A Study in Inelastic Electron Tunnelling Spectroscopy

by

Philip Nicholas Shott

A thesis submitted for the degree of

Doctor of Philosophy

Department of Chemistry,
The City University, London.

September 1979

Contents

Location of Figures

Acknowledgement

I would like to extend my sincere appreciation and gratitude to my supervisor, Mr. Brian Field, for his personal supervision, enthusiasm and continuing interest throughout the course of this study.

I have been helped in this work by many colleagues in both the Department of Chemistry and the Department of Physics at The City University. Among these people, I am particularly indebted to Mr. R.E. Noakes and Mr. T.A. Harvey, for the provision of invaluable advice, assistance and discussion in all aspects of the experimental work.

I am grateful for the support of the Science Research Council during the course of this study.

Abstract

An inelastic electron tunnelling spectrometer has been constructed incorporating novel improvements in design, resulting in improvements in resolution and reduction in background noise levels relative to previous instrumentation. In particular, the use of epoxy resin circuit board in place of glass as the junction substrate, and the careful definition of oxidation parameters has resulted in a considerable improvement in junction integrity. The utilisation of a two terminal instead of a four terminal bridge arrangement, combined with electronic circuitry specifically designed to minimise noise levels has resulted in a significant enhancement of signal to noise ratio. The design of an epoxy resin junction support of minimal thermal capacity to enable direct insertion of the junction into a cryogenic storage dewar has resulted in the maximum utilisation of refrigerants.

Infusion and liquid doping methods have been investigated, and the effects of dopant concentration and spurious contamination upon junction performance have been elucidated. High resolution tunnelling spectra have been obtained for benzaldehyde, m-bromobenzaldehyde, N-methylaniline and N^3-benzoyl-3^1, 5^1-diacetyl-2^1-bromo-2^1-deoxyuridine. Comparison with infrared and Raman data has allowed the interpretation of molecular conformation of these molecules on the aluminium oxide surface. The inability of other molecules including water, benzene and chromium hexacarbonyl to produce satisfactory tunnelling spectra has been observed.

A survey of the present and potential applications of IETS as an analytical and spectroscopic technique has been compiled.

"There is no higher or lower knowledge, but one only,
flowing out of experimentation".

Leonardo da Vinci.

Chapter 1 - Introduction

1.1. The Discovery of Inelastic Electron Tunnelling.

Electron tunnelling between metals separated by a thin oxide barrier has been studied extensively over the last twenty years, mainly in the field of solid state superconductors and their applications in electronics[1-4]. It was not until 1966, however, that Jaklevic and Lambe first pointed out the potential of inelastic, as opposed to elastic, tunnelling[5,6]. They observed that the conductance of metal-metal oxide-metal (M-I-M) tunnelling junctions increased at certain characteristic voltages, and they surmised that tunnelling electrons were interacting with vibrational states of molecules included at a metal-metal oxide interface. They proposed that this interaction results in the transfer of energy from electrons to the vibrational states of molecular species contained within the barrier.

The experiments performed by Jaklevic and Lambe can be understood in terms of a simple model in which the height of the potential barrier is modified by molecular species in such a way as to couple electrons to molecular vibrations. Scalapino and Marcus showed that one type of coupling led to the usual infrared dipole selection rules[7] and their calculation was extended by Jaklevic and Lambe[6] to yield Raman-like selection rules.

1.2. The Theory of Inelastic Electron Tunnelling.

The theory of inelastic electron tunnelling

spectroscopy (IETS) must predict the position, widths and intensities of the peaks in tunnelling spectra. Fortunately the positions and widths can be predicted on very general grounds, independent of the details of the electron-molecule interaction. The peak intensities are much more difficult to predict because they do depend on the details of the interaction.

All the the theoretical results of this section will be for d^2I/dV^2 because it is easiest to calculate. Although d^2I/dV^2 can be measured with modern bridge circuits[8], d^2V/dI^2 can be measured more easily, and thus almost all experimental results have recorded the latter quantity. This is not a major disadvantage, however, since the two values are related by the expression

$$\frac{d^2V}{dI^2} = -\left(\frac{dV}{dI}\right)^3 \left(\frac{d^2I}{dV^2}\right)$$

and(dV/dI) is relatively constant over the voltage range of interest. Furthermore, dV/dI has no sharp structure in this range so that the peak shapes are essentially identical for the two second derivatives.

1.2.1.Peak position

An energy diagram of the two metal electrodes of a tunnel junction separated by a thin insulating layer is shown in Fig.1

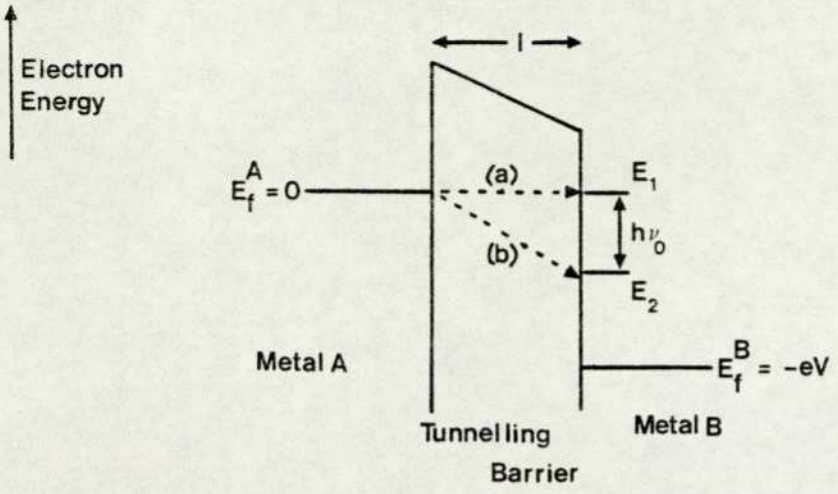


Fig.1 Schematic energy-level diagram for tunnelling between normal metals

The two Fermi levels of the metals are separated by the energy eV where V is the applied voltage. According to energy conservation, the electrons in metal A are capable of making horizontal transitions into empty states in the right-hand metal provided a voltage V is applied to the junction. Such a transition is represented by path (a), and is termed an elastic transition since it involves neither gain nor loss of energy.

In addition to this elastic conduction path, there may exist an inelastic conduction path whereby an electron can tunnel from left to right and at the same time give energy to a local impurity state. This new contribution can only occur if an empty state on the right is open for the tunnelling electron, that is if $eV \geq h\nu_o$, where $h\nu_o$ is the excitation energy of the impurity centre. As V increases, current from this inelastic process will continue to increase since more and more candidates for inelastic tunnelling become apparent. At

the temperatures and excitation energies of interest, it may be assumed that all of the impurities are in their ground state, and thus backward inelastic processes from right to left will not be significant.

The combined effect of these two tunnelling processes is shown in Fig. 2.

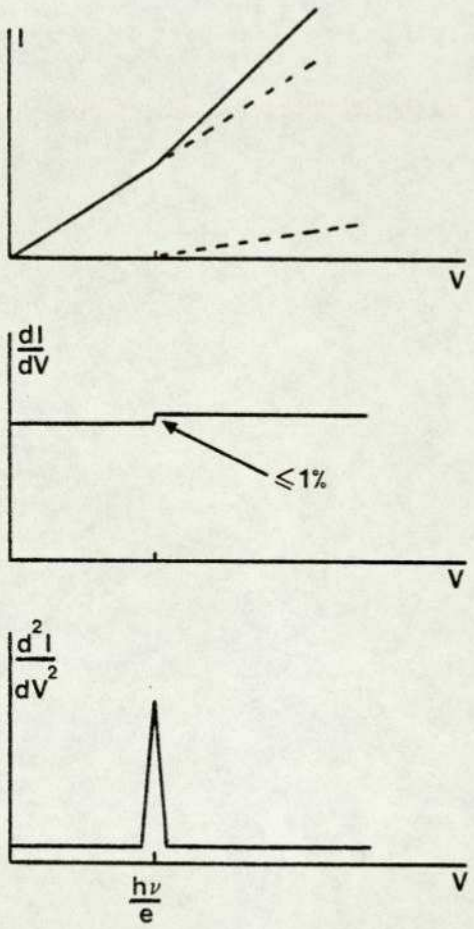


Fig.2. The current-voltage curve and its derivaties for the tunnelling process

The elastic current increases linearly with the applied voltage, V, as indicated by the initial part of the current-voltage curve. Since tunnelling electrons can excite a molecular vibration of energy $h\nu$ only if $eV>h\nu$, for low voltages there are no vacant final states for the electron to tunnel into. Thus the inelastic current has a threshold at $V=h\nu/e$, and above this value the current-voltage curve

represents the sum of the elastic and inelastic components. Thus the total current, shown by the solid line in the upper diagram, has a kink at $V=h\nu/e$.

The middle plot shows the conductance versus voltage curve in which the kink becomes a step. Since only a small fraction of the electrons (typically 1% or less) tunnel inelastically, the height of the conductance increase due to the onset of the inelastic tunnelling channel is too small to be observed. In practice, the curve is electronically differentiated to obtain the bottom curve in which the step becomes a peak, and it is this plot of d^2I/dV^2 against V which constitutes an inelastic electron tunnelling spectrum.

Thus the position of the peak corresponding to a vibrational mode of energy $h\nu$ is $V=h\nu/e$. This conclusion is based on energy conservation and is independent of the mechanism for the electron-molecule interaction. Small shifts, of order 1% or smaller, due to superconductivity and interactions with the top metal electrode will be discussed in a later section.

1.2.2. Peak width

In IETS there are only two contributions to vibrational peak width in addition to the natural width: thermal broadening and modulation broadening.

The effect of thermal broadening was first discussed by Lambe and Jaklevic[6]. Following their treatment, assume there is a vibrational mode of negligible natural width at an energy $h\nu$. The inelastic electron current due to the corresponding conductance channel can be written as:

$$I_i = C \int_{-\infty}^{\infty} dE\ f(E)(1-f(E + eV-h\nu)) \qquad (2.1)$$

Here all the details of the electron-molecule coupling are contained in the parameter C, and the remainder of the function represents the integral over the Fermi functions for the two normal metal electrodes. This integral simply expresses analytically the condition discussed qualitatively in the previous subsection : electrons must tunnel from a filled state in the one electrode into an empty state in the other electrode. It can be shown that this function may be written:

$$I_i = C \int_{-\infty}^{\infty} dE \left(\frac{1}{1+\exp(E/kT)} \right) \left(1 - \frac{1}{1+\exp\{(E+eV-h\nu)/kT\}} \right) \qquad (2.2)$$

$$= C(eV - h\nu) \frac{e^x}{e^x - 1} \qquad (2.3)$$

where $x \equiv (eV-h\nu)/kT$

Taking derivatives leads to the expression:

$$\frac{d^2I}{dV^2} = C \cdot \frac{e^2}{kT} \cdot e^x \left(\frac{(x-2)e^x + (x+2)}{(e^x-1)^3} \right) \qquad (2.4)$$

This function is plotted in Fig. 3. It is observed that the predicted linewidth at half maximum is 5.4kT/e. This observation was confirmed by the experimental work of Jennings and Merrill[9].

The effect of modulation broadening was first discussed analytically by Klein et al[10]. Following their treatment, assume that the modulation voltage is $eV_\omega \cos\omega t$ and that $f^{11}(eV)$ is the exact second derivative (with no modulation voltage broadening). The current can then be written as:

$$I = f(eV_o + eV_\omega \cos\omega t) \qquad (2.5)$$

where V_o is the slowly changing bias voltage.

To obtain d^2I/dV^2, the current $I_{2\omega}$, the value at the second harmonic frequency can be determined as:

$$I_{2\omega} = \frac{2}{\tau} \int_{\tau} f(eV_o + eV_\omega \cos\omega t) \cos 2\omega t . d\tau \qquad (2.6)$$

Two partial integrations give:

$$I_{2\omega} \propto \int_{-eV_\omega}^{eV_\omega} f^{11}\{eV_o + E\}\left\{(eV_\omega)^2 - E^2\right\}^{3/2} . dE \qquad (2.7)$$

where $E = eV_\omega \cos\omega t$. This function is also plotted in Fig.3, and would be the complete peak shape for a vibrational mode of negligible natural width in the absence of thermal broadening. It is plotted in terms of $\Delta V = 0.707\ V_\omega$, the root-mean-squared value of the modulation voltage, since this is usually the experimentally measured quantity.

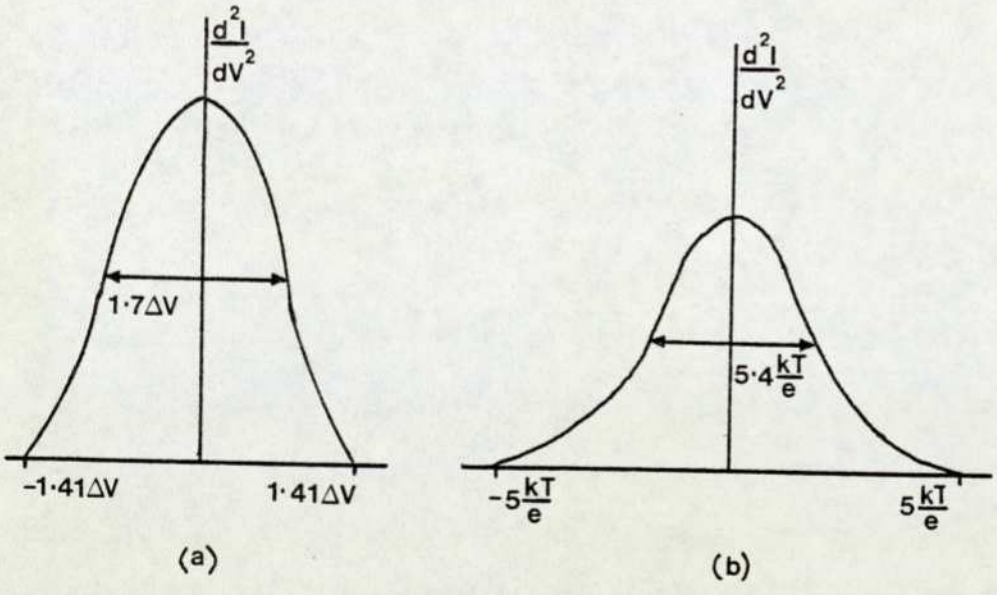


Fig.3 The line shapes due to modulation voltage broadening (a) and thermal broadening (b)

Of these two contributions to peak width, the modulation voltage broadening is perhaps the more serious. The thermal

broadening at 1K, even using normal electrodes, the root-mean-squared modulation voltage would need to be 0.27mV. Even at $\Delta V = 0.7$mV, the smallest modulation voltage reported in the literature, the second harmonic signal is so small (<50μV for even the largest peaks) that it takes overnight to obtain a complete tunnelling spectrum. Since the second harmonic signal varies roughly with the square of the modulation voltage, and since the signal to noise improvement varies with the square root of averaging time, it would take over two weeks to obtain a spectrum with the same signal to noise ratio if $\Delta V = 0.27$mV.

1.2.3. Peak intensity

The first theory for IETS intensities was developed by Scalapino and Marcus[7]. It approximates the electron-molecule interaction as a Coulomb interaction between the electron and the dipole moment of the molecule. Thus the interaction potential for an electron at z and $r_\perp$ (coordinates defined by Fig.4) with a molecule at the origin with z component of dipole moment p_z was assumed to be

$$U_{int}(z) = \frac{2ep_z Z}{(Z^2+r_\perp^2)^{3/2}} \qquad (3.1)$$

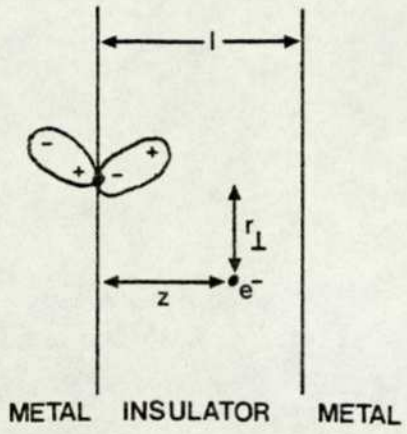


Fig. 4 An electron at position $(z,r_\perp)$ interacts with a molecular dipole moment at the origin

The induced image dipole in the metal electrode approximately cancels the perpendicular component of the molecular dipole moment $p_\perp$, and enhances the z component, p_z, by a factor of two. Thus the effective interaction is with a dipole of magnitude $2p_z$ in the z direction.

Scalapino and Marcus then treat U_{int} as a perturbation on the barrier potential U(z), which they assume to be rectangular : $U(z)= U$ if $0 \leqslant z \leqslant 1$, and U(z) is zero otherwise. If the inelastic energy loss is neglected, then the WKB approximation, an approximate procedure for solving Schrödinger's equation in one dimension, for the electron part of the tunnelling matrix is[11]:

$$|M_{1,2}| \propto \exp\left\{ - \int_0^l dz \left[\frac{2m}{h^2}(U(z)+U_{int}(z)-(E-E_\perp)) \right]^{\frac{1}{2}} \right\} \quad (3.2)$$

where E is the total energy of the electron and $E_\perp$ is the kinetic energy associated with its motion perpendicular to the z direction. In this expression U_{int} still contains the operator p_z.

To find the lowest order effect $U_{int}(z)$, they approximate $U(z)-(E-E_\perp) \propto \phi$, a constant, and expand to lowest order in the small parameter U_{int}/ϕ, obtaining

$$\left|M_{1,2}\right| \propto \left[1+\left(\frac{2m}{\phi}\right)^{1/2} \frac{ep_z}{h1}\, g\left(\frac{r_\perp}{1}\right)\right] \exp\left\{-\left(\frac{2m\phi}{h^2}\right)^{1/2} 1\right\} \quad (3.3)$$

where $g(x) = 1/x - 1/(1+x^2)^{1/2}$.

Considerable mathematical complications are avoided and the final result is conveniently obtained by making the following observations at this point:

(1) What is required as a final result is $(dI_i/dV)/(dI_e/dV)$, the ratio of the inelastic conductance for a given mode to the elastic conductance.

(2) The relevant matrix element for inelastic electron tunnelling is the second term of equation (3.3) taken between the ground state $|0\rangle$ and the first excited state $\langle 1|$ of energy $h\omega_1$ of the molecule. (At the temperatures of interest essentially all molecules are in the ground state for all vibrations).

(3) The relevant matrix element for elastic tunnelling is the first term of equation (3.3).

(4) If it is assumed that the density of states of the metal electrodes are independent of energy and that transverse electron momentum is conserved for inelastic tunnelling, just as for elastic tunnelling, then the rest of the quantities multiplying square of the matrix element are the same for the inelastic and elastic current.

Thus the ratio of the inelastic conductance a distance $r_\perp$ from the impurity to the elastic conductance is the square of the matrix element of the second term in parenthesis of equation (3.3),

$$\frac{(dI_i(r_\perp)/dV)}{(dI_e/dV)} = \frac{2m}{\phi}\left(\frac{e}{h1}\right)^2 |\langle 1|p_z|0\rangle|^2 g^2\left(\frac{r_\perp}{1}\right)\theta\left(V-\frac{h\nu_1}{e}\right) \quad (3.4)$$

where the θ function has been included ($\theta(x)=1$ for $x \geq 0$, $\theta(x)=0$ for $x<0$) to express the energy conservation condition: electrons cannot tunnel inelastically unless they have enough energy to excite the vibration i.e. $eV \geq h\nu_o$. To obtain the total inelastic conductance due to one vibrational mode of one impurity, equation (3.4) must be integrated over $r_\perp$ using:

$$2\pi\int_{r_o}^{1} dr_\perp\, r_\perp\, g^2\left(\frac{r_\perp}{1}\right) \simeq 2\pi\, 1^2 \ln\left|\frac{1}{r_o}\right| \quad (3.5)$$

If the integrated form of equation (3.4) is multiplied by N impurities per unit area and is summed over all the possible vibrations of each impurity, the final result is obtained:

$$\frac{dI_i/dV}{dI_e/dV} = N.\frac{4\pi me^2}{\phi h^2} \cdot \ln\left|\frac{1}{r_o}\right| |\Sigma|\langle m|P_z|0\rangle|^2\, \theta\left(V-\frac{h\nu_m}{e}\right) \quad (3.6)$$

where $\langle m|$ denotes the first excited state of the m^{th} vibrational mode and $h\nu_m$ denotes its energy.

Two observations can reasonably be made from equation (3.6) and its solutions:

(1) The calculated magnitude is about 1% for typical parameters with the assumption that the cutoff value of $r_\perp$, $r_o = 1/30^7$. This value of 1% corresponds with the experimentally determined value of the ratio.

(2) The size of a conductance increase is proportional to $|\langle m|P_z|0\rangle|^2$, the square of the dipole matrix element of the corresponding vibrational mode. This is the same quantity that determines the intensity of the peaks in an infrared absorption spectrum. Experimentally it is

observed that although intense peaks in infrared spectra usually correspond to intense peaks in tunnelling spectra, the proportionality is not always exact. Further peaks appear in tunnelling spectra that are completely absent from infrared spectra.

Lambe and Jaklevic pointed out that there are other mechanisms for electron-molecule interaction[6], generalising the Scalapino and Marcus theory to include the interaction via the polarisability of the molecule. The electron induces a dipole moment in the molecule and interacts with this induced dipole. If the treatment outlined previously is repeated, the ratio of inelastic to elastic conductance can be shown to be:

$$\frac{dI_i/dV}{dI_e/dV} = N.\frac{4\pi me^2}{\phi h^2}.\frac{e^2}{16l}6\int_{r_o}^{l} t^2\left(\frac{r_\perp}{l}\right) r^2 dr_\perp \Sigma|\langle m|\alpha|0\rangle|^2 \Theta\left(V-\frac{h\nu_m}{e}\right) \quad (3.7)$$

where α is the polarisability of the molecule and

$$t(x) = \frac{1}{x^2}\left[\frac{1-x^2}{(1+x^2)}2 + \frac{1}{x}\tan^{-1}\left(\frac{1}{x}\right)\right]$$

Unfortunately t(x) is strongly divergent for small x. Nevertheless, estimates for reasonable values of parameters give inelastic conductance changes of nearly the same order of magnitude as for the electron-dipole moment interaction considered previously. Jaklevic and Lambe calculate conductance changes of 0.1 to 0.5% for this interaction[12]. Thus the theory predicts that it is reasonable to expect both infrared active and Raman active vibrational modes in tunnelling spectra.

Although the theory of Scalapino and Marcus as extended

by Lambe and Jaklevic is successful in predicting magnitudes in the range found experimentally, it has several difficulties. Firstly, it assumes that the transverse electron momentum is conserved. In fact, inelastic scattering by an impurity usually changes the transverse electron momentum. Secondly, it does not include the effect of the electron energy lost in exciting a molecular vibration on the tunnelling probability of that electron. Thirdly, it depends upon a cutoff r_o which is difficult to determine. Finally, it implicitly assumes localised tunnelling electrons by first computing $I_i(r_\perp)$ and then integrating over $r_\perp$.

Thus a second approach to a theoretical explanation of inelastic tunnelling has been attempted using the transfer-Hamiltonian formalism originated by Bardeen[13]. This approach had previously been used by Cohen, Falicov and Phillips[14] to explain the presence of energy gaps in tunnel junctions with a superconducting electrode, and by Josephson who won the Nobel Prize for using it to predict the Josephson effect[15]. It has been applied to inelastic electron tunnelling by Bennett, Duke and Silverstein[16].

The transfer-Hamiltonian formalism begins with the assumption that the Hamiltonian for a Metal-Insulator-Metal junction can be written as a sum of three terms:

$$H = H_L + H_R + H_T \qquad (3.8)$$

where H_L and H_R are the Hamiltonians that describe the electrons in the left and right metal, while H_T is the transfer Hamiltonian which describes the tunnelling of electrons from one electrode to the other. The wave functions for the

electrons in each metal electrode are solutions to H_L and H_R.

Consider the process of tunnelling from an initial state in the left electrode to a final state in the right electrode. In the usual, free-electron approximation, the wavefunction for the electrons in the initial state can be written as[13]:

$$\Psi_i \propto \exp\left\{i(k_x x + k_y y)\right\} \sin(k_z z + \gamma) \quad z < 0 \text{ (in the metal)}$$

$$\Psi_i \propto \exp\left\{i(k_x x + k_y y)\right\} \exp\left\{-|K_z| z\right\} \quad 0 < z < 1 \text{ (in the oxide)} \tag{3.9}$$

where, for a rectangular barrier (U(z)=U for 0<z<1, or 0 otherwise):

$$|K_z| = \left[\frac{2m}{h^2}\left(U - \frac{h^2 k_z^2}{2m}\right)\right]^{\frac{1}{2}} \tag{3.10}$$

The electrons in the right hand metal electrode are described by a similar wavefunction, as shown in Fig.5:

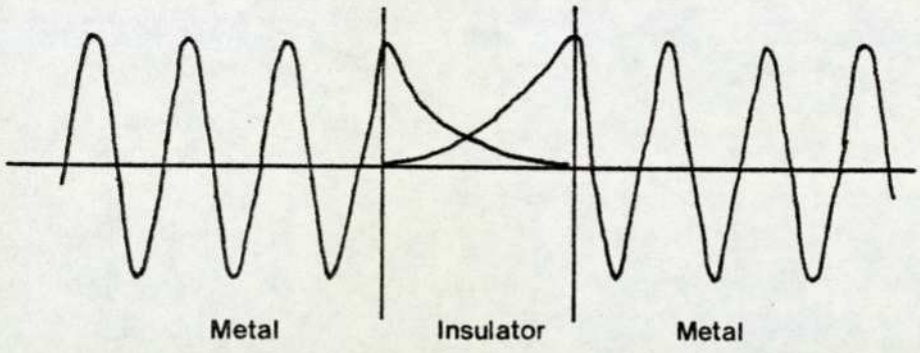


Fig. 5 Schematic illustration of the initial and final states of the transfer-Hamiltonian formalism.

The probability per unit time that H_T will induce a transfer of an electron from an initial state Ψ_i to a set of final states ψ_f is given by the golden rule[17]:

$$\Gamma_{fi} = \frac{2\pi}{h} \sum_f \left|(H_T)_{f_i}\right|^2 fi(1-f_f)\ \delta(E_f - E_i) \tag{3.11}$$

where $(H_T)_{fi} \equiv \int d^3r\, \Psi_f^* H_T \Psi_i$. The Fermi functions, $f(x) \equiv (\exp(\beta x)-1)^{-1}$, result from the requirement that the initial state be occupied and that the final state be vacant. The delta function ensures conservation of total energy. For inelastic electron tunnelling the total energy includes both the electron energy and the molecular vibration energy because the states Ψ_i and Ψ_f include both the electron wavefunction and the molecular vibration wavefunction.

If the transfer-Hamiltonian is a sum of terms, M_e describing elastic and M_i describing inelastic tunnelling, cross terms will not appear in the matrix element $|(M_e + M_i)_{fi}|^2$ because these two types of tunnelling share no common final states; the final states for inelastic electron tunnelling include an excited molecular vibration that is not present for elastic electron tunnelling. Thus it is possible to calculate both the elastic current, which is proportional to $|(M_e)_{fi}|^2$ and the inelastic current, which is proportional to $|(M_i)_{fi}|^2$.

Using this approach, Kirtley, Scalapino and Hansma have obtained agreement with experiment for the percentage conductance change for the O-H stretching mode[18]. Specifically, they obtain 0.5% while a typical experimental value is 0.4%. Whether or not the conductance changes for more complex molecular vibrations can be calculated with this accuracy remains to be seen.

They also calculate the ratio of conductance changes for tunnelling in opposite bias directions. In one bias direction (Al_{neg}) the electron tunnels and then excites the molecule, while in the other (Al_{pos}) it excites the molecule and then tunnels. Since tunnelling probability increases with electron energy, the first process is favoured.

Further, they show that the inclusion of tunnelling processes in which the electron changes transverse momentum results in coupling to infrared inactive modes as well as infrared active modes; even without the inclusion of any polarisability interaction. In fact it results in coupling even to modes that are both infrared and Raman inactive. Such modes have not yet been reported experimentally.

It is too early to tell if this theory will be successful in actually predicting the intensities in the measured tunnelling spectra of complex molecules. Perhaps a more sophisticated interaction will be required, or the transfer-Hamiltonian approach itself will prove impractical. The initial results, however, are very promising.

Chapter 2. Experimental Methods

2.1. Introduction

The essential features of a tunnel junction are displayed in Fig.6.

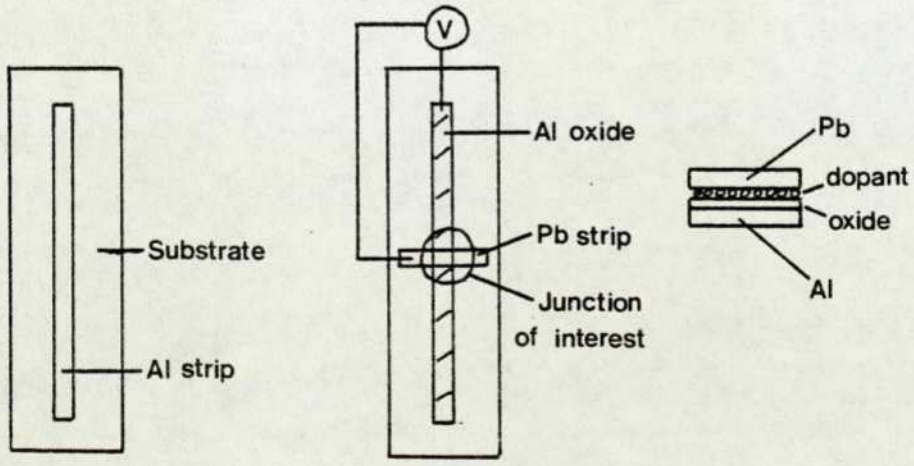


Fig. 6 The tunnel junction

The junction occurs at the intersection of two metal electrodes separated by an insulating layer, which in the simplest case is the oxide of the lower electrode. Both metal strips are deposited by evaporation of pure metal under high vacuum conditions, while the insulating layer is obtained either by exposure to a suitable vapour, or by the operation of a glow discharge technique. Finally, the doping of the oxide layer may be achieved in one of several ways described below.

2.2. Vacuum system and sample preparation

The vacuum system used during sample preparation is schematically represented in Fig.7. The basic requirement is for a compact system with a base pressure of around 10^{-8} Torr and a pumping speed capable of attaining around 10^{-6} Torr from atmospheric pressure in a period of around ninety minutes. The rapidity of the turn round time, and the fact that the system is let up to atmospheric pressure regularly (between five and seven times a day) preclude the use of sorption pumps and ion pumps, which would provide considerably less potential contamination than the oil-based system used throughout the course of this study.

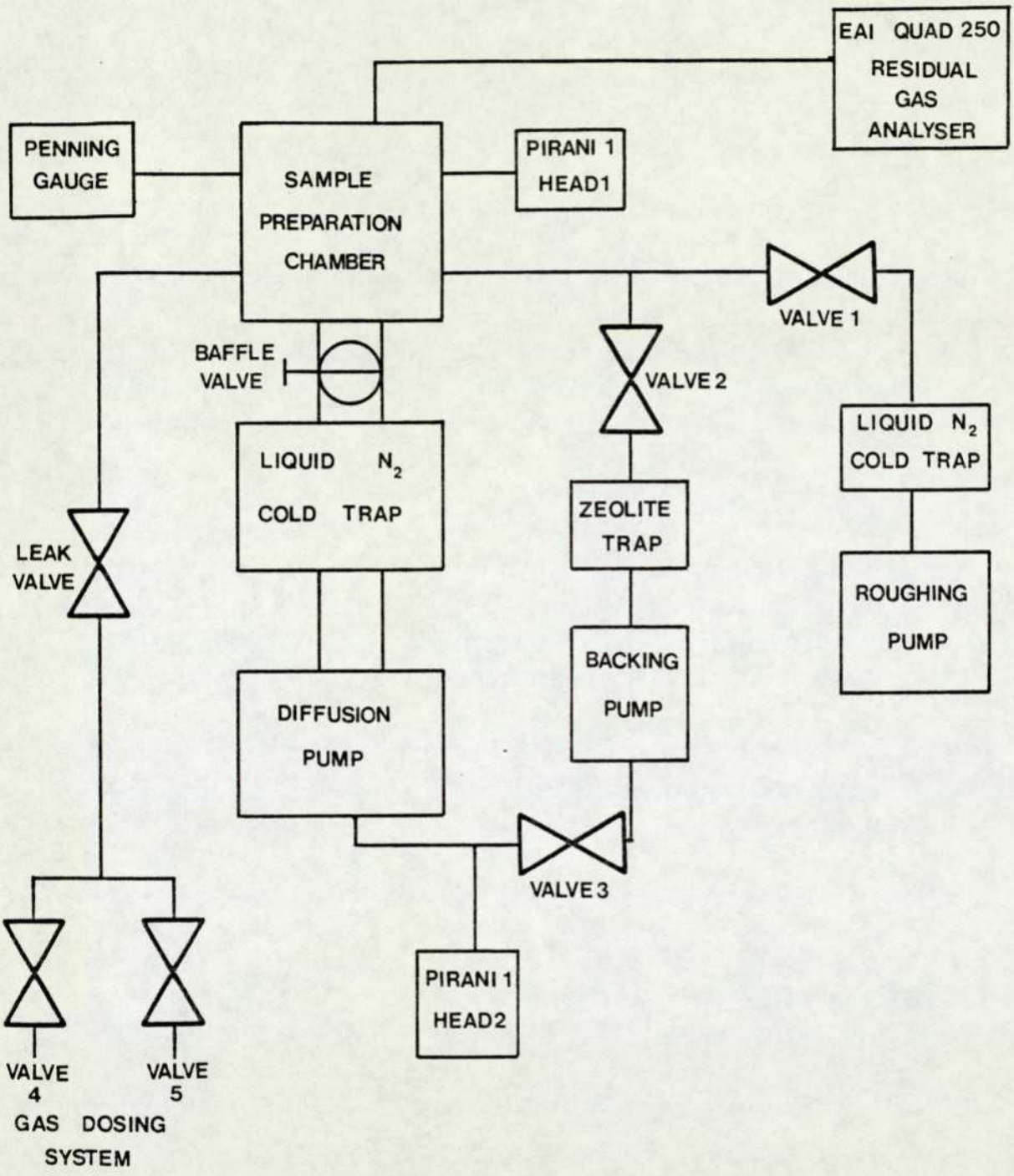


Fig. 7. Schematic representation of the vacuum system used in sample preparation.

2.2.1. Vacuum system construction

As a result of the requirement that the system be regularly vented to the atmosphere, an Edwards E04 oil vapour diffusion pump was selected as the major pumping unit. This particular model was utilised since it was already available within the department as a result of a previous study on vacuum deposition. The diffusion pump is backed by a Metrovac D.R.I.K. rotary pump, chosen for its ultimate vacuum of 5×10^{-3} torr. This rotary pump is also used as the major means of evacuating the chamber to a sufficiently low pressure, around 10^{-2} torr, to enable efficient operation of the diffusion pump. It was found early in the study, however, that the use of this rotary pump alone to rough out the deposition chamber made the system inefficient, due to saturation of the rotary pump oil by the large volumes of air which must be removed from the chamber at atmospheric pressure. In order to avoid this saturation, a second rotary pump, an Edwards 1SC 450B, having a poor ultimate pressure of around 1 torr but a very rapid pumping speed of 450 ls^{-1} is utilised to evacuate the pressure from atmospheric to around 3 torr. This reduces the amount of bulk vapour passing through the backing pump, and considerably decreases the overall pumping time for chamber evacuation.

Although this system performs admirably with respect to ultimate vacuum and pumping speed, the presence of various oils as the pumping fluids provide major contamination risks, especially since the tunnel junctions prepared in this study are exceedingly sensitive to contamination by organic

vapours. Thus the oils used, particularly in the diffusion pump, must be chosen with consideration of vapour pressures a priority, and to this end the oil used in the diffusion pump is Santovac 5, a commercially available polyphenyl ether. This precaution alone, however, was found to be insufficient, as backstreaming of oil from the rotary pumps together with passage of diffusion pump oil through the conventional water-cooled chevron-baffle assembly still resulted in the presence of organic vapours in the preparation chamber during junction fabrication.

In order to totally eliminate this contamination from the pumps, it was found necessary to resort to the use of the three traps shown in Fig. 7. By a process of elimination, it was observed that the least significant source of organic impurity arose from the backing pump via valve 2. Thus in order to remove this, an Edwards FL25 zeolite trap, consisting simply of a container tightly packed with molecular sieve material, which acts as an absorber of backstreaming rotary pump oil, preventing it reaching the chamber, was employed. Regular replacement of the molecular sieve material every three months was found to reduce backstreaming through this valve to an acceptable level.

A far more serious contamination risk was found to exist as a result of backstreaming from the roughing pump through valve 1. This arises because firstly the pump is exposed to a volume of vapour far in excess of that pumped by the backing pump, and secondly because, due to the roughing pump's low ultimate vacuum, it had to be used to the limits of its specification in order to maximise overall pumping performance.

Thus it was found necessary to employ a Vacuum Generator's NCT4 liquid nitrogen trap in order to remove all organic vapours from this source.

The final, and most significant, source of organic vapour is the oil used as a pumping agent in the diffusion pump. A diffusion pump operates by boiling the pump oil and spraying the vapour out of a series of cones to "trap" molecules and transport them to the base of the pump, whence they are removed by the backing pump, as shown in Fig.8.

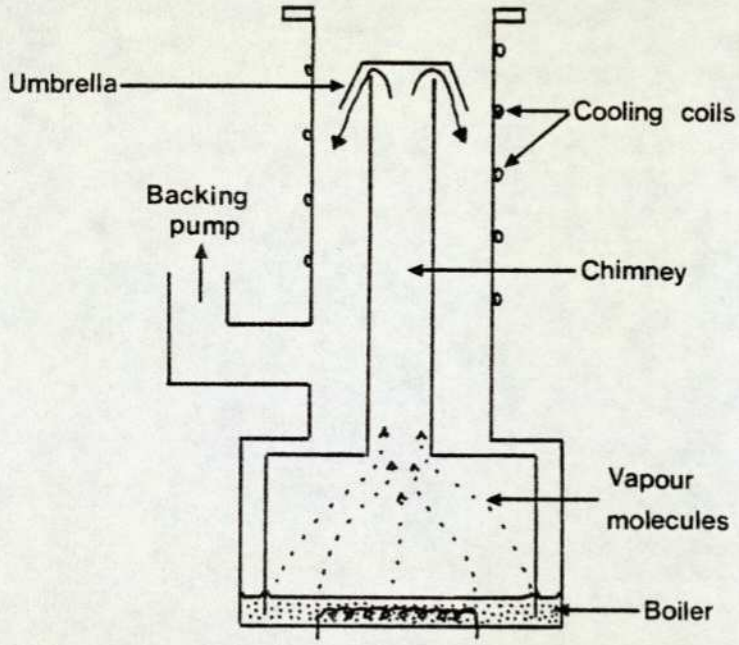


Fig. 8. The essential features of diffusion pump operation

The top of the diffusion pump is generally protected by a water-cooled chevron-baffle assembly, which condenses a large proportion of the escaping vapour, but by no means all. This residue of escaping vapour necessitated the use in this study of a Vacuum Generator's NCT4 liquid nitrogen cold trap. The use of this cold trap provides an additional bonus in that it acts as a sorption pump, further reducing the pumping time, and is particularly effective in the pumping of water molecules

produced during the glow-discharge cleanup described in section (2.2.5.)

