Tunneling of Defects in Metals

Norman O. Birge and Brage Golding
Department of Physics and Astronomy
Michigan State University, East Lansing, MI 48824, USA

1. INTRODUCTION

In this review, we discuss the quantum mechanical tunneling of particles of atomic or near-atomic mass in condensed matter systems. Since the probability of tunneling decreases exponentially with mass, this subject might appear to have few experimental consequences. As it turns out, however, the subject has already generated an extensive history, producing significant impact on a surprising number of subfields in physics. Examples of atomic tunneling, brought to focus in the present article, concern the behavior of defects in crystalline and amorphous solids at very low temperatures.

Until recently, virtually all experimental inferences concerning the nature of the defects and their environment were derived from measurements on relatively large, bulk quantities of material. Recent advances have made it possible to detect a few, or even a single, microscopic tunneling defect. In the examples discussed herein, experiments have relied on the sensitivity of the electrical conductivity of small metallic samples to defect motion at temperatures near and below 1 K. These experiments have allowed experimentalists to pick apart the mysteries of the distributions of defects, allowing a detailed examination of the energetics and dynamics of essentially a single atom as it interacts with its environment.

We shall be concerned primarily with a tunneling defect that is coupled to electrons and phonons, both of which may lead to departures from the coherent, periodic oscillations expected from a simple tunneling picture. The study of these couplings expands the utility of single defect studies, focusing attention on many-body interactions. Once a detailed description of the interacting system becomes available, one can then view the tunneling entity as a highly sensitive local probe of its environment.

1.1 Background and scope of article

The dynamical properties of defects have been studied since the beginnings of condensed matter research and the subject has proven to be a rich and useful area of investigation. In a crystal, atoms execute small oscillations about their equilibrium positions as a result of zero-point and thermal vibrational excitations. By adding defects, we insert quasi-static irregularities into the otherwise perfectly ordered (time-averaged) state of the ideal crystal. Defects in crystals encompass the spectrum of atomic perturbations associated with substitutional and interstitial impurities, vacancies, dislocations, grain boundaries, and other more complex spatial arrangements. Glassy or amorphous substances, which do not possess long range atomic order, also possess defects, although they may be less obviously classified.

Defects display dynamics that are influenced by their interactions with their host. Otherwise perfect crystals will always contain a small concentration of vacancies and interstitials with some amount of translational mobility, i.e., they migrate randomly through the host lattice. At another
extreme, defects in a specific substance may become trapped in a local environment, fluctuating in a more restricted space, for example by hopping between two closely separated, nearly equivalent sites. At high temperatures, nearly free diffusion of defects is likely; at low temperatures, the defects become trapped and undergo localized oscillations. For the most part, we shall restrict discussion to defects that undergo fluctuations in a bistable, or double well, potential formed by the surrounding framework of atoms. Thus, we are primarily concerned with local motions rather than transport phenomena such as diffusion. Under many circumstances, the description of an object trapped in a bistable potential can be reduced to a model of a two-state, or two-level, system.

The two-state system coupled to fluctuating vibrational and electronic degrees of freedom is a simple, but non-trivial, model for a surprisingly wide range of physical phenomena [1, 2]. It is the basis for constructing a quantitative picture of the elementary processes of a chemical reaction [3]. It is the first step in understanding mass diffusion in solids, particularly for light interstitial particles such as hydrogen and its isotopes [4]. Very dilute concentrations of two-state systems are known to be responsible for many thermal properties of solids, the most striking example being the low-temperature properties of insulating glasses [5]. There, the Debye theory of the vibrational heat capacity provides an inadequate description of these disordered solids at temperatures below about 10 K, where the defects give rise to a quasi-linear temperature-dependent term in the specific heat and a quadratic temperature dependence to the thermal conductivity.

It is a common misperception that two-state tunneling systems are restricted to highly disordered materials such as glasses. Virtually all crystalline substances, if examined carefully, exhibit low temperature thermal anomalies, indicating the pervasive character of these defects [6]. In metals, fluctuating two-state systems are now believed to underlie the 1/f noise observed in the electrical resistance at low frequencies [7, 8]. Related phenomena, originating at semiconductor-insulator or metal-insulator interfaces, contribute to noise in electronic devices such as field-effect transistors [9].

Since defects exist at relatively low concentrations relative to the number of normal atoms or sites, relatively sensitive spectroscopies are required to differentiate their properties from the background. In addition, because the presence of defects implies disorder, the environment of a defect is unlikely to be unique, i.e., the environment creates an ensemble of states. An understanding of such distributions is necessary to calculate most observables, which represent an average of a physical property over the distribution.

The tunneling of defects in glasses provides a good example of the difficulty in extracting the physics of a tunneling entity from the properties of an ensemble. Tunneling depends sensitively on several physical properties: the mass of the tunneling object, the tunneling potential, the asymmetry of the end state energies, and the coupling to the environment [5]. In the presence of disorder, a statistical model of the distributions of all of these quantities is necessary to extract thermodynamic, transport, or spectral properties. In many instances, the most essential element in a theory may be the assumption about the distribution of tunneling parameters. Of course, a well-designed experiment or theory may require few assumptions about the character of the distributions.

The recent ability to examine individual defects has altered this picture. Most, if not all, of the properties of a tunneling system (TS) can be extracted from observations of the dynamics of a single tunneling object. Furthermore, at least in principle, the distribution functions can be constructed from a sufficiently large set of single defect measurements, although in practice this
may be a daunting task. In the near future, there may be additional spectroscopic techniques brought to bear on defect dynamics, which parallel recent advances in the optical spectroscopy of single chromophores [10, 11].

The structure of this review is the following. First, we shall provide a theoretical framework for understanding the behavior of a massive tunneling object coupled to its environment of conduction electrons. Since there are a number of excellent recent reviews of theory [2, 12, 13], we shall only provide enough detail to clarify the description of experiments that follow. The main part of the review will discuss experiments that have allowed the study of the dynamics of single, or small numbers, of defects in great detail.

2. THEORY

In this section, we highlight the major theoretical results, with emphasis on the physical assumptions and approximations made in obtaining them.

2.1 Isolated tunneling systems

Consider first the case of an isolated particle in a double-well potential \( U(r) \), as shown in Fig. 1. One can think of the particle either as a single atom, in the case of hydrogen tunneling in Nb, or it could be a group of atoms moving cooperatively, as in the case of a TS in a highly-disordered solid. In the latter case the horizontal axis of Fig. 1 is a collective coordinate running through two nearby local minima in the configuration space. In general the potential is not symmetric; we denote by \( \varepsilon \) the energy difference between the lowest vibrational states in each well. The vibrational frequencies in the two wells are in general unequal; for simplicity we take them to be the same and call them \( \omega \). The average barrier height is \( V \), the distance between minima is \( \delta R \), and the particle mass is \( M \). The Hamiltonian for the particle is:

\[
H^0 = \frac{p^2}{2M} + U(R).
\]  

(1)

If \( \hbar \omega \ll V \), it is convenient to take as basis states the states of the particle in the left and right wells, \( \phi_{1,n}(R) \) and \( \phi_{2,n}(R) \), where the subscripts 1 and 2 refer to the left and right well, respectively, and \( n \) is the index for the vibrational level in the well. At temperatures such that \( k_B T \ll \hbar \omega \), only the lowest vibrational state is significantly populated in each well. In that case, the system described in (1) can be truncated to a two-state system with Hamiltonian given by

\[
H^0_{\text{trunc}} = \frac{1}{2} \varepsilon \sigma_z - \frac{1}{2} \hbar \Delta \sigma_x.
\]  

(2)

where \( \sigma_z \) and \( \sigma_x \) are Pauli spin matrices, and \( \Delta \) is the tunneling matrix element due to overlap of the wavefunctions in the left and right wells (\( \hbar \Delta = \langle \phi_1 | H^0 | \phi_2 \rangle \propto \langle \phi_1 | \phi_2 \rangle \)). For a one-dimensional potential, \( \Delta \) is calculated in the WKB approximation with the result
$$ \Delta \approx \omega \exp \left( -\frac{\sqrt{2mV}}{\hbar} \delta R \right) \approx \omega \exp \left( -\frac{B V}{\hbar \omega} \right). $$

where B is a constant that depends on the shape of the potential [14]. From (3) we see that the condition $\hbar \omega \ll V$ implies that $\Delta \ll \omega$. Truncation to a two-state system requires $\varepsilon \ll \hbar \omega$ in addition to $k_B T \ll \hbar \omega$. Note, however, that there are no conditions on the relative magnitudes of $k_B T$, $\varepsilon$, and $\hbar \Delta$.

