ELECTRON TUNNELLING STUDY
OF HIGH-TEMPERATURE
SUPERCONDUCTORS

S.J. CHANDLER

Jesus College and the Cavendish Laboratory,
Cambridge, U. K.

Thesis submitted for the degree of Doctor of Philosophy
at the University of Cambridge, August 1993
TO MY FAMILY
PREFACE

I would like to thank my supervisor, Dr C.J. Adkins, and my fellow students at the Cavendish laboratory, especially Anthony Holden, Alison Speakman and Simon Dunkley, for their friendship and help over the past six years. I would also like to thank the workshop and electronics department technicians for their assistance and patient guidance, and Ken Gadsby and Pete Flaxman for supplying the cold stuff. Special thanks to Dave Swainston and Paul Booth of the LTP workshop, and Dave Powell of the student workshop for (hopefully) imparting to me some of their considerable skill and knowledge. Thanks also to my eagle-eyed copy editor, Sarah Hamilton, and to my colleagues at Hewlett-Packard (CPB) for help with the computing and DTP facilities.

I would like to thank the Science and Engineering Research Council, the Cambridge Philosophical Society and the Steeple Bumpstead Educational Charity for financial assistance.

Thanks also to my family for their support over the years and to Louise and Isobel for their love and patience - ‘it’ is finally finished.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where reference or acknowledgement is made in the text. The dissertation is not substantially the same as any that I have submitted for a degree, diploma or other qualification at any other University. I further state that no part of my dissertation has already been or is being concurrently submitted for any such degree, diploma or qualification.

Simon Chandler
August 14, 1993
ABSTRACT

This dissertation describes work carried out between June 1987 and October 1991, in the Low Temperature Physics Group at the Cavendish Laboratory, Cambridge.

The aim of this work was to use electron tunnelling spectroscopy to measure the density of excitation states of the recently discovered high-temperature superconductors. Tunnelling is the most sensitive method for measuring a superconductor’s energy gap, and historically has provided important evidence for the microscopic mechanism of superconductivity in conventional metals. It was hoped that electron tunnelling would prove equally successful in revealing the mechanism of superconductivity in these new materials.

Preliminary experiments showed that a thick, degraded surface layer prevented preparation of high-quality tunnel junctions by conventional evaporation techniques. For this reason, apparatus for the formation and fine control of low-temperature point-contact junctions was constructed, together with new measurement electronics and a computer-controlled data-acquisition system.

To test this apparatus, point-contact junctions were formed on conventional superconductors. By increasing pressure of the tip on the sample the junction could be moved from the tunnelling to the metallic regime.

Point-contact measurements were then made on a number of ceramic, single-crystal and thin-film high-temperature superconducting materials; some not previously investigated by tunnelling. Although ‘gap-like’ structure was occasionally observed, anomalous features (e.g., voltage-dependent background, broadening, large zero-bias conductance) were always present and usually dominated the tunnelling characteristics. These complicate estimation of the energy gap and preclude measurement of more subtle properties such as gap anisotropy or the effective phonon spectrum, $\alpha^2 F$. The origins of these non-ideal features have been much debated in the literature and are reviewed in this dissertation.

In the case of thin films deposited by laser ablation the tunnelling characteristics were dominated by single-electron tunnelling effects (Coulomb gap and staircase structure). The results suggest that the surface region consists of numerous, isolated normal metal particles.
4.3 Broadening .................................................................................................................. 76
  4.3.1 Inhomogeneous Gap .................................................................................................. 77
  4.3.2 Anisotropic Gap ........................................................................................................ 81
  4.3.3 Lifetime Broadening ................................................................................................. 87
4.4 Charging Effects: Coulomb Gaps and Staircases .......................................................... 88
4.5 Surface Layers .............................................................................................................. 92
4.6 $2\Delta/k_B T_c$ : The Reduced Gap Value of High-$T_c$ Superconductors ......................... 97
4.7 Phonon Effects .............................................................................................................. 101
4.8 Halbritter’s Model ........................................................................................................ 102

Chapter 5  Experimental Details: I. Measurement Electronics .............................................. 105
5.1 A.C. Measurement of Conductance ............................................................................... 105
  5.1.1 Conductance-Bridge Measurements ........................................................................ 108
5.2 D.C. Measurement of Conductance .............................................................................. 110
  5.2.1 Calibration of Raw Data ....................................................................................... 111
  5.2.2 Numerical Differentiation and Smoothing .............................................................. 115
  5.2.3 Additional Features of the D.C. Measurement System ........................................... 118
  5.2.4 Possible Improvements to the D.C. Measurement System ...................................... 118

Chapter 6  Experimental Details: II. Junction Fabrication .................................................. 121
6.1 Review of Junction Types .............................................................................................. 121
  6.1.1 Planar Junctions ....................................................................................................... 121
  6.1.2 Point-Contact Junctions ......................................................................................... 123
  6.1.3 Break Junctions ...................................................................................................... 125
  6.1.4 Scanning Tunnelling Microscopes ......................................................................... 125
6.2 Planar Junctions ........................................................................................................... 126
6.3 Point-Contact Junctions .............................................................................................. 127
  6.3.1 Design Requirements ............................................................................................ 128
  6.3.2 Vibration Isolation ................................................................................................. 128
  6.3.3 Fine Control of the Tip–Sample Separation ............................................................ 134
  6.3.4 Course Control of the Tip–Sample Separation ....................................................... 142
  6.3.5 Positioning the Sample ......................................................................................... 148
  6.3.6 Preparation of the Tunnelling Tip .......................................................................... 149
  6.3.7 Operation of the Point-Contact Insert .................................................................. 152
  6.3.8 Suggested Improvements ..................................................................................... 153
Chapter 7  Results and Discussion .................................................................................................... 155
  7.1 Conventional Superconductor Point-Contact Junctions .............................................................. 155
    7.1.1 Pb SIN Junctions ............................................................................................................. 155
    7.1.2 Pb ScN Junctions .......................................................................................................... 160
    7.1.3 Nb ScN Junctions .......................................................................................................... 161
  7.2 High-\(T_c\) Superconductor Planar Junctions ........................................................................ 162
    7.2.1 \(YBa_2Cu_3O_y\) Ceramic ............................................................................................. 162
  7.3 High-\(T_c\) Superconductor Point-Contact Junctions .............................................................. 163
    7.3.1 \(YBa_2Cu_3O_y\) Ceramic ............................................................................................. 163
    7.3.2 Bi-Sr-Ca-Cu-O Single Crystals .................................................................................... 171
    7.3.3 \((Ti,Pb)(Ca,Pr)Sr_2Cu_2O_y\) Ceramic ......................................................................... 176
    7.3.4 \(YBa_2Cu_3O_y\) Thin Films ........................................................................................ 182
Chapter 8  Conclusion .................................................................................................................. 193
  8.1 Summary ..................................................................................................................................... 193
  8.2 Recommendations for Future Work ........................................................................................... 194
Appendix A  Electronic Circuit Details .......................................................................................... 197
  A.1 Measurement of Conductance–Voltage Characteristics .......................................................... 197
    A.1.1 Circuit Design ............................................................................................................. 197
    A.1.2 Modulation Broadening .............................................................................................. 203
  A.2 Servo Control of the Tunnelling Tip ......................................................................................... 204
Appendix B  Low-\(T_c\) Oxide Superconductors ............................................................................ 207
  B.1 Bismuthate Superconductors .................................................................................................. 207
  B.2 \(Nd_{2-x}Ce_xCuO_{4-y}\) .......................................................................................................... 209
References ........................................................................................................................................ 211
CHAPTER 1
INTRODUCTION

1.1 High-Temperature Superconductors

In 1986 a revolution occurred in the field of superconductivity. After a three-year search for high-temperature superconductivity in oxide materials, J.Georg Bednorz and K.Alex Müller were rewarded with the discovery of a complex Ba-doped La–Cu oxide that had a transition temperature \(T_c\) of 30 K*. C.W.Chu replaced La with Y and produced a material with a staggering critical temperature of 92 K - overturning a widely held belief, substantiated by theoretical arguments, that the previous highest \(T_c\) of 23.2 K was close to a physical upper bound. There followed an explosion of interest in high-temperature superconductivity with huge world-wide effort directed towards raising \(T_c\) still further and understanding the properties of these materials.

The many high-temperature superconductors that are now known (table 1.1) have a common basic structure consisting of stacked layers of square-planar \(\text{CuO}_2\)†. These are partitioned into groups of \(n\) closely spaced planes with larger distances between groups. \(\text{CuO}_2\) planes within a group are separated by a sparsely occupied layer of metal atoms whilst the inter-group space is occupied by denser oxide ‘isolation layers’ (figure 1.1). This highly anisotropic structure is reflected in anisotropy of many physical properties - suggesting a quasi-two-dimensional character in which the charge carriers and superconductivity are confined to the \(\text{CuO}_2\) planes.

A second common property is that superconductivity occurs in the cuprate materials close to a metal–insulator transition, with the Cu ion magnetic moments antiferromagnetically ordered in the insulating phase. The superconducting metallic phase is obtained by doping the insulator with excess charge, for example by adding oxygen to \(\text{YBa}_2\text{Cu}_3\text{O}_6\), or by replacing \(\text{La}^{3+}\) with \(\text{Sr}^{2+}\) in \(\text{La}_{2-x}\text{Sr}_x\text{CuO}_4\).

---

* The full story was told in the 1987 Nobel prize in Physics award speech, see Bednorz and Müller (1988) for a transcript.

† In this dissertation the classification ‘high-temperature superconductor’ refers only to the layered cuprate materials. In particular, it does not include the bismuthate or buckminsterfullerene superconductors (e.g., \(\text{Ba}_1\text{K}_2\text{BiO}_3\) with \(T_c = 30\) or \(\text{Rb}_{2.7}\text{TI}_{2.2}\text{C}_{60}\) with \(T_c = 42.5\) K). ‘Conventional superconductors’ include the elemental and metal-alloy superconductors.
<table>
<thead>
<tr>
<th>Material</th>
<th>$T_c$ (K)</th>
<th>$n$</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>La$_{2-x}$Sr$_x$CuO$_4$</td>
<td>38</td>
<td>1</td>
<td>LSCO or 214</td>
</tr>
<tr>
<td>La$_{2-x}$Sr$_x$CaCu$_2$O$_6$</td>
<td>60</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Nd$_{2-x}$Ce$_x$CuO$_4$</td>
<td>30</td>
<td>1</td>
<td>NCCO</td>
</tr>
<tr>
<td>YBa$_2$Cu$_3$O$_7$</td>
<td>92</td>
<td>2</td>
<td>YBCO or 123</td>
</tr>
<tr>
<td>YBa$_2$Cu$_4$O$_8$</td>
<td>80</td>
<td>2</td>
<td>124</td>
</tr>
<tr>
<td>Y$_2$Ba$_4$Cu$<em>7$O$</em>{14}$</td>
<td>40</td>
<td>2</td>
<td>247</td>
</tr>
<tr>
<td>Bi$_2$Sr$_2$CuO$_6$</td>
<td>0–20</td>
<td>1</td>
<td>Bi-2201</td>
</tr>
<tr>
<td>Bi$_2$Sr$_2$CaCu$_2$O$_6$</td>
<td>85</td>
<td>2</td>
<td>BSCCO or Bi-2212</td>
</tr>
<tr>
<td>Bi$_2$Sr$_2$Ca$_2$Cu$<em>3$O$</em>{10}$</td>
<td>110</td>
<td>3</td>
<td>Bi-2223</td>
</tr>
<tr>
<td>TlBa$_2$CuO$_5$</td>
<td>0–50</td>
<td>1</td>
<td>Tl-1201</td>
</tr>
<tr>
<td>TlBa$_2$CaCu$_2$O$_7$</td>
<td>80</td>
<td>2</td>
<td>Tl-1212</td>
</tr>
<tr>
<td>TlBa$_2$Ca$_2$Cu$_3$O$_9$</td>
<td>110</td>
<td>3</td>
<td>Tl-1223</td>
</tr>
<tr>
<td>TlBa$_2$Ca$_3$Cu$<em>4$O$</em>{11}$</td>
<td>122</td>
<td>4</td>
<td>Tl-1234</td>
</tr>
<tr>
<td>Tl$_2$Ba$_2$CuO$_6$</td>
<td>0–80</td>
<td>1</td>
<td>Tl-2201</td>
</tr>
<tr>
<td>Tl$_2$Ba$_2$CaCu$_2$O$_8$</td>
<td>108</td>
<td>2</td>
<td>Tl-2212</td>
</tr>
<tr>
<td>Tl$_2$Ba$_2$Ca$_2$Cu$<em>3$O$</em>{10}$</td>
<td>125</td>
<td>3</td>
<td>Tl-2223</td>
</tr>
<tr>
<td>HgBa$_2$CuO$_5$</td>
<td>94</td>
<td>1</td>
<td>Hg-1201 †</td>
</tr>
<tr>
<td>HgBa$_2$CaCu$_2$O$_7$</td>
<td>112</td>
<td>2</td>
<td>Hg-1212 †</td>
</tr>
<tr>
<td>HgBa$_2$Ca$_2$Cu$_3$O$_9$</td>
<td>133.5</td>
<td>3</td>
<td>Hg-1223 †</td>
</tr>
<tr>
<td>Sr$_{1-y}$Nd$_y$CuO</td>
<td>40</td>
<td>$\infty$</td>
<td>all-layer phase ‡</td>
</tr>
</tbody>
</table>

Table 1.1. Transition temperatures of the principal high-$T_c$ superconductors. Note that $T_c$ generally increases with $n$. Also shown are commonly used names for these materials. Where $T_c$ depends on stoichiometry, higher values are shown. After Burns (1992) except † (Edwards 1993) and ‡ (Cava 1991).
The existence of very high transition temperatures immediately oppugned the well established BCS theory of superconductivity with phonon-mediated electron pairing. To account for the observed high $T_c$ values BCS-like theories with higher-energy bosonic interactions (e.g., magnons) were proposed or brought out of mothballs. Other more exotic theories also flourished amidst the excitement and lack of clear guidance from experimental data; it was often said that there were ‘as many theories as theorists’. This chaotic situation highlights the importance of experimental results that are able to place constraints on possible theories (Little 1988). An important class of such experiments are those that measure the energy gap, $\Delta$, in the excitation spectrum of a new superconductor. The most sensitive (and historically the most important) method for measuring this gap is by electron tunnelling.

1.2 Electron Tunnelling

Tunnelling is a phenomenon of quantum mechanics whereby an electron, as a consequence of its wave-like character, may pass from one conducting electrode to another through a thin ($\leq 1$ nm) intervening potential barrier that would confine a classical particle. An artificial structure in which the conductance is predominantly
due to tunnelling electrons is called a ‘tunnel junction’. The voltage applied to a
junction determines the range of energies of tunnelling electrons whilst the current
depends on properties of the barrier and electrode materials. This forms the basis of a
simple but powerful energy spectroscopy, with high resolution of order $k_B T$, which is
only 0.36 meV at the temperature of liquid helium. In particular, when one of the
electrodes is superconducting the tunnelling electrons probe important low-lying
excitations and the dynamic conductance is proportional to the density of excitation
states.

The importance of electron tunnelling as a tool for investigating the origin of
superconductivity in conventional metals and alloys may be illustrated by reference to
its role in the progress of BCS theory. Tunnelling was first used to investigate the
superconducting state in the early 1960s. These experiments demonstrated in a very
direct way the existence of an energy gap and verified detailed predictions of the 1957
BCS theory of superconductivity. Many materials were found to have a gap ratio,
$2\Delta/k_B T_c$, in agreement with the BCS value of 3.52. Further work, however, revealed
anomalous behaviour in Pb and Hg, manifested as a larger gap ratio ($\approx 4.6$) and
structure in the tunnelling conductance ‘above the gap’ (i.e., at energies in the range
4–8$\Delta$). It was soon realised that these deviations occurred at phonon energies,
supporting the idea of a phonon-mediated pairing interaction, and a modified BCS
theory was developed that took into account the effects of strong electron–phonon
coupling.

With this distinguished record it was hoped that electron tunnelling would prove
equally successful in revealing the mechanism of high-$T_c$ superconductivity.
However, formidable problems due to the short coherence length in these materials
and their propensity to form a non-conducting surface layer have hindered progress.
The tunnelling characteristics of cuprate superconductors contain many non-ideal
features when compared with conventional superconductor tunnel junctions (e.g.,
voltage-dependent background, broadening, ‘multiple’ energy gaps). Most of these
have been observed (to a lesser degree) with conventional materials and have usually
been explained as extrinsic effects. Their presence complicates estimation of the
energy gap and precludes measurement of more subtle properties such as gap
anisotropy or the effective phonon spectrum, $\alpha^2 F$. Despite an enormous amount of
experimental work over the last six years the interpretation of tunnelling results and
the mechanism of high-$T_c$ superconductivity remain controversial.
CHAPTER 2

MICROSCOPIC THEORY OF SUPERCONDUCTIVITY

This section outlines the BCS model of superconductivity with emphasis on those topics that are important to the interpretation of tunnelling measurements. BCS theory is based on an assumption that under certain conditions an attractive interaction can exist between electrons near the Fermi surface of a metal. This binds them into pairs and results in the dramatic changes in thermodynamic and electrical properties that are characteristic of the superconducting state. We first discuss a possible mechanism for the seemingly paradoxical attraction between electrons, then a simple argument will be used to show that once this attraction exists the ‘Fermi sea’ of the normal metal can no longer be the ground state. The BCS derivation of the actual superconducting ground state will then be summarised and finally the properties of single-particle excitations from this state will be discussed.

2.1 An Attractive Electron–Electron Interaction

BCS theory does not stipulate the nature of the attractive interaction between electrons; however, discussion of the theory is made clearer by reference to a particular mechanism. In 1950 Fröhlich proposed that such an interaction could occur by the exchange of virtual phonons. This is believed to be the origin of superconductivity in most metals and alloys.

Fig. 2.1. Electron–electron interaction by exchange of a virtual phonon. Conservation of momentum requires \( l = l' + q \) and \( k + q = k' \) but because the phonon is short lived the uncertainty principle allows \( E_{l'} - E_k \neq \hbar \omega_q \).
An attractive interaction may be considered to arise when an electron polarises a lattice of positive ions to such an extent that it is *overscreened*, resulting in an outwardly pointing \( \mathbf{E} \)-field, which exerts an attractive force on a second electron. One would not expect such a field to persist since it would obviously attempt to disperse the positively charged ions that created it; however, there are two ways in which overscreening can occur. Firstly, a lattice that is unstable due to internal mechanical forces may be stabilised by the additional electric force created by overscreening. Secondly, even in a stable lattice, an oscillating charge density will be overscreened if it is vibrating just below the lattice resonant frequency, \( \omega_q \), where the response of the ions is large (de Gennes 1966, p.102). The charge oscillation may arise if one electron creates a virtual phonon that is later absorbed by a second electron (figure 2.1). Whilst the phonon is being transferred, state \( \mathbf{k} \) is mixed with \( \mathbf{k}' \) so the wavefunction is

\[
\Psi_k(\mathbf{r}, t) = e^{i(\mathbf{k} \cdot \mathbf{r} - E_k t / \hbar)} + b e^{i(\mathbf{k}' \cdot \mathbf{r} - E_{k'} t / \hbar)}
\]

resulting in a charge density change,

\[
\delta \rho = e \sum_k (\Psi_k' \cdot \Psi_k - 1)
\]

which contains oscillatory terms with frequency \( \omega = (E_k - E_{k'})/\hbar \). If the difference in energy between the initial and final states is smaller than the energy of lattice vibrations, \( \hbar \omega_q \), then the scattering matrix element,

\[
V_{kk'} = \langle \Psi_k' | V(\mathbf{r}_1 - \mathbf{r}_2) | \Psi_k \rangle
\]

is negative and the interaction is attractive.

The maximum change in momentum of a scattered electron is the momentum of the exchanged phonon, \( \mathbf{q} \), which, if the metal is well below the Debye temperature, is much smaller than the Fermi wavevector \( k_F \). In addition, the electron must be scattered from a full state into an empty state so all scattering events take place in a thin shell of width \( \hbar \omega_D \) around the Fermi surface. BCS theory and the Cooper pair model use an extremely simplified interaction \( V_{kk'} \), which is assumed to be constant and attractive up to the Debye frequency (i.e., for \( |E_k - E_{k'}| < \hbar \omega_D \)) and zero elsewhere. The model interaction is simplified further by requiring that both \( |\epsilon_k| \) and \( |\epsilon_{k'}| \) are separately smaller than the Debye energy, where \( \epsilon \) is the energy relative to the Fermi energy, \( E_F \).

---

* The following conventions will be used: \( \epsilon \) is an energy measured with respect to a Fermi energy, \( \xi \) is an excitation energy and \( E \) is a general energy, measured with respect to zero at infinity.
i.e.,

\[
V_{kk'} = \begin{cases} 
    -V & |\varepsilon_k'| < \hbar \omega_D, \\
    0 & \text{otherwise}
\end{cases}
\]  

(2.1)

This simplified model will be acceptable only if electrons scattered by \( V_{kk'} \) are restricted to a very narrow energy range that is well away from the cut-off energy \( \hbar \omega_D \); the true interaction would then appear to be energy independent and the introduction of an artificial cut-off in the model interaction will have little effect. As will be shown, scattering electrons in a superconductor are confined to a thin shell of width \( \Delta \) around \( E_F \) so (2.1) may be used if \( \Delta \ll \hbar \omega_D \). In most superconductors this is reasonably well satisfied and the BCS weak-coupling model accounts well for the superconducting properties.

### 2.2 Cooper Pairs

In 1956 Cooper considered what would happen if two electrons were added to a Fermi sea at \( T = 0 \). The two electrons were allowed to interact with each other via the potential \( V(r_i - r_j) \) but not with the free electron gas (other than by Pauli exclusion, which forced them to occupy states above the Fermi level). The added electrons were considered to be in states with opposite momenta, \( \mathbf{k} \) and \( -\mathbf{k} \), so that their centre of mass was at rest. The wavefunction of the two electrons is then

\[
\Psi(r_i - r_j) = \sum_k g(k) \exp[i\mathbf{k} \cdot (r_i - r_j)]
\]  

(2.2)

With the simplified interaction the probability of occupancy of the pair state \((\mathbf{k}, -\mathbf{k})\), \( g(k) \), is independent of the direction of the wavevector so the spatial part of \( \Psi(r_i - r_j) \) is symmetric with respect to exchange of the electrons. To satisfy Pauli exclusion and the indistinguishability of the pair electrons (Eisberg and Resnick 1985, p.308) the total wavefunction must be antisymmetric so the electrons must be in an antisymmetric singlet spin state, \( \frac{1}{\sqrt{2}} (\uparrow \downarrow - \downarrow \uparrow) \), with oppositely directed spins. The electron pair is represented by \((\mathbf{k}, -\mathbf{k})\).

Treating \( V(r_i - r_j) \) as a perturbation of the Hamiltonian for the non-interacting pair and using the simplified potential (2.1) it can be shown (see, e.g., van Duzer and Turner 1981, p.42) that the wavefunction of the pair forms a stationary state with a single energy eigenvalue (relative to \( E_F \)) of

\[
W = \frac{2\hbar \omega_D}{[1 - \exp(2/N(0)V)]}
\]
This is negative for any $V > 0$ and means that an attractive interaction, *however weak*, will cause the two electrons to form a bound state, which will lower the system energy. The Fermi sea will then be unstable since perturbations that place two electrons above the Fermi level are energetically favoured; electrons will continue to condense into pairs until the system is so greatly changed that further pairing does not lower its energy.

The assumption that electrons pair up with oppositely directed momenta is justified by noting that the binding energy of a pair increases with the number of states available for scattering ($\propto N(0)$, the density of states of free electrons at the Fermi surface). As illustrated in figure 2.2, this number is sharply peaked for pairs with $k_1 = -k_2$, so pairs with zero total momentum are energetically most favoured.

**Fig. 2.2.** Graphical construction to show that the number of available scattering states is greatest for $K = k_1 + k_2 = 0$. Consider initial states $k_1$ and $k_2$ that lie on the Fermi surface as shown. (i) If $k_1$ is elastically scattered it must still lie on the Fermi surface and so, to satisfy conservation of momentum, $k_2$ must lie somewhere on the displaced Fermi surface $K - k_F$. In addition, the change in $|k_2|$ cannot be greater than $|q_D|$ so $k_2$ must lie on the line labelled $\ell-\ell'$. (ii) If $k_1$ is also allowed to change by up to $q_D$ then $k_1$ and $k_2$ must be scattered into states in the shaded annular regions. (iii) The volume of this ring is greatest when $K = 0$ since it then includes the complete shell of width $2|q_D|$ around the Fermi surface $\Rightarrow$ the number of states available for $k_1$ and $k_2$ to scatter into is greatest for $k_1 = -k_2$. 


2.3 BCS Theory

Cooper’s analysis showed that a Fermi sea cannot be the superconducting ground state and, without calculating it explicitly, indicated some important properties of the correct ground state. In 1957 Bardeen, Cooper and Schrieffer presented a complete microscopic theory of superconductivity. Their model was based on the following assumptions:

- In the ground state at $T = 0$ all electrons form Cooper pairs, so states $\mathbf{k} \uparrow$ and $-\mathbf{k} \downarrow$ are either both occupied or both vacant.
- Only the net attractive interaction between electrons in a pair needs to be considered. BCS theory does not actually specify that the interaction is phonon mediated, however, it must have the form of (2.1), i.e., be constant up to some cut-off and zero thereafter.

BCS used a trial wavefunction${^*}$,

$$\Psi_G = \prod_{\mathbf{k}} (u_{\mathbf{k}} + e^{i\theta} v_{\mathbf{k}} b_{\mathbf{k},S}^\dagger)|0\rangle$$

(2.3)

which depicts the construction of the ground state from an initial ‘vacuum state’, $|0\rangle$, in which all $\mathbf{k}$-states are empty, by repeated application of the operator $b_{\mathbf{k},S}^\dagger$. This operator represents the creation of the Cooper pair state, $(+\mathbf{k}+\mathbf{S} \uparrow,-\mathbf{k}+\mathbf{S} \downarrow)$, or equivalently the combined creation of two electrons, i.e.,

$$b_{\mathbf{k},S}^\dagger = c_{\mathbf{k}+\mathbf{S},\uparrow}^* c_{-\mathbf{k}+\mathbf{S},\downarrow}^*$$

where $c_{\mathbf{k},\uparrow}$ is the usual creation operator for an electron with momentum $\mathbf{k}$ and spin $\frac{1}{2}$. In (2.3) $v_{\mathbf{k}}^2$ is the probability that the pair state $(+\mathbf{k}+\mathbf{S} \uparrow,-\mathbf{k}+\mathbf{S} \downarrow)$ is occupied whilst $u_{\mathbf{k}}^2$ is the probability that it is vacant, so $v_{\mathbf{k}}^2 + u_{\mathbf{k}}^2 = 1$.

In (2.3) every electron pair has a total momentum $2\hbar \mathbf{S}$ and so a current is flowing. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discussion consider the situation where no current flows. To simplify the following discuss...
2.3.1 The BCS Ground State

To find an explicit expression for the ground state the occupation factors are calculated by a variational method. To do this a Hamiltonian must be formulated and then the expectation value of the system energy minimised by adjusting \( \nu_k^2 \) under the restriction that \( \nu_k^2 + u_k^2 = 1 \).

The full Hamiltonian for the interacting electrons is

\[
H = \sum_{k,\sigma} \varepsilon_k c_{k,\sigma}^* c_{k,\sigma} + \frac{1}{2} \sum_{k,\sigma_1,\sigma_2, q} c_{k+q,\sigma_1}^* c_{k-q,\sigma_2}^* V_{k,\sigma_1,\sigma_2} c_{k,\sigma_1} c_{k,\sigma_2} \tag{2.5}
\]

In the first term \( \varepsilon_k \) is the kinetic energy of state \( k \) relative to the Fermi energy (i.e., \( \hbar^2 k^2 / 2m - E_F \)) and \( c_{k,\sigma}^* c_{k,\sigma} \) is the number of electrons in state \( (k,\sigma) \). The second term represents the interaction between electrons being scattered from states \( k \) and \( l \) by exchange of a phonon* with wavevector \( q \) (figure 2.1). When the expectation value, \( W \), of the system’s energy is calculated from

\[
W = \langle \psi_G | H | \psi_G \rangle
\]

most of the terms in the interaction part of (2.5) drop out. This allows the full Hamiltonian to be replaced by an equivalent ‘reduced Hamiltonian’

\[
H_{\text{red}} = \sum_{k,\sigma} \varepsilon_k c_{k,\sigma}^* c_{k,\sigma} + \frac{1}{2} \sum_{k,k', \sigma} c_{k,\uparrow}^* c_{k',\downarrow} V_{k,k'} c_{k',\downarrow} c_{k,\uparrow}
\]

To find \( W \) the two terms may be considered separately. Because \( \nu_k^2 \) is the probability that states \( k \) and \( -k \) are occupied then the kinetic energy will be

\[
\langle \text{KE} \rangle = 2 \sum_k \varepsilon_k \nu_k^2 \tag{2.6}
\]

Similarly, the probability of scattering from \((+k, -k)\) to \((+k', -k')\) is the probability that initially \( k \) is occupied and \( k' \) is unoccupied (i.e., \( u_k \nu_k \)), and that after scattering \( k \) is unoccupied and \( k' \) is occupied (i.e., \( u_k \nu_{k'} \)). The potential energy term is therefore

\[
\langle \text{PE} \rangle = \sum_{k,k'} u_k \nu_k u_{k'} \nu_{k'} V_{kk'} \tag{2.7}
\]

Minimising the total energy with respect to \( \nu_k^2 \) leads to an expression for the Cooper pair occupation (Tinkham 1980, p.26),

* In more exotic models of superconductivity this is replaced by some other excitation or system of particles.
\[ v_k^2 = (1 - u_k^2) = \frac{1}{2} \left( 1 - \frac{\epsilon_k}{\xi_k} \right) \]  

(2.8)

where \( \xi_k \) (which turns out to be the excitation energy) is defined as

\[ \xi_k = (\epsilon_k^2 + \Delta_k^2) \]  

(2.9)

and \( \Delta_k \) (which will be identified as the energy gap) is defined as

\[ \Delta_k = -\sum_k u_k v_k V_{kk'} \]  

(2.10)

Substituting (2.8) and (2.9) into (2.10) gives the BCS self-consistent equation for \( \Delta_k \)

\[ \Delta_k = \sum_{k'} \frac{-\Delta_{k'}}{2 \sqrt{(\epsilon_{k'}^2 + \Delta_{k'}^2)}} V_{kk'} \]  

(2.11)

In principle, detailed knowledge of the interaction \( V_{kk'} \) would allow this to be solved to give \( \Delta_k \) for arbitrary wavevector. Two solutions may be found quite simply.

- If \( V_{kk'} = 0 \) then there is no interaction and \( \Delta_k = 0 \). In this case \( v_k \) is zero for \( \epsilon_k > 0 \) and \( v_k = 1 \) for \( \epsilon_k < 0 \). This just describes the Fermi-sea ground state of the non-interacting electron gas.

- If \( V_{kk'} \) is the highly simplified BCS interaction (2.1) then a solution exists if \( \Delta_k \) is isotropic and given by

\[ \Delta_k = \begin{cases} 
\Delta & \text{for } |\epsilon_k| < \hbar \omega_c \\
0 & \text{otherwise} 
\end{cases} \]  

(2.12)

In this case the Cooper pair occupancy \( v_k^2 \) is simply

\[ v_k^2 = \frac{1}{2} \left( 1 - \frac{\epsilon_k}{\sqrt{\epsilon_k^2 + \Delta^2}} \right) \]

Figure 2.3 shows that this differs from the occupancy of the ground state of a normal metal (i.e., the Fermi–Dirac distribution) only over a range of energies of width \( \approx \Delta \) around \( E_F \). Even at \( T = 0 \) some electrons gain kinetic energy and move above the Fermi level. In doing so the number of states available for scattering increases and the potential energy is lowered by more than the increase in kinetic energy. The
equilibrium distribution is reached when further movement of electron pairs into higher $k$-states no longer gives a net drop in the system’s energy. In a conventional superconductor $\Delta$ is only a few meV in magnitude and is much smaller than the Fermi energy (e.g., for Sn $\Delta = 0.57$ meV and $E_F = 10.03$ eV). Therefore only a very small change occurs in the energy distribution of electrons on entering the superconducting state.

The dramatic change that does occur when the superconductor is cooled below its transition temperature is the onset of long-range phase coherence. To see how this arises examine the potential energy (2.7). The only terms that contribute to this summation are those with both $v_k$ and $u_k$ not equal to zero. Therefore, the states that are involved in lowering of the system energy by condensation into Cooper pairs are those that lie within approximately $\Delta$ of the Fermi surface. If the wavefunction of a single pair is assumed to have the form of (2.2) then the finite range of contributing $k$-values, $\delta k$, centred around the Fermi wavevector results in $\Psi(r_1-r_2)$ having a spatial extent of about $1/\delta k$, i.e.,

$$\delta r \approx \frac{1}{\delta k} = \left(\frac{dE}{dk}\right)_{k_F} \frac{1}{\delta E} = \left(\frac{2\hbar^2}{m}\right)^{1/2} \frac{E_F^{1/2}}{\Delta}$$

A rigorous derivation by BCS of this coherence length, $\xi_0$, introduces an additional factor of $\pi$ into the denominator. The coherence length* is usually written in terms of

---

* To distinguish this characteristic length from the temperature-dependent coherence length, $\xi(T)$, introduced in Ginzburg–Landau theory $\xi_0$ is referred to as the intrinsic or Pippard coherence length.
the Fermi velocity,

$$\xi_0 = \frac{\hbar v_F}{\pi \Delta_0} \quad (2.13)$$

$\Delta_0$ is the value of $\Delta$ at $T = 0$. For typical values for the Fermi energy and $\Delta$ in a conventional superconductor the pair wavefunctions have a spatial extent of about 1 μm. This is much larger than the spacing between pairs so many wavefunctions overlap and the wavefunction of any one pair is not independent of the others. An ordered state is formed in which the total momentum of each pair is identical and the random phases of electrons in the normal state have been replaced by a common phase for each of the pair wavefunctions. In addition, the lowering of energy when any one pair forms depends on the number of other pairs in the system and condensation becomes a co-operative phenomenon.

### 2.3.2 Single-Particle Excitations from the Ground State

In the superconductor ground state all electrons are in pairs. Consider now a situation that differs from this in just one respect; the superconductor contains a single unpaired electron with wavevector $\mathbf{l}$ and spin $+\frac{1}{2}$. The new wavefunction may be written by considering it to be built up from the vacuum state, however, rather than creating a pair in the state $(+\mathbf{l}\uparrow, -\mathbf{l}\downarrow)$ by application of the pair creation operator $b^*_l$, a single electron is created by the operator $c^*_l$,

$$|\psi_l\rangle = c^*_l \prod_{k \neq l} (u_k + v_k b^*_k) |0\rangle$$

Since no operation occurs on the state $-\mathbf{l}$ it is left vacant as in the vacuum state. In summary:

- State $+\mathbf{l}\uparrow$ is now definitely occupied by an electron; the occupancy is 1.
- State $-\mathbf{l}\downarrow$ is definitely unoccupied by an electron; the occupancy is 0.
- All other pairs of states, $+\mathbf{k}\uparrow$ and $-\mathbf{k}\downarrow$ with $\mathbf{k} \neq \mathbf{l}$, are simultaneously occupied by electrons with probability $v^2_k$, or simultaneously vacant with probability $1 - v^2_k$. $v^2_k$ is just the occupancy of the Cooper pair state $(+\mathbf{k}\uparrow, -\mathbf{k}\downarrow)$.

To create the situation described above, a single electron may be injected into $+\mathbf{l}\uparrow$ from an external reservoir of electrons, as in a tunnelling experiment. Because the Cooper pair electrons are blocked from entering both $+\mathbf{l}\uparrow$ and $-\mathbf{l}\downarrow$ this automatically causes state $-\mathbf{l}\downarrow$ to be vacant, i.e., creates a hole in $-\mathbf{l}\downarrow$. An identical system is created if a hole is injected into $-\mathbf{l}\downarrow$. This immediately shows that the $I$–$V$ characteristics of a superconducting tunnel junction will be antisymmetric.
The effect of the unpaired particle is to prevent any of the remaining Cooper pairs from scattering into the state \((+l, -l)\), which, therefore, no longer participates in creating an attractive electron–electron interaction. Referring to (2.7) it is clear that the potential energy of the system will increase by

\[
\delta(\text{PE}) = -2v_l u_l \sum_k V_{lk} u_k v_k
\]

To find the total energy of \(|\psi_l\rangle\) relative to the ground state the change in kinetic energy must also be considered. This will be the kinetic energy of the single electron in state \(l\), less the expectation value of kinetic energy of the two usurped pair electrons in states \(+l\) and \(-l\),

\[
\delta(\text{KE}) = \epsilon_l - 2v_l^2 \epsilon_l
\]

The summation in the expression for \(\delta(\text{PE})\) is exactly the quantity defined in (2.10) as \(\Delta_l\), so substituting for \(u_l\) and \(v_l\) from (2.8) and using (2.9), the definition of \(\xi_l\), gives

\[
\delta(\text{PE} + \text{KE}) = \xi_l
\]

This is a real positive quantity, therefore, to create the first excited state \(|\psi_l\rangle\) from the ground state an excitation energy, \(\xi_l\), must be added to the system. This discussion has emphasised that an excitation should strictly be thought of as the existence of an unpaired electron with its associated hole of opposite spin and momentum. Moreover, because the unpaired electron reduces the scattering of every Cooper pair the excitation energy should be considered as distributed amongst all electrons in the system. However, it is convenient to think of this energy as belonging totally to a single-particle excitation called a quasiparticle. For an isotropic BCS superconductor the quasiparticle energy \(\xi_k\) is defined by (2.9) with \(\Delta_k = \Delta\),

\[
\xi_k^2 = (\epsilon_k^2 + \Delta^2)
\]

This is plotted in figure 2.4 as a function of \(\epsilon_k\) - the energy of a free electron with the same wavevector.

Before considering the nature of the superconductor’s quasiparticles, excitations in the normal state will be discussed. These are simply electrons if \(k > k_F\) or holes if \(k < k_F\), and the excitation energy is just the absolute value of the energy of the Bloch state associated with the excitation. To clarify this description, consider a situation in which we start with the Fermi sea at \(T = 0\). If an electron with \(k_l < k_F\) is moved above the Fermi level to state \(k_2 > k_F\) then the total energy required is \((E_2 - E_1)\), or equivalently \((E_2 - E_F) + (E_F - E_1)\). This second form suggests that this process can be
considered to be the creation of an electron excitation above the Fermi level with energy \((E_2 - E_F) = \varepsilon_2\), plus the creation of a hole excitation below the Fermi level with energy \((E_F - E_1) = |\varepsilon_1|\). The excitation energy is

\[
\xi_{k,n} = |\varepsilon_k| = |E_k - E_F|
\]

with a discontinuous change in the nature of the excitations at the Fermi level from holes to electrons; the curve is distinctly divided into a ‘hole branch’ and an ‘electron branch’. Because \(\xi\) goes to zero at the Fermi level an excitation may be created from the ground state by the addition of any energy, however small. In contrast, there are no excitations in the superconductor with energy smaller than \(\Delta\) so an energy gap exists at the Fermi level in the excitation energy spectrum; an excitation can be created only if the added energy exceeds this minimum amount.

In a normal metal the excitations are simply electrons or holes and so the density of excitations is just the density of Bloch states, \(N_n(E)\). When the material becomes superconducting the opening of a gap at the Fermi energy dramatically changes the distribution of excitation states, ‘pushing’ those that were previously inside the gap into a narrow region just outside. The density of superconducting excitation states is, therefore,

\[
N_s(\xi_k) = \frac{dn}{d\xi_k} = \frac{dn}{d\xi} \frac{d\xi_k}{d\xi} = N_n(\varepsilon_k) \frac{d\xi_k}{d\xi}
\]
The derivative may be calculated from (2.9) to give

\[ N_s(\xi_k) = N_n(\xi_k) n_s(\xi_k) \]

where

\[ n_s(\xi_k) = \begin{cases} \frac{\xi_k}{\sqrt{\xi_k^2 - \Delta_k^2}} & \xi_k > \Delta_k \\ 0 & \xi_k < \Delta_k \end{cases} \tag{2.14} \]

To clarify the nature of the quasiparticle excitations consider the creation of a second excited state. One would hope to be able to describe this in a similar way to the first excitation with a wavefunction of the form

\[ |\psi_{l,m}\rangle = c_m^* c_l^* \prod_{k \neq l,m} (u_k + v_k b_k^*) |0\rangle \]

However, de Gennes (1966, p.115) shows that this is not necessarily orthogonal to the ground state \( |\psi_G\rangle \) so this is not a valid description of the second excited state. BCS originally overcame this problem by writing

\[ |\psi_{l,m}\rangle = c_m^* c_l^* \prod_{k \neq l,m} (u_k + v_k b_k^*) |0\rangle + \lambda |\psi_G\rangle \]

and choosing the factor \( \lambda \) in such a way that \( \langle \psi_G | \psi_{l,m} \rangle = 0 \).

A more sophisticated method of representing excitations from the BCS ground state was proposed in 1958 by Bogoliubov who introduced the quasiparticle excitation creation and annihilation operators, \( \gamma^* \) and \( \gamma \), respectively. They possess the property that all states obtained by applying these operators successively to the ground state will be both orthogonal to that ground state and mutually orthonormal; hence, they are suited to describing multiple excitations. To behave in this way \( \gamma^* \) and \( \gamma \) must satisfy the anticommutation relations for fermions, i.e.,

\[ [\gamma_s, \gamma_{s'}]^* = \gamma_s, \gamma_{s'} + \gamma_{s'}^*, \gamma_s = \delta_{s,s'} \]
\[ [\gamma_s, \gamma_{s'}]^* = [\gamma_{s'}, \gamma_s]^* = 0 \tag{2.15} \]

where \( s \) represents both the spin and momentum of the state. In addition, when acting on the ground state (which has no excitations) the annihilation operator must give zero, i.e.,

\[ \gamma |\psi_G\rangle = 0 \]
Bogoliubov showed that suitable operators could be written as a linear combination of the electron creation and annihilation operators

\[
\gamma_{k\uparrow} = u_k c_{k\uparrow}^* - v_k c_{-k\downarrow}
\]
\[
\gamma_{k\downarrow} = u_k c_{-k\downarrow}^* + v_k c_{k\uparrow}
\]
\[
\gamma_{k\uparrow} = u_k c_{-k\uparrow} - v_k c_{-k\downarrow}
\]
\[
\gamma_{k\downarrow} = u_k c_{-k\downarrow} + v_k c_{k\uparrow}
\]

Examination of these operators reveals some apparently peculiar properties. When a quasiparticle is created well above the Fermi level \((k \gg k_F)\) then from (2.8) \(v_k \approx 0\) and \(u_k \approx 1\). This means that \(\gamma_{k\uparrow} \approx c_{k\uparrow}^*\) and \(\gamma_{k\downarrow} \approx c_{-k\downarrow}^*\), so the operator creates an excitation that is essentially an electron. Similarly for \(k \ll k_F\) the excitation creation operator reduces to an electron annihilation operator and so may be considered to produce an excitation that is strongly hole like. At intermediate wavevectors, with energies within approximately \(\Delta\) of the Fermi level, both \(u_k\) and \(v_k\) are significantly different from both zero and one so the excitation seems to be a linear combination of a hole and an electron. This behaviour was suggested by the excitation-energy curve in figure 2.4, which continuously links the hole-like and electron-like branches with no discontinuity at the Fermi level as in the normal state.

When a quasiparticle creation operator acts on a system the expectation value of the number of particles changes by \(u_k^2 - v_k^2\) (Tinkham 1980, p.39). This ranges from \(-1\) well below the Fermi energy to \(+1\) well above the Fermi energy, corresponding to the creation of a hole and an electron respectively. Again the behaviour near the Fermi surface seems strange because the particle number changes by some fractional amount between \(-1\) and \(+1\). What is meant, for example, by a quasiparticle that is \(-0.4\) electrons? To answer this question it must be remembered that \(v_k^2\) is an expectation value and so is the average occupancy of the Cooper pair state \((+k,-k)\) over a large number of identical systems. In addition, the concept of a single particle excitation was introduced as a convenient device for describing excitations; however, it was emphasised that the excitation should strictly be thought of as an unpaired electron with an associated hole of opposite spin and momentum. Consider the creation of an excitation by injecting an electron just below the Fermi surface with wavevector \(k\), such that \(v_k^2 = +0.7\). Before addition of this unpaired electron, states \(+k\) and \(-k\) were both occupied by electrons with probability \(+0.7\). Afterwards excitation \(+k\) is definitely occupied so the change in its occupation is \(+0.3\). In addition, \(-k\) is now definitely vacant so its occupation has changed by \(-0.7\). The overall change in the average occupation of \((+k,-k)\) is, therefore, \(-0.4\). This is interpreted in the single-
particle excitation picture as a quasiparticle with 30% electron character and 70% hole character - or an overall loss of 0.4 electrons. Clearly the mixed electron-like and hole-like character is an artefact of trying to describe the excitation as a single particle. So long as the true meaning of this behaviour is borne in mind the quasiparticle description remains a useful tool. A popular misconception that arises from careless use of the semiconductor model (see section 3.2) is that the excitation spectrum in figure 2.4 represents electrons for \( k > k_F \), holes for \( k < k_F \) and some combination for \( k \approx k_F \). The explicit example given above has shown that electrons (and holes) may be injected into all states both above and below \( k_F \), thus figure 2.4 should be thought of as the excitation spectrum for both holes and electrons.

When an excitation is created in a tunnelling experiment electrons or holes are injected into the superconductor from an external electrode so quasiparticles are created singly. However, if the superconductor is isolated and the excitations are created by breaking a Cooper pair, such as in an infra-red absorption experiment, then excitations must be created in pairs to conserve the total number of electrons. From (2.8), \( v^2(-\epsilon) = u^2(\epsilon) \), so particle conservation is satisfied if the two excitations are symmetrically located with respect to the Fermi energy. The minimum energy required to excite the superconductor in such an experiment (and so give rise to absorption of the incident photons) is \( 2\Delta \). This quantity, sometimes referred to as the ‘spectroscopic gap’, is much larger than the energy required simply to overcome the weak attractive interaction between the pair electrons. This is because creation of the two unpaired electrons lowers the number of scattering states available to every remaining Cooper pair and so raises the potential energy of the whole assembly.

### 2.3.3 Finite Temperatures

The creation operators for quasiparticles obey the fermion anticommutation relations (2.15) and so describe a set of particles that are thermally excited with probability given by the Fermi–Dirac function

\[
 f_k = f(\xi_k) = \frac{1}{1 + \exp(\xi_k / k_B T)}
\]

At finite temperatures the presence of these single-particle excitations will change the distribution of pair states and therefore the energy of the system. For example, the kinetic energy will be decreased by the loss of some pair states but is increased by the energy of the excitations

\[
 \langle KE \rangle = 2 \sum_k \epsilon_k v^2_k (1 - 2f_k) + 2 \sum_k \epsilon_k f_k
\]
Comparison with (2.6), the kinetic energy at \( T = 0 \), shows that the first summation is over pair states and that each term is reduced by the probability that \((+\mathbf{k},-\mathbf{k})\) is ‘blocked’ by a quasiparticle. The second summation is over the kinetic energies of the excitations weighted by the occupancy of each state. Quasiparticles will also prevent some pair states from participating in scattering and so will reduce the magnitude of the potential energy term

\[
\langle \text{PE} \rangle = \sum_{k,k'} u_k^* v_k u_{k'} v_{k'} (1 - 2 f_k) (1 - 2 f_{k'}) V_{kk'}.
\]

Again this differs from (2.7), the \( T = 0 \) expression, due to the occupancy factors, which are the averaged probability that pair state \((+\mathbf{k},-\mathbf{k})\) is available to scattering (or equivalently that neither \(+\mathbf{k}\) nor \(-\mathbf{k}\) are occupied by a quasiparticle).

