha\alpha'}^{r\,-p})$ is that part of the total rate $\Sigma_{\alpha\alpha'}$ that involves the coherent transfer of a total of p electrons onto (from) the grain from (onto) lead r. Of course, charge conservation ensures that $I_L = I_R$.

Eqs. (17) and (18) have intuitively plausible forms, but the calculation of the rates $\Sigma_{\alpha\alpha'}$ and $\Sigma_{\alpha\alpha'}^{r\,\pm\,p}$ is in general a highly nontrivial task. The standard strategy is to perform an expansion in powers of $\hat{H}_{\text{tun}}$, the most general and systematic formulation of which is the so-called real-time diagrammatic approach developed by Schoeller, Schön and König [72–78]. Fortunately, for our present purpose of analysing tunnel-spectroscopic measurements on ultrasmall grains, we may restrict ourselves to the simplest possible situation: the tunnel barriers are so large and the tunneling current so small that it suffices to consider only *sequential* tunneling[3] of electrons, described by lowest-order perturbation theory in H_{tun}. For this case one has [77,78]

$$\Sigma_{\alpha\alpha'} = \sum_{r=L,R} \sum_{p=\pm} \Sigma_{\alpha\alpha'}^{rp} \;, \tag{19}$$

where $\Sigma_{\alpha\alpha'}^{r\pm} \equiv \Sigma_{\alpha\alpha'}^{r\,\pm 1}$ can be calculated using the golden rule

$$\Sigma_{\alpha\alpha'}^{r+} = \frac{2\pi}{\hbar} \sum_{k\sigma} f(\varepsilon_{kr}) |T_{k\sigma,\alpha'\alpha}^{r*}|^2 \delta(\mathscr{E}_\alpha - \mathscr{E}_{\alpha'} + eV_D - \varepsilon_{kr} - eV_r) \tag{20}$$

$$= f(\mathscr{E}_\alpha - \mathscr{E}_{\alpha'} - e\bar{V}_r)\Gamma_{\alpha\alpha'}^{r+} \;, \tag{21}$$

$$\Sigma_{\alpha\alpha'}^{r-} = \frac{2\pi}{\hbar} \sum_{k\sigma} [1 - f(\varepsilon_{kr})] |T_{k\sigma,\alpha\alpha'}^{r}|^2 \delta(\mathscr{E}_\alpha - \mathscr{E}_{\alpha'} - eV_D + \varepsilon_{kr} + eV_r) \tag{22}$$

$$= f(\mathscr{E}_\alpha - \mathscr{E}_{\alpha'} + e\bar{V}_r)\Gamma_{\alpha\alpha'}^{r-} \;, \tag{23}$$

$$\Gamma_{\alpha\alpha'}^{r+} = \sum_{l\sigma} \Gamma_{l\sigma}^{r} |\langle\alpha| c_{l\sigma D}^{\dagger} |\alpha'\rangle|^2, \qquad \Gamma_{\alpha\alpha'}^{r-} = \sum_{l\sigma} \Gamma_{l\sigma}^{r} |\langle\alpha| c_{l\sigma D} |\alpha'\rangle|^2 \;. \tag{24}$$

Here $f(E) = 1/(e^{E/k_B T} + 1)$ is the Fermi function and $\bar{V}_r = V_r - V_D$ is the voltage drop (electrostatic potential difference) between lead r and the grain. For the standard case of a "symmetric" circuit with $V_L = -V_R = V/2$, which we shall henceforth adopt, it is given by

$$\bar{V}_{L/R} = [\pm V(C_{R/L} + C_g/2) - V_g C_g - Q_0]/C \;, \tag{25}$$

which reflects the capacitative division of the bias voltage $V(= V_L - V_R = \bar{V}_L - \bar{V}_R)$ across the two barriers. Eqs. (21) and (23) show that when V is swept, at a fixed gate voltage V_g and at temperatures much lower than the typical spacing d between the eigenenergies $\mathscr{E}_\alpha$, a rate $\Sigma_{\alpha\alpha'}^{r\pm}$ will be switched from "off" (exponentially small) to "on" (or order $\Gamma_{\alpha\alpha'}^{r\pm}$) each time $e\bar{V}_r$ passes through

---

[3] An exception will be encountered in Section 6.2, where cotunneling needs to be considered.

a threshold at which one of the inequalities

$$\pm e\bar{V}_r \geq \mathscr{E}_\alpha - \mathscr{E}_{\alpha'} \tag{26}$$

becomes true. Intuitively speaking, this occurs each time the energy gained by an electron when leaving (entering) lead r, namely $\pm eV_r$, becomes greater than the energy needed to enter (leave) the grain while inducing the transition $|\alpha'\rangle \to |\alpha\rangle$, namely $\mathscr{E}_\alpha - \mathscr{E}_{\alpha'} \pm eV_\mathrm{D}$. Figs. 4(a) and (b) illustrate this condition for the rates $\Sigma^{\mathrm{L}+}$ and $\Sigma^{\mathrm{L}-}$, respectively, and the case that the initial state is the ground state, $|\alpha'\rangle = |G\rangle$.

The standard expressions for a *meso*scopic (i.e. large) SET with a normal island, described by $\hat{H}_\mathrm{D}^{\mathrm{normal}}$ of Eq. (13), can be recovered as follows from the above formulas: firstly, one uses a factorized form for the island's density matrix, with $P_\alpha = P_N P_{\alpha_N}$ (and $\sum_{\alpha_N} P_{\alpha_N} = 1$). Secondly, one makes the replacement

$$\sum_{\alpha_{N\pm1}\alpha'_N} (\Sigma^{r+}_{\alpha_{N\pm1}\alpha'_N}) P_{\alpha'_N} \to \sum_{l\sigma} f(\pm\varepsilon_{l\sigma\mathrm{D}} + \delta E^\pm_{\mathrm{pot}} \mp eV_r) f(\mp\varepsilon_{l\sigma\mathrm{D}}) \Gamma^r_{l\sigma} \tag{27}$$

$$= 2\Gamma^r \mathscr{N}(\varepsilon_\mathrm{F}) \left[\frac{-(\delta E^\pm_{\mathrm{pot}} \mp eV_r)}{1 - \mathrm{e}^{(\delta E^\pm_{\mathrm{pot}} \mp eV_r)/k_\mathrm{B}T}} \right], \tag{28}$$

where $\delta E^\pm_{\mathrm{pot}}$ [Eq. (7)] is the change in electrostatic energy for the transition $N_{\mathrm{ex}} \to N_{\mathrm{ex}} \pm 1$. For the second line we took the continuum limit for the energies $\varepsilon_{l\sigma\mathrm{D}}$, and assumed that $\Gamma^r_{l\sigma} = \Gamma^r$ is independent of l and σ.

2.3.4. Extracting the fixed-N excitation spectrum

The very large charging energies ($E_\mathrm{C} \gtrsim 5\,\mathrm{meV}$, $\gg T, eV$) of ultrasmall grains ensure that for a fixed value of gate voltage V_g, at most two kinds of charge states are involved in electron
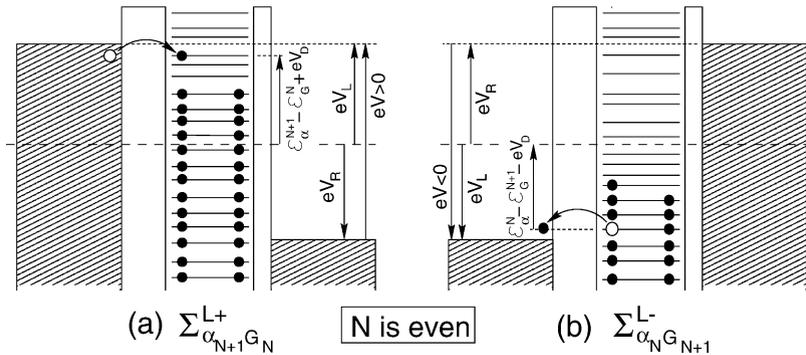
Fig. 4. Schematic illustration of the threshold conditions [see Eq. (26)] for bottleneck rates to become nonzero: (a) $eV_r \geq \mathscr{E}^{N+1}_\alpha - \mathscr{E}^N_G + eV_\mathrm{D}$ for $\Sigma^{r+}_{\alpha_{N+1},G_N}$ and (b) $-eV_r \geq \mathscr{E}^N_\alpha - \mathscr{E}^{N+1}_G - eV_\mathrm{D}$ for $\Sigma^{r-}_{\alpha_N G_{N+1}}$, for $r = L$ and the case that the left barrier is larger than the right ($\Gamma^L \gg \Gamma^R$). The long-dashed line indicates the equilibrium, $V = 0$ chemical potential of the L and R leads relative to which eV_r and $\mathscr{E}_\alpha - \mathscr{E}_G \pm eV_\mathrm{D}$ are drawn, using upward or downward arrows for positive and negative energies, respectively. Note that excitation energies are always positive: (a) $\mathscr{E}^{N+1}_\alpha - \mathscr{E}^N_G + eV_\mathrm{D}$ for particle-like excitations produced when an electron tunnels onto the grain; and (b) $\mathscr{E}^N_\alpha - \mathscr{E}^{N+1}_G - eV_\mathrm{D}$ for hole-like excitations when it tunnels off. Filled circles depict the *final* electron configuration $|\alpha\rangle$ *after* such a tunneling process, i.e. the electron number in the initial ground state before tunneling is drawn to be even (N) in (a) and odd ($N + 1$) in (b).

transport through the grain, i.e. $P_\alpha = 0$ unless N_α equals N or $N + 1$ (say). To get a nonzero current, one needs nonzero rates for tunneling both onto and off the grain, through different barriers, say $\Sigma^{r+}_{\alpha_{N+1}\alpha'_N}$ and $\Sigma^{r'-}_{\alpha_N\alpha'_{N+1}}$. To be specific, these will be of type

$$\Sigma^{\mathrm{L}+}, \Sigma^{\mathrm{R}-} \text{ for } eV > 0 \quad \text{and} \quad \Sigma^{\mathrm{L}-}, \Sigma^{\mathrm{R}+} \text{ for } eV < 0 \quad [\text{with } e < 0] \,, \tag{29}$$

since the direction of electron tunneling is always from the higher toward the lower chemical potential. One can distinguish two situations, depending on whether (a) only one or (b) both barriers are bottlenecks that limit transport.

(a) *One bottleneck*: If all the tunneling-on rates Σ^{r+} are generally, say, much smaller than any of the tunneling-off rates $\Sigma^{r'-}$, then barrier r is the bottleneck for transport. The probability to find N or $N + 1$ electrons on the grain will then be close to unity or zero, respectively. Moreover, for $T \ll d$ (and neglecting nonequilibrium effects, to be discussed in Section 6), the N-electron ground state $|G_N\rangle$ will be the overwhelmingly most probable, i.e. $P_{G_N} \simeq 1$ and $P_\alpha \simeq 0$ for all other α. In this case the current through barrier r [Eq. (18)] is simply $I_r \simeq e\sum_{\alpha_{N+1}}(\Sigma^{r+}_{\alpha_{N+1}G_N})$. If V is swept at fixed V_g, the current will thus show a step-like increase and the differential conductance a peak each time a rate $\Sigma^{r+}_{\alpha_{N+1}G_N}$ is switched on, i.e. each time $e\bar{V}_r$ increases past one of the threshold values $\mathscr{E}^{N+1}_\alpha - \mathscr{E}^N_G$ of (26), see Fig. 4(a). Denote this threshold value by $\bar{V}^\alpha_r$ and the corresponding threshold bias voltage [related to it via (25)] by V_α. The voltage differences $V_{\alpha\tilde{\alpha}} = V_\alpha - V_{\tilde{\alpha}}$ between any two conductance peaks can directly be translated into one of the excitation energies of the *fixed-$(N + 1)$* excitation spectrum, namely (for bottleneck barrier $r = \mathrm{L/R}$)

$$\mathscr{E}^{N+1}_\alpha - \mathscr{E}^{N+1}_{\tilde{\alpha}} = e(\bar{V}^\alpha_{\mathrm{L/R}} - \bar{V}^{\tilde{\alpha}}_{\mathrm{L/R}}) = \pm eV_{\alpha\tilde{\alpha}}(C_{\mathrm{R/L}} + C_\mathrm{g}/2)/C \,. \tag{30}$$

The case in which all rates Σ^{r-} are much smaller than any rate $\Sigma^{r'+}$ [Fig. 4(b)] is completely analogous to the one just discussed. Then the current is $I_r \simeq -e\sum_{\alpha_N}(\Sigma^{r-}_{\alpha_N G_{N+1}})$, and the fixed-$N$ excitation energies are given by (for $r = \mathrm{L/R}$)

$$\mathscr{E}^N_\alpha - \mathscr{E}^N_{\tilde{\alpha}} = -e(\bar{V}^\alpha_{\mathrm{L/R}} - \bar{V}^{\tilde{\alpha}}_{\mathrm{L/R}}) = \mp eV_{\alpha\tilde{\alpha}}(C_{\mathrm{R/L}} + C_\mathrm{g}/2)/C \,. \tag{31}$$

Eqs. (30) or (31) form the basis of RBT's tunnel-spectroscopic measurements of discrete spectra: a conductance curve showing a series of peaks as function of voltage can be replotted as function of energy by rescaling the horizontal axis using the voltage-to-energy conversion factors[4] $e(C_{\mathrm{R/L}} + C_\mathrm{g}/2)/C$ in Eqs. (30) or (31). The rescaled curves, which we shall henceforth generically call "fixed-N excitation spectra", where N is the number of electrons in the *final* state *after* the bottleneck tunneling process, then allow the corresponding excitation energies to be read off directly from the peak spacings.

(b) *Two bottlenecks*: The situation is more involved if the tunneling-on and tunneling-off rates Σ^{r+} and Σ^{r-} are comparable in magnitude so that both barriers are bottlenecks: Then some of the P_{α_N} and $P_{\alpha_{N+1}}$ will be comparable too, implying the same for the total probabilities to find N or $N + 1$ electrons on the grain. The conductance will then show peaks corresponding to the switching-on of both Σ^{r+} and $\Sigma^{r'-}$ rates, at both $e\bar{V}_r$ and $e\bar{V}_{r'}$ values corresponding to both $\mathscr{E}^{N+1}_\alpha - \mathscr{E}^N_{\alpha'}$ and $\mathscr{E}^N_\alpha - \mathscr{E}^{N+1}_{\alpha'}$ energy differences. It would thus be rather difficult to extract a purely fixed-N or fixed-$(N + 1)$ excitation spectrum from the resulting mixed set of peak spacings (though

---

[4] If no gate is present ($C_\mathrm{g} = 0$) or if $C_\mathrm{g} \ll C_\mathrm{L}, C_\mathrm{R}$ (as in the device of Fig. 2), this factor reduces to $eC_{\mathrm{R/L}}/(C_\mathrm{L} + C_\mathrm{R})$.

in principle not impossible: if a magnetic field is used to switch the leads from being superconducting to normal, peaks due to Σ^{r+} or $\Sigma^{r'-}$ shift by different amounts, see Eq. (34) in Section 2.4.3; or, if a gate is present, peaks due to Σ^{r+} or $\Sigma^{r'-}$ move in opposite V-directions if V_g is changed, see Eq. (35) in Section 2.4.4).

2.4. Experimental details

This section is devoted to experimental details such as device fabrication and characterization, and what effects superconducting leads or a gate voltage have on the tunneling spectra. In particular, we explain how one can tell whether all conductance peaks correspond to tunneling across the *same* grain, whether a given peak has barrier L or R as bottleneck, and how the corresponding capacitance ratios can be determined very accurately [this knowledge is needed in Eqs. (30) or (31)].

2.4.1. Device fabrication
The devices of RBT shown schematically in Fig. 1(a) were fabricated as follows [10]: First electron-beam lithography and reactive-ion etching were used to make a bowl-shaped hole in a suspended silicon nitride membrane, with an orifice between 5 and 10 nm in diameter [79]. The gate electrode was formed by evaporating 12 nm of Al onto the flat [bottom in Fig. 1(a)] side of the membrane. Plasma anodization and deposition of insulating SiO were then used to provide electrical isolation for the gate. Next, an Al electrode which fills the bowl-shaped side [top in Fig. 1(a)] of the nitride membrane was formed by evaporation of 100 nm of Al, followed by oxidation in 50 mTorr O_2 for 45 s to form a tunnel barrier near the lower opening of the bowl-shaped hole. A layer of nm-scale grains was created by depositing 2.5 nm of Al onto the lower side of the device; due to surface tension the metal beaded up into separate grains less than 10 nm in diameter [80,81]. In approximately 25% of the samples (determined as those exhibiting the typical SET "Coulomb-staircase" structure exemplified by Fig. 2), a single grain formed under the nm-scale tunnel junction to contact the top Al electrode. Finally, after a second oxidation step to form a tunnel junction on the exposed surface of the grain, a lower electrode is formed by evaporating 100 nm of Al to cover the grain.

The resulting device is an SET whose central island is a nm-scale grain fully coated by an insulating oxide layer, sitting on the lower lead electrode, encircled by a gate electrode, with the bowl part of the upper lead electrode at a fixed distance above it like "an STM tip cast in concrete".

Recently, Davidović and Tinkham succeeded in contacting leads to an ultra-small grain with a radius as small as $r \approx 1$ nm, using a somewhat different fabrication technique, described in detail in [82,15].

2.4.2. Device characterization
Only those devices are studied in detail whose I–V characteristics display the large-scale Coulomb-staircase (Fig. 2) expected for a SET, since its presence indicates that the tunneling current flows through a single grain only (as opposed to several). The basic parameters of the SET (R_L, R_R, C_L, C_R, C_g and E_C) can be determined by fitting this Coulomb-staircase to the standard expressions for mesoscopic SETs (with $C_g = 0$ for devices without a gate), as described in detail by Hanna and Tinkham [62]. To estimate the grain radius, RBT crudely assumed [5–7,10] the grain
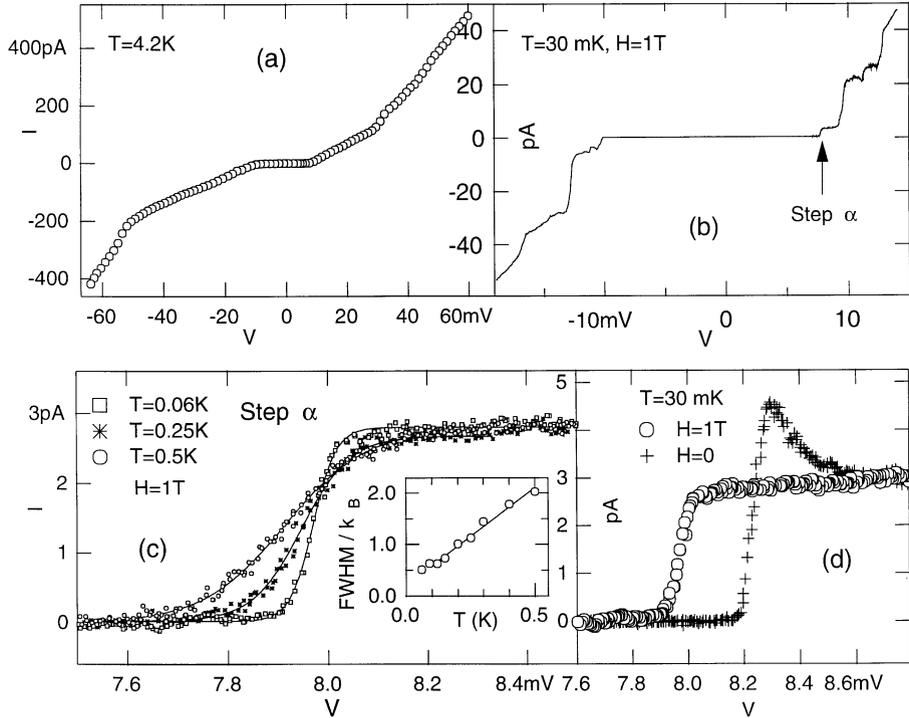
Fig. 5. Current–voltage characteristics of an Au grain with radius $r \approx 4.5$ nm, from [15] (their C_1, C_2, q_0 correspond to our C_L, C_R, Q_0). (a) I–V curve of a typical device at 4.2 K. (b). I–V curve of the device at 30 mK. (c) First current step (α) of (b), at three different refrigerator temperatures. The curves between the points are fits to the Fermi distribution. After correcting for the capacitive division of voltage, the full-width-half-maximum (FWHM) of the peak, shown in inset, is linear with the refrigerator temperature, with a slope of FWHM/$k_B T$ = 3.8, which is close to the expected slope of 3.5. (d) Lineshape of the level when the leads are superconducting ($H = 0$) compared to the lineshape when the leads are normal ($H = 1$ T).

to be a hemisphere with radius r and Vol $= 2\pi r^3/3$ (actually their grains were more pancake-shaped) and determined r by equating the (larger of the) grain-to-lead capacitances to $\pi r^2 C_\square$; here $C_\square$ is the capacitance per unit area of the tunnel junctions. In Ref. [5], RBT estimated that $C_\square \simeq 0.075$ aF/nm^2 for the Al$_2$O$_3$ layers insulating their grains from the leads and gates. More detailed subsequent studies [63] gave a slightly revised value of $C_\square \simeq 0.050$ aF/nm^2. Finally, the grain's single-particle mean level spacing d can be estimated using Eq. (2).

2.4.3. Superconducting leads

In many of RBT's samples, the leads are made from Al. They thus become superconducting below $T_c = 1.21$ K, but can be driven normal by turning on a magnetic field (the critical field is very small, about $H \simeq 20$ mT, since RBT's leads are in effect disordered thin films). Fig. 6 [and also Fig. 5(d)] show how the current and conductance curves change when the leads switch from normal to superconducting upon removal of this magnetic field: each current step is shifted to a higher $|V|$, relative to the normal-lead data, and takes the form of a spike, with, correspondingly, a region of
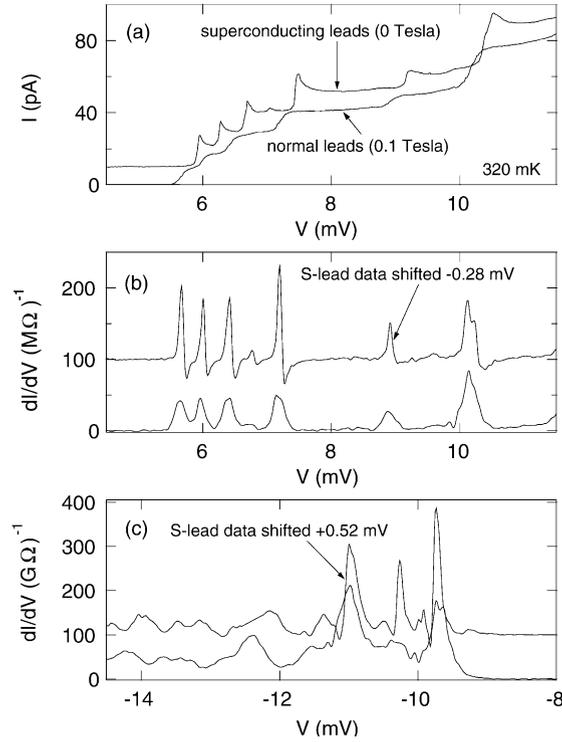
Fig. 6. Current–voltage characteristics for an Al grain connected to superconducting and normal leads [5], without a gate electrode (i.e. $C_g = 0$). The device parameters were $C_L = 4.9 \pm 0.5$ aF, $C_R = 8 \pm 1$ aF, $R_L/R_R = 8 \pm 2$, $R_L + R_R = 9 \pm 1$ MΩ, $Q_0/e = 0.20 \pm 0.03$, and $E_C \sim 6$ meV. (a) I vs. V curves; the N-lead curve shows steps, the S-lead curve (displaced 10 pA in I) shows spikes, reflecting the BCS quasiparticle density of states in the leads [cf. Eq. (33)]. (b) and (c) dI/dV vs. V for positive and negative V, with the S-lead data shifted in V, as labeled, so as to align the maxima of dI/dV with the N-lead data. For ease of comparison, the amplitude of the S-lead data is reduced by a factor of 2 and offset on the dI/dV axis in (b) and (c).

negative dI/dV. This can readily be understood by redoing, for superconducting leads, the golden-rule calculations of Section 2.3.3 (see Appendix A): at $H = 0$ and temperatures $T \ll \Delta_r$ (the BCS gap in lead r), Eqs. (21) and (23) for the rates $\Sigma^{r\pm}$ are modified to

$$\Sigma^{r\pm}_{\alpha\alpha'} = f(\mathscr{E}_\alpha - \mathscr{E}_{\alpha'} \mp e\bar{V}_r)\mathscr{N}_{r,\mathrm{qp}}(\mathscr{E}_\alpha - \mathscr{E}_{\alpha'} \mp e\bar{V}_r)\Gamma^{r\pm}_{\alpha\alpha'} \;, \tag{32}$$

$$\mathscr{N}_{r,\mathrm{qp}}(E) = \begin{cases} |E|/\sqrt{E^2 - \Delta_r^2} & \text{for } |E| > \Delta_r \;, \\ 0 & \text{for } |E| \leq \Delta_r \;. \end{cases} \tag{33}$$

The novel feature in Eq. (32) is the appearance of $\mathscr{N}_{r,\mathrm{qp}}(E)$, the BCS quasiparticle density of states in lead r, which reflects the fact that a quasiparticle is created each time an electron enters or leaves a lead. A detailed analysis [5] showed that the shape of the spikes in the superconducting-lead data in Fig. 6(a) [and also Fig. 5(d)] quantitatively agrees rather well with the prediction of Eq. (32), up to a slight amount of broadening near the threshold (cf. Fig. 3 of [5]). A possible reason

for this broadening was suggested by Levy-Yeyati et al. [83]: as a discrete state is swept past the threshold, its lifetime broadening becomes anomalously large, since this broadening is proportional to the leads' density of states and hence reflects the BCS peak in the latter.

Despite this slight amount of anomalous broadening, the conductance peaks in the dI/dV curve still are much sharper for superconducting than for normal leads, as is evident in Fig. 6(b). This is a very useful feature that can be exploited to enhance spectroscopic resolution.

The shift in V-thresholds for superconducting vs. normal leads can be used to determine which barrier ($r = L$ or R) acts as bottleneck for a given conductance peak: for superconducting leads, the threshold value for $e\bar{V}_r$ at which a rate $\Sigma^{r\pm}$ is switched on, namely $\pm e\bar{V}_r \geq (\mathscr{E}_\alpha - \mathscr{E}_{\alpha'}) + \Delta_r$ [by Eq. (32)], is shifted by $\delta(e\bar{V}_r) = \pm \Delta_r$ relative to the threshold of Eq. (26) for normal leads, reflecting the added energy cost for removing or adding an electron from or to a superconductor. By (25), this corresponds to a shift in bias voltage of

$$\delta V = \frac{\mp C\Delta_r}{|e|(C_R + C_g/2)} \quad \text{if } r = L, \qquad \delta V = \frac{\pm C\Delta_r}{|e|(C_L + C_g/2)} \quad \text{if } r = R . \tag{34}$$

Whereas the sign of δV always equals that of V [as follows by combining Eqs. (34) and (29), with $e < 0$], its magnitude evidently depends on r, allowing the bottleneck barrier r to be identified. Incidentally, once r is known, Eq. (29) can be used to determine whether bottleneck transitions for a given sign of bias voltage involve tunneling onto or off the grain.

Figs. 6(b) and (c) illustrate these ideas: they show two kinds of shifts, whose magnitudes 0.28 and 0.52 mV agree well with the predictions of Eq. (34) for L and R, using the sample parameters listed in the figure caption. In Fig. 6(b), *all* peaks are shifted by the same amount of 0.28 mV, thus L is the bottleneck barrier, with $\Sigma^{L-}_{\alpha_N G_{N+1}}$ as bottleneck rates [by Eq. (29), since $eV < 0$ for Fig. 6(b)]. Thus, this is an example of the "one-bottleneck" situation described in Section 2.3.4(a), from which fixed-N excitation spectra can be extracted. In Fig. 6(c), both types of shifts occur, implying that this is the "two-bottleneck" situation described in Section 2.3.4(b): the peaks between -9 and -12 mV are shifted by -0.52 mV and hence have bottleneck rates $\Sigma^{R-}_{\alpha_N G_{N+1}}$, whereas the peaks between -12 and -15 mV with shifts of -0.28 mV have bottleneck rates $\Sigma^{L+}_{\alpha_{N+1} G_N}$. Finally, the fact that *only two* values of δV shifts were observed confirms that all current steps are due to tunneling through the *same* grain.

2.4.4. Effects of gate voltage on tunneling spectrum

In devices with a gate electrode, V_g can be used to controllably change the average number of electrons on the grain, and to maximize or minimize nonequilibrium effects, as desired. This is illustrated in Fig. 7(a), which shows a series of differential conductance curves for a range of different gate voltages, for the same device as in Fig. 2 (whose caption gives the device parameters). The bottleneck processes in its various regimes are depicted schematically in Fig. 7(b), which one can construct by the following arguments: First note in Fig. 7(a) that as V_g is increased, the extent of the Coulomb blockade region at low $|V|$, in which there are no conductance peaks, decreases, goes to 0, and then increases. At this zero crossing the SET is at a Coulomb blockade degeneracy point, at which, say, the N and $(N + 1)$-electron ground states are degenerate. When passing this point while increasing V_g, which lowers eV_D (since $e < 0$) and hence [by Eq. (6)] favors larger N_{ex}, the grain's average electron number changes from N to $N + 1$. Moreover, nonequilibrium effects
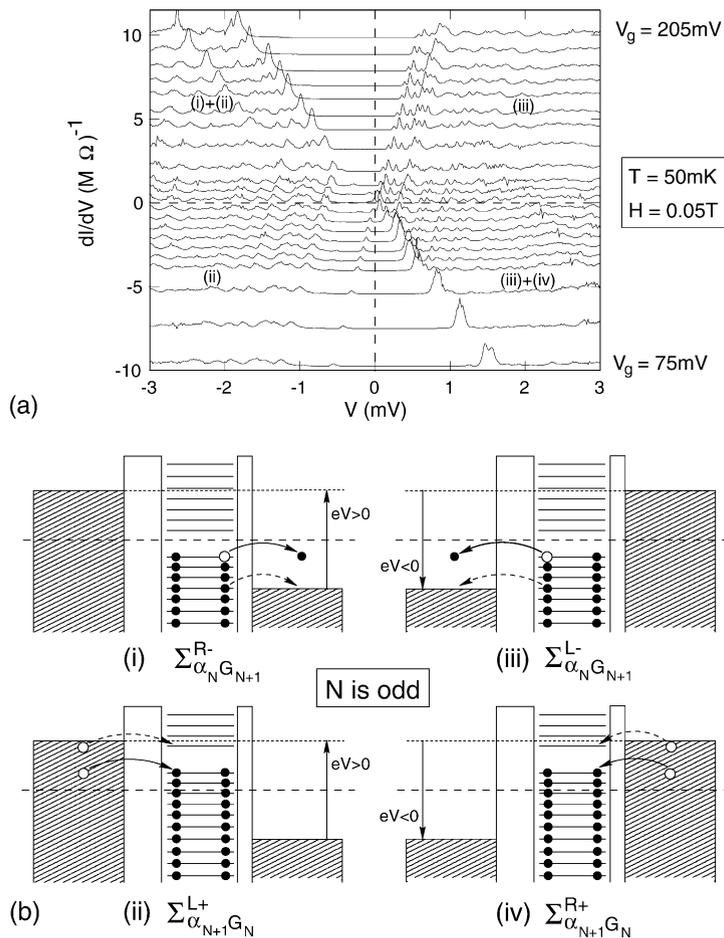
Fig. 7. (a) dI/dV vs. bias voltage V for the same sample [10] as in Fig. 2, plotted with different vertical offsets for a set of V_g-values ranging from 75 mV (bottom) to 205 mV (top). All data are taken at $T = 50\,\text{mK}$ and $H = 0.05\,\text{T}$ (to drive the Al leads normal). (b) Schematic depiction (using the conventions of Fig. 4) of the processes acting as bottlenecks in the four quadrants of (a), for N being odd, with rates: (i) $\Sigma^{R-}_{\alpha_N, G_{N+1}}$ (upper left quadrant), (ii) $\Sigma^{L+}_{\alpha_{N+1}, G_N}$ (upper and lower left quadrants), (iii) $\Sigma^{L-}_{\alpha_N, G_{N+1}}$ (upper and lower right quadrants), (iv) $\Sigma^{R+}_{\alpha_{N+1}, G_N}$ (lower right quadrant). Solid (dashed) arrows depict bottleneck tunneling transitions into the lowest- (highest) energy *final* states accessible for the chosen value of V, and filled circles represent the electron configuration of the lowest-energy *final* state. Increasing V_g decreases eV_D (recall: $e < 0$) and shifts the drawn grain levels down; the degeneracy point between the N- and $(N+1)$-electron ground states, at which the Coulomb blockade regime has shrunk to zero, is reached when the topmost non-empty level of the final ground state coincides with the equilibrium, $V = 0$ chemical potential of the L and R leads, indicated by the horizontal long-dashed line.

(discussed in detail in Section 6) are weak close the zero crossing but become stronger as V_g is tuned away from it, since their strength is governed, roughly speaking, by the size of the threshold voltage $|V|$ at which the first peak occurs. If one wants to make an "ideal" measurement of the "isolated" grain's properties in which the grain is disturbed as little as possible by the measuring process, one has to minimize nonequilibrium effects by tuning V_g close to a zero crossing.

Next, note from Eq. (29) that the left half of Fig. 7(a) (where $eV > 0$) is governed by rates of the type $\Sigma^{R-}_{\alpha_N \alpha'_{N+1}}$ and $\Sigma^{L+}_{\alpha_{N+1} \alpha'_N}$, and the right half by $\Sigma^{L-}_{\alpha_N, \alpha'_{N+1}}$ and $\Sigma^{R+}_{\alpha_{N+1}, \alpha'_N}$, as illustrated in the left and right halves of Fig. 7(b), respectively. To find out which of these rates ($r = $ L or R) are bottlenecks to transport for a given V_g, consider the change δV in the position of a conductance peak upon a small change δV_g in gate voltage; these changes are related by the condition that the corresponding threshold voltage $\bar{V}_r$ in (26) remain unchanged [by Eq. (25)]:

$$\frac{\delta V}{\delta V_g} = \frac{C_g}{C_R + C_g/2} \quad \text{if } r = \text{L}, \quad \frac{\delta V}{\delta V_g} = \frac{-C_g}{C_L + C_g/2} \quad \text{if } r = \text{R} . \tag{35}$$

A peak that has barrier L or R as bottleneck will thus move with slope $\delta V/\delta V_g > 0$ or < 0, respectively. In Fig. 7(a), all peaks indeed do shift uniformly with V_g, which is another proof that all are due to tunneling through the *same* grain. In the upper left and lower right quadrants of Fig. 7(a), slopes of both signs occur (with magnitudes 0.009 and -0.025, consistent with the capacitances listed in the caption of Fig. 2); each of these quadrants thus corresponds to a "two-bottleneck" situation [see Section 2.3.4(b)], in which *both* of the rates shown in the corresponding halves of Fig. 7(b) act as bottlenecks. In contrast, in the upper right and lower left quadrants *all* peaks move with slope $\delta V/\delta V_g > 0$, implying that only the left barrier acts as bottleneck (consistent with the fact that $R_L/R_R \simeq 18 \gg 1$); each of these quadrants thus corresponds to a "one-bottleneck" situation [see Section 2.3.4(a)], from which excitation spectra can be extracted. The upper right quadrant has $\Sigma^{L-}_{\alpha_N, \alpha'_{N+1}}$ as bottleneck rate and hence yields the fixed-N excitation spectrum [cf. Eq. (31)], while the lower left quadrant has $\Sigma^{L+}_{\alpha_{N+1}, \alpha'_N}$ as bottleneck rate and hence yields the fixed-$(N + 1)$ excitation spectrum [cf. Eq. (30)].

Excitation spectra such as the latter two are the central measurement results of SET tunneling spectroscopy. One can proceed to extract from them all sorts of interesting information about the grain and the correlations which its electrons experience. For example, a very remarkable property of the fixed-$(N + 1)$ excitation spectrum in the lower left quadrant of Fig. 7(a) is the sizeable distance between the first and second peaks (counted from $|V| = 0$). As will be discussed in more detail in Section 4.1, this "spectroscopic gap" is evidence for superconducting pairing correlations on the grain: roughly speaking, it reflects the gap of $2\Delta_{\text{grain}}$ (the pairbreaking energy cost) in the excitation spectrum of a superconducting grain with an *even* number of electrons. This also implies that N is odd. The behavior of these excitation spectra in a magnetic field, discussed in Sections 4.1, 4.5 and 4.6, confirms these conclusions.

2.4.5. Capacitance ratios

The accuracy of tunnel-spectroscopic measurements depends on the accuracy with which one knows the capacitance ratios $(C_{R/L} + C_g/2)/C$ [or $C_{R/L}/(C_L + C_R)$ if no gate is present]. These occur, e.g., in the voltages $\bar{V}_{L/R}$ [Eq. (25)] that enter in the threshold condition (26), or as voltage-to-energy conversion factors in Eqs. (30) and (31). These capacitance ratios can of course be determined from the large-scale Coulomb-blockade I–V curves of the grain. However, two additional methods are available, which are more accurate (typically to within 1% [18]), since they exploit the properties of the discrete tunneling spectrum itself:

(i) If the leads are made from a superconducting material, then the ratios $(C_{R/L} + C_g/2)/C$ can be read off directly from the δV shifts [Eq. (34)] in peak positions that occur when superconductivity in the leads is switched off by switching on a magnetic field.

(ii) One can compare the positions of two tunneling peaks occurring for opposite signs of the bias voltage but involving the same transition, e.g. $|G_N\rangle \to |\alpha_{N+1}\rangle$ (the case $|G_{N+1}\rangle \to |\alpha_N\rangle$ is analogous). If $eV > 0$, this transition will [by Eq. (29)] occur by tunneling across barrier L with rate $\Sigma_{\alpha G}^{L,+}$, and if $eV < 0$ across barrier R with rate $\Sigma_{\alpha G}^{R,+}$. Denote the bias voltages at which the two conductance peaks occur by $V^{\alpha,>}$ and $V^{\alpha,<}$, respectively, and use Eq. (25) to define from these two corresponding $\bar{V}_r$-voltages, denoted by $\bar{V}_L^{\alpha,>}$ and $\bar{V}_R^{\alpha,<}$. By Eq. (26), these in fact are equal, $\bar{V}_L^{\alpha,>} = \bar{V}_R^{\alpha,<} (= \mathscr{E}_\alpha^{N+1} - \mathscr{E}_G^N)$. Using Eq. (25) to rewrite this equality in terms of $V^{\alpha,>}$ and $V^{\alpha,<}$, one immediately finds that $|V^{\alpha,>}/V^{\alpha,<}| = (C_L + C_g/2)/(C_R + C_g/2)$. Adding unity to this or to its inverse gives the desired capacitance ratios.

2.4.6. Ground state energy differences are currently not measurable

According to Eq. (26), the threshold values of $e\bar{V}_r^\alpha$ and eV_α at which the *first* conductance peak occurs as $|V|$ is increased from 0 should allow one, in principle, to also determine the ground state energy difference $E_{G_{N\pm1}} - E_{G_N}$ between grains with adjacent electron numbers. In practice, however, it has so far *not* been possible to do this accurately: by Eq. (25), $e\bar{V}_r^\alpha$ depends through $e(V_g C_g + Q_0)/C$ on the gate voltage and the random offset charge Q_0, and in the devices studied so far, eQ_0/C could not be determined with sufficiently high accuracy, i.e. with an uncertainty smaller than the scale of the grain's mean level spacing d, typically $\simeq 0.1$ meV, for the following reason:

Q_0 can of course be determined reasonably accurately by studying the large-scale Coulomb oscillations of the I–V curve that occur as functions of V_g at fixed V, a procedure that is well established for mesoscopic SETs, for which indeed it has been possible to measure the ground state energy difference between a superconducting island with an even or odd number of electrons [62,84–88]. However, a complication arises for the nanoscopic grains of present interest, due to the smallness of their gate capacitances (typically $\simeq 0.1$ aF): to sweep V_g through one Coulomb oscillation, the gate voltage V_g must be swept through a range so large (namely $e/C_g \simeq 1$ V) that during the sweep, RBT routinely observed small "rigid" shifts of the entire tunneling spectrum at random values of V_g. (Similar spontaneous shifts are discussed, e.g., in [88].) Likewise, shifts are also observed when eV is swept over a range of order E_C. These shifts presumably are due to single-electron changes in the charges contained in *other* metal grains[5] in the neighborhood of the grain of interest. These changes produce sudden shifts in the electrostatic potential energy of the grain, i.e. shifts in eQ_0/C by a few %, which are comparable in magnitude to the mean level spacing d and hence spoil any attempts to determine eQ_0/C with an uncertainty smaller than d. The V_g and Q_0 dependencies are subtracted out, however, when one considers the distances $V_{\alpha\tilde{\alpha}}$ *between* peaks, as in Eqs. (30) and (31). Moreover, to measure only the first few discrete states, the range over which V needs to be varied typically is small enough that (with some luck) no shifts occur.

3. Normal grains in an applied magnetic field

In this section we describe how the ability to resolve individual eigenenergies allowed RBT to determine the parity (even or odd) of the number of electrons, say N, in the $V = 0$ ground state of

---

[5] This interpretation is bolstered by the fact that the shifts usually occur at approximately periodic values of V_g.

a normal-state metal grain [5] (Section 3.1). We also explain why orbital magnetism is negligible in ultrasmall grains (Section 3.2).

3.1. Breaking of Kramers-degeneracy by applied magnetic field

In the absence of an applied magnetic field ($H = 0$), a normal-state grain will have time-reversal symmetry. For an even-N grain, the many-electron wave function for the ground state will be a spin singlet, in order that the orbital energy be minimized. In contrast, the ground state of an odd-N grain for $H = 0$ necessarily is two-fold degenerate, by Kramers' theorem, forming a Kramers doublet. When H is turned on, this doublet is Zeeman-split by $\pm \frac{1}{2}\mu_B gH$. Therefore, for an *even-N* grain at small H, the lowest-lying tunneling excitations correspond to transitions from the even-grain ground state singlet to the odd-grain ground state doublet, i.e. to two states split by H, so that the lowest-V conductance peak should exhibit Zeeman splitting in an applied field. On the other hand, for an *odd-N* grain with $T \ll \mu_B gH/k_B$, the odd-grain ground state will be the lower-energy state of the Kramers doublet; the lowest-lying tunneling excitation will thus consist only of a single transition from this odd-grain ground state to the even-grain ground state singlet, so that the lowest-V conductance peak should *not* split into two as a function of H.

RBT observed both kinds of behavior, first in several different ungated Al grains [5,7], in each of which N has a fixed, random value, and subsequently also in a given gated Al grain, in which N could be tuned via V_g [10,9]. Bulk Al superconducts, but the present grains were so small ($r \lesssim 3$ nm) that superconducting pairing correlations on the grain were negligible, since the mean level spacings ($\simeq 0.5$ meV) were substantially larger than the bulk gap $\tilde{\Delta}(\simeq 0.18$ meV for Al). Figs. 8(a) and (b) show the lowest-V conductance peaks for an even-N and an odd-N grain, respectively. The lack of splitting of *both* of the prominent peaks in Fig. 8(b) indicates that the first two even-grain excited states for this grain are both spin singlets. The small peak in Fig. 8(b) (visible below the second large peak), which moves to lower V with increasing H, is attributable to the nonequilibrium occupation of the higher-energy level of the odd-grain initial-state Kramers doublet.

By measuring the difference in V between the Zeeman-split peaks and converting them to energy using Eqs. (30) or (31), RBT determined the g-factors for the states in Fig. 8(a) and (b) to be $g = 1.87 \pm 0.04$ and $g = 1.96 \pm 0.05$, respectively. In the former case the deviation from the value expected for pure Al, namely $g^{\text{pure}} = 2$, is significant, and can be attributed to spin–orbit scattering from the surface or impurities [20,61], a subject that will be discussed in detail in Section 7.

3.2. Why orbital diamagnetism is negligible in ultrasmall grains

Of course, a magnetic field in principle also couples to the orbital motion of the conduction electrons. Orbital effects in spherical and cylindrical superconductors whose dimensions are smaller than the penetration depth were first considered by Larkin [89]. However, in grains as small as those of RBT, orbital diagmagnetic effects are negligible [14], as can be seen from the following argument [90]: Orbital diamagnetism becomes important when the splitting which it produces in the eigenenergies of two time-reversed states (analogous to a Zeeman splitting) becomes comparable to the mean level spacing. This can be shown to occur (see Section 7.4.1) when the flux enclosed by an electron traversing a "closed" trajectory corresponding to a discrete
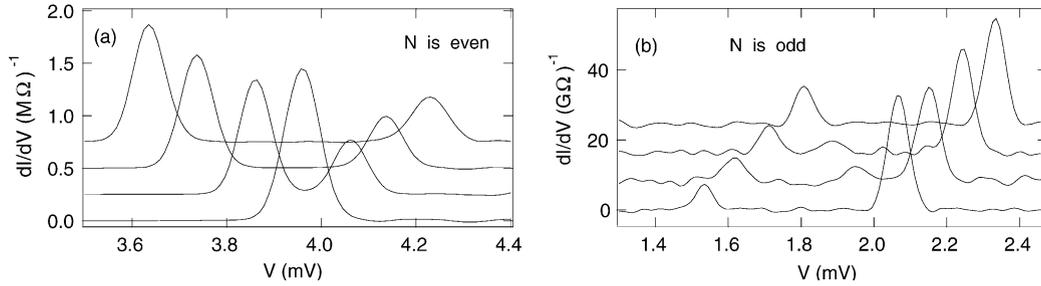
Fig. 8. dI/dV vs. V at $50\,\text{mK}$ and $H = 0.03, 1, 2, 3$ T from bottom to top, for two different samples [5]. The first transition above the Coulomb-blockade threshold exhibits Zeeman splitting in (a) but not in (b), hence the electron number N of the $V = 0$ ground state is even for grain (a) and odd for grain (b).

quantum state exceeds one flux quantum $\Phi_0 (= hc/2e)$, i.e. when the magnetic field exceeds the scale $H_{\text{orb}} = \Phi_0/A_{\text{typ}}$, where A_{typ} is the typical (directed) area covered by the electron's trajectory during the "period" of its motion, $\hbar/d$. Since the number of bounces off the grain's boundaries during this time is roughly $g_{\text{dim}} = E_{\text{Thouless}}/d$, where $\hbar/E_{\text{Thouless}}$ is the time to cross the grain once, the directed area is $A_{\text{typ}} \approx r^2 \sqrt{g_{\text{dim}}}$, where the square root accounts for the fact that the direction of motion after each bounce is random [55]. It follows that

$$H_{\text{orb}} \approx \frac{\Phi_0}{r^2\sqrt{g_{\text{dim}}}} = \frac{2067\,\text{T}\,\text{nm}^2}{r^2\sqrt{g_{\text{dim}}}} \ . \tag{36}$$

Thus, H_{orb} grows with decreasing grain size. Using d from Eq. (2) and the ballistic estimate of Eqs. (4) and (5) for g_{dim}, with $a = 3$, we find that hemispherical Al grains with radii of (say) $r \approx 3$ or $5\,\text{nm}$ have $H_{\text{orb}} \approx 19$ or 7 T, respectively. If larger values are used for a, as would be appropriate for more pancake-shaped grains (cf. footnote 1 and footnote 2), H_{orb} would be even larger.

We may thus conclude that orbital diamagnetic effects only begin to play a role for largish grains ($\gtrsim 5\,\text{nm}$), and then only for the highest fields (of 7 T) studied by RBT.

4. Superconductivity: experiment and phenomenological theory

Among RBT's most striking experimental results are those for Al grains: they found a significant spectroscopic gap indicative of superconducting pairing correlations in the fixed-N excitation spectra of largish Al grains ($r \gtrsim 5\,\text{nm}$) if N is even, but not if it is odd (see Figs. 9 and 10 below); however, in their smallest grains ($r \lesssim 3\,\text{nm}$) no such parity-dependent gap could be discerned.

These results are of rather general interest and significance, since they invite reconsideration of an old but fundamental question: *what is the lower size limit for the existence of superconductivity in small grains?* Anderson [45] addressed this question already in 1959: he argued that if the sample is so small that its electronic eigenspectrum becomes discrete, "superconductivity would no longer be possible" when its mean level spacing d becomes larger than the bulk gap, which we shall denote by $\tilde{\Delta}$. Heuristically, this is obvious (see Fig. 12 below): $\tilde{\Delta}/d$ is the number of free-electron states that
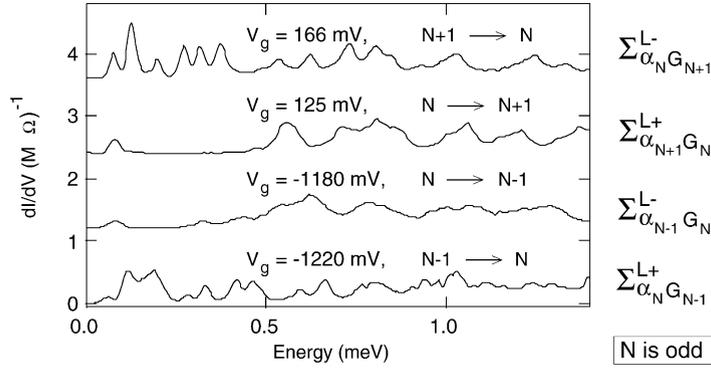
Fig. 9. Excitation spectra for the same sample [10] as Figs. 2 and 7(a), measured at $T = 50\,\text{mK}$ and $H = 0.05$ T (to drive the Al leads normal), for four different V_g-values, corresponding to different values for the grain's average electron number (from top to bottom: $N + 1, N, N, N - 1$). The curves are artificially offset on the vertical axis and each is labeled by the associated bottleneck tunneling rate $\Sigma^{\text{L}\pm}$ depicted in Fig. 7(b), the bottleneck barrier being $r = L$ in this case. Plotted is dI/dV vs. energy, where the latter is given by $|eV|(C_{\text{R}} + C_{\text{g}}/2)/C = 0.73|eV|$ [see Eqs. (30) and (31)]; the voltage-to-energy conversion factor reflects the voltage drop across barrier L. The sizeable spectroscopic gap between the first two peaks in the middle two curves, and its absence in the top and bottom curves, reflects the pairbreaking energy cost in the excitation spectrum of a superconducting grain with an *even* number of electrons, and implies that N is odd.

pair-correlate (those with energies within $\tilde{\Delta}$ of ε_{F}), i.e. the "number of Cooper pairs" in the system; when this becomes $\lesssim 1$, it clearly no longer makes sense to call the system "superconducting".