Solution of the Hamiltonian (2) is straightforward. The energy eigenstates are linear superpositions of the left and right well states, with eigenvalues

$$ E = \pm \frac{1}{2} \sqrt{\varepsilon^2 + \hbar^2 \Delta^2}. $$

In the special case where $\varepsilon=0$, the energy eigenstates are just the symmetric and antisymmetric linear combinations of the two well states. As $\varepsilon$ increases, the two energy eigenstates become more localized in the left and right well, respectively.

The tunneling model of glasses is based on the above picture, along with physically plausible assumptions about the distributions of the parameters $\varepsilon$ and $\Delta$ in highly disordered solids [15, 16].

To calculate the effect of the TS on thermal conductivity or acoustic absorption and dispersion,
one must also take into account interactions between the TS and the phonons. The phonons perturb the TS in two ways: they either induce resonant transitions between the ground and excited states, or they cause a slow variation in $\varepsilon$ which leads to relaxational absorption and dispersion of sound. The predictions of the tunneling model have been extraordinarily successful in explaining the thermal, acoustic, and dielectric properties of insulating glasses at temperatures below about 1 K [5].

The underlying assumption in the tunneling model is that interactions with the bath do not change the qualitative picture of isolated TS. That assumption breaks down in metals, however, with the addition of the electronic degrees of freedom to the heat bath. The first signs of this breakdown appeared in measurements of sound propagation in metallic glasses [17]. Some of the experimental results could be explained by a very fast TS relaxation rate, due to interaction with the conduction electrons. Other aspects, however, were quantitatively inconsistent with predictions of the tunneling model [18].

### 2.2 Tunneling systems in metals

The presence of the conduction electron bath changes qualitatively the tunneling dynamics of the defect [2, 12, 13]. One can view the bath as influencing the tunneling in two rather distinct ways, namely dephasing (or decoherence) and screening. A sufficiently high dephasing rate of the tunneling particle's wavefunction leads to overdamping of the tunneling process, resulting in a semi-classical picture where the defect can be viewed as being in a particular well at any given time. More interestingly, dephasing by bath fluctuations also causes the tunneling rate to decrease with increasing temperature under certain conditions. Screening of the tunneling particle by the conduction electrons leads to a reduction of the tunneling rate relative to its bare value in the absence of screening. The effect of screening grows at low temperature, and gives rise to an anomalous power-law temperature dependence of the tunneling rates.

The interaction of the tunneling defect with the conduction electron bath has been studied by many workers. There are two main approaches to the problem: direct determination of the overlap integral between the initial and final many-body wavefunctions [1, 19], and functional integral techniques using instantons to describe the tunneling process [2]. In the former approach, one treats the conduction electron bath as a bath of true fermions. In the latter approach, one has the option of replacing the conduction electron system with a bath of harmonic oscillators, i.e. bosons, with an appropriate spectrum of low-energy excitations. Caldeira and Leggett have shown that such a replacement is sufficient to describe the effect of the dissipative bath on the dynamics of the tunneling particle [20]. In this approach, the original problem is mapped onto the spin-boson Hamiltonian, whose solution is now known over a very wide portion of its parameter space [2]. These two approaches yield results that are completely consistent with one another [21-23].

Even though atomic defects in a metal interact both with conduction electrons and with phonons, most of the phenomena described in this chapter can be understood by considering only the former. This is because the density of states of low-energy excitations is much larger for electron-hole pairs than for phonons. (We will return to the issue of phonon-TS coupling at the end of this section.) The physical picture of the problem is as follows. Imagine the tunneling particle in one well surrounded by its screening cloud of electrons. The electron density surrounding the particle exhibits Friedel oscillations, which decay slowly with increasing distance...
from the particle. When the tunneling particle moves to the other well, the screening cloud tries to follow it. The spatial mismatch of the Friedel oscillations between these two electronic states extends to large distances from the tunneling particle, hence it involves electron-hole excitations of low energies. The inability of the conduction electron system to screen the charged particle instantaneously gives rise to most of the phenomena to be discussed in this chapter.

The Hamiltonian for the TS-electron system is:

$$H = H^0 + H^{el} + H^{\text{int}}.$$  \hspace{1cm} (5)

$H^0$ is given in (1), and $H^{el}$ is the Hamiltonian of the free-electron system given by

$$H^{el} = \sum_{k} \frac{p_{ik}^2}{2m} = \sum_{k\eta} \xi_{k\eta} c_{k\eta}^+ c_{k\eta},$$ \hspace{1cm} (6)

where $c_{k\eta}$ is the annihilation operator for an electron of wavevector $k$ and spin $\eta$, and $\xi$ is the electron energy measured from the Fermi surface. The interaction term is given by

$$H^{\text{int}} = \sum_{i} V(r_i - R) = \sum_{kk'\eta} V_{kk'} e^{i(k-k')R} c_{k\eta}^+ c_{k'\eta}. \hspace{1cm} (7)$$

There are two approximations that are often made immediately to the Hamiltonian given by Eqs. (1), (5)-(7). The first is the truncation of the original double-well potential of the tunneling particle to a two-state system, i.e. the replacement of (1) with (2). The danger of making such a truncation is that one loses information about the structure of the wells, and hence the vibrational frequency of the tunneling particle. Kagan and Prokof'ev have argued that this loss of information leads to an incorrect assignment of the high-frequency cut-off of the tunneling matrix element renormalization to be discussed below [19]. In the two-state system basis, the interaction term (7) is expressed in terms of the Pauli matrices of the two-state system, rather than the particle position $R$. The truncated Hamiltonian is:

$$H_{\text{trunc}} = H^0_{\text{trunc}} + H^{el} + \sigma \sum_{kk'\eta} V_{kk'} c_{k\eta}^+ c_{k'\eta} + \sigma \sum_{kk'\eta} V_{kk'} c_{k\eta}^+ c_{k'\eta}, \hspace{1cm} (8)$$

where $H^0_{\text{trunc}}$ is given by (2) and $H^{el}$ is given by (6). A second approximation often made is to ignore the last term in Eq. (8). In general, one expects the matrix elements that are diagonal in the left and right well states, $V^z$, to be much larger than the off-diagonal elements, $V^x$, since the latter are proportional to the wavefunction overlap, $\Delta$ [24]. The phenomena involving explicit detection of slowly tunneling centers to be discussed in this chapter can be explained with only the diagonal term in Eq. (8). The off-diagonal term is responsible for electron-assisted tunneling phenomena implicated in non-linear conductance characteristics of ballistic point-contacts. Both classes of experiments will be discussed in Section 3.
2.3 Adiabatic and nonadiabatic electron-hole excitations

Theoretical treatment of the Hamiltonian given by (8) is difficult even in the absence of the last term. At first glance, one might guess that because the electrons are so much lighter than the tunneling particle, one could treat them in the adiabatic approximation. That is, one writes the total wavefunction of the system as a product:

$$\Psi_{\text{adiabatic}}(R, \vec{r}) = \phi(R)\psi(r, R)$$  \hspace{1cm} (9)

where $\psi(r, R)$ is the solution of the Schrodinger equation for the electronic Hamiltonian in Eqs. (6) and (7) with the position of the tunneling particle fixed at position $R$. $\phi(R)$ is then obtained by solving the Schrodinger equation for the tunneling particle with the Hamiltonian given by (1) or (2), with the addition of an extra term due to the electronic energy. For a homogeneous electron gas, the electron energy does not depend on the position of the tunneling particle, hence the extra term is an additive constant. The net result is that the electrons have little effect on the dynamics of the tunneling particle; they cause only a slight enhancement of the particle mass in Eq. (3).

