STUDIES OF THE WORK FUNCTION OF RARE EARTH AND OTHER METAL FILMS

A thesis submitted to the Faculty of Science of the University of London for the Degree of Doctor of Philosophy

by

Nicholas Chandler B.Sc. (Lond)

Department of Physics
Royal Holloway College

October 1978

Studies of the Work Function of Rare Earth and other Metal Films

by

Nicholas Chandler

A B S T R A C T

Rare earth metal films have been prepared in Ultra-high Vacuum (U.H.V.) and work function measurements made, in situ, by a vibrating capacitor technique. The effects of aging in U.H.V. and of atmospheric contamination have been studied.

Measurements have been made on Gd above and below the Curie temperature, before and after annealing at 390°K by means of a novel nichrome-film heater (developed for U.H.V. use).

The resistivities of the rare earth films were determined by an in situ potential probe method in conjunction with interferometric thickness measurements.

Work function measurements have also been made on other metal films and on silicon surfaces to provide data relevant to theoretical work on point-contact detectors for far infra-red radiation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS USED

This list excludes those symbols used only in the appendices and those which only appear once in the text where they are explained.

A = $4\pi mk^2e/h^3$ = Richardson constant

A_w = whisker tip radius

B = $[4\pi\Delta s/h](2m)^{\frac{1}{2}}$ in Simmons' equation for tunnel resistivity

B_w = Bandwidth (of electrical measurements)

C = a constant (photoelectric emission, equation (2.6))

C_B = barrier capacitance of a point-contact diode

C_o = Kelvin capacitance, at rest

C_R = Input capacitance of detecting circuit, including lead capacitance

E = electric field strength

E_F = Fermi Level

E_{FA} = Fermi Level of metal A

E_{FB} = Fermi Level of metal B

E_{F1} = Fermi Level of electrode 1

E_{F2} = Fermi Level of electrode 2

F = Helmholtz Free Energy

$H(x)$ = step function = $0, x \leq 0$
= $1, x > 0$

I = first ionisation potential (of a gas)

I_D = d.c. component of current in the rectifying diode

J = saturated current density

J_o = saturated zero-field current density

$J_{1,2}$ = Tunnel currents with electrode of lower work function negatively and positively biased respectively

$J'_{1,2}$ = Thermionic currents with electrode of lower work function negatively and positively biased respectively

K = a constant in the evaluation of the Laplace equation

M = dipole moment (on surface) due to adsorbed atom and its image force

N = no. of unit cells in the repeating segment of a region in the Brillouin Zone

P	=	R.F. power coupled into a point contact diode
P_B	=	power absorbed in the barrier resistance
Q	=	total charge in circuit
Q_B	=	charge in bias circuit
R	=	radius of annulus contributing to current flow at thickness d (Figure 7.8)
R_B	=	barrier resistance of a point-contact diode (Figure 7.12)
R_I	=	input resistance of detecting circuit
R_s	=	spreading resistance in series with the barrier of a point-contact diode (Figure 7.12)
S	=	entropy
T	=	absolute temperature
U	=	total internal energy
V	=	bias applied to the diode
V_a, V_A, V_B	=	electrostatic potentials just outside metals, in vacuo
V_{AB}	=	c.p.d. between conductors A and B
V_{ABi}	=	c.p.d. between opposing patches
V_c	=	electrostatic potential just inside a metal
V_D	=	detected voltage
V_E	=	external (bucking) potential
V_{eff}	=	effective potential near metal surface ("jellium", Linear Potential Model)
V_{es}	=	electrostatic potential
V_i	=	electrostatic potential at a point just outside the i^{th} patch of a metal
V^l	=	electrostatic potential at a distance large compared to the patch size
V_R	=	voltage developed across detector input
V_R^ω	=	component of V_R at the fundamental frequency (ω)
W_v	=	vacuum level, relative to bottom of conduction band
X, Y	=	periods (in x and y) of periodically distributed work functions on the surface of the Kelvin electrode (s)
a, b, c	=	constants (Equation(7.12))
a'	=	position of metal surface, for overall charge neutrality (Linear Potential Model)
a_c	=	common area of the vibrating and fixed electrodes

a_n, b_n, c_m, d_m, g = constants in the evaluation of the Laplace equation

b' = a constant, (photoelectric emission, equation (2.7))

$\underline{b_1}, \underline{b_2}, \underline{b_3}$ = lattice vectors in the Brillouin Zone

c_i, c_{io} = capacitances of capacitor formed by the common area of patches on opposing electrodes, dynamic and static respectively

d_i, d_{io} = spacings of same

d = thickness of insulator layer

d_o = separation of Kelvin capacitor, at rest

e = electronic charge

f = a universal function of $(h\nu - h\nu_o)/kT$ used for photoelectric emission

f_i = fractional area of the i^{th} patch

h = Planck constant

i_s = saturation current (i.e. all emitted electrons are collected)

k = Boltzmann constant

k_F = wave number at the Fermi level

l_1, l_2 = dimensions of an electrode

m = electron mass

n = number of electrons in a body

n_a = surface density of adsorbed ions

n_e = electron density

$\bar{n}_e$ = mean electron density

n_t = total charge density

n_+ = charge density of positive uniform background ("jellium" model)

r = radius of annulus contributing to current flow at thickness t (Figure 7.8)

$\bar{r}_e$ = zero-field reflection coefficient for incident electrons

r_s = Wigner-Seitz radius

s = insulator layer thickness of plane-parallel pair and minimum barrier thickness of hemisphere/plane pair

s_i = actual area of patches of the i^{th} type

s_1, s_2 = limits of the barrier at the Fermi level (see Appendix I)

t	=	thickness of the insulator layer for an incremental ring radius, r, on the hemisphere (Figure 7.8)
v	=	volume of a body
w	=	fraction of total zero-field thermionic emission from all patches with $\phi_i^* < \overline{\phi}$
w_i	=	fractional contribution of the i^{th} patch to total zero-field thermionic emission current
Σ	=	electron affinity of a elemental gas
α_+	=	positive ionisation current
β	=	field factor (field emission, paragraph 2.4)
γ	=	voltage responsivity of a detector diode
ε_o	=	absolute permittivity of free space
ε_r	=	dielectric constant (relative permittivity)
ε_{xc}	=	exchange-correlation energy per particle for the uniform electron gas
η	=	rectification ratio of a diode
$\theta(x)$	=	Nordheim elliptic function
λ	=	a parameter associated with the image forces, $= e^2 \ln 2/8\pi\varepsilon_r\varepsilon_o s$
μ	=	chemical potential
$\overline{\mu}$	=	electrochemical potential
μ_{xc}	=	exchange-correlation contribution to chemical potential
ϕ, ϕ_1, ϕ_2	=	work functions
ϕ_a	=	work function with adsorbed ion
$\overline{\phi}$	=	average work function of a patchy surface
ϕ_i	=	work function of the i^{th} patch
ϕ_{eff}	=	effective value of ϕ measured by the Kelvin technique
ϕ^*	=	apparent work function measured by thermionic emission
ϕ^{**}	=	apparent work function of a patchy surface at high field
$\overline{\phi}^{**}$	=	apparent work function of a patchy surface at weak field
ϕ_p^*	=	apparent work function measured by positive surface ionisation, for a uniform surface

χ = electron affinity of a bulk semiconductor

$\psi(x)$ = barrier height as a function of distance x from one electrode

$\psi_{1,2}$ = barrier heights at interfaces between insulator and metals 1,2 (image forces omitted)

$\overline{\psi}_{1,2}$ = mean barrier heights in the equations for J_1, J_2 respectively

ω = angular frequency

ω_c = corner frequency (upper -3dB frequency) in the response of the detector (equation(7.18))

Δs = barrier thickness at Fermi level = $s_2 - s_1$

ΔV = difference in electrostatic potential between surface and bulk metal

$\Delta \phi$ = $\phi_2 - \phi_1$

δr = width of annulus at radius r

δz = amplitude of oscillation of Kelvin capacitor

$\delta \phi$ = surface dipole barrier

Introductory Remarks

The principal objective of the work described here was to extend the study of rare earth films to measurement of their work functions. Previous values had been obtained from only a few samples of gadolinium and by techniques which were subject to significant errors. Following the introduction to the work function and its role in physics in Chapter I, the range of available measurement techniques is described in Chapter II, and reasons are given for the choice of the vibrating capacitor method used in this work. Chapter III details the experimental arrangement and considers the sources of error. The preparation of the specimens is described in Chapter IV and the development of a novel film heater in Chapter V. The results are reported and discussed in Chapter VI. A secondary objective arose within the laboratory to construct and explain the behaviour of point-contact diodes as detectors of far infra-red radiation. The resulting model and the importance of the work functions of the materials, together with the experimental results, are described in Chapter VII. Chapter VIII concludes the thesis and includes some suggestions for further work.

Chapter I The Role of the Work Function

1.1 Introduction.

The term "work function" appears first to have been used in the early part of the century in connection with the thermionic emission of electrons from a heated metal cathode.[1] Even then the concept of a certain amount of work done in causing an electron to escape from a cathode was not novel, and most of the earliest work concentrated either on improving the life and efficiency of electric lamps or on gaining experimental knowledge of the electron physics of metals to compare with the emerging theories. Other means of emitting electrons were conceived, such as applying high electric fields or electromagnetic radiation. It was the interpretation of the photoemission of electrons which helped establish the quantum theory of light. Classical physics could not explain the failure to emit electrons when the incident radiation was of too low a wavelength, whereas the localised and quantised photon energy was able to do so, in terms of the photon energy being smaller than the work function of the emitter. The developments in several other technological and scientific fields have depended on a knowledge of the work functions of various materials and their variations under different conditions. These will be considered after the work function and some related parameters have been more fully described.

The methods of measuring the work function have been as numerous as the topics to which it is important. Chapter II describes the major techniques in brief, with particular reference to conditions suitable for the highly reactive rare earth metals. With the advances in Ultra High Vacuum (U.H.V.) technology and the development of experimental surface techniques, there has been a renewal of interest in work function as one of the basic properties of surface physics. Specimens of ultra-high purity may now be examined under unprecedentedly clean conditions and the effects of specific treatments or contaminants monitored by a variety of techniques.

1.2 Definition of the Work Function.

The work function depends on some fundamental physical parameters which must first be explained. The Helmholtz free energy, F, of an isolated body containing electrons is given by

$$F = U - TS \qquad \text{equation (1.1)}$$

where U = total internal energy
T = temperature
S = entropy.

A convenient thermodynamic parameter to describe the system is termed the "electrochemical potential" of an electron therein and is defined by[2]

$$\overline{\mu} = \left(\frac{\partial F}{\partial n}\right)_{T,v} \qquad \text{equation (1.2)}$$

where n = no. of electrons
and v = volume of system.

To avoid ambiguity, the conventional definitions are followed where S = 0 at T = 0 and the energy of an electron at rest at infinite distance is also zero. Qualitatively, the electrochemical potential is the work done when an electron is transferred from a point in vacuo at infinity to a point in the bulk of the real material. This definition is not dependent solely on the internal state of the body, but also on surface and external conditions. For example, a change in electrostatic potential, V_{es}, of the body will change the electrochemical potential, so a second term, the "chemical potential", μ, is defined by[2]

$$\mu = \overline{\mu} + eV_{es} \qquad \text{equation (1.3)}$$

where e = electronic charge.

This quantity is a function only of the internal state of the body

and is independent of surface and external conditions. Qualitatively, it is the work done in transferring an electron from a point in vacuo at infinity to a point in an (unreal) homogeneous bulk of material equal to the real material except for the absence of a surface dipole[3]. It is due entirely to the interactions of the electron with the bulk of the material.

The work function ϕ of a uniform surface of an electronic conductor, conceptually described as the work required to remove an electron from within the material to a point just outside it, and at rest, can now be defined in terms of the electrochemical potential just inside and the electrostatic potential energy just outside, V_a,[2]

i.e. $$\phi = -V_a - \frac{\bar{\mu}}{e} = V_c - V_a - \frac{\mu}{e} \qquad \text{equation (1.4)}$$

where V_c = electrostatic potential inside the conductor

Since μ is solely a volume property, its contribution to the total work function of a material is reflected in its alternative name "inner work function".

From equation (1.4) the total work function may be regarded as comprising two parts: the inner work function and a surface effect. The latter results from the discontinuity of the lattice at the surface. The electrons are no longer distributed symmetrically about the ion cores of the surface atoms, but spread out into the vacuum. The redistribution of electrons occurs both normal to the surface and parallel to it, in such a way as to minimise the overall energy of the whole system. The former has been referred to as "spreading"and the latter as "smoothing"[4]. The spreading of the electrons outwards from the surface leaves a corresponding positive charge, hence a dipole or "double layer" which increases the work function since it tends to prevent electrons escaping from the bulk of the material.On

an atomic scale the lattice surface is, of course, not smooth, and the electrons are redistributed so as to minimise the surface area, and hence the energy, of the electron cloud just outside the surface. This is equivalent to a reduction in the polarisation caused by the spreading, and therefore tends to reduce the work function. Numerical calculations for (b.c.c.) tungsten[4] showed the spreading effect to be essentially isotropic, but that the smoothing caused a significant variation of work function with crystal orientation. The results agreed qualitatively with the experimental observation that there is a strong tendency for the crystal faces on which the surface atoms are more densely packed to have the higher work functions.

In addition to the surface contribution anisotropy, the volume contribution, i.e. the inner work function, may vary with crystal orientation. This is due to the electronic band structure[2], which determines the electron density and hence the exchange and correlation energies. Further discussion will appear in Chapter VI, where the work function anisotropy of gadolinium will be quantitatively estimated from the band structure.

It is to be expected that all materials for which the work function can be defined will exhibit some anisotropy. This has indeed been observed experimentally for a wide range of metals, tungsten, molybdenum, nickel, zinc, silver, copper and bismuth[2] to mention but a few, and for the semiconductors silicon and germanium[1].

A polycrystalline specimen will therefore consist of regions of different work function, depending on the crystal orientation at the surface of each grain, or "patch". The effect on the experimentally observed values for different methods of measurement will be discussed further in Chapters II and III.