Although Anderson's answer is correct in general, it generates further questions: What, precisely, does "superconductivity" mean in ultrasmall grains, for which many of the standard criteria such as zero resistivity, Meissner effect and Josephson effect, are not relevant?[6] How is one to modify the grand-canonical BCS theory to obtain a fixed-N theory appropriate for ultrasmall grains, whose charging energy suppresses number fluctuations? What happens in the regime $d \gtrsim \tilde{\Delta}$ in which superconductivity has broken down? Is the breakdown parity dependent? How is it influenced by a magnetic field? Sections 4 and 5 are concerned with providing detailed answers to these and related questions.

Section 4 is devoted to the experiments themselves. We analyze and qualitatively explain them in the framework of a phenomenological theory by Braun et al. [22,23,91], which offers a simple intuitive picture for visualizing the pairing correlations and the changes these incur when the grain size is decreased. Further theoretical developments, inspired by RBT's experiments but not directly concerned with their interpretation, will be discussed in Section 5.

Section 4 is organized as follows: Section 4.1 presents the experimental results of RBT; Section 4.2 proposes a phenomenological model for an isolated ultrasmall grain; Section 4.3 discusses how pairing correlations can be visualized in a fixed-N system and explains when and in what sense it can be called "superconducting"; Section 4.4 presents a generalized variational BCS

---

[6] For an isolated nm-scale grain, (i) its resistivity is not defined, since electron motion is ballistic and the mean-free path is boundary-limited; (ii) the grain radius is smaller than the penetration depth, so that no Meissner effect occurs; and (iii) the electron number is fixed, so that the order parameter cannot have a well-defined phase.
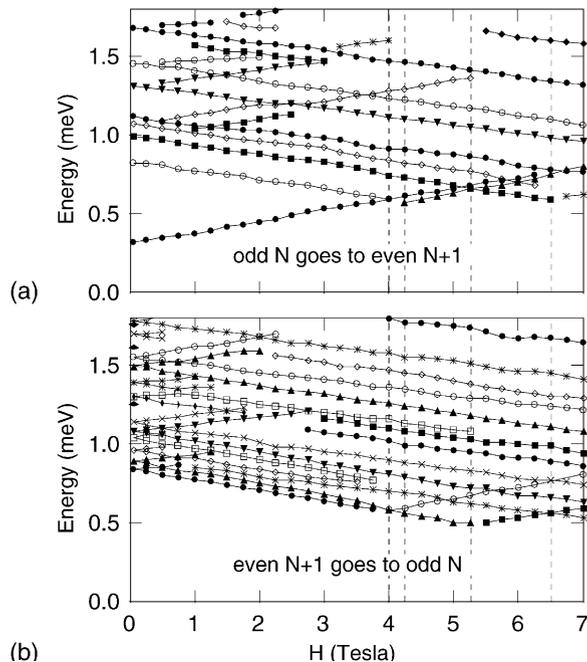
Fig. 10. Magnetic field dependence [10] of excitation spectra such as those of Figs. 7 and 9 and taken for the same grain, at (a) $V_g \approx 110$ mV and (b) $V_g \simeq 180$ mV [lower left and upper right quadrants of Figs. 7(a), respectively]. Each line represents a distinct conductance peak in the dI/dV curve and traces how its energy changes with H. Upward-moving peaks are broader and less distinct than downward-moving ones (for reasons poorly understood) and can be followed only for a limited range of H before they are lost in the background. The distances between lines directly given the grain's (a) fixed-$(N+1)$ and (b) fixed-N excitation spectrum, where N is odd; the corresponding rate-limiting processes are schematically illustrated in Fig. 7(b,ii) for (a) and in Fig. 7(b,iii) for (b). The vertical dashed lines indicate the first four level-crossing fields $H_{s,s'}$ (defined in Eq. (71) and assigned by comparison with Fig. 17, see Section 4.6), namely $H_{0,1} = 4$T, $H_{1/2,3/2} = 4.25$ T, $H_{1,2} = 5.25$ T and $H_{3/2,5/2} = 6.5$ T with uncertainty ± 0.13 T (half the H-resolution of 0.25 T).

approach for calculating the eigenenergies of various variational eigenstates of general spin $|s\rangle$; Section 4.5 discusses how an increasing magnetic field induces a transition from a pair-correlated to a normal paramagnetic state; Section 4.6 presents theoretical tunneling spectra of the RBT type, which are in qualitative agreement with RBT's measurements; Section 4.7 explains how RBT's experiments give direct evidence for the dominance of purely time-reversed states in the pairing interaction; Section 4.8 discusses various parity effects that are expected to occur in ultrasmall grains.

4.1. A gap in the excitation spectrum

RBT's results for largish Al grains ($r \gtrsim 5$ nm) are exemplified by Figs. 9 and 10: if the number of electrons on the grain in the final state after the bottleneck tunneling process is even (middle two curves of Fig. 9), the excitation spectra display a spectroscopic gap between the first two conductance peaks that is significantly larger than the mean spacing between subsequent peaks, whereas

no such gap occurs if the final-state electron number is odd (top and bottom curves of Fig. 9). In other words, even-N excitation spectra (number parity $p = 0$) are gapped, whereas odd-N excitation spectra ($p = 1$) are not. This is even more clearly apparent when comparing Figs. 10(a) and (b), which give the magnetic-field (H) dependence of an even- and odd-N excitation spectrum, respectively. However, in their smallest grains ($r \lesssim 3$ nm) no such clear parity-dependent spectroscopic gap could be discerned.

BRT interpreted these observations as evidence for *superconducting pairing correlations* in their larger grains, using notions from the BCS theory of superconductivity: in an even grain, all *excited* states involve at least one broken Cooper pair, i.e. two BCS quasiparticles, and hence lie at least 2Δ above the fully paired BCS ground state; in contrast, in an odd grain *all* states have at least one unpaired electron, i.e. at least one quasiparticle, and hence no significant gap exists between ground- and excited states. Fig. 7(b) is a cartoon illustration of this interpretation, if one imagines two electrons drawn on the same energy level to represent a "Cooper pair" (making this cartoon precise will be one of the main goals of this chapter): in Figs. 7(b,ii) and (b,iv) the final electron number is even, and all final excited states (reached via dashed arrows) have one less "Cooper pair" than the final ground state (reached via the solid arrow); in contrast, in Figs. 7(b,i) and (b,iii) the final electron number is odd, and the final ground and excited states have the same number of "Cooper pairs".

The approximately linear H-dependence of the excitation spectra in Fig. 10 was attributed by RBT to the Zeeman energy shifts of discrete levels in a magnetic field, as discussed in Section 3.1. Indeed, the method described there for determining the grain's number parity applies here too: the fact that the lowest state in Figs. 10(a) or (b) does not or does display Zeeman splitting, respectively, confirms that N is odd. The reduction of the spectroscopic gap in Fig. 10(a) therefore is purely due to Zeeman energy shifts and has nothing to do with the reduction of the BCS gap parameter due to pair-breaking that occurs in bulk samples in a magnetic field [92]. A detailed discussion of the spectra's magnetic field dependence will be given in Sections 4.5 and 4.6.

For completeness, it should be remarked that a spectral gap in ultrasmall superconducting grains was observed as long ago as 1968 by Giaever and Zeller [80,81], who studied tunneling through granular thin films containing electrically insulated Sn grains. They found gaps for grain sizes right down to the critical size estimated by Anderson (radii of 2.5 nm in this case), but were unable to prove that smaller particles are always "normal". RBT's experiments are similar in spirit to this pioneering work, but their ability to focus on *individual* grains makes a much more detailed study possible.

4.2. A model for ultrasmall grains with pairing correlations

In this section we construct a model for an isolated ultrasmall grain with pairing correlations, using phenomenological arguments valid for the regime $d \lesssim \tilde{\Delta}$. The model, which we shall call "discrete BCS model", allows a rather detailed qualitative understanding of the measurements of RBT [22,23] and hence is "phenomenologically successful". For $d \gg \tilde{\Delta}$ it is unrealistically simple, however, and should rather be viewed as a toy model for studying how pairing correlations change as the grain is made smaller and smaller.

4.2.1. A simple reduced BCS interaction plus a Zeeman term
Following the philosophy of the "orthodox theory" for Coulomb blockade phenomena mentioned in Section 2.3.1, we assume that the only effect of the Coulomb interaction is to contribute

an amount $E_C N_{ex}^2$ [cf. Eq. (6)] to the eigenenergy of each eigenstate of the grain. Since the charging energy is huge (5–50 meV) in ultrasmall grains, this term strongly suppresses electron number fluctuations, so that, to an excellent approximation, all energy eigenstates will also be number eigenstates.[7] Since $E_C N_{ex}^2$ is constant within every fixed-N Hilbert space, we shall henceforth ignore it, with the understanding that the model we are about to construct should always be solve in a fixed-N Hilbert space (and that the use of grand-canonical approaches below, after having dropped $E_C N_{ex}^2$, simply serves as a first approximation to the desired canonical solution).

The only symmetry expected to hold in realistic, irregularly shaped ultrasmall grains at zero magnetic field is time-reversal symmetry. We therefore adopt a single-particle basis of pairs of time-reversed states $|j \pm \rangle$, enumerated by a discrete index j. Their discrete energies ε_j are assumed to already incorporate the effects of impurity scattering and the average of electron–electron interactions, etc. As simplest conceivable model that incorporates pairing interactions and a Zeeman coupling to a magnetic field, we adopt a Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_{red}$ of the following reduced BCS form:

$$\hat{H}_0 = \sum_{j,\sigma = \pm} (\varepsilon_j - \mu - \sigma h)c_{j\sigma}^\dagger c_{j\sigma}, \qquad \hat{H}_{red} = -\lambda d \sum_{ij} c_{i+}^\dagger c_{i-}^\dagger c_{j-} c_{j+} \ . \tag{37}$$

Here $-\sigma h \equiv \sigma \frac{1}{2} \mu_B g H$ is the Zeeman energy of a spin σ electron in a magnetic field H, and we shall take $h > 0$ below. Models of this kind had previously been studied by Richardson [46–54], Strongin et al. [93], Mühlschlegel et al. [94,95] and Kawataba [96,97]. The first application to RBT's grains for $h = 0$ was by von Delft et al. [21] and for $h \neq 0$ by Braun et al. [21–23].

Due to level repulsion the ε_j's will, to first approximation, be uniformly spaced. Unless otherwise specified, we shall for simplicity always (except in Section 5.4) take a completely uniform spectrum with level spacing d. Fluctuations in the level spacings have been studied with methods of random matrix theory [24,39], with qualitatively similar results (see Section 5.4). For a system with a total of N electrons, where the *electron number parity* $p \equiv N \bmod 2$ is equal to 0 for even N and 1 for odd N, we use the label $j = 0$ for the lowest-lying non-doubly-occupied level (with occupation number p) in the $T = 0$ Fermi sea, which we shall denote by $|F_N\rangle$. We choose the Fermi energy at $\varepsilon_F \equiv 0$ to write[8]

$$\varepsilon_j = jd + (1 - p)d/2 \ , \tag{38}$$

thereby taking the doubly-occupied and empty levels of $|F_N\rangle$ to lie symmetrically above and below ε_F (see e.g. Fig. 13 below). The parameter μ in Eq. (37) is, in g.c. theories, the chemical potential, whose value[10] determines the average particle number. For canonical theories, which make no reference to a chemical potential, μ is not needed and can be dropped (i.e. set equal to 0).

The pairing interaction is of the reduced BCS form, in that it scatters a pair of electrons from one pair of time-reversed states into another. It is taken to include only states whose energy separation from the Fermi energy lies within the cutoff given by the Debye frequency: $|\varepsilon_j| < \omega_D$. Experimental

---

[7] An exception occurs *at* a degeneracy point, where $E_{pot}(N_{ex}) = E_{pot}(N_{ex} + 1)$ [cf. Eq. (6)]; the resulting complications will not be considered here.

[8] This convention differs slightly from that used in [21–23], namely $\varepsilon_j = jd + \varepsilon_0$. The latter is a little less convenient, resulting, e.g., in a p-dependent chemical potential for the variational BCS ground states discussed below, $\mu_p^{BCS} = \varepsilon_0 + (p - 1)d/2$, whereas (38) results simply in $\mu_p^{BCS} = 0$.

support for the sufficiency of neglecting couplings between nontime-reversed pairs of states,[9] i.e. of using only a *reduced* BCS-Hamiltonian, is given in Section 4.7. The pair-coupling constant in Eq. (37) is written as λd, where λ is a dimensionless parameter independent of the grain's volume, to make it explicit that both $\hat{H}_0$ and $\hat{H}_{\text{red}}$ make extensive ($\propto$ Vol) contributions to the ground state energy (since the number of terms in each sum $\sum_j$ in Eq. (37) scales with N, and $d \propto \varepsilon_{\text{F}}/N$). The "bulk gap" of the model, obtained by solving the standard BCS gap equation [Eq. (57)] at $T = 0$ in the bulk limit, thus is

$$\tilde{\Delta} = \omega_{\text{D}}/\sinh(1/\lambda) . \tag{39}$$

To be precise, by "bulk limit" we shall always mean $d/\tilde{\Delta} \to 0$ and $N \to \infty$ while the product Nd is kept fixed, and use $d\sum_j \to \int d\varepsilon_j$.

An applied magnetic field will completely penetrate an ultrasmall grain, since its radius (typically $r \lesssim 5$ nm) is much smaller than the penetration length of 50 nm for bulk Al. The Zeeman term in Eq. (37) models the fact that RBT's measured tunnel spectra of Fig. 10 evolve approximately linearly as a function of magnetic field, with g factors between[10] 1.95 and 2 (determined from the differences between measured slopes of up- and downward-moving lines). Deviations from $g = 2$ probably result from spin–orbit scattering, known to be small but nonzero in thin Al films [98,99], but neglected below (where $g = 2$ is used). Furthermore, orbital diamagnetism is also negligible, just as for thin films in a parallel magnetic field [98,99] but in marked contrast to bulk samples where it causes the Meissner effect: as explained in Section 3.2, a magnetic field only begins to significantly affect the orbital motion of the electrons once it exceeds the scale $H_{\text{orb}} \simeq \Phi_0/r^2\sqrt{g_{\text{dim}}}$ [90,55]; but according to the estimates mentioned after Eq. (36), RBT's ballistic grains with $r \lesssim 5$ nm have $H_{\text{orb}} \gtrsim 7$ T, hence orbital effects should set in only near the highest fields of 7 T studied by RBT. Indeed, some larger grains do show slight deviations from H-linearity [6] for large fields, which probably reflect the onset of such orbital effects [100]; however these are much smaller than Zeeman effects in the grains of present interest, and will be neglected here. Thus, the model assumes that Pauli paramagnetism due to the Zeeman energy completely dominates orbital diamagnetism, similarly to the case of thin films in parallel magnetic fields [98,99].

Intuitively speaking, it is clear that the discrete BCS model introduced above contains all ingredients necessary to make contact with the spectra of Fig. 10: it is formulated in terms of discrete levels, it contains a pairing interaction which is known, from bulk BCS theory [101,92], to cause a gap in the excitation spectrum, and it contains a Zeeman term that will cause eigenenergies to linearly depend on an applied magnetic field. Indeed, we shall see in Section 4.6 that it can be used to obtain a rather detailed qualitative understanding of the spectra of Fig. 10.

4.2.2. Choice of numerical values for model parameters

When doing numerical calculations for this model, some choices must be made for the numerical values of its parameters (though slight changes in their values will not change the results

---

[9] A theoretical motivation for the simple form assumed for the pairing interaction, based on random matrix theory, may be found in Ref. [13] and is briefly mentioned at the beginning of Section 6.1.3.

[10] Claims of smaller g factors made in Ref. [6] are wrong, the result of confusing different orbital states as Zeeman-split spin states. This was made clear in Ref. [10], where it was observed that upward-trending Zeeman states can have significantly smaller amplitude than downward-trending states, making them difficult to observe.

qualitatively). We shall follow the choices made by Braun et al. [23], since these led to reasonable agreement between experimental and theoretical excitation spectra. For the Debye frequency they used the textbook value [102] for Al of $\omega_D = 34\,\mathrm{meV}$. Making an appropriate choice for the "bulk gap" $\tilde{\Delta}$ is less straightforward, since its experimental value for systems of reduced dimensionality often differs from that of a truly bulk system, presumably due to (poorly understood) changes in the phonon spectrum and the effective electron–phonon coupling. For example, for thin Al films [93,103] it is known that $\tilde{\Delta}_{\mathrm{thin\ film}} \simeq 0.38\,\mathrm{meV}$, which is about twice as large as the gap of a truly bulk system, $\tilde{\Delta}_{\mathrm{bulk}} = 0.18\,\mathrm{meV}$. (This increase in $\tilde{\Delta}$ is not universal, though; e.g., for Nb $\tilde{\Delta}$ is smaller in thin films than in the bulk.) Since ultrasmall grains are in many ways analogous to thin films in a parallel magnetic field [see Section 4.5], Braun et al. adopted the thin-film value for grains too, i.e. used $\tilde{\Delta} \simeq 0.38\,\mathrm{meV}$. These choices imply that the dimensionless pair-coupling constant $\lambda = [\sinh^{-1}(\omega_D/\tilde{\Delta})]^{-1}$ [cf. Eq. (39)] has the value $\lambda = 0.194$. (In Section 4.6 we shall see, *a posteriori*, that the choices $\tilde{\Delta} = 0.34$ and $\lambda = 0.189$ would have been slightly more appropriate.) Finally, for those numerical calculations that are explicitly cut-off dependent, Braun et al. smeared the cutoff of the BCS interaction over two single-electron levels; this smooths out small discontinuities that would otherwise occur in d-dependent quantities each time the energy $|\varepsilon_j|$ of some large-$|j|$ level moves beyond the cutoff ω_D when d is increased.

Note that the above way of choosing λ lumps into a single phenomenological constant all the poorly understood effects of reduced dimensionality [93] on the phonons that mediate the attractive electron–electron interaction. Studying these effects in detail would be interesting in its own right, but would require systematic investigations with grains of well-controlled shapes and sizes. For the case of RBT's irregularly shaped grains, using a phenomenological coupling constant seems the best one can do. Note, though, that the precise value of λ is not very important as long as all energies are measured in units of $\tilde{\Delta}$ (as we shall do for all numerical calculations), since most of the λ-dependence is thereby normalized away. Therefore, the slight difference between the λ-values proposed above and those used in [35,91,39] (namely 0.224) hardly matters.

4.2.3. Some general properties of the eigenstates – the blocking effect

Th eigenstates of the discrete BCS model of Eq. (37) have some simple but general properties that are worth stating at the outset.

Firstly, every eigenstate of $\hat{H}$ will also be an eigenstate of the number operator $\hat{N} = \sum_{j\sigma} c^{\dagger}_{j\sigma} c_{j\sigma}$, since $[\hat{H}, \hat{N}] = 0$.

Secondly, since the interaction only involves levels within the cutoff energy ω_D of ε_F, the dynamics of those lying outside this range is trivial. We shall thus ignore them henceforth and focus only on the remaining set of *interacting* levels, denoting this set by I.

Thirdly, *singly occupied* levels do not participate in the pair scattering described by $\hat{H}$: "unpaired" electrons in such levels are not scattered to other levels, hence the labels of singly occupied levels are good quantum numbers. Moreover, every unpaired electron Pauli-blocks the scattering of other pairs into its own singly occupied level, i.e. it restricts the phase space available to pair scattering and thereby weakens the amount of pairing correlations, as we shall see in detail later. This was called the "*blocking effect*" by Soloviev [104], who discussed it extensively in the early 1960s in the context of nuclear physics. The eigenstates $|\alpha\rangle$ and corresponding eigenenergies $\mathscr{E}_{\alpha}$ of

$\hat{H}$ thus have the following general forms:

$$|\alpha\rangle = |\Psi_n, B\rangle = \prod_{i \in B} c_{i\sigma_i}^{\dagger} |\Psi_n\rangle \,, \tag{40}$$

$$|\Psi_n\rangle = \sum_{j_1, \ldots, j_n}^{U} \psi(j_1, \ldots, j_n) \prod_{\nu=1}^{n} b_{j\nu}^{\dagger} |\text{Vac}\rangle \,, \tag{41}$$

$$\mathcal{E}_{\alpha} = \mathcal{E}_n + \mathcal{E}_B(h), \quad \mathcal{E}_B(h) = \sum_{i \in B} (\varepsilon_i - \mu - \sigma_i h) \,. \tag{42}$$

This describes $N = 2n + b$ electrons, b of which are unpaired and sit in a set B of singly occupied, blocked levels, making a contribution $\mathcal{E}_B(h)$ to the total eigenenergy. The remaining n pairs of electrons, created by the pair operators $b_j^{\dagger} = c_{j+}^{\dagger} c_{j-}^{\dagger}$, are distributed among the remaining set $U = I \backslash B$ of *unblocked* levels, with wavefunction $\psi(j_1, \ldots, j_n)$ ($\sum_j^U \equiv \sum_{j \in I \backslash B}$ denotes a sum over all *unblocked* levels in I). The corresponding state $|\Psi_n\rangle$ is an eigenstate of the pair number operator and a Hamiltonian $\hat{H}_U$ involving only pair operators:

$$\sum_{j}^{U} b_j^{\dagger} b_j |\Psi_n\rangle = n|\Psi_n\rangle, \quad \hat{H}_U |\Psi_n\rangle = \mathcal{E}_n |\Psi_n\rangle \,, \tag{43}$$

$$\hat{H}_U = \sum_{ij}^{U} [2(\varepsilon_j - \mu)\delta_{ij} - \lambda d] b_i^{\dagger} b_j \,. \tag{44}$$

Each eigenstate $|\Psi_n, B\rangle$ may be visualized as a coherent superposition of eigenstates of $\hat{H}_0$ that all lie in the same fixed-N Hilbert space, and in all of which each pair of unblocked ($j \in U$), time-reversed levels $|j \pm \rangle$ is either doubly occupied or empty. This is illustrated in Figs. 11(a) and (b), which schematically depict the exact ground states for even- and odd-N, respectively. The odd ground state has a single blocked level, at the Fermi energy, containing an unpaired electron. The latter somewhat weakens pairing correlations relative to the even ground state and hence leads to parity effects, which will be extensively discussed in later sections.

A useful measure for the amount of energy gained by $|\alpha\rangle$ via its correlations is its "condensation energy" relative to the uncorrelated state $|\alpha\rangle_0$,

$$E_{\alpha}^{\text{cond}} = \mathcal{E}_{\alpha} - {}_0\langle\alpha|\hat{H}|\alpha\rangle_0, \quad \text{where} \quad |\alpha\rangle_0 = \prod_{i \in B} c_{i\sigma_i}^{\dagger} |U\rangle_0 \tag{45}$$

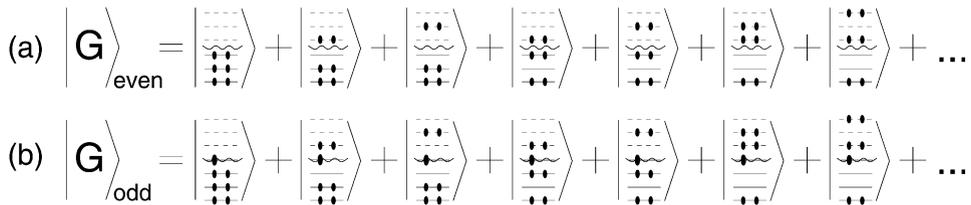

Fig. 11. A cartoon depiction of the exact ground state for a reduced BCS Hamiltonian, for N being even (a) or odd (b): they are coherent superpositions of eigenstates of $\hat{H}_0$ (whose respective amplitudes are not depicted) that all have the same N; the leftmost of these is (a) the even or (b) the odd Fermi ground state $|F_N\rangle$, whose Fermi energy is indicated by the wavy line.

and $|U\rangle_0$ is the "Fermi ground state" in U, for which the n pairs occupy the n lowest-lying levels in U.

Note that $\hat{H}_U$ is h-independent, since the total Zeeman energy of any *pair* of electrons is zero. Hence the full h-dependence of the eigenenergies resides in the rather trivial contribution $\mathcal{E}_B(h)$ of the blocked levels, which is a very important and useful simplification.

Diagonalizing $\hat{H}_U$ would be trivial if the b's were true bosons. However, they are not, and in the subspace spanned by the set U of all non-singly-occupied levels instead satisfy the "hard-core boson" relations,

$$b_j^{\dagger 2} = 0, \quad [b_j, b_{j'}^\dagger] = \delta_{jj'}(1 - 2b_j^\dagger b_j), \quad [b_j^\dagger b_j, b_{j'}^\dagger] = \delta_{jj'} b_j^\dagger \, , \tag{46}$$

which reflect the Pauli principle for the fermions from which the b's are constructed. In particular, $b_j^{\dagger 2} = 0$ implies that only those terms in (41) are nonzero for which the indices $j_1, \ldots, j_n$ are all distinct.

The task of finding the eigenstates $|\Psi_n\rangle$ is thus highly nontrivial. Nevertheless, an exact solution does exist. Unbeknownst to most of the condensed-matter physics community, it was found and studied extensively by Richardson in the 1960s and will be presented in Section 5.1. Throughout the present Section 4, however, we shall use more well-known approaches based on the variational wavefunctions introduced by BCS [101], and that had been used to study the discrete BCS model before Richardson's solution was revived towards the end of 1998.

4.3. Canonical characterization of pairing correlations

Since the discrete BCS model has the standard reduced BCS form, the most natural first step toward understanding its $T = 0$ properties is to use BCS-like variational wavefunctions (or equivalently Bogoljubov's mean-field approach), and indeed this will be done in Sections 4.4–4.8. However, the limitations of such an approach should be realized at the outset: the spectra measured by RBT are excitation spectra for a grain with a *fixed electron number N*, and hence should be calculated for a grain with definite electron number N (i.e. completely isolated from the rest of the world, e.g. by infinitely thick oxide barriers). In contrast, the variational wavefunctions of BCS [Eq. (47) below] do not have the fixed-N form [Eq. (41)] which any true eigenstate should have, but instead are formulated in a *grand-canonical* (g.c.) framework (as is the Bogoljubov mean-field approach to which they are intimately related).

When considering a truly isolated superconductor such as a perfectly insulated grain (another example would be a superconductor levitating in a magnetic field due to the Meissner effect), one therefore needs to address the following question, which will be the main theme of the present section: *how is one to incorporate the fixed-N condition into BCS theory, and how important is it to do so?* This issue is well understood and was discussed at length in the early days of BCS theory (Rickayzen's book gives a beautiful discussion [105]), in particular in its application to pairing correlations in nuclei [106, p. 439] (see also the general remarks in [107]). Nevertheless, for pedagogical reasons the arguments are well worth recapitulating in the present context.

We shall first remind the reader that the use of a g.c. framework is only a matter of convenience, since the essence of the pairing correlations that lie at the heart of BCS theory is by no means inherently g.c. and can easily be formulated in canonical language [21]. We then show how standard BCS theory fits into this scheme, point out that the differences between results obtained

using g.c. and canonical wavefunctions are negligible for $d \ll \tilde{\Delta}$, and conclude that for the purposes of gaining a phenomenological understanding of the experimental data, standard grand-canonical BCS theory should be sufficient. Nevertheless, the fundamental question of how to improve on this theory, in order to achieve a truly canonical description and to properly treat fluctuation effects, which become important for $d \gtrsim \tilde{\Delta}$ [108,109,28], is interesting and important in its own right and will be addressed at length in Section 5.

For simplicity, throughout the present Section 4.3 we shall consider only the even ground state in the thermodynamic limit (in which even–odd differences are negligible), so that $U = I$ and blocking effects need not be worried about.

4.3.1. The grand-canonical BCS wavefunction

Conventional BCS theory describes the pairing correlations induced by an attractive pairing interaction such as $\hat{H}_{\text{red}}$ of Eq. (37) within a g.c. ensemble, formulated on a Fock space of states in which the total particle number N is not fixed. This is illustrated by BCS's famous variational ground state ansatz

$$|\text{BCS}\rangle = \prod_j (u_j + e^{i\phi_j} v_j b_j^\dagger)|\text{Vac}\rangle \quad \text{with } u_j^2 + v_j^2 = 1 , \qquad (47)$$

where the variational parameters u_j and v_j are real and ϕ_j is a phase (which, it turns out, must be j-independent, for reasons discussed below). $|\text{BCS}\rangle$ is not an eigenstate of $\hat{N}$ and its particle number is fixed only on the average by the condition $\langle \hat{N} \rangle_{\text{BCS}} = N$, which determines the g.c. chemical potential μ. Likewise, the commonly used g.c. definition

$$\Delta_{\text{gc}} \equiv \lambda d \sum_j \langle b_j \rangle_{\text{BCS}} = \lambda d \sum_j u_j v_j e^{i\phi_j} \qquad (48)$$

for the superconducting pairing parameter only makes sense in a g.c. ensemble, since $\langle b_j \rangle$ would trivially give zero when evaluated in a canonical ensemble, formulated on a strictly fixed-N Hilbert space of states. (We shall use the term "pairing parameter" instead of "order parameter", since the latter carries the connotation of a phase transition, which would require the thermodynamic limit $N \to \infty$, which is not applicable for ultrasmall grains.)

4.3.2. Canonically meaningful definition for the pairing parameter

A theory of strictly fixed-N superconductivity must evidently entail modifications of conventional BCS theory. However, these are only of technical, not of conceptual nature, since the essence of the pairing correlations discovered by BCS can easily be formulated in a canonically meaningful way, including a definition for the pairing parameter. We shall now attempt to explain, in intuitive, nontechnical terms, how this may be done (our discussion is indebted to that of Rickayzen [105]). Readers with a preference for rigor may consult Section 5 for a corroboration, using Richardson's exact solution, of the arguments presented below.

Let $|G\rangle$ be the exact even ground state of the system, depicted in Fig. 11(a). As explained in Section 4.2.3, it is a coherent superposition of eigenstates of $\hat{H}_0$ that all have the same N and in all of which each pair of time-reversed levels $|j \pm \rangle$ is either doubly occupied or empty. Due to this coherent superposition, $|G\rangle$ entails strong pairing correlations, whose essential properties may be

understood by investigating how they modify the correlators

$$C_{ij} \equiv \langle b_i^\dagger b_j \rangle, \quad \bar{v}_j^2 \equiv C_{jj} = \langle b_j^\dagger b_j \rangle, \quad \bar{u}_j^2 \equiv \langle b_j b_j^\dagger \rangle, \tag{49}$$

relative to the form these take on for the Fermi ground state $|F_N\rangle$:

$$(C_{ij})_\mathrm{F} = \delta_{ij}(\bar{v}_j^2)_\mathrm{F}, \quad (\bar{v}_j^2)_\mathrm{F} = \theta(-\varepsilon_j), \quad (\bar{u}_j^2)_\mathrm{F} = \theta(\varepsilon_j). \tag{50}$$

$C_{ij}(= C_{ji}^*)$ is the matrix element for the interaction to be able to scatter a pair of electrons from level j to i, and $\bar{v}_j^2$ and $\bar{u}_j^2$ are the probabilities to find level j doubly occupied or empty, respectively. The pairing correlations in $|G\rangle$ must be such that $\hat{H}_\mathrm{red}$ lowers the ground state energy below that of the uncorrelated Fermi sea $|F_N\rangle$ by an amount that is extensive ($\propto N \propto d^{-1}$) in the thermodynamic limit. Clearly, this requires that $\langle \hat{H}_\mathrm{red}\rangle_G - \langle \hat{H}_\mathrm{red}\rangle_\mathrm{F}$ is negative and extensive, i.e. that

$$\lambda d \sum_{ij} [C_{ij} - (C_{ij})_\mathrm{F}] \simeq \lambda d \sum_i \sum_{j<i} 2\,\mathrm{Re}(C_{ij}) \propto N \quad (\text{and positive}). \tag{51}$$

In the second expression we neglected the diagonal terms, since their number is so small (only $\propto N$) that $\lambda d \sum_j [\bar{v}_j^2 - (\bar{v}_j^2)_\mathrm{F}]$ is at best of order unity in the thermodynamic limit. For Eq. (51) to hold, $|G\rangle$ must have two properties:

(i) the number of C_{ij}'s that differ significantly from zero (i.e. are of order unity) should scale like N^2, i.e. one power of N per index [105, p. 167];

(ii) most or all of the C_{ij} for $i < j$ should have the same phase, since a sum over random phases would average out to zero.

Since a suitable pairing parameter should vanish in the thermodynamic limit unless both these conditions hold, the definition

$$\Delta_\mathrm{can}^2 \equiv (\lambda d)^2 \sum_{ij} (C_{ij} - \langle c_{i+}^\dagger c_{j+} \rangle \langle c_{i-}^\dagger c_{j-} \rangle) \tag{52}$$

(or its square root) suggests itself, where the subscript emphasizes that (52) is meaningful in a canonical ensemble too, and we subtracted[11] the "normal-state contribution to C_{ij}". If (i) and (ii) hold, Δ_can will take on a finite value; its relation to a gap in the spectrum will become clear below. In the bulk limit, Δ_can can be shown (see Section 5.3) to reduce to the "bulk pairing parameter" $\tilde{\Delta}$ of Eq. (39).

---

[11] This subtraction was suggested to us by Moshe Schechter, who pointed out that then Eq. (52) has a natural generalization to position space: it is the spatial average, $\Delta_\mathrm{can}^2 \equiv (\lambda d)^2 \int d\vec{r}_1\, d\vec{r}_2\; \mathscr{F}(\vec{r}_1, \vec{r}_2)$, of the two-point function

$$\mathscr{F}(\vec{r}_1, \vec{r}_2) \equiv \langle \psi_+^\dagger(\vec{r}_1)\psi_-^\dagger(\vec{r}_1)\psi_-(\vec{r}_2)\psi_+(\vec{r}_2)\rangle - \langle \psi_+^\dagger(\vec{r}_1)\psi_+(\vec{r}_2)\rangle\langle \psi_-^\dagger(\vec{r}_1)\psi_-(\vec{r}_2)\rangle$$

(with $\psi_\sigma(\vec{r}) \equiv \mathrm{Vol}^{-1/2}\sum_{\vec{k}}\mathrm{e}^{i\vec{k}\cdot\vec{r}}c_{\vec{k}\sigma}$), which evidently measures the amplitude for the propagation of *pairs* as opposed to uncorrelated electrons. Other definitions for a canonically meaningful pairing parameter have been suggested [21,35,23], such as $\lambda d\sum_j \bar{u}_j\bar{v}_j$ or $\lambda d\sum_j [\langle b_j^\dagger b_j\rangle - \langle c_{j+}^\dagger c_{j+}\rangle\langle c_{j-}^\dagger c_{j-}\rangle]^{1/2}$, but these focus only on requirement (i) and fail to incorporate requirement (ii). A quantity very similar to Eq. (52) was recently proposed in Eq. (55) of Ref. [30], namely $(\lambda d)^2\sum_{ij}[C_{ij} - (C_{ij})_{\lambda=0}]$.

4.3.3. Redistribution of occupation probability across ε_F

Now, property (i) can be realized if all C_{ij} in a finite (d-independent) range of ε_i's and ε_j's around the Fermi energy ε_F differ significantly from zero; the width of this range will evidently determine the magnitude of Δ_{can} (provided (ii) also holds), which conversely can be viewed as a measure of this width. But a nonzero C_{ij} evidently requires *both* $b_i^\dagger b_j |G\rangle \neq 0$, implying $(\bar{v}_j)_G \neq 0$ and $(\bar{u}_i)_G \neq 0$, *and also* $\langle G|b_i^\dagger b_j \neq 0$, implying $(\bar{v}_i)_G \neq 0$ and $(\bar{u}_j)_G \neq 0$. The product $(\bar{u}_j \bar{v}_j)_G$ must thus be different from zero [in contrast to $(\bar{u}_j \bar{v}_j)_F = 0$] for all ε_j within a finite range around ε_F (cf. Fig. 12). This can be achieved by smearing out the sharp steps of the θ-functions of $(\bar{v}_j)_F$ and $(\bar{u}_j)_F$, so that $(\bar{v}_j)_G$ [or $(\bar{u}_j)_G$] is nonzero also for a finite range of ε_j above [or below] ε_F. In other words, for $|G\rangle$ some occupation probability must be redistributed (relative to $|F_N\rangle$) from below to above ε_F, as illustrated in Fig. 11. This redistribution, which was called pair mixing in [21,23], frees up phase space for pair scattering and so achieves a gain in interaction energy (provided (ii) also holds) that more than compensates for the kinetic energy cost incurred thereby.

Furthermore, note that properties (i) and (ii) imply, even without detailed calculations, that the spectrum will be gapped. Consider, for example, a "blocking excitation" that disrupts pairing correlations by having $|j+\rangle$ definitely occupied and $|j-\rangle$ definitely empty; since pair scattering involving level j is blocked, the energy cost is

$$(\varepsilon_j - \mu) - [(\varepsilon_j - \mu)2\langle b_j^\dagger b_j \rangle) - \lambda d \sum_{i(\neq j)} \langle b_i^\dagger b_j + b_j^\dagger b_i \rangle] \tag{53}$$

$$= (\varepsilon_j - \mu)(1 - 2\bar{v}_j^2) + \lambda d \sum_{i(\neq j)} (C_{ij} + C_{ji}) , \tag{54}$$
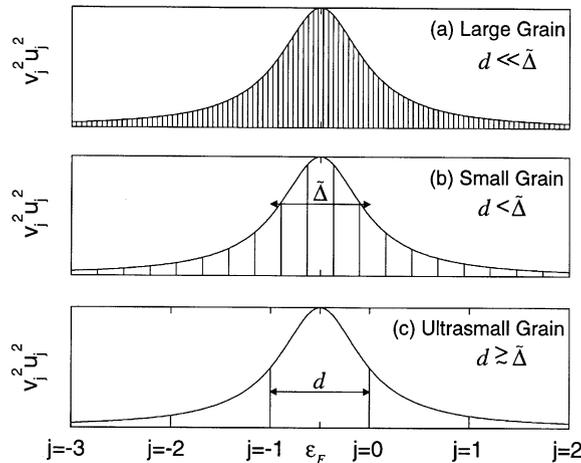
Fig. 12. A cartoon depiction of why "superconductivity breaks down" when the sample becomes sufficiently small. Vertical lines are drawn at each single-particle energy ε_j, spaced with a mean level spacing d corresponding to (a) a "large" grain ($d \ll \tilde{\Delta}$); (b) a "small" grain ($d \simeq 0.25\tilde{\Delta}$); (c) an "ultrasmall" grain ($d \simeq \tilde{\Delta}$). The lines' height represents the function $u_j^2 v_j^2 = \frac{1}{4}\tilde{\Delta}^2/(\varepsilon_j^2 + \tilde{\Delta}^2)$ of standard bulk BCS theory, to illustrate the energy range (of width $\tilde{\Delta}$ around ε_F) within which pairing correlations are strongest. Loosely speaking, the number of single-particle levels in this regime, namely $\tilde{\Delta}/d$, corresponds to "the number of Cooper pairs" of the system. Evidently, this number becomes less than one when $d \gtrsim \tilde{\Delta}$ as in (c), so that it then no longer makes sense to call the system "superconducting" (cf. Section 4.4.2).

in which the restriction on the sum reflects the blocking of scattering events involving level j. Since the first term of (54) is positive definite (particle–hole symmetry ensures that $(\frac{1}{2} - \bar{v}_j^2) \gtrless 0$ if $\varepsilon_j - \mu \gtrless 0$) and the second of order Δ_{can}, the excitation energy will be *finite* even for $d \to 0$, implying the existence of a gap order Δ_{can}. Similarly, "phase-breaking excitations" that violate the fixed-phase condition (ii) are gapped too: for example, if $(C_{ij})_{\mathrm{excited}} = -(C_{ij})_{\mathrm{ground}}$ for a given j and all $i(\neq j)$, the energy cost is $-\lambda d \sum_{i(\neq j)}[(C_{ij} + C_{ji})_{\mathrm{excited}} - (C_{ij} + C_{ji})_{\mathrm{ground}}]$, which is at least of order $2\Delta_{\mathrm{can}}$.

We see, therefore, that the essence of pairing correlations can readily be formulated in a canonical framework: (i) a redistribution of occupation probability across ε_F occurs, such that each level j in a finite range around ε_F has a finite probability of both being doubly occupied or empty, and (ii) any two components of the ground state wavefunction that differ only by the exchange of a pair of electrons between two levels i and j have the same phase.

Since pairing correlations with these properties are the microscopic property at the heart of all manifestations of "superconductivity", it seems reasonable to call a sample "superconducting" as long as it exhibits pairing correlations with measurable consequences. And by this criterion the gap observed in the even grains of RBT certainly qualifies.

4.3.4. Gauge symmetry breaking

Note that property (ii) will be preserved under the gauge transformation $c_{j\sigma} \to e^{i\phi'_j} c_{j\sigma}$, i.e. $C_{ij} \to e^{-2i(\phi'_i - \phi'_j)} C_{ij}$, only if all ϕ'_i are equal, say $\phi'_j = \phi'$. Property (ii), and likewise the pairing parameter Δ_{can}, therefore (a) are not gauge invariant "locally" in j-space, but (b) are gauge invariant globally. These are obvious consequences of the facts that (a) a *correlated fixed-N* state consists of a *phase-coherent* superposition of many different components, and hence cannot be invariant under arbitrary changes of the phases of individual components; and that (b) all of these components contain the *same* number of electrons N and hence under a global gauge transformation all pick up the *same* phase factor $e^{iN\phi'}$. Obviously, global gauge symmetry can therefore never be broken in a canonical ensemble. In contrast, the breaking of global gauge symmetry by the g.c. pairing parameter Δ_{gc} of Eq. (48), which transforms as $\Delta_{\mathrm{gc}} \to e^{2i\phi'} \Delta_{\mathrm{gc}}$, is an inevitable consequence or artefact of its g.c. definition ([105], p. 142).

4.3.5. Making contact wiht standard BCS theory

One of the breakthrough achievements of BCS was, of course, to propose a simple variational ground state which has precisely properties (i) and (ii) described above: when evaluating the correlators of Eq. (49) using $|\mathrm{BCS}\rangle$ of Eq. (47), one finds

$$(\bar{u}_j)^2_{\mathrm{BCS}} = u_j^2, \quad (\bar{v}_j)^2_{\mathrm{BCS}} = v_j^2, \quad (C_{ij})_{\mathrm{BCS}} = u_i v_i u_j v_j e^{-i(\phi_i - \phi_j)} \tag{55}$$

and also $(\Delta_{\mathrm{can}}^2)_{\mathrm{BCS}} = |\Delta_{\mathrm{gc}}|^2$. The definite-phase requirement (ii) can thus be implemented by choosing all the phases ϕ_j to be the same, say $\phi_j = \phi$ for all j, thereby breaking local gauge invariance (usually one simply takes $\phi = 0$); and requirement (i) is fulfilled automatically when minimizing the expectation value $\langle \hat{H} \rangle_{\mathrm{BCS}}$ w.r.t. u_j and v_j, since this does yield smeared-out step functions, namely [101,92]

$$v_j^2 = \tfrac{1}{2}[1 - (\varepsilon_j - \mu)/E_j], \quad E_j \equiv \sqrt{(\varepsilon_j - \mu)^2 + |\Delta_{\mathrm{gc}}|^2} \,. \tag{56}$$

Here we neglected terms that vanish for $d \to 0$, and Δ_{gc} is determined by the famous gap equation (for $T = 0$)

$$\frac{1}{\lambda} = d \sum_{|\varepsilon_j| < \omega_D} \frac{1}{2E_j} \; . \tag{57}$$

The BCS wavefunction instructively illustrates some of the general properties discussed above. Firstly, the product $u_j^2 v_j^2$, shown in Fig. 12, has a bell-shaped form with a well-developed peak around ε_F of width $\simeq |\Delta_{gc}|$, illustrating that pairing correlations are strongest within a region of width $|\Delta_{gc}|$ around the Fermi surface. Secondly, the energy of a blocking excitation [Eq. (54)] reduces to $(\varepsilon_j - \mu)(1 - 2v_j^2) + 2u_j v_j |\Delta_{gc}| = E_j$, which is just the well-known energy of the Bogoljubov quasiparticle state $\gamma_{j+}^\dagger |\mathrm{BCS}\rangle$, where

$$\gamma_{j\sigma} = u_j c_{j\sigma} - \sigma v_j e^{i\phi} c_{j-\sigma}^\dagger \; . \tag{58}$$

Thirdly, an example of a phase-breaking excitation is

$$\gamma_{j+}^\dagger \gamma_{j-}^\dagger |\mathrm{BCS}\rangle = (-v_j e^{-i\phi} + u_j b_j^\dagger) \prod_{i(\neq j)} (u_i + v_i e^{i\phi} b_i^\dagger) |\mathrm{Vac}\rangle \; , \tag{59}$$

which has $(C_{ij})_{\mathrm{excited}} = -u_i v_i u_j v_j$ and energy $2E_j$.

It should be appreciated, however, that BCS chose a *grand-canonical* construction purely for calculational convenience (as is made clear on p. 1180 of their original paper [101]): the trick of using a factorized form of *commuting* products in (47), at the cost of N-indefiniteness, makes it brilliantly easy to determine the variational parameters u_j and v_j. In fact, BCS proposed themselves to use the projection of $|\mathrm{BCS}\rangle$ to fixed N as the actual variational ground state, namely [105]

$$|\mathrm{PBCS}\rangle \equiv \int_0^{2\pi} d\phi \, e^{-i\phi N} \prod_j (u_j + e^{2i\phi} v_j b_j^\dagger) |\mathrm{Vac}\rangle \tag{60}$$

$$= \frac{1}{(N/2)!} \left(\prod_j u_j \right) \left(\sum_j \frac{v_j}{u_j} b_j^\dagger \right)^{N/2} |\mathrm{Vac}\rangle \tag{61}$$

(PBCS for Projected BCS), which *is* of the general form of Eq. (41). In the bulk limit ($d/\tilde{\Delta} \ll 1$), however, it is completely adequate to use $|\mathrm{BCS}\rangle$: firstly, the relative error which its factorized form causes, by taking the occupation amplitude of level j to be independent of that of level i, scales like $1/N$ [105, pp. 150,163]; and secondly, the fluctuations in its particle number, $(\Delta N^2)_{\mathrm{BCS}} = \langle N^2 \rangle_{\mathrm{BCS}} - N^2 = \sum_j (2u_j v_j)^2$, are equal to $\pi \tilde{\Delta}/d$ in the bulk limit, in which the relative fluctuations $(\Delta N^2)_{\mathrm{BCS}}/N^2 \propto d\tilde{\Delta}/\varepsilon_F^2$ therefore vanish. Thus, bulk results obtained from $|\mathrm{PBCS}\rangle$ or $|\mathrm{BCS}\rangle$ are essentially identical. In fact, Braun [35,91] checked by explicit calculation that the functions $(\bar{v}_j^2)_G$, $(\bar{v}_j^2)_{\mathrm{PBCS}}$ and v_j^2 are practically indistinguishable even for $d/\tilde{\Delta}$ as large as 0.5 (see Section 5.3). Significant differences *do* develop between them once $d/\tilde{\Delta}$ increases past 0.5, however, as will be discussed in Section 5.

To end this section, note that Fig. 12 offers a very simple intuitive picture for why pairing correlations weaken with increasing level spacing until, in Anderson's words [45], "superconductivity is no longer is possible" when $d \gtrsim \tilde{\Delta}$: an increase in level spacing implies a decrease in the

number of levels within $\tilde{\Delta}$ of ε_F for which $u_j^2 v_j^2$ differs significantly from zero, i.e. a decrease in the number of pairs with significant pairing correlations. This number, namely $\tilde{\Delta}/d$, can roughly speaking be viewed as the "number of Cooper pairs" of the system, and when it becomes less than one, as in Fig. 12(c), it no longer makes sense to call the system "superconducting". However, this should not be taken to imply that pairing correlations cease altogether in this regime; remnants of them do persist, in the form of fluctuations, up to arbitrarily large $d/\tilde{\Delta}$, as will be discussed in detail in Section 5.3.

4.4. Generalized variational BCS approach

In the next several sections we review the generalized variational BCS approach used by Braun et al. [35,23,91] to describe the paramagnetic breakdown of superconductivity in nm-scale grains in a magnetic field. This theory produces theoretical excitation spectra that are in good qualitative agreement with the measurements of BRT shown in Fig. 10 and thereby yields the most direct confirmation available of the relevance to experiment of the discrete BCS model. Moreover, it sheds considerable light on how "superconductivity breaks down" (more precisely, how pairing correlations weaken) with increasing d and h: As mentioned in the previous paragraph, in grains with $d \simeq \tilde{\Delta}$ (bulk gap), near the lower size limit [45] of observable superconductivity, the number of free-electron states with strong pairing correlations (those within $\tilde{\Delta}$ of ε_F) is of order one. Thus, even in grains in which a spectral gap can still be observed, pairing correlations are expected to become so weak that they might be destroyed by the presence of a single unpaired electron [21]. This can be probed directly by turning on a magnetic field, since its Zeeman energy favors paramagnetic states with nonzero total spin.

The theory reviewed below exploits analogies to thin films in a parallel magnetic field [98,99], but explicitly takes account of the discreteness of the grain's spectrum. Since in RBT's experiments the temperature $T = 50\,\text{mK}$ is much smaller than all other energy scales $(d, \tilde{\Delta})$, we shall neglect finite-temperature effects and set $T = 0$. In Section 4.4 the eigenenergies $\mathcal{E}_\alpha$ of the grain's lowest-lying eigenstates $|\alpha\rangle$ are calculated approximately using a *generalized g.c. variational BCS approach* that goes beyond standard mean-field theory by using a different pairing parameter Δ_α for each $|\alpha\rangle$. The $\mathcal{E}_\alpha$ are then used to discuss various observable quantities, such as h-dependent excitation spectra (Section 4.6), direct experimental evidence for the dominance of purely *time-reversed* states in the pairing interaction (Section 4.7), and various parity effects (Section 4.8).

The reasons for deciding to calculate the excitation spectra, despite their *fixed-N* nature, within a *grand-canonical* framework are as follows: Firstly, its simplicity. Secondly and perhaps most importantly, the exact eigenenergies have the general form $\mathcal{E}_\alpha = \mathcal{E}_n + \mathcal{E}_B(h)$ [Eq. (42)], in which *all h-dependence resides in the exactly known contribution $\mathcal{E}_B(h)$ from the blocked levels*. The choice of approximation scheme therefore only affects $\mathcal{E}_n$, which determines the $h = 0$ properties of the spectrum, such as the size of the zero-field spectral gap, etc., but not the qualitative features of the h-dependence. In particular, this means that all of the analysis below could easily be "made exact" by simply replacing the g.c. approximations for $\mathcal{E}_n$ by the exact values from Richardson's solution. However, this is expected to cause only slight quantitative differences, since, thirdly, canonical calculations (mentioned after Eq. (61) and discussed in Section 5) yield very similar results to g.c. ones as long as $d/\tilde{\Delta} \lesssim 0.5$, which, by inspection of Fig. 10, does seem to be the case for the grain in question (the analysis of Section 4.6 yields $d/\tilde{\Delta} \simeq 0.67$).