The intial effect of fitting this cold trap, in conjunction with the other traps described above, was the total disappearance of organic contamination derived from pump oils combined with a significant improvement in pumping speed. This situation maintained over a period of a month, after which the vacuum system's performance declined rapidly. This prompted a prolonged search for leaks, as described in a later section, and the ultimate discovery of a leak at one of the fabrication welds within the cold trap. Due to the fact that the leak was caused by contraction of the cold trap away from a poor weld, it was only apparent when the trap was at liquid nitrogen temperatures, and did not appear when the trap was at room temperature. It was only finally assigned to the cold trap after the elimination of all other possibilities, and by the application of acetone to the interior of the trap followed by the cooling of the interior with a stream of liquid nitrogen. The final detection of the leak necessitated the purchase of a further NCT4 coldtrap which has fortunately proved effective throughout the remainder of the study.

A further source of contamination within the system, other than general atmospheric leaks which will be discussed in section (2.2.5.) is outgassing of adsorbed vapours from internal fittings and of materials with high vapour pressures. This is a prime consideration in the choice of sealing agents in the numerous joints incorporated within the system. The joints in the central pumping stack, namely between the base-

plate and the baffle, the baffle and the coldtrap, and the coldtrap and the diffusion pump are all sealed using compressed indium wire, while flanged connections from the leak valve and mass spectrometer to the ports on the baffle assembly are sealed with copper gaskets compressed between two knife edges. These are the most effective sealing methods in high vacuum systems, combining a negligible leak rate with a negligible outgassing rate, providing that both the indium and the copper gasket are thoroughly cleaned in a solvent such a trichloroethane in order to remove all traces of grease before fitting. It is not always possible, however, to utilise copper gaskets or indium wire as a sealant, especially when joining together the various pieces of piping required for the lines to the rotary pumps, for example. In such cases the joints are sealed using a compressed 'O'-ring.

The 'O'-rings utilised in these joints on the high vacuum side of the diffusion pumps, for example between the Penning head and the chamber port, and between Pirani head 1 and the chamber port, are composed of Viton. Viton is a compressible synthetic rubber, designed for use in high vacuum systems to provide a leak-free seal and having a sufficiently low vapour pressure to reduce any outgassing to a level such that ultimate pressures of 10^{-8} torr may be attained. The diaphragms which provide the operational base of valves 1 and 2 are also composed of Viton, as are the gaskets at the top and base of the pyrex cylinder which provides the walls of the sample preparation chamber.

The various gases required in the chamber during junction fabrication are introduced via valves 4 and 5 and the leak valve

of Fig.7. Valves 4 and 5 are conventional glassware high vacuum taps, consisting of a glass plug with a polytetrafluoro-ethylene (PTFE) sealing ring which screws in and out of a glass barrel. PTFE is chosen as a tough compressible inert material, thus providing a good seal, having a low vapour pressure and thus a low outgassing rate. The leak valve used is an Edwards LB2B which seals via a Viton Dowty 'O'-ring as shown in Fig.9.

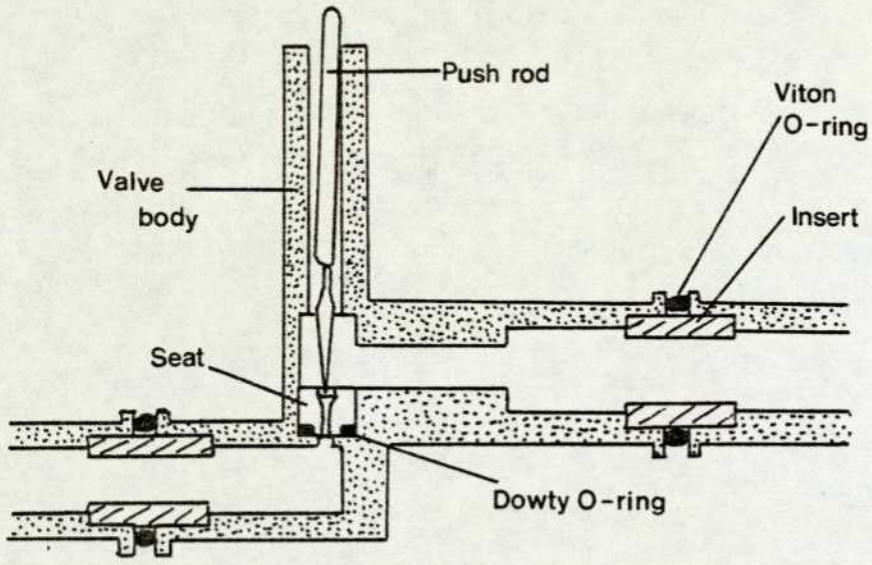


Fig.9. The sealing mechanism
of the Edwards' LB2B leak valve

The final section of joints are those in the backing line from the diffusion pump to the rotary pump, and from valves 1 and 2 to the rotary pump. This line is composed of lengths of 1" diameter copper piping, constructed in a suitable geometry and silver-soldered together. All flanged joints in this part of the system are coupled in the same way as in Fig.9, except that the 'O'-rings are composed of neoprene instead of viton. Neoprene is a compressible rubber having a

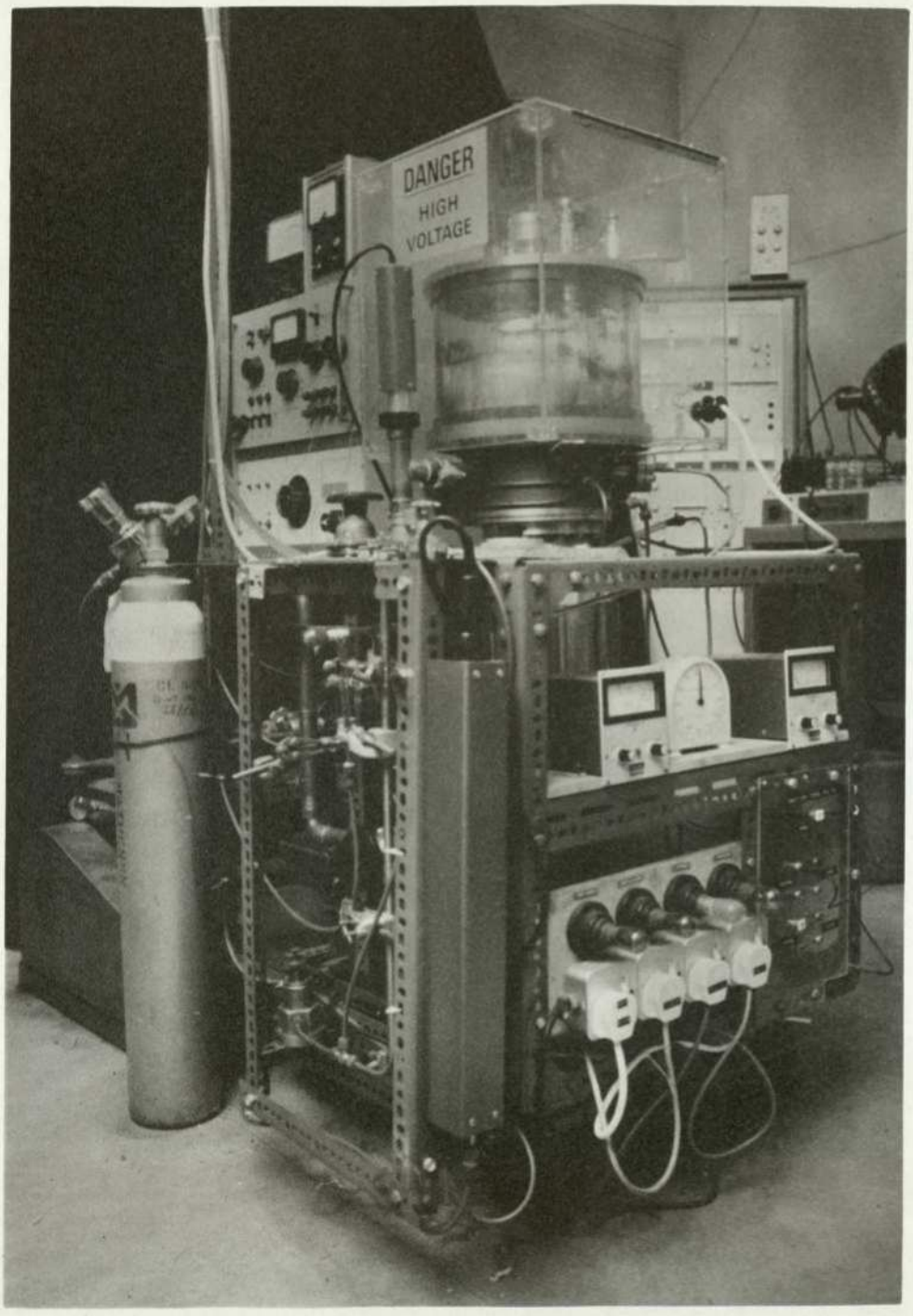


Plate 1. The vacuum preparation chamber and pumping system

much higher vapour pressure than viton, but is suitable for use in low-vacuum apparatus where pressures of above 10^{-4} torr are required. Similarly, the diaphragm in valve 3 is composed of neoprene.

2.2.2. Principles of pressure measurement

In the majority of vacuum systems a measurement of the pressure at various points is required, and a wide variety of gauges is available for this purpose. The pressure ranges covered by vacuum gauges which can be manufactured in the average laboratory or which are commercially available are displayed in Fig.10.

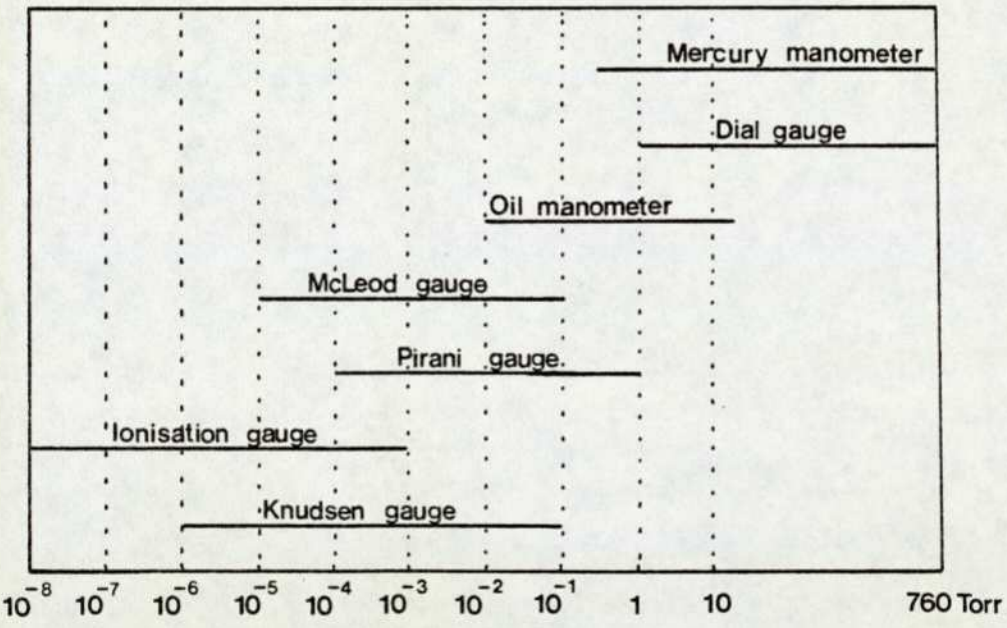


Fig.10. Pressure ranges of vacuum gauges

The dial gauge, diaphragm manometer, liquid manometer, McLeod gauge and Knudsen gauge measure pressure directly as the force exerted by the gas on a solid surface or liquid column. They are absolute gauges whose sensitivities are the same for all gases. All the other gauges mentioned above measure pressure

indirectly in terms of some property of the gas which is pressure-dependent. Their sensitivities are different for different gases and they must be calibrated against an absolute gauge.

It is from this latter group of gauges that those used for pressure measurement in this study are drawn. At the pressures of interest in the evaporation chamber, which range from 10^{-8} Torr to around 10^{-1} Torr, there is no single gauge which will operate, and thus a combination of the Penning and Pirani gauges is utilised.

2.2.2.1 Pirani Gauge

The Pirani gauge[19] is one of the most widely-used vacuum instruments in the range of medium high vacua. Its action depends on the variation in thermal conductivity of a gas with pressure, since at low pressures the thermal conductivity of a gas is a linear function of the pressure.

If a metal filament is heated by an electric current the wire reaches an equilibrium temperature at which the heat generated by the electric current is exactly balanced by the heat losses due to conduction, convection and radiation. Radiation losses are small unless the temperature is very high; at low pressures when the mean free path is comparable with the size of the vessel the distinction between convection and conduction disappears. Gas molecules on striking the heated filament gain kinetic energy which they transfer by collision either to other gas molecules or to the walls of the vessel. At high pressures the number of molecules striking the filament in unit time is large and the heat losses are also large.

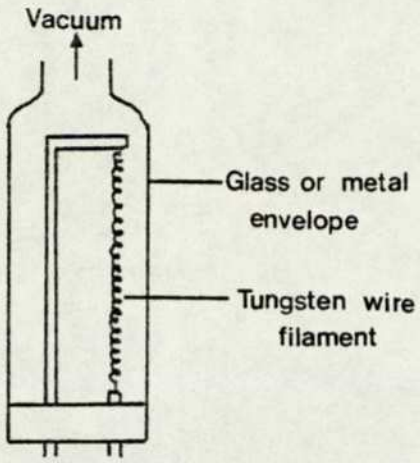


Fig.11. The Pirani Gauge Head.

At high pressure, when heat transfer is all by intermolecular collisions, the heat losses are independent of pressure. As the pressure falls a mixture of intermolecular and molecule-wall collisions is produced and the loss of heat is a complex function of the pressure. As the pressure continues to fall the molecule-wall collision process becomes dominant and the loss of heat becomes directly proportional to the pressure. In general then, the heat losses become smaller as the pressure decreases and in consequence the temperature of the filament increases as the pressure falls. A rise in temperature also produces a rise in electrical resistance and a new equilibrium condition is established. Hence a measurement of the electrical resistance of the filament provides a means of determining the pressure.

The pressure limits to the range of the Pirani gauge are set by the heat loss rates due to conduction and radiation. The heat loss due to radiation is small but independent of pressure. At the high pressure end (~500 Torr) the heat loss due to conduction is also constant with pressure, and thus the

reading of the gauge becomes independent of pressure. At about 10^{-3} Torr the heat loss due to conduction becomes similar in magnitude to the heat loss due to radiation. Consequently, changes in pressure cause only minor alterations in heat losses and hence in filament resistance; the gauge becomes too insensitive for practical use.

The usual method for measuring resistance is by some form of Wheatstone bridge. Fig.12 shows the Wheatstone bridge together with the power unit circuit:

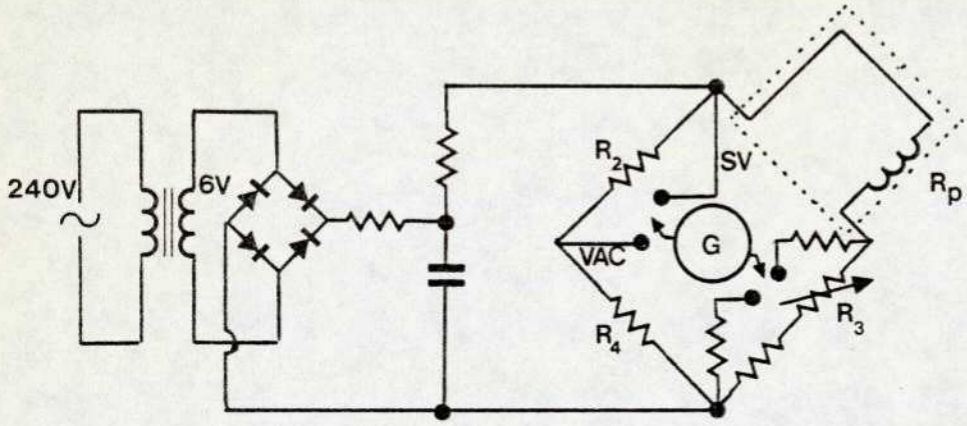


Fig.12. Pirani gauge control unit.

R_p is the resistance of the Pirani gauge filament, R_2 and R_4 are fixed resistances and R_3 is an adjustable resistance. With the milliammeter connected in the VAC position the balance condition is

$$\frac{R_p}{R_2} = \frac{R_3}{R_4} \quad \text{or} \quad R_p = \frac{R_2 R_3}{R_4}$$

One method of measuring the pressure is to balance the bridge by varying R_3 and calculate the resistance of R_p, a previous calibration enabling this resistance to be converted into pressure. It is more usual, however, to keep R_2 and R_4 constant and preset R_3, and to measure the out-of-balance current through G, which can be a milliammeter whose scale is calibrated in pressure units. In this case it is essential to keep the voltage across the bridge constant. The bridge may be balanced initially at atmospheric pressure; then an increase in the resistance of R_p causes an increase in out-of-balance current so that the lowest pressures correspond to full-scale readings of the milliammeter. The scale at the high pressure end is very compressed and becomes progressively more open as the pressure decreases. The normal form of the gauge is useful over a range of 1 Torr to 5×10^{-3} Torr.

Commercial forms of the Pirani gauge are generally supplied with a control unit, a typical example of which is also illustrated in Fig. 12. The bridge voltage is supplied by a direct current derived from rectified a.c., and can be varied by means of a "set voltage" rheostat. By switching to the position SV the milliammeter can be used as a voltmeter and the voltage standardised against a "set voltage" mark on the scale. In some instruments the bridge voltage varies with the pressure in the gauge head and standardisation must be carried out before each pressure observation.

Each control unit can be used with different heads of nominally the same type; in order to match exactly the resistances

of the gauge heads to the bridge circuit (and hence to the scale calibration) it is necessary to have a supplementary rheostat in the Wheatstone bridge network. In Fig.12 this is the variable part of R_3 and is adjusted by the manufacturer during calibration and then sealed in a small unit which is attached to the gauge head.

There are a number of modifications by which a Pirani gauge may be made more accurate and also more sensitive. The most important of these concerns the variations in the resistance of the gauge head caused by ambient temperature fluctuations. In order to overcome this a dummy head is provided which has a resistance element of the same size and material as the active gauge head but is sealed in a glass envelope at a pressure much lower than 10^{-3} Torr equal to that at which the bridge was balanced. The dummy head becomes the resistance R_2 and ambient temperature changes then produce equal resistance changes in R_p and R_2. Thus this source of error is eliminated.

Another form of the gauge is the four-head model. In this, two arms of the Wheatstone bridge network (R_p and R_4) are made variable, and the other pair (R_2 and R_3) are kept fixed. R_2 and R_3 are dummy heads as described above. R_p and R_4 are made into idential gauge filaments and both are placed in the vacuum system so that pressure changes affect both of them equally, and hence approximately double the out-of-balance current for a given pressure. This means of increasing the sensitivity of the bridge is particularly useful in the pressure range 5×10^{-3} to 1×10^{-3} Torr where the sensitivity of the single-head type falls off.

2.2.2.2. The Penning Gauge

The most commonly used gauge for measuring pressures below 10^{-3} Torr is the ionisation gauge, which exists in a number of forms all of which operate on the same basic principle. The residual gas in the gauge head is subjected to ionising radiation and some of the gas molecules become ionised. The positive ions are attracted towards a negatively charged electrode placed nearby, and a very small electric current flows in an external circuit in order to maintain the charge on the electrode. The magnitude of the positive ion current depends on a number of factors including the intensity of the ionising radiation, the nature of the gas and the number of gas molecules per unit volume, i.e. the pressure. For a given set of conditions the ion current is directly proportional to the pressure and the measurement of pressure becomes the measurement of this current. The ion current is only of the order of microamperes even at pressures as high as 10^{-3} Torr, and at lower pressures amplification is required to make measurements possible. In theory at least, the gauge should be useful down to zero pressure, although in practice modifications are necessary for use below 10^{-6} Torr.

The cold cathode ionisation gauge as developed by Penning[20] is illustrated in Fig.13. Two cathodes are used in the form of parallel plates and midway between them is placed the anode consisting of a loop of metal wire whose plane is parallel to that of the cathodes. A potential difference of about 2KV is maintained between the anode and the cathodes. In addition, a magnetic field of the order of 400gauss is applied at right-angles to the plane of the electrodes by a permanent magnet.

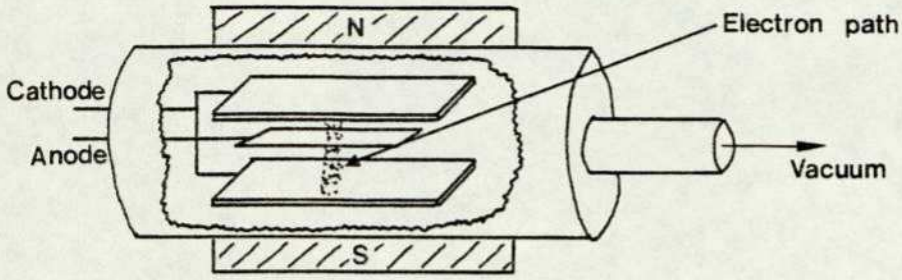


Fig. 13. The Penning gauge head

Consider an electron at a position near one of the cathodes. It is accelerated towards the anode by the electric field, but the action of the magnetic field causes its path to be in the form of a tight helix about the direction of the magnetic field. The electron generally passes through the plane of the anode loop until its path is reversed by the electric field due to the second cathode. The electron continues to oscillate in this manner about the plane of the anode loop, thus following a very long path so that the chance of striking a gas molecule is high even at low pressures. The kinetic energy of the electron gained from the electric field is generally large enough to cause ionisation should such a collision occur. The secondary electrons produced by ionisation themselves perform similar oscillations and the rate of ionisation increases rapidly. Eventually, the electrons are captured by the anode and equilibrium is reached when the number of electrons produced per second by ionisation equals the number of those captured per second by the anode.

The positive ions created in this process are captured by the cathode and hence cause a current to flow in the external circuit. The drift energy of the positive ions is not large enough to contribute significantly to the ionisation process. The ion current is measured by a milliammeter in the cathode circuit since the current is sufficiently large to be measured directly without amplification. This is the major advantage of the cold-cathode type of ion gauge over the hot-cathode gauge, which use a heated filament to produce the electrons, and require substantial amplification of the cathode current.

The number of ions produced per unit time is proportioned to the molecular density and hence to the pressure in the gauge; thus there should be a linear relationship between current and pressure although in practice the current increases at a smaller rate than the pressure. The calibration of the gauge, however, depends on the type of gas used in the system because of the different ionisation energies of different gas molecules.

The electrical circuit associated with the Penning gauge is very simple and is shown in Fig.14:

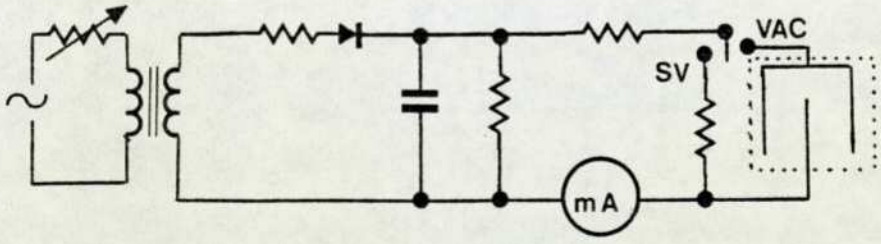


Fig.14. Simplified Penning gauge control unit

The gauge is operated from a control unit consisting of a rectified but unstabilised a.c. power supply. The voltage applied across the gauge head may be standardised by

using the milliammeter as a voltmeter and bringing the pointer to a fixed mark on the scale.

The range of the Penning gauge is from about 10^{-2} to 10^{-6} Torr, the upper limit being set by the onset of a glow discharge and the lower limit by the smallness of the ion current generated. One disadvantage of the Penning gauge is the difficulty of initiating the build-up of ionisation, especially at low pressures. In order to start the ionisation process it is necessary to produce a few electrons near one of the cathodes. At higher pressures, of the order of 10^{-2} Torr, general background radiation in the atmosphere, such as cosmic radiation, which passes through the gauge head usually provides sufficient ionisation to provide these electrons. However at lower pressures the probability of their production by this means becomes smaller, and eventually is insufficient to start the process. Several methods may be used to overcome this difficulty:

(a) The Gauge may be switched on at 10^{-2} Torr and allowed to remain on throughout the pumpdown.

(b) For low pressure starting a radioactive gamma or beta source may be brought near the head, the gamma or beta rays producing sufficient initial ionisation.

(c) A small filament may be built into the head through which an electric current is passed for a second or so, the thermionically produced electrons initiating the ionisation.

2.2.2.3. Interpretation of gauge readings

Most commercial non-absolute gauges have scales which are calibrated for clean dry air, and for many purposes this

calibration is sufficiently accurate. However, it must be borne in mind that as the pressure in a system falls the gas composition changes. For example, in an unbaked system the predominant gas at pressures down to 10^{-3} Torr will be air, but below this pressure the proportions of water vapour and hydrogen, due to desorption from the walls, will increase rapidly. In a baked ultra-high vacuum system the predominant gas is most likely to be hydrogen. The composition of the residual gas is generally unknown and thus the gauge reading cannot be corrected.

The reading of the gauge also depends on its position in the system. For example, if a cold trap is employed a gauge will indicate a lower pressure on the pump side than on the vessel side. This is due to the removal of condensable vapours by the trap and also because of the pressure drop across the trap impedance. In addition, a gauge "looking" at a cold trap will read a lower pressure than one, at the same location, but "looking" into the vessel. This is due to (i) the condensable components being frozen at the trap and hence not being reflected into the gauge and (ii) the non-condensable gases reflected from the cold trap having velocities corresponding to the cold trap temperature, and hence causing thermal transpiration from the gauge head.

2.2.3. The sample preparation chamber

The essential composition of the sample preparation chamber is schematically represented in Fig. 15.

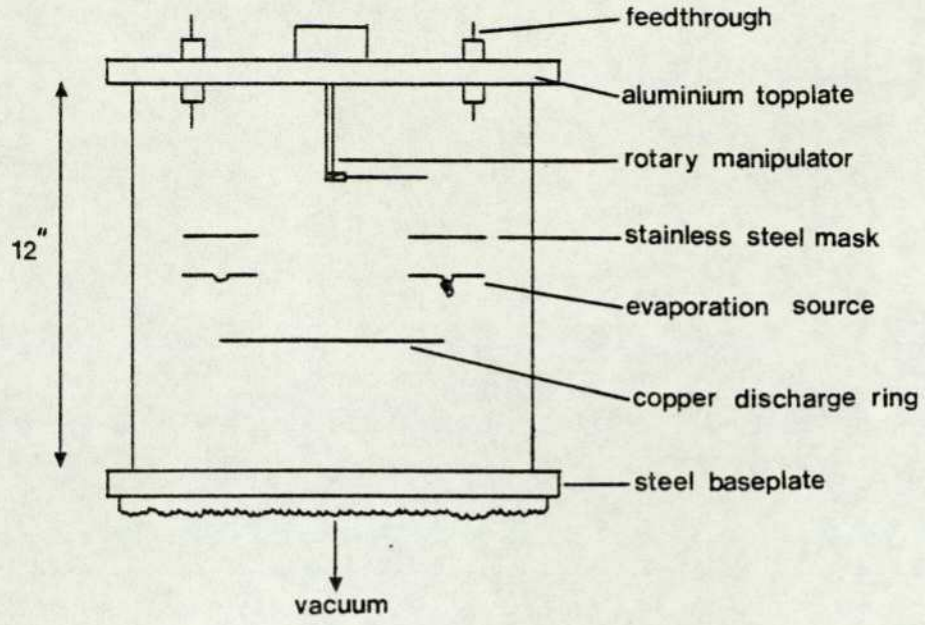


Fig. 15. Schematic representation of the sample preparation chamber

The chamber is composed of a pyrex glass cylinder, cut to a height of twelve inches, with ground glass edges which are sealed to the top and base plates by means of viton L-section gaskets. The base plate is connected to the baffle assembly of the pumping system, and all internal fittings are suspended from the aluminium topplate to facilitate removal. The internal masking and evaporation assemblies are rigidly suspended, and this necessitates the transfer of the junction substrate from one mask to another during junction preparation.

This requires the presence of some form of rotary feedthrough. The model originally employed was an Edwards rotary shaft vacuum seal, and this was utilised in the early part of the study. It enabled substrate motion vertically as well as horizontally, and thus had the advantage of enabling the masks and evaporation equipment to be supported by a central strut, since the specimen could be moved from one mask to

another by lowering under the mask assembly. However, this type of feedthrough achieves its vacuum seal by the use of double Wilson type vacuum seals, which require the presence around the central shaft of Apiezon grease. As the sensitivity of junction manufacture to the presence of organic contamination was realised, it was found that small amounts of this grease sealant were being transferred to the chamber by the vertical movement of the shaft, creating impurity problems in the completed junctions.

Thus the grease-sealed rotary feedthrough was replaced by a Vacuum Generators RD3 rotary feedthrough, which seals by a combination of oil-free bearings and bellows, and removes any possibility of organic contamination from this source. Unfortunately, however, this type of feedthrough is capable only of rotational motion, and thus the central support for the masks and evaporation assembly had to be removed in order to provide an unhindered arc of rotation.

Since pumping speed of any vacuum chamber is severely restricted by outgassing, and again because of the sensitivity of the technique to low levels of organic contamination, it is imperative that fingerprints are excluded from the chamber. To this end all input to the chamber, such as aluminium, lead and substrate for each preparation, is loaded using tweezers, and all adjustment to the internal fittings, such as the replacement of the tungsten basket used in aluminium evaporation, is accomplished using nylon gloves. In this way no junction contamination occurs from this source and the pumping speed is sufficient to allow the production time of a junction to be reduced to around 90 minutes.

The metals used in the production of tunnel junctions, most commonly aluminium and lead, are required as strips approximately 1 mm wide and are thus evaporated through masks. These masks must be rigid, must not outgas and must be immune to attack, especially important since vaporised aluminium is extremely reactive. To this end the masks were machined from 3mm thick stainless steel. Slots were machined, in the required geometry, which were 0.5mm wide, and, allowing for the separation between the substrate and mask, these were found to produce metal strips of the required width.

During the production of a junction it is necessary to produce a glow discharge in the chamber, both as an aid to cleanup before preparation and as an integral part of the oxidation stage. Several different discharge electrodes have been utilised in the course of this study. The original design was a circular loop of multistrand aluminium wire. Since the current sustained during the operation of the discharge is very much a function of the system geometry, this ring had to be approximately eight inches in diameter, and was eventually found to be too difficult to support. This was replaced by a titanium sheet cylinder, which was far more structurally robust, but would sustain a maximum current of only 10mA. The final design was a six inch circle of copper wire, thick enough to be structurally robust, and yet of a geometry which will sustain a current of up to 200mA.

All other fittings within the chamber are composed of materials which will not outgas at any appreciable rate at pressures down to 10^{-8} Torr. The support scaffolding is stainless steel studding tapped into the aluminium topplate;

the clamping for the evaporation sources was constructed from stainless steel plate; the feedthrough insulation is PTFE; the masking to prevent sideways evaporation of the metals obscuring vision through the glass is aluminium sheet. All thevacuum seals through the topplate are either by means of viton O-rings or compressed indium wire.

2.2.4. Principles of glow discharge

The process of forming oxide films on metals such as aluminium, tantalum, niobium, etc. by means of anodisation in an electrolytic cell employing an aqueous solution is well established in science and industry, the use of such films extending from protective coatings to capacitor dielectrics. In 1963, Miles and Smith[21] performed experiments in which the anodisation process employs either a gaseous or a solid electrolyte rather than the usual aqueous solution. Oxide films formed by these methods are similar, and in certain respects superior to those formed in the usual manner.

The techniques they describe possess certain advantages over "wet" anodisation, since they may be applied to the growing field of microminiaturisation of electronic circuitry through the application of thin film technology. In particular the process of gaseous anodisation provides a simple, convenient, and reliable method of forming tunnelling barriers and thin film dielectrics for use in both active and passive electronic circuit elements. These barriers may be formed over a wide range of thickness and circuits may be fabricated without removing the substrate from the vacuum system.

According to the theory of oxidation presented by Mott[22], two conditions are required in order for the growth of an oxide film to proceed on a metal such as aluminium. One is the availability of oxygen and the second is the existance of an electric field of proper magnitude and sign at the metal-oxide boundary. It is this field which is responsible for dissolving positive metal ions in the oxide already formed, thus allowing them to proceed to the oxide-oxygen interface by a combination of drift in the field and diffusion. Once they reach this interface they react chemically with the oxygen and hence growth of the oxide takes place.

The plasma anodisation described by Miles and Smith is carried out inside an ordinary oil diffusion pumped high vacuum system capable of achieving pressures in the 10^{-6} Torr range during normal operation. The system is shown schematically in Fig.16 :

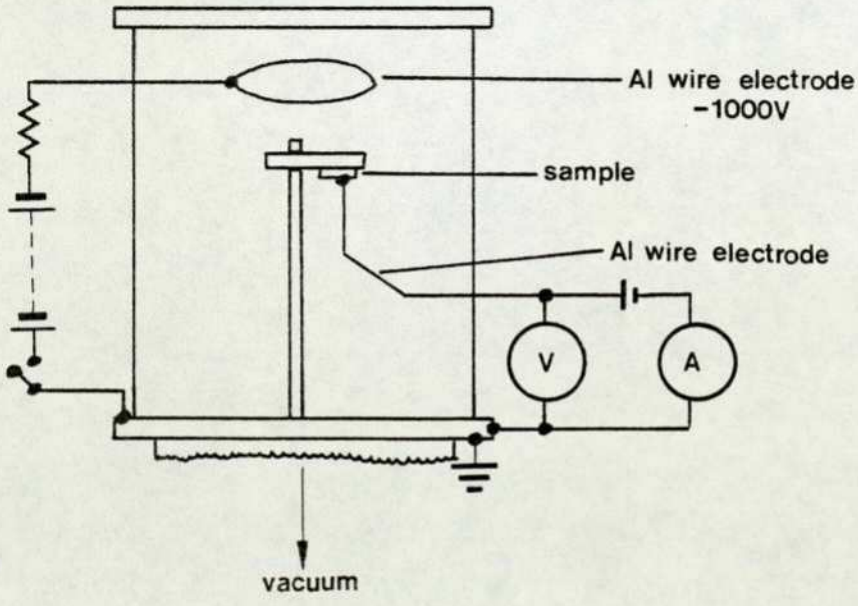


Fig. 16. Schematic representation of the gaseous anodising system.

When the system has pumped down into the 10^{-6} Torr range, a film of aluminium approximately 1000Å thick is formed on an ordinary glass microscope slide by evaporating a charge of 99.999% pure material through a suitable mask. After the evaporation is completed, the diffusion pump valve is closed and oxygen gas is admitted through a calcium hydride drying tube to bring the system to a pressure of approximately 5×10^{-2} Torr. A glow discharge is then initiated in the system by applying several hundred to one thousand volts negative potential to an aluminium wire ring mounted in the bell jar. The metal base plate serves as the ground electrode.

Next, one of two courses is followed. If a relatively thin oxide is desired, the glow is maintained for a few seconds to some tens of minutes; the discharge is then shut off and the diffusion pump valve opened to reduce the pressure to 10^{-6} Torr again. A second metal electrode is then evaporated across the first to form a diode structure in the form of a cross. It is possible in this way to produce barriers of up to 50Å in thickness, and it is this method which has been adopted in the course of the present study in IETS.

If on the other hand, a somewhat thicker oxide film is required, a direct electrical connection is made to the aluminium film in the vacuum system by means of a probe formed of pure aluminium wire. This probe passes outside the system through a vacuum seal and a small positive d.c. potential is applied to it with respect to the system ground. A current is observed to flow in this circuit once contact with the film is established, and this is seen to decay with time in a manner similar to that which prevails in an ordinary aqueous

anodising bath. When the current has decayed, connection is broken, the discharge shut off, and the system pumped down for the evaporation of the second metal electrode as before.

Oxide films formed by the first of these two methods have been found to be particularly useful in tunnelling studies as they are extremely reproducible, allow a very fine control of thickness to be achieved, are quite robust when formed in a dry atmosphere and obey predictions based on simple electron tunnelling theory. Aluminium oxide films formed in this manner have been observed to carry current densities of 1000 amp cm^{-2} at 4 volts before breaking down[21].

2.2.5. Leaks and leak detection

The limitation on the ultimate pressure attainable in a vacuum system is set by the continuous appearance of gas in the system. The general term leak is used to describe this phenomenon and may be considered under two headings, namely true and virtual leaks. A true leak is due to gas entering the system through a hole, whereas a virtual leak is caused by the outgassing of the inner surfaces of the system. Both these processes produce the same result, i.e. a rise in pressure in a system when it is isolated from the pumps, or a finite limit to the pressure attainable in a continuously pumped system. Virtual leaks can only be reduced by degassing methods, while true leak rates may by finding and sealing the hole.

In the system used for the present study, and in particular since the tunnel junctions prepared in the vacuum

system are extremely sensitive to the presence of impurity, it is essential that the leak rate be minimised. Unfortunately, since the chamber itself is a pyrex cylinder, it is not possible to reduce virtual leaks by baking-out the system as is the practice in conventional all-metal high-vacuum systems. To minimise virtual leaks, therefore, the following precautions are taken:

(i) All internal fittings, including sealing gaskets and O-rings, are composed of materials selected to minimise their outgassing.

(ii) All material is loaded into the chamber using forceps and gloves in order to prevent any contamination from fingerprints.

(iii) The system is vented to atmospheric pressure under dry nitrogen. This minimises adsorption on to the internal surfaces, and thus increases pumping speed.

The incorporation of these precautions in the operation of the vacuum system reduces virtual leaks to an acceptable level. The remainder of this section is devoted to a discussion of true leak detection.

Leak detection methods may be divided into two broad groups: those in which the system under test is filled with gas to slightly over atmospheric pressure and the outflow is detected, and those in which a device within a low pressure system detects the inflow of an externally applied probe gas. Due to the volume and complexity of the present system, it is only the latter group which is applicable.

In these low pressure methods the detector is essentially a pressure gauge mounted within a continuously pumped system, a leak being indicated by an apparent change in pressure when the leaking air is replaced by a suitable probe gas. For sensitive leak detection a number of essential factors must be taken into account:

(a) The pressure in the region of the detector head must be stable so that small pressure changes due to probe gas can be detected.

(b) The rate of flow of gas past the detector must be sufficiently slow for the detector to be able to respond to a small proportion of probe gas, but must be sufficiently fast for the time lag between probe application and detector response to be fairly short.

(c) The detector head must be positioned so that gas from all possible leaks flows past it.

A summary of the properties of leak detectors utilised in the present study is given in Table 1:

Detector head	Useful pressure range (Torr)	Detectable leak rates (Torr 1 sec^{-1})	Probe
Pirani gauge	$1 - 10^{-3}$	10^{-5}	Acetone
		10^{-4}	Hydrogen
Penning gauge	$10^{-2} - 10^{-6}$	10^{-8}	Acetone
		10^{-7}	Hydrogen
Mass spectrometer	below 10^{-3}	10^{-11}	Hydrogen
		10^{-12}	Helium

Table 1. Summary of the properties of low pressure leak detectors

The larger leaks in the system were detected using acetone and the Pirani and Penning gauges. Liquids, however, have the disadvantage that they can completely or partially block a leak, and this can lead to large response times. However, as an initial check on the integrity of joints and seals this method proves acceptable. Smaller leaks are then detected by the use of low molecular weight gases, eventually with the aid of a quadrupole mass spectrometer. For use as a leak detector, the mass spectrometer is set up so that only ions of the probe gas are collected.