1.3 The Relations between Work Function and Other Properties.

Although the term work function is used in connection with

several electron emission properties of materials, the quantities measured by observing the various phenomena are not identical. The discrepancies between the true and the apparent work functions of metals and semiconductors measured by various techniques will be discussed in Chapter II. One group of techniques which does not suffer from such discrepancies makes use of the identity of the electrochemical potential in two conductors, or semiconductors, which are in thermal and electrical equilibrium. By inspection of equation (1.4) it is evident that there will exist a difference of potential between a point just outside material A and one just outside material B given by

$$V_A - V_B = \phi_B - \phi_A = V_{AB} \qquad \text{equation (1.5)}$$

and which is called the "contact potential difference" (c.p.d.). Some early work[5, 6] refers to "Volta potential" to describe the same effect, after its discoverer. Due to the anisotropy of the work function there will be a c.p.d. between various crystal faces of a single material. The consequences of such variations on the measurements between polycrystalline surfaces will be discussed in detail in Chapter III. Equation (1.5) is valid for contacts between metals and semiconductors. Indeed, techniques based on c.p.d. measurements are unique in measuring the true work function of semiconductors[1].

In many texts the work function is derived in terms of the familiar Fermi level. At absolute zero, when the most energetic electron in a metal resides at the Fermi level, E_F, the work function is simply the difference between that level and the energy of an electron at rest just outside the surface, the vacuum level, W_V, both energies being measured from the bottom of the conduction band rather than from infinity[7], as shown in Figure (1.1)

$$\phi = W_V - E_F \qquad \text{equation (1.6)}$$

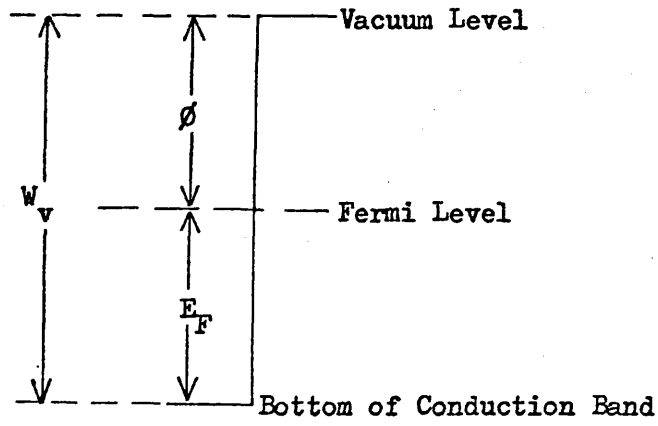


Figure 1.1 Energy Level Diagram of Electrons in a Metal

In this scheme it is the Fermi levels of the two metals (A and B) which become identical when they are in thermal and electrical equilibrium, as shown in Figure 1.2.

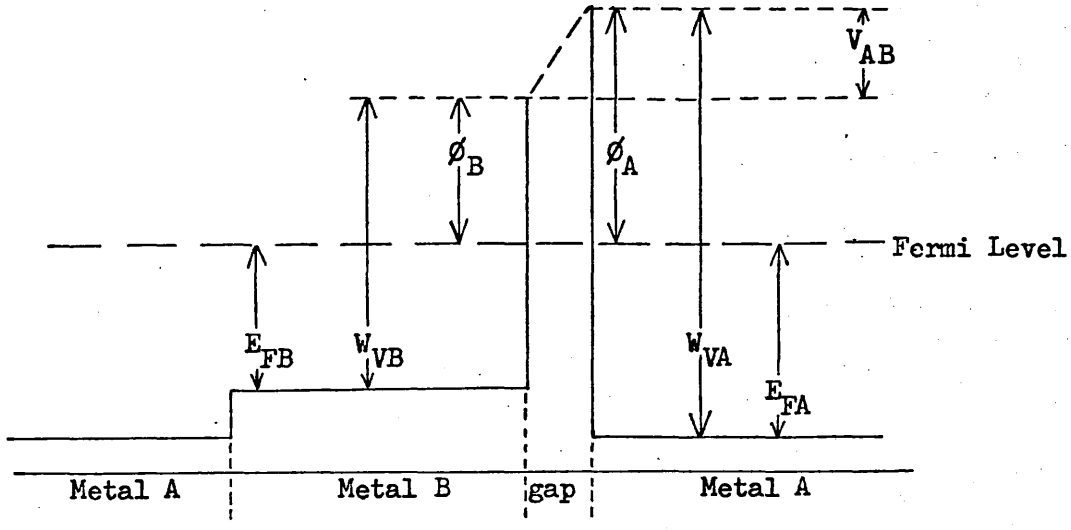


Figure 1.2 Energy Level Diagram of Electrons in Two Metals in Equilibrium

The work function of a semiconductor may be derived analagously, with the same reference point, but the energy just outside the surface is now termed the "electron affinity", χ , and the Fermi

level is below the conduction band, and therefore negative. The relation is[7]

$$\phi = \chi - E_F \qquad \text{equation (1.7)}$$

as shown in Figure 1.3.

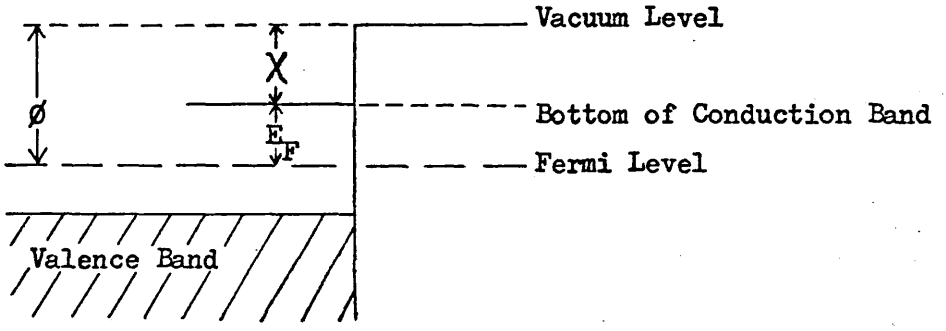


Figure 1.3 Energy Level Diagram of Electrons in a Semiconductor

The above electron affinity, χ, should not be confused with the elemental electron affinity, Σ, which is defined as the energy required to neutralize a stable, singly negatively ionised atom, often free but sometimes in an ionic lattice[8]. The elemental electron affinity, Σ, and the first ionisation potential, I, which is the energy required to produce an ion-electron pair, are interesting parameters in that they form the bases of experimental methods for measuring work function. The methods involve measuring the positive and/or negative ionisation currents emitted from a solid surface in equilibrium with a gas for which I and/or Σ are known. However, these techniques are far from simple in practice, especially for semiconductors, due to their large band-gaps[1].

The change in the work function of a surface caused by adsorption of a foreign species is called the surface potential (of that combination). The magnitude of the surface potential depends on the number of dipoles per unit area of the surface and on their magnitude

If the adsorbed species is ionised, then a dipole of moment M is produced by the ion and its image charge. For a surface density of adsorbed ions n_a, the work function of the material after adsorption is given by[7]

$$\phi_a = \phi - \frac{n_a M}{\varepsilon_o} \qquad \text{equation (1.8)}$$

For positive ions $M > 0$, so the work function is reduced and for negative adsorbate ions it is increased. This result is strictly valid only for coverages significantly less than a monolayer, since as the coverage increases it is energetically more favourable for the adsorption to be as neutral atoms rather than as ions[7].

1.4 The Relevance of the Work Function to Operating Devices and Equipment.

As mentioned earlier, a knowledge of the work function has been useful in many spheres of technological and scientific activity. Although it is beyond the scope of this work to examine all the difficulties and ambiguities of these topics, it is interesting to consider the role of the work function in application.

Probably the oldest and most familiar phenomenon to which the work function relates is thermionic emission. The current density for thermionic emission is dependent on the work function and on the temperature (see Chapter II). Unfortunately the work functions of metals are quite high, especially those metals which will withstand the high temperatures required for high current density. The work function may be lowered by the adsorption of positive ions on the surface, as described by equation (1.8). A greater reduction in work function is achieved by the inclusion of a monolayer of oxygen between the base material and the foreign species, with the added advantage of better stability at high temperatures. Barium, caesium and thorium are commonly used in this way to reduce the work function of tungsten cathodes. For example, a cathode with a W-O-Th surface has a work func-

tion of only 2.7 eV, compared to 4.5 eV for clean tungsten, and is over 200 times as emissive at 2000°K[9]. Still further improvements have been made by using a mixture of barium, calcium and strontium oxides on nickel, resulting in a work function as low as 1.4 eV for the optimum mixture[7]. Apart from the obvious improvements in stability and efficiency to be gained by operating cathodes at lower temperatures, the reduction in noise for electronic applications was useful, although this is now less important due to the development of solid state devices. The improved stability and reliability of cathodes at higher operating pressures has benefited mass spectrometers and other users of ion sources.

The efficiency of the direct conversion of thermal energy to electricity depends on the work functions of both cathode and anode of a vacuum diode[7]. To minimise heat losses by radiation and conduction, the electrodes should be as cool as possible, consistent with sufficient net current, so the cathode work function should be low but the difference between the two work functions should be large. Obviously a compromise must be reached in the design of an efficient converter.

Photoemission is also dependent on the work function of the emitter. The work functions of most metals are too high for photoemission in the visible region. The alkali metals have low work functions, but even these are not generally used since their quantum efficiency is low due to the high reflectivity of the metal surfaces. By methods very similar to those used for thermionic emitters, the efficiency and range of applicability (i.e. wavelengths at which photoemission occurs) may be improved. However, it is more common to use photoconductive materials for detectors, although of course photomultipliers must still use emitters. In this application only a very small photoemitted current is required, albeit at the expense of a long chain of dynodes. The secondary-electron emission from the dynodes is dependent on the work function of the material used[7].

The emission of electrons in the presence of a high field strength permits [illegible] crystallographic studies of surfaces at low temperatures, whereas thermionic emission requires relatively very high temperatures which would in many cases destroy or prevent the feature which it is desired to observe. The Field Emission Microscope consists basically of a fine tip, mostly tungsten to date, in order to produce a high field strength, and a phosphorescent screen on which to observe the pattern of emission. The emission current from each crystal grain on the tip surface depends on its work function (see Chapter II), so that a pattern is produced which copies the work function distribution of the surface, allowing observation of the crystal structure, its change due to heat treatment, and the adsorption of foreign atoms or molecules, including surface diffusion and nucleation studies of deposited films[3]. An alternative technique for surface studies of adsorption mechanisms from work function measurement is based on the surface potential described by equation (1.8). The value obtained is that averaged over the whole surface rather than showing the variation from point to point, so that either the degree of coverage of an adsorbant by an adsorbate or the dipole moment of the adsorbant-adsorbate bond may be measured provided that the other parameter is known. At low coverages the value measured by techniques which employ an electric field for electron collection cannot simply be related to the dipole moment[10]. This is an example where c.p.d. can provide a true value measurement and a simple, versatile technique unlimited by otherwise common difficulties such as are associated with filaments, e.g. high temperatures, non-zero field, and perturbation of the system by the process of measurement.

The c.p.d. between metals or between a metal and a semiconductor is also important to the behaviour of a third type of photodetector. The rectifying characteristics of a device consisting of metal-insulator-metal (M.I.M.) or metal-insulator-semiconductor (M.I.S.) depend on the absolute and relative values of the work functions of the two electrodes, since they determine the tunnel and thermionic currents through the intermediate barrier. Detectors of this type are operational up to very high frequencies due to the extremely low

junction capacitance afforded by point-contact construction. Further description of the fabrication and performance of such diodes for use as far infra-red detectors and a new model which explains their behaviour will be given in Chapter VII.

The c.p.d. between a metal and a semiconductor plays a significant role in the performance of many other M.I.S. (more usually called MOS) or metal-semiconductor (M.S.) contacts. To connect the semiconductor to external circuitry it is desirable to use a contact which is ohmic, i.e. which produces no rectification or other non-linear effects, by allowing free transfer of charge between the two materials. In order to equalise the Fermi levels in the two materials when contact is made, electrons flow from one to the other, the direction being dependent on the relative magnitude of their work functions. If the electron flow is in such a direction as to increase the number of majority carriers (i.e. from metal to n-type semiconductor or from p-type semiconductor into metal) then the barrier between the materials is negligibly low and thin, resulting in an ohmic contact. Electron flow of the opposite polarity produces a depletion layer wholly in the semiconductor, resulting in rectification. This simple theory predicts ohmic contacts when the metal work function is smaller than that of an n-type or larger than that of a p-type semiconductor. Practical results do not always follow this rule because the physical discontinuity produces energy levels at the surface of the semiconductor, called surface states. In some cases the electron re-distribution cannot neutralize the effect of these states, so a barrier remains and causes rectification. There are, fortunately, several means of ensuring ohmic contact, including producing a high impurity concentration at the semiconductor surface (either directly or by diffusion of the contacting metal) and producing a high rate of electron-hole recombination near the semiconductor surface by mechanically or chemically damaging the surface before metallisation[11].

Chapter II Methods of Measuring the Work Function

2.1 Introduction.

This chapter describes briefly the major experimental techniques for measuring true or apparent work function, preliminary to the detailed description of the method used for the present work, given in Chapter III. Reasons are given for the exclusion of specific techniques from consideration for use for rare earth metals.

2.2 Thermionic Emission.

The current density J emitted from a conductor of uniform surface at temperature T is given by the Richardson-Dushman equation

$$J = A(1-\bar{r}_e)T^2 \exp(-e\phi/kT) \qquad \text{equation (2.1)}$$

where

$A = 4\pi mk^2e/h^3$ = Richardson constant = $120.4\ A/cm^2/^oK^2$
m = electronic mass
k = Boltzmann constant
e = electronic charge
h = Planck constant
$\bar{r}_e$ = zero-field reflection co-efficient for incident electrons
ϕ = true work function,

assuming that the applied field necessary to produce saturation (i.e. all the emitted electrons are collected) is so weak that its value approaches zero. When the field E is not weak, the work function is reduced[1] by an amount

$$\Delta\phi = 3.79 \cdot 10^{-4} E^{\frac{1}{2}} \text{ eV} \qquad \text{equation (2.2)}$$

This effect has been named after Schottky.

