An Introduction to the Kondo Effect

Tim Child, Martin Cross, Alexandra Tully, Max Werner
(Dated: November 23, 2018)

The following paper discusses the resistance minima found at low temperatures in metals doped with magnetic impurities, known today as the Kondo effect[1]. It provides a brief introduction to the history and main principles behind the effect, touches on the theory that gave this effect the Kondo name, presents two experimental examples of the Kondo effect, and succinctly reviews its relevance in the physics world today. Any fault, flaws, or logical fallacies are solely attributable to the authors of this paper (and please do let us know if you find any!).

I. INTRODUCTION

In 1934, at the Kamerlingh Onnes Laboratory in the Netherlands, W. J. de Haas, J. de Boer and G.J. van den Berg discovered that the resistance curve of impure gold wires has a minimum at low temperatures (see Fig. 1) [2]. At the time, resistivity was understood to be modelled by Matthiessen’s rule describing the scattering of electrons by lattice vibrations and impurities; resistivity, therefore, was expected to decrease monotonically with temperature[3]. De Haas’ gold wires, however, were not displaying the expected behaviour at low temperatures. In subsequent years, the phenomenon was discovered in a wide range of metals. From the beginning, it was linked to impurities; later, it was found these impurities needed to be magnetic[3]. All in all, it would take 30 years for someone to understand and explain the theory behind this phenomenon. In 1964, Jun Kondo showed that an alternative electron scattering process was taking place: at temperatures on the order of a few Kelvin, magnetic impurities give rise to a scattering process which involves a temporary exchange of spin state between the conduction electron and magnetic impurity. Kondo discovered that when using perturbation theory to model the scattering from a magnetic impurity interacting with the spins of conducting electrons, the second order term in the calculation can be bigger than the first, resulting in the resistance of a metal increasing logarithmically as the temperature is lowered[4].

An intuitive explanation for this phenomenon is as follows: imagine a magnetic impurity (an ion) sitting in a metal lattice. The impurity’s highest energy electron has an energy less than that of the Fermi sea of the lattice. In some circumstances this electron can tunnel to an unoccupied state near the surface of the Fermi sea, however, this would require a substantial amount of energy (on the order of several eV)[4]. This scenario can occur spontaneously, by virtue of Heisenberg’s uncertainty principle: the magnetic impurity’s electron can spontaneously tunnel from the impurity to the lattice’s Fermi sea if, within the timescale limited by the uncertainty principle, an electron from the Fermi sea tunnels to the available state in the impurity. In the latter case, however, the spins of these two electrons do not need to be identical - we can have a spin exchange. The spin exchange changes the overarching energy spectrum of the system, and as a many-body phenomenon (i.e. if we take enough of these spin exchanges together) this renders the system in a new state. This new state, with the same energy as the Fermi level, is called a Kondo resonance, and it effectively scatters electrons with energies also near the Fermi level - conduction electrons. As these conduction electrons are scattered, we see a commensurate increase in resistivity in the bulk state of the material[4]. This phenomenon only occurs at low temperatures because the Fermi distribution (of filled vs. unfilled states) broadens as temperature increases. In order for the spin exchange to take place, and have a significant effect compared to other scattering process, the conduction electrons must have energies close to that of the Fermi energy.

Although the Kondo effect was explained over 50 years ago, it remains of interest as a well-defined many-body problem comprising a strongly coupled system. The Kondo effect serves as a testing ground for newly developed analytical and numerical techniques, and it can provide insight into strongly-correlated systems such as high-T_c superconductors. Additionally, recent developments in nanotechnology have led to unprecedented levels of control in these experiments, making them all the more appealing and hopefully useful. This paper will touch on the theory behind the Kondo effect, provide two examples of experimental techniques frequently used to create and probe these systems (quantum dots and scanning tunnelling microscopy), and discuss the role of Kondo effect experiments in the world of physics today.
II. THEORY