4.4.1. The generalized variational ansatz

The Zeeman term in the Hamiltonian of Eq. (37) favors states with a nonzero total z-component of the total spin, $s = \frac{1}{2}\sum_{j\sigma} \sigma c^\dagger_{j\sigma} c_{j\sigma}$ (henceforth simply called "spin"). Increasing h will thus eventually lead to a series of ground state changes to states with successively larger spins. In general, therefore, we are interested in pair-correlated states with nonzero spin, and in particular in their eigenenergies. Following Braun et al. [22,23,91], we now show how this can be calculated variationally, using the following general BCS ansatz for a state $|s, B\rangle$ with $N = 2n + 2s$ electrons and a definite total spin $s \geq 0$ (first introduced by Soloviev for application in nuclei [104]):

$$|s, B\rangle = \prod_{i \in B} c^\dagger_{i+} \prod_j^U (u_j^{(s,B)} + v_j^{(s,B)} b_j^\dagger)|\text{Vac}\rangle . \tag{62}$$

If the spin is nonzero, it is built up by placing $2s$ unpaired spin-up electrons in a set B of $b = 2s$ single-particle levels [cf. Eq. (40)] while the remaining single-particle levels have BCS-like amplitudes to be either empty ($u_j^{(s,B)}$) or doubly occupied by a pair ($v_j^{(s,B)}$), with $(u_j^{(s,B)})^2 + (v_j^{(s,B)})^2 = 1$. The product $\prod_j^U$ thus constitutes a g.c. approximation to the state $|\Psi_n\rangle$ of Eq. (41). The superscript U over products (and over sums below) indicates exclusion of the singly occupied levels in B, for which $u^{(s,B)}$, $v^{(s,B)}$ are not defined.

More specifically, in a given spin-s sector of Hilbert space the following two types of specializations of Eq. (62) were studied in detail ($p = 2s \bmod 2$):

$$|s\rangle = \prod_{i=-s+p/2}^{s-1+p/2} c^\dagger_{i+} \prod_j^U (u_j^s + v_j^s b_j^\dagger)|\text{Vac}\rangle , \tag{63}$$

$$|s, k\rangle = c^\dagger_{(s-1+p/2+k)+} c_{(s-1+p/2)+}|s\rangle , \tag{64}$$

where $|s\rangle$ is the spin-s state with the lowest energy, i.e. the "variational spin-s ground state", obtained by placing the $2s$ unpaired electrons as close as possible to ε_F [Figs. 13(b) and (c)], in order to minimize the kinetic energy cost of having more spin ups than downs. $|s, k\rangle$ is a particular type of excited spin-s state, obtained from $|s\rangle$ by moving one electron from its topmost occupied level ($s - 1 + p/2$) upwards by k units of d into a higher level ($s - 1 + p/2 + k$). These constructions are illustrated in Fig. 13, of which (a) and (b) represent the variational ground states of a grain with an even or odd number of electrons, respectively.

The orthogonality of the wavefunctions, $\langle s, B|s', B'\rangle = \delta_{ss'}\delta_{BB'}$, implies that the variational parameters $v_j^{(s,B)}$ and $u_j^{(s,B)}$ must be found *anew* for each (s, B) (hence the superscript), by minimizing the variational "eigenenergies"

$$\mathscr{E}_{s,B}^{\text{BCS}}(h, d) \equiv \langle s, B|\hat{H}|s, B\rangle \tag{65}$$

$$= -2sh + \sum_{i \in B} (\varepsilon_i - \mu) + \sum_j^U [2(\varepsilon_j - \mu)(v_j^{(s,B)})^2 + \lambda d(v_j^{(s,B)})^4]$$

$$- \lambda d \left(\sum_j^U u_j^{(s,B)} v_j^{(s,B)} \right)^2 , \tag{66}$$

which we use as approximations to the exact eigenenergies $\mathscr{E}_{s,B}^{\text{exact}}(h, d)$. The v_j^4 term is not extensive and hence neglected in the bulk case where only effects proportional to the system volume are of interest. Here it is retained, since in ultrasmall systems it is nonnegligible (but not dominant either)
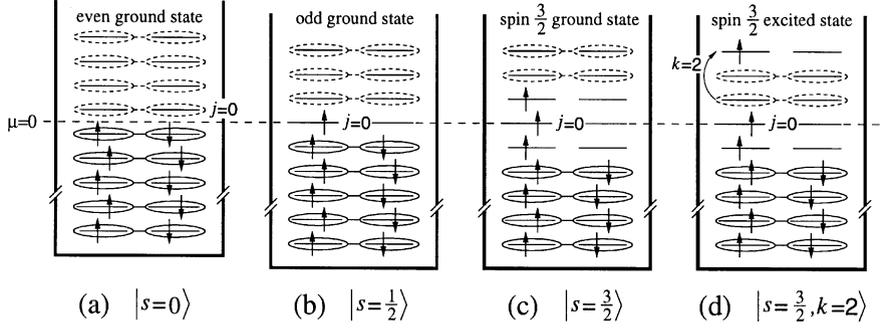
Fig. 13. Cartoon of four typical variational states, labeled using the notation of Eq. (63) for (a–c) and Eq. (64) for (d). They represent (a) the even ground state $|0\rangle$; (b) the odd ground state $|\frac{1}{2}\rangle$; (c) the spin-$\frac{3}{2}$ ground state $|\frac{3}{2}\rangle$; (d) a spin-$\frac{3}{2}$ excited state $|\frac{3}{2}, 2\rangle$. The single-particle levels are drawn for $h = 0$, with the chemical potential half-way between levels 0 and 1 for even systems (a), but exactly on level 0 for odd ones (b,c,d). The ellipses joining states on the same level are meant to represent a "Cooper pair", and signify its being empty or doubly occupied with amplitude $(u_j + v_j b_j^\dagger)$; solid (dashed) ellipses are used for levels that would be completely filled (empty) in the absence of pairing correlations.

[22,23]. Solving the variational conditions $\partial \mathscr{E}_{s,B}^{\mathrm{BCS}} / \partial v_j^{(s,B)} = 0$ in standard BCS fashion yields

$$(v_j^{(s,B)})^2 = \tfrac{1}{2}(1 - \xi_j/[\xi_j^2 + \Delta_{s,B}^2]^{1/2}), \quad \xi_j \equiv \varepsilon_j - \mu - \lambda d (v_j^{(s,B)})^2 \,, \tag{67}$$

where the "pairing parameter" $\Delta_{s,B}$ is defined by the relation

$$\Delta_{s,B} \equiv \lambda d \sum_j^U u_j^{(s,B)} v_j^{(s,B)} \quad \text{or} \quad \frac{1}{\lambda} = d \sum_j^U \frac{1}{2\sqrt{\xi_j^2 + \Delta_{s,B}^2}} \,, \tag{68}$$

which in the limit $d/\tilde{\Delta} \to 0$ reduces to the standard bulk $T = 0$ gap equation. Note that it is h-independent, because it involves only unblocked levels $j \in U$, which are populated by pairs with zero total Zeeman energy. Note also that in Eq. (67) the $\lambda d (v_j^{(s,B)})^2$ shift in ξ_j, usually neglected because it simply renormalizes the bare energies, is retained, since for large d it somewhat increases the effective level spacing near ε_F (and its neglect turns out to produce a significant upward shift in the $\mathscr{E}_{s,B}^{\mathrm{BCS}}(h,d)$'s, which one is trying to minimize).

The chemical potential μ is fixed by requiring that

$$2n + 2s = \langle s, B|\hat{N}|s, B\rangle = 2s + 2\sum_J^U (v_j^{(s,B)})^2 \,. \tag{69}$$

In contrast to conventional BCS theory, the pairing parameter $\Delta_{s,B}$ can, in general, not be interpreted as an energy gap and is *not* an observable. It should be viewed simply as a mathematical auxiliary quantity which was introduced to conveniently solve the variational conditions. However, by parameterizing $v_j^{(s,B)}$ and $u_j^{(s,B)}$, $\Delta_{s,B}$ does serve as a measure of the pairing correlations present in $|s, B\rangle$: for vanishing $\Delta_{s,B}$ the latter reduces to an uncorrelated paramagnetic state $|s, B\rangle_0$ with spin s and energy $\mathscr{E}_{s,B}^0$, namely

$$|s, B\rangle_0 \equiv \prod_{i \in B} c_{i+}^\dagger \prod_{j<0}^U b_j^\dagger |0\rangle \quad \text{with} \quad \mathscr{E}_{s,B}^0 \equiv {}_0\langle s, B|\hat{H}|s, B\rangle_0 \tag{70}$$

and the condensation energy $E_{s,B}^{\mathrm{cond}} \equiv \mathscr{E}_{s,B}^{\mathrm{BCS}} - \mathscr{E}_{s,B}^0$ of $|s, B\rangle$ reduces to zero.

4.4.2. General numerical solution — illustration of the blocking effect

The simultaneous solution of Eqs. (67)–(69) is a straightforward numerical exercise which Braun and von Delft performed [22,23], for the sake of "numerical consistency", without further approximations. (Analytical solutions can be found only in the limits $d \ll \tilde{\Delta}$ and $d \gg \Delta_{s,B}$, see Appendix A of [23].) The numerical results are summarized in Fig. 14, which shows the pairing parameters $\Delta_{s,B}$ [Figs. 14(a) and (c)] and energies $\mathscr{E}_{s,B}^{\mathrm{BCS}}$ [Figs. 14(b) and (d), solid lines] of some selected variational states $|s, B\rangle$, as well as the energies $\mathscr{E}_{s,B}^{0}$ of the corresponding uncorrelated states $|s, B\rangle_0$ [Figs. 14(b) and (d), dashed lines]; both $\mathscr{E}_{s,B}^{\mathrm{BCS}}$ and $\mathscr{E}_{s,B}^{0}$ are plotted relative to the energy $\mathscr{E}_{p/2}^{0}$ of the uncorrelated spin-p Fermi sea $|p/2\rangle$. The results have a number of salient features.

(i) In the bulk limit $d/\tilde{\Delta} \to 0$, all of the pairing parameters $\Delta_{s,B}$ reduce to $\tilde{\Delta}$, as expected, and the energy differences $\mathscr{E}_{s,B}^{\mathrm{BCS}} - \mathscr{E}_{p/2}^{0}$ between the correlated states $|s, B\rangle$ and the uncorrelated Fermi sea $|p/2\rangle$ reduce to $-\frac{1}{2}\tilde{\Delta}^2/d = -\frac{1}{2}\mathscr{N}(\varepsilon_{\mathrm{F}})\tilde{\Delta}^2$, which is the standard bulk result for the condensation energy.

(ii) Each $\Delta_{s,B}$ in Figs. 14(a) and (c) decreases with increasing d. This reflects the fact that with increasing d, the number of pair-correlated states within $\tilde{\Delta}$ of ε_{F} decreases (cf. Fig. 12 and the last paragraph of Section 4.3.5), so that the amount of pairing correlations, for which $\Delta_{s,B}$ is a measure, decreases too.

(iii) Each $\Delta_{s,B}$ vanishes abruptly at a critical level spacing $d_{s,B}^{\mathrm{BCS}}$ (whose precise numerical value depends sensitively on model assumptions such as the value of λ and the use of uniformly spaced levels [24]). For $d > d_{s,B}^{\mathrm{BCS}}$ no pairing correlations exist at this level of approximation, so that the
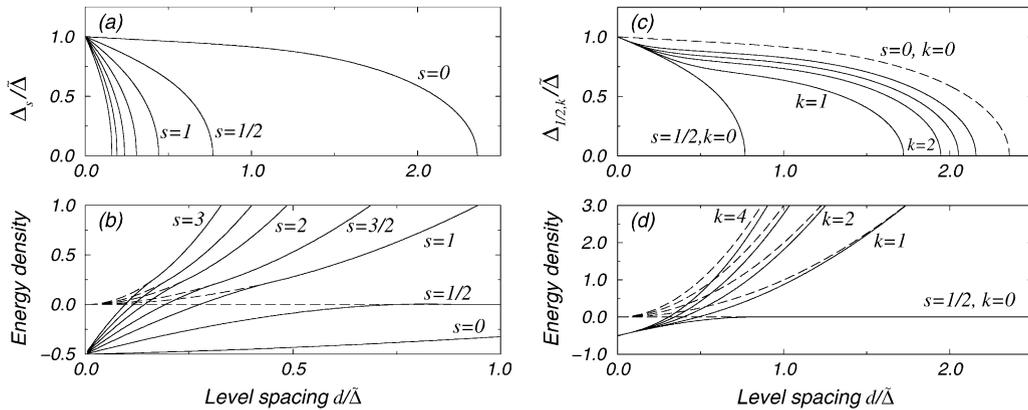

Fig. 14. Properties of (a,b) spin-s ground states $/s\rangle = |s,0\rangle$ [Eq. (63)] and (c,d) spin-$\frac{1}{2}$ excited states $|\frac{1}{2}, k\rangle$ (for $k = 0, \ldots, 4$) [Eq. (64)], as functions of $d/\tilde{\Delta}$ (i.e. decreasing grain size), calculated for $\lambda = 0.194$. (a) The pairing parameters $\Delta_s(d)/\tilde{\Delta}$, which vanish at critical level spacings d_s^{BCS} of $2.36, 0.77, 0.44, 0.31, \ldots$ for $s = 0, 1/2, 1, 3/2, \ldots$, respectively. (c) The pairing parameters $\Delta_{1/2,k}$, together with their $k \to \infty$ limit, which equals Δ_0 (dashed line). (b,d) show relative energy *densities* (since normalized by $d/\tilde{\Delta}^2 \propto \mathrm{Vol}^{-1}$) at $h = 0$ for both correlated (solid) and uncorrelated (dashed) states, the latter obtained by setting $\Delta_{s,B} = 0$ in the former. (b) shows $(\mathscr{E}_s^{\mathrm{BCS}} - \mathscr{E}_{p/2}^{0})d/\tilde{\Delta}^2$ (solid) and $(\mathscr{E}_s^{0} - \mathscr{E}_{p/2}^{0})d/\tilde{\Delta}^2$ (dashed), the energy differences of $|s\rangle$ and $|s\rangle_0$ relative to the uncorrelated spin-$p/2$ Fermi sea $|p/2\rangle_0$. (d) shows $(\mathscr{E}_{1/2,k}^{\mathrm{BCS}} - \mathscr{E}_{1/2}^{0})d/\tilde{\Delta}^2$ (solid) and $(\mathscr{E}_{1/2,k}^{0} - \mathscr{E}_{1/2}^{0})d/\tilde{\Delta}^2$ (dashed), the energy differences of $|\frac{1}{2}, k\rangle$ and $|\frac{1}{2}, k\rangle_0$ relative to the uncorrelated spin-$\frac{1}{2}$ ground state $|\frac{1}{2}, 0\rangle_0 = |\frac{1}{2}\rangle$. Solid and dashed lines meet at the critical level spacing $d_{s,k}^{\mathrm{BCS}}$ at which $\Delta_{s,k}$ becomes 0 and the condensation energy $E_{s,k}^{\mathrm{cond}} = \mathscr{E}_{s,k}^{\mathrm{BCS}} - \mathscr{E}_{s,k}^{0}$ vanishes.

condensation energy $E_{s,B}^{\mathrm{cond}}$ (difference between solid and dashed lines) vanishes and the solid and dashed lines in Figs. 14(b) and (d) meet.

(iv) In Fig. 14(a), the pairing parameters $\varDelta_s$ for the spin-s ground states decrease rapidly with increasing s at fixed d (and $d_s^{\mathrm{BCS}} < d_{s'}^{\mathrm{BCS}}$ if $s > s'$). [This is a generalization of a parity effect discussed by von Delft et al. [21], who studied only ground state pairing correlations and found that these are weaker in odd ($s = 1/2$) grains than in even ($s = 0$) grains, $\varDelta_{1/2} < \varDelta_0$, cf. Section 4.8.] This tendency is a direct consequence of the *blocking effect* described in Section 4.2.3 and is independent of model details: larger s means more unpaired electrons, more terms missing from the sum $\sum_j^U$, less correlated pairs and hence smaller $\varDelta_{s,B}$.

(v) As d increases the blocking effect described in (iv) becomes stronger, i.e. the difference between the various $\varDelta_s$ for different s becomes more pronounced, since then the relative weight of each term missing in the sum $\sum_j^U$ increases. The blocking effect is most dramatic in the regime $d/\tilde{\varDelta} \in [0.77, 2.36]$ in which $\varDelta_0 \neq 0$ but $\varDelta_{s \neq 0} = 0$. This is a regime of "minimal superconductivity" [22,23], in the sense that all pairing correlations that still exist in the *even* variational ground state $|0\rangle$ (since $\varDelta_0 \neq 0$) are completely destroyed by the addition of a single electron or the flipping of a single spin (since $\varDelta_{s \neq 0} = 0$).

(vi) Considering the spin-$\frac{1}{2}$ excited states $|\frac{1}{2}, k\rangle$ of Figs. 14(b) and (d), one finds that the larger k, the longer the pairing correlations survive with increasing d: the critical spacings $d_{1/2,k}^{\mathrm{BCS}}$ increase with k, approaching the value d_0^{BCS} of the spin-0 case as $k \to \infty$; correspondingly, the larger k, the larger the d-value at which the condensation energies $E_{1/2,k}^{\mathrm{cond}}$ [differences between solid and dashed lines in Fig. 14(d)] vanish. The intuitive reason why the amount of pairing correlations in an excited $|s, k\rangle$ increases with k is of course quite simple: the further the unpaired electron sits from the Fermi surface where pairing correlations are strongest, the less it disrupts the latter (since $u_k v_k$ becomes very small for large k, see Fig. 12). In fact, the state $|\frac{1}{2}, k \to \infty\rangle$ will have just about the same amount of pairing correlations as the even ground state $|0\rangle$ ($\varDelta_{1/2, k \to \infty} \simeq \varDelta_0$).

(vii) Similar effects hold for excited states in other spin sectors (not shown): The higher the excitation, the larger the pairing parameter $\varDelta_{s,B}$. However, the concomitant gain in correlation energy is always less than the kinetic-energy cost of having an unpaired electron far from ε_{F}.

(viii) The strong dependence of $\varDelta_{s,B}$ on s and d for $d \gtrsim \tilde{\varDelta}$ illustrates why in this regime a conventional mean-field treatment is no longer sufficient: *the system cannot be characterized by a* single *pairing parameter, since the amount of pairing correlations vary from state to state, each of which is characterized by its own pairing parameter*. Instead, the present variational approach is, roughly speaking, equivalent to a doing a *separate* mean-field calculation for each new choice U of unblocked levels within the Fock space spanned by them (i.e. replacing $b_j \to \{b_j - \langle b_j \rangle\} + \langle b_j \rangle$ and neglecting terms quadratic in the fluctuations $\{b_j - \langle b_j \rangle\}$). Indeed, the behavior of $\varDelta_{s,B}(d)$ near $d_{s,B}^{\mathrm{BCS}}$ has the standard mean-field form $\sqrt{1 - d/d_{s,B}^{\mathrm{BCS}}}$, as can be shown analytically ([23], Appendix A).

To summarize: pairing correlations decrease with increasing d and s and decreasing k. These features survive also in more accurate canonical calculations. This is not the case, however, for the abrupt vanishing of $\varDelta_{s,B}$ at $d_{s,B}^{\mathrm{BCS}}$, which signals the breakdown of the g.c. approach once d becomes of order $\tilde{\varDelta}$: canonical methods show that, regardless how large d becomes, some remnants of pairing correlations survive and the pairing parameters $(\varDelta_{s,B})_{\mathrm{can}}$ do not vanish (Section 5), in accordance with the rule of thumb that "in a *finite system* no abrupt phase transition can occur between a zero and nonzero order parameter."

4.5. Softening of the H-induced transition to a paramagnetic state

Since states with nonzero spin are favored by the Zeeman energy but have smaller correlation energy due to the blocking effect, a competition arises between Zeeman energy and correlation energy. The manifestations of the blocking effect can thus be probed by turning on a magnetic field; if it becomes large enough to enforce a large spin, excessive blocking will destroy all pairing correlations.

The situation is analogous to ultra-thin films in a parallel magnetic field [98,99], where orbital diamagnetism is negligible for geometrical reasons and *superconductivity is destroyed at sufficiently large h by Pauli paramagnetism*. This occurs via a first-order transition to a paramagnetic state, as predicted by Clogston and Chandrasekhar (CC) [110,111] by the following argument (for bulk systems): a pure Pauli paramagnet chooses its spin s such that the sum of the kinetic and Zeeman energies, $s^2/\mathcal{N}(\varepsilon_F) - 2hs$, is minimized, and hence has spin $s = h\mathcal{N}(\varepsilon_F)$ and ground state energy $-h^2\mathcal{N}(\varepsilon_F)$. When this energy drops below the bulk correlation energy $-\frac{1}{2}\tilde{\Delta}^2\mathcal{N}(\varepsilon_F)$ of the superconducting ground state, which happens at the critical field $h_{CC} = \tilde{\Delta}/\sqrt{2}$, a transition will occur from the superconducting to the paramagnetic ground state. The transition is first order, since the change in spin, from 0 to $s_{CC} = h_{CC}\mathcal{N}(\varepsilon_F) = \tilde{\Delta}/(d\sqrt{2})$, is macroscopically large ($\mathcal{N}(\varepsilon_F) = 1/d \propto \mathrm{Vol}$).

This transition has been directly observed by Meservey and Tedrow [98,99] in ultra-thin (5 nm) superconducting Al films ($\tilde{\Delta} = 0.38$ meV), whose density of states [Fig. 15(a)] they measured via the tunnel conductance through an oxide layer between a normal metal and the film. They found that in a magnetic field the BCS quasiparticle peak splits up into two subpeaks, separated in energy by $2\mu_B H$ [Fig. 15(b)], which simply reflects the Zeeman splitting of quasiparticles states[12] with spin-up or -down (and $g = 2$). Remarkably, the tunneling threshold abruptly dropped to zero at a field of 4.7 T [Fig. 15(b)], which they associated with the field H_{CC} at which the phase transition from the superconducting to the paramagnetic ground state occurs. Indeed, Fig. 15(b) demonstrates clearly that the transition to the normal state is first order: *the mean of the spin-up and spin-down spectral gaps, i.e. the pairing parameter $\tilde{\Delta}$, is constant until the critical field H_{CC} is reached, at which it abruptly drops to zero*.

For the case of isolated ultrasmall grains, the above picture of the transition needs to be rethought in two respects due to the discreteness of the electronic spectrum: Firstly, the spin must be treated as a discrete (instead of continuous) variable, whose changes with increasing h can only take on (parity-conserving) integer values. Secondly, one needs to consider more carefully the possibility of h-induced transitions to nonzero spin states that are still *pair-correlated* (instead of being purely paramagnetic), such as the variational states $|s, B\rangle$ discussed above. (In the bulk case, it is obvious that such states play no role: the lowest pair-correlated state with nonzero spin obtainable from the ground state by spin flips is a two-quasiparticle state, costing energy $2\tilde{\Delta} - 2h$; when h is increased from 0, the paramagnetic transition at $h_{CC} = \tilde{\Delta}/\sqrt{2}$ thus occurs before a transition to this state, which would require $h = \tilde{\Delta}$, can occur.)

---

[12] Recall that the BCS quasiparticles $\gamma_{j,\sigma}^\dagger = u_j c_{j,\sigma}^\dagger - \sigma v_j c_{j,-\sigma}$ have well-defined spins.
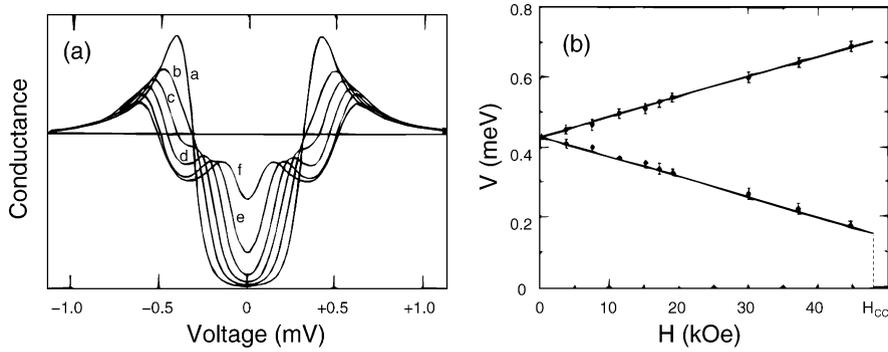
Fig. 15. Thin flims in a magnetic field (Figs. 11 and 12 of [98]). (a) Tunneling conductance from a normal metal through a tunnel barrier into a thin superconducting Al film, as function of voltage, for several magnetic fields labeled in increasing order "a" to "f". The conductance reflects the BCS quasiparticle density of states, whose single peak (for a given sign of V) for $H = 0$ splits into two separate peaks for $H \neq 0$, corresponding to the Zeeman energy difference between quasiparticles with spin-up and -down. (b) Voltage corresponding to the maxima of spin-up and spin-down density of states as a function of magnetic field. At the critical field H_{CC} superconductivity is destroyed and the tunneling threshold drops abruptly to zero.

Quite generally, the effect of increasing h from 0 can be analyzed as follows: At given d and h, the grain's ground state is the lowest-energy state among all possible spin-s ground states $|s\rangle$ having the correct parity $p = 2s \bmod 2$. Since $\mathscr{E}_s(h,d) = \mathscr{E}_s(0,d) - 2hs$, level crossings occur with increasing h, with $\mathscr{E}_{s'}$ dropping below $\mathscr{E}_s$ at the *level crossing field*

$$h_{s,s'}(d) = \frac{\mathscr{E}_{s'}(0,d) - \mathscr{E}_s(0,d)}{2(s' - s)} \ . \tag{71}$$

Therefore, as h is slowly turned on from zero with initial ground state $|s_0 = p/2\rangle$, a cascade of successive ground state changes (GSCs) to new ground states $|s_1\rangle, |s_2\rangle, \ldots$ will occur at the fields $h_{s_0,s_1}, h_{s_1,s_2}, \ldots$. Let us denote this cascade by $(s_0, s_1); (s_1, s_2); \ldots$; for each of its GSCs the corresponding levelcrossing field $h_{s,s'}(d)$ is shown in Fig. 16. Generalizing CC's critical field to nonzero d, let us denote the (parity-dependent) field at which the *first* transition (s_0, s_1) occurs by $h_{CC}(d, p) \equiv h_{s_0,s_1}(d)$, which simply is the lower envelope of the level-crossing fields h_{s_0,s_1} in Fig. 16 (shown as bold solid and dashed lines for $s_0 = 0$ and $s_0 = \frac{1}{2}$, respectively). In the limit $d \to 0$ it is numerically found to reduce to the Clogston–Chandrasekhar value, i.e. $h_{CC}(0, p) = \tilde{\Delta}/\sqrt{2}$, as expected.

In general, the order in which the GSCs occur with increasing h within a cascade (i.e. the order of $h_{s,s'}$ lines encountered when moving vertically upward in Fig. 16) depends sensitively on d, and an infinite number of distinct regimes (cascades) I,II,III, … can be distinguished: starting at large d we find the typical normal-grain behaviour $(0, 1); (1, 2); (2, 3); \ldots$ for even grains and $(\frac{1}{2}, \frac{3}{2}); (\frac{3}{2}, \frac{5}{2}); \ldots$ for odd grains, with $h_{0,1} < (\text{or} >) h_{1/2,3/2}$ in regimes I (or II). In regimes III and IV of somewhat smaller d, the order of GSCs is $(0, 2); (2, 3); \ldots$ and $(\frac{1}{2}, \frac{3}{2}); (\frac{3}{2}, \frac{5}{2}); \ldots$, etc., i.e. the spin s_1 attained after the first GSC (s_0, s_1) has increased to 2 in the even case. This illustrates a general trend: *the spin $s_1(d)$ after the first transition increases with decreasing d and becomes macroscopically large in the $d \to 0$ limit*, where $s_1 = h_{CC}/d = \tilde{\Delta}/(d\sqrt{2})$, as explained in recounting CC's argument above.
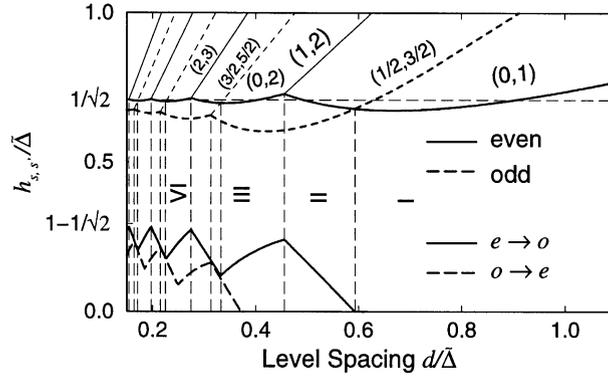
Fig. 16. d-dependence of the level-crossing fields $h_{s,s'}(d)/\tilde{\Delta}$ [Eq. (71)] at which $\mathscr{E}_{s'}^{\text{BCS}}$ drops below $\mathscr{E}_{s}^{\text{BCS}}$ with increasing h. Only those level crossing fields are shown that belong to the cascade of (*fixed-N*) ground state changes (GSCs) $(s_0, s_1); (s_1, s_2); \ldots$ that occur as h increases from 0 at given d. Solid (dashed) lines are used for even (odd) grains with integer (half-interger) spins, and some are labeled by the associated GSC (s, s'). (In contrast, in Fig. 17 the *N-changing* tunneling transitions in are labeled by $|s_i\rangle \rightarrow |s_f\rangle$.) The size $|\Delta\mathscr{E}_{s_1, f'} - \Delta\mathscr{E}_{s_0, f}|$ of the first jump (occurring at the level-crossing field $h_{\text{CC}}(p, d) = h_{s_0, s_1}$) in the lowest line of the tunneling spectra of Fig. 17 is shown by the lowest two (jagged) curves (solid for $e \rightarrow o$ and dashed for $o \rightarrow e$ tunneling spectra), which both approach the CC value $1 - 1/\sqrt{2}$ as $d \rightarrow 0$.

Furthermore, it turns out that $\Delta_{s_1}(d) = 0$ and therefore $\mathscr{E}_{s_1}^{\text{BCS}} = 0$ for *all d*, implying that after the first GSC the new ground state $|s_1\rangle$ is *always* (not only in CC's bulk limit) a purely paramagnetic state, i.e. without any pairing correlations in the g.c. framework (canonical calculations would yield some weak remnant pairing correlations in the form of fluctuations). In this regard, CC's picture of the transition remains valid throughout as d is increased: at $h_{\text{CC}}(d, p)$, a transition occurs from the superconducting ground state to a paramagnetic, *uncorrelated* state $|s_1\rangle_0$, the transition being first order in the sense that $\Delta_{s_1}(d) = 0$; however, *the first-order transition is "softened" with increasing d, in the sense that the size of the spin change, $s_1 - s_0$, decreases from being macroscopically large in the bulk to being equal 1 at $d \gg \tilde{\Delta}$ (regimes I and II)*.

To conclude this section, we mention that the above analysis of the paramagnetic breakdown of superconductivity has recently been generalized to finite temperatures [32], using the so-called static path approximation (explained in Section 5.5.3) to treat fluctuation effects properly.

4.6. Excitation spectrum in a magnetic field

In this section we compare the theoretical tunneling spectra for a grain coupled to leads, calculated as functions of h and d [22,23,91], and compare these to RBT's measurements of Fig. 10.

The form of the tunneling spectrum depends in a distinct way on the specific choice of level spacing d and on the electron number parity p of the final states $|f\rangle$ of the bottleneck tunneling processes $|i\rangle \rightarrow |f\rangle$ (or $|\alpha'\rangle \rightarrow |\alpha\rangle$ in the notation of Section 2.3). However, for the uniformly spaced ε_j-levels used here, particle–hole symmetry ensures that there is no difference between electron addition or removal spectra $|i_{N\mp 1}\rangle \rightarrow |f_N\rangle$. To calculate the spectrum *for given d and p*, Braun et al. [22,23,91] proceeded as follows: they first analyzed at each magnetic field h which tunneling processes $|i\rangle \rightarrow |f\rangle$ are possible, then calculated the corresponding tunneling energy thresholds

$\Delta\mathscr{E}_{if}(h) \equiv \mathscr{E}_f(h) - \mathscr{E}_i(h)$ [cf. Eq. (26)] and plotted $\Delta\mathscr{E}_{if}(h) - \Delta\mathscr{E}_{\min}(0)$ as functions of h for various combinations of $|i\rangle$ and $|f\rangle$, each of which gives a line in the spectrum. Since the selection rule $s_f - s_i = \pm 1/2$ holds, only slopes of ± 1 can occur. The reason for subtracting $\Delta\mathscr{E}_{\min}(0)$, the $h = 0$ threshold energy cost for the *first* (lowest-lying) transition, is that in experiment, this energy depends on V_g and hence yields no significant information, as explained in Section 2.4.6. Neglecting nonequilibrium effects [10–13] (which were minimized in the present experiment by tuning V_g, cf. Section 2.4.4, and which are discussed in Section 6.2), the initial state is always taken to be the ground state of a given spin-s sector. The appropriate $s_i(h,d)$ must be determined from Fig. 16.

Fig. 17 shows four typical examples of such theoretical tunneling spectra, with some lines labeled by the corresponding $|i\rangle \rightarrow |f\rangle$ transitions. Whenever, h passes through one of the level-crossing fields $h_{s_i,s_{i'}}$ of Eq. (71), the grain experiences a ground state change $(s_i, s_{i'})$, at which the set of allowed tunneling transitions changes from $|s_i\rangle \rightarrow \{|s_f\rangle\}$ to $|s_{i'}\rangle \rightarrow \{|s_{f'}\rangle\}$. Therefore, at $h_{s_i,s_{i'}}$ one set of lines in the tunneling spectrum ends and another begins, producing kinks or discontinuities. A *kink* occurs if one of the new final states coincides with one of the old ones, $|f'\rangle = |f\rangle$, meaning that it can be reached from both $|s_i\rangle$ and $|s_{i'}\rangle$ [i.e. $s_f - s_i = -(s_f - s_{i'})$], in which case $\Delta\mathscr{E}_{if}(h)$ and $\Delta\mathscr{E}_{if'}(h)$ have slopes of opposite sign. However, for most lines this is not the case, so that at $h_{s_i,s_i'}$ the line $|s_i\rangle \rightarrow |f\rangle$ simply ends while new lines $|s_{i'}\rangle \rightarrow |f'\rangle$ begin. This results in *discontinuities*
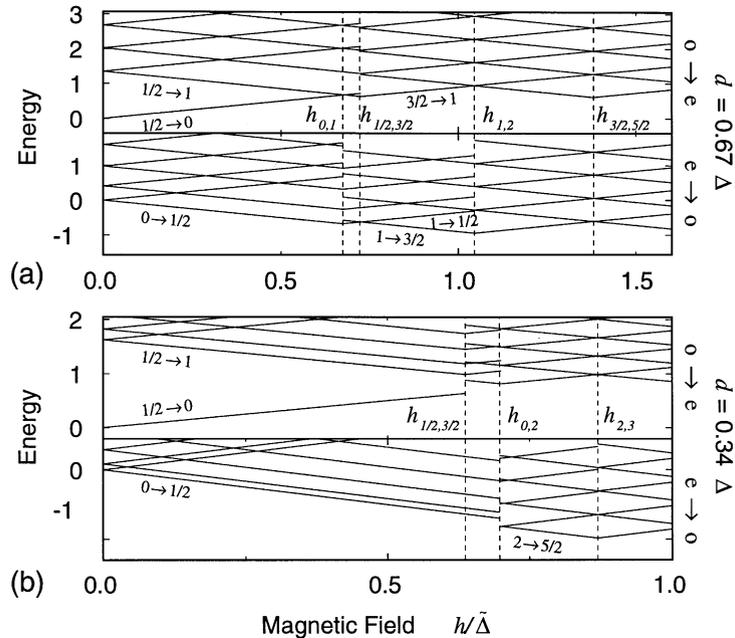
Fig. 17. The theoretical odd-to-even and even-to-odd tunneling spectra $(\Delta\mathscr{E}_{if} - \Delta\mathscr{E}_{\min}(0))/\tilde{\Delta}$ predicted for an ultrasmall superconducting grain as a function of magnetic field h, for two different level spacings: (a) $d = 0.67\tilde{\Delta}$ and (b) $d = 0.34\tilde{\Delta}$ (corresponding to regimes I and III of Fig. 16, respectively). Some lines are labeled by the corresponding $s_i \rightarrow s_i'$ tunneling transitions. Not all possible higher lines (corresponding to excited final states $|s,j\rangle$) are shown. Vertical dashed lines indicate those level-crossing fields $h_{s,s'}$ [Eq. (71)] at which kinks or jumps occur, with $h_{0,1} < h_{1/2,3/2} < h_{1,2} < h_{3/2,5/2}$ in (a) and $h_{1/2,3/2} < h_{0,2} < h_{2,3}$ in (b).

(or "jumps") in the spectrum at $h_{s_i,s_i'}$ of size $(\Delta\mathscr{E}_{i'f'} - \Delta\mathscr{E}_{if})(h_{s_i,s_i'})$, unless by chance some other final state $|f'\rangle$ happens to exist for which this difference equals zero.

Since the order in which the GSCs $(s_i, s_{i'})$ occur as functions of increasing h depend on d and p, as indicated by the distinct regimes I, II, III, ... in Fig. 16, one finds a distinct kind of tunneling spectrum for each regime, differing from the others in the positions of its jumps and kinks. In regime I, where the order of occurrence of GSCs with increasing h is $(0,1); (\frac{1}{2},\frac{3}{2}); (1,2); (\frac{3}{2},\frac{5}{2}); \ldots$, there are no discontinuities in the evolution of the lowest line [see Fig. 17(a)]. For example, for the $e \to o$ spectrum, the lowest $|0\rangle \to |1/2\rangle$ line changes *continuously* to $|1\rangle \to |1/2\rangle$ at $h_{0,1}$, since $|s_f - s_i'| = 1/2$. However, in all other regimes the first change in ground state spin (at h_{0,s_1} from 0 to s_1) is > 1, implying a *jump* (though possibly small) in all $e \to o$ lines, as illustrated by Fig. 17(b).

The jump's magnitude for the tunneling thresholds, i.e. the *lowest* $e \to o$ and $o \to e$ lines, is shown as function of d in the lower part of Fig. 16. It starts at $d = 0$ from the CC value $\tilde{\Delta}(1 - 1/\sqrt{2})$ measured for thin Al films [98,99], and with increasing d decreases to 0 (nonmonotonically, due to the discrete spectrum). This *decrease of the size of the jump in the tunneling threshold* reflects the fact, discussed in Section 4.5, that the change in spin at the first ground state change (s_0, s_1) decreases with increasing d (as $s_1 - s_0 \sim h_{CC}/d$), and signals the softening of the first-order superconducting-to-paramagnetic transition.

The fact that the measured tunneling thresholds in Fig. 10 show no jumps at all, which might at first seem surprising when contrasted to the threshold jumps seen at h_{CC} in Fig. 15 for thin films in a parallel field [98,99], can therefore naturally be explained [22,23] by assuming the grain to lie in the "minimal superconductivity" regime I of Fig. 16 (where the jump size predicted in Fig. 16 is zero). Indeed, *the overall evolution (i.e. order and position of kinks, etc.) of the lowest lines of Fig. 10 qualitatively agrees with those of a regime* I *tunneling spectrum* (*Fig. 17(a)*). This important result rather convincingly establishes the phenomenological success of the discrete BCS model. It also allows one to deduce the following values for the level-crossing fields $H_{s_i,s_i'}$ (indicated by vertical dashed lines in Figs. 10 and 17): $H_{0,1} = 4$ T, $H_{1/2,3/2} = 4.25$ T, $H_{1,2} = 5.25$ T and $H_{3/2,5/2} = 6.5$ T. As corresponding uncertainties we take $\Delta H_{s_i,s_i'} = 0.13$ T, which is half the H resolution of 0.25 T used in experiment.

By combining the above $H_{s_i,s_i'}$ values with Fig. 16, some of the grain's less-well-known parameters can be determined somewhat more precisely:

(i) To estimate the grain's "bulk H_{CC}", note that since $H_{1/2,3/2}/H_{0,1} \simeq 1.06$, this grain lies just to the right of the boundary between regions II and I in Fig. 16 where $d/\tilde{\Delta} \simeq 0.63$, at which we have $h_{0,1}/h_{CC} \simeq 0.95$, so that $H_{CC} = H_{0,1}/0.95 \simeq 4.2$ T. This is quite close to the value $H_{CC} \simeq 4.7$ T found experimentally [98,99] in thin films in a parallel field, confirming our expectation that these correspond to the "bulk limit" of ultrasmall grains as far as paramagnetism is concerned.

(ii) The grain's corresponding bulk gap is $\tilde{\Delta} = \sqrt{2}\mu_B H_{CC} \simeq 0.34$ meV, implying a coupling constant of $\lambda = 0.189$ [by Eq. (39)]. A posteriori, these values can be regarded as being more appropriate for the present grain than the choices $\tilde{\Delta} = 0.38$ meV and $\lambda = 0.194$ made in Section 4.4.2, though the differences are evidently not significant (12% for $\tilde{\Delta}$ and 3% for λ).

(iii) The mean level spacing implied by $d/\tilde{\Delta} \simeq 0.63$ is $d \simeq 0.21$ meV. The crude volume-based value $d \simeq 0.45$ meV cited in the caption of Fig. 2 thus seems to have been an overestimate. It would

be useful if this determination of d could be checked via an independent accurate experimental determination of d directly from the spacing of lines in the excitation spectrum. Regretably, this is not possible: the measured levels are shifted together by pairing interactions, implying that their spacing does not reflect the mean *independent*-electron level spacing d. Nevertheless, note that the measured spacing of 0.05 meV between the lowest two states of the odd grain agrees quite well with the crude BCS estimate $\sqrt{\tilde{\Delta}^2 + d^2} - \tilde{\Delta}$ [cf. Eq. (56)], which gives 0.06 meV when evaluated for $d = 0.21$ meV and $\tilde{\Delta} = 0.34$ meV.

The higher lines plotted in Fig. 17 correspond to transitions into spin-s_f state of the form $|s_f, k\rangle$ [cf. Eq. (64) and Fig. 13(d)] (for simplicity these were the only ones considered in [22,23,91], though in general others are expected to occur too). The jumps in these lines, e.g. in Fig. 17(a) at $h_{1,2}$, occur whenever the two final excited states $|s_f, k_f\rangle$ and $|s_{f'}, k_{f'}\rangle$ before and after the GSC at $h_{s_i, s_i'}$ have different correlation energies. (Recall that the correlation energy of an excited state $|s_f, B_f\rangle$ can be nonzero even if that of the corresponding ground state $|s_f\rangle$ is zero, since the former's unpaired electrons are further away from ε_F, so that $\Delta_{s_f, B_f} > \Delta_{s_f}$, see point (vi) of Section 4.4.) Experimentally, these jumps have not been observed. This may be because up-moving resonances lose amplitude and are difficult to follow [10] with increasing h, or because the widths of the excited resonances ($\simeq 0.13\tilde{\Delta}$) limit energy resolution [11–13].

For somewhat larger grains, the present theory predicts jumps even in the lowest line, as illustrated in Fig. 17(b). It remains to be investigated, though, whether orbital effects, which rapidly increase with the grain size, would not smooth out such jumps.

To conclude this section, we emphasize once again that more than qualitative agreement between theory and experiment cannot be expected, since both the model and our variational treatment thereof are very crude: the model neglects, for instance, fluctuations in level spacing and in pair-coupling constants, and the g.c. wavefunctions become unreliable for $d/\tilde{\Delta} \gtrsim 0.5$. Furthermore, we neglected nonequilibrium effects in the tunneling process and assumed equal tunneling matrix elements for all processes. In reality, though, random variations of tunneling matrix elements could suppress some tunneling processes which would otherwise be expected theoretically.

4.7. Time-reversal symmetry

In this section we argue that RBT's spectra give direct support for the dominance of purely time-reversed states in the pairing interaction, implying that the sufficiency of using only a reduced BCS Hamiltonian, well established for bulk systems and dirty superconductors, holds for ultra-small grains, too.

When defining the discrete BCS model in Eq. (37), we adopted a *reduced* BCS Hamiltonian, in analogy to that conventionally used for macroscopic systems. In doing so, we neglected interaction terms of the form

$$-d \sum_{iji'j'} \lambda(i,j,i',j') c_{i+}^{\dagger} c_{j-}^{\dagger} c_{i'-} c_{j'+} \qquad (72)$$

between nontime-reversed pairs $c_{i+}^{\dagger} c_{j-}^{\dagger}$, following Anderson's argument [45] that for a short-ranged interaction, the matrix elements involving time-reversed states $c_{j+}^{\dagger} c_{j-}^{\dagger}$ are much larger than all others, since their orbital wavefunctions interfere constructively. (This argument can be substantiated using $1/g_{\mathrm{dim}}$ as small parameter [13], where g_{dim} of Eq. (4) is the grain's dimensionless

conductance, see Section 6.1.3.) Interestingly, the experimental results of RBT provide strikingly direct support for the correctness of neglecting interactions between nontime-reversed pairs of the form (72) at $h = 0$: Suppose the opposite, namely that the matrix elements $\lambda(j + k, j, j' + k', j')$ were all roughly equal to λ for a finite range of k- and k'-values [instead of being negligible for k or $k' \neq 0$, as assumed in the reduced BCS Hamiltonian (37)]. Then for $2s < k$, one could construct a spin-s state $|s\rangle'$ with manifestly lower energy ($\mathscr{E}'$) than that ($\mathscr{E}$) of the state $|s\rangle$ of Eq. (63), namely

$$|s\rangle' = \prod_{i=j_{\min}}^{j_{\min}+2s-1} c_{i+}^{\dagger} \prod_{j=j_{\min}}^{\infty} (u_j^{(s)} + v_j^{(s)} c_{(j+2s)+}^{\dagger} c_{j-}^{\dagger})|\mathrm{Vac}\rangle \,, \tag{73}$$

where $j_{\min}$ labels the lowest lying of the interacting levels (i.e. the bottom of the interacting band). Whereas in $|s\rangle$ pairing correlations involve only time-reversed partners, in $|s\rangle'$ we have allowed correlations between *non*time-reversed partners, while choosing the $2s$ unpaired spin-up electrons that occupy their levels with unit amplitude to sit at the band's *bottom* (see Fig. 18). To see that $|s\rangle'$ has lower energy than $|s\rangle$,

$$\mathscr{E}_s' = E_s'^{\text{cond}} + \mathscr{E}_s'^{0} < E_s^{\text{cond}} + \mathscr{E}_s^{0} = \mathscr{E}_s \,, \tag{74}$$

we argue as follows: Firstly, $\mathscr{E}_s'^{0} = \mathscr{E}_s^{0}$, since the corresponding uncorrelated states $|s\rangle_0'$ and $|s\rangle_0$ are identical (and given by Eq. (70), where B is the set of levels $i = -s + p/2, \ldots, s - 1 + p/2$). Secondly, $\Delta_s' = \Delta_0$ (implying $E_s'^{\text{cond}} = E_0^{\text{cond}}$), because the $2s$ unpaired electrons in $|s\rangle'$ sit at the band's bottom, i.e. so far away from ε_F that their blocking effect is negligible [cf. point (vi) of Section 4.4.2]. In contrast, $\Delta_s < \Delta_0$ (implying $E_s^{\text{cond}} > E_0^{\text{cond}}$), because the $2s$ unpaired electrons in $|s\rangle$ sit around ε_F and cause significant blocking. Thus the condensation energies satisfy $E_s'^{\text{cond}} < E_s^{\text{cond}}(\leq 0)$, so that Eq. (74) holds, implying that $|s\rangle'$ would be a better variational ground state for interaction (72) than $|s\rangle$.

Now, the fact that $E_s'^{\text{cond}} = E_0^{\text{cond}}$ is *independent* of s means that flipping spins in $|s\rangle'$ does not cost correlation energy. Thus, the energy cost for turning $|0\rangle'$ into $|1\rangle'$ by flipping one spin is simply the kinetic energy cost d, implying a threshold field $h_{0,1}' = d/2$ [see Eq. (71)]; in contrast, the cost for turning $|0\rangle$ into $|1\rangle$, namely ($\mathscr{E}_1 - \mathscr{E}_0$), implies a threshold field $h_{0,1} = \frac{1}{2}(\mathscr{E}_1 - \mathscr{E}_0)$, which (in the regime $d \lesssim \tilde{\Delta}$) is rather larger than $d/2$. [($\mathscr{E}_1 - \mathscr{E}_0$) is the $h = 0$ "spectral gap" between the first and second lines of Figs. 10(a) and 17(a)]. The fact that RBT's experiments [Fig. 10(b)] clearly show a threshold field $h_{0,1}$ significantly larger than $d/2$ shows that the actual spin-1 ground state which
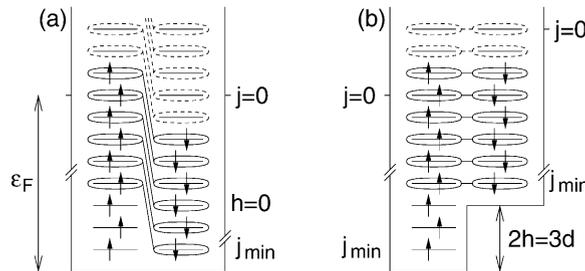
Fig. 18. Schematic representations of the nontime-reversed-pairing state $|3/2\rangle'$ defined in Eq. (73). The energies $\varepsilon_j \mp h$ of the single-particle states $|j, \pm \rangle$ are drawn (a) for $h = 0$ and (b) for $2h = 3d$. We indicated schematically how nontime-reversed states are paired according to $(u_i + v_i c_{(i+3)+}^{\dagger} c_{i-}^{\dagger})$ in the BCS-like ansatz (73), with solid or dashed ellipses encircling states that would be completely filled or empty in the absence of pairing correlations.

nature chooses is better approximated by $|1\rangle$ than by $|1\rangle'$, in spite of the fact that $\mathcal{E}_1' < \mathcal{E}_1$. Thus the premise of the above argument was wrong, and we can conclude that those terms in Eq. (72) not contained in the reduced BCS Hamiltonian can indeed be neglected, as done throughout in this review.