The mass spectrometer was initially fitted to the system in order to try to accurately detect the source of organic vapour contaminants presumed to be leaking from the atmosphere. Having been successful in this respect, a second useful facet became obvious in that the mass spectrometer can operate as a residual gas analyser. This is particularly useful in assessing the efficiency of methods used for removing organic contamination from the system.

Early workers in this field such as Magno and Adler[23] advocated cleanup of residual impurities using a high current argon discharge. The mass spectra of the residual gases at low mass numbers before and after the operation of this discharge are represented in Fig.17. For the sake of simplicity only the major peaks are considered, and peaks due to isotope effects are not represented.

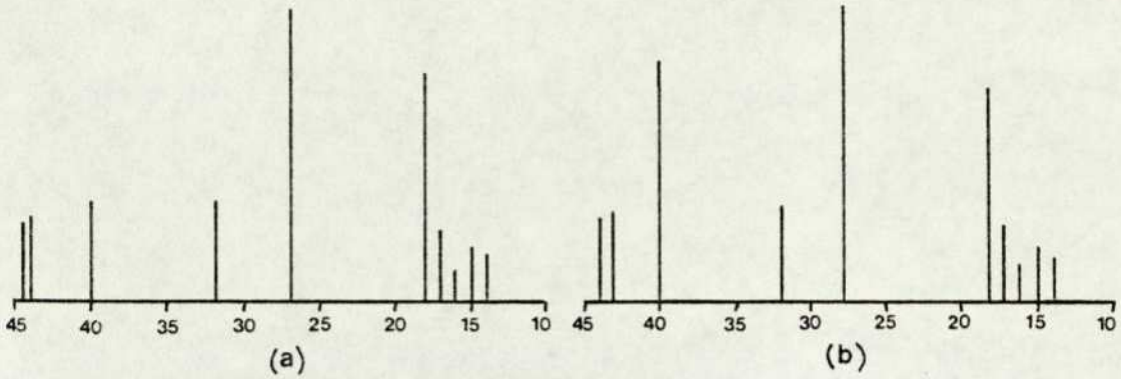


Fig. 17. Residual mass spectra (a) before and (b) after the operation of a high current argon discharge

The peaks can be assigned to the ion species as shown in Table 2. It can be observed that the only significant difference between the two spectra is the increase in the peak at mass 40 after the discharge. This peak is due to the presence of argon as an enhanced residual in the vacuum chamber. Thus as a method of removing the organic contamination in the chamber the high current argon discharge is

m/e	Groups commonly associated with the mass
14	CH_2^+, N^+
15	CH_3^+
16	O^+
17	OH^+
18	H_2O^+
28	CO^+, $C_2H_4^+$, N_2^+
32	CH_4O^+, O_2^+
40	Ar^+, $C_3H_4^+$
43	CH_3CO^+
44	CO_2^+

Table 2. Assignment of the peaks appearing in the residual mass spectra.

ineffective, in particular with respect to peaks such as that at mass 43 which is derived from the acetone used during leak testing. Thus oxygen was substituted for argon as the cleanup discharge agent, and the mass spectra produced before and after operation of the discharge are represented in Fig. 18:

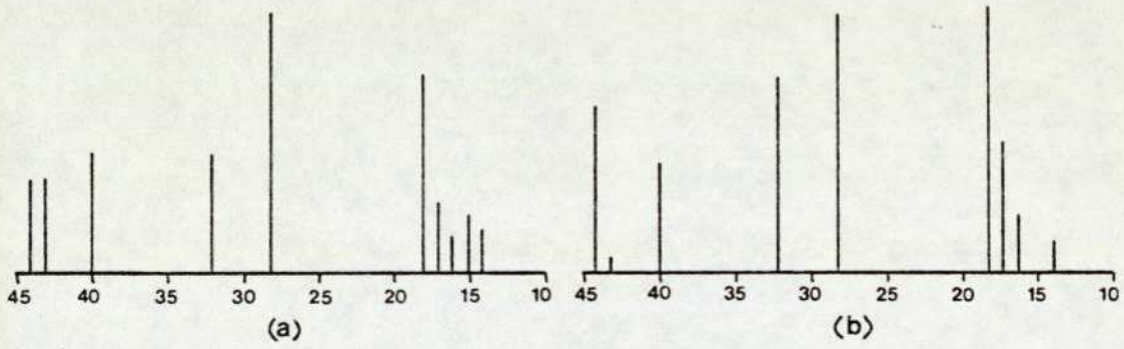


Fig. 18. Residual mass spectra
(a) before and (b) after the operation
of a high current oxygen discharge

As is expected, the peaks due to organic impurities at mass 43 and 15 are significantly reduced, while those due to the oxidation products at mass 44, 18 and 17, and those due to oxygen itself at mass 32 and 16 are significantly enhanced. This indicates that a high current oxygen discharge is extremely efficient in removing organic contamination by oxidation to carbon dioxide and water. The introduction of the high current discharge as a cleanup procedure prior to junction fabrication removes all inherent sources of contamination of the oxide layer, and thus facilitates consistent, reproducible oxide film production.

2.2.6. The junction substrate

In their early papers on IETS [5,6] Jaklevic and Lambe

elected to prepare their tunnel junctions on glass microscope slides, and this technique has been adopted by all other workers in this field. The tunnel junctions studied in IETS have a characteristic metal-insulator-metal sandwich structure where the metals are ~500 to 3000Å in thickness while the insulator is thinner, only ~30Å. Not surprisingly such structures are fragile and are always mounted on an insulating substrate such as quartz, glass or silica for stability and ease of handling.

The substrate dimensions are critical only in that the samples must be immersed in liquid helium during experiments; hence the samples must fit the opening of the helium dewar employed. Of major importance, however, is the cleanliness of the substrate. Improperly cleaned substrates can present a major frustration in terms of reproducibility of good junctions. Small amounts of contaminants such as finger grease, dust particles or soap residue generate considerable difficulties. Poorly cleaned substrates almost inevitably produce tunnel junctions with shorts through the insulator. Proper cleaning reduces the frequency of these shorts but does not eliminate them altogether.

The early junctions prepared in the course of the present study were prepared on glass microscope slides. The slides were cleaned in an ultrasonic bath using a moderately dilute glass cleanser, Deacon 90, for around 15 minutes, followed by a thorough rinsing in hot tap water for several minutes and then in deionised water for around an hour. Finally the substrates were dried in a hot isopropyl alcohol vapour. This method of cleaning was found to produce substrates on which

there was about a 95% probability of forming useful junctions.

Having prepared the junction, however, it is necessary to make electrical connections to the metal strips, and this was found to be an even greater problem. Again, all other workers in the field use the same two methods. Prefired silver paint electrodes applied to the substrate before evaporation allows leads to be soldered to the strips afterwards. This was the first method attempted, and led to two inherent problems. Firstly, it requires the use of aluminium solder, and this spits flux, which can cause breaks in the metal strips, and pulls off under any strain. Secondly, the application of the silver paste must be carried out after the substrate has been cleaned, and this causes contamination of the substrate and thus shorts through the insulator. The second method of attaching connections employed involved the use of screw clamps. Again this required prefired silver paste, and thus neither method proved satisfactory.

At this point it appeared far simpler to find a substrate to which electrical contacts could be applied using ordinary tin-lead solder. The method of Tillery[24] was utilised in an attempt of deposit a copper film on the glass slides, which could then be etched away to form copper pads to fulfil the function of the prefired silver paste. Early formulations for electroless coppering solutions date back to the beginning of the present century when baths were designed for the copper mirroring of glass. Hoffman and Schumpelt[25] were pioneers in this field, using copper formate as the reducing agent.

The solutions used in the modern electroless process are those described by Upton in the mid-1940's[26]. The slides

were initially scrubbed with precipitated chalk, which was then removed by immersion in 10% v/v hydrochloric acid until all effervescence ceased. The slide was then sensitised by the deposition of a thin film of a catalytic metal, usually gold, platinum, palladium or silver. In this case the slide was immersed for 1 minute in a 0.1% w/v solution of palladous chloride with 5 ml/l hydrochloric acid. The addition of 0.01% sodium lauryl sulphate to this activating bath was found to be advantageous in ensuring complete wetting of the glass surface.

The electroless deposition was then performed using a solution which was essentially Fehling's solution to which sodium carbonate and EDTA had been added to improve the stability of the bath and the adhesion of the plating. This solution was mixed with a 25% v/v formaldehyde solution in a ratio of 7:1, with the addition of 0.01% w/v sodium lauryl sulphate in order to reduce the surface tension of the solution and enable the hydrogen to escape as smaller bubbles, thus reducing the roughness of the deposited copper. A 10 minute immersion was found to deposit a film thick enough to allow subsequent electroplating.

However, it was found that, as the slide was immersed in a conventional electroplating solution, the electroless film floated off the glass. This was due to the smooth surface of the pyrex slides, and in order to prevent this occurrence the slides were first washed in a variety of solutions in order to reduce the surface smoothness. These solutions included HNO_3, NaOH, scouring powder and finally HF, but, no matter how the slide was prepared, the copper film would not adhere.

Finally, the initial copper film was deposited by evaporation in a vacuum chamber, but this produced identical results when transferred to the bath for electroplating.

Having attempted to lay a copper film on glass slides in order to produce good electrical connections, the next development was the use of conventional resin-based printed circuit board. This was successful in the case of the lead strip which was of a suitable resistance, but failed for the aluminium, which caused a reaction at the circuit board surface. There was also a considerable amount of impurity generated by outgassing of organics from the resin base below 10^{-5} Torr.

However the solution was finally achieved by the use of fibreglass based printed circuit board. This was found to be stable at pressures down to 10^{-8} Torr, and unaltered by attack from aluminium vapour. The initial circuit boards were cut to the 3" by 1" dimensions of the microscope slides previously utilised and are schematically represented in Fig. 19(a). On preparing the initial junctions on the fibreglass board, it was discovered that problems associated with shorts through the insulator were minor compared to those previously experienced, and thus only one junction needed to be prepared on a given substrate. This allowed the operation of the junction as a two terminal as opposed to a four terminal device, and substrates of the type shown in Fig. 19(b) were utilised. The resistance of the aluminium strip between the copper pad and the lead intersection however was found to be of the order of 40Ω, causing a drop in both a.c. and d.c. voltages across its length. In order to overcome this

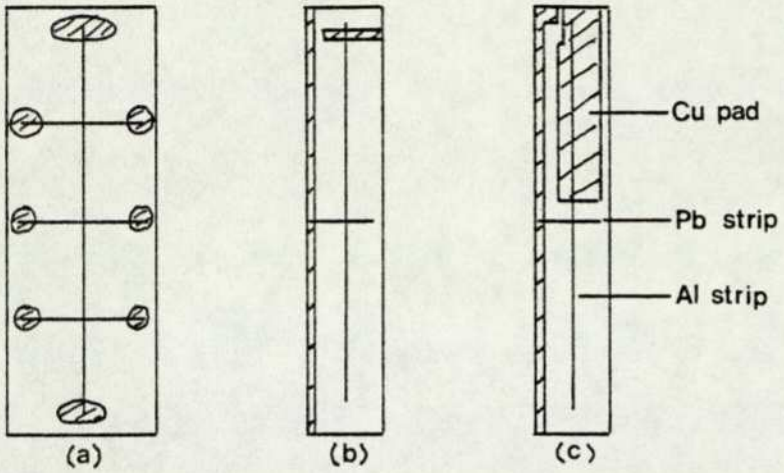


Fig. 19. The various glass fibre circuit board substrates utilized in IETS

voltage drop, circuit boards of the type shown in Fig. 19(c) were employed.

This geometry allows optimum operation of the junction as a two terminal device since the resistance of the aluminium strip was found to be of the order of 2Ω. The use of photosensitive printed circuit board allows the production of 24 of these substrates from a single sheet by exposure to a u.v. source through a suitable mask before etching. The application of this type of substrate to IETS has been found to eliminate all problems caused by poor electrical contacts and the cracking of glass substrates at cryogenic temperatures, as well as increasing the probability of obtaining junctions without shorts through the insulator to around 97%.

2.2.7. Methods of metal evaporation

From the outset of the present study, only one method has been used for the evaporation of the lead top electrode, namely direct resistance heating of a small pellet of lead in a molybdenum boat. Due to the low melting point and

boiling point of lead, and the inert nature of lead vapour, a current of around 30A at a potential of 8V has been found ample. Unfortunately, considerable problems were encountered during the evaporation of aluminium.

The initial method involved the evaporation of a pellet of aluminium from a tungsten wire basket by resistance heating using a current of around 100A at a potential of 8V. It was observed that after seven evaporations, the resistance of the deposited aluminium strip rose considerably, and a black deposit was discovered both on the strips and the masking. The appearance of this black deposit was correlated with corrosion of the tungsten wire, and the deposit itself was subsequently identified by atomic absorbtion spectroscopy as an aluminium-tungsten alloy.

After consultation with several commercial bodies including Nanotech, V.G. and A.E.I. specialising in production of evaporation plant, it was decided to use boron nitride crucibles to hold the aluminium. These crucibles were suspended inside tungsten baskets, and heated by direct contact with the tungsten wire. However the contact between the basket and the crucible was found to be too poor to achieve more than melting the aluminium. Thus the crucible and basket were encased in pyruma fire cement. This considerably enhanced the thermal transmission, and aluminium was evaporated without attacking the boron nitride crucible. Unfortunately the aluminium vapour wetted the pyruma cement, and attacked the tungsten basket producing the same black deposit observed previously.

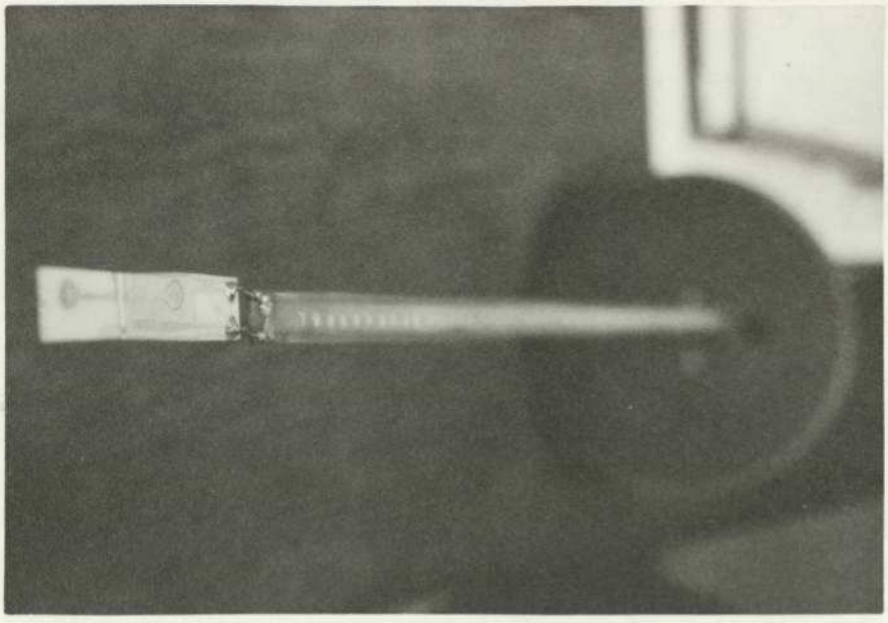


Plate 2a. Completed junction attached to cryogenic probe.

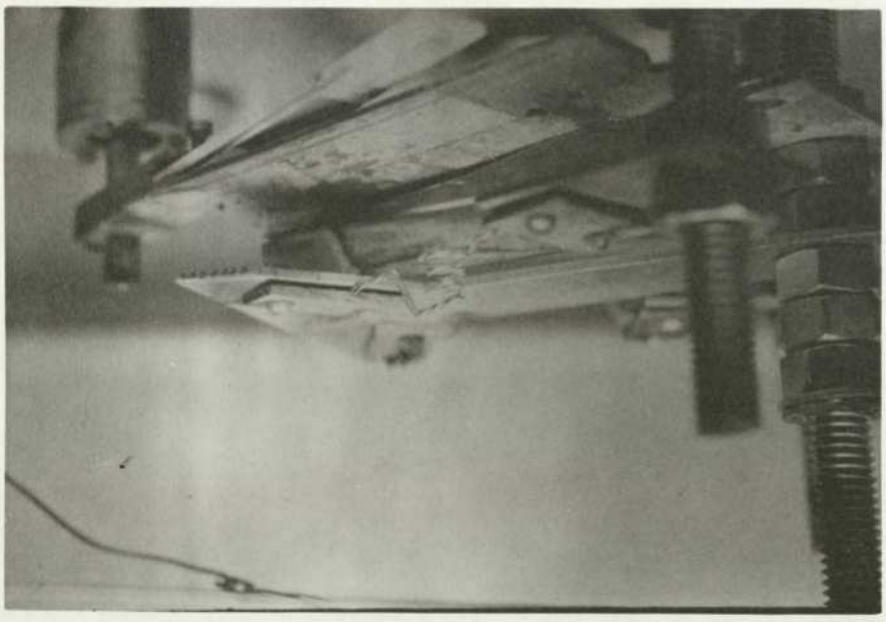


Plate 2b. Tungsten basket and masks for aluminium evaporation

Subsequent methods involved the use of tantalum and pyrophyllite boats, both of which were corroded by aluminium in a very short time; a carbon hearth, which was attacked to produce yellow aluminium carbide; various ceramics, all of which were attacked to varying degrees; alumina thermocouple tubing, internally heated by tungsten wire, which was not attacked, but would do no more than melt the aluminium because of insufficient thermal contact; and a silica crucible heated in a carbon block, which again enabled the production of molten aluminium without sufficient heat transfer for evaporation.

Finally it was decided to return to the initial method utilising a tungsten basket. By careful monitoring of the resistance of the produced aluminium film it was observed that five evaporations could safely be carried out using any basket before attack on the tungsten occurred, and thus a new basket was introduced after every five evaporations. This method is inconvenient only in that the regular removal and refitting of tungsten baskets requires careful manipulation using gloves in order to avoid the introduction of any impurity to the preparation chamber.

2.2.8. System operation and junction fabrication

To produce a junction, the system is loaded with aluminium, lead and substrate and is evacuated to 10^{-6} Torr. Clean-up of residual impurities in the chamber is promoted by the operation, for about 15 min., of an oxygen glow discharge, using an oxygen pressure of 0.5 Torr and a discharge current of around 150mA. As demonstrated previously in Section 2.2.5. this intial clean-up has been found necessary and sufficient to remove impurities from the system to produce a clean, undoped junction.

The system is evacuated to 10^{-6} Torr and a film of aluminium approximately 1000Å thick and 1 mm wide is formed on the substrate by evaporating a charge of 99.99% pure metal through a stainless steel mask. The insulator is then grown on the aluminium strip by the operation of a glow dishcarge of 4mA for 27 sec. in an oxygen pressure 3×10^{-2} Torr. Several features of this discharge are critically important. Firstly, the value of the current established during the discharge, which itself is dependent on the system and electrode geometries, must be constant in order to obtain reproducible results. A value higher than 4 mA necessitates the use of a shorter discharge time or a lower oxygen pressure in the present chamber.

The position of the electrode relative to the substrate is also important. The glow discharge electrode must not be placed along a direct optical path to the oxidising surface in order to protect the oxide surface from pinholing. To this end, the aluminium strip is positioned slightly away from the mask slit, but above the mask. At the same time, it is vital that the copper wire ring electrode is not displaced either horizontally or vertically from its standard position. Vertical displacement alters the distance to the earthed base-plate, and therefore alters the value of the current. Horizontal displacement causes a reduction in the concentration of activated oxygen molecules around the substrate, and thus the formation of thinner oxides and, in the limit, electrical shorts between the aluminium and lead strips.

Finally, the purity of oxygen used is of critical importance. Early workers in the field[6,18] used mixtures of oxygen and water at partial pressure ratios of 2:1 in order to grow

hydrated oxides. It has been found, however, that these are less electrically stable than anhydrous oxides, which are therefore less prone to shorts and more reliable. The growth of anhydrous oxides, however, requires oxygen with very low levels of organic impurity (typically less than 0.1 ppm). The initial oxygen used was grade X supplied by B.O.C. Special Gases in 1 litre sealed glass flasks. This was found to introduce considerable organic contamination to the oxide layer after four or five oxidations.

This was due to backstreaming of organic impurities into the oxygen feed line by diffusion processes caused by dropping pressure differentials. This supply was replaced by "ultra high purity" oxygen supplied by Cambrian Gases, which not only has a lower level of inherent organic impurity, but also is supplied at 0.3 $kgcm^{-2}$ positive pressure, which resists contamination by back-diffusion. Under the present system, described above, oxide barriers are produced with a success rate of greater than 97%.

The introduction of the material to be studied to the oxide surface, doping, may now be carried out by one of several methods. Firstly, the vapour doping technique described by Jaklevic and Lambe[5,6] may be used simply by exposing the oxide to the dopant in the vacuum system at a pressure of 10^{-1} Torr for several minutes. In the case of liquids, such as acetic acid, ethanol and water, exposure of the liquid to the vacuum system via the leak valve will produce sufficient vapour to raise the pressure in the chamber to 10^{-1} Torr. A resistance heater must be used to aid evaporation of less volatile materials such as cyanoacetic

acid. This method produces an oriented deposit of dopant molecules on the oxide surface. It suffers, however, from the necessity of introducing deliberate organic contamination, or water vapour, into the clean preparation chamber, and this can be difficult to remove.

An extension of this method which overcomes this contamination problem has been described by Simonsen and Coleman[27]. They used a twin-chamber system, pumped separately by two nitrogen-trapped diffusion pumps, in which the two chambers were connected via a 2 inch gate valve. A roller-drawn carriage was mounted in the manifold connecting the two systems, operated by means of a rotary vacuum feedthrough. Using this carriage, substrates could be transferred between the separate chambers and positioned for either electrode deposition and barrier preparation or for doping with the molecular "impurity".

After allowing 5-10 min for the films to cool after oxidation, the substrates are transferred without breaking vacuum to the centre of the doping chamber which was surrounded by a liquid-nitrogen jacket. A vapour of the dopant chemical is then introduced to the chamber for 10-15 mins. without pumping. Following doping, the chamber is pumped through liquid-nitrogen traps back to the original pressure of 10^{-6} Torr, and the substrates are then returned through the gate value to the metal-deposition chamber for completion.

A second method of doping has been described by Simonsen, Coleman and Hansma[28]. This technique, termed liquid doping, considerably simplifies the doping procedure. After oxidation, the system is vented to dry nitrogen and the substrate is placed,

face up, on a spinner similar to those used for applying photoresist to integrated circuit chips. A drop of the dopant solution is placed on the junction and the spinner is immediately accelerated to approximately 3000rpm to spin off excess solution. The size of the drops of dopant solution is not critical, conventional disposable pipettes serving adequately to transfer the solution to the junction. After doping, the slides are returned to the vacuum system which is pumped down to 10^{-5} Torr for evaporation of the top lead electrode.

Liquid doping has the advantage of maintaining the low levels of organic contamination within the fabrication chamber, reducing the cleaning time between successive junction production and considerably increasing pumping speeds. However it does suffer from the necessity of exposing the junction to the atmosphere before and during doping. The method is applicable to the complete range of compounds that can be dissolved in solvents such as water, ethanol, benzene and chloroform.

The correct concentration of dopant solution, that is, the concentration giving resistances in the desired range (50 to 1500Ω) is determined by trial and error, although similar compounds usually require similar concentrations. The general range of any dopant is 0.2 to 4 mg/g. The solvents noted above were used since they are sufficiently volatile to be pumped from the surface and contribute only very small, if any, peaks.

After either of these methods of doping, the chamber is reevacuated to 10^{-5} Torr and the junction is completed by evaporation of a lead top electrode of width 1 mm and thickness around 3000Å. After approximately 5 min for cooling, the

chamber is vented to dry nitrogen and the completed junctions are removed and tested with a low power ohmmeter[29]. It has been found that the resistance of undoped junctions varies from 0.1 to 10Ω while that of doped junctions varies from 50 to 1500Ω. Junctions outside this resistance range produce inferior, though sometimes acceptable, spectra. The higher resistance junctions are usually electrically noisy, while the lower resistance junctions yield spectra of poorer resolution.

Both the doping techniques described above require the inclusion of the molecules to be studied into the junction during fabrication. Recently, Jaklevic and Gaertnner[30,31] have introduced a third method which enables doping of completed junctions. The tunnel junctions are fabricated as described above without the inclusion of the molecular layer and may be stored in a dry box without any substantial change in their properties, the IETS spectrum remaining stable and uncontaminated for at least three months. Doping is accomplished by infusion from vapour, the junctions being exposed to the organic molecules of interest while in the presence of water vapour and atmospheric air. The substrate is placed on a pedestal inside a small Pyrex container containing a few cm^3 of water and is covered with a close fitting glass lid. All glassware is flame cleaned to remove volatile organic contamination. The container rests in a dish of shallow water which, by evaporation, serves to keep the cup at a temperature slightly below ambient. Thus the relative humidity at the substrate is kept slightly below 100% so that a fluctation in room temperature cannot cause the formation of dew on the substrate.

A certain amount of the organic dopant of interest can be

mixed with the water inside the humidifier so that the tunnel junction is exposed to a mixture of water and organic vapour for a given time period. In the case of very low vapour pressure organic molecules, the organic is first dissolved in a suitable solvent and a thin layer is deposited over the tunnel junction by dipping in the solution and evaporating away the solvent. The junction is then exposed to water vapour for a period of time. Lastly, infusion directly from the liquid state is accomplished simply by dipping the substrate in a dilute solution of pure water and the organic of interest.

It has been observed that lead films can survive fifteen minutes or more of immersion in pure water and yet can be damaged in tens of seconds if dew is allowed to form on them in a vapour atmosphere. This change in structure is evident from the increase in film sheet resistance of 10% or more. Certain organics such as formic acid appear to cause changes in the overlay metal much more than others (e.g. phenol) while even pure water vapour can cause easily observable changes.

As infusion progresses, the resistances of the junctions are found to increase with time with a rate determined by relative humidity, the organics present and the overlay metal film. Typically, exposure to carefully purified water at about 95% relative humidity causes the resistance of an Al-Pb junction to increase from 0.1 to 10^3 ohm in a few minutes while 70% relative humidity exposure requires several hours to cause the same resistance change. For low relative humidities, less than about 35%, no change in the resistance or IETS spectra are observed for long periods of time. During infusion the

resistance of the junction is continuously monitored so that the process can be stopped in the range 100 to 1000 ohm. If left in the infusion vapour the resistance will continue to increase to values of 10^6 ohm or more even if no organics are used.

The critical role of water vapour as a "carrier" is believed to result from its ability to adsorb on to the surfaces of the film comprising the tunnel junction. The aluminium is covered with a thin layer of oxide, similar to γ-alumina[32] and presumably a thin layer of oxide would exist on the lead film as well. The adsorption behaviour of water on various surfaces, especially oxides, has been the subject of extensive study and is of commercial importance[33]. For the case of porous γ-alumina it is found that, for relative humidity varying from 0 to about 50% the average coverage of adsorbed water ranges from zero to about one monolayer (Fig. 20).

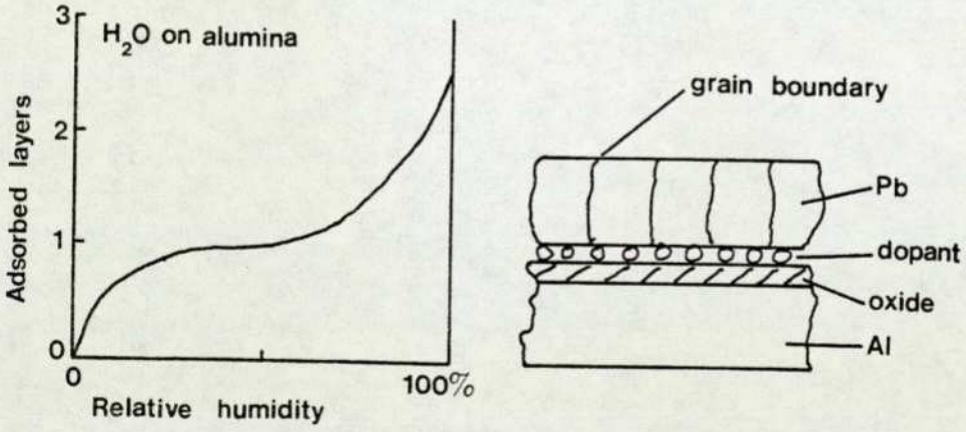


Fig. 20. Schematic showing the qualitative behaviour of water adsorbed on an alumina surface, and a cross section of a tunnel junction.

Above 50% relative humidity the adsorbed water can attain even higher average coverage reaching several monolayers above 90% relative humidity depending on the detailed structure of the adsorbate. The initial layer is bound more strongly than subsequent ones, which are bound with energies similar to those in liquid water. There is also evidence that the water in such multilayers is mobile[34]. The fact that organic molecules are only infused into a tunnel junction in the presence of water vapour indicates that their mobility is greatly enhanced by the presence of the adsorbed water layer.

Jaklevic and Gaerttner have assumed that a similar type of behaviour occurs on the surface of the lead film. The observation that the organic species penetrate the lead film at a relatively rapid rate suggests that there is sufficient space along the lead grain boundaries to admit molecules as large as 12Å. Although some change, corresponding to enlarged grain size and surface roughness has been observed in most cases during infusion, the fact that this effect can also occur for a molecule such as phenol with only minimal change in lead surface appearance leaves in doubt whether infusion must necessarily be accompanied by a gross change in the structure of the overlay film.

2.3. Cryogenics

The linewidth of the d^2V/dI^2 versus V curve of a molecular excitation, and hence the resolution attainable in tunnelling spectroscopy, is determined by natural width, thermal broadening, modulation broadening and electrode (density of states) effects. The modulation circuitry described in Section 2.4. is designed

with the particular objective of minimising modulation broadening. The other two experimentally controlled broadening effects are a function of the temperature at which the experiment is performed.

For normal metal electrodes at finite temperatures, there exist filled electron levels above and unfilled electron levels below the Fermi level, E_f. As a consequence, the bias voltage required to cause a vibrational transition may be greater or less than the $h\nu/e$ depending on the energy of the initial and final energy states of the electron, measured with respect to the Fermi energy. It can be shown that this variation of energy levels leads to a thermal smearing, which in turn broadens the linewidth to a width at half peak height equal to $5.4kT$[9]. This broadening due to the thermal smearing of the energy levels ranges from about 150meV at room temperature to about 2 meV($16cm^{-1}$) at liquid helium temperatures. Thus to produce the most highly resolved spectra, liquid helium temperatures are required. Spectra can be recorded up to liquid nitrogen temperatures, 77K, but at this temperature the spectra merely consists of an envelope of the fine structure observed at lower temperatures. The broadening effects caused by temperature evaluation are schematically represented in Fig. 21.

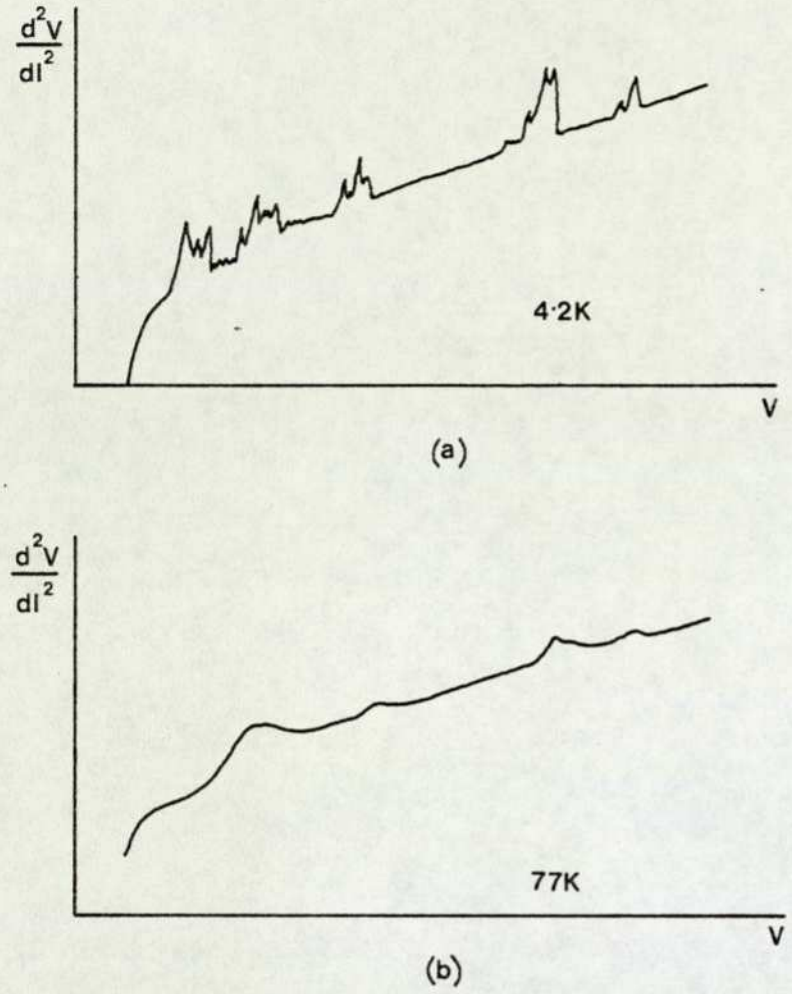


Fig. 21. Schematic representation of IETS spectra at (a) 4.2K and (b) 77K.

The thermal smearing referred to above is considerably less in superconductors because of an energy gap of 2Δ centred on the Fermi level. At liquid helium temperatures, lead, which is the usual upper electrode, is superconducting and this again helps to reduce the broadening. Giaever et al[35] have reported an increase in resolution of a factor of 5 in going from normal to superconducting electrodes.

The experimental problems involved in the handling of liquid helium are well documented. Initially a cryostat was considered for holding the liquid helium while the junction was

immersed, but this was dismissed as it would need to be emptied to remove the junction, and was thus prohibitively expensive with regard to the liquid helium consumption. The original system utilised involved two glass dewars, one inside the other. The outer dewar was filled with liquid nitrogen, and with several spacers to protect the inner dewar, which was filled with liquid helium. The inner dewar was fitted with a tap attached to a rotary pump, since helium vapour passes through glass and thus reduces the vacuum seal to atmospheric pressure in a very short time. The pair of dewars then stood inside a μ-metal chamber, and the junction was suspended by copper rods from the bridge network.

This system performed adequately during the early stages of junction fabrication when samples were run at liquid nitrogen temperatures. However when the system was operated with liquid helium in the inner dewar, the boil-off rate was found to be too high to allow equilibrium to be established at 4.2K without continuous syphoning of liquid helium into the inner dewar. This again resulted in prohibitively expensive running costs,and the system was finally abandoned in favour of direct immersion in the liquid helium storage dewar.

The completed junctions are mounted on a specially designed strip of epoxy-glass circuit board and are immersed directly into the standard liquid helium storage dewar. The circuit board mounting system consists of a single strip, 3ft. long and 0.3in wide, along which run two copper strips which are used to make electrical connections with the junction substrate. Circuit board is utilised size it provides a non-fragile, rigid, insulating conduit for the connections to the junction. The strip of circuit board has holes drilled

Plate 3. Bridge and junction probe inserted into liquid helium dewar

throughout its length to reduce its thermal capacity and reduce helium loss through excess boil-off.

2.4. Modulation Techniques and circuitry

2.4.1 Introduction

As discussed in Section 1.2., the excitation of a vibrational mode of an adsorbed molecule by electrons traversing the insulating barrier is manifested by a break in slope of the I-V characteristic of the junction. The requisite derivatives of the I-V curve can be obtained by suitably processing the I-V curve electronically or with a computer[36]. However, the more usual way is by modulating the current through or the voltage across the junction which has a weakly non-linear I-V characteristic. This gives rise to an alternating signal which, because of the nonlinearity, contains harmonics of the fundamental frequency, the magnitudes of which are proportional to the derivative of the characteristic. The modulation approach has become standard in NMR and EPR spectroscopy[37] and is extensively used in optical spectroscopy[38]. The signal and resolution techniques used in these spectroscopies[39] will be adapted to IETS in the near future.

2.4.2. Theory of modulation techniques

The modulation technique as applied to tunnelling spectroscopy is discussed by several authors[40,41], but the circuitry utilised in the present study is modelled on that of Adler and Jackson[42]. The basic measurements required in a tunnelling experiment are the I-V curve, the dynamic conductance $G = dI/dV$, and $dG/dV = d^2I/dV^2$. However, for reasons of sensitivity and simplicity, most spectra have been obtained using instrumentation which gives the derivatives

$dV/dI - R_T$, the dynamic tunnel resistance, and d^2V/dI^2. The derivatives are related by the identities $dI/dV = 1/(dV/dI)$, using which:

$$\frac{d^2I}{dV^2} = \frac{d}{dV}\left(\frac{dI}{dV}\right) = \frac{dI}{dV}\,\frac{d}{dI}\left[\frac{1}{dV/dI}\right]$$

$$= -\frac{dI}{dV} \cdot \frac{1}{(dV/dI)^2} \cdot \frac{d^2V}{dI^2}$$

$$= -G^3\left(\frac{d^2V}{dI^2}\right)$$

Since both dV/dI and d^2V/dI^2 can be obtained, the conversion to G and dG/dV can be effected using ratio or computer techniques. However, when studying molecular excitations, it has not been the practice to determine the absolute values of dI/dV and d^2I/dV^2 since, firstly, good theoretical calculations of intensity are not available for fruitful comparison, and secondly the conductance in the regions of interest, away from zero bias where there are rapid changes in conductance due primarily to superconducting effects, is very nearly constant so that $d^2I/dV^2 \propto d^2V/dI^2$. Thus the peak positions and shapes of d^2V/dI^2 and d^2I/dV^2 versus bias voltage characteristics are identical, and only the absolute intensity is lacking[43].

The application of the modulation technique to tunnelling spectroscopy is most clearly expounded by Adler and Jackson[42]. Consider the simplified constant current circuit of Fig. 22.

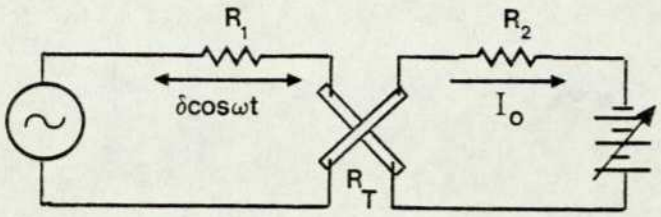


Fig. 22. Simplified circuit showing a junction with dc bias and ac modulation

The resistances R_1, $R_2 >> R_T$ so that the bias current I_o and modulation current amplitude δ are effectively constant as the tunnel resistance R_T changes. Fig. 23 shows the effect of sine wave modulation of the current at a hypothetical non-linear I-V curve. The resulting output voltage is distorted due to the nonlinearity, and will contain harmonics of the fundamental frequency $\omega/2\pi$. The voltage across the junction, according to Ohm's law, is:

$$V(I) = V(I_o + \delta\cos\omega t)$$

This can be written as a Taylor series expansion as:

$$V(I) = V(I_o) + (dV/dI)_{I_o}\ \delta\cos\omega t + \tfrac{1}{2}(d^2V/dI^2)_{I_o}\ \delta^2\cos^2\omega t + \ldots$$

and expanding multiples of $\cos\omega t$:

$$V(I) = V(I_o) + (dV/dI)_{I_o}\ \delta\cos\omega t + \tfrac{1}{4}(d^2V/dI^2)_{I_o}\ (1+\cos 2\omega t)+\ldots$$

Thus the voltages proportional to the derivatives

$$V_\omega = \text{constant.}\ (dV/dI)_{I_o}\ \delta\cos\omega t$$

and $$V_{2\omega} = \text{constant.}\ (d^2V/dI^2)_{I_o}\ \delta^2\cos 2\omega t$$

are available across the junction and can be obtained using a lock-in amplifier[44]. Obviously, higher harmonic terms could be retained in the expansion and detected by the lock-in although at reduced signal levels.