As the temperature of a solid-state system is lowered, the dimensionality of k-space that conduction electrons occupy reduces to a quasi-2D layer above the Fermi surface. This reduction in k-space leads to a greater probability of interaction - in BCS theory we have electron-phonon interaction leading to the formation of Cooper pairs, similarly, the Kondo effect results from an increased interaction between the spins of localized (on an impurity) and free electrons, which ultimately leads to the formation of net-zero-spin quasi-particles. In this low-temperature limit, we cannot ignore higher-order perturbations. That is, in addition to single-step scattering events, we must consider other pathways, where the electron briefly occupies an intermediate state. We will follow Kondo’s derivation [1], starting with the Hamiltonian of the unperturbed system:

\[ H_0 = \sum_{k\sigma} \epsilon_k a_{k\sigma}^\dagger a_{k\sigma}, \]

where \( \epsilon_k \) is the energy of a conduction electron with momentum \( k \) and spin \( \sigma \), with associated creation and annihilation operators \( a_{k\sigma}^\dagger \) and \( a_{k\sigma} \). We consider a s-d exchange model with the following perturbative Hamiltonian:

\[ H' = -(J/N) \sum_{nkk'} \exp[i(k - k') \cdot R_n] \]

\[ \times \left[ (a_{k\uparrow}^\dagger a_{k\uparrow} - a_{k'\uparrow}^\dagger a_{k'\uparrow}) S_{nz} \cdots \right. \]

\[ + a_{k'\downarrow}^\dagger a_{k\downarrow} S_{n-} + a_{k\downarrow}^\dagger a_{k'\downarrow} S_{n+} \].

Here, \( R_n \) is the position of the \( n \)-th impurity, with associated z-axis spin projection \( S_{nz} \), and raising/lowering operators \( S_{n\pm} \); \( N \) is the total number of atoms, and \( J \) is an experimentally determined constant relating the interaction strength. The transition probability per unit time from an initial state \( a \) to a final state \( b \) is given to second order in the Born approximation by

\[ W(a \to b) = (2\pi/\hbar) \delta(E_a - E_b) \]

\[ \times \left[ H'_{ab} H'_{ba} + \sum_{c \neq a} (H'_{ac} H'_{cb} H'_{ba} + \text{h.c.})/(E_a - E_c) \right] \]

where \( a \), \( b \) and \( c \) are total states of the system, with respective energies \( E_a \), \( E_b \) and \( E_c \). The term of order \( H'^2 \) (first order in the Born approximation) describes a single-step scattering process from \( a \) to \( b \), while the latter term, of order \( H'^3 \), includes processes with an intermediate state \( c \). It was known that the first order correction, while notable at low temperatures, was not temperature dependent; Kondo’s significant finding was that expanding to second order reveals terms with \( \log(T) \) temperature dependence.

Focusing on the second order terms, let’s consider the possible two-step scattering processes for a spin-up electron with momentum \( k \). There are two main branches, one in which the spin of the electron is conserved, and the other where the spin is flipped. Within these branches, the impurity can intermediate to conserve the net system-spin. There are of course many permutations, but these four processes broadly capture the themes:

1. The electron with \( k\uparrow \) is first scattered to the unoccupied state \( q'\uparrow \), and then to \( k'\uparrow \).

2. Another electron, \( q\uparrow \), is scattered to \( k'\uparrow \), and then replaced by \( k\uparrow \).

3. There is a temporary spin-flip such that \( k\uparrow \) scatters to \( q'\downarrow \) and then to \( k'\uparrow \). In the intermediate stage, the spin of an impurity located at \( R_n \), denoted \( M_n \), is increased by one to \( M_n + 1 \). The final step then returns this spin back to its original value, \( M_n \).

4. There is a permanent spin-flip such that \( k\uparrow \) scatters to \( q'\downarrow \) and then to \( k'\downarrow \). In this process the intermediate impurity retains its indexed spin \( M_n + 1 \).

The terms in the rate equation relating to these processes, in the above order, are as follows:

\[ \sum_{q'} H'_{k\uparrow,q'\uparrow} H'_{q'\uparrow,k'\uparrow} H'_{k'\uparrow,k\uparrow}(1 - f_{q'\uparrow}^0)/(\epsilon_k - \epsilon_{q'\uparrow}) \]

\[ - \sum_{q} H'_{k\uparrow,q\uparrow} H'_{k\uparrow,q\uparrow} H'_{k'\uparrow,q\uparrow}(f_{q\uparrow}^0)/(\epsilon_q - \epsilon_{k\uparrow}) \]