The wavefunction may still be written as in (2.4) but the \( v_k^2 \) factors will differ from those in the ground state. As before, the equilibrium pair occupancy may be found by minimising the system’s energy with respect to \( v_k^2 \) (van Duzer and Turner 1981, sec.6.04). When this is done (de Gennes 1966, p.121) the form of the occupation factors and excitation energy is unchanged from (2.8) and (2.9). However, the definition of the gap parameter is changed to

\[
\Delta_k = -\sum_{k'} u_k v_{k'} (1 - f_{k'}) V_{kk'}.
\]

and the BCS self-consistent equation is

\[
\Delta_k = \sum_{k'} (1 - 2 f_{k'}) \frac{-\Delta_{k'}}{2 \sqrt{\epsilon_{k'}^2 + \Delta_{k'}^2}} V_{kk'}.
\] (2.17)

In principle the energy gap could be calculated for any \( \mathbf{k} \) and at all temperatures; however, this would be an extremely complicated procedure and would require detailed information about the interaction \( V_{kk'} \). The simplified BCS interaction (2.1) can be used if the energy gap is much smaller than the cut-off energy in the model interaction (i.e., \( \Delta \ll h\omega_D \)). This leads to a solution with an isotropic energy gap satisfying (2.12). \( \Delta_k \) then cancels from both sides of (2.17) and the summation can be converted to an integral over energy

\[
1 = V \int_0^{h\omega_D} \left( 1 - 2 f \left( \sqrt{\epsilon^2 + \Delta(T)^2} \right) \right) \frac{N(\epsilon) d\epsilon}{\sqrt{\epsilon^2 + \Delta(T)^2}}.
\] (2.18)

If the density of single electron states, \( N(\epsilon) \), is slowly varying over the range of integration it may be assumed to be constant, i.e., \( N(\epsilon) \approx N(0) \). This simplification is
frequently made when calculating the gap parameter and will be used in the calculations below. It should be borne in mind that if the density of states contains van Hove singularities in this range, as may be the case with high-temperature superconductors (section 4.2.1), $N(\varepsilon)$ will be rapidly changing and this approximation will not be valid.

Expression (2.18) implicitly contains the temperature dependence of the energy gap of a BCS superconductor; however, an analytical solution cannot generally be found and $\Delta(T)$ must be calculated numerically (Mühlschlegel 1959). As temperature is increased the energy gap decreases monotonically from its value at $T = 0$, denoted by $\Delta_0$ or $\Delta(0)$, and vanishes at the critical temperature, $T_c$ (figure 2.5). Above $T_c$ the BCS self-consistent equation is satisfied only for $\Delta = 0$, so the pair states are unstable and the material remains in its normal state.

The gap expression (2.18) may be solved analytically in two special cases. At $T = 0$ it may be integrated exactly to give

$$\frac{1}{N(0)V} = \text{arcsinh} \left( \frac{\hbar \omega_d}{\Delta_0} \right)$$

which for $V > 0$ has a real solution for $\Delta_0$. This expression was derived using the BCS model interaction and so is valid only if $\Delta \ll \hbar \omega_d$. From (2.19) this condition is equivalent to requiring $N(0)V \ll 1$, hence the simplified BCS interaction is valid only if the electron–electron interaction is weak. In this weak-coupling limit the gap

![Fig. 2.5. Temperature dependence of the energy gap according to BCS weak-coupling theory. Also shown (dotted curve) is $\Delta(T)/\Delta_0 = 1.8(1 - T/T_c)^{1/2}$, an approximation valid near $T_c$. Data from table by Mühlschlegel (1959).](image-url)
parameter may be approximated by

$$\Delta_0 \approx 2\hbar \omega_0 e^{-1/N(0)V} \quad (2.20)$$

The second simple solution to (2.18) occurs at the transition temperature where

$$\Delta(T_c) = 0.$$ On inserting $$\Delta = 0$$ and changing the integration variable (2.18) reduces to

$$\frac{1}{N(0)V} = \int_0^{\hbar \omega_0/2k_B T_c} \frac{\tanh(x)}{x} dx$$

which may be solved to give

$$k_B T_c = \left( \frac{2\gamma}{\pi} \right) \hbar \omega_0 e^{-1/N(0)V} \quad (2.21)$$

$$\gamma$$ is Euler’s constant so the prefactor is approximately 1.13. This expression is useful in several ways. Firstly, if the Debye temperature is known from measurements of the specific heat* then the coupling parameter, $$N(0)V$$, of the superconductor may be estimated from (2.21). For most conventional superconductors this value turns out to be smaller than 0.3 so the weak-coupling limit is justified. Detailed theoretical considerations suggest that a coupling parameter larger than 0.5 is unlikely to occur without the lattice becoming unstable. In addition, the Debye temperature of most materials is in the range 100–300 K so the highest transition temperature that can be expected from a phonon-mediated interaction is about 45 K. Indeed, until the discovery of high-temperature superconductivity in 1987 the highest known transition temperature was 23.2 K in Nb₃ Ge and superconductivity was generally considered to be limited to very low temperatures.

### 2.3.4 The Gap Ratio and Strong-Coupling Theory

Combining (2.20) and (2.21) to eliminate the Debye frequency gives the gap ratio or reduced gap

$$\frac{2\Delta_0}{k_B T_c} = 3.52 \quad (2.22)$$

---

* At low temperatures the heat capacity is given by $$C_v = aT + bT^3$$, where the first term is the electronic heat capacity with $$a = \pi^2 k_B^2 N(0)/3 J K^{-2}$$ and the second is the lattice heat capacity with $$b = 234nk_B/\theta_D^2 J K^{-4}$$, where $$n$$ is the number of atoms in the sample (Kittel 1986, p.108 and p.139). Plotting $$C_v/T$$ allows the Debye temperature to be found from the gradient and the density of states at the Fermi level to be found from the intercept.
Figure 2.6 plots this parameter using experimentally determined values. For most superconductors this ratio is within 30% of the weak-coupling value and generally increases with the strength of the coupling, as predicted by more elaborate models of superconductivity.

The gap ratio is often regarded as a ‘benchmark’ for testing new theoretical models. Combescot (1988) repeated the weak coupling calculation of $2\Delta_0/k_B T_c$ but allowed the density of states in (2.18) to vary rather than be held fixed. He showed that:

- If $N(\varepsilon)$ is sharply peaked at the Fermi level with width smaller than $\Delta_0$ then $2\Delta_0/k_B T_c$ is greater than the BCS value.
- If this peak is wider than $\Delta_0$ the gap ratio is close to its BCS value.
- If the density of states is peaked at an energy that differs from $E_F$ by more than $\Delta_0$ then $2\Delta_0/k_B T_c$ is smaller than the BCS value.
- Whatever the exact form of the density of states, if the superconductor is weakly coupled then the ratio satisfies $2\Delta_0/k_B T_c \leq 4.0$. For $T > 0$ this is generalised to $\Delta(T) \leq 1.13\Delta_{\text{BCS}}(T)$, where $\Delta_{\text{BCS}}(T)$ is the usual result calculated from (2.18) with a constant density of states. In addition, if the energy gap is anisotropic then the inequalities are satisfied by the minimum gap.

These results have been spelled out in detail for two reasons. Firstly, they provide a simple test that provides information about the size of the cut-off and hence the possible origin of the electron–electron interaction; if the inequalities are not satisfied then $\Delta \ll h\omega_c$ and the material must be strongly coupled. Secondly, this analysis demonstrates that many approximations have been made in deriving the weak-coupling model and that lifting any one of these restrictions may lead to significantly different results. In fact Combescot’s result is remarkable because there is relatively little freedom in the gap ratio even when $N(\varepsilon)$ is allowed to take any form.

An obvious extension to BCS theory is to consider strong electron–phonon coupling. As a first step Thouless (1960) considered the same interaction as used in the weak-coupling model (2.1) but allowed $N(0)V > 1$. He obtained a gap ratio $2\Delta_0/k_B T_c = 4.0$, confirming that stronger coupling increases the gap ratio, with only minor modifications to the tunnelling characteristics.

BCS theory is based on the pairing of electrons due to an attractive interaction that arises when they are scattered by exchange of a virtual phonon. One would expect the strength of $V_{kk'}$ to depend on the density of phonons available for scattering an electron from $k$ to $k'$. In the weak-coupling model the energy range of states that
Fig. 2.6. Experimentally measured gap ratio plotted as function of ‘coupling strength’. Weak-coupled superconductors have $T_c \ll \theta_D$. Data from Kittel (1986).

participate in pair formation is assumed to be narrow compared to the phonon energies, so the phonon density of states and hence the matrix element $V_{kk}$ are independent of energy. This allows the energy gap to be calculated using the highly simplified BCS interaction (2.1). No account is taken of the spatial extent of the interaction and both the quasiparticles and pairs are considered to have infinite lifetimes. Consequently they have well defined energies and the interaction occurs instantaneously.

In contrast, strong-coupling theories use a local interaction that is ‘retarded’ in time to account for the delay in development of overscreening. This leads to a description of the superconductor by an energy-dependent complex gap, $\Delta(E) = \Delta_1(E) + i\Delta_2(E)$, the imaginary part of which is inversely proportional to the quasiparticle’s finite lifetime. The excitations decay by creation or absorption of real phonons - a process which occurs most rapidly if the quasiparticle can readily couple to a large number of phonon states. Therefore, a peak occurs in $\Delta_2(E)$ when the electron–phonon spectral function, $\alpha^2 F(E)$, has a maximum. $F(E)$ is the phonon density of states and $\alpha^2(E)$ is the electron–phonon matrix element, which measures the strength of coupling.

At very low temperatures the conductance of a superconducting tunnel junction is closely related (see section 3.2) to the density of quasiparticle excitations (2.14). In the case of strong coupling the energy gap is complex and the normalised tunnel conductance is given by

$$\sigma_{\text{sc}}(E) = \text{Re} \left\{ \frac{|E|}{\sqrt{E^2 - (\Delta_1(E) + i\Delta_2(E))^2}} \right\}$$

(2.23)
Differences between the tunnelling conductance of a strong-coupled and a weak-coupled superconductor are rather small. In both cases the conductance is low when $|eV| < \Delta_0$ and rises rapidly to a maximum at this value, where $V$ is the junction bias. In a weak-coupled superconductor $\Delta_0$ is simply the energy gap parameter, which is a constant, whilst in a strong-coupled superconductor it is the energy satisfying

$$\text{Re}[\Delta(\Delta_0)] = \Delta_0$$

(2.24)

For biases larger than the gap edge the conductance of a strong-coupled superconductor tunnel junction deviates slightly from the predictions of weak-coupled BCS theory. This is caused by the energy dependence of $\Delta(E)$ which in turn is caused by the energy dependence of the phonon spectrum, $\alpha^2(E)F(E)$. These deviations are small (<10%) and are most easily seen in the second-derivative tunnelling spectrum where a phonon peak is manifested as a dip in $d^2I/dV^2$. This phonon structure was first observed by Giaever et al. (1962) in lead.

Eliashberg’s (1960) strong-coupling theory contains a set of coupled integral equations equivalent to the BCS self-consistent equation for $\Delta_k$ (2.11). The superconductor’s material properties are introduced through two parameters; the $\alpha^2(E)F(E)$ function, discussed above, and a constant, $\mu^*$, called the Coulomb pseudopotential. This represents the screened residual Coulomb repulsion between quasiparticles. A further useful parameter of the material is the dimensionless coupling strength

$$\lambda = 2\int_0^\infty \frac{\alpha^2(\omega)F(\omega)}{\omega} d\omega$$

(2.25)

which is analogous to $N(0)V$ in the weak-coupling BCS theory. With these input parameters the energy-dependent complex gap may be found by numerically solving the Eliashberg equations. This may then be used in (2.23) to calculate the expected tunnelling characteristics. McMillan and Rowell developed their ‘inversion technique’ for reversing this procedure to extract the phonon spectrum from the BCS-reduced tunnelling conductance, $\sigma(V)/\sigma_{\text{BCS}}(V) - 1$. Because the Eliashberg equations are non-linear in their parameters this must be done numerically using an iterative fitting procedure. $\alpha^2(E)F(E)$ and $\mu^*$ are adjusted to obtain the best fit between the calculated and experimentally determined conductance over the range of phonon energies (i.e., those with $\alpha^2(E)F(E) \neq 0$). The validity of the inversion method and of the Eliashberg model are confirmed by the good agreement obtained between the model conductance and experimental results even outside this fitting range, where the conductance still deviates from the BCS result but there are no phonons. Once the
best-fit $\alpha^2(E)F(E)$ and $\mu^*$ parameters are known they may be fed back into the Eliashberg equations to obtain the superconductor’s gap function, $\Delta(E)$. Equation (2.24) may then be used to calculate the Eliashberg theory prediction of zero-temperature gap and hence $2\Delta_0/k_BT_c$.

This analysis was carried out by Mitrovic et al. (1984) for many superconductors and the calculated gap ratio was found to agree with experimental results to within a few percent (figure 2.7). Mitrovic et al. then fitted the calculated gap ratios to an equation of the type introduced by Geilikman and Kresin (1972),

$$\frac{2\Delta_0}{k_BT_c} = 3.53 \left[ 1 + \alpha \left( \frac{T_c}{\omega_{in}} \right)^2 \ln \left( \frac{\omega_{in}}{\beta T_c} \right) \right]$$

obtaining best-fit parameters $\alpha = 12.5$ and $\beta = 2.0$. In this expression $\omega_{in}$ is the Allen–Dynes average phonon energy in kelvin, defined as

$$\omega_{in} = \exp \left( \frac{2}{\lambda} \int_0^{\infty} \frac{\alpha^2(\omega)F(\omega)}{\omega} \ln(\omega) \, d\omega \right)$$

where $\lambda$ is defined by (2.25). This phenomenological equation gives a good description of the general trend in the gap ratio with coupling strength and is accurate to about 10% (figure 2.7). Similar curves of $\Delta C/\lambda T_c$ and $\lambda T_c^2/H_c(0)^2$ against $T_c/\omega_{in}$

![Figure 2.7](image_url)

**Fig. 2.7.** Gap ratio vs $T_c/\omega_{in}$ for strong-coupled superconductors. Circles are theoretical results from numerical solutions of the Eliashberg equations. Bars show the error from the measured gap ratio and the curve is the phenomenological model found by fitting to the numerical results. Discrepancies between the theoretical and measured gap ratios for Nb$_3$Sn, a-Pb$_{0.75}$Bi$_{0.5}$ and a-Ga were attributed to experimental difficulties in correctly measuring $\Delta_0$. Data from Mitrovic et al. (1984).
have been constructed from measured properties. These may be used to assess whether a new material is strongly or weakly coupled and so offer valuable guidance to theoretical work.
Electron tunnelling is a phenomenon of quantum mechanics in which a particle, as a consequence of its wave-like character, may pass through a potential barrier that would confine a classical particle of the same energy. The probability of transmission depends on the barrier shape and properties of the electrode and barrier materials, but in all cases it falls rapidly and approximately exponentially with electrode separation. At low voltages the current density through a rectangular potential barrier between planar electrodes is

$$J = \left( \frac{3e^2}{8\pi^2\hbar} \right) V \frac{\alpha}{w} \exp(-2\alpha w)$$  \hspace{1cm} (3.1)

where $w$ and $\phi$ are the barrier width and height, $V$ is the bias, $m$ is the mass of the tunnelling particle and $\alpha = \sqrt{2m\phi}/\hbar$ (Simmons 1963). With typical barrier heights ($\approx 5$ eV) significant current is observed only if the barrier has a width of a few nanometres.

The transfer Hamiltonian description of tunnelling (Wolf 1985, p.23) may be used when the probability of tunnelling is small. The electrodes are then only weakly interacting* and may be regarded as two nearly independent subsystems with wavefunctions $\psi_l$ and $\psi_r$ that satisfy

$$H_l |\psi_l\rangle = E_l |\psi_l\rangle \hspace{1cm} \text{in the left electrode and}$$

$$H_r |\psi_r\rangle = E_r |\psi_r\rangle \hspace{1cm} \text{in the right electrode.}$$

Tunnelling is modelled as a weak coupling of the wavefunctions (i.e., little perturbation of $\psi_l$ and $\psi_r$ from the non-tunnelling state) with a Hamiltonian of the form

$$H_T = \sum_{\sigma kq} \left( T_{kq} c_{k\sigma}^* c_{q\sigma} + T_{qk} c_{q\sigma}^* c_{k\sigma} \right)$$  \hspace{1cm} (3.2)

where the sum is over all wavevectors in the left and right electrodes, $q$ and $k$, and over both spin states. The first term contains an electron annihilation operator, $c_{q\sigma}$,
and a creation operator, $c_{k\sigma}^*$, and so describes tunnelling from left to right with no change in spin. The second term describes tunnelling in the reverse direction. Details of the barrier transmission probabilities are included in the tunnelling matrix elements, $T$. The probability per unit time of transmission from an initial discrete state $\psi_i$ into a final state $\psi_r$ is proportional to the square of the matrix element and is given by Fermi’s golden rule (Merzbacher 1970, p.479)

$$P_{i \rightarrow r} = \frac{2\pi}{\hbar} |M_{i r}|^2 \tilde{n}_r(E_r) \delta(E_r - E_i)$$  \hspace{1cm} (3.3)

Here, $\tilde{n}_r(E)$ is the density of states in the right electrode that may contribute to tunnelling from left to right. Exponential behaviour in (3.1) occurs because the tunnelling matrix element, $M_{i r} = \langle \psi_r | H_T | \psi_i \rangle$, depends on overlap between $\psi_i$ and $\psi_r$, which decay exponentially outside their electrodes.

The current from the single state $\psi_i$ into all energy-conserving states in the right electrode is given by multiplying (3.3) by the electronic charge. To calculate the total current density due to electrons tunnelling from left to right we sum over all states from which tunnelling may take place, thus in integral form

$$J_{i \rightarrow r} = \frac{-2\pi e}{\hbar} \left| M_{i r} \right|^2 \tilde{n}_r(E) n_i(E) dE$$  \hspace{1cm} (3.4)

where $n_i(E)$ is the density of states in the left electrode that may contribute to the tunnel current from left to right. A similar expression is found for tunnelling from the right electrode to the left,

$$J_{r \rightarrow i} = \frac{-2\pi e}{\hbar} \left| M_{r i} \right|^2 n_i(E) \tilde{n}_r(E) dE$$  \hspace{1cm} (3.5)

and the net current is then

$$J = J_{r \rightarrow i} - J_{i \rightarrow r}$$  \hspace{1cm} (3.6)

### 3.1 MIM Junctions

Figure 3.1 depicts the tunnelling of a single electron between two normal electrodes as the creation of a hole excitation in one electrode and an electron excitation in the other. If no energy loss occurs the tunnelling event is elastic and the sum of the excitation energies must equal the energy provided by the bias.

From (3.4) the current depends on the density of excitations that can participate in
tunnelling. This is more clearly seen in figure 3.2, which plots the density of excitation states as a function of excitation energy. In this type of diagram the excitation energy is positive in both directions from $\xi = 0$. Below this level the excitations are holes and the curve plots the density of hole excitations $N_h(\xi)$, whilst in the other direction the excitations are electrons and $N_e(\xi)$ is plotted. In a normal-metal the density of excitations is simply the density of Bloch states.

With no applied bias the system is in equilibrium and the Fermi levels in the two electrodes are identical. Applying a bias to the positive electrode lowers its Fermi energy (and hence its zero excitation level) by the energy of the external bias, $eV$, resulting in a rigid shift of the density-of-states curve towards lower energies. Figure 3.2 shows a single tunnelling event in which an electron moves out of the left electrode, creating a hole excitation of energy $\xi_l$, and is injected into the right electrode as an electron excitation of energy $\xi_r$. With this type of diagram the energy conservation condition, $\xi_l + \xi_r = eV$, is satisfied if the hole excitation on the left and the electron excitation on the right are at the same height, thus elastic tunnelling is represented by a horizontal line connecting the two electrodes. This simple depiction of elastic tunnelling allows the functions $n_l(E)$ and $\tilde{n}_l(E)$ that appear in the calculation of the partial tunnel current $J_{1\rightarrow r}$ to be evaluated easily. This will be done in two ways. The first considers the tunnelling process to consist of transfer of a Bloch state from one electrode to the other. The second method emphasises the excitation description of tunnelling in which tunnelling is regarded as the creation of excitations in both electrodes. Although the latter method is not a natural way of describing tunnelling in an MIM junction, it is more general and may also be used to model tunnelling between superconducting electrodes.
3.1.1 Transfer of Bloch States

In this section tunnelling will be described as the transfer of an electron across the barrier. Consequently it is useful to measure all energies with respect to the Fermi energy in the positive electrode, i.e., by $\varepsilon = E - E_F$. In the derivation of (3.4) $n_l(\varepsilon,T)$ was defined as the density of states in the left electrode that contribute to the tunnelling of electrons with energy $\varepsilon$ from the left electrode to the right. Since a tunnelling electron must come from a filled Bloch state on the left $n_l(\varepsilon,T)$ is simply the density of these states and is given by

$$n_l(\varepsilon,T) = N_{e,l}(\varepsilon - eV) f(\varepsilon - eV,T)$$

where $N_{e,l}$ is the density of electron states in the left electrode and $f(\varepsilon,T)$ is the Fermi function

$$f(\varepsilon,T) = \frac{1}{1 + \exp(\varepsilon / k_BT)}$$

$n_l(\varepsilon,T)$ is represented in figure 3.2 by the shaded part of the curve for the left electrode.

Similarly $\bar{n}_l(\varepsilon,T)$ is the density of states in the right electrode that contribute to tunnelling from left to the right and so in a normal-metal is simply the density of unfilled electron states. i.e.,

$$\bar{n}_l(\varepsilon,T) = N_{e,r}(\varepsilon)[1 - f(\varepsilon,T)]$$

represented in figure 3.2 by the unshaded portion of the curve for the right electrode.

Elastic tunnelling from left to right may be envisaged as horizontal movement of filled (shaded) states on the left to unfilled (unshaded) states on the right. All such contributions to the current may be added together by integrating over energy to include all possible horizontal transitions, thus (3.4) may be written

$$J_{l\rightarrow r}(V,T) = -\frac{2\pi e}{\hbar} \int_{-E_F}^{\infty} |M_{rl}|^2 N_{e,r}(\varepsilon - eV) N_{e,l}(\varepsilon) f(\varepsilon - eV,T)[1 - f(\varepsilon,T)] d\varepsilon$$  \hspace{1cm} (3.7)

The arguments used above may be repeated to find the functions $n_r(\varepsilon,T)$ and $\bar{n}_r(\varepsilon,T)$ that describe tunnelling from right to left and (3.5) may then be evaluated,

$$J_{r\rightarrow l}(V,T) = -\frac{2\pi e}{\hbar} \int_{-E_F}^{\infty} |M_{lr}|^2 N_{e,r}(\varepsilon) N_{e,l}(\varepsilon - eV) f(\varepsilon,T)[1 - f(\varepsilon - eV,T)] d\varepsilon$$

From (3.6) the total tunnel current flowing from the positive to negative electrode is
\[ I(V, T) = \frac{2\pi eA}{h} \int_{-E_F}^{\infty} |M|^2 N_{e,\ell}(\epsilon) N_{e,\ell}(\epsilon - eV)(f(\epsilon - eV, T) - f(\epsilon, T)) \, d\epsilon \] (3.8)

where the junction area is \( A \) and symmetry of the matrix element has been used to set \( |M_{lr}|^2 = |M_{rl}|^2 = |M|^2 \). The tunnelling probability may be calculated exactly only for simple rectangular barriers (Merzbacher 1970, p.93) and more realistic barriers (e.g., trapezoidal) must be treated by approximate methods (Simmons 1963) or numerically (Brinkman et al. 1970). If the barrier is very thin and high the tunnelling probability is approximately independent of energy and applied bias so \( |M|^2 \) may be removed from the integral. For small biases the density of states may also be considered constant. When \( \epsilon < -E_F \), \( f(\epsilon) \approx f(\epsilon - eV) \) so the lower limit in (3.8) may be changed to \(-\infty\) with negligible error. The remaining simple integral gives

\[ I = G_n V \]

where

\[ G_n = \frac{2\pi e^2 A}{h} |M|^2 N_{e,\ell}(0) N_{e,\ell}(0) \] (3.9)

**Fig. 3.2.** Electron tunnelling at \( T > 0 \) described by a density-of-excitation-states diagram. Shading represents filled electron states. The horizontal arrow represents the same tunnelling event as shown in figure 3.1.
This result was found by assuming that $N_{e,1}(E)$, $N_{e,r}(E)$ and $|M|^2$ are slowly varying over the range of energies for which $f(\varepsilon - eV) - f(\varepsilon)$ is significant, and so may be treated as constant. Harrison (1961) used a WKB approximation to evaluate the tunnelling matrix elements in (3.2) and found that

$$|T_{ik}|^2 = \frac{C}{N_{e,i}(E)N_{e,r}(E)}$$

where $C$ is a constant, so the product $|M|^2 N_{e,i}(E)N_{e,r}(E)$ is always independent of energy and may be removed from the integration in (3.8) even for large applied biases. This suggests that electron tunnelling cannot be used to measure the density of states in normal metals (Solymar 1972, p.25).

Although this is true in general, features due to the band structure and density of states may appear in the tunnelling characteristics under certain conditions. If the wavelength of tunnelling electrons is comparable with the width of the tunnel barrier then the barrier is no longer slowly varying and the WKB approximation is invalid. It has been suggested (Wolf 1985, sec.2.4.2) that $C$ will then be $k$ dependent and may strongly weight the transmission probability to favour tunnelling with small transverse component of wavevector. The tunnel current then depends on $N(E,k_x)$ at the electrode’s surface, where $N(E,k_x)$ is the local density of states with total energy $E$ and a component $k_x$ of wavevector in the direction of tunnelling. This situation probably arises in the oxide superconductors where the Fermi wavelength is about 2 nm (Wolf and Kresin 1991) - approximately equal to a typical barrier width. Features originating from the normal density of states have been observed in studies of conventional superconductors with low Fermi energies (e.g., Esaki and Stiles 1965).

The density of states may also appear in the conductance–voltage characteristics if tunnelling electrons enter states with discrete values of wavevector $k_x$. These states are bound in the $x$-direction (zero group velocity) so the electron’s motion in the second electrode is confined to the two transverse directions (Wolf 1985, sec.2.4.1). The tunnelling probability is then determined by the bound-state wavefunction, which is independent of the transverse wavevector. In the model analysed by BenDaniel and Duke (1967) this leads to a tunnelling conductance proportional to the product of the two-dimensional density of states $N(E_t)$ and a barrier transmission coefficient $D$, similar to $|M|^2$ in (3.8). Bound states may arise because the tunnelling electron sets up a standing wave in a very thin (<100 nm) electrode or in an accumulation-layer potential well at the interface between an $n$-type semiconductor and an oxide tunnelling barrier (Wolf 1985, sec.8.2.2–3). The underlying cause of these ‘final-state effects’ is electronic low dimensionality in the junction electrodes. This prompted Que
and Kirczenow (1988) to suggest that the two-dimensional character of high-$T_c$ superconductors could also lead to the appearance of structure due to the normal-state density of states in the tunnelling conductance. Indeed, such structure has been invoked by several authors to explain anomalous features in the $I$–$V$ characteristics of tunnel junctions made from oxide superconductors (section 4.2.1).

### 3.1.2 Excitation Description of Tunnelling

A tunnelling event may be very generally depicted as the creation of excitations in the two electrodes. For example, tunnelling in an MIM junction is equivalent to the creation of a hole excitation in one electrode and an electron excitation in the other. In the case of an SIN junction the excitation created in the superconductor is a quasiparticle.

Guided by this description, the tunnel current will be derived from (3.4). Rather than specifying the nature of the excitations in this electrode we will just say they have a density $N_i(\xi)$ with probability of occupation $f_i(\xi,T)$. In the ground state no excitations are present at $T = 0$, thus $f_i(\xi,0) = 0$ and tunnelling may occur by creating any of the $N_i(\xi)$ available excitation states. If $T > 0$ a fraction $f_i(\xi,T)$ of excitation states is already filled by thermal activation and so cannot contribute to the tunnel current. The density of states remaining for participation in tunnelling, represented by $n_i(\xi,T)$ in (3.4), is reduced to $(1 - f_i(\xi,T))N_i(\xi)$. Defining similar quantities in the right electrode,

$$n_i(\xi,T) = N_i(\xi)(1 - f_i(\xi,T))$$

and

$$\tilde{n}_i(\xi,T) = N_i(\xi)(1 - f_i(\xi,T))$$

The partial tunnel current from left to right is found by integrating over all possible excitation energies in the left electrode,

$$J_{l\rightarrow r} = \frac{-2\pi e}{h} \int_{-\infty}^{+\infty} |M_{rl}|^2 N_i(\xi_1)N_i(\xi_2)(1 - f(\xi_1))(1 - f(\xi_2))d\xi_1$$

where $\xi_1 + \xi_2 = eV$. In all cases (quasiparticles, electrons and holes) the excitations are fermions so $f(\xi)$ is the Fermi function (2.16) with $1 - f(\xi) = f(-\xi)$, thus

$$J_{l\rightarrow r} = \frac{-2\pi e}{h} \int_{-\infty}^{+\infty} |M_{rl}|^2 N_i(eV - \xi_2)N_i(\xi_2)f(\xi_2 - eV)(1 - f(\xi_2))d\xi_2$$

Clearly this expression will apply to MIM, SIN and SIS junctions when the appropriate density of excitation states is substituted. For an MIM junction this approach must lead
to a partial current identical to (3.7). In this case, excitations created in the right electrode are electrons so \( N_r(\xi) = N_e(\xi) \), whilst those in the left are holes so \( N_l(\xi) = N_h(\xi) \). Since the excitation energy of an electron is simply its energy relative to the Fermi energy then \( \xi_r \equiv \varepsilon = (E - E_F) \). Substituting these expressions into (3.10) and using \( N_h(\xi) \approx N_e(-\xi) \) yields (3.7) as required.

A partial current in the reverse direction results from the annihilation of pre-existing thermal excitations. In this case,

\[
n_r(\xi) = N_r(\xi) f_r(\xi, T)
\]

and

\[
\tilde{n}_l(\xi) = N_l(\xi) f_l(\xi, T)
\]

so from (3.5) and rearranging as above

\[
J_{r \rightarrow l} = -\frac{2\pi e}{\hbar} \int_{-\infty}^{\infty} N_r(eV - \xi_r) N_l(\xi_r) f_r(\xi_r) (1 - f(\xi_r - eV)) d\xi_r
\]

When combined with (3.10) the net tunnel current in the junction may be calculated.

### 3.2 SIN Junctions

Figure 3.3 depicts tunnelling in an SIN junction as the creation of a hole excitation in the normal electrode and a quasiparticle excitation in the superconducting electrode. Consider a hypothetical situation in which the excitation of quasiparticles (other than those created by tunnelling) has somehow been suppressed. Since \( \xi_r \geq \Delta \) the energy conservation condition cannot be satisfied for \( eV < \Delta \) so no tunnelling occurs. In contrast, if \( eV > \Delta \) then for each value of \( \xi_l \) there are potentially two excitation states available in the superconductor that may participate in the tunnelling process. In figure 3.3 these are labelled with wavevectors \( k' \) and \( k'' \). The probability that an electron will be able to tunnel into a state \( k \) is the probability \( u_k^2 \) that it is not occupied by a Cooper pair, which at all temperatures is given in BCS theory by

\[
u_k^2 = 1 - v_k^2 = \frac{1}{2} \left( 1 + \frac{\varepsilon_k}{\xi_k} \right)
\]

where \( \xi_k \) is the excitation energy of a quasiparticle with wavevector \( k \) (2.9) and \( \varepsilon_k \) is the kinetic energy of the electrons that form the Cooper pair relative to the Fermi energy, (i.e., \( \hbar^2 k^2 / 2m - E_F \)). If the energy gap is isotropic (\( \Delta_k = \Delta_{k'} \)) then \( \varepsilon_{k'} = -\varepsilon_k \), and the combined probability of tunnelling into state \( k' \) or \( k'' \) is
This means that if the effect of thermally excited quasiparticles is ignored the electron will definitely be able to tunnel into one of the two states. Cancellation of the coherence factors $u_k$ and $v_k$ associated with quasiparticles allows calculation of the tunnel current in an SIN junction to be simplified. This is done by replacing the real distribution of Cooper pairs, $v_k^2$, with one in which pair occupation is unity below the Fermi energy and zero above. In this simplified model tunnelling occurs by movement of electrons from the normal electrode into empty quasielectron states above $k_F$ in the superconducting electrode. Tunnelling in the reverse direction is represented by movement of a quasielectron from below the Fermi level of the superconductor into an empty electron state in the normal-metal. Clearly this model ignores to a large extent the true nature of excitations in a superconductor.

When thermal excitations are present some of the quasielectron states above $k_F$ will already be filled so will no longer be available to an electron tunnelling from the normal electrode. Similarly, in creating a thermally excited quasiparticle a Cooper pair is annihilated and a quasihole state below $k_F$ is created, opening an additional channel for tunnelling into the superconductor. This model closely parallels the one used for tunnelling between normal-metals in which the process was regarded as transfer of Bloch states across the barrier. The argument presented above has shown that in terms of simply counting the number of tunnelling events that occur this model is equivalent to the true tunnelling process. However, it is only a representation that preserves the counting statistics and its artificial nature should be kept in mind when calculating the tunnel current in this way.
Having drawn the analogy with MIM tunnelling it can be seen that the SIN tunnel current is given by (3.8) but in this case the density of excitation states in the left electrode is simply that of a normal-metal, $N_{nl}(\xi)$, whilst in the superconducting electrode it is given by (2.14), i.e.,

$$I_\text{s} = \frac{2\pi eA}{\hbar} \int_{-\infty}^{\infty} |M|^2 N_{nl}(E) N_{nl}(E - eV)n_s(e)(f(e - eV) - f(e))de$$  \hspace{1cm} (3.11)$$

$N_{nl}(E)$ is the density of Bloch states in the right electrode and $E$ is the energy of the Bloch state with the same wavevector as a quasiparticle of excitation energy $\xi$ (i.e., $E = E_F + \sqrt{\xi^2 - \Delta^2}$). The normalised density of excitation states, $n_s(\xi)$, is often loosely called the superconductor’s density of states or the tunnelling density of states.

A diagram equivalent to figure 3.2 may be used to describe SIN tunnelling (figure 3.4). Because of the similarity between the right-hand side of figure 3.4 and the density of states plotted for a semiconductor this simplified description of tunnelling is often called the semiconductor model.

![Diagram](image)

**Fig. 3.4.** Quasiparticle tunnelling in an SIN junction for $T > 0$. The size of the energy gap and the number of thermal excitations have been greatly exaggerated for clarity.
At $T = 0$ there are no thermal excitations, so both electrodes are in their ground state with all states below the Fermi level filled and all states above $E_F$ empty (this is the simplified distribution in the superconducting electrode). With no applied bias the system is in equilibrium and the Fermi levels in the two electrodes are at the same energy. Because there are no vacant states at the same energy as occupied states in the other electrode (i.e., at the same horizontal level in the diagram) then elastic tunnelling cannot occur and current does not flow. As the voltage between the electrodes is increased the right-hand side of this diagram is shifted rigidly downwards by $eV$. The current remains at zero until at $eV = \Delta$ filled states on the left line up with the large number of empty states just above the gap in the right electrode and the current rises rapidly (figure 3.5). The physical interpretation is that an energy $\Delta$ must be supplied by the bias before an excitation can be created in the superconductor. Further increase in bias creates a greater overlap of the shaded and unshaded regions and so the current increases. At high biases the increase in overlap is proportional to the change in bias so the $I$–$V$ characteristics are linear.

At finite temperatures thermal excitations create filled states above $E_F$ in the left electrode and empty states below the gap in the right electrode. Overlap of filled and unfilled states occurs, therefore, for applied biases with $eV < \Delta$ and so the current–voltage curve is broadened.

Quantitative expressions for the tunnel current and conductance of an SIN tunnel junction can be calculated based on an assumption that the superconductor may be modelled by weak-coupling BCS theory (i.e., the energy gap is isotropic and satisfies $\Delta \ll \hbar \omega_p \ll E_F$). This is true of many conventional metals; for example, tin has $\Delta = 0.57$ meV, $\hbar \omega_p = 17.2$ meV and $E_F = 10.03$ eV. As described at the end of section 3.1.1, in an ideal junction $|M|^2 N_{nr}(E)N_{nl}(E)$ is approximately constant so may be removed from the integral in (3.11) to give

$$I_s = \frac{G_n}{e} \int_{-\infty}^{\infty} n_s(\varepsilon) \left( f(\varepsilon - eV) - f(\varepsilon) \right) d\varepsilon$$

(3.12)

$G_n$ is the conductance of the junction when both electrodes are normal (3.9). Using the BCS density of states the current at $T = 0$ is exactly

$$I_s(V) = \begin{cases} 
\frac{G_n}{e} \sqrt{eV^2 - \Delta^2} & eV > \Delta \\
0 & eV < \Delta
\end{cases}$$

(3.13)

however, at finite temperatures (3.12) must be integrated numerically.
At very low temperatures the energy gap may be determined by measuring the onset voltage of the tunnel current. More information may be obtained, however, from the differential conductance of the junction. From (3.12)

\[
G_e (V) = \frac{dI}{dV} = \frac{G_a}{e} \int_{-\infty}^{+\infty} n_s(\epsilon) \frac{df(\epsilon - eV)}{d\epsilon} d\epsilon
\]

(3.14)

This may be written

\[
\sigma(V) = \int_{-\infty}^{+\infty} n_s(\epsilon) t(\epsilon - eV) d\epsilon = n_s(eV) * t(eV)
\]

(3.15)

Hence the normalised conductance, \(\sigma(V)\), defined as the differential conductance of the SIN junction divided by the conductance when both electrodes are normal, is a convolution of \(n_s(V)\) with a thermal broadening function

\[
t(x) = \frac{1}{4k_B T} \text{sech}^2 \left( \frac{x}{2k_B T} \right)
\]

(3.16)

This is a bell-shaped function with unit area, height \((4k_B T)^{-1}\) and a full width at height 
\((4k_B T)^{-1}/c\) of

\[
w_{1/c} = 2k_B T \ln \left[ (2c - 1) + 2c^{1/2}(c - 1)^{1/2} \right]
\]

i.e., the full width at half maximum (FWHM) is \(3.52 \ k_B T\). In the low-temperature limit the thermal broadening function is a delta function

\[
\delta(x) = \lim_{T \to 0} t(x)
\]

so

\[
n_s(eV) = \lim_{T \to 0} \sigma(V)
\]

(3.17)

At sufficiently low temperatures the normalised conductance is a direct measurement of the density of excitation states in the superconductor. In particular, the gap edge, \(\Delta_0\), may be found simply by measuring the position of the peak in conductance (figure 3.6).

This approximation may be made when the width of thermal broadening is much smaller than the energy gap, which for a BCS superconductor with \(\Delta_0 = 3.52 \ k_B T_c\) is equivalent to \(T/T_c << \frac{1}{2}\). This condition is clearly satisfied for tunnelling measurements on the high-temperature superconductors at liquid-helium temperature \((T/T_c \approx 0.05)\) so one would hope to be able to find the excitation spectrum and energy gap very easily. However, this simple result assumes that the energy gap in the superconductor is isotropic and homogeneous and that the junction is ‘ideal’.
Fig. 3.5. Calculated current–voltage characteristics of an SIN junction at $T = 0$ (continuous curve), $T/T_c = 0.05$ (dotted curve) and $T/T_c = 0.85$ (dashed curve). The latter two curves correspond to a superconductor with $T_c$ of 91 K at liquid-helium and liquid-nitrogen temperatures respectively. Also shown is the $I$–$V$ characteristic obtained when both electrodes are normal (dot-dashed curve).

Fig. 3.6. Normalised conductance–voltage characteristics corresponding to figure 3.5.
Fig. 3.7. Normalised conductance–voltage characteristics of an ideal BCS SIN tunnel junction as a function of temperature. This surface was calculated from (3.14) using $\Delta(T)/\Delta(0)$ values taken from Mühlschlegel (1959). As temperature is increased the contour of $G_s/G_n = 1$ remains at roughly the same voltage whilst the conductance peak increases gradually in voltage. This behaviour can be seen more easily in figure 3.16.
As temperature is increased the conductance peak decreases in height and becomes broader, giving rise to finite conductance in the gap region. In addition the conductance maximum moves to higher values of $eV/\Delta(T)$ so evaluation of the energy gap from the tunnelling characteristics is more complicated. In addition to these changes (caused by thermal broadening) the energy gap is itself temperature dependent (figure 2.5). When these effects are combined the conductance–voltage characteristics behave as shown in figure 3.7.

### 3.3 SIS Junctions

At $T = 0$ both superconductors are in the ground state so all electrons are in pairs. Tunnelling may occur by annihilation of a Cooper pair state in one superconducting electrode with the creation of a quasiparticle excitation in each electrode (figure 3.8).

The minimum energy required for this process is the sum of the energy gaps in the two superconductors, $\Delta_i + \Delta_r$. Since this energy must be supplied by the external bias no current can flow for $eV < \Delta_i + \Delta_r$. At larger biases the current is given by (3.8) with a density of excitation states in both electrodes given by (2.14),

$$I_{ss} = \frac{2\pi eA}{\hbar} \int_{-\infty}^{+\infty} |M|^2 N_{ml}(E) N_{nl}(E - eV) \frac{|\epsilon - eV|}{\sqrt{(E - eV)^2 - \Delta_i^2}} \frac{|\epsilon|}{\sqrt{\epsilon^2 - \Delta_r^2}} (f(\epsilon - eV) - f(\epsilon)) d\epsilon$$

(3.18)

At $T = 0$ this may be written

$$I_{ss}(T = 0) = \frac{G_n}{e} \left( \beta^{-1} E(\gamma) - 2\Delta_i \Delta_r \beta K(\gamma) \right)$$

![Excitation energy diagram of tunnelling in an SIS junction at $T = 0$. Solid double circles represent a Cooper pair, the shaded circles represent quasiparticles.](image)
where
\[
\beta^{-1} = \sqrt{(eV)^2 - (\Delta_i - \Delta_r)^2} \quad \gamma = \beta \sqrt{(eV)^2 - (\Delta_i + \Delta_r)^2}
\]

\[
K(\gamma) = \int_0^1 \frac{dt}{(1-t^2)^{1/2}(1 - \gamma^2 t^2)^{1/2}} \quad \text{and} \quad E(\gamma) = \int_0^1 \frac{(1 - \gamma^2 t^2)^{1/2}}{(1-t^2)^{1/2}} dt
\]

As \( eV \rightarrow \Delta_i + \Delta_r \), \( I_n \) remains finite so the current is discontinuous at this bias (figure 3.9). At higher temperatures the current must be calculated by numerical integration of (3.18). This results in non-zero current for biases \( eV < \Delta_i + \Delta_r \) with a region of negative differential resistance between \( eV = |\Delta_i - \Delta_r| \) and \( eV = \Delta_i + \Delta_r \).

The curve shown in figure 3.10 may be obtained by measurement only if the junction is driven by a constant-voltage generator, since a finite source impedance will lead to current–voltage characteristics that are determined by the load line of the source. If the impedance is too high the junction voltage will jump from just below \( eV = |\Delta_i - \Delta_r| \) to a point on the curve near \( eV = \Delta_i + \Delta_r \) as the current is increased. The voltage jump that takes place as current is decreased will occur at a different position on the curve so measured characteristics will be hysteretic and determination of the energy gap is further complicated (Douglass and Meservey, 1964).

The negative differential conductance region may be explained qualitatively by the semiconductor model. When the applied bias is smaller than the sum of the energy gaps in the two electrodes then the tunnelling process illustrated in figure 3.8 cannot occur; however, current may still flow due to tunnelling of thermal excitations (figure 3.11a). As the bias is increased from zero to \( |\Delta_i - \Delta_r| \) (figure 3.11a–b) there is

---

**Fig. 3.9.** Normalised \( I-V \) characteristics of an SIS junction at \( T = 0 \). At \( eV = \Delta_i + \Delta_r \) the current jumps from zero to \( \pi \sqrt{\Delta_i \Delta_r} \frac{I_n}{2(\Delta_i + \Delta_r)} \), where \( I_n \) is the current flowing when both electrodes are normal (dashed line).
an increase in the number of excitations in the left electrode that are able to tunnel into empty states in the right electrode so the current increases. When the bias is further increased from \(|\Delta_i - \Delta_r|\) to \((\Delta_i + \Delta_r)\) (figure 3.11b–c) there is no change in the number of excitations in the left electrode that are potentially capable of tunnelling; however, there is a decrease in the density of empty states in the right electrode resulting in an incremental decrease in current and negative differential resistance.

When the bias is increased beyond \((\Delta_i + \Delta_r)/e\) excitations may be created in both electrodes and tunnelling may occur by breaking a Cooper pair. The sudden onset of this new tunnelling mechanism results in a discontinuous increase in current with magnitude

\[
\Delta I = \frac{\pi G_n \sqrt{|\Delta_i \Delta_r|}}{2e} \cdot \frac{\sinh[(\Delta_i + \Delta_r)/k_B T]}{2 \cosh(\Delta_i/k_B T) \cosh(\Delta_r/k_B T)}
\]

In figure 3.11 this increase occurs as the large number of occupied states just below the gap in the left electrode line up with the large number of empty states just above the gap in the right electrode. Theoretically the discontinuity persists even at finite temperatures; however, a number of non-ideal properties of the tunnel junction and superconducting electrodes may broaden the corresponding peak in differential conductance (see section 4.3).

![Fig. 3.10. Sketch of normalised I–V characteristics of an SIS junction for \(T > 0\), showing the 'sum and difference features' at \(eV = |\Delta_i - \Delta_r|\) and \(\Delta_i + \Delta_r\).]
Fig. 3.11. Semiconductor model of the normalised density of excitation states in an SIS tunnel junction for $T > 0$. Horizontal arrows indicate elastic tunnelling events described in the text.
3.4 ScN Junctions

The tunnelling barrier between a superconductor and normal-metal point-contact may be made thinner by increasing the force between the spear and anvil. As this occurs there is a continuous and appreciable change in the conductance–voltage characteristics from those of an SIN tunnel junction to those of a superconductor–constriction–normal-metal (ScN) contact. The latter is formed when the barrier is completely removed and the two electrodes are strongly coupled.

Rather than exhibiting reduced conductance for biases smaller than $\Delta/e$, as in the case of an SIN junction, the conductance of a perfect ScN contact is twice that of the normal state, $G_n$. Above $eV = \Delta$ the conductance falls asymptotically and monotonically towards $G_n$ so at high biases

$$I = G_n V + I_{ex} V/|V|$$

where $I_{ex}$ is the excess current (Pankove 1966). The source of $I_{ex}$ and the conductance for $eV < 0$ is the conversion of a single-particle current in the normal-metal into a paired current in the superconductor. This occurs when an electron of momentum $k'$ combines with an electron at the SN interface to form a Cooper pair and a hole with momentum $-k'$. Having the opposite momentum to the incident electron the hole is retro-reflected back into the normal electrode. This process, called Andreev reflection, is illustrated in figure 3.12. Above $eV = \Delta$ the probability of direct quasiparticle tunnelling rises rapidly because quasiparticle excitations may be created in the superconductor. In contrast the probability of Andreev reflection falls and the excess current saturates.