Experimentally the saturated current density is measured as a function of field strength for a series of fixed emitter temperatures and plotted as ln J vs. $E^{\frac{1}{2}}$. Extrapolation to zero-field yields values of J_o, the saturation zero-field current density. A graph of $\ln(J_o/T^2)$ vs. $(1/T)$ is plotted, the gradient, by analogy with equation (2.1) is

$$\frac{d \ln(J_o/T^2)}{d(1/T)} = -e\phi^*/k, \qquad \text{equation (2.3)}$$

where ϕ^* is the "apparent work function" and is related to ϕ since J_o must satisfy equation (2.1). Substitution of (2.3) into (2.1) gives

$$\phi^* = \phi - T\left(\frac{d\phi}{dT}\right) - \frac{kT^2}{e(1-\bar{r}_e)} \cdot \frac{d\bar{r}_e}{dT} \qquad \text{equation (2.4)}$$

The values for $\bar{r}_e$ and $\frac{d\bar{r}_e}{dT}$ have been shown to be small for tungsten[2], but cannot necessarily be assumed to be so for rare earth metals.

In general, the surface of an emitter will not be uniform, as has so far been assumed, but will have patches of different work functions. The contribution of an individual patch to the total emitted current will vary in a complex way, depending on its relative size and its work function relative to the average over the whole surface.

Riviere[1] has shown that the apparent work function at high field strenth is

$$\phi^{**} = \sum_i w_i \phi_i^* \qquad \text{equation (2.5a)}$$

where the subscript denotes the i^{th} patch, and w_i is the fraction of

the total zero-field current carried by that patch, and at weak field strength is

$$\overline{\phi}^{**} = \sum_{i \geqslant j} w_i \phi_i^{*} + w\,\overline{\phi} - T\frac{d\overline{\phi}}{dT} \qquad \text{equation (2.5b)}$$

where $\phi_i^{*} \leqslant \overline{\phi}$ for $i < j$

$$\overline{\phi} = \sum_i f_i \phi_i$$

f_i is the fractional area occupied by the i^{th} patch, and
w is the fraction of the total zero-field emission from all patches with $\phi_i^{*} < \overline{\phi}$

The inestimability of w_i renders an exact relation between ϕ^{**} or $\overline{\phi}^{**}$ and $\overline{\phi}$ impossible, so that the comparison of $\overline{\phi}$ with the results of other workers is only possible for known crystal orientations in single crystal samples. For such samples, the work function temperature coefficient, the electron-reflection coefficient and its temperature dependence (see equation (2.5b)) must be measured or shown to be negligible.

Because of the extreme reactivity of the rare earth metals, the chosen technique would be required to be as rapid as possible, which would possibly also be helpful in determining the effects of aging. High purity specimens would be achieved by vacuum distillation, i.e. thermal evaporation onto a clean substrate in U.H.V., and would be unlikely to be in single crystal form.

By evaluation from equation (2.1), emission current densities at room temperature would be of the order of $10^{-25} A/cm^2$, so that temperatures significantly above the Curie temperature of gadolinium would be required to produce practicable currents.

Because of the uncontrollability of the errors in the measurement of $\overline{\phi}$, and the temperature and time required for this technique, thermionic emission measurements were judged inappropriate for rare earth metal films.

2.3 Photoelectric Emission

When photons of energy $h\nu$ are incident on the uniform surface of a conductor of work function ϕ at a temperature T, the zero-field photocurrent in the region of $h_o = e\phi$ is given by [12]

$$J = C(kT)^2 f\left[(h\nu - h\nu_o)/kT\right] \qquad \text{equation(2.6)}$$

where f is a universal function of $\left[(h\nu - h\nu_o)/kT\right]$. The value of $h\nu_o$ is that of the photoelectric threshold, which is the energy required to emit an electron from the Fermi surface, i.e. at T = 0. The quantum yield, defined as the number of electrons emitted per absorbed photon[9], I_q, is given by

$$I_q = b'T^2 f\left[(h\nu - h\nu_o)/kT\right] \qquad \text{equation (2.7)}$$

where b' is assumed constant over the range of values considered here, although it is a function of quantum mechanical selection rules for the probability of photoemission and of the optical properties of the material. Experimentally, the material is illuminated monochromatically, and photoemitted electrons are collected by an electrode maintained at a small positive potential to the subject.

For a more accurate estimate of the threshold energy $h\nu_o$, and hence ϕ, than would be obtained from the intercept on a graph of I_q vs. ν, a plot is made of

$$\ln \frac{I_q}{T^2} \text{ vs. } \frac{h\nu}{kT},$$

together with a plot of

$$\ln f\left[(h\nu - h\nu_o)/kT\right] \text{ vs. } \frac{h\nu}{kT},$$

and the two curves made to coincide by a shift parallel to one or both axes. The shift parallel to the $\left(\frac{h\nu}{kT}\right)$ axis is equal to $\frac{h\nu_o}{kT} = \frac{e\phi}{kT}$. In practice, however, because a non-zero field is required to collect the photocurrent, a modified universal function is used,

yielding the apparent photoelectric threshold, $h\nu_o^*$[1]. The effects of surface patchiness may spoil the fit of the two curves in the Fowler plot described above. As in the case of thermionic emission, the sum of the contributions from individual patches bears an extremely complicated, and generally inestimable, relationship to the true physical quantities. The work function value so derived is therefore only applicable for a particular cathode and in particular conditions[1].

A further disadvantage of the photoelectric technique is its inapplicability to semiconductors as a true work function technique. Not only is the free-electron model totally invalid for semiconductors, it has been shown both experimentally and theoretically[1] that photoemission from semiconductors is a combination of surface and volume effects, so the threshold energy is not equal to the true work function[1].

2.4 Field Emission

Electron emission from a metal or semiconductor in the presence of a very high electric field, which reduces the width of the potential barrier at the surface, has been discussed by Fowler and Nordheim[14,15]. The current density

$$J = (1.55 \cdot 10^{-6}\ E^2/\phi)\ \exp(-\ 6.86 \cdot 10^7 \phi^{\frac{3}{2}}\ \theta(x)/E) \qquad \text{equation (2.8)}$$

where

$$x = 3.62 \cdot 10^{-4}\ E^{\frac{1}{2}}/\phi$$

and $\theta(x)$ is the Nordheim elliptic function.

On a graph of $\ln(J/E^2)$ vs. $1/E$, the gradient would be

$$\frac{d\ \ln(J/E^2)}{d(1/E)} = -\ 2.98 \cdot 10^7\ \phi^{\frac{3}{2}} \left[\theta - \frac{x}{2}\left(\frac{d\theta}{dx}\right)\right]$$

$$= -\ 2.98 \cdot 10^7\ \phi^{\frac{3}{2}}\ s(x).$$

Tabulation of $S(x)$[16] has shown it to be slowly varying, so that the gradient $\frac{d \ln(J/E^2)}{d(1/E)}$ depends principally on ϕ.

In practice the applied voltage, V, and collected current, i, are related to E and J by $V = \beta E$ and $i = AJ$. Absolute determination of ϕ depends on an evaluation of β. A technique for numerical evaluation of ϕ and β has been given by Yelinson et al[17], but suffers the disadvantages of requiring measurements at two different temperatures, and that a narrow tip specimen is required in order to achieve a suitable field strength, which necessitates high temperatures to produce measureable currents. This would cause annealing, possibly a rise in vacuum pressure and therefore contamination local to the specimen, and possibly re-evaporation from the specimen. Obviously it would also preclude measurement below the Curie temperature.

A field-emission technique for measuring the average work function of a patchy surface, without temperature restrictions and without reference to any other work function value has been given by Holscher[18]. However, the consideration of the effect of patchiness was only qualitative, and the effect of non-normal electron incidence can be expected to alter the measurement of work function. With careful tube design, manufacture and operation, particularly to minimize non-normal electron incidence at the collecter, Holscher estimated the accuracy of the work function values to be $\pm$ 0.01eV. Later work[19] using this technique for single crystal samples estimated the error to be from $\pm$ 0.02 to $\pm$ 0.05eV.

It was also shown that to allow validity of the free-electron condition on which the technique is theoretically based, certain crystal directions must be avoided. For polycrystalline samples this cannot be guaranteed, so errors would occur due to the departure of electron band structure from the free-electron model. This would be especially true for materials with band structures as complex as the rare earths.

2.5 Surface Ionisation

The similarity between the definitions of work function, ionisation potential and electron affinity provides two techniques for measurement.

2.5.1 Positive Surface Ionisation

The first technique involves measurement of the positive ionisation current, α_+, emitted from a conductor which is in equilibrium with an elemental gas of known first ionisation potential, I. A graph of

$$\ln \alpha_+ \quad \text{vs.} \quad 1/T$$

includes a linear portion whose gradient is equal to

$$e(\phi_p^* - I)/k$$

where ϕ_p^* is the apparent work function for a uniform surface, measured by this technique[1], which suffers from the same sources of error as field emission measurements. The presence of the gas, and the uncontrollability of the temperature at which the relevant linear slope occurs render this technique unsuitable for rare earth metals.

2.5.2 Ratio of Positive to Negative Surface Ionisation.

The second technique requires the metal surface to be heated and to be in equilibrium with a gas which can provide both positive and negative ions and which does not condense onto the metal. The ratio of the positive to negative ionisation currents is determined by the ionisation potential of the atom species, I, the electron affinity of the negative ions, Σ, and the apparent work function of the metal, ϕ_p^*[1]. Limited current knowledge of electron affinity values restricts the gas to the alkali halides. The same comments apply as to unsuitability for measurements on rare earths only more so.

2.6 Contact Potential Difference.

If two conductors A and B of true work functions ϕ_A and ϕ_B at the same temperature are connected electrically, then electron flow will occur until the equilibrium state is reached where the two Fermi levels are the same, resulting in a potential difference between points just outside the two unconnected surfaces of V_{AB}, given by

$$V_A - V_B = -(\phi_A - \phi_B) = V_{AB} \qquad \text{equation (2.9)}$$

called the contact potential difference, (c.p.d.) see Figure 1.2.

For a patchy surface, the electrostatic potential at a point just outside the i^{th} patch, of work function ϕ_i, is

$$V_i = -(\phi_i + \bar{\mu}/e)$$

where $\bar{\mu}$ is the electrochemical potential. For zero applied field, the potential at a distance large compared to the patch size is

$$V^1 = -(\bar{\phi} + \bar{\mu}/e)$$

(see paragraph 3.2.4) where $\bar{\phi} = \sum_i f_i \phi_i$, which is the area-weighted average work function for a patchy surface. If both conductors have a patchy surface then

$$\bar{V}_{AB} = -(\bar{\phi}_A - \bar{\phi}_B) \qquad \text{equation (2.10)}$$

If an external potential V_E is introduced into the electrical circuit between A and B then

$$V_A^1 - V_B^1 = -(\bar{\phi}_A - \bar{\phi}_B) + V_E = \bar{V}_{AB} + V_E \qquad \text{equation (2.11)}$$

Several experimental techniques exploit equation (2.11) including

1 Kelvin (Vibrating Capacitor)
2 Static capacitor
3 Magnetron
4 Saturated diode
5 Breakdown field
6 Electron beam
7 Space-charge-limited diode

and 8 Thin-film sandwich diode methods.

2.6.1 Kelvin Method

The two materials form a parallel-plate capacitor, C, the charge on the capacitor, Q, being given by

$$Q = C(\overline{V}_{AB} + V_E) ,$$

and alteration of the capacitance will cause a charge flow

$$\Delta Q = \Delta C(\overline{V}_{AB} + V_E) \qquad \text{equation (2.12)}$$

If V_E is adjusted so that

$$\overline{V}_{AB} + V_E = 0 ,$$

then no charge will flow on alteration of C, and the c.p.d. is evaluated [20]. The capacitance may be varied either by plate oscillation

or by changing the dielectric between the plates, in either case a repeated modulation produces an a.c. signal for easy detection and measurement at low level. In practice, surface patchiness and plate non-parallelism blur the point where zero charge flow should occur into a region of minimum non-zero current.

2.6.2 Static Capacitor Method

This is similar to the previous method. A well screened, cylindrical capacitor is used, the tendency for charge flow upon electrical connection is detected and a compensating p.d. applied to oppose it[21], in a time shorter than the time constant of the circuit. The claimed sensitivity is 0.25mV.

2.6.3 Magnetron Method

The two materials form a fine cathode in the axis of a cylindrical anode, respectively, and an external voltage V is applied. A uniform magnetic field, parallel to the cathode, is applied, and the field strength H_o required to halve the saturation current from its zero-field value is noted, then

$$V + \overline{V}_{AB} + V_T = H_o^2 R^2 e/8m \qquad \text{equation (2.13)}$$

where e/m is the electronic specific charge, and V_T is a temperature dependent term, independent of V and H_o, derived by Oatley[22] and R is the anode radius. $\overline{V}_{AB}$ is obtained from the intercept of a plot of H_o^2 vs. V, but very accurate alignment of cathode and magnetic field coaxially with the anode is necessary to prevent non-linearity of the graph. Oatley obtained a repeatability of only ± 0.02 volts.

2.6.4 Saturated Diode Method

A coaxial diode is used, as in 2.6.3, but at such a low current that space-charge effects are negligible, so that

$$i = i_s \exp\left[e(V + \overline{V}_{AB})/kT\right] \quad \text{for} \quad (V + \overline{V}_{AB}) < 0$$

equation (2.14)

$$i = i_s = \text{saturation current for} \quad (V + \overline{V}_{AB}) \geqslant 0.$$

A plot of ln i vs. V will therefore have 2 linear sections, intersecting sharply where $V = -\overline{V}_{AB}$. In practice, the sharp intersection is blurred by patchy surfaces, geometric misalignment of the electrodes, and the influence of the applied potential on the electron trajectories, although these difficulties can be mollified. Shelton[23] has described a combination of electric and magnetic fields with a mechanical aperture, arranged to form an electron gun. This ensured that all the collected electrons originated from a known area of the emitting surfaces, and that the relevant emitting and collecting areas were known to be smaller than individual patches on the two surfaces, thus eliminating the effects both of electrode patchiness and of electron-trajectory dependence on the applied potential. It also enabled plane-parallel alignment between the two surfaces, reducing as far as possible the geometric errors of the earlier coaxial diodes.