\[ + \sum_{nq'} H'_{k\uparrow,M_n,q'\downarrow,M_n+1} H'_{q'\downarrow,M_n+1,k'\uparrow,M_n+1} H'_{k'\uparrow,k\uparrow} \cdots \]

\[ \times (1 - f_{q'\uparrow}^0)/(\epsilon_k - \epsilon_{q'\uparrow}) \]

\[ + \sum_{nq} H'_{k\uparrow,M_n,q'\downarrow,M_n+1} H'_{q'\downarrow,M_n+1,k'\downarrow,M_n+1} H'_{k'\downarrow,k\downarrow} \cdots \]

\[ \times (1 - f_{q\uparrow}^0)/(\epsilon_k - \epsilon_{q\uparrow}) \]

where \( f_{q\uparrow}^0 \) is the Fermi distribution for an electron with energy \( \epsilon_q \). The addition of Fermi statistics is necessary to prevent scattering into occupied states, or out of empty ones. As it turns out, this requirement is what introduces temperature dependence into the rate equations. Also of note is the sign difference between the first and second terms, which is due to the anti-symmetry of fermion particle exchange.

The matrix elements \( \langle j | H' | i \rangle \) simplify if we consider the concentration of impurities to be low enough that they do not interact with one another. Taking the third
the formation of a quasi-particle with net zero spin [6].

by interacting (scattered) electrons [5] - in other words,

If we further assume that the scattering is elastic ($\epsilon_k = \epsilon_k'$), and if we ignore terms that do not involve $J_{q'}$, it can be shown that the total scattering rate for the spin preserving processes is

$$W(k^{\uparrow}/\downarrow \rightarrow k'^{\uparrow}/\downarrow) = \frac{2\pi J^2 S(S + 1)c/3\hbar N}{\epsilon - \epsilon_k},$$

and that the spin-flip processes have twice this rate

$$W(k^{\uparrow}/\downarrow \rightarrow k'^{\downarrow}/\uparrow) = 2W(k^{\uparrow}/\downarrow \rightarrow k'^{\uparrow}/\downarrow),$$

where $S$ is the total spin of the impurity, $c$ is the concentration of impurities, and the density of states is

$$g(\epsilon_k) = (1/N) \sum_q J_{q'}/(\epsilon_q - \epsilon_k).$$

Using these scattering rates, we determine the rate of change in the occupation of a particular state $k^{\uparrow}/\downarrow$ due to collisions with the impurities as follows:

$$\frac{\partial f_k^{\uparrow}/\downarrow}{\partial t} \text{coll} = \sum_{k'} W(k^{\uparrow}/\downarrow \rightarrow k'^{\uparrow}/\downarrow)(f_{k'}^{\uparrow}/\downarrow - f_k^{\uparrow}/\downarrow) \cdots$$

It can be shown that the resistivity is then given by

$$\rho_{\text{spin}} = c\rho_M \left[ 1 - \frac{h^2J/\pi mk_0}{f(\epsilon_k)df/\epsilon df} \right],$$

where $\rho_M$ is defined as

$$\rho_M = 3\pi mJ^2 S(S + 1)(V/N)/2e^2\hbar F.$$

Carrying out the integration produces a term proportional to $\log(T)$, giving the final result

$$\rho_{\text{spin}} = c\rho_M \left[ 1 + 3zJ/F \right] \log(T) .$$

If $J$ is negative, this correction accurately describes the minimum and subsequent increase in resistivity as the temperature is lowered. Note, however, that the result is singular as $T$ approaches 0. This characteristic is not observed in real materials, rather, the resistivity plateaus at a temperature now known as the “Kondo temperature,” $T_K$. This discrepancy was dubbed the “Kondo problem,” and only later explained as a screening of the impurity by interacting (scattered) electrons [5] - in other words, the formation of a quasi-particle with net zero spin [6].