**Fig. 3.12.** Excitation-diagram description of Andreev reflection. One open circle in the left (normal) electrode represents the removal of the incident electron whilst the other is the retro-reflected hole. The solid circles represent the Cooper pair at the chemical potential.
If a residual tunnelling barrier remains at the interface, incident electrons will also be normally reflected leading to a zero-bias local minimum in the conductance and small peaks at $eV = \pm \Delta$. Blonder, Tinkham and Klapwijk (BTK, 1982) characterised scattering at the interface (e.g., due to oxide layer, roughness or mismatch of $v_F$ in the two electrodes) by a delta-function repulsive potential with amplitude $\hbar v_F Z^*$. The Bogoliubov equations were used to calculate the probability of Andreev reflection, $A(E)$, ordinary reflection, $B(E)$, and quasiparticle tunnelling $(1 - A - B)$ of electrons incident on the interface. At all temperatures

$$A(E) = \frac{\Delta^2}{E^2 + (\Delta^2 - E^2)(1 + 2Z^2)^2} \quad \text{for } E < \Delta \quad (3.19a)$$

$$B(E) = 1 - A(E)$$

and

$$A(E) = \frac{u_0^2 v_0^2}{\gamma^2} \quad \text{for } E > \Delta \quad (3.19b)$$

$$B(E) = \frac{Z^2(1 + Z^2)(u_0^2 - v_0^2)^2}{\gamma^2}$$

where $\Delta$ is the energy gap at the junction temperature, $u_0^2$ and $v_0^2$ are the Cooper pair occupation factors ((2.8) and (2.9) with $\xi = E$) and $\gamma^2 = [u_0^2 + Z^2 (u_0^2 - v_0^2)]^2$. The $I-V$ characteristics (figure 3.13) can then be calculated from

$$I = (1 + Z^2)G_0/e \int_{-\infty}^{\infty} [1 + A(E) + B(E)][f(E - eV) - f(E)]dE \quad (3.20)$$

The single scalar parameter, $Z$, enables the model to describe the full range of behaviour from the limiting case of a pure metallic point contact ($Z = 0$) to a tunnel junction ($Z \geq 10$).

At finite temperatures the conductance–voltage characteristics (figure 3.14) are smeared by convolution with the elastic-tunnelling thermal broadening function (3.16). This ‘fills in’ the central conductance minimum, reduces the height of the conductance peaks and shifts the position of the conductance peaks to higher energies for large $Z$, and lower energies for low $Z$.

---

* The assumption of a delta-function potential is reasonable only if the mfp > contact diameter, $d$, so that electron transport is ballistic, and if $\xi(T=0) > d$. The latter could be invalid for HTSC ScN contacts.
Fig. 3.13. BTK model current–voltage characteristics calculated for various barrier strengths, $Z$, at $T = 0$. The normalised current is $Ie/\Delta G_n$.

Fig. 3.14. Conductance–voltage characteristics corresponding to figure 3.13.
In general the energy gap must be determined by fitting the results of BTK theory to conductance–voltage characteristics measured at several temperatures - to obtain both $Z$ and $\Delta$. Several other strategies exist for estimating $Z$ from more limited data (Smith et al. 1993).

Experimental results for contacts with $R_n \geq 0.5 \, \Omega$ are well explained by the BTK model with $Z > 0.3$ (Blonder and Tinkham 1983, Reinertson et al. 1990). Lower-resistance contacts, however, are distorted by large spurious structure and locally depressed $\Delta$ caused by Joule heating of the contact region. The onset of quasiparticle tunnelling may totally quench superconductivity resulting in peaks in $\frac{\text{d}V}{\text{d}I}$ at $eV = \pm \Delta$ (McLean 1984, p.103).

Most ScN experiments with high-$T_c$ superconductors have used a bilayer configuration (figure 3.15) where, due to the thick normal layer, only retro-reflected holes return through the orifice and contribute to the excess current (van Kempen et al. 1989). An advantage of using a thick overlayer is that Andreev reflection may be positively identified by measuring a decrease in the excess current with magnetic field, due to trajectory bending by the Lorentz force.

BTK theory was extended by van Son, van Kempen and Wyder (SKW 1988a) to accommodate the normal layer and, in particular, a gradually varying order parameter near the NS interface due to a proximity-effect-induced gap in N or a depressed gap at the surface of S. This predicts that if scattering around the interface remains weak the probability of Andreev reflection is only slightly modified and the bulk value of $\Delta$ may still be determined. This may be an advantage over SIN tunnelling, which will measure the severely depressed $\Delta$ at the surface (van Kempen et al. 1990). However, SKW theory has been only partially successful in explaining experimental results on Ag–Pb bilayers (SKW 1988b). In addition, non-ideal retro-reflection, temperature broadening and energy-dependent electron–phonon scattering in the normal-metal may

---

**Fig. 3.15.** SnScN bilayer configuration. Incident electrons are labelled $e$ and retro-reflected holes, created at the superconductor–normal-metal interface, are labelled $h$. Lines represent particle trajectories with (solid) and without (dashed) the magnetic field, $B$. 

---
reduce the excess current and make the inferred energy gap smaller than that in the bulk (van Bentum et al. 1989).

### 3.5 Inelastic Tunnelling

This section considers inelastic tunnelling, in which a tunnelling particle changes energy by interacting with a bosonic excitation (e.g., a phonon) in the barrier or electrodes. The inelastic current arising from a spectrum, $F(\omega)$, of bosonic excitations is

$$I_\text{i}(V) \propto \int_{-\infty}^{\infty} d\omega F(\omega)n(\omega) \int_{-\infty}^{\infty} dE \left\{ f(E)[1 - f(E + eV + \hbar \omega)] - f(E + eV - \hbar \omega)[1 - f(E)] \right\}$$

$$+ \int_{-\infty}^{\infty} d\omega F(\omega)[n(\omega) + 1] \int_{-\infty}^{\infty} dE \left\{ f(E)[1 - f(E + eV - \hbar \omega)] - f(E + eV + \hbar \omega)[1 - f(E)] \right\}$$

where $f(E)$ and $n(\omega)$ are the Fermi and Bose distributions. The two terms describe absorption and emission of an excitation with energy $\hbar \omega$ respectively. At zero temperature and with symmetrical $F(\omega)$ the conductance simplifies to

$$\frac{dI_\text{i}(V)}{d(eV)} \propto \int_{0}^{\mid eV \mid} d\omega F(\omega)$$

(3.21)

If there is a single possible excitation, with energy $\hbar \omega_1$, then $F(\omega) = c_1 \delta(\omega)$. In this case the inelastic conductance is

$$\frac{dI_\text{i}(V)}{d(eV)} = \begin{cases} 
0 & |eV| < \hbar \omega_1 \\
\frac{c_1}{2\hbar} & |eV| \geq \hbar \omega_1
\end{cases}$$

For $|eV| < \hbar \omega_1$ there is no inelastic conductance because the bias energy is too small to excite the vibration. At the threshold voltage, $|eV| = \hbar \omega_1$, an ‘inelastic channel opens’ as the highest energy electrons are able to excite a single phonon and tunnel to the Fermi level in the other electrode. At this point there is a step increase in the inelastic conductance. With a spectrum of discrete modes,

$$F(\omega) = \sum_j c_j \delta(\omega_j)$$

a step occurs whenever the bias exceeds a characteristic energy $\hbar \omega_j$. At zero temperature the excitation spectrum may be determined by measuring the tunnelling conductance derivative, since $d^2I/d(eV)^2 = F(eV/\hbar)$. This forms the basis of a number of ‘threshold spectroscopies’ (Wolf 1985, sec.8.3), including inelastic electron
tunnelling spectroscopy (IETS), which is the study of vibrational modes of organic molecules adsorbed onto the insulating barrier (Adkins and Phillips 1985). Although the increase in conductance when an inelastic channel opens is much smaller than the elastic ‘background conductance’ (typically < 0.5%), the latter is smoothly varying so does not significantly affect the second-derivative spectrum; IETS may detect as little as 1% of a monolayer of adsorbed molecules.

Thermal broadening at finite temperatures may be represented by a convolution,

\[
\frac{d^2I}{d(eV)^2} = F(eV/\hbar) * t(eV)
\]

where

\[
t(x) = \frac{1}{k_B T} \exp(z) \frac{(z-2)\exp(z) + z + 2}{(\exp(z) - 1)^3}
\]

and \(z = x/k_B T\). This bell-shaped thermal broadening function has a FWHM of 5.4 \(k_B T\) (Lambe and Jaklevic 1968). This is wider than the elastic-tunnelling thermal broadening (3.5 \(k_B T\)) because the current in the latter case is proportional to the difference between two Fermi functions whilst for inelastic tunnelling it is proportional to two Fermi functions convolved together. To increase spectral resolution inelastic tunnelling measurements are carried out at 4.2 K.

### 3.6 Indirect Tunnelling

The conductance of a tunnel junction may be increased by several orders of magnitude if the barrier contains intermediate states with similar energy to the tunnelling electrons. These states may arise from impurities or defects in the surface or oxide layer. *Two-step tunnelling* occurs if the lifetime \(\tau\) of a barrier state is large enough for a tunnelling electron to become localized for some time, before tunnelling on to the second electrode or back to the first electrode. This occurs when

\[
\Gamma = \frac{\hbar}{\tau} \ll \frac{\hbar^2}{2m^*a^2}
\]

where \(\Gamma\) is the energy width of the intermediate state, \(m^*\) is the electron effective mass and \(a\) is the radius of the local state’s wavefunction (Wolf 1985, sec.2.5.3).

If \(\tau\) is small the electron is not bound by the barrier state; however the transmission probability may still be increased by resonant enhancement of the tunnelling particle’s wavefunction in the vicinity of the state; this is *resonant tunnelling*.

These elastic indirect tunnelling processes begin when the Fermi level of one electrode aligns with the barrier state’s energy level. Because the barrier potential is distorted
by an applied bias the voltage \( V_t \) at which this occurs depends on the location of the state in the barrier. A simple trapezoidal barrier has threshold voltages,

\[
e V_t = \begin{cases} 
-wE_0/(w-z) & V < 0 \\
wE_0/z & V > 0 
\end{cases}
\]  

(3.23)

where \( w \) is the barrier width, \( z \) is the location of the localised state in the barrier and \( E_0 \) is its zero-bias energy level. Structure arising from indirect tunnelling will only be symmetrical about zero bias if the state is located in the centre of the barrier. If the localised state is very close to one electrode then resonant tunnelling may not take place at all when that electrode is negatively biased.

The transmission probability is most enhanced by resonant or two-step tunnelling when the wavefunctions from the two electrodes have equal amplitude at the intermediate state. With planar electrodes this occurs when the state is located at the centre of the barrier (i.e., \( z = w/2 \)). The current then varies exponentially with barrier width, but with half the coefficient of direct tunnelling. i.e.,

\[
j \propto \begin{cases} 
n_s \exp(-2\alpha w) & \text{direct} \\
n_s n_L \exp(-\alpha w) & \text{indirect} 
\end{cases}
\]

where \( n_s \) is the density of states in the electrodes, \( n_L \) is the density of intermediate states in the centre of the barrier, \( \alpha = \sqrt{2m\phi/h} \) and \( \phi \) is the barrier height. Because indirect tunnelling has the smaller exponential coefficient its relative importance will increase with barrier thickness.

The wavefunction of a small-radius tip decays more rapidly than that of a plane electrode (by an extra factor or \( r^{-1} \)) so, in contrast with the planar junction case, resonant tunnelling in an STM will be dominated by barrier states closer to the tip (Berthe and Halbritter 1991).

Resonant tunnelling results in a peak in the conductance–voltage characteristics centred at the threshold voltage (3.23). However, discrete peaks can only resolved in small–area tunnel junctions, such as the 0.1 \( \mu \text{m}^2 \) amorphous–Si/SiO\(_x\) barrier junctions investigated by Naito and Beasley (1987). In larger-area junctions conductance peaks from many intermediate states with a distribution of energies and positions in the barrier combine to give a conductance step or a continuous background (Cockburn 1988).

Elastic two-step tunnelling also results in a conductance peak. However, when a
tunnelling electron is localized on the intermediate state the lattice distorts and energy is lost to phonons. In this case inelastic tunnelling occurs via the intermediate state and there is a step increase in conductance. In other words, at the threshold voltage there is both a step increase in current and an increase in slope of the $I$–$V$ curve. This effect was also observed by Naito and Beasley (1987).

At high voltages the probability of transmission of an electron through the barrier may be greatest if it tunnels from trap to trap, rather than directly or via a single intermediate state. The multiple intermediate states are bridged by emission of a barrier excitation with each ‘hop’. Hopping conductance is most favoured by high temperatures, high voltages, wide barriers and large $n_t$ (Halbritter 1992).

If a minimum energy, $U$, is required to place an electron in the intermediate state then resonant tunnelling via that state is obstructed. This effect becomes important when $U$ is larger than the energy width of the intermediate state, i.e., $U > \Gamma$ (Halbritter 1982, 1985). The intrinsic energy $E_0$ of the state is effectively increased by $U$ and the threshold voltage (3.23) is increased. Indirect tunnelling can then occur only if $U$ is provided by the bias voltage, thermal energy or both. The ‘blocking energy’ $U$ can arise from either capacitance of the intermediate state, if it is a small physical particle (see section 4.2.4), or from correlation energy, if the intermediate state is one of a number of localized states in the barrier (section 4.8). In either case $U$ is referred to as a ‘Coulomb blockade’ or barrier and the reduced conductance at low voltages is a ‘Coulomb gap’.

### 3.7 Methods for Estimating the Energy Gap

This chapter has attempted to outline theories that predict the tunnelling characteristics for a given energy gap and other junction and material parameters. The (harder) inverse problem is to determine the energy gap and density of states from experimental tunnelling characteristics. The experimental data will be complicated by thermal broadening, noise and possibly broadening due to anisotropy or inhomogeneity in the sample (section 4.3). Different analysis methods lead to significant ($\approx 8\%$) variation in the estimated gap even from a single result (Edwards 1990).

Various methods may be used to estimate the energy gap. These are discussed briefly in turn:

(i) Section 3.2 showed that if $T < \frac{1}{2} T_c$ then thermal broadening may be neglected. If an SIN junction is ideal then $n_j(eV) \approx \sigma(V)$ and $2\Delta$ is found from the separation of the conductance peaks. The position of the sum and difference features in the tunnelling characteristics of an SIS’ junction are sharper and less affected by thermal broadening so may be used to estimate the energy gap in both electrodes to higher temperatures.
(ii) A useful estimate of \( \Delta(0) \) for an ideal BCS SIN junction may be obtained from the bias at which the normalised conductance equals unity. Figure 3.16 shows that this differs from \( \Delta(0)/e \) by less than 10\% over the entire temperature range. Furthermore, if the reduced temperature is known the \( \sigma = 1 \) contour may be used to determine \( \Delta(0) \) exactly. Similarly, if the position of the conductance peak is measured \( \Delta(0) \) may be found from the short dashed line in figure 3.16. The conductance peak’s voltage increases approximately linearly with temperature, and \( \Delta(0) \) may be estimated from

\[
eV_{\text{peak}} \approx \Delta(0) \left( 1 + 0.836 \frac{T}{T_c} \right)
\]

with less than 2\% error over the full temperature range (maximum error at \( T/T_c = 0.2 \) and 0.7, zero error at 0 and 0.5).

**Fig. 3.16.** Normalised conductance–voltage characteristics of an ideal BCS SIN tunnel junction as a function of temperature. This plot is an alternative representation of figure 3.7. Solid lines are contours of normalised conductance (labelled with their \( \sigma = G/G_n \) value), the short-dashed line (labelled ‘max’) shows the position of the conductance peak.
(iii) From (3.13) plotting $I(V)^2$ against $V^2$ yields a straight line, which intercepts the $I = 0$ axis at $(\Delta e)^2$. At finite temperatures the curve will be rounded near this point so an intercept is obtained by extrapolating from the straight region at higher biases. An additional upward curvature occurs in the $I-V$ characteristics of high-temperature superconductors so the intercept cannot be determined unambiguously. Nevertheless, this method has been used to estimate the energy gap in these materials (see, e.g., Hawley et al. 1987).

(iv) The normalised zero-bias conductance of an ideal BCS SIN junction decreases monotonically with increasing $\Delta(T)/k_BT$ so measuring both $\sigma(0)$ and temperature allows $\Delta(T)$ to be found from pre-calculated tables derived from (3.15)–(3.16). At low bias values the tunnelling current may be written

$$I_s = 2G_n \Delta e \sum_{m=0}^{\infty} (-1)^{m+1} K_1(m \Delta / k_BT) \sinh(meV/k_BT)$$

where $K_1(x)$ is the first-order modified Bessel function of the second kind. This series, which converges rapidly for $eV < \Delta$, may be obtained by expansion of the Fermi functions in (3.12) (Solymar 1972, app.3). At sufficiently low temperatures and at low bias it has a limiting form

$$\lim_{T \to 0} \left( \frac{I_s}{I_n} \right) = \frac{G_s(0)}{G_n(0)} = \sqrt{\frac{2\pi \Delta(0)}{k_BT}} \exp(-\Delta(0)/k_BT)$$

By varying the single adjustable parameter, $\Delta(0)$, this may fitted to experimentally determined $\sigma(V=0,T)$ values using a suitable non-linear optimisation routine. Alternatively the square-root term may be assumed constant and $\Delta(0)$ estimated from linear regression of $\ln \sigma(0)$ versus $-1/k_BT$; however this increases the error by about 20%. Methods that use $\sigma(0)$ are subject to error from leakage currents so are not suitable for analysing high-$T_c$ superconductor tunnel junctions.

(v) The objective of the four methods outlined above is to nullify the effects of thermal broadening in order to estimate the energy gap. The methods were derived from known properties of weak-coupled BCS superconductors in an ideal tunnel junction, and so are rather suspect when used to analyse junctions that do not clearly satisfy these conditions. In principle, the effects of thermal broadening may be completely cancelled by deconvolving the broadening function (3.16) from the data using Fourier transform techniques. i.e.,

$$\sigma(V,0) = F^{-1} \left[ \frac{F[\sigma(V,T)]}{F[f(eV,T)]} \right]$$
where $F$ is the Fourier transform functional operator. If the SIN junction were ideal then from (3.15) the deconvolved conductance is identically the density of excitation states in the superconductor. This method has the advantage that no \textit{a priori} information about the form of $n_s(\varepsilon,T)$ is required. For example, if the superconductor is homogeneous then deconvolution would yield the sample’s density of states spatially averaged over the junction area; the additional broadening resulting from this mechanism would lead to incorrect gap values from methods (i)–(iv) above. Unfortunately, direct deconvolution usually fails due to the sensitivity of this technique to the combination of zeros in $F(t)$ and noise in the input data. These difficulties may be partially overcome using elaborate numerical techniques (e.g., optimal filtering (Press \textit{et al}. 1987)).

(vi) The most versatile way of analysing tunnelling characteristics is to fit appropriate models of the tunnelling conductance to the data using non-linear optimisation methods*. These permit one to account for multiple broadening mechanisms (e.g., thermal and anisotropy or inhomogeneity), junction non-idealities (e.g., energy-dependent tunnelling probabilities) and a non-BCS tunnelling density of states. In view of the presence of these non-idealities in high-$T_c$ superconductor tunnel junctions, fitting techniques appear to be the most suitable method for estimating the energy gap.

Chapter 3 described the conductance of highly idealised superconducting tunnel junctions. The tunnelling probability was independent of energy and the electrodes were modelled as weakly-coupled BCS superconductors with a homogeneous and isotropic energy gap. Many junctions do not satisfy these conditions, resulting in discrepancies between the observed conductance and theoretical predictions. This chapter describes the acutely non-ideal characteristics exhibited by high-temperature superconductor tunnel junctions.

4.1 An Overview

High-$T_c$ superconductor tunnelling characteristics vary greatly from sample to sample and contact to contact (in the case of point-contact work); however, several features are seen in nearly all experiments. Superimposed upon an asymmetric V-shaped or U-shaped background is a gap-like feature, consisting of a region of depressed conductance flanked by regions in which the conductance exceeds the extrapolated background (termed ‘conductance overshoots’). The conductance minimum at zero bias is rarely as low as predicted by BCS theory, whilst the conductance on either side of the gap is only occasionally large enough to give local maxima. The gap and conductance overshoots (when they occur) are significantly broader than expected from thermal broadening alone.

In figure 4.1 the width of the gap region is roughly that predicted by BCS theory; however, there have been reports in the literature of reduced gap values $(2\Delta/k_BT_c)$ ranging from less than 1.0 (Aminov et al. 1987) to greater than 10 (see, e.g., Bulaevskii et al. 1988), with this variation occasionally seen in a single sample (Gallagher et al. 1988). Theories of tunnelling in these materials must account for this wide variation in reported energy gap.

High-resistance point-contact junctions occasionally show a very clear gap-like feature with prominent conductance peaks at the gap edge. This is often highly asymmetric, appears to have a very large ‘reduced-gap’ value and is frequently accompanied by uniformly spaced, multiple conductance peaks. These features indicate that the conductance may be dominated by charging effects, which complicate the interpretation of tunnelling measurements (section 4.4).
The remainder of this chapter describes the non-ideal characteristics seen in high-$T_c$ superconducting tunnel junctions, and briefly outlines theories published in the literature to explain them. In addition, it will discuss the acute experimental difficulties associated with fabrication of high-$T_c$ tunnel junctions, and the origin of the wide range of energy-gap values reported for these materials.

4.2 The Background Conductance

4.2.1 Experimental Observations

All electron tunnelling experiments performed on YBa$_2$Cu$_3$O$_y$ have exhibited a steep v-shaped background conductance, on which any gap-like feature is superimposed. The conductance overshoots usually associated with the superconducting energy gap are severely distorted and may be swamped by the rapidly rising background. The linear region extends to high voltages (e.g., ±500 mV, Svistunov et al. 1989) and is usually asymmetric, with larger slope when the cuprate superconductor is biased negatively with respect to the counterelectrode$^*$. The linear conductance region may extend to zero bias. Kirtley and Scalapino (1990) investigated junctions formed by pressing bulk indium onto the surface of ceramic La$_{2-x}$Sr$_x$CuO$_4$. The superconducting gap of indium (at $T < 3$ K) showed that tunnelling

$^*$ Ba$_{1-x}$K$_x$BiO$_3$ and other bismuth oxide superconductors have the reverse asymmetry.
was the principal conductance mechanism; however, no features were seen that could be associated with a gap in the cuprate - presumably due to a non-superconducting surface layer. The conductance–voltage characteristics were highly linear, with a thermally broadened V-shaped discontinuity at zero bias. By measuring the FWHM of their experimental \(d^3I/dV^3-V\) curve, this broadening was shown to increase roughly as \(6k_BT\) (section 4.2.3).

The presence of linear conductance in the absence of superconducting structure suggests that it is a property of the normal state, or is extrinsic to the oxide materials - possibly due to a degraded surface layer. Evidence for this comes from STM tunnelling experiments on ceramic \(\text{La}_{2-x}\text{Sr}_x\text{CuO}_{4-y}\) (Pan et al. 1987). With an as-prepared sample current could only be detected when the tunnelling tip was embedded in the surface, resulting in conductance–voltage characteristics containing an indistinct gap-like dip, superimposed on a V-shaped background. Vacuum tunnelling was achieved by cleaving the sample \textit{in situ} at 77 K and advancing the tip towards the freshly exposed surface. This resulted in clear tunnelling characteristics with a similar gap but without the V-shaped background. Attempts by Fein \textit{et al.} (1988) to repeat this experiment failed, probably because their samples cleaved along pre-existing grain boundaries to expose more degraded material.

The conductance background seen in \(\text{YBa}_2\text{Cu}_3\text{O}_y\) and \(\text{La}_{2-x}\text{Sr}_x\text{CuO}_{4-y}\) tunnel junctions is consistently reported as being linear or V-shaped. There is wider variation in the background observed in Bi-Sr-Ca-Cu-O tunnel junctions. It has been suggested that flatter backgrounds are due to tunnelling in the \(ab\)-plane, whilst the V-shape results from \(c\)-axis tunnelling (see section 4.2.5).

The conductance of a point-contact between an aluminium tip and an as-grown \(\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8\) crystal was measured by Tulina \textit{et al.} (1990). A strongly increasing background made the characteristics appear highly non-ideal and prevented accurate determination of the superconductor’s energy gap. On breaking the sample at 4.2 K the background was eliminated, and the tunnelling characteristics were described fairly well by the BTK model with \(Z = 1.6\) (section 3.4). Moreover, most contacts exhibited characteristics of a normal-metal–superconductor microconstriction \((Z \ll 0.5)\), suggesting an exceptionally clean interface between the tip and superconductor. This again indicates that the voltage-dependent background arises from a degraded surface layer on the oxide superconductor.

Although the results described above suggest that linear conductance has an extrinsic origin or is associated with the normal state, there is also evidence for some link with the superconducting mechanism. Tunnel junctions of \(\text{Ba}_{1-x}\text{K}_x\text{BiO}_3\) and other low-\(T_c\) bismuthate superconductors display V-shaped conductance backgrounds identical to
those observed in high-\(T_c\) junctions. Sharifi et al. (1991) observed a remarkable linear correlation between the slope of the linear conductance (normalised by the zero-bias conductance) and the superconducting transition temperature of these materials* , i.e.,

\[
\frac{d^2I}{dV^2} \left/ \left( \frac{dI}{dV} \right)_{V=0} \right. \propto T_c
\]  

(4.1)

A linear relationship between the conductance slope and zero-bias conductance has also been observed in \(\text{YBa}_2\text{Cu}_3\text{O}_y\), \(\text{Bi-Sr-Ca-Cu-O}\) and \(\text{Na}_{0.3}\text{WO}_3\) samples, with a roughly linear increase in the constant of proportionality with transition temperature (Srikanth 1992). Linear conductance was also observed in non-superconducting tungsten bronzes, though with much smaller slope.

A series of point-contact experiments on both \(\text{Tl}_2\text{Ba}_2\text{CaCu}_2\text{O}_x\) and \(\text{Pb}_{0.3}\text{Bi}_{1.7}\text{Sr}_2\text{CaCu}_2\text{O}_x\) (Huang et al. 1989b,c) revealed a systematic dependence of the shape of the conductance background on the normal state resistance, \(R_n\). The conductance of high-resistance junctions \((R_n > 1 \text{k}\Omega)\) was seen to increase with bias, as described above, whilst low-resistance junctions exhibited a flatter or even decreasing background. Matsumoto et al. (1992) observed identical behaviour in planar tunnel junctions fabricated on \(\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8\) single crystals; the conductance background became flatter with decreasing barrier width and actually decreased in the absence of an artificial barrier.

Huang et al. interpreted their observations as the result of two competing effects. The first, predominant in high-resistance \(\text{Pb}_{0.3}\text{Bi}_{1.7}\text{Sr}_2\text{CaCu}_2\text{O}_x\) junctions and in all \(\text{YBa}_2\text{Cu}_3\text{O}_y\) junctions, yields the increasing conductance. The origin of this widespread phenomenon will be discussed in detail in the following section. The second process dominates when \(R_n\) is low and produces a decreasing conductance, occasionally accompanied by a small zero-bias conductance peak. The authors discussed several possible mechanisms for this, including zero-bias anomalies and localised superconducting regions, either in parallel or in series with the junction, that are driven normal at high currents. Details in their data led them to reject these explanations and to speculate that the normal-state density of states is strongly energy dependent, with a sharp maximum at the Fermi energy (see section 3.1.1). The authors were unable to measure the normal-state tunnelling conductance of their samples; however, the behaviour predicted by their hypothesis was observed in a thin-film junction fabricated from the \(\text{Ba}_{1-x}\text{K}_x\text{BiO}_3\) (Zasadzinski et al. 1989). These

* Details of this experiment are discussed in appendix B.
experimentally derived ideas tie in nicely with a recent theory of high-temperature superconductivity in strongly interacting quasi-2D electron systems. This proposes that at the maximum $T_c$ of a given material, a van Hove singularity in the density of states is coincident with the Fermi level (Wheatley 1993). The high polarisability of electrons at the singularity means that they are susceptible to a pairing interaction and high transition temperatures can be achieved by a low-energy electron–mediated electron–electron attraction (Little 1988). Changing the level of doping (e.g., $\delta$ in YBa$_2$Cu$_3$O$_{6+x}$) shifts the van Hove singularity away from the optimum position at $E_F$ and the transition temperature is reduced.

### 4.2.2 Barrier Effects

The tunnelling probability in a junction with a narrow, high potential barrier may be considered independent of particle energy and applied bias. This leads to a constant conductance. Many real junctions have an approximately parabolic conductance because the barrier distorts with increasing bias and transmission probability increases (Simmons 1963).

The model constructed in chapter 3 may be extended to take into account these barrier effects. The current may be written

$$I = \frac{qA}{4\pi^2\hbar} \int_0^{E_m} dE_z P(E_z, V) \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y N_l(E, k) N_r(E-eV, k) \left( f(E-eV-E_z) - f(E-E_z) \right)$$

(4.2)

Here the kinetic energy and wavevector are resolved into components parallel (in the $xy$-plane) and normal to the junction. $q$ is the charge of the tunnelling particles, $E_m$ is their maximum energy and $A$ is the junction area. For $T > T_c$ the density of excitation states in the left electrode is simply the density of Bloch states, $N_nl(E, k)$, whilst for $T < T_c$ it is given by (2.14). For isotropic electrodes the wavevector dependence may be dropped.

If the MIM system’s band structure varies slowly compared to the electron wavelength then the WKB approximation holds inside the barrier and the tunnelling probability may be written

$$P(E_z, V) = D(E_z) \exp \left( -\frac{2}{\hbar} \int_{s_1}^{s_2} \sqrt{2m(\varphi(z, V) - E_z)} dz \right)$$

(4.3)

where $\varphi(z, V)$ is the barrier height and $s_{1,2}$ are the classical turning points for energy $E_z$. The product $D(E) N_m(E)$ is energy independent (section 3.1.1) so may be removed from the integral.
Fig. 4.2. Trapezoidal potential barrier. Parameters of the model are junction width \( w \), area \( A \), and barrier heights relative to \( E_F \) with no applied voltage, \( \varphi_1 \) and \( \varphi_2 \).

Using a trapezoidal barrier model and the WKB expression, Brinkman, Dynes and Rowell (BDR 1970) numerically evaluated (4.2) for normal electrodes at \( T = 0 \) and showed that ‘reasonable’ barrier asymmetries could account for commonly observed parabolic characteristics. They investigated a region of parameter space defined by

- \( \bar{\varphi} = \frac{1}{2}[\varphi_1 + \varphi_2] \) from 1.5 to 2.5 eV
- \( \Delta \varphi = \varphi_1 - \varphi_2 \) from 0 to 3 eV
- \( \Delta \varphi / \bar{\varphi} < 1 \)
- \( w \) from 10 to 25 Å
- \( V \) from −250 to +250 mV

and found that the calculated conductance per unit area (in \( \Omega^{-1} \text{cm}^{-2} \)) could be approximated by

\[
\frac{G(V)}{G(0)} = 1 - \left( \frac{A_0 \Delta \varphi}{16 \bar{\varphi}^{\frac{3}{2}}} \right) eV + \left( \frac{9 A_0^2}{128 \bar{\varphi}^2} \right)(eV)^2
\]

(4.4)

where \( A_0 = 4(2m)^{1/2}w/3\hbar \) and \( G(0) = (3.16 \times 10^{10} \bar{\varphi}^{1/2} / w) \exp(-1.025w \bar{\varphi}^{1/2}) \) †.

---

* The units were not explicitly stated in the original paper so (4.4) has been incorrectly used by some authors. In addition, \( w \) is in angstroms and \( \varphi \) is in volts.

† BDR claim their expression is accurate to 10% in the region of parameter space investigated. Calculations for this thesis, however, indicate that the range of validity may be more limited.
This model may be used to estimate barrier parameters by fitting a quadratic curve to the parabolic background of an experimental tunnelling conductance. The best-fit coefficients are then equated to those of the model and the resulting simultaneous equations solved.

Barrier effects may explain the parabolic conductance background occasionally observed with Bi-Sr-Ca-Cu-O junctions. Kirtley (1990) showed that they may also account for V-shaped backgrounds if the work functions are smaller those considered by BDR. From (4.2)–(4.4) the tunnel current at $T = 0$ may be written

$$I(V) = \begin{cases} 
\text{Sgn}(qV) \frac{4\pi mqA}{h^3} \left[ \sqrt{(qV)^2 - \Delta^2} \right]_{E_z+qV}^{E_z} P(E_z, V) \, dE_z \\
+ \int_{E_z+qV}^{E_z+\Delta \text{Sgn}(qV)} \sqrt{(E_z - E_z')^2 - \Delta^2} \, P(E_z', V) \, dE_z' 
\end{cases} \quad \text{for } |qV| < \Delta
$$

(4.5)

where, for a trapezoidal barrier (figure 4.2),

$$\varphi(z, V) = E_z + \varphi_1 + \frac{z}{w}(\varphi_2 + qV - \varphi_1) \quad \text{(4.6)}$$

$$s_1(E_z, V) = \begin{cases} 
0 & E_z \leq E_z + \varphi_1 \\
\frac{w}{[\varphi_2 + qV - \varphi_1]} & \text{otherwise}
\end{cases} \quad \text{(4.7a)}$$

and

$$s_2(E_z, V) = \begin{cases} 
w & E_z \leq E_z + \varphi_2 + qV \\
w\left(1 - \frac{E_z - E_z - \varphi_2 - qV}{[\varphi_2 + qV - \varphi_1]}\right) & \text{otherwise}
\end{cases} \quad \text{(4.7b)}$$

When $\Delta = 0$ the model outlined in (4.5)–(4.7) is identical to the numerical model invesigated by BDR.

As barrier height is reduced the conductance changes from flat to upwardly curving and then to linear (figures 4.3 and 4.4), so variation of a single parameter may explain the backgrounds observed with different junctions.
Fig. 4.3. Effect of reducing barrier height on the conductance of a WKB model MIM junction. To emphasise the change in shape, conductances are normalised by their values at 1 V. Parameters are: $w = 15 \, \text{Å}$, $q = +|e|$ and $E_F = 1 \, \text{eV}$ (Millis et al. 1990). The legend shows the value of $\phi_1 = \phi_2$.

Fig. 4.4. Effect of reducing barrier height on the conductance of a WKB SIN junction. The conductance for each barrier height has been normalised by its value at $+14 \, \text{mV}$ so the gap structure appears to have a similar amplitude. Parameters are as in figure 4.3 but with $\Delta = 13.7 \, \text{meV}$.
Fig. 4.5. Effect of reducing barrier width on the background conductance of a WKB SIN junction. \( \Delta = 13.7 \text{ meV}, \varphi_1 = 10 \text{ meV}, \varphi_2 = 30 \text{ meV}, \hat{E}_c = 1 \text{ eV} \) and \( q = +e \). The legend shows the junction width. As before, the conductance is normalised by its value at +14 mV.

As barrier width is increased the model normalised conductance background remains linear but increases in steepness and so becomes a more dominant feature (figure 4.5). This qualitatively agrees with the experimental results of Huang et al. described in the previous section.

The analysis presented above indicates that V-shaped conductance backgrounds may be due to the prevalence of low, wide barriers; possibly due to a degraded or contaminated surface layer (section 4.5). This would also explain the voltage-independent backgrounds occasionally obtained by cleaving the sample at cryogenic temperatures. However, the barrier heights required to fit the data are now considered too small to be realistic and should result in conductances with larger temperature dependence than is observed (Valles et al. 1991). In addition, the barrier-effect model cannot reproduce the sharp change in slope at zero bias observed at low temperature in the absence of superconducting structure (Kirtley et al. 1992), or the linear increase in slope with zero-bias conductance observed by Sharifi et al. (4.1). More significantly, it is difficult to see how the linear relationship between conductance slope and \( T_c \) can be attributed to barrier effects.

### 4.2.3 Inelastic Tunnelling

A further argument against the barrier-effect explanation for linear conductance is the observed thermal broadening width of \( 6k_B T \) (section 4.2.1). This is larger than the \( 3.5 k_B T \) expected for an elastic tunnelling mechanism and is closer to the \( 5.4 k_B T \) width for inelastic tunnelling (3.22). Although the increased width may be attributed to a
temperature-dependent density of states (Valles et al. 1991), Kirtley and Scalapino (1990) argue that it is evidence for inelastic tunnelling, and that the V-shaped background occurs because there is a uniform distribution of excitations, i.e., \( F(\omega) = F_c = \text{constant} \). In this case (3.21) predicts

\[
\frac{dI(V)}{d(eV)} \propto F_c \frac{|eV|}{\hbar}
\]

This explanation has previously been used to explain linear conductance backgrounds observed in several conventional tunnelling junctions (e.g., Cr–Cr\(_2\)O\(_3\)–Ag). Kirtley et al. proposed that magnons are the excitations responsible for linear conductance in this system and thoroughly tested this hypothesis in an extensive series of experiments on Al–Al\(_2\)O\(_3\)–Cr\(_2\)O\(_3\)–Pb junctions (Kirtley et al. 1992). More controversially, Kirtley and Scalapino suggested that magnons are also responsible for the linear conductance background observed in high-temperature superconductor tunnel junctions. They found good agreement between the calculated inelastic conductance due to a phenomenological \( F(\omega) \), previously used to explain NMR spin-relaxation data, and the experimental temperature dependence of the conductance and the temperature-dependent shape and width of the \( d^3I/dV^3–V \) curve.

Antiferromagnetic ordering occurs in La\(_2\)CuO\(_4\) and YBa\(_2\)Cu\(_3\)O\(_6\) when holes localised on adjacent Cu\(_{2+}\) 3d\(^9\) ions (in the CuO\(_2\) planes) take opposite spin. This minimises energy by reducing overlap of their wavefunctions. Fluctuations in this ordered state give rise to spin waves with quantised energy; the associated elementary excitations are called magnons. High-\( T_c \) superconductors with transition temperatures close to their maximum values are not antiferromagnetically ordered. Hole doping the insulating ‘parent’ compound (with Sr for La\(_2\)CuO\(_4\) and by increasing \( x \) in YBa\(_2\)Cu\(_3\)O\(_{6+x}\)) suppresses antiferromagnetism by ‘magnetic frustration’. At a critical density antiferromagnetism is completely suppressed and the system changes from an insulating to a superconducting metallic state (Gough 1991). Linear conductance will only be observed when tunnelling into antiferromagnetically ordered, hole-deficient material that can support spin fluctuations (Phillips 1989, p.137). This may explain why linear conductance backgrounds can occasionally be removed by cleaning the sample’s surface, since in this case all antiferromagnetic degraded material is removed.

In both high-\( T_c \) superconductor and Cr\(_2\)O\(_3\) junctions the linear-conductance region extends to energies higher than those of excitations that could participate in inelastic tunnelling. Multiple scattering processes were invoked to explain this apparent discrepancy, and a quantitative model was developed to show that the observed range
of linear conductance is possible. This model also accounts for asymmetry in the V-shaped conductance background.

Hasegawa et al. (1992a, p.60) comment that the inelastic tunnelling model described above cannot explain linear conductance in the bismuthate superconductors, since these materials are non-magnetic. Kirtley argues, however, that flat broad-band excitation spectra may be rather common in nature, and that other excitations may be present (e.g., charge-density waves) with the required energy distribution.

The linear relationship between transition temperature and slope of the linear conductance region seen by Sharifi et al. (4.1) could be explained by a superconducting mechanism that is somehow linked to the bosonic excitation spectrum, such that $T_c \propto F_c$. However, Kirtley does not discuss any possible links between his model and either this observation, or the relationship between conductance slope and zero-bias conductance (section B.1).

### 4.2.4 The Zeller–Giaever Model and Charging Effects

Zeller and Giaever (1969) investigated properties of evaporated Al–Al$_2$O$_3$–Al tunnel junctions with small Sn particles embedded in the oxide. The conductance increased linearly at low voltages but levelled off to a constant value when junction bias exceeded a saturation voltage, which depended on particle size. Their observations were explained by a model in which an isolated particle formed the central electrode of two tunnel junctions in series.

When the charging energy of a junction is greater than the thermal energy (i.e.,

![Fig. 4.6. Schematic voltage-biased double junction.](image-url)
The junction is said to be mesoscopic∗, and its conductance–voltage characteristics are dominated by single-electron tunnelling (SET) effects. The behaviour of two such junctions when coupled in series and driven by a voltage source depends on the electrostatic energy required to transfer charge to the inter-junction region.

Consider a double junction with normal-metal electrodes and a small inter-electrode particle (figure 4.6). With no applied voltage (i.e., equivalent to shorting the junction) the Fermi levels in these three regions will try to align by exchanging electrons. If the particle’s capacitance \( C = C_1 + C_2 \) is very small then discreteness of charge will be manifested as clearly quantised values of the particle’s voltage, and movement of its Fermi level can only occur in voltage steps of \( e/C \). Exact alignment of the central electrode’s Fermi level with the Fermi level of the outer two electrodes may be impossible and a residual polarisation voltage, \( V_p \), will then remain between the particle and electrodes. This voltage may be considered to arise from a fractional polarisation charge on the central electrode, \( \delta Q = CV_p \), which can take values from 0 to \( \pm e/2 \)†.

If an electron tunnels from L to M the charge on the middle electrode instantaneously changes from \( \delta Q \) to \( \delta Q - e \), the charge on junction 1 changes to \( Q_1 - e \) and the charge on junction 2 remains at \( Q_2 \). To keep the total voltage across the junctions constant, a charge of \( q' = -eC_2/(C_1 + C_2) \) is then moved from R to L through the source potential \( V \). This changes the charges on junctions 1 and 2 to \( Q_1 - e - q' \) and \( Q_2 - q' \) respectively. The change in energy due to work done in moving this charge and due to increased Coulomb energy on the central electrode is

\[
\delta E_1 = \frac{e}{C_1 + C_2} \left( \frac{e}{2} - \delta Q - C_2 V \right)
\]

∗ Although this meaning has been adopted by the literature (e.g., Wilkins et al. 1990a), ‘mesoscopic’ literally means ‘medium sized’. The isolated conducting particles are small enough for a change in charge of only \( \pm e \) to produce a large change in electrostatic energy (due to the small capacitance), but are large enough for their electron energies to be continuous in comparison with other energy scales. Quantum-mechanical effects can then be ignored and a semi-classical model can be used.

† The fractional charge is a ‘convenient fiction’ that is used to represent the initial misalignment of Fermi levels. \( \delta Q \) is really the difference between the charge that would reside on the central electrode at equilibrium if charge were infinitesimally divisible, and the equilibrium charge that actually occurs because it can only be changed by multiples of \( e \) (Adkins 1993).
The tunnelling event just described can only take place if \( V > (e/2 - \delta Q)/C_1 \), so that \( \delta E_1 < 0 \). Once the bias has been increased beyond this threshold, tunnelling may take place across junction 1 to place a single additional electron on the central electrode (changing the charge to \( \delta Q - e \)). In addition, tunnelling across junction 2 to remove this electron (returning the charge to \( \delta Q \)), is also allowed* so a current channel has opened. The same argument can be repeated for an initial tunnelling event across junction 2 that removes a single electron from the central electrode, changing the charge to \( \delta Q + e \). This can only occur for \( V > (e/2 + \delta Q)/C_1 \).

In summary, at positive biases tunnelling is suppressed for

\[
V < \min \left\{ \frac{(e/2 + \delta Q)}{C_1}, \frac{(e/2 - \delta Q)}{C_2} \right\}
\]

resulting in a region of zero conductance that is close to, but offset from, zero bias. This is called a Coulomb gap or blockade. This phenomenon is really just an example of two-step tunnelling via an intermediate state (section 3.6).

Above the threshold voltage, Zeller and Giaever assumed that current flows as in a normal junction, so there is a step increase in conductance (figure 4.7a). They then showed that suitable distributions of \( V_p \) and particle capacitance can lead to a smooth transition from linear conductance to constant conductance at high bias (figure 4.7c), as observed in their experiments†.

There have been several suggestions that small particles are present in high-temperature superconductor tunnel junctions and may lead to Zeller–Giaever-like charging effects (Kirtley et al. 1987a). This hypothesis is (superficially) supported by the granular composition of ceramic samples used in many experiments. The experimentally observed thermal broadening width in these materials, of \( 6k_B T \), is also consistent with this explanation; Medina et al. (1990) showed that tunnelling via a broad uniform distribution of capacitances gives a linear conductance near zero bias of

* By following a similar argument for this change in charge on the central electrode, one obtains the condition \( V > (\delta Q - e/2)/C_1 \). This must be satisfied since \( \delta Q < e/2 \).

† More recent experimental and theoretical work has shown that for a single inter-electrode particle there is a conductance peak at the threshold voltage (section 4.4), so figure 4.7a is incorrect. However, with many inter-electrode particles the distribution of capacitance and \( V_p \) values may again yield conductance–voltage characteristics resembling figure 4.7c.
Fig. 4.7. Conductance–voltage characteristics of a Zeller–Giaever tunnelling system: (a) single particle with fixed polarisation voltage, $V_p$, and capacitance, $C$; (b) assembly of particles with uniform distribution of $V_p$ and fixed $C$; (c) particles with uniform distribution of $V_p$ and Gaussian distribution of $C$. 
\[ \frac{\mathrm{d}I}{\mathrm{d}V} \propto \frac{V}{\tanh(eV/2k_B T)} \]

which has a bell-shaped $\frac{\mathrm{d}^3I}{\mathrm{d}V^3} - V$ curve with FWHM of 5.44 $k_B T_c$.

High saturation voltages of 100–500 mV indicate that the hypothetical particles must have very small capacitances. With reasonable assumptions about the dielectric constant, the inferred particle diameter is 1–10 nm - much smaller than physical grains in ceramic samples (= 1 μm). Observation of V-shaped conductances in single crystals also suggests that the granularity is somehow ‘built in’ (Moreland and Clark 1988). This size issue suggests that the picture of small isolated particles should be replaced by a more general one of microscopic localised states in the tunnelling barrier (Hasegawa et al. 1992a, p.63). In this case the Coulomb gap arises from correlation energy due to interactions between neighbouring localised states (section 3.6). Suitable distributions of trap locations in the barrier (analogous to $V_p$ in the Zeller–Giaever model) and correlation energies will again result in a linear conductance. This model is discussed further in section 4.8.

Several other pieces of evidence point away from the Zeller–Giaever model. It does not explain the parabolic or decreasing backgrounds that occur with other high-$T_c$ superconductors, the commonly observed asymmetry in the magnitude of the conductance slopes (Kirtley 1990) or the systematic dependence on junction resistance observed by Huang et al. (section 4.2.1). In addition, Barone et al. (1989) found that the zero-bias conductance of a high-$T_c$ planar junction increased gradually with temperature rather than linearly as predicted by the Zeller–Giaever model.

### 4.2.5 Intrinsic Causes for Linear Conductance

Although there are strong indications that the linear conductance background has an extrinsic origin, several intrinsic mechanisms have been proposed.