4.8. Measurable consequences of the blocking effect: parity effects

This section is devoted to various measurable manifestations of the blocking effect, in the form of parity effects, i.e. differences between a grain with an even or odd number of electrons.

4.8.1. Bulk consequences of blocking

The most obvious measurable manifestation of the blocking effect is the very existence of a spectral gap: "breaking a pair" and placing the two newly unpaired electrons in two singly occupied levels costs a significant amount of correlation energy, because the unpaired electrons loose pairing energy themselves and also disrupt the pairing correlations of the other pairs. This, of course, is already present in standard bulk mean-field BCS theory via the energy cost of at least $2\tilde{\Delta}$ involved in creating two quasiparticles, and is one of the hallmarks of superconductivity.

In the context of ultrasmall grains, let us denote the *pair-breaking energies* for an even (odd) grain, i.e. the minimum energy cost per electron for breaking a pair by flipping a single spin at $h = 0$, by Ω_{e} (Ω_{o}):

$$\Omega_{\mathrm{e}} \equiv \tfrac{1}{2}(\mathcal{E}_1 - \mathcal{E}_0)_{h=0}, \quad \Omega_{\mathrm{o}} \equiv \tfrac{1}{2}(\mathcal{E}_{3/2} - \mathcal{E}_{1/2})_{h=0} \ . \tag{75}$$

The even pair-breaking gap Ω_{e} is of course strikingly visible in RBT's $h = 0$ spectra as a large spectral gap for even grains (cf. Figs. 9 and 10; the latter gives $\Omega_{\mathrm{e}} = 0.26\,\mathrm{meV}$). Its presence is direct evidence for the existence of pairing correlations in the grain, which in that sense can still be called "superconducting".

In contrast, the odd pair-breaking gap Ω_{o} cannot be obtained from $h = 0$ spectra, since in an odd grain the lowest excitation does not involve breaking a pair, but simply exciting the unpaired electron, which does not require a correlation-induced gap to be overcome. To measure Ω_{o}, a finite field is needed: by Eq. (71), $\Omega_{\mathrm{e}} = h_{0,1}$ and $\Omega_{\mathrm{o}} = h_{1/2,3/2}$, hence both spin-flip gaps are equal to level-crossing fields that can be deduced from $h \neq 0$ data, as explained in Section 4.6. For Fig. 10 this yields $\Omega_{\mathrm{e}} = 0.23 \pm 0.01\,\mathrm{meV}$ and $\Omega_{\mathrm{o}} = 0.24 \pm 0.01\,\mathrm{meV}$ (a result further discussed in Section 4.8.4). The reason that the Ω_{e}-value determined in this way is somewhat *smaller* than the above-mentioned $0.26\,\mathrm{meV}$ determined at $h = 0$ is presumably that the experimental spectral lines are not perfectly linear in h (having a small h^2-contribution due to orbital diamagnetism, which should cause the spectroscopic gap to decrease faster with h than in our model).

Another consequence of the blocking effect is that the condensation energies $E_{p/2}^{\mathrm{cond}} = \mathcal{E}_{p/2} - \mathcal{E}_{p/2}^0$ for an even and odd grain differ: the unpaired electron of an odd grain weakens its pairing correlations relative to an even grain, so that $E_{1/2}^{\mathrm{cond}}$ is less negative than E_0^{cond}. In the bulk limit their difference approaches $E_{1/2}^{\mathrm{cond}} - E_0^{\mathrm{cond}} \to \tilde{\Delta}$, the energy of a single quasiparticle. For large mesoscopic islands (with $d/\tilde{\Delta} \ll 1$) this energy difference has indeed been directly observed: it causes a change from e- to $2e$-periodicity in the gate-voltage dependence of Coulomb oscillations [62,84–88]. For ultrasmall grains, however, ground state energy differences are currently not directly measurable, due to experimental difficulties explained in Section 2.4.6.

The parity effects discussed above survive in the bulk limit. Let us now turn to parity effects that result from even-odd differences in the *d-dependence* of various quantities.

4.8.2. Parity-dependent pairing parameters

As is evident from Figs. 14(a) and (b), not only the condensation energies $E_{p/2}$ are parity dependent; as soon as one leaves the bulk regime, the pairing parameters $\Delta_{p/2}$ become parity-dependent too, with $\Delta_0 > \Delta_{1/2}$. In the context of ultrasmall grains this was first emphasized by von Delft et al. [21], but it had been anticipated before by Janko et al. [109] and Golubev and Zaikin [112], who had studied the first correction to the bulk limit, finding $\Delta_0 - \Delta_{1/2} = d/2$ to leading order in $d/\tilde{\Delta}$; and this result, in turn, had already been published by Soloviev in the nuclear physics literature as long ago as 1961 [104].

The g.c. results of Fig. 14(a), in particular the fact that the critical level spacing $d_{p/2}^{\mathrm{BCS}}$ at which $\Delta_{p/2}$ vanishes is smaller for odd than even grains ($d_{1/2}^{\mathrm{BCS}} < d_0^{\mathrm{BCS}}$), suggest that "pairing correlations break down sooner in odd than even grains" [21]. However, it should be remembered that the vanishing of $\Delta_{p/2}$ signals the breakdown of the g.c. approach. A more accurate statement, that is borne out by the canonical calculations reviewed in Section 5, is that the inequality $\mathscr{E}_{1/2} > \mathscr{E}_0$ persists for arbitrarily large d (see Fig. 20 in Section 5.2), i.e. pairing correlations are always weaker for odd than even grains, although they never vanish altogether in either.

4.8.3. Matveev–Larkin parity parameter

To capture the difference between correlations in even and odd grains in terms of measurable quantities (which $\Delta_{p/2}$ are not), Matveev and Larkin [28] proposed the parameter (sometimes called "pairing energy" in nuclear physics [49])

$$\Delta_{\mathrm{P}}^{\mathrm{ML}} \equiv \mathscr{E}_{1/2}^{N+1} - \tfrac{1}{2}(\mathscr{E}_0^N + \mathscr{E}_0^{N+2}) \quad \text{(where } N \text{ is even)} , \tag{76}$$
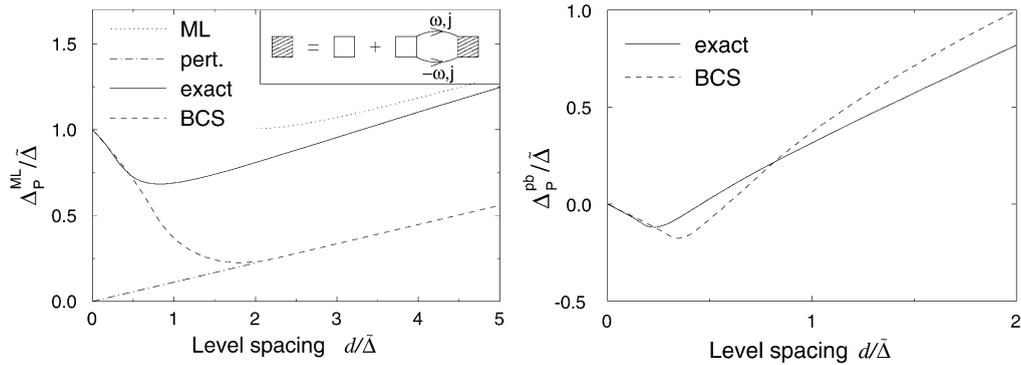
Fig. 19. The parity parameters (a) $\Delta_{\mathrm{P}}^{\mathrm{ML}}$ of Matveev–Larkin [Eq. (76)] and (b) $\Delta_{\mathrm{P}}^{\mathrm{pb}}$ for the pair-breaking energies [Eq. (81)], as functions of $d/\tilde{\Delta}$, calculated using the g.c. variational BCS approach of Section 4.4.1 (dashed lines), and Richardson's exact solution of Section 5.1 (solid lines). In (a), we also show the perturbative result for the uncorrelated Fermi sea, $(\Delta_{\mathrm{P}}^{\mathrm{ML}})_{\mathrm{pert}} = \tfrac{1}{2}\lambda d$ (straight dash–dotted line); and the renormalized result $(\Delta_{\mathrm{P}}^{\mathrm{ML}})_{\mathrm{ren}} \simeq d/[2\ln(ad/\tilde{\Delta})]$ of Eq. (80), in its range of validity $d/\tilde{\Delta} \gg 1$ (dotted line). The parameter $a = 1.35$ is chosen to ensure quantitative agreement with the exact result in the limit $d/\tilde{\Delta} \gg 1$. For a summary of the results of various other canonical calculations of $\Delta_{\mathrm{P}}^{\mathrm{ML}}$, see Fig. 20. The inset of (a) shows the Dyson equation used to calculate the renormalized coupling $\tilde{\lambda}$ in Eq. (78).

i.e. the difference between the ground state energy of an odd grain and the mean of the ground state energies of the even grains obtained by removing or adding one electron. Fig. 19(a) shows its behavior as function of $d/\tilde{\Delta}$. In the bulk limit we have $\mathscr{E}_0^N \simeq \mathscr{E}_0^{N+2}$ and $\Delta_{\mathrm{P}}^{\mathrm{ML}} \simeq \tilde{\Delta}$, which is simply the energy cost for having an unpaired electron on the odd grain. With increasing $d/\tilde{\Delta}$, this energy cost decreases since pairing correlations get weaker, hence $\Delta_{\mathrm{P}}^{\mathrm{ML}}$ initially decreases. It begins to increase again for $d \gtrsim \tilde{\Delta}$, since then pairing correlations are so weak that the behavior of $\Delta_{\mathrm{P}}^{\mathrm{ML}}$ is governed by the "self-energy" of the one extra pair in $\mathscr{E}_0^{N+2}$ relative to $\mathscr{E}_0^N$. For example, in the g.c. variational BCS result for $\Delta_{\mathrm{P}}^{\mathrm{ML}}$, namely

$$(\Delta_{\mathrm{P}}^{\mathrm{ML}})_{\mathrm{BCS}} = \mathscr{E}_{1/2}^{\mathrm{BCS}} - \mathscr{E}_0^{\mathrm{BCS}} + \lambda d/2 \; , \tag{77}$$

it is this self energy which produces the $\lambda d/2$ contribution.

A more careful calculation for the regime $d \gg \tilde{\Delta}$ was performed by Matveev and Larkin [28], who considered the renormalization of λ due to "pairing fluctuations" about the uncorrelated Fermi ground state $|p/2\rangle_0$. Summing up the leading logarithmic vertex corrections [113] [see inset of Fig. 19(a)], they obtained a renormalized coupling $\tilde{\lambda}$ given, with logarithmic accuracy, by

$$\tilde{\lambda} = \lambda + \lambda d \left[\sum_j^U \int \frac{\mathrm{d}\omega}{2\pi} \frac{1}{[\mathrm{i}\omega - (\varepsilon_j - \mu)]} \frac{1}{[-\mathrm{i}\omega - (\varepsilon_j - \mu)]} \right] \tilde{\lambda} \; , \tag{78}$$

$$\tilde{\lambda} = \frac{\lambda}{1 - \lambda d \sum_j^U \frac{1}{2|\varepsilon_j - \mu|}} \simeq \frac{\lambda}{1 - \lambda \log(\omega_{\mathrm{D}}/d)} \; . \tag{79}$$

This result evidently is valid only if $d \gg \omega_{\mathrm{D}} e^{-1/\lambda} \simeq \tilde{\Delta}/2$ (which, incidentally, is another way of seeing that $d \simeq \tilde{\Delta}$ defines the crossover between the fluctuation-dominated and bulk regimes). Matveev and Larkin concluded that

$$(\Delta_{\mathrm{P}}^{\mathrm{ML}})_{\mathrm{ren}} \simeq \tilde{\lambda} d/2 = d/(2 \log d/\tilde{\Delta}) \quad \text{for } d \gg \tilde{\Delta} \; . \tag{80}$$

This logarithmic renormalization is beyond the reach of the g.c. variational BCS method, but was confirmed using exact methods [33,91,39] (see Section 5.2). Its occurrence, in a regime that in g.c. variational calculations appears to be "uncorrelated", can be regarded as the "first sign of pairing correlations", in particular since, by Eq. (79), the interaction strength *increases* upon renormalization only if the interaction is attractive ($\lambda < 0$ would imply $|\tilde{\lambda}| < |\lambda|$). The pairing fluctuations responsible for this renormalization will be discussed in more detail in Sections 5.2 and 5.3.

Unfortunately, $\Delta_{\mathrm{P}}^{\mathrm{ML}}$ is at present not measurable in ultrasmall grains, for the same experimental reasons as apply to $\mathscr{E}_{1/2} - \mathscr{E}_0$, see Section 2.4.6.

4.8.4. Parity effect for pairbreaking energies

Braun and von Delft [35,23,91] discussed yet another parity effect, based on

$$\Delta_{\mathrm{P}}^{\mathrm{pb}} = \Omega_{\mathrm{o}} - \Omega_{\mathrm{e}} \; , \tag{81}$$

the difference between the *pair-breaking energies* of an even and on odd grain [see Eq. (75)]. Fig. 19(b) shows its behavior as function of d. In the bulk limit $\Omega_{\mathrm{e}} \simeq \Omega_{\mathrm{o}} \simeq \tilde{\Delta}$ and $\Delta_{\mathrm{P}}^{\mathrm{pb}} \simeq 0$. The most interesting feature of $\Delta_{\mathrm{P}}^{\mathrm{pb}}$ is that it initially becomes negative as $d/\tilde{\Delta}$ increases; this occurs because in an odd grain pairing correlations are weaker and hence breaking a pair costs less energy than in an even grain. $\Delta_{\mathrm{P}}^{\mathrm{pb}}$ becomes positive again for $d/\tilde{\Delta} \gtrsim 0.5$, since then pairing correlations are so weak

that $\Delta_{\mathrm{P}}^{\mathrm{pb}}$ is governed by the kinetic energy cost of flipping a spin, which is $2d$ for an odd grain but only d for an even grain.

$\Delta_{\mathrm{P}}^{\mathrm{pb}}$ is directly measurable in RBT's grains, via the level-crossing fields $h_{0,1} = \Omega_{\mathrm{e}}$ and $h_{1/2,3/2} = \Omega_{\mathrm{o}}$ [Eq. (71)]. The measured values $\Omega_{\mathrm{e}} = 0.23 \pm 0.01\,\mathrm{meV}$ and $\Omega_{\mathrm{o}} = 0.24 \pm 0.01\,\mathrm{meV}$ cited in Section 4.8.1 give a positive value of $\Delta_{\mathrm{P}}^{\mathrm{pb}} = 0.1\,\mathrm{meV}$, implying that the grain under study was too small to fall in the most interesting regime where $\Delta_{\mathrm{P}}^{\mathrm{pb}}$ is negative. Braun and von Delft suggested that the latter should be observable in a somewhat larger grain with $h_{1/2,3/2} < h_{0,1}$, i.e. in regime II of Fig. 16. (This suggestion assumes that despite the increased grain size, the complicating effect of orbital diamagnetism is still nondominant in regime II.) *To look for negative $\Delta_{\mathrm{P}}^{\mathrm{pb}}$ experimentally would thus require good control of the ratio $d/\tilde{\Delta}$, i.e. grain size.* This might be achievable if recently reported new fabrication methods, which allow systematic control of grain sizes by using colloidal chemistry techniques [114,186,187], could be applied to Al grains.

5. Superconductivity: crossover from the bulk to the limit of a few electrons

This section is devoted to the question: *How do pairing correlations change when the size of a superconductor is decreased from the bulk to the limit of only a few electrons?* In particular, we shall attempt to refine the answer given by Anderson [45], namely that superconductivity as we know it breaks down for $d \gtrsim \tilde{\Delta}$.

First steps towards a more detailed answer were taken in the early 1970s by Strongin et al. [93] and by Mühlschlegel et al. [94], who calculated the thermodynamic properties of ensembles small superconducting grains. Experimental realizations of such ensembles were, e.g., the granular films studied by Giaver and Zeller [80,81]. The interest of theorists was rekindled in 1995 by RBT's success in probing *individual* superconducting grains. Apart from motivating the phenomenological theory of Braun et al. reviewed in Section 4, these experiments also inspired a substantial and still growing number of theoretical studies [21–44] of how superconducting pairing correlations in such grains are affected by reducing the grains' size, or equivalently by increasing its mean level spacing $d \propto \mathrm{Vol}^{-1}$ until it exceeds the bulk gap $\tilde{\Delta}$.

In the earliest of these, von Delft et al. studied the discrete BCS model of Section 4.2 within a parity-projected g.c. BCS approach [21] closely related to the variational BCS method of Section 4.4. Their g.c. results suggested that pairing correlations, as measured by the pairing parameter or the condensation energy, vanish abruptly once d exceeds a critical level spacing $d_{p/2}^{\mathrm{BCS}}$ *that depends on the parity ($p = 0$ or 1) of the number of electrons on the grain*, being smaller for odd grains ($d_{1/2}^{\mathrm{BCS}} \simeq 0.89\tilde{\Delta}$) than even grains ($d_0^{\mathrm{BCS}} \simeq 3.6\tilde{\Delta}$). Parity effects were also found in a number of subsequent papers that used parity-projected g.c. methods to study the behavior of the BCS mean-field gap parameter Δ_{gc} and related quantities as functions of level spacing [22–32], temperature and magnetic field. All these parity effects are consequences of the blocking effect (cf. Section 4.2.3): for odd grains, the unpaired electron somewhat disrupts the pairing correlations of the remaining paired ones, by reducing the phase space available for pair scattering.

A series of more sophisticated canonical approaches [33–41] (summarized in Section 5.2) *confirmed the parity dependence of pairing correlations*, but established that the abrupt vanishing of pairing correlations at $d_{p/2}^{\mathrm{BCS}}$ is an artifact of g.c. treatments: *pairing correlations do persist, in the form of so-called fluctuations, to arbitrarily large level spacings* [28], and the crossover between the

bulk superconducting (SC) regime ($d \ll \tilde{\Delta}$) and the fluctuation-dominated (FD) regime ($d \gg \tilde{\Delta}$) is completely smooth [36–40]. Nevertheless, these two regimes are qualitatively very different [35–40]: the condensation energy, e.g., is an extensive function of volume in the former and almost intensive in the latter, and pairing correlations are quite strongly localized around the Fermi energy ε_F, or more spread out in energy, respectively. Very recently, Di Lorenzo et al. [44] suggested that the remnant pairing correlations in the FD regime might be detectable via susceptibility measurements.

Toward the end of 1998 and after the appearance of most of these works, R.W. Richardson pointed out [115] to their various authors that the discrete BCS Hamiltonian on which they are based actually has an exact solution, discovered by him in 1963 [46] (and independently by Gaudin in 1968 [116]). Richardson published his solution in the context of nuclear physics in a series of papers between 1963 and 1977 [46–54] which seem to have completely escaped the attention of the condensed matter community. Very recently, the model was also shown to be integrable [117,118]. The revival of this remarkably simple exact solution after such a long and undeserved period of neglect is perhaps one of the most important consequences of RBT's experimental breakthrough: Richardson's solution allows the elucidation and illustration by *exact* means of many important conceptual ingredients of the standard BCS theory of superconductivity, such as the nature of pairing correlations, the importance of phase coherence, the validity of using a mean-field approximation and a grand-canonical formulation for bulk systems, and the limitations of the latter approaches for ultrasmall systems. Moreover, it allows the exact calculation of essentially all quantities of interest for ultrasmall grains.

We shall therefore start this section by discussing the exact solution (Section 5.1). We then summarize the other canonical approaches somewhat more briefly than they perhaps would have deserved had an exact solution not existed, and compare their results to those of the exact solution (Section 5.2). Next, we analyze the qualitative differences between the bulk and FD regimes (Section 5.3), then discuss the case of randomly (as opposed to uniformly) spaced energy levels ε_j (Section 5.4), and finally discuss finite-temperature parity effects (Section 5.5). Throughout this section we set $\mu = 0$, since canonical treatments make no reference to a chemical potential.

5.1. Richardson's exact solution

In this section we summarize some of the central results of Richardson's exact solution of the discrete BCS model.

5.1.1. General eigenstates

Consider $N = 2n + b$ electrons, b of which are unpaired, as in Section 4.2.3. According to the general discussion there, the nontrivial aspect of solving the model is finding the eigenenergies $\mathscr{E}_n$ and corresponding eigenstates $|\Psi_n\rangle$ [Eq. (43)] of the pair Hamiltonian [Eq. (44), in which we set $\mu = 0$ below]

$$\hat{H}_U = \sum_{ij}^{U} (2\varepsilon_j \delta_{ij} - \lambda d) b_i^\dagger b_j \ , \tag{82}$$

in the Hilbert space of all states containing exactly n pairs $b_j^\dagger = c_{j+}^\dagger c_{j-}^\dagger$ of electrons, where j runs over the set of all unblocked single-particle levels, $U = I \backslash B$ (I is the set of all interacting levels, B the set of all blocked levels). In general, degenerate levels are allowed in I, but are to be distinguished by distinct j-labels, i.e. they have $\varepsilon_i = \varepsilon_j$ for $i \neq j$.

Richardson showed that the sought-after eigenstates (with normalization $\langle \Psi_n | \Psi_n \rangle = 1$) and eigenenergies have the general form

$$|\Psi_n\rangle = \mathscr{N} \prod_{v=1}^{n} B_v^\dagger |0\rangle, \quad \mathscr{E}_n = \sum_{v=1}^{n} E_v \quad \text{with } B_v^\dagger = \sum_j^U \frac{b_j^\dagger}{2\varepsilon_j - E_v}. \tag{83}$$

Here $\mathscr{N}$ is a normalization constant and the n parameters E_v ($v = 1, \ldots, n$) are a solution of the set of n coupled algebraic equations

$$\frac{1}{\lambda d} - \sum_j^U \frac{1}{2\varepsilon_j - E_v} + \sum_{\mu=1(\neq v)}^{n} \frac{2}{E_\mu - E_v} = 0 \quad \text{for } v = 1, \ldots, n, \tag{84}$$

which are to be solved (numerically, see Appendix B.2) subject to the restrictions $E_\mu \neq E_v$ if $\mu \neq v$. Richardson originally derived this remarkably simple result by solving the Schrödinger equation for the wavefunction $\psi(j_1, \ldots, j_n)$ of Eq. (41). A simpler proof, also due to Richardson [119], may be found in Refs. [40,120] and in Appendix B; its strategy is to verify that $(\hat{H}_U - \mathscr{E}_n)|\Psi_n\rangle = 0$ by simply commuting $\hat{H}_U$ past the $B_v^\dagger$ operators in (83).

Below we shall always assume the ε_j's to be all distinct (the more general case that degeneracies are present is discussed by Gaudin [116]). Then it can be shown explicitly [116] that (i) the number of distinct solutions of Eq. (84) is equal to the dimension of the n-pair Hilbert space defined on the set of unblocked levels U, namely $\binom{N_U}{n}$, where N_U is the number of unblocked levels; and (ii) that the corresponding eigenstates (83) are mutually orthogonal to each other, thus forming an eigenbasis for this Hilbert space. This can easily be understood intuitively, since there exists a simple relation between the bare pair energies $2\varepsilon_j$ and the solutions of Eqs. (84): as λ is reduced to 0, it follows by inspection that each solution $\{E_1, \ldots, E_n\}$ reduces smoothly to a certain set of n bare pair energies, say $\{2\varepsilon_{j_1}, \ldots, 2\varepsilon_{j_n}\}$; this particular solution may thus be labeled by the indices $j_1, \ldots, j_n$, and the corresponding eigenstate (83) written as $|\Psi_n\rangle \equiv |j_1, \ldots, j_n\rangle_U$. By inspection, its $\lambda \to 0$ limit is the state $|j_1, \ldots j_n\rangle_{U,0} \equiv \prod_{v=1}^{n} b_{j_v}^\dagger |0\rangle$, thus there is a one-to-one correspondence between the sets of all states $\{|j_1, \ldots, j_n\rangle_U\}$ and $\{|j_1, \ldots, j_n\rangle_{U,0}\}$. But the latter constitute a complete eigenbasis for the n-pair Hilbert space defined on the set of unblocked levels U, thus the former do too.

5.1.2. Ground state

For a given set of blocked levels B, the lowest-lying of all states $|\Psi_n, B\rangle$ of the form (40), say $|\Psi_n, B\rangle_G$, is obtained by using that particular solution $|j_1, \ldots j_n\rangle_U$ for which the total "pair energy" $\mathscr{E}_n$ takes its lowest possible value. The lowest lying of all eigenstates with n pairs, b blocked levels and total spin $s = b/2$, say $|n, s\rangle_G$ with energy $\mathscr{E}_s^G(n)$, is that $|\Psi_n, B\rangle_G$ for which the blocked levels in B all contain spin-up electrons and are all as close as possible to ε_F, the Fermi energy of the uncorrelated N-electron Fermi sea $|F_N\rangle$. The E_v for the ground state $|n, s\rangle_G$ coincide at $\lambda = 0$ with the lowest n energies $2\varepsilon_v$ ($v = 1, \ldots, n$), and smoothly evolve toward (initially) lower values when λ is turned on, a fact that can be exploited during the numerical solution of Eq. (84). As λ is increased further, some of the E_v's become complex; however, they always occur in complex

conjugate pairs, so that $\mathscr{E}_n$ remains real [51]. For details, see Ref. [51] and our Appendix B.2, where some algebraic transformations are introduced that render the equations less singular and hence simplify their numerical solution considerably.

5.1.3. General comments

Since the exact solution provides us with wavefunctions, it is, in principle, straightforward to calculate arbitrary correlation functions of the form $\langle \Psi_n | b_i^\dagger b_j^\dagger \cdots b_{i'} b_{j'} | \Psi_n \rangle$, by simply commuting all b's to the right of all $b^\dagger$'s. However, due to the hard-core boson commutation relations (46) of the b's, the combinatorics is rather involved. Nevertheless, Richardson succeeded to derive [50] explicit results for the normalization constant $\mathscr{N}$ of (83) and the occupation probabilities $\bar{v}_j^2$ and correlators C_{ij} of Eq. (49), which we summarize in Appendix B.3. The exact result for the C_{ij}'s show that they are *all* positive, in agreement with requirement (ii) formulated in Section 4.3.2. It is also natural to ask whether in the bulk limit ($d \to 0$ at fixed nd), the standard BCS results can be extracted from the exact solution. Indeed they can, as Richardson showed in [54] (following unpublished work by Gaudin [116]), by interpreting the problem of solving the eigenvalue equations (84) for the E_ν as a problem in two-dimensional electrostatics (see Appendix B.2). Exploiting this analogy, he showed that in the bulk limit, Eqs. (84) reduce to the well-known BCS equations determining the gap and chemical potential at $T = 0$ [Eqs. (68) and (69)], and the ground state condensation energy $\mathscr{E}_0^{\text{cond}}(n)$ [Eq. (45)] to its BCS result, namely $-\tilde{\Delta}^2/2d$.

Finally, let us mention that Cambiaggio, Rivas and Saraceno have recently shown that the discrete BCS model is integrable and have constructed explicit expressions for all its constants of the motion [117]. The latter's relation to Richardson's solution was clarified by Sierra [118], who has also explored possible connections between the exact solution and conformal field theory. It would be an interesting challenge for mathematical physicists to try to exploit this integrability to calculate finite-temperature properties exactly – although these can, in principle, be obtained from Richardson's solution by "simply" computing the partition function over all states, this is forbiddingly tedious in practice for large temperatures, since the eigenenergy of *each* state requires a separate (nontrivial) numerical calculation.

5.2. Comparison of other canonical methods with the exact solution

In this section we briefly mention the various canonical methods by which the discrete BCS model had been investigated prior to the revival of Richardson's exact solution in 1999. All of these studies used a half-filled band with fixed width $2\omega_D$ of $N = 2n + p$ uniformly spaced levels [i.e. $\varepsilon_j = jd + (1 - p)d/2$, with $p = 0$ or 1, as in Eq. (38)], containing N electrons. Then the level spacing is $d = 2\omega_D/N$ and the bulk gap is $\tilde{\Delta} = \omega_D/\sinh(1/\lambda)$. Following [35], we take $\lambda = 0.224$ throughout this section. To judge the quality of the various approaches, we compare in Fig. 20 the results which they yield with those from Richardson's solution, for the even and odd ($s = 0, 1/2$) condensation energies E_s^{cond} and the Matveev–Larkin parity parameter Δ_P^{ML} (cf. Section 4.8.3). In the notation of Section 5.1.2, these are given by

$$E_s^{\text{cond}}(n) = \mathscr{E}_s^{\text{G}}(n) - \langle F_N | \hat{H} | F_N \rangle , \tag{85}$$

$$\Delta_P^{\text{ML}}(n) = \mathscr{E}_{1/2}^{\text{G}}(n) - [\mathscr{E}_0^{\text{G}}(n) + \mathscr{E}_0^{\text{G}}(n + 1)]/2 . \tag{86}$$
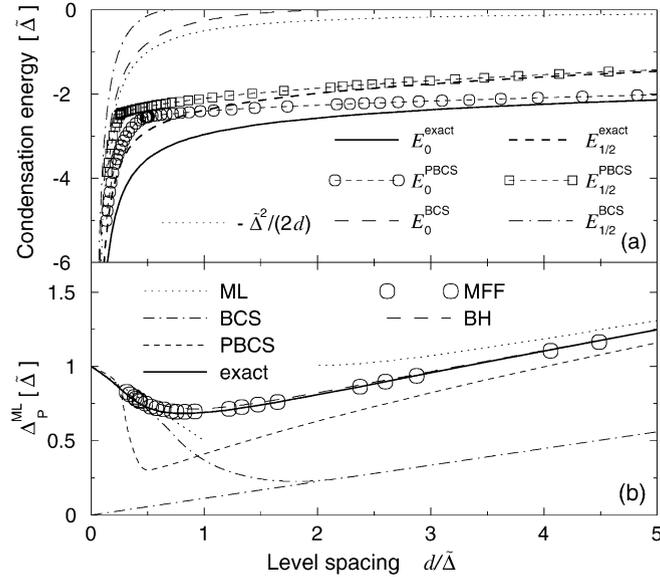
Fig. 20. (a) The even and odd ($s = 0, 1/2$) condensation energies E_s^{cond} of Eq. (85) [in units of $\tilde{\Delta}$], calculated with BCS, PBCS and exact wave functions [39], as functions of $d/\tilde{\Delta} = 2\sinh(1/\lambda)/(2n + 2s)$, for $\lambda = 0.224$. For comparison, the dotted line gives the "bulk" result $E_0^{bulk} = -\tilde{\Delta}^2/(2d)$. (b) Comparison [39] of the parity parameters Δ_P^{ML} [28] of Eq. (86) [in units of $\tilde{\Delta}$] obtained by various authors: ML's analytical result (dotted lines) [$\tilde{\Delta}(1 - d/2\tilde{\Delta})$ for $d \ll \tilde{\Delta}$, and $d/2\log(ad/\tilde{\Delta})$ for $d \gg \tilde{\Delta}$, with $a = 1.35$ adjusted to give asymptotic agreement with the exact result]; grand-canonical BCS approach (dash–dotted line) [the naive perturbative result $\frac{1}{2}\lambda d$ is continued to the origin]; PCBS approach (short-dashed line); Richardson's exact solution (solid line); exact diagonalization and scaling by MFF (open circles) and BH (long-dashed line).

Following the initial g.c. studies [21–24,28] of the discrete BCS model, the first purely canonical study was that of Mastellone, Falci and Fazio (MFF) [33], who used Lanczos exact diagonalization. Despite being limited to $n \leq 12$, they managed to reach reasonably small ratios of $d/\tilde{\Delta}$ by using an ingenious scaling approach: for a given level spacing d, they increased the coupling constant λ to about 0.5, thereby decreasing $d/\tilde{\Delta} = d/\omega_D \sinh(1/\lambda)$ to values as small as 0.5. This allowed them to probe, coming from the few-electron side, a remarkably large part of the crossover to bulk limit. They found, i.a., that the condensation energies are negative for *all* d, showing that the system can *always* gain energy by allowing pairing correlations, even for arbitrarily large d.

Berger and Halperin (BH) [34] showed that almost identical results can be achieved with less than 6 pairs, thus significantly reducing the calculational effort involved, by first performing a "poor man's scaling" renormalization: they reduce the bandwidth from $\omega_D \approx nd$ to, say, $\bar{\omega}_D \approx \bar{n}d$ (with $\bar{n} \leq 6$) and incorporate the effect of the removed levels by using a renormalized coupling constant

$$\bar{\lambda} = \lambda \left[1 - \sum_{\bar{\omega}_D < |\varepsilon_j| < \omega_D} \frac{\lambda}{2|\varepsilon_j|} \right]^{-1}. \tag{87}$$

The reduced system is then diagonalized exactly. Note that the renormalization of Matveev and Larkin [Eq. (79)] corresponds to taking $\bar{\omega}_D \simeq d$ in Eq. (87), i.e. to integrating out the entire band.

Also note that the renormalization prescription of (87) has the property that it would leave the bulk gap invariant in the limit $d/\tilde{\Delta} \to 0$, for which Eq. (87) would imply $\bar{\omega}_D e^{-1/\tilde{\lambda}} \simeq \omega_D e^{-1/\lambda} \simeq \tilde{\Delta}$.

To access larger values of n, Braun and von Delft [35] used a *fixed-n projected* BCS approach (PBCS), in which BCS-like variational wavefunctions are projected to fixed particle number, as in Eq. (60). The projection integrals occurring in Eq. (60) were evaluated numerically for $n \le 600$, using tricks developed in the nuclear physics literature by Bayman [121], Dietrich et al. [122] and Ma and Rasmussen [123], and summarized in part in the book of Ring and Schuck [106]. (A much simpler way of dealing with the projection, using recursion relations, was recently found by Dukelsky and Sierra [37].) The PBCS method gives condensation energies that (i) are significantly lower than the grand-canonical ones (see Fig. 20), thus the projection much improves the variational ansatz, and that (ii) are negative for *all d*, confirming that the abrupt vanishing of the g.c. condensation energies is indeed an artifact of the g.c. treatment. The PBCS method is able to fully recover the bulk limit, but the crossover is not completely smooth and shows a remnant of the g.c. breakdown of pairing correlations: the d-dependence of the condensation energy $(E_s^{\text{cond}})^{\text{PBCS}}$ changes rather abruptly [kinks in the short-dashed lines in Fig. 20(a)] from being *extensive* ($\sim 1/d$) to being practically *intensive* (almost d independent).

It should be mentioned here that a generalization of the PBCS method to finite temperatures has been worked out by Essebag and Egido in the context of nuclear physics [124]. The PBCS method has recently also been applied to the attractive Hubbard model in one dimension by Tanaka and Marsiglio [43], who found even–odd and super-even effects. The latter consist of differences between the number of pairs being equal to $n = 2m$ or $2m + 1$, and arise if boundary conditions are used that produce doubly degenerate levels ($\varepsilon_{\bar{k}} = \varepsilon_{-\bar{k}}$) near the Fermi surface.

Dukelsky and Sierra [36,37] used the density matrix renormalization group (DMRG) (with $n \le 400$) to achieve significant improvements over the PBCS results for the discrete BCS model, in particular in the regime of the crossover, which they found to be *completely smooth*. In general, the DMRG approach is applicable to systems that can be divided into two pieces, called block and environment, which interact via a preferably rather small number of states. One starts with a small block and environment, computes their combined density matrix, then enlarges both and recomputes the density matrix, etc., until a large part of the system has been treated. Dukelsky and Sierra chose the block and environment to consist, respectively, of all particle or hole states relative to the Fermi sea (for a detailed description of the method, see [37]). Since the pairing correlations involving coherent superpositions of particle and hole states are peaked in a rather small regime of width $\tilde{\Delta}$ around the Fermi energy (cf. Figs. 12 or 22), the "interaction" between block and environment is "localized", so that the DMRG can a priori be expected to work rather well for this problem.

Finally, Dukelsky and Schuck [41] showed that a self-consistent RPA approach, which, in principle, can be extended to finite temperatures, describes the FD regime rather well (though not as well as the DMRG).

To check the quality of the above methods, Braun [91,39] computed E_s^{cond} (for $s = 0, 1/2$) and Δ_P^{ML} using Richardson's solution (Fig. 20). The exact results

(i) quantitatively agree, for $d \to 0$, with the leading $-\tilde{\Delta}^2/2d$ behavior for E_s^{cond} obtained in the g.c. BCS approach [21,35,23], which in this sense is exact in the bulk limit, corrections being of order d^0;

(ii) confirm that the even ground state energy *always* lies below the odd one (this had independently been proven rigorously by Tian and Tang [42]);

(iii) confirm that a completely smooth [36,37] crossover occurs around the scale $d \simeq \tilde{\Delta}$ at which the g.c. BCS approach breaks down;

(iv) show that the PBCS crossover [35] is qualitatively correct, but not quantitatively, being somewhat too abrupt;

(v) are reproduced remarkably well by the approaches of MFF [33] and BH [34];

(vi) are fully reproduced by the DMRG of [36,37] with a relative error of $< 10^{-4}$ for $n \leq 400$; our figures do not show DMRG curves, since they are indistinghuishable from the exact ones and are discussed in detail in [36,37].

The main conclusion we can draw from these comparisons is that the two approaches based on renormalization group ideas work very well: the DMRG is essentially exact for this model, but the bandwidth rescaling method of BH also gives remarkably (though not quite as) good results with rather less effort. In contrast, the PBCS approach is rather unreliable in the crossover region. To study generalizations of the discrete BCS model, e.g. using state-dependent couplings of the form $d \sum_{ij} \lambda_{ij} b_i^\dagger b_j$, the DMRG would thus be the method of choice.

5.3. Qualitative differences between the bulk and the few-electron regimes

Does the fact that the exact condensation energy E_s^{cond} is always negative, even for arbitrarily large $d/\tilde{\Delta}$, mean that the system stays "superconducting" even if it is arbitrarily small? The answer is certainly no, since in the fluctuation-dominated (FD) regime, the pairing correlations are qualitatively different than in the bulk, superconducting regime. In this section we shall try to make this statement more precise by analyzing the qualitative differences between the two regimes, with regard to the λ- and d-dependence of E_s^{cond}, and the behavior of the occupation probabilities $\bar{v}_j^2$.

Fig. 21(a) shows, on a log–log plot, the d-dependence of the even condensation energy $E_0^{\text{cond}}(d)$. Note that even on the log–log plot, the crossover of the exact E_0^{cond} from the bulk to the FD regime is completely smooth. According to Sierra and Dukelsky [37], the exact result for $E_0^{\text{cond}}(d)$ can be fitted very well to the form

$$E_0^{\text{cond}}(d) = -\tilde{\Delta}^2/(2d) - \eta_0(\ln 2)\omega_{\text{D}}\lambda^2 + \gamma_0(\tilde{\Delta}d/2\omega_{\text{D}})\log(2\omega_{\text{D}}/d) , \tag{88}$$
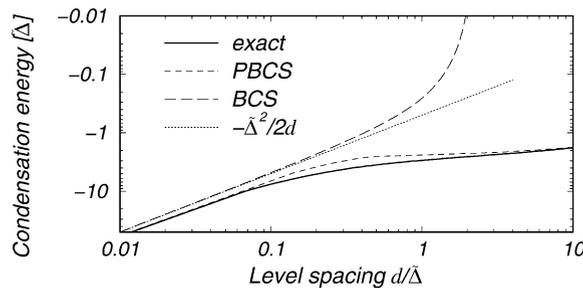

Fig. 21. Log–log plot [91] of some of the curves of Fig. 20(a) for the even condensation energy E_0^{cond} [in units of $\tilde{\Delta}$], for $\lambda = 0.224$; its asymptotic $-\tilde{\Delta}^2/(2d)$ behavior for $d/\tilde{\Delta} \to 0$ is shown by the dotted line.

where η_0 and γ_0 are constants of order unity [37]. The first term is *extensive* ($\propto$ Vol) and dominates in the bulk limit; its standard heuristic interpretation [92] is that roughly $\tilde{\Delta}/d$ levels (those within $\tilde{\Delta}$ of ε_F) are strongly affected by pairing, with an average energy gain per level of $-\tilde{\Delta}/2$. The second term, which is *intensive* and dominates in the FD limit, is equal (up to the numerical factor η_0) to the result from second-order perturbation theory [37], namely $(\lambda d)^2 \sum_{i=1}^{n} \sum_{j=n+1}^{2n} (2\varepsilon_i - 2\varepsilon_j)^{-1}$. This subleading term's *d*-independence (which was anticipated in [45,108]) may be interpreted by arguing that in the FD regime, the number of levels that contribute significantly to E_0^{cond} is no longer of order $\tilde{\Delta}/d$: instead, fluctuations affect *all* $n \simeq 2\omega_D/d$ unblocked levels within ω_D of ε_F (this is made more precise below), and each of these levels contributes an amount of order $-(\lambda d)^2/d$ (corresponding, in a way, to its selfenergy). Finally, the third term contains the small parameter $\tilde{\Delta}/\omega_D$ and thus represents a very small correction.

The λ- and volume dependencies of E_0^{cond} in Eq. (88) strikingly illustrate the qualitative differences between the bulk and FD regimes: in the bulk regime, dominated by the first term, E_0^{cond} is nonperturbative in λ (since $\tilde{\Delta} \simeq 2\omega_D e^{-1/\lambda}$) and extensive, as expected for a strongly correlated state; in constrast, in the FD regime, dominated by the second term, E_0^{cond} is perturbative in λ and practically intensive (up to the weak $\log d$ dependence of the third term).

Perhaps the most vivid way of illustrating the qualitative difference between the bulk and FD regimes is to study properties of the ground state wavefunction. We shall consider here the correlators [35]

$$\bar{C}_j^2(d) = \langle b_j^\dagger b_j \rangle \langle b_j b_j^\dagger \rangle \,, \tag{89}$$

which measure the probability that a level can be "both occupied and empty", and vanish identically for states without pairing correlations. For the discrete BCS model $\bar{C}_j^2$ identically equals $\langle b_j^\dagger b_j \rangle - \langle b_j^\dagger b_j \rangle^2 = \bar{v}_j^2 - \bar{v}_j^4$ [by Eqs. (46) and (49)], which measures the fluctuations in the pair occupation number of level j, and it vanishes for any blocked single-particle level. Note that $\bar{C}_j^2$ also equals $\langle b_j^\dagger b_j \rangle - \langle c_{j+}^\dagger c_{j+} \rangle \langle c_{j-}^\dagger c_{j-} \rangle$; this form, which was used in [35] and corresponds to the diagonal terms under the sum in Eq. (52) for Δ_{can}, can be interpreted as the probability enhancement for finding a *pair* of electrons instead of two uncorrelated electrons in a single-particle level $|j, \pm \rangle$.

When evaluated using the grand-canonical BCS wavefunction, $(\bar{C}_j^2)_{\text{BCS}}$ is equal to $u_j^2 v_j^2 = \frac{1}{4}\tilde{\Delta}^2/(\varepsilon_j^2 + \tilde{\Delta}^2)$ [thick solid lines in Fig. 22, the same function as that plotted in Fig. 12]. The $(\bar{C}_j)_{\text{BCS}}$'s thus have a characteristic peak of width $\propto \tilde{\Delta}$ around ε_F, implying that pairing correlations are "localized around ε_F in energy space", which may be taken to be the defining property of "BCS-like correlations". Moreover, in the bulk regime $d \ll \tilde{\Delta}$, and $(\bar{C}_j)_{\text{BCS}}$ are virtually identical to $(\bar{C}_j)_{\text{exact}}$ [long-dashed line of Fig. 22(a)], *vividly illustrating why the grand-canonical BCS approximation is so successful: not performing a canonical projection hardly affects the parameters $\bar{u}_j$ and $\bar{v}_j$ if $d \ll \tilde{\Delta}$, but tremendously simplifies their calculation.*

As one enters the FD regime $d \gtrsim \tilde{\Delta}$, the character of the correlator $(\bar{C}_j)_{\text{exact}}$ changes [Fig. 22(b), circles]: weight is shifted into the tails far from ε_F at the expense of the vicinity of the Fermi energy. Thus, *pairing correlations become delocalized in energy space* (as also found in [33,36,37]), so that referring to them as mere "fluctuations" is quite appropriate. In the extreme case $d \gg \tilde{\Delta}$, the $(\bar{C}_j)_{\text{exact}}$ for all interacting levels are roughly equal.

Richardson's solution can also be used to calculate, for a given set B of blocked levels, the *d*-dependence of the canonical order parameter $\Delta_{\text{can}}^B(d)$ of Eq. (52). Schechter has found [125] that it
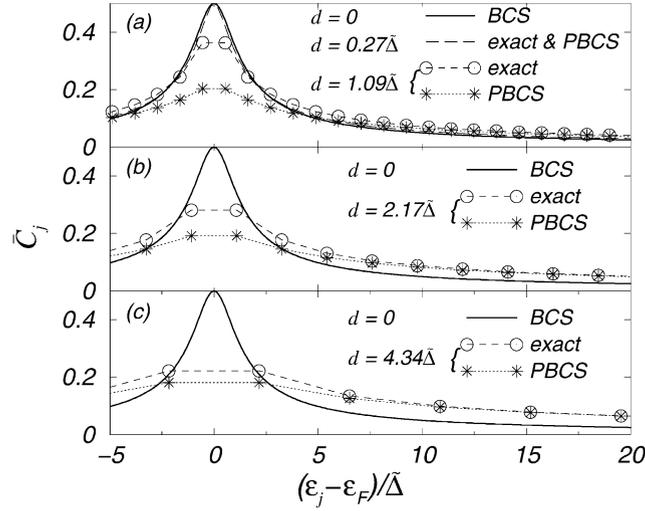
Fig. 22. The occupation probabilities $\bar{C}_j$ of Eq. (89) for $d/\tilde{\Delta} = 0$, 0.27, 1.09, 2.17 and 4.34 [91]. In all three figures, the thick solid lines give the $d = 0$ bulk BCS result, whereas circles and stars represent $\bar{C}_j$-values evaluated for discrete j's using the exact solution and PBCS method, respectively. For $d = 0.27\tilde{\Delta}$, the PBCS and exact results are indistinguishable, and are shown in (a) as a single long-dashed line, which is also virtually identical to the bulk curve. For small d, pairing correlations are evidently localized within a few $\tilde{\Delta}$ of ε_F. With increasing d more and more weight is shifted away from ε_F into the tails; compared to the exact results, the PBCS method somewhat overemphasizes this delocalization, which is one of the reasons why it produces a somewhat too abrupt crossover.

can be fit to the form $\Delta_{\text{can}}^B(d) = \tilde{\Delta}(1 + \tilde{\gamma}_B d/\tilde{\Delta})$, where $\tilde{\gamma}_B$ is a positive numerical constant, and the linear term essentially reflects the factor of d in the definition of Δ_{can}^B. The fact that Δ_{can}^B is a strictly increasing function of d is in very striking contrast to the behavior of the grand-canonical pairing parameters $\Delta_s(d)$ shown in Fig. 14(a).

5.4. Effect of level statistics

Smith and Ambegaokar investigated the effect of level statistics on the crossover between the bulk and FD regimes [24]. In contrast to the uniform level spacing used in previous works, they employed a random spacing of levels, distributed according to the Gaussian orthogonal ensemble. Using a g.c. mean-field BCS approach, they found, interestingly, that *randomness enhances pairing correlations*: compared to uniform spacings (u.s.) [21], it (i) on average *lowers* the condensation energy E_s^{cond} to more negative values, $\langle E_s^{\text{cond}} \rangle < E_s^{\text{cond}}(\text{u.s.})$, but (ii) these still are parity dependent, $\langle E_0^{\text{cond}} \rangle < \langle E_{1/2}^{\text{cond}} \rangle$. These results can readily be understood intuitively: pairing correlations become stronger the higher the density of levels around ε_F, where pair-mixing costs the least energy. When determining the amount of pairing correlations for a set of randomly spaced levels, fluctuations that increase the level density near ε_F are thus weighted more than those that decrease it, so that randomness enhances pairing correlations.

Although the g.c. mean-field treatment of Smith and Ambegaokar breaks down for mean level spacings much larger than $\tilde{\Delta}$, just as was the case in [21–23], their main conclusions (i) and (ii) are robust. Indeed, these were recently confirmed by Sierra et al. [39], who used Richardson's exact
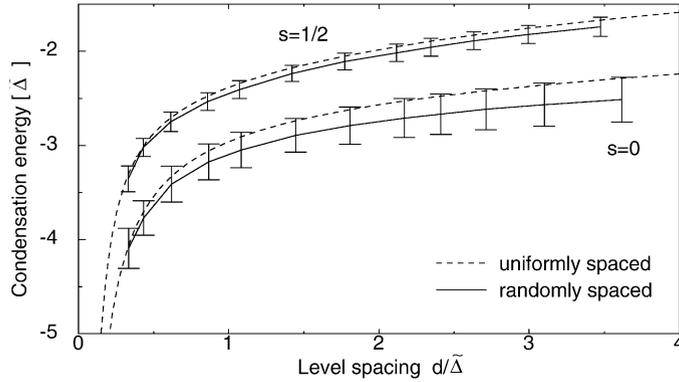
Fig. 23. Exact even and odd condensation energies E_s^{cond} [in units of $\tilde{\Delta}$] for equally spaced levels (dashed line), and the ensemble-average $\langle E_s^{\mathrm{cond}}\rangle$ for randomly spaced levels (solid line) [39]. The height of the fluctuation bars gives the variances $\delta E_s^{\mathrm{cond}}$.

solution to calculate E_s^{cond} for ensembles of random levels (Fig. 23). Moreover, they found that the blocking effect responsible for (ii) manifests itself in the fluctuations too, which likewise are parity dependent: for example, Fig. 23 shows that both the variances $\delta E_s^{\mathrm{cond}} \equiv [\langle (E_s^{\mathrm{cond}})^2\rangle - \langle E_s^{\mathrm{cond}}\rangle^2]^{1/2}$ and the randomness-induced changes in condensation energies $|\langle E_s^{\mathrm{cond}}\rangle - E_s^{\mathrm{cond}}(\mathrm{u.c.})|$ were larger for even than for odd grains.

5.5. Finite-temperature parity effects

Although finite-temperature studies of the discrete BCS model are not of direct relevance for spectroscopic measurements of the BRT-type (a finite T would simply smear out the discrete spectra, thereby blurring their most interesting features), they are important in their own right for extending our understanding of superconductivity in ultrasmall grains. We hence review several recent finite-T developments below.