Kondo was the first to show that the adiabatic approximation breaks down in metals even though the electron mass is much smaller than the mass of the tunneling particle [1, 21, 25]. The reason is that there is a high density of states of low-energy electron-hole excitations with energies smaller than the characteristic energy scale of the tunneling process. Because of this fact, Kondo considered the opposite limit to the adiabatic approximation, i.e. the limit where the particle tunnels so quickly that the electrons do not have time to adjust their wave functions. In that limit, one can approximate the total system wavefunction as [12]

$$\Psi_{\text{nonadiabatic}}(R, \vec{r}) = \frac{1}{\sqrt{2}} \left( \phi_1(R)\psi_1(\vec{r}, R_1) + \phi_2(R)\psi_2(\vec{r}, R_2) \right)$$  \hspace{1cm} (10)

where $R_1$ and $R_2$ refer to the stable positions of the left and right wells, and we have assumed zero energy asymmetry between the wells, $\varepsilon=0$, for simplicity. In this nonadiabatic approximation, the tunneling matrix element between the initial and final states contains not only the overlap between the initial and final tunneling particle wavefunctions, but also the overlap between the initial and final electronic states. The latter term vanishes in the limit of large system size at zero temperature, a result known as the "Anderson orthogonality catastrophe" [26]. The expression for the effective tunneling matrix element in a finite system is:

$$\Delta_{\text{eff}} \approx \Delta \cdot \langle \psi(\vec{r}, R_1) | \psi(\vec{r}, R_2) \rangle \approx \Delta \cdot \left( \frac{\delta E}{E_F} \right)^K$$  \hspace{1cm} (11)

where $\delta E$ is the electronic energy level spacing and $E_F$ is the Fermi energy. The coupling parameter $K$ depends on the potential $V(r)$ and the distance between the two wells. $K$ is related to the scattering phase shifts of the electronic wavefunctions due to the presence of the charged defect.

Yamada and coworkers have calculated $K$ exactly for the case where $V(r)$ is spherically symmetric, so that only the s-wave phase shift is nonzero. In that case, the result is [27]:

...
where $\delta$ is the phase shift, and

$$x = \frac{\sin^2 k_x \delta R}{(k_F \delta R)^2}, \quad \delta R = R_1 - R_2.$$

For the case of weak coupling, i.e. $N(E_F)V^z << 1$, with $N(E_F)$ the electronic density of the states near the Fermi level, $K$ is approximated by:

$$K \approx 2[N(E_F)V^z]^2(1 - x).$$

The fact that $K \leq \frac{1}{2}$ in this model is significant, as it will lead to a tunneling rate that decreases with increasing temperature in the incoherent tunneling regime.

It appears that the extreme nonadiabatic approximation of Eq. (10) leads to a vanishingly small tunneling matrix element given by Eq. (11). As it turns out, both the numerator and denominator of the tunneling matrix element renormalization term in Eq. (11) are incorrect for this problem. The energy range of electron-hole excitations that contribute to the reduction of the overlap integral is much smaller than indicated in Eq. (11). Consider first the upper limit, i.e. the denominator $E_F$. The intrawell motion of the tunneling particle occurs at frequency $\omega$. The tunneling event between the wells also has its own characteristic frequency scale, given by the small-oscillation frequency of the particle in the inverted potential [33]. These two frequencies are generally of the same order of magnitude; for simplicity we will consider them to be the same $\omega$. One can divide the electron-hole excitations into two classes: 1) those with energy greater than $\hbar \omega$ adjust adiabatically to the motion of the tunneling particle, hence they have little effect on the tunneling dynamics [19]; 2) those with energies smaller than $\hbar \omega$ do not follow the particle's motion instantaneously. Only the latter contribute to the overlap integral, and the high-energy cut-off in Eq. (11) should be $\hbar \omega$ rather than $E_F$.

That still leaves the problem of the infra-red divergence in Eq. (11) associated with the electronic level spacing, $\delta E$. Let us consider first the limit of weak coupling between the tunneling particle and the conduction electrons. This limit is achieved when $K$ is very small or the temperature is very low -- the exact criterion will be spelled out later. There we expect the solution of the problem to be similar to that of the isolated TS, i.e. the particle tunnels coherently back and forth between the two wells. We can use Eq. (11) to estimate a renormalized tunneling frequency, $\Delta_r$, but we should not include the effect of electron-hole excitations with energies lower than $\Delta_r$ itself. Such low-energy excitations are so slow that they do not follow the motion of

* Yamada's proof that $K \leq \frac{1}{2}$ assumes that $\delta \leq \pi/2$, a result obtained directly from the Friedel sum rule if the tunneling particle has charge $e$ [28]. Until recently, it was thought that values of $K>1$ would lead to localization of the tunneling particle in a single well [29-31]. Such values seem possible for a defect of charge $2e$ or more, or even for a singly-charged defect with a non-spherical potential. Moustakas and Fisher have shown recently, however, that the particle can tunnel even if its charge is $2e$ or more [32].
the tunneling particle at all. We therefore replace the low-energy cut-off, $\delta E$, in Eq. (11) by the renormalized tunneling frequency, $\Delta_r$, and solve for the latter self-consistently:

$$\Delta_r = \Delta \cdot \left( \frac{\Delta_r}{\omega} \right)^{\kappa} = \Delta \cdot \left( \frac{\Delta}{\omega} \right)^{\frac{\kappa}{1-\kappa}}$$

(15)

Rigorous calculations confirm that at sufficiently low temperature the particle tunnels coherently with frequency $\Delta_r$ given by Eq. (15) [34]. Scattering of conduction electrons from the tunneling particle excites transitions between the two tunneling eigenstates, and leads to dephasing or damping of the coherent oscillations. We can estimate the coherence or damping time, $\tau_c$, by treating the $V^2$ term in Eq. (8) in perturbation theory. Such a calculation was first performed in 1978 to explain the fast relaxation time of TS in metallic glasses [17, 35]. The result is analogous to the Korringa relaxation of nuclear spins in NMR, and is given by:

$$\hbar \tau_c^{-1} = \frac{\pi}{2} [N(E_F)V^2]^2 E \coth \left( \frac{E}{2k_BT} \right)$$

(16)

where $E$ is the energy splitting of the TS given by Eq. (4). The term in brackets is essentially the weak-coupling limit of the constant $K$, given by Eq. (14). In the case of a symmetric ($\varepsilon=0$) double-well potential, it turns out that $E$ should be replaced by $\Delta_r$, implying that $\tau_c^{-1} \propto \Delta_r$ at low temperature [2]. A convenient way to conceptualize the motion of the particle is through the quantity $P_{12}(t)$, defined as the probability to find the particle in the right well at time $t$ if the system was prepared initially with the particle in the left well. Fig. 2a shows $P_{12}(t)$ for the case of weakly-damped coherent oscillations, as described above. An analytical form for $P_{12}(t)$ has been derived only at zero temperature [2, 12].

### 2.4 Incoherent tunneling

If the damping becomes large enough, then the tunneling particle no longer oscillates between the two wells. The crossover to overdamped behavior occurs when either $\tau_c^{-1} > \Delta_r$ or $\varepsilon > \hbar \Delta_r$. The high-temperature limit of the damping rate given by Eq. (16) is:

$$\hbar \tau_c^{-1} = \pi [N(E_F)V^2]^2 k_B T \approx \pi K k_B T .$$

(17)
Figure 2: Probability $P_{12}(t)$ to find the tunneling particle in well 2 at time $t$, given that it started in well 1 at time 0, for: (a) coherent, or underdamped motion; (c) incoherent, or overdamped motion. Panel (b) presents a conceptual picture of the transition from coherent to incoherent motion when the damping or interruption rate $\tau_c^{-1}$ is greater than the tunneling frequency $\Delta$. 
For the purpose of this discussion, we restrict ourselves to the case \( \varepsilon = 0 \), so the criterion for overdamped tunneling is \( \pi K k_B T > h \Delta \). (The full result for arbitrary \( \varepsilon \) will be given below.) In this limit, we must modify the argument leading to Eq. (15) for the renormalized tunneling matrix element. The low-energy cut-off of the electron-hole excitations leading to reduction of the tunneling matrix element is now determined by thermal smearing of the electronic wavefunctions \[ \] [1]. The result is:

\[
\Delta_{\text{eff}} \approx \Delta \left( \frac{k_B T}{h \omega} \right)^K.
\]  