2.6.5 Breakdown Field Method

The intrinsic electric field E_I across an insulator of thickness d between conductors of average work function $\overline{\phi}_A$ and $\overline{\phi}_B$ is

$$E_I = (\overline{\phi}_A - \overline{\phi}_B)/ed \ .$$

If an external voltage V is applied until breakdown occurs through the insulator, then for one polarity

$$V_1 = (E_c + E_I)d$$

and the other

$$V_2 = (E_c - E_I)d$$

where E_c is the critical breakdown field strength. If two identical structures are exposed to dielectric breakdown of opposite polarisation, then

$$V_1 - V_2 = 2E_I d = 2(\bar{\phi}_B - \bar{\phi}_A)/e = 2\bar{V}_{AB}/e$$

equation (2.15)

In practice a sandwich incorporating a very thin insulating film is used, but spurious results will arise if the film is not truly insulating, due to premature breadkown via impurity levels near the band edges. Patchiness will affect results, since breakdown will occur between opposite patches of lowest local c.p.d. The physical contact between the insulating layer and the counterelectrodes renders this method unsuitable for rare earth metals.

2.6.6 Electron Beam Method

A low velocity, collimated, narrow beam of electrons impinges on an area which is small compared to the specimen, but large atomically[24]. A curve is plotted of the target current vs. the voltage applied between target and emitter, characteristic of the specimen. A similar curve plotted for another specimen of known work function, will, ideally, be different from the first only in displacement along the voltage axis, by an amount equal to the c.p.d. between the two. In addition to the usual perturbing effect of patchy surfaces, poor specimen alignment and less than total electron collection by the specimen will contribute to the non-parallelism of the displaced curves, and hence uncertainty in assigning a c.p.d. for their coincidence.

2.6.7 Space-charge-limited Diode Method

In this method, the anode current-voltage characteristics in the region where the current is space-charge limited are plotted for different anodes, which can be in the form of a cylinder, wire, or ribbon, and the displacement between the curves, along the voltage

axis, again gives the c.p.d. directly. The same disadvantages apply as to the electron-beam method, aggravated in particular by changes in cathode temperature causing geometrical changes, and the possibility of some of the electrons emitted from the cathode not being collected at the anode, but by some other element of the equipment. The proximity of the hot cathode to the specimen further militates against the use of this method for rare earth metals.

2.6.8 Thin-film Sandwich Diode Methods

A thin-film sandwich diode, consisting of a base electrode, a thin insulating layer, often formed from the base electrode, and a counterelectrode of the material under investigation, can be used to determine the c.p.d. between the two metals. For temperature, voltage and insulating layer thickness, s, conditions where thermionic conduction through the insulator dominates the tunnel conduction through it, Simmons' equation[25], may be used for the forward current, i.e. when the electrode of lowest work function is negatively biased,

$$J_1 = AT^2 \exp\left(\frac{-\psi_1}{kT}\right) \exp\left[\frac{14.4(eV + \Delta\phi)}{\varepsilon_r s (kT)^2}\right]^{\frac{1}{2}} \quad \text{for} \quad V > \frac{\Delta\phi}{e}$$

equation (2.16)

where ε_r is the dielectric constant of the insulating layer

ψ_1 is the barrier height between the base electrode and the insulator

$\Delta\phi$ is the work function difference (c.p.d.) between the electrodes

and A is the Richardson constant.

To avoid the need for absolute determination of ε_r, s and T, comparison may be made between diodes which are identical but for counterelectrodes of materials of different work function, one of which is used as a reference. Plots of J_1 vs. V produce exponential

curves whose separations parallel to the voltage axis are equal to the c.p.d. between the respective counterelectrodes. Distinction between forward and reverse bias may be made because in reverse bias, the curves of J_2 vs. V are separated to a far greater extent than the forward bias curves[25]. This apparent increased sensitivity is not usable, however, since the relation between the reverse current density, J_2, and the c.p.d. is not simple.

For conditions where tunnel conduction predominates, the analysis of Simmons[26] enables an estimate of the work function difference between the two electrodes to be made, by curve fitting to the theoretical dependence of tunnel-resistivity on either insulator thickness or applied voltage, e.g. at low voltage, the tunnel-resistivity,

$$\sigma = 1/(J_L \overline{\psi}_L^{\frac{1}{2}} \exp(- B\overline{\psi}_L^{\frac{1}{2}})) \qquad \text{equation (2.17)}$$

where $$J_L = \left[(2m)^{\frac{1}{2}} / \Delta s\right](e/h)^2$$

$$B = 4\pi \Delta s/h \, (2m)^{\frac{1}{2}}$$

Δs = barrier thickness at the Fermi level

and $\overline{\psi}_L$ = mean barrier height above the Fermi level of the negatively biased electrode, at low voltage. The latter parameter is dependent on the c.p.d. between the two electrodes. For both thermal and tunnel conduction it has been assumed that the use of common base electrode and insulator renders all variations attributable to the counterelectrode work function. This may not be so if there are variations in the insulator permittivity or effective thickness, which may be affected by the deposition temperature or the atomic radius of the counterelectrode material[27]. No account has been taken of the effect of patchiness, and it is apparent from the non-linearity of the equations that the area-weighted average work function, $\overline{\phi}$, would not be measured. The use of a sandwich diode was considered unsuitable for the highly chemically reactive rare earth metals, due to the need for physical contact between the electrodes and the insulating

layer. The use of a vacuum for the potential barrier is precluded by the practical difficulty in obtaining a uniform, constant yet very small separation.

2.7 Summary and Choice of Method.

The Surface Ionisation, Breakdown Field and Thin-film Sandwich Diode methods were all rejected because of their need for contact with foreign material. Thermionic Emission was rejected due to theoretical uncertainty in the effect of surface patches, the need for high temperatures, significantly above the magnetic transitions of all rare earths, and the lengthy procedure. The photoelectric work function is not the true work function, but an "apparent work function", due to the non-zero field required for electron collection; the effect of patches produces a differently weighted average from that of, for example, contact potential difference methods, and causes experimental uncertainty in the evaluation of even the apparent work function. It is unsuitable for semiconductor materials, being based on a free-electron approximation. This group of methods was therefore rejected. Most Field Emission techniques require knowledge of the relationship between a measured potential and the field strength at the surface, incidentally non-zero. For the reasons given in paragraph 2.4, this is not simple, and unsuitable for rare earth samples. The Field Emission Retarding Potential method of Holscher, although very attractive in being an absolute measurement, and rapid in use once set up, was considered unsuitable for the polycrystalline rare earth films likely to be encountered, (see paragraph 2.2). Only a qualitative description of the errors caused by patch-fields was given by Holscher; and uncertainty is introduced in interpretation of the voltage at which the field-emitted current sets in, due to emission of electrons from above the Fermi level. The technique also requires a complex electron-optical arrangement for optimum accuracy, the design and construction of which was outside the experience of the author.

Of the remaining contact potential difference methods, the Magnetron and Saturated Diode techniques require magnetic fields, which would affect the behaviour of rare earth specimens. In addition there are theoretical and experimental difficulties with patchy surfaces. The Magnetron requires a hot cathode, which is undesirable in the proximity of the specimen, and an uncertainty in its temperature produces error in the measurement of the work functions, via the term V_T. The Saturated Diode technique requires an electric field, which, in addition to affecting the measured work - function, may influence the electron trajectories in the diode. The modifications used by Shelton[23], although alleviating most problems, only effectively measured single crystals, and produced a c.p.d. of 0.015eV between two supposedly identical surfaces.

The Electron Beam method takes no account of the distribution of patches over the small areas of the two specimens which are compared, which could cause variations in measurements over a period of time if the movements of the samples were not exactly repeatable, either for a single specimen or between different specimens of a single material. Imperfect electron collection also produces errors. Similar reasons for the rejection of the Space-charge-limited Diode method have been enumerated in paragraph 2.6.7.

The remaining methods, the Vibrating Capacitor and the Static Capacitor are very similar. Because they are c.p.d. techniques, they measure the true work function for semiconductors as well as metals, which was desirable for the measurements described in Chapter VII. The choice of specimen temperature is not affected by the experimental requirements, so may be made at one or more convenient, and non-harmful, points. The electric field at the sample (and reference) surface is due only to the patch fields, so the measured contact potential difference remains the area-weighted average, and is not affected by the arbitrary value of electric field strength obtained, for example, in the Thermionic method. Nor are estimates of the electron reflection coefficients at the surfaces required.

It is of interest to note that if sub-monolayer contamination of the surface of a patchy sample were to occur, then a technique which measures the area-weighted average work function would be expected to measure the effect as soon as it became significant compared to the basic sensitivity of the method, no matter where on the sample the contamination occurred. A technique which is biased towards areas of a particular type, e.g. low work function, would be less sensitive to the effects of contamination unless it occurred at sites to which the bias was already directed. Even then the effect may be masked by causing the bias to be directed to uncontaminated areas. Therefore if it is desired sensitively to study the effects on work function of contamination on a patchy surface, a method which measures the area-weighted average would generally be the most suitable. The speed of electrical adjustment required by the static capacitor method necessitates particularly effective electromagnetic and electrostatic screening, and a sophisticated servo-amplifier adjustment system. The screening, as illustrated by Delchar et al[21], would prevent the possibility of in situ optical constant measurements. Although the speed of the latter method is very attractive, that of the former was considered adequate. In view of the possibility of optical measurements, the ease of heating and cooling the specimen in the available ultra-high vacuum equipment, the relative simplicity of potential-probe resistance measurements in parallel plate rather than cylindrical form, and the complexity of the electronic control and measurement circuitry required for the static capacitor, the Kelvin vibrating capacitor method was chosen. Further description of the technique, and the design and operation of the experimental equipment are given in Chapter III.

Chapter III. The Vibrating Capacitor Technique

3.1 Introduction

The basic circuit used for the experiments is shown in Figure 3.1:

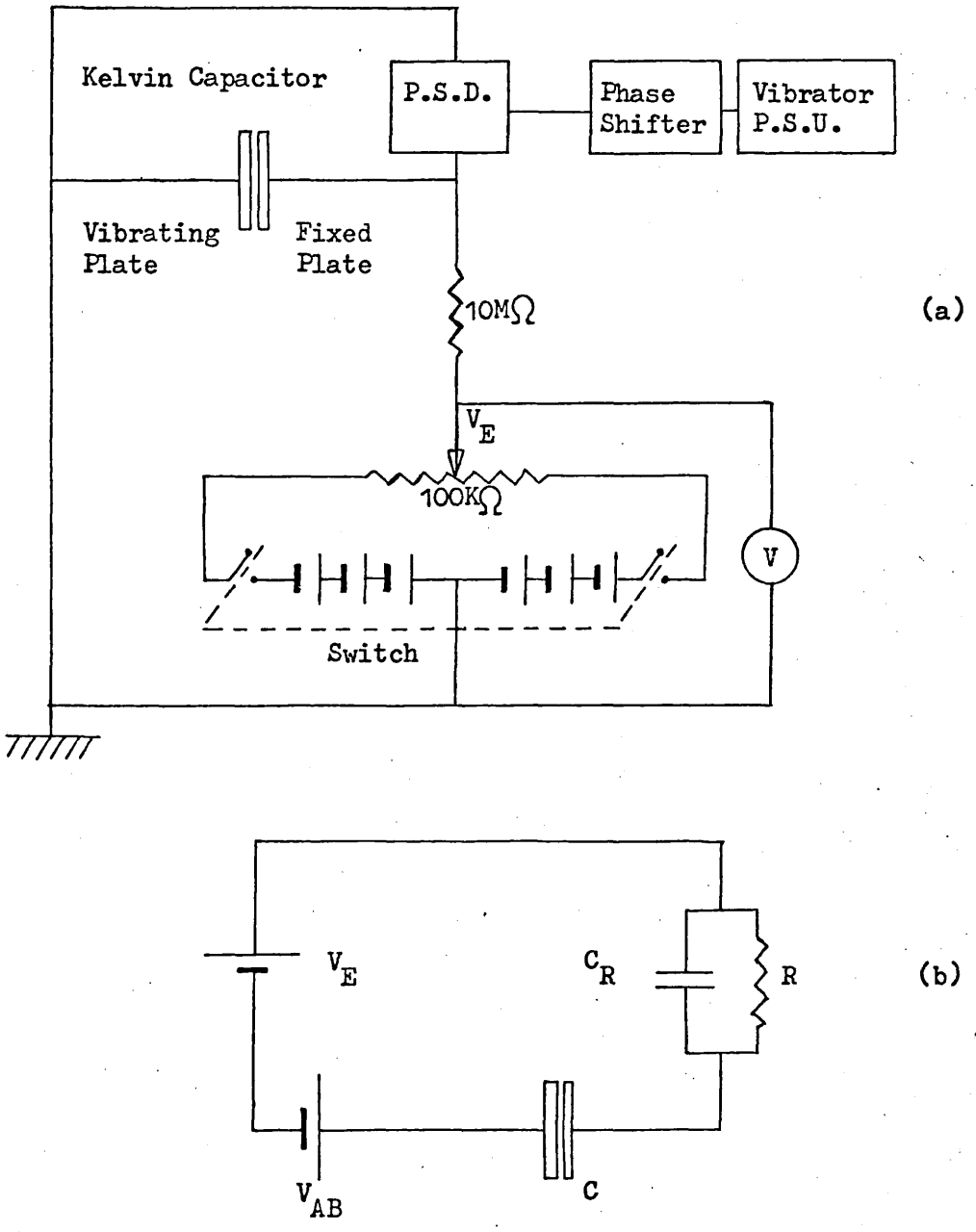


Figure 3.1 (a) Experimental Circuit (b) Equivalent Circuit.

Variation of the capacitance was achieved by driving one plate by means of an electromagnetic coil, so that the spacing, d, between the parallel plates varied at the frequency of the drive unit. The alternative method, of varying the capacitance by changing the dielectric constant, was rejected because of the possibility of errors due to an accumulation of charge on the dielectric[28]. The driving waveform was also fed to the reference input of the phase-sensitive detector (p.s.d.) which was used to detect the output signal from the capacitor. The external bucking voltage, V_E, supplied from the potentiometer across a grid bias battery, was adjusted for null-detection at the p.s.d. and measured with a high impedance voltmeter.

Although based on a very simple experimental method, not involving complicated calculations, analysis, or graph techniques, the technique requires in practice considerable care to avoid errors and inaccuracies in measurement. The factors affecting the precision of measurement and the steps taken to minimise them are described below. The choice of tungsten as a reference surface is explained with the details of its preparation. An electromechanical vibrator was designed as an accessory to be fitted externally to the vacuum chamber in which the specimen preparation and measurements were made. The assembly did not undergo the 250°C bake of the vacuum pumpdown cycle, and it was a simple matter to adjust the parallelism and spacing of the two plates which made the vibrating capacitor and to align the vibrator for smooth simple-harmonic operation.