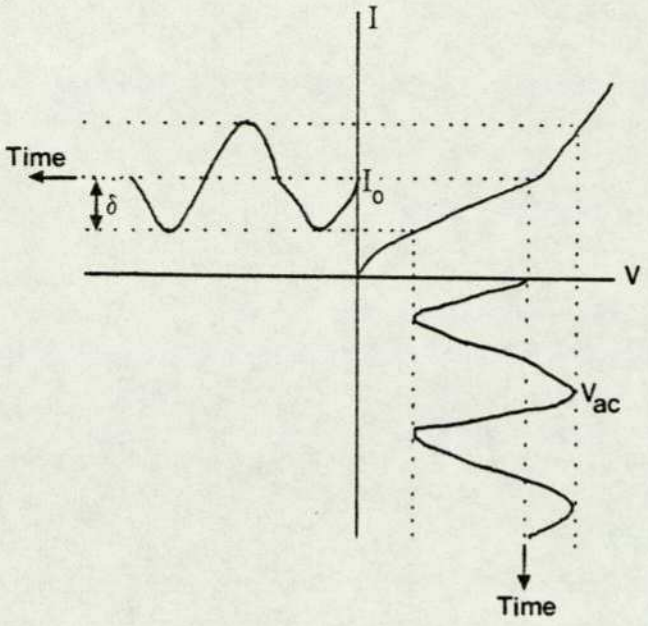


Fig. 23. The distortion of the output voltage of a sinusoidal modulation voltage across a nonlinear I-V curve

As an alternative, the experiment can be performed so that G and dG/dV are obtained directly. This corresponds to maintaining constant voltage conditions and observing the expansion of $I=I(V_o + v\cos\omega t)$. The current through the junction, which can be sampled by measuring the voltage across a resistor in series with the junction, will then have frequency components proportional to the desired derivatives[45]. Thus, depending on the details of the circuity, the junction conductance or resistance and their derivatives can be obtained.

2.4.3. The use of modulation techniques in IETS

The method of measurement outlined above has been used in conjunction with synchronous detection using a lock-in amplifier in some early studies in tunnelling spectroscopy[35,41]. Unfortunately, direct application of this method using a preamplifier and a lock-in amplifier cannot provide the required high resolution. Since the junctions of interest are only weakly nonlinear, the

variation of interest comprises only a small part of the total signal across the junction, so that most of the lock-in amplifier output signal must be bucked out in order to display these small variations on an X-Y plotter or similar recording device. Two difficulties immediately arise: firstly the high gains required cannot always be attained because of the limited dynamic range of the lock-in amplifier, and secondly significant errors may be introduced due to gain variations of the amplifier and instabilities in the bucking voltage. These two difficulties may be overcome by bucking out most of the signal prior to amplification.

Such a method has been applied[42] by incorporating the tunnel junction as one arm of an ac Wheatstone bridge. An added advantage of this approach is that it provides rejection of the fundamental frequency when measurements of d^2V/dI^2 are made. This cancellation of the fundamental is complete when the bridge is in balance, while the signal proportional to d^2V/dI^2 (i.e. the second harmonic) is unaffected, since the tunnel junction is the only nonlinear element of the bridge.

Consider first the simplified bridge configuration shown in Fig. 24. One arm of the bridge contains the tunnel junction R_T which is treated as a nonlinear passive element, while the other arm has R_d set so that $R_d \approx R_T$. The inductance L and resistance R_a are large enough that the ac shunting of the tunnel junction by the dc bias supply may be neglected, and R_b is large enough so that shunting by the ac modulation supply is negligible. Under these conditions the current in the two bridge arms is the same.

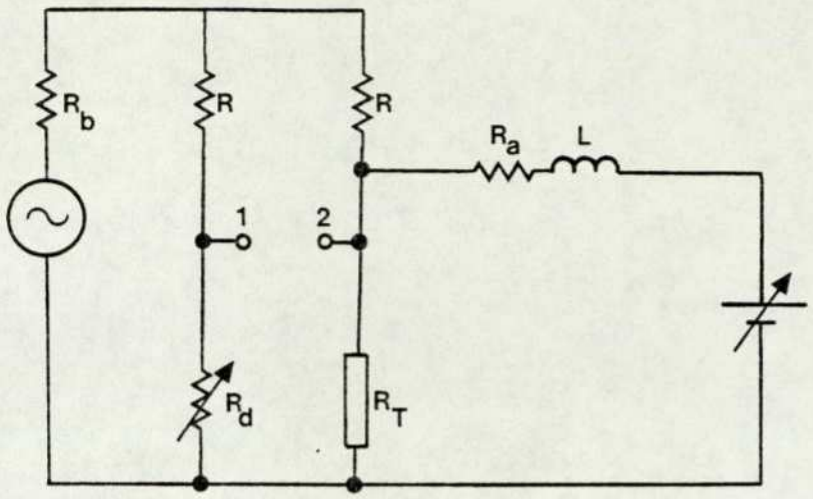


Fig. 24. Simplified constant current bridge configuration

If the current in each arm is $\delta\cos\omega t$, and V_1 and V_2 are the voltages at points 1 and 2 respectively then from section 2.4.2.:

$$V_1 = R_d\delta\cos\omega t$$

$$V_2 \simeq I_o R_d + (dV/dI)_{I_o}\,\delta\cos\omega t + \tfrac{1}{4}(d^2V/dI^2)_{I_o}\,\delta^2(1 + \cos 2\omega t) + \ldots$$

where I_o is the dc bias current flowing through the junction when the bridge is balanced so that $R_d \simeq R_T$. Thus the difference in potential between points 1 and 2, V_{12}, has components at the modulation frequency ω and at its various harmonics. The only functions of interest in tunnelling spectroscopy are dV/dI and d^2V/dI^2 and the relevant components of V_{12}, by subtraction, are:

$$V_{12}(\omega) = \delta\,\left|R_d - (dV/dI)_{I_o}\right|\cos\omega t$$

$$V_{12}(2\omega) = \tfrac{1}{4}\delta^2(d^2V/dI^2)_{I_o}\cos 2\omega t$$

Accurate determination of R_d enables the use of the full scale of the recording device to display the small variation in dV/dI of the tunnel junction rather than dV/dI itself. In this way the problems of dynamic range and gain stability of the lock-in amplifier are circumvented. Similarly, and more significantly for inelastic tunnelling where d^2V/dI^2 is the most useful function to obtain, if R_d is accurately selected so that the bridge is balanced, a good deal of fundamental rejection occurs when $V_{12}(2\omega)$ is measured.

2.4.4. The design of a high resolution bridge

The basic circuit can be represented by the block diagram in Fig. 25.

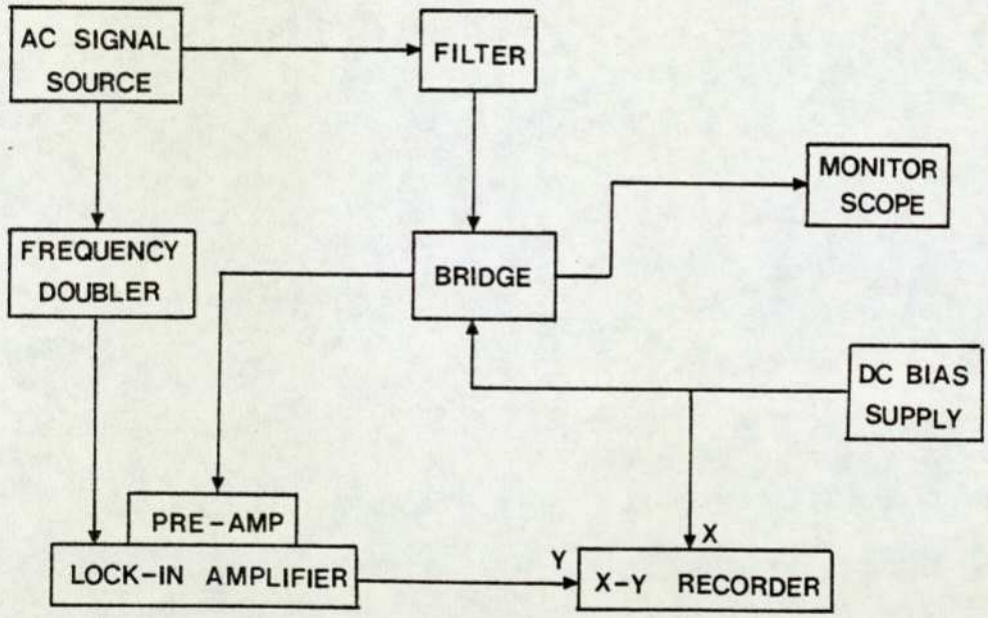


Fig. 25. Block diagram of a bridge spectrometer

The initial design for the bridge and the bias circuit, as well as the purchased instruments, are shown in Fig.26. The ac source must provide a coherent in-phase signal both to the bridge and to the reference channel of the lock-in. Thus the single output oscillator must be fitted with a coax T-piece

adaptor in order to derive the two signals from the same source. It was initially decided, mainly for convenience, but also with consideration of the lock-in characteristics, to operate with a fundamental frequency of 500 Hz, and thus the frequency for ultimate detection is 1000 Hz, the second harmonic. In the Brookdeal system purchased for the present study there is a pushbutton controlled, low-distortion frequency doubler incorporated, thus obviating the necessity for the construction of a separate frequency doubler. The ac signal to the junction must have a second harmonic content which is less than the level of the harmonic generated by the nonlinearity of the junction. This necessitates using a stable, low distortion ac source which can be followed by tuning circuitry to supress the second harmonic. This was the major factor in the choice of oscillator, the present one being chosen since its second harmonic content in the output signal is less than 0.005%. At this level, the distortion is so low that further filtering is unnecessary.

The inductor-capacitor tuned circuits marked (A) and (B) in Fig. 26 are designed to prevent any second harmonic feedback from the bridge to the oscillator, and hence to the reference of the lock-in. (B) has a high impedence to 1000Hz and hence prevents return from the bridge network by rejection, while (A) has a low impedence to 1000Hz and thus allows any second harmonic which overcomes (B) through to ground. Both the 2.5H coils and the 6.5H coil had to be ordered as components and wound as they are of too high inductance to be purchased in a convenient size. The coils were wound using the following number of turns of wire on Mullard LA1411 cores:

2.5H 2030 turns 36 swg (0.2 mm) insulated copper wire

6.5H 3273 turns 38 swg (0.16mm) insulated copper wire

The inductance of a wound coil is related to the number of turns of wire on a given core by the formula:

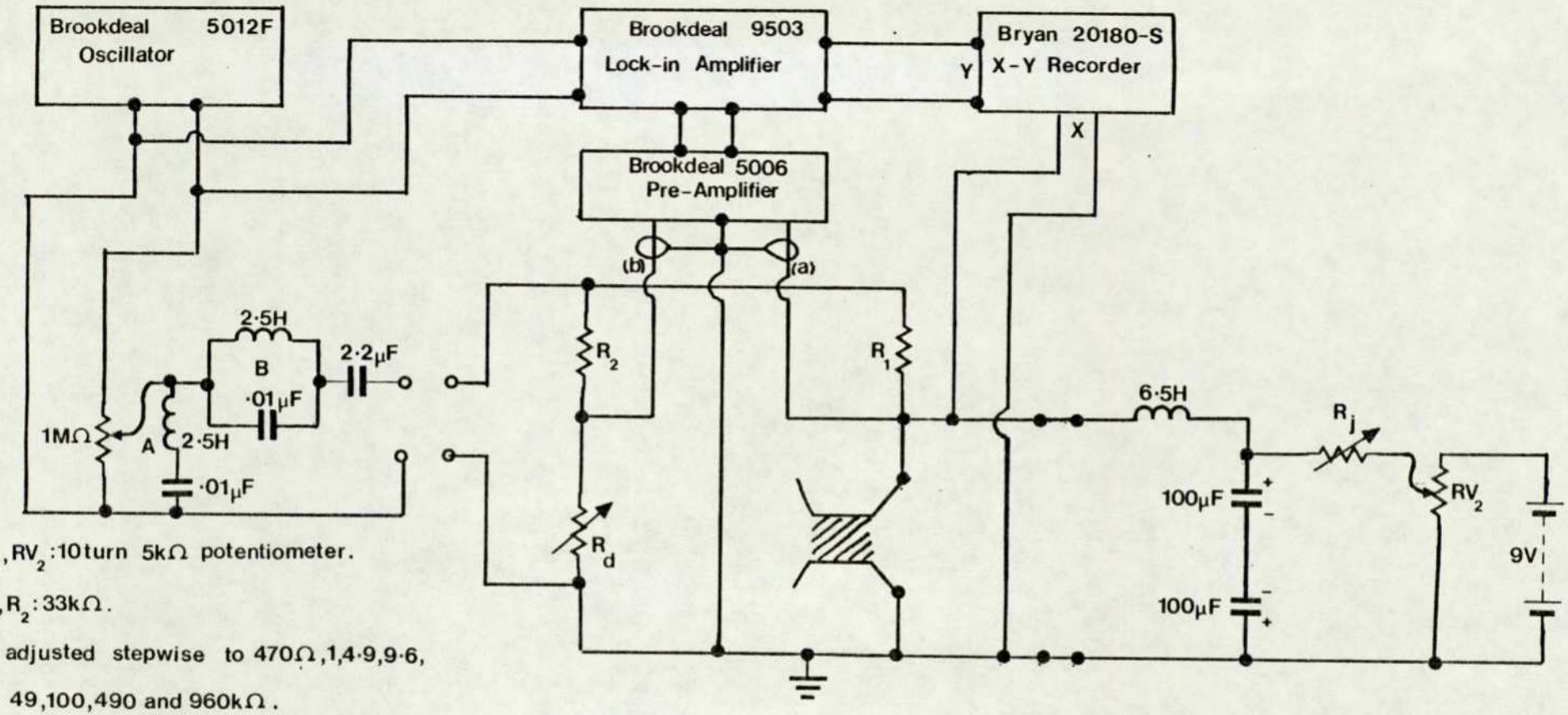
$$L = \frac{3.2 \times N^2 \times \mu \times A}{1 \times 10^8}$$

where L is the inductance in Henrys, N is the number of turns of wire, μ is the incremental ac permeability of the iron core material, A is the cross sectional area of the winding limb in square inches and l is the length of the magnetic path in inches[45]. An initial coil was wound using 1000 turns of insulated copper wire, and its induction calculated by measuring the rejection frequency in a tuned circuit using a 0.1μF capacitor and then using the standard formula:

$$F = \frac{1}{2\Pi\sqrt{LC}}$$

where F is the frequency in hertz, L is the inductance in henrys and C is the capacitance in farads. The number of turns required for the correct inductance was then calculated by a simple proportion based on $L \propto N^2$.

The pre-amp used is differential, and this further aids the bucking out of unwanted signal; since only signals which appear on arm (a) in excess of those on arm (b) are passed on to the lock-in. The band width of the lock-in is precise around 1000Hz, and thus no 500Hz filter is required before the amplification stage. The two reversed 100μF capacitors are to prevent the dc bias reaching the amplifiers, while the



R_d, RV_2 : 10 turn 5kΩ potentiometer.

R_1, R_2 : 33kΩ.

R_j adjusted stepwise to 470Ω, 1, 4·9, 9·6, 49, 100, 490 and 960kΩ.

Fig. 26. Initial bridge and bias network

6.5H choke shunts ac from the dc bias supply.

The dc bias itself is produced by a battery circuit modelled on that of Blackford[47]. RV_2 is a ten-turn motor driven potentiometer complete with suitable gearing to enable a range of scan times between 1 minute and 80 minutes to be attained. R_J is a switch resistor so that the voltage scan range can be adjusted to suitable values.

Since signal levels are less than 1μV in many instances, care must be taken in shielding and properly grounding the system components to keep the noise at a low level. To this end, both the main bridge network and the dc supply with the filters are screened in μ-metal boxes, while all leads consist of screened coax cable. The system was initially grounded at the main 13A supply, with the exception of the oscillator and dc bias supply which float.

On constructing and operating this sytem several difficulties became immediately apparent. The first, and by far the most significant problem was the amount of mains-borne and air-borne electrical noise within the system. It is usually found that noise is composed of randomly occurring voltages which are unrelated in phase or frequency and may sometimes be of a very spiky nature. On close examination, noise voltages show pulse-like waveforms, some with large peaks, which occur quite randomly and frequently. When viewed on an oscilloscope noise gives a "spiky" impression, where on average the peaks are about a microsecond in duration and therefore have high frequency components. It is these high frequency components which are particularly troublesome in the present study, since they are

generally unresponsive to conventional tuned-circuit filtering, and thus must be eliminated at source.

As there are many sources which produce electrical noise, they may be broadly classified as natural or artificial. Artificial or man-made noise arises mainly from electrical equipment such as commutator motors, sparking plugs in ignition systems, faulty switches, electric shavers, etc. They produce noise voltages which very often have regular properties and may also be regarded as a form of interference. The noise effect of sparking plugs is clearly seen on a domestic TV screen as a set of bright dots which cover the whole picture as a car drives past the house. A study of man-made noise reveals that it can be minimised or eliminated at the source by the use of suppressors or by improved component design[48].

The natural forms of noise are those due to cosmic radiation, atmospherics and the more familiar circuit noise associated with electronic circuits. Cosmic noise and atmospherics mainly comprise electromagnetic radiation from solar and galactic sources, and usually enter the system via a receiving antenna. Thus these sources are relatively unimportant in the circuits used in the present study, since all cables were initially designed as earth-shielded coaxial leads.

The most troublesome form of naturally occurring noise is circuit noise, generally divided into thermal noise and shot noise[49]. Thermal noise is produced by the random motion of free electrons in a conductor. Due to collisions with vibrating atoms in the conductor, the electron motion depends on temperature and produces a fluctuating current in the conductor.

The average current in the conductor is zero, but the random fluctuations produce a noise voltage across the conductor on open circuit. Shot noise on the other hand, is due to the random fluctuation in the emission of electrons from a cathode surface, which produces a shot noise current in a valve. Random variations in a semiconductor diode or transistor also give rise to a shot noise current in the device.

Having briefly considered the main sources of noise in general, the methods of elimination and reduction utilised in this study will now be discussed. The first problem encountered in the original circuit design is the presence of earth loops. Just as alternating magnetic fields will induce noise into exposed coil windings and their cores, so they may also interact with unintentional "coils" existing as loops in the signal wiring. An example of how superfluous earth connections can create a loop into which noise may be induced is described in Fig. 27.

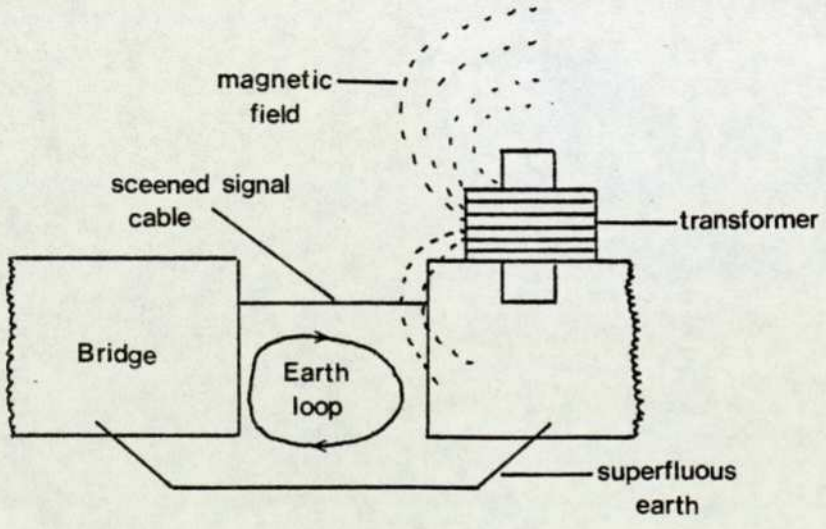


Fig. 27. Creation of an earth loop in which noise may be induced by interaction with adjacent magnetic field

An extra earthing wire duplicates the function of the screened signal cable's outer conductor, at the same time creating a loop or single-turn coil that can interact with a magnetic field creating noise. In Fig. 27 part of the field from the amplifier mains transformer is seen passing through the earth loop formed by the wiring, the effect being to induce a tiny circulating current in the loop which, because of the minute ac voltage resulting across the earth side of the signal feed, causes noise generation. Thus ideally interconnecting cables in the system should be arranged to convey signals by one path only between any two points, that path comprising one live and one earthed connection in the form of a screened lead.

Thus, initially, it is necessary to determine which terminals on the oscillator, preamplifier, amplifier, dc supply, bridge, recorder and oscilloscope are floating, that is not earthed to the case, and which are case-earthed. This is simply established using an ohmmeter between the outside of the terminal and an established earth. Having decided which terminals are floating, the entire system is mapped so that single-connection earths are established to a common point. By consulting the wiring diagram of the amplifier system, this point is established as the case of the lock-in amplifier, which is earthed via the power lead, while the preamplifier and oscillator are internally floated at their power supplies at 10Ω above earth, so as to avoid internal earth loops. Thus all connections are made to the case of the lock-in amplifier which is then earthed via the mains. This procedure alone reduced the apparent noise level by a factor of a half.

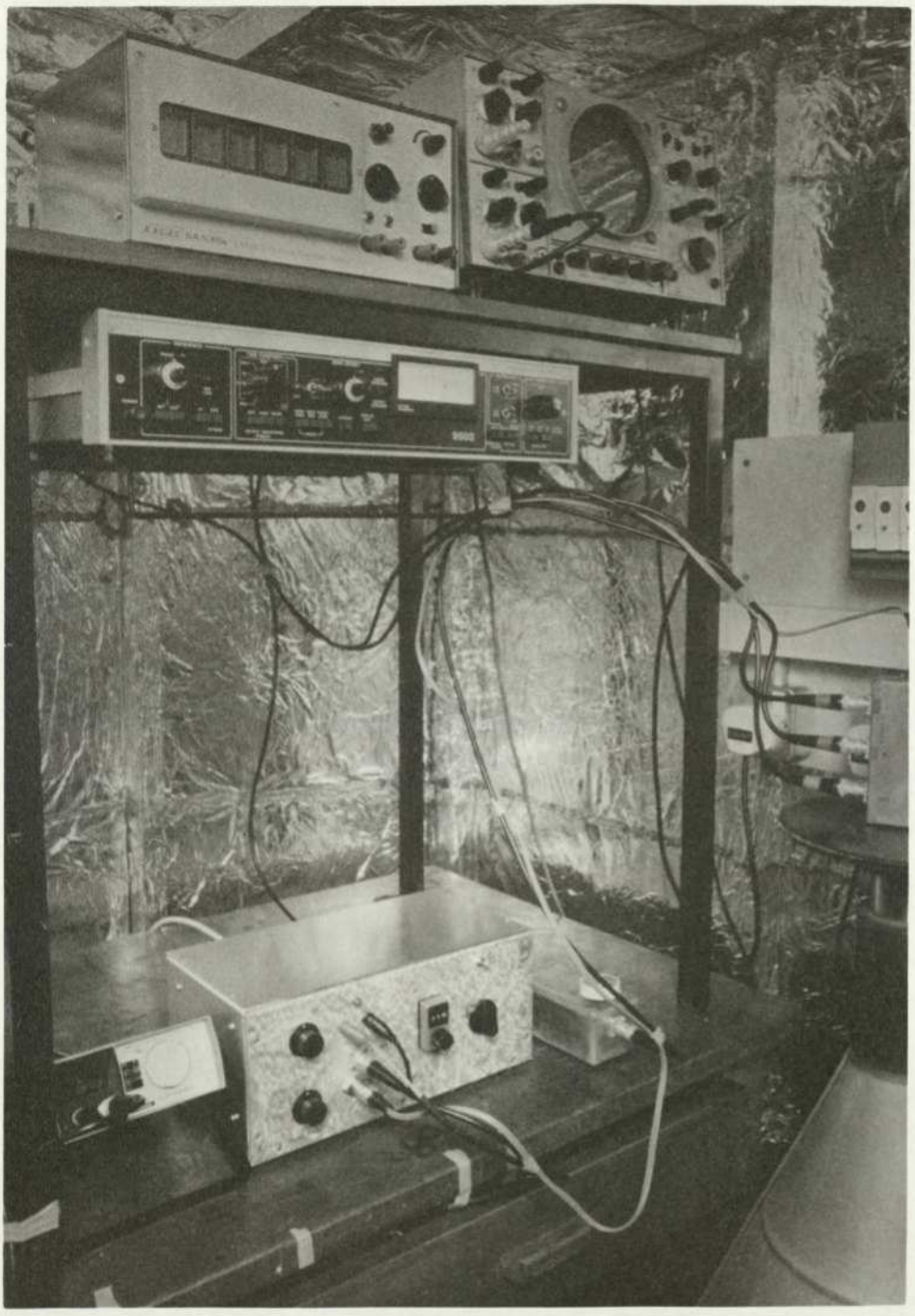


Plate 4. Faraday cage with electronics

The next apparent noise source to be tackled is a function of the location of the equipment. The Chemistry Department in The City University is situated on a busy road with a considerable volume of traffic. This leads to a high noise level from motor ignition systems. In addition, the laboratory in which the tunnelling spectrometer is operated is situated opposite the main lift, and this again leads to a considerable level of airborne noise. As a result, when the equipment is operated in the open laboratory, the noise level is excessive, and signal detection is impossible, despite the use of μ-metal shielding and screened leads. This problem is overcome by operating the electronics of the spectrometer inside a Faraday cage. The cage was constructed from a garden shed, lined with aluminium foil forming a complete screen which is then earthed, thus diverting all airborne noise to ground. It is imperative that the aluminium lining is complete: to this end the pieces are all joined with aluminium solder to ensure electrical integrity.

Furthermore, on monitoring the noise level of the mains power supply and earth, a high noise level of spikes is observed, derived from the operation of thermostats and relays in adjacent laboratories. This noise source is removed firstly by ensuring earthing of all equipment to a separate, quiet ground. A convenient water pipe which runs from the laboratory to the ground outside is used in the present system. The noise level on the live and neutral terminals of the power lines themselves requires a second solution, and the present spectrometer is run off the output of a commercial generator, model F7649, supplied by Newton-Derby Ltd. This generator runs off the 3-phase power

supply, which, having considerably less use within the department, has a far lower inherent noise level. The generator has an output power of 5KVA, which is sufficient to run all the electronics using around half the available output. Moreover, and more importantly, the frequency and voltage of the output is stable to within 0.5%, and this is essential to the accurate operation of the lock-in amplifier, which becomes unstable under varying mains power conditions.

Having taken the precautions outlined above, the circuit of Fig. 26 provides a system which in the first instance proved acceptable. However, it suffers from two major design faults. Firstly, there is no protection for the junction from noise generated within the valve-based X-Y recorder. This is inevitable using this circuit, since any conventional filtering system has an inherent impedence which will alter the absolute value of the dc voltage reaching the X-Y recorder, and thus will remove the exact relationship between the voltage across the junction and the voltage on the X-axis of the recorder.

The second problem is a result of an even more fundamental design error. Since the choke which bucks ac from the dc bias has to be of the order of 6.5H if the modulation frequency is 500Hz, it has a resistance of 650Ω, even using the most efficient core and coil winding techniques. This leads to the circuit developing into the classic example of a potential divider, in that the voltage appearing across the junction is determined by the ratio of the junction resistance and the choke resistance. Thus only very low voltages are obtainable across junctions having resistances in the range 0-80Ω. This in itself is a drawback, but is not particularly important since most junctions

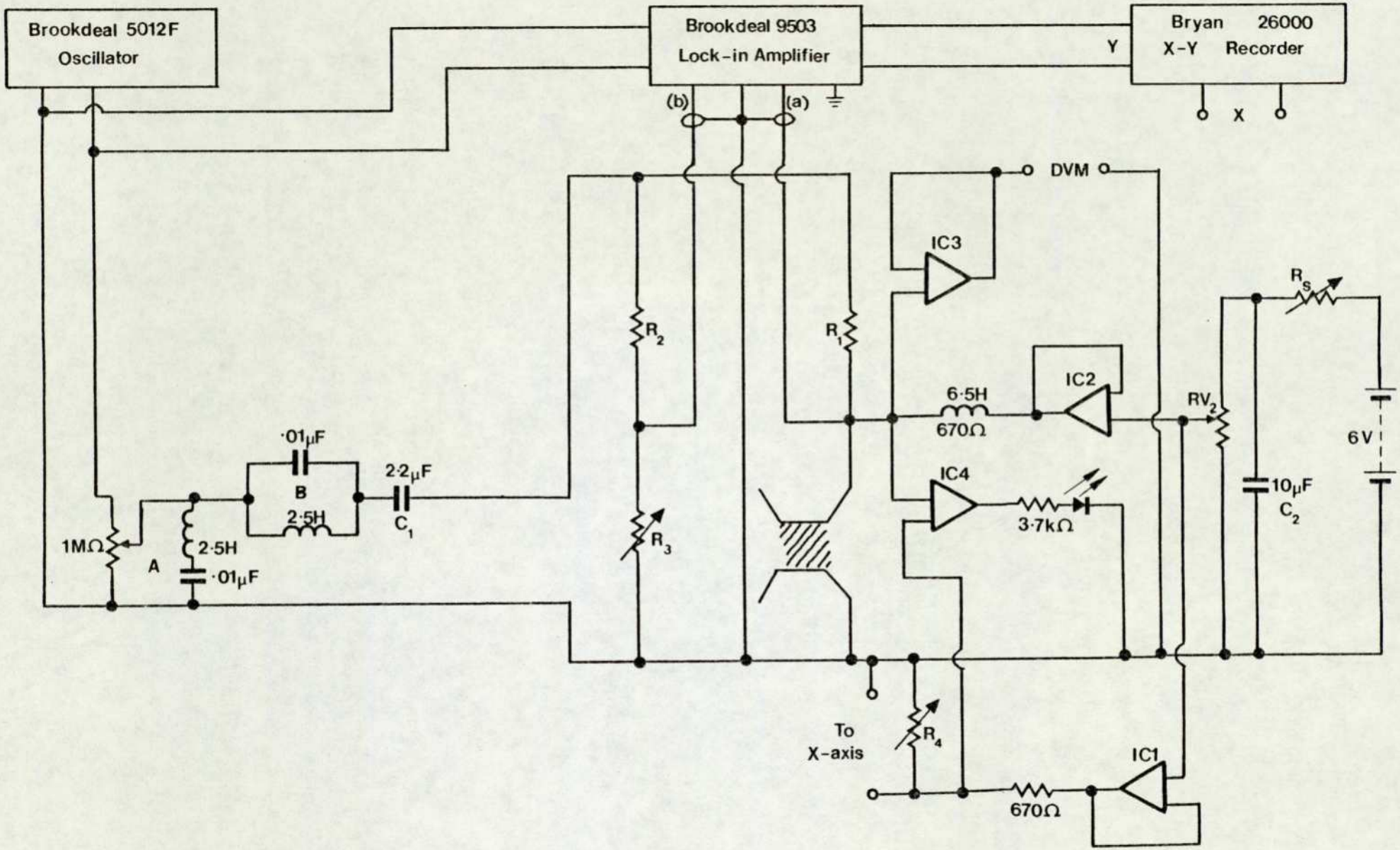


Fig.28. Mk II Bridge and dc bias supply

are in the range 400-1000Ω. However, this combined with the fact that the system produces a logarithmic output of voltage with time i.e. the voltage range 0-100mV scans in approximately the same time as the range 100mV-1V, provided an annoying and frustrating drawback.

The solution to these problems involves the use of integrated circuits. Thus the second circuit, described in Fig. 28 is constructed to remove the inherent problems in the first design[50]. This circuit differs in several ways from that of Fig. 26. Firstly, the initial valve-based chart recorder used in the early part of the study is replaced by the transistor-based Bryans 26000 recorder, having a considerably lower noise output. Secondly, the preamplifier used in the initial circuit is removed. This coincided with the purchase of a Brookdeal 5011F active Filter. The performance of the lock-in is considerably improved by the utilisation of this filter in its notch mode, set at the fundamental frequency prior to amplification. This removes the majority of the fundamental signal, which comprised the largest part of the differential output from the preamp, and combined with the operation of the lock-in itself as a differential amplifier, enables sufficient amplification to be attained without the use of the preamplifier. Finally, a digital voltmeter is included to provide a digital monitor of the dc voltage as the spectrum is recorded.

The changes within the circuitry, however, are more fundamental, and are concerned with the introduction of the four integrated circuits into the dc bias supply. As mentioned previously, these fulfil two distinct purposes. 1C1, 1C2 and 1C3 are used as non-inverting voltage followers. The

noise generated within the X-Y recorder necessitates protection of the junction from any contact with the recorder, and thus 1C1 is used to provide a separate dc bias for the X-axis of the recorder. 1C2 is used to isolate the potentiometer RV_2 from the loading effect of the junction, providing a resistance independent voltage across the combination of the junction and the shunt choke, and thus removing the non-linearity with respect to time from the dc ramp. 1C3 is used, in the same way as 1C1 to prevent noise generated within the DVM reaching the junction. Finally 1C4 is operating as a comparator to ensure that the voltage reaching the X-axis is identical to that across the junction itself. To this end, the 670Ω resistor is used to equalise the voltage drop across the 6.5H choke, and R_4 is a 10kΩ 10-turn potentiometer which derives the bias for the X-axis. The comparator circuit includes a LED to facilitate determination of the point where the value of R_4 is equal to the junction resistance.

This circuit was utilised extensively for a period of around a year, but still suffered from two disadvantages. Although 1C2 isolates the potentiometer from the loading effects of the junction, the choke and junction still behave as a potential divider. Thus the low resistance junctions, typically those below 10Ω, necessitate the operation of the integrated circuit beyond its specification, and instabilities in voltage supply occur in this resistance range. At the same time, the inconvenience involved in the comparator circuit, and the slight errors involved in determining precisely the correct value of R_4 in each case become significant at low junction resistances.

The effect of the choke as a potential divider is simply removed by increasing the modulation voltage fundamental frequency by a factor of ten to 5000Hz. It is experimentally determined that this frequency can be shunted from the dc supply using a choke of inductance 0.3H. This value of inductance can be wound on the same former and core with a resistance of only 8Ω. This removes the effect of the choke in the potential divider, but does not remedy the instability encountered with the integrated circuit at low junction resistances. This problem is ultimately overcome using the final circuit design of Fig. 29.

This circuit overcomes the low junction resistance voltage instability by the incorporation of a series control transistor. This circuit provides a stable dc bias supply across resistances above 1Ω. Integrated circuits 1C1 and 1C2 are again acting as non-inverting voltage followers, protecting the bridge and bias supply from noise generated within the X-Y recorder and DVM. 1C3 is acting as a differential amplifier comparator, in that it compares the voltage provided by the output of the transistor with that provided by the ramping reference supply, and operates the transistor to ensure that these voltages are identical.

Thus the circuit described in Fig. 29 is that which is incorporated in the present design, and is that which has provided most of the spectral information discussed later. As far as can be ascertained by the comparison of the spectra obtained with those previously published in the literature[6,27], it provides a cleaner signal with a lower noise level. This is seen as a result of the considerable pains taken to eliminate

noise sources during the development stage, and is vindication of the time spent over the design and testing of the early circuits.

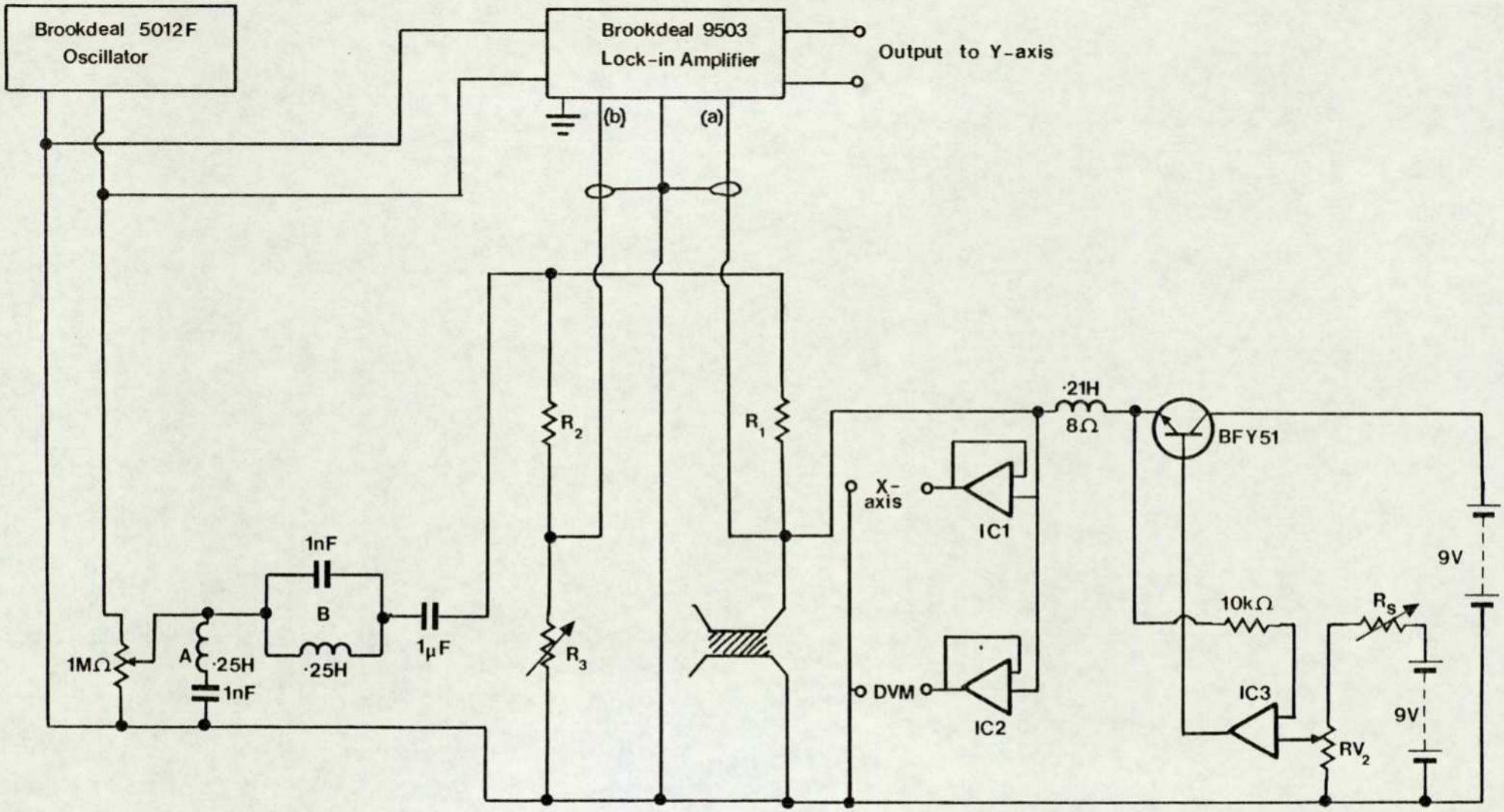


Fig. 29. Final bridge, bias and filter network

Chapter 3. Physical parameters influencing junction performance in IETS

3.1. Introduction

There is at present agreement between all workers in this field on the appearance of the spectrum of an undoped Al-oxide-Pb junction. However, there is a major point of disagreement concerning the nature of the oxide itself, and this will be discussed in this chapter. Several systems other than aluminium-oxide-lead junctions have also been utilised, both with regard to varying the base and upper metal electrodes and also varying the insulator itself. These other systems, although not utilised in the present study, will be briefly reviewed, since they provide information relevant to the overall view of tunnelling spectroscopy.