A helpful way to visualize the tunneling behavior in the overdamped regime is to consider again the quantity \( P_{12}(t) \) \[12\]. Imagine that the particle starts in the left well at time \( t=0 \). It starts to tunnel coherently to the right well with tunneling matrix element \( \Delta_{\text{eff}} \), so \( P_{12}(t) = \sin^2(\Delta_{\text{eff}} t/2) \). But the buildup of probability amplitude in the right well is interrupted after a time \( \tau_c < \Delta_{\text{eff}}^{-1} \) by scattering from conduction electrons. Every time an interruption occurs, the wavefunction of the tunneling particle loses its phase coherence, so the tunneling process must start again with zero slope. This state of affairs is depicted in Fig. 2b, following Kondo \[12\]. The net effect of a large number of such interruptions is that the probability to tunnel to the right well grows linearly with time, rather than quadratically, and with a slope \( \gamma = \Delta_{\text{eff}}^2 \tau_c \). This linear time dependence is simply the initial rise of the function \( P_{12}(t)=[1-\exp(-\gamma t)]/2 \), shown in Fig. 2c, which characterizes the incoherent tunneling process. Using the expressions for \( \Delta_{\text{eff}} \) and \( \tau_c \) in Eqs. (17) and (18) we find \[25, 34\]:

\[
\gamma (\varepsilon = 0) \approx \Delta_{\text{eff}}^2 \tau_c = \frac{h \Delta_{\text{eff}}^2}{\pi K k_B T} \propto T^{2K-1}.
\]  

Since \( K < 1/2 \) always, the preceding argument leads to the surprising conclusion that the incoherent tunneling rate decreases with increasing temperature! The origin of this effect is clear from the above discussion: as the interruption rate increases, the ability of the system to build up probability amplitude in the opposite well decreases. This effect appears in other aspects of physics, and is referred to as "Zeno's Paradox", "Turing's Paradox", or others \[36\].

Returning to the general case of an asymmetric double well, the total transition rate \( \gamma \) is the sum of the individual rates from one well to the other: \( \gamma = \gamma_{12} + \gamma_{21} \). Since the double-well system is in equilibrium with a heat bath, the two individual rates must obey the detailed-balance relation:

\[
\frac{\gamma_{12}}{\gamma_{21}} = e^{\varepsilon / k_B T}.
\]  

A full calculation of \( \gamma \) for an asymmetric double-well potential was first performed by Fisher and Dorsey \[37\] and Grabert and Weiss \[38\]. The result is:
\[
\gamma = \frac{\Delta^2}{2\omega} \left( \frac{2\pi k_B T}{\hbar} \right)^{2K-1} \frac{\cosh(\varepsilon / 2k_B T)}{\Gamma(2K)} \cdot \left( \frac{\Gamma + i\varepsilon / 2\pi k_B T}{\Gamma(2K + i\varepsilon / 2\pi k_B T)} \right)^2.
\]

(21)

where is \( \Gamma \) is the gamma function. Eq. (21) has two limits. In the limit \( \varepsilon \ll k_B T \), Eq. (21) reduces to Eq. (19) with \( \Delta_{\text{eff}} \) defined in (18), to within a factor of order unity. In the opposite limit, \( \varepsilon \gg k_B T \), the faster of the two individual rates is nearly temperature independent, while the slower rate decreases rapidly with decreasing temperature due to the Boltzmann factor in Eq. (20). In this limit we can think of the TS undergoing spontaneous emission from the higher-energy well to the lower-energy well, and stimulated absorption from the lower well to the higher.

Interestingly, one can reproduce Eq. (21) by treating the tunneling term in the Hamiltonian (the \( \sigma_z \) term in Eq. (2)) in second-order perturbation theory using the "Golden Rule". Such a calculation has been carried out treating the electrons as Fermions [12] or as Bosons [2], with the same result.

From the point of view of fitting experimental data, Eq. (21) appears to have four free parameters, namely \( K, \varepsilon, \Delta, \) and \( \omega \). In fact, the last of these can be eliminated by rewriting Eq. (21) in terms of the renormalized tunneling matrix element given by Eq. (15). Doing so gives the result:

\[
\gamma = \frac{\Delta^2}{2} \left( \frac{2\pi k_B T}{\hbar\Delta_r} \right)^{2K-1} \frac{\cosh(\varepsilon / 2k_B T)}{\Gamma(2K)} \cdot \left( \frac{\Gamma + i\varepsilon / 2\pi k_B T}{\Gamma(2K + i\varepsilon / 2\pi k_B T)} \right)^2.
\]

(22)

2.5 Phonons

At sufficiently low temperature, the presence of phonons in a metal does not change the picture outlined above, due to the small density of states of phonons at low energy. In fact, the damping rate due to phonons must be small at low temperature, given the observation of coherent phenomena in insulating glasses [39, 40]. The phonons do, however, contribute to the overlap of the wavefunctions in the two wells, hence they renormalize the tunneling matrix element. This is the "polaron" problem that has been studied for many years [41].

At temperatures well below the Debye temperature \( \Theta_D \), the influence of the phonons and the electrons on the tunneling rate can be treated separately [12, 42]. The phonons cause the incoherent tunneling rate to increase dramatically above a characteristic temperature \( T_{ph} \) that varies inversely with the coupling strength between the tunneling defect and the long-wavelength acoustic phonons. For \( T_{ph} \ll T \ll \Theta_D \), the tunneling rate increases exponentially with temperature: \( \gamma \propto \exp \left[ 2T^2 / 3T_{ph}^2 \right] \). At higher temperatures, the electrons and phonons must be treated together, and the tunneling rate must be calculated numerically [43].
3. EXPERIMENTS

3.1 Macroscopic phenomena: multiple defects

3.1.1 Muon Diffusion

A manifestation of quantum tunneling arises in the transport properties of the positive muon, $\mu^+$, implanted into crystals by high-energy particle accelerators. Following the initial work of Gurevich et al. [44] demonstrating that measurement of muon diffusion rates in metals was possible with muon spin rotation (µSR), numerous studies of muon diffusion in different metallic hosts have been performed. A diffusion rate that increased as the temperature was decreased was reported for muons in copper, iron and aluminum [45-47]. These measurements did not derive the diffusion coefficient directly from the motional narrowing but rather from the rate of trapping of the muons at impurities or vacancies. In later measurements, it was found that at low temperatures (below $\sim 50$K in copper and $\sim 10$K in aluminum) the diffusion rate $D$ in metals depended on temperature as $D \sim T^{-\delta}$, with $\delta$ in the range 0.6-0.7. An early suggestion that quantum diffusion in metals is fundamentally different than in insulators was made by Andreev and Lifshitz [48] in discussing the motion of hypothetical zero-point vacancies in quantum crystals. They predicted $D \sim T^{-1}$ in metals as opposed to $D \sim T^{-9}$ for insulators.

Agreement between theory and experiment was unsatisfactory until Kondo [21] showed that at low temperatures the diffusion constant $D$ varies as $T^{\delta}$ with $0 < \delta < 1$. This theory applies when the temperature is much greater than the energy difference between the initial and final states of the diffusing particle. At the lowest temperatures impurity effects trap the muon and cause $D$ to saturate. At higher temperatures a crossover to $D \sim T^{\gamma}$ behavior is observed, consistent with a phonon-assisted tunneling process (hopping) [41]. The ability of Kondo's ideas to explain the unusual power-law temperature dependences of light particle diffusion in metals at low temperatures must be viewed as a remarkable achievement. A review of muon diffusion experiment and theory has recently appeared [49].

3.1.2 Metallic glasses: acoustics

Power law temperature dependences of the heat capacity and thermal conductivity of insulating glasses were observed in the early 1970's [50]. Although somewhat dependent on impurity content, the features were recognized as an intrinsic feature of the glassy state. The observations were explained by the tunneling of some of the glasses' atoms or molecules in a double well potential [15, 16]. The power laws, observed over several decades of temperature below 1 K, result from a broad, nearly constant, density of two-level TS. The low thermal conductivity of glasses arises from the resonant scattering of thermal phonons by the states. Experiments with monochromatic phonons confirmed the two-level character of the tunneling states by observation of an intensity-dependent absorption, attributed to population equalization, or saturation, of the systems [51, 52]. The picture was consistent with coherent, or nearly coherent, tunneling between the lowest eigenstates of a small number of localized two-level entities, with relatively weak interactions with other TS and phonons.