The V-shaped discontinuity at zero bias is reminiscent of the tunnelling characteristics of disordered or granular conductors close to the metal–insulator transition. In these systems, strong electron–electron interactions lead to breakdown of Fermi liquid theory and, on the metal side of the transition, the normal density of states is modified at the Fermi energy. In 3-dimensional systems

\[ N(E) - N(E_F) \propto |E - E_F|^{1/2} \]

so at $T = 0$

\[ G(V) = G(0)(1 + |V/V^*|^{1/2}) \]

where $G(V)$ is the tunnelling conductance and $V^*$ is a constant. In 2-dimensional systems the conductance is logarithmic in voltage (White et al. 1985). In either case
there is cusp-like discontinuity at zero bias which is thermally broadened at finite 
temperature. The broadening width is greater than $3.5k_B T$ due to temperature 
dependence of the correction to the density of states (Valles et al. 1991). It has been 
suggested that the cuprate’s linear conductance is also a manifestation of non-Fermi 
liquid behaviour arising from strong correlation effects (Valles and Dynes 1990). 
Sharifi et al (1991) argue that this cannot be so because $|V|^{1/2}$ or $\log(|V|)$ behaviour 
and scaling in strength of the background conductance with level of disorder does not 
occur in the oxide superconductors. However, in some systems (e.g., Au$_{1-x}$Ge,$_x$) the 
square-root voltage dependence seen well into the metallic state does not extend up to 
the M–I transition (Srikanth 1992). In these materials the tunnelling conductance can 
be written

$$G(V) = G(0) \left( 1 + \frac{|V|^n}{V^*} \right)$$

where $n$ changes from 0.5 to 1 as the transition is approached from the metallic side. 
Srikanth et al. (1992) observed this behaviour in the low-$T_c$ oxide superconductor 
Na$_{0.3}$WO$_3$ and also in several non-superconducting perovskite oxides. They also 
showed that the linear background conductance of Au–Bi$_2$Sr$_2$CaCu$_2$O$_8$ point-contact 
junctions could be made sub-linear ($n = 0.6$) by doping the Ca sites with Y. Srikanth 
(1992) believes that the linear conductance is a property of the normal state of the 
oxide superconductors, and that its origin is associated with the nearby metal–insulator 
transition, disorder–induced localisation and correlation effects.

The anomalous normal-state properties of high-$T_c$ superconductors have also lead to 
suggestions that these materials have a non-Fermi liquid ground state. Specifically: 
the $ab$-plane resistivity is linear to much lower temperatures than expected for a Fermi 
liquid*, $c$-axis conductivity is approximately linear in many cuprates, spin 
susceptibility and Hall coefficient are not temperature independent and the temperature 
dependence of specific heat, NMR relaxation rate and thermal conductivity are unlike 
those seen in other metals. Theoreticians also raise objections to a Fermi-liquid 
ground state based on a number of unusual properties of the cuprates, for example 
their low dimensionality, low carrier concentration and the (apparent) coexistence of 
antiferromagnetism and superconductivity. Resonating valence bond (RVB) theory is 
an exotic description of the cuprate’s ground state, largely motivated by these 
objections.

* The cuprates and bismuthates have Debye temperatures in the range 300–400 K, so linear resistivity 
extends well below the $T > 0.2 \theta_D$ range expected for a Fermi liquid. For example, linear resistance has 
been seen from 15 K to 700 K in low-$T_c$ (9 K) Bi$_{2+y}$Sr$_{2-x}$CuO$_{6+\delta}$ (Fiory et al. 1989).
In the RVB model, electrons in the CuO$_2$ planes are strongly correlated and pair up into singlet pairs. The ground state consists of the superposition of all possible electron pairs with the constraint that each site on the 2D lattice may host at most one electron. There are two types of quasiparticle excitation from this ground state; spinons arise from unpaired electrons and have spin $\frac{1}{2}$ but no charge, whilst holons are located at unoccupied sites and have charge $+e$ and no spin. The central theme of RVB theory is that the conduction mechanism (the charge degrees of freedom) is separated from magnetic fluctuations (the spin degrees of freedom) allowing metallic superconductivity and insulating antiferromagnetic behaviour to coexist$^\dagger$.

$^\dagger$ It is now believed that antiferromagnetism must be suppressed before superconductivity appears and that apparent coexistence of the two properties indicates a mixed-phase sample. This led Phillips (1989b, p.171) to comment that ‘The RVB idea represents an ingenious and resourceful approach to a problem that seems not to exist’.

Fig. 4.8. Conductance–voltage characteristics of a normal-metal–RVB tunnel junction. $V$ is the potential of the metal electrode with respect to the RVB electrode. (a) Both electrodes normal. (b) RVB electrode superconducting. The solid line shows the conductance when spinons are fermions and holons are bosons, the dashed line shows the conductance for ‘reversed’ statistics (Flensberg et al. 1989).
Three linear phenomena observed in the cuprate superconductors - tunnelling conductance, conductivity in the $c$-direction and resistivity in the $ab$-plane - were originally proclaimed as ‘unequivocal’ evidence in favour of RVB theory (Anderson and Zou 1988). However, further investigation (Flensberg et al. 1988) revealed that the tunnelling conductance of a normal-metal–normal-RVB junction would be linear for one bias polarity but constant for the other (figure 4.8a). After modification of the theory* to remove this discrepancy (figure 4.8b) RVB theory remained inconsistent with experimental results on SIN junctions, since the conductance peaks seen in all high-quality junctions are not predicted to occur (Flensberg and Hedegård 1989).

Although data from early PES experiments supported suggestions that the cuprates were not Fermi liquids, surface-preparation techniques have improved and yielded results consistent with a more conventional description. In addition, theories of Fermi liquids with strong correlations have been re-examined and it is now believed that the experimental observations are not totally incompatible with a Fermi-liquid ground state.

Comparisons have been made with other strongly correlated systems. For example, it has been reasonably well established that heavy-fermion metals (e.g., UPt$_3$) are Fermi liquids at $T = 0$, however, at modest temperatures (1–10 K) their f-electron states become increasingly localised in nature. This results in deviations from Fermi liquid behaviour that are similar to those seen in the oxide superconductors. Levin et al. (1991) suggest that the cuprates’ 3d electrons may also be ‘nearly localised’.

Varma et al. (1989) proposed a phenomenological modification to Fermi liquid theory that quantitatively explains many of the anomalous properties. Their marginal Fermi liquid theory results in an energy-dependent quasiparticle self-energy,

$$\Sigma(\omega) = \lambda [\omega \log(\frac{\omega}{\omega_c}) + i\pi x]$$

(4.9)

for $x \leq \omega_c$, where $\omega_c$ is a high-energy cut-off, $\lambda$ is a coupling constant and $x = \max(\omega,T)$. The important point for tunnelling is that the quasiparticle scattering rate, $\text{Im}(\Sigma)$, is proportional to energy. Although this anomalous self-energy does not modify the density of states, it does affect the tunnelling conductance through momentum dependence of the tunnelling matrix element.

Littlewood and Varma (1992) showed that the contribution to tunnelling conductance

* By reversing the statistics to make holons (charge $+e$, no spin) be fermions and spinons (no charge, spin $-1/2$) be bosons. This allocation is valid because these are composite particles immersed in a many-body background (Gunn 1988).
from states with momentum $k_F$ results in a (nearly) voltage-independent conductance, $dI/dV \approx g_0$, where

$$g_0 \propto \lambda$$

(4.10)

In contrast, the contribution from states away from $k_F$ is proportional to the quasiparticle scattering rate, so a tunnel junction that preferentially selects such states will have a V-shaped conductance–voltage characteristic. For this type of tunnelling Littlewood and Varma give the expression

$$\frac{dI(V)}{dV} = g_0 \frac{|eV|}{\hbar \omega_c}$$

(4.11)

The overall conductance may be written

$$dI/dV = g_0 \left(1 + \alpha |V|\right)$$

(4.12)

where $\alpha$ is a parameter that depends on the distribution of $k$-states sampled by the junction. Full numerical modelling showed that $\alpha$ actually differs by a factor of about two between positive and negative voltages. In addition, temperature dependence of the self-energy (4.9) adds to the Fermi function’s thermal smearing, so the linear cusp at zero bias is broadened with an FWHM larger than $3.5 k_B T^*$. Another important prediction of this model is that the conductance slope is proportional to the zero-bias conductance. These three details are in agreement with experimental observations (section 4.2.1).

Littlewood and Varma also argued that the ‘strength’ of the V-shaped conductance will be sensitive to the orientation of a cuprate tunnel junction. As discussed in section 4.3.2 (with reference to figure 4.12), if the distribution of tunnelling angles is strongly weighted to the junction normal, $c$-axis tunnelling selects states with $k_T = (k_x^2 + k_y^2)^{1/2} \approx 0$. These lie away from $k_F$ so the conductance–voltage characteristics will be V shaped. Increasing the barrier height or decreasing its width broadens the distribution of tunnelling angles (figure 4.10), leading to greater contribution from states near $k_F$ and hence a flatter conductance–voltage characteristic. This is in qualitative agreement with the results of Huang et al. discussed in section 4.2.1. Tunnelling in the $ab$-plane will also be dominated by states near the Fermi surface, resulting in a very flat conductance–voltage characteristic.

The predicted orientation dependence of the background agrees with observations made by Mandrus et al. (section 4.3.2). It may also explain the occasional elimination

* The FWHM estimated from figure 1 of Littlewood and Varma (1992) is $4.5 k_B T$. 
of a background by cleaving samples, since this may be required to form a well defined interface with ideal ab-plane tunnelling. Littlewood and Varma state that a ‘clean’ gap-like feature will only be seen in the contribution from $k_F$-states, so observations of a clear BCS-like gap should be correlated with flatter backgrounds. This does seem to be supported by the literature; however, it may just be that gaps are easier to see when the background is flatter.

Observation of linear conductance in the cubic bismuthate materials can also be attributed to marginal Fermi liquid behaviour. In this case barrier transmission peaked at states away from $k_F$ could arise through details of the band structure. Littlewood and Varma claim that correlation between the slope of the linear-conductance region and $T_c$, observed in these materials by Sharifi et al., is evidence that the anomalous self-energy (4.9) and superconducting pairing interaction share a common mechanism. They draw this conclusion by noting that the conductance derivative increases with $\lambda$ through (4.10) and (4.11), and, presumably, are suggesting that if $T_c \propto \lambda$ the experimental observations are reproduced. However, this is a misinterpretation of Sharifi’s results, which actually show a linear relationship between $T_c$ and the slope of the conductance normalised by the zero-bias conductance. This indicates that $T_c$ dependence of the slope occurs in $\alpha$ in (4.12), not $g_0$ (see appendix B).

4.3 Broadening

The tunnelling characteristics of junctions containing high-$T_c$ superconductors are invariably broadened in voltage compared with the predictions of BCS theory and have greater conductance inside the gap region. This broadening is seen even on the rare occasions when it has been possible to ‘remove’ the voltage-dependent conductance background by cleaving the sample at low temperature. Although broadening is always present, its magnitude varies from sample to sample and from point to point on the surface of a single sample when examined with a point contact or STM. These junction types generally display more BCS-like characteristics than planar or break junctions and have occasionally resulted in zero sub-gap conductance. This section will discuss several possible explanations for the broadening.

One explanation for the broadening is that the total conductance is an integral over many tunnelling events in which different electrons sample different energy gaps, either because the energy gap is inhomogeneous ($\Delta(r) \neq \text{constant}$) or because it is anisotropic ($\Delta(k) \neq \text{constant}$). An excess sub-gap conductance could then be attributed to normal regions with $\Delta(r) = 0$, or the presence of states at some points on the Fermi surface (i.e., $\Delta(k) = 0$ for some $k$).
4.3.1 Inhomogeneous Gap

For broadening to appear in the conductance–voltage characteristics the energy gap must vary on a scale smaller than the volume sampled by the tunnelling particles. The *temperature-dependent coherence length*, $\xi(T)$, was introduced in Ginzburg-Landau theory as the shortest distance over which the order parameter may change*. At low temperature this is identical to the intrinsic coherence length,

$$\xi_0 = 0.18 \frac{\hbar v_F}{k_B T_c}$$

(from 2.13 and 2.22). In conventional superconductors $\xi_0$ is typically 0.1–1.0 μm but it is much smaller in high-$T_c$ superconductors: 1–3 nm in the $ab$-plane and 0.2–0.5 nm in the $c$-direction (Worthington et al. 1987, Batlogg et al. 1988). This is due partly to the higher $T_c$, but mostly to the Fermi velocity, which is two orders of magnitude smaller (Wolf and Kresin 1991).

Because the area of any planar junction will be much greater than the coherence length of a high-temperature superconductor, inhomogeneity is likely to lead to considerably broadened tunnelling characteristics. Decreasing the junction area will reduce the broadening, so one would expect to see improved characteristics from point contacts and edge junctions (sections 6.1.1–2). The smallest effective junction area is obtained by using an STM tunnel junction with the tip separated from the sample by a vacuum gap, $d$. The theory of Tersoff and Hamann (1985) predicts wavefunction decay lengths of $0.5\kappa^{-1}$ into the surface and $\sqrt{0.5\kappa^{-1}(R+d)}$ laterally, where $R$ is the tip radius, $\kappa = h^{-1} \sqrt{2m\phi}$, $\phi$ is the workfunction and $m$ is the mass of the tunnelling particle. Using $\phi = 1–5$ eV, $R = 0.9$ nm, $d = 1$ nm and $m = m_e$ (Burns 1992, p.106) the

---

* In Ginzburg–Landau (GL) theory the density $n_s$ of superconducting electrons is given by the squared amplitude of a pseudo-wavefunction order parameter, $\psi$. The behaviour of $\psi$ is described by a differential equation obtained by minimising the Gibbs free energy of the superconducting state expressed in terms of $\psi$ and a vector potential $A$. The distance over which the order parameter may vary without undue increase in energy is constrained to be greater than a characteristic *temperature-dependent coherence length*, $\xi(T)$, due to a kinetic term in the free energy, which increases with the gradient $\nabla \psi$. Although GL theory was introduced several years before BCS theory and was phenomenological in nature Gor’kov showed that it is a limiting form of the microscopic theory, valid near $T_c$, and that the order parameter is directly proportional to the energy gap, $\psi \propto \Delta$.

† $\xi_0$ is estimated from measurements of the upper critical field, $H_{c2}(T)$, near $T_c$ via $H_{c2}(0) = \Phi_0 / 2\pi \xi_0^2$ (Burns 1992, p.154).
decay lengths are 0.05–0.1 nm and 0.3–0.5 nm respectively. These values are slightly smaller than $\xi_0$ so suggest that broadening could just be eliminated.

Inhomogeneity of the energy gap has been investigated in a number of STM experiments. Ramos and Vieira (1989) produced a histogram of gap values found at different locations on the surface of a ceramic sample. Volodin and Khaikin (1987) used an STM to ‘map’ the energy gap over the surface of a single crystallite of ceramic YBa$_2$Cu$_3$O$_y$. They observed a smooth variation in $2\Delta/k_B T_c$ from 2.4 to 7 over most of the area examined and a rapid drop to $\Delta = 0$ in one normal region. Aleksandrov et al. (1989) found that 40% of the surface of a YBa$_2$Cu$_3$O$_y$ single crystal was normal. Tunnel current from these regions could be the origin of excess sub-gap conductance commonly observed with planar and large-area point-contact junctions (Gijs et al. 1988).

A distribution of energy gap values may be incorporated into models for the tunnelling characteristics of an SIN junction simply by replacing the tunnelling density of states for a single energy gap, $n_s(E;\Delta)$, by

$$n_d(E) = \int_0^\infty g(\Delta)n_s(E;\Delta) d\Delta$$

(4.13)

where $g(\Delta)$ is the normalised distribution function. In practice $g(\Delta)$ is unknown so a Gaussian distribution is used (Kirtley et al. 1988):

$$g(\Delta;\bar{\Delta},\delta\Delta) = \frac{e^{-(\Delta-\bar{\Delta})^2/2\delta\Delta^2}}{\delta\Delta\sqrt{\pi}}$$

(4.14)

Fitting model tunnelling characteristics to the data then gives an estimate of the mean gap $\bar{\Delta}$ sampled by the junction and the spread $\delta\Delta$ of gap values. This quantifies the level of broadening and indicates the sample and junction quality.

One obvious source of energy-gap inhomogeneity is inhomogeneity in chemical composition of the high-temperature superconductor. The properties of these materials are sensitive to their composition - the complexity of which makes preparation of pure samples exceedingly difficulty. Moreover, the surfaces of high-$T_c$ superconductors are frequently degraded and may be normal, semiconducting or insulating (section 4.5). Clearly, an irregularly depressed surface energy gap could give rise to the observed broadening. Edgar et al. (1987) suggested that surface material is only proximity coupled to the underlying bulk superconductor and that variation in surface-layer thickness over the contact area will lead to a distribution of sampled gaps with a cut-off at the true gap value - corresponding to strongly thinned layers. The high-energy ‘tail’ in their experimental gap distribution was attributed to
uncontrolled potential drops across resistive grain boundaries in the ceramic sample. Yoshikawa and Sugahara (1991) claimed that even a uniform proximity-coupled normal layer has extremely smeared tunnelling conductance, and fitted theoretical calculations to Kirtley’s experimental results. Their model, however, assumes that the superconductor is in the dirty limit and applies to temperatures just below $T_c$, so its validity is questionable.

Gap inhomogeneity could be extrinsically introduced by the tunnelling experiment. The magnetic field due to tunnelling particles ($< 10^{-5}$ T) is too small to have any noticeable effect on the energy gap of a high-temperature superconductor. However, at about at $10^{10}$ A cm$^{-2}$ the velocity of electron pairs, $v_s = J_s/(2en_s)$, will exceed the *depairing velocity*, $|\Delta(0)|/\hbar k_F$, resulting in a reduced energy gap and possible broadening (Pan *et al*. 1987). If the energy gap is pressure dependent and there is a non-uniform pressure over the point-contact junction area then the conductance–voltage characteristics will be broadened. However, van Bentum *et al*. (1989) investigated this effect in $\text{YBa}_2\text{Cu}_3\text{O}_y$ and $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4-y$ junctions, concluding that it was significant only in the latter case.

The $c$-direction coherence length of a high-$T_c$ superconductor is smaller than the lattice parameter and approximately equals the inter-layer spacing. An intriguing possibility is that $\Delta$ varies from layer to layer within the unit cell so is *intrinsically* inhomogeneous (Briceno and Zettl 1989). Some evidence for this has been obtained by using an STM to tunnel (in the $ab$-plane) into individual atomic layers of a $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ crystal at 4.2 K (Hasegawa *et al*. 1991). As the tip was moved along the $c$-axis they observed alternating metallic and semiconducting tunnelling characteristics. These were assigned to the $\text{CuO}_2$ and $\text{BiO}$ layers respectively because the semiconducting characteristics resembled those obtained when tunnelling (in the $c$-direction with large tip–sample separation) into the $\text{BiO}$ layer exposed by cleaving*. This evidence is contradicted by data from angle-resolved photoemission spectroscopy (ARPES), which indicated that superconductivity extends throughout the unit cell (Wells *et al*. 1990). These conflicting results may be due to a difference in oxygen content of crystals used in these experiments; the semiconducting as-cleaved surface may be made metallic by annealing in oxygen (Wei *et al*. 1991).

Several theoretical models of a layered superconductor have recently been investigated. These model the cuprates as an intrinsic superlattice of superconducting $\text{CuO}_2$

---

* The same technique was recently used to show that $\text{CuO}_2$ layers in $\text{YBa}_2\text{Cu}_3\text{O}_y$ are superconducting whilst the $\text{CuO}$ ‘chain layers’ are metallic (Tanaka *et al*. 1993). However, these conclusions are more uncertain because atomic resolution was not achieved.
layers, separated by normal and insulating planes and weakly coupled together by single-particle tunnelling. For example, Takahashi and Tachiki (1990a,b) assumed that the pairing interaction operated only between oxygen p-orbital holes in the CuO$_2$ layer. Relative strengths of transfer integrals between oxygen sites were estimated (based on an assumption that these were weakest between layers) and the local density of states in each layer was calculated (figure 4.9). Proximity coupling induces small gaps in the intrinsically normal layers (BaO and CuO in YBa$_2$Cu$_3$O$_y$ and BiO in Bi-Sr-Ca-Cu-O) and broadens the density of states in the CuO$_2$ layer to give a finite number of states in the gap region. The 0 K conductance of an SIN junction formed with the CuO$_2$ layer is approximately proportional to the density of states. This is not true of the normal layers, which may yield complex, multiple-peaked, V-shaped characteristics; however, the probability of tunnelling into these layers is considerably smaller (Abrikosov 1992). At finite temperature the CuO$_2$ layer tunnelling conductance is rapidly smeared, resulting in characteristics resembling a highly broadened BCS result with large zero-bias conductance (Frick and Schneider 1992).

Chang et al. (1992) claim to have indirectly observed an internal proximity effect in Bi$_2$Sr$_2$CaCu$_2$O$_y$. Using a vacuum-tunnelling STM at 4.2 K, large spatial variation was seen in the conductance–voltage characteristics of the BiO surface exposed by cleaving. Regions with stronger than average gap-like overshoots but smaller than average energy gap occurred at the same location as apparent peaks in the surface topography, even though these regions were believed to be atomically flat. They proposed that this behaviour is due to a locally increased ‘metallicity’ (i.e., Fermi-level

---

**Fig. 4.9.** Local density of states at each oxygen site in YBa$_2$Cu$_3$O$_y$ according to the layered model of Takahashi and Tachiki (1991). O(2,3) refers to sites in the CuO$_2$ layer, O(1) the CuO layer and O(4) the BaO layer.
density of states) in the surface BiO layer (possibly because of increased oxygen content) and that in such regions tunnelling probes superconductivity induced in the BiO by an internal proximity effect from the underlying CuO$_2$ layer. The associated topographic structure occurs because tunnelling current increases due to the larger density of states, $N(E_F)$, even though the vacuum gap and work function are constant. Where the BiO remains insulating, tunnelling takes place directly into the CuO$_2$ plane so the observed gap is larger but less clear.

4.3.2 Anisotropic Gap

Simplifying assumptions made in BCS theory result in an isotropic energy gap (2.12). However, if the electron-pairing interaction is $k$ dependent or the sum over $k'$ in the self-consistent gap equation (2.11) is anisotropic (e.g., due to a non-spherical Fermi surface) then the energy gap will also depend on $k$. Many properties of the high-temperature superconductors are highly anisotropic, including resistivity (Tozer et al. 1987), critical current (Dinger et al. 1987), penetration depth (Burns 1992, p.156), the critical fields (Koike et al. 1988), coherence length (Worthington et al. 1987) and possibly the proximity effect (Lee et al. 1990). The two-dimensional nature of these materials suggests that the energy gap will also be anisotropic.

Tunnelling appears to be the ideal technique with which to study gap anisotropy. One simply needs specular tunnelling, in which the transverse momenta of tunnelling particles is conserved, in a series of well defined directions on a single-crystal or oriented thin-film sample (Solymar 1972, p.90). In practice, however, a non-ideal interface may perturb the nominal tunnelling direction. For example, transverse momenta are not conserved if the interface is rough or disordered. Tunnelling is then said to be diffuse and the conductance depends less on the orientation of the interface.

Anisotropy studies of high-temperature superconductors are particularly difficult due to the non-conducting surface layer. Point-contacts and STM tips must be pushed through this layer before a tunnelling current is detected, disturbing the crystal structure and rendering the actual tunnelling direction unknown (Kirtley et al. 1987b). Attempts to remove this layer by chemical etching results in numerous etch pits, the walls of which provide alternative, uncontrolled tunnelling directions (Gurvitch et al. 1989). Indeed, it was suggested that tunnelling characteristics were improved because etching permits ‘sideways’ tunnelling into the $ab$-plane where the coherence length is longer. The growth mechanism of thin films also produces large pits and terraced hill-like structures and hence an intrinsically rough surface (Lang et al. 1992). The inevitable multi-directional tunnelling may be the source of the broadening and multiple gap features observed in thin-film junctions (Beasley 1991).
Beuermann (1981) showed that even in an ideal junction the usual assumption of strong directional selectivity cannot be taken for granted* and that significant current may arise from angles as large as 30–40° away from the junction normal (figure 4.10). Similar results were obtained in calculations of the current distribution in an STM - with implications for the interpretation of topographic images (van de Leemput and van Kempen 1992, sec.3.1.2).

Despite the difficulties outlined above, evidence has emerged of anisotropy in the energy gap of these materials. In addition to broadened tunnelling characteristics and observation of multiple (usually two) gap-like features in a single conductance–voltage characteristic (see, e.g., Gurvitch et al. 1989), both of which may be explained by anisotropy or inhomogeneity, there have been several attempts to compare the tunnelling characteristics of $c$-direction junctions with $ab$-plane oriented junctions.

Many results show that a gap may be observed in the $ab$-plane but not in the $c$-direction (see, e.g., Boguslavskii and Rudenko 1989). This may be due to genuine lack of a gap; however, the smaller coherence length in this direction may simply be

---

* The conductance is weighted by a distribution function of tunnelling angles, $f(\varphi) = c T(\varphi) \sin \varphi$, where $c$ is the normalisation constant. Directional selectivity is impaired by the $\sin \varphi$ term and because the transmission probability $T(\varphi)$ (Merzbacher 1970, p.93) is not as strongly peaked at $\varphi = 0$ as is usually assumed.
preventing formation of a suitable tunnel junction. Mandrus et al. (1991) reported measurements of ‘giant tunnelling anisotropy’ in single crystal Bi$_2$Sr$_2$CaCu$_2$O$_8$ break-junctions. With $T < T_c$ a BCS-like gap was seen for tunnelling parallel to the CuO$_2$ layers, with large conductance peaks at ± 55 mV. Above $T_c$ the conductance was almost constant. In contrast ‘no well defined structure’ was seen in c-axis tunnelling characteristics, which were modelled, above and below $T_c$, by a phenomenological expression

$$\frac{dI}{dV} = \gamma \log(V^2 + \alpha T^2) - \beta$$

with $\alpha$, $\beta$ and $\gamma$ as fitting parameters. Mandrus et al. do not attempt to wrest a gap value from the non-ideal c-axis tunnelling characteristics. Other studies, particularly those using planar junctions, are less careful with interpretation of their data and infer gap values from the slightest hint of structure (e.g., Iguchi and Wen 1991). The most convincing gap-anisotropy results have come from a study in which oriented YBa$_2$Cu$_3$O$_y$ thin films were broken in liquid helium and a Pb tip pressed against the freshly exposed film edge (Tsai et al. 1989a). The intrinsic insulating surface layer on the top face ensured that tunnelling occurred only in the plane of the film. Despite wide variation in $T_c$ values (determined from temperature dependence of the observed gaps) $2\Delta(0)/k_B T_c$ was consistently found to be 5.9 ± 0.2 parallel to CuO$_2$ layers and 3.6 ± 0.2 in the c-direction. Raman light-scattering experiments on YBa$_2$Cu$_3$O$_y$ single crystals gave similar values but also suggested that $\Delta(k) = 0$ for some $k$-values (Cooper et al. 1988).

There have been several theoretical studies of the effects of gap anisotropy on the tunnelling characteristics of high-temperature superconductors. Kirtley (1990) crudely calculated the conductance–voltage characteristics of SIN junctions with gap anisotropy,

$$\Delta(\hat{k}) = \Delta_{ab} \sin^2 \theta + \Delta_c \cos^2 \theta$$

where $\cos \theta = \hat{k} \cdot \hat{c}$, for several junction orientations and for an un-weighted average over all $\theta$. The latter characteristics were broad, with the onset of conduction occurring at the smallest gap, $\Delta_c$, and the conductance overshoot occurring at the larger of the gaps, $\Delta_{ab}$. This model was successfully fitted to their (embedded) point-contact results on single-crystal YBa$_2$Cu$_3$O$_{y}$ and La$_{2-x}$Sr$_x$CuO$_{4-y}$.

Maekawa et al. (1987) showed that an anisotropic gap leads to broadening in both the tunnelling conductance and far infra-red (FIR) absorption spectra, $\sigma(\omega)$, of polycrystalline samples. As above, the conductance maximum occurs at $eV = \text{max}(\Delta(k))$ but the peak in $d\sigma/d\omega$ occurs at lower energies, leading to smaller estimates of the gap from FIR measurements. This explains the observed discrepancy between gap estimates from these techniques in ceramic samples.
The most complete theoretical investigation into the origin and effects of gap anisotropy in layered copper-oxide superconductors has been carried out Schneider et al. (Schneider et al. 1989, Frick and Schneider 1990a,b, Schneider and Sörensen 1990). The tight-binding approximation with intralayer and interlayer matrix elements was used to derive the band structure and Fermi surface (Kittel 1986, 229; Burns 1992, p.97). The anisotropic gap, transition temperature, density of states and tunnelling characteristics for both diffuse and specular tunnelling were then calculated for two types of electron-pairing interaction. The resulting gaps were

\[ \Delta(k) = \Delta_0 + 2\Delta_1 \cos(k_z s) \]  

(4.15)

and

\[ \Delta(k) = \Delta_0 + 2\Delta_2 (\cos k_x a + \cos k_y a) \]  

(4.16)

respectively, where \( a \) is the lattice parameter of the CuO\(_2\) layers and \( s \) is the spacing between these layers. These anisotropic gaps cause the density of states to differ markedly from the BCS case, with a smooth rise from zero at the minimum gap and a van Hove singularity at the maximum gap (figures 4.11c and 4.13b). Both diffuse and specular tunnelling characteristics were broadened considerably in comparison with the BCS result. The former is approximately proportional to the density of states and depends only slightly on tunnelling direction, whilst specular tunnelling at low temperature was shown to be very sensitive to the junction orientation.

In contradiction to experimental observations, ‘uniaxial’ gap anisotropy given by (4.15) was shown to result in \( c \)-direction tunnelling with a conductance peak at energies greater than or equal to the conductance peak’s energy for \( ab \)-plane tunnelling (figure 4.11). This can be seen by considering the \( k \)-states that contribute to the tunnelling current. If the distribution function of tunnelling angles is strongly peaked at the junction normal, \( c \)-axis tunnelling is dominated by electrons with \( k_T = (k_x + k_y) \sqrt{2} = 0 \) and the observed gap will be determined by states labelled 1 and 2 in figure 4.12. Weak dispersion in the \( c \)-direction means that all \( k_z \) values, and hence all \( \Delta(k) \), will be involved so the gap-like feature is broadened with a peak near the maximum value. Tunnelling in the \( ab \)-plane, however, will involve states at the Fermi surface with \( k_T = k_z = 0 \). This is a stronger selection rule and results in a sharp gap at either the maximum or minimum value of \( \Delta(k) \) (arising from points 3 and 4), depending on the sign of \( \Delta_1 \).

Schneider et al. showed that in-plane anisotropy given by (4.16) does qualitatively reproduce the observed experimental behaviour (figure 4.13) and, in addition, showed it to be consistent with the results of selected ARPES experiments on Bi\(_2\)Sr\(_2\)CaCu\(_2\)O\(_8^*\)*.  

* Other ARPES experiments have indicated that the gap is isotropic (Burns 1992, p.123).
Fig. 4.11. Sketched normalised conductance–voltage characteristics (at $T = 0$) of an SIN tunnel junction with gap anisotropy (4.15). Solid lines represent tunnelling parallel to the layers, i.e., in the $ab$-plane, dashed lines represent tunnelling perpendicular to the layers, i.e., in the $c$-direction. (a) Specular tunnelling with $\Delta_1 < 0$. (b) Specular tunnelling with $\Delta_1 > 0$. (c) Diffuse tunnelling. The solid curve in (c) is very similar to the superconductor’s density of states.
**Fig. 4.12.** Sketch of the Fermi surface from the model of Schneider *et al.* with gap (4.16). The solid curve is for $k_z s = 0$, and the dashed curve for $k_z s = \pi$, i.e., the cross section varies periodically with $k_z$, so the Fermi surface resembles a corrugated column. The Cartesian axes, $x, y$ and $z$ are parallel to the crystal axes $a, b$ and $c$ respectively.

**Fig. 4.13.** Conductance–voltage characteristics of an SIN junction with gap anisotropy (4.16): (a) specular and (b) diffuse tunnelling in the $y$-direction i.e., in the $ab$-plane (solid lines) and in the $c$-direction (dashed lines). Minimum and maximum gaps originate from points 1 and 4 on the Fermi surface (figure 4.12) respectively. The solid curve in (b) is similar to the density of states and shows features due to van Hove singularities arising from points 2 and 3.
4.3.3 Lifetime Broadening

There have been several other explanations for the broadened tunnelling characteristics of high-temperature superconductors. Most prominent has been the lifetime-broadening model of Dynes et al. (1984), first used to account for broadened characteristics of Al–Al₂O₃–(granular)Al tunnel junctions, with the granular film close to the metal–insulator transition. The tunnelling density of states of the granular aluminium was accurately described by a density of states,

\[
n_{s}(\xi, \Gamma) = \text{Re}\left[ \frac{\xi - i\Gamma}{\sqrt{\xi^2 - 4\Delta^2}} \right]
\]

The finite linewidth, \(\Gamma\), is due to reduced quasiparticle lifetime, \(\tau\),

\[
\Gamma = \frac{\hbar}{\tau} = \hbar \sum_{m} \frac{1}{\tau_{m}}
\]

and the \(\tau_{m}\) arise from individual lifetime-shortening processes. Dynes et al. proposed that the limiting mechanism in granular aluminium junctions was inelastic electron–phonon scattering.

Dynes’ modified density of states has frequently been invoked to explain high-\(T_c\) superconductor tunnelling characteristics (see, e.g., Hinks et al. 1989). Zasadzinski et al. (1992) point out that this is not because there is a clear theoretical basis for its use, but because it conveniently allows the gap (and an estimate of the broadening) to be extracted from problematic data. Srikanth (1992) comments that the large \(G/D\) ratios that are often required to fit the data call into question the physical meaning of the model. Nevertheless, Dynes’ model was combined with BTK theory to explain broadening in ‘leaky’ junctions with \(Z \approx 0.5–1.0\) (Srikanth and Raychaudhuri 1992).

Hasegawa et al. (1992a) comment that \(G\)-values reported in the literature seem to be decreasing - pointing to an extrinsic origin for broadening, such as sample inhomogeneity.

Wolf et al. (1991) showed that broadening in planar tunnel junctions fabricated on cleaved Bi₂Sr₂CaCu₂O₈₊\(y\) crystals is given by

\[
\Gamma(T) = a + b\left(\frac{T}{T_c}\right)^3
\]

with \(a = 15\) meV and \(b = 2.3 \ k_B T_c\) meV. The temperature-dependent term is in agreement with the model of pair breaking by spin fluctuations proposed by Coffey (1990) to explain anomalous nuclear relaxation rates in YBa₂Cu₃O₇.

Several authors report that the lifetime-broadening model predicts excess sub-gap
conductance and that the data are more accurately described by a Gaussian distribution of energy gaps (see, e.g., Kirtley et al. 1988). For the same reason Zhao et al. (1988) use the normalised tunnel conductance of a strong-coupled superconductor (2.23) with energy-independent gap, $\Delta = \Delta_1 + i \Delta_2$.

### 4.4 Charging Effects: Coulomb Gaps and Staircases

Periodic multiple conductance peaks seen in point-contact experiments on high-temperature superconductors were initially interpreted as evidence for a multiple energy gap. Theoretical investigations demonstrated that multiple singularities could occur in the normal density of states and tunnelling conductance of low-dimensional superconductors with an unretarded pairing interaction (Que and Kirczenow 1988, Frick and Schneider 1990a). Further experiments, however, revealed that the peak structure may change if the tip position is adjusted or vary in detail with time with no intentional tip movement (Wan et al. 1990b). Barner and Ruggiero (1987) observed similar structure due to charging of small, isolated, metallic grains embedded in the barrier of an MIM junction and suggested that particles were also present between point-contact tips and high-$T_c$ superconductor samples.

Section 4.2.4 considered Zeller and Giaever’s semi-classical model of two mesoscopic MIM tunnel junctions in series (an MMM or NNN junction). When driven by a voltage source, a ‘Coulomb gap’ occurs below the threshold voltage $V_t = \min(e/2C_1, e/2C_2)^*$

*Assuming $\delta Q$ is zero.*
and no current can flow. At this voltage the central electrode may be charged by one more electron and a current channel opens.

As the applied bias is further increased the number of electrons that may be accommodated on the central electrode periodically increments. By numerically simulating a double junction (using stochastic calculations), Mullen et al. (1988a) showed that if the junction is asymmetric (e.g., $R_1C_1 \gg R_2C_2$) then sharp current steps occur at each threshold voltage. The resulting $I$–$V$ curve has steps at uniformly spaced voltages,

$$V_{s,k} = \pm (2k - 1)e/2C_m$$

where $C_m = \max(C_1, C_2)$ and $k$ is the set of positive integers*. This striking $I$–$V$ characteristic is often referred to as a Coulomb staircase (figure 4.15d). These simulations also showed that a double junction with similar resistances or capacitances has strongly smeared current steps, so only the Coulomb gap is clear (figure 4.15c).

Modelling NNN junctions with arbitrary parameters is considerably simplified by use of the analytical expression presented by Laikhtman (1991)†.

Wilkins et al. (1990b) showed that the multiple peaked conductance–voltage characteristics observed with YBa$_2$Cu$_3$O$_y$ point-contact junctions could be successfully fitted by these semi-classical charging-effect models. They suggested that an isolated normal-metal grain was present on the surface of the sample forming two junctions in series; the first between the STM tip and the grain, the second between the grain and the bulk sample. By adjusting the position of the STM tip, and hence the capacitance of that junction, van Bentum et al. (1988b) confirmed that the peak separation was inversely proportional to $C$. The theory was also tested using normal-metals in a planar array of parallel junctions (Ruggiero and Barner 1991) and in the STM configuration (Wilkins et al. 1989). Coulomb staircase structure has recently been observed at room temperature using granular NbN STM tips and films (Reeve et al. 1992).

In both the numerical simulations and the analytical model the opening of an additional current channel results in a much sharper step than is observed experimentally. The experimental broadening is much larger than thermal broadening alone and was attributed by Mullen et al. (1988b) to averaging over a number of double junctions with a distribution of parameters.

* i.e., $k \in \mathbb{Z}^+$, where $\mathbb{Z}^+$ is standard notation for the set $(1, 2, 3, \ldots)$.

† See also Amman et al. (1991) for a method that may be extended to an arbitrary junction configuration, e.g., SSN.
Fig. 4.15. (a) Measured Coulomb staircase from a W–YBa$_2$Cu$_3$O$_y$ point contact (see figure 7.33). (b) Laikhtman’s model $I$–$V$ characteristics with $R_1 = 4059 \, \Omega$, $C_1 = 3.1$ aF, $R_2 = 974 \, \Omega$, $C_2 = 0.4$ aF. These parameters were obtained by least-squares fitting to curve $a$. (c) as curve $b$ but with identical capacitances, $C = 3.1$ aF (a similar curve results from using identical resistances). (d) as curve $b$ but with very different resistances, $R_1 = 4059 \, \Omega$ and $R_2 = 10 \, \Omega$. Curves $a$–$c$ are offset for clarity.
Mullen et al. (1988b) extended the stochastic calculations to two series SIS junctions (SSS configuration) and showed that current steps occur for a wider range of junction parameters; even identical junctions exhibit multiple conductance peaks. In this case, the first step occurs at $4\Delta/e + e/2C$ and subsequent steps are separated by $e/C$. When the junctions are not identical two series of steps occur at

$$V_{s,k} = \pm \left[ (2k-1)e/2C_j + 2\Delta(T)(C_1 + C_2)/(eC_j) \right]$$

where $k \in \mathbb{Z}^+$ and $j = 1$ and 2. If the junction time constants are very different the slower junction’s steps will dominate; however, step height is in general a complicated function of $V_j$, $C_j$ and $R_j$ and may be zero - so steps may be ‘missing’. This behaviour has been verified in series junctions formed between an STM tip and an isolated grain in a Pb film (McGreer et al. 1989).

$I$–$V$ characteristics of other junction configurations were modelled by Barner et al. (1989). Symmetrical SNS and NSN junctions were found to have almost identical characteristics to those of SSS junctions but with half the energy gap:

$$V_{s,k} = \pm \left[ (2k-1)e/2C_j + \Delta(T)(C_1 + C_2)/(eC_j) \right]$$

This also gives the position of the current steps of an asymmetric NSS junction; however, the step heights in this case are different; in general, those associated with the SIS junction are sharper than those of the SIN junction. Experimental tunnelling characteristics of planar Cu/Al$_2$O$_3$/Pb/Al$_2$O$_3$/Cu NSN junctions were found to be consistent with the charging model.

The SET structure is usually offset from zero bias. This may be explained by a non-zero value of fractional charge, $\delta Q$, in (4.8). If the offset is close to $\pm e/2C$, the step that has been shifted close to zero bias is too small to be seen and the central gap appears twice as large as expected. $\delta Q$ has been observed to jump uncontrollably between traces from 0 to $\pm e/2$, leading to an apparently bistable central step width (van Bentum et al. 1988b). Changes also occur during a voltage sweep resulting in anomalously short or long steps (McGreer et al. 1989). $\delta Q$, and hence the offset, has been controlled in experiments on lithographic junctions by applying an electric field to the central electrode (Fulton and Dolan 1987) and in an STM junction by changing the tip–grain spacing (Rong et al. 1992). The polarisation may also be instantly removed by shining light on the junction (Wilkins et al. 1990b).

The SET effects described above seriously complicate interpretation of high-temperature superconductor tunnelling characteristics. For example, if the junctions are similar only the first step may be clear and the central Coulomb gap may be mistaken for a large superconducting energy gap (van Bentum et al. 1988b).
4.5 Surface Layers

The coherence lengths of high-\(T_c\) superconductors are extremely small (section 4.3.1) so their properties may vary over very short distances. Exceptionally high-quality surfaces and interfaces are required if surface-sensitive techniques, such as tunnelling or photoemission spectroscopy, are to prove useful tools for probing bulk properties*. Unfortunately, as-prepared surfaces are rarely representative of the bulk and may be insulating, normal, or superconducting with a reduced transition temperature. Intrinsic insulating surfaces usually prevent vacuum tunnelling by an STM and the tip must be pushed into the sample in order to detect a tunnel current (Volodin et al. 1988).† When vacuum tunnelling is achieved at room temperature the barrier height deduced from \(d\ln I/dV\) is extremely low (e.g., 0.3 eV, Yang et al. 1992) - probably due to a damaged or contaminated surface or residual tip–sample contact via a non-conducting particle (Coombs 1987, sec.6.4). Intrinsic surface layers were used as the tunnelling barrier in many early junctions fabricated on high-\(T_c\) superconductors (see, e.g., Gijs et al. 1988).

Auger depth profiling has shown that the surfaces of as-deposited amorphous YBa\(_2\)Cu\(_3\)O\(_y\) thin-films are Y and Cu poor to a depth of 150 nm (Gavaler and Braginski 1988). Angle-resolved x-ray photoelectron spectroscopy (ARXPS) identified this surface material as insulating BaCuO\(_2\) (Chang et al. 1990), although the estimated thickness was only 1 nm. ARXPS also showed that the surface layer on Bi-Sr-Ca-Cu-O is thinner than 0.1 nm. Surface atomic segregation is exacerbated by the post-deposition anneal in oxygen usually required to bring about superconductivity in thin-film samples (Gavaler et al. 1989), and may result in wild variation in composition to depths of 50–150 nm (Bakunin et al. 1989).

Zandbergen et al. (1988) demonstrated that YBa\(_2\)Cu\(_3\)O\(_y\) free surfaces are unstable and that decomposition observed at \(T > 150^\circ\text{C}\) also occurs slowly at room temperature. High-resolution electron microscopy and XPS showed that decomposition occurred by removal of Cu and O from the near-surface region and formation of (CuO)\(_2\) double

* The wavefunction of a tunnelling particle undergoes rapid exponential decay so current is determined predominantly by the wavefunction at the surface of the junction’s electrodes. However, the properties of a superconductor vary in a non-local way so tunnelling effectively ‘probes material to a depth equal to the coherence length’ - as is often stated in the literature.

† Vacuum tunnelling and atomic-resolution imaging has recently been achieved, mostly with cleaved Bi-Sr-Ca-Cu-O crystals (Chang et al. 1992) but also with YBa\(_2\)Cu\(_3\)O\(_y\) thin films (Hasegawa et al. 1992b) and crystals cleaved at 20 K (Edwards et al. 1992).
layers between the BaO planes. The rate of defect formation is limited by the mobility and availability of the copper and oxygen so does not take place in vacuum.

The surface of YBa$_2$Cu$_3$O$_y$ is also degraded by reaction with CO$_2$ catalysed by water vapour (Egdell et al. 1989)

$$2\text{YBa}_2\text{Cu}_3\text{O}_y + 3\text{H}_2\text{O} \rightarrow 3\text{Ba(OH)}_2 + \text{Y}_2\text{BaCuO}_5 + 5\text{CuO} + \frac{1}{2}\text{O}_2$$

$$\downarrow$$

$$3\text{Ba(OH)}_2 + 3\text{CO}_2 \rightarrow 3\text{H}_2\text{O} + 3\text{BaCO}_3$$

To a lesser extent Bi-Sr-Ca-Cu-O and La$_{2-x}$Sr$_x$CuO$_{4-y}$ surfaces are degraded by atmospheric exposure, producing surface carbonates and hydroxides (Flavell et al. 1989, Kurtz et al. 1988). Flavell et al. (1990) report that in UHV Bi$_2$Sr$_2$CaCu$_2$O$_8$ crystals are very stable to water exposure.

The well known sensitivity of YBa$_2$Cu$_3$O$_{7-\delta}$ to oxygen non-stoichiometry has prompted frequent suggestions that the non-conducting surface layer is due to oxygen deficiency. There have been several studies of the dependence of high-temperature superconductor tunnelling characteristics, and the observed energy gap, on bulk oxygen content of the sample (e.g., Al’tfeder et al. 1989, Tsai et al. 1989b, Kotyuzhaanskii 1991). These all show a decrease in energy gap and $T_c$ with reduced oxygen stoichiometry (with $2\Delta/k_B T_c$ approximately constant), suggesting that loss of oxygen from the sample’s surface would result in a locally reduced transition temperature and energy gap.

Fully oxygenated YBa$_2$Cu$_3$O$_{7-\delta}$ (i.e., $\delta = 0$) is orthorhombic and superconducting with $T_c = 90$ K. If $\delta$ is increased the transition temperature continuously decreases until at $\delta \approx 0.56$ the material loses superconductivity (Kwok et al. 1988). Structurally the material becomes less orthorhombic as ‘chain’ O($0,\frac{1}{2},0$) atoms are depleted by overall loss and transfer to O($\frac{1}{2},0,0$) sites. The material becomes fully tetragonal at approximately the same point that superconductivity is lost, although the exact correlation between these events is still unclear. In the tetragonal phase further loss of oxygen results in gradual appearance of semiconducting behaviour (i.e., a region of metallic non-superconductivity does exist). Transmission electron diffraction (TED) has revealed that to a depth of 1 µm on the surface of YBa$_2$Cu$_3$O$_y$ single crystals, orthorhombic splitting is reduced by 10–30% from the bulk value (Werder et al. 1989). Moreover, ‘pockets of disorder’, polycrystalline regions and a high defect density were present in the surface layers. The reduced splitting could be due to either lower oxygen concentration at the surface or cation disorder.

The anomalously low Fermi energy density of states measured by XPS and the
consistent failure of this technique to observe a cut-off on cooling the sample below $T_c$ (Egdell et al. 1990) have also been attributed to an oxygen-deficient surface. List et al. (1988) and Arko et al. (1989) showed that a ‘Fermi edge’ is present in well oxygenated YBa$_2$Cu$_3$O$_y$ single crystals cleaved in vacuo at 20 K. Subsequent XPS spectra remained stable whilst the sample was maintained at 20 K (suggesting that surface contamination is not the problem) but on warming above 80 K the Fermi edge irreversibly disappeared. Warming also resulted in charging of the sample surface by incident electrons. These results are widely regarded as evidence for (chain) oxygen loss leading to degraded superconducting or even insulating surface layers.