3.2. The aluminium-oxide-lead junction

The majority of junction structures studied using inelastic electron tunnelling spectroscopy consist of aluminium-oxide-lead strips. This is essentially because the metals themselves are relatively easily deposited; because aluminium forms a compact, thin oxide rather easily; and because a lead overlayer is convenient due to its chemical inertness towards molecular adsorbates on the oxide, its large ionic diameter which retards its diffusion into the oxide under the influence of an applied bias voltage, and its rather high superconducting transition temperature ($7.2^{o}K$). The spectrum of an undoped Al-oxide-Pb junction is shown in Fig. 30.

This result was originally published in the definitive paper by Lambe and Jaklevic[6]. The junction was prepared in a pure oxygen discharge and thus the -OH bending and stretching modes at 120mV and 450mV respectively are due to the presence of water vapour adsorbed on the chamber walls and fittings. Thus two possibilities arise: the water molecules are either introduced during the process of oxide formation, and are thus incorporated within the bulk oxide, or are adsorbed on to the oxide surface before deposition of the second electrode.

Geiger et al[51] assigned the peak at 118mV to OH vibration in alumina hydrate, rather than a free-OH bending mode, and that at 450mV to the stretching mode of free-OH groups. They compared the spectrum of Fig. 30 with that of an Al-oxide-Al junction, which contains only the bending mode at 118mV, and its first harmonic, but shows no evidence of the stretching mode at 450mV, and suggested that this implies that the two bands arise from different sources. The absence of the stretching mode at 450mV was apparent even when twin Pb junctions formed on the same oxidised Al base electrode showed definite free-OH stretching modes at 450mV.

Furthermore, they claimed that information was forthcoming from examination of IR absorption studies of alumina films, quoting the results of Dorsey on alumina films formed by anodising pure Al sheet in different electrolytic solutions[52]. In most electrolytes alumina is slightly soluble and two types of oxide film may form. The first layer, closest to the metal, is called the barrier layer and is very compact, with good dielectric properties. The second type, which may be

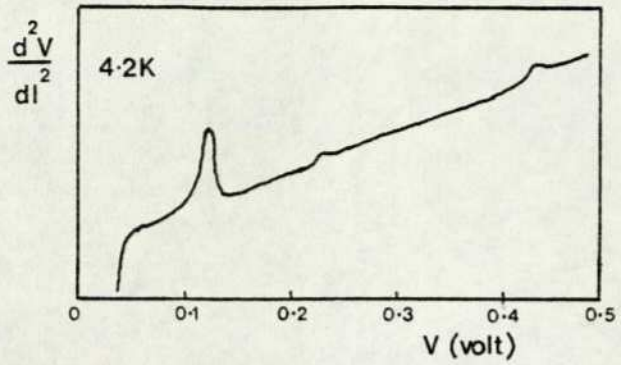


Fig. 30. Tunnelling spectrum of an undoped Al-oxide-Pb junction at 4.2K

formed on top of the barrier layer, is a porous layer, with poorer dielectric properties made by the simultaneous building up the oxide in some parts of the film and its dissolution in other parts. Geiger et al assumed that the type of oxide formed with these electrolytes was similar to the film formed in the gaseous anodisation process, except that the acidically anodised films may be more completely hydrated.

The reflectance spectrum of the barrier layer films has only one absorption band in the energy range 4-500mV: at 955cm^{-1} (118mV), exactly the location of the prominent peak seen in Al-Al and Al-Pb junctions. They claim that this similarity between the tunnelling spectrum of Al-Al junctions and the IR spectrum of barrier-layer alumina supports the hypothesis that in Al-Al junctions the tunnelling spectrum is displaying the vibrational modes of the insulating oxide itself, and these modes only. In Al-Pb junctions, the oxide modes are again present, but in addition vibrational modes of molecular impurities not associated with the barrier structure appear. By comparison with IR spectra of known mineral forms, Dorsey identifies the 118mV peak as an OH bending mode of an alumina hydrate, specifically a trihydrate. The existence of only this bending mode presents an anomaly, because alumina hydrates

normally have an OH stretching band near 420mV. The absence of this band in the barrier-layer alumina is explained by the existence of strong hydrogen bonding, which reduces the intensity of such a band by broadening it[53].

The absence of vibrational modes in the Al-Al junction for free-OH stretching vibrations at 450mV can be explained by the atomic size of the top electrode. As the top electrode is evaporated on to the oxide, the penetration of the metal into the oxide is governed by the atomic radius of the metal. For a metal of large radius such as Au or Pb, there is very little penetration so that the M - oxide interface is very close to the initial surface of the oxide. In this case, most of the molecular impurities on the oxide surface are present in the tunnelling region of the junction so they contribute to inelastic tunnelling. For a metal of small radius like Al or Mg, there is considerable penetration of metal into the oxide so that the M - oxide interface for tunnelling is in the interior of the original oxide film. In such cases, the surface impurities are not in the tunnelling region of the junction, and only oxide vibrations contribute to inelastic tunnelling.

With consideration of these results, Gieger et al proposed that the oxide consists of a polymeric, hydrogen-bonded hydrate, of the type shown in Fig. 31.

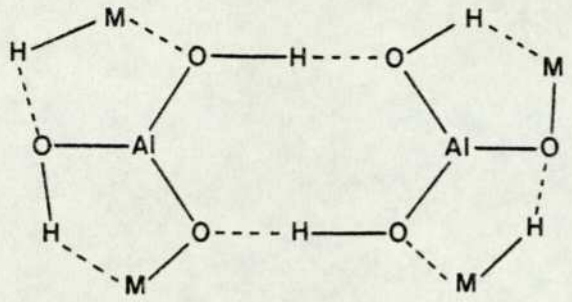


Fig. 31. Polymeric, hydrogen-bonded alumina hydrate

Fig. 31 represents a unit section of a long polymer, and M represents an identical unit within the same chain. This, then, is the evidence leading to consideration of the tunnelling oxide layer as an alumina hydrate, and this was the accepted solution in the early stages of the development of tunnelling spectroscopy.

In 1974, an alternative explanation was proposed by Lewis et al[54]. They proposed that the intense peak at 118mV could be assigned to stretching Al-O vibrational frequencies, as opposed to bending O-H frequencies. This would therefore explain the presence of this peak in an anhydrous oxide where penetration of the upper electrode occurs to an extent where surface adsorbed species, particularly water, are no longer present within the tunnelling region i.e. in the Al-oxide-Al junctions. At the same time, surface adsorbed water molecules would give rise to the peak at 450mV in the case where little or no top-electrode penetration occurs, i.e. in the Al-oxide-Pb junctions. Thus they conclude that the barrier is in fact anhydrous aluminium oxide, and not the hydrate as proposed by Gieger et al.

In a later study, Bowser and Weinberg[55] produced results supporting those of Lewis et al. They formed the insulator in discharges of three different gases, O_2, H_2O and D_2O at various

temperatures. The temperature was controlled by resistive heating of the aluminium strip over a range from room temperature to around 200°C. All samples prepared in O_2 plasma discharges at room temperature, 60°C and 116°C yielded tunnelling spectra essentially identical to the spectra previously published for clean tunnel junctions fabricated at room temperature. The variation in spectra with temperature was found to be considerably more marked when H_2O was used as the oxidising gas. Those junctions prepared in H_2O at low temperatures ($\lesssim$100°C) had spectra quite different from those junctions prepared in O_2, the spectra having a marked dependence on the temperature at which the oxide was formed. The spectra of those samples prepared in H_2O at higher temperatures, however, are identical to the spectra of samples prepared in O_2. Furthermore, junctions prepared with D_2O as the oxidising gas produced very similar results to those of H_2O, with an additional peak at 334mV due to the O-D stretching mode.

Bowser and Weinberg pointed out that if the peak at 117mV was associated with a bending mode of a hydrated oxide, then there would be a peak at 83mV for any sample prepared in D_2O. This was not found to be evident. They further argued that it is unlikely that aluminium reacts exclusively with H_2O in forming a bulk hydroxide while still absorbing deuteroxyl groups on the surface. These workers therefore argued against the inclusion of hydroxyl groups within the bulk oxide and proposed, in agreement with Lewis et al, that the insulating barrier consisted of anhydrous aluminium oxide.

Thus clearly a dichotomy occurs in the interpretation of

the experimental results with regard to the exact nature of the insulating substrate for oxidised aluminium.

3.3. Barriers other than oxide formed on aluminium.

In a study by Magno and Adler[56], IETS was used to study the molecular structure of barriers formed in a glow discharge of ethylene or benzene. These barriers were found to contain many molecular groups which may be associated with substituted forms of benzene and ethylene, together with many which must arise from breaking of these molecules. This established inelastic electron tunnelling spectroscopy as a tool for studying materials produced by glow discharge polymerisation. Further than this, however, it served to indicate that the barriers utilised in tunnelling spectroscopy are by no means restricted to oxides.

This is an important observation in relation to the use of IETS in the study of electronic transitions. The lowest energy electronic transitions commonly recognised are those of the cyanines and phthalocyanines[57], occurring at between 1.2V and 1.8V (9,600 cm^{-1} to~ 14,000 cm^{-1}). However, most common transitions occur at potentials considerably in excess of these levels. Aluminium oxide breaks down at potentials around 1.5V, and suffers from very high noise levels in the region above 1.3V, and thus is only of limited use in the study of electronic transitions. Furthermore, the dielectrics formed by the discharges in benzene and ethylene have too great a background absorption level to be useful. Thus any investigation of other than the lowest level electronic transitions requires the development of barriers of higher dielectric strength.

3.4. The influence of the top electrode

The influence of the top metal electrode on the tunnelling spectra of Al-Al oxide-metal junctions was first investigated systematically by Geiger et al[51]. They prepared a series of junctions where the base electrode and insulating barrier were identical, so that differences in the tunnelling characteristics depended solely on the properties of the top electrode. The metals used as top electrodes were gold, silver, indium, tin, aluminium, magnesium and lead. By equating the intensity of the 118mV oxide peak, an order of relative intensities of molecular vibrations was obtained, and this was found to be

$$Au > Ag > Pb > In > Sn > Mg \approx Al$$

In magnesium and aluminium the impurity bands are not detectable at all.

If a single property of the upper electrode is determining the differences present in the characteristics of the junctions, the variations in tunnelling characteristics should be correlated with variation of this property. Geiger correlated these observed variations with variation in the ionic radius of the metal, as given by Pauling[58]. This radius decreases in the same order that is seen in the size of the impurity bands relative to the oxide band in the tunnelling spectra of junctions whose top electrodes are composed of these metals. This correlation suggests that the controlling mechanism which accounts for the spectral variations may be the penetration or diffusion of the metal atoms into the oxide film. A similar type of correlation between atomic radius and junction resistance has been explained by this mechanism[59].

Kirtley and Hansma[60] have conducted experiments which investigate the bandwidth and peak position, as opposed to intensity, of impurity bands contained in junctions with varying upper electrodes. They considered the stretching modes of hydroxyl and deuteroxyl ions adsorbed on aluminium and magnesium oxide and studied the shifts and broadening due to four different upper electrodes : Pb, Sn, Ag and Au. The shift across this series was found to be≈ 3% while the broadening was≈ 50%. With the hydroxyl ions in close proximity to the top metal electrode, two mechanisms for the shifting and broadening of the hydroxyl stretch mode are apparent. The first, and the simpler, is hydrogen bonding between the hydroxyl ion and the metal atoms. Such a bond would allow the metal to lower the electron density in the O-H bond, decreasing the bond strength and thus the frequency of the OH stretching mode. This accounts for the shift occurring, but it is difficult to correlate the degree of shift meaningfully with any property of the particular upper electrode.

Therefore Kirtley and Hansma postulated a second mechanism for producing shifts and broadening involving an image-dipole effect. An oscillating charge outside a plane metal surface induces an image charge inside the metal. The attractive interaction potential between the charge and its image is proportional to 1/d, where d is the separation between the charge and the surface of the metal. This attractive potential shifts the resonance frequency down. They explain the broadening of the impurity bands by assuming that the hydroxyl ions occupy non-equivalent sites on the surface of the oxide,

leading to a distribution of values of d and thus a larger shift for some ions than for others. The upper electrodes that produce larger overall peak shifts can be expected to have larger distributions in peak shifts and thus larger bandwidths.

This study also compared the shifts observed with various upper electrodes to the values of adsorbed molecules on aluminium oxide with no covering electrode obtained using I.R. spectroscopy. Again considering the O-H stretching mode, for the series lead, tin, silver and gold the energies are 446, 444, 438 and 433mV respectively. With no top metal electrode, the energy is approximately 445±3mV for various aluminas. Thus the tunnelling results with a lead top electrode are the least perturbed with a shift of order 2%. This value is large enough to cause considerable difficulty in a comparison of vibrational energies with those of conventional vibrational spectroscopy, in that it would be difficult to correct for shifts as large as 2% in the region below 200meV where typical organic molecules have a large number of modes.

Fortunately the mode shifts appear to be much smaller for these modes. The modes below 200mV can be divided into two general classes: those which involve C-H bending motion and those that don't. Considering first the group which do involve C-H bending, the major different from the standpoint of the image dipole theory between the C-H and O-H modes is that the C-H dipole derivative is approximately three times smaller[61,62]. The image dipole effect scales as the square of the dipole derivative, and thus shifts would be expected to be of the order of tenths of a percent rather than several

percent. The spectrum of benzoic acid, when compared with the infrared spectrum shows a shift of less than 0.1%[63], which verifies the theoretical prediction.

Secondly, consider the modes which do not involve CH motion, for example ring deformations. The major difference between these modes and the C-H and O-H modes is that d should be larger. Assuming that the value of d was thus increased by around 1Å (the C-H bond length), the $1/d^3$ dependence of the mode shifts would reduce these shifts by a factor of roughly 0.1. The ratio of these shifts to the O-H shifts would be of order 5×10^{-4}, leading to shifts less than 0.01mV, well below the resolution of around 0.2mV. Hansma and Kirtley have observed no shifts to within this resolution for these modes. Thus shifts produced by the presence of lead as an upper electrode can be discounted, and direct comparison may be obtained between vibrational frequencies in IETS and I.R./Raman techniques.

3.5. The use of alternative base electrodes

As mentioned previously, the vast majority of work in IETS to date has been carried out using aluminium-oxide-lead junctions. This owes as much to the historical importance of this junction as the tool for Jaklevic and Lambe's initial discovery as to the ease of preparation. There have, however, been some studies carried out involving junctions prepared on alternative base electrodes, and of these the work of a group at the University of Paris has been the most extensive. The initial interest of Klein, Léger, Belin and Defourneau centred upon the portion of the tunnelling spectrum which supplies information relevant to phonon vibrations within the

metal electrodes[10].

Crystalline solids are, in general, good conductors of sound. In a gas, sound travels as a longitudinal wave of compression. In a solid, sound waves can be carried by vibrations of the ions or atoms. These vibrations can be divided into two groups: acoustical vibrations, which behave like sound waves and are therefore called phonons in the crystal, and optical vibrations, which behave like light waves and are called optical phonons. The basic distinction is that acoustical vibrations do not cause an oscillating electric field - they are mechanical but not electrical vibrations. Optical vibrations, however, are accompanied by oscillating electric fields, and thus are susceptible to excitation by the tunnelling electron. These phonon vibrations occur at the low energy end of the tunnelling spectrum, predominantly in the region from zero bias to 40mV.

Clearly the frequency of these vibrations is dependent upon the nature of the solid in which they are occurring, and thus after Lambe and Jaklevic had explored the use of IETS as a method for studying phonon vibrations in their aluminium based junctions[64], Klein et al[10] set about observing corresponding vibrations in an alternative system, based on magnesium as the lower electrode. They observed both longitudinal and transverse phonons within the magnesium electrode, and correlated the phonon density deduced from their results with that deduced from neutron scattering experiments. Furthermore, they found the same pattern of dopant vibrational transitions in the range from 100 to 500mV as had been observed for aluminium based junctions.

Subsequently, however, their interests diversified to

include the extension of IETS to the electronic transition region, and they became pioneers in the observation of low energy electronic transitions using tunnelling spectroscopy[65]. As briefly mentioned earlier, this extension proved viable for only those molecules having low energy transition because both aluminium and magnesium oxides break down at values between 2 and 3 volts. Thus in order to provide dielectrics with a higher break-down volume, Klein et al began producing junctions based on different lower electrodes. They failed to find any combinations which would significantly extend the voltage range, but succeeded in observing low energy transitions in the oxides of rare earth metals by producing Er-ErOx-Pb and Ho-HoOx-Pb junctions[66].

Thus various other combinations of junctions have been explored outside the scope of the standard AL-Al_2O_3-Pb system. However, although the results obtained have been interesting with regard to the properties of the lower electrode and its associated barrier, little or no effect has been observed upon the region of the spectrum associated with vibrations of the dopant molecules.

3.6. The sensitivity of inelastic electron tunnelling spectroscopy

The most important advantage of inelastic electron tunnelling over infrared and Raman spectroscopy is its sensitivity; electrons couple more strongly to molecular vibrations than do photons. Workers in the field ever since Jaklevic and Lambe have speculated that the dopant layers involved in IETS are of the order of one monolayer, and there is a good deal of indirect evidence for this. Specifically, many compounds react with the oxide layer to form new surface species. The tunnelling spectra

for these compounds usually show no trace of the unreacted compound, and thus it may be inferred that only a monolayer or less is present in the junction.

In 1975, Langan and Hansma performed quantitative measurements using benzoic acid as dopant which confirm this qualitative conclusion and yield a good deal of additional information[67]. They observed a trend of decreasing peak height with decreasing concentration of dopant solution, as expected, and moreover found that peak height rises roughly linearly with dopant solution concentration up to a plateau for solution concentrations greater than 0.2mg/ml. They also measured surface concentration as a function of dopant solution concentration using solutions of radioactively labelled benzoic acid and measuring their activity with a scintillation counter. The conversion between activity and surface concentration was then effected by a calibration run in which measured volumes of benzoic acid solutions of known activity were dropped on to glass substrates and evaporated to dryness.

Again they observed that the surface concentration also rises roughly linearly to a plateau which occurs at roughly 7×10^{14} molecules cm^{-2}, about 1 molecule per $15A^2$. The results of this study indicate that coverages down to the order of 1/30 of a monolayer can be detected with inelastic electron tunnelling. It is this extreme sensitivity that is of key importance to many of the applications of this technique which will be discussed in Chapter 5.

They also observed that peak height decreases more rapidly than the surface concentration as solutions fall

below saturation coverage. For example, as the surface concentration decreases by a factor of 10 from 5×10^{14} to 5×10^{13} molecules cm^{-2}, the peak height decreases by a factor of 40. They provide a simple qualitative explanation for this effect. Junctions doped with a monolayer of the benzoate ion typically have resistances two orders of magnitude higher than undoped junctions. Thus the tunnelling conductance through the doped insulating layer is much lower than through the insulating layer alone. As the surface concentration falls below a monolayer, microscopic regions in which the upper electrode is in direct contact with the insulating layer develop. Since the tunnelling conductance per unit area is higher for these regions, the tunnelling current density will be higher in them. Thus the fraction of current flowing through the insulating layer plus benzoate ion will fall more rapidly than the fraction of the surface covered by the ion. Assuming that the fraction of the current flowing through the insulating layer plus the benzoate ion is primarily responsible for inelastic excitation of the benzoate ion, the fraction of inelastic tunnelling events and thus the measured peak height will fall more rapidly than the surface coverage.

3.7. Experimental evidence for selection rules

The experimental data on selection rules is scanty and easy to summarise: there is no evidence for any strong selection rules except for orientation selection rules. This is in accordance with the theoretical predictions of section 1.2.3. The first attempt to investigate whether both infrared and Raman active modes appear in tunnelling spectra was made by Bogatina, Yanson, Verkin and Batrak[68]. As part of their

investigation of the spectra of organic solvents they examined the spectrum of benzene in Pb-oxide-Pb junctions. Bogatina analysed this spectrum using group theory[69].

This was the first molecule studied which had sufficient symmetry to have separate Raman and infrared modes since in molecules with insufficient symmetry both the dipole moment and the polarisability oscillate during its vibrations and thus there are no vibrational modes that are only Raman or only infrared active. The spectrum was puzzling since many vibrational modes were not present, particularly between 200 and 300mV. Nevertheless, they found no evidence that infrared active modes differed in intensity from Raman active modes in any systematic way.

Also in 1974, Simonsen,Coleman and Hansma[28] obtained the tunnelling spectrum of anthracene, another molecule having a centre of inversion symmetry. Both infrared and Raman active vibrational modes were found to be present, again with comparable intensities. However, certain modes were absent as in the benzene spectrum.

The accepted explanation for these missing bands was proposed in 1976 by Kirtley, Scalapino and Hansma[18]. They used the theory of Scalapino and Marcus to predict that there should be no IETS intensity for a vibrational mode which results in an oscillating dipole moment parallel to the metal surface. These workers provided experimental support for the presence of this selection rule by consideration of the tunnelling spectrum of benzoic acid on alumina. They observed one band at 1560cm^{-1} in the infrared which is markedly smaller in the tunnelling spectrum, as shown in Fig.32.

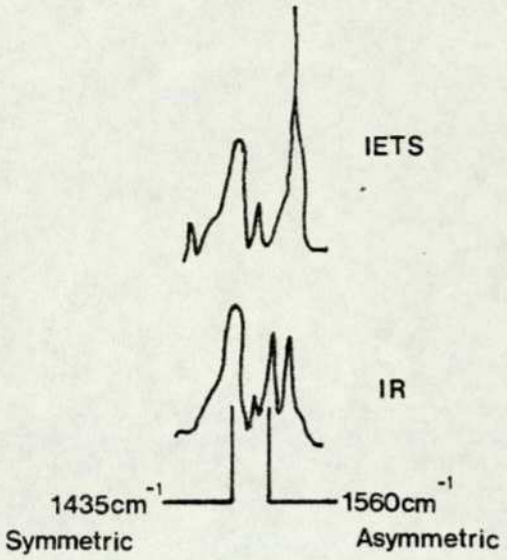


Fig. 32. Comparison of the tunnelling and infrared spectra of benzoic acid on alumina

Kirtley et al[18] attributed the missing band to the asymmetric stretching mode of the carboxylate group. Benzoic acid is thought to absorb on alumina as a benzoate anion with the two oxygen atoms in the group in equivalent sites, equidistant from the oxide surface. The resultant COO^- group has a symmetric and asymmetric vibrational mode, which have their oscillating dipole moments at right angles to each other. Based on the above assignment of orientation, the symmetric mode has its oscillating dipole moment perpendicular to the oxide surface while the asymmetric mode has its oscillating dipole moment parallel to the oxide surface. Thus they propose that the asymmetric mode should have a much lower intensity than the symmetric mode, an observation confirmed by the appropriate spectrum.

3.8. Opposite voltage bias asymmetries of IETS intensities

Kirtley, Scalapino and Hansma[18] also obtained plots of

second harmonic voltage signal as a function of bias voltage when junctions were biased in the two opposite polarities of voltage. All curves were obtained at 4.2K and 2mV, and they observed that the inelastic electron tunnelling response was larger when the electrons tunnel from the aluminium to the lead (i.e. aluminium biased negative) than vice versa.

The physical reason for the asymmetry in inelastic electron tunnelling intensities is that the doped molecules are closest to the lead surface. When electrons tunnel from aluminium to lead, they cross more of the barrier before losing energy to a vibrational mode of the impurity than when they tunnel in the opposite direction. Thus since electrons with higher energy are more likely to cross the barrier, it is expected that the intensities of inelastic electron tunnelling spectroscopy should be greater for aluminium negative with respect to lead than vice versa.

Chapter 4. Results and discussion

4.1. Introduction

The original objective of this study to investigate the chemical, analytical and spectroscopic potential of IETS has been only partially achieved due to difficulties experienced in the construction of the spectrometer and problems associated with defining the preparation of the junction and methods of doping. These parameters were not defined by the previous workers in this field. Nevertheless, in the later stages of the study the factors affecting junction performance and preparation have been evaluated and have produced several interesting phenomena. In addition, spectra of high resolution have been obtained for a variety of molecules with a noise level consistently lower than that present in other published spectra. These results are discussed in this chapter.

4.2. Obtaining spectra

The detailed effects of varying oxidation and doping parameters will be discussed in sections 4.3 and 4.4 respectively. In this section, the methods used to obtain a spectrum from a completed junction, together with the effects of varying the instrumental parameters will be discussed. The junctions are initially prepared as described in section 2.2.8. Briefly, this involves evaporation of the aluminium, oxidation under controlled conditions, doping and evaporation of the lead upper electrode.

Once the junction has been prepared and removed from the vacuum system, resistance measurements are taken to ensure

the integrity of the metal electrode strips, and to ascertain the resistance of the junction itself. Since the spectrometer incorporates a bridge which uses the junction as a two terminal device, it is imperative that the strip resistances are zero, since finite electrode resistances lead to a volt drop across the strips themselves, thus decreasing the bias potential across the junction relative to that appearing on the X-axis of the recorder. This leads to an apparent shift of the peaks to higher energies. In practice, no volt drop has been found to occur across the lead electrode, but occasionally the aluminium electrode has been found to have a significant resistance. This, however, has been localised to slight strip discontinuities at the point where the aluminium crosses the edge of the raised copper pad, and in these rare cases, this can be remedied by the application of a small quantity of conducting silver paint at this point. By deliberate application to junctions which already have aluminium strips of zero resistance, the presence of this silver paint has been found not to affect the subsequently obtained spectra.

Having produced junctions with zero resistance base and upper electrodes, the junction resistance is itself obtained. It is imperative to utilise a resistance measuring device with a low current flow so as not to impair the integrity of the oxide barrier, causing shorts and rendering the junction useless. The required resistance range has been found to be between 1Ω and 10Ω for undoped junction, and between around 50Ω and 1500Ω for doped junctions. It is possible to record spectra at resistances outside this range, especially at higher resistances. These spectra, however, are of inferior resolution, with

considerably higher noise levels, and, at resistance around 3KΩ, with noise levels comparable to the signal level, thus producing extraneous peaks in the spectrum.

When a junction is produced which appears, on initial inspection, to be in the correct resistance range to enable a spectrum to be obtained, it is attached to the bridge arrangement and immersed in a storage dewar of liquid nitrogen. The spectrum is always obtained at 77K prior to immersion in liquid helium for three reasons. Firstly, any junction which produces an envelope in the spectrum at 77K which is noise free will produce a reasonable spectrum at 4.2K, and thus this initial check prevents immersion of unusable junctions in the liquid helium, with its associated boil-off. Liquid nitrogen is readily available within the department, while liquid helium is an expensive commodity and must be purchased externally.

The second reason for prior immersion in liquid nitrogen is that this greatly reduces the temperature of the junction and bridge arm before immersion in the liquid helium. This again reduces considerably the amount of boil-off on helium immersion, and increases the usage of a 17 litre dewar by a factor of a half, relative to immersion of junctions at room temperature. Finally, in some junctions an effect termed "boosting" is seen to occur. This is discussed more fully in section 4.3, but it is important to note that this boosting is considerably more efficient at liquid nitrogen temperature, than at liquid helium. This is presumed to be an ohmic effect governed by current density across the insulator, which depends upon its resistance.

Once the junction is immersed in the liquid nitrogen dewar, it is transferred to the Faraday cage and incorporated in the spectrometer electronics. The various controls on the lock-in amplifier must then be adjusted to produce the spectrum. Since the lock-in settings can have a considerable bearing on the quality of the spectrum produced, it is worth discussing these at some length. As discussed in section 2.4.3, the junction is incorporated as one arm of an a.c. Wheatstone bridge, with two identical 33KΩ resistors and an adjustable potentiometer to balance the bridge. When the bridge is in balance, the lock-in is used in its differential mode to subtract the signal across the balance potentiometer from that across the junction and amplify the remainder.

When the spectrum is obtained, the lock-in is amplifying the second harmonic of the fundamental modulation frequency, and considers all other signals as noise. Thus, since the fundamental frequency is around four orders of magnitude more intense than the second harmonics generated within the junction, as much of the fundamental as possible must be removed before amplification. Thus the electronics are turned on, the d.c. bias is set to zero, the modulation voltage is set to 2mV, the sensitivity is set to 1mV and the bridge is balanced by adjusting the balance potentiometer until the lock-in meter shows zero output. The sensitivity is then further increased by a factor of ten and fine adjustments are made to the potentiometer to ensure an accurately determined null point.

At this point, the 2f pushbutton on the lock-in is selected, which operates an interior frequency doubler in the reference channel so that the frequency sensitive detector is

now amplifying the second harmonic. Since harmonics are only generated whenever non-linearities appear in the junction current-voltage characteristic, a voltage where this occurs is applied to the junction. In Al-oxide-Pb systems, an intense peak is always obtained for the aluminium phonons at 40mV, and this has been used as a reference for setting up the system. Thus 40mV is applied to the junction, and the sensitivity is adjusted to the microvolt range in order to provide a deflection on the output meter.

A lock-in amplifier, in addition to being frequency-selective, also incorporates a phase-sensitive detector to minimise interference from other phase-shifted signals. Thus the phase of the reference voltage must be adjusted until it is in phase with the signal in order to maximise the response. This is achieved by adjusting the phase control until the signal is reduced to zero output (i.e. until the signal and reference voltages are exactly 90° out of phase), and then altering the reference phase by 90° using the appropriate push-button control.

The lock-in is now set to produce the maximum second harmonic frequency output. It is customary in tunnelling spectroscopy to produce spectra where increasing intensity is plotted in the positive Y-direction, and thus the phase may now be adjusted by 180° to ensure that the 40mV phonon absorption peaks up as opposed to down. This is accomplished by slowly ramping the dc through the 40mV range, observing the direction of movement of the output meter, and changing phase by 180° if necessary using the appropriate push-button control.

The lock-in is now set to produce the spectrum. The final adjustment of sensitivity is achieved to maximise peak height, and the range is set using the output-zero adjustment so that the range from zero to 500mV always appears on the output meter without any indication of noise or signal overload. These adjustments must always be carried out as described in order to maximise signal strength, and thus intensity and resolution of the spectrum.

The final settings are a function of each individual junction, and are determined solely with consideration of the noise characteristics of the input signal to the lock-in from the bridge. The optimum response time of the output detector system, in this case the X-Y recorder, is 100ms. Thus if the signal is noise-free, the lock-in time constant would be set to 100ms to obtain optimum scan-time conditions. However, there is always an inherent noise level generated within the junction, and the ratio of this level to the signal level governs the time constant selection.

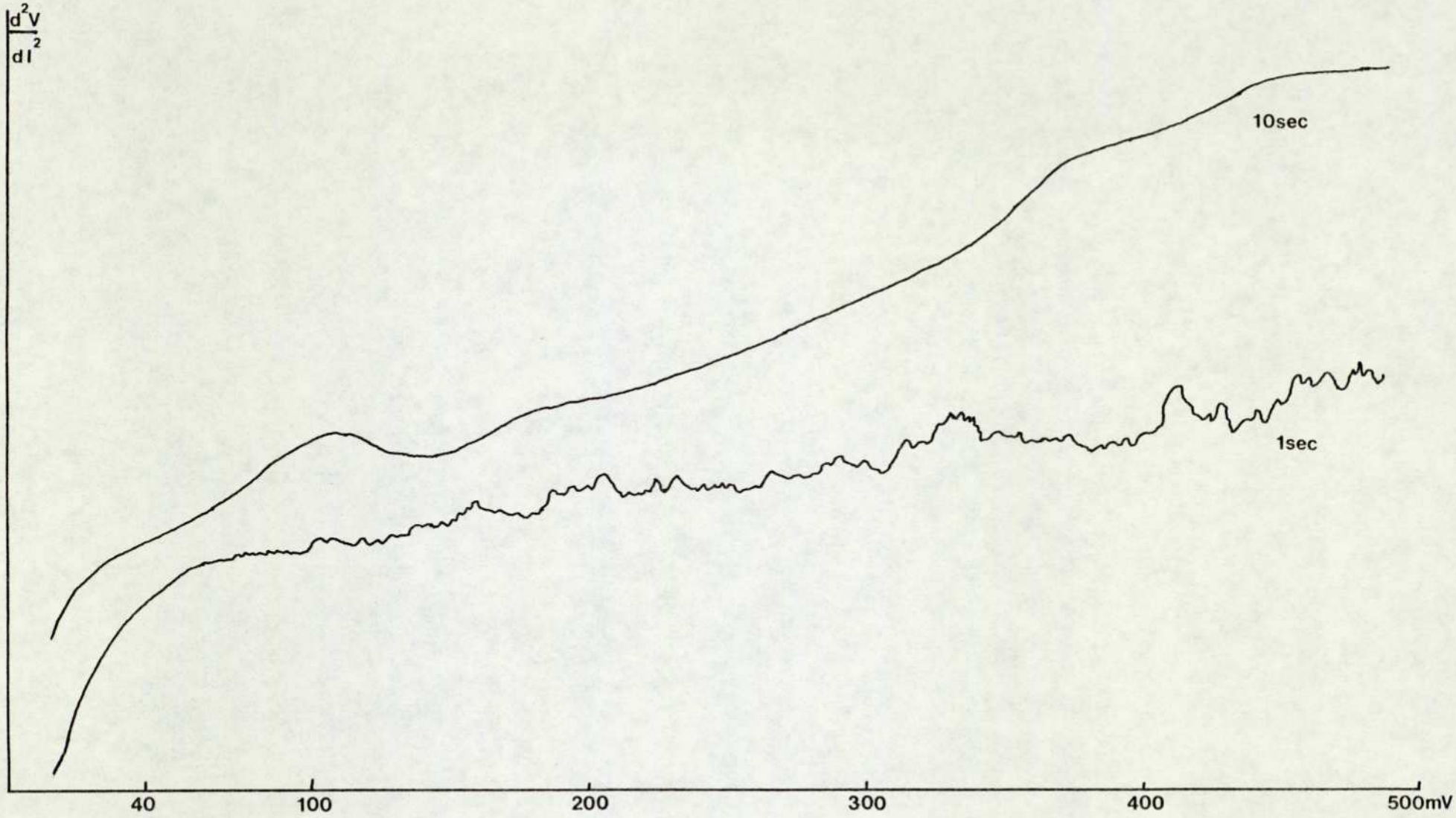


Fig.33. The effect of time constant. Tunnelling spectra at 77K of tapwater at time constants of 10 sec. (upper trace) and 1 sec. (lower trace)

The effect of time constant on noise levels within the signal is demonstrated by Fig.33. This represents one of the early junctions made during the course of this study, Al-oxide-Pb based, doped with tapwater and run at 77K. The spectrum was obtained using a 5mV modulation voltage. The lower trace was plotted with a 5 minute scan time and a 0.1 sec. time constant, while the upper trace was plotted using the same scan time but a 10 sec. time constant. It is apparent that a considerable increase in resolution has occurred in the latter case. This is because the time constant setting of the lock-in directly affects the response time of the phase-sensitive detector. Since the noise detected on tunnelling spectroscopy is generally high-frequency shot noise, the detector will ignore this if the time constant setting is high, provided that the intensity of the noise level is of the same order of magnitude as the signal intensity.

Increasing the time constant without changing the scan speed, however, leads to shifts within the spectrum, since the lock-in amplifier output lags behind the d.c. ramp supplied to the X-axis of the recorder. Thus, the higher the time constant, the longer must be scan time. However, this increase in scan time itself generates more noise since the lock-in amplifier has longer exposure to any given d.c. value. Thus there is a trade off in noise levels between the decreased sensitivity of the lock-in detection system and the increased noise level inherent in a slower scan speed. The optimum values for junctions which fall within the resistance range 50-1500Ω have been found to be a 10 sec time constant and a

40-minute scan time, and it is these settings which have been used to produce the spectra discussed in this chapter.

A further trade-off which is noticeable at this stage is that between peak intensity and peak width, relative to the value of the modulation voltage. As discussed in section 1.2.2., the two major contributors to peak width in tunnelling spectroscopy are thermal broadening and modulation broadening. Now at 4.2K, modulation broadening is significant, and the value of 2mV used as the modulation voltage in the course of this study is the minimum voltage which will produce spectra of sufficient peak intensity. At 77K, however, the modulation broadening is insignificant by comparison with temperature effects which alone give a peak width at half peak height of 36mV (c.f. 2mV at 4.2K, 138mV at 298K).

Fig. 34. The effect of modulation voltage. Tunnelling spectra at 77K of benzoic acid in water, at modulation voltages of 2mV (upper trace) and 5mV (lower trace)

Thus at 77K the effect of modulation voltage can be observed without consideration of broadening effects. The value of the modulation voltage broadening is given by $1.7\Delta V$ where ΔV is the root-mean-squared modulation voltage. Thus at 77K, for a 2mV modulation voltage, the broadening is only 3.4mV, compared with the 36mV attributed to thermal effects. The effect of increasing the modulation voltage from 2mV to 5mV at 77K is shown in Fig. 34, which represents another early junction, Al-oxide-Pb based and doped with a dilute solution of benzoic acid in water (approximately 1mg/ml).

As can be seen, the peak intensity improves significantly, and the noise level again decreases, this time because its ratio to the signal level has been decreased. The resolution of peaks has improved, although the peaks at low energy, which were separately resolved at 2mV have become less so due to overlap because of increased broadening effects.

Finally in this section, because of the broadening associated with temperature rises, it is very important to be able to predict when the supply of liquid helium has evaporated to an extent when the temperature rises above 4.2K. In many cases this is straightforward, in that the superconducting band gap of the lead electrode disappears above 5.8K, and the non-appearance of this band gap can therefore act as an indication that the temperatures have risen above the required level. This is not totally reliable however, because junctions which are too thick to produce meaningful tunnelling spectra do not produce this band gap.

Furthermore, this method gives no quantitative measure of temperature. The immediately obvious solution to the

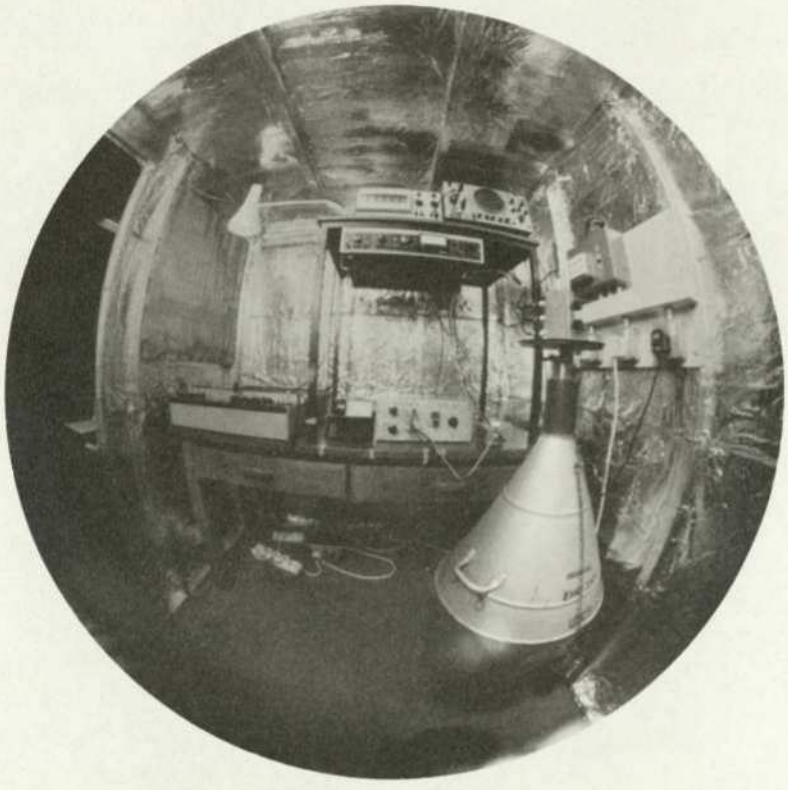


Plate 5. The spectrometer

in operation

problem is to utilise a thermocouple as a temperature probe. However, most commercially available thermocouples only operate down to liquid nitrogen temperatures, the exception being a gold-chromel alloy thermocouple which was prohibitively expensive. However, a far simpler and cheaper solution was eventually obtained by simply placing an OA202 diode in the liquid helium. This particular diode has a general temperature dependent operation, with a significant increase (commonly from around 3mV/K to 80mV/K) below 5K, and provides an excellent and immediate response to increasing temperature. Thus by using double-sided circuit board for the probe which is inserted into the dewar, and affixing this diode to the side which is not used for junction measurement, an immediate quantitative measurement of temperature was obtained.