3.2 Factors Affecting the Precision of Measurement.

The causes of measurement error may be broadly grouped as:

1. Noise, including electromagnetic pick-up, microphony, detector input noise and the generation of harmonics of the fundamental oscillation frequency.
2. Stray capacitance.
3. Thermoelectric effects.
4. Surface patches of different work function.

5. Surface roughness and non-parallelism of the capacitor plates.
6. The reference surface.
7. Gas in the capacitor space.

3.2.1 Noise

As described above the Kelvin technique avoids inherent electric and magnetic fields, but stray fields must be reduced as far as possible. Electromagnetic screening was provided by the stainless steel chamber in which the vibrating capacitor was situated and by the aluminium box in which the bias circuitry was assembled. Earth loops were avoided in the detection and bias circuitry. The phase-sensitive detector used for null-detection of the Kelvin Signal reduced any remaining electromagnetic pick-up.

Microphony was reduced as far as possible by using bare wires, inside the screened vacuum chamber, for electrical connection to the vibrating electrode, so as to eliminate tribological generation; and stainless steel flexible bellows between the vibrator and the rigid parts of the vacuum chamber to which it was mounted, to minimise the vibration transmitted to the stationary electrode. The inaccuracy due to pick-up and microphony, together with effects of stray capacitance, can be estimated by varying the separation of the electrode plates. In particular, for large plate separations, the fractional change in Kelvin capacitance becomes very small, whereas the stray capacitances, pick-up and microphony are unchanged, so virtually all signal, and noise, will be due to error sources. In these experiments, the remaining signal at large plate spacings (over $\frac{1}{2}$") was totally buried in the non-synchronous noise, even at high d.c. bias ($\gtrsim 4$V). This also indicated that there was negligible e.m.f. induced by the movement of the plate in any remaining stray magnetic field.

Amplifier noise was minimised by the high input impedance of the phase-sensitive detector, which was $10^7\Omega$ or $10^8\Omega$ depending on which of several available was being used at the time.

The generation of harmonics may be caused either by non-sinusoidal oscillation of the capacitor spacing, by the reciprocal relationship between capacitance and plate-spacing, or because the capacitor is time-varying. The equation describing the complete circuit can be expressed in integral form, and Fourier analysed. Macdonald and Edmondson[29] have given the analysis for the case where the input capacitance of the detection amplifier is negligibly smaller than the Kelvin capacitor at rest. In this work that is not the case, since the Kelvin capacitance at rest,

$$C_o = \frac{a_c \varepsilon_o}{d_o} \simeq \frac{7. \ 10^{-4} \times 8.854 \ \ 10^{-12}}{10^{-3}} \simeq 6. \ 10^{-12} \text{ Farads}$$

equation (3.1)

and the input capacitance of the amplifier(s) was of the same order. The actual harmonic generation could be observed in practice by C.R.T. comparison of the driving waveform and the capacitor waveform with the capacitor d.c. biased so that the other sources of noise were relatively negligibly small. The difference between the two waveforms was not discernible.

3.2.2 Stray Capacitance

In addition to the capacitance between the sample and the reference surface, there were capacitances between each surface and the surrounding equipment, in this case the vacuum chamber walls, the supports for each plate and the electrical connecting wires. Therefore their contributions to the detected signal had to be eliminated. This was achieved by electrically connecting the vibrating electrode and the assembly around the stationary electrode to the metal vacuum chamber, and the whole to the earth of the detector circuit. The stationary electrode was therefore connected to the "signal" connection of the detector. Therefore no current was generated by the relative motion between the vibrating electrode and either the vacuum chamber or the support for the stationary electrode. The capacitance between the vib-

rating electrode and the gold wires connected to the stationary electrode must have been vanishingly small, and the support for the vibrating electrode, though stout to avoid non-sinusoidal and sideways motion, was screened from the stationary electrode by the vibrating electrode itself. The relatively inert tungsten reference surface was chosen to be the vibrating electrode, so that any remaining stray capacitance would be as constant as possible during the measured life of each sample, and from sample to sample of similar and different materials.

As mentioned in 3.2.1, the signal obtained with the Kelvin capacitance reduced by over an order of magnitude by increased plate separation, was buried in non-synchronous noise, indicating the negligible contribution from stray capacitance. Further evidence of the elimination of stray capacitance was afforded by the independence of the measured c.p.d. of both the plate separation at rest, d_o [30], and the vibration amplitude[31] because changing of these would change the proportional effect of any stray capacitance and hence the apparent c.p.d.

3.2.3 Thermoelectric Effects

The three major thermoelectric effects are named after Seebeck, Peltier and Thomson[13]. The Seebeck effect is the generation of an e.m.f. by having two metallic junctions (thermocouples) at different temperatures, the Peltier effect is the absorption or liberation of heat when current flows through the junction of two dissimilar metals, and the Thomson effect is absorption or evolution of heat when a current flows in a (homogeneous) conductor along which a temperature gradient exists. For common metals the Seebeck effect causes only a few microvolts e.m.f. per ^{o}K of temperature difference, and in the present work the overall temperature differences were generally as close to zero as possible, except for the experiments with heated and cooled specimens, in which cases the temperature differences were no more than $100^{o}K$. The consequent errors were therefore likely to be less than a millivolt in the worst case and only microvolts in general. Near

to, or at, the null point of the c.p.d. measurement virtually no current flowed in the circuit, so the Peltier and Thomson effects, being current dependent, were not expected to affect the temperature of any parts of the circuit, and therefore not to generate e.m.f., by the Seebeck effect.

3.2.4 Surface Patches of Different Work Function

Consider an electrode whose surface lies in the xy plane. The electrostatic potential at a point (x,y,z) outside the surface, generally expressed as

$$V_{es} = X(x) \cdot Y(y) \cdot Z(z) + K \qquad \text{equation (3.2)}$$

where K is a constant, will satisfy Laplace's equation, when no charge density exists in the space, so that

$$\nabla^2 V_{es} = 0 \qquad \text{equation (3.3)}$$

Expressed in Cartesian co-ordinates, and divided by $X(x) \cdot Y(y) \cdot Z(z)$ this becomes

$$\nabla^2 V_{es} = \frac{1}{X(x)} \frac{d^2X}{dx^2} + \frac{1}{Y(y)} \frac{d^2Y}{dy^2} + \frac{1}{Z(z)} \frac{d^2Z}{dz^2} = 0 \qquad \text{equation (3.4)}$$

Consider a surface made up of patches of different work function arranged periodically in x and y[2], the solution

$$V_{es} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \left(a_n \sin\frac{2\pi nx}{X} + b_n \cos\frac{2\pi nx}{X} \right) \left(c_m \sin\frac{2\pi my}{Y} + d_m \cos\frac{2\pi my}{Y} \right) \exp(-gz) + K \qquad \text{equation (3.5)}$$

will satisfy equation (3.4) where a_n, b_n, c_m, d_m, g, X, Y, are constants. X and Y are the periods in the x and y directions respectively and are small compared to the dimensions of the electrode.

End effects, in the xy plane, have been neglected. Substituting equation (3.5) into equation (3.4),

$$\nabla^2 V_{es} = \frac{-4\pi^2 n^2}{X^2} - \frac{4\pi^2 m^2}{Y^2} + g^2 = 0$$

Hence $$g = 2\pi\left[\frac{n^2}{X^2} + \frac{m^2}{Y^2}\right]^{\frac{1}{2}}$$ equation (3.6)

In the absence of applied fields and at distances large compared to X and Y, the electrostatic potential reaches a constant value, V^1. At a position just outside the i^{th} patch

$$V_i = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left(a_n \sin\frac{2\pi n x_i}{X} + b_n \cos\frac{2\pi n x_i}{X}\right)\left(c_m \sin\frac{2\pi m y_i}{Y} + d_m \cos\frac{2\pi m y_i}{Y}\right) + V^1$$

equation (3.7)

This holds for all (x_i, y_i) of the i^{th} patch, and similarly for all patches, so the sum over all points on the surface gives

$$\sum_i s_i V_i = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\left(a_n \int_{-l_1/2}^{l_1/2} \sin\frac{2\pi n x}{X}dx + b_n \int_{-l_1/2}^{l_1/2} \cos\frac{2\pi n x}{X}dx\right)\right.$$

$$\left.\left(c_m \int_{-l_2/2}^{l_2/2} \sin\frac{2\pi m y}{Y}dy + d_m \int_{-l_2/2}^{l_2/2} \cos\frac{2\pi m y}{Y}dy\right)\right] + aV^1$$

equation (3.8)

where s_i is the area of the i^{th} type of patch

a is the total area of the electrode $= \int_{-l_2/2}^{l_2/2}\int_{-l_1/2}^{l_1/2} dxdy$

The first integrals in each bracket equal zero, because the cosine is an even function so that

$$\sum_i s_i V_i = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{b_n X d_m Y}{4\pi^2 nm}\left[\sin\frac{2\pi n x}{X}\right]_{-l_1/2}^{l_1/2}\left[\sin\frac{2\pi m y}{Y}\right]_{-l_2/2}^{l_2/2} + aV^1$$

equation (3.9)

Now the two trigonometric terms must each be $\leqslant 2$, X and Y have been assumed much less than l_1 and l_2, and for a reasonable physical sample the total variation of V_{es} from patch to patch will be a fraction of the absolute value of V^1, so that

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} b_n d_m < V^1$$

$$\therefore \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{b_n d_m XY}{4\pi^2 nm} \ll aV^1$$

So $\sum_i s_i V_i \doteq aV^1$

$$\therefore \sum_i f_i V_i \doteq V^1 \qquad \text{equation (3.10)}$$

where $f_i = \frac{s_i}{a}$, the fraction of the surface area occupied by patches of the i^{th} type. The effective work function ϕ_{eff} and electrostatic potential are related by the electrochemical potential, $\bar{\mu}$

$$\left.\begin{aligned} \text{i.e. } V^1 &= -\frac{\bar{\mu}}{e} - \phi_{eff} \qquad \text{for the overall surface} \\ \text{and } V_i &= -\frac{\bar{\mu}}{e} - \phi_i \qquad \text{for the } i^{th} \text{ patch} \end{aligned}\right\} \text{equation (3.11)}$$

$$\phi_{eff} = -\frac{\bar{\mu}}{e} - \sum_i f_i V_i = -\frac{\bar{\mu}}{e} - \sum_i f_i\left(-\frac{\bar{\mu}}{e} - \phi_i\right)$$

$$\phi_{eff} = \sum_i f_i \phi_i = \bar{\phi} \qquad \text{equation (3.12)}$$

since $\bar{\mu}$ is a bulk property of the sample and hence independent of $\sum_i$, and $\bar{\phi}$ is the area-weighted average of the work functions over the patchy surface. In practice, a perfectly periodic patch distribution of electrostatic potential will not obtain, and V_{es} will decrease more slowly with Z than indicated by equation (3.5).

If an external field is applied by a parallel uniform electrode at a distance which is large compared to the patch size but small compared to the electrode size, then an extra term, linear in Z, should be added to equation (3.5). If the second electrode is also patchy, and with patch sizes small compared to the electrode separation, then similar terms in x and y should be added to the equation. The general result remains the same, however, [2] so that

$$V_A^1 = -\frac{\bar{\mu}_A}{e} - \bar{\phi}_A$$

and

$$V_B^1 = -\frac{\bar{\mu}_B}{e} - \bar{\phi}_B$$

If the two electrodes, A and B, are at the same temperature and electrically connected, so that $\bar{\mu}_A = \bar{\mu}_B$, then the contact potential difference,

$$\bar{V}_{AB} = V_A^1 - V_B^1 = -\left(\bar{\phi}_A - \bar{\phi}_B\right) \qquad \text{equation (3.13)}$$

If an external potential, V_E, is applied between the electrodes, reducing the potential between them to zero, then

$$V_A^1 - V_B^1 = \bar{V}_{AB} + V_E = -\left(\bar{\phi}_A - \bar{\phi}_B\right) + V_E = 0 \qquad \text{equation (3.14)}$$

i.e. $$\bar{\phi}_B - \bar{\phi}_A = V_E \qquad \text{equation (3.15)}$$

If $\overline{\phi}_A < \overline{\phi}_B$ then the positive terminal of the external source must be connected to electrode B if a null is to be achieved. Equation (3.15) shows that for perfectly parallel patchy electrodes, separated by a distance large compared to the patch sizes, the signal developed across the detector resistor is nulled when the bucking potential is equal to the difference between the area-weighted work function averages of the two electrodes. The case where the electrodes are not perfectly parallel is considered in the next paragraph.

3.2.5 Surface Roughness and Non-Parallelism of the Capacitor Plates

The two patchy electrodes may be considered as a collection of small capacitors, $\sum_i c_i$, each formed by the common area of two opposing patches on the two electrodes. Each capacitor is vibrated by an amount δz, so the time varying capacitance,

$$c_i = c_{io} \Big/ \left(1 + \frac{\delta z \cos\omega t}{d_{io}}\right) \qquad \text{equation (3.16)}$$

where c_{io} and d_{io} are the capacitance and separation at rest, respectively. The total charge in the circuit, (see Figure 3.1)

$$Q = \sum_i c_i (V_{ABi} - V_E) + C_R V_R + \int \frac{V_R}{R_I} dt + Q_B \qquad \text{equation (3.17)}$$

where V_{ABi} is the c.p.d. between the opposing patches of c_i, R_I and C_R are the input resistance and capacitance of the detecting circuit, Q_B is the charge in the bias circuit (viz. the external battery, the potentiometer and series resistor). Ideally, the resistance of the bias circuit may be made so large that the current flowing in it may be ignored. The total charge in the circuit must be constant

i.e. $$\frac{dQ}{dt} = 0$$

i.e. $$\frac{dQ}{dt} = \sum_i (V_{ABi} - V_E) \left(\frac{c_{io}\omega\delta z \sin\omega t}{d_{io}}\right)\left(1 + \frac{\delta z \cos\omega t}{d_{io}}\right)^{-2} + C_R \frac{dV_R}{dy} + \frac{V_R}{R_I} = 0$$

equation (3.18)

Now for minimum current in the detector, R_I, $\frac{dV_R}{dt} = 0$

$$\therefore \frac{V_R}{R_I} = \sum_i (V_E - V_{ABi}) \left(\frac{c_{io}\omega\delta z \sin\omega t}{d_{io}}\right)\left(1 + \frac{\delta z \cos\omega t}{d_{io}}\right)^{-2} \qquad \text{equation (3.19)}$$

$$= \sum_i \left(V_E - V_{ABi}\right) \left(\frac{c_{io}\omega\delta z \sin\omega t}{d_{io}}\right) \left[1 - \frac{2\,\delta z \cos\omega t}{d_{io}} + 3\left(\frac{\delta z \cos\omega t}{d_{io}}\right)^2 - 4\left(\frac{\delta z \cos\omega t}{d_{io}}\right)^3\right] \quad \text{equation (3.20)}$$

The binomial series is convergent since $\frac{\delta z}{d_{io}} \cos\omega t < 1$

Hence

(i) If V_{ABi} is the same for all patch pairs, and equal to V_E then $V_R = 0$, i.e. a null is achieved. This will hold for non-patchy electrodes.