In view of the substantial differences in structure between insulating and metallic glasses it is at first glance surprising that glassy metals contain about the same densities of two level systems
as their insulating counterparts [53]. Instead of a relatively open, predominantly covalent bonded random network like SiO₂ and chalcogenide glasses, metallic glasses can be pictured as dense random packings of spherical atoms. The lack of open structures, the absence of bridging chalcogens, and the lack of highly directional bonding means that the TS derive from more fundamental considerations. The existence of conduction electrons in a metallic glass makes the observation of TS via thermodynamic and transport properties difficult. For example, the linear specific heat due to conduction electrons overshadows the tunneling contribution by about an order of magnitude. However, the discovery of superconducting metallic glasses [54, 55] allows one to isolate tunneling properties from the electronic terms. At temperatures well below the superconducting transition, quasiparticle contributions to specific heat and heat transport are suppressed exponentially by Δ/k_B T, where Δ is the superconducting energy gap. In this range, a linear tunneling term in the specific heat was observed, with a magnitude close to insulators [56].

The tunneling of glassy defects in the presence of conduction electrons was discovered to be significantly different from tunneling in insulators. Ultrasonic saturation studies of metallic glasses at low temperatures revealed a strong coupling between TS and conduction electrons. A high ultrasonic saturation threshold implied that relaxation times of tunneling states in metals were some 4 to 5 orders of magnitude faster than in insulators [17]. A weak coupling, Korringa-like theory was developed, as noted in the previous section, and comparison with experiment allowed the determination of the TS-conduction electron coupling coefficient K [35, 57]. The theory views the electrons as scattering inelastically from the TS, thereby enhancing the transition rates between the two TS energy levels. A value of K = 0.2 was obtained in the metallic glass PdSiCu.

3.1.3 Hydrogen tunneling in transition metals

Hydrogen incorporated into transition metals at low concentrations exhibits many of the features associated with tunneling in glasses. When octahedral interstitial impurities such as nitrogen or oxygen are introduced into niobium, they locally expand the lattice, creating traps for hydrogen. The tetrahedral trap sites occur in nearly degenerate pairs, forming double-well potentials for hydrogen tunneling.

The study of hydrogen in crystalline transition metals is attractive since the tunneling species is known (isotope effects are observable) and tunneling distances are reasonable [58]. Evidence for low temperature hydrogen tunneling in Nb:O was obtained from specific heat measurements by Sellers et al. [59]. The tunneling contribution appears as a broadened Schottky anomaly. This contribution depends upon the concentration of hydrogen, or deuterium, in the Nb. The presence of a large isotope effect in deuterated Nb lent strong support to a tunneling model.

Subsequent ultrasonic measurements led to a model [60] in which hydrogen tunneling occurs between two sites separated by a uniform barrier and distance. Disorder is introduced by a distribution in the ground state energy in each well introduced by random strains. Experiments by Magerl et al. [61] were used to extract the electron-TS coupling parameter K which had a value similar to that found in metallic glasses, typically 0.1.

Inelastic neutron scattering experiments from TS in Nb:H by Wipf et al. [62] and by Steinbinder et al. [63] have provided a comprehensive insight into the dynamics. For example, by examining the inelastic peak position, it was observed that the average tunneling energy differs between the normal and superconducting states at 0.2K, confirming the renormalization of the
tunneling energy by conduction electrons. Additional measurements have obtained the
temperature dependent energy and damping rate. The damping rate provides a particularly useful
view of the transition between coherent and incoherent tunneling. In the superconducting state,
the tunneling is coherent at temperatures as high as 4.2K. In the normal state, however, the
damping is large, leading to a quasielastic peak in the structure factor indicative of incoherent
tunneling [64]. The measurements also determine the average transition rate as a function of
temperature. The power-law observed below 30 K is consistent with a rate varying as $T^{-0.9}$ leading
to a Kondo exponent $K = 0.05$.

3.2 Mesoscopic phenomena: single defects

The thermodynamic, transport, acoustic, and other measurements described above reflect an
average over a large ensemble of TS. With reasonable assumptions about the character of the excitation contributing to the specific heat of glasses, as an example, it can be estimated that
about $10^{17}$ states per cm$^3$ exist with energy less than 1 K. Much fewer states can be detected in
resonant experiments; for example, in acoustic echo experiments signals from $10^{12}$ states can be
detected. Nonetheless, these are still large numbers and any sensible interpretation requires that the
distribution of TS parameters be considered as part of theoretical models. In addition to physical
parameters, one now is obliged to justify assumptions about the tunneling distribution functions.
Since there is no a priori knowledge about the distributions, it necessarily weakens many
conclusions about the low temperature properties of disordered materials.

As we shall now discuss, recent advances, both technical and conceptual, have revolutionized
the study of low temperature defect tunneling phenomena. With the advent of mesoscopic
physics, it is possible to devise experiments with sufficient sensitivity, and immunity from noise,
so that the behavior of a single microscopic TS can be examined in detail. Historically, the study of single defects began in junction devices with investigations of their role in carrier trapping. We
describe that work, as well as more recent advances, as background for what follows.

3.2.1 Electron tunneling in semiconductor structures

A substantial amount of research has dealt with noise and discrete switching events in
microfabricated MOS (metal-oxide-semiconductor) or MIM (metal-insulator-metal) structures.
The conductance of a MIM device is modulated by the capture and emission of charges at a
defect residing in the insulator. In a MOS structure, the conductance-modulating defect must be
near the conducting channel, either in the oxide or the semiconductor. When the devices are large,
charge trapping events give rise to 1/f noise, but when device dimensions approach μm scales, discrete conductance fluctuators are readily observed. The subject has been reviewed by Kirton
and Uren [9].

In a silicon MOSFET, the Si energy bands near the Fermi level are bent at the oxide interface.
An inversion layer is created at the interface under appropriate biasing conditions. If a localized
defect state whose energy $E_d$ is within a few $k_BT$ of $E_F$ is located nearby in the oxide, tunneling of
carriers between the 2-D electron gas and the defect can occur. In practice, a bias is applied by a
gate to tune the difference $E_d-E_F$ into the appropriate region. The change in the charge state of the
defect modifies the potential seen by the transport channel and can lead to relatively large changes, \( \sim 10^{-3} \), in the conductance of the device.

Ralls et al. [65] demonstrated discrete trapping events in a Si MOSFET and showed that the traps must lie extremely close to the conducting channel. Rogers and Buhrman [66, 67] used a small-area Nb-Nb\(_2\)O\(_5\)-PbBi tunnel junction to show that traps located in the oxide were able to capture and release electrons tunneling through the oxide barrier. They showed that 1/f noise arose from the superposition of independent fluctuators with Lorentzian spectra. At higher temperatures, the discrete spectrum transformed smoothly over to a continuous 1/f spectrum. At temperatures below 20 K or so, the total trap and release rate became nearly independent, which was interpreted as evidence that a tunneling process was controlling the rates.

Evidence for dissipative tunneling of electrons coupled to defect motion has been observed in MOS devices. In n-channel MOSFETs with 1 \( \mu \)m\(^2\) active areas, discrete conductance fluctuations attributed to single defects were observed subsequent to various electrical stressing treatments [68]. In stressed devices, a single fluctuator with a gate-voltage dependent transition rate was observed. The transition rate exhibited a maximum at a bias consistent with alignment of defect energy with the Fermi level. The width of the resonance was strongly temperature dependent, narrowing as the temperature was decreased from 4.2 to 1.2 K.

The results were interpreted as a two-state system interacting with an electron bath under conditions of extremely weak coupling, i.e., \( K \ll 1 \). Under this assumption, \( \Delta r \approx \Delta \), \( \Gamma(K) \approx 1/K \), and provided \( \varepsilon \leq k_B T \), Eq. 22 reduces to

\[
\gamma(\varepsilon) \approx \frac{(h\Delta_r^2 / 2\pi K k_B T)(1 + (\varepsilon / 2\pi K k_B T)^2)^{-1}}{
\pi \varepsilon \sqrt{2}\Delta_r \langle \varepsilon \rangle^{3/2}}
\]

This result for the transition rate takes the form of a Lorentzian, with a level width \( 2\pi K k_B T \) determined by coupling to an electron reservoir at temperature \( T \). For a particular defect, it was determined that \( K = 0.007 \) and \( h\Delta_r = 3.7 \times 10^{-5} \) K [68]. It is noteworthy that for this particular system, the coupling is extremely weak relative to defects studied in disordered metals, although the renormalized tunneling matrix element is large. The weakness of the coupling may derive from a defect that resides in the gate oxide, with small overlap with the 2-D electron gas wave function.