4.3. Parameters affecting oxidation of the aluminium electrode

In general, the parameters which determine the resistance of a junction can be subdivided into those which are a function of the insulator, in this case an oxide of aluminium, and those which are a function of the dopant. Early in the study, it became apparent that the combination of these factors produces an almost insoluble problem, and it became imperative to find a dopant which would dope consistently as a pure liquid, thus removing any solution concentration and solvent effects. Having found such a dopant, the correct oxidation parameters may be determined with fewer variables, and once these parameters have been assigned values, these should be applicable to other dopant systems.

Initially, a considerable amount of time was devoted to the production of undoped junctions, in the hope that the correct

doping procedures could be introduced after the oxidation conditions had been established. This ideal solution to what has proved a very difficult problem has been found to be impracticable for two reasons.

Firstly, because of the extremely low resistance of undoped junctions, the electronics at their present level of sophistication would not allow the determination of spectra. Because the d.c. ramping supply contains voltage followers and transistors to remove loading effects, no problem occurred in supplying the junction with the necessary level of d.c. However, voltage followers in a.c. circuity invariably introduce signal distortion and harmonic generation, which would make any measurements of second harmonic ambiguous. The maximum a.c. voltage obtainable across an undoped junction without the inclusion of voltage followers was of the order of 0.1mV,much too low to enable spectra to be obtained. The problem of loading effects on the a.c. circuitry is one which is still receiving attention.

Secondly, the resistance of undoped junctions varies very little with increasing oxide thickness. The type of plot obtained is shown in Fig.35.

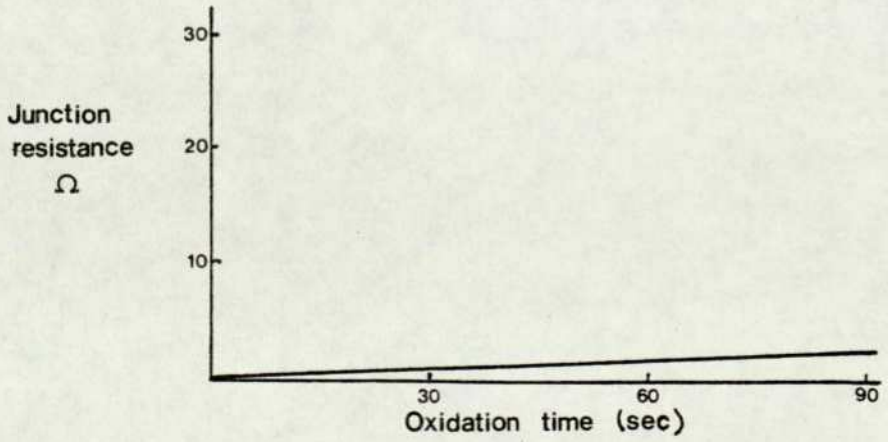


Fig. 35 variation of junction resistance with oxidation time at constant oxygen pressure and discharge current.

Very little variation is noted over a range of oxidation times varying from tens of seconds to tens of minutes. Thus junctions were prepared by varying oxidation conditions and doping with benzaldehyde, until a spectrum of recognisable shape was produced at liquid nitrogen temperatures. The duration of this work was seriously extended by two main problems, both of which were features of the vacuum preparation chamber.

The early junctions were prepared using oxygen supplies which were sealed in 1 litre glass flasks and were very prone to contamination since they were not at a positive head pressure relative to atmospheric. At the same time, contamination through leaks within the vacuum system seals and components leads to the presence of hydrocarbons in the oxygen supply. Secondly, the original oxidation electrode was supported from the top-plate of the vacuum chamber and was prone to distortion leading to inconsistent results. However, once the oxygen had been replaced with a supply having a positive head pressure, the leaks had been sealed and the electrode had been stabilised a pattern began to appear.

The thickness of an oxide grown in a plasma discharge depends upon three parameters: the oxygen pressure, the current flowing in the circuit, and the time for which the discharge is maintained. Of these, the most manageable and easily varied is the oxidation time. Thus the current, which is mainly dependent upon applied voltage and the geometry of the electrode and chamber, was fixed at 4mA. The oxidation pressure was fixed at 3×10^{-2}T. If a graph is then plotted of junction resistance against oxidation time, an unexpected plot occurs, as shown in Fig.36:

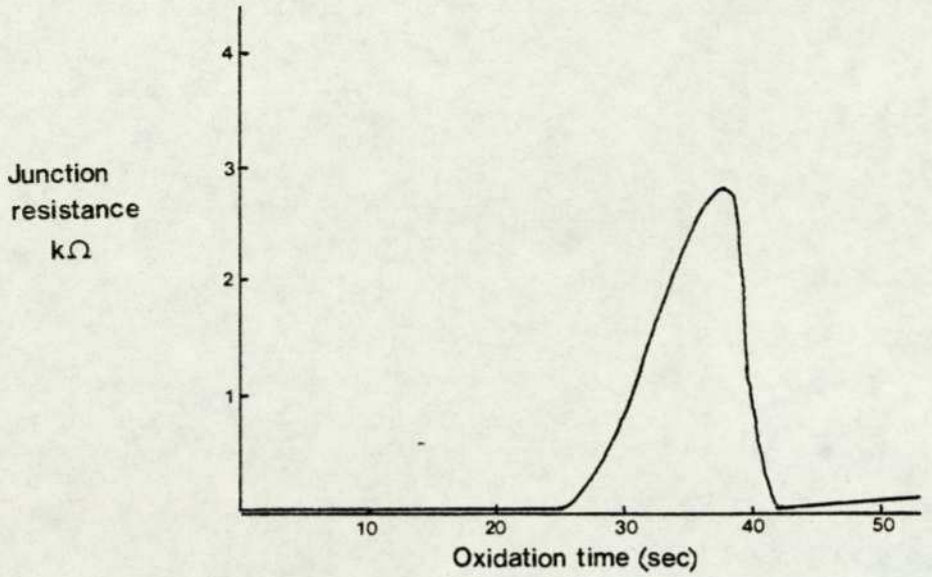


Fig. 36. Plot of junction resistance versus oxidation time for an Al-oxide-Pb junction doped with benzaldehyde

This can be explained as follows. Below 25 seconds, the oxide thickness is too low to sustain any voltage measurement, and shorts through the insulator occur on testing. Thereafter, the oxide thickness increases and the effect of the dopant becomes apparent in raising the resistance by three orders of magnitude until around 35 seconds. At oxidation times above 35 seconds, the oxide thickness becomes such that the dopant starts to lose its effect - the tunnelling electrons give up so much of their energy in tunnelling through the oxide that the numbers reaching the dopant declines. Above 40 seconds, the oxide is now too thick to allow inelastic tunnelling to occur and thus all effects of doping are removed. The increase in resistance is steady from this point on, up to a maximum value of around 100Ω after ten minutes.

This effect has not been previously reported, almost certainly because all other workers are using oxidation times of around 5 to 10 minutes. This is a function of the vacuum system preparation chamber layout, and depends on the distance of the electrode from the base, the distance of the substrate

from the electrode etc. Thus other workers plots would be spread out over a much wider time scale, such that their region of experiment would extend from perhaps five to ten minutes instead of 25 to 28 seconds as in the case in this study. Thus if they were to make adjustments of, for example, 10 seconds in their oxidation time it would have a considerably more gradual effect than a similar adjustment in the present system.

The oxidation time which has been found to produce the optimum results in terms of siganl to noise ratio and spectral resolution is 27 sec. Thus the oxidation conditions established for this study were the operation of a current of 4mA for 27 seconds at an oxygen pressure of 3×10^{-2} Torr. If these conditions are rigorously applied, they produce oxide barriers of high integrity with a frequency of around 97%. This compares very favourably with results reported by other workers, whose oxide integrity varies between 20 and 90%.

The lack of shorts through the insulating barrier appears initially difficult to explain. It is unlikely that the effect is due to the different orders of time involved, since a very large alteration in conditions would be required over the 10 min oxidation to bring about changes equivalent to a very small alteration in the short time scale oxidation. There is, however, one major difference between the oxides grown in the course of this study and those previously reported. Almost all the other workers in this field who have produced oxide barriers on aluminium have used either water vapour-oxygen mixtures or pure water vapour. The oxides grown in the course of this study have been prepared in an anhydrous environment. This is the

only significant difference, and it is therefore postulated that it is this factor which accounts for the higher success rate of barrier formation currently achieved. There is one further factor which adds weight to this argument.

Another feature of this study which has received very little previous comment is the occurrence of what, in section 4.2, was termed boosting. This effect was first detected just after the oxidation conditions had been established. Some junctions were prepared using identical conditions and the same dopant, benzaldehyde, which were of very low resistance. However, when these junctions were placed in the circuitry and a dc bias was applied, as the bias increased the junctions would boost i.e. their resistance would suddenly increase to a value comparable with junctions of the same material having optimum initial resistance. When the spectra of these junctions were compared with those which had not been boosted, they were found to be identical.

The origin of low resistance necessitating this boosting was eventually traced, in the case of benzaldehyde doped junctions, to a displacement of the oxidising electrode which was producing a plasma oriented differently with respect to the substrate. When the electrode was returned to its original position, junctions were again produced which did not require boosting. Thus the boosting itself appears to be a property of the oxide, rather than the dopant. This effect has been previously reported only once, by Skarlatos et al[70]. They report this increase in resistance at applied bias of several hundred millivolts, but have carried these experiments out at room temperature with the junction open to attack over a period

of hours.

Given the sudden increase in resistance, and given the shape of the plot of Fig. 36 where an over-oxidised junction loses its doping effect, it seems reasonable to adopt the explanation of Dorsey[52] previously mentioned in section 3.2. It seems likely that the original oxide can, under some oxidation conditions, exist as two layers, the first being of high dielectric strength and the second with a far poorer dielectric strength. When the two layers exist together, the oxide is too thick to allow the effect of doping to be observed. However, at biases of above 150mV, this second diffuse layer can short through, thus bringing the dopant molecules into the tunnelling region and causing sudden and large increases in resistance. It is clear that further information is required about these boosted and unboosted systems, and more attention must be directed to the nature of the oxide in tunnelling spectroscopy.

4.4. Parameters affecting doping of the aluminium oxide

As in the case of oxidation, there are a number of experimental parameters which influence the doping of the aluminium oxide. Of these, several may be maintained as constants to enable a single variable to be studied. The size and number of drops applied to the oxide, the time they sit on the oxide before spinning off and the speed of rotation of the spinner have all been maintained as constants throughout the study. This leaves as the major variable the concentration of dopant solution.

Spin doping was selected as the doping method because it was considered that introducing organics to the vacuum

chamber as required by vapour doping would seriously impair the turn-around time of the pumping system. The 150mA oxygen glow discharge as used at present will remove all traces of organics which are introduced into the vacuum system from the atmosphere. However, the glass walls of the preparation chamber inevitably become coated with aluminium and its oxide, and the introduction of relatively high pressures of around 10^{-1} Torr of high purity organics would seriously affect the cleanup times and procedures required to guarantee the production of undoped junctions. Thus vapour doping, which was the principal doping method previously reported in the literature, was discounted in this study.

There remained a choice of two doping methods, spin doping and infusion doping of completed junctions. Of the two, the latter is of primary importance as a future development to allow the ready use of IETS as an analytical technique. Experiments were performed over a period of several months using infusion doping following the example of Jaklevic and Gaerttner[31] with very little success. In retrospect, these experiments were performed before the spin-doping with benzaldeyde had been utilised to determine the correct oxidation conditions, and this was probably the main reason for the lack of success. They were also performed at a time when the oxidation electrode was of a different geometry from that used in the more recent part of the study which has produced high resolution spectra.

It is likely that the oxidation time used during the trials with infusion doping was considerably in excess of the required value. Thus the lack of apparent doping was probably due to

the formation of oxides which were too thick to allow the dopant molecules to appear within the tunnelling region. At this point in the study, the lack of appearance of doping of oxides which are too thick was not appreciated, and the solution applied was to increase the oxide thickness on the assumption that insufficient oxide was present to allow doping to occur. Thus every increase in oxide thickness removed the produced junction further from the required range.

Thus after this unfruitful experiment with infusion doping, spin doping was adopted. This technique has since been used with varying degrees of success. The oxidation conditions presently employed were determined, as described in section 4.3, using benzaldehyde as a dopant. Having established these conditions, a variety of dopant molecules were used to attempt to obtain spectra. Following the work of Simonsen et Al[28], it was observed that ethanol, chloroform and benzene produce very little increase in resistance of the junction. Typical increases were observed to be in the area of around 10Ω. These junctions were of insufficient resistance to obtain spectra, and thus these liquids were utilised as solvents.

The effect of concentration of the dopant solution is nicely illustrated by a series of spectra, obtained at 4.2K for solutions of various concentrations of N-methylaniline in benzene, as shown in Figs. 37-41. At high solution concentrations, above 10mg/ml, no fine structure is revealed on the spectrum. Fig.37 is a typical plot of an overdoped junction - the electrons lose their energy by multiple collisions within the insulating barrier, and second harmonic detection becomes impossible. As the dopant concentration decreases, modulated noise levels appear at

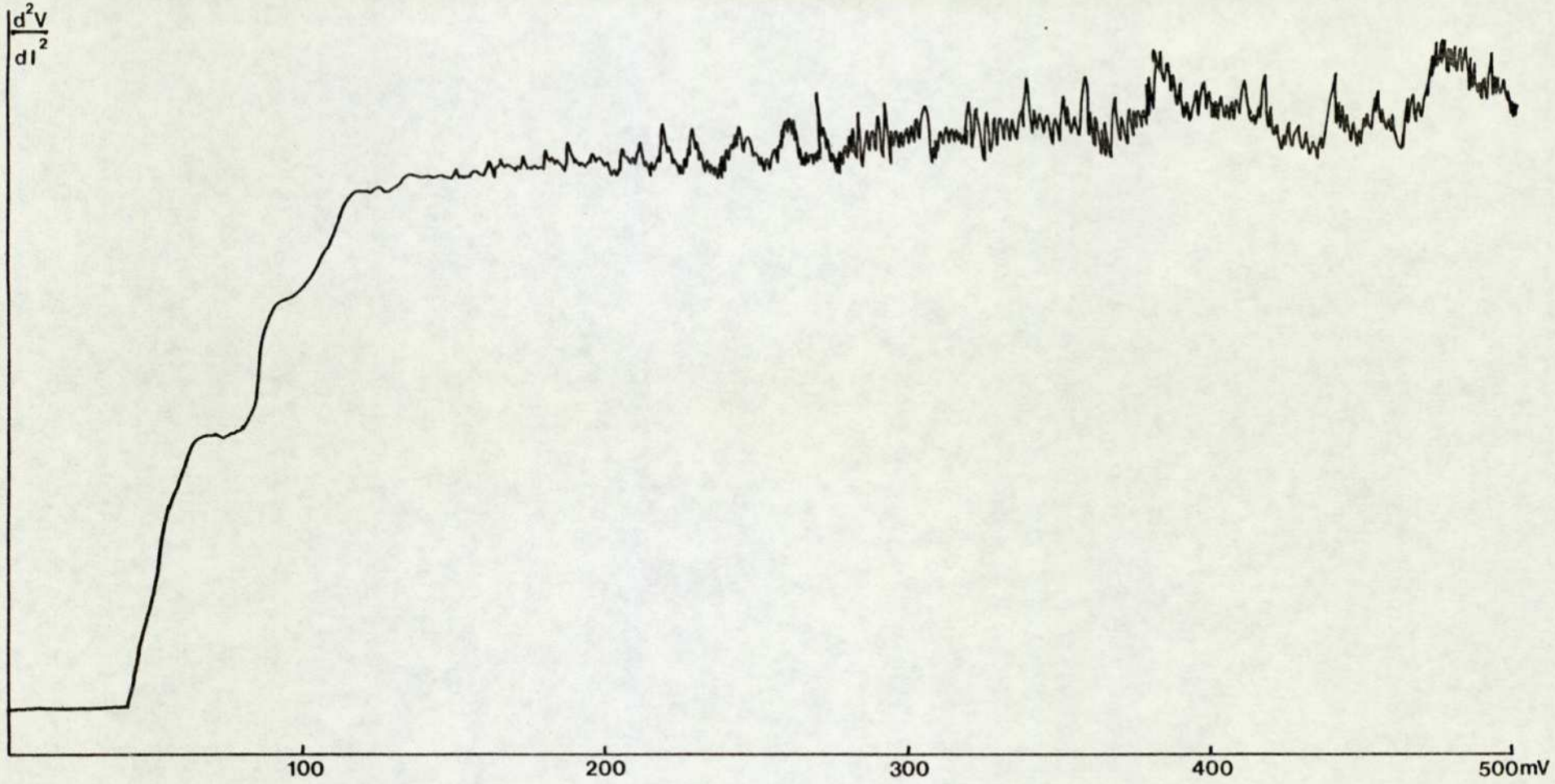


Fig. 37. Tunnelling spectrum of a 20 mg/ml solution of N-methylaniline in benzene. Time constant 100 sec; scan time 40 min; 4.2K.

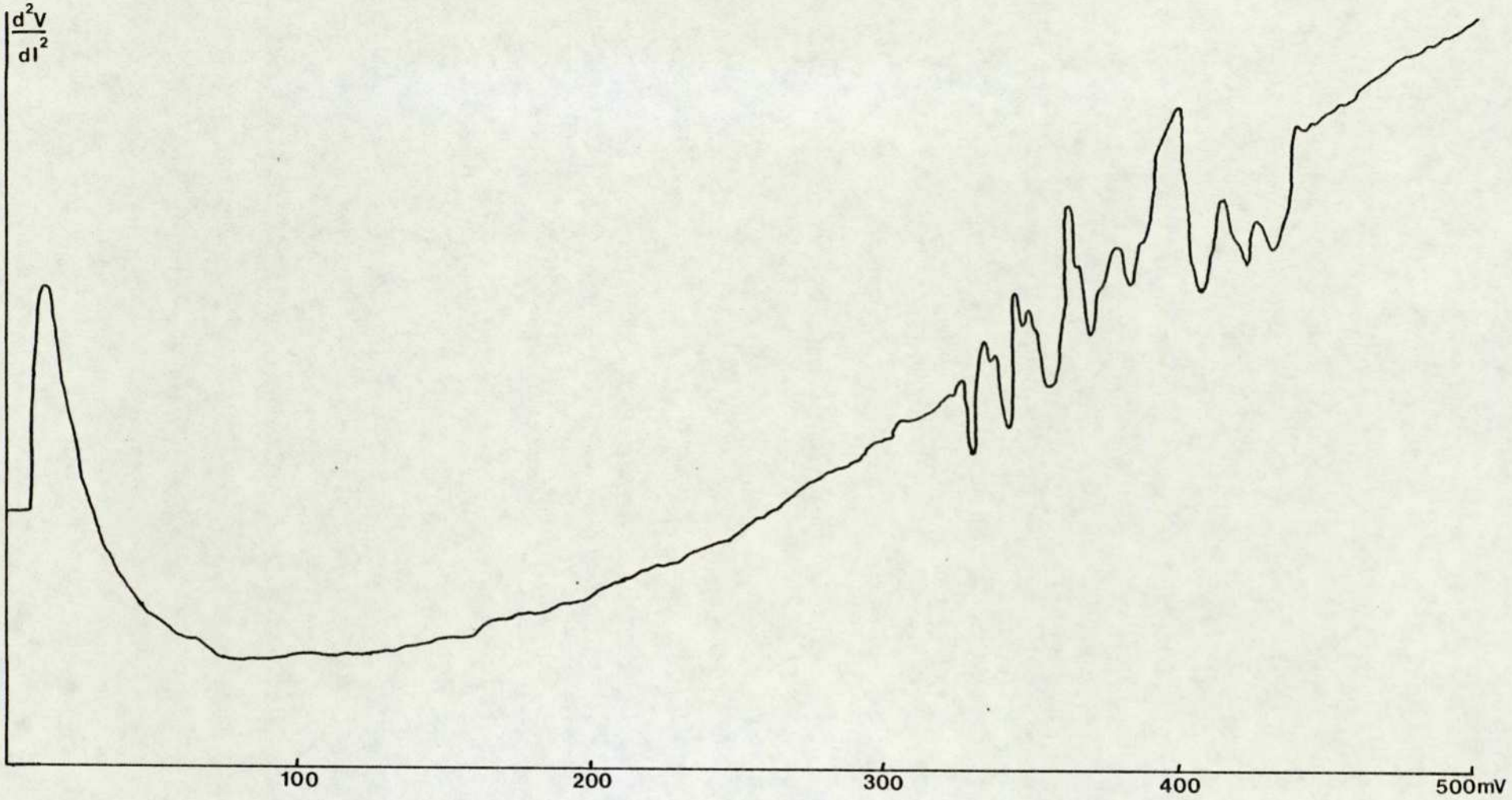


Fig. 38. Tunnelling spectrum of a 6mg/ml solution of N-methylaniline in benzene. Time constant 100 sec.; scan time 80 min.; 4.2K.

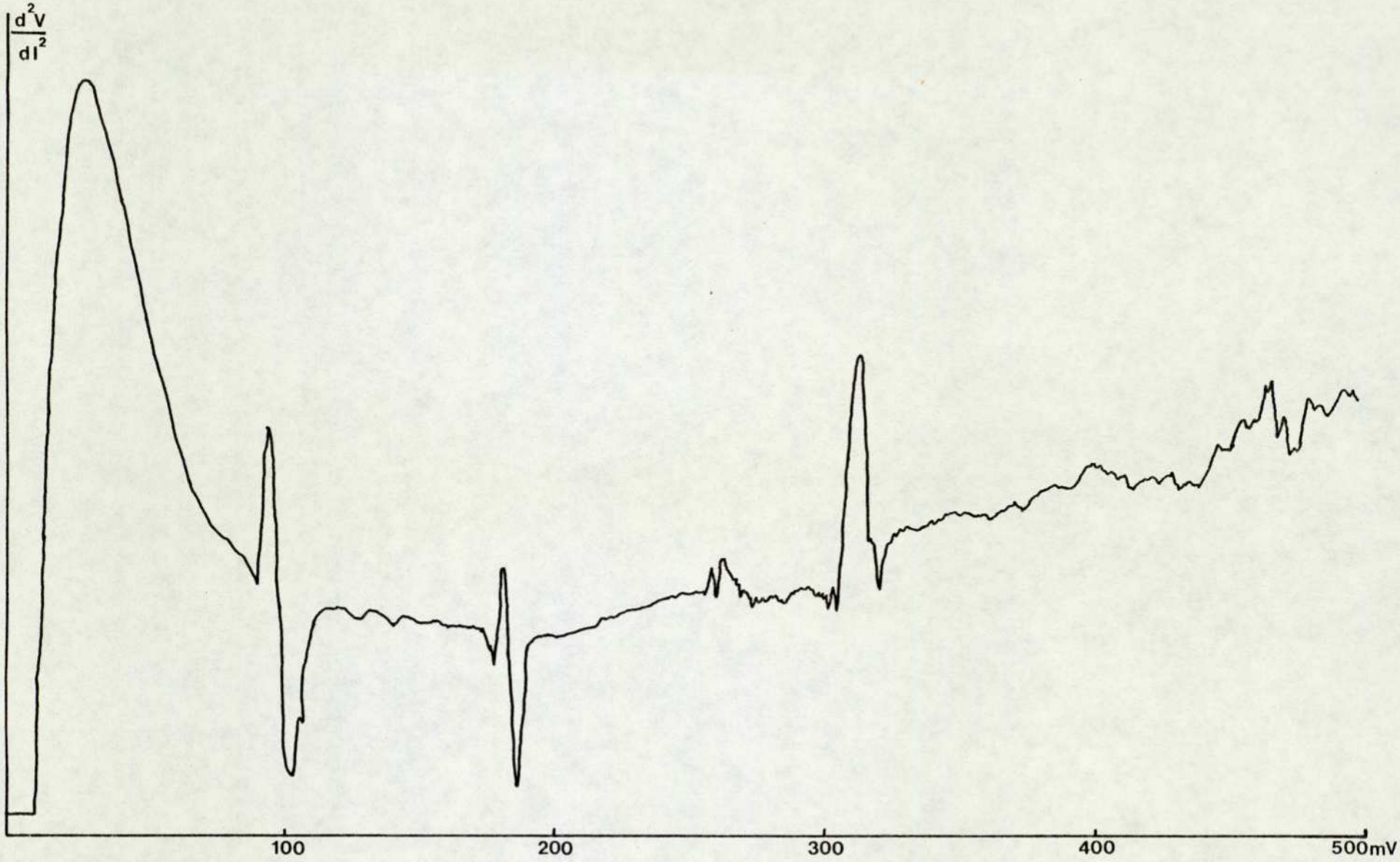


Fig. 39. Tunnelling spectrum of a 3mg/ml solution of N-methylaniline in benzene. Time constant 10 sec.; scan time 30 min.; 4.2K.

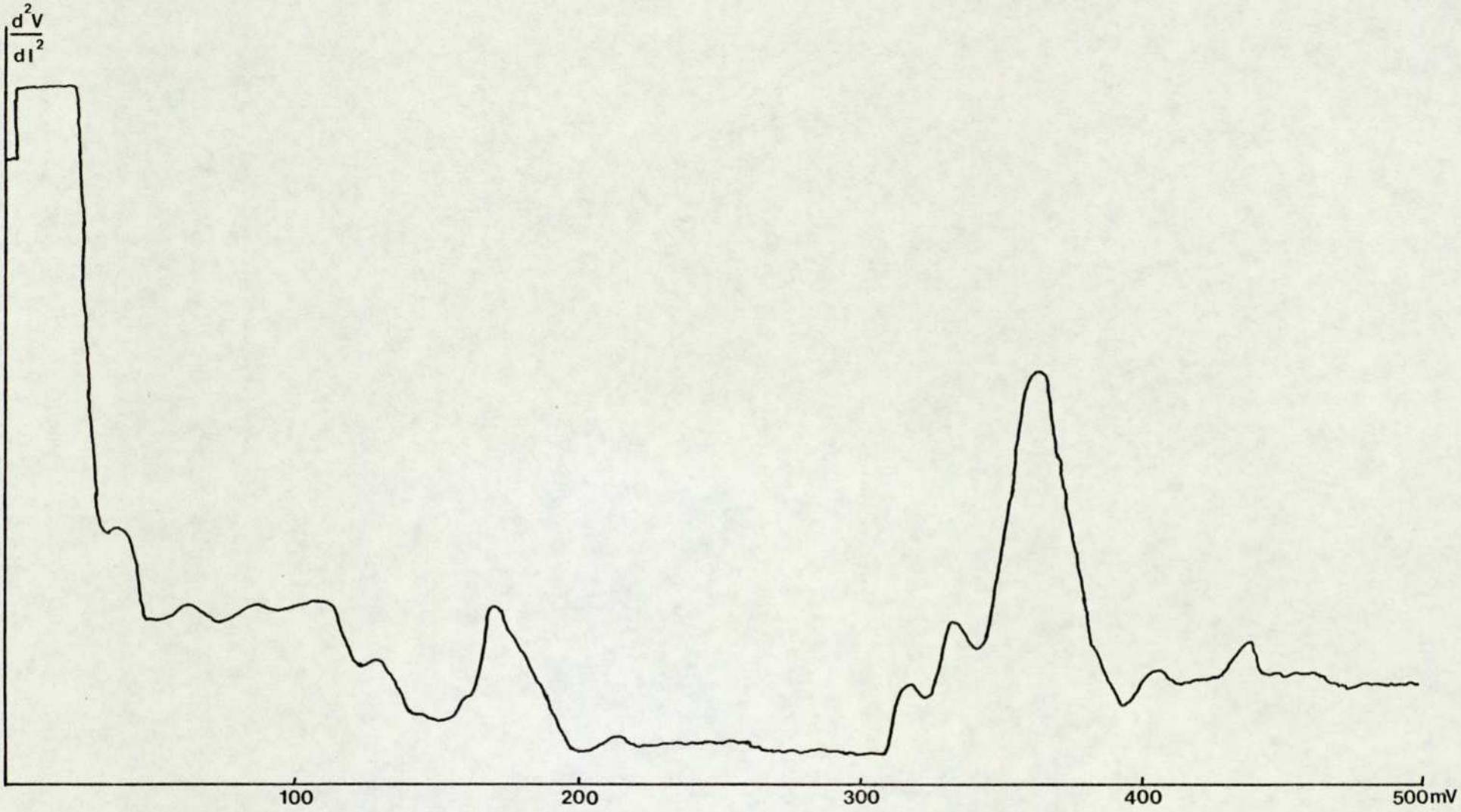


Fig. 40. Tunnelling spectrum of a 1mg/ml solution of N-methylaniline in benzene. Time constant 10 sec.; scan time 13 min.; 4.2K.

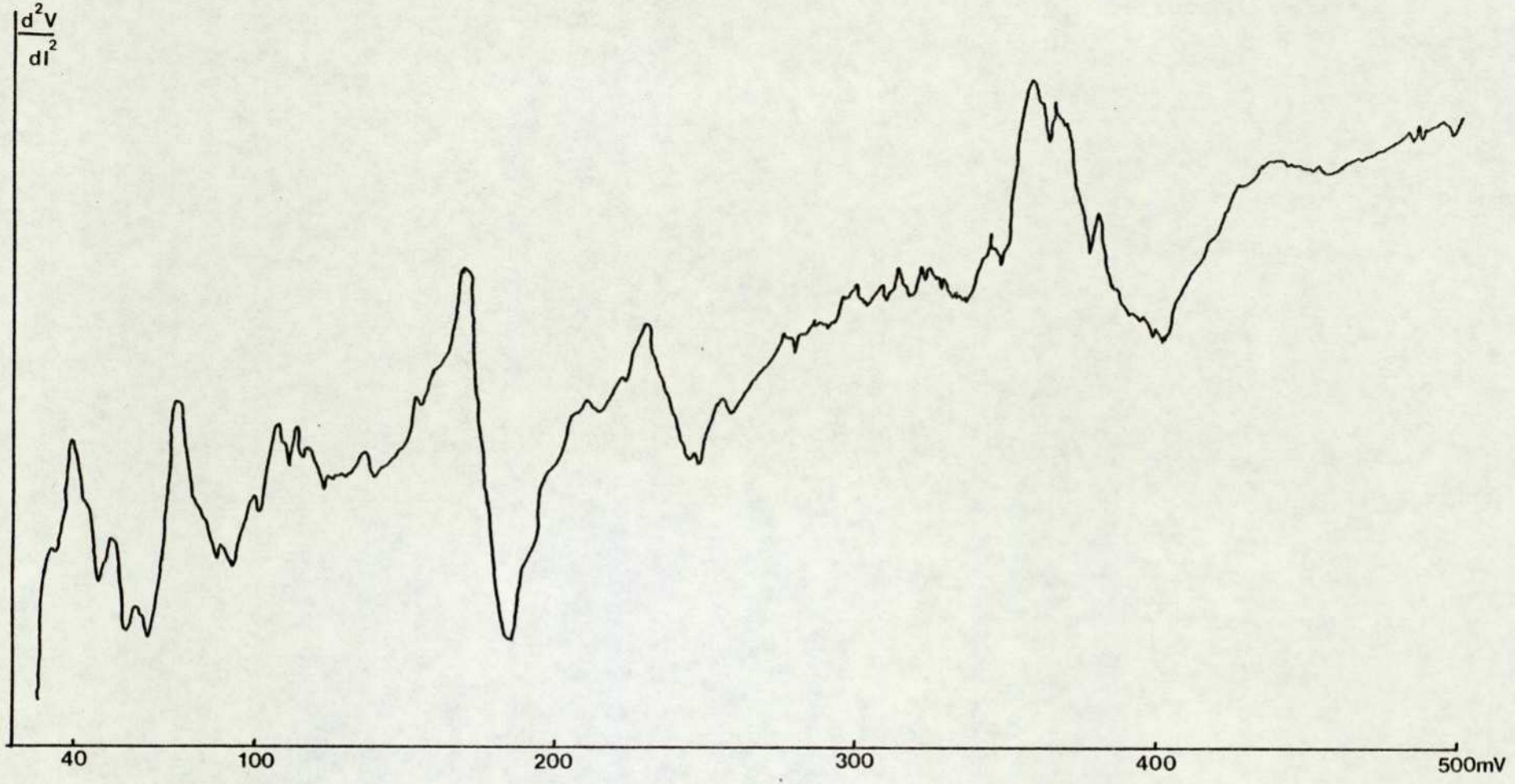


Fig. 41. Tunnelling spectrum of a 0.4mg/ml solution of N-methylaniline in benzene. Time constant 10 sec.; scan time 80 mins.: 4.2K.

selected points of the spectrum which correspond to the peaks of highest intensity in the final spectrum. These always include a broadened phonon peak combined with the appearing superconducting gap, and invariably begin in the higher energy region with the ν_{C-H} at around 360mV, as shown in Fig. 38. At this point in the progression, shifts can occur because the background noise level is so intense that time constants of the order of 100 sec must be employed.

As the solution concentration is further reduced, as in Fig. 39, these modulated noise levels appear in parts of the spectrum corresponding to other strong transitions. The last characteristic spectrum before production of the useful plot is shown in Fig. 40. At this point, peaks are beginning to appear in the high energy spectral range. This is thought to be because only the electrons of higher energy have sufficient to tunnel through the barrier. Fig. 40 has a strong but broad band at the correct frequencies for ν_{C-H} and δ_{C-H}. A further reduction in dopant concentration will produce the spectrum displayed in Fig. 41.

This series is characteristic of overdoping by all the molecular species used in the course of this study, and can therefore be utilised as a type of calibration curve to provide information on the magnitude of concentration reductions required to produce spectra of good resolution. The concentrations used vary considerably from dopant to dopant, for example benzaldehyde and its derivatives which require no dilution, but is generally within the range 01.-2 mg/ml. Correct concentrations are generally determined by trial and error, using the plots of Fig. 37-41 as a guide.

Spectra have been obtained at 4.2K for benzaldehyde, m-bromobenzaldehyde, N-methylaniline, chromium hexacarbonyl and N^3-benzoyl-3^1,5^1-diacetyl-2^1-bromo-2^1-deoxyuridine, and these will be discussed in section 4.5. All these spectra have been produced using the standard oxidation conditions of an oxygen pressure of 3×10^{-2} Torr, a current of 4mA for 27 sec. However certain other solutions have been attempted without success, and prominent among these are all solutions which are made up in water.

All the compounds which have been run as aqueous solutions, including neat D_2O, tapwater and triply distilled water have produced spectra similar to that shown in Fig. 37. This is postulated as being due to the enhancement of oxide thickness, by the formation of a hydrated layer on top of the anhydrous oxide. Preliminary experiments have shown that improvement in the type of spectra produced using aqueous systems is observed by reducing the oxidation time, and therefore the oxide thickness. This indicates, as would be expected, that the type of spectra produced is dependent upon a complex function of oxidation conditions and doping conditions. There does not appear to be one set of oxidation conditions which is suitable or all dopants, nor one set of dopant conditions, especially concentration, which is suitable for a large range of oxide thicknesses.

As mentioned above, aqueous systems have not proved successful. These are, however, by no means the only dopants which have not been successfully applied to the surface. Nitrobenzene, acetonitrile, butyric acid, methyl ethyl ketone, acetaldehyde and benzoic acid have all been tried in various

concentrations in benzene with no success. Certainly a considerable amount of further research is required before the parameters affecting the interaction of the dopant with the insulator are understood.

4.5. Spectra obtained using IETS

4.5.1. Vacuum system impurities

The first spectrum produced in the course of this study at liquid helium temperatures in shown in Fig.42. This spectrum was obtained in the early stages of spectrometer development by running the preparation chamber with an empty liquid nitrogen cold trap, and represents hydrocarbon contamination in the vacuum chamber during sample preparation. The resolution obtained is rather poor, the only peaks present being due to the aluminium phonon (38mV), the aluminium oxide stretching mode (118mV) and carbon hydrogen bending and stretching modes (184mV and 372mV respectively).

The lack of resolution could conceivably be attributed to the early stage of development of the spectrometer electronics, but this is unlikely since the only major modifications made after this spectrum was obtained were concerned with aspects of noise reduction. Therefore it is thought more likely that these peaks are due not only to hydrocarbon impurities adsorbed on the oxide surface, but also to impurities incorporated into the insulator itself. Spectra of comparable resolution have previously been reported using air as the oxidising medium[10], where peaks have been assigned to impurities incorporated within the insulator. This incorporation leads to different environments for the hydrocarbon bonds leading in turn to a broadening inherent in the sample, and unaffected by instrumental

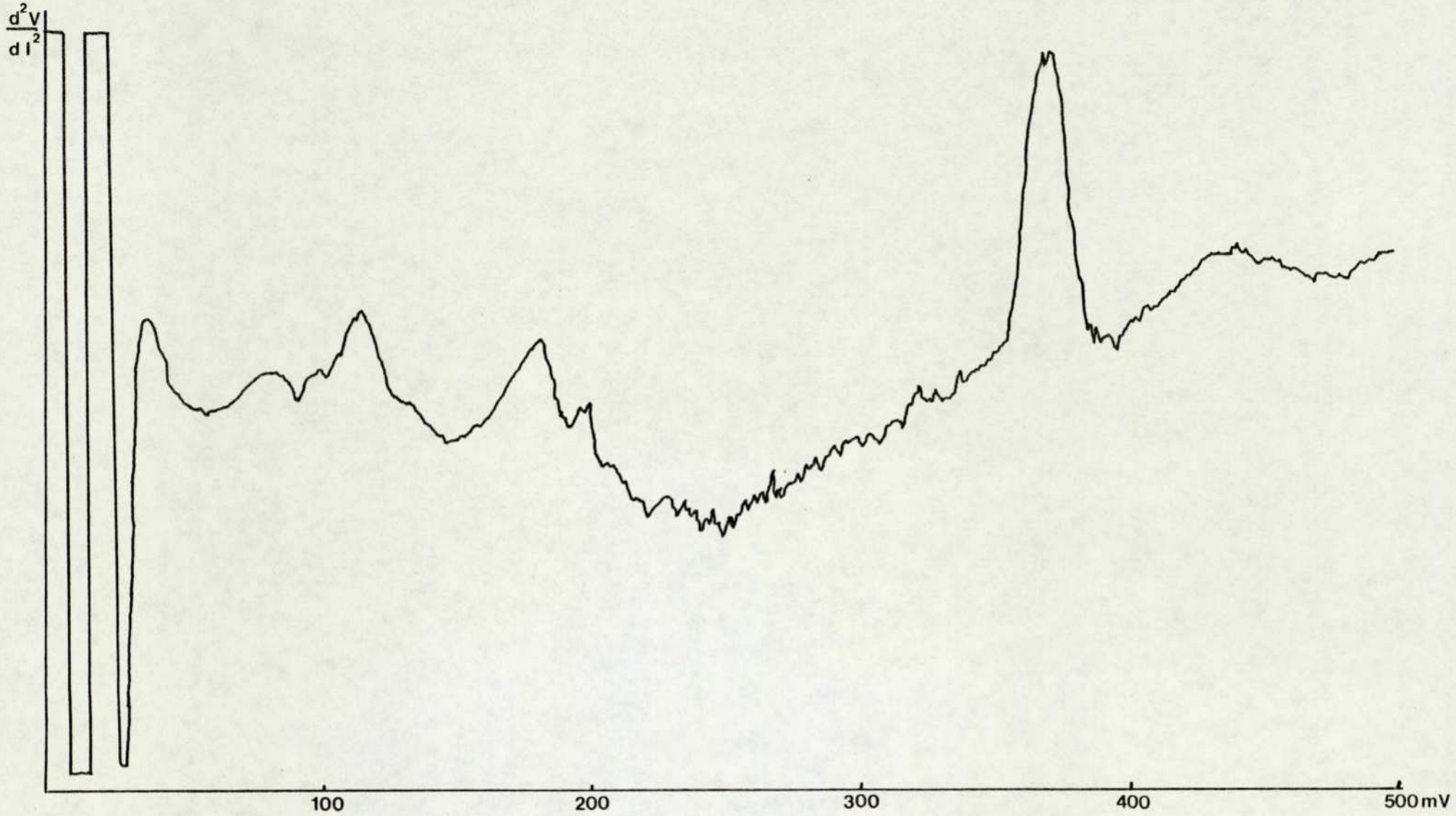


Fig.42. Tunnelling spectrum of vacuum chamber impurities. Time constant 3 sec.; scan time 20 min.; 4.2K.

parameters.