Disruption of the CuO chains on warming a YBa$_2$Cu$_3$O$_{y}$ single crystal has been observed directly in atomic-resolution topographic STM images. Edwards et al. (1992) found that vacuum tunnelling was possible only if the sample was cleaved in situ at temperatures lower than 60 K (and in UHV) prior to measurement, resulting in a clear energy gap ($2\Delta/k_B T_c = 6–8$) and atomic images of a corrugated surface with period equal to the a-direction lattice constant*. On warming the sample to 70 K these features irreversibly disappeared and the tunnelling barrier height, estimated from the logarithmic slope of $I(z)$, was reduced from 1 eV to less than 0.5 eV.

YBa$_2$Cu$_3$O$_{y}$ is also degraded by contact with evaporated non-noble metals (e.g., Pb, In). Contact resistance studies and XPS depth profiling have shown that oxygen is leeched from the ceramic to form the metal oxide and a semiconducting surface layer (Gao et al. 1987, Suzuki et al. 1989). If the metal is deposited at low temperature an insulating surface layer only forms when the junction is warmed (to 200 K for Pb) and oxygen diffusion commences (Beasley 1991). Ageing occurs as the insulating barrier grows thicker, and fine details of the tunnelling characteristics, including the lead gap, are ‘washed out’ (Tao et al. 1991). The development of the contact resistance as a function of temperature and time has been investigated by Grajcar et al. (1991) and found to be consistent with diffusion-limited oxygen depletion of near-surface YBa$_2$Cu$_3$O$_{y}$ material.

Noble-metal contacts on YBa$_2$Cu$_3$O$_{y}$ have lower resistance with metallic temperature dependence (Adkins et al. 1990). However, a degraded interface layer may still remain (Tarte et al. 1991b), possibly a 1 nm thick insulating layer containing conducting pin holes (Takeuchi et al. 1990, Tsai et al. 1990). Luo et al. (1992) investigated disruption of the surface of Bi$_2$Sr$_2$CaCu$_2$O$_8$ crystals by sub-monolayer deposition of an inert metal. They proposed that the Bi–O layer exposed by cleaving

* Twin boundaries were also seen as a rotation of the corrugation axis by 90°. These observations indicate that YBa$_2$Cu$_3$O$_{y}$ cleaves between the CuO chain layer and the BaO layer.
is metastable, and that a lower-energy state is created by conversion of the surface to more bulk-like Bi$_2$O$_3$. STM observation of 2.5 nm diameter metal clusters, arranged in rows along the $b$-axis, suggests that the activation energy for this process is provided by energy released by agglomeration of mobile metal atoms. The uniform arrangement of clusters occurs because they preferentially form along lines of greatest instability, which are correlated with superstructure intrinsic to the Bi–O layer*.

Oxygen required for the conversion of BiO to Bi$_2$O$_3$ is usually obtained by reduction of near-surface CuO$_2$ planes - probably resulting in loss of superconductivity in this region - so deposition of the metal in an oxygen atmosphere has been proposed as a method for minimising surface disruption. When non-noble metals are deposited the physical disturbance is augmented by strong chemical reactions.

The ubiquitous nature of the surface layer has prompted suggestions that high-$T_c$ superconductors have intrinsically degraded surfaces, either through reconstruction (Allan et al. 1989, Jagannadham and Narayan 1990) or because their small coherence lengths results in a naturally depressed energy gaps at line defects and surfaces (Deutscher and Müller 1987, Halbritter 1992). The latter has been used to explain observations of SNS or SIS weak-link-type behaviour at twin boundaries†, despite the absence of any microscopic evidence for an insulating or normal layer (Dwir 1990). Halbritter’s model for intrinsic degradation of superconductivity at interfaces, and its wide-ranging effects, are discussed in section 4.8.

The presence of an uncontrolled surface layer is a major impediment to high-quality tunnelling experiments on copper-oxide superconductors. Low-temperature cleaving has already been discussed as a technique for exposing bulk material and appears to be successful with crystals and thin films. Ceramic materials, however, may cleave along pre-existing grain boundaries resulting in no improvement (Fein et al. 1987). Barone et al. (1987) found that Josephson supercurrent was only present in Nb/ceramic YBa$_2$Cu$_3$O$_y$ point contacts if the sample’s surface was first cleaned by ion-beam milling.

---

* Bi–O planes in Bi$_2$Sr$_2$CaCu$_2$O$_8$ are corrugated with a modulation period of 4–5 unit cells and an amplitude of approximately 1 Å, due to the presence of an extra oxygen atom for every 9–10 Bi atoms. This superstructure has been observed in the bulk by x-ray diffraction and at cleaved surfaces by STM.

† Formed under very controlled conditions by thin-film deposition of YBa$_2$Cu$_3$O$_y$ on an SrTiO$_3$ bicrystal with known grain misorientation (Mannhart et al. 1988). Fitting an SIS weak-link model to the data indicated a much reduced energy gap immediately adjacent to the grain boundary.
Several etchants have been used to remove surface layers from high-temperature superconductors. Most successful have been 1.0% bromine in ethanol (Geerk et al. 1991) and a mixture of 10 mM HClO₄ and 1 M NaClO₄ in water (Gurvitch et al. 1989). These have been shown to remove contaminating hydroxides and carbonates and restore orthorhombic structure to the surface (Werder et al. 1989). TEM and XPS of etched surfaces indicated that no further degradation occurred over a period of seven months. In contrast, STM studies (Balzarotti et al. 1992) showed that surface modification occurred on a time scale of several hours after etching, resulting in an insulating layer that could be modelled as a 2.5 nm thick trapezoidal potential barrier with heights 1.5 eV and 4.7 eV on the superconducting and normal sides respectively. This recontamination was much slower than observed with air-cleaved surfaces and was interpreted as efficient passivation of the surface by the etchant. Valles et al. (1991) accelerate formation of a barrier by heating the etched YBa₂Cu₃O₇ crystal surface in air for 5–10 min at 100°C. Planar tunnel junctions formed by evaporating a 0.1 × 1.0 mm² metal counterelectrode then exhibit reproducible characteristics that are believed to be representative of the bulk. One disadvantage of etching is that large square pits form at dislocations and defects, introducing additional surface roughness, which disturbs anisotropy experiments. The presence of etch pits, however, may ‘improve’ the tunnelling characteristics on a surface nominally oriented perpendicular to the c-axis (i.e., the cleavage surface of single crystals) by allowing tunnelling in the ab-plane where the coherence length is much larger.

Attempts have been made to encapsulate cleaned or freshly prepared surfaces with a deposited metal passivation layer on which a proximity-effect junction may then be fabricated (Moreland et al. 1989). Silver is a particularly good overlayer because it acts as a switchable filter to oxygen; at high temperature oxygen diffuses through the silver, so samples may be annealed as usual (Ekin et al. 1988). The main success of this technique has been to reduce significantly the sub-gap conductance; however, the tunnelling characteristics are still broadened, have a V-shaped background, lack conductance overshoots and contain complex fine structure (Gijs et al. 1990, Lee et al. 1989b, Golubov and Kupriyanov 1989). To extract properties of the underlying superconductor (i.e., Δ, α²F) it is necessary to ‘deconvolve’ the effect of the proximity layer. The theory and technique of proximity-effect tunnelling spectroscopy (PETS) was developed by Arnold and Wolf (Arnold 1978, Arnold et al. 1980, Wolf et al. 1980a,b) and subsequently used by Zasadzinski et al. (1982) in studies of vanadium*.

---

* Complex oxidation behaviour and strong reactivity with oxygen leads to degraded surface layers that prevent fabrication of SIN junctions.
The problematic nature of natural barriers has led to interest in artificially deposited barriers, which are potentially more controllable. Many barrier materials have been tried, usually resulting in highly non-ideal conductance–voltage characteristics (Iguchi and Wen 1991). Recently great progress has been made using laser ablation to grow thin films of insulating Bi$_2$Sr$_2$CuO$_6$ epitaxially on single crystals of Bi$_2$Sr$_2$CaCu$_2$O$_8$ (section 6.1.1). Similar success was achieved by Kasiviswanathan and Rangarajan (1992) using a sputtered indium oxide barrier and evaporated Pb$_{0.5}$I$_{0.5}$ counter electrode on NdBa$_2$Cu$_3$O$_{7-y}$. The conductance–voltage characteristics of this junction exhibited a well defined gap-like feature with $\Delta = 20$ meV ($2\Delta/k_B T_c = 5.1$) and a very low level of broadening of $\Gamma = 1.5$ meV.

4.6 $2\Delta/k_B T_c$ : The Reduced Gap Value of High-$T_c$ Superconductors

Several reviews contain excellent summaries of reduced gap values, $2\Delta/k_B T_c$, determined by tunnelling studies (e.g., Kirtley 1990, Hasegawa et al. 1992a). Hasegawa claims that ‘the results showed considerable scatter in the early period, but have been converging’, and this opinion is often repeated. Although there is a trend towards more ‘ideal’ tunnelling characteristics, as junction and sample quality improves, there is little evidence for convergence towards an underlying true $2\Delta/k_B T_c$.

Figures 4.16–18 summarise reduced gap values measured by tunnelling and reported in the literature from 1987 to 1992. All results are treated with equal merit and no extra weighting is given to more ‘ideal-looking’ tunnelling characteristics. All that can be said with certainty is that $2\Delta/k_B T_c$ seems to be larger than the BCS value of 3.52.

After five years of intensive effort by the world-wide tunnelling community why does so much scatter remain in published reduced gap values? Possible reasons for this have already been discussed in connection with broadening of the conductance–voltage characteristics, since each of the mechanisms considered could also lead to scatter in the mean observed gap. In addition to true variation in energy gap from sample to sample, significant differences can arise even from a single conductance–voltage characteristic due to the way in which the conductance background is handled and the method used to estimate the gap (section 3.7, also cf. Matsumoto et al. 1991, 1992).

Exceptionally large values of $2\Delta/k_B T_c$ may be erroneous for a number of reasons. The authors may be overlooking charging effects, which increase separation of the central conductance peaks, or may be misinterpreting the type of junction being measured. The conductance–voltage characteristics may be dominated by one or more ‘rogue’ tunnel junctions that exist in series with the intended junction. For example, if a normal tunnelling tip is in metallic contact with a grain of superconducting material
Fig. 4.16. Box-whisker diagram of reduced energy-gap values measured by tunnelling and reported in the literature between 1987 and 1992. Shaded boxes represent the range of ±25% of the sample about median (central line). Bars show the range of remaining samples but exclude those lying outside ±2σ, which are plotted individually (circles). Bracketed numbers indicate sample size. Summary information is given in the table below.

<table>
<thead>
<tr>
<th>material</th>
<th>mean reported $2\Delta/k_B T_c$</th>
<th>standard deviation σ in reported $2\Delta/k_B T_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>YBa$_2$Cu$_3$O$_y$</td>
<td>5.0</td>
<td>2.8</td>
</tr>
<tr>
<td>Bi$_2$Sr$_2$CaCu$_2$O$_8$</td>
<td>6.7</td>
<td>2.2</td>
</tr>
</tbody>
</table>
Fig. 4.17a. Histogram of reduced gap values of YBa$_2$Cu$_3$O$_y$ measured by tunnelling and reported in the literature between 1987 and 1992.

Fig. 4.17b. The trend in reduced gap values of YBa$_2$Cu$_3$O$_y$ measured by tunnelling.
Fig. 4.18a. Histogram of Bi$_2$Sr$_2$CaCu$_2$O$_8$ reduced gap values measured by tunnelling.

Fig. 4.18b. Reduced gap values of Bi$_2$Sr$_2$CaCu$_2$O$_8$ measured by tunnelling.
that is separated from the bulk superconductor by a thin insulating layer, then the characteristics will be those of an SIS junction rather than the nominal SIN junction; the estimated energy gap will be too large by a factor of two (Rodrigo et al. 1991). Errors of this type may occur in thin-film, single-crystal and ceramic samples due to the presence of weak links inside grains at crystallographic defects as well as at grain boundaries (England et al. 1988, Dimos et al. 1990, Dulcic et al. 1990). Theoretical calculations have shown that diffuse tunnelling characteristics of a two-dimensional anisotropic superconductor may contain only subtle structure associated with the energy gap, and that even at \( T = 0 \) the SIN and SIS cases may be almost indistinguishable (Wyosinski and Dziura 1990).

The broadening present in most tunnelling characteristics also tends to result in overestimated energy gaps. Zasadzinski et al. (1992) observed a range of behaviour with one Ba\(_{1-x}\)K\(_x\)BiO\(_3\) sample, from ideal BCS characteristics with a point contact to severely broadened characteristics with a planar junction. The conductance maxima in the latter case occurred at voltages much larger than \( \Delta / e \), indicating that these cannot be used to estimate the energy gap. Moreover, even when broadening was taken into account by fitting Dyne’s model (4.17) to the data, the resulting gap estimate was too large by 50%.

Exceptionally low reduced gap values may result from using the bulk transition temperature to calculate \( 2\Delta_0 / k_B T_c \), since this does not take into account depression of the transition temperature at the degraded surface. Reduced gap values should only be believed if \( T_c \) has been determined from the ‘gap opening temperature’ found by measuring \( \Delta(T) \); however, this experiment is very difficult with point contacts due to changes in the junction brought about by thermal expansion. On the other hand, planar junctions usually have extremely non-ideal conductance–voltage characteristics so \( \Delta_0 \) cannot be confidently estimated.

Despite these difficulties several groups have investigated the temperature dependence of the energy gap. These efforts are reviewed by Hasegawa et al. (1992a, sec.II E) and Kirtley (1990, sec.VI). Briefly, YBa\(_2\)Cu\(_3\)O\(_y\) has yielded conflicting results; some \( \Delta(T) \) curves are BCS like, whilst others show a decrease in amplitude of the conductance dip at zero-bias with increased temperature, but almost no change in the estimated gap width. Most studies of Bi-Sr-Ca-Cu-O have been interpreted as agreeing with the BCS temperature dependence.

### 4.7 Phonon Effects

In view of the experimental difficulties discussed in this chapter it is not surprising that fine structure above the energy gap has yet to be measured repeatably in any high-temperature superconductor. The McMillan–Rowell inversion technique has been
applied to some conductance–voltage characteristics in an attempt to estimate the
However, the results often disagree with the phonon density of states, $F(E)$, observed
by neutron scattering and the $T_c$-value calculated from the measured $\alpha F(E)$ is not
always consistent with the sample’s transition temperature. Some success has been
achieved with the low-$T_c$ oxide superconductors, Ba$_{1-x}$K$_x$BiO$_3$ and Nd$_{2-x}$Ce$_x$CuO$_{4-y}$ (see
appendix B).

4.8 Halbritter’s model

Halbritter (1992) has proposed a model for intrinsic degradation of superconductivity
at the internal and external surfaces of a high-temperature superconductor. This theory
is particularly attractive because it explains many of the anomalous phenomena that
have been discussed earlier in this chapter, notably the linear background conductance,
the origin of gap inhomogeneity and hence broadening of the tunnelling characteristics, leakage currents, insulating and normal surface layers, and scatter in
reported gap values*. In addition to being wide ranging the model agrees well with
details of experimental observations.

Halbritter suggests that intrinsic relaxation of dangling bonds at surfaces creates
disorder (particularly of O in the CuO chains of YBa$_2$Cu$_3$O$_y$) and a tetragonal
insulating phase containing a very high density of localised states. Evidence for
surface-derived localised states has been seen in ARPES measurements of cleaved
YBa$_2$Cu$_3$O$_y$ single crystals (Claessen et al. 1990) and in specific-heat measurements of
radiation-damaged YBa$_2$Cu$_3$O$_y$. The surfaces of Bi-Sr-Ca-Cu-O superconductors are
more inert and less prone to disorder, so the density of localised states is smaller and
the tunnelling characteristics more ideal.

The energy gap at a localised state in the insulator is drastically suppressed by Coulomb
repulsion. This suppression increases rapidly with increasing distance from the
superconductor so that states further than 0.4 nm from the interface are normal. Only
the nearest layer of localised states to the superconductor will retain a non-zero energy
gap, though this will be reduced to 40–90% of the bulk value. Pair weakening also
occurs on the superconducting side of the interface, resulting in a locally depressed $\Delta$.
Becase the energy gap is very inhomogeneous the tunnelling characteristics will be
severely broadened and measured gap values will be very scattered.

* The model also predicts a number of detailed properties of Josephson currents across weak links (e.g.,
$J_c$ depression, dependance on temperature and grain-boundary resistance, excess noise and rf losses).
These will not be discussed here.
The presence of localised states in the barrier means that direct tunnelling will be augmented by the indirect tunnelling processes described in section 3.6 (i.e., resonant and two-step tunnelling). Halbritter suggests that the large sub-gap ‘leakage currents’ of high-$T_c$ superconductor tunnel junctions are due to resonant tunnelling via states well into the insulating layer, which have $\Delta = 0$. In addition, Halbritter proposes that the linear (i.e. V-shaped) conductance also results from resonant tunnelling via these intrinsic localised states. He states that this will yield a conductance slope of

$$\left| \frac{d^2I}{dV^2} \right| \propto \left( \frac{dI}{dV} \right)_{V=0} \propto n_s n_L \exp(-w \sqrt{2m\phi/h})$$

(4.18)

where $w$ is the barrier width, $\phi$ is the barrier height and $n_s$ and $n_L$ are the carrier density and density of localised states respectively. This predicted linear relationship between conductance slope and zero-bias conductance agrees with experimental observations made by Sharifi et al. on bismuthate tunnel junctions (section 4.2.1). Halbritter also claims that if $n_s$ increases with transition temperature then (4.18) also explains the experimentally observed $T_c$ dependence, since then

$$\left( \frac{dI}{dV} \right)_{V=0} \propto T_c$$

Unfortunately, this is the same misinterpretation of Sharifi’s experimental data as was made by Varma et al. (section 4.2.5). The data actually shows a linear relationship between $T_c$ and the conductance slope normalised by the zero-bias conductance - so $T_c$ dependence occurs in the unwritten constant of proportionality in (4.18), not in $(dI/dV)_{V=0}$.

Halbritter shows that the mixture of direct and indirect tunnelling channels can explain the observed dependence of tunnel current on voltage, temperature and barrier width. When these parameters are large indirect tunnelling is more important so, for example, the V-shaped conductance will dominate the tunnelling characteristics. When the barrier is thin the increasing background is less severe and subtle features in the direct tunnelling characteristics will be visible (e.g., a normal density of states that intrinsically decreases away from $E_F$). This qualitatively explains the observed change in conductance background with junction resistance (or tip–sample separation) described in section 4.2.1.

† This is the same mechanism that was discussed in section 4.2.4.

‡ In this case the localized states are created extrinsically by ion-milling damage used to controllably form the tunnelling barrier, however, the consequences of their presence in the junction will be the same as for the cuprate’s intrinsic states.
CHAPTER 5
EXPERIMENTAL DETAILS
I. MEASUREMENT ELECTRONICS

5.1 A.C. Measurement of Conductance

As described in section 3.2, the most direct way of investigating a superconductor’s density of excitation states is to measure the bias-dependent dynamic conductance, \( \frac{dI(V)}{dV} \), of an SIN junction. A standard technique for measuring this quantity is to apply to the sample a d.c. bias, \( V_0 \), on which is superimposed a small sinusoidal modulation of angular frequency \( \omega \) and r.m.s. amplitude \( \delta V \). The current flowing through the sample will contain a component of frequency \( \omega \), which is proportional to

\[
V = V_0 + \sqrt{2} \delta V \cos \omega t
\]

Fig. 5.1. Schematic diagram of the circuit used to measure conductance–voltage characteristics. The sample is represented by impedance \( Z_s \). Practical design details of this circuit are discussed in appendix A.
dI(V₀)/dV and may be detected with a phase sensitive detector (PSD, or equivalently ‘lock-in amplifier’). Sweeping the d.c. bias whilst recording the PSD output yields the sample’s conductance–voltage characteristic. A non-ohmic sample will also produce harmonics of this fundamental frequency from which the higher-order derivatives may be found. The design of spectrometers using this a.c. modulation technique is reviewed by Wolf (1985, p.505). In this study the sample conductance was measured using an active voltage-biasing spectrometer (figure 5.1), based on a design by Edgar (1987).

The modulation voltage and d.c. bias were combined by an audio-frequency transformer. The resulting signal, \( V₀ + \sqrt{2} \delta V \cos \omega t \), was applied to the non-inverting input of an op-amp (OP1) which held one side (A) of the sample at this voltage, independent of sample or lead resistances. The other side (B) of the sample was maintained at the circuit’s ground potential by a similar op-amp loop (OP2). Voltage biasing the sample ensures that the modulation amplitude is constant and results in a simple relationship between the output of the PSD and the dynamic conductance.

Op-amps with extremely low input bias currents (< 10 pA) were used so that when operating in four-terminal mode (with the ‘2T’ links removed) all current flowing through the sample, \( Z_s \), passed through the sensing resistor, \( R_b \). This current may be expressed as a Taylor series expansion about the d.c. bias, \( V₀ \):

\[
I(V) = I(V₀) + \left( \frac{dI}{dV} \right)_{V₀} \sqrt{2} \delta V \cos \omega t + \frac{1}{2!} \left( \frac{d^2I}{dV^2} \right)_{V₀} (\sqrt{2} \delta V \cos \omega t)^2 + O(\delta V^3) 
\]

\[
≈ I(V₀) + \left( \frac{dI}{dV} \right)_{V₀} \sqrt{2} \delta V \cos \omega t + \frac{1}{2!} \left( \frac{d^2I}{dV^2} \right)_{V₀} \delta V^2 \cos 2\omega t 
\]

The voltage drop across \( R_b \) was amplified by an instrumentation amplifier (OP3) with gain \( A \) and input to both a voltmeter and a lock-in amplifier. These measure a voltage proportional to each term in the series expansion and so allow the current and derivatives to be calculated from

\[
I(V₀) = \frac{V_{dc}}{R_b A} 
\]

\[
\left( \frac{dI}{dV} \right)_{V₀} = \frac{V_ω (V₀)}{\delta V R_b A} 
\]

\[
\left( \frac{d^2I}{dV^2} \right)_{V₀} = \frac{2\sqrt{2} V_2ω (V₀)}{\delta V^2 R_b A} 
\]
where $V_{dc}$ is the voltmeter reading and $V_{\omega}$ and $V_{2\omega}$ are the r.m.s. fundamental and second-harmonic voltages measured by the lock-in. If the sample is purely resistive then (5.3) and (5.4) are the dynamic conductance and its derivative respectively, and the PSD is simply used to measure the amplitude of the input signal. If the sample has some capacitance then the total current is

$$I(V) = I_{dc}(V) + C(V) \frac{dV}{dt}$$

where $C$ is the sample’s dynamic capacitance, $dq/dV$, and $dV/dt = \omega \sqrt{2} \delta V \sin \omega t$. The dynamic conductance, $G$, is $dI_{dc}/dV$ so

$$\frac{dI}{dV} = G(V) + C'(V) \frac{dV}{dt}$$

and

$$\frac{d^2I}{dV^2} = G'(V) + C''(V) \frac{dV}{dt}$$

Substituting these into (5.1) gives expressions for the dynamic conductance and capacitance and their derivatives, in terms of the in-phase ($V_{||}$) and quadrature ($V_{\perp}$) voltages measured by the lock-in amplifier,

$$G(V_0) = \frac{V_{\omega}(V_0)}{\delta VR_b A} \quad \text{and} \quad C(V_0) = \frac{V_{\omega}'(V_0)}{\delta VR_b A \omega}$$

$$\left( \frac{dG}{dV} \right)_{V_0} = \frac{2 \sqrt{2} V_{2\omega}(V_0)}{\delta V^2 R_b A} \quad \text{and} \quad \left( \frac{dC}{dV} \right)_{V_0} = \frac{\sqrt{2} V_{2\omega}'(V_0)}{\delta V^2 R_b A \omega}$$

The phase is with respect to the voltage across the sample so this information must be supplied to the lock-in amplifier. This was done by adjusting the phase control to zero the quadrature voltage measured at the input to OP1*. The in-phase voltage was then $\delta V$.

The d.c. bias was provided by a function generator and input to both OP1 and the $x$-axis of a chart recorder. The $y$-axis of the chart recorder was connected to either the OP3 output, to measure current, or the PSD output to measure $dI(V)/dV$ or $d^2I(V)/dV^2$. The recorded traces were later digitised and calibrated. The modulation was supplied by the internal oscillator of the lock-in amplifier (EG&G PAR model

---

* This method will not work if the op-amps or differential amplifier introduce a phase shift or attenuate the modulation signal. See appendix A.
or by a Krohn-Hite 4400A oscillator. A modulation frequency of about 1 kHz was used, this being well within the bandwidth of the op-amps and a compromise between reducing sample $1/f$ noise, reducing noise in the electronics (increasing with frequency) and reducing the capacitive current. Harmonics of the mains frequency (50 Hz) were avoided. The modulation amplitude was chosen to reduce the instrumentation broadening inherent in the modulation technique (see appendix A) to the same level as thermal broadening, (i.e., for measurements of elastic tunnelling conductance, $\sqrt{3} \delta V < 3.52 k_B T$). Reducing the modulation amplitude decreases the signal-to-noise ratio and so requires a longer lock-in time constant and hence a longer measurement time.

Care must be taken to ensure that the sample leads are properly connected to all four terminals of the sample. A break in one of the leads will open a feedback loop causing the op-amp’s output to saturate at 15 V, possibly damaging the junction*. To guard against this the feedback loops were initially kept closed by switchable links (illustrated by the dashed lines labelled ‘2T’), which allowed the circuit to make two-terminal measurements. Before opening the links the condition of the sample leads and contacts was monitored by measuring an $I(V)$ curve with only two of the leads connected (e.g., $I+$ and $I-$); clearly if any current was observed the leads and contacts were continuous. This was repeated for the other pair of leads. Usually the lead resistance was much smaller than the sample resistance and there was very little difference between characteristics obtained in the two-terminal and four-terminal configurations.

### 5.1.1 Conductance-Bridge Measurements

The circuit shown in figure 5.1 was used to measure samples in which the conductance changed significantly over the bias range of interest. Many junctions made in the early stages of this study had much more linear $I-V$ characteristics in which the conductance changed by only a few percent. To improve the accuracy of measurements on such junctions the voltage-biasing electronics was incorporated into a conductance bridge (figure 5.2).

When the bridge is balanced by a parallel resistor and capacitor ($R_b$ and $C_b$) the sample conductance and capacitance are given by†

---

* To ensure that the lower feedback loop remains closed, switches used to select the value of $R_b$ must be ‘make before break’.

† A derivation of the calibration equations and an analysis of errors can be found in the thesis of Speakman (1992, ch.3).
\[ G(V) = G(V_b) \left[ 1 - \frac{\beta}{\delta V} \left( V_{\omega}^\parallel(V) - R_b C_b \omega V_{\omega}^\bot(V) \right) \right] \]

and

\[ C(V) = C(V_b) \left[ 1 - \frac{\beta}{\delta V} \left( V_{\omega}^\parallel(V) + \frac{V_{\omega}^\bot(V)}{R_b C_b \omega} \right) \right] \]  \hspace{1cm} \text{(5.5)}

where \( \beta = 1 + \kappa \), \( \kappa = \frac{R_3}{R_4} \) and the conductance and capacitance at balance are

\[ G(V_b) = \frac{1}{\kappa R_b} \quad \text{and} \quad C(V_b) = C_b / \kappa \]

The capacitance and conductance derivatives can also be found from the second-harmonic signal. Most superconducting tunnel junctions have a negligible (and unimportant) capacitance so balance was achieved using just a resistor \( (C_b = 0) \). In this case

\[ G(V) = G(V_b) \left[ 1 - \frac{\beta}{\delta V} V_{\omega}^\parallel(V) \right] \]

and

\[ C(V) = -G(V_b) \left[ \frac{\beta}{\delta V \omega} V_{\omega}^\bot(V) \right] \]  \hspace{1cm} \text{(5.6)}

\[ V = V_0 + \sqrt{2} \delta V \cos \omega t \]

\[ \text{Fig. 5.2. Schematics of the conductance bridge used to measure junction characteristics.} \]
In the bridge configuration the lock-in amplifier output is proportional to the conductance change from the balance condition, so the dynamic range is increased and changes of less than 1% may be detected. An additional advantage of the bridge configuration is that noise generated in the voltage source appears across both arms of the bridge and does not affect the lock-in, which measures the difference signal.

5.2 D.C. Measurement of Conductance

The a.c. modulation method was very convenient for measurement of the conductance of stable tunnel junctions. An analogue voltage source could be swept slowly through the bias range of interest and the output of the lock-in amplifier recorded on a chart recorder. Noise in the measured conductance was reduced to acceptable levels simply by raising the lock-in’s time constant and sweeping more slowly. Typically a single conductance curve would take several minutes to collect.

This was not a practical method, however, for measuring the conductance–voltage characteristics of vibration-sensitive point-contact junctions; a particular $I$–$V$ characteristic would remain stable for several minutes and then change unpredictably. Initially the experimental procedure was to sweep the sample bias using a 30 Hz triangular waveform and monitor the $I$–$V$ curve of the contact on an oscilloscope whilst manipulating the point. When interesting characteristics were obtained the sensitivity of the lock-in amplifier was set to an appropriate level and a slow sweep to measure the conductance–voltage curve was started. On most attempts good data would be collected from only part of the voltage range of interest and then the junction would be disturbed resulting in a dramatic jump in the lock-in amplifier’s output.

It would be better to obtain data from the full bias range of interest even if signal-to-noise ratio was compromised to achieve this. A large amount of potentially useful data-acquisition time was also wasted in setting up the lock-in amplifier, replacing chart-recorder paper and placing calibration marks on the recorded trace.

To solve these problems the d.c. $I$–$V$ characteristics of point-contact junctions were measured using a computer-controlled data-acquisition system, designed and built by the author. This consisted of a digital-to-analogue converter (DAC) generating a voltage ramp which was input to OP1 in figure 5.1, and an analogue-to-digital converter (ADC) recording the instrumentation amplifier’s output voltage (i.e., the current signal). The lock-in amplifier was not used and the differential conductance was found by numerical differentiation of the $I$–$V$ curve.

The DAC and ADC were part of a CIL6580 data-acquisition unit*. This also incorpo-

* Manufactured by CIL Microsystems, Lancing, Sussex.
rated variable-gain differential amplifiers, digital filtering, an RS-232 (or optionally an IEEE-488) interface and a low-level programming language. The DAC and ADC were 16 bit and digitisation noise was not found to be a problem.

The CIL6580 unit was controlled by a host computer† running a custom-written Pascal program. This allowed the user to define a table of voltages, which would be output sequentially by the DAC to the sample-biasing electronics. For each element in the table the program output the required voltage, waited for a user-defined period of time, then measured the output of the instrumentation amplifier. In this way the $I$–$V$ curve for the sample was measured point by point and displayed by the computer, which effectively acted as an oscilloscope. Whilst monitoring the display the point contact was adjusted and, when interesting characteristics were obtained, data acquisition could be started simply by pressing a button. No time was wasted in preparing the chart recorder and useful data was collected from even the most transient of contacts. On completing a voltage sweep the recorded $I$–$V$ points were saved in a simple ASCII format data file, which could later be read by a number of custom-written and standard data-processing packages. The spectrometer control program was optimised for rapid data gathering with minimum loss of data-acquisition time due to setting measurement parameters etc.

5.2.1 Calibration of Raw Data

Computer control also allowed the data to be digitised and calibrated automatically. This was done laboriously by hand for curves previously recorded on paper by a chart recorder. The calibration models are schematically illustrated in figure 5.3.

$V_{\text{nom}}$ is the voltage defined by the user in the look-up table and $N$ is the 16-bit number output by the ADC. These are the two items of raw data ‘known’ by the computer and input to the calibration routines of the program. $AF$ is the attenuation factor of a potential divider used to reduce the ± 10 V output of the DAC to the required sample voltage.

The gain and offset parameters of the DAC and ADCs were found by fitting a straight line to 50 output and input voltages measured with a Thurlby 1905a multimeter. Using these linear calibration models the CIL6502 unit was capable of measuring and outputting voltages with an error of less than 0.1%. After six months these calibration parameters were rechecked and found to be identical to within 0.05%.

The sensing resistors, $R_{\text{sense}}$ ($R_b$ in figure 5.1), which ranged from 2 Ω to 20 MΩ, were measured with an error of less than 0.4%, also with the Thurlby 1905a multimeter.

† Apple Macintosh Plus, 2 MHz 68000 processor, 2 Mb memory, 40 Mb hard disk.
Fig. 5.3. Calibration of the measurement system. $V_{\text{nom}}$ and $N$ are the quantities ‘known’ by the computer. Currents and voltages at other points in the system can be calculated from these via the linear calibration models defined here.

The calibration parameters of other elements in the measurement system (i.e., $AF$, $BIAS_{\text{gain}}$, $BIAS_{\text{offset}}$, $AMP_{\text{gain}}$) were found by using the CIL6580 unit (plus calibration models for its DAC and ADCs) to apply a voltage to the element and to measure its response. Several hundred voltages could be output and the parameters found by fitting a straight line. The d.c. gain of the sample-biasing electronics was always unity and the instrumentation-amplifier gain was exactly the value expected from its circuit-board connections. In addition, the instrumentation amplifier’s offset was trimmed on the circuit board to negligible levels and was not included in the calibration models*.

Because the calibration parameters were expected to change (either through modification of the circuitry or by long-term drift) they were stored in the ‘resource fork’ of the program and read into memory when the program was first launched. They could be changed with a resource editor (e.g., ResEdit, Apple Computer, Inc.) and modification of source code was unnecessary. One parameter, $V_{\text{correction}}$, could be set whilst the program was running. No current can be flowing when the tip is retracted from the sample so any voltage measured at the ADC input must arise through

* Any error in this assumption could be accommodated by the run-time offset correction, $V_{\text{correction}}$. 
an instrumentation-amplifier offset or through thermoelectrically induced voltages. A routine in the program allowed this stray voltage to be measured and subsequently subtracted from all following measurements.

There were three principal sources of error in the calibrated data - noise and residual offset in the electronics and errors in the calibration model. The $I-V$ characteristics of known resistances were measured and analysed to determine empirically the total error appearing in the current and the numerically calculated conductance. Acquisition and processing of this dummy data was done in exactly the same way as for real data. The results of these tests are summarised in figures 5.4–6. The position of conductance peaks in the tunnelling characteristics of conventional SIS tunnel junctions indicated that residual offset in the sample bias was smaller than 0.05 mV. Sources of noise and calibration errors are discussed in appendix A.

![Graph](image)

**Fig. 5.4.** Noise and offset in current for each value of the sensing resistor. Measurements were made by taking 101 readings of current with ADC range set to 100 mV and no sample connected. The offset was taken from the mean of these readings and the noise from the standard deviation. Lines show full-scale current for each ADC range. In these measurements the run-time offset correction, $V_{\text{correction}}$, was not used, so this data shows the raw, worst-case offset.
Fig. 5.5. Error and noise in the numerically calculated conductance of a 1 MΩ dummy sample. The conductance was calculated from $I$–$V$ characteristics measured over a range of bias voltages - each sweep containing 101 $I(V)$ points.

Fig. 5.6. Error and noise in the numerically calculated conductance of known resistances. Each conductance was calculated from an $I$–$V$ sweep of 101 points over ±100 mV.
5.2.2 Numerical Differentiation and Smoothing

A numerical differentiation routine, derived from Stirling’s interpolation formula, was used to calculate the dynamic conductance from the measured $I$–$V$ curves (Hosking et al. 1986). If $f_m = f(x_m)$ is a function tabulated at points $x_0, x_1, \ldots, x_N$, which are evenly spaced with separation $h$, then the derivative is approximately

$$f'(x_m) = \frac{1}{h} \sum_{k=0}^{N} \left( \frac{(-1)^k (k!)^2}{(2k+1)!} \mu \delta^{2k+1} f_m \right)$$  \hspace{1cm} (5.7)

where, for odd $r$

$$\mu \delta f_m = \frac{1}{2} \sum_{j=0}^{(r+1)/2-1} (-1)^j \left( \binom{r}{j} - \binom{r}{j-1} \right) \varepsilon_{(r+1)/2-j,f_m}$$

$$\binom{r}{j} = \begin{cases} \frac{r!}{j!(r-j)!} & j \geq 0 \\ 0 & j < 0 \end{cases}$$

and the ‘range difference operator’ is defined as $\varepsilon_{r,f_m} = f_{m+r} - f_{m-r}$. If just the first term of (5.7) is used then

$$f'(x_m) = \frac{1}{2h} \varepsilon_1 f_m = \frac{1}{2h} (f_{m+1} - f_{m-1})$$

This is the familiar central difference expression, which only uses nearest-neighbour tabulated points. Retaining two terms in (5.7) uses the nearest two neighbouring points from each side, etc.

One might expect that use of $f(x_m \pm 3)$ to calculate $f'(x_m)$ would lead to excessive broadening in the derivative; however, testing the routine on dummy data showed that using higher terms in (5.7) results in greater accuracy. The increased accuracy is rapidly lost however, as noise is introduced into the data; if noise is greater than 1% no advantage is gained from using more than one term.

Numerically calculated conductance–voltage curves of point-contact junctions were frequently very noisy, with the level of noise increasing with bias voltage. The noise was much larger than observed in resistors with the same value and is attributed mainly to vibrational disturbance of the delicate contact. Consider an ohmic contact with a voltage-independent fluctuating conductance, $G \pm \Delta G$. If this conductance is found by numerical differentiation of a measured $I$–$V$ curve then the observed noise is

$$\delta G(V) = \frac{V}{h} \Delta G$$
This expression shows that the d.c. measurement technique is intrinsically more noisy than a direct measurement of conductance, since $V/h \gg 1$, and that attempts to increase voltage resolution by decreasing $h$ will further increase the noise.

An alternative source of noise was discussed by Cockburn (1988). He observed that noise in the conductance of point-contact tunnel junctions did not increase uniformly but rose rapidly above a threshold voltage of about 100 mV. This phenomenon was attributed to fluctuation in tunnelling barrier height due to changes in the electronic population of defect or impurity states in the barrier. This effect has been seen in many small-area tunnelling systems and so probably contributes to the noise observed in the present study. Indeed, some point contacts did show a dramatic increase in noise above a particular threshold voltage (figure 5.7).

Some useful information might still be present in noisy conductance–voltage characteristics, since fitting a model to the data is perfectly legitimate if errors are used correctly. A numerical routine may be used to construct a smoothed curve, which helps with selection of data sets that are suitable for fitting. This curve should simply serve to guide the eye through the scattered raw data points, each with a large error bar, and should always be presented together with the original data.

The simplest smoothing routine is an unweighted average of the nearest neighbouring points. The data is listed in order of increasing abscissa value ($x_0, x_1, x_2,$ …) and the ordinates of the nearest $r$ data points on either side of $x_m$ are averaged to give the smoothed value, i.e.,

$$\bar{f}(x_m) = \frac{1}{2r+1} \sum_{j=0}^{2r} f(x_{m-r+j})$$

---

**Fig. 5.7.** Characteristics of a W→(Tl,Pb)(Ca,Ce)Sr$_2$Cu$_2$O$_y$ point-contact.
The width of the averaging window, \(2r+1\), sets the level of smoothing. This method is particularly suited to smoothing of data with an error distribution that has a broad tail (Press et al. 1987).

Several other smoothing routines were tried including low-pass filtering of a fast fourier transform (FFT) of the raw data and ‘windowed averaging’ with various weighting functions over the window width. It was hoped that these would smooth the data equally well but with reduced broadening of sharp features. However, neither of these significantly improved upon smoothing by the unweighted averaging method. In addition there was some evidence that peak-like structure was left in the smoothed data as an artefact of the FFT procedure.

The standard numerical analysis procedure employed throughout this work was to differentiate using three terms in Stirling’s numerical expression (5.7) and then smooth with an unweighted ‘window’ of width five (i.e., \(r = 2\)\(^*\)). The analysis was performed by a program which batch processed all files stored in a specified directory and automatically produced conductance–voltage plots. This allowed a large number of current-voltage curves to be analysed rapidly and in a consistent manner.

Some signal averaging was available at the time of data acquisition to reduce noise in the measured \(I–V\) characteristics. The CIL6580 unit contained a digital filter which could be programmed to return the average of \(2^n\) ADC readings \((n = 0–15)\). This was carried out rapidly by the CIL6580’s hardware each time a data acquisition was requested and required no further communication with the controlling computer. In this way each point of the \(I–V\) curve was averaged as it was collected.

Averaging could also be carried out by combining complete \(I–V\) curves. To indicate any changes in the characteristics the set of points from the most recent voltage sweep was displayed together with a rolling average of all curves taken since initiating the acquisition. After each sweep the user was asked whether to add the new data to the rolling average and initiate another sweep, or ignore it and abort the acquisition process. This choice was assisted by the program, which warned if any point in the new data set differed from the rolling average by more than a pre-set amount. The program could also be configured to test the new data set automatically, include it in the rolling average if every point was valid and trigger the next voltage sweep. Acquisition was stopped automatically when any point in the new set was too far from the rolling average.

\* As pointed out by Speakman (1993), if the first term in (5.7) dominates then this procedure actually loses all information from points \(f_m, f_{m-1}\) and \(f_{m+1}\)!
5.2.3 Additional Features of the D.C. Measurement System.

The CIL6580 and control program may be used to simultaneously measure the current, conductance and capacitance of a sample. The latter two quantities are read, via additional ADCs, from the in-phase and quadrature outputs of a lock-in amplifier set up as shown in figure 5.1. Linear calibration models enable these readings to be stored as accurately calibrated voltages. Alternatively, by entering the lock-in sensitivity and the modulation amplitude and frequency, the conductance and capacitance values are calculated immediately. Figure 5.8 is an example of data taken in this way. It also compares conductance–voltage characteristics obtained by the a.c. and d.c. methods.

The program may be configured to make multiple voltage sweeps - either averaging the readings (as described above) or storing them individually - and then to wait for a pre-defined period of time before making another set of sweeps. During this waiting period the CIL6580 outputs an adjustable d.c. voltage. The program may also be set to start a new sweep immediately or wait until prompted by the user to continue. These facilities have recently been used to investigate time- and voltage-dependent annealing effects in MIM tunnel junctions (Holden 1992).

5.2.4 Possible Improvements to the D.C. Measurement System

The d.c. measurement system is too slow. Measurement, processing and plotting of a single current–voltage point took 45 ms, of which 16 ms was spent communicating data between the CIL6580 unit and the computer via the serial RS-232 interface (at 9600 baud). Communication time could have been reduced to less than 0.5 ms by

![Fig. 5.8. Comparison of conductance–voltage characteristics measured by the a.c. modulation technique (single sweep) and by numerical differentiation of the average of 382 I–V sweeps. The sample is an evaporated Pb/Al–Al₂O₃–Al proximity-effect junction.](image-url)
using a parallel IEEE-488 interface; however, this required additional hardware which was unavailable. A further 18 ms could be gained simply by using a faster host computer.

The computer program that controls the CIL6580 unit operates in a local way. For each measured point on the I–V curve an ‘output voltage’ command must be sent to the CIL6580 from the computer. The computer then delays before sending a ‘trigger ADC’ command which prompts the unit to read a voltage and send the resulting number back to the Macintosh. It should be possible to run this simple source-sense cycle from a program written with the CIL6580’s built-in programming language. A table of output voltages would be pre-defined in the unit’s RAM together with controlling code. On receiving a single ‘trigger acquisition’ command the internal program would be executed and data would be collected from the appropriate ADCs and stored in the unit’s own memory. After collecting data from a complete voltage sweep the whole set of points could be transmitted to the computer. It is difficult to estimate the speed advantage that would be gained by operating the CIL6580 unit in this remote way. A pseudo-parallel-processing arrangement could possibly be used, with the host computer analysing previous data sweeps whilst the CIL6580 unit collected a new set. In addition, overheads due to communication of small chunks of data would be avoided.

Preliminary investigations indicate that the simplicity of the CIL6580’s internal language would make this a formidable programming problem. Difficulties are envisaged with attempts at synchronising the two programs and hardware handshaking would probably be required. Documentation of the CIL6580’s hardware and software is extremely poor and the manufacturer is reluctant to answer technical enquiries.

Rather less fundamental changes could be made to the program currently used. The program is only able to calibrate the output of a lock-in amplifier positioned in the circuit as shown in figure 5.1. An option to calibrate a lock-in placed across the arms of the bridge circuit would also be useful and would be a simple modification. The original intention was to read voltages from only one ADC. Routines to read an additional two channels were ‘tacked on’ rather than logically integrated into the code. It should be possible to write a more general data-logging program that is able to use all eight ADCs and four DACs and that can apply a general calibration procedure to each channel. One way of doing this would be to use a higher-level programming language, such as LabView*, which is specifically designed to simplify computer interfacing.

* Manufactured by National Instruments, Newbury, Berkshire.
CHAPTER 6
EXPERIMENTAL DETAILS
II. JUNCTION FABRICATION

The literature describes several types of tunnel junction used to study high-temperature superconductors. These junction fabrication methods are briefly reviewed in this chapter. Details are then presented of the planar and point-contact junctions that were used to obtain the results described in chapter 7.

6.1 Review of Junction Types

6.1.1 Planar Junctions

A planar tunnel junction is a rigid ‘sandwich’ structure consisting of two conducting electrodes separated by a thin ($\approx$1 nm) insulating layer. Electrodes of many metals can be formed simply by thermal evaporation of bulk material from a resistively heated crucible. If this is not possible other thin film deposition techniques (e.g., sputtering) can be used. After deposition of the first electrode onto a suitable substrate the insulating tunnel barrier must be formed. Aluminium is often used as the first electrode because it readily forms a thin continuous oxide layer of suitable thickness for tunnelling, simply by exposure to air. With less convenient base electrode materials an ‘artificial’ barrier, consisting of a thin film of a suitable insulating material (e.g., MgO), must be used. When the second electrode is deposited the tunnelling interface is encapsulated. This makes the junction robust and resistant to atmospheric degradation. Because the device is rigid the conductance–voltage characteristics are not disturbed by mechanical vibrations, or by changes in geometry due to thermal contraction or expansion when the junction temperature is changed.

The biggest problem with this type of tunnel junction is obtaining a uniform barrier and electrodes over the whole junction area - which tends to be large ($>0.05$ mm$^2$). The insulating layer may be shorted by pin holes and the electrodes may be inhomogeneous on a length scale smaller than the junction dimensions. Complications will also be caused by traps in the barrier (section 3.6) or by degradation of the electrode material at its interface with the barrier. These difficulties are particularly acute with high-$T_c$ superconductor electrodes (sections 4.3–5).

Early high-$T_c$ superconductor planar junctions used the intrinsic insulating surface layer that forms on these materials as the tunnelling barrier. Junctions were formed
either by evaporating a metal counter electrode onto the as-prepared surface of the superconductor, or simply by pressing small blobs of indium onto the surface (Geerk et al. 1988). Probably for the reasons outlined above, these junctions had anomalous, irreproducible tunnelling characteristics which contained little or no convincing evidence of a superconducting gap (section 6.2). By etching the sample’s surface and allowing a new surface layer to form (by exposure to air for one hour), Geerk et al. (1991) improved the reproducibility of the tunnelling characteristics and reduced the strength of the voltage dependent background. The gap-like feature, however, remained weak and had a very un-BCS-like appearance.

Many attempts have been made to prepare more controlled artificial tunnelling barriers. These were mostly unsuccessful, yielding no improvement; however, Matsumoto et al. (1992) recently made a significant breakthrough. A Bi$_2$Sr$_2$CuO$_6$ thin film barrier was deposited by laser ablation onto a large Bi$_2$Sr$_2$CaCu$_2$O$_8$ single crystal (figure 6.1). Because these materials have well matched lattice constants the interface is epitaxial and disturbance of the high-temperature superconductor is minimised. This may be important to prevent formation of a degraded non-superconducting layer. A further advantage of Bi$_2$Sr$_2$CuO$_6$ is that it is semiconducting with a low barrier height. This allows use of thick barriers (≈ 29 nm), which reduces the possibility of pin holes.