Evidence for MOS defects that couple strongly to the electron gas has been uncovered more recently [69]. In some samples that were electrically unstressed, substantial deviations of the fluctuation rates from a weakly interacting picture were observed. A theory, developed by Matveev and Larkin [70], specifically for electrons tunneling out of a 2D electron gas, yields a singularity in the tunneling rate when the defect energy and Fermi energy coincide. The origin is similar to the picture developed for defects tunneling in metals: a tunneling electron loses energy to low energy electron-hole pairs created by the change in defect potential.

**3.2.2 Thermally activated defects in metals**

A significant advance in the study of individual defects in metals was made by the introduction of lithographic point contact spectroscopy by Ralls and Buhrman [71]. By etching holes of 40-nm diameter in a thin Si\(_3\)N\(_4\) membrane and evaporating metals on both sides, a metallic point contact nanoconstriction is formed. This is a two-terminal device that is extremely
sensitive to scattering events in the small volume of the constriction where the current density is greatest. At temperatures near 20 K individual fluctuators were observed in Cu, Al, and Pd junctions, all of which showed thermally activated behavior [72]. As the temperature was raised above 150 K, increasing numbers of discrete events appeared, suggesting that the junction contained many more active defects. It was proposed that here, and up to room temperature, the defects interacted strongly with one another, leading to a dynamic liquid-like state.

3.2.3 Quantum transport and universal conductance fluctuations

The point contact spectroscopy discussed above has the advantage that the sample conductance is extremely sensitive to the motion of defects lying within the contact area, although only a very small sample volume is probed in a single experiment. In larger samples, such as metal wires fabricated on a substrate using lithographic techniques, it is not clear a priori that the conductance change due to the motion of a single defect is large enough to measure successfully. Following earlier work [73-75], Pelz and Clarke [76] estimated the change in resistivity due to the change in scattering cross section of a single defect. Their "local interference" model predicts for the fractional change in resistivity [76],

\[
\frac{\delta \rho}{\rho} = \beta \sigma \frac{l_c}{V},
\]

where \(\beta \sigma\) is the change in the cross section due to the displacement of the defect, \(l_c\) is the mean free path and \(V\) is the sample volume. With typical values for Bi wires (see sections 3.2.5-3.2.7), the fractional change in sample resistance is only a few parts per million. Such a small change is not easily detectable at very low temperature.

Fortuitously, under some circumstances the sensitivity of the conductance to the motion of single defects is greatly enhanced over the local interference result. In disordered metals at low temperature, the electron motion is diffusive with a step size of order the elastic mean free path, typically a few nm. If inelastic scattering processes are weak, however, the electrons maintain phase coherence over much longer distances, of order a \(\mu\)m at 1 Kelvin. Phase coherence is eventually destroyed either by inelastic scattering or by spin-flip scattering due to magnetic impurities. The sample conductance can be viewed as arising from the interference between all the possible paths an electron can take in traversing the sample. The motion of a single scatterer then affects the phases of all the paths that intersect it. In one- or two-dimensional samples, a significant fraction of the electron paths through the sample intersect any given scatterer.

The ideas expressed above form the basis of the existing quantitative theory of "universal conductance fluctuations", or UCF [77-79]. Originally, UCF theory arose to explain sample to sample fluctuations of the conductance of an ensemble of macroscopically identical, but microscopically distinct, samples. The term "universal" refers to the fact that the conductance fluctuations of sufficiently small samples are of order \(e^2/h\), and are approximately independent of the sample size and shape. The fluctuations are observable experimentally by varying either the Fermi level (in semiconducting samples) or an external magnetic field. Later, UCF ideas were applied to the motion of scattering centers by Feng, Lee and Stone [80], and by Altshuler and Spivak [81]. Those authors showed that the motion of a single scattering center can bring about a conductance change \(\delta G_1\) comparable to \(e^2/h\) in a sample with dimensions \(L\) smaller than the phase-breaking length, \(L_\phi\). The theoretical result is:
\[(\delta G_i)^2 = C \left( \frac{e^2}{\hbar} \right)^2 \frac{(1-x)}{(k_\text{F}l_e)^2} \left( \frac{L}{l_e} \right)^{2-d} \] (24)

where \(C\) is a constant of order unity which depends on sample dimensionality. The factor \((1-x)\) accounts for the phase shift of the electronic wavefunction due to the moving scatterer, c.f. Eq. (13). The last factor, \((L/l_e)^{2-d}\), takes into account the fraction of paths traversing the sample that scatter from the atom in question. In the 1D limit, this term can exceed unity, indicating that electron paths are likely to intersect any given scatterer more than once.

As the temperature decreases, the sensitivity of the sample conductance to the motion of a single defect increases due to the increase of \(L_\varphi\). For the Bi wires to be discussed later, Eq. (24) predicts conductance changes on the order of 0.1 e\(^2\)/h at 1 K. For a sample resistance of 1 k\(\Omega\), that corresponds to a fractional change of a few parts per thousand, i.e. a thousand times larger than the local interference result [76, 82].

Initial studies implicating UCF in single defect fluctuations were carried out by Giordano and coworkers, who observed conductance jumps in thin sputtered films of Bi, Pt and Ag [83, 84]. As the temperature was lowered, they saw large resistance jumps due to the motion of discrete defects. In addition, the overall background noise increased, due to the motion of a large number of defects whose individual signals were too small to resolve in the experiment. The observed increase in both components of the signal with decreasing temperature suggested UCF enhancement.

More quantitative verification of UCF ideas came from studies of 1/f-noise, which results from the fluctuations of an ensemble of defects. Garfunkel et al. [85] measured 1/f noise in C-Cu metal-insulator composites, and found that the conductance noise was temperature independent below about 70 K, consistent with the UCF mechanism. Simultaneous measurements of 1/f noise in polycrystalline Bi films found that the noise increased dramatically below about 70 K as a power law, in agreement with UCF ideas and the tunneling distribution [86]. The different temperature dependence of the noise observed in the two materials may have resulted from the small value of \(L_\varphi\) in the composites. Fig. 3 shows noise versus temperature for three different Bi samples. The most convincing evidence of UCF in these experiments was the observation of a strong magnetic field dependence of the 1/f noise in the Bi samples [86]. UCF theory predicts that the noise should drop by a factor of two in the presence of a magnetic field large enough to break time-reversal symmetry of a diffusive electron trajectory of length \(L_\varphi\). The reduction was observed at a field scale of 0.02 T at 0.5 K, corresponding to a value \(L_\varphi \approx 0.2 \mu m\), in agreement with an independent measurement of \(L_\varphi\) in the same sample [86].

3.2.4 Dynamics of single tunneling systems

The experiments discussed in the preceding section demonstrate the role of UCF in low-temperature transport experiments of defect motion. None of those experiments, however, gave direct information about the dynamics of the moving defects. A first step in this direction was taken by Garfunkel et al. [87], who measured the dynamics of single fluctuators in C-Cu and Si-Au composites over the temperature range 4-30K. Most of the observed fluctuators obeyed
thermally-activated dynamics over this temperature range, although a few showed weakly
temperature-dependent jump rates near 4 K.

Figure 3: Normalized 1/f noise power, evaluated at 1 Hz, vs. temperature, for three Bi samples.
Sample dimensions are: 10 μm x 100 μm x 90 nm (■), 2 μm x 20 μm x 90 nm (▲), and 10 μm x
100 μm x 11 nm (●). The noise increases below about 70 K due to the increase in $L_\phi$. Taken from
[86].