To begin, let us note that parity effects are of course not restricted to the $T = 0$ limit discussed so far. To be observable [84–88], they only require the temperature to be smaller than the free energy difference $\delta\mathscr{F} \simeq \tilde{\Delta} - k_{\mathrm{B}} T \ln[N_{\mathrm{eff}}(T)]$ between an odd and even grain. Here $N_{\mathrm{eff}}(T)$ is the effective number of states available for quasiparticle excitations at temperature T, and for $d \ll \tilde{\Delta}$ is given by $N_{\mathrm{eff}}(T) = \sqrt{8\pi T\tilde{\Delta}/d^2}$ [84]. Below the corresponding crossover temperature where $\delta\mathscr{F} = 0$, determined by $k_{\mathrm{B}} T_{\mathrm{cr}}^* = \tilde{\Delta}/\ln[N_{\mathrm{eff}}(T_{\mathrm{cr}}^*)]$ and roughly equal to $\tilde{\Delta}/\ln\sqrt{8\pi\tilde{\Delta}^2/d^2}$, the single unpaired electron begins to matter: it causes a crossover from e-periodicity to $2e$-periodicity in the I–V characteristics of mesoscopic superconducing SET's [84–88], due to the ground state energy difference $\mathscr{E}_{1/2} - \mathscr{E}_0 \simeq \tilde{\Delta}$. Since T_{cr}^* becomes of order $\tilde{\Delta}$ in nanoscopic grains with $d \simeq \tilde{\Delta}$, parity effects should survive to temperatures as high as the (bulk) superconducting transition temperature T_{c} itself.

Regrettably, the canonical methods discussed in the preceding sections become impractical at finite temperatures, since the number of states that need to be considered increases rapidly for $T \gtrsim d, \tilde{\Delta}$. On the other hand, g.c. finite-T methods, some of which we review below, are, in principle,

inherently unreliable for $d \gtrsim \tilde{\Delta}$. This applies in particular to the simplest of these, parity-projected mean-field theory [109,112] (Section 5.5.1) and certain variational generalizations thereof [25,26] (Section 5.5.2): they yield the same sharp phase transition as function of temperature for finite systems as for bulk systems, whereas on general ground no sharp transitions are possible in finite systems. The reason for this problem is that they neglect fluctuations in the order parameter, which become very important in the transition region. The sharp transition is smoothed out once fluctuations are included. A rather efficient way of doing this is the so-called static path approximation (Section 5.5.3). Its use is illustrated in Section 5.5.4 for a calculation of the spin susceptibility, which shows an interesting parity effect that should be measurable in ensembles of ultrasmall grains.

5.5.1. Parity-projected mean-field theory

The simplest finite-T approach that is able to keep track of parity effects is parity-projected mean-field theory, first used in nuclear physics by Tanabe et al. [126], and, independently, introduced to the condensed-matter community by Jankó et al. [109] and Golubev and Zaikin [112]. One projects the g.c. partition function exactly onto a subspace of Fock space containing only even or odd ($p = 0, 1$) numbers of particles, using the parity-projector $\hat{P}_p$:

$$Z_p^{\mathrm{gc}} \equiv \mathrm{Tr}^{\mathrm{gc}} \hat{P}_p \mathrm{e}^{-\beta(\hat{H} - \mu_p \hat{N})}, \quad \hat{P}_{0,1} \equiv \tfrac{1}{2}[1 \pm (-1)^{\hat{N}}] \,. \tag{90}$$

One then makes the mean-field replacement $b_j \to \{b_j - \langle b_j \rangle_p\} + \langle b_j \rangle_p$, neglects terms quadratic in the fluctuations represented by $\{\ \}$, and diagonalizes $\hat{H}$ in terms of the Bogoljubov quasiparticle operators $\gamma_{j\sigma}$ of Eq. (58). The self-consistency condition $\Delta_{p/2} \equiv \lambda d \sum_j \langle b_j \rangle_p$, evaluated in a parity-projected g.c. ensemble according to Eq. (90), leads to a gap equation of the standard form

$$\frac{1}{\lambda} = d \sum_{|\varepsilon_j| < \omega_{\mathrm{D}}} \frac{1}{2E_j} \left(1 - \sum_\sigma f_{pj\sigma} \right), \quad E_j \equiv \sqrt{(\varepsilon_j - \mu)^2 + |\Delta_{p/2}|^2} \tag{91}$$

which is parity-dependent, via the occupation function $f_{pj\sigma} = \langle \gamma_{j\sigma}^\dagger \gamma_{j\sigma} \rangle_p$ for quasiparticles. Since their number parity is restricted to be p, $f_{pj\sigma}$ differs from the usual Fermi function $f_{j\sigma}^0$. The condition $2n + p = \langle \hat{N} \rangle_p$ fixes the chemical potential μ to lie exactly half-way between the last filled and first empty levels if $p = 0$, and exactly on the singly occupied level if $p = 1$, implying $\mu = 0$ in both cases [by Eq. (38)].

von Delft et al. [21] applied this approach to the discrete BCS model with uniformly spaced levels, and solved Eq. (91) for the parity parameter $\Delta_{p/2}(d, T)$ as function of both level spacing and temperature. Fig. 24 summarizes their results. At zero temperature, $\Delta_{p/2}(d, 0)$ corresponds to the spin-dependent parity parameters $\Delta_{s=p/2}$ discussed in Section 4.4.2 [cf. Fig. 14(a)], and drops to zero at a critical level spacing d_s^{BCS}. The $\Delta_{p/2} \to 0$ limit of Eq. (91) defines the parity-dependent "critical temperature" $T_{\mathrm{c},p}(d)$, which can be viewed as another measure of how rapidly pair-mixing correlations break down as function of level spacing (although ultrasmall grains, of course, cannot undergo a sharp thermodynamic phase transition, which can only occur if $n \to \infty$). In both the even and odd cases, the behavior of $T_{\mathrm{c},P}(d)$ shows direct traces of the parity projection:

In the even case, $T_{\mathrm{c},0}(d)$ (Fig. 24, curve D) is nonmonotonic as function of increasing d, initially increasing slightly before dropping to zero very rapidly as $d \to d_0^{\mathrm{BCS}}$. The intuitive reasons for the initial increase is that the difference between the actual and usual quasiparticle occupation functions is $f_{pj\sigma} - f_{j\sigma}^0 < 0$ for an even grain (becoming significant when $d \simeq \tilde{\Delta}$), reflecting the fact
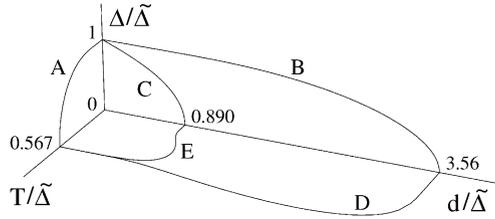
Fig. 24. d- and T-dependence of the pairing parameter $\Delta_{p/2}(d,T)$, calculated using parity-projected mean-field theory [21]. Curve A gives the bulk gap $\Delta(0,T)$, with $\Delta(0,0) \equiv \tilde{\Delta}$; curves B–E give $\Delta_{p/2}(d,T)/\tilde{\Delta}$ as a function of $d/\tilde{\Delta}$ and $T/\tilde{\Delta}$ for $p = 0$ (B, D) and $p = 1$ (C, E). The critical spacings $d_0^{\mathrm{BCS}} = 3.56\tilde{\Delta}$ and $d_{1/2}^{\mathrm{BCS}} = 0.890\tilde{\Delta}$ given here differ somewhat from those in Fig. 14(a), because the present mean-field approach differs in minor details (via terms that vanish when $d \to 0$) from the variational approach of Section 4.4.2.

that exciting quasiparticles two at a time is more difficult than one at a time. Therefore, the quasiparticle-induced weakening of pairing correlations with increasing T will set in at slightly higher T if $d \simeq \tilde{\Delta}$.

In the odd case, the critical level spacing $d_{1/2}^{\mathrm{BCS}}(T)$ (Fig. 24, curve E) is nonmonotonic as a function of increasing T, first increasing to a maximum before beginning to decrease toward $d_{1/2}^{\mathrm{BCS}}(T_c) = 0$. The intuitive reason for this is that for $0 < \Delta_{1/2} \ll T$, d, the odd $j = 0$ function $f_{p0\sigma}(T)$ becomes somewhat smaller than its $T = 0$ value of $1/2$, because with increasing T some of the probability for finding a quasiparticle in state j "leaks" from $j = 0$ to higher states with $j \neq 0$, for which $E_j^{-1} < E_0^{-1}$ in Eq. (91). Thus, the blocking-of-pair-scattering effect of the odd quasiparticle becomes slightly less dramatic as T is increased, so that $d_{1/2}^{\mathrm{BCS}}$ increases slightly.

It should be noted, however, that although the nonmonotonicities of $T_{c,0}(d)$ and $d_{1/2}^{\mathrm{BCS}}(T)$ are intuitively plausible within the g.c. framework in which they were derived, their physical significance is doubtful, since they fall in the regime where $d/\Delta_s \gtrsim 1$ and the g.c. approach is unreliable, due to its neglect of fluctuations.

5.5.2. Variational extensions of BCS theory

The above-mentioned results of von Delft et al. [21] were reproduced and extended to finite magnetic fields by Balian, Flocard and Vénéroni, using a more general g.c. variational BCS approach [25,26]. It is designed to optimize the characteristic function $\varphi(\xi) \equiv \ln \mathrm{Tr}\, \hat{P}_p e^{-\beta(\hat{H}-\mu\hat{N})}\hat{A}(\xi)$, where $\hat{P}_p$ is the parity projector of Eq. (90) and $\hat{A}(\xi) \equiv \exp(-\sum_\gamma \xi_\gamma \hat{Q}_\gamma)$, $\hat{Q}_\gamma$ being observables of interest (e.g. the total spin) and ξ_γ the associated sources (e.g. the magnetic field). This approach goes beyond the usual minimization of the free energy [101], since it optimizes not only thermodynamic quantities but also equilibrium correlation functions, which can be obtained by differentiating $e^{\varphi(\xi)}$ with respect to ξ_γ. However, its g.c. version also suffers from the drawback of yielding abrupt, spurious phase transitions even though the systems are finite. Presumably, this problem would be cured if an exact projection to fixed particle number were incorporated into this approach, but this is technically difficult and has not yet been worked out.

5.5.3. Static path approximation

For finite systems, in contrast to infinite ones, *fluctuations* of the order parameter about its mean-field value are very important in the critical regime, causing the phase transition to be
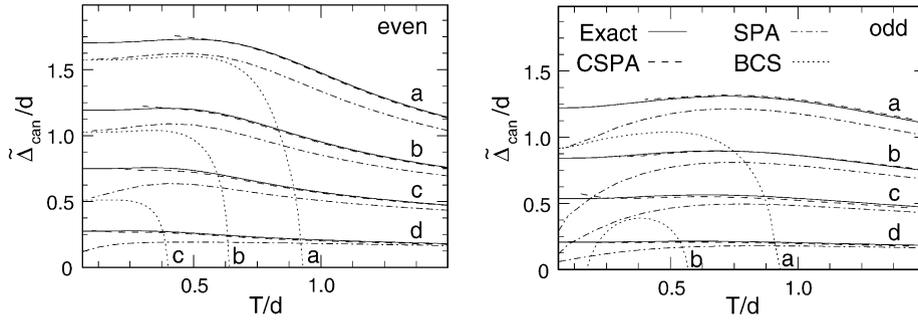
Fig. 25. Temperature dependence of the pairing correlation energy $\tilde{\Delta}_{can}$ of Eq. (92) [in units of level spacing d], as calculated in [30] using parity-projected mean-field BCS theory (dotted lines), SPA (dash–dotted), CSPA (dashed) and exact diagonalization (solid lines). A system of 10 equally spaced, doubly degenerate, levels was studied, containing 10 (left panel) or 11 (right panel) electrons. $\tilde{\Delta}_{can}(T)$ is shown at four fixed values of $d/\tilde{\Delta}$ (thus this figure elaborates Fig. 24), namely 0.60, 0.91, 1.7, and 15, labeled by a, b, c and d, respectively (calculated using $\lambda = 0.55, 0.45, 0.35$ and 0.2). CSPA data are shown only above the CPSA breakdown, which occurs at $T_{CSPA} < \tilde{\Delta}/4$ for the cases considered. The absence of dotted lines for the cases d (even) and c, d (odd) means that for these no nontrivial mean-field BCS solution exists. Of course, $(\tilde{\Delta}_{can})_{exact}$ is nonzero nevertheless. The abrupt BCS transition is completely smeared out for the SPA, CSPA and exact results, for which the asymptotic decay at $T \gg \tilde{\Delta}$ can be shown [30] to be $\tilde{\Delta}_{can} \sim \frac{1}{4}(\lambda^3 d/T)^{1/2}$.

smeared out; conversely, the spurious sharp transition found in the g.c. approaches above is a direct consequence of the neglect of such fluctuations. A rather successful way of including fluctuations is the so-called *static path approximation* (SPA), pioneered by Mühlschlegel et al. [94] and developed by various nuclear theorists [127–143], while recently an exact parity projection has also been incorporated [29–32]. A detailed and general discussion, including a complete list of relevant references, was given very recently by Rossignoli et al. [30]. We therefore confine ourselves below to stating the main strategies of the SPA and illustrating its capabilities by showing its results (Fig. 25) for the quantity

$$\tilde{\Delta}_{can}^2 = (\lambda d)^2 \sum_{ij} \left[C_{ij} - (C_{ij})_{\lambda=0} \right] . \tag{92}$$

$\tilde{\Delta}_{can}$ is reminiscent[13] of Δ_{can} of Eq. (52), and measures the increase in pairing correlation energy due to a nonzero coupling strength λ.

One starts by decoupling the quartic interaction $\hat{H}_{red}$ of Eq. (37) into a quadratic form using a Hubbard–Stratonovich transformation with a complex auxiliary field $\Delta(\tau) = \Delta^+(\tau) + i\Delta^-(\tau)$, with Matsubara-expansion $\Delta^{\pm}(\tau) = \sum_n \Delta_n^{\pm} e^{-i2\pi n\tau/\beta}$ in the interval $\tau \in (0, \beta)$. The parity-projected partition function of Eq. (90) then has the following path integral representation (our notation is deliberately schematic; see [30] for a precise version):

$$Z_p \propto \int \prod_n d\Delta_n^+ \, d\Delta_n^- \, \mathscr{Z}_p[\Delta], \quad \mathscr{Z}_p[\Delta] = \text{Tr}\{\hat{P}_p \hat{\mathscr{T}} e^{-\int_0^\beta d\tau \, h[\Delta(\tau)]}\} , \tag{93}$$

---

[13] The definitions of $\tilde{\Delta}_{can}$ and Δ_{can} differ by terms of order $(d/\omega_D)^{1/2}$; for example, when evaluating both using $|\text{BCS}\rangle$ of Eq. (47) and comparing to Δ_{gc} of Eq. (48), one finds $(\Delta_{can})_{BCS} = \Delta_{gc} = (\tilde{\Delta}_{can})_{BCS} + \mathcal{O}[(d/\omega_D)^{1/2}]$.

$$h[\Delta] = \sum_{j\sigma} (\varepsilon_j - \mu - \lambda d/2) c_{j\sigma}^\dagger c_{j\sigma} - \sum_j (b_j^\dagger \Delta + \Delta^* b_j) + \frac{|\Delta|^2}{\lambda d} . \tag{94}$$

The path integral can be treated at several levels of sophistication:

(i) In the simplest, one uses a "fixed-phase saddle-point approximation" for the "static" $n = 0$ modes and neglects all $n \neq 0$ modes, i.e. one fixes the phase of $\Delta_0^+ + i\Delta_0^- = |\Delta_0|e^{i\phi_0}$ by, say, setting $\phi_0 = 0$, so that $\int d\Delta_0^+ \, d\Delta_0^-$ is replaced by $\int d|\Delta_0|$, and approximates this integral by its saddle-point value. The saddle-point condition for maximizing $\mathscr{Z}_p[|\Delta_0|]$ then yields the gap equation (91), thus this approach simply reproduces the *parity-projected mean-field* approach of Section 5.5.1, including its sharp phase transition (Fig. 25, dotted lines).

(ii) The next-best approximation is obtained if one writes $\int d\Delta_0^+ \, d\Delta_0^- = \int_0^\infty |\Delta_0| d|\Delta_0| \int_0^{2\pi} d\phi_0$ and performs the phase integral fully. Remarkably, "liberating" the phase degree of freedom in this way already suffices to smooth out the phase transition [141,30], even if the $\int d|\Delta_0|$ integral is again replaced by its saddle-point value, provided that the latter is found by now maximizing $|\Delta_0| \mathscr{Z}_p[|\Delta_0|]$ (i.e. including the factor $|\Delta_0|$ from the integration measure). This yields a modified gap equation with a nontrivial solution for arbitrarily large T, i.e. no abrupt transition.

(iii) For finite systems, fluctuations about the saddle become large in critical regions. To obtain an improved description of the latter (Fig. 25, dash–dotted lines), the *static path approximation* (SPA) [94,127–133] incorporates all *static* fluctuations exactly, via a (numerical) evaluation of the full integral $\int_0^\infty |\Delta_0| d|\Delta_0| \int_0^{2\pi} d\phi_0$ over all "static paths".

(iv) In the so-called *correlated static path approximation* (CSPA) (also called SPA + RPA), small-amplitude quantum fluctuations around each static path are included too, by performing the remaining $\int d\Delta_{n \neq 0}^\pm$ integrals in the Gaussian approximation [134–140,29–32]. The CSPA yields qualitatively similar but quantitatively more reliable results (Fig. 25, dashed lines) than the SPA, but breaks down below a temperature T_{CSPA}, below which the fluctuations of the $\Delta_{n \neq 0}^\pm$ modes become large at unstable values of $|\Delta_0|$, causing the Gaussian approximation to fail.

(v) Finally, in the so-called *canonical CSPA* one projects the partition function not only to fixed number parity (as done throughout above) but also to fixed particle number, by performing an integration over the chemical potential (*before* performing any of the $\Delta_n^\pm$ integrals) [124,131–133]. However, this too is usually done only in the Gaussian approximation (and would produce negligible corrections to the CPSA results for the quantities shown in Fig. 25).

Comparisons with exact diagonalization results [142] (Fig. 25, solid lines) show that in its regime of validity ($T > T_{\mathrm{CSPA}}$), the CSPA produces results that are qualitatively completely similar and also quantitatively very close to the exact ones, whereas the quantitative agreement is significantly worse if only the SPA is used. Since the CSPA is conceptually simple, well-documented [30] and straightforward to implement, it seems to be the method of choice for not too low temperatures. A possible alternative is a quantum Monte Carlo evaluation of the path integral (93) [144,145], but the numerics is much more demanding than for the CPSA, while the convergence at low T is in general rather poor, due to the familiar sign problem of Monte Carlo methods.

The development of canonical finite-T methods that remain quantitatively reliable for $d \gtrsim \tilde{\Delta}$ and arbitrarily small T is one of the open challenges in this field. It would be very interesting if progress in this direction could be made by exploiting the integrability [117,118] of the model, using

Bethe ansatz techniques. For the FD regime, another possibility would be to develop a finite-T generalization of the self-consistent RPA approach of Dukelsky and Schuck [41].

5.5.4. Re-entrant spin susceptibility

For grains so small that $d \gg \tilde{\Delta}$, the spectroscopic *transport* measurements of BRT are not able, in principle, to reliably detect the effect of pairing correlations, since in this regime these cause only small changes to the eigenspectrum of a normal metallic grain, whose spectrum is, however, irregular to begin with. In contrast, *thermodynamic* quantities do have the potential to measurably reveal the existence of pairing correlations for $d \gg \tilde{\Delta}$. Since very recently parity effects for the *spin susceptibility* have been observed experimentally for an ensemble of small, normal metallic grains [146], it is an interesting and experimentally relevant question to investigate how pairing correlations affect its behavior in superconducting grains.

This question was worked out in detail by Di Lorenzo et al. [44], whose results are summarized in Fig. 26. The spin susceptibility for an isolated grain is defined as

$$\chi_p(T) = \left. -\frac{\partial^2 \mathscr{F}_p(T,H)}{\partial H^2} \right|_{H=0}, \tag{95}$$

where $\mathscr{F}_p = -k_B T \ln Z_p^{\mathrm{can}}$ is the free energy of a grain with parity p and Z_p^{can} is the canonical partition function.

In the bulk limit [Fig. 26(a)], it is well known [147] that the spin susceptibility decreases below its Pauli value $\chi_{\mathrm{Pauli}} = 2\mu_B^2/d$ once T drops below the superconducting transition temperature T_c, since the electrons tend to bind into Cooper pairs, which are spin singlets and do not contribute to the spin susceptibility. Interestingly, however, the spin susceptibility becomes parity dependent as T is lowered below the crossover temperature T_{cr}^* mentioned in the opening paragraphs of
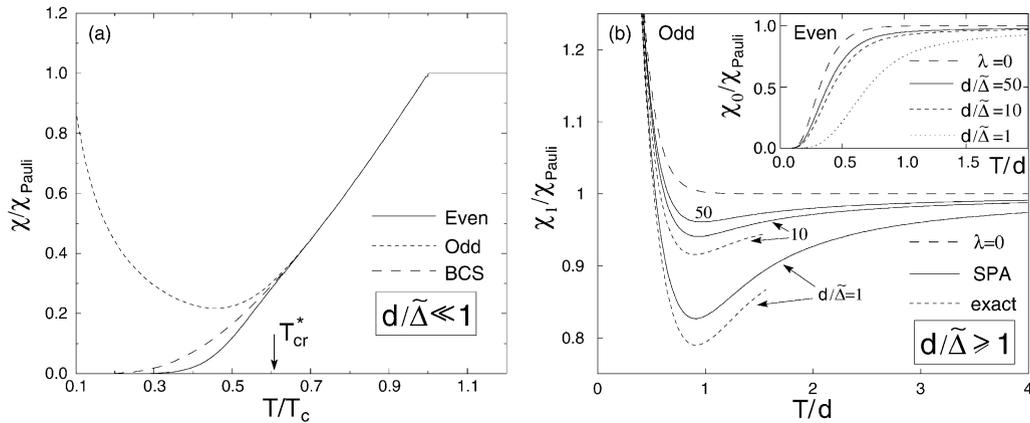

Fig. 26. Spin susceptibility χ_0 (χ_1) of an even (odd) superconducting grain as function of T, plotted in units of its bulk high-T value $\chi_{\mathrm{Pauli}} = 2\mu_B^2/d$. (a) Bulk limit ($d/\tilde{\Delta} \ll 1$): the even (solid) and odd (short-dashed) curves were calculated using parity-projected mean-field theory, the long-dashed curve using standard (unprojected) BCS theory. (b) Fluctuation-dominated regime ($d/\tilde{\Delta} \gtrsim 1$) for χ_1 (main figure) and χ_0 (inset): All curves were calculated using the static path approximation, except the short-dashed curves in the main figure, calculated using Richardson's exact solution, and the long-dashed curves for the non-interacting case ($\lambda = 0$).

Section 5.5: In the even case, χ_0 exponentially drops to zero for sufficiently small temperatures, $T \ll \max(\tilde{\Delta}, d)$, for reasons that are intuitively obvious in the two limits $\tilde{\Delta} \gg d$ (all electrons bound into Cooper pairs) and $\tilde{\Delta} \ll d$ (no Cooper pairs, but all levels doubly occupied). In contrast, in the odd case χ_1 shows a *re-entrant* behavior, in that it increases as μ_B^2/T for low temperatures, due to a Curie-like contribution from the unpaired odd electron. As a result, $\chi_1(T)$ has a *minimum* somewhat below T_{cr}^*, which can be viewed as a "smoking gun" for pairing correlations, since it is absent for odd normal grains. For the latter, $\chi_1(T)$ [long-dashed $\lambda = 0$ curve in Fig. 26(b)] also has the Curie-like increase at very low T, but lacks the initial pairing-induced decrease as T is reduced below T_c.

Remarkably, Di Lorenzo et al. found that *this re-entrance of χ_1 survives also for $d \gtrsim \tilde{\Delta}$* [Fig. 26(b)]: although pairing correlations survive here only as fluctuations, these are evidently sufficiently strong to still significantly reduce $\chi_1(T)$ relative to χ_{Pauli} [by several percent even for $d/\tilde{\Delta} \simeq 50$ (!)], before the Curie-like increase sets in at low T. Di Lorenzo et al. established this result by considering the limits $T \gg d$ and $T \ll 2d$ analytically, using a static path approximation to capture the crossover numerically, and checking the results for $T \lesssim d$ using Richardson's exact solution (they considered all eigenstates with excitation energy up to a cutoff $\Lambda \sim 40d$, for grains with $N \leq 100$ electrons). This check shows that the static path approximation somewhat underestimates the amount of pairing correlations (its minima for $\chi_1(T)$ are too shallow), but in general is in good qualitative agreement with the exact results, confirming that it is a useful and qualitatively reliable tool for describing the crossover regime.

6. Nonequilibrium effects

So far, we have always assumed that the gate voltage has been tuned such that the system is close to a Coulomb-blockade degeneracy point, where the Coulomb-blockade barrier $\min[\delta E_{pot}^\pm]$ of Eq. (7) is small (of order d) and nonequilibrium effects can be neglected. However, since this requires fine-tuning V_g, it is a rather nongeneric situation. In this section, we discuss the opposite case in which this Coulomb barrier is so large that transport occurs only in the nonequilibrium regime, for which several interesting new phenomena have been observed [10–13,15]. Among these, we focus (i) in Section 6.1 on the observation of *clusters* of closely spaced levels in normal grains; (ii) in Section 6.2 on the occurrence of *prethreshold structures* in the excitation spectra of even superconducting grains; and (iii) in Section 6.3 on a direct observation of the crossover from a discrete to a continuous excitation spectrum at energies of order $E_{Thouless}$.

All of these can be understood only by going beyond the simple theoretical framework presented in Section 2.3: (i) and (iii) require consideration of corrections to the "orthodox" model for treating the Coulomb interaction, which are of order d/g_{dim} and hence become important for sufficiently small grains; and (ii) requires consideration of cotunneling processes, whereas we had hitherto restricted our attention to sequential tunneling.

The general significance of the results presented below is that they provide a very direct confirmation and illustration of the general theoretical picture [66,64] (summarized in Section 6.3) for the nature of excitations in disordered interacting systems that has emerged during the last decade. The development of this picture was initiated by studies of semiconductor quantum dots [66], but is here beautifully confirmed for metallic grains too.

6.1. Clusters of resonances

6.1.1. Experimental results

Figs. 27(a) and (b) display the excitation spectrum, dI/dV vs. energy, of two different *normal* metallic grains, of radii roughly 2.7 and 3.6 nm (if assumed hemispherical). The spectra display four clear features:

(1) The low resonances of the differential conductance are grouped in *clusters*.
(2) The distance between nearby clusters is of order the mean level spacing d of the noninteracting electrons in the dot.
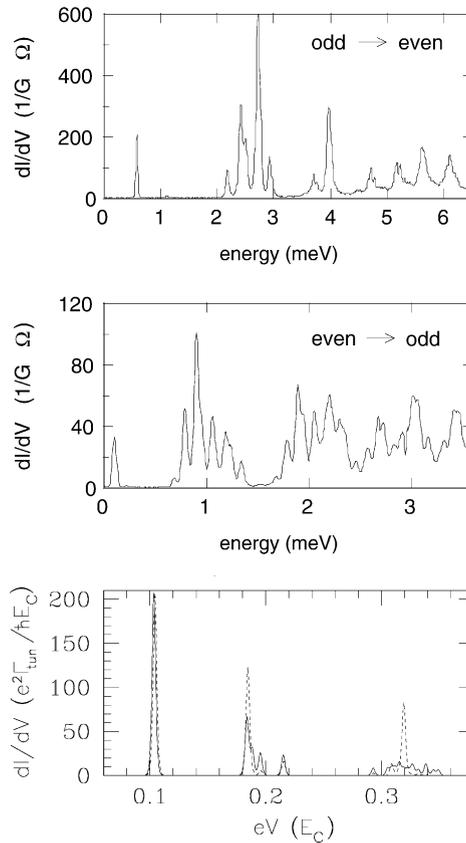

Fig. 27. (a,b) The excitation spectra (at $T = 30$ mK, $H = 0$) of two ultrasmall A1 grains with volumes ≈ 40 nm^3 (a) and ≈ 100 nm^3 (b) [11]. The first resonance is isloated while subsequent resonances are clustered in groups. The distance between nearby groups of resonances is approximately the single-particle mean level spacing d. (c) Differential conductance for a model system of 7 equally spaced levels, occupied alternately by 4 or 5 electrons in a current-carrying steady state, calculated [11] from Eq. (18) by solving the master equation (17) numerically. The tunneling rate Γ_α^r into each level was chosen to be uniform, $\Gamma_\alpha^r \equiv \Gamma_{\text{tun}}$, and a temperature $T = d/100$ and $\delta U = d/5$ were used, where δU is the variance of the fluctuations [Eq. (110)] in the interaction energy, see text. Dashed line: inelastic relaxation rate Γ^{inel} of each level was chosen to be larger than the tunneling rate into it ($\Gamma^{\text{inel}} = 5\Gamma_{\text{tun}}$); solid line: no inelastic processes ($\Gamma^{\text{inel}} = 0$).

(3) The first cluster contains only a single resonance.
(4) Higher clusters consist of several resonances *spaced much more closely than d.*

Agam et al. [11] have shown that these features are manifestations of the interplay between nonequilibrium effects and electron–electron interactions, where the latter must be treated more carefully than in the usual orthodox model for charging effects. The general idea is simple: if the threshold voltage for the onset of current through the grain is large, the transfer of one electron through the grain by sequential tunneling can leave behind several possible particle–hole excitations on the grain. If sufficiently long lived, each of these excitations will modify the threshold energy for the next electron that tunnels into the grain in a slightly different way, and thus cause what would have been a single conductance resonance to split up into a cluster of subresonances.

6.1.2. Reaching nonequilibrium via sequential tunneling

To explain the interpretation of Agam et al. in more detail, we shall begin our analysis within the framework of the orthodox model defined by Eqs. (8)–(10) of Section 2.3, where the notation used below is introduced and explained. The present grains are so small that the influence of superconducting pairing may be neglected, since the single-particle mean level spacing, $d \simeq 1\,\text{meV}$, is much larger than the BCS superconducting gap for A1, which is $0.18\,\text{meV}$. We thus describe the grain by the normal-state single-particle Hamiltonian of Eqs. (13) and (14),

$$\hat{H}_{\text{D,orth}}^{\text{normal}} = E_{\text{pot}}(\hat{N}_{\text{ex}}) + \sum_p \varepsilon_p c_p^\dagger c_p = \sum_\alpha (eV_{\text{D}} N_{\text{ex},\alpha} + \mathscr{E}_\alpha)|\alpha\rangle\langle\alpha| \,, \tag{96}$$

$$|\alpha\rangle = |\{n^p\}\rangle, \quad \mathscr{E}_\alpha = E_{\text{C}} N_{\text{ex},\alpha}^2 + \sum_p \varepsilon_p n_p^{(\alpha)} \,, \tag{97}$$

where, for notational simplicity, we have here used the roman label p for the combined labels $l\sigma\text{D}$ of Eq. (13). $\hat{H}_{\text{D,orth}}^{\text{normal}}$ incorporates the effect of electron–electron interactions only through the electrostatic contribution $E_{\text{pot}}(\hat{N}_{\text{ex}})$ [defined in Eq. (6)], which, as we shall see below, is too simplistic to explain the clustering properties (1)–(4). Corrections to the orthodox model will be introduced (Section 6.1.3 below) once its insufficiency has become apparent.

For definiteness, suppose that the overall ground state of the grain has N electrons and that the measured excitation spectra arise via tunneling transitions, with the left barrier acting as bottleneck, between N and $(N + 1)$ electron states (see Fig. 28). Denoting their respective ground states by $|G\rangle_N$ and $|G\rangle_{N+1} \equiv c_0^\dagger|G\rangle_N$, where $c_0^\dagger$ is the creation operator for the lowest unfilled level of $|G\rangle_N$, the corresponding ground state energy difference is

$$\mathscr{E}_{\text{G}}^{N+1} - \mathscr{E}_{\text{G}}^N + eV_{\text{D}} = \delta E_{\text{pot}}^+ + \varepsilon_0 \,, \tag{98}$$

where $\delta E_{\text{pot}}^+ = E_{\text{pot}}(N_{\text{ex}} + 1) - E_{\text{pot}}(N_{\text{ex}})$ is the electrostatic energy cost for entering the grain. All N-electron excited states are particle–hole excitations w.r.t. $|G\rangle_N$. We shall denote them by $|p\bar{p}\rangle_N \equiv c_p^\dagger c_{\bar{p}}|G\rangle_N$ and $|q\bar{q}, p\bar{p}\rangle_N \equiv c_q^\dagger c_{\bar{q}} c_p^\dagger c_{\bar{p}}|G\rangle_N$, etc., and within the orthodox model of Eqs. (96) and (97), their excitation energies are

$$(\mathscr{E}_{p\bar{p}}^N - \mathscr{E}_{\text{G}}^N)_{\text{orth}} = \varepsilon_p - \varepsilon_{\bar{p}}, \quad (\mathscr{E}_{q\bar{q},p\bar{p}}^N - \mathscr{E}_{\text{G}}^N)_{\text{orth}} = \varepsilon_q - \varepsilon_{\bar{q}} + \varepsilon_p - \varepsilon_{\bar{p}} \,, \tag{99}$$
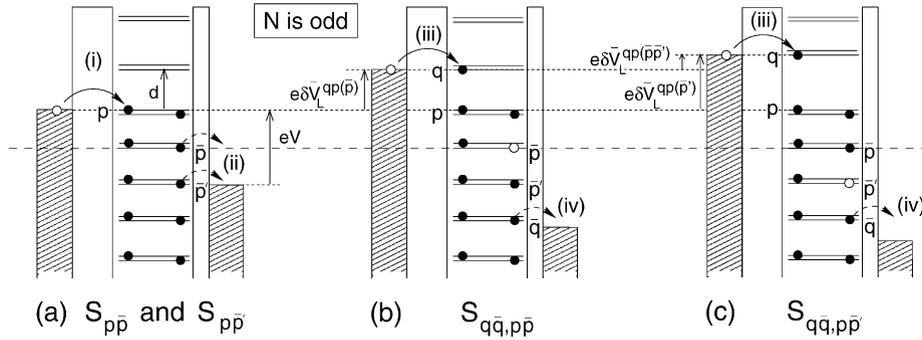
Fig. 28. Schematic depiction (using the conventions of Fig. 4) of the tunneling sequences, starting from the ground state $|G\rangle_N$, that lead to clusters of resonances in nonequilibrium transport. Filled circles depict the electron configuration of the intermediate $(N + 1)$-particle states (with N odd), solid arrows the bottleneck tunneling transitions into them; levels drawn close together represent time-reversed pairs that are degenerate in the absence of a magnetic field. (a) $S_{p\bar{p}}$ and $S_{p\bar{p}'}$ [Eq. (100)] are two possible sequences via level p that leave the grain in an excited final state, $|p\bar{p}\rangle_N$ or $|p\bar{p}'\rangle_N$. (b) $S_{q\bar{q},p\bar{p}}$ or (c) $S_{q\bar{q},p\bar{p}'}$ [Eq. (104)] are two possibilities, differing only in the hole position ($\bar{p}$ vs. $\bar{p}'$), for the next tunneling sequence via level q; their thresholds differ by an amount $e\delta V_L^{pq(\bar{p}\bar{p}')}$ [Eq. (109)], since the interaction energies of two electrons in levels q and $\bar{p}'$ or q and $\bar{p}$ differ [Eq. (110)]. The figure is drawn such that $p = 0$ and $q = 1$ are the two lowest-lying empty levels of $|G\rangle_N$.

etc. Analogous definitions hold for $N + 1$. Now, sequential tunneling through the grain occurs via sequences of transitions of the type [Fig. 28(a)]

$$S_{p\bar{p}}: \quad |G\rangle_N \overset{(i)}{\to} |p0\rangle_{N+1} \overset{(ii)}{\to} |p\bar{p}\rangle_N \ . \tag{100}$$

In step (i), an electron tunnels via the left (bottleneck) barrier into an empty level p (with $\varepsilon_p \geq \varepsilon_0$), leaving the grain in an intermediate state $|p0\rangle_{N+1}$, which we shall call a "quasiparticle state" (following [64]); and in step (ii), one tunnels out via the right barrier from a filled level $\bar{p}$, leaving the grain in the excited (if $\bar{p} \neq p$) final particle–hole state $|p\bar{p}\rangle_N$. According to Eq. (26), steps (i) and (ii) become possible only once the following inequalities hold, respectively (with $\bar{V}_r \equiv V_r - V_D$):

$$(i) \quad e\bar{V}_L \geq \mathscr{E}_{p0}^{N+1} - \mathscr{E}_G^N, \qquad (ii) \quad -e\bar{V}_R \geq \mathscr{E}_{p\bar{p}}^N - \mathscr{E}_{p0}^{N+1} \ . \tag{101}$$

Addings Eqs. (101(i)) + (98) and Eqs. (101(i)) + (101(ii)) gives, respectively,

$$(i)' \quad eV/2 = eV_L \geq \delta E_{pot}^+ + \varepsilon_p, \qquad (ii)' \quad eV \geq \mathscr{E}_{p\bar{p}}^N - \mathscr{E}_G^N \ , \tag{102}$$

which have obvious physical interpretations: Condition (i)′ makes explicit that transport is possible only if the bias voltage is large enough that an electron leaving the left lead can overcome the Coulomb blockade barrier δE_{pot}^+; the lowest-threshold conductance resonance evidently occurs for $p = 0$. Condition (ii)′ states that an excited final state $|p\bar{p}\rangle_N$ can be produced whenever the energy supplied by the bias voltage exceeds the corresponding excitation energy $\mathscr{E}_{p\bar{p}}^N - \mathscr{E}_G^N$. Now, (i)′ implies that if the Coulomb blockade barrier is large ($\delta E_{pot}^+ \gg d$), i.e. for transport far away from a degeneracy point, the threshold bias voltage for the current to begin to flow will be large ($eV \gg d$); but this would automatically guarantee that (ii)′ is fulfilled, at least for low-lying final-state excitations with an energy cost of only a few d. We can thus draw the following important

conclusion: *Since sequential tunneling transport away from a degeneracy point requires a large threshold bias voltage, it will always generate excited final states once this threshold is reached.*

Whether an excited final state $|p\bar{p}\rangle_N$ affects transport or not depends on the ratio between $\Gamma_{p\bar{p}}^{\mathrm{inel}}$, its inelastic decay rate, and Γ_{tun}^r, the average tunneling rate across the bottleneck barrier r into an individual discrete level near ε_F. In all previous sections we had assumed that $\Gamma_{p\bar{p}}^{\mathrm{inel}} \gg \Gamma_{\mathrm{tun}}^r$, i.e. that $|p\bar{p}\rangle_N$ will decay to $|G\rangle_N$ long before the next electron tunnels in across the left barrier. In this case, the distance between two so-called "quasiparticle conductance resonances", associated with the onset of sequential tunneling through two different intermediate quasiparticle state $|q0\rangle_{N+1}$ and $|p0\rangle_{N+1}$, is determined by the bottleneck condition (101(i)) [or equivalently from Eq. (30)] to be

$$e\delta\bar{V}_{\mathrm{L}}^{qp} \equiv \mathscr{E}_{q0}^{N+1} - \mathscr{E}_{p0}^{N+1} = \varepsilon_q - \varepsilon_p \ . \tag{103}$$

However, estimates of $\Gamma_{p\bar{p}}^{\mathrm{inel}}$ (discussed in Section 6.3) show that it is quite possible for an ultrasmall grain to have $\Gamma_{p\bar{p}}^{\mathrm{inel}} \lesssim \Gamma_{\mathrm{tun}}^r$. In this case, the grain will still be in nonequilibrium when the next electron tunnels in, so that new tunneling sequences of the type [Fig. 28(b)]

$$S_{q\bar{q},p\bar{p}}: |p\bar{p}\rangle_N \overset{\mathrm{(iii)}}{\to} |q0,p\bar{p}\rangle_{N+1} \overset{\mathrm{(iv)}}{\to} |q\bar{q},p\bar{p}\rangle_N \ , \tag{104}$$

$$\mathrm{(iii)} \quad e\bar{V}_{\mathrm{L}} \geq \mathscr{E}_{q0,p\bar{p}}^{N+1} - \mathscr{E}_{p\bar{p}}^N, \qquad \mathrm{(iv)} \quad -e\bar{V}_{\mathrm{R}} \geq \mathscr{E}_{q\bar{q},p\bar{p}}^N - \mathscr{E}_{q0,p\bar{p}}^{N+1} \ , \tag{105}$$

become possible once the corresponding conditions (iii) and (iv) are met. These differ from $S_{q\bar{q}}$ by the presence of additional particle–hole excitations $p\bar{p}$. The change in $e\bar{V}_{\mathrm{L}}$ needed to reach threshold (105(iii)) from threshold (101(i)) is

$$e\delta\bar{V}_{\mathrm{L}}^{qp(\bar{p})} \equiv (\mathscr{E}_{q0,p\bar{p}}^{N+1} - \mathscr{E}_{p\bar{p}}^N) - (\mathscr{E}_{p0}^{N+1} - \mathscr{E}_{G}^N) \tag{106}$$

[see Fig. 28(b)], and each time this is positive, the conductance should show another resonance (if $e\delta\bar{V}_{\mathrm{L}}^{qp(\bar{p})} < 0$, the sequence $S_{q\bar{q},p\bar{p}}$ simply increases the peak height associated with the sequences $S_{p\bar{p}}$ through level p, without causing another peak, i.e. these *do* contribute to the current, although they are spectroscopically "hidden"). But within the orthodox model, $e\delta\bar{V}_{\mathrm{L}}^{qp(\bar{p})} = \varepsilon_q - \varepsilon_p$ [by Eqs. (99)], which is the *same* as $e\delta\bar{V}_{\mathrm{L}}^{qp}$ of Eq. (103). Thus, the orthodox model predicts that even in the presence of nonequilibrium final states, the spacing of conductance peaks will be of order d. It is therefore unable to explain the observed occurrence in Figs. 27(a) and (b) of *clusters* with mean inter-cluster spacing d and an intra-cluster spacing much less than d.

6.1.3. Beyond the orthodox model: interaction fluctuations

The reason for this failure, identified by Agam et al. [11,13], is that the orthodox model treats the electron–electron interaction too simplistically. Its general form is

$$\hat{H}_{\mathrm{el-el}} = \frac{1}{2} \sum_{ijkl} \sum_{\sigma\sigma'} U_{ijkl} c_{i\sigma}^\dagger c_{j\sigma'}^\dagger c_{k\sigma'} c_{l\sigma} \ , \tag{107}$$

$$U_{ijkl} = \int d\mathbf{r}_1 \, d\mathbf{r}_2 \, U(\mathbf{r}_1, \mathbf{r}_2) \psi_i^*(\mathbf{r}_1) \psi_j^*(\mathbf{r}_2) \psi_l(\mathbf{r}_1) \psi_k(\mathbf{r}_2) \ , \tag{108}$$

where $U(\mathbf{r}_1, \mathbf{r}_2)$ is the (screened) interaction potential. The orthodox model keeps only "diagonal" terms in which two pairs of indices are identical (U_{ijji}, U_{iijj}, U_{ijij}), namely

$$\hat{H}_{\text{el-el}}^{\text{orth}} = E_C \hat{N}_{\text{ex}}^2 - \lambda d \sum_{ij} c_{i+}^\dagger c_{i-}^\dagger c_{j-} c_{j+} + \lambda_s d \sum_{ij,\sigma\sigma'} c_{i\sigma}^\dagger c_{j\sigma'}^\dagger c_{i\sigma'} c_{j\sigma} + \mathcal{O}(d/g_{\text{dim}})$$

and neglects both "*off-diagonal* terms" U_{ijkl} (no equal indices) and ij-dependent *fluctuations* in diagonal terms. (λ and λ_s are dimensionless, volume-independent interaction constants in the pair and spin channels; we took $\lambda = \lambda_s = 0$ in this section, since we are not interested in superconducting or magnetic correlations here). Now, by studying the correlations of eigenfunctions in chaotic systems (and an irregularly shaped ultrasmall grain may be viewed as chaotic), the neglected terms can be shown [13,67–71,148] to be small as d/g_{dim}, where $g_{\text{dim}} = E_{\text{Thouless}}/d$ is the dimensionless conductance[14] and E_{Thouless} the Thouless energy [cf. Eqs. (4) and 5]. Intuitively, one may say that the smaller $1/g_{\text{dim}}$, the more uniform are the wavefunctions and the smaller the fluctuations in the interaction energy. Indeed, the neglect of these fluctuations is parametrically justified for bulk samples ($1/g_{\text{dim}} \simeq 0$) and in mesoscopic samples ($1/g_{\text{dim}} \ll 1$). For ultrasmall grains, however, $1/g_{\text{dim}}$ can become of order unity, so that interaction fluctuations are expected to become measurable.

Agam et al. [11,13] attributed the above-mentioned clustering properties (1)–(4) to a combination of precisely such interaction fluctuations and the presence of nonequilibrium excitations on the grain: they argued that the energy of the added electron in level q of the intermediate state $|q0, p\bar{p}\rangle_{N+1}$ depends, in general, on precisely which particle–hole excitation $p\bar{p}$ is present. In particular, two intermediate states that differ only by the interchange of a hole between levels $\bar{p}$ and $\bar{p}'$, say $|q0, p\bar{p}\rangle_{N+1}$ and $|q0, p\bar{p}'\rangle_{N+1}$, will have thresholds that differ, according to Eq. (105(iii)) by [see Figs. 28(b) and (c)]

$$e\delta\bar{V}_{\text{L}}^{qp(\bar{p}\bar{p}')} \equiv (\mathcal{E}_{q0,p\bar{p}}^{N+1} - \mathcal{E}_{p\bar{p}}^N) - (\mathcal{E}_{q0,p\bar{p}'}^{N+1} - \mathcal{E}_{p\bar{p}'}^N) \,. \tag{109}$$

Agam et al. argued that the magnitude of this expression may be estimated using the Hartree term of interaction (107), which gives $U_{q\bar{p}'\bar{p}'q} - U_{q\bar{p}\bar{p}q}$, i.e.

$$e\delta\bar{V}_{\text{L}}^{qp(\bar{p}\bar{p}')} \simeq \int d\mathbf{r}_1 \int d\mathbf{r}_2 |\psi_q(\mathbf{r}_1)|^2 U(\mathbf{r}_1, \mathbf{r}_2)(|\psi_{\bar{p}'}(\mathbf{r}_2)|^2 - |\psi_{\bar{p}}(\mathbf{r}_2)|^2) \,. \tag{110}$$

Let δU denote the variance of (110), evaluated either by ensemble or energy averaging. If diffusive dynamics is assumed, δU can be shown [68] to be of order cd/g_{dim},[14,15] where c is a geometry-dependent constant of order unity. Thus, *each quasiparticle resonance associated with the set of sequential tunneling sequences $S_{q0,p\bar{p}}$ through a level q should split up into a cluster of subresonances, spaced by d/g_{dim}, associated with different nonequilibrium particle–hole configurations $p\bar{p}$.* Since in Figs. 27(a) and (b) there typically are about 5 subresonances per cluster, we can take $g_{\text{dim}} \simeq 5$ to be

---

[14] Strictly speaking, in this context one should use $g_{\text{dim}} = \hbar\gamma_1/d$, where γ_1 is the first non-vanishing Perron–Frobenius eigenvalue [67,13], which depends on the geometry and the amount of disorder present. E.g. for a pancake-shaped grain with thickness z, radius r and diffusive dynamics, Agam [13] gives $g_{\text{dim}} \propto (\hbar v_F/2rd)(z/r)^2$.

[15] This estimate does not take into account a change in the electrostatic potential due to the extra electron in level q. Blanter et al. [69] showed that the latter effect will lead to even stronger fluctuations, namely $\delta U \sim d/\sqrt{g_{\text{dim}}}$.

an "experimental estimate" for the dimensionless conductance, which is entirely reasonable for grains of the present size.[16]

Whether the *lowest*-lying quasiparticle resonance splits into subresonances or not depends on the threshold voltage for the onset of transport and on the grain's number parity. Let ε_0 and ε_1 denote the two lowest-lying empty levels in $|G\rangle_N$. The lowest-lying resonance involves tunneling sequences $S_{0\bar{p}}$ via level $p = 0$. A necessary condition for it to split into a cluster is that the level $\bar{p}$ from which the second electron leaves the grain [step (ii) of $S_{0\bar{p}}$] should come from a Kramers doublet not involving the level ε_0 into which the first electron tunneled [step (i) of $S_{0\bar{p}}$], implying $\varepsilon_0 - \varepsilon_{\bar{p}} \gtrsim d$, which requires a threshold [by Eq. (102(ii))] of $eV \geq \mathscr{E}_{0\bar{p}}^N - \mathscr{E}_G^N \gtrsim d$. This condition is evidently not met in either of Figs. 27(a) and (b), for which the threshold voltage at the first peak is significantly smaller than the inter-cluster spacing, which explains in both figures why the first peak not split into a cluster.

If this condition *is* met, then the lowest-lying peak will or will not split into subresonances, depending on whether ε_0 and ε_1 are degenerate or not, respectively, as can be seen by choosing $p = 0$ and $q = 1$: If N is even so that $\varepsilon_0 = \varepsilon_1$ (by Kramers degeneracy, see Section 3.1), the lowest quasiparticle resonance will have a cluster of subresonances on its upper shoulder, generated by those tunneling sequences $S_{1\bar{q},0\bar{p}}$ through level 1 for which $e\delta\bar{V}_L^{10(\bar{p})} > 0$. On the other hand, if N is odd so that $\varepsilon_1 - \varepsilon_0 \simeq d$, the threshold difference $e\delta\bar{V}_L^{10(\bar{p})}$ is of order d too, hence the $S_{1\bar{q},0\bar{p}}$ cluster will be separated from the lowest-lying resonance $S_{0\bar{p}}$, which remains unsplit, by an energy of order d.

In general, when M available states below the highest accessible energy level (including spin) are occupied by $M' < M$ electrons, there are $\binom{M}{M'}$ different occupancy configurations. The typical width of a cluster of resonances in this case is of order $W^{1/2}d/g_{\text{dim}}$, where $W = \min(M - M', M')$. The width of a cluster of resonances therefore *increases* with the bias voltage. The distance between nearby resonances of the cluster, on the other hand, *decreases* as $(W^{1/2}d/g_{\text{dim}})/\binom{M}{M'}$. This behaviour can be seen in Fig. 27(c), which shows the differential conductance obtained in a model calculation by Agam et al. [11] and explicitly demonstrates the splitting of resonances induced by fixed fluctuations δU in the interaction energy.