**Fig. 6.1.** Epitaxial planar tunnel junction with artificial barrier.
Using this fabrication process Matsumoto reproducibly obtained very BCS-like tunnelling characteristics. These could be modelled extremely well by Dynes’ lifetime-broadening model* with $2\Delta/k_B T_c = 4.7$ and a very low broadening value of $\Gamma = 2$ meV. The temperature dependence of the energy gap agreed well with the BCS model of $\Delta(T)$ for a transition temperature of 84 K. This ‘gap-opening’ $T_c$ is identical to that estimated from resistivity measurements of the crystal - indicating complete absence of degradation of the superconductor at the junction interface.

An alternative ‘edge-junction’ configuration, depicted in figure 6.2, has the stability of a planar junction but a much smaller area of less than $0.1 \mu m^2$. The junction is made by depositing a thick MgO layer onto a YBa$_2$Cu$_3$O$_y$ thin film. Ion milling at an angle exposes a sloping edge, onto which a thinner MgO tunnelling barrier is deposited, followed by the YBa$_2$Cu$_3$O$_y$ counterelectrode and silver contacts. The YBa$_2$Cu$_3$O$_y$ and MgO layers are prepared by laser ablation and sputtering respectively, with all operations carried out in situ. Junction quality is not yet ideal and the tunnelling characteristics display many of the usual anomalous features of high-$T_c$ planar tunnel junctions. In particular the gap-opening temperature is well below the film’s bulk transition temperature, indicating degraded superconductivity at the interface (possibly due to ion-milling damage). However, control over junction properties, including crystal orientation of the high-$T_c$ superconductor films, has been achieved and is improving. In particular, Ying et al. (1992) obtained very low conductance in the gap region.

### 6.1.2 Point-Contact Junctions

The first reliable measurements of energy gaps in the A15 superconductors (e.g., Nb$_3$Ga) were achieved by Levinstein and Kunzler (1966) using point-contact

---

* This model produced a marginally better fit than a Gaussian distribution of gaps.
junctons. In these experiments a rod of aluminium was etched to a sharp point and heavily anodised. By pressing this point onto the surface of a superconducting sample the thick oxide layer was mechanically deformed until it was thin enough for a tunnel current to flow. Junction resistances varied from $10^2 \, \Omega$ to $10^5 \, \Omega$, roughly controlled by pressure on the point. The point-contact method permits tunnelling measurements on materials that cannot be incorporated into planar junctions, either because they cannot be prepared in thin film form or because continuous tunnel barriers cannot be prepared on their surfaces.

An additional advantage of point-contact tunnelling is that the small contact area results in a local probe of electrode properties, so the technique is suitable for measurements on small crystals and inhomogeneous samples. Hoffstein and Cohen (1969) successfully measured the gap anisotropy of an Nb$_3$Sn single crystal. Large-area planar junctions made from vapour-deposited films of this material had previously resulted in broadened conductance–voltage characteristics, due to averaging over randomly oriented crystallites within the film. A similar story with Ba$_{1-x}$K$_x$BiO$_3$ samples is described in appendix B.

Huang et al. (1990a) compared the tunnelling characteristics of gold point contacts on an Nb crystal, where the tunnelling barrier was native oxide, with those of Nb planar junctions. They concluded that phonon structure can be observed by point contacts, and that anomalous effects such as depression of the energy gap and shifting of phonon peaks, which have been reported for Pb (Tsui 1970), are not intrinsic to the point-contact method.

If the point contact is pressed firmly into the sample the surface layer may be disrupted. Although this may be useful in cleaning away degraded material, damage-induced crystalline disorder will prevent anisotropy measurements.

The importance of measuring a gap-opening transition temperature was stressed in section 4.6. Unfortunately, thermal expansion renders the spear–anvil type of point contact unstable to temperature changes, resulting in an uncontrolled contact force, mechanical instability and possibly loss of contact. For this reason there are only a few examples of point-contact measurements of an energy gap’s temperature dependence (e.g., Edgar et al. 1987). To overcome this problem Reinertson et al. (1992) developed the wire–wedge geometry, in which a thin wire is held in tension across a wedge-shaped superconductor. An increase in length of the wire caused by heating is compensated by thermal expansion of the wedge, so the contact force is held nearly constant. In addition, expansion of the wire on either side of the wedge is equal so the contact position does not change.
6.1.3 Break Junctions

Many of the early successful tunnelling experiments on ceramic La$_{2-x}$Sr$_x$CuO$_{4-y}$ and YBa$_2$Cu$_3$O$_y$ samples were performed by Moreland et al. (1987a,b) using break junctions. These are basically crude point-contact junctions formed by breaking a sample in two and pushing the pieces back together. This is usually achieved by gluing the sample to a flexible substrate that can be bent and relaxed under the control of a screw mechanism. The only advantage of a break junction is that the tunnel junction is formed between freshly exposed bulk material, which is free from troublesome surface layers. Despite the rather uncontrolled nature of junction formation, Mandrus et al. (1991) successfully investigated anisotropy in the tunnelling characteristics of single-crystal Bi$_2$Sr$_2$CaCu$_2$O$_8$ break junctions (section 4.3.2).

6.1.4 Scanning Tunnelling Microscopes

The advantages of point-contact tunnel junctions are enhanced if contact is actually avoided. The tunnelling barrier at the heart of a scanning tunnelling microscope* (STM) is a vacuum gap formed between two electrodes, one of which is a sharp metal tip while the other is a flat conducting sample. The vacuum gap is controlled by moving the tip with a piezoelectric transducer, and is held constant by a feedback loop formed by driving the piezo with a voltage derived from the tunnel current. Because the current is extremely sensitive to barrier width the tip–sample spacing is held constant with sub-Ångström resolution; using (3.1), a typical barrier height of 5 eV gives an order of magnitude decrease in current for a width increase of only 0.1 nm. The junction width is determined by the bias applied to the junction (typically a few millivolts to a few volts) and the tunnel current set point (typically $10^{-11}$–$10^{-8}$ A).

If the tip is rastered over the surface of the sample using an additional piezoelectric element, the tip must be moved up and down in sympathy with topographic variation to maintain a constant current (and hence tip–separation). The feedback signal (i.e., the piezo-voltage) is proportional to the height variation†, so an atomic-resolution map of the surface topography is produced.

At selected points on the surface the current–voltage characteristics of the tunnel junction may be measured. Scanning is halted and the feedback loop broken whilst the

---

* See van de Leemput and van Kempen (1992) for an excellent review of scanning tunnelling microscopy.

† This assumes that the workfunction and density of states are constant. Variation in either of these quantities will result in ‘false’ topographic structure (see section 4.3.1).
junction voltage is ramped rapidly through the bias range of interest and the tip–sample current recorded. An STM with appropriate controlling software can simultaneously acquire topographic and spectroscopic information about the sample, making possible studies of the correlation of conductance–voltage characteristics with surface features or even ‘atomic-site tunnelling spectroscopy’.

With a superconducting sample an STM should be able to spatially map the energy gap and measure the local density of excitation states. The area sampled by an STM tunnel junction is even more localised than that of a point contact (see section 4.3.1) and surface damage is completely avoided if the tip and sample do not touch. For these reasons an STM is superior to planar junctions and point contacts in studies of highly anisotropic or inhomogeneous materials (e.g., NbN thin films, Kashiwaya et al. 1991).

Unfortunately, vacuum tunnelling was not achieved in early STM studies of high-temperature superconductors. To detect a tunnel current it was necessary to push the tip through an intrinsic insulating surface layer (section 4.5), so the STM was being used merely as a well controlled point contact. Reports of true vacuum tunnelling have increased however, as surface-cleaning techniques have improved (section 4.3.1).

The only disadvantage of scanning tunnelling microscopy as a junction fabrication method (apart from the increased complexity and expense of the equipment) is the instability of the junction to temperature changes. As with point-contact junctions, measurement of the temperature dependence of a superconductor’s energy gap is extremely difficult.

### 6.2 Planar Junctions

In this study conventional Al/Al₂O₃/Pb S1N junctions were used to test the $I$–$V$ measuring equipment. An aluminium strip 1 mm wide and 100 nm thick was deposited by thermal evaporation of bulk metal onto a clean microscope slide. The electrode shape was defined by a mask, which could be completely covered with a shutter to stop deposition. When the bulk metal was heated the pressure in the vacuum chamber rose, initially to a high level, but by carefully outgassing the metal prior to opening the shutter the pressure during deposition could be maintained below $2 \times 10^{-6}$ torr. The thickness of the deposited metal was monitored by measuring the drop in frequency of a quartz crystal oscillator as metal became attached to the crystal and its mass increased.

The aluminium electrode was oxidised by exposing it to air for 30 minutes. The

* Pumped to a base pressure of $3 \times 10^{-7}$ torr by a liquid-nitrogen-trapped, oil diffusion pump.
evaporation chamber was then pumped back down and a Pb strip of similar dimensions was deposited across the Al strip to form a junction of area 1 mm$^2$. In practice several cross strips of Pb were deposited simultaneously to produce a device with more than one junction. Four-terminal connections were made by attaching copper wire to the ends of each electrode with pressed indium.

Preliminary attempts at fabricating high-temperature superconductor tunnel junctions consisted of direct deposition of Pb electrodes onto the surface of a ceramic YBa$_2$Cu$_3$O$_y$ pellet. Before being placed in the evaporator the ceramic was abraded with emery paper then blown with compressed air to remove small particles of material. The surface was still quite rough and a 1 μm thick Pb layer had to be deposited before the electrode appeared continuous under an optical microscope.

Four-terminal connections to the Pb and YBa$_2$Cu$_3$O$_y$ were made with silver dag; however, this produced junctions that did not show SIN- or SIS-type characteristics even when the lead was superconducting. Presumably the silver dag was making direct contact with the ceramic through holes that remained in the Pb layer and conduction was via a metallic short rather than tunnelling. To overcome this problem gallium metal was used as a solder. This required no heating and wetted the Pb but not the ceramic, resulting in junctions which at 4.2 K had a conductance minimum at zero bias, as expected for tunnelling from a superconducting electrode. The temperature dependence of this feature confirmed that an SIN junction had been formed between superconducting Pb and a normal material on the surface of the bulk superconducting YBa$_2$Cu$_3$O$_y$ pellet†. This result also indicates the existence of an intrinsic tunnelling barrier at the sample’s surface and suggests that junctions could be formed by the point-contact technique.

### 6.3 Point-Contact Junctions

Preliminary investigations showed that BCS-like conductance–voltage characteristics could be obtained by pressing a sharp aluminium or tungsten point into the surface of a ceramic YBa$_2$Cu$_3$O$_y$ sample (Edgar et al. 1987). The sample was mounted on the end of a modified dip stick and inserted into an helium storage dewar which sat on the laboratory floor. The tunnelling tip was attached to the end of a ‘push rod’, which ran from the sample to the top of the dewar, through the dip stick’s copper-nickel tube. Helium gas was prevented from escaping from the dewar (other than via the return line) by a Plasticine seal at the top of the tube. The pressure and position of the point on the sample was controlled very crudely by hand. A new sample could rapidly be

† In a separate experiment the Ga/Pb interface was shown to be ohmic.
prepared for investigation and, once damaged, the tunnelling tip could be replaced without removing the sample from the dewar, simply by withdrawing the push rod from the dip stick. The junction temperature could be crudely controlled by lifting the end of the dip stick out of the helium into the neck of the dewar. To reduce temperature fluctuations the sample was mounted on a copper block and the end of the dip stick was enclosed in a copper can. Temperature was measured using a calibrated carbon-glass resistor mounted in a hole in the copper block close to the sample. This arrangement produced a large quantity of useful results; however, the contacts were unstable and could not be controlled with any accuracy. In an attempt to improve the quality of results a cryostat insert was built, which could manipulate the point contact with nanometre precision.

6.3.1 Design Requirements

To form a point-contact tunnel junction a sharp metal tip must be moved in a finely controlled way into contact with a sample’s surface. Great finesse is required to prevent damage to the tip or sample when they make contact and to allow the contact resistance to be adjusted. Although final positioning of the tip must be done very precisely the initial tip–sample separation must be large enough to allow the sample and tip to be changed easily without damage. A major objective of the design is a mechanism for tip manipulation with a precision : range ratio of about 1 : $10^8$.

For the point-contact measurement system to be efficient some method of moving the tunnelling tip parallel to the surface must be available. Several positions on a surface may then be examined without having to bring the whole insert back up to room temperature and the tip may be moved away from regions that are contaminated or damaged by contact with the metal point. It was also hoped that lateral movement of the tip would enable spatial mapping of the superconductor’s energy gap.

Once in position the contact must be stable enough to prevent damage through excessive, uncontrolled movement and to allow noise-free measurement of the conductance between the tip and sample. The design must therefore minimise the coupling of vibrations with the spacing between the tip and sample.

6.3.2 Vibration Isolation

The principal source of vibration that could disturb the point contact is movement of the laboratory walls and floors (Pohl 1986). These structures are excited by machinery operating at mains frequency and give rise to vibrations of typically 10–100 Hz, with displacements of up to 1 μm at the resonant frequency of about 20 Hz. Problems also arise due to acoustic vibrations, which span a wide frequency range (10 Hz–10 kHz).
and have large amplitude. Disturbances of less than 3 Hz are produced by general laboratory activity. The preliminary point-contact experiments were frequently disturbed by this background of vibrational noise and some form of isolation was considered necessary.

The objective is to reduce the amplitude of vibrations appearing in the gap between the tip and sample. In a point-contact experiment the tip is in contact with the sample and an oxide layer (or other insulator) intrinsic to one of the two surfaces acts as the tunnelling barrier. If this layer is thick it may need to be thinned by pressure from the tip, or the tip may need to be embedded firmly in the surface, in order for a tunnel current to be observed. Although the specifications of the vibration isolation are clearly less stringent than those required for an atomic-resolution STM the problems faced are similar. The first line of defence against ambient vibrations is to attenuate high-frequency components by weakly coupling the body of the STM to the outside world, via some form of elastic damping element. The effect of residual low-frequency disturbances that do manage to couple with the body of the STM are then minimised by making the mechanical path from the tunnelling tip to the sample as rigid as possible.

To clarify this description a simple model may be employed. Consider the response of the body of the STM assembly to an external vibration when the two are connected by a spring with undamped resonant frequency $\omega_i = \sqrt{k/m}$ and a damping element (figure 6.3). Solving the equations of motion for this system gives the transfer function of the coupling,

$$F = \left( \frac{B}{A} \right) \exp(-j\phi) = \frac{1 + j\Omega/Q}{(1 - \Omega^2) + j\Omega/Q} \quad (6.1)$$

**Fig. 6.3.** Spring and dashpot model of vibration isolation.
where \[ \Omega = \left( \frac{\omega}{\omega_i} \right) \] and \[ Q = \frac{m\omega_i}{\alpha} = \left( \frac{\text{Im} F}{\text{Re} F} \right)_{\omega = \omega_i}. \]

The amplitude of this function, \(|F_i|\), is plotted in figure 6.4. The important properties of this behaviour are:

- Well below the resonant frequency (\(\Omega \ll 1\)) there is no attenuation.
- Near the resonant frequency vibrations are amplified, with a maximum at resonance of \(\sqrt{1 + Q^2}\), i.e., amplitude is decreased by damping.
- Vibration is attenuated at frequencies above \(\sqrt{2}\omega_i\) and at high frequencies falls as \((Q\Omega)^{-1}\); so damping reduces the efficiency of isolation.

To optimise isolation of the STM from external vibrations the resonant frequency of the coupling should be as low as possible. However, low resonant frequencies require inconveniently large displacements and the lowest practical resonant frequency is about 2 Hz. If possible, the level of damping should be adjusted to find a compromise between reducing the amplitude of disturbances at \(\omega_i\) and maintaining sufficient isolation at high frequencies.

The point-contact insert used in this work was crudely isolated from building vibrations by connecting the cryostat to a heavy metal frame via an inflated car inner-tube (Cockburn 1988). The vibrational characteristics of this system were measured

![Fig. 6.4. Frequency characteristics of the damped-spring vibration isolation, |F_i|, the tip–sample gap, |F_g|, and the total response, |F_T|=|F_iF_g|, using experimentally determined values for the quality factors and resonant frequencies: \(f_i = 5.3\) Hz, \(Q_i = 15.2\), \(f_p = 816\) Hz, \(f_s = 1130\) Hz, \(Q_p = Q_s = 30\).](image)
with an accelerometer and spectrum analyser. The dominant modes were a vertical motion with resonant frequency 5.28 Hz and $Q = 15.2$, and a swinging movement with natural frequency 1.96 Hz and quality factor 51.2. Figure 6.4 shows that this simple arrangement should be effective at attenuating the important 20 Hz building vibrations. However, the cryostat was exposed to the laboratory air so acoustic vibrations were able to by-pass the inner tube and couple directly with the point-contact mechanism. In addition, the vibrations of less than 3 Hz due to random vibrations are below the resonant frequency so remain unattenuated. To avoid these sources of noise all point-contact experiments were carried out at night when the laboratory was empty. An improvement would be to complete the external vibration isolation by placing the cryostat inside an airtight chamber with anechoic panelling. All cables that linked the point-contact insert to the measurement and control electronics were suspended on weak elastic (resonant frequency < 1 Hz) before being attached to the insert. This attenuated vibrations that would otherwise have by-passed the vibration isolation by travelling along the wiring.

In a tunnelling experiment it is the relative movement of the tip with respect to the sample that is the important parameter. The response to residual vibrations in the body of the STM may be examined by considering the mechanical path between the tip and sample to consist of two damped-spring elements (figure 6.5).

The tip–sample gap is $z_g = z_p - z_s$ and vibrational noise in $z_g$ is related to vibrations in the base by the transfer function $F_g = F_p - F_s$. Both $F_p$ and $F_s$ are given by (6.1) with appropriate values for the resonant frequency and damping coefficient. $|F_g|$ is plotted in figure 6.4 for resonant frequencies of 816 Hz and 1130 Hz, both elements having an estimated $Q$ of 30. In the limit of high and low frequency,
\[
\lim_{\omega \to 0} |F| = -\frac{1}{\omega^2} + \frac{1}{\omega_s^2} - \frac{1}{\omega_p^2}
\]

and
\[
\lim_{\omega \to \infty} |F| = -\frac{1}{\omega} \left( \frac{\omega_p}{Q_p} - \frac{\omega_s}{Q_s} \right)
\]

Clearly low-frequency vibrations move both the tip and sample by the same amount and no differential movement occurs.

The total response \(|F_g|\) of the gap to external vibrations is given by the product of the response of the vibration isolation and the response of the gap. The amplitude of vibrations appearing in the gap would be reduced by lowering \(\omega_q\) and by increasing the structural rigidity of the tip–sample path to raise the lowest resonant frequency. The resonant frequency of a rectangular beam of length \(l\) and thickness \(t\) constructed from material with Young’s modulus \(E\) and density \(\rho\) is given by

\[
f \propto \frac{t}{l^2} \sqrt{\frac{E}{\rho}}
\]

This shows that to achieve a high resonant frequency components of an STM must be short (low \(l\)) and thick (large \(t\)). The current trend in STM design is to use small rigid components and to simplify the design in order to reduce the number of elements in the system.

Despite considering the principles outlined above whilst designing and building the point-contact mechanism the lowest resonant frequency of the finished insert was disappointingly small. Two methods were used to measure this important parameter. Firstly the electronics used to control a commercially manufactured STM* was operated in constant current mode to stabilise the tunnelling current. Variations in tunnelling distance due to vibrational disturbance resulted in noise in the feedback signal, which was input to a spectrum analyser. The second method used a small SmCo magnet mounted on a tunnelling tip and placed inside a pick-up coil. Vibrational motion of the magnet in the coil induced a small signal, which was amplified and input to the spectrum analyser. In both cases measurements were made with the point-contact mechanism mounted in the cryostat to simulate the experimental configuration. Positive feedback was used to emphasise resonant frequencies by exciting the point contact with a small vibrating motor or loudspeaker driven by the noise signal. This method resulted in power spectra that were sensitive to the

* Loaned by W.A. Technology, Cambridge.
Fig. 6.6. Vibrations appearing at the point-contact tip, measured by the magnetic method with white-noise excitation.

Fig. 6.7. Estimate of the amplitude of vibrations appearing at the point contact based on a two-component model of the tip–sample path and using experimentally determined values for the quality factors and resonant frequencies.
position of the speaker in relation to the cryostat, and was abandoned in favour of stimulation by a white-noise signal. Power spectra measured in this way were reproducible and results from the magnetic and tunnel current methods agreed very well (figure 6.6).

An estimate of the amplitude of vibration appearing in the gap may be made by multiplying the total response $|F_r|$ of the system by the ‘typical’ noise spectrum suggested by Pohl (1986). This model spectrum consists of a $1/f$ background with superimposed characteristic resonances. Figure 6.7 predicts that successful imaging of atoms is beyond the capabilities of this instrument; the vibrations at 20, 50 and 100 Hz all exceed 1 pm, which is generally regarded as the required level. In practice the amplitudes are probably even larger than calculated, because acoustic vibrations are able to by-pass the isolation.

To ensure that all vibrations are below 1 pm in amplitude the lowest resonance frequency must be greater than 3 kHz. An STM with this specification will be able to image atoms but must be operated in a moderately quiet environment. By raising the resonance to 100 kHz rejection of vibrations due to rigidity of the structure is so high that no vibration isolation is necessary and the unit could be operated on the laboratory bench. Most recent STM designs have resonant frequencies between 5 kHz and 15 kHz.

Because point-contact experiments are carried out with the tip in passive contact with the sample surface the discussion above has been restricted to analysing the amplitude of response to external vibrations. If the instrument is to be operated with active feedback to keep the tunnelling current constant then the stability of the response must also be considered. This subject is beyond the scope of the present work and the interested reader is referred to a review by Grafström et al. (1990).

**6.3.3 Fine Control of the Tip–Sample Separation**

The need for precise control combined with a large range of movement resulted in a design that separated the positioning of the tunnelling tip into three stages (table 6.1). This scheme was motivated in part by the prevailing fashions in STM design when the point-contact insert was built.

Without detailed information about the properties of the surface layer found on the high-$T_c$ superconductors it was impossible to predict the resolution needed to make a high-quality tunnel junction. Observation of tunnelling with only crude control of the tip (in preliminary experiments) suggested that a fine-control stage was unnecessary. However, should this level of control be required at a later date then adaptation of the insert would be greatly simplified by the inclusion of a piezoelectric element in the
Table 6.1. Movement of tunnelling tip towards the sample.

<table>
<thead>
<tr>
<th>Description</th>
<th>Range</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial positioning</td>
<td>Sample moved at room temperature by M3 screw</td>
<td>15 mm – 200 µm</td>
</tr>
<tr>
<td>Coarse control</td>
<td>Tip moved at 4.2 K by combined differential screw and differential spring</td>
<td>200 µm – 80 nm</td>
</tr>
<tr>
<td>Fine control</td>
<td>Tip moved at 4.2 K by piezoelectric tube</td>
<td>260 nm – contact</td>
</tr>
</tbody>
</table>

original design. The piezoelectric stage would also provide a mechanism for transverse movement of the tip, and fine control of the contact force would allow investigation of pressure effects. In addition, it was hoped that the point-contact mechanism would be rigid enough to be operated as an STM.

The piezoelectric element used in this work was a radially polarised PZT-5A ceramic tube. Voltage applied between electrodes on the inner and outer surfaces of the cylinder caused the tube to change length, so moving the tunnelling tip, mounted axially on an end face, towards or away from the sample’s surface. The tensor component of piezoelectric response that relates changes in the tube’s length to radial electric field is $d_{31}$, defined as strain per unit field when there is no stress on the material. If the tube has a free length $l$ and wall thickness $t$ then

$$\frac{\Delta l}{l} = -d_{31} \frac{(V_o - V_i)}{t}$$

(6.2)

Where $\Delta l$ is the change in length and $V_i$ and $V_o$ are the voltages on the inner and outer electrodes. A piezoelectric crystal in this form was used in a previous point-contact insert built by the Low Temperature Physics Group for spectroscopic studies of magnetic metals (McLean 1984).

If the outer electrode is divided into four quadrants by removing metal along axial lines at 0, 90, 180 and 270 degrees (figure 6.8) then the length of each segment may be independently controlled. A voltage applied to one of the electrodes will cause the tube to both extend and bend, so a suitably mounted tip may be moved with three degrees of freedom (Binnig and Smith 1986). ‘Piezotubes’ have several advantages over piezoelectric tripod structures previously used as scanning elements in STMs. They have higher resonant frequencies (10 kHz compared to less than 5 kHz for tripods) and are more compact - leading to smaller STMs with greater rigidity. They are also more sensitive and so have a greater range of movement. For these reasons piezotubes have replaced tripods in most recent STM designs.
To estimate the range and precision of movement of a segmented piezotube, consider a simple model in which the length of each quadrant depends only on the electric field between its outer electrode and the inner electrode. The voltages on the two \( y \)-electrodes are identical but otherwise undefined, so as electrodes \( p \) and \( n \) change length the top of the tube will move only in the \( x \)-direction and \( z \)-direction. If the distorted tube is modelled as a segment of a torus, as shown in figure 6.9, then the centre of the top of the tube is displaced by

\[
\delta x = \frac{(\delta p - \delta n)}{2} \left( \frac{l + \delta m}{\phi} \right)
\]

\[
\delta z = \frac{(\delta p + \delta n)}{2}
\]

where \( \phi \) is the tube diameter, \( \delta m = (\delta p + \delta n)/2 \) and \( \delta p \) and \( \delta n \) are the length changes of the \( p \) and \( n \) electrodes respectively. Calculating \( \delta p \) and \( \delta n \) from (6.2) and assuming \( \delta m \ll l \) gives

\[
\delta x = \left| \frac{d_{31}}{2t} \right| \frac{l^2}{\phi} (V_p - V_n)
\]

\[
\delta z = \left| \frac{d_{31}}{2t} \right| l (V_p + V_n - 2V_z)
\]

(6.3)

The end of the tunnelling tip has an additional \( x \)-displacement of \( T \sin \theta \) or

\[
\delta x_{tip} = \left| \frac{d_{31}}{2t} \right| \frac{2T}{\phi} (V_p - V_n)
\]

(6.4)
Fig. 6.9. Geometry of the piezotube movement.

Table 6.2. Properties of piezotube.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material</td>
<td>PZT-5A</td>
</tr>
<tr>
<td>$d_{31}$ at 300 K</td>
<td>$-171$ pm V$^{-1}$</td>
</tr>
<tr>
<td>Length</td>
<td>12.7 mm</td>
</tr>
<tr>
<td>Outside diameter</td>
<td>6.35 mm</td>
</tr>
<tr>
<td>Wall thickness</td>
<td>0.51 mm</td>
</tr>
<tr>
<td>$Y_{1j}^E$ at 300 K</td>
<td>61 GN m$^{-2}$</td>
</tr>
</tbody>
</table>
An increase in $V_z$ always decreases the length of each quadrant by the same amount and so moves the tip away from the sample without any transverse motion. In contrast, a change in $V_p$ will in general move the tip in both the $z$- and $x$-directions. This coupling may be cancelled by using an antisymmetric power supply to make $V_p = -V_n$.

The range of tip movement is limited by the maximum electric field that may be applied across the piezoelectric ceramic without depolarizing the material. No information was available from the manufacturer of the piezotube* about the low-temperature value of this field, which at room temperature is 1 kV mm$^{-1}$; however, the tube was regularly operated with fields of up to 0.8 kV mm$^{-1}$ without any obvious decrease in sensitivity.

A unipolar HT voltage source was built by the Cavendish Laboratory Electronics Department† to control the piezotube deflection. This provided two voltages in a 0 to +300 V range, which were applied to the $X_p$ and $Y_p$ electrodes for movement in the $x$ and $y$ direction, and a voltage of 0 to −500 V, which was applied to the central electrode for movement of the tip towards the sample. The $X_n$ and $Y_n$ electrodes were connected to the HT supply’s ground. A better configuration would be to apply an antisymmetric voltage to opposite elements of the outer electrodes and ground the central electrode, which could then shield wires going through the tube to the tip. Movement in the $z$-direction would be achieved by adding an identical voltage to each of the four outer electrodes.

The piezotubes used in this work (table 6.2) resulted in a sensitivity in each direction of 0.87 nm V$^{-1}$ and a range of 260 nm. This calculation has included a reduction in $d_{31}$ at 4.2 K to approximately 20% of the room-temperature value (Gerson 1962). The outputs of the high-voltage supply contained 6 mV noise (peak–peak over 40 MHz bandwidth) in a 100 V d.c. output. The resulting tip disturbance of less than 3 pm is just below the amplitude of vibration expected due to coupling with external vibrational noise (figure 6.7).

The expressions for piezotube movement derived from geometrical arguments may be compared with the finite-element-analysis calculations by Carr (1988). This numerical model divided a piezotube into 218 identical elements and by considering the mechanical stress, stiffness, strain and piezoelectric stress on each element calculated the displacement of important points on the tube as one electrode was energised. The

* Vernitron Ltd., Thornhill, Southampton.
† Job number J473
results were found to be linear on the first iteration and higher-order corrections were very small even for large electrode voltages. Carr’s results for the $x$- and $z$-movement of a range of piezotubes are shown in figure 6.10 together with the predictions of the geometrical model and some experimentally measured values. The agreement between these results suggest that (6.3) is a reasonable estimation of the actual piezotube movement.

![Graph showing $x$- and $z$-movement of a PZT-5H piezotube at 300 K ($d_{31} = -2.74 \text{ Å V}^{-1}$), calculated by finite-element analysis (Carr 1988) and from the geometrical model (6.3). Also shown are some experimentally measured values from Carr’s survey of published data. $l =$ tube length; $\phi = $ diameter; $t = $ wall width.]

**Fig. 6.10.** $x$- and $z$-movement of a PZT-5H piezotube at 300 K ($d_{31} = -2.74 \text{ Å V}^{-1}$), calculated by finite-element analysis (Carr 1988) and from the geometrical model (6.3). Also shown are some experimentally measured values from Carr’s survey of published data. $l = $ tube length; $\phi = $ diameter; $t = $ wall width.
Fig. 6.11. Tilt angle of a tip attached to the top of a PZT-5H piezotube at 300 K, calculated by finite-element analysis (Carr 1988) and from (6.4).

Fig. 6.12. Difference in the tilt angle predicted by finite-element analysis and the geometrical model. The geometrical model constrains the top of the tube to lie on the dotted line, i.e., perpendicular to the walls. In the finite-element calculation the top is free to move and distorts to the shape shown by the solid line.
Carr also calculates the additional $x$-displacement of the tip and presents it in the form of a tilt angle (i.e., $\delta x_{\text{tip}}/T$). The discrepancy between this angle and that calculated from the geometrical analysis (figure 6.11) is caused by the different behaviour of the top of the tube in the two models; in the geometrical case the top is always perpendicular to the tube walls whereas in the finite-element analysis it is free to move and so distorts (figure 6.12). In this work a plastic ‘cap’ with a hole tapped in its centre was fixed rigidly to the end of the piezoelectric cylinder with Araldite. Tips were attached to the tube by screwing the 8BA screws on which they were mounted into the hole. The solid plastic and epoxy probably lead to behaviour more like that predicted by the geometrical model.

The plastic material used to make the piezotube’s cap was Dexter Hysol electrical cast product CP3-4285. This has a similar thermal contraction to the PZT-5A and may be glued to it with Araldite standard epoxy without causing any damage on cooling to 4.2 K. Hysol was also used as an insulating connector between the piezotube and the metal to which it was attached; however, they could not simply be glued together because the large difference in thermal contraction caused the join to fail. The connection was made by sandwiching the 1 mm thick BeCu beam between two pieces of Hysol as shown in figure 6.13. The recess in the lowest block of Hysol is thinner than the BeCu beam so as the epoxy cures and contracts the metal is clamped between the two pieces of plastic. This connection survived repeated thermal cycles.

**Fig. 6.13.** Tip, piezotube and beam assembly. (not to scale). All parts are glued together with Araldite standard epoxy. Diagonally hatched parts are Hysol plastic. The sample wires, $I+$ and $V+$, are connected to the tip by a solder tag.
6.3.4 Coarse Control of the Tip–Sample Separation

A cryostat insert used in a previous point-contact experiment (Cockburn 1988) was modified to provide coarse control of the tip. The insert included a differential screw with pitch 17–20 \( \mu \text{m} \), which could be turned from the top of the insert to move a plunger inside a vacuum can (figure 6.14). Non-linear movement of the differential screw (figure 6.15) was quite useful because contact between the tip and the sample could be arranged to occur in the higher-precision region. At 4.2 K the full range of the piezoelectric extension was equivalent to only 5° rotation of the differential screw. This is too coarse to be controlled easily by hand and further attenuation of movement was considered necessary. The enhanced precision would also increase the chances of operating the point-contact mechanism successfully without using the piezoelectric control.

Space in the vacuum can was restricted by the pre-existing arrangement of the differential screw, so a differential spring rather than a lever was chosen as the attenuation mechanism. The plunger pushed on a helical spring, which, in turn, pushed on a cantilever, which was fixed to the insert at both ends (figure 6.16). Movement of the centre of the beam, to which the piezotube was attached, is given by

\[
\frac{z_b}{z_s} = \frac{k_s}{k_s + k_b}
\]

Both the beam and compression spring were made from suitably annealed BeCu so that temperature variations in the spring constants did not affect the movement. An early trial using a steel spring failed due to an irreversible change in the material’s elasticity when it was cooled to 4.2 K. The spring constant of a cantilever is

\[
k_b = C \frac{E w t^3}{l^3}
\]

where \( E \) is the Young’s modulus, \( w \) is the width of the beam, \( l \) its length and \( t \) its thickness. For a beam fixed rigidly at both ends \( C \) is 16 whilst if the ends are free to move \( C = 4 \) (Timeshenko and Gere 1987). The cantilever used in the point-contact mechanism was nominally fixed; however, the large forces involved give rise to some movement of the ends so the actual value of \( C \) will be somewhere between these extreme cases. The length and width of the beam were restricted by the available space in the vacuum can so the actual spring constant was determined largely by the thickness of BeCu sheet from which the beam was machined. The thickest material readily available was 1 mm - giving a spring constant of about \( 4 \times 10^5 \text{N m}^{-1} \). With an estimated load (due to the piezotube) of 3 g this would give a resonant frequency of 2 kHz.
Fig. 6.14. Lower section of the point-contact insert showing details of the tip-positioning mechanism. This was enclosed in a vacuum can to protect the sample and piezotube.
blank page - reverse of photograph
Fig. 6.15. Differential screw movement measured with a dial gauge.

Fig. 6.16. Reduction in movement by differential spring.
**Fig. 6.17.** Reduction of movement by the differential spring. The curve is calculated from the individual force–displacement curves of the helical spring and the beam; points are measured using a dial gauge as the differential screw is advanced and retracted.

**Fig. 6.18.** Tip movement as the differential screw is rotated, measured with a dial gauge.
Helical springs with $k_s = 1.7 - 3.6 \times 10^5 \text{ N m}^{-1}$ were manufactured by W.H.Rooke & Co. (Redditch) Ltd and several beams were made by the author. To help select the most appropriate helical spring and beam combination the individual force–displacement curves of the springs and beams, $\delta s(F)$ and $\delta b(F)$, were measured. The helical springs were found to be rather non-linear due to poorly machined end surfaces so the simple linear relationship (6.5) was not applicable. To calculate the combined behaviour we note that the piezo displacement is given by $\delta b(F)$ whilst the displacement of the plunger which presses on the helical spring is given by $\delta b(F) + \delta s(F)$. In equilibrium the force on the two springs is equal so $z_b(z_s)$ can be found from a parametric plot of $\delta b(F)$ and $\delta b(F) + \delta s(F)$ against $F$ (figure 6.17). Once the springs were selected the point-contact mechanism was constructed and the actual displacement measured with a dial gauge. Figure 6.18 shows that the piezotube and tip may be moved over a distance of 200 $\mu$m with a resolution of about 7 $\mu$m/turn, so the range of the piezotube is equivalent to about 14° rotation of the screw. This level of control allowed the tip to be moved into contact with the sample by turning the differential screw by hand whilst monitoring the tunnelling current on an oscilloscope. When contact occurred the current increased rapidly, the coarse control was locked in position and all further manipulation of the tip done with the piezotube.

To reduce initial contact between the tip and sample a simple servomechanism was used (appendix A). During the coarse–approach stage the piezotube was extended and the differential screw turned by a low-geared d.c. motor*. The tunnel current was monitored and when it rose above a pre-set level the voltage to the piezotube and the motor were turned off automatically. This rapidly retracted the tip from the sample’s surface and stopped the differential screw from advancing. A latch kept the system in this state until the user was ready to by-pass the servocontrol and manipulate the tip manually with the piezotube. The current rise on initial contact was monitored with a digital storage scope. This allowed spurious triggering of the system by noise in the current signal to be distinguished from the more uniform current rise when contact really occurred. By simultaneously monitoring the tunnel current and the voltage to the d.c. motor it was found that the tip was retracted 1.1 ms after current appeared, and that the motor was stopped a further 30 ms later. With 1 V applied to the motor the tip advanced towards the sample at 27 nm sec$^{-1}$, suggesting an initial penetration of the tip into the sample of less than 0.03 nm. When the 100 V on the piezotube’s inner electrode was removed the tip retracted a distance of 76 nm. Because this separation was far larger than the amplitude of vibrations appearing in the tip–sample gap no further contact was made until the servo control was by-passed.

* Precision 12 V d.c. motor (RS-336-292) with 2560:1 gearbox (RS-336-286).
To make the next contact the tip was moved transversely by about 40 nm and the $z$-voltage gradually increased by hand whilst monitoring the tunnel current. Because the tip–sample separation after retraction was roughly known ($\approx 80$ nm) the $z$-voltage required for the next contact could be predicted to $\pm 20$ V. In this way, contact was made to an undamaged region of the sample’s surface with very high resolution in the tip movement. This method could be used to limit the initial current rise to less than 0.5 nA and SEM photographs of the tunnelling tip before and after the experiment showed no visible damage on a scale of 200 nm.

6.3.5 Positioning the Sample

The differential spring successfully increased resolution of the coarse–approach stage; however, it also reduced the range of movement to only 200 $\mu$m. To allow an initial room-temperature adjustment of the tip–sample separation the sample was mounted on a ‘table’, which could be moved through several millimetres by turning the M3 threaded ‘legs’ (figure 6.19). The sample was positioned at the correct distance from the tip by comparing the tip–sample gap, seen through a $\times 30$ optical microscope, with a 200 $\mu$m thick copper wire held in the field of view. An initial tip–sample separation of 200 $\mu$m at room temperature was reduced by thermal contraction to 100 $\mu$m at 4.2 K; this is roughly in the middle of the coarse–adjustment range and was fortunately in the region of differential screw movement with the highest resolution.

![Design of the sample table. Scale 3:2. The sample and copper table were separated by an insulating strip of thin plastic or cigarette paper soaked in GE varnish.](image)
To increase positioning accuracy, two of the M3 screw legs were closer to the centre of the sample than the third, resulting in a lever reduction in movement of 5:1 when the latter was turned. To improve the smoothness of movement, contact between the rotating screws and sample table was made via 4 mm diameter steel balls. For the pivotal screws these were silver soldered to the sample table and engaged in conical recesses in the end of the M3 screws. A steel ball was fixed rigidly to the end of the third screw and pushed against the flat surface of a stainless steel insert in the sample table.

Between the M3 screws was a 6BA steel bolt. This had a hexagonal head, which engaged in a T-shaped slot cut along the length of the sample table. When the table was being positioned the nut was loose and the table was moved against the resistance of the small compression spring. Once the sample was in position the lock nut was tightened and opposing pressure between the central bolt and the three screw legs clamped the table rigidly in place. This arrangement allowed the sample to be positioned vertically with an accuracy of ±20 μm and also, by tilting the table, to be moved transversely with respect to the tip by ±2 mm.

The sample could be changed rapidly by unplugging its wires from a socket on the insert and sliding the whole sample table off the clamping bolt. The speed with which a new sample may be mounted will be important if surface contamination or oxide formation is to be minimised. With sensible preparation a sample could be removed from the system in which it was made, attached to the insert and cooled to 77 K within 30 minutes.

### 6.3.6 Preparation of the Tunnelling Tip

Sharp tunnelling tips were made by electrochemical etching of rods of tungsten or tantalum. At low current densities, anodes of these metals are readily attacked by solutions of KOH with the metal passing directly into solution (i.e., \( M \rightarrow M^{z+} + ze^- \)). The surface of the metal rod is etched uniformly along its length at a fairly rapid rate resulting in the complete removal of all the metal that is immersed in the electrolyte. A tip formed in this way typically has a radius of 10 μm and an irregular surface.

At higher current densities ‘passivity’ occurs due to the formation of a continuous oxide film on the metal (e.g., \( \text{WO}_2 \)). This inhibits direct anodic dissolution and the corrosion rate is greatly reduced. Metal ions pass into solution by an indirect process in which new oxide material is formed at the interface between the bulk metal and the

---

* If the table was not protected with this insert the steel ball embedded itself in the soft copper when the table was clamped into position.
passive layer, whilst outer layers of oxide in contact with the electrolyte dissolve. When the current is stable the rates of formation and dissolution of the passive layer are equal and an equilibrium state is reached with an oxide thickness of typically 100 nm (Vetter 1967). If the current is briefly interrupted the passivity is destroyed and the oxide layer is rapidly removed resulting in a shiny surface on the metal (Evans 1923).

Passivation regulates the dissolution of the metal and produces a much smoother surface by polishing away surface irregularities. The rate at which this occurs is sensitive to the chemical and physical environment near the anode. At the meniscus formed where the anode enters the solution only a thin layer of electrolyte is in contact with the metal and growth of the passivating layer is inhibited. A possible explanation for this is that as the oxide layer on the anode’s surface is formed, hydroxide ions are removed from the electrolyte near the metal and a concentration gradient is set up. Hydroxide ions diffuse out of the bulk of the electrolyte towards the anode; however, in the crevice formed by the meniscus diffusion is impeded and the concentration of hydroxide is locally depressed. No further oxide will be able to form to replace that which continues to dissolve. In this region the metal can pass directly into solution so dissolution occurs more rapidly than on the rest of the metal rod, which is still passivated. Etching continues until the notch cut into the metal at the meniscus extends right through the rod and the immersed part falls off. This is an example of the widespread phenomenon of crevice corrosion (West 1986).

If the anode voltage is increased to about 20 V oxygen gas is produced by the electrode reaction, \(4\text{OH}^- \rightarrow \text{O}_2 + 2\text{H}_2\text{O} + 4\text{e}^-\). This does not occur immediately because the passivation layer forms a continuous protective coating on the metal and hydroxide ions are unable to reach the anode. Tapping the electrolytic cell (or leaving the voltage turned up for a few seconds) results in a sudden onset of this reaction as hydroxide ions breach the oxide layer. The violent production of gas bubbles disturbs the remaining oxide leading to very rapid corrosion along the entire rod and resulting in an irregular and rather blunt tip.

To prepare the tunnelling tip a 20 mm × 1 mm diameter rod of tungsten or tantalum was inserted into a hole drilled along the axis of an 8BA brass screw and permanently fixed in place by crimping the screw head. The rod was suspended from a weak spring and placed in an electrolytic cell in the centre of a circular stainless steel cathode. The cylindrical geometry ensured that the metal rod etched uniformly. Freshly prepared 2 M KOH solution was pumped into the cell from a 500 ml syringe until about 5 mm of rod remained above the liquid surface. The anode voltage was briefly turned up until oxygen gas was evolved, as described above, and was then reduced to allow a
thin passive layer to form. This rapid etch removed the thick oxide layer, which had formed on the metal as a result of long-term atmospheric exposure, to leave a shiny metal surface that was more suited to etching.

The voltage ranges in which the three types of etching occur are clearly seen in the \( I-V \) characteristics of the electrolytic cell (figure 6.20). Measuring the current through the cell allowed the plateau region to be identified so the tip could be etched under optimum conditions with a passivating layer covering the metal rod.

After approximately 15 minutes the rod etched through at the meniscus and, as the lower 15 mm of the rod fell away, the spring pulled the tip away from the electrolytic solution. With a minimum of delay the electrolyte was drawn out of the cell by the syringe and the tip was flushed with distilled water to remove traces of KOH. This procedure prevented further etching which would just round off the sharp point. The tip was then removed from the cell, examined with a scanning electron microscope and, if sufficiently sharp, attached to the piezotube.

In apparent agreement with Huang et al. (1990b), who claimed that the ‘best’ tunnelling characteristics were obtained with very sharp points, it was noted that the clearest non-linearities and gap-like conductance peaks were obtained with sharpest tips. Consequently all tips with a radius of curvature greater than 500 nm were rejected. However, this strategy for tip fabrication and selection may have been an unfortunate choice. Wilkins et al. (1990a) found that the energy gap of lead, clearly observed by tunnelling from either a cleaved platinum or tungsten tip, was never

![Fig. 6.20. Current–voltage characteristics of a KOH electrolytic cell with tungsten anode. A passivating oxide layer forms at voltages above the Flade potential, \( V_F \), and is indicated by a region of constant current.](image-url)
obtained when an electrochemically etched tip was used; the conductance–voltage curves were more characteristic of single-electron tunnelling (SET) in a mesoscopic junction (section 4.4). Berthe and Halbritter (1991) showed that this may be due to resonant tunnelling via localised states in a 1–3 nm thick adsorbate layer* on the wet-etched tungsten tip. These states have very small effective capacitances with large Coulomb charging energies. Consequently, spurious SET effects arising from the tip can dominate the tunnelling characteristics even at room temperature. It is possible that the process used to prepare tips for the work described in this thesis was incorrectly optimised to produce SET effects rather than SIN tunnelling. Clearly, future work should use mechanically sharpened tips.

6.3.7 Operation of the Point-Contact Insert

The sample was glued to the insulated copper table with GE varnish and the current and voltage wires, $I$– and $V$–, were connected to the sample with either pressed indium or silver dag. The table was then attached to the insert and positioned at the correct distance from the tip, as described above. The copper vacuum can, sealed with a polythene gasket, was evacuated and filled with helium gas for good thermal contact between the sample and liquid-helium bath. The insert was placed in the cryostat which was then filled with liquid nitrogen and left to precool for 4 hours.

After precooling, any remaining liquid nitrogen was forced out of the cryostat through a siphon tube by an overpressure of dry helium gas. Once the bath space had been purged of residual nitrogen (by pumping with a rotary pump) liquid helium was transferred into the cryostat from a storage dewar by the standard technique. Care was taken to ensure that liquid helium was delivered below the large thermal mass of the vacuum can by providing a dog-legged extension to the standard transfer tube. This maximised heat exchange between the cold outflowing gas and the warmer can, reduced the quantity of helium lost to the return line during the transfer and prevented excessive ice formation on top of the cryostat. This last point was important because of the danger of shorting in the high-voltage feedthrough and the possibility of electrocution. As an additional precaution the feedthrough’s plug and socket were dried thoroughly with a heat gun before connection and then covered with a tightly fitting rubber jacket. During helium transfer the volume of liquid in the cryostat was monitored continually using a simple comparator circuit. A column of LEDs represented the level of helium in the vessel, each LED being lit when a corresponding Allen–Bradley resistor, attached to the insert, was cooler than 10 K.

* Probably OH, H$_2$O or hydrocarbons bonded to the WO$_3$. 
A super-insulated cryostat with no nitrogen jacket was used to avoid disturbance of the tip–sample contact by boiling liquid nitrogen. To maximise the length of time available for each low-temperature run the cryostat’s vacuum jacket was re-evacuated between experiments. With a jacket pressure of $10^{-5}$ torr* the bottom of the insert would remain covered by liquid helium for about 9 hours. Attempts were made to eliminate vibrations caused by boiling helium by cooling the liquid below its $\lambda$-point. Due to the lack of vibration isolation of the pumping line this was found to be ineffective and merely increased the rate of loss of liquid helium.