3.2.5 Magnetic field tuning of individual tunneling systems in metals

Studies of single TS in small Bi wires were carried out by Zimmerman et al. [88] at
temperatures below 1K. The utility of thermally evaporated Bi films derives from their fairly
large resistivity. In contrast to semi-metallic crystalline Bi, polycrystalline films are metallic with
carrier densities near $10^{19}$ cm$^{-3}$. This makes it feasible to generate sample resistances ~ $10^3$ ohms.
The experiments were carried out with a relatively high degree of isolation from external
perturbations, using extensive filtering of wires introduced into the cryostat. Conductance
fluctuations arising from the motion of a single defect were stationary; i.e., showed reproducible
average transition rates for time periods exceeding a month. By using electron beam lithography
and thermal evaporation, small wires of length 300 nm and cross-section 80 x 30 nm$^2$ were
fabricated. After cooling the samples below 1 K, it was possible to measure individual defect
random telegraph signals as a function of temperature and external magnetic field.
An example of data from a specific defect is shown in Fig. 4. The conductance (in units of $e^2/h$) shows jumps between two states, with instrumental noise superimposed. At 0.58 K, the conductance spends most of its time in one state, with occasional brief transitions to the other state. As the temperature is increased, two changes occur: the amplitude of the fluctuations decreases but their frequency increases. The first effect arises from the loss of electron phase coherence as the temperature increases. The second effect on the transition rates, which is of primary interest here, occurs because the two-state defect is in equilibrium with a thermal reservoir of conduction electrons whose temperature is raised.

![Figure 4: Conductance fluctuations due to a single tunneling defect in a mesoscopic Bi wire. As the temperature increases, the amplitude of the fluctuations decreases since $L_\phi$ decreases. The time spent in each state also shortens as $T$ increases. By analyzing time series such as these, the transition rates $\gamma_2$ and $\gamma_1$ are obtained.](image)

It is convenient to characterize the dynamics of the fluctuations by computing the mean time in each of the two states, $\tau_1$ and $\tau_2$. Histograms of dwell times in each state are computed and, since they obey Poisson statistics, show an exponential distribution as a function of time characterized by time constant $\tau_i$. If one assumes that only two states exist, separated by an energy $\varepsilon$, by detailed balance the transition rates $\gamma_i = \tau_i^{-1}$ are related by $\gamma_2 = \gamma_1 \exp(\varepsilon/k_B T)$, where $\gamma_2$ is the rate out of the excited state.
Although the energy splitting can be obtained at a single temperature, because parasitic heating by the measuring current or by external sources may be present it is generally useful to measure the rates as a function of temperature. Fig. 5 shows that a plot of $\frac{\gamma_2}{\gamma_1}$ vs. $1/T$ is linear, with slope $\frac{\varepsilon}{k_B}$. The plot shows a surprising effect: the energy splitting of the defect is strongly dependent on the magnetic field $B$, applied perpendicular to the plane of the substrate. Results for two defects are shown in Fig. 6 for field values between $\pm 9$ T. The energy is tuned over a range of $1K$ by the field, is symmetric about $B = 0$, and is nonmonotonic. This is clearly not the result of a Zeeman interaction in which the energy would be linear in $B$.

Figure 5: Plot of switching rate ratios ($\tau_i^{-1}=\gamma_i$) for 3 magnetic fields as a function of inverse temperature. The straight lines show that detailed balance is obeyed, but that $\varepsilon$ is a strong function of magnetic field. The slopes yield defect energy splittings $\varepsilon/k_B = 0.9$ K ($\bigcirc$), 0.4 K ($\Delta$), and 0.1 K ($\square$) [88].

An explanation for the random dependence of $E(B)$ was given by Al’tshuler and Spivak [89]. In a coherent mesoscopic metal there are random accumulations of charge brought about by interference of the conduction electron wavefunctions in the weakly disordered metal. The magnetic field rearranges the charge distribution at the defect site, which depends randomly on $B$. If a charge coupling is assumed between the defect and its surrounding medium, the defect energy is modulated by its environment. It should be noted that conductance fluctuations may not be observable at certain $B$ values at which the defect does not change the UCF amplitude [80].
general, however, B-tuning offers a useful practical advantage in the study of defect dynamics, by allowing manipulation of the environment of the defect so that the conductance fluctuation from a specific defect can be maximized.

Figure 6: Defect energy splitting $\varepsilon$ as a function of applied perpendicular magnetic field at 0.5 K [88]. The two curves correspond to two different defects in the same mesoscopic sample, measured over different time periods. Note the symmetric increase of $\varepsilon(H)$, the tuning effect, with reproducible nonmonotonic features predicted by Altshuler and Spivak [89].

3.2.6 Tunneling dynamics of single defects at low temperatures

With the developments in improved stability and field-tuning mentioned above, it has been possible to confirm several aspects of dissipative quantum tunneling by following the time evolution of a single tunneling defect below 1 K. Perhaps the most striking observation is the increase in tunneling rates with decreasing temperature. Experiments were carried out with Bi wires using the same methods described in 3.2.5. The initial observation [90] noted a minimum in the transition rate of a nearly symmetric defect at 1 K, followed by a rate increasing by more than 50% just below 0.5 K. When the asymmetry energy $\varepsilon$ of the defect was increased by a magnetic field, the anomalous temperature-dependent rate was suppressed.

With magnetic field tuning, the energy of the defect was varied from 0.08 to 0.9 K, so that $y = k_B T / \varepsilon$ varied over a substantial range, allowing a quantitative comparison with theory. Taking as a starting point Eq. 22, the expression for the fast transition rate can be cast in a scaled form, so that data from different energies can be represented on a single universal plot. Fig. 7 shows $(T/T_o)^{1-2K} \gamma_f$ plotted vs. $y = k_B T / \varepsilon$. The data indeed fall on a single curve, in which the dissipative tunneling theory is shown by the solid line, with parameters $K = 0.24$ and $\hbar \Delta / k_B = 4.4 \times 10^{-7}$ K. The peak in the scaled rate occurs near $y = 1$ and separates the behavior into two regimes. For $y<1$, the slope is positive, so that increasing temperature corresponds to increasing tunneling rate.
However, for $y > 1$, the scaling function is nearly constant and corresponds to the region for which $\gamma_f$ varies as $T^{2K-1}$, or $T^{-1.52}$ for this specific defect.

Figure 7: Plot of scaled fast transition rate vs $k_BT/\varepsilon$, where $T_0=1$ K. The solid line is the prediction of the dissipative quantum tunneling model with conduction electron-defect coupling $K=0.24$. The data are for a single defect whose energy has been tuned into different $k_BT/\varepsilon$ regions by the external magnetic field. Note the sign change of the slope at $y=1$ [90].

A more comprehensive test of the dissipative tunneling model was provided by experiments on an identical physical system [91]. Single defects were field-tuned but were followed over a wider temperature region. This allowed the theoretical model (Eqs. 20 & 22) to be fit directly to the temperature-dependent tunneling rates. Fig. 8 shows data from a single defect at three different values of magnetic field, with values of the energy asymmetry $\varepsilon/k_B$ between 92 and 402 mK. The data show three different temperature regimes. The low-temperature regime is seen most clearly at $B=6.99$ T. Here the faster rate is nearly temperature-independent, while the slower rate decreases rapidly due to the detailed balance relation, Eq. (20). In the intermediate temperature regime, seen most clearly in the $B=2.286$ T data, the rates decrease with increasing temperature, and obey the Kondo power law, $T^{2K-1}$. Above 1.5 K, the rates increase rapidly, due to thermal excitation to excited vibrational states in the wells [92], or coupling of the TS to phonons [93, 94]. (The data in Fig. 8 do not extend to sufficiently high temperature to distinguish between these two possibilities.) The solid lines are fits to the data below 1.2 K only, where the model of a two-level system coupled only to the conduction electrons is appropriate. The data from all values of magnetic field were fit simultaneously to Eqs. (20) and (22) with a unique value of $K$, but with the renormalized tunneling rate allowed to vary with field [95]. The fit yields the values $K = 0.195$ and $\hbar \Delta_r/k_B = 1.0 \times 10^{-7}$ K, with the latter quantity varying by only a few percent with field. An alternative fitting procedure, allowing $K$ also to vary with field, produces fits that are only
slightly better than those shown in Fig. 8. Data over this relatively wide temperature range are insufficient to determine if $K$ is field-dependent.