4.5.2. Benzaldehyde

The spectrum of benzaldehyde, obtained after several months of experiments optimising the oxidation conditions, is displayed in Fig.43. The spectrum was obtained using a 3 second time constant, a modulation voltage of 2mV, and a scan time of 20 minutes. The junction resistance at 4.2K was 540Ω. The peaks are assigned in Table 3 and compared with infrared and Raman results for benzaldehyde on alumina, for which recent values are available[71].

Tunnelling Spectrum		IR	Raman	Assignment
mV	cm^{-1}	cm^{-1}	cm^{-1}	
38	307			Al phonon
54	436			
79	637	unavailable	617	
87	702		697	δ_{C-H} o.o.p.
106	855		858	
118	952			ν_{Al-O}
124	1000		1005	
134	1081	1070		
137	1105		1100	δ_{C-H} i.p.
147	1178	1170	1160	δ_{C-H} i.p.
165	1331			ν_{C-O}
180	1452	1450	1460	$\nu_{C=C}$
201	1621	1620	1602	$\nu_{C=C}$
360	2904			ν_{C-H}
382	3082	3070	3080	ν_{C-H}

Table 3. Vibrational assignments for benzaldehyde on alumina

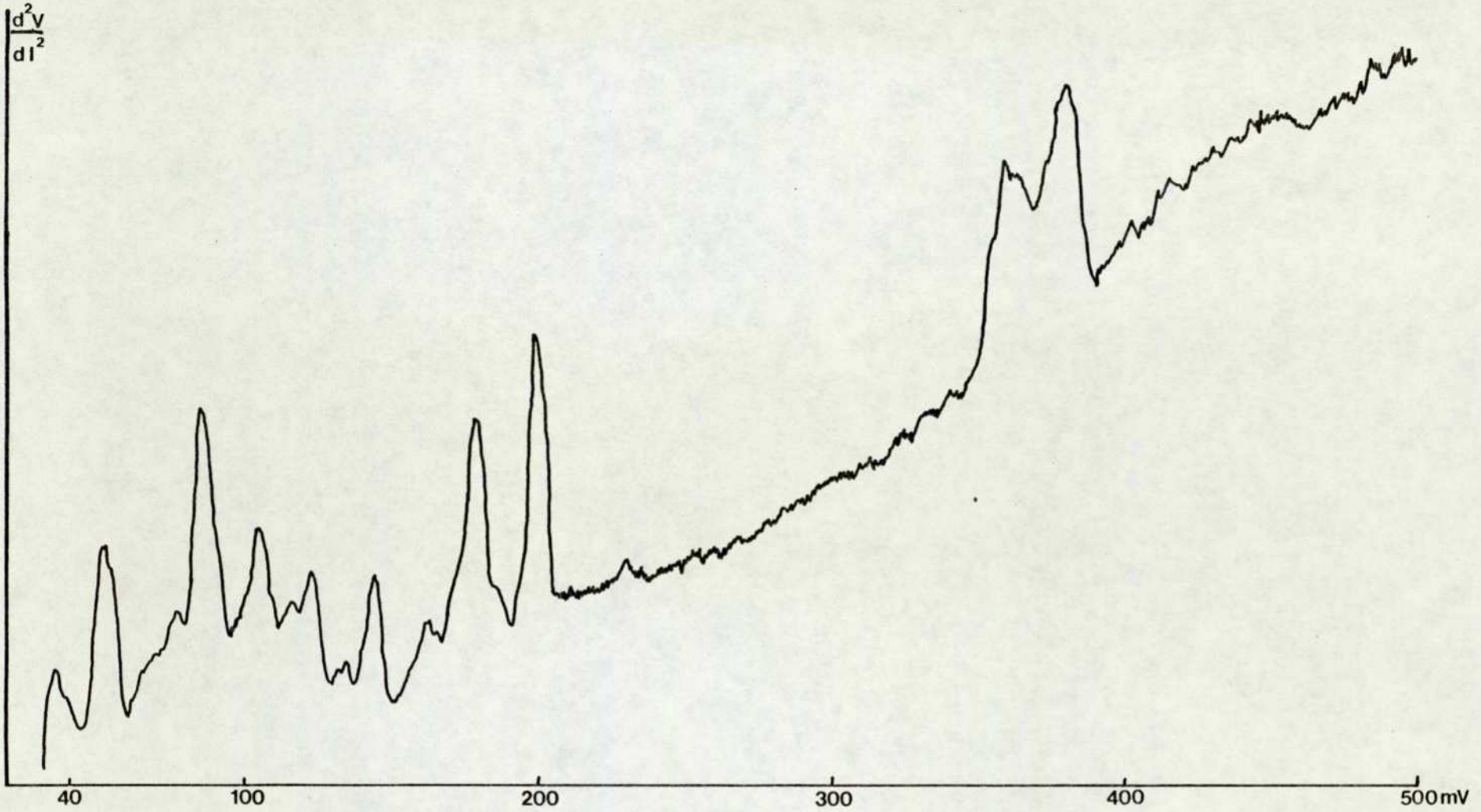
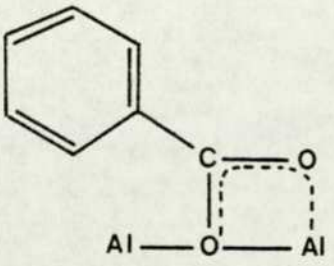


Fig. 43. Tunnelling spectrum of benzaldehyde

Time constant 3 sec.; scan time 20 mins.; 4.2K.

It can be seen that the tunnelling results extend into frequency ranges inaccessible to infrared spectroscopy because of the strong infrared absorption of alumina below 1000 cm^{-1}. Similarly, difficulties arise with Raman because of alumina's strong fluorescence. The work of Kuiper et al[71] indicated that benzaldehyde is oxidised on an alumina surface to give a species with a benzoate structure. This conclusion is supported by the tunnelling spectra of benzaldehyde. In particular, it is notable that the C=O stretching mode of benzaldehyde at around 1700 cm^{-1} is totally absent.

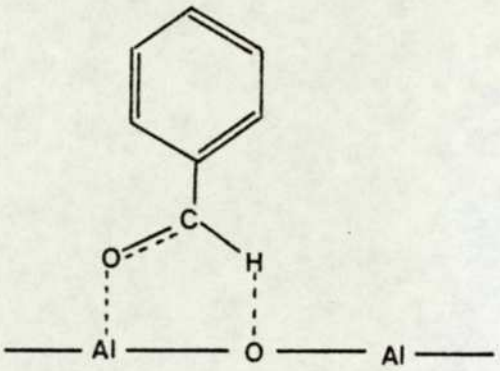
However, the carbon hydrogen stretch associated with an aryl aldehyde is present at 2904 cm^{-1}, significantly higher than the 2820 cm^{-1} in the IR spectrum of benzaldehyde itself. The explanation of Kuiper et al for this peak in their infrared studies was the inclusion of small amounts of alcoholate in addition to the benzoate. Thus the structure indicated by Kuiper et al is:



This is an acceptable formulation, and the peak at 2904 cm^{-1} can be explained by the inclusion of small amounts of alcoholate. However, it is not the only possibility which would produce the tunnelling spectrum in Fig. 43. In fact the infrared spectrum

of aluminium benzoate indicates an asymmetric stretch for the C-O bonds at 1560 cm^{-1} and a symmetric stretch at 1435 cm^{-1}. Both these bands are absent from our tunnelling spectrum. There is, however, a band at 1331 cm^{-1} which does not occur in the IR spectrum of aluminium benzoate. This could be indicative of a structure in which the C-O bond is weakened to a greater extent than that which would occur in the benzoate, and the C-H bond electron density is increased.

This situation would occur if the benzaldehyde was arranged on the surface as follows:



In the light of the evidence of tunnelling spectroscopy, this structure appears at least as likely as the oxidised benzoate structure of Kuiper et al. It does not require the presence of alcoholate to explain the occurrence of the peak at 2904 cm^{-1}, and it explains more adequately the discrepancies with regard to the C-O bond.

4.5.3. m-bromobenzaldehyde

The tunnelling spectrum of m-bromobenzaldehyde, doped by using one drop of the neat chemical, and obtained using a 3 second time constant and a 30 minute scan time is displayed in Fig. 44. The peak positions here may be compared to those of benzaldehyde. The positions are in fact very similar, with the exception of a new peak at 62mV (500 cm^{-1}), due to the C-Br stretching mode. The remaining frequencies are tabulated in Table 4.

mV	cm^{-1}	assignment
38	307	Al phonon
54	436	
62	500	ν_{C-Br}
83	670	
95	766	δ_{C-H}
103	831	δ_{C-H} m-substitution
118	952	ν_{Al-O}
123	992	
134	1081	δ_{C-H} l.p.
144	1161	δ_{C-H} i.p.
161	1298	ν_{C-O}
179	1443	$\nu_{C=C}$
198	1597	$\nu_{C=C}$
358	2888	ν_{C-H}
381	3074	ν_{C-H}

Table 4. Vibrational assignments for m-bromobenzaldehyde on alumina

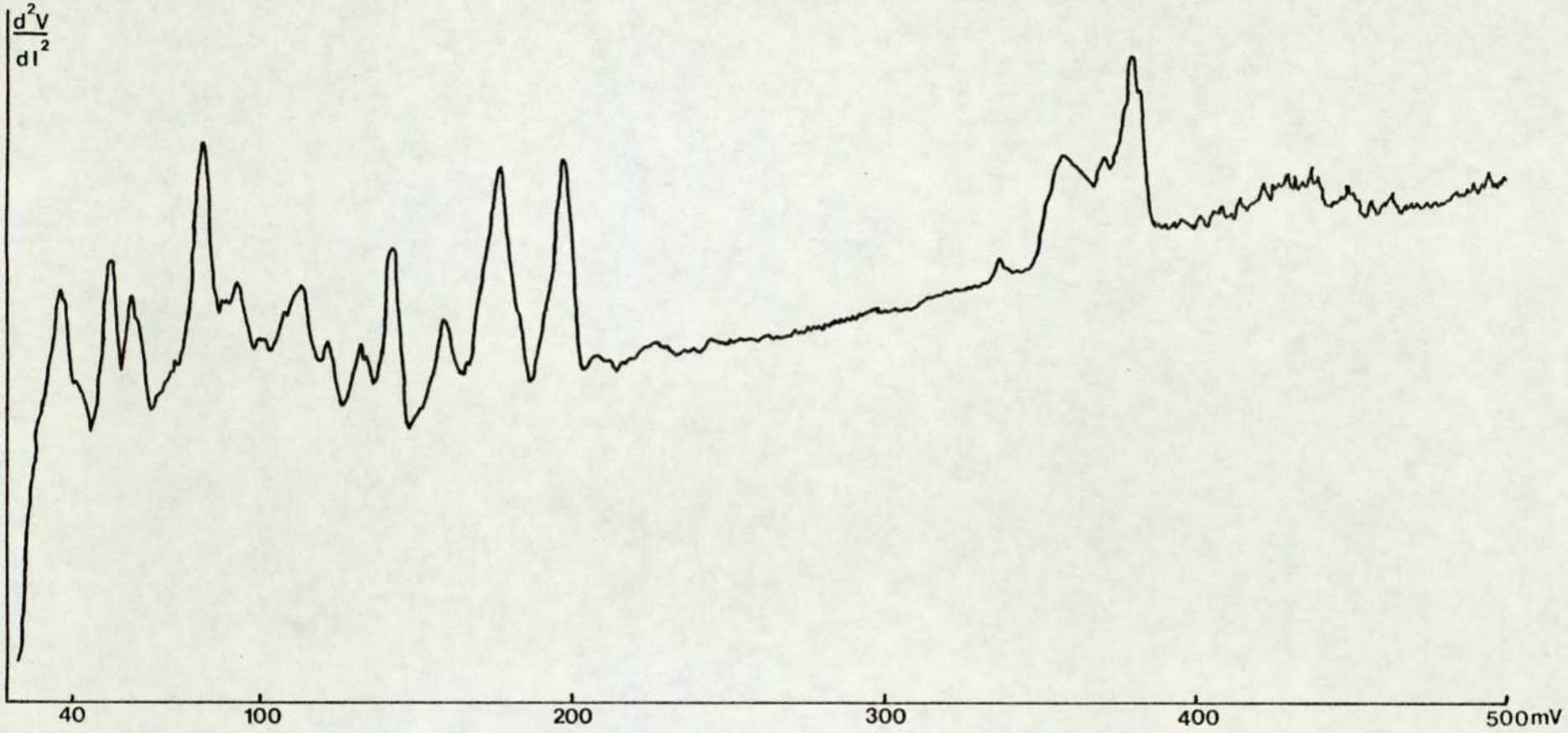
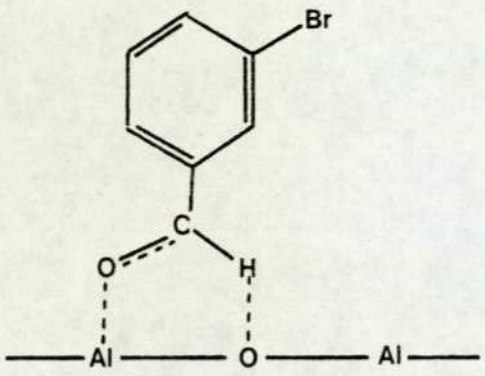


Fig.44. Tunnelling spectrum of m-bromobenzaldehyde. Time constant 3 sec.; scan time 30 min.; 4.2K.

As noted on previous page, the positions are very similar to those of benzaldehyde as one would expect. The major differences arise in the lower part of the spectrum where the C-Br stretching frequency is apparent at 62mV, and where the peaks at 95 and 103 mV are indicative of meta-substitution. In general, the remaining peaks are shifted to slightly lower frequency, which is to be expected with an electron withdrawing ring substituent. The peaks which were assigned to the carbonyl group of benzaldehyde are again present, at 1298 cm^{-1} (ν_{C-O}) and 2888 cm^{-1}(ν_{C-H}). Thus it is postulated that the bromobenzaldehyde is adsorbed on the surface in a similar fashion, namely:



4.5.4. N-methylaniline

The tunnelling spectrum of N-methylaniline is depicted in Fig.45. This spectrum was obtained using a time constant of 10 sec and a scan time of 80 minutes. N-methylaniline dopes alumina too strongly if applied undiluted, and the dopant was therefore applied in a solution in benzene of concentration 0.4 mg/ml. The bands present in the spectrum are assigned in Table 5.

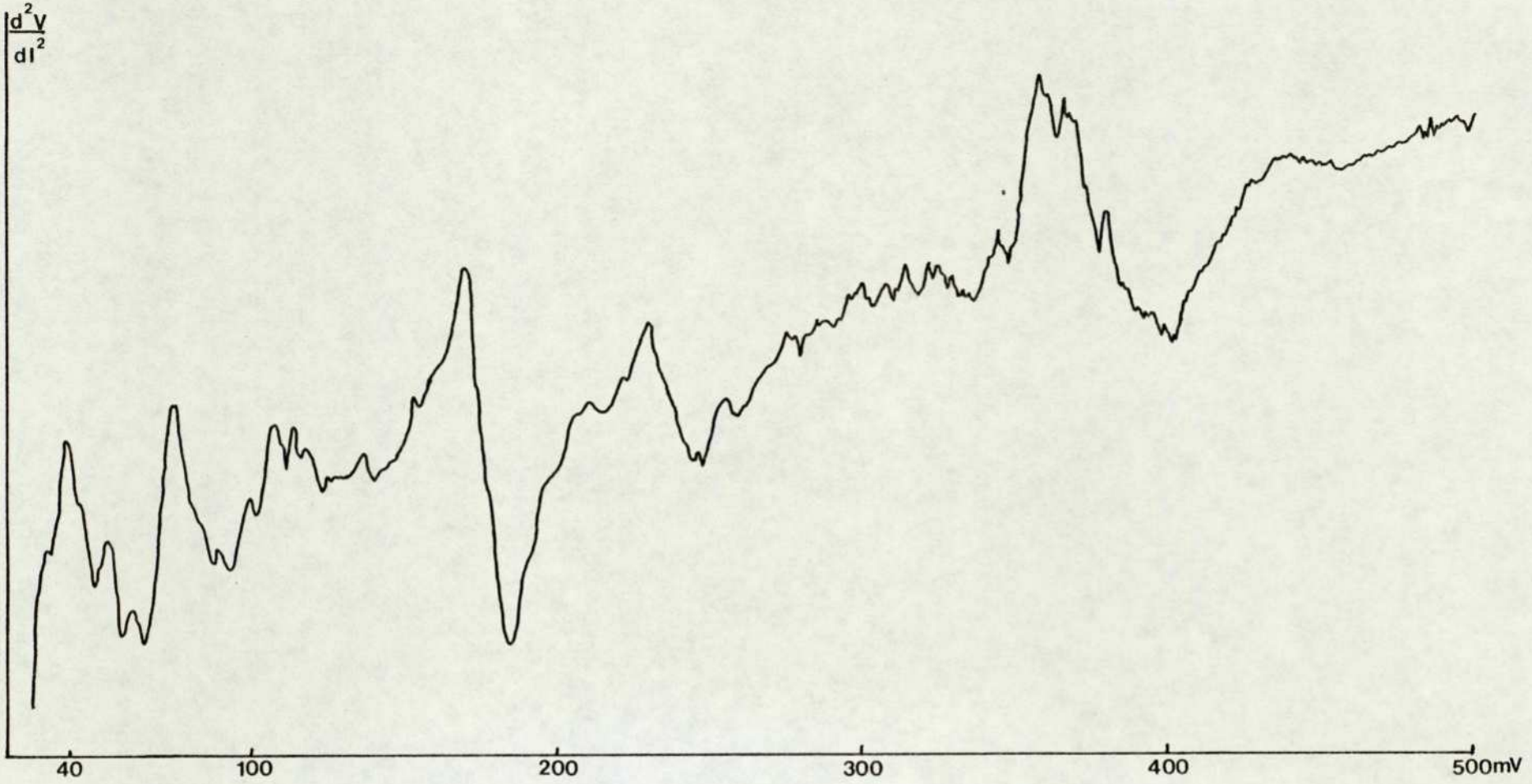
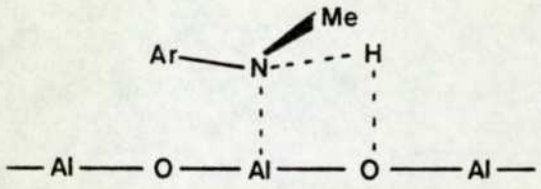


Fig. 45. Tunnelling spectrum of a 0.4mg/ml solution of N-methylaniline in benzene. Time constant 10 sec.; scan time 80 mins.; 4.2K.

The only strong bands in the molecular vibration region are those corresponding to the methyl group. The bands corresponding to vibrations of the aromatic ring are very weak, compared with their intensity in the benzaldehyde and substituted benzaldehyde spectra. The bands associated with the N-H bond are totally absent. Moreover the C-N stretching mode at 1242 cm^{-1} is far lower than the value of the IR of benzene solution of 1320 cm^{-1}. This leads to the tentative assignment of the bonding to the surface as being via the N-H bond, and leading to a conformation such that the benzene ring is almost parallel to the surface:



If the molecule is attached in this fasion, then the bands at 1694 cm^{-1} and 1855 cm^{-1} could represent the stretching frequencies of the weakened N-H bond and the bond formed between the alumina oxygen atom and the hydrogen atom of the amine. The lower of these frequencies, which corresponds to the N-H bond is of considerably lower intensity than that of the O-H bond, which would be perpendicular to the surface, and therefore in the optimum position for IETS interaction. This would indicate that the N-H bond exists at an angle to the oxide surface in the same way as the aromatic ring orientates. The bond from the nitrogen atom to the aluminium atom would utilise the lone pair on the

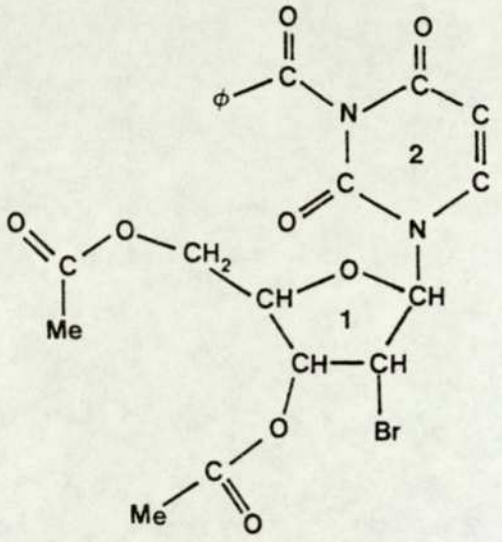
nitrogen, which would allow the N-H and N-Ar bonds to lie closer to the surface.

mV	cm^{-1}	assignment
39	315	aluminium phonon
53	428	
62	500	
75	605	
91	734	δ_{C-H} o.o.p.
100	807	
108	871	
114	920	
118	952	ν_{Al-O}
137	1105	δ_{C-H} i.p.
154	1242	ν_{C-N}
170	1371	δ_{C-H}, sym,bend,CH_3 group
198	1597	$\nu_{C=C}$
210	1694	$(\nu_{N..H})$
230	1855	$(\nu_{O...H})$
358	2888	ν_{C-H} sym,stretch,CH_3group
366	2953	ν_{C-H} asym.stretch,CH_3group
381	3073	$\nu_{=C-H}$

Table 5. Assignment of the peaks in the spectrum of N-methylaniline.

4.5.5. N^3-benzoyl-3^1, 5^1-diacetyl-2^1-bromo-2^1-deoxyuridine

The tunnelling spectra of this molecule, hereafter referred to as A, is shown in Fig. 46. It was obtained using a 0.7 mg/ml solution of A in benzene, and was plotted using at 10 sec time constant and a 40 minute scan time. The molecular structure of A is:



This compound was selected in an attempt to obtain a strong absorption for a carbonyl bond. It is stereochemically interesting in this respect because the two rings, (1) and (2), are arranged perpendicularly. There are three different carbonyl absorptions observed in the IR, the carbonyls of the ring (2) at 1760 cm^{-1}, those of the acetyl group at 1620 cm^{-1}, and those of the benzoyl group at 1670 cm^{-1}. By considering the assignments of the spectrum in Table 6, it may be observed that the vibrations associated with the acetyl and benzoyl carbonyl groups are observed at 202 and 208 mV respectively, as strong but unresolved absorptions.

The carbonyl groups of ring (2) however are not observed.

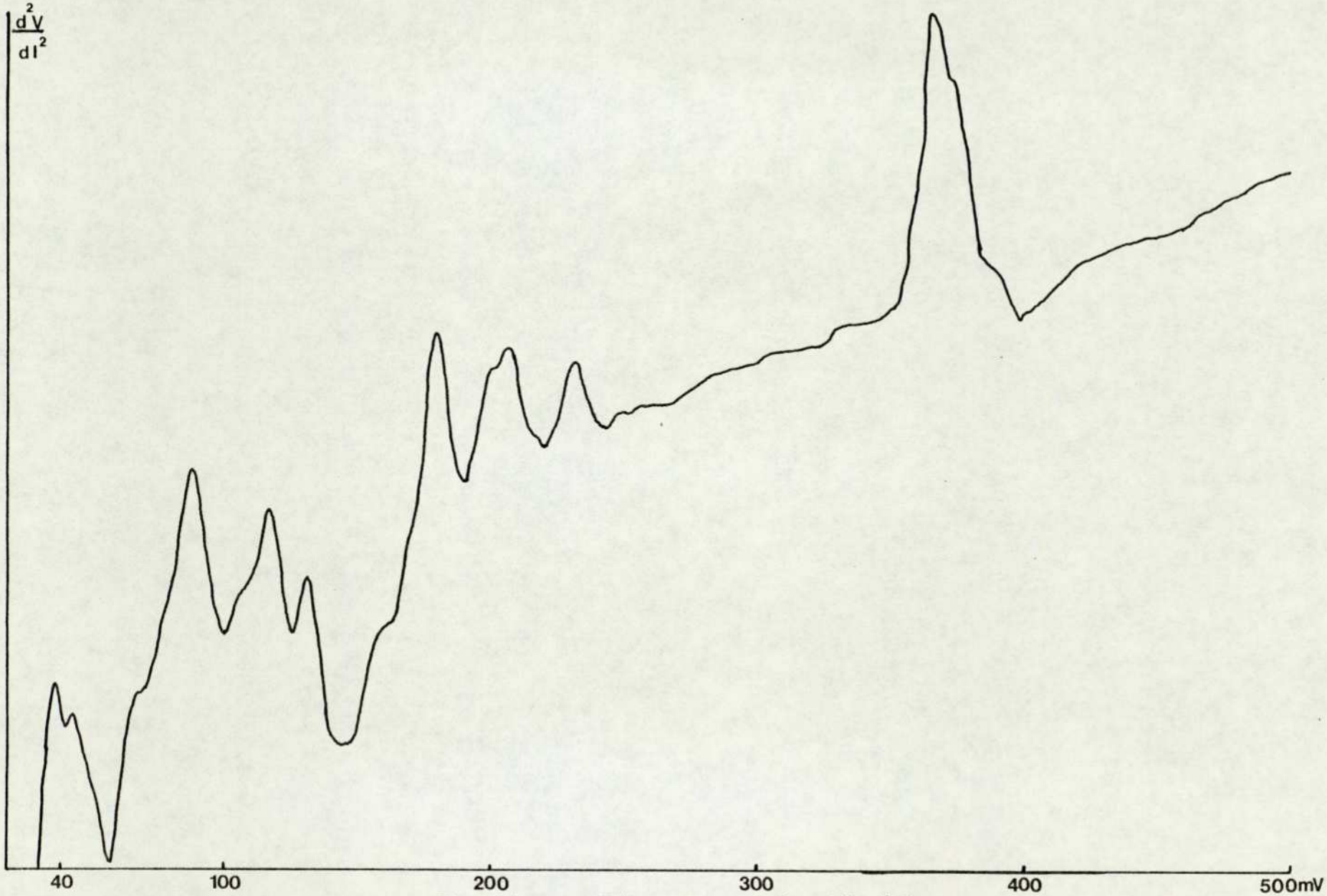


Fig. 46. Tunnelling spectrum of a 0.7mg/ml solution of N^3-benzoyl-3^1-5^1-diacetyl-2^1- bromo-2^1-deoxyuridine in benzene. Time constant 10 sec.; scan time 40 min.; 4.2K.

Neither are the C-N bands, or the skeletal vibrations of that ring. However, the bands associated with ring (1) are observed in the spectrum, especially the band at 1072 cm^{-1} associated with the cyclic ether. This leads to the conclusion that this molecule is adsorbed on to the alumina surface with ring (2) parallel to the surface, presumably using the carbonyl groups to bond to the alumina. This leaves ring (1) perpendicular to the surface, the optimum position for the appearance of strong bands in the spectrum. Thus, by suitable choice of dopant molecule, it is possible to observe carbonyl vibrations in tunnelling spectroscopy.

mV	cm^{-1}	assignment
38	307	Al phonon
45	363	
69	557	ν_{C-BR}
90	726	
118	952	ν_{AL-O}
133	1072	ν_{C-O} cyclic ether
161	1299	δ_{C-H} ring (1)
181	1460	δ_{C-H} methyl group
202	1629	$\nu_{C=O}$ acetyl group
208	1677	$\nu_{C=O}$ benzoyl group
233	1879	? (Al-O overtone)
364	2936	ν_{C-H}

Table 6 Assignment of the bands in the spectrum of N^3-benzoyl-3^1-5^1-diacetyl-2^1-bromo-2^1-deoxyuridine

4.5.6. Chromium carbonyl

Following the success of obtaining carbonyl absorptions for

molecule A, it was decided to attempt to obtain the spectrum of chromium hexacarbonyl. This molecule is an octahedral array of six carbonyl ligands around a central chromium atom. Thus however it is arranged on the surface, there should be carbonyl groups which have a large dipole moment component perpendicular to the oxide surface.

As a progression from the previous results, a solution of concentration 0.7 mg/ml was made up using carbon tetrachloride as the solvent. When a junction doped with this solution was prepared and run, the spectrum shown in Fig. 47 was obtained. This spectrum was at first a little disappointing in that it is of low resolution, and the only bands which appear strongly are the stretching and bending modes of a hydrocarbon, at 361mV and at 181 mV respectively. At first this proved perplexing, since the chromium carbonyl used was a freshly obtained sample of high purity.

Thus, as a precautionary measure, the tunnelling spectrum of the CCl_4 used as the solvent was obtained, and is displayed in Fig. 48. The junction obtained was of low resistance, around 40Ω, which is characteristic of underdoping, as would be expected from a non polar molecule such as carbon tetrachloride. The spectrum was again of poor resolution, but maintained the presence of large bands associated with hydrocarbons. This extreme sensitivity of tunnelling spectroscopy for hydrocarbons can be explained by the observation that a tetrahedrally arranged configuration is always likely to produce a strong dipole moment component perpendicular to the oxide surface.

As an indication of the comparative sensitivity of IETS over IR, the infrared spectrum of this impure carbon tetrachloride was

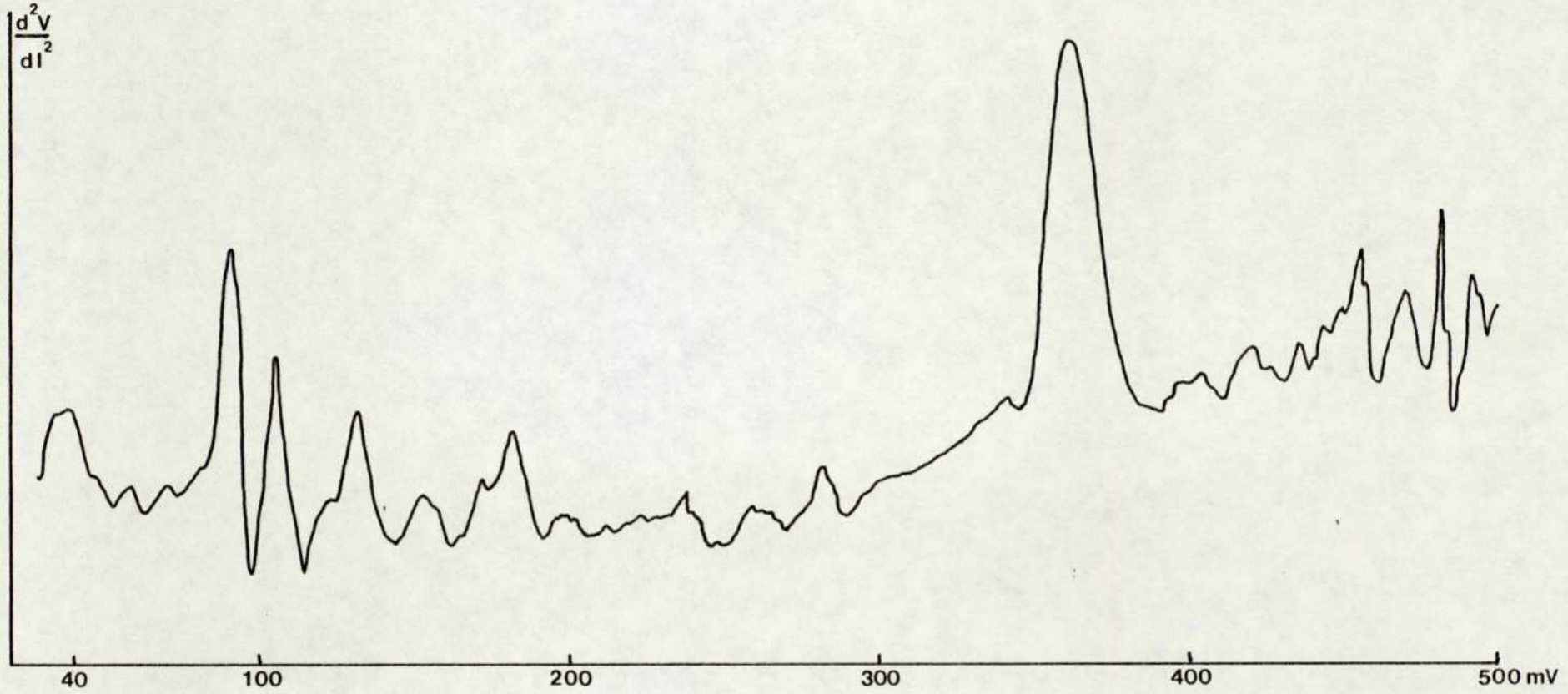


Fig.47. Tunnelling spectrum of a 0.7mg/ml solution of chromium hexacarbonyl in impure CCl_4. Time constant 3 sec.; scan time 20 min.; 4.2K.

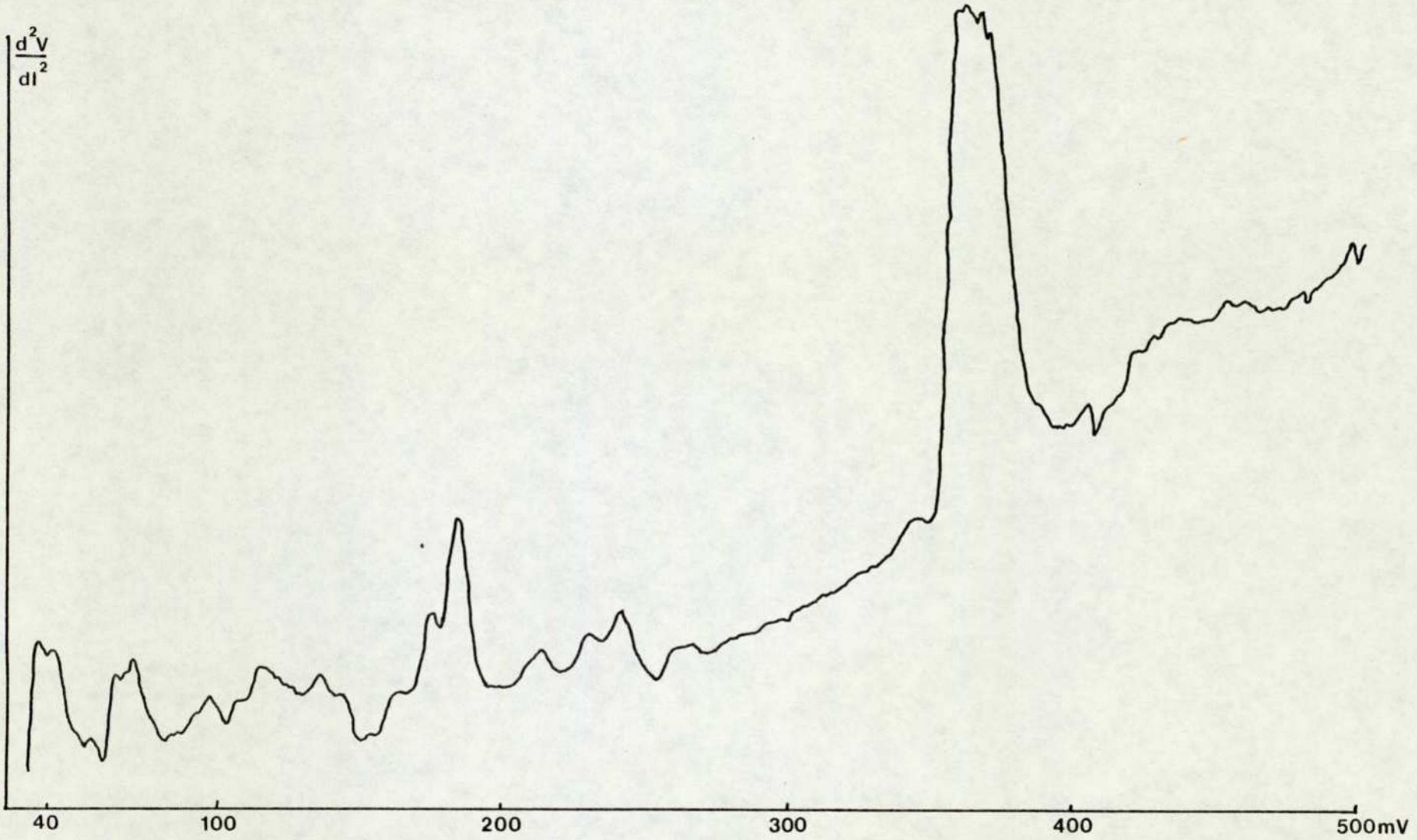


Fig. 48. Tunnelling spectrum of impure carbon tetrachloride. Time constant 3 sec.; scan time 40 min.; 4.2K.

obtained and is shown in Fig. 49. The spectrum was obtained using a 1 mm pathlength cell, and the hydrocarbon stretching frequency is observed as a minute peak at 2975 cm^{-1}.

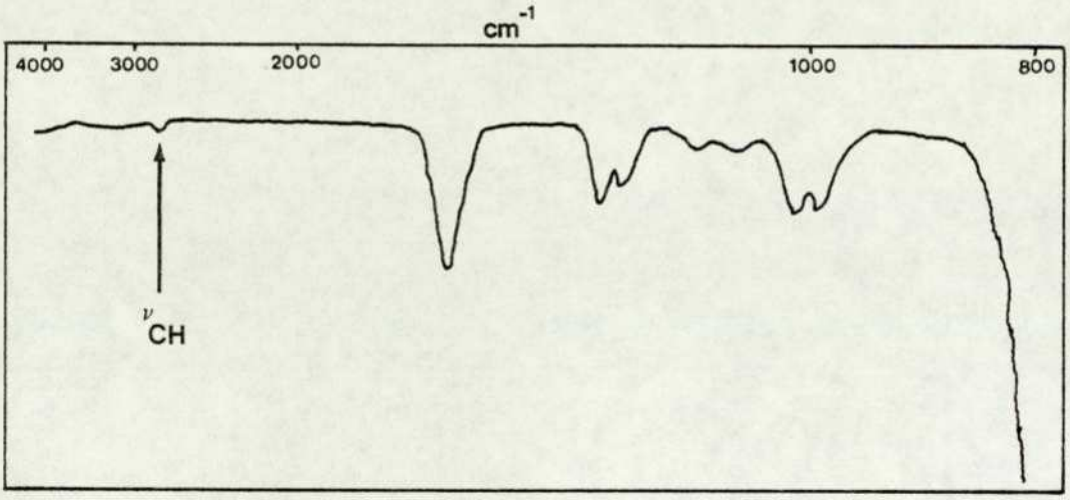


Fig. 49. The infrared spectrum of impure CCl_4.