(ii) If d_{io} is the same for all pairs, then a null will be achieved if

$$\sum_i c_{io} (V_{ABi} - V_E) = 0$$

$$\text{i.e.} \quad V_E = \frac{\sum_i c_i V_{ABi}}{\sum_i c_i} = \frac{\sum_i s_i V_{ABi}}{a} = \sum_i f_i V_{ABi}$$

which is the same result as in 3.2.4.

(iii) If neither V_{ABi} nor d_{io} are constant, i.e. the electrodes are patchy and non-planar and/or non-parallel, then $V_R = 0$ can only be achieved for special combinations of all the variables, i.e. a minimum value of V_R, given by equation (3.19) will be achieved but will not generally be zero. If only the terms in equation (3.20) which are at the fundamental frequency are considered, then the detected signal amplitude

$$\left|V_R^{\omega}\right| = R_I \sum_i \left(V_E - V_{ABi}\right) \left(\frac{c_{io}\omega\delta z}{d_{io}}\right) \left[1 + \frac{3}{4}\left(\frac{\delta z}{d_{io}}\right)^2\right]$$

Hence

$$V_E = \frac{\frac{\left|V_R^{\omega}\right|}{R_I} + \sum_i \frac{V_{ABi}\, c_{io}\omega\delta z}{d_{io}} \left[1 + \frac{3}{4}\left(\frac{\delta z}{d_{io}}\right)^2\right]}{\sum_i \frac{c_{io}\omega\delta z}{d_{io}} \left[1 + \frac{3}{4}\left(\frac{\delta z}{d_{io}}\right)^2\right]} \quad \text{equation (3.21)}$$

In practice the fractional variation of d_{io} was very small ($\lesssim 10^{-2}$), so the terms in d_{io} may be extracted from the summations,

viz.

$$V_E \simeq \frac{\left|V_R^{\omega}\right|}{R_I \frac{\omega\delta z}{d_{io}}\left[1 + \frac{3}{4}\left(\frac{\delta z}{d_{io}}\right)^2\right]\sum_i c_{io}} + \frac{\sum_i V_{ABi} c_{io}}{\sum_i c_{io}}$$

$$V_E \simeq \frac{\left|V_R^{\omega}\right|}{R_I \frac{\omega\delta z}{d_{io}}\left[1 + \frac{3}{4}\left(\frac{\delta z}{d_{io}}\right)^2\right]C_o} + \overline{V}_{AB} \qquad \text{equation (3.22)}$$

Now the minimised value of

$$\left|V_R^{\omega}\right| \lesssim 10^{-6} \text{ Volts}$$
$$\omega = 22\pi$$
$$d_{io} \simeq 10^{-3} \text{ metres}$$
$$C_o \simeq 6.\ 10^{-12} \text{ Farads}$$
$$\delta z \simeq 0.75.\ 10^{-3} \text{ metres}$$
$$R_I \geqslant 10^7 \Omega$$

Hence the measurement error due to small variations in d_{io} was

$$V_E - \overline{V}_{AB} \lesssim 2.\ 10^{-4} \text{ Volts}$$

3.2.6 The Reference Surface

Since the Kelvin method measures the c.p.d. between two electrodes, the accuracy with which the work function of one of them can be measured depends on the accuracy to which that of the other is known. The ideal reference electrode would be inert to changes of its environment during and between successive measurements, flat, and have uniform work function, either well documented or measured by a technique compatible with c.p.d. measurements. The reference electrode chosen for this work was a rectangle of sintered polycrystalline tungsten, cut from an X-ray target and ground flat, which was expected to be inert under vacuum conditions and repeatable from run to run since a system bake occurred between runs. It did not have uniform work function, however, since it was not single-crystal, but due to its stability the area-weighted average of work function, as measured by this technique, would be constant. The work function of tungsten in many forms is very heavily documented and polycrystalline tungsten has been recommended as a work function standard, at 4.545 eV [32,33]. As a check the c.p.d. between the reference electrode and a vacuum-evaporated film of tungsten on pyrex was measured.[34] Further details are given in paragraph 3.3

3.2.7. Gas in the Capacitor Space

The presence of gas in the capacitor space can lead to errors in three ways. Firstly, the necessarily high input impedance of the detector may allow charge accum ulation on the virtually ungrounded electrode, in this case the sample.[35] The absence of such a build-up may be verified by the equality of c.p.d. measurements before and after short-circuiting the capacitor.

Secondly, the presence of charged particles between the electrodes would produce a signal at the same frequency and phase as the signal due to the c.p.d.[36], due to the collision between them and the electrode. This would occur whatever the electrode spacing, so its effect would be included during the check for stray-capacitance effects (see Paragraph 3.2.2) The gas, and therefore charge, density in the capacitor space may be reduced when the spacing is very small, i.e. the effect will be even smaller under the c.p.d. measuring conditions. For measurements to be made at gas pressures where such an effect would be significant, a method has been described [36] which avoids the error due to charge impact, by driving the capacitor at two frequencies. The signal from the capacitor would be at the higher frequency, amplitude modulated at the lower. The modulation would not be affected by the linear mechanism of charge collection by collisions at the moving surface. This technique would be useful for adsorption studies in relatively low vacuum.

The third form of error is the changing of work function of either electrode, or both, due to reaction with or adsorption of gas[37]. The rare earth and tungsten films were prepared and measured in U.H.V. equipment, to reduce these effects as far as possible. Two films were prepared and measured at O.H.V. in the same equipment for comparison. Further details are given in Chapters IV and VI.

3.3 The Reference Surface

As mentioned in 3.2.6 it has been recommended[32] that a common reference surface be used for all c.p.d. measurements, that this should be well-aged polycrystalline tungsten foil, and that its assumed true work function value at 300°K be 4.545 eV. Comparisons of tungsten foil

with vapour deposited films of tungsten on glass [33, 34] have shown equality of work functions, so that the latter form may equally be used as a reference. For this work a block of sintered tungsten, cut from an X-ray target disc, was used as a reference surface, comparison being made between it and a vapour-deposited film of tungsten. The purity quoted by the manufacturers (Metalwerk Plansee of Austria) was 99.95% minimum. The maximal impurities, in p.p.m., were quoted as oxygen 50, hydrogen 10, nitrogen 50, carbon 80 and iron 50.[38] The use of a solid plate of tungsten, rather than a foil, enabled a good and repeatable plane-parallel gap to be maintained between the reference and specimen surfaces. The tungsten was diamond-sawn into a rectangle 3.8 cm x 1.9 cm and was ground flat on its front surface, using diamond dust. A tongue was machined on its rear surface to enable a stainless steel holder to be spot-welded to the rear without affecting the front. This intimate metal contact was then the only connection between dissimilar metals between the reference surface and the earthed metal vacuum chamber.

The reference surface was cleaned by electropolishing, which consisted of the application of a low alternating voltage ($\sim$6V) between it and a stainless steel counter-electrode, while immersed in a 1N sodium hydroxide solution at room temperature. The voltage was increased till gas evolution just occurred at the electrodes, which were spaced sufficiently far apart to prevent streaming of the bubbles between them. The procedure was repeated till a uniformly bright and clean surface was obtained, the total polishing time being approximately 10 minutes. The block was then thoroughly rinsed in distilled water to remove the sodium hydroxide, then baked to remove the water. Observations in an optical microscope, at 600 X magnification, revealed grain sizes typically 25 microns, with extremes of $< \frac{1}{2}$ and up to 200 microns. Figure 3.2 shows the tungsten block unetched and etched, and a tungsten ribbon all at 1300 X magnification. Back-reflection Laue X-ray diffraction patterns of the unetched, etched and ribbon samples are shown in Figure 3.3. The sintered block samples had no obvious symmetry (other than circular) because they were randomly polycrystalline, whereas the ribbon showed preferred orientation. The etched sample pattern, Figure 3.3b, consisted of a larger number of smaller spots than the

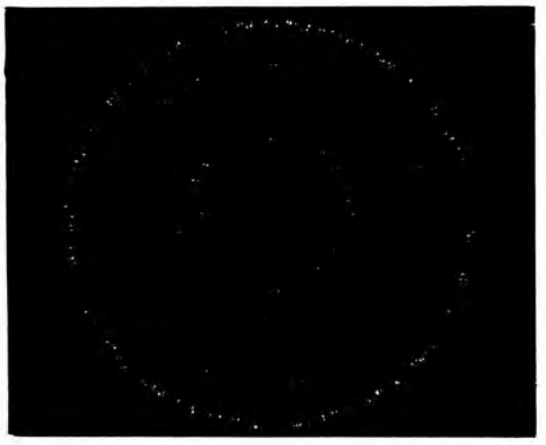
(a)

(b)

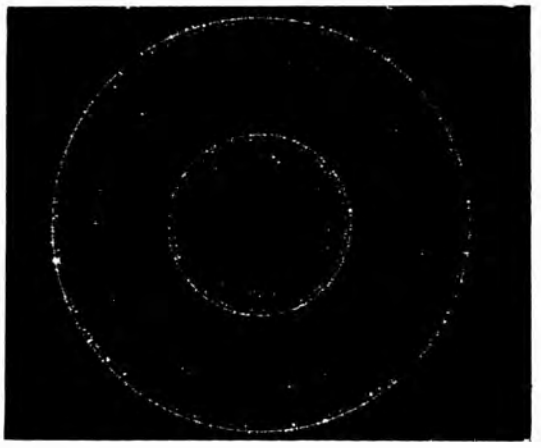
(c)

10 μm

Figure 3.2 Surface Detail of Tungsten at 1300X Magnification
(a) unetched block (b) etched block (c) ribbon



(a)



(b)

(c)

Figure 3.3 X-ray Diffraction Patterns of the Samples Shown in Fig. 3.2

unetched, showing that the exposed grain sizes were smaller after etching, although the optical micrographs suggested that the grain sizes were smaller on the unetched sample. The smaller size of the surface features of the unetched sample must therefore have been irregularities on the surface of the grains, possibly caused by the surface grinding. These would be removed by the etching, leaving fresh grain boundaries exposed on the etched sample. One possible explanation for the larger actual grain size at the surface of the unetched sample is that the previous use of the disc as an electron-target for X-ray generation caused grain growth by annealing, and that the etching removed those enlarged grains that had not already been ground away.

The vapour-deposited tungsten film was evaporated onto flame-polished 'pyrex' borosilicate glass at room temperature and in U.H.V., after Dobson and Hopkins.[34] The purity of the film was checked by X-ray fluorescence, at the Fulmer Research Institute. Further details are given in Chapter IV. Reflection electron diffraction patterns of the film consisted of semi-circles, broadened by the non-constant distance from specimen to plate, showing polycrystallinity of the film. Electronmicroscopy showed small grain sizes, $\lesssim 0.2\mu m$. The c.p.d. between the two tungsten surfaces in U.H.V. was measured at -0.003V ($\pm$0.0015V) which was less than the error in the c.p.d. measurements in the remainder of the work. It has therefore been assumed that the tungsten reference surface had an area-weighted average work function of 4.542eV ($\pm$0.0015eV) and that this value was invariant from run to run because its treatment was invariant and included a bake at 250°C during the vacuum pump-down cycle.

3.4 The Design and Operation of the Vibrator.

The initial experiments to determine the dimensions of the electrodes and oscillation amplitude required to generate a signal of sufficient magnitude for sensitive measurement were conducted in O.H.V. using evaporated films of gold and copper. With a common electrode area of around 13cm^2, separation of less than 1mm. and oscillation amplitude of around 0.5mm., it was possible to observe a signal null with a sensitivity in the bucking voltage, V_E, of around 1mV. It was decided to build a vibrator which fitted outside the metal U.H.V. chamber used

for rare earth measurements, and transmitted its oscillation into the chamber via a vacuum-tight feedthrough, for the following reasons:

(i) Contamination of the U.H.V. would be minimised by the exclusion of all but the most appropriate materials, i.e. stainless steel and tungsten, while allowing free use of the materials suitable for the vibrator.

(ii) Repeated outgassing would be minimised by keeping the extra included surface area to a minimum.

(iii) Since the use of an electro-mechanical drive-mechanism was desirable, for easy and rapid variation of oscillation frequency and amplitude, the metal vacuum chamber would screen the Kelvin capacitor from the electromagnetic radiation, which would be at the same frequency as the signal from the capacitor.

(iv) The adjustments for parallelism and spacing of the capacitor plates, and, if necessary, for smooth running of the vibrator, would be facilitated. This was particularly important in view of the lengthy pump-down requirements for U.H.V.

The substrate carrying the electrode of unknown work function was mounted on a stage carried by a Vacuum Generators goniometer feedthrough, type R.D.1, which could be moved in all three linear directions and rotated about its own near-vertical axis and about two orthogonal horizontal axes, using a Vacuum Generators Universal Motion Drive U.M.D.1. The centres of rotation of all three were approximately at the centre of the electrode surface. An external view of the feedthrough is given in Figure 3.4. An out-of-contact mask, made of stainless steel, served two purposes. Firstly, it allowed the specimen to be electrically isolated from its metallic holder by defining an evaporation area smaller than the substrate. Secondly, it provided an extra screen, against the small amount of electromagnetic radiation which may have entered the chamber through the small glass windows. Figures 3.5(a) and (b) show the specimen holder assembled and exploded, including the film heater (see Chapter V), the electrical connections for potential probe measurements (see Chapter VI) and the small tubular feedthrough used for cooling.