III. EXPERIMENTAL TECHNIQUES

A. Quantum Dots

Developments in nanoscale lithography have enabled researchers to create tuneable Kondo effect devices by trapping small numbers of electrons (< 50) in so called quantum dots (QDs). These QDs are effective models of a single magnetic impurity in a bulk metal system while providing the ability to tune (in situ) the voltage bias, magnetic field, temperature, the energy level of the impurity, and the coupling strength between the impurity and the source of the conduction electrons. [7] In order to simulate a single magnetic impurity, QDs must both be distinct from a surrounding material replete with conduction electrons (the Fermi sea) and also have a net spin. A QD comprises a 2D electron gas (2DEG), at an interface of two semiconducting layers of a heterostructure, contained by a series of gold electrical contacts (gates). Voltage applied to these electrical contacts dynamically controls the number of electrons from the 2DEG trapped within a specific QD. Researchers can accurately increase or decrease the number of electrons confined in the dot one at a time, allowing the number of electrons to be changed between odd and even. The surrounding 2DEG is equivalent to the conduction electrons in the bulk metal system. The mechanism for controlling the overall spin of the system is as follows: the electrons fill the energy bands below the fermi level with pairs of spin up and spin down electrons. At low temperatures, an even number of electrons results in no net spin on the dot. An odd number of electrons, on the other hand, results in the highest accessible energy level being occupied by only one electron, giving a net spin of ±1/2 (dependent on the spin of that single electron). With a net spin, and controllably connected to the 2DEG, the quantum dot effectively replicates a magnetic impurity caused by ions in a metal such as the classic dilute copper cobalt alloy system.

In addition to controlling for spin, QDs allow for adjustment of several parameters of the Kondo problem including the coupling strength, voltage bias, and energy level of the impurity. The coupling strength between the dot and the source (and drain) determines the probability of tunnelling from one to another and is set by applying a potential to the electrical contacts coloured in gold in Fig. 2. This forms quantum point contacts or QPCs. The voltage bias is managed through Ohmic electrical contacts that connect to the source and drain of the QD, and, with the coupling strength fixed, it enables conductance measurements of the QD itself. The energy level of the impurity is regulated by the plunger gate a gold electrical contact that uses an applied negative potential (of varying magnitude) to control the number of electrons in the QD. The number of electrons in the dot can be varied from 0 to on the order of 50 and is fixed when the system is at a low enough temperature due to the Coulomb blockade. The Coulomb blockade describes a situation in
which the highest occupied energy state of the quantum dot lies significantly below the Fermi energy of the system, while the next energy level corresponding to adding another electron to the dot (at a cost of coulomb repulsion energy $U$) lies significantly above the Fermi energy (see Fig. 3). Because energy is required to add or remove an electron, the transition is unfavourable and the number of electrons confined to the dot is fixed. Thus, the probability of an electron tunnelling through the dot is very low. Once the number of electrons in the QD is fixed, the plunger gate can make additional minute changes to the energy level of the impurity, by applying a non-zero change in potential too small to result in addition or removal of an electron. In addition to controlling the coupling strength, voltage bias, and energy level of the impurity, researchers working on QDs can change the temperature and magnetic field of the system using heaters and magnets.

The devices themselves are made by manipulating a 2D electron gas (2DEG) which exists between semiconducting layers of a heterostructure often GaAs/AlGaAs, grown using molecular beam epitaxy. The 2DEG is confined by placing electrical contacts (gates) on the surface of the heterostructure which can be negatively charged to deplete the 2DEG below it through coulomb repulsion. These gates are written onto the heterostructure using e-beam lithography, with a resolution limit around 10nm, allowing for nanoscale structures to be defined. A quantum dot is typically defined by enclosing a small $\sim 0.5 \mu m^2$ area between several gates, as shown by the grey and gold in Fig. 2. The dot can be completely closed off by increasing the negative potential on the gates surrounding it, or quantum point contacts can be formed by reducing the negative potential until electrons are able to tunnel into or out of the dot at a controllable rate as shown by the gold contacts. A small current can be passed through such a dot by connecting external leads to the 2DEG using ohmic contacts as shown in Fig. 4. In this way, the conductance through the QD can be measured whilst varying numerous parameters. The number of electrons on the dot can be varied by increasing or decreasing the potential on the plunger gate. Varying the plunger gate voltage whilst measuring the conductance through the dot and repeating the test for a range of temperatures around the Kondo temperature ($T_K$) reveals the Kondo effect as shown in Fig. 5. In the case of a QD, the Kondo effect increases the conductance through the dot in contrast to the increase in resistivity it causes in dilute magnetic alloys. This is due to the fact that the QD creates a 1D channel for the electrons to pass through, and increased interaction between the spins of the electrons in the leads and dot creates a Kondo resonance which allows electrons to tunnel through; in a dilute magnetic alloy, on the other hand, the problem

![FIG. 2.](image)

**FIG. 2.** [7] A false coloured SEM image of QD depicted in Fig. 4 showing the 2DEG in the in blue, the gates responsible for establishing the QPCs to the dot in gold, and the gates which only confine the dot in light grey. The plunger gate can be varied separately in order to change the size of the dot (and hence the energy levels of it).