6.2. Prethreshold structures in superconducting grains

Nonequilibrium effects also cause interesting anomalies in Al grains large enough for superconducting pairing correlations to be important [10]. Figs. 29(a) and (b) show the evolution with gate voltage V_g of the odd-to-even and even-to-odd tunneling spectra, respectively, of the same grain as that discussed in Section 4.1. The top curves in both (a) and (b) correspond to V_g at a degeneracy point, and show features familiar from Section 4.1 and indicative of pairing correlations: a distinct gap (of 0.55 meV) between the first and second main peaks in the odd-to-even case, and its absence in the even-to-odd case [cf. Figs. 9 and 10]. As V_g is changed away from the degeneracy point, we

---

[16] For comparison, note that a similar order of magnitude can be obtained from Eqs. (2), (4) and (5(ii)), for a pancake-shaped grain with Volume $\approx 4\pi r^2 z$ and $a \approx (r/z)^2$ [cf. footnote 1 and footnote 2], for which $g_{\text{dim}} \approx 50z^3/r(\text{nm})^{-2}$: for radius $r \approx 3$ nm and thickness $z \approx 0.7$ nm this would give $g_{\text{dim}} \approx 6$. It is clear, though, that such estimates depend very sensitively on the relative magnitudes of r and z.
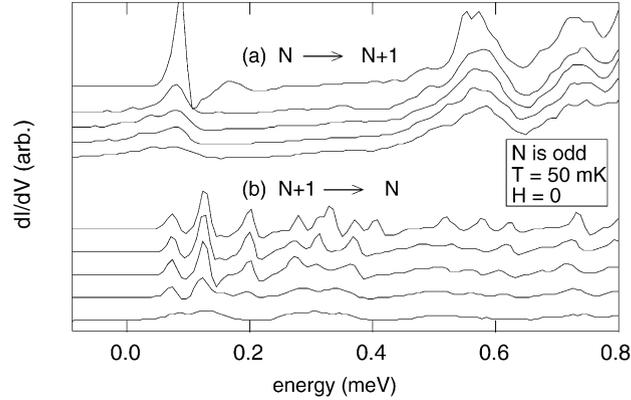
Fig. 29. The tunneling spectra [10], for five different values of the gate voltage V_g, of a largish Al grain with (a) an odd and (b) an even number of electrons in its $V = 0$ ground state. The top curve in each case corresponds to V_g at a degeneracy point, i.e. tuned such that the peaks occur at the lowest possible values for the bias voltage $|V|$. As V_g is changed to shift them to larger $|V|$ values (by amounts $\delta|V| = 0.2$, 0.4, 0.6 and 1.2 mV), the peaks broaden and develop substructure. The corresponding curves have been artificially shifted in energy to align peaks due to the same eigenstates. The sample is the same as in Figs. 2, 7, 9 and 10. All curves were taken at $T = 50$ mK and, to improve spectroscopic resolution, at $H = 0$, so that the Al leads were superconducting [this is the reason for the dip to the right of the first peak in the topmost curve, cf. Section 2.4.3 and Fig. 6(b)].

note the following salient features:

(a)   *Odd-to-even spectrum* [Fig. 29(a), $N \to N + 1$ transitions, with N odd]:
   (a1)   The height of the first peak rapidly decreases and a structure of "*prethreshold*" subresonances develops on its low-voltage shoulder, i.e. *below* the threshold of the first main peak.
   (a2)   The *characteristic spacing* between these subresonances is comparable to that of the first few peaks of the even-to-odd spectrum in Fig. 29(b).
(b)   *Even-to-odd spectrum* [Fig. 29(b), $N + 1 \to N$ transitions]:
   (b1)   For V_g sufficiently close to the degeneracy point (top three curves), the first few peaks of the differential conductance are *merely shifted* by increasing V_g, without changing their shape.
   (b2)   For sufficiently large changes of V_g (lowest two curves), the resonances are rapidly *smeared out* by increasing V_g.

Properties (b1) and (b2) are easily understood: due to the large spectral gap of 0.55 meV for creating an excitation on an even grain, the even-to-odd spectrum will be affected by nonequilibrium excitations only once the bias voltage $e|V|$ exceeds this value. Therefore, the low-lying peaks of the even-to-odd spectrum are smeared out by V_g only for those curves for which V_g has shifted the bias threshold to $\delta|eV| > 0.55$ meV, namely the lowest two of Fig. 29(b).

   Property (a1), however, stands in direct contrast to the predictions of the sequential tunneling mechanism of Section 6.1.3, namely that for odd N, the lowest peak of the spectrum does *not* split into subresonances [cf. third paragraph after Eq. (110)]. Moreover, the existence of the *prethreshold* peaks mentioned in (a1) suggests that nonequilibrium is reached even while $|eV|$ still is *below* the threshold for the onset of equilibrium sequential tunneling through the bottleneck barrier (say L),

namely

$$e\bar{V}_{\mathrm{L}} \geq \mathscr{E}_{\mathrm{G}}^{N+1} - \mathscr{E}_{\mathrm{G}}^{N} \ . \tag{111}$$

Agam and Aleiner [12] have therefore proposed that in the present case, nonequilibrium is established not by sequential tunneling, but instead by *cotunneling*. A cotunneling process with $|G\rangle_N$ as initial state involves a coherent, second-order transition of the form

$$|G\rangle_N \overset{c_p^\dagger}{\to} |\beta\rangle_{N+1} \overset{c_{\bar{p}}}{\to} |\alpha\rangle_N \ , \tag{112}$$

in which an electron tunnels into level p via the bottleneck barrier (say L), producing a virtual intermediate state $|\beta\rangle_{N+1}$, and one tunnels out of level $\bar{p}$ via the other barrier (R), leaving the grain in the final state $|\alpha\rangle_N$. (Since the grain is pair-correlated, $|\alpha\rangle_N$ and $|\beta\rangle_{N+1}$ are, in general, of the form of Eq. (40).) The energy conditions for cotunneling are

$$\text{(i)} \quad e\bar{V}_{\mathrm{L}} < \mathscr{E}_{\beta}^{N+1} - \mathscr{E}_{\mathrm{G}}^{N}, \qquad \text{(ii)} \quad eV \geq \mathscr{E}_{\alpha}^{N} - \mathscr{E}_{\mathrm{G}}^{N} \ . \tag{113}$$

Here (ii) is the analogue of Eq. (102(ii)′) for sequential tunneling, and (i), which should be contrasted to Eq. (101(i)), is the condition that the $|G\rangle_N \to |\beta\rangle_{N+1}$ transition is virtual, not real (once the inequality (113(i)) is violated, cotunneling turns into sequential tunneling). Since conditions (113) are much less stringent than (111), cotunneling sets in long before equilibrium sequential tunneling.

The current due to pure cotunneling is of course typically very small, since the cotunneling rate $\Gamma_{p\bar{p}} \propto \Gamma_p^{\mathrm{L}} \Gamma_{\bar{p}}^{\mathrm{R}}$ is smaller than that for either step of a sequential tunneling sequence by one factor of Γ [the latter are defined in Eq. (15)]. However, if the excited state $|\alpha\rangle_N$ produced by cotunneling is sufficiently long-lived, it can enable the next electron to enter the grain via *sequential* tunneling whenever the condition

$$e\bar{V}_{\mathrm{L}} \geq \mathscr{E}_{\mathrm{G}}^{N+1} - \mathscr{E}_{\alpha}^{N} \tag{114}$$

is satisfied, which will cause a step-like increase in current as eV is tuned past this "nonequilibrium sequential tunneling threshold". Now, since the latter lies *below* the threshold of Eq. (111) for equilibrium sequential tunneling by an amount $\mathscr{E}_{\alpha}^{N} - \mathscr{E}_{\mathrm{G}}^{N} > 0$, a set of subresonances will occur on the low-bias shoulder of the first main resonance, precisely as observed in (a1). Moreover, the spacing between these subresonances will reflect that between consecutive excited states $\mathscr{E}_{\alpha}^{N}$ of the odd grain, i.e. between the peaks measured in the even-to-odd excitation spectrum, in accordance with property (a2).

Agam and Aleiner [12,13] worked out these ideas quantitatively and concluded that the rate of creation of nonequilibrium excitations by cotunneling is indeed sufficient to explain the data of Fig. 29(a): they showed that the smallness of the cotunneling rates due to their being quadratic in Γ_p^r is offset to a large extent by the increase in the density of states near ε_{F} of both the grain and the leads, due to superconducting pairing correlations in both.

6.3. Discussion of relaxation rates

As emphasized above, nonequilibrium can only be established if relaxation processes are slower than tunneling processes, $\Gamma_\varepsilon^{\mathrm{inel}} \lesssim \Gamma_{\mathrm{tun}}$ for $\varepsilon \simeq d$, hence it is important to understand the mechanism

contributing to $\Gamma_\varepsilon^{\text{inel}}$. Agam et al. [11] discussed several possibilities: (i) The electron–phonon interaction; (ii) the electron–electron interaction beyond the Hartree–Fock approximation; (iii) Auger processes in which an electron in the dot relaxes while another one in the lead is excited; (iv) the relaxation of an electron in the dot as another electron tunnels out to the lead; and (v) thermalization with the leads via tunneling. Of these, it is easy to see that (iii)–(v) are at best small corrections: (iii) is exponentially small in the ratio w/χ of the barrier width w to the screening length χ, and moreover the interaction between electrons on both sides of the tunnel junction can take place only within a very limited volume; and (iv) and (v) are small, since they clearly happen on time scales larger than the tunneling time. Conditions (i) and (ii), however, require a careful discussion.

6.3.1. Electron–phonon interaction

Since the temperature of $T \simeq 30\,\text{mK}$ at which the experiments are being done is much smaller than the mean level spacing, the probability of phonon absorption is negligible, and only emission may take place. The sound velocity in aluminium is $v_{\text{sound}} = 6420\,\text{m/s}$, therefore the wavelength of a phonon associated with relaxation of energy $\varepsilon \sim d = 1\,\text{meV}$ is approximately 5 nm, the same as the system size. In this regime, Agam et al. estimate the photon emission rate to be

$$\Gamma^{\text{el}-\text{ph}} \sim \left(\frac{2}{3}\varepsilon_{\text{F}}\right)^2 \frac{\varepsilon^3 \tau_{\text{el}} d}{2\rho\hbar^4 v_{\text{sound}}^5} \, , \tag{115}$$

where ε_{F} is the Fermi energy (11.7 eV in Al), and ρ is the ion mass density (2.7 g/cm^2 in Al). This rate is that of a clean metal but reduced by a factor of $\tau_{\text{el}}\,d/\hbar$, where τ_{el} is the elastic mean free time [149] which, in ballistic systems, equals the traversal time across the system of an electron at the Fermi level. Assuming ballistic motion this factor is of order 10^{-3}. The resulting relaxation rate for $\varepsilon \simeq d$ is therefore of order $\Gamma_{\varepsilon\simeq d}^{\text{el}-\text{ph}} \approx 10^8\,\text{s}^{-1}$ which is similar to the tunneling rate $\Gamma_{\text{tun}} \approx 6 \times 10^8\,\text{s}^{-1}$ (corresponding to an increase in current of 10^{-10} A at the first current step). Thus, by increasing the resistance of the tunnel junctions, it should be possible to cross over to the near-equilibrium regime shown by the dashed line in Fig. 27(c).

6.3.2. Electron–electron interaction beyond Hartree–Fock

The effect of the electron–electron interaction on the resonance widths measured in a tunneling experiment, or more generally on the lifetime of quasiparticle excitations in a finite, disordered system of interacting electrons, has been elucidated by Altshuler, Gefen, Kamenev and Levitov (AGKL) [64]. Since their conclusions are directly relevant for the interpretation of nonequilibrium tunneling spectra, we briefly review them here.

Consider again tunneling transitions from the N to $(N + 1)$ particle Hilbert spaces, which probe the excitations of the latter. In Section 6.1.2, we took these excitations to be simply particle–hole excitations with respect to the Fermi ground state $|G\rangle_{N+1}$, such as $|p\bar{p}\rangle_{N+1}$, $|q\bar{q}, p\bar{p}\rangle_{N+1}$, etc.; in the nomenclature of AGKL, these are called Hartree–Fock states, $|p0\rangle_{N+1}$ is called a "quasiparticle", $|q0, p\bar{p}\rangle_{N+1}$ two quasiparticles and a quasihole, etc. Now, while these are exact eigenstates of the orthodox model Hamiltonian $H_{\text{D,orth}}^{\text{normal}}$ [Eq. (96)], they will decay in the presence of the more general interaction $H_{\text{el}-\text{el}}$ of Eq. (107), e.g., a quasiparticle will decay into two quasiparticles and a quasihole, etc.

A quasiparticle state with energy ε will thus acquire a lifetime $1/\Gamma_{\varepsilon}^{\text{inel}}$, and in general should be viewed as a wave packet, with packet width $\Gamma_{\varepsilon}^{\text{inel}}$ (we set $\hbar = 1$ in this section), constructed from the exact eigenstates of the full Hamiltonian. In a finite-sized system with a discrete spectrum, a quasiparticle will be well defined if $\Gamma_{\varepsilon}^{\text{inel}} < d$ (note that this condition is more stringent than the one for infinite systems, namely $\Gamma_{\varepsilon}^{\text{inel}} \ll \varepsilon$). Moreover, the quasiparticle wave packet width $\Gamma_{\varepsilon}^{\text{inel}}$ can be identified with the measured width of a resonance in the single-particle tunneling density of states (DOS) probed by an incident tunneling electron, since the mathematical description of a single-electron tunneling process involves projecting a single-particle state onto exact eigenstates of the system [this is made explicit in Eq. (12)]. In principle, in a finite system *every* exact eigenstate should produce an infinitely sharp conductance resonance. However, in reality only a small fraction of the exact eigenstates overlap sufficiently strongly with the one-particle excitations produced by the incoming electron to be detected as strong peaks by a measurement of finite sensitivity. The nature and spacing of these strong peaks depends on the energy ε of the incoming electron; under certain conditions, clarified by AGKL, they group into clusters of combined width $\Gamma_{\varepsilon}^{\text{inel}}$ that can be interpreted as quasiparticle peaks.

The standard golden rule (GR) approach for estimating $\Gamma_{\varepsilon}^{\text{inel}}$ gives $\Gamma_{\varepsilon,\text{GR}}^{\text{bulk}} \sim \varepsilon^2/\varepsilon_{\text{F}}$ for bulk systems [150] and $\Gamma_{\varepsilon,\text{GR}}^{\text{finite}} \approx d(\varepsilon/dg_{\text{dim}})^2$ for finite systems [66], but is valid only when the quasiparticle effectively decays into a *continuum* of final states (otherwise the GR will not give the decay rate, but rather just a first-order perturbation correction to the energy of a given eigenstate). For sufficiently low ε, the accessible final states will not form a continuum; for example, the density of final states for the decay of a quasiparticle into two quasiparticles and a quasihole, namely $\varepsilon^2/2d^3$, is much smaller than that of all many-body states. To analyze how matters change with decreasing ε, AGKL considered the decay of a quasiparticle into two quasiparticles and a quasihole, which in turn can decay into five quasiparticles and three quasiholes, etc., and thus mapped the problem onto one of the delocalization in Fock space, in which the "distance" between two states is measured, roughly speaking, by the number of particle–hole excitations by which they differ. A state that is "localized" or "delocalized" in Fock space is associated with a quasiparticle peak of zero and finite width, respectively. Assuming the dimensionless conductance to be large, $g_{\text{dim}} \gg 1$, AGKL exploited analogies to Anderson localization to study the localization–delocalization crossover quantitatively, and constructed the following picture.

The nature of an excitation with energy ε depends on the relation of ε to the hierarchy $E^{**} < E^* < E_{\text{Thouless}}$ of energy scales, where

$$E^{**} \equiv d\sqrt{g_{\text{dim}}/\ln(g_{\text{dim}})}, \quad E^* \equiv d\sqrt{g_{\text{dim}}}, \quad E_{\text{Thouless}} \equiv dg_{\text{dim}} \ . \tag{116}$$

E^{**} separates the localized from the delocalized regime, and altogether four cases be distinguished:

(i) $\varepsilon < E^{**}$ (localized phase): Here the true many-body excited states are just slightly perturbed Hartree–Fock states, i.e. their overlap with simple particle–hole excitations of the Fermi ground state is very close to unity. In other words, quasiparticles do not decay significantly, and the single-particle DOS will consist of a few δ-function-like resolved peaks, which may have weak satellites due to coupling to many-particle states involving more particle–hole pairs. If the latter are purposefully created by driving the grain into nonequilibrium, the satellites will become strong and each quasiparticle peak will appear as a cluster. Since $g_{\text{dim}} \simeq 5$ for Fig. 27, at least its first few quasiparticle excitations fall into this regime; note that

AGKL's conclusion that these are well described by simple particle–hole states fully justifies using only such states in the analysis in Sections 6.1.2 and 6.1.3.

(ii) $E^{**} < \varepsilon < E^*$ (delocalized phase with non-Lorentzian peaks): A given quasiparticle state is connected to several others far away from it in Fock space, so that the number of satellites in a quasiparticle peak rapidly increases with ε. This causes the quasiparticle peaks in the DOS to have non-Lorentzian shapes, but their widths are still $\ll \Gamma_{\varepsilon,\mathrm{GR}}^{\mathrm{finite}}$.

(iii) $E^* < \varepsilon < E_{\mathrm{Thouless}}$ (delocalized phase with Lorentzian peaks): A given quasiparticle state is connected to so many others that it effectively decays into a continuum, so its decay rate is given by the GR result, $\Gamma_{\varepsilon,\mathrm{GR}}^{\mathrm{finite}} \approx d(\varepsilon/dg_{\mathrm{dim}})^2$. Since this is still $< d$ for $\varepsilon < E_{\mathrm{Thouless}}$, the peaks in the DOS can still be well-resolved, and have a Lorentzian shape.

(iv) $E_{\mathrm{Thouless}} < \varepsilon$ (quasiparticles not defined): In this regime $\Gamma_{\varepsilon,\mathrm{GR}}^{\mathrm{finite}} > d$, hence the quasiparticles are not well defined, and the DOS is a featureless continuum.

Experimental results strikingly reminiscent of this picture have recently been obtained by Davidović and Tinkham [15], see Fig. 30. Fig. 30(b), in particular, seems to show the progression, with increasing energy, of the measured tunneling spectrum through all the crossovers. Estimates
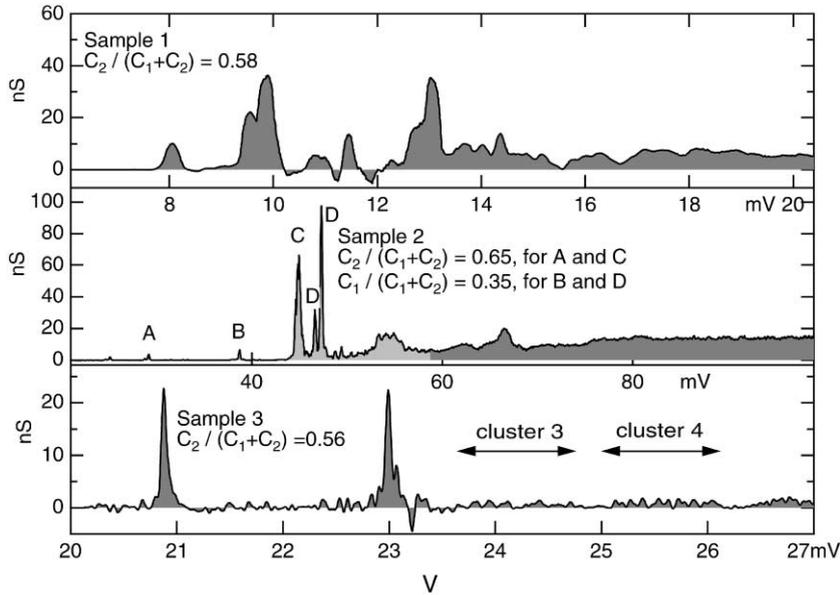

Fig. 30. Excitation spectra of three different nm-scale Au grains measured by Davidović and Tinkham [15], at $T = 30\,\mathrm{mK}$ and $H = 1\,\mathrm{T}$. In order to convert from bias voltage to excitation energy, the V must be multiplied by e times the capacitative conversion factor [cf. Eqs. (30) and (31)] indicated for each graph. Assuming the grains to be hemispherical, the grains parameters were estimated to be as follows for samples 1, 2 and 3, respectively: grain radius (estimated from $C_1 + C_2$ and a capacitance per unit area of $0.05\,\mathrm{aF/nm^2}$): $r = 4.5$, 2.4, and 3 nm; estimated single-electron level spacing (from Eq. (2), using r as input): $\bar{d} = 0.65$, 4.6, and 2.1 meV; measured single-electron level spacing d (taken as distance between the two lowest energy peaks): $d = 1, 7$, and 1.2 meV; Thouless energy (estimated as $E_{\mathrm{Thouless}} = \hbar v_F/6r'$ [Eq. (5) with $a = 3$, cf. footnote 1], where r' is the grain radius estimated from Eq. (2), using the *measured d*-values as input): $E_{\mathrm{Thouless}} = 37, 75$, and 40 meV.

for E_{Thouless}, cited in the caption of Fig. 30, agree within a factor of 2 with the voltage[17] eV at which the measured spectra become continuous. Note that in order to observe the entire regime up to and beyond the Thouless energy, all within the first step of Coulomb blockade, one needs to have $E_{\text{Thouless}} \ll E_C$. In this respect ultrasmall metallic grains, with their huge charging energies, have an edge over disordered semiconducting quantum dots, studies of which had initiated the analysis of the role of E_{Thouless} as the upper scale for the observability of level discreteness [66].

7. Spin–orbit interaction

For several of the normal-state grains investigated so far, the magnetic-field dependence of the excitation spectra showed marked deviations from the simple behavior described in Section 3.1: g factors differing from 2 were measured [7,14,15] (in Au grains, some were as small as 0.28 [15]), and deviations from linear H-dependence, in the form of avoided level crossings, were observed [7,14]. These features were attributed to the spin–orbit interaction, which is the subject of this section. A summary of early work on the spin–orbit interaction in small metallic grains, mainly with regard to thermodynamical properties, may be found in the reviews by Buttet [151], Perenboom et al. [19] and Halperin [20].

The information on spin–orbit scattering obtained from the analysis summarized below is complementary to that obtained from studies of bulk systems. There, the quantity of primary interest is the average scattering rate (Γ_{so}) for an electron, assumed to be initially in a pure spin-up or spin-down state, to be scattered into a *continuum* of states with opposite spin. This rate can be measured in studies of weak localization in disordered metals [152] or of tunneling between thin superconducting films in a parallel magnetic field [99], or it can be related to the effective g factors [153] measured by electron spin resonance. In ultrasmall grains, spin–orbit scattering can be studied at a more detailed level via its effects on individual, discrete eigenstates of the grain, such as the above-mentioned occurrences of anomalous g factors and avoided crossings. Moreover, the effects of mesoscopic fluctuations [55] on phenomena governed by spin–orbit scattering can be observed directly.

7.1. Experimental results for weak spin–orbit interaction

Fig. 31 shows the results of measurements published by Salinas et al. [7,14] on a grain of estimated radius $r \simeq 3$ nm, made from nominally pure Al. It presumably contained an unintended defect or impurity, because its behaviour was similar to that observed in grains intentionally doped with 4% of Au impurities. The following salient features can be identified:

(i) This is an *even-to-odd* excitation spectrum, since all three the lowest peaks in Fig. 31(a) Zeeman-split in an applied magnetic field H.

---

[17] Davidović and Tinkham argued that E_{Thouless} should be compared to the voltage eV itself and not to the corresponding (capacitatively corrected) excitation energy, because the number of possible nonequilibrium excitations is controlled by eV, cf. Eq. (102(ii)′).
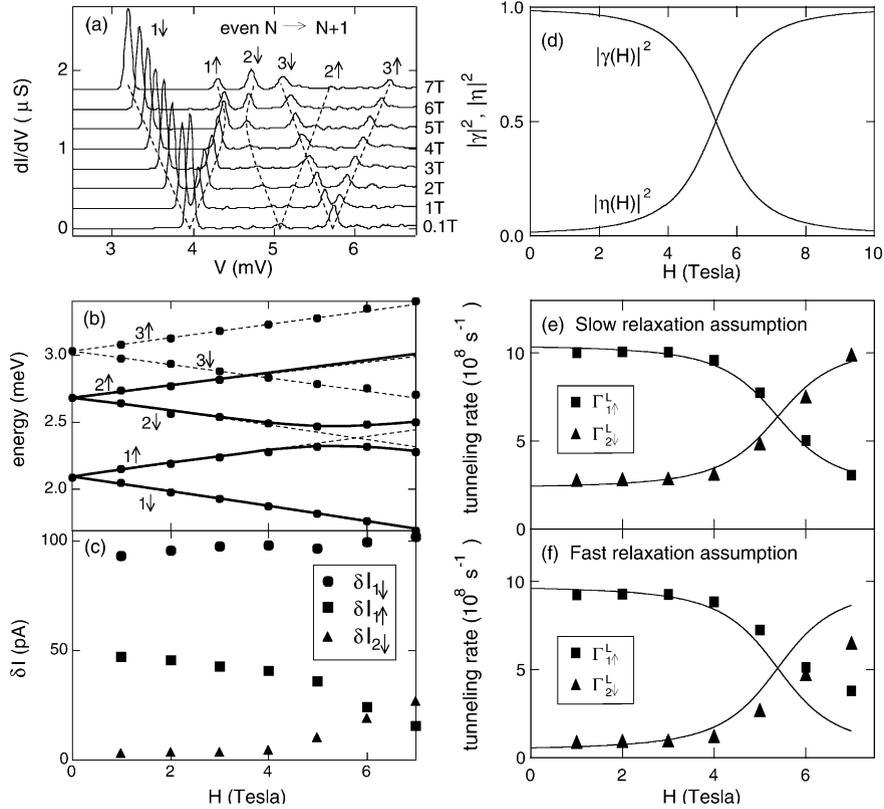
Fig. 31. The effect of the spin–orbit interaction on the excitation spectrum of an Al grain of radius $r \simeq 3$ nm [14]: (a) Differential conductance for a range of applied magnetic fields, at $T = 50$ mK. The curves are offset in dI/dV for visibility. Orbital state no. 2 gives small but visible resonances at low H. Small changes in offset charge (Q_0) occurred between the 0.1 and 1 T scans and between the 6 and 7 T scans, shifting peak positions. The 0.1 and 7 T scans have therefore been shifted along the voltage axis, to give the best fit to a linear dependence for peak 1↓. Dotted lines are guides to the eye. (b) Magnetic-field dependence of the peak positions of (a), which have been converted from voltage to energy by a conversion factor of $eC_L/(C_L + C_R) = e(0.53 \pm 0.01)$, R being the bottleneck barrier here [cf. Eqs. (30)]. Dots represent data points. Dashed lines are extensions of the low-field linear dependence of the energies on H. Solid lines show the result of the spin–orbit model of Section 7.2, describing the avoided crossing between levels $|1\uparrow\rangle$ and $|2\downarrow\rangle$. (c) Areas under the first three conductance peaks, which are equal to the current steps δI at the corresponding resonances. Note the exchange in weight between $\delta I_{1\uparrow}$ and $\delta I_{2\downarrow}$, which occurs near the avoided crossing between the eigenstates $|1\uparrow\rangle$ and $|2\downarrow\rangle$ in (b). (d) Magnetic-field dependence of the coefficients $\chi(H)$ and $\eta(H)$ in Eq. (125), which relate the eigenstates $|1\uparrow\rangle$ and $|2\downarrow\rangle$ to the basis states $|1\uparrow\rangle_{\tilde{p}}$ and $|2\downarrow\rangle_{\tilde{p}}$. (e) and (f): Markers: Tunneling rates $\Gamma^L_{1\uparrow}$ and $\Gamma^L_{2\downarrow}$ for transitions out of the grain's eigenstates $|1\uparrow\rangle$ and $|2\downarrow\rangle$, extracted from the measured current as discussed in text, assuming that the relaxation rate of nonequilibrium excitations of the grain is either smaller (e) or larger (f) than the tunneling rates. Lines: fits using the predictions of the spin–orbit model.

(ii) *Orbital effects are negligible:* For each pair of peaks the splitting is symmetric about their $H = 0$ resonance energy, and the average of their resonance energies shows hardly any H-dependence (ftting the latter to a linear form, $g^{\mathrm{ave}} \mu_B H$, yields $g^{\mathrm{ave}} < 0.1$, which is negligible). This was to be expected, since the largest value of H used here (7T) is much smaller than the

scale H_{orb} [Eq. (36)] beyond which orbital effects become significant. The latter will thus be neglected below.

(iii) The *effective g_j^{eff} factor* (for low fields) of each Zeeman-split resonance is defined by writing its Zeeman splitting, to *linear* order in H, as $\delta E_j = g_j^{eff} \mu_B H$ [cf. Eq. (119)]. Then Fig. 31(b) gives $g_1^{eff} = 1.84 \pm 0.03$, $g_2^{eff} = 1.68 \pm 0.08$, and $g_3^{eff} = 1.76 \pm 0.05$.

(iv) An *avoided crossing* occurs between resonances 1↑ and 2↓ in Figs. 31(a) and (b);

(v) the *areas* under the corresponding peaks 1↑ and 2↓ in Fig. 31(a), plotted in Fig. 31(c), are *strongly H-dependent*: the smaller one (2↓) grows at the expense of the larger one (1↑) in the avoided-crossing region, while their sum remains approximately constant.

Properties (iii)–(v) contrast strongly with the behaviour observed for pure Al grains (cf. Fig. 10); the latter typically have g-factors close to $g^{pure} \equiv 2$ (see also footnote 10 in Section 4.2) that vary only weakly from level to level (if at all), their spectra show no clear avoided crossings, and they have essentially H-independent peak areas.

7.2. A simple model for the spin–orbit interaction

Salinas et al. [14] have attributed anomalies (iii)–(v) to spin–orbit scattering due to some defect in the grain, and have constructed a simple model within which their measurements could be analyzed quantitatively. Using a somewhat more general notation than they did and guided by the reviews of Refs. [19,20], we shall consider a Hamiltonian $\hat{H}_{tot} = \hat{H}_0 + \hat{H}_{so}$, to be called the *spin–orbit model*, of the following simple form:

$$\hat{H}_0 = \sum_{j\sigma} \varepsilon_{j\sigma,0} c_{j\sigma,0}^\dagger c_{j\sigma,0}, \quad \varepsilon_{j\sigma,0} = \varepsilon_{j,0} + \tfrac{1}{2}\sigma g^{pure}\mu_B H , \tag{117}$$

$$\hat{H}_{so} = \sum_{ij\sigma} v_{ij}^\sigma c_{i\sigma,0}^\dagger c_{j\bar{\sigma},0}, \quad \text{with } v_{ij}^\sigma = (v_{ji}^{\bar{\sigma}})^* = -v_{ji}^\sigma, \quad v_{jj}^\sigma \equiv 0 . \tag{118}$$

Since time-reversal symmetry guarantees each orbital level to be two-fold degenerate in the absence of an applied magnetic field ($H = 0$), we have adopted a single-particle basis of pairs of time-reversed states $|j\sigma\rangle_0$, labeled by a discrete orbital index j and a spin index $\sigma = (\uparrow, \downarrow) = (+, -)$. The orbital eigenenergies $\varepsilon_{j,0}$ are meant to include the effects of all spin-independent forces and interactions, and are taken to have a mean spacing of d. A magnetic field Zeeman-splits these levels by $\tfrac{1}{2}\sigma g^{pure}\mu_B H$. In principle, it can also produce orbital diamagnetic effects, which could cause $\varepsilon_{j,0}$ to be H-dependent and, via spin–orbit coupling, contribute to reducing the effective g factors [154,55]. We shall neglect these here, however, in accordance with property (ii). $\hat{H}_{so}$ describes the spin–orbit interaction and is taken to couple only states of *opposite* spin, σ and $\bar{\sigma} \equiv -\sigma$. The conditions specified in Eq. (118) for the matrix elements v_{ij}^σ ensure that $\hat{H}_{so}$ is hermitian and invariant under time reversal [which maps $c_{j\sigma}^\dagger \to \sigma c_{j\bar{\sigma}}^\dagger$ and $v_{ij}^\sigma \to (v_{ij}^\sigma)^*$], i.e. that $\hat{H}_{so}$ represents a symplectic matrix, as required on general grounds for the spin–orbit interaction [19,20].

The average spin–orbit scattering rate in this model, expressed in terms of the average (over i, j) of the matrix elements in $\hat{H}_{so}$, is $\Gamma_{so} = 2\pi\overline{|v_{ij}^\sigma|^2}/(\hbar d)$. The effects of spin–orbit scattering on an individual eigenstate of $\hat{H}_{tot}$ will be "weak" or "strong" depending on whether the dimensionless parameter $\lambda_{so} \equiv \hbar\Gamma_{so}/d$ is $\ll 1$ or $\gg 1$, respectively. Since the rate Γ_{so} is an intensive (volume-independent) quantity (implying that v_{ij}^σ must scale as $Vol^{-1/2}$), it should be possible to reach the

regime of weak spin–orbit scattering in grains with sufficiently large d. Indeed, the grain of Fig. 31 is an example of this case, since the minimum splitting at the avoided crossing, which is a measure of $\overline{|v^\sigma_{ij}|}$, is much less than the mean level separation, d.

Since $\hat{H}_{\text{tot}}$ is quadratic, it can in principle be brought into the diagonal form $\hat{H}_{\text{tot}} = \sum_{j\sigma}\varepsilon_{j\sigma}c^\dagger_{j\sigma}c_{j\sigma}$ by a linear transformation. The time-reversal symmetry of $\hat{H}_{\text{tot}}$ for $H = 0$ implies that the exact eigenstates again come in time-reversed pairs, say $|j\sigma\rangle$, degenerate at $H = 0$. Since this degeneracy is lifted by the magnetic field,

$$\varepsilon_{j\pm} = \varepsilon_j \pm \tfrac{1}{2}g^{\text{eff}}_j\mu_{\text{B}}H + \mathcal{O}(H^2) \,, \tag{119}$$

the labels $(+ , -) = (\uparrow, \downarrow)$ identify states with "predominantly spin-up or down" character (though no longer eigenstates of $\hat{\sigma}_z$). The linear term in Eq. (119) defines the effective g factor g^{eff}_j for the pair of states $|j\sigma\rangle$, which, being superpositions of spin-up and spin-down states, will have[18] $g^{\text{eff}}_j < g^{\text{pure}}$. For weak spin–orbit scattering ($\lambda_{\text{so}} \ll 1$), this reduction can be calculated perturbatively: using the lowest-order perturbative approximations (designated by the subscript p) to $|j\sigma\rangle$ and $\varepsilon_{j\sigma}$ to evaluate g^{eff}_j, one finds [20]

$$|j\sigma\rangle_{\text{p}} = \left[1 - \frac{1}{2}\sum_i \frac{|v^{\bar\sigma}_{ij}|^2}{(\varepsilon_{j\sigma,0} - \varepsilon_{i\bar\sigma,0})^2}\right]|j\sigma\rangle_0 + \sum_i \frac{v^{\bar\sigma}_{ij}}{\varepsilon_{j\sigma,0} - \varepsilon_{i\bar\sigma,0}}|i\bar\sigma\rangle_0 \,, \tag{120}$$

$$\varepsilon_{j\sigma,\text{p}} = \varepsilon_{j\sigma,0} + \sum_i \frac{|v^{\bar\sigma}_{ij}|^2}{\varepsilon_{j\sigma,0} - \varepsilon_{i\bar\sigma,0}} \,, \tag{121}$$

$$g^{\text{eff}}_j = g^{\text{pure}}\left[1 - \sum_i \frac{2|v^{\bar\sigma}_{ij}|^2}{(\varepsilon_{j,0} - \varepsilon_{i,0})^2}\right] \,. \tag{122}$$

(Eq. (122) follows by expanding (121) to linear order in H.) Since the bare energies $\varepsilon_{j,0}$ and matrix elements v^σ_{ij} are subject to mesoscopic fluctuations, g^{eff}_j will be too. This is in accord with the observation of *j-dependent*, *reduced* g^{eff}_j values cited in property (iii) above.

7.3. Detailed analysis of an avoided crossing

In order to analyse the avoided crossing in Figs. 31(a) and (b) quantitatively, Salinas et al. truncated the problem to an effective four-state system, associated with the four lowest levels in Fig. 31(b). The idea is to treat the effect of all other levels on the lowest four using perturbative expressions evaluated in the low-field limit, while treating the spin–orbit interaction among the lowest four fully, for arbitrary H. To this end, one introduces four "bare" basis states for the truncated problem $|1\!\downarrow\rangle_{\tilde{\text{p}}}$, $|1\!\uparrow\rangle_{\tilde{\text{p}}}$, $|2\!\downarrow\rangle_{\tilde{\text{p}}}$, $|2\!\uparrow\rangle_{\tilde{\text{p}}}$, defined by the $H = 0$ limit of Eqs. (120), with "bare" eigenenergies $\varepsilon_{j\sigma,\tilde{\text{p}}}$ given by (121), expanded to linear order in H. In order not to overcount the interaction among these states, the $\sum_i$ sums in Eqs. (120) and (121) should now *exclude* $i = 1$ and 2 (the tilde on the label $\tilde{\text{p}}$ indicates this restriction). The effective Hamiltonian for the truncated

---

[18] In situations where orbital effects are important, such as semiconductor quantum dots, they may reverse this trend and give $g^{\text{eff}}_j > g^{\text{pure}}$, see [55].

system therefore is[19]

$$\hat{H}_{\text{tot}}^{\tilde{p}} = \sum_{j=1,2} \sum_{\sigma} \varepsilon_{j\sigma,\tilde{p}} c_{j\sigma,\tilde{p}}^{\dagger} c_{j\sigma,\tilde{p}} + \sum_{\sigma} v_{12}^{\sigma} c_{1\sigma,\tilde{p}}^{\dagger} c_{2\bar{\sigma},\tilde{p}} + \text{h.c.} , \tag{123}$$

$$\varepsilon_{j\sigma,\tilde{p}} = \varepsilon_{j,\tilde{p}} + (g_j^{\text{ave}} + \tfrac{1}{2}\sigma \tilde{g}_j^{\text{eff}})\mu_{\text{B}} H, \qquad \tilde{g}_j^{\text{eff}} \equiv g_j^{\text{eff}} + \frac{g^{\text{pure}} 2|v_{12}^{\downarrow}|^2}{(\varepsilon_{2,0} - \varepsilon_{1,0})^2} . \tag{124}$$

Note that $\tilde{g}_j^{\text{eff}}$ is the contribution of the $i \neq 1,2$ levels to the total effective g_j^{eff} factor of state $|j\sigma\rangle_{\text{p}}$. Moreover, simply for convenience in the fitting, a linear term $g_j^{\text{ave}}\mu_{\text{B}} H$ has been allowed too, to model any shift in the average energy of the Zeeman-split pairs [the fits turn out to give $g_j^{\text{ave}} \approx 0$, cf. property (ii)].

It is now straightforward to solve this truncated problem explicitly and to find the eigenstates $|1\downarrow\rangle$, $|1\uparrow\rangle$, $|2\downarrow\rangle$, $|2\uparrow\rangle$ of $\hat{H}_{\text{tot}}^{\tilde{p}}$. Fitting the corresponding eigenenergies to the lowest four lines of Fig. 31(b), the level repulsion between $|1\uparrow\rangle$, $|2\downarrow\rangle$, is reproduced and an excellent fit obtained for $|v_{12}^{\downarrow}| = 73 \pm 4\,\mu\text{eV}$, $g_1^{\text{ave}} = -0.03 \pm 0.04$, $g_2^{\text{ave}} = -0.10 \pm 0.06$, $\tilde{g}_1^{\text{eff}} = 1.90 \pm 0.04$, and $\tilde{g}_2^{\text{eff}} = 1.74 \pm 0.04$. This yields a difference of $\tilde{g}_j^{\text{eff}} - g_j^{\text{eff}} = 0.06$ for $j = 1,2$, which is consistent with Eq. (124), since $4|v_{12}^{\downarrow}|^2/(\varepsilon_{2,0} - \varepsilon_{1,0})^2 = 0.06$. This amount is the contribution to g_j^{eff} of spin–orbit coupling between states $|1\sigma\rangle$ and $|2\bar{\sigma}\rangle$, and can be used as estimate for $4\overline{|v_{ij}^{\sigma}|^2}/d^2 = (2/\pi)\lambda_{\text{so}}$; the result, $\lambda_{\text{so}} \simeq 0.09$, confirms that we are in the regime of weak spin–orbit scattering. Since the amount of 0.06 accounts for only 40% (or 20%) of the total reduction $g^{\text{pure}} - g_j^{\text{eff}}$ of 0.16 (or 0.32) for orbital state 1 (or 2), the remainder must come from spin–orbit coupling to other ($i \neq 1,2$) states, whose contribution is therefore very significant.

The eigenstates of $\hat{H}_{\text{tot}}^{\tilde{p}}$ have the general form

$$|1\uparrow\rangle = \gamma|1\uparrow\rangle_{\tilde{p}} + \eta|2\downarrow\rangle_{\tilde{p}}, \quad |2\downarrow\rangle = -\eta^*|1\uparrow\rangle_{\tilde{p}} + \gamma^*|2\downarrow\rangle_{\tilde{p}} \tag{125}$$

with $|\gamma|^2 + |\eta|^2 = 1$ (and similarly for $|1\downarrow\rangle$, $|2\uparrow\rangle$). The coefficients $\gamma(H)$ and $\eta(H)$ are fully determined by the parameters obtained from the above-mentioned fit, and are plotted in Fig. 31(d). Evidently, their relative magnitudes are "interchanged" as H is tuned through the avoided crossing, and the character of, e.g., $|1\uparrow\rangle$ changes from being predominantly $|1\uparrow\rangle_{\tilde{p}}$ for small H to predominantly $|2\downarrow\rangle_{\tilde{p}}$ for large H.

This feature is also the reason for property (v). Salinas et al. were able to confirm this quantitatively by analyzing the H-dependence of, say, the rates $\Gamma_{1\uparrow}^{\text{L}} \equiv \Gamma_{G,1\uparrow}^{\text{L}-}$ and $\Gamma_{2\downarrow}^{\text{L}} \equiv \Gamma_{G,2\downarrow}^{\text{L}-}$ for a tunneling-off transition via barrier L from one of the odd eigenstates $|1\uparrow\rangle_{N+1}$ or $|2\downarrow\rangle_{N+1}$ into the even ground state $|G\rangle_N$. According to Eq. (24), these rates have the general form[20]

$$\Gamma_{1\uparrow}^{\text{L}} = \sum_{l\sigma} \Gamma_{l\sigma}^{\text{L}} |\langle G|c_{l\sigma}|1\uparrow\rangle|^2 \simeq |\gamma|^2 \tilde{\Gamma}_{1\uparrow}^{\text{L}} + |\eta|^2 \tilde{\Gamma}_{2\downarrow}^{\text{L}} , \tag{126}$$

---

[19] When deriving $\hat{H}_{\text{tot}}^{\tilde{p}}$ by reexpressing $\hat{H}_{\text{tot}}$ in terms of $c_{j\sigma,\tilde{p}}$-operators, off-diagonal terms beyond linear order in v_{ij}^{σ} were dropped, since they contain further factors of the small quantity $v_{ij}^{\sigma}/(\varepsilon_{j\sigma,0} - \varepsilon_{i\bar{\sigma},0})$, with $i = 1,2$ and $j \neq 1,2$.

[20] The rightmost expressions in Eqs. (126) and (127) were obtained by neglecting terms containing factors of the form $_{\tilde{p}}\langle j\sigma|c_{l\sigma}^{\dagger}|G\rangle$. These are not completely zero (because, by Eq. (120), $|j\sigma\rangle_{\tilde{p}}$ is not a pure spin-σ state and contains some spin $\bar{\sigma}$ components), but are smaller than those kept by a factor of order $v_{ij}^{\sigma}/(\varepsilon_{j\sigma,0} - \varepsilon_{i\bar{\sigma},0})$.

$$\Gamma_{2\downarrow}^{\mathrm{L}} = \sum_{l\sigma} \Gamma_{l\sigma}^{\mathrm{L}} |\langle G|c_{l\sigma}|2\downarrow\rangle|^2 \simeq |\eta|^2 \tilde{\Gamma}_{1\uparrow}^{\mathrm{L}} + |\gamma|^2 \tilde{\Gamma}_{2\downarrow}^{\mathrm{L}} \ , \tag{127}$$

$$\tilde{\Gamma}_{1\uparrow}^{\mathrm{L}} \equiv \sum_{l} \Gamma_{l\uparrow}^{\mathrm{L}} |\langle G|c_{l\uparrow}|1\uparrow\rangle_{\tilde{\mathrm{p}}}|^2, \quad \tilde{\Gamma}_{2\downarrow}^{\mathrm{L}} \equiv \sum_{l} \Gamma_{l\downarrow}^{\mathrm{L}} |\langle G|c_{l\downarrow}|2\downarrow\rangle_{\tilde{\mathrm{p}}}|^2 \ . \tag{128}$$

The H-dependence of $\Gamma_{1\uparrow}^{\mathrm{L}}$ and $\Gamma_{2\downarrow}^{\mathrm{L}}$ is therefore contained completely in the factors $|\gamma|^2$ and $|\eta|^2$, since $\tilde{\Gamma}_{1\uparrow}^{\mathrm{L}}$ and $\tilde{\Gamma}_{2\downarrow}^{\mathrm{L}}$ are H-independent (and fully determined by $H = 0$ data). Salinas et al. compared the predictions of the spin–orbit model for $\Gamma_{1\uparrow}^{\mathrm{L}}$ and $\Gamma_{2\downarrow}^{\mathrm{L}}$ [solid lines in Figs. 31(e) and (f)] with estimates [markers in Figs. 31 (e) and (f)] that were extracted from the measured conductance peak areas mentioned in property (v). These estimates involved inverting Eq. (18) for the current in order to express the rates $\Gamma_{1\uparrow}^{\mathrm{L}}$ and $\Gamma_{2\downarrow}^{\mathrm{L}}$ occurring therein in terms of the (measured) current and the probabilities P_α that the grain is in eigenstate $|\alpha\rangle$, and determining the P_α by solving a corresponding master equation [Eq. (17)]. The analysis is complicated by the fact *all* accessible states $|\alpha\rangle$ must be considered, not only $|1\uparrow\rangle$ and $|2\downarrow\rangle$, including nonequilibrium excitations[21] that can be created when the bias voltage $|eV|$ is sufficiently large. To keep the model tractable nevertheless, some simplifying assumptions were made about unknown model parameters, e.g. that the ratios $\Gamma_\alpha^{\mathrm{L}}/\Gamma_\alpha^{\mathrm{R}}$ are the same for all α (see Ref. [14] for details). The results depend, in particular, on whether the relaxation rates Γ^{inel} for nonequilibrium excitations are assumed smaller or larger than the average tunneling rate Γ_{tun}. The "slow-relaxation" assumption [$\Gamma^{\mathrm{inel}} \ll \Gamma_{\mathrm{tun}}$, Fig. 31(e)] yields good agreement between the prediction (solid lines) of the spin–orbit model and the experimental estimates (markers). For the "fast-relaxation" assumption [$\Gamma^{\mathrm{inel}} \gg \Gamma_{\mathrm{tun}}$, Fig. 31(f)] the agreement is not so good, though qualitatively still reasonable. The fact that the slow-relaxation assumption works better is in accord with estimates [11] (see Section 6.3) that the energy relaxation rate due to phonons is $\Gamma^{\mathrm{inel}} \approx 10^8\,\mathrm{s}^{-1}$, an order of magnitude less than the estimates of $\Gamma_{\mathrm{tun}} \simeq 10^9\,\mathrm{s}^{-1}$ for the tunneling rates extracted from the detailed analysis.

7.4. Distributions for the effective g factors

The statistics of the fluctuations in g_j^{eff} have recently been investigated theoretically by Matveev et al. [55], and independently by Brouwer et al. [56], who studied the dependence of g_j^{eff} on the direction of the applied field. In this section, we summarize those aspects of these results that are relevant for ultrasmall metallic grains.