Figure 8: Tunneling rates vs. temperature of a single defect in a mesoscopic Bi wire. The lines are from a global fit of Eqs. (20) & (21) with a single value of the conduction electron-defect coupling parameter, $K=0.195$ [91, 95]. Above about 1.5 K, the tunneling rates increase rapidly, probably due to coupling of the tunneling defect to phonons [93].
3.2.7 Strain modulation of tunneling systems

The magnetic field tuning of single tunneling defects discussed earlier proved to be a valuable tool, allowing one to explore a wider range of $k_B T/\varepsilon$ than was possible by varying the temperature alone. One can modulate the double-well potential asymmetry in an alternate way by applying strain to the sample. It is known from phonon echo experiments that TS in glasses have deformation potentials of order 1 eV [39]. Experiments probing the coupling of a single TS to static strain were carried out on Bi wires similar to those described earlier [96, 97]. Strain was imposed by fabricating the samples on thin Si chips attached to piezoelectric substrates, allowing application of isotropic expansion or contraction in the plane of the substrate. Fig. 9 shows the asymmetry energy $\varepsilon$ of a single defect in a Bi wire as a function of strain. Although the maximum applied strain was only in the range of $10^{-5}$, the effect on $\varepsilon$ is quite noticeable. The deformation potential, $d\varepsilon/d(\text{strain})$, for this particular defect was 3.4 eV, which was the highest value observed among 6 different defects studied.

This defect produced large enough conductance jumps to be observable at temperatures as high as 3 K. Since the lower temperature limit for this particular experiment was 1.1 K, the electron-dominated tunneling regime was not observed. Over this narrow temperature range, the total tunneling rate increased by 3 orders of magnitude, providing enough information to determine the mechanism governing the rapid upturn in the tunneling rates observed previously for similar defects above about 1 K. Rather than an Arrhenius form, as would arise from a thermally activated process [92], the tunneling rate is exponential in $T^2$, consistent with the phonon coupling model discussed at the end of Section 2 [42, 93]. Fig. 10 shows that a plot of $\log(\gamma)$ versus $T^2$ produces a straight line whose slope gives a value of the characteristic temperature $T_{\text{ph}} \approx 0.8$ K, corresponding to a deformation potential $d\varepsilon/d(\text{strain}) \approx 3.2$ eV. Since this result represents an average over all phonon modes and polarizations, while the previous result concerns a particular static deformation, this good agreement may be somewhat fortuitous.

3.2.8 Electron scattering from tunneling systems

The off-diagonal term in Eq. 8 allows for electron-assisted tunneling between the two positional eigenstates of the TS. In the non-commutative model introduced by Zawadowski and co-workers [98, 99], the effective matrix element $V'$, which is normally much smaller than the bare $V$ coupling, is renormalized upwards as the temperature decreases below a characteristic Kondo temperature $T_K$. The state formed by the coupled conduction electron-TS has the characteristics of the two-channel Kondo system, with a symmetric TS providing a spin-like degree of freedom. In the two-channel Kondo model, the enhanced electron scattering is manifested by an electrical resistivity $\rho(0)-\rho(T)$ proportional to $\log T$ for $T >> T_K$, whereas for $T << T_K$ the resistivity is expected to vary as $T^{1/2}$. The reader is referred to Ref. [100] for a comprehensive discussion of the theory.
Evidence for the presence of a tunneling-system-mediated two-channel Kondo state at temperatures near 1 K has been presented by Ralph and Buhrman [101, 102]. Using point-contact spectroscopy, they measured the current-voltage (I-V) characteristics of nominally pure copper, evaporated on both sides of Si$_3$N$_4$ membrane punctured by a small hole with diameter $d \sim 3$ nm. The small diameter of the aperture results in ballistic transport in the Cu film in the vicinity of the hole, since $l_e > d$, and provides the geometrical enhancement that makes the point contact sensitive to scattering by a small number of centers. Multiple elastic scattering by static scatterers is therefore not necessary for observation of the Kondo state.

The features observed in conductance-voltage (dI/dV vs. V) spectra that strongly indicate the formation of a two-channel Kondo state below 10 K are (a) a minimum in the conductance at zero-bias ($V = 0$), (b) a conductance enhanced over the normal state, i.e. defect-free, conductance at moderate bias (a few mV), and (c) the presence of abrupt conductance spikes, interpreted as a transition between the Kondo and a normal state at yet higher bias. At very low bias, the normalized differential conductance $[G(V, T) - G(0, T)] / T^{1/2}$ scales as $f(eV/k_B T)$, where $f$ is a scaling function consistent with the two-channel Kondo theory [103].

Figure 9: Energy asymmetry of a single defect in a Bi wire vs. applied static strain. The strain is isotropic in the 2D plane of the substrate. The slope of the strain response corresponds to a deformation potential of $d\varepsilon/d$ (strain) = 3.4 eV [97].
Figure 10: Tunneling rate vs. temperature for the Bi defect shown in Fig. 9 [97]. Over the temperature range shown, the rate increases with temperature as $\gamma \propto \exp[(T/T_{ph})^2]$, as predicted theoretically for coupling to phonons [93].

It is worth pointing out that the evidence implicating two-level tunneling states as the source of the Kondo-like scattering is circumstantial. Since the TS are presumed “fast”, i.e., they are symmetric states of energy ~ $T_K$ (0.5 to 5 K), they are not directly observed, as was possible for the “slow” TS discussed above. However, the anomalous G-V characteristics occur only in those samples cooled to liquid helium temperatures soon after evaporation, suggesting that the states necessary for the Kondo scattering anneal out quickly at room temperature, implicating structural disorder. Samples intentionally prepared with dilute magnetic impurities do not exhibit similar features. Application of a high magnetic field in nominally pure samples destroys the Kondo state but does not yield characteristics associated with Zeeman splitting. All of these features require an entity with an internal degree of freedom that is similar to a true spin but depends upon the presence of structural disorder, i.e., a two-state TS.

However, alternative explanations of the tunneling data have been proposed. Wingreen et al. have suggested that electron-electron interactions in the presence of disorder, which leads to a suppression of the density of states at $E_F$, can be important [104, 105]. The interactions will give rise to a $V^{1/2}$ dependence of the tunneling current near zero bias. More recently, Aleiner and Controzzi [106] have argued that, after correctly mapping the TS problem onto the two-channel Kondo model, it is not possible to reach the strong coupling regime where $E, T < T_K$. Even for
symmetric TS, $T_k$ is always smaller than $\Delta$. Clearly, more definitive tests are required, possibly utilizing TS with known tunneling matrix elements.

4. CONCLUSIONS AND OUTLOOK

In this review we have shown how measurements of mesoscopic samples have been used to extract detailed information on the many-body interactions that are intrinsic to atomic tunneling in conductors. Such studies have taken advantage of the sensitivity of electrical transport in this regime to examine single tunneling entities, eliminating the uncertainty that accompanies averaging over distributions of states. As a result, the experiments have yielded dramatic confirmation of theoretical conjectures regarding screening and ground states of the coupled electron-atom system.

On a practical note, with the development of smaller and smaller device dimensions, mobile defects are assuming greater importance in defining the limits of device performance and reliability. In MOS structures, the dynamics of trapping states in the oxide barrier or at its interfaces, influences the speed at which the devices can operate. Telegraph and 1/f noise associated with defects are clearly undesirable features in such devices. In metal interconnects, large current densities are responsible for electromigration processes that ultimately result in catastrophic failures. These current-driven diffusive events are also intimately associated with lattice defects such as vacancies and dislocations.

In closing, we remark that we have necessarily neglected a number of parallel developments made possible by increasingly powerful scanning probe microscopies. Scanning tunneling microscopy, in particular, has been used to study the conformations, dynamics, and spectral characteristics of a variety of atoms and molecules on surfaces. The ability to create spatial images of atomic configurations is an enormous advantage in deciphering structural arrangements and their slow dynamics [107]. Likewise, fluorescent emission from single molecules has been used to examine the spectral characteristics of two-level systems associated with tunneling processes [11]. In the future, we expect that experiments invoking combinations of these local probes will yield unprecedented detail about the dynamics of a much wider variety of individual tunneling systems.

ACKNOWLEDGMENTS

Our own experimental work on dissipative quantum tunneling would not have been possible without the hard work of our students, post-docs, and other collaborators over the past several years. We gratefully acknowledge the contributions of G.B. Alers, K. Chun, S. Coppersmith, W.H. Haemmerle, D. Hoadley, J.S. Moon, and N. Zimmerman. We have also benefited from numerous discussions of the theory with many people. We especially thank B.L. Altshuler, R.I. Cukier, H. Grabert, A.J. Leggett, and A. Würger.

We are grateful to AT&T Bell Laboratories, the Michigan State University Center for Fundamental Materials Research, and the National Science Foundation for support of this research, the latter under grants DMR-9023458, DMR-9312544, DMR-9321850, DMR-9801841, and DMR-9971537.
REFERENCES