![FIG. 3.](image)

**FIG. 3.** [7] The grey column on the left is the source (sea of electrons), on the right is the drain. (Note that the energy level of the source is slightly higher than that of the drain. This is due to applied potential bias which directs the flow of electrons.) $U$ is the energy required to add an electron to the dot due to coulomb repulsion and $\varepsilon_0$ is the energy of the electron in the highest occupied energy state. The state at $a_1$ won’t spontaneously change given only 1st order quantum effects. However, 2nd order quantum effects provide the possibility of the $\varepsilon_0$ electron tunnelling out of the QD into the Fermi sea ($a_2$) as long as, within the timescale dictated by the uncertainty principle, an electron tunnels from the Fermi sea back into the QD ($a_3$). This is one example out of 8 virtual tunnelling processes which constitute the Kondo effect. [8]. This process can result in a spin flip of the impurity, leading to a singlet state and increases the DOS close to the Fermi level (b).
is 3D in nature, and the increase interaction with the magnetic impurity increase electron scattering, therefore increasing resistance.

There are many ways in which the Kondo effect can be observed in a QD by varying any combination of the variables previously described, or by creating more complex circuits. Here we look at one example of a measurement taken on a simple QD to investigate the Kondo effect. The results are shown in Fig. 5. This measurement is carried out by recording the conductance through the dot (where the conductance is limited by the QPCs to the quantum of conductance, \((2e^2/h = 12906\omega)\) whilst varying the voltage on the plunger gate in the QD. This measurement is then repeated at a range of temperatures. In this case the plunger gate voltage is swept from -500mV (enough to fully deplete the QD) to -350mV (resulting in 6 electrons confined) by varying the energy levels of the dot. When the energy levels are far from the Fermi energy, a dip in conductance is observed in the red trace (800mK). This is explained by the fact that there are fixed numbers of electrons confined at each of these points due to the coulomb blockade effect, making electron tunnelling into and out of the dot very unlikely. When the energy levels of the dot are shifted such that the highest occupied/lowest unoccupied energy level lies very close to the Fermi energy, the number of electrons on the dot can fluctuate between \(N\) and \(N+1\). This dramatically increases the probability of an electron passing through the dot (conductance) when a bias is applied, as seen by the peaks in Fig. 5. As temperature decreases towards zero, one would expect the conductance to drop further in each valley: the energies of the electrons being less thermally broadened increasingly prevents tunnelling into or out of the dot. This is indeed seen in each valley that corresponds to an even number of confined electrons (the lower temperature lines have deeper valleys in Fig. 5). However, for each valley with an odd number of electrons, the Kondo effect actually increases the conductance due to the increased DOS around the Fermi energy caused by Kondo resonance. As temperature decreases, the conductance through the dot increases towards the unitary limit of conductance through the QPCs (illustrated by the T=15mK black line in Fig. 5). By repeating the measurements for a range of temperatures, the logarithmic temperature dependence of the effect explained by Kondo [1] can be seen in the bottom right inset to Fig. 5.

![FIG. 4. [7] A diagram of a QD device where the 2DEG is shown in pale blue, the gold contacts (gates) are shown in grey, and the Ohmic contacts which connect directly to the 2DEG are shown in gold.](image)

![FIG. 5. conductivity through the QD.[10] A graph showing the conductance (G) through a QD against varying gate voltage (Vgl) for a range of temperatures from 15mK (Black) to 800mK (Red). Inset in the bottom right shows the conductance against temperature for a gate Voltage of -418mV where the fitted red line shows the Log(T) dependence of conductance through the QD.](image)

B. Scanning Tunnelling Microscopy

While quantum dots allow for significant control over the parameters of an artificial magnetic impurity, a scanning tunneling microscope (STM) provides a means to directly measure the effects of a real magnetic impurity deposited on a metal substrate. Similar to the quantum dot experiments, the STM is used to observe the Kondo resonance - a many-body state consisting of the magnetic impurity and a cloud of spin-screening conduction electrons [11][12].