Thus the hydrocarbon impurity was assigned to impurities within the solvent. A second sample of CCl_4 was therefore obtained. This still contained hydrocarbon impurities, but the level was less than that of the previous batch, almost undetectable in the I.R. To try to counteract the presence of these hydrocarbon impurities, the concentration of chromium carbonyl was increased to 1 mg/ml, and the resultant spectrum is displayed in Fig. 50. The spectrum was obtained using a 10 second time constant and a scan time of 40 minutes.

The resolution remains rather poor. Little is to be gained from detailed assignment of the peaks since the contribution of impurity within the solvent is unknown, However this spectrum may be compared with that of the CCl_4 containing

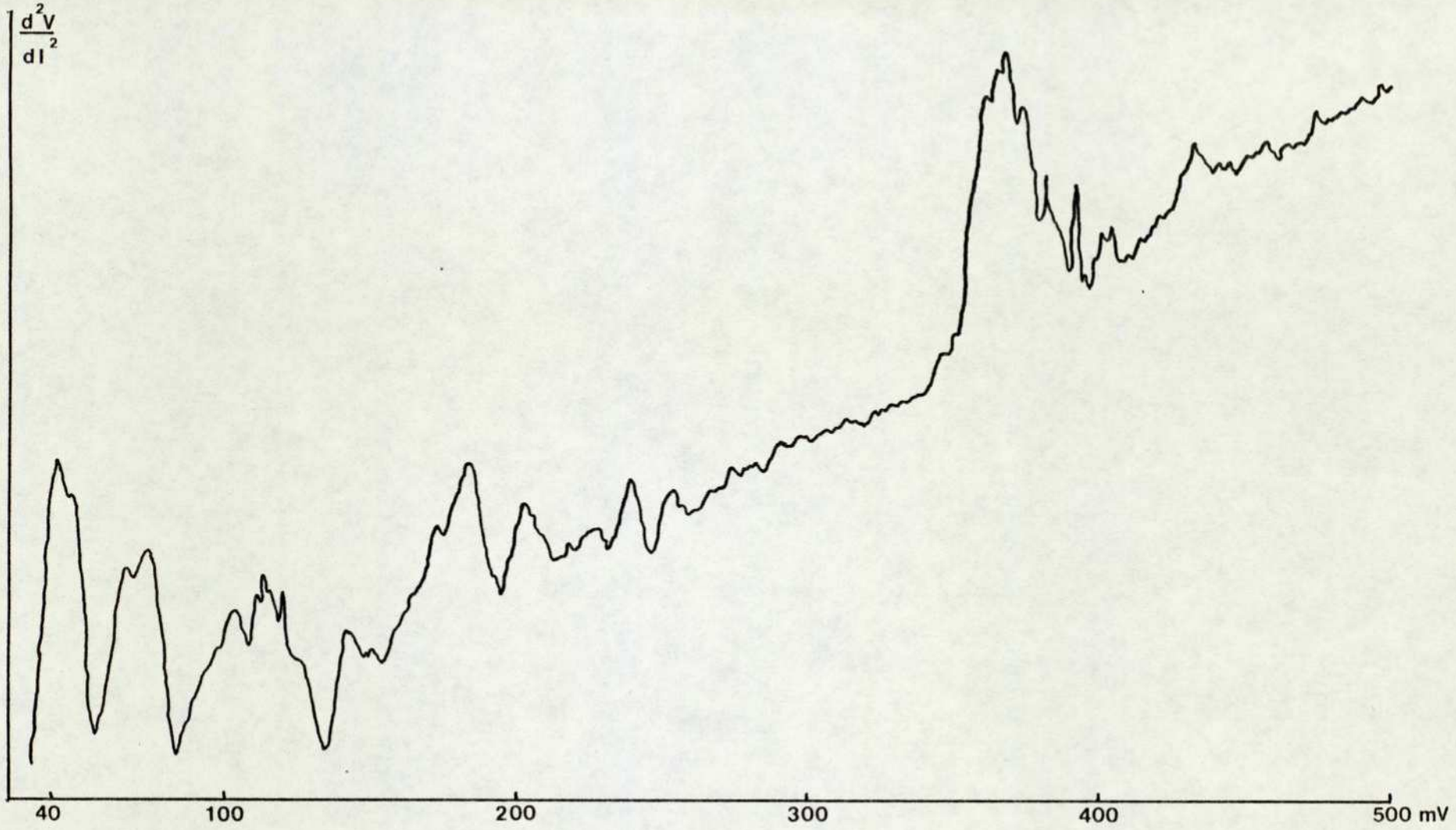


Fig. 50. Tunnelling spectrum of a 1mg/ml solution of chromium hexacarbonyl in chloroform
Time constant 10 sec.; scan time 40 min.; 4.2K.

impurity but no carbonyl, in order to ascertain whether any extra peaks which can be attributed to the $Cr(CO)_6$ can be observed. A comparison of Figs. 50 and 48 reveals the presence of no additional bands. Whether this apparent inability of chromium hexacarbonyl to dope aluminium oxide is due purely to effects of the dopant and substrate, or whether the degree of contamination of the carbon tetrachloride used as solvent was sufficient to mask any effects due to the carbonyl via competitive doping is unknown. The experiment did serve, however, to illustrate the relative sensitivity of IETS and IR spectroscopy.

It is of interest to note that the only previously reported spectrum of any compounds other than those which may be termed organic is by Hansma et al[72]. This will be discussed in Chapter 5, but briefly it involved the detection of carbon monoxide adsorbed not on alumina, but on a thin film of rhodium supported on the alumina insulator. The effect of this rhodium film was to strongly suppress the vibrational modes of the insulator region, and it may be that this is necessary to prevent selective doping by polar organic molecules from masking the effects of compounds such as metal carbonyls.

4.6. Summary

This chapter has attempted to illustrate, using the examples studied to date, the potential of IETS as a surface spectroscopy. The present state of junction preparation and thus the ability to obtain spectra is of ahighly empirical nature. The effects of oxidation and doping parameters are at present understood only at a qualitative level, making

the preparation of suitable junctionsvery much a case of trial and error. It is of interest to note that the work described here in fixing oxidation conditions and obtaining the spectra described involved the preparation of some 580 junctions over a period of around eighteen months.

It is considered, however, that the present determination of oxidation conditions which appear successful for a range of dopant molecules will facilitate preparation of suitable junctions in the future.

Chapter 5. An appraisal of the present position and future potential of IETS.

5.1. Introduction

Inelastic electron tunnelling spectroscopy permits vibrational and electronic modes of adsorbed molecules to be determined with high resolution at picogramme concentrations. This information can be used to define adsorbant-substrate interactions of importance in fields such as heterogeneous catalysis, adhesion and lubrication. It will also allow the determination of very low concentrations of a wide variety of molecules in gas phase or in solution. This attribute is likely to find application in environmental pollution monitoring as well as in biological chemistry. Finally, IETS data can be used to characterise molecular vibrations and excitations with high precision.

In this section, the present position of this new spectroscopic technique, as well as potential applications which may prove to be of some importance in the near future, are considered.

5.2. Range of excitation potentials observable with IETS

The following excitations have been observed in IETS, where $1mV = 8.067\ cm^{-1}$.

Type of excitation	Excitation potential
Phonons in metal electrode at the metal-insulator interface[73]	Up to 50mV
Molecular or lattice vibrations of the insulating barrier[74]	50 - 120mV

Type of excitation	Excitation potential
Vibrations of molecules contained within the insulator region[27]	120 - 500mV
Electronic transitions[65]	above IV

Of these four groups of excitation, the first is of interest mainly to the solid state physicist in studies of metal structures, and in such fields as superconductivity measurements and strain modulation techniques. The other three regions, however, are of far more interest to the chemist, and it is applications involving these regions of the IETS spectrum which will form the basis for the remainder of this chapter.

5.3. Surface chemistry and catalysis

Henry Weinberg and his co-workers have pioneered the application of IETS to surface chemistry[75]. Both the sensitivity and the resolution of IETS make it a highly attractive alternative to infrared techniques in the investigation of adsorption and heterogeneously catalysed surface reactions[76]. Benzoic acid[77] and acetic acid[54] have been shown to adsorb on to aluminium oxide as the benzoate and acetate ion respectively, as evidenced by the loss of the OH stretching and bending mode in the observed spectra. Benzoyl chloride has also been found to adsorb on to aluminium oxide as the benzoate ion, suggesting a surface reaction during or following the adsorption process[76].

Molecular orientations of the adsorbate species relative to the substrate surface have also been deduced in IETS from

a study of sulphonic acids adsorbed on aluminium oxide[78]. The sulphonic acids CH_3SO_3H and CF_3SO_3H have been found to chemisorb anionically on to alumina by losing a proton, just as found previously for carboxylic acids. The S-OH free stretch, associated stretch, wag and bend modes which would appear near 830,900,480 and 1120 cm^{-1} have all been found to be absent, and the characteristic sulphonate mode of a symmetric stretch has been observed near 1040 cm^{-1}. By consideration of the symmetry properties of the various possible anionic arrangements of sulphonates on alumina, it has been deduced that the ion adsorbs in a totally symmetrical tripod arrangement such that all oxygen atoms are in equivalent sites on the surface. This system is best attained when the sulphur atom is directly over an aluminium atom.

This orientation would involve the C-S bond being perpendicular to the oxide surface, and this is supported by the observation that the modes associated with this bond are responsible for the strongest absorptions in the IETS spectra.

Hansma et al have extended the application of IETS to catalytic studies by their investigation of rhodium supported on aluminium oxide and its interaction with carbon monoxide[72]. Their results, depicted in Fig. 51, show a series of tunnelling spectra for various coverages of rhodium metal on alumina. The coverage was determined with an oscillating quartz crystal microbalance and thus represents an average over the surface. The spectrum with no rhodium coverage is free of any vibrational peaks due to CO. The peaks present are due to vibrational modes of the alumina, OH groups bound to the surface of the alumina and the aluminium electrode.

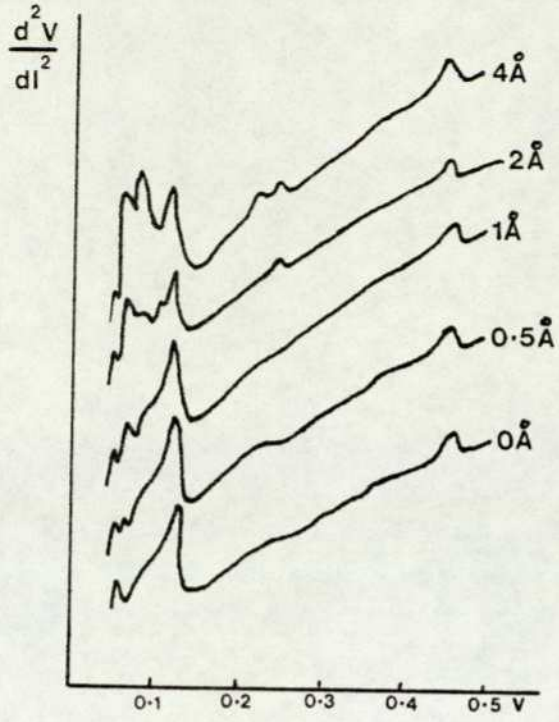


Fig. 51. Tunnelling spectra of CO adsorbed on various thicknesses of Rh evaporated on alumina (After P.K. Hansma et al in ref. 72)

Previous studies[79] on rhodium carbonyls such as $Rh_4(CO)_{12}$ have suggested that the rhodium-carbonyl stretching modes should be in the range 400-600 cm^{-1}. At low rhodium coverages, only one peak is observed in this range at 408 ± 5 cm^{-1}. As coverage increases, this peak grows in size and two new peaks appear: a sharp one at 454 ± 5 cm^{-1} and a broad one at 580 ± 5 cm^{-1}. Hansma has assigned the original peak to carbon monoxide which is linearly bound to a single rhodium atom, and the two growing peaks to a bridged species.

The intensity of the peaks in the carbon-oxygen stretching region is surprisingly small, and partially obscured by the second harmonic of the strongest peak in the background spectrum. This is possibly due to an orientation effect,

in that the peaks would be expected to be weak if the CO were oriented parallel to the oxide surface. However, the lower peak, at 1730±10 cm^{-1}, appears only at higher rhodium coverages. This suggests that the adsorbed species present at higher rhodium coverages has a weaker carbon-oxygen bond, and thus provides additional evidence for the presence of a bridging carbon monoxide molecule.

Another application in this field has been documented in the work of Skarlatos et al[70]. They obtained the spectrum of partially ozonated phenol solutions containing intermediate reaction complexes, as shown by the yellow colouration of the solution. By comparing their tunnelling spectra with the predictions of two alternative theories of the mechanism of the ozonolysis of phenol[80,81], they were able to provide corroborative evidence for one of the possible mechanisms, involving the formation of zwitterions as intermediates[80].

5.4. Analytical Applications

The extreme sensitivity of IETS indicates that this new technique will have many important analytical applications. As all of the earliest work in this field involved adsorption from the vapour phase, it was natural that the first trace substance analysis should be in a similar mode. By 1970, Klein et al[10] had detected a large number of pollutants in Paris air by this method. The constituents of tobacco smoke have also been studied[82].

The greatest analytical advantage of IETS should be realised, however, in aqueous phase analysis. The first analytical results on solutions in liquid phases were reported in 1973 and 1974 by Bogatina et al[68], Skarlatos et al[83], and Hansma et al[84]. These

results were obtained by exposing the oxidised Al electrode of the incompleted junction to the solution under investigation prior to eventual completion of the junction by evaporation of the lead electrode. Aqueous solutions of acetic acid[83], formic acid[77], phenol[70], benzoic acid[85] and stearic acid[27] have been studied in this way. Skarlatos has studied the phenolic pollution of the Paris water supply and monitored its treatment by ozonisation using IETS[70], and has shown that less than 10ppm of a wide variety of organic pollutants in water produce a recognisable spectrum.

From a quantitative viewpoint, much more evidence on a wide variety of molecular species is needed to support the observed quantitative relationship for aqueous benzoic acid between the concentration in the aqueous phase and the peak height[83], in particular, it is still necessary in all cases to empirically establish the relationship between solution concentration and surface concentration. This entails considerable effort, particularly if two different organic molecules in solution are competing for the same adsorption site on the surface of the junction insulator.

5.5. Applications in Adhesive Science

The majority of adhesive research has been concerned with how the adhesive sets i.e. its polymerisation, and not with how the adhesive adheres to the surface. This situation reflects the lack of tools for effectively studying the monolayer of adhesive molecules actually in contact with the substrate. The proposition that an adhesive interface consists of a molecular layer bonded to an oxide suggests the possibility of inelastic electron tunnelling spectroscopy as a useful tool.

This application of IETS has been preliminary studied by White, Godwin and Wolfram[86].

Fig.52 shows a schematic representation of a typical adhesive bond between two pieces of aluminium metal. The adhesive is in contact with the aluminium oxide rather than the metal since an oxide quickly forms on aluminium when it is exposed to air or cleaned with etching solutions prior to bonding. The chemical nature of the interface is very important to the integrity and service life of an adhesive bond. The broken line in Fig. 52 shows a typical fracture line, and it often happens that the fracture line originates at the adhesive-oxide interface.

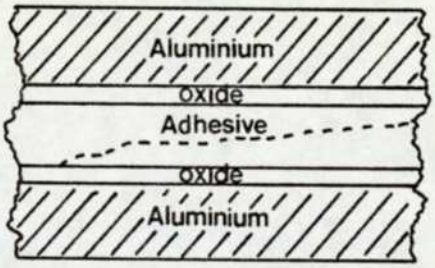
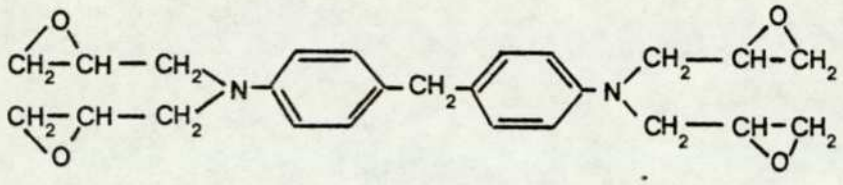


Fig. 52. Schematic representation of an adhesive bond between two pieces of aluminium (After ref. 86)

The epoxy system chosen for study was a mixture of the two molecular compounds diamino diphenyl sulphone (DPS) and tetraglycidyl 4,4^{1} diamino diphenyl methane(DPM), and their structures are shown in Fig. 53. The IETS spectra of both these components in THF have been obtained, and vibrational modes have been assigned by comparison with computer calculations and existing infrared optical absorption spectra. Some evidence for an ageing effect over a period of nine days has

been observed for the adsorbed DPS. This effect appears as a dramatic change in low frequency vibrational modes and may be associated with the formation of hydrogen bonding between the NH_2 groups and the oxide layer.

Tetraglycidyl 4,4[1] diaminophenyl methane, DPM.



Diaminodiphenylsulphone, DPS.

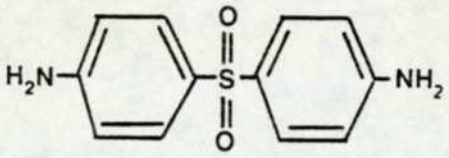


Fig. 53. Structures of the two components in the Hercules 3501 epoxy system (after ref. 86).

5.6. Radiation Damage Studies

Radiation damage to thin organic films is an area where several of the capabilities of IETS are important. The sensitivity allows measurement on small surface area samples which can be uniformily exposed to moderate beam intensities. The resolution of IETS provides detailed information on the breaking of individual bonds within relatively complex molecules, while the spectral range may permit examination of disordering in polymers.

Hansma and Parikh[87] have investigated the radiation effects of junctions doped with β-D-fructose, and have shown that decrease in peak intensity can be correlated with damage to the corresponding functional groups, and that the growth of new

peaks can be attributed to the generation of new functional groups during radiation.

More recently, Hall et al[88] have studied radiation effects on junctions doped with benzene solutions of a variety of organic acids. Damage cross sections for various bonds can be calculated from the rate of decrease of the corresponding peak in the tunnelling spectrum. For example, the ν_{C-H} in 2,4-hexadienoic acid decreased at a rate equivalent to a damage cross section of $0.26\pm0.04\text{Å}^2$/incident 30KeV electron. For comparison, hexanoate ions had damage cross sections five times larger, while benzoate ions had values seven times smaller. Qualitatively, this is explained by the concept that molecules with more delocalised electrons are more resistant to radiation damage.

5.7. Systems of biological interest

The high sensitivity, wide spectral range and relative lack of selection rules may enable tunnelling spectroscopy to become a useful tool for the identification of biologically important molecules. The majority of the work reported so far in this field is by R.V. Coleman and his co-workers[89]. They have shown that nucleic acid derivatives can be differentiated by IETS and this observation may lead to the possibility of sequencing nucleic acids by this technique. Some of the molecules whose spectra have been observed include D.N.A, R.N.A. and haemoglobin.

5.8. Electronic spectra studies

The Al-oxide-Pb junction, which is by far the most commonly employed system used in obtaining tunnelling spectra,

has a dielectric breakdown voltage under optimum conditions of about 2.5V, which corresponds to 20,000 cm^{-1}. Hence only the lowest energy electronic transitions may be studied using this particular junction system. However, Klein et al have investigated the low energy electronic transitions which occur in the rare earth metal oxides[66] and in some large transition metal complexes such as the phthalocyanins[65]. Even below the 2.5V dielectric breakdown threshold, the inelastic signal has been found to be seriously influenced by three effects[90]:

(a) The junction current-voltage relationship at higher voltage ceases to be linear and becomes approximately exponential beyond 0.6V. This causes the small increase in current due to the onset of inelastic phenomena to become insignificant with respect to the elastic background.

(b) The large increase in barrier height at higher voltages strongly attenuates the inelastic current.

(c) Under high applied bias, the junctions become increasingly noisy and electrically unstable.

Klein et al[90] have overcome these difficulties by the use of a logarithmic derivative coupled with the substitution of a multi-channel analyser for the lock-in amplifier as the detection device.

It has been found that optically forbidden and allowed transitions appear in IETS with roughly the same intensity[57]. This lifting of optical selection rules is particularly useful; for example, the S_o-T_1 transition in carotene has been

observed at around 1.3V. IETS electronic transitions may be correlated more directly with optical spectra of molecules in a thin film situation; the peak broadening observed in both these cases is interpreted as arising from the disordered environment of the molecule, and not from an intrinsic property of the technique.

5.9. The future potential of IETS techniques

As outlined in the previous sections, the thirteen years which have elapsed since the discovery of IETS by Jaklevic and Lambe have resulted in work which indicates the wide ranging potential applications of the method, although very few systems have been thoroughly investigated. In this section, some important fields of study are described which may lend themselves to investigation by IETS methods.

In the field of heterogeneous catalysis, the interaction between molecules adsorbed on to catalytic surfaces and the molecular orientation of the adsorbate relative to the substrate are ideal systems for study by IETS, particularly if the catalyst itself can be formulated on a metal electrode surface. However this condition need not be limiting. Hansma et al[72] have shown that metal-CO interactions can be studied with the metal catalyst distributed on the junction insulator surface. Magno et al[91] have shown that the deliberate oxidation of the lower electrode is not essential and have studied the interaction between the Al electrode and ethylene in this way. Hence a wide variety of heterogeneous catalytic systems, metallic, non-metallic, electrically insulating or conducting can now be studied by IETS, and catalytic studies are expected to provide one of the major applications of this technique in the near future.

Hansma[92] has suggested that useful information about lubrication in general and lubrication of aluminium in particular could be obtained from tunnelling spectroscopy. The molecules used as lubricant additives are typically in the range of 20 to 30 atoms which is ideal for high resolution spectra to be obtained, so defining the interaction of these additives with the surfaces to be lubricated. A wide variety of oxide supported metal films could also be investigated. Since the definition of lubricant-surface interaction obtained by IETS can be expected to be of a high order and because of the growing importance of efficient lubrication to an energy-conscious technology, this no doubt will provide another area of rapid development in the field of tunnelling spectroscopy.

Further developments may confidently be predicted in the use of IETS as an analytical tool. Some preliminary environmental pollution monitoring applications have already been reported in both air[10] and water phases[68,70,84]. The development of this application is critically dependent upon the possibility of doping completed junctions. An extension of the work of Jaklevic and Gaerttner[30,31] on the doping of completed junctions to a wider range of molecular species and to water-free systems would permit the mass production of undoped junctions similar in concept to disposable IR cells. This would allow the removal of the junction fabrication vacuum facility from the spectrometer itself, and would greatly facilitate the development of IETS as a routine analytical tool. Further applications may be foreseen in using IETS probes to trap interstellar molecules and to monitor the ozone layer, which may have the advantage

over other analytical methods as the IETS junction probe could be made millimeters in size and a few milligrams in weight.

Finally, work has already been carried out to show the application of IETS to adhesion science[86]. A logical development of these studies would be to investigate corrosion inhibition and protective paints on metal structures by similar IETS methods. The metal to be investigated could be supported on the tunnelling junction insulating barrier. By this means the interaction of the inhibitor or protective paint with the metal to be protected could be defined, as well as its long-term stability under a variety of conditions.

5.10. Comparison of IETS with some other spectroscopic techniques

Infrared spectroscopy has certainly been the most widely used technique for vibrational spectroscopy of adsorbed species[93]. Its advantages over tunnelling spectroscopy are that it utilises well-developed technology, it does not require cryogenic temperatures, and it can be applied to a wide variety of substrates. It has the disadvantages that, in general, only limited spectral regions can be studied, specifically almost no work has been done in the very important range below 1000 cm^{-1} because of the very strong absorptions of the substrates below this frequency; and its sensitivity is of the order of 10 times less than IETS. Raman spectroscopy has not been extensively used because of difficulties associated with fluorescence of the substrate[94].

A further technique applicable in this respect is electron energy loss spectroscopy[95]. In essense it involves measuring, to a precision of around 10meV, the energy loss of low energy electrons (typically 5eV) to the vibrations of adsorbed molecules. The advantages of EELS are that it can be used in-situ on metal surfaces that can also be studied by a variety of other surface spectroscopies such as LEED and Auger Electron Spectroscopy; and also there is no top metal electrode to influence the behaviour of the adsorbed species. It has disadvantages, however, that it is an extremely difficult technical task to achieve an energy resolution an order of magnitude poorer than that possible with tunnelling spectroscopy; and that charging effects make it difficult or impossible to apply EELS to adsorption on insulating surfaces.

It has already been shown that IETS could be applied to a variety of surface and analytical problems of some importance. The question must therefore be posed as to why this technique is not more widely known and utilised. The answer to this question probably lies in the relatively wide interdisciplinary expertise required at the present stage of development to construct the requisite electronic, cryogenic and high vacuum facilities. With a growing awareness of the potential applications of this technique coupled with a realisation that the hardware of IETS is well within the understanding and constructional capabilities of chemical laboratories, a rapid development of the chemical applications of inelastic electron tunnelling spectroscopy may now be envisaged.

Chapter 6. Some unresolved aspects of IETS

From the scope of this study, it becomes apparent that, since the discovery of inelastic electron tunnelling by Jaklevic and Lambe in 1966, a small group of laboratories, about ten in number, have developed the expertise and equipment to obtain tunnelling spectra. The technique is in its infancy, both in terms of development and interpretation, and, by the very nature of the way it has developed over the last thirteen years, is highly empirical. Thus it may appear ambitious to entitle a chapter "Some unresolved aspects of IETS". Nevertheless, there are certain areas where obvious developments may be expected and should be sought. These areas are outlined below.

As indicated in section 3.2., there is a dichotomy of opinion concerning the nature of the oxide formed in aluminium-oxide-lead junctions. This is unlikely to be resolved by tunnelling methods, where ambiguities in interpretation of frequencies can occur. The situation could be clarified however by the application of other surface techniques to the junction structures. In particular, the application of low energy electron diffraction and Auger electron spectroscopies to oxides of comparable thickness grown by glow discharge techniques in high vacuum systems on single crystals of aluminium should enable distinction between anhydrous and hydrated oxides. The incorporation into such a study of argon-ion bombardment equipment commonly used for oxide stripping under controlled conditions should enable depth

profiling of penetration by dopant molecules, and water molecules, into the bulk oxide to be observed, thus indicating whether IETS is a true surface technique.

Improvements in peak width may only be achieved at the present time by the reduction of the modulation voltage below 2mV, with the associated lowering of the signal to noise ratio, or by further lowering of the temperature below 4.2K. This involves pumping the liquid helium, and using this technique it is possible to reduce the temperature to 1.1K. However, this decrease in temperature is accompanied by a severe increase in expenditure due to extra losses of helium. It would therefore be of interest to ascertain whether a decrease in peak width could be achieved by any other method. One possibility is an investigation of the effect of a magnetic field upon the junction. The presence of such a field could lead to an alignment of dopant molecules on the surface, and a reduction in peak width due to differences in molecular environment.

There are several apparent improvements which could be introduced to the instrumentation itself. The introduction of an a.c. voltage follower or similar device to offset the loading effects of the junction on the oscillator would enable the range of junction resistances which are suitable for measurement to be extended to a level of a few ohms. This in itself would improve junction preparation reliability by enabling the observation of meaningful spectra from junctions whose resistances are too low for the present unprotected a.c. supply. The upper end of the resistance range could be extended by the incorporation of a signal averaging and enhancement system as well as or instead of

the lock-in as the detector. Such a system would result in the averaging out of noise signals over a large number of scans, thus enabling signals to be isolated from noise levels greater than those which limit the technique at present.

In terms of junction preparation, it seems unlikely that the reliability of junction electrode and oxide formation can be improved beyond the level attained in this study. It is desirable, however, to improve upon the production time of a junction, which is commonly of the order of 90 minutes. The majority of this time is accounted for by evacuation time of the chamber by the diffusion and rotary pump. It seems likely that the substitution of a turbomolecular pumping system for the present rotary-diffusion pump combination would result in a considerably improved turnaround time, enabling more junctions to be prepared in a given time period.

Since the construction and operating costs of this spectrometer are considerably less than the combined cost of an infrared and Raman spectrometer, there is a probability of the development of IETS as a commercially available technique. However, before this can become a reality, it would be necessary to find a method of mass producing junctions so that the vacuum equipment could be eliminated from a commercial prototype. This development would need to be accompanied by an extension of the infusion doping technique of Jaklevic and Gaerttner[30,31] to a wider range of dopant molecules, especially to non-aqueous systems.

Another aspect of tunnelling spectroscopy which should yield further information in the near future is the extension of the energy range above the levels attainable in present

junctions i.e. above 2.5V. This will only become practical with the growth and development of insulating barriers suitable for tunnelling spectroscopy and having a higher dielectric strength than those currently utilised. Thus experimentation is required not only with various barriers grown on aluminium, but also on different base electrodes, in order to enable the energy range to be extended further into the electronic transition region.

Only thirteen years have elapsed since the discovery of IETS. In that very short time, sufficient developments of the technique, and numerous applications have been reported to justify its claim to a place in the expanding field of surface spectroscopy.

--ooOoo--

References

1. I.Giaever and K. Megerle, Phys. Rev. 122 (1961) 1101.

2. J. Bardeen, L.N. Cooper and J.R. Schrieffer, ibid 108 (1957) 1175.

3. B.D. Josephson, Phys. Lett. 1 (1962) 251.

4. P.W. Anderson and J.M. Rowell, Phys. Rev. Lett. 10 (1963) 230.

5. R.C. Jaklevic and J. Lambe, Phys Rev. Lett. 17 (1966) 1139.

6. J. Lambe and R.C. Jaklevic, Phys. Rev. 165 (1968) 821.

7. D.J. Scalapino and S.M. Marcus, Phys. Rev. Lett. 18 (1967) 459.

8. A.F. Hebard and P.W. Shumate, Rev. Sci. Instr. 45 (1974) 529.

9. R.J. Jennings and J.R. Merrill, J. Phys. Chem. Solids 33 (1972) 1261.

10. J. Klein, A. Leger, M. Belin, D. Defourneau and M.J.L. Sangster, Phys. Rev. B, 7 (1973) 2336.

11. W.A. Harrison, Phys. Rev. 123 (1961) 85.

12. R.C. Jaklevic and J. Lambe, Tunnelling Phenomena in Solids (Plenum Press, 1969) Chapter 18.

13. J. Bardeen, Phys. Rev. Lett. 6 (1961) 57.

14. M.H. Cohen, L.M. Falikov and J.C. Phillips, Phys. Rev. Lett. 8 (1962) 316.

15. B.D. Josephson, Phys. Lett. 1 (1962) 251.

16. A.J. Bennett, C.B. Duke and S.D. Silverstein, Phys. Rev. 176 (1968) 969.

17. L.I. Schiff, Quantum Mechanics (McGraw Hill, 3rd Edn. 1968) p.314.

18. J. Kirtley, D.J. Scalapino and P.K. Hansma, Phys. Rev. B 14 (1976) 3177.

19. M. Pirani, Verk. dt. Phys. Ges., 8 (1906) 686.

20. F.M. Penning, Physica, Eindhoven 4 (1937) 71.

21. J.L. Miles and P.H. Smith, J. Electrochem Soc., 110 (1963) 1240.

22. N.F. Mott, Trans. Faraday Soc.43 (1947) 429.

23. R. Magno and J.G. Adler, Phys. Rev. B., 13 (1976) 2262.

24. J.T. Tillery, Trans. Inst. Metal Finishing, 40 (1963) 28.

25. P. Hoffman and J.R. Schumpelt, Ber. Deut. Chem Gesells, 49 (1916) 303.

26. N.J. Upton, J. Electrodepositors Tech. Soc., 22 (1946) 64.

27. M.G. Simonsen and R.V. Coleman, Phys Rev. B., 8 (1973) 5875.

28. M.G. Simonsen, R.V. Coleman and P.K. Hansma, J. Chem. Phys., 61 (1974) 3789.

29. B. Wallace, Electronic Design, 14 (1974) 110.

30. R.C. Jaklevic and M.R. Gaerttner, Appl. Phys. Lett., 30 (1977) 646.

31. R.C. Jaklevic and M.R. Gaerttner, Appl. Surf. Sci., 1 (1978) 479.

32. P.K. Hansma, D.A. Hickson and J.A. Schwartz, J. Catalysis, 48 (1977) 377.

33. J.W. Newsone, H.W. Heiser, A.S. Russell and H.C. Stumpf, Alumina Properties (Aluminium Company of America, Pittsburgh, Pa., 1960).

34. A.B. Kurzaev, S.N. Kozloz and V.F. Kiselev, Doklady Phys. Chem., 228 (1976) 526.

35. I. Giaever, N. Hart and K. Megerle, Phys. Rev., 126 (1962) 941.

36. J.G. Adler, T.T. Chen and J. Straus, Rev. Sci. Instr. 42 (1971) 362.

37. C.P. Poole, Electron Spin Resonance (Interscience, N.Y., 1967) pp. 387,775.

38. T.C. O'Haver and G.L. Green, Amer Lab. 7 (1975), 15.

39. J.E. Houston, Rev. Sci. Instr. 45 (1974) 897.

40. W.R. Patterson and J. Shrewchun, Rev. Sci. Instr. 35 (1964) 1704.

41. D.E. Thomas and J.M. Rowell, Rev. Sci. Instr. 36 (1965) 1301.

42. J.G. Adler and J.E. Jackson, Rev. Sci. Instr. 37 (1966) 1049.

43. R.J. Jennings and J.R. Merrill, J. Phys. Chem. Solids 33 (1972) 1261.

44. R.E. Simpson, Introduction to Electronics for Scientists and Engineers, (Allyn and Bacon, Boston, 1974) p. 394.

45. A.F. Hebard and P.W. Shumate, Rev. Sci. Instr. 45 (1974), 529.

46. "Coil design and construction manual", Bernards (Publishers) Limited, London 1960 p.42.

47. R. Blackford, Rev. Sci. Instr. 42 (1971) 1198.

48. D. Lewin, Theory and Design of Digital Computers, (Nelson, London, 1972) p. 279.

49. D.W. Hill, Principles of Electronics in Medical Research, (Butterworths, London, 1965) p.315.

50. P.N. Shott and B.O. Field, Spectrochem. Acta 35A (1979) 301.

51. A.L. Geiger, B.S. Chandrasekher and J.G. Adler, Phys. Rev. 188 (1969) 1130.

52. G.A. Dorsey, Jr., J. Electrochem. Soc. 113 (1966) 169.

53. G.A. Dorsey, Jr. J. Electrochem. Soc. 113 (1966) 284.

54. B.F. Lewis, M. Mosesman and W.H. Weinberg, Surf. Sci. 41 (1974) 142.

55. W.M. Bowser and W.H. Weinberg, Surf. Sci. 64 (1977) 377.

56. R. Magno and J.G. Adler, Thin Solid Films 42 (1977) 237.

57. S.de Cheveigne, J. Klein, A. Leger, M. Belin and D. Defourneau, Phys. Rev. B. 15 (1977) 750.

58. L. Pauling, "The Nature of the Chemical Bond", Cornell University Press, Ithaca, New York (1960) p. 514.

59. R.M. Handy, Phys. Rev. 126 (1962) 1968.

60. J. Kirtley and P.K. Hansma, Phys. Rev. B. 12 (1975) 531.

61. P.E. Cade, J. Chem. Phys. 17 (1967) 2390.

62. H. Spedding and D.H. Whiffen, Proc. Roy. Soc. A238 (1956) 245.

63. A.E.T. Kuiper, J. Medema, and J.J.G.M. Van Bokhoven, J. Catalysis 29 (1973) 40.

64. R.C. Jaklevic and J. Lambe, Surf. Sci. 37 (1973) 922.

65. A. Leger, J. Klein, M. Belin and D. Defourneau, Solid State Commun. 11 (1972) 1331.

66. A. Adane, A. Fauconnet, J. Klein, A. Leger, M. Belin and D. Defourneau, Solid State Commun. 16 (1975) 1071.

67. J.D. Langan and P.K. Hansma, Surf. Sci. 52 (1975) 211.

68. N.I. Bogatina, I.K. Yanson, B.I. Verkin and A.G. Batrak, Sov. Phys. JETP 38 (1974) 1162.

69. N.I. Bogatina, Opt. Spectrosc. 38 (1974) 43.

70. Y. Skarlatos, R.C. Barker, G.L. Haller and A. Yelon, J. Phys. Chem. 79 (1975) 2587.

71. A.E.T. Kuiper, J. Medema and J.J.G.M. Van Bokhoven, J. Catal. 29 (1973) 40.

72. P.K. Hansma, W.C. Kaska and R.M. Laine, J. Am. Chem Soc. 98 (1976) 6064.

73. J.G. Adler, Solid State Commun. 7 (1969) 1635.

74. R.C. Jaklevic and J. Lambe, Phys Rev B2 (1970) 808.

75. W.M. Bowser and W.H. Weinberg, Surf. Sci. 64 (1977) 377.

76. D.G. Walmsley, I.W.N. McMorris and N.M.D. Brown, Solid State Commun. 16 (1975) 663.

77. S. Hayashi and J. Umemura, J. Chem. Phys. 60 (1974) 2630

78. J.T. Hall and P.K. Hansma, Surf. Sci. 71 (1978) 1.

79. W.P. Griffith and A.J. Wickham, J. Chem Soc. A(1) (1969) 834.

80. P.S. Bailey, S.S. Bath and J.B. Ashton, Adv. Chem. Ser. 21 (1959) 143.

81. H.R. Eisenhauer, Water Pollut. Control Fed. 40 (1962) 1887.

82. J. Walachova, in "Prac. Konf. Cezk. Fyz. Pr., 3rd, 1973", M. Matyas (Ed.), Acedemia Prague (1974) p. 154.

83. Y. Skarlatos, R.C. Barker, G.L. Haller and A. Yelon, Surf. Sci. 43 (1974) 353.

84. P.K. Hansma and R.V. Coleman, Science 184 (1974) 1369.

85. C.S. Korman and R.V. Coleman, Phys. Rev. B 15 (1977) 1877.

86. H.W. White, L.M. Godwin and T. Wolfram, J. Adhesion 9 (1978) 237.

87. P.K. Hansma and M. Parikh, Science 188 (1975) 1304.

88. J.T. Hall, P.K. Hansma and M. Parikh, Phys. Rev. A. 14 (1976) 1437.

89. J.M. Clark and R.V. Coleman, Proc. Natl. Acad. Sci. 73 (1976) 1598.

90. A. Leger, B. Delmas, J. Klein and S.de Cheveigne, Rev. Phys. Appl. 11 (1976) 307.

91. R. Magno, M.K. Konkin and J.G. Adler, Surf. Sci. 69 (1977) 437.

92. P.K. Hansma, in "Inelastic Electron Tunnelling Spectroscopy", T. Wolfram (Ed.), Springer-Verlag, Berlin, p. 20.

93. L.H. Little, "Infrared Spectra of Adsorbed Species", Academic Press. N.Y. (1966).

94. T.A. Egerton, A.H. Hardin, Y. Kozirovski and N. Shepard, J. Catalysis 32 (1974) 343.

95. H. Ibach, J. Vac Sci. Technol. 9 (1972) 713.