The reference electrode was mounted on the end of a horizontal rod, inside a flange connected to the main chamber by a welded stainless steel bellows, as shown in Figure 3.6. The remainder of the vibrator was external to the U.H.V. and attached as in Figure 3.7. A schematic

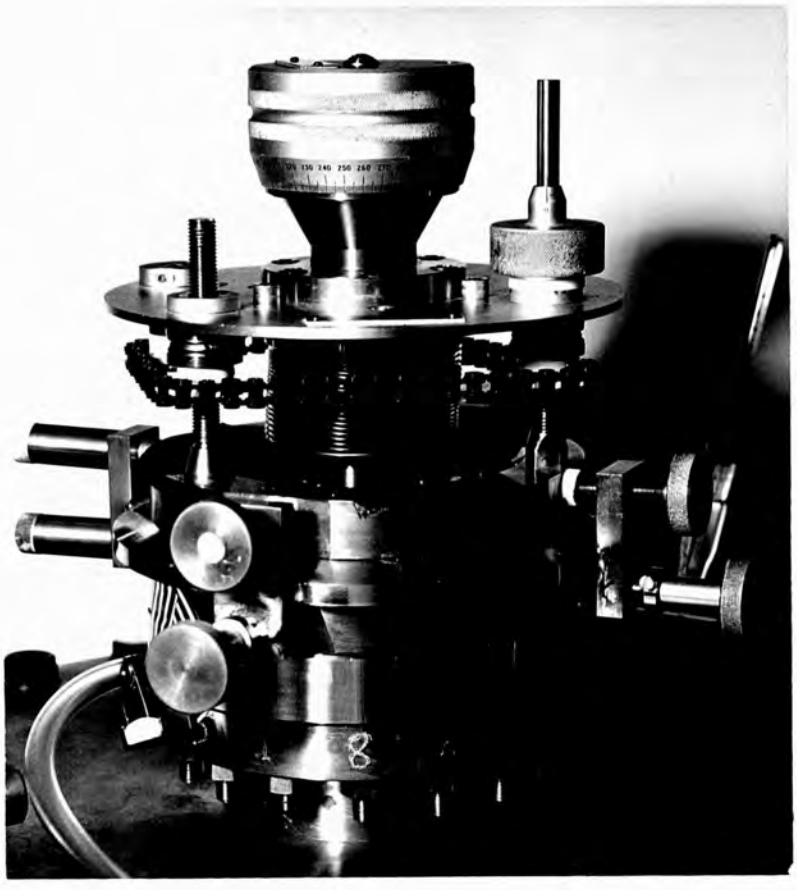


Figure 3.4 External View of the Goniometer Feedthrough

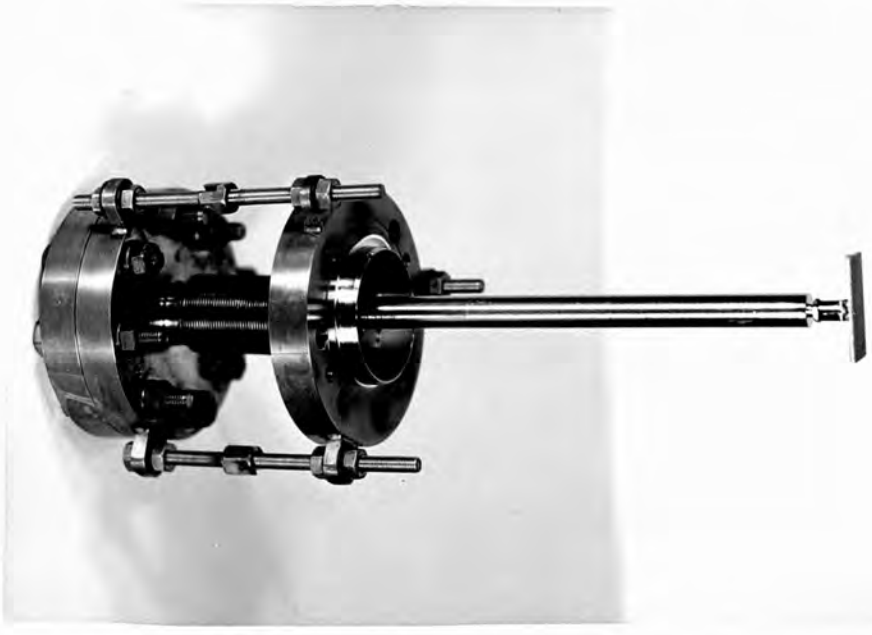
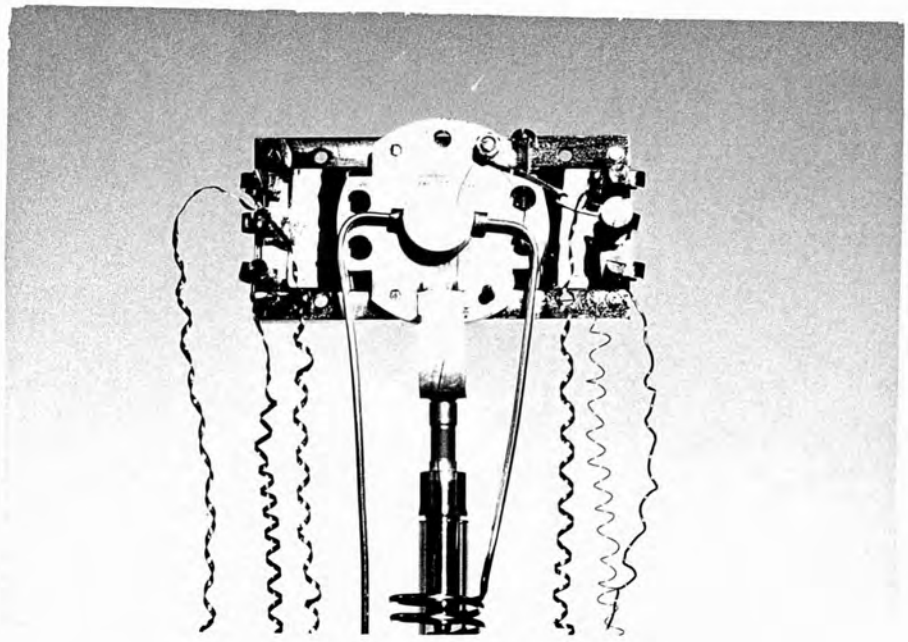
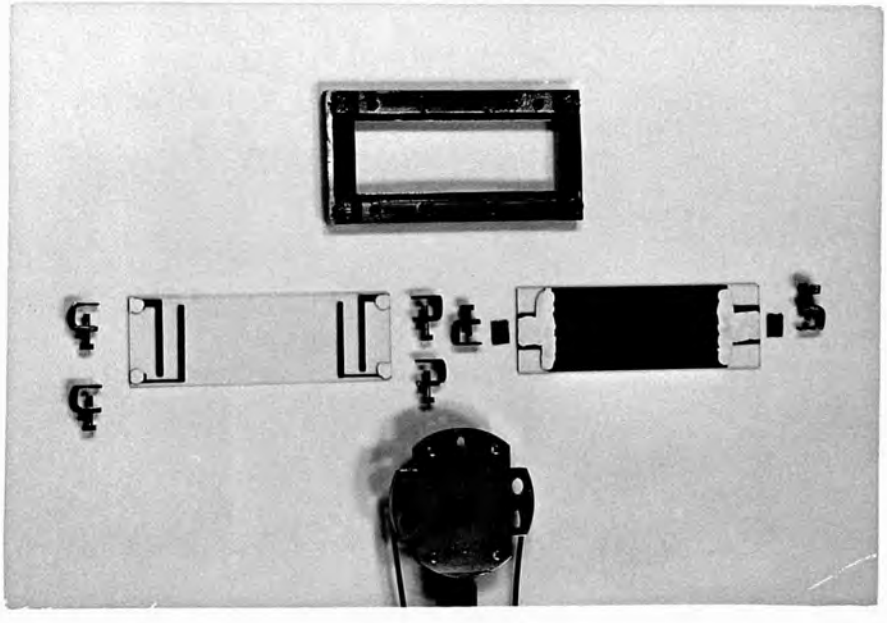


Figure 3.6 Reference Electrode, Support Rod and Bellows



(a)



(b)

Figure 3.5 Substrate Holder (a) assembled, rear view (b) exploded

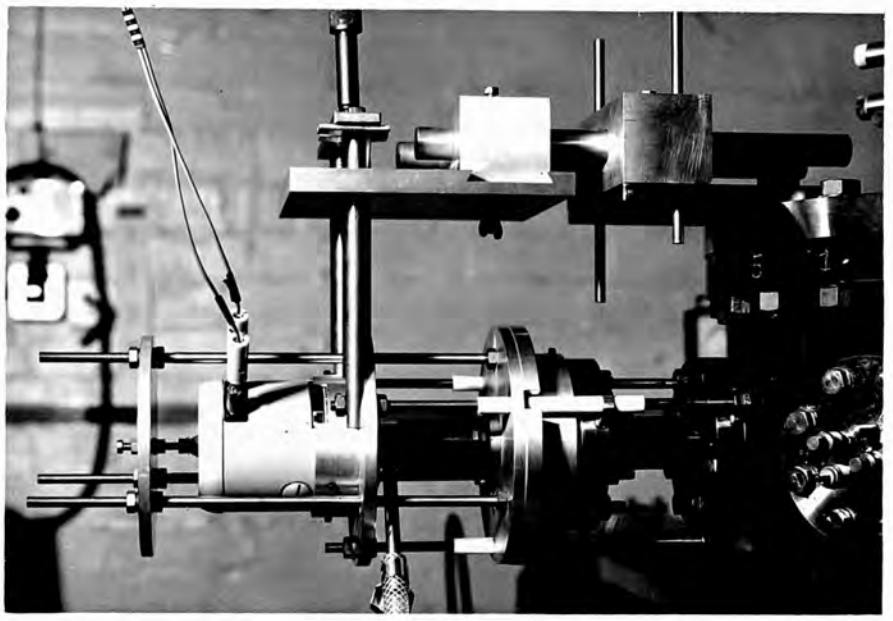


Figure 3.7 External View of the Vibrator Assembly

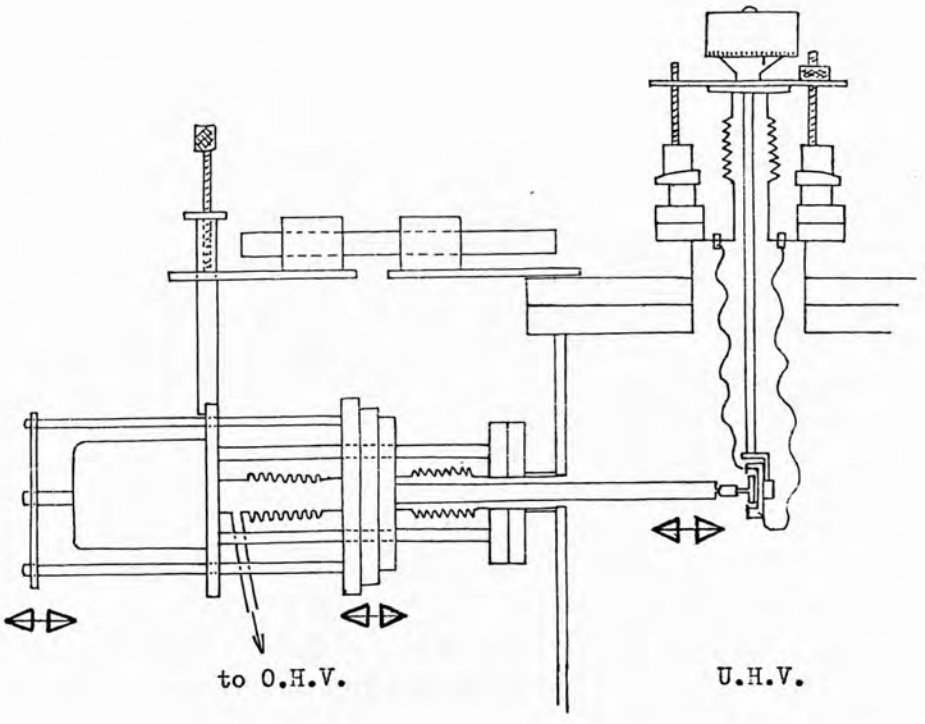


Figure 3.8 Schematic View of the Vibrating Capacitor Apparatus

view of the whole is given in Figure 3.8. The natural tendency for the atmospheric pressure to force the bellows to full compression was overcome by the use of a second, demountable, bellows on the opposite side of the flange, which was evacuated using a rotary pump. The electromagnetic unit, a Ling Dynamic Type 202 driven by a Ling Dynamic T.P.O. 20 oscillator/amplifier, was suspended from an adjustable solid stationary demountable platform. It drove the plate most distant from the chamber, which in turn drove the flange separating the two bellows via three rods and a demountable collar. The weight of the flange and its collar were supported on three long p.t.f.e. bushes. Coarse adjustments to the capacitor spacing and the position of the reference electrode relative to the larger surface area of the specimen electrode were made by moving the main platform, while fine adjustments, especially for parallelism, were effected by moving the specimen.

The electrode separation might also have been adjusted by accurate regulation of the vacuum pressure in the outer bellows, but this was not necessary with the available specimen mount.

The main support platform was assembled after the system bake (see Chapter IV) but before the evaporation of the specimen. During the vacuum pump-down and the evaporation the reference electrode was parked facing slightly downwards and away from the evaporation source. The specimen electrode was then rotated to be in a plane parallel to that in which it would be measured, but slightly further from the vibrator feedthrough, and the vibrator unit and moving parts assembled so that the reference electrode was in the measurement position. The final adjustment to the capacitor spacing and parallelism were then made, using the goniometer feedthrough. This assembly procedure caused a delay between evaporation and first measurement of over an hour, but this was not thought to be excessively long in terms of contamination time (see paragraph 4.8) and the value of the c.p.d. was monitored for up to 30 hours in U.H.V. to observe changes including those due to gas collision with the specimen. It was assumed that the reference surface had stabilised during the long vacuum pump-down and rapid return to base pressure following the slightly higher evaporation pressure. Nor is the delay an inherent feature of the technique.

The delay between deposition and first measurement could have been greatly reduced, to only a few minutes, by redesigning the vibrator and specimen mount so that the vibrator could be fully assembled prior to deposition of the specimen. This would have necessitated a few mechanical changes to the vibrator, and significant changes to the ways in which the many electrical connections from outside the chamber to the specimen and its holder were made, in order that it might be moved and rotated from deposition position to measurement position without fouling any of the electrical connections. Further details are given in Chapter VIII. The changes were too extensive to be made in the time available.