7.4.1. Distribution of g_j^{eff} for a random field direction

Matveev et al. [55] studied the statistics of g_j^{eff}, as defined by Eq. (119), for a magnetic field applied in a random direction, in the two limits that the spin–orbit interaction is weak or strong. Their results are as follows:

(i) For *weak spin–orbit scattering* ($\lambda_{\mathrm{so}} \ll 1$), the average $\langle\!\langle g\rangle\!\rangle$ of the effective g_j^{eff} factors, averaged either over an ensemble of grains or over different levels of a single grain, can be calculated from the perturbative expression (122). MGL's calculation corresponds to replacing $|v_{ij}^\sigma|^2$ by $\lambda_{\mathrm{so}}d^2/2\pi$ and performing the $\sum_i$ sum by assuming the bare energy levels $\varepsilon_{j,0}$ to obey Wigner–Dyson statistics in

---

[21] The shifting of single-particle excitation energies due to nonequilibrium excitations discussed in Section 6.1.3 could not be resolved for the present grain, however.
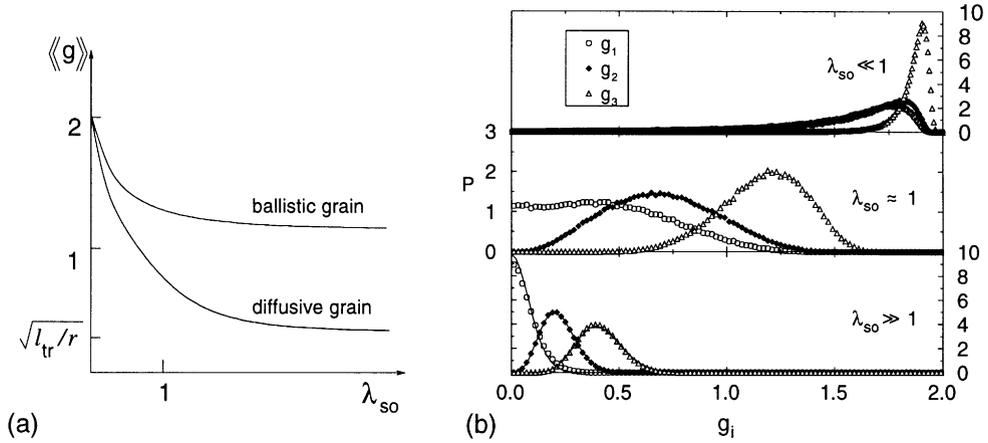
Fig. 32. Properties of ensemble-averaged g factors of ultrasmall metallic grains, as function of the strength $\lambda_{so} = \hbar\Gamma_{so}/d$ of the spin–orbit interaction. (a) Sketch [55] of $\langle\!\langle g\rangle\!\rangle$ as function of λ_{so}, for a magnetic field applied in a random direction; the limits $\lambda_{so} \ll 1$ and $\gg 1$ illustrate Eqs. (129) and (131), respectively. (b) Distributions of the principal g factors g_1, g_2, g_3 for weak, intermediate and strong spin–orbit coupling [56]. The data points were obtained from numerical simulations of a random-matrix version of the spin–orbit model of Section 7.2 (see [56] for details); the parameter λ_{so} corresponds to λ^2 there, up to numerical coefficients).

the orthogonal ensemble. This leads to the following result, in which the logarithm arises due to level repulsion:[22]

$$\langle\!\langle g\rangle\!\rangle = g^{\text{pure}}\left[1 - \frac{\pi}{12}\lambda_{so}\ln\frac{1}{\lambda_{so}}\right], \tag{129}$$

(ii) For *strong spin–orbit scattering* ($\lambda_{so} \gg 1$), Eq. (122) is not an adequate starting point. Instead, MGL calculated the linear term in Eq. (119) perturbatively in the magnetic field H, exploiting general properties of the exact $H = 0$ eigenstates $|j\sigma\rangle$, which were treated using random matrix theory for an ensemble of symplectic matrices. They found the distribution $P(g)$ of g_i^{eff} factors to be given by

$$P(g) = 3\sqrt{\frac{6}{\pi}}\frac{g^2}{\langle\!\langle g^2\rangle\!\rangle^{3/2}}\exp\left(-\frac{3g^2}{2\langle\!\langle g^2\rangle\!\rangle}\right), \quad \langle\!\langle g\rangle\!\rangle = (8/3\pi)^{1/2}\langle\!\langle g^2\rangle\!\rangle^{1/2}, \tag{130}$$

$$\langle\!\langle g^2\rangle\!\rangle = \frac{6}{\pi}\frac{1}{\lambda_{so}} + b\frac{l_{tr}}{r}. \tag{131}$$

Here b is a geometrical constant of order unity and r is the system size (e.g. radius for a spherical or hemispherical grain); for a diffusive grain, $l_{tr} = 3D_{\text{diff}}/v_F$ is the transport mean-free path, and for a ballistic grain, l_{tr} should, roughly speaking, be replaced by r. Fig. 32(a) schematically illustrates the behavior of $\langle\!\langle g\rangle\!\rangle$ as function of λ_{so}. The parametric dependencies of the two contributions to

---

[22] Level repulsion suppresses the probability p_0 of two levels being very close by a factor $p_0(\varepsilon_{i,0} - \varepsilon_{j,0}) = \pi^2|\varepsilon_i - \varepsilon_j|/6d^2$; the logarithmic divergence in the $\sum_j$ sum is cut off above by d, and below at the energy scale $|\varepsilon_{i,0} - \varepsilon_{j,0}| \sim |v_{ij}^\sigma| \propto d\sqrt{\lambda_{so}}$, because of the additional level repulsion caused by the weak spin–orbit coupling.

$\langle\!\langle g^2\rangle\!\rangle$ in Eq. (131) can be interpreted intuitively as follows [55]: the linear-in-H contribution to the energies $\varepsilon_{j\pm}$ of Eq. (119) of the time-reversed states $|j\pm\rangle$ is given by $\langle j\pm|\hat{M}_z|j\pm\rangle\mu_{\mathrm{B}}H\equiv\pm\langle\hat{M}_z\rangle_j\mu_{\mathrm{B}}H$, where $\hat{M}_z=\hat{l}_z+g^{\mathrm{pure}}\hat{s}_z$ is the total magnetic moment, and $\hat{l}_z$ and $\hat{s}_z$ are the orbital angular momentum and spin, respectively (all in units of $\hbar$). Therefore, by Eq. (119), $g_j^{\mathrm{eff}}=2\langle\hat{M}_z\rangle_j$. Now, the angular momentum of an electron traversing the "closed trajectory" corresponding to a discrete quantum level can be estimated as the typical (directed) area A_{typ} covered by its trajectory divided by the period $\hbar/d$ of its motion, $\langle\hat{l}_z\rangle\approx mA_{\mathrm{typ}}d/\hbar^2$, and its spin as $\langle\hat{s}_z\rangle\approx 1/\sqrt{N}$, where $N\approx(\hbar/d)\Gamma_{\mathrm{so}}$ is the number of spin-flips due to spin–orbit scattering which it undergoes during one period. It follows that $\langle\hat{s}_z\rangle^2\approx 1/\lambda_{\mathrm{so}}$, which corresponds to the first term of Eq. (131). Moreover, using $A_{\mathrm{typ}}\approx r^2\sqrt{g_{\mathrm{dim}}}$ (cf. Section 3.2) and Eqs. (4) and (5) for the dimensionless conductance g_{dim}, one finds that $\langle\hat{l}_z\rangle^2\approx l_{\mathrm{tr}}/r$ [or $\mathcal{O}(1)$] for the diffusive (or ballistic) case, which corresponds to the second term of Eq. (131).

Note that, in contrast to the case of weak spin–orbit scatterring considered in most of this section, the orbital diamagnetic contribution, though small, cannot be neglected for a grain with $\lambda_{\mathrm{so}}\gg 1$, since the spin contribution to $\langle\!\langle g\rangle\!\rangle^2$ is small as $1/\lambda_{\mathrm{so}}$. MGL suggested that for such a grain the parametrically small factor l_{tr}/r distinguishing the diffusive from the ballistic case [as illustrated in Fig. 32(a)] can be used to judge which of the two cases applies, depending on whether its tunneling resonances are measured to have effective g factors of $g_j^{\mathrm{eff}}\ll 1$ or $1\lesssim g_j^{\mathrm{eff}}\leq 2$, respectively.

Note also that if one defines the correlation field (H_{orb}) at which orbital diamagnetism becomes dominant by equating the orbital level splitting to the mean level spacing, $2\langle\hat{l}_z\rangle_j\mu_{\mathrm{B}}H_{\mathrm{orb}}\approx d$, and substitutes $\langle\hat{l}_z\rangle\approx mA_{\mathrm{typ}}d/\hbar^2$, one arrives at $H_{\mathrm{orb}}\approx\Phi_0/(\pi A_{\mathrm{typ}})$, which is (up to a factor of π) the criterion used in Section 3.2 to derive Eq. (36). If, more formally, H_{orb} is associated with the field at which the crossover between the symplectic and unitary ensembles, driven by the orbital effects of the magnetic field, is complete [154], the result is again Eq. (36).

7.4.2. Distribution for anisotropic tensor of g factors

Very recently, Brouwer, Waintal and Halperin (BWH) [56] pointed out that since spin–orbit coupling provides a mechanism for the spin to "notice" anisotropies in the orbital wavefunctions, the splitting of a Kramers doublet $\varepsilon_{j\sigma}$, and hence the effective g_j^{eff} factor of a given conductance resonance, should measurably depend on the *direction* of the applied magnetic field $\vec{H}$. In general, the splitting of the Kramers doublet has the form [155]

$$(\varepsilon_{j+}-\varepsilon_{j-})^2=\mu_{\mathrm{B}}^2\vec{H}\cdot\mathscr{G}_j\cdot\vec{H}=\mu_{\mathrm{B}}^2(g_1^2H_1^2+g_2^2H_2^2+g_3^2H_3^2)\ . \tag{132}$$

Here ($\mathscr{G}_j$) is a 3×3 tensor, the g_a^2 ($a=1,2,3$) are its eigenvalues, called "principal g factors" of level j, and the H_a are the components of $\vec{H}$ along the level's "principal axes" (defined as the coordinate axes which diagonalize $\mathscr{G}_j$). $\mathscr{G}_j$ is isotropic (all three $g_a=2$) only in the absence of spin–orbit scattering. In its presence, it becomes anisotropic, is subject to mesoscopic fluctuations, and for large λ_{so} decreases as $1/\lambda_{\mathrm{so}}$. BWH calculated the distribution $P(g_1,g_2,g_3)$ with respect to an ensemble of small metallic grains of roughly equal size, neglecting orbital effects (i.e. assuming $\langle\hat{l}_z\rangle=0$, as applicable for a diffusive grain with $l_{\mathrm{tr}}/r\ll 1$). For sufficiently strong spin–orbit scattering ($\lambda_{\mathrm{so}}\gg 1$), they found it to be of the form

$$P(g_1,g_2,g_3)\propto\prod_{a<b}|g_a^2-g_b^2|\prod_a \mathrm{e}^{-3g_a^2/2\langle\!\langle\bar{g}^2\rangle\!\rangle}\ . \tag{133}$$

Here $\bar{g}^2 = \frac{1}{3}(g_1^2 + g_2^2 + g_3^2)$ is the average of $(\varepsilon_{j+} - \varepsilon_{j-})^2/(\mu_{\rm B}|H|)^2$ over all directions of $\vec{H}$, and its average over the ensemble of grains, $\langle\!\langle \bar{g}^2 \rangle\!\rangle$, is $\propto 1/\lambda_{\rm so}$ [in agreement with Eq. (131)]. Importantly, the factor $|g_a^2 - g_b^2|$ in Eq. (133) suppresses the probability for the $\mathcal{G}_j$ tensor to be isotropic (i.e. g_a's all equal), and hence *favors an anisotropic response* of a given level $\varepsilon_{j\sigma}$ to an applied magnetic field. Fig. 32(b), which shows the distributions for the g_a's for weak, intermediate and strong spin–orbit coupling [56], illustrates both how increasing $\lambda_{\rm so}$ leads to smaller g_a's, and the tendency for g_1, g_2 and g_3 to be unequal. By changing variables from g_1, g_2 and g_3 (with the convention that $g_1 < g_2 < g_3$) to $\bar{g}$, $r_{12} = |g_1/g_2|$ and $r_{23} = |g_2/g_3|$, in terms of which Eq. (133) becomes

$$P(\bar{g}, r_{12}, r_{23}) \propto \frac{r_{23}^3(1 - r_{23}^2)(1 - r_{23}^2 r_{12}^2)(1 - r_{12}^2)}{(1 + r_{23}^2 + r_{23}^2 r_{12}^2)^{9/2}} \bar{g}^8 e^{-9\bar{g}^2/2\langle\!\langle \bar{g}^2 \rangle\!\rangle} \,, \tag{134}$$

the anisotropy is seen to be rather strong: the typical value for r_{12} (or r_{23}) is of order 1/3 (or 1/2), so if the direction of $\vec{H}$ is varied arbitrarily, changes in $g_j^{\rm eff}$ by a factor of 8–10 can be expected (independently of the actual value of $\lambda_{\rm so}$, as long as it is $\gg 1$).

Eq. (133) can be used to recover the results of MGL [55] by associating their g with $[(\mathcal{G}_j)_{zz}]^{1/2}$, the $g_j^{\rm eff}$-factor for a magnetic field in the z-direction (which is a random direction with respect to the grain's principal axes). For g thus defined, the resulting distribution $P(g)$ is found to be given by Eq. (130).

7.5. Experimental results for strong spin–orbit interaction

Experimentally, three reports of grains in the regime $\lambda_{\rm so} \gg 1$ have been published to date, all involving Au, whose large atomic number leads to a much stronger spin–orbit interaction than for Al:

(1) Salinas et al. [14] observed values of $g_j^{\rm eff}$ in the range 0.5–0.8, as well as the occurrence of avoided crossings, in large Al grains doped with 4% of Au impurities.
(2) Davidović and Tinkham [15] observed $g_j^{\rm eff} = 0.28$ and 0.44 in two Au grains whose estimated radii (assuming hemispherical shapes) were 4.5 and 3 nm, respectively (these grains correspond to samples 1 and 3 in Fig. 30).
(3) Davidović and Tinkham [16] observed $g_j^{\rm eff}$ values between 0.2 and 0.3 in an Au grain whose estimated radius (assuming hemispherical shape) was 1.5 nm.

The grain of case (3) had the further interesting feature that the measured level spacing was very much larger than the free-electron estimate, and that the spacing between three subsequent levels *increased* with applied field in the large-H regime, in a way reminiscent of a spin multiplet. Davidović and Tinkham have suggested the possibility that this might perhaps reflect a ground state having a *total spin s larger than* 1/2, favored by the Coulomb interaction: if, e.g., the spacing between two orbital energy levels at $\varepsilon_{\rm F}$ is smaller than the Coulomb interaction, say u, between two electrons in the same orbital state, then, in analogy to Hund's first rule in atomic physics, the ground state would be a $s = 1$ spin triplet in which both orbital states are singly occupied by electrons with parallel spins, because this allows their coordinate wavefunctions to be antisymmetrized, reducing their Coulomb interaction energy. This is well known to occur in cylindrically shaped semiconductor quantum dots [156], and was predicted to be possible in metallic grains too by Brouwer et al. [157]: they calculated the statistical distribution $P(s)$ of the ground state spin s for

an ensemble of small normal metallic grains without spin–orbit coupling, and found an appreciable probability for values other than 0 or 1/2 already for interaction strengths well below the Stoner criterion, i.e. u well below d (see also [158] for similar work for quantum dots). For example, already at the quite modest interaction strength of $u/d \simeq 0.4$, a ground state spin of $s = 1$ was found to be more likely than $s = 0$. The probability to find nonminimal s is reduced if spin–orbit coupling is present, but may still be appreciable if the Coulomb interaction is strong [159]. More work is required to demonstrate definitively that the clustering observed in [16] is due to excitations within spin multiplet and not a nonequilibrium effect (of the sort described in Section 6).

7.6. Spin–orbit interaction in superconducting grains

The Au-doped Al grain mentioned in case (1) of the previous section was sufficiently large that its odd-to-even tunneling spectrum, shown in Fig. 33(a), exhibited the sizeable spectroscopic gap typical of superconducting pairing correlations (cf. Figs. 9 and 10). Interestingly, this allowed Salinas et al. [14] to study how pairing correlations are affected by spin–orbit scattering, by comparing the magnetic-field dependence of this spectrum to that of a pure Al grain of similar size, shown in Fig. 33(b) (the same grain as that discussed in Sections 4.1 and 6.2). We note the following features:

(i) *Nonequilibrium broadening*: For the Au-doped grain (a), which had no gate, the resonance peaks are somewhat broader than in the undoped grain (b). This is believed *not* to be related to the Au impurities, but is instead a nonequilibrium effect, since the spectrum of grain (b), which had a gate, showed a similar broadening when purposefully tuned into nonequilibrium [cf. Fig. 29(a)].

(ii) *Reduced g^{eff} factors*: In contrast to grain (b), for which $g^{eff} = 2 \pm 0.05$, for grain (a) the first two peaks move at low H with slopes $g_1^{eff}/2 + g_1^{ave} = 0.41 \pm 0.03$ and $-g_2^{eff}/2 + g_2^{ave} = -0.27 \pm 0.03$, suggesting values for g^{eff} in the range 0.5–0.8. This behavior is typical of strong spin–orbit scattering.

(iii) *Spectroscopic gap unchanged*: At $H = 0$, the spectroscopic gaps between the first two states are very similar for grains (a) and (b), namely $2\Omega_e \simeq 0.26$ and 0.25 meV, respectively. This nicely illustrates that spin–orbit coupling does not disrupt the superconducting pairing of time-reversed states, since it does not break time-reversal symmetry.
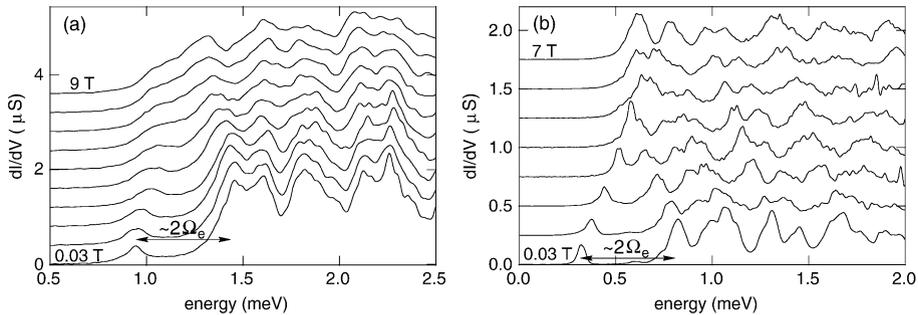

Fig. 33. (a) Odd-to-even tunneling spectrum [14] for an Al grain containing 4% Au impurities, for a sequence of magnetic fields from 0.03 to 9 T in 1 T increments, at $T = 15$ mK. (b) Comparison data for a pure Al grain [10], the same as in Figs. 2, 7, 9, 10 and 29. The curves are artificially offset for visibility.

(iv) *Increased critical field*: Whereas for grain (b) the spectroscopic gap between the lowest two states has disappeared already at 4 T, for grain (a) it decreases much slower with H, due to the reduced g^{eff} factors, and is nonzero even at 9 T. This implies that the critical field for the paramagnetic breakdown of superconductivity (cf. Section 4.5) is much increased by spin–orbit scattering, an effect familiar from thin films in a parallel magnetic field [99].

(v) *Avoided crossing*: At fields above 6 T, the slope of the energy vs. H curve of the first peak of grain (a) changes sign (with the energy decreasing with increasing H at high fields), suggesting an avoided crossing (with a minimum gap of 130 μeV) with the higher-lying levels [similar to Fig. 31(a)]. Interestingly, this implies that spin–orbit scattering modifies the details of the paramagnetic breakdown of pairing correlations: in contrast to a pure grain, where this breakdown is expected to occur rather abruptly when the Zeeman energy of a spin-1 state $|1\rangle$ (with one broken pair) crosses below that of the spin-0 state (with no broken pairs), here the first two levels never really cross, and are not pure spin states to begin with. Therefore, the disruption of pairing correlations must occur gradually, as the spin content of the particle's ground state changes continuously in the avoided crossing region.

8. Ferromagnetic grains

If the leads and/or island of a single-electron transistor are made from an itinerant ferromagnetic material such as Co, Fe or Ni, which have different densities of states for spin-up (majority-band) or spin-down (minority-band) electrons, transport becomes *spin-dependent*. This leads to a number of interesting new phenomena [160], such as a tunneling magnetoconductance (the conductance of the SET depends on the relative orientation of the magnetic moments of leads and island) or Coulomb oscillations as a function of the applied magnetic field (which can shift the chemical potential of the island, analogous to the effect of V_g, if spin-up and spin-down electrons have different density of states near ε_F). Such effects had previously been studied in micronsize ferromagnetic islands [160–169], and also in nm-scale cobalt grains [170,171] at or above helium temperatures.

This work has recently been extended to smaller size and/or lower temperatures by Guéron et al. [18], who studied individual nm-scale Co grains using tunneling spectroscopy. They found that these differed in several interesting ways from nonmagnetic grains, showing, in particular, hysteretic behavior and a larger-than-expected number of low-energy excitations. Since a complete understanding of the observed phenomena and a reliable theoretical framework for their analysis is still lacking, we shall here just briefly mention the most important experimental features, referring the reader to Ref. [18] for details.

8.1. Experimental results

Guéron et al. studied Co grains connected to Al leads via Al_2O_3 tunnel junctions, without a gate. Grain radii were estimated to vary in the range 0.5–2 nm, which, assuming hemispherical grains and a magnetic moment of $1.7\mu_B$/atom (the value for bulk Co), would imply magnetic moments of 90–6000μ_B. The tunneling spectra showed well-resolved peaks due to tunneling via discrete electronic levels within each particle, qualitatively similar to those of Al and Au grains, but with
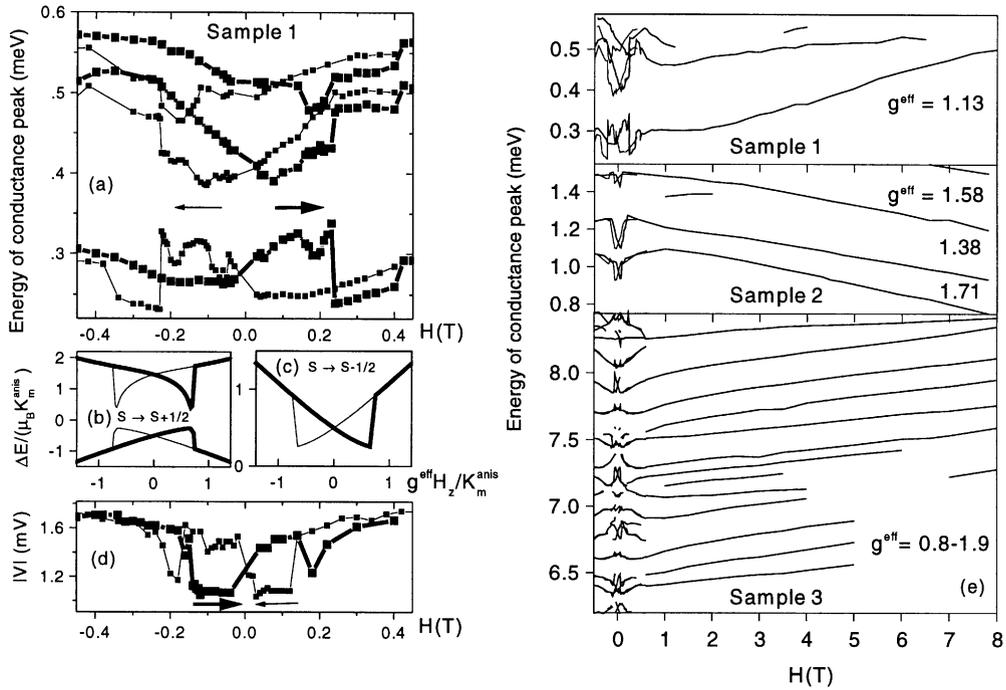
Fig. 34. Magnetic-field dependence of the tunneling spectra of several Co grains [18]. (a) Hysteresis curves showing the dependence of tunneling energies on H for sample 1, at $T = 20\,\text{mK}$. (b, c) Lowest-energy transitions calculated using the Hamiltonian $\hat{H}_{\text{spin}}$ of Eq. (135), for $S = 50$ and H oriented $45°$ from the easy axis, for the case where S increases during tunneling (b) and decreases (c). The qualitative features are independent of the value of S. (d) Voltage threshold for tunneling in one sample which exhibits antihysteretic behavior. (e) Tunneling energies over a larger range of H, for three different samples.

a very different magnetic-field dependence. The latter is summarized in Fig. 34, which exhibits the following representative features:

(i) *Hysteresis*: Fig. 34(a) shows the small-H behavior of the transition energies for the lowest three conductance resonances of a Co grain (sample 1). Each line shows hysteresis as H is swept from -0.45 to $0.45\,\text{T}$ (thick lines) and back (thin lines); it changes smoothly with H, except at two "switching-fields" $\pm H_{\text{sw}}$, at which all three transition energies simultaneously show sudden jumps. The qualitative explanation for these features is that the direction of the magnetic moment vector $\vec{m}$ of the grain changes as H is increased from $-0.45\,\text{T}$: initially $\vec{m}$ is parallel to $\vec{H}$, then rotates toward and past the easy direction (say $\hat{z}$) as the field approaches zero and changes sign, and at H_{sw} abruptly changes direction to align itself with $\vec{H}$ again. We shall call this abrupt directional change "magnetization reversal".

(ii) *Anti-hysteresis*: Fig. 34(d) shows data for a different grain which exhibits anti-hysteresis, i.e. magnetization reversal occurs before H changes sign. This can be explained by the influence of a dipolar magnetic field oriented opposite to the applied H, produced by a second magnetic grain adjacent to the one through which electron tunneling occurs. The reversed field from the second particle can shift the hysteresis curve of the first so that its value of H_{sw} can be negative, while the nonmonotonic shifts at large positive H ($\simeq 0.2\,\text{T}$) are understood as the magnetization reversal of the second particle.

(iii) *Strong spin-up–down asymmetry*: Fig. 34(e) shows the large-H behavior of the Co tunneling spectra. Once H is significantly larger than H_{sw} (say $H > 2\,\text{T}$), all measurable transition energies within a given sample have the *same sign of slope* for $\delta E(H)$ vs. H. This contrasts with Al and Au grains, which have lines of both slopes, with comparable conductance amplitudes, and with a degeneracy at $H = 0$. Thus, for Co there is a strong asymmetry between the tunneling probability for spin-up or spin-down electrons.

(iv) g^{eff} *factors*: The effective g_j^{eff} factors for individual transitions vary between 0.8 and 1.9 and fluctuate quite strongly from level to level.

(v) *Many low-energy excitations*: The energy spacing between tunneling peaks is somewhat less than $0.2\,\text{meV}$, much smaller than expected from the mean level spacing d for noninteracting electrons: For grains of radii 0.5–$2\,\text{nm}$, the calculated density of states (including both sp and d bands) in Co, namely $0.88\,\text{eV}^{-1}\,\text{atom}^{-1}$ [172], implies d-values between 0.75 and $40\,\text{meV}$. The level spacing for individual sp or d bands would be even larger. Note also that the peak spacings of the three samples in Fig. 34 are surprisingly similar to each other.

8.2. In search of a model

At present, a theoretical framework within which all these features can be consistently understood is still lacking. The hysteretic behavior of point (i) can be understood qualitatively in terms of a simple model Hamiltonian [18]:

$$\hat{H}_{\text{spin}} = -g^{\text{eff}}\mu_B \vec{H} \cdot \vec{S} - K_m^{\text{anis}}\mu_B S_z^2 / \sqrt{S(S+1)} \,. \tag{135}$$

It describes a quantum-mechanical spin $\vec{S}$ (representing $\vec{m}$), Zeeman-coupled to a magnetic field, with an easy-axis anisotropy in the $\hat{z}$ direction, where K_m^{anis} is an anisotropy energy per unit $|\vec{m}|$. The magnitude of S of the spin is assumed not to change with H, since the exchange energy U in Co is huge (several eV), and U determines the energy cost for changing S (analogous to the role of the charging energy E_C for changing N). By diagonalizing $\hat{H}_{\text{spin}}$ for a grain with N or $N \pm 1$ electrons (and spin S or $S \pm 1/2$), the tunneling spectrum $\delta\mathscr{E}(H) \equiv \mathscr{E}_{\alpha'}^{N+1} - \mathscr{E}_{\alpha'}^N$ can be calculated for spin-increasing and -decreasing transitions (see [18] for details). The results, shown in Fig. 34(b) and (c), do exhibit hysteresis and jumps reminiscent of the measured ones. In particular, they show that the scale of the anisotropy constant K_m^{anis} can be estimated as $\mu_B K_m^{\text{anis}} \approx \mathscr{E}^{\text{jump}}$, where $\mathscr{E}^{\text{jump}}$ is the size of the jump in the tunneling energy at H_{sw}. This yields $\mu_B K_m^{\text{anis}} \simeq 0.05\,\text{meV}$ for the jumps of Fig. 34(a). Thus, for ultrasmall Co grains the "characteristic magnetic energy scale", which governs the low-lying excitation spectrum in the meV range (and in that sense plays a role analogous to the pairing parameter $\tilde{\Delta}$ for superconducting grains), is the *anisotropy* energy (and not the exchange energy U).

This model fails to capture an important detail, however: it predicts that the jumps occurring for $\delta\mathscr{E}(H)$ as $|H|$ increases through H_{sw} will always have the same sign ($\pm$) as the ensuing large-$|H|$ slope of $\delta\mathscr{E}(H)$, but the lowest line of Fig. 34(a) is a counterexample (jump goes downward, then line moves upward). Guéron et al. proposed that the latter type of behavior could be explained by assuming either that K_m^{anis} changes significantly, due to mesoscopic fluctuations, if the grain's electron number changes, or that the magnitude of $\vec{S}$ (classically $\vec{m}$) does in fact change with H (despite the largeness of U), but no theoretical estimates for the likelihood of these possibilities

exist yet. Making such estimates would have to involve a somewhat more microscopic approach [173], perhaps in the spirit of a Stoner model, in which the degrees of freedom are individual electrons in discrete energy levels, not just their total spin $\vec{S}$.

Next, note that the spin asymmetry of point (ii) is not surprising for Co, since the exchange field breaks the symmetry between spin-up and -down. However, since Co is not fully spin polarized ($P \approx 30\%$, [99]), a priori lines of both slopes would be expected to occur, even if in unequal numbers and with different amplitudes. Recent data (unpublished) do occasionally show a line with a slope quite different in magnitude from that of the majority of lines, although the sign of the slope is still the same. The reason why all observed transitions shift as a function of magnetic field with the same sign of the slope is still a mystery. Its resolution may involve a better understanding of whether the states involved in tunneling are primarily sp or primarily d states (whose density of states near ε_F differ vastly), whether there is a difference in the nature of the matrix elements for spin-up or spin-down electrons to enter the many-body eigenstates of the grain, or whether spin–orbit coupling within the magnetic nanoparticle is sufficiently strong that the states involved in tunneling are coherent mixtures of spin states (rather than being purely spin-up or spin-down).

A possible candidate for explaining the small level spacing of point (iv), apart from nonequilibrium effects (which cannot be ruled out in ungated devices), are "spatially uniform[23] ($\vec{k} = 0$) spin-wave modes": classically speaking, these involve fluctuations in the direction of $\vec{m}$ about its ground state direction (say $\hat{m}_0$), quantum-mechanically speaking, they involve different s_m eigenvalues of the operator $\hat{m}_0 \cdot \vec{S}$. The spin-wave excitation energy can be estimated as $\approx 2\mu_B K_m^{anis} \approx 0.1$ meV, which is comparable to the observed inter-peak spacing. Note, though, that spin waves of different quantum numbers s_m could be expected to have s_m-dependent slopes for $\delta\mathscr{E}(H)$ vs. H, whereas no systematic tendencies for the slopes are discernable in Fig. 34(e). This might be held against the spin-wave interpretation, but might also be a consequence of mesoscopic fluctuations in g_j^{eff} factors [cf. point (iv)].

8.3. Dynamics of magnetization reversal

As the above discussion indicates, a number of features of the tunneling spectra of Co grains still await detailed clarification. This reflects the fact that the electronic structure of Co is much more involved than that of Al or Au. However, independently of whether all open questions can be resolved or not, the fact that magnetic anisotropy causes each resonance energy in the tunnel spectrum to shift reproducibly by on the order of 0.1 meV as H is swept about the hysteresis loop has a very interesting potential application: it could be used as a tool to perform detailed studies of the *dynamics of magnetization reversal* in individual nm-scale grains, since the jumps occurring in the tunneling spectra upon reversal of the magnetization allow one to monitor precisely when, as function of ramped applied field or of waiting time, this magnetization reversal occurs. Thus this method would be complementary to magnetic force microscopy [174], Hall magnetometry [175], and SQUID techniques [176,177] for studying magnetization reversal.

---

[23] The contribution of exchange energy to the lowest-energy *nonuniform* spin-wave modes can be estimated by quantizing the spin-wave dispersion curve of Co within the size of a nanoparticle. This gives an energy (300 meV) $(a/2r)^2$ where a is a lattice spacing and r is the grain radius [102], or ≈ 1 meV for a 2-nm grain.

Overcoming the energy barrier between two different directions for the magnetization occurs by thermal activation for large temperatures, or quantum tunneling for sufficiently small T. For Co grains of the present size ($m \approx 90$–$6000\mu_B$), the latter case would be an example of *macroscopic quantum tunneling* (MQT) of the magnetization, which has been studied in great detail theoretically [178]. It should be very interesting to use the relevant information in the literature in order to predict the feasibility for observing MQT in an individual Co grain in a SET geometry, and to estimate what grain sizes would be optimal for this purpose. Wernsdorfer et al. [176] found deviations from purely thermally activated behavior when studying an individual Co grain with a diameter of 20 nm, using highly sensitive SQUID techniques. Since the grains of Guéron et al. are much smaller, it is quite likely that they would show such deviations too. It will be important, however, to clarify to what extent damping, due to the coupling to the leads, reduces the chances for seeing MQT. Conversely, if it turns out that MQT can indeed be observed, it might be possible to investigate the effects of damping on MQT in a controlled way by purposefully tuning the grain out of equilibrium.

9. The Kondo box: a magnetic impurity in an ultrasmall grain

In this section we briefly discuss what is expected to happen if a normal ultrasmall grain contains a single magnetic impurity. This system, to be called a "Kondo box", was studied theoretically by Thimm, Kroha and von Delft (TKvD) [57]. The Kondo box is another example of the general rule of thumb that we repeatedly encountered in previous sections: when the mean level spacing d becomes larger than the energy scale characterizing the system's correlations, in this case the Kondo temperature T_K, interesting new effects occur. TKvD showed that for $d \gtrsim T_K$, level discreteness strongly affects the Kondo resonance, in a parity-dependent way, and predicted that this should lead to measurable anomalies in the conductance through the grain.

For the impurity concentrations of 0.01–0.001% that yield a detectable Kondo effect in bulk alloys, an ultrasmall grain of typically 10^4–10^5 atoms will contain only a single impurity, so that inter-impurity interactions need not be considered. TKvD thus considered a single impurity in an ultrasmall metallic grain, described by the impurity Anderson model [179,180] with a discrete conduction band:

$$H = H_0 + \varepsilon_d \sum_{\sigma} c_{d\sigma}^{\dagger} c_{d\sigma} + v \sum_{j,\sigma} (c_{j\sigma}^{\dagger} c_{d\sigma} + c_{d\sigma}^{\dagger} c_{j\sigma}) + U c_{d+}^{\dagger} c_{d+} c_{d-}^{\dagger} c_{d-} \ , \qquad (136)$$

where $H_0 = \sum_{j,\sigma} \varepsilon_j c_{j\sigma}^{\dagger} c_{j\sigma}$. Here σ denotes spin and the $c_{j\sigma}^{\dagger}$ create conduction electrons in the discrete, delocalized eigenstates $|j\sigma\rangle$ of the "free" system (i.e. without impurity). Their energies, measured relative to the chemical potential μ, are taken uniformly spaced for simplicity: $\varepsilon_j = jd + \bar{\varepsilon}_0 - \mu$. $c_{d\sigma}^{\dagger}$ creates a spin-σ electron in the localized level of the magnetic impurity, which has bare energy ε_d far below ε_F, and a Coulomb energy cost $U(\rightarrow \infty)$ for being doubly occupied. The hybridization matrix element v between the local impurity level and the conduction band is an overlap integral between a localized and a delocalized wavefunction, and, due to the normalization of the latter, scales as $v \propto \text{Vol}^{-1/2}$. Thus, the effective width of the impurity level, $\Gamma = \pi v^2 / d$, is volume-independent, as is the bulk Kondo temperature, $T_K = \sqrt{2\Gamma D / \pi} \exp(-\pi\varepsilon_d / 2\Gamma)$, where D is a high-energy band cutoff. To distinguish, within the grand canonical formalism, grains for which

the *total* number of ($c_{j\sigma}$ *and* $c_{d\sigma}$) electrons is even or odd, μ is chosen either on ($\mu = \bar{\varepsilon}_0$) or half-way between two ($\mu = \bar{\varepsilon}_0 + d/2$) single-particle levels, respectively [21,22].

The system's correlations can be characterized in terms of the spectral density $A_{d\sigma}(\omega)$ of the impurity Green's function $G_{d\sigma}(t) = -i\theta(t)\langle\{c_{d\sigma}(t), c_{d\sigma}^\dagger(0)\}\rangle$, which TKvD calculated using the standard noncrossing approximation (NCA) [181,182]. Their results are summarized in Fig. 35.

(i) *Bulk limit*: For $d \ll T$, the impurity's spectral density is indistinguishable from the bulk case ($d \to 0$) [lowest lines in Figs. 35 (a) and (b)]: it shows the familiar Kondo resonance near $\omega = 0$, with width of order T_K when $T \ll T_K$, which arises due to coherent virtual transitions between the impurity level and the conduction band.

(ii) *Subresonances*: When d is increased well beyond T, however, the Kondo resonance splits up into a set of individual sub-peaks, reflecting the discreteness of the conduction band [higher
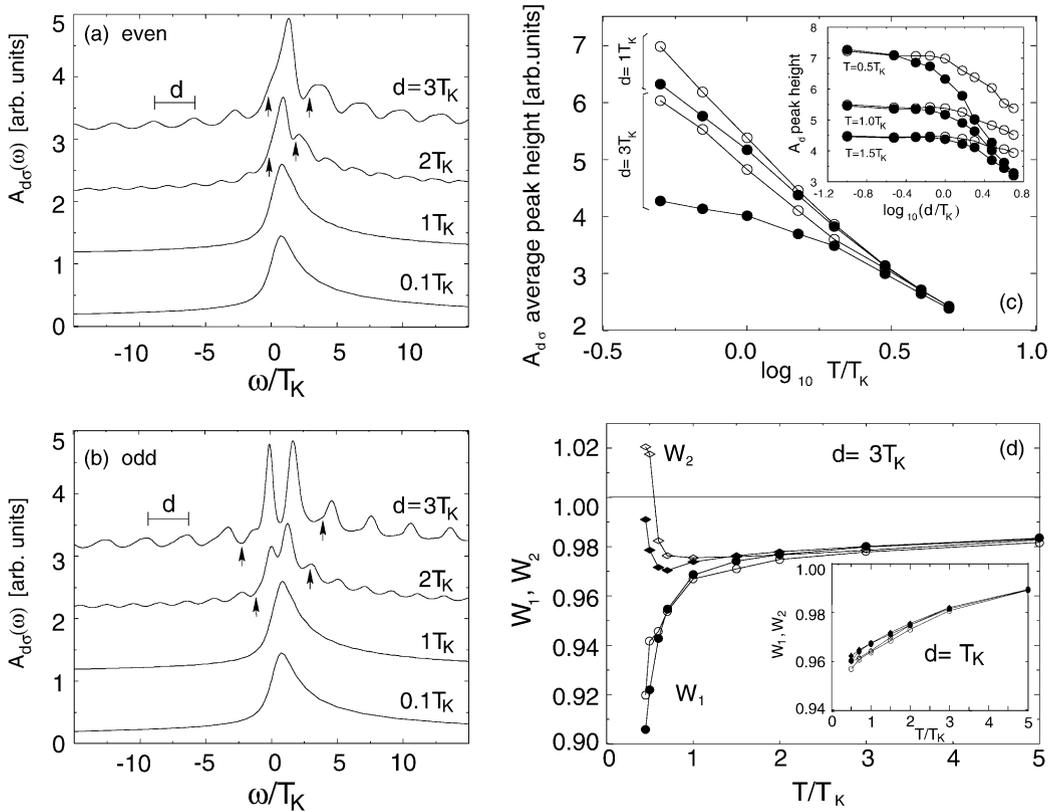
Fig. 35. Results for a magnetic impurity in an ultrasmall metallic grain. (a, b) Impurity spectral function $A_{d\sigma}(\omega)$ for various values of d at $T = 0.5T_K$; (a) even, (b) odd total number of electrons. The individual curves are vertically offset by one unit each. (c) Even/odd dependence of the average peak height of the Kondo resonance, as function of T/T_K. For an even box ($\circ$), $A_{d\sigma}$ was averaged over a range d centered on its central sub-peak, for an odd box ($\bullet$) over a range $2d$ centered on its central two sub-peaks [as indicated by arrows in (a) and (b)]. The inset shows the same quantity as function of d/T_K. Numerical uncertainties are smaller than the symbol sizes. (d) Anomalous temperature dependence of the weights W_1 (circles) and W_2 (diamonds) of the first two conductance peaks of an even grain ($\circ,\diamond$) and an odd grain ($\bullet,\blacklozenge$), for $d = 3T_K$ and T_K (inset).

lines of Figs. 35(a) and (b)]. Nevertheless, the Kondo resonance retains its main feature, namely significant spectral weight within a few T_K around the Fermi energy, up to the largest ratios of $d/\max(T, T_K)$ ($\simeq 5$) that were considered. This implies that the Kondo correlations induced by the spin-flip transitions between the impurity level and the lowest-lying unoccupied *j*-levels persist up to remarkably large values[24] of $d/\max(T, T_K)$.

(iii) The *Kondo correlations* do *weaken* systematically with increasing *d*, however, as can be seen in the inset of Fig. 35(c), which shows the average peak height of the Kondo resonance (which quantifies the "strength" of the Kondo correlations) as function of *d* at fixed *T*: the peak height drops logarithmically with increasing *d* once *d* becomes larger than about *T*. Conversely, at fixed *d*, it drops logarithmically with increasing *T* once *T* becomes larger than about $0.5d$ [main part of Fig. 35(c)], thus reproducing the familiar bulk behavior. Qualitatively, these features are readily understood in perturbation theory, where the logarithmic divergence of the spin flip amplitude, $t(\omega) \propto \sum_{j \neq \omega} f(\varepsilon_j)/(\omega - \varepsilon_j)$, is cut off by either *T* or *d*, whichever is largest.

(iv) *Parity effects*: For $d \gg T$, the even and odd spectral functions $A_{d\sigma}$ in Figs. 35(a) and (b) differ strikingly: the former has a single central main peak, whereas the latter has two main peaks of roughly equal weight. This can be understood as follows: for an even grain, spin-flip transitions lower the energy by roughly T_K by binding the *d* electron and the conduction electrons into a Kondo singlet, in which the topmost, *singly* occupied *j* level of the free Fermi sea carries the dominant weight, hence the single dominant peak in $A_{d\sigma}$. For an odd grain, in contrast, the free Fermi sea's topmost *j* level is *doubly* occupied, blocking energy-lowering spin-flip transitions. To allow the latter to occur, the two electrons in this topmost level are redistributed with roughly equal weights between this and the next-higher-lying level, causing two main peaks in $A_{d\sigma}$ and reducing the net energy gain from T_K by an amount of order *d*. This energy penalty intrinsically weakens Kondo correlations in odd relative to even grains; indeed, the average $A_{d\sigma}$ peak heights in Fig. 35(c) are systematically lower in odd than in even grains, and more so the larger *d* and the smaller *T*.

(v) *Anomalous conductance*: The Kondo correlations for $A_{d\sigma}(\omega)$ affect the properties of the conduction electron density of states on the grain, in a way which should be detectable by using single-electron-tunneling spectroscopy and measuring the differential conductance $G(V)$ through the grain. TKvD showed that contributions to $G(V)$ due to the bare conduction electron density of states interfere with contributions coming from the Kondo resonance in $A_{d\sigma}(\omega)$, in a way reminiscent of a Fano resonance [185]. This interference causes the conductance to acquire a distinct, anomalous *T* dependence: the *weights* W_j under the individual peaks (labeled by *j*) of $G(V)$ become *T* dependent. (In contrast, the weight W_0 under an individual peak of the bare conductance, say $G_0(V)$, is *T* independent, since the *T* dependence of the peak shapes of G_0 is determined solely by the derivative of the Fermi function.) This anomalous behavior is illustrated in Fig. 35(d), which shows the *T* dependence of the weights W_1 and W_2 of the first and second conductance peaks that occur when *V* is increased from 0: When *T* decreases at fixed $d = T_K$, both W_1 and W_2 decrease, while at fixed $d = 3T_K$, W_1 decreases whereas W_2 increases. The fact that the weights can either increase or decrease with decreasing *T* results from the constructive or destructive Fano-like interference effects

---

[24] Actually, remnants of Kondo correlations have been demonstrated to exist even in molecular systems [183,184].

mentioned above. Moreover, at the larger value for d, both W_1 and W_2 develop a *parity effect in the strength of their T dependence*.

(vi) *Strength of anomalies*: TKvD estimated that the Kondo-induced T dependence in peak weights in Fig. 35 should be strong enough to be experimentally detectable. For e.g. an Fe impurity in an even Cu grain of size $(3\,\text{nm})^3$ ($d = 30\,\text{K}$, $T_K \simeq 10\,\text{K}$), cooling from $T = 2T_K$–$0.5T_K$ should change W_1 by $\simeq 7\%$. (TKvD expect W_1 to change some more as T is lowered further, but their numerics became unreliable in this regime.)

To conclude this section, we note that since the effects discussed above result, above all, from the discrete density of states near ε_F, they should be generic for ultrasmall grains. In other words, they should be robust against including randomness in the model, like j-dependent level spacings d_j and hybridization matrix elements v_j.

10. Summary and outlook

The technique of single-electron-tunneling spectroscopy on ultrasmall metallic grains has proved to be a very fruitful way of probing electron correlations in metals, and the way in which these are modified by mesoscopic fluctuations and level discreteness. Let us briefly summarize the main conclusions reached in each of the preceding sections:

Section 3: For normal grains, the effect of an applied magnetic field is simply to lift the Kramers degeneracy of time-reversed states. This can be used to determine the grain's *number parity*. In ultrasmall grains, the effect of a magnetic field on orbital motion is negligible.

Section 4: For largish Al grains, the observation of a distinct spectral gap in even grains and its absence in odd grains is clear evidence for the presence of *superconducting pairing correlations*. These can be satisfactorily described using the simple discrete BCS model introduced in Section 4.2. The blocking of some levels by unpaired electrons leads to various measurable parity effects; among these, a pairbreaking-energy parity effect should be observable in experiments of the present kind, provided the grain size can be better controlled. The dominant mechanism by which a magnetic field destroys pairing correlations in ultrasmall grains is Pauli paramagnetism. Decreasing the grain size softens the first-order transition observed for thin films in a parallel field, by reducing the number of spins flipped from being macroscopically large for $d \ll \tilde{\Delta}$ to being of order one for $d \simeq \tilde{\Delta}$. The grand-canonical variational BCS approach fails for $d \gtrsim \tilde{\Delta}$; nevertheless, it yields a useful framework for a qualitative analysis of the experiments, which had $d \lesssim \tilde{\Delta}$.

Section 5: The *crossover* of the behavior of superconducting pairing correlations from the bulk limit ($d \ll \tilde{\Delta}$) to the fluctuation-dominated regime ($d \gg \tilde{\Delta}$) is parity dependent and completely smooth. This remains true for systems with nonuniform rather than uniform level spacings. Very remarkably, the discrete BCS model has an exact solution, due to Richardson, with which $T = 0$ properties can be calculated rather easily. Finite-temperature properties for finite-sized systems can be calculated quite reliably with the correlated static-path approximation (provided $T > T_{\text{CSPA}}$). Thus, the only remaining open problem is the development of canonical finite-T methods that remain quantitatively reliable for $d \gtrsim \tilde{\Delta}$. The spin susceptibility $\chi(T)$ of an odd grain shows an interesting re-entrant behavior even for $d \gg \tilde{\Delta}$, which might be a way to detect remnants of pairing correlations in the fluctuation-dominated regime.

Section 6: The excitation spectra of *nonequilibrium* grains show *clusters of resonances*. The spacing between clusters is comparable to the mean level spacing, but the spacings between subresonances of the same cluster are much smaller than d. Each cluster of resonances can be identified with one excited single-electron state of the metal grain. Subresonances reflect energy shifts (of order d/g_{dim}) due to different nonequilibrium occupancy configurations of other single-electron states. These energy shifts are manifestations of mesoscopic fluctuations of the matrix elements of the electron–electron interaction. Superconducting grains that are tuned out of equilibrium by a gate voltage show subgap structures in odd-to-even but not in even-to-odd tunneling spectra; these features can be attributed to nonequilibrium excitations generated by *cotunneling* processes. As a function of increasing excitation energy, the nature of the excitation spectrum has been observed to change from consisting of discrete peaks to being continuous, the crossover scale being $\varepsilon \sim E_{Thouless}$; the observed details of this crossover are in accord with theoretical expectations for the behavior of quasiparticle lifetimes in disordered interacting systems.

Section 7: The *spin–orbit interaction* leads to reduced g_j^{eff} factors for the splitting of Kramers doublets in a magnetic field, and to avoided crossings. For grains with superconducting pairing correlations, the critical field for which these are expected to break down is increased significantly relative to pure samples; the $H = 0$ spectroscopic gap remains unchanged, however, as expected due to time-reversal symmetry. The g_j^{eff} factors exhibit mesoscopic fluctuations, which can be studied theoretically using random matrix theory. Interestingly, the Zeeman-response to a magnetic field has been predicted to be anisotropic in the presence of spin–orbit scattering, implying that upon changing the direction of the applied field, the measured g_j^{eff} factors should vary strongly (by a factor of up to 8–10).

Section 8: *Ferromagnetic Co grains* have hysteretic tunneling spectra. These can be used as a tool to measure the switching field at which the grain's magnetization undergoes a sudden change in direction as an applied magnetic field is ramped. There is a strong asymmetry for the tunneling probability of spin-up vs. spin-down electrons. The spacing between low-lying excitations is much smaller than expected from the estimated free-electron mean level spacing d (and remarkably uniform from one grain to the next). However, it is comparable to the grains' anisotropy energy (per unit magnetization), which for ultrasmall Co grains is the "characteristic magnetic energy scale" governing the low-lying excitation spectrum in the meV range. Many features of these measurements are not yet understood in detail, and further work, both experimental and theoretical, will be needed to clarify them.

Section 9: For an ultrasmall grain containing a *magnetic Kondo impurity*, the Kondo resonance is strongly affected, in a parity-dependent way, when the mean level spacing in the grain becomes larger than the Kondo temperature. The tunneling spectrum is predicted to show Kondo-induced Fano-type resonances of measurable size, with an anomalous dependence on temperature and level spacing.

Prospects for future work:

Experiment: In the current generation of experiments, the grain's actual size and shape cannot be determined very accurately. It would be a great advance if fabrication techniques could be developed to the point that grains can be used which have been custom-made, by chemical techniques, to have well-defined sizes and shapes (e.g. spherical). This would significantly reduce

the uncertainties which one presently encounters when estimating characteristic parameters of the grain, such as the single-particle mean level spacing d, the dimensionless conductance g_{dim} or the Thouless energy. Moreover, it would allow systematic studies of the dependence of various quantities on grain size or mean level spacing [for example, it would be interesting to try to do this for the pairbreaking energies Ω_{e}, Ω_{o} of Fig. 19(b)]. Encouragingly, the feasibility of using chemically prepared grains in SETs has already been demonstrated several times [114,186,187], though the resulting devices have not yet been used for single-electron-tunneling spectroscopy.

Moreover, most of the devices that have been studied so far had no gate, because it is technically very challenging to add one. It would, however, be very useful if ways could be found for routinely fabricating gates, because that allows both the grain's number parity and the amount of nonequilibrium to be adjusted controllably.

Theory: The behavior of superconducting pairing correlations in an individual ultrasmall grain can now be regarded as a subject that is well understood. It would be interesting to try to use the insights that have been gained for a single grain in order to now study systems of several coupled grains: what, for example, is the fate of the Josephson effect between two coupled grains as their sizes are reduced to the point that $d \sim \tilde{\Delta}$?

The experiments on nonequilibrium effects and spin–orbit interactions revealed the importance of mesoscopic fluctuations in remarkably direct ways, and are well understood in terms of theories that treat these fluctuations via random matrix theory. The success of the latter approach stems from the fact that for chaotic systems such as irregularly shaped grains, "details do not matter" much, so that a description in terms of random matrices is appropriate.