STM spectroscopy relies on a measurement of the differential conductance \(dI/dV\), which is proportional to the local density of states (LDOS). To acquire this, the conducting tip of the STM is held at a constant distance from a sample surface. The voltage difference between the tip and sample is then swept and the resulting current fluctuations between the tip and sample are measured. These \(dI/dV\) spectra are taken at various lateral points across the sample, and the Kondo resonance is observed as a dip in the spectrum when the STM tip is positioned near a magnetic impurity. This dip in the \(dI/dV\) curve corresponds to a decrease in the LDOS near the impurity.

Madhavan et al. [11] report observations of the Kondo resonance due to a single magnetic Co atom on the surface of a Au crystal. The measured \(dI/dV\) curve, plotted as a function of bias voltage can be seen in Fig. 6. The different curves show the spectrum at various distances from the Co impurity. Within \(\approx 10\AA\) of the impurity, there is a clear dip in the spectra near \(V = 0\), corresponding to a resonance at \(E_f\) as predicted for the Kondo resonance. This decrease in the DOS is explained in terms of...
FIG. 6. A plot of the $dI/dV$ spectra as a function of bias voltage at various distances from a Co impurity on a Au substrate. The Kondo resonance can be seen centered around 0V (corresponding to $E_f$) within $\approx 10\,\text{Å}$ of the impurity.

The two scattering pathways available to electrons coming from the STM tip. The electron can scatter off the magnetic impurity or the screening electron cloud, and these two possibilities destructively interfere. Thus the formation of the Kondo resonance results in an apparent reduction in the DOS. It is important to note that the temperature of the sample in this experiment was $4\,\text{K}$, well below the Kondo temperature for the system of $T_K = 19\,\text{K}$. At temperatures above $T_K$, this resonance does not appear.

Li et al. present similar findings for a cesium impurity on a silver substrate. As a confirmation of the requirement that the impurity is magnetic, they show that a non-magnetic Ag impurity on the substrate does not produce the same resonance.

Although Kondo’s initial calculations do not account for the formation of the screened impurity, both Madhavan et al. and Li et al. were able to match their data with the Anderson model, an updated calculation that accounts for screening. Fig. 7 shows a plot of the measured resonance around a Ce impurity along with the theoretical prediction.

As an interesting application of the Kondo effect, Manoharan et al. at IBM created an elliptical cavity of atoms with a magnetic impurity at one of the foci. An ellipse is defined such that the sum of the distances to each focus is constant for every point on the ellipse. Thus electrons scattering off of the magnetic impurity at one focus will constructively interfere at the other focus. In Fig. 8 a topographical image of the cavity and impurity can be seen for ellipses of two different eccentricities, along with the measured and calculated $dI/dV$ spectra.

In addition to the Kondo resonance apparent at the location of the impurity, another peak shows up at the second focus of the ellipses.
IV. CONCLUSION

While the single channel Kondo effect has been the subject of many studies, and the phenomenon in the single channel case are now well understood in terms of both theory and experiment, it remains a crucial [critical/useful] gateway to studying many interesting and unexplained phenomena. The strongly correlated electron effects in the single and multi-channel Kondo effects provides accessible testing grounds for both Fermi and non-Fermi liquid theories, and potential insights into other phenomena showing these behaviours, such as heavy Fermion and Cuprate high-$T_c$ superconductivity.

A far cry from the first experiments on the Kondo effect that were carried out on fabricated nanoscale devices[9], experiments today benefit from advances in nanotechnology over the past 20 years and realise more complex systems involving Kondo physics. Siero et al., for example, have recent work on Kondo Lattices[13] which may provide an opening to study quantum criticality a phase transition with the potential to help interpret a wide variety of experiments[14].

Here at UBC, Josh Folk’s group incorporates the Kondo effect into their research by utilising a Kondo device to test a new technique allowing for direct measurement of the entropy of mesoscopic (few electron) systems[15]. This could provide a method to accurately identify and reveal information about Majorana bound states through the two-channel Kondo effect. Such work could have applications as far-reaching as quantum computers, since Majorana fermions (quasi-particles which are their own anti-particles) are contenders for the qubits of stable quantum computing.