After completing the helium transfer the servomechanism was used to move the tip towards the sample. This took approximately 1 hour, during which time the liquid helium stopped boiling vigorously and vibration of the cryostat was reduced to an acceptable level. Once the coarse-approach stage had been completed the metal tip could be moved into contact with the sample by manually increasing the voltage to the piezotube’s central electrode. The current flowing between the tip and sample was monitored continuously using the oscilloscope-like display of the data-acquisition program. The tip–sample contact could be adjusted and, when desired, the $I$–$V$ characteristics could be recorded and saved for later analysis. By adjusting the voltage on the piezotube’s outer electrodes the metal tip could be moved to a new region on the sample’s surface and more data could be taken. This process was repeated until either the tip became damaged, in which case only low-resistance linear $I$–$V$ characteristics could be obtained, or the liquid-helium level dropped below the vacuum can and the sample warmed up.

### 6.3.8 Suggested Improvements

The biggest problem with the point-contact insert is the very low resonant frequency of only 816 Hz. Disturbance of the tunnelling tip by external vibrations seriously limited the time available to collect data from a single contact and the data acquisition rate become an important factor. Failure to obtain a higher resonant frequency stems from a decision to modify a pre-existing insert to produce the coarse tip movement, resulting in a rather complicated hybrid of differential screws and springs. In an effort to make sample mounting easier and to allow the tip and sample to be inspected with an optical microscope, an ‘open’ framework was built to which the differential springs etc were attached. This consisted of three brass platforms connected by struts (figure 6.14). Clearly this rather slender structure contains many possible sources of weakness and should not have been used.

---

* Using a pump with base pressure of $10^{-7}$ torr and taking into account the pressure drop along the pumping line.
Many of the problems and limitations of the point-contact insert described in this work have recently been overcome in the Low Temperature Physics Group by development of a cryogenic scanning tunnelling microscope. This takes advantage of recent advances in STM design by using a piezoelectric inertial slider for coarse movement (Renner et al. 1990) and piezotubes for fine movement. The lowest resonant frequency of the STM is about 5 kHz and topographic images of graphite atoms have been observed at 4.2 K using commercial electronics and software. The electronics also allows rapid acquisition of a junction’s $I-V$ characteristics, so correlation of electronic and topographical features should be possible. In particular, spatial dependence of energy gap over the surface of a sample and at different depths could be measured (see, e.g., Volodin and Khaikin 1987, Al’tfeder et al. 1989). The STM is mounted in a continuous-flow cryostat so studies of temperature dependence of the gap should also be possible.
7.1 Conventional Superconductor Point-Contact Junctions

The point-contact insert and electronics were tested by measuring the tunnelling characteristics of junctions containing conventional superconductors. The aim was to compare the results obtained with those predicted by theory or obtained by other techniques in order to verify correct operation*, to demonstrate the potential of the apparatus and to improve experimental technique without complications introduced by use of poorly understood samples.

7.1.1 Pb SIN Junctions

Pb films were prepared by the evaporation technique described in section 6.2. Initial contact between the tungsten tip and the sample was always of high resistance (> 10 kΩ at 10 mV) and displayed SIN-type conductance–voltage characteristics (figure 7.1).

To compare the experimental tunnelling characteristics with those predicted by BCS theory the normalised conductance,

\[ \sigma(V) = \frac{G_s(V)}{G_n(V)} = \frac{(dI/dV)_s}{(dI/dV)_n} \]

must be calculated. Conventionally \( G_n \) is obtained by applying a magnetic field that exceeds the critical field \( H_c \), driving the Pb normal and allowing \( G_s(V) \) and \( G_n(V) \) to be measured at the same temperature. High-\( T_c \) superconductors, however, are extreme type-II superconductors with upper critical fields larger than 100 T (Burns 1992, p.154), so no attempt was made to fit the point-contact insert with a coil or permanent magnet†. If the normal-state conductance does not vary too much with temperature \( G_n \) may be found by raising the sample above its transition temperature. Unfortunately

* The conductance of a ‘standard’ non-linear device was also regularly measured to test calibration of the electronics (Edgar and Zyskowski 1985).

† A magnetic field would still be useful when using conventional superconducting counterelectrodes. This allows one to switch from SIS to SIN tunnelling to estimate leakage currents (Valles et al. 1991).
thermal expansion of the insert usually resulted in sporadic changes in contact so this technique could not be used. For both high-$T_c$ and conventional superconductors the normal-state conductance was estimated by fitting a suitable function to the tunnelling characteristics at high biases ($|eV| > 3\Delta$) and interpolating this ‘background’ into the gap region.

The measured conductance of an evaporated Al/$\text{Al}_2\text{O}_3$/Pb junction (dashed curve in figure 7.1) was normalised by a fitted low-order polynomial - resulting in good agreement with the BCS prediction in the gap region (i.e., for $eV < \Delta$). However, to obtain similar agreement between data from a point-contact junction onto Pb and the BCS result a large constant *leakage conductance* of 1.24 S (60% of the raw zero-bias conductance) had to be subtracted from the point contact’s conductance before normalising. Leakage may be caused by large-area microshorts (contact radius $\gtrsim 0.1\text{\,\mu m}$) across the junction or resonant tunnelling via localised states caused by surface contamination (Thompson and von Molnar 1970). Either explanation is possible because junction formation relied on the intrinsic insulating surface layer on the lead.

The clear ‘satellite’ peaks at $\pm 4.48\text{\,mV}$ in the point contact’s conductance–voltage characteristics may also be due to resonant tunnelling (section 3.6). However, the symmetrical peak positions would require that the localised state is located, coincidentally, in the centre of the barrier. In addition, the peaks are much larger
(approx. 0.1 \( \Omega^{-1} \)) than the maximum \( \frac{dI}{dV} \) of approximately \( e^2/h \) predicted by models of resonant tunnelling (Speakman 1992, sec.4.5.5.3). An alternative explanation is suggested by the presence of additional smaller peaks at \( \pm 7.24 \text{ mV} \) and \(-10.6 \text{ mV}\); a series of conductance peaks spaced at intervals of approximately 3.0 \text{ mV} \) appears to be superimposed upon the usual SIN conductance–voltage characteristics. Periodic multiple conductance peaks are characteristic of two mesoscopic tunnel junctions in series (section 4.4). The double-junction configuration could result from an isolated conducting particle positioned between the point-contact tip and the bulk Pb electrode. When voltage biased this double-junction system will have current steps at

\[
V_{s,k} = \begin{cases} 
\pm \left( k - \frac{1}{2} \right) \frac{e}{C_i} & \text{for NNN junctions} \\
\pm \left[ \left( k - \frac{1}{2} \right) \frac{e}{C_i} + \frac{\Delta(T)}{e} \left( \frac{C_1 + C_2}{C_i} \right) \right] & \text{for NSN, SNS, NSS and SSN} \\
\pm \left[ \left( k - \frac{1}{2} \right) \frac{e}{C_i} + \frac{2\Delta(T)}{e} \left( \frac{C_1 + C_2}{C_i} \right) \right] & \text{for SSS junctions}
\end{cases}
\] 

(7.1)

where \( k \in \mathbb{Z}^+:1,2,3,... \). \( C_1 \) and \( C_2 \) are the junction capacitances and \( \Delta \) is the energy gap of the superconducting electrode. \( C_i \) represents both \( C_1 \) and \( C_2 \) since two series of current steps are superimposed; however, if the junction time constants are very different the slower junction’s steps will dominate.

If the main SIN junction producing the gap-like feature in figure 7.1 were one of the two series junctions then the central peak separation would be greater than in an isolated SIN junction by at least the spacing of subsequent peaks. Comparison with the BCS and evaporated junction curves shows that the peak separation has not increased so the main SIN junction is not in series.

\[
\text{Fig. 7.2. Parallel SIN and SSN tunnelling.}
\]
The series junctions may, however, be in parallel with the SIN junction. This could occur if the tip were in contact with several regions of metallic and superconducting material - which would also explain the large leakage currents. Fitting the observed peak positions to (7.1) results in two possible interpretations.

i. If the conductance peaks at ± 4.5 mV are due to the $k = 1$ current steps then the junction is NSS, SNS or NSN with $\Delta(T)$ being that of Pb at 4.2 K and with capacitances of 53 aF and 80 aF. Alternatively it could be an SSS junction with capacitances 53 aF and 14 aF.

ii. If the ± 4.5 mV peaks are for $k = 2$ then the series junction is NNN with $C_i = 53$ aF. The other capacitance cannot be determined from the step position. In this configuration the $k = 1$ steps occur at ± 1.51 mV so are obscured by the SIN junction’s conductance peaks.

The conductance–voltage characteristics of a single, isolated SIN junction are clearly observed indicating that the tip is normal and the sample is superconducting and so suggesting that the SSS, SNS, NSN and NNN configurations are unlikely. The most probable configuration is sketched in figure 7.2.

The normalised conductance of the evaporated junction and the point contact depart from the weak-coupled BCS prediction for $|eV| > \Delta$. These deviations, emphasised by plotting the BCS-reduced conductance, $\sigma(V)/\sigma_{BCS}(V)$, (figure 7.3) are due to strong electron–phonon coupling. The effective phonon spectrum, $a^2F(\omega)$, may be obtained from the BCS-reduced conductance curve and the gap edge value, $\Delta_0$, by use of the

![Fig. 7.3.](image)

Fig. 7.3. The BCS-reduced conductance of data from figure 7.1. Also shown (MR) is the result of McMillan and Rowell (1969) taken at 0.3 K.
Fig. 7.4. Derivative of the BCS-reduced conductance from figure 7.3. Peaks at 4.5 meV and 8.6 meV indicate the transverse and longitudinal phonon energies (Wolf 1985, p.166).

Fig. 7.5. Eliashberg function for Pb calculated from the reduced conductance data of McMillan and Rowell shown in figure 7.3 (Wolf 1985, p.169).
numerical algorithms developed by McMillan and Rowell (Wolf 1985, appendix B). The location of peaks in $\hat{F}(\omega)$ (figure 7.5) may be estimated from the energy of the most negative slopes in the reduced conductance (figure 7.4).

7.1.2 Pb ScN Junctions

High-resistance SIN junctions formed by initial contact between the W tip and Pb sample could be controllably reduced in resistance by extending the piezotube. Slightly retracting the tip at any stage resulted in immediate loss of conductance, suggesting that either the tip or a thick insulating surface layer on the Pb was being deformed plastically. SEM images of the tungsten tip taken before and after these experiments showed no evidence of deformation on a scale of 10 nm.

The conductance of a low-resistance W–Pb contact is shown in figure 7.6. An increased conductance for $|eV| < 2\Delta$ with superimposed dip for $|eV| < \Delta$ is characteristic of Andreev reflection at an NS interface with some normal electron reflection (i.e., $0 < Z \leq 0.5$, see section 3.4). Also shown in figure 7.6 is the BTK model conductance obtained by adjusting $Z$ to fit the energy of the conductance maxima. In common with other published results (e.g., Reinertson et al. 1990) the conductance peaks (normalised by an estimated background) are smaller than predicted by the model. The conductance minima at $|eV| > \Delta$ are also anomalous, and probably associated with quenching of superconductivity by the onset of quasiparticle current in the Pb (McLean 1984). Conductance dips such as these, or sometimes

Fig. 7.6. Conductance–voltage characteristics of a W–Pb point contact at 4.2 K (solid curve). Conductance maxima and minima are at $\pm 1.0 \pm 0.1$ mV and $\pm 3.9 \pm 0.1$ mV respectively. Also shown (dashed curve) is the thermally broadened BTK model conductance for $Z = 0.45$ and $\Delta = 1.23$ meV, which has peaks at $\pm 1.01$ mV. The energy gap of Pb at 4.2 K measured by tunnelling is 1.23 meV and is shown by the vertical lines.
complex multiple-dip structure at the same location, are frequently observed in very low-resistance contacts. Once the contact region has been driven normal by the quasiparticle current, increasing voltage causes the normal region to grow and results in a gradually decreasing conductance background, as observed here.

7.1.3 Nb ScN Junctions

Point contacts formed between a tungsten tip and Nb foil usually resulted in ohmic conductance. SIN-like conductance–voltage characteristics were never observed and on only one occasion was an ScN-like characteristic obtained (figure 7.7). A single, sharp peak at zero bias with flanking conductance minima at $eV \approx \pm \Delta$ is not predicted by BTK theory (see figure 3.14). In addition, the peak in figure 7.7 is larger than the maximum normalised conductance allowed by BTK theory (i.e., two). Similar structure has been often been observed with low-resistance ScN contacts, for example, with Sn–Cu (McLean 1984, fig.54), Au–Pb and Au–YBa$_2$Cu$_3$O$_y$ (Srikanth 1992, figs 4.6 and 5.4). There appears to be no clear explanation in the literature for these anomalous features; they are usually attributed to heating or proximity effects, though little detail is given. Similar structure was seen recently in the conductance–voltage characteristics of quantum wells with superconducting Nb electrodes (Nguyen et al. 1992). In this case, a detailed explanation was proposed that involved multiple normal and Andreev reflections at interfaces between various layers in their device.

![Fig. 7.7](image-url) Conductance–voltage characteristics of a W–Nb point contact at 4.2 K. The conductance dips are located at $-1.78 \pm 0.32$ mV and $+1.62 \pm 0.12$ mV. The energy gap of Nb at 4.2 K (1.51 meV measured by tunnelling) is shown by the vertical lines.
7.2 High-\(T_c\) Superconductor Planar Junctions

7.2.1 \(\text{YBa}_2\text{Cu}_3\text{O}_\text{y}\) Ceramic

Figure 7.8 shows typical conductance–voltage characteristics of a \(\text{Pb–YBa}_2\text{Cu}_3\text{O}_\text{y}\) planar tunnel junction, prepared as described in section 6.2. The temperature dependence of the zero-bias conductance minimum indicates that this is an \(\text{SIN}\) junction with superconducting \(\text{Pb}\) and a normal counterelectrode and hence suggests that the \(\text{YBa}_2\text{Cu}_3\text{O}_\text{y}\) surface is covered with a non-superconducting layer. No features were ever observed that could be confidently associated with an energy gap in the high-temperature superconductor. Using ion-beam milling, glow-discharge cleaning and sputtering with \(\text{O}\) and \(\text{Ar}\) ions, attempts were made to clean the sample’s surface before deposition of the \(\text{Pb}\). However, these all resulted in a rust-brown discoloration and the surface remained normal. Annealing the junction in oxygen after deposition of the second electrode was also unsuccessful.

Narrow conductance dips were often present at high bias. These generally decreased in energy with increasing temperature but their exact positions were irreproducible and hysteretic with bias and temperature. Similar structure has been attributed to abrupt redistribution of current, as weak links in a percolation path network within the sintered sample are driven normal (Goodrich \textit{et al}. 1991). Alternatively, a conductance dip occurs when a superconducting short through the tunnelling barrier (with a series resistance) is driven normal at its critical current (Speakman 1992, sec.4.5.4.2). The observed temperature dependence of the position of the dips is qualitatively explained by decreasing critical current with increasing temperature.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{conductance_v_bias.png}
\caption{\(\text{Pb–YBa}_2\text{Cu}_3\text{O}_\text{y}\) planar tunnel junction at 4.2 K. Identical results were obtained by Gijs \textit{et al}. (1988). Broad conductance dips, like that seen here between \(\pm 30\) mV, have been tenuously attributed by some authors to the \(\text{YBa}_2\text{Cu}_3\text{O}_\text{y}\) energy gap.}
\end{figure}
Similarly, magnetic field dependence of the sharp dip structure seen by Hasegawa et al. (1992, fig.24) can be attributed to decreasing critical current with increasing field. In view of the uncertain nature of the tunnelling barrier used in this type of junction the presence of pin holes shorted by Pb is unsurprising.

### 7.3 High-\(T_c\) Superconductor Point-Contact Junctions

#### 7.3.1 \(\text{YBa}_2\text{Cu}_3\text{O}_y\) Ceramic

This section presents the results of point-contact tunnelling measurements on samples of \(\text{YBa}_2\text{Cu}_3\text{O}_y\) prepared in ceramic form by the solid-state reaction method (Chen et al. 1987). These measurements were made at an early stage in this investigation using the crude push-rod apparatus described in section 6.3*. This lacked the stability and fine control of the cryostat insert but did allow the junction temperature to be changed by raising the dip stick out of the liquid helium into the neck of the dewar.

Even at 4.2 K most contacts showed no evidence of a gap-like feature and the conductance–voltage characteristics consisted merely of an asymmetric V shape or U shape. Occasionally weak gap-like structure was observed, superimposed upon the background curve (figure 7.9). The amplitude of the gap feature decreased with increasing temperature, disappearing entirely at a temperature typically much lower than the bulk \(T_c\) of approximately 90 K. The conductance remains V shaped at temperatures well above \(T_c\) indicating that this behaviour arises from a normal-state property of the sample.

The junction temperature could not usually be changed without disturbing the junction, which prevented measurement of both the superconducting and normal-state conductances of a single contact. For this reason it was usually necessary to estimate the normal-state conductance from the high-bias region of the superconducting characteristics. A method for doing this was developed using data from a junction whose conductance was reproducible on thermal cycling, indicating that mechanical disturbance had been minimised (figure 7.10).

Both above and below \(T_c\) the conductance at high biases could be fairly well modelled by

\[
\frac{dI(V)}{dV} = a + bV + cV^2 + d|V|
\]

(7.2)

* These measurements and the apparatus used were based on earlier work by Edgar et al. (1987).
Fig. 7.9. Temperature dependence of the tunnelling characteristics of an Al–YBa$_2$Cu$_3$O$_y$ point contact. Numbers indicate the junction temperature in kelvins. Curves are offset for clarity.
Fig. 7.10. Measured conductance–voltage characteristics of a single Al–YBa$_2$Cu$_3$O$_y$ point contact at 4.2 K (circles) and 100 K (crosses). The continuous curve is an estimate of the normal-state ‘background’ conductance found from the 4.2 K characteristics by the process outlined below. The 100 K background conductance (dashed curve) was obtained by thermally smearing the 4.2 K estimate.

where $a$, $b$, $c$ and $d$ are adjustable parameters. The normal-state conductance was estimated from the measured low-temperature conductance by least-squares fitting (7.2) to the high-bias data, i.e., ignoring data with $|V|$ smaller than some $V_{\text{cutoff}}$. To test this method the estimated normal-state conductance was convolved with the thermal broadening function for 100 K and compared with the measured conductance at that temperature. Initially $V_{\text{cutoff}}$ was chosen to be large ($\approx 50$ meV) so that fitting was unaffected by conductance in the gap region. However, agreement between the estimated and true normal conductances improved as $V_{\text{cutoff}}$ was reduced to zero so ultimately all data, even that in the gap region, was used in the fit. Figure 7.10 shows the normal-state ‘background’ conductance estimated by this process at both 4.2 and 100 K. Figure 7.11 shows the normalised conductance obtained by dividing the measured conductance at 4.2 K by the estimated normal-state conductance.

This procedure to obtain the normalised conductance is somewhat arbitrary. Moreover, the practice of normalising by some estimated background of unknown origin in order to ‘extract’ an excitation density of states is itself rather dubious. Normalisation for this purpose is suggested by the ideal model of a BCS SIN tunnel junction (2.14 and 3.15–16); however, the technique is not valid in (even only slightly) more complicated systems. This is demonstrated in figure 7.12.

In the presence of a large background of uncertain origin the normalised conductance cannot be confidently used as an estimate of the density of excitation states and should
**Fig. 7.11.** Estimated normalised conductance of the 4.2 K data from figure 7.10 (solid curve). The dashed curve is ideal BCS model at 4.2 K for $\Delta = 13.7$ meV.

**Fig. 7.12.** This figure shows that normalisation of an SIN junction’s conductance–voltage characteristic by its normal-state conductance–voltage characteristic will not, in general, yield the superconductor’s density of states. The curves are calculated using the WKB BCS model outlined in section 4.2.2 with junction area $= 10^{-13}$ m$^2$, barrier width $= 1$ nm, $q = \pm |e|$, $E_F = 3$ eV, $\phi_1 = 100$ meV and $\phi_2 = 120$ meV. The gap-like feature is superimposed on a voltage-dependent background; however, normalisation would clearly not yield the ideal BCS density of states used to calculate these curves.
not strictly be used for quantitative analysis. It is useful, however, as a visualisation tool to enhance qualitative details of the gap-like feature.

The normalised conductance shown in figure 7.11 is clearly broadened in comparison with the thermally smeared BCS model conductance and has a large leakage conductance at zero bias. This complicates determination of the sample’s energy gap. One method for estimating \( \Delta \) and for quantifying the level of broadening is to fit to the data the tunnelling conductance from a metal electrode into a BCS superconducting electrode with a Gaussian distribution of energy gaps (4.13–14). The effects of leakage conductance may be minimised by using the *scaled conductance*

\[
g(V) = \frac{G_S(V) - G_S(0)}{G_N(V) - G_S(0)}
\]

where \( G_S(V) \) is the differential conductance at bias \( V \) in the superconducting state and \( G_N(V) \) is the actual or estimated normal-state differential conductance. Figure 7.13 shows the scaled conductance of a different point-contact junction formed between the same tip and sample used to obtain the data shown in figures 7.10–11. Also shown is the best-fit Gaussian-model curve.

There is some indication of fine structure at 40–50 mV so data in this region were not used in fitting the model curve. The residual \( \chi^2 \) value is too large to be attributed to random errors in the data, indicating that the model or normalisation scheme used is

**Fig. 7.13.** Scaled conductance (solid curve) of an Al–YBa\(_2\)Cu\(_3\)O\(_y\) point contact and best-fit Gaussian model (dashed curve). With the parameters shown, and an estimated error of 3\% in the 54 points used in the fit (circles), \( \chi^2 = 218 \). With large \( \delta \Delta \), as here, thermal broadening of the Gaussian model may be ignored with negligible error.
incorrect; nevertheless, the best-fit parameters are a useful indication of the energy gap in this sample.

The scaled conductance in figure 7.13 was calculated from the raw conductance–voltage characteristic shown in figure 7.14a. Also shown in figure 7.14 are four other conductance–voltage characteristics taken immediately afterwards (over period of 3 hours) of nominally the same junction. There has clearly been a large uncontrolled change in the junction, probably due to vibration-induced movement of the tip. Sharp conductance peaks occur at approximately twice the energy of those in curve a, suggesting that an SIS tunnel junction has been formed with the aluminium tip in good contact with one of the superconducting electrodes. The tunnel barrier is probably formed at an internal grain boundary. From the peak separation of curve b the energy gap may be estimated to be $13.3 \pm 0.7$ meV. It is tempting to use the bulk transition temperature of $90.7 \pm 0.5$ K (four-terminal d.c. zero resistance) to conclude that the sample has a reduced energy gap of $3.40 \pm 0.18$; in good agreement with the BCS value.

**Fig. 7.14.** Sequence of conductance–voltage characteristics of an Al–YBa$_2$Cu$_3$O$_y$ point-contact at 4.2 K., displaced vertically for clarity.
of 3.52. However, it is clear from figure 7.9 that the local transition temperature may be much lower than the bulk value so the reduced gap could be larger.

Smaller differences between curves $b - e$ could be due to further tip movement resulting in a change in the sampled inhomogeneous gap distribution (particularly if $\Delta$ varies within the unit cell) or gap anisotropy. Erratic movement of structure has been observed in similar samples (see, e.g., Wan et al. 1990b), usually accompanied by evidence of charging effects.

Curves $b$ and $c$, particularly their V shape at low bias, are remarkably similar to the tunnelling conductance predicted by the layered model of Takahashi and Tachiki (figure 7.15). Simple matching of peak positions suggests that $\Delta = 13.04 \pm 0.3$.

Attempts were made to form ScN junctions by increasing the tip–sample contact force from the fairly low values used in the experiments discussed above. This generally resulted in almost linear conductance–voltage characteristics with small, complex structure containing narrow conductance dips (figure 7.16). These were found to decrease in voltage with increasing force on the contact. Similar structure and behaviour were seen by Srikanth (1992, p.131), who points out that sensitivity to tip pressure suggests that the conductance dips originate close to the contact region. In addition, this structure was largely removed when the sample’s surface was cleaned (by cleaving). The narrow conductance dips resemble those seen in high-$T_c$ superconductor planar tunnel junctions (section 7.2.1) and can be attributed to a

---

Fig. 7.15. Comparison of the SIS model conductance (points) of Takahashi and Tachiki (1990a,b, 1991) with measured conductance (curve $b$ of figure 7.14) normalised by an estimated background. The bias has been scaled by $2\Delta = 26.08$ meV to fit peak positions to those of the model.
Fig. 7.16. High-contact-force Al–YBa$_2$Cu$_3$O$_y$ point contact at 4.2 K.

Fig. 7.17. High-contact-force Pb–YBa$_2$Cu$_3$O$_y$ point contact at 4.2 K. The conductance minima separation is 1.3 meV. The dashed curve is the current, the solid curve is the conductance.
similar origin - an internal superconducting weak link (possibly between grains of ceramic) in series with the resistive tip–sample contact. By increasing the contact force the series resistance, $R$, goes down and the dip structure, which occurs at $I_c R$, where $I_c$ is the weak link’s critical current, moves to lower voltages.

The tunnelling characteristics frequently contained a narrow zero-bias conductance maximum. It has been suggested that this feature is caused by Andreev reflection, although the curve does not resemble those obtained with conventional superconductors or predicted by theory. This feature is more probably evidence of supercurrent at an internal grain-boundary Josephson junction (or an array of such junctions) in series with the point contact (Walsh 1992).

The high-contact-force conductance–voltage characteristics shown in figure 7.17 resemble those of a low-barrier-strength $\text{SN}$ microshort (figure 7.7). This suggests that current is flowing from superconducting Pb into normal material on the surface of bulk superconducting $\text{YBa}_2\text{Cu}_3\text{O}_y$.

### 7.3.2 Bi-Sr-Ca-Cu-O Single Crystals

Samples were prepared in the following way. Powdered $\text{Bi}_2\text{O}_3$, CuO, SrCO$_3$ and CaCO$_3$ were mixed with molar ratio Bi:Sr:Ca:Cu=1.17:1.00:1.00:2.37 and heaped at the upper end of an alumina boat tipped at 5° to the horizontal. This was heated to 920 °C at 50 °C h$^{-1}$, held at this temperature for 2 h, cooled to 820 °C at 4 °C h$^{-1}$ and finally cooled to room temperature at 50 °C h$^{-1}$. Throughout this process oxygen was flowing through the furnace and no additional oxygen anneal was used. If the boat was horizontal during heating then a dense solid, clearly containing several phases, formed in the bottom. Embedded in this ‘clinker’ were regions with a crystal-like structure; however, these could not successfully be separated from the matrix and were always too small to be used in tunnelling experiments. If the boat was tipped before heating then the melt flowed and a pool of shiny black material containing numerous cavities formed at the lower end. Single-crystal* platelets of typically 1 mm $\times$ 1 mm $\times$ 0.02–0.1 mm were attached loosely to the cavity walls and could be removed with a scalpel. When cooled to liquid-nitrogen temperature these were repelled strongly by a permanent magnet and aligned their plane parallel to the field. This indication of high-temperature superconductivity was used to identify the most suitable samples for further experiments. Several unidentified black and green phases remained in the upper half of the boat and could not be removed.

* Confirmed by X-ray diffraction (Pigram 1988).
Fig. 7.18. *ab*-plane resistance of Bi-Sr-Ca-Cu-O single crystal. $T_c$ (zero resistance) is at 78.0 ± 0.5 K with a 10%–90% transition width of 8 K. The resistance is approximately linear above 90 K.

The transition temperature of each crystal was determined from four-terminal a.c. resistivity measurements (figure 7.18). Low-accuracy* electron probe microanalysis (EPMA) indicated a Bi:Sr:Ca:Cu ratio of 2.2:1.3:1.0:1.7, with variation of 5% over the surface. Comparison of this information with the properties of known high-temperature superconductors (Burns 1992, p.57) suggests that the crystal was Bi$_2$Sr$_2$CaCu$_2$O$_8$.

To assist with their handling, the small crystal platelets were glued flat on a glass cover slip using GE varnish. The exposed face of the platelet was parallel to the (001) plane (Hazen *et al.* 1988) hence the nominal tunnelling direction was parallel to the $c$-axis. Attempts were made to bond 50 μm gold wires to the sample by spark discharge of a capacitor (Iye *et al.* 1988). This was unsuccessful, however, and large-area silver dag contacts were eventually used. This limited the area available on the sample for tunnelling to less than 0.5 mm$^2$ so the simple push-rod apparatus could not be used. The sample table and differential screw of the point-contact insert were built to allow the tip to be placed within this area; however, the piezotube was not yet completed so fine control and lateral movement of the tip was unavailable; this limited the quantity of tunnelling data that could be obtained from these samples and only one gap-like conductance–voltage characteristic was ever obtained (figure 7.19).

* Approximately 20% error in ratios using energy dispersive spectrometer and ‘ZAF4’ data analysis package with software standards.
Fig. 7.19. Conductance–voltage characteristics of W–Bi$_2$Sr$_2$CaCu$_2$O$_8$ single-crystal point contact (solid curve) and BDR–Kirtley model with Gaussian distribution of gap values with mean 37.7 meV and width 24 meV (dashed curve). Estimating junction area \( \approx 10^{-9} \) m$^2$, \( E_F = 1.0 \) eV and \( q = +|\epsilon| \) gives fitted parameters \( \phi_1 = 2.7 \) eV, \( \phi_2 = 3.7 \) eV and width = 10.9 Å. Conductance maxima occur at ± 49 ± 2 mV and minima occur at ± 86 ± 2 mV.

This conductance–voltage characteristic exhibits a much clearer gap than was ever observed with a YBa$_2$Cu$_3$O$_y$ sample. The large symmetrical conductance peaks, low sub-gap conductance and parabolic background are far closer to conventional results; however, the characteristics are still strongly smeared. These results are typical of tunnelling measurements on Bi$_2$Sr$_2$CaCu$_2$O$_8$ (Hasegawa et al. 1992a, p.27). Combined with the broad resistive transition and the non-stoichiometric EPMA results, the broadened characteristics suggest that Bi$_2$Sr$_2$CaCu$_2$O$_8$ ‘single crystals’ have an inhomogeneous composition - possibly due to defects and stacking faults. The extrinsic nature of this broadening is also suggested by the near-ideal results observed in an SIN planar junction fabricated on the cleaved surface of a Bi$_2$Sr$_2$CaCu$_2$O$_8$ single crystal (Matsumoto et al. 1992).

In contrast to all YBa$_2$Cu$_3$O$_y$ results, the background conductance of the W–Bi$_2$Sr$_2$CaCu$_2$O$_8$ point contact is similar to those seen in conventional tunnel junctions, where the parabolic dependence on voltage is caused by changes in tunnelling probability as the barrier is distorted by the applied bias. Fitting the BDR model to the background conductance gave the barrier parameters indicated in figure 7.19. Using these parameters, and combining Kirtley’s extension to the BDR model (see section 4.2.2) with a Gaussian distribution of energy gaps, results in the dashed curve shown in the figure.
Forro *et al.* (1992) obtained characteristics very similar to figure 7.19 using a single-crystal break junction with SIS tunnelling in the *ab*-plane*. In contrast, no well defined structure was observed for tunnelling in the *c*-direction (section 4.3.2). Several other reports also claim that gap-like features can only be observed when tunnelling in the *ab*-plane (e.g., Wnuk *et al.* 1991).

Figure 7.19 contains very strong conductance dips or ‘undershoots’ just outside the gap region. These have been seen in many tunnelling experiments on Bi$_2$Sr$_2$CaCu$_2$O$_8$, particularly in those with tunnelling in the *ab*-plane. For example, although Tulina *et al.* (1990) observed a gap with SIN tunnelling in both the *c*-direction and *ab*-plane on a cleaved single crystal, undershoots were present only in the latter case. Conductance dips are far less prominent in SIN tunnelling characteristics than SIS characteristics and may be too weak to be seen as such against a voltage-dependent background and with thermal broadening. Wnuk *et al.* (1991) observed undershoots in the conductance–voltage characteristics of an *ab*-plane SIS junction, formed when the edges of two (Pb$_x$Bi$_{1-x}$)$_2$Sr$_2$CaCu$_2$O$_8$ crystals were brought into contact. An *ab*-plane SIN point contact on the same sample, however, did not seem to show these features and had a more prominent V-shaped background conductance. However, Zasadzinski *et al.* (1992) showed that if these SIN data are used to calculate the associated SIS curve, strong dips are generated, which agree with the experimental SIS characteristics in both position and amplitude.

These observations suggest that the characteristics shown in figure 7.19 may be due to SIS tunnelling in the *ab*-plane of the crystal, despite the nominally *c*-direction SIN configuration. This hypothesis is supported by the highly symmetrical nature of the conductance–voltage curve in comparison with most SIN characteristics, and the reduced gap value† appropriate for this interpretation of the data of 5–7.5, which agrees with most published values for this material. The equivalent SIN values ($2\Delta/k_BT_c \approx 10–15$) fall in the upper 8% of published data (figure 4.18a). An SIS tunnel junction with this orientation could have formed because the tip was pushed into the crystal, making metallic contact with superconducting material adjacent to an internal defect or crack that provided the tunnel barrier.

The undershoot feature has been likened (see, e.g., Tulina *et al.* 1990) to similar structure observed in conventional superconductor proximity-effect junctions (Adkins and Kington 1969). Theoretical work by Arnold (1978), showing that this

---

* A gap was only observed when the junction was fabricated both at low temperature and in a vacuum.

† Using peak and $\sigma = 1$ energies as bounds on $2\Delta$ (see section 3.7).
A conductance dip is caused by a gap in the local density of states at the tunnelling interface (figure 7.20), has been supported by detailed experiments with Nb–Zr alloy/Al junctions (Wolf et al. 1980a,b).

Undershoots observed in Bi$_2$Sr$_2$CaCu$_2$O$_8$ tunnel junctions have also been attributed to strong electron–phonon interaction (Aminov et al. 1990, Huang et al. 1989a) and combined effects of a parabolic background and normal density of states with a peak at the Fermi level (Huang et al. 1989b). Hasegawa et al. (1992a, p.29) claim that observation of very similar structure in photoemission spectra indicates that this is not an extrinsic junction effect but a genuine feature in the density of states. An intrinsic origin, either from the bulk or a reproducible surface effect, is also suggested by observation of dips in SIN tunnelling characteristics of Nd$_{2-x}$Ce$_x$CuO$_{4-y}$; a cuprate superconductor with very different properties (Zasadzinski et al. 1992, see appendix B).

Recently, Coffey and Coffey (1993) postulated that the undershoot arises as a consequence of deviations from weak-coupling mean-field behaviour and occurs roughly at the threshold energy for spontaneous quasiparticle decay. If the superconductor has an order parameter with s-wave symmetry the onset of this decay is at $\omega = 3\Delta_0$, whilst for a d-wave superconductor decay rises rapidly at $2\Delta_0$. Weak (barely perceptable) dips will occur at these energies in the conductance–voltage characteristics of an SIN tunnel junction (i.e., at $eV = 3\Delta_0$ for s wave and $eV = 2\Delta_0$ for d wave). In the SIS case, however, two superconductor densities of states are

![Fig. 7.20. Sketch of local density of states at a thin proximity-effect junction interface. The energy scale is highly expanded about the superconductor’s energy gap, $\Delta_S$. The density of states is zero for $E \leq (\Delta_S \Delta_N)^{1/2}$ then rises to a maximum at the bound-state energy $E_0 \approx \Delta_S \left[1 - \frac{1}{2} R^2 (\Delta_S - \Delta_N)^2\right]$. $\Delta_N$ is the energy gap induced in the normal layer. $R \approx d_N/v_{F,N}$ is typically $< 0.1$ and $d_N$ is the normal-layer thickness. Above $\Delta_S$ there is a continuum of states, which merges at high energies with the conventional BCS density of states for gap $\Delta_N$. There is no BCS-like singularity at $\Delta_N$.](image-url)
convolved together, which strengthens the feature and moves it to $eV \geq 4\Delta_0$ for s-wave superconductivity and $eV = 3\Delta_0 - 4\Delta_0$ for d-wave superconductivity. In agreement with most published tunnelling data the position of the undershoots in figure 7.19 indicate d-wave rather than s-wave pairing. However, this conclusion is at odds with many other experimental results (e.g., penetration-depth studies, Josephson and Andreev tunnelling), which indicate that high-temperature superconductors contain singlet s-state pairs, as in BCS theory.

Zero-bias conductance peaks, like that seen in figure 7.19, are often reported in the literature and their origin has been examined in detail by Walsh (1992). With an SIS junction the *prima facie* explanation for this feature is d.c. Josephson supercurrent; however, attempts to verify this assignment by looking for the characteristic magnetic field dependence of a Josephson junction have been inconclusive (Lee *et al.* 1989a, Mandrus *et al.* 1991). Zero-bias conductance peaks have also been observed with high-temperature superconductor SIN junctions; in particular, in the near-ideal (planar-junction) tunnelling characteristics observed by Matsumoto *et al.* (1992).

### 7.3.3 (Tl,Pb)(Ca,Pr)Sr$_2$Cu$_2$O$_y$ Ceramic

Point-contact measurements were performed on a single-phase, polycrystalline ceramic sample of (Tl$_{0.5}$Pb$_{0.5}$)(Ca$_{0.8}$Pr$_{0.2}$)Sr$_2$Cu$_2$O$_y$ with $T_c$ (zero resistance) of 106 K (Liu *et al.* 1989). This sample was the first to be used in the point-contact insert with fine piezo-control of the tip (sections 6.3.1–8). The junction’s conductance–voltage characteristics often varied smoothly with voltage with no evidence of gap-like structure. Occasionally however, fine adjustment of the contact could produce gap-like characteristics that resembled those of a superconductor tunnel junction (figure 7.21). In particular, these curves display the attributes used by Kirtley (1990) for selecting ‘good’ tunnelling characteristics, i.e., a low-conductance region about zero bias, conductance overshoots at the edge of this region and symmetry in position and size of the overshoots.

If the tunnelling characteristics of figure 7.21 are those of an SIS' junction formed between the sample and the tantalum tip, then the bias at which the conductance crosses the estimated background indicates a reduced energy gap in the high-temperature superconductors of $9.4 \pm 0.3$. Because this is a factor of two larger than reported in most tunnelling experiments on these materials an alternative explanation was suggested (Chandler and Adkins 1991); the tip may be in good contact with an isolated grain of high-$T_c$ superconducting material and we are observing SIS tunnelling.

---

* These experiments are reviewed by Little (1988) and Burns (1992, sec.5-2d).
between this grain and the bulk ceramic with $\Delta = 21.5 \text{ meV}$ and $2\Delta/k_B T_c = 4.7 \pm 0.3$. This explanation was also used by Vieira et al. (1989) to account for anomalously large apparent gaps.

Taken in isolation this interpretation of figure 7.21 seems reasonable; however, there is some circumstantial evidence that suggests it may be incorrect. The $I$–$V$ characteristics were acquired from nominally the same point contact as those shown in figures 7.22 and 7.23. The latter display highly asymmetric, non-linear conductance with a ‘gap’ region of reduced conductance but poorly developed or no conductance overshoots. Similar features are often seen together with Coulomb gap and staircase

![Tunnelling characteristics of a Ta–(Tl,Pb)(Ca,Pr)Sr$_2$Cu$_2$O$_y$ point contact at 4.2 K.](image)

The dynamic conductance (solid curve) is calculated from the $I$–$V$ data by numerical differentiation and smoothing.

**Fig. 7.21.** Tunnelling characteristics of a Ta–(Tl,Pb)(Ca,Pr)Sr$_2$Cu$_2$O$_y$ point contact at 4.2 K. The dynamic conductance (solid curve) is calculated from the $I$–$V$ data by numerical differentiation and smoothing.
**Fig. 7.22.** $I-V$ characteristics acquired immediately before (dashed curve) and after (solid curve) data shown in figure 7.21, from nominally the same contact. Although both curves were acquired with increasing bias their asymmetry is almost exactly reversed, i.e., the curves are almost coincident when one of them is rotated by 180° about (0,0). This phenomenon was also observed by Moog *et al.* (1988).

**Fig. 7.23.** $I-V$ characteristic obtained from the same point on the sample as the data shown in figure 7.21 but with a less extended piezotube.
Fig. 7.24. $I-V$ characteristics of a W–(Tl,Pb)(Ca,Pr)Sr$_2$Cu$_2$O$_3$ point contact. Labels refer to the curves’ position in the data-acquisition sequence.
structure due to single-electron tunnelling effects (Wilkins et al. 1990b). This behaviour was clearly observed in a series of contacts made at a different point on the same sample (figure 7.24).

The first three curves in figure 7.24, labelled 5i, 12iii and 23vi, are similar to those shown in figures 7.22 and 7.23, with a region of reduced conductance about zero bias and weak (or no) conductance overshoots (figure 7.25). Slightly adjusting the tip’s vertical position by retracting the piezotube leads to structure (curve 27i) very similar to figure 7.21 that could be mistaken for a single SIS tunnel junction. However, taking a wider voltage sweep (curves 28, 29vi) reveals multiple current steps. These are characteristic of two mesoscopic tunnel junctions in series (section 4.4), such as might be formed by an isolated conducting particle positioned between the point-contact tip and the bulk sample. The isolated particle could be a small indigenous grain on the surface of the ceramic, defect material or a chip broken from the fragile oxide by the tip. Further adjustment of the contact weakened the step structure (30i) and led to huge asymmetry in the $I$–$V$ characteristics (31i). The diversity in structure seen from contact to contact is caused by sensitivity of the current steps’ position and strength to the resistance and capacitance of both junctions. This may be aggravated by a network of double junctions formed between multiple particles.

![Conductance–voltage characteristics](image)

**Fig. 7.25.** Conductance–voltage characteristics, calculated from curve 5i in figure 7.24. This displays a large Coulomb gap but no staircase structure.
Fig. 7.26. Conductance–voltage characteristics calculated from curve 29vi in figure 7.24. The staircase is shifted by a polarization voltage so that the step close to zero bias is almost suppressed; when this occurs the central peak separation appears wider by $e/C$.

Fig. 7.27. Conductance–voltage characteristics calculated from curve 28 in figure 7.24.
The particularly uniform multiple-peak structure of figure 7.27 was used to estimate some of the junction’s properties. Only one series of peaks was evident, indicating either identical junctions or a strongly dominant series of peaks arising from the slower junction. The peak separations indicate a capacitance of $3.4 \pm 0.4 \, \text{aF}$. There is no obvious offset in the staircase structure and the estimated voltage shift is $0.0 \pm 6.9 \, \text{mV}$. The separation between the central peaks is wider than expected for an NNN junction by $81.6 \pm 14.6 \, \text{mV}$. This is too large to be attributed to a suppressed peak (as in figure 7.26) so indicates that one or more of the junction electrodes are superconducting. From (7.1) the first current step occurs at $e/2C_i + r\Delta(T)/e$, where $r$ may vary continuously from one to four depending on the ratio of junction capacitances and junction configuration (e.g., SSN). Without further information about both junctions’ capacitances the zero-bias peak separation cannot be used to estimate $\Delta(T)$ unambiguously. If one (rashly) assumes that $r = 2$, as in an asymmetric SSS configuration or an SSN configuration with similar junction time constants, then $\Delta \approx 20.4 \pm 3.7 \, \text{meV}$ and $2\Delta/k_B T_c = 4.5 \pm 0.8$. Analytical expressions for the $I$–$V$ characteristics of an arbitrary junction configuration (e.g., NSS) can be derived using the method developed by Amman et al. (1991). These may then be fitted to experimental data to extract $\Delta$, $\delta Q$ and the junction capacitances and resistances. This detailed analysis was not carried out with the data presented here, but should be pursued in future work.

### 7.3.4 YBa$_2$Cu$_3$O$_y$ Thin Films

Point-contact measurements were performed on $c$-axis-oriented thin films of YBa$_2$Cu$_3$O$_y$ prepared by laser ablation (Scott 1992). With this technique a pulse of high-energy laser radiation is used to evaporate material from a target onto a substrate positioned a few centimetres away in the vacuum chamber. If the laser pulse is short ($< 50 \, \text{ns}$) and the beam intensity profile narrow, most of the laser’s energy is converted into kinetic energy and a plume of material is ejected perpendicular to the target. Any energy absorbed by the target as heat, for example from lower-intensity radiation at the beam’s periphery, may result in damage or unwanted phase changes. The advantages of laser ablation are that deposition occurs rapidly ($\approx 1 \, \text{nm s}^{-1}$), the apparatus is easy to set up and the deposited film has the same composition as the target. In addition, the high energy of ejected material permits evaporation in an oxygen atmosphere, resulting in films with the correct oxygen stoichiometry and obviating the need for a physically damaging post-deposition oxygen-anneal.

Samples used in this work were prepared using an XeCl excimer laser ($40 \, \text{ns}$ pulse, $\lambda = 308 \, \text{nm}$, $3 \, \text{Hz}$ pulse rate) focused to an energy density of $1.3$–$1.55 \, \text{J cm}^{-2}$. Deposition occurred in an oxygen partial pressure of $30 \, \text{Pa}$ onto substrates of (100)
SrTiO$_3$ placed 4.5 cm from target and heated to 760 °C*. The target of melt-processed YBa$_2$Cu$_3$O$_y$ was rotated to minimise heating. Approximately 1000 laser shots were used to produce films with thicknesses from 200 to 300 nm. After deposition the films were cooled to 480 °C and annealed for 15 minutes in 1 atm oxygen. On being removed from the vacuum chamber the sample was immediately mounted in the helium-filled vacuum can of the point-contact insert. There was always less than 20 minutes exposure to the ambient laboratory atmosphere and samples were cooled to 77 K within 1 hour of being removed from the deposition chamber.

When tunnelling measurements were completed the sample was left in the vacuum can until the apparatus had warmed to room temperature. This avoided degradation of the sample caused by exposure to frost. The sample’s transition temperature was then found by a four-terminal, low-frequency resistance measurement. All films examined had a $T_c$ (zero resistance) of 89.0–90.5 K with transition width of 1 K.

Point contacts on all the samples studied exhibited structure characteristic of a voltage-biased double-junction Coulomb-blockade system (section 4.4). The Coulomb staircase was very pronounced and was observed over a wide voltage range (figure 7.28).

![Coulomb staircase tunnelling characteristics](image)

**Fig. 7.28.** Coulomb staircase tunnelling characteristics of a W–YBa$_2$Cu$_3$O$_y$ thin-film point contact at 4.2 K.

---

* This produces films with the $c$-axis oriented perpendicular to the plane of the substrate (Scott 1992). Identical tunnelling characteristics were obtained from one other film deposited on (110) SrTiO$_3$ at 740 °C, with the $c$-axis oriented at roughly 45° to the plane of the substrate.
Fig. 7.29. Three successive measurements of a single contact. The same sample and tip were used to obtain figure 7.28.

By fitting the models discussed in section 4.4 to the tunnelling characteristics all the junctions’ parameters could be estimated. These numerical calculations were not attempted due to lack of time. However, if the junctions’ time constants are very different the conductance–voltage characteristics will be dominated by a single series of peaks at voltages

\[ V_{s,k} = \frac{\delta Q}{C_i} \pm \left[ \left( k - \frac{1}{2} \right) \frac{e}{C_i} + \frac{r\Delta(T)}{e} \right] \tag{7.3} \]

where \( k \in Z^+ \), \( \delta Q \) is a polarisation-induced fractional charge, \( C_i \) is the capacitance of the slower junction and \( r \) may take any value between zero and four depending on the ratio of junction capacitances and the junction configuration. In this case some of the junction’s parameters may be found without fitting by using the following procedure.