At present, the most intruiging open theoretical problem seems to be that of finding a suitable framework within which to analyze and interpret the experiments on ferromagnetic Co grains. The observed phenomena are rather rich and complex, and it is currently not at all clear which details "matter" and which do not. More experimental studies on this system will be needed as guidelines for theoretical attempts to construct a useful model for it.

A further stimulus for theoretical work on ferromagnetic grains might come from the prospect of exploiting the hysteretic properties of their tunneling spectra to study the *dynamics of magnetization reversal*: it would be interesting to estimate the feasibility for seeing macroscopic quantum tunneling in this system. In particular, one should try to determine whether (or to what extent) this phenomenon can survive when a current flows through the grain.

Acknowledgements

Appendix A. Superconducting leads

For the case of superconducting leads in the absence of an external magnetic field, the analysis of Sections 2.3.2 and 2.3.3 has to be modified as follows (for generality, we continue to allow for the presence of a gate; if none is present, set $C_g = 0$ below): Firstly, the lead Hamiltonian (8) must be replaced by the BCS Hamiltonian for Bogoljubov quasiparticles,

$$H_r = \sum_{k\sigma} (E_{kr} + eV_r)\gamma_{k\sigma r}^{\dagger}\gamma_{k\sigma r}, \qquad E_{kr} = \sqrt{\varepsilon_{kr}^2 + \Delta_r^2}, \tag{A.1}$$

$$\gamma_{k\uparrow r} = u_{kr}c_{k\uparrow r} - v_{kr}c_{-k\downarrow r}^{\dagger}, \qquad \gamma_{k\downarrow r} = u_{kr}c_{k\downarrow r} + v_{kr}c_{-k\uparrow r}^{\dagger}, \tag{A.2}$$

$$u_{kr} = [\tfrac{1}{2}(1 + \varepsilon_{kr}/E_{kr})]^{1/2}, \qquad v_{kr} = [\tfrac{1}{2}(1 - \varepsilon_{kr}/E_{kr})]^{1/2} \tag{A.3}$$

and H_{tun} of (11) should be rewritten in terms of the γ's. Secondly, the golden rule expressions for the tunneling rates now are [instead of (20) to (23)]

$$\Sigma_{\alpha\alpha'}^{r+} = \frac{2\pi}{\hbar}\sum_{k\sigma} |T_{k\sigma,\alpha\alpha'}^{r*}|^2 [f(E_{kr})|u_{kr}|^2\delta(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} - E_{kr} - e\bar{V}_r) \tag{A.4}$$

$$+ (1 - f(E_{-kr}))|v_{kr}|^2\delta(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} + E_{-kr} - e\bar{V}_r)] \tag{A.5}$$

$$\simeq f(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} - e\bar{V}_r)\mathscr{N}_{r,\mathrm{qp}}(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} - e\bar{V}_r)\Gamma_{\alpha\alpha'}^{r+}, \tag{A.6}$$

$$\Sigma_{\alpha\alpha'}^{r-} = \frac{2\pi}{\hbar}\sum_{k\sigma} |T_{k\sigma,\alpha\alpha'}^{r}|^2 [(1 - f(E_{kr}))|u_{kr}|^2\delta(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} + E_{kr} - e\bar{V}_r) \tag{A.7}$$

$$+ f(E_{-kr})|v_{kr}|^2\delta(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} - E_{-kr} - e\bar{V}_r)] \tag{A.8}$$

$$= f(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} - e\bar{V}_r)]\mathscr{N}_{r,\mathrm{qp}}(\mathscr{E}_{\alpha} - \mathscr{E}_{\alpha'} - e\bar{V}_r)\Gamma_{\alpha\alpha'}^{r-}, \tag{A.9}$$

where $\mathcal{N}_{r,\mathrm{qp}}(\omega)$ is the BCS quasiparticle density of states in lead r [see Eq. (33)]. To arrive at (A.6) and (A.9), we took the limit $T \ll \Delta_r$ [so $f(E_{kr}) \simeq 0$] and used [instead of Eq. (15)]

$$\frac{2\pi}{\hbar} \sum_k T^{r*}_{kl\sigma} T^r_{kl'\sigma} \begin{Bmatrix} |u_{kr}|^2 \\ |v_{kr}|^2 \end{Bmatrix} \delta(\omega - E_{kr}) = \delta_{ll'} \Gamma^r_{l\sigma} \mathcal{N}_{r,\mathrm{qp}}(\omega) \tag{A.10}$$

(the ε-anti-symmetric part of $|u_{kr}|^2$ and $|u_{kr}|^2$ does not contribute).

Eqs. (A.6) and (A.9) are the results used in Eq. (32) of Section 2.4.3.

Appendix B. Richardson's exact solution of discrete BCS model

This appendix summarizes some of the main features of Richardson's exact solution of the discrete BCS model. We begin in Appendix B.1 by deriving in pedagogical detail, following [40], some of the key results of Richardson's exact solution of the discrete BCS model of Section 4.2. In Appendix B.2 we transform the eigenvalue equation derived below [Eq. (B.17)] into a less singular form more convenient for numerical solution, and in Appendix B.3 we present analytic formulas needed to calculate the correlators $\langle b_j^\dagger b_j \rangle$ and $\langle b_i^\dagger b_j \rangle$ exactly.

B.1. Derivation of eigenstates and eigenvalues

B.1.1. The hard-core boson problem

For presented purposes it suffices to consider the pair Hamiltonian[25] [cf. Eq. (44)]

$$H_U = \sum_{ij}^U (2\varepsilon_j \delta_{ij} - g) b_i^\dagger b_j \quad (\text{where } g = \lambda d) , \tag{B.1}$$

where j runs over a set U of N_U non-singly-occupied single-particle levels (degenerate levels with $\varepsilon_i = \varepsilon_j$ for $i \neq j$ are allowed), on which the pair-creation operators $b_j^\dagger = c_{j+}^\dagger c_{j-}^\dagger$ satisfy "hard-core boson" relations

$$b_j^{\dagger 2} = 0, \quad [b_j, b_{j'}^\dagger] = \delta_{jj'}(1 - 2b_j^\dagger b_j), \quad [b_j^\dagger b_j, b_{j'}^\dagger] = \delta_{jj'} b_j^\dagger . \tag{B.2}$$

We seek eigenstates that contain n pairs, satisfy $(H_U - \mathscr{E}_n)|\Psi_n\rangle$ and are of the general form

$$|\Psi_n\rangle = \sum_{j_1,\ldots,j_n}^U \psi(j_1,\ldots,j_n) \prod_{v=1}^n b_{j_v}^\dagger |\mathrm{Vac}\rangle . \tag{B.3}$$

In his original publications [46–48], Richardson derived a Schrödinger equation for $\psi(j_1,\ldots,j_n)$ and showed that its exact solution was simply a generalization of the form that $\psi(j_1,\ldots,j_n)$ would have had if the b's has been true (not hard-core) bosons. With the benefit of hindsight, we shall here follow an alternative, somewhat shorter root, also due to Richardson [119]: we first consider the related but much simpler case of *true* bosons and write down the generic form of its eigenstates; we then clarify why this form fails to produce eigenstates of the *hard-core* boson Hamiltonian; and

---

[25] In this appendix the shorthand $g = \lambda d$ is useful, since Richardson's solution depends only on this combination.

having identified the reason for the failure, we show that (remarkably) only a slight generalization is needed to repair it and to obtain the sought-after hard-core-boson eigenstates.

B.1.2. True bosons

Let $\tilde{b}_j$ denote a set of true bosons (i.e. $[\tilde{b}_j, \tilde{b}_{j'}^{\dagger}] = \delta_{jj'}$), governed by a Hamiltonian $\tilde{H}_U$ of precisely the form (44), with $b_j \to \tilde{b}_j$. This problem, being quadratic, can be solved straightforwardly by any number of methods. The solution is as follows: $\tilde{H}_U$ can be written as

$$\tilde{H}_U = \sum_J \tilde{E}_J \tilde{B}_J^{\dagger} \tilde{B}_J + \text{const.} , \tag{B.4}$$

where the new bosons $\tilde{B}_J^{\dagger}$ (with normalization constants C_J) are given by

$$\tilde{B}_J^{\dagger} = g C_J \sum_j^U \frac{\tilde{b}_j^{\dagger}}{2\varepsilon_j - \tilde{E}_J}, \quad \frac{1}{(g C_J)^2} = \sum_j^U \frac{1}{(2\varepsilon_j - \tilde{E}_J)^2} , \tag{B.5}$$

and the boson eigenenergies $\tilde{E}_J$ are the roots of the eigenvalue equation

$$1 - \sum_j^U \frac{g}{2\varepsilon_j - \tilde{E}_J} = 0 . \tag{B.6}$$

This is an equation of order N_U in $\tilde{E}_J$. It thus has N_U roots, so that the label J runs from 1 to N_U. As the coupling g is turned to 0, each E_J smoothly evolves to one of the bare eigenenergies $2\varepsilon_j$. A general n-boson eigenstate of $\tilde{H}_U$ and its eigenenergy $\tilde{\mathscr{E}}_n$ thus have the form

$$|\tilde{\Psi}_n\rangle = \prod_{v=1}^n \tilde{B}_{J_v}^{\dagger}|0\rangle, \quad \tilde{\mathscr{E}}_n = \sum_{v=1}^n \tilde{E}_{J_v} , \tag{B.7}$$

where the n indices $J_1, \ldots, J_n$ that characterize this state need not all be distinct, since the $B_J^{\dagger}$ are true bosons.

B.1.3. Complications arising for hard-core bosons

Let us now return to the hard-core boson Hamiltonian H_U. Its eigenstates will obviously *not* be identical to the true-boson eigenstates just discussed, since matters are changed considerably by the hard-core properties of b_j. To find out exactly *what* changes they produce, it is very instructive to take an Ansatz for $|\Psi_n\rangle$ similar to (B.7) (but suppressing the normalization constants and taking all boson indices to be distinct), namely

$$|\Psi_n\rangle = \prod_{v=1}^n B_v^{\dagger}|0\rangle \quad \text{with} \quad B_v^{\dagger} = \sum_j^U \frac{b_j^{\dagger}}{2\varepsilon_j - E_v} \tag{B.8}$$

and to check explicitly whether or not it could be an eigenstate of H_U, i.e. to check under what conditions $(H_U - \mathscr{E}_n)|\Psi_n\rangle$ would equal zero, where $\mathscr{E}_n = \sum_v^n E_v$. To this end, we commute H_U to the right past all the $B_v^{\dagger}$ operators in $|\Psi_n\rangle$, using

$$\left[H_U, \prod_{v=1}^n B_v^{\dagger} \right] = \sum_{v=1}^n \left\{ \left(\prod_{\eta=1}^{v-1} B_\eta^{\dagger} \right) [H_U, B_v^{\dagger}] \left(\prod_{\mu=v+1}^n B_\mu^{\dagger} \right) \right\} . \tag{B.9}$$

To evaluate the commutators appearing here, we write H_U as

$$H_U = \sum_j^U 2\varepsilon_j b_j^\dagger b_j - g B_0^\dagger B_0 \quad \text{where} \quad B_0^\dagger = \sum_j^U b_j^\dagger \tag{B.10}$$

and use the following relations:

$$[b_j^\dagger b_j, B_\nu^\dagger] = \frac{b_j^\dagger}{2\varepsilon_j - E_\nu}, \qquad [B_0, B_\nu^\dagger] = \sum_j^U \frac{1 - 2b_j^\dagger b_j}{2\varepsilon_j - E_\nu}, \tag{B.11}$$

$$[H_U, B_\nu^\dagger] = E_\nu B_\nu^\dagger + B_0^\dagger \left[1 - g \sum_j^U \frac{1 - 2b_j^\dagger b_j}{2\varepsilon_j - E_\nu} \right]. \tag{B.12}$$

Inserting these into (B.9) and using $H_U|0\rangle = 0$ and $\mathscr{E}_n = \sum_\nu^n E_\nu$, we find

$$H_U|\Psi_n\rangle = \mathscr{E}_n|\Psi_n\rangle + \sum_{\nu=1}^n \left[1 - \sum_j^U \frac{g}{2\varepsilon_j - E_\nu} \right] B_0^\dagger \left(\prod_{\eta=1(\neq\nu)}^n B_\eta^\dagger \right) |0\rangle$$

$$+ \sum_{\nu=1}^n \left\{ \left(\prod_{\eta=1}^{\nu-1} B_\eta^\dagger \right) \left[\sum_j^U \frac{2g B_0^\dagger b_j^\dagger b_j}{2\varepsilon_j - E_\nu} \right] \left(\prod_{\mu=\nu+1}^n B_\mu^\dagger \right) \right\} |0\rangle. \tag{B.13}$$

Now, suppose we do the same calculation for true instead of hard-core bosons (i.e. run through the same steps, but place a $\tilde{\ }$ on H_U, b_j, E_ν and $\mathscr{E}_n$). Then the second line of (B.13) would be absent (because the $b_j^\dagger b_j$ terms in the second of Eqs. (B.2) and (B.11) and in (B.12) would be absent); and the first line of (B.13) would imply that $(\tilde{H}_U - \tilde{\mathscr{E}}_n)|\tilde{\Psi}_n\rangle = 0$ provided that the term in square brackets vanishes, which is nothing but the condition that the $\tilde{E}_\nu$ satisfy the true-boson eigenvalue equation of (B.6)! In other words, we have just verified explicitly that all true-boson states of the form (B.7) are indeed eigenstates of $\tilde{H}_U$, provided that the $\tilde{E}_\nu$ satisfy (B.6). Moreover, we have identified the term in second line of (B.13) as the extra complication that arises for hard-core bosons.

B.1.4. The cure: a generalized eigenvalue equation

Fortunately, this extra complication is tractable: first, we note that

$$\left[\sum_j^U \frac{2g B_0^\dagger b_j^\dagger b_j}{2\varepsilon_j - E_\nu}, B_\mu^\dagger \right] = \sum_j^U \frac{2g B_0^\dagger}{2\varepsilon_j - E_\nu} \frac{b_j^\dagger}{2\varepsilon_j - E_\mu} \tag{B.14}$$

$$= 2g B_0^\dagger \frac{B_\nu^\dagger - B_\mu^\dagger}{E_\nu - E_\mu}. \tag{B.15}$$

The second line follows via a partial fraction expansion, and remarkably, contains only $B^\dagger$ operators and no more $b_j^\dagger b_j$s. This enables us to eliminate the $b_j^\dagger b_j$s from the second line of (B.13), by rewriting it as follows (we commute its term in square brackets to the right, using a relation similar to (B.9), but with the commutator (B.14) instead of $[H_U, B_\mu^\dagger]$):

$$\sum_{\nu=1}^n \left\{ \left(\prod_{\eta=1}^{\nu-1} B_\eta^\dagger \right) \sum_{\mu=\nu+1}^n \left\{ \left(\prod_{\eta'=\nu+1}^{\mu-1} B_{\eta'}^\dagger \right) \left[2g B_0^\dagger \frac{B_\nu^\dagger - B_\mu^\dagger}{E_\nu - E_\mu} \right] \left(\prod_{\mu'=\mu+1}^n B_{\mu'}^\dagger \right) \right\} \right\} |0\rangle$$

$$= \sum_{\mu=1}^n \left[\sum_{\nu=1}^{\mu-1} \frac{2g}{E_\nu - E_\mu} \right] B_0^\dagger \left(\prod_{\eta=1(\neq\mu)}^n B_\eta^\dagger \right) |0\rangle$$

$$- \sum_{\nu=1}^{n} \left[\sum_{\mu=\nu+1}^{n} \frac{2g}{E_\nu - E_\mu} \right] B_0^\dagger \left(\prod_{\eta=1(\neq\nu)}^{n} B_\eta^\dagger \right) |0\rangle$$

$$= \sum_{\nu=1}^{n} \left[\sum_{\mu=1(\neq\nu)}^{n} \frac{2g}{E_\mu - E_\nu} \right] B_0^\dagger \left(\prod_{\eta=1(\neq\nu)}^{n} B_\mu^\dagger \right) |0\rangle \ . \tag{B.16}$$

(The last line follows by renaming the dummy indices $\nu \leftrightarrow \mu$ in the second line.) Substituting (B.16) for the second line of (B.13), we conclude that $(H_U - \mathscr{E}_n)|\Psi_n\rangle$ will equal zero provided that Eq. (84) holds, namely:

$$1 - \sum_{j}^{U} \frac{g}{2\varepsilon_j - E_\nu} + \sum_{\mu=1(\neq\nu)}^{n} \frac{2g}{E_\mu - E_\nu} = 0, \quad \text{for } \nu = 1, \dots, n \ . \tag{B.17}$$

This constitutes a set of n coupled equations for the n parameters $E_1, \dots, E_n$, which may be thought of as self-consistently determined pair energies. Eq. (B.17) can be regarded as a generalization of the true-boson eigenvalue equation (B.6). It is truly remarkable that the exact eigenstates of a complicated many-body problem can be constructed by such a simple generalization of the solution of a quadratic (i.e. noninteracting) true-boson Hamiltonian!

B.2. Solving the eigenvalue equation numerically

In this section we discuss some algebraic transformations invented by Richardson to render Eqs. (84) or (B.17) less singular and thereby simplify their numerical solution.

First note that the solutions of Eqs. (84) can be given an electrostatic interpretation [116,54], since they are the extrema w.r.t. E_ν of the function

$$W[\{E_\nu\}] \equiv \sum_\nu \frac{E_\nu}{2g} + \frac{1}{2} \sum_\nu \sum_j^{U} \log|2\varepsilon_j - E_\nu| - \sum_{\nu < \mu} \log|E_\mu - E_\nu| \ , \tag{B.18}$$

where E_ν is regarded as a complex number, say $E_{\nu x} + iE_{\nu y}$, and $|E_\nu| = \sqrt{E_{\nu x}^2 + E_{\nu y}^2}$, etc. $W[\{E_\nu\}]$ can thus be interpreted as the potential energy of n free unit charges at the locations $(E_{\nu x}, E_{\nu y})$ in the xy-plane (actually lines of charge perpendicular to the plane), under the combined influence of a uniform external field $-(1/2g)\hat{x}$ parallel to the x-axis, and the field of a set of fixed charges of strength $-1/2$, located at the points $2\varepsilon_j$ on the x-axis. The task at hand is to find the (unstable) equilibrium positions of n free unit charges.

We shall consider only the case that all the ε_j are nondegenerate (degeneracies are considered in [116]). Then every solution of Eqs. (84), i.e. every set $\{E_\nu\}$, coincides at $g = 0$ with a set of n bare energies, say $\{2\varepsilon_{j_\nu}\}$ (by inspection; or in the electrostatic analogy, the field $-1/2g$ is so strong that the only stable configurations are those where each positive unit charges sits infinitely close to some fixed $-1/2$ "host" charge). Richardson showed [51] that as g is turned on, all E_ν's initially smoothly evolve toward lower values (as the field weakens, the free charges leave their hosts, pulled in the negative x direction by the field); but as g increases further, singular points are reached at which two roots approach both each other (and a bare energy), say $E_\mu = E_\nu(=2\varepsilon_{\bar{j}})$, and then turn complex, becoming a complex conjugate pair, so that $\mathscr{E}_n$ remains real (as the field weakens some more, the repulsion between the positive charges begins to become important; it tends to push some of them off the x-axis into the xy plane, but does so without breaking reflection symmetry about the x-axis).

Due to this complication, it is convenient to parametrize the roots in terms of purely real variables, proceeding as follows: Denote roots destined to become conjugates by (E_{2a-1}, E_{2a}) [with $g = 0$ values $(\varepsilon_{j2a-1}, \varepsilon_{j2a})$, say], with $a = 1, \ldots, n/2$ for even n, with one further purely real root, say E_0, for odd n. Let us write $E_{2a-1} = \xi_a - i\eta_a$, $E_{2a} = \xi_a + i\eta_a$, where ξ_a and η_a^2 are purely real, with $\eta_a^2 \lessgtr 0$ for $g \lessgtr g_a$, the critical value where this pair of roots becomes complex. Rewriting Richardson's eigenvlaue equation (84) in terms of these purely real variables by adding and subtracting the equations for $v = 2a - 1$ and $2a$, one readily finds, for each a,

$$\frac{1}{g} - \sum_{j \neq 2a, 2a-1}^{U} \frac{2\varepsilon_j - \xi_a}{(2\varepsilon_j - \xi_a)^2 + y_a(x_a^2 - \rho_a^2)} - \frac{2x_a(1 + y_a)}{\rho_a^2(1 - y_a)^2 - x_a^2(1 + y_a)^2}$$
$$+ \frac{2(E_0 - \xi_a)}{(E_0 - \xi_a)^2 + \eta_a^2} + \sum_{b \neq a}^{n/2} \frac{4\xi_{ba}(\xi_{ba}^2 + \eta_b^2 + \eta_a^2)}{(\xi_{ba}^2 + \eta_b^2 + \eta_a^2)^2 - 4\eta_b^2\eta_a^2} = 0 \,, \tag{B.19}$$

$$\frac{(1 - y_a^2)}{\rho_a^2(1 - y_a)^2 - x_a^2(1 + y_a)^2} + \sum_{j \neq 2a, 2a-1} \frac{y_a}{(2\varepsilon_j - \xi_a)^2 + \eta_a^2}$$
$$- \frac{y_a}{(E_0 - \xi_a)^2 + \eta_a^2} - \sum_{b \neq a} \frac{(\xi_{ba}^2 - \eta_b^2 + \eta_a^2)4ya}{(\xi_{ba}^2 + \eta_b^2 + \eta_a^2)^2 - 4\eta_b^2\eta_a^2} = 0 \,, \tag{B.20}$$

where we introduced the further real variables

$$x_a = \xi_a - \varepsilon_{2a-1} - \varepsilon_{2a}, \qquad y_a = \eta_a^2/(x_a^2 - \rho_a^2) \,, \tag{B.21}$$

$$\rho_a = \varepsilon_{2a} - \varepsilon_{2a-1}, \qquad \xi_{ba} \equiv \xi_b - \xi_a \tag{B.22}$$

and factored out an overall factor of $(x_a^2 - \rho_a^2)$ from Eq. (B.20). Moreover, the terms involving E_0 occur only if n is odd and should be omitted if n is even.

Compared to Eqs. (84), Eqs. (B.19) and (B.20) have the advantage that their roots are always real and that no singularities occur any more. They can thus be solved straightforwardly for the x_a's and y_a's by standard numerical techniques (e.g. Broyden's algorithm [188] in the form given in Ref. [189]), using the set $\mathcal{R} = \{(\varepsilon_{j2a-1}, \varepsilon_{j2a}), \varepsilon_{j0}\}$ as "initial solution". Note, however, that the choice of initial pairings in $\mathcal{R}$ is crucial: an incorrect choice results in the failure of solutions with real x_a, y_a to exist beyond a certain g-value. Since making this choice requires knowledge of which roots end up as conjugate pairs, some trial and error may be involved in finding the correct $\mathcal{R}$ (see, e.g., Ref. [15] of [39]).

B.3. Correlation functions

In Ref. [50], Richardson derived the following explicit results for the normalization constant $\mathcal{N}$ of (83) and the occupation probabilities $\bar{v}_j^2$ and correlators C_{ij} of Eq. (49):

$$\mathcal{N} = [\det(M)]^{-1/2} \,, \tag{B.23}$$

$$\bar{v}_j^2 \equiv \langle \Psi_n | b_j^\dagger b_j | \Psi_n \rangle = \sum_\mu \alpha_\mu E_{j\mu}^2, \quad \alpha_\mu \equiv -g^2(\mathrm{d}E_\mu/\mathrm{d}g) \,, \tag{B.24}$$

$$C_{ij} = \mathcal{N}^2 \sum_{\mu\nu} E_{i\mu} E_{j\nu} I_{\mu\nu}(ij) \,, \tag{B.25}$$

where we used the following abbreviations:

$$M_{\mu v} = \begin{cases} C_{\mu^2} - 2S_{\mu^2} & \text{for } \mu = v \,, \\ 2E_{\mu v}^2 & \text{for } \mu \neq v \,, \end{cases} \tag{B.26}$$

$$E_{j\mu} \equiv (2\varepsilon_j - E_\mu)^{-1}, \quad E_{\mu v} \equiv (E_\mu - E_v)^{-1} \,, \tag{B.27}$$

$$C_{\mu^2} \equiv \sum_j E_{j\mu}^2, \quad S_\mu \equiv \sum_\eta^{n\backslash\mu} E_{\eta\mu}, \quad S_{\mu^2} \equiv \sum_\eta^{n\backslash\mu} E_{\eta\mu}^2 \,, \tag{B.28}$$

$$I_{11'}(ij) \equiv S'_{n-1} \begin{vmatrix} 1 & 2E_{12}E_{1'2'} & \cdots & 2E_{1n}E_{1'n'} \\ 1 & \tilde{D}_{22'}(ij) - 2S_2S_{2'} & \cdots & 2E_{2n}E_{2'n'} \\ \vdots & & & \\ 1 & 2E_{n2}E_{n'2'} & \cdots & \tilde{D}_{nn'}(ij) - 2S_nS_{n'} \end{vmatrix} \tag{B.29}$$

$$\tilde{D}_{\mu v}(ij) \equiv D_{\mu v} + (E_{i\mu}E_{jv} - E_{i\mu}E_{iv} - E_{j\mu}E_{jv}) \,, \tag{B.30}$$

$$D_{\mu v} \equiv \sum_j E_{j\mu}E_{jv} = -4E_{\mu v}^2 + \sum_\eta^{n\backslash\mu,v} 2E_{\eta\mu}E_{\eta v} \,. \tag{B.31}$$

Here M is a matrix with matrix indices μ, v. The notation $\sum_\mu^{n\backslash v}$ means a sum over all $\mu = 1, \ldots, n$, excluding v. The coefficients α_μ in Eq. (B.24) can most conveniently be found by solving the set of n algebraic equations

$$(C_{\mu^2} - 2S_{\mu^2})\alpha_\mu + \sum_\eta^{n\backslash\mu} E_{\eta\mu}^2\alpha_\eta = 1 \quad (\mu = 1, \ldots, n) \,, \tag{B.32}$$

which follow from differentiating Eq. (84) with respect to g. In Eq. (B.29), S'_{n-1} is an operator which symmetrizes the $(n-1)$ primed indices $\eta' = 2', \ldots, n'$ occurring in the determinant for $I_{11'}(ij)$, and then sets $\eta' = \eta$; in other words, S' stands for the sum over the $(n-1)!$ permutations of these indices, and for $\eta' = 2', \ldots, n'$, the prime on η' is only a mark to distinguish it from η for the purposes of symmetrization. (To obtain $I_{\mu\mu'}(ij)$, simply make the replacements $1 \leftrightarrow \mu$ and $1' \leftrightarrow \mu'$ in the determinant written in Eq. (B.29) for $I_{11'}(ij)$, then symmetrize over the indices $\eta' = 1', \ldots, n'$, excluding μ'.) The need to evaluate $(n-1)!$ permutations makes C_{ij} practically impossible to calculate numerically for n much larger than, say, 10. However, Richardson showed that Eq. (B.25) simplifies to the (tractable)

$$C_{ij} \simeq \sum_{\mu v} \alpha_\mu E_{i\mu}E_{j\mu} \,, \tag{B.33}$$

if $\tilde{D}_{\mu v}(ij)$ is approximated by $D_{\mu v}$ in Eq. (B.29). The accuracy of this approximation can be judged by using the fact [51] that when used to evaluate $\mathscr{E}_n - \langle \Psi_n|H_U|\Psi_n \rangle$, it produces, instead of zero, the result $(1/g)\sum_v^n \alpha_v[g^2(\sum_j^U E_{jv})^2 - 1]$, which can easily be evaluated numerically.

References

[1] A.T. Johnson, L.P. Kouwenhoven, W. de Jong, N.C. van der Vaart, C.J.P.M. Harmans, C.T. Foxon, Phys. Rev. Lett. 69 (1992) 1592.
[2] V. Goldman, B. Su, J.E. Cunningham, Int. Nat. J. Mod. Phys. B 6 (1992) 2321.
[3] L.L. Sohn, L.P. Kouwenhoven, G. Schön (Eds.), in: Mesoscopic Electron Transport, NATO ASI Series, Vol. 345, Kluwer Academic Publishers, Dordrecht, 1997.
[4] L.P. Kouwenhoven, C.M. Marcus, P.L. McEuen, S. Tarucha, R.M. Westervelft, N.S. Wingreen, in: L.L. Sohn, L.P. Kouwenhoven, G. Schön (Eds.), Mesoscopic Electron Transport, NATO ASI Series, Vol. 345, Kluwer Academic Publishers, Dordrecht, 1997, p. 105.
[5] D.C. Ralph, C.T. Black, M. Tinkham, Phys. Rev. Lett. 74 (1995) 3241. ∗∗∗
[6] C.T. Black, D.C. Ralph, M. Tinkham, Phys. Rev. Lett. 76 (1996) 688. ∗
[7] D.C. Ralph, C.T. Black, M. Tinkham, Physica (Amsterdam) 218B (1996) 258. ∗
[8] C.T. Black, Ph.D. Thesis, Harvard University, 1996.
[9] D.C. Ralph, C.T. Black, J.M. Hergenrother, J.G. Lu, M. Tinkham, in: L.L. Sohn, L.P. Kouwenhoven, G. Schön (Eds.), Mesoscopic Electron Transport, NATO ASI Series, Vol. 345, Kluwer Academic Publishers, Dordrecht, 1997, p. 447.
[10] D.C. Ralph, C.T. Black, M. Tinkham, Phys. Rev. Lett. 78 (1997) 4087. ∗∗∗
[11] O. Agam, N.S. Wingreen, B.L. Altshuler, D.C. Ralph, M. Tinkham, Phys. Rev. Lett. 78 (1997) 1956. ∗∗∗
[12] O. Agam, I.L. Aleiner, Phys. Rev. B 56 (1997) 5759. ∗∗
[13] O. Agam, in: I.V. Lerner, J.P. Keating, D.E. Khmelnitskii (Eds.), Supersymmetry and Trace Formulae: Chaos and Disorder, Plenum Press, New York, 1999, p. 13.
[14] D.G. Salinas, S. Guéron, D.C. Ralph, C.T. Black, M. Tinkham, Phys. Rev. B 60 (1999) 6137. ∗∗∗
[15] D. Davidović, M. Tinkham, Phys. Rev. Lett. 83 (1999) 1644. ∗∗∗
[16] D. Davidović, M. Tinkham, cond-mat/9910396.
[17] R. Desmicht, G. Faini, V. Cros, A. Fert, F. Petroff, A. Vaurés, Appl. Phys. Lett. 72 (1998) 386.
[18] S. Guéron, M.M. Deshmukh, E.B. Myers, D.C. Ralph, Phys. Rev. Lett. 83 (1999) 4148. ∗∗∗
[19] J.A.A.J. Perenboom, P. Wyder, F. Meier, Phys. Rep. 78 (1981) 173.
[20] W.P. Halperin, Rev. Mod. Phys. 58 (1986) 533.
[21] J. von Delft, A.D. Zaikin, D.S. Golubev, W. Tichy, Phys. Rev. Lett. 77 (1996) 3189. ∗
[22] F. Braun, J. von Delft, D.C. Ralph, M. Tinkham, Phys. Rev. Lett. 79 (1997) 921. ∗∗
[23] F. Braun, J. von Delft, Phys. Rev. B 59 (1999) 9527. ∗
[24] R.A. Smith, V. Ambegaokar, Phys. Rev. Lett. 77 (1996) 4962. ∗
[25] R. Balian, H. Flocard, M. Vénéroni, cond-mat/9802006.
[26] R. Balian, H. Flocard, M. Vénéroni, Phys. Rep. 317 (1999) 251.
[27] M.C. Bønsager, A.H. MacDonald, cond-mat/9812042.
[28] K.A. Matveev, A.I. Larkin, Phys. Rev. Lett. 78 (1997) 3749. ∗∗
[29] R. Rossignoli, N. Canosa, P. Ring, Phys. Rev. Lett. 80 (1998) 1853.
[30] R. Rossignoli, N. Canosa, P. Ring, Ann. Phys. (NY) 275 (1999) 1. ∗∗
[31] R. Rossignoli, J.P. Zagorodny, N. Canosa, Phys. Lett. A 258 (1999) 188.
[32] R. Rossignoli, private communication, 2000.
[33] A. Mastellone, G. Falci, R. Fazio, Phys. Rev. Lett. 80 (1998) 4542. ∗
[34] S.D. Berger, B.I. Halperin, Phys. Rev. B 58 (1998) 5213. ∗
[35] F. Braun, J. von Delft, Phys. Rev. Lett. 81 (1998) 4712. ∗
[36] J. Dukelsky, G. Sierra, Phys. Rev. Lett. 83 (1999) 172.
[37] J. Dukelsky, G. Sierra, Phys. Rev. B 61 (2000) 12302. ∗∗
[38] F. Braun, J. von Delft, in: B. Kramer (Ed.), Advances in Solid State Physics, Vol. 39, Vieweg, Braunschweig, 1999, p. 341.
[39] G. Sierra, J. Dukelsky, G.G. Dussel, J. von Delft, F. Braun, Phys. Rev. B 61 (2000) 11890. ∗∗
[40] J. von Delft, F. Braun, cond-mat/9911058, in: I.O. Kulik, R. Ellialtioğlu (Eds.), Proceedings of the NATO ASI, Quantum Mesoscopic Phenomena and Mesoscopic Devices in Microelectronics, Ankara/Antalya, Turkey, June 1999, Kluwer Academic Publishers, Dordrecht, 2000, p. 361. ∗∗∗

[41] J. Dukelsky, P. Schuck, Phys. Lett. B 464 (1999) 164.
[42] G.-S. Tian, L.-H. Tang, Phys. Rev. B 58 (1998) 12333.
[43] K. Tanaka, F. Marsiglio, cond-mat/9902173.
[44] A. Di Lorenzo, R. Fazio, F.W. Hekking, G. Falci, A. Mastellone, G. Giaquinta, Phys. Rev. Lett. 84 (2000) 550. ∗∗
[45] P.W. Anderson, J. Phys. Chem. Solids 11 (1959) 28. ∗
[46] R.W. Richardson, Phys. Lett. 3 (1963) 277.
[47] R.W. Richardson, Phys. Lett. 5 (1963) 82.
[48] R.W. Richardson, N. Sherman, Nucl. Phys. 52 (1964) 221. ∗∗∗
[49] R.W. Richardson, Phys. Lett. 14 (1965) 325.
[50] R.W. Richardson, J. Math. Phys. 6 (1965) 1034. ∗∗
[51] R.W. Richardson, Phys. Rev. 141 (1966) 949.
[52] R.W. Richardson, Phys. Rev. 144 (1966) 874.
[53] R.W. Richardson, Phys. Rev. 159 (1967) 792.
[54] R.W. Richardson, J. Math. Phys. 18 (1977) 1802.
[55] K.A. Matveev, L.I. Glazman, A.I. Larkin, Phys. Rev. Lett. 85 (2000) 2789. ∗∗
[56] P.W. Brouwer, X. Waintal, B.I. Halperin, Phys. Rev. Lett. 85 (2000) 369. ∗∗
[57] W.B. Thimm, J. Kroha, J. von Delft Phys. Rev. Lett. 82 (1999) 2143. ∗∗
[58] D.V. Averin, K.K. Likharev, in: B.L. Altshuler, P.A. Lee, R.A. Webb (Eds.), Mesoscopic Phenomena in Solids, Elsevier, New York, 1991, p. 169.
[59] H. Grabert, M.H. Devoret (Eds.), Single Charge Tunneling, Plenum Press, New York, 1992.
[60] G. Schön, in: T. Dittrich, P. Hänggi, G. Ingold, B. Kramer, G. Schön, W. Zwerger (Eds.), Quantum Transport and Dissipation, Wiley-VCH Verlag, New York, 1998.
[61] D.V. Averin, A.N. Korotkov, J. Low Temp. Phys. 80 (1990) 173. ∗∗
[62] A.E. Hanna, M. Tinkham, Phys. Rev. 41 (1991) 5919.
[63] J.G. Lu, J.M. Hergenrother, M. Tinkham, Phys. Rev. B 57 (1998) 4591.
[64] B.L. Altshuler, Y. Gefen, A. Kamenev, L.S. Levitov, Phys. Rev. Lett. 78 (1997) 2803. ∗∗
[65] V. Falko, K. Efetov, Phys. Rev. B 50 (1994) 11267.
[66] U. Sivan, Y. Imry, A.G. Aronov, Europhys. Lett. 28 (1994) 115.
[67] O. Agam, B.L. Altshuler, A.V. Andreev, Phys. Rev. Lett. 75 (1995) 4389.
[68] Ya.M. Blanter, Phys. Rev. B 54 (1996) 12807.
[69] Ya.M. Blanter, A.D. Mirlin, B.A. Muzykanskii, Phys. Rev. Lett. 78 (1997) 2449.
[70] Ya.M. Blanter, A.D. Mirlin, Phys. Rev. E 55 (1997) 6514.
[71] I.L. Aleiner, L.I. Glazman, Phys. Rev. B 57 (1998) 9608.
[72] H. Schoeller, in: L.L. Sohn, L.P. Kouwenhoven, G. Schön (Eds.), Mesoscopic Electron Transport, NATO ASI Series, Vol. 345, Kluwer Academic Publishers, Dordrecht, 1997, p. 291.
[73] H. Schoeller, G. Schön, Phys. Rev. B 50 (1994) 18,436.
[74] J. König, H. Schoeller, G. Schön, Europhys. Lett. 31 (1995) 31.
[75] J. König, H. Schoeller, G. Schön, Phys. Rev. Lett. 76 (1996) 1715.
[76] J. König, J. Schmid, H. Schoeller, G. Schön, Phys. Rev. B 54 (1996) 16820.
[77] H. Schoeller, Habilitationsthesis, Universität Karlsruhe, 1997.
[78] J. König, Ph.D. Thesis, Universität Karlsruhe, 1998.
[79] K.S. Ralls, R.A. Buhrman, R.C. Tiberio, Appl. Phys. Lett. 55 (1989) 2459.
[80] I. Giaever, H.R. Zeller, Phys. Rev. Lett. 20 (1968) 1504.
[81] H.R. Zeller, I. Giaever, Phys. Rev. 181 (1969) 789.
[82] D. Davidović, M. Tinkham, Appl. Phys. Lett. 73 (1998) 3959.
[83] A. Levy Yeyati et al., Phys. Rev. B 55 (1997) R6137.
[84] M.T. Tuominen, J.M. Hergenrother, T.S. Tighe, M. Tinkham, Phys. Rev. Lett. 69 (1992) 1997.
[85] M.T. Tuominen, J.M. Hergenrother, T.S. Tighe, M. Tinkham, Phys. Rev. B 47 (1993) 11599.
[86] M. Tinkham, J.M. Hergenrother, J.G. Lu, Phys. Rev. B 51 (1995) 12649.
[87] P. Lafarge, P. Joyez, D. Esteve, C. Urbina, M.H. Devoret, Phys. Rev. Lett. 70 (1993) 994.
[88] T.M. Eiles, J.M. Martinis, M.H. Devoret, Phys. Rev. Lett. 70 (1993) 1862.

[89] A.I. Larkin, Sov. Phys. JETP 21 (1965) 153 [JETP (USSR) 48 (1965) 232].
[90] L.I. Glazman, private communication.
[91] F. Braun, Ph.D. Thesis, Universität Karlsruhe, 1999. *
[92] M. Tinkham, Introduction to Superconductivity, 2nd Edition, McGraw-Hill, New York, 1996.
[93] M. Strongin, R.S. Thompson, O.F. Kammerer, J.E. Crow, Phys. Rev. B 1 (1970) 1078.
[94] B. Mühlschlegel, D.J. Scalapino, R. Denton, Phys. Rev. B 6 (1972) 1767. *
[95] B. Mühlschlegel, Physica B 203 (1994) 444.
[96] A. Kawabata, J. Low Temp. Phys. 38 (1980) 191.
[97] A. Kawataba, Surf. Sci. 106 (1981) 358.
[98] R. Meservey, P.M. Tedrow, P. Fulde, Phys. Rev. Lett. 25 (1970) 1270.
[99] R. Meservey, P.M. Tedrow, Phys. Rep. 238 (1994) 173. *
[100] S. Bahcall, unpublished, 1996.
[101] L.N. Cooper, J. Bardeen, J.R. Schrieffer, Phys. Rev. 108 (1957) 1175.
[102] N.W. Ashcroft, N.D. Mermin, Solid State Physics, Saunders, Philadelphia, 1976.
[103] J.W. Garland, K.H. Bennemann, F.M. Mueller, Phys. Rev. Lett. 21 (1968) 1315.
[104] V.G. Soloviev, Mat. Fys. Skrif. Kong. Dan. Vid. Selsk. 1 (1961) 1.
[105] G. Rickayzen, Theory of Superconductivity, Wiley, New York, 1965 (Chapter 4).
[106] P. Ring, P. Schuck, The Nuclear Many-Body Problem, Springer, Berlin, 1980.
[107] H.J. Lipkin, Ann. Phys. (NY) 9 (1960) 272.
[108] B. Mühlschlegel, J. Math. Phys. 3 (1962) 522.
[109] B. Jankó, A. Smith, V. Ambegaokar, Phys. Rev. B 50 (1994) 1152.
[110] A.M. Clogston, Phys. Rev. Lett. 9 (1962) 266.
[111] B.S. Chandrasekhar, Appl. Phys. Lett. 1 (1962) 7.
[112] D.S. Golubev, A.D. Zaikin, Phys. Lett. A 195 (1994) 380; in: H.A. Cerdeira et al. (Eds.), Quantum Dynamics of Submicron Structures, Kluwer Academic Publishers, Dordrecht, 1995, p. 473.
[113] A.A. Abrikosov, L.P. Gorkov, I.E. Dzyaloshinksii, Methods of Quantum Field Theory in Statistical Physics, Prentice-Hall, Englewood Cliffs, NJ, 1963.
[114] D.L. Klein, R. Roth, A.K.L. Lim, A.P. Alivisatos, P.L. McEuen, Nature 389 (1997) 699.
[115] R.W. Richardson, private communication, 1998.
[116] M. Gaudin, Etats et Valeurs Propres de l'Hamiltonien d'Appariement, Internal Report D.Ph.T/DOC-11/DD, Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, BP. no 2-92-GIF-sur-YVETTE, 1968.
[117] M.C. Cambiaggio, A.M.F. Rivas, M. Saraceno, Nucl. Phys. A 624 (1997) 157.
[118] G. Sierra, Nucl. Phys. B 572 (2000) 517.
[119] R.W. Richardson, private communication, 1999.
[120] J. von Delft, Ann. Physik (Leipzig) 10 (3) (2001), in press.
[121] B.F. Bayman, Nucl. Phys. 15 (1960) 33.
[122] K. Dietrich, H.J. Mang, J.H. Pradal, Phys. Rev. 135 (1964) B22.
[123] C.W. Ma, J.O. Rasmussen, Phys. Rev. C 16 (1977) 16.
[124] C. Essebag, J.L. Egido, Nucl. Phys. A 522 (1993) 205.
[125] Moshe Schechter, private communication, 2000.
[126] K. Tanabe, S. Tanabe, H.J. Mang, Nucl. Phys. A 357 (1981) 20.
[127] Y. Alhassid, J. Zingman, Phys. Rev. C 30 (1984) 684.
[128] Y. Alhassid, B. Bush, Nucl. Phys. A 549 (1992) 43.
[129] P. Arve, G.F. Bertsch, B. Lauritzen, G. Puddu, Ann. Phys. (NY) 183 (1988) 309.
[130] B. Lauritzen, P. Arve, G.F. Bertsch, Phys. Rev. Lett. 61 (1988) 2835.
[131] R. Rossignoli, P. Ring, N. Dang, Phys. Lett. B 297 (1992) 1061.
[132] R. Rossignoli, A. Ansari, P. Ring, Phys. Rev. Lett. B 70 (1993) 9.
[133] R. Rossignoli, P. Ring, Ann. Phys. (NY) 235 (1994) 350.
[134] G. Puddu, P.F. Bortignon, R.A. Broglia, Phys. Rev. C 42 (1990) 1830.
[135] B. Lauritzen et al., Phys. Lett. B 246 (1990) 329.
[136] G. Puddu, P.F. Bortignon, R.A. Broglia, Ann. Phys. (NY) 206 (1991) 409.

[137] G. Puddu, Phys. Rev. B 45 (1992) 9882.
[138] H. Attias, Y. Alhassid, Nucl. Phys. A 625 (1997) 565.
[139] R. Rossignoli, N. Canosa, Phys. Lett. B 394 (1997) 242.
[140] R. Rossignoli, N. Canosa, Phys. Rev. C 56 (1997) 791.
[141] R. Rossignoli, N. Canosa, P. Ring, Nucl. Phys. A 591 (1995) 15.
[142] R. Rossignoli, Phys. Rev. C 54 (1996) 1230.
[143] R. Rossignoli, N. Canosa, J.L. Egido, Nucl. Phys. A 605 (1996) 1.
[144] G.H. Lang, C.W. Johnson, S.E. Koonin, W.E. Ormand, Phys. Rev. C 48 (1993) 1518.
[145] S.E. Koonin, D.J. Dean, K. Langanke, Phys. Rep. 278 (1997) 1.
[146] Y. Volotikin et al., Nature 384 (1996) 621.
[147] K. Yosida, Phys. Rev. 110 (1958) 769.
[148] Ya.M. Blanter, A.D. Mirlin, Phys. Rev. B 53 (1996) 12601.
[149] M.Yu. Reizer, A.V. Sergeyev, Zh. Eksp. Teor. Fiz. 90 (1986) 1056 [Sov. Phys. JETP 63 (1986) 616].
[150] D. Pines, P. Noziéres, The Theory of Quantum Liquids, Vol. 1, Addison-Wesley, Reading, MA, 1989, p. 309.
[151] J. Buttet, in: Entre l'atome et le cristal: les agrégats, 4ième École d'Été Méditerranéene, Les Editions de Physiques, Paris, 1981, p. 287.
[152] G. Bergmann, Phys. Rep. 107 (1984) 1.
[153] P. Monod, A. Janossy, J. Low Temp. Phys. 26 (1977) 311.
[154] V.E. Kravtsov, M.R. Zirnbauer, Phys. Rev. B 46 (1992) 4332.
[155] C.P. Slichter, Principles of Magnetic Resonance, Springer, Berlin, 1980.
[156] S. Tarucha, D.G. Austing, T. Honda, R.J. van der Hage, L.P. Kouwenhoven, Phys. Rev. Lett. 77 (1996) 3613.
[157] P.W. Brouwer, Y. Oreg, B.I. Halperin, Phys. Rev. B 60 (1999) R13 977.
[158] H.U. Baranger, D. Ullmo, L.I. Glazman, Phys. Rev. B 61 (2000) R2425.
[159] P.W. Brouwer, Y. Oreg, B.I. Halperin, Phys. Rev. B 60 (1999) 13977.
[160] K. Ono, H. Shimada, Y. Ootuka, J. Phys. Soc. Japan 66 (1997) 1261.
[161] J. Barnaś, A. Fert, Phys. Rev. Lett. 80 (1998) 1058.
[162] A. Takahashi, S. Maekawa, Phys. Rev. Lett. 80 (1998) 1758.
[163] K. Majumdar, S. Hershfield, Phys. Rev. B 57 (1998) 11521.
[164] H. Shimada, K. Ono, Y. Ootuka, J. Phys. Soc. Japan 67 (1998) 2852, 1359.
[165] A.N. Korotkov, V.I. Safarov, Phys. Rev. B 59 (1999) 89.
[166] J. Barnaś, A. Fert, Europhys. Lett. 44 (1998) 85.
[167] A. Brataas, Y.V. Nazarov, U. Inoue, G.E.W. Bauer, Phys. Rev. B 59 (1999) 93.
[168] H. Imamura, S. Takahashi, S. Maekawa, Phys. Rev. Lett. 82 (1999) 3911.
[169] S. Takahashi, H. Imamura, S. Maekawa, Phys. Rev. B 59 (1999) 6017.
[170] L.F. Schelp, A. Fert, F. Fettar, P. Holody, S.F. Lee, J.L. Maurice, F. Petroff, A. Vaurès, Phys. Rev. B 56 (1997) R5747.
[171] R. Desmicht, G. Faini, V. Cros, A. Fert, F. Petroff, A. Vaurès, Appl. Phys. Lett. 72 (1998) 386.
[172] D.A. Papaconstantopoulos, Handbook of the Band Structure of Elemental Solids, Plenum, New York, 1986.
[173] S. Kleff, Diplomthesis, Universität Karlsruhe, 2000.
[174] M. Lederman, S. Schultz, M. Ozaki, Phys. Rev. Lett. 73 (1994) 1986.
[175] J.G.S. Lok, A.K. Geim, J.C. Maan, S.V. Dubonos, L.T. Kuhn, P.E. Lindelof, Phys. Rev. B 58 (1998) 12201.
[176] W. Wernsdorfer, E. Bonet Orozco, K. Hasselbach, A. Benoît, B. Barbara, N. Demoncy, A. Loiseau, Phys. Rev. Lett. 78 (1997) 1791.
[177] E. Bonet, W. Wernsdorfer, B. Barbara, A. Benoît, D. Mailly, A. Thiaville, Phys. Rev. Lett. 83 (1999) 4188.
[178] E.M. Chudnovsky, J. Tejada, Macroscopic Quantum Tunneling of the Magnetic Moment, Cambridge Studies in Magnetism, Cambridge University Press, Cambridge, 1998.
[179] P.W. Anderson, Phys. Rev. B 124 (1961) 41.
[180] A.C. Hewson, The Kondo Problem to Heavy Fermions, Cambridge University Press, Cambridge, 1993.
[181] N. Grewe, H. Keiter, Phys. Rev. B 24 (1981) 4420.
[182] Y. Kuramoto, Z. Phys. B 53 (1983) 37.
[183] M. Dolg et al., J. Chem. Phys. 94 (1991) 3011.

[184] M. Dolg et al., Chem. Phys. 195 (1995) 71.
[185] U. Fano, Phys. Rev. 124 (1961) 1866.
[186] S. Sun, C.B. Murray, J. Appl. Phys. 85 (1999) 4325. ∗
[187] A. Bezryadin, C. Dekker, G. Schmid, Appl. Phys. Lett. 71 (1997) 1273.∗
[188] C.G. Broyden, Math. Comput. 19 (1965) 577.
[189] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipies in C, The Art of Scientific
       Computing, 2nd Edition, Cambridge University Press, Cambridge, 1992.