From (7.3) the position of conductance peaks at positive and negative biases may be written,

\[ V_{s+,n} = \left(a_n + b_+ \right) = \frac{\delta Q}{C_i} + \left( n - \frac{1}{2} \right) \frac{e}{C_i} + \frac{r\Delta(T)}{e} \quad n \in (+1,+2,+3,\ldots) \]

\[ V_{s-,n} = \left(a_n + b_- \right) = \frac{\delta Q}{C_i} + \left( n + \frac{1}{2} \right) \frac{e}{C_i} - \frac{r\Delta(T)}{e} \quad n \in (-1,-2,-3,\ldots) \tag{7.4} \]

If the observed peak positions are plotted against the ‘peak number’, \( n \), then linear regression may be used to find \( a_\pm, b_\pm \). These may then be used to find the following characteristic voltages:
\[ V_{\text{sep}} = a_+ = a_- = \frac{e}{C_i} \]
\[ V_{\text{shift}} = \frac{(b_+ + b_-)}{2} = \frac{\delta Q}{C_i} \]
\[ V_{\text{gap}} = \frac{(b_+ - b_- + V_{\text{sep}})}{2} = \frac{r\Delta(T)}{e} \]  

(7.5)

which represent the peak separation, \( V_{\text{sep}} \), the structure’s offset \( V_{\text{shift}} \) from zero bias and any additional spacing \( V_{\text{gap}} \) between the central two peaks. As indicated, these may then be used to estimate \( C_i \), \( \delta Q \) and \( r\Delta(T) \). Care must be taken with this analysis; if the true value of \( \delta Q \) is approximately \( \pm e/2 \), then a peak will be shifted close to zero bias and will be suppressed. The peak numbers will then be incorrectly assigned, \( V_{\text{shift}} \) will appear to be zero and the calculated value of \( V_{\text{gap}} \) will be too large by \( V_{\text{sep}}/2 \). The occurrence of peak suppression is flagged by \( b_+ \) or \( b_- \) values close to zero, in which case the \( V_{\text{shift}} \) and \( V_{\text{gap}} \) values may be corrected. However, \( b_+ \) or \( b_- \) will also be zero if \( V_{\text{gap}} \) is coincidentally equal to \( V_{\text{sep}}/2 \pm V_{\text{shift}} \). This ambiguity is usually resolved by examining associated tunnelling characteristics (see, e.g., figure 7.33).

All periodic multiple-peak tunnelling characteristics observed with YBa\(_2\)Cu\(_3\)O\(_y\) thin films were analysed as described above. In all but one case (figure 7.30) \( V_{\text{gap}} \) was clearly zero to within experimental error, indicating that much of the surface of the sample was normal.

![Graph](image)

**Fig. 7.30.** Tunnelling characteristics of a W–YBa\(_2\)Cu\(_3\)O\(_y\) thin-film point contact at 4.2 K. Conductance was measured by the modulation technique. Analysing the multiple peak structure by the process outlined above gives \( C = 5.3 \pm 0.5 \text{ aF}, \delta Q/e = -0.16 \pm 0.2 \text{ and } r\Delta(T) = 48 \pm 6 \text{ meV}. \) However, weak structure at \( \approx \pm 30 \text{ mV} \) may be evidence of missing (or unresolved) \( n = 1 \) peaks. In that case the gap estimate changes to \( r\Delta(T) = 7 \pm 6 \text{ meV}. \)
The capacitance of some junctions (figure 7.31) was extremely small, resulting in a very large gap between the first \( (n = 1) \) conductance peaks. A YBa\(_2\)Cu\(_3\)O\(_y\) ‘energy gap’ with \( 2\Delta/k_BT_C \approx 13 \) was reported by Gallagher \textit{et al.} (1988) with a caveat that this may be due to tunnelling between a series of superconducting regions within the sample. Their observation of large variation and asymmetry in the position and size of the peaks, extremely sharp peaks and erratic unstable structure, suggests that SET effects were also being observed.

Attempts were made to look for systematic changes in the junction parameters as the tip–sample separation was varied. After taking several conductance–voltage characteristics at a particular tip location the voltage on the central electrode of the piezotube was adjusted whilst continuously monitoring the \( I-V \) trace. This usually resulted in a sudden change in current, indicating that there had been a gross change in the double-junction system; however, gradual changes were occasionally obtained (figure 7.32). An ability to control the spacing of conductance peaks in this way indicates that the dominant capacitance is between the tip and the isolated particle. In particular, the stray capacitance of leads far from the junction area does not seem to interfere with the Coulomb blockade.

The decreasing peak separation with increasing piezo-voltage indicates an increasing capacitance - as would be expected for decreasing tip–particle separation. One may attempt to quantitatively model this system as a parallel-plate capacitor with plate spacing \( s = s_0 - \alpha V_z \), where \( V_z \) is the piezo-voltage, \( s_0 \) is the tip–sample separation when no voltage is applied and \( \alpha \) is the extension per volt of the piezotube \((\approx -0.9 \text{ nm V}^{-1})\). Fitting a straight line to the points in figure 7.32 then yields

\[\text{current / nA} \quad \begin{array}{c}
\text{bias / mV} \\
\end{array}\]

\[\begin{array}{c}
\text{conductance / mS} \\
\end{array}\]

\[\begin{array}{c}
-500 & -250 & 0 & 250 & 500 \\
0 & 0.1 & 0.2 & 0.3 & 0.4 \\
-80 & -40 & 0 & 40 & 80 \\
\end{array}\]
Fig. 7.32. Conductance peak spacing vs voltage applied to the piezotube z-electrode. Conductance–voltage characteristics were from a W–YBa$_2$Cu$_3$O$_y$ point-contact at 4.2 K.

$\varepsilon A = (80 \pm 8 \text{ nm})^2$ and $s_0 = 68 \pm 17 \text{ nm}$, where $\varepsilon$ is the dielectric constant of material between the tip and sample and $A$ is the effective area of the capacitor plates.

These figures indicate that the tunnel-junction width is about $18.9 \pm 16 \text{ nm}$ (for a piezo-voltage of $-54.6 \text{ V}$), which is probably too large to be physically realistic (junction widths are typically $< 4 \text{ nm}$). However, this analysis is very crude and can be criticised on a number of points. Firstly, the contact geometry is uncertain and a hemisphere-plane model for the capacitance may be more realistic (Reeve et al. 1992). Unfortunately the error bars in figure 7.32 are very large so fitting to the data does not help with selection of the most appropriate model. Secondly, the value of $\alpha$ used above strictly applies to free expansion of the piezotube. Because the tip is in contact with the sample $\alpha$ may be significantly reduced so $\varepsilon A$ and the junction width will be smaller than the figures given above. Changes in contact area, and hence the capacitance, may also be occurring due to deformation of the tip or sample. Wan et al. (1990a) reported that even with a well defined junction geometry the capacitance calculated from the size of the inter-electrode particle (observed with an STM) may differ by an order of magnitude from the value estimated from its SET-type tunnelling characteristics.

The distribution of $\delta Q/e$ values observed during these experiments spanned the range $[-\frac{1}{2}, +\frac{1}{2}]$ with some evidence for peaks at $-\frac{1}{2}$, 0 and $+\frac{1}{2}$. $\delta Q/e$ did not change systematically with piezo-voltage but changed uncontrollably between voltage sweeps, often between two of the ‘favoured’ values (figure 7.33). On several occasions $\delta Q/e$ changed as the voltage applied to the junction passed through zero bias, as did the capacitance (figure 7.34).
**Fig. 7.33.** Successive NNN junction tunnelling characteristics from a W–YBa$_2$Cu$_3$O$_y$ thin-film point contact. The curves have equal peak separation but differ in their voltage shift. The dashed curve is unshifted ($\delta Q/e = 0$) whilst the solid curve is shifted by half the peak separation ($\delta Q/e = +\frac{1}{2}$) leading to suppression of the lowest-energy conductance peak and a wider central-peak separation. The $I$–$V$ characteristic corresponding to the dashed curve is plotted in figure 4.15a, together with a fitted model curve.

**Fig. 7.34.** Change in capacitance, and hence conductance peak spacing (i.e., gradient), as the junction voltage passed through zero bias (from $-$ve to $+$ve). The spacing was $60.3 \pm 1.9 \text{ mV}$ for $V < 0$ (and was this value for all voltages in the previous sweep) but $46.8 \pm 1.3 \text{ mV}$ for $V > 0$ (and for all voltages in the next sweep). On this occasion no change occurred in $\delta Q/e$. The dotted and dashed lines are least-squares fits to the $V < 0$ and $V > 0$ data respectively.
The multiple-peak tunnelling characteristics of \( \text{YBa}_2\text{Cu}_3\text{O}_y \) thin-film point contacts often contained irregularities, such as anomalously large or small peaks, additional peaks or missing peaks (figure 7.35). These effects were sometimes dependent on the direction of voltage sweep (figure 7.36). Occasionally fine structure or regions of negative differential conductance were observed (figure 7.37). Wilkins et al. (1990b) proposed three possible mechanisms to explain the latter: resonant tunnelling via states in a quantum well formed by an extremely small isolated particle, the formation of an Esaki diode in a semiconducting surface layer on the sample and tunnelling at localised surface states. The last explanation was suggested by similar results obtained when tunnelling at boron impurity sites on silicon.

It is interesting to consider the origin of the isolated ‘particle’ that forms the central electrode of the double junction. The prevalence of (MMM) Coulomb staircase structure suggests a very high density of such regions, many of which are normal rather than superconducting. SEM examination of films used in this work showed that they were not smooth (Scott 1992, fig.5.19), for although the surface was free of boulders* even the highest-quality films appeared to have a fine granular structure. In addition, the surface was covered with irregularly shaped pits with a mean diameter of about 150 nm and a high density of approximately \( 2 \times 10^8 \text{ cm}^{-2} \). STM

![Graph](image)

**Fig. 7.35.** Irregular multi-peak structure of a Ta–\( \text{YBa}_2\text{Cu}_3\text{O}_y \) point contact.

* Lumps of material often seen on the surface of laser-ablated films. They are possibly fully intact grains ejected from the target. Use of a dense target of melt-processed \( \text{YBa}_2\text{Cu}_3\text{O}_y \) may have avoided their formation.
**Fig. 7.36.** Conductance of a W–YBa$_2$Cu$_3$O$_y$ point contact at 4.2 K, showing hysteretic dependence of the first conductance peak’s position and height on direction of voltage sweep (shown by arrows). Each curve was measured several times.

**Fig. 7.37.** Coulomb-staircase structure of a very stable Ta–YBa$_2$Cu$_3$O$_y$ point contact, with regions of negative differential conductance and fine structure.
examination of thin films similar to those used in this work reveal that their surfaces are not continuous but completely covered with a high density ($10^9 \text{ cm}^{-2}$) of concentric step and spiral growth hills (see, e.g., Lang et al. 1992, Gerber et al. 1991). These form by preferential nucleation of the growing film at the emergence point of a screw dislocation. This structure has been seen by several groups using various deposition methods (including laser ablation) and is believed to be an intrinsic property of $c$-axis-oriented, high-$T_c$ superconductor thin films (Chapman et al. 1992). Screw dislocations and defects at boundaries between growth hills are thought to act as strong pinning sites for magnetic flux lines; the high defect density could explain the extremely high critical current densities of thin films when compared to bulk material. The thin films described in this section also had high critical current densities of about $5 \times 10^6 \text{ A cm}^{-2}$ at 77 K (Scott 1992).

This evidence suggests that the Coulomb staircase tunnelling characteristics were obtained from samples with a complex surface structure, containing many internal and external grain boundaries. If YBa$_2$Cu$_3$O$_y$ is intrinsically degraded at surfaces and line defects, as proposed by Halbritter (section 4.8), then this could divide the near-surface region into a mosaic of normal or superconducting ‘islands’ separated by seams of insulating material. We speculate that this is the mechanism by which the isolated normal metal particles were formed.
CHAPTER 8

CONCLUSIONS

Sadly, little information about the physics of high-temperature superconductors can be concluded from the original work described in this dissertation. The conclusions that can be drawn are more in the nature of suggested improvements for future experiments. Before presenting these suggestions the results of tunnelling experiments on high-$T_c$ superconductors will be briefly summarised. These results fall a long way short of the stated aim of this work - to measure the excitation density of states of high-temperature superconductors - and are not unique, since all characteristics seen in this work have also been observed and reported by other groups.

8.1 Summary

To start on a positive note, evidence for an energy gap has been observed, if infrequently, in all the high-$T_c$ superconducting materials studied. However, the tunnelling characteristics were all complicated by the presence of non-ideal features that prevented confident quantitative estimation of the gap value. In particular, samples of ceramic YBa$_2$Cu$_3$O$_y$ were dominated by a V-shaped conductance background and only weak gap-like features were seen. The gap was clearer and the background less prominent in the tunnelling characteristics of a Bi$_2$Sr$_2$CaCu$_2$O$_8$ single crystal. However, samples of this type were difficult to work with due to their small size, and only one crystal out of the many examined exhibited gap-like tunnelling characteristics. This single successful result appears to have arisen through the serendipitous formation of an SIS junction rather than by improved experimental technique. The relatively more ideal characteristics of this single-crystal point contact permitted estimation of a gap ratio of 5–7.5. Although there is little confidence in this value, due to use of the sample’s bulk transition temperature and uncertainty in the exact junction configuration, it is larger than predicted by BCS theory and agrees with most reports in the literature (figure 4.18).

The tunnelling characteristics of both ceramic (Tl,Pb)(Ca,Pr)Sr$_2$Cu$_2$O$_y$ and thin-film YBa$_2$Cu$_3$O$_y$ samples were dominated by Coulomb gap and staircase structure (i.e., charging or SET effects). This is believed to be due to the formation, by an isolated conducting particle positioned between the point-contact tip and the bulk electrode, of two small tunnel junctions in series. Although much of the surface of these samples
was normal, evidence was seen for an energy gap. However, an unambiguous estimation of the gap value was not possible with the simplified models used in this work.

### 8.2 Recommendations for Future Work

In discussing possible improvements to the point-contact apparatus at the end of chapter 7, the advantages of a recently built STM were described. It would clearly be sensible to use this, and its associated control and data-acquisition electronics, in any future point-contact tunnelling studies of high-temperature superconductors. Following the discussion in section 4.5, the author believes that a facility to cleave samples *in situ* at low temperature (< 70 K) is essential in such work; this appears to be the only way to guarantee that surfaces examined by a point contact are representative of the bulk.

The discussion in section 6.3.6 suggests that electrochemically etched tungsten is susceptible to charging effects and should not be used as a point-contact or STM tip in spectroscopic studies. Although the Coulomb staircase behaviour described in sections 7.3.3–4 is believed to be due to isolated particles at the surface of the sample, one cannot rule out the possibility that these are also tip-induced artefacts. This uncertainty should be resolved by repeating the experiments with more suitable, mechanically sharpened tips. In retrospect, it is clear that whilst performing the majority of experiments described in this dissertation the author was insufficiently aware of the importance of charging effects. In particular, the voltage ramp applied to the junction was usually too small, so only the first (positive and negative) steps of a Coulomb staircase were seen and measured. This resulted in incorrect or ambiguous interpretation of some data (e.g., figure 7.21). To help distinguish between simple SIN tunnelling and charging effects, and to permit full analysis of the latter, the junction voltage should be swept over at least three times the bias of the lowest-energy conductance peaks. This is the practice recommended for future work.

To avoid SET effects ceramic samples should not be used. These are also very inhomogeneous in composition and cleaving may expose defect-rich grain boundaries rather than bulk material. The results of section 7.3.4 indicate that YBa$_2$Cu$_3$O$_y$ thin films also suffer from SET effects. For this reason it is suggested that future work should concentrate on single-crystal samples. Moreover, attempts should be made to fabricate junctions with tunnelling in the $ab$-plane of these crystals*, since results are generally superior to those of $c$-axis tunnelling, probably due to the longer coherence

---

* Possibly using the crossed-crystal technique of Wnuk *et al.* (1991) and variants thereof.
The Bi-Sr-Ca-Cu-O family of high-$T_c$ superconductors would be a suitable initial focus for these studies, since these seem to produce more BCS-like tunnelling characteristics and single crystals can now be grown to a considerable size (e.g., $5\times5\times0.5$ mm$^3$, Tanaka et al. 1989). Establishing a source of high-quality single-crystal samples should be given higher priority in future work.
APPENDIX A

ELECTRONIC CIRCUIT DETAILS

A.1 Measurement of Conductance–Voltage Characteristics

A.1.1 Circuit Design

Circuits used to measure current–voltage and conductance–voltage characteristics are illustrated in figures 5.1 and 5.2. This section discusses their design and construction.

A.1.1.1 Sample-Biasing Electronics

OPA111 operational amplifiers were used to voltage bias the sample (OP1 and OP2 in chapter 5). These were selected because of their low noise, low input bias currents (1 pA) and low offset and drift characteristics. To prevent degradation of the drift performance offset adjustment was not used; however, the resulting 0.25 mV offset at the sample was subtracted by software when calibrating the raw data.

A very low phase-shift OPA633 buffer amplifier was placed in the OPA111’s feedback loop (figure A.1). This enabled the circuit to produce large currents of 100 mA at the maximum sample bias of 1 V, removed load-dependent voltage drift and prevented feedback instabilities when driving capacitive loads (Horowitz and Hill 1980, p.277). The OPA633 was mounted in a DIP socket for easy replacement if damaged by excessive currents, whilst output-protection circuitry was used to minimise the possibility of damage occurring. These precautions prevented any damage to the OPA111 op-amps and allowed these to be soldered directly to the PCB as recommended by the manufacturer (Burr-Brown).

The d.c. voltage input to the sample-biasing electronics was provided by a DAC in the CIL6580 unit, controlled by computer via an RS-232 interface. To use the full 16-bit resolution of the DAC, the output voltage ramp was always defined to be between ±10 V and was attenuated to the required full-scale sample bias by an adjustable potential divider.

A simple low-pass $RC$ filter was added to the potential divider to attenuate high-frequency noise (>10 kHz) generated by the DAC (figure A.2). The resulting attenuation was given by

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{1}{(1 + R_1/R_{\text{atten}}) + j\omega C (R_1 R_2/R_{\text{atten}} + R_1 + R_2)}$$
Fig. A.1. Construction of the sample-biasing op-amps, OP1 and OP2. This shows the high-speed buffer amplifier and its output protection circuitry.
where $V_{in}$ is the voltage applied to the $RC$ network and $V_{out}$ is the voltage at the input to OP1. With $R_1 = R_2 = 10\,\text{k}\Omega$ and $R_{atten}$ values from 1 $\Omega$ to 1 $\text{k}\Omega$ (in a 1–3–5 sequence), the full-scale sample bias for maximum DAC output of 10 V could be varied from 1 mV to 0.9 V. The filter was used primarily to reduce noise in the d.c. measurement of current, so the time constant was chosen to be much smaller than the time between voltage steps (45 ms, but potentially < 1 ms). In addition, this attenuation was applied to the a.c. modulation, which was superimposed on the d.c. voltage before the potential divider, so the cut-off frequency was set even higher to prevent filtering of the 1 kHz modulation signal. With $C = 2\,\text{nF}$ the $-3\,\text{dB}$ frequency changes from 7.3 kHz to 8.0 kHz as $R_{atten}$ is adjusted over its full range. A better arrangement would be to attenuate and filter the d.c. bias before adding the modulation and to set the filter cut-off frequency to remove DAC noise at the modulation frequency.

If OP1 acts as an ideal follower the voltage across the sample is identical to the voltage at the non-inverting input to OP1, with no phase shift or change in amplitude between the two. Measurement of the modulation amplitude and adjustment of the phase may then be more conveniently done using a single-ended input lock-in amplifier connected to the input of OP1. By measuring the voltage at this point and across the sample it was found that the phase shift and amplitude change were negligible for frequencies below 10 kHz and for typical sample impedances. When the circuit is operated above this limit additional phase shifts of unknown size cause the in-phase and quadrature outputs of the lock-in to be an inseparable mixture of the capacitance and conductance signals. An improvement to the circuit suggested by Holden (1992) would be to include a two-pole reed-relay switch controlled by the CIL6580 that allows the differential amplifier to be switched from across the sample to across the sensing resistor. This would allow the conductance and the capacitance of the sample to be measured separately even at high frequencies.
A.C. stability of the active voltage-biasing electronics is analysed by Speakman (1992, sec.3.5.4.1). The circuits described in this appendix have a larger phase margin, but a 15 pF capacitance was still used across the OPA111 op-amps to provide additional stabilising feedback at high frequencies (> 1 MHz). No difficulties were experienced with instability, even with dummy samples of large capacitance.

A.1.1.2 Differential Amplifier

An INA110 instrumentation amplifier was used to amplify the voltage drop across the sensing resistor and to isolate the sample-biasing electronics from the lower-quality input amplifiers of the CIL6580. The INA110 has low bias current (50 pA), very low gain drift and low offset drift so allows accurate measurements of small nanoamp currents. The manufacturer’s recommended PCB layout was used and the gain was set to 200. The input leads were actively shielded by an OPA121 buffer driven by an average of the input voltages, available on pin 3*. The offset voltage was nulled using an OPA27 buffer to the reference (pin 6). This prevents degradation of offset voltage drift. The offset current, measured with no sample connected, was typically smaller than 0.1% of the full-scale current and had a standard deviation (a measure of the noise in the current) of less than 0.01%.

A.1.1.3 Noise and Shielding

Most experiments carried out in this work had sample resistances of 10 kΩ – 10 MΩ and measurements were made with the sample biased between ± 200 mV. Typical currents were less than 10 nA so effort was made to minimise noise pickup and to reduce current offsets to pA levels.

All electronic circuits were constructed on PCBs, which had lower noise levels than prototype circuits built on veroboard. All PCB signal tracks were placed next to wide ground tracks and extensive use was made of ground planes. Wide tracks were also used for the power supply lines, which were connected to ground close to each op-amp by decoupling capacitors. Connections to the PCB were made using miniature BNC cable via SMC connectors. In addition to reducing pickup, these connectors allowed the board to be removed easily from its diecast aluminium box for repair or adjustment.

An outline of the grounding and shielding of the electronics is shown in figure A.3. This shows that there are no ground loops and that the signal and signal reference voltages have been carefully shielded. Low-noise co-axial cable was used for the

* This is actually the average plus a large offset of −1.1 V. Because the a.c. voltage follows the input the active shields are still effective.
Appendix A 10k W OP1 10k W 2nF Aux. Output ADC DAC

RC filter: 800Hz – 8MHz

INA110 KP inst. amp. (x200)

Krohn-Hite Osc.

CIL6580 DAC/ADC unit

Bias In

240 : 6.3

Cryostat

Fig. A.3. Outline of the electronics shielding and grounding.

 INA110 KP inst. amp. (x200)

RC filter: 800Hz – 8MHz

CIL input: channel 0

 mains earth

INA110 KP inst. amp. (x200)
sample leads that connected the electronics to the cryostat. Unfortunately there was not sufficient space on the cryostat’s top plate for four BNC sockets so the shielding had to be broken over a short distance to admit the wires into the cryostat via a nine-pin feedthrough. Inside the cryostat the sample wires, in addition to being fed through grounded metal tubing and routed as near to the sample and tip as possible, were made from low-temperature co-axial wire*. A few millimetres of thin unshielded copper wire made the final connection between the co-ax cable and the sample. High-voltage leads to the piezotube entered the cryostat through a separate feedthrough and were fed down a separate metal tube. At all times the signal and high-voltage leads were kept far apart to minimise pickup.

Circuit ground of the sample-biasing electronics was defined at the source of the bias voltage (i.e., ground of the CIL6580’s DAC) and was connected to the source by screened twisted-pair cable. Similarly, circuit ground of the instrumentation amplifier was defined by the ground of the recording electronics (i.e., ground of the CIL6580’s input amplifier). The sample-biasing electronics and instrumentation amplifier were powered by separate floating-earth power supply units. The op-amps used ‘star-point’ grounding and power lines wherever possible and the shielding was connected to earth at only one point. The principles guiding this choice of circuit layout are described by Morrison (1986).

A.1.1.4 Calibration Errors

The calibration expressions, calculated in chapter 5, were derived by assuming that all current flowing through the sample also passed through the balance (or sense) impedance. In practice, the instrumentation amplifier, op-amps and lock-in amplifier have finite input impedances and input bias currents, which become significant with large impedance samples (>1 MΩ). The d.c. input bias current of OP3 (< 50 pA) causes an offset in the measured I–V characteristic (i.e., there appears to be a small current at zero sample bias); however, this could be compensated for at run time as described in section 5.2.1. There is a smaller d.c. leakage current into the V+ and V–sense leads due to the input bias currents of OP1 and OP2 (< 1 pA). This contributes slightly to the current offset but also affects the bias across the sample due to the resulting voltage drop between the inverting input of the op-amps and the sample. Empirical measurements of current offset, noise and errors in known conductances measured with the d.c. measurement system are described in section 5.2.1.

* Lake Shore Cryotronics, Inc. Type C.
A.1.2 Modulation Broadening

This section analyses the method of measuring derivatives of a function $f(v)$ by superimposing a small sinusoidal modulation, $A \cos \omega t$, on a d.c. bias, $v_0$, and time averaging the resulting signal using a lock-in amplifier. The lock-in actually measures the r.m.s. amplitude of the $\omega$ or $2\omega$ component of a Fourier series expansion of $f(v_0 + A \cos \omega t)$. This may be represented by

$$s_m(v_0) = \sqrt{2} \frac{T}{T/2} \int_{-T/2}^{T/2} f(v_0 + A \cos \omega t) \cos(m \omega t) \, dt$$  \hspace{1cm} (A.1)$$

where $s_1(v_0)$ and $s_2(v_0)$ are the $\omega$ and $2\omega$ measurements respectively. To show how $s_m(v_0)$ is related to the derivatives consider a Taylor-series expansion of $f(v)$ about the d.c. bias voltage:

$$f(v) = f(v_0 + A \cos \omega t)$$

$$= f(v_0) + f'(v_0)A \cos \omega t + \frac{1}{2!} f''(v_0)(A \cos \omega t)^2 + \ldots + \frac{1}{n!} f^{(n)}(v_0)(A \cos \omega t)^n + \ldots$$

Here, $f^{(n)}(v)$ is the exact $n^{th}$ derivative of $f(v)$. Expanding powers of $A \cos \omega t$ and collecting terms with the same frequency gives, for $A \ll 1$

$$f(v) = f(v_0) + \frac{A^2}{4} f''(v_0) + \sum_{n>0} \frac{A^n \cos n \omega t}{2^{n-1} n!} \left( f^{(n)}(v_0) + O(A^2) \right)$$

Substituting this expression into (A.1), using the orthogonality of cosines and rearranging gives

$$f^{(m)}(v_0) \approx F^{(m)}(v_0) = \frac{m!}{\sqrt{2} \, 2^{m}} s_m(v_0) \quad m > 0$$  \hspace{1cm} (A.2)$$

$F^{(m)}(v_0)$ is the approximate derivative calculated from the lock-in amplifier’s output. For small $A$,

$$\lim_{A \to 0} F^{(m)}(v_0) = f^{(m)}(v_0)$$  \hspace{1cm} (A.3)$$

But how are $f^{(m)}(v_0)$ and $F^{(m)}(v_0)$ related for finite $A$? Clearly the modulation should cause the time-average function, $s_m(v_0)$, to depend on the behaviour of $f(v)$ not just at $v = v_0$ but for values of $v$ roughly in the range $v_0 - A$ to $v_0 + A$. $F^{(m)}(v_0)$ may be represented by a convolution of the exact derivative function with a broadening function, $b_m(v)$.

$$F^{(m)}(v) = f^{(m)}(v) * b_m(v)$$  \hspace{1cm} (A.4)$$
To calculate $b_m(v)$ consider a test function

$$f(v) = \begin{cases} 
\frac{v^{m-1}}{(m-1)!} & v \geq 0 \\
0 & v < 0
\end{cases}$$

For this particular function, substituting (A.1) into (A.2) gives

$$F^{(m)}(v) = \frac{2^m m!}{\pi A} \int_0^{\cos^{-1}(v/A)} \left( \frac{v}{A} + \cos \theta \right)^{m-1} \cos(m\theta) d\theta$$

(A.5)

The $m^{th}$ derivative of the test function is the delta function so (A.4) simplifies to

$$F^{(m)}(v) = f^{(m)}(v) * b_m(v)$$

$$= \delta(v) * b_m(v)$$

$$= b_m(v)$$

and (A.5) is exactly the broadening function, $b_m(v)$. This may easily be evaluated to give

$$b_1(v) = \frac{2}{\pi A^2} \text{Re}\left(\sqrt{A^2 - v^2}\right)$$

(A.6)

and

$$b_2(v) = \frac{8}{3\pi A^4} \text{Re}\left(\sqrt{A^2 - v^2}\right)^3$$

(A.7)

Although these were calculated for a special case, the broadening function should be independent of $f(v)$ and (A.6) and (A.7) are true in general. The FWHM of $b_1(v)$ and $b_2(v)$ are $A\sqrt{3}$ and 1.22A respectively. $A$ is usually chosen so that modulation broadening is equal to thermal broadening, which has width $5.4 \ k_B T$ for inelastic tunnelling (3.22) and $3.5 \ k_B T$ for elastic tunnelling (3.15–17).

Modulation-like additional broadening, observed by Speakman (1992, sec.3.7.4.3) in high-resistance samples (>10 kΩ) and attributed to high-frequency noise from the voltage source, was not noted as a problem.

**A.2 Servo Control of the Tunnelling Tip**

The tip was advanced towards the sample by a motor-driven differential screw, controlled by the circuit shown in figure A.4. During this advance the piezotube was extended by applying $-100$ V to the inner electrode. The motor and piezotube voltages were applied via reed relays which were energised if the input to the relay
Fig. A.4. Threshold detector and latch circuit used to control the coarse stage of tip–sample approach.
driver was high. This state was pre-set by briefly closing switch SW1. Initially the tip was far from the sample’s surface so no tunnel current flowed and the output of the instrumentation amplifier in the current–voltage measuring electronics was well below the Schmitt trigger’s comparison voltage, $V_{\text{thresh}}$. The output of the Schmitt trigger was then high and the NOR gate give a low at the clock input of the D-type flip-flop. When the tip made contact with the sample the current rose above the threshold voltage and the Schmitt trigger output dropped. The NOR gate gave a rising edge at the clock input so the low at the D-input was passed to the output and the relay driver turned off. Both the piezotube and motor voltages were cut, retracting the tip and stopping the motor. The system stayed latched in this state until SW1 was closed again. By closing switch SW2 the servomechanism was bypassed (i.e., the piezotube was permanently energised and motor was off) so the piezotube could be manually controlled.

The Schmitt trigger was built with a 311 comparator (Horowitz and Hill 1980, fig.3.54) and had a fixed threshold voltage of 4 V. The corresponding threshold current could be changed by adjusting $R_{\text{sense}}$. Audio and visual indicators were incorporated into the circuit to warn of triggering events and to show clearly which relays were turned on.

Figure A.4 illustrates the servocontrol circuit just after a triggering event. The piezotube is retracted, the motor and LEDs are off and the buzzer is on.
APPENDIX B

LOW-\(T_c\) OXIDE SUPERCONDUCTORS

B.1 Bismuthate Superconductors

Prior to 1986 the highest known transition temperature in an oxide superconductor was 13 K in BaPb\(_{1-x}\)Bi\(_x\)O\(_3\) with \(x = 0.75\). The discovery of high-temperature superconductivity in the cuprates stimulated further investigation of bismuthate materials culminating in the discovery (in 1988) of \(T_c \approx 30\) K in Ba\(_{1-x}\)K\(_x\)BiO\(_3\) with \(x \approx 0.4\).

The ‘parent’ compounds of both the cuprate and bismuthate superconductors (e.g., BaBiO\(_3\)) have half-filled bands so would be metallic according to simple band theory; however, these are energetically unstable so a phase transition to an insulating state occurs (Hinks 1990). In the cuprates this is an antiferromagnetic ordering; however, the bismuthates are non-magnetic and the change is structural. As doping of the parent compounds is increased they remain insulating until at a critical level there is a transition to a superconducting metallic state. In the bismuthates the highest transition temperature occurs at this point and further doping reduces \(T_c\).

Bismuthate superconductors lack the CuO\(_2\) layers that are widely believed to play a role in the origin of high-\(T_c\) superconductivity. In addition, they are three-dimensional in comparison with the quasi-2D cuprates, having a cubic crystal structure and more isotropic properties. Being chemically simpler than the cuprates, sample preparation has improved more rapidly and pure materials are now available even in ceramic form - resulting in elimination of extrinsic effects that led to anomalous results from early experiments. This is particularly true of tunnelling, where the much larger coherence length \((\approx 4\) nm) also reduces sensitivity to sample inhomogeneity and surface layers.

Early tunnelling results on bismuthate superconductors were similar to those still observed in the cuprates, with excessive broadening, a wide range of voltage-dependent backgrounds and scattered energy gap values. Zasadzinski et al. (1989) prepared planar junctions by evaporating indium onto ceramic Ba\(_{1-x}\)K\(_x\)BiO\(_3\), with the tunnelling barrier provided by an intrinsic surface layer. The conductance–voltage characteristics were much broader than the BCS result and had a large zero-bias conductance of 50% of the above the gap value. Dyne’s lifetime-broadening model with \(\Delta = 7.0\) meV and \(\Gamma = 5.0\) meV was used to explain these features.

The same sample was later investigated using a gold point contact (Huang et al. 1990b), yielding normalised conductance–voltage characteristics in perfect agreement with a thermally smeared BCS result. No additional broadening was
necessary and there was no leakage current in the gap region. A large variation in energy gap measured at different points on the sample’s surface (3.6–4.6 mV) and a broadened superconducting transition (28.5–27.5 K) were attributed to inhomogeneity in potassium doping within the sample. Evidently, the point contact was able to tunnel into individual grains whilst the planar junction could only measure averaged properties. Similar results have been obtained by several other groups, e.g., Valles and Dynes (1990).

Sample quality has improved still further. Hou et al. (1992) recently fabricated large-area planar SIN junctions on thin films of Ba$_{1-x}$K$_x$BiO$_3$ and Ba$_{1-x}$Rb$_x$BiO$_3$. Their narrow superconducting transition widths and near-ideal BCS-like tunnelling characteristics (with $2\Delta/k_B T_c = 3.5 \pm 0.2$) indicate a very homogeneous material.

Attempts have been made to invert the tunnelling data from Ba$_{1-x}$K$_x$BiO$_3$ to obtain the effective phonon spectrum, $\alpha^2 F(\omega)$. Huang et al. (1990c) claim that structure in their tunnelling spectrum correlates with the phonon spectrum determined by inelastic neutron scattering. Moreover, the transition temperature calculated from $\alpha^2 F (24.5 \text{ K})$ was in good agreement with the measured $T_c$ (20 K). To obtain this agreement however, it was necessary to use a modified McMillan–Rowell procedure that includes the effect of a thin surface proximity layer. The electron–phonon coupling strength ($\lambda = 1.2 \pm 0.3$) and reduced gap value ($2\Delta/k_B T_c = 3.6–4.3$) determined by this analysis are consistent with trends found in conventional superconductors and suggest that superconductivity in these materials may be explained by a conventional phonon-mediated pairing interaction. Hinks (1990) draws the same conclusion in his review of the properties of Ba$_{1-x}$K$_x$BiO$_3$. Although $\lambda$ is low, indicating that Ba$_{1-x}$K$_x$BiO$_3$ is weakly coupled, electrons appear to be coupled to high-energy phonons, resulting in a large average frequency of about 40 meV and a high $T_c$. Although the tunnelling-inversion results from Ba$_{1-x}$K$_x$BiO$_3$ are more reproducible than those of the cuprates, some experimenters remain unconvinced that the bismuthates are conventional, phonon-coupled BCS superconductors (Dynes et al. 1991).

One anomalous feature that persists even in ‘ideal’ bismuthate tunnelling characteristics is a V-shaped conductance background - identical to that seen with the cuprates. The presence of this phenomenon in the non-magnetic and cubic bismuthates casts doubt on explanations invoking spin fluctuations (section 4.2.3) or reduced dimensionality, unless one is prepared to ignore Occam’s razor. Because high-quality tunnel junctions can be prepared on bismuthate materials with a wide range of transition temperatures (1 K–30 K), Sharifi et al. (1991) used them to investigate the $T_c$ dependence of the linear conductance.
Firstly, using $\text{Ba}_{1-x}\text{K}_{x}\text{BiO}_3$ junctions with fixed area and $T_c$ but with barriers prepared under different conditions*, they found that the linear conductance slope was proportional to the zero-bias conductance. This indicates that

$$G(V) = G(0)(1 + kV)$$

where $G(V)$ is the dynamic conductance, $dI/dV$, and $k$ is a constant. Secondly, using junctions made from different bismuthate superconductors, with transition temperatures ranging from 1 K to 30 K, they observed a remarkable linear correlation between the slope of $G(V)/G(0)$ and $T_c$, i.e.,

$$\frac{1}{G(0)}\left(\frac{dG(V)}{dV}\right)_{V>0} = k \propto T_c$$

They also noted a consistent asymmetry in the conductance slope of all the junctions studied, i.e.,

$$\left|\frac{dG(V)}{dV}\right|_{V>0} \approx \frac{1}{2} \left|\frac{dG(V)}{dV}\right|_{V<0}$$

These results have been spelt out here because this $T_c$ dependence has been misinterpreted by some authors (see sections 4.2.5 and 4.8).

**B.2 Nd$_{2-x}$Ce$_x$CuO$_{4-y}$**

This is a relatively low-$T_c$ cuprate with a maximum value of 25 K when $x \approx 0.15$ and $y \approx 0.02$. $\text{Nd}_{2-x}\text{Ce}_x\text{CuO}_{4-y}$ is discussed in this appendix because, as with $\text{Ba}_{1-x}\text{K}_{x}\text{BiO}_3$, high-quality BCS-like tunnelling characteristics have been obtained and inversion has yielded an effective phonon spectrum in reasonable agreement with neutron-scattering data (Zasadzinski *et al.* 1992). Reduced gap values estimated from the tunnelling data are lower than in other cuprates and similar to the BCS weak-coupled value (e.g., $2\Delta/k_B T_c \approx 3.9$, Huang *et al.* 1990c).

Although this material does contain CuO$_2$ layers, several properties distinguish it from other cuprates. Most importantly the antiferromagnetic insulating parent compound, $\text{Nd}_{2-x}\text{CuO}_{4-y}$, is electron doped by Ce substitution. Secondly, although the coherence lengths are small ($\xi_c = 1.5$ nm, $\xi_{ab} = 7$ nm) they are larger than in other cuprates - possibly explaining the superior tunnelling results. Other differences are discussed in the review by Maple (1990).

* By ion-milling damage to the film's surface. Junction resistance could be controlled by changing the beam voltage, current and milling time.
REFERENCES

Particularly useful sources of information are indicated by bold type.

Abrikosov, A.A.

Adkins, C.J.

Adkins, C.J. and B.W.Kington.


1989. JETP Lett. 49. 327.


1987. JETP Lett. 46. 547.


Amman, M., R.Wilkins, E.Ben-Jacob, P.D.Maker and R.C.Jaklevic.

Anderson, P.W. and Z.Zou.

Arko, A.J., R.S.List, R.J.Bartlet, S.W.Cheong, Z.Fisk, J.D.Thompson, C.G.Olson,
A.-B.Yang, C.Gu, B.W.Veal, J.Z.Liu, A.P.Paulikas, K.Vandervoort, H.Claus,
J.C.Campuzano, J.E.Schriber and N.D.Shinn.

Arnold, G.B.

Bakunin, O.M., S.M.Klotsman, S.A.Matveev and K.A.Steanov.  

Balzarotti, A., F.Patella, N.Motta, A.Cucola, R.Di Leo and M.De Crescenzi.  


Barone, A., A.Di Chiara, G.Peluso, U.Scotti di Uccio, A.M.Cucolo, R.Vaglio,  
F.C.Matacotta and E.Olzi.  

Barone, A., A.Di Chiara, F.Fontana, G.Paterno, L.Maritato, G.Peluso, G.Pepe and  
U.Scotti di Uccio.  


Beasley, M.R.  

Bednorz, J.G. and K.A. Müller.  


Berthe, R. and J.Halbritter.  

Beuermann, G.  

Binnig, G. and D.P.E.Smith.  


Briceno, G. and A.Zettl.  


Burns, G.  

Carr, R.G.  

Cava, R.J.  


Cockburn, J.W.  

Coffey, D., and L.Coffey.  

Coffey, L.  

Combescot, R.  

Coombs, J.H.  
Cooper, S.L., F.Slakey, M.V.Klein, J.P.Rice, E.D.Bukowski and D.M.Ginsberg.  

de Gennes, P.G.  

Deutsher, G. and K.A.Müller.  

Dimos, D., P.Chaudhari and J.Mannhart.  


Dwir, B.  


Dynes, R.C.  

Edgar, A. and A.Zyskowski.  

Edgar, A.  


Edwards, N.H.  

Edwards, P.  

Egdell, R.G., W.R.Flavell and M.S.Golden.


Eliashberg, G.M.


Esaki, L. and P.J.Stiles.

Evans, U.R.

Evettts, J., ed.


Fein, A.P., J.R.Kirtley and M.W.Shafer.


Flensberg, K., P.Hedegård and M.Brix.

Flensberg, K., P.Hedegård and M.Brix Pedersen.

Flensberg, K. and P.Hedegård.

Frick, M. and T.Schneider.

Frick, M. and T.Schneider.

Frick, M. and T.Schneider.

Fulton, T.A. and G.J.Dolan.

Gallagher, M.C., J.G.Adler, J.Jung and J.P.Franck.


Gao, Y., T.J.Wagener, J.H.Weaver, B.Flandermeyer and D.W.Capone II.

Gavaler, J.R. and A.I.Braginski.


Geilikman, B.T. and V.Z.Kresin.


Gerson, R.

Giaever, I., H.R.Hart Jr and K.Megerle.


Gough, C.  

Grafström, S., J.Kowalski and R.Neumann.  

Grajcar, M., A.Plecenik, M.Darula and S.Benacka.  

Gunn, M.  


Halbritter, J.  

Halbritter, J.  

Halbritter, J.  

Harrison, W.A.  

Hasegawa, T., M.Nantoh, H.Ikuta and K.Kitazawa.  

Hasegawa, T., H.Ikuta and K.Kitazawa.  

Hasegawa, T., M.Nantoh, A.Takagi, H.Ikuta, M.Kawazaki, H.Koinuma and K.Kitazawa.  

Hawley, M.E., K.E.Gray, D.W.Capone II and D.G.Hinks.  


Kirtley, J.R.

Kirtley, J.R. and D.J. Scalapino.

Kirtley, J.R., S. Washburn and D.J. Scalapino.

Kittel, C.

Koike, Y., T. Nakanomo and T. Fukase.

Kotyuzhanskii, B.Ya.


Laikhtman, B.

Lambe, J. and R.C. Jaklevic.


Levin, K., J.H.Kim, J.P.Lu and Q.Si.  

Levinstein, H.J. and J.E.Kunzler.  

List, R.S., A.J.Arko, Z.Fisk, S.-W.Cheong, S.D.Conradson, J.D.Thompson, 
C.B.Pierce, D.E.Peterson, R.J.Bartlett, N.D.Shinn, J.E.Schriber, B.W.Veal, 
A.P.Paulikas and J.C.Campuzano.  

Little, W.A.  

Littlewood, P.B. and C.M.Varma.  


Luo, Y.S., Y.N.Yang and J.H.Weaver.  

Maekawa, S., H.Ebisawa and Y.Isawa.  

Mandrus, D., L.Forro, D.Koller and L.Mihaly.  


Maple, M.B.  
1990. MRS Bull. 15. 60.

Matsumoto, T., T.Kawai, K.Kitahama and S.Kawai.  

Matsumoto, T., T.Kawai, K.Kitahama and S.Kawai.  


McLean, A.B.  
McMillan, W.L. and J.M.Rowell.  

Medina, R., J.Aponte and M.Octavio.  

Merzbacher, E.  


Mitrovic, B., H.G.Zarate and J.P.Carbotte.  


Moreland, J. and A.F.Clark.  


Morrison, R.  

Mühlschlegel, B.  

Mullen, K., E.Ben-Jacob, R.C.Jaklevic and Z.Schuss.  

Mullen, K., E.Ben-Jacob and S.Ruggiero.  


Nguyen, C., H.Kroemer and E.L.Hu.  

Pankove, J.I.  

Phillips, J.C.  

Phillips, J.C.  

Pigram, D.  

Pohl, D.W.  


Que, W.-M. and G.Kirczenow.  

Ramos, M.A. and S.Vieira.  

Reeve, M.D., O.G.Symko and R.Li.  


Renner, Ch., Ph.Niedermann, A.D.Kent and O.Fischer.  


Schneider, T., H.De Raedt and M.Frick.  

Schneider, T. and M.P.Sörensen.  
Scott, K.

Sharifi, F., A.Pargellis and R.C.Dynes.

Simmons, J.G.

Smith, C.W., R.C.Reinertson and P.J.Dolan, Jr.

Smith, D.P.E. and G.Binnig.

**Solymar, L.**

Speakman, A.

Speakman, A.

Srikanth, H.

Srikanth, H. and A.K.Raychaudhuri.


Svistunov, V.M., S.Benacka, A.Plecenik, S.Chromic, M.A.Belogolovskii and A.I.Khachaturov.

Takahashi, S. and M.Tachiki.

Takahashi, S. and M.Tachiki.

Takahashi, S. and M.Tachiki.

Tanaka, M., T.Takahashi, H.Katayama-Yoshida, S.Yamazaki, M.Fujinami, Y.Okabe,
W.Mizutani, M.Ono and K.Sajimura.


Tanaka, M., S.Takebayashi, M.Hahimoto, S.Kashiwaya, F.Hirayama and
M.Koyanagi.

Tao, H.J., Farun Lu and E.L.Wolf.


Tersoff, J. and D.R.Hamann.

Thompson, W.A. and S.von Molnar.

Thouless, D.J.

Timeshenko, S.P. and J.M.Gere.

Tinkham, M.

Tozer, S.W., A.W.Kleinsasser, T.Penney, D.Kaiser and F.Holtzberg.

Tsai, J.S., I.Takeuchi, J.Fujita, T.Yoshitake, S.Miura, S.Tanaka, T.Terashima,
Y.Bando, K.Iijima and K.Yamamoto.

Tsai, J.S., I.Takeuchi, J.Fujita, T.Yoshitake, S.Miura, S.Tanaka, T.Terashima,
Y.Bando, K.Iijima and K.Yamamoto.


Tsui, D.C.

and G.S.Karabashev.

Valles, J.M. Jr. and R.C.Dynes.
1990. MRS Bull. 15. 44.


van Bentum, P.J.M., R.T.M.Smokers and H.van Kempen.


van de Leemput, L.E.C. and H.van Kempen.

van Duzer, T. and C.W.Turner.

van Kempen, H., H.F.C.Hoervers, P.J.M.van Bentum, A.J.G.Schellingerhout and D.van der Marel.


van Son, P.C., H.van Kempen and P.Wyder.

van Son, P.C., H.van Kempen and P.Wyder.


Vetter, K.J.


Volodin, A.P. and M.S.Khaikin.
1987. JETP Lett. 46. 588.


Wyosinski, K.I. and T.Dziura.


Ying, Q.Y., C.Hilbert and H.Kroger.

Yoshikawa, N. and M.Sugahara.


Zasadzinski, J.F., N,Tralshwala, P.Romano, Q.Huang, J.Chen and K.E.Gray.


Zhao, S., H.Tao, Y.Chen, Y.Yan and Q.Yang.