The oscillation frequency of the U.H.V. vibrator was limited, by the natural frequency of the mechanical assembly and the power capability of the drive unit, to less than about 20Hz. At higher frequencies only very small amplitudes were possible without generating harmonics. In general 11Hz was used, since it gives good rejection of 50Hz interference when detected by phase-sensitive techniques. The phase-sensitive detectors available were Brookdeal Models 629, 401, 9402 and 9501. The detector in use was set for best sensitivity by adjusting the phase of the reference signal, derived from the waveform driving the vibrator, so that the output from the p.s.d. was maximised. For convenience, this was done with a large bias voltage, V_E, to provide an artificially large signal. The phase-lag between the Kelvin signal and the drive waveform was confirmed by displaying them on a double-beam oscilloscope. This also allowed comparison of the waveforms, to inspect for harmonics and for purity of oscillation of the vibrator. As predicted by equation (3.19), the signal lagged the drive by $\pi/2$ radians. Low- and high-pass filters were used to reduce interference before the signal entered the p.s.d. proper. The time-constant of the integrator circuits, which converted the phase-detected signal to a d.c. level for easy reading, was chosen to be not less than 100 cycles of the Kelvin signal, and the output displayed on the meter was only judged acceptably accurate if it had remained constant for 10 times that. In practice the reading usually stabilized within a fraction of the time-constant generally used, 100 seconds, as long as the adjustments were small. As the signal null, or minimum, was approached, indicating

equality of c.p.d. and V_E, the adjustments of V_E were small, of the order of 1mV, and the condition was easily satisfied. The balance point was approached from both increasing and decreasing values of V_E, and for each the value of V_E was noted at which the p.s.d. meter reading moved away from its null, or minimum. The c.p.d. was taken as the mean of the two values so obtained. The sensitivity range of the p.s.d. at balance was typically 10μV, with the signal typically less than 0.5μV. A change in V_E of less than 1mV was possible, and a change of a few millivolts or less was necessary to determine motion of the Kelvin signal away from its balance point. The difference between the two values of V_E for each c.p.d. determination was typically less than 10mV, sometimes zero. The sensitivity of the measurements was therefore of the same order.

The separation between the two electrodes of the capacitor was around 1mm, at rest, with the oscillation amplitude being between 0.5 and 0.7mm for good signal strength (i.e. large proportional change in the Kelvin capacitance). During the initial setting-up checks were usually carried out for the effects of stray capacitance, microphony, pick-up and charge collection at the ungrounded electrode, as described in paragraph 3.2. On some films, the c.p.d. was measured at several frequencies, between 9 and 15Hz. The results were equal within the experimental errors.

Chapter IV Specimen Preparation

4.1 Introduction.

All the work function measurements were made on vapour-deposited films, the rare earths having been grown and measured in a U.H.V. chamber while the others were grown and measured in O.H.V.

The vacuum systems are described briefly, and the preparation of the evaporation sources, the substrates, the vapour-deposited electrical contacts and the films themselves are detailed. The results of residual gas analysis in the U.H.V. system are discussed.

The film thicknesses were measured by multiple-beam interferometry. A résumé of the technique is given, together with the results.

4.2 The Evaporation Systems.

4.2.1 The O.H.V. Systems

The O.H.V. systems were used for the deposition of the electrical contacts onto the experimental substrates; the initial tests for the feasibility of the Kelvin technique; the initial outgassing of the evaporation sources for the U.H.V.; the manufacture of the film heaters; the c.p.d. measurements after aging the rare earth films in air; and the deposition of the silver films for the multiple-beam interferometry. Most of this work was conducted in an Edwards type 12EA, although two Bir-Vac 12" systems were also used. All three systems were evacuated by rotary pumps and oil-diffusion pumps, all containing silicone-based oils, and had water-cooled baffles and baffle-valves. The pressure in the Edwards system was generally in the 10^{-5} torr range, while the Bir-Vacs achieved 10^{-6} torr. The evaporations were effected from directly heated sources. The silver for the multiple-beam interferometry was deposited in a vacuum system reserved for this purpose in order to achieve high-purity films for best resolution.

4.2.2 The U.H.V. System

The U.H.V. system was built by the Edwards High Vacuum Company to a design by R.F.Miller and A.J.Taylor[39], primarily for optical studies. It was constructed of stainless steel EN58B, with an internal diameter of 12" and height of 22.5". The chamber was roughed out by an Edwards ED250 Double Stage rotary pump, acting through a foreline trap filled with activated alumina balls, Edwards type 20-H260-A7, in order to prevent oil-vapour ingress into the chamber. The alumina was reactivated every run or two and replaced at the first sign of discolouration. The second pump was a Varian type 912-7008 Diode Ion pump, of pumping speed 140 litres per second, which was connected to the chamber via a bakeable valve whose sealing action was a circular knife-edge biting into a copper gasket. The chamber and ion pump were supported by a massive concrete plinth, to reduce vibration transmission from the ground. Within the chamber was a Titanium Sublimation Pump (T.S.P.), any one of whose four elements could be activated manually or automatically to pump at up to 4000 litres per second. The final pumping action was provided by a cryo-pump consisting of a hollow body which could be cooled by filling with liquid nitrogen to condense vapours and gases in the chamber. All the pumps were situated in or connected to, the lower half of the chamber.

A series of I.S.O. ports in a horizontal ring in the upper half of the chamber were used for the ionisation gauge; the evaporation source; the shutter, to mask the substrate from the source during outgassing; the residual gas analyser head; three optical windows and the vibrator feedthrough.

The ionisation gauge was an Edwards IG4M, a nude Bayard-Alpert type, with an Edwards model 5 power supply. It was possible accurately to measure to 10^{-11} torr using the "modulator"facility. Measurement of the total ion current with the modulator collector firstly at earth then at the same potential as the standard collector allowed the elimination of the "x-ray" current [40].

The evaporation sources were multiwire tungsten helical filaments directly heated, as described in paragraph 4.3, and mounted on an electrical feedthrough as shown in Figure 4.1. A cylinder projected into the chamber, protecting all but the substrate region from the vapour source.

The shutter was a circular sheet holding a glass slide, to enable the filament to be viewed from the opposite side of the chamber while still masking the chamber and the viewing window from the source. It was mounted on a rotary feedthrough for mobility.

The residual-gas analyser was an A.E.I. Minimass. It allowed detection of partial pressure of hydrogen and helium and all ions with masses between 12 and 240 a.m.u. Its main use was for the detection of hydrogen and the elements and molecules between 12 and 60 a.m.u., including the principal components of rotary-pump contamination, if any. Its use and results will be described in paragraph 4.8.

The optical windows were used merely as viewing ports for this work, but care was taken to avoid contamination by vapour-deposition.

The vibrator feedthrough, as described in Chapter III, included a bellows, which was supported by removable rods during the vacuum pump-down cycle, until the vibrator unit was assembled around it. The stainless-steel rod holding the reference surface is shown projecting through the I.S.O. flange in Figure 3.6. All the flanges were I.S.O. types, using gold-wire seals; except for the ion-pump-to-chamber, the residual gas analyser and the rotating feedthrough in the chamber lid, which were sealed by copper compression gaskets.

The flat lid of the chamber was removable for access to the chamber, generally when engineering changes were required. The substrate carrier was attached to a rotary feedthrough in the lid, described in Chapter III and in paragraph 4.6.

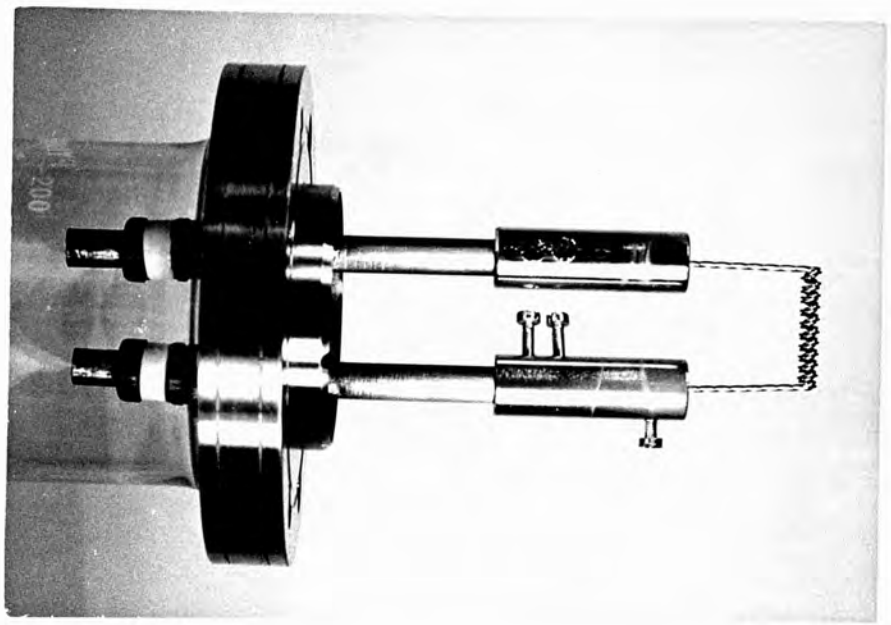


Figure 4.1 Filament Source as Used in U.H.V.

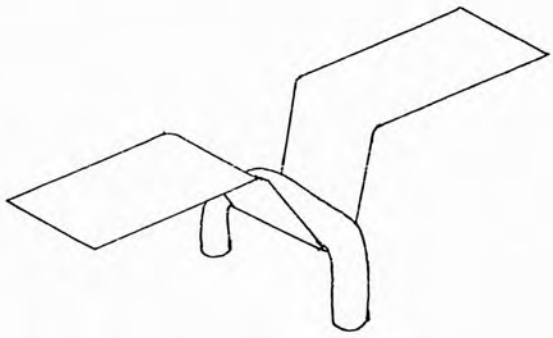


Figure 4.3 Evaporation Source for Aluminium

The materials used for parts assembled into the U.H.V. chamber were selected to minimise outgassing. Stainless steel, oxygen-free, high conductivity (O.F.H.C.) copper and gold wire were used, except as otherwise described. The steel and copper parts were electropolished before assembly, to reduce their surface areas and to remove the outer layers of the metals which had been contaminated by the cooling fluids of the machine tools.

The electrolyte was phosphoric acid saturated with chromium trioxide, heated to approximately 50°C, contained in a stainless-steel beaker which acted as the cathode. The current between the polished part and the beaker was adjusted so that effervescence at the anode just occurred, which was found to give the best results. This point is confirmed by Tegart[41]. In the present work the optimum current density was found to be approximately 1 Amp / cm^2.

Final cleaning of all components before entry into the U.H.V. consisted of an ultrasonically agitated wash in diluted Decon 75 or Decon 90, which are proprietary mildly-alkaline "surfactant" detergents, followed by thorough flushing by distilled water and either evaporation drying or blowing dry with "White Spot" nitrogen gas or hot air.

The U.H.V. chamber was always let up to "White Spot" nitrogen when access was required for renewal of substrate and vapour source etc. This was to reduce the adsorption of condensable vapours and gases of low-pumping speed in the walls and other surfaces in the chamber.

The vacuum pump-down procedure included a bake of the whole system, by assembling an oven of shutters around all but the rotary pump (and electronics). The bake temperature of 250°C was limited by the permanent magnet of the ion pump. The bake was continued till a pressure of $\lesssim 10^{-7}$ torr registered on the ion-pump power supply, which was known by experience to lead to low ultimate pressures. The total bake time was usually around 30 to 35 hours. The ion bombard-

ment of the whole chamber when the ion-pump was initially switched on and lasting for a few minutes, followed by the high temperature of the bake, effectively comprised final cleaning stages within the chamber. The total pump-down, from atmosphere to evaporation pressures, usually took approximately 90 hours. Although the working base pressures were generally in the range 1 to 3. 10^{-10} torr, pressures as low as 5. 10^{-11} torr were sometimes measured.

4.3 Filament Preparation.

The filaments used as evaporation sources were of two types. Both were constructed from three strands of tungsten wire, each 0.025" diameter, twisted together then formed into helical coils, by the R.D.Mathis Company (U.S.A.) and supplied via Nordiko Ltd., (Hants U.K.). The fluting of the filament surface, due to the triple-wire twist, promoted wetting due to surface tension effects, and an increased capacity to hold a molten charge.

Type number F12 3x .025W was used for nichrome and other metals in O.H.V., the helix measured 1" x $\frac{1}{4}$" I.D. and had 5 coils. The deposition charge was wound tightly round the helix for good thermal contact.

Type number F1 3x .025W was used for the U.H.V. work, with the leads modified to be perpendicular to the helix axis, as shown in Figure 4.1. The helix was 1" long, $^{1}/_{16}$" I.D. and had 10 turns. Prior to loading with rare earth charges, the filaments were washed in Decon, dried, then outgassed in O.H.V. at a current higher than that required for use in the U.H.V. evaporation (viz. 60 Amps). For gadolinium or terbium films, a length of rare earth wire, 1mm. diameter, was then placed along the helix axis. Initially, the filament was then loaded into the U.H.V. for outgassing and deposition. In order to increase the evaporant charge the melting was later done in O.H.V. and was repeated 2 or 3 times till the filament was heavily charged and generally totally blocked in 2 or 3 places.

The filaments were further outgassed in U.H.V. prior to film deposition, although at a slightly lower current than that used to deposit the films, in order to conserve the charge.

The outgassing caused an initial rise in pressure, typically from $\sim 3.10^{-10}$ torr to $\sim 2.10^{-8}$ torr, followed by a reduction as the filament heating was continued. It was not certain whether this fall was due to gettering of the gases by rare-earth vapour or by adsorption into the melt, but it was thought to be the former because some evaporation was often apparent by viewing the charge through the shutter-windows in situ.

For the dysprosium deposition, the pilot tests in O.H.V. showed sublimation before melting, so that the charge could not be increased by successive load and melt cycles. The charge in the U.H.V. was therefore only a single length of 1mm. diameter wire, and was not outgassed prior to U.H.V. insertion. Outgassing in the U.H.V. consisted of repeated heating for a few minutes at a time, maintaining system pressure below $\sim 10^{-7}$ torr, until a long period ($\sim 1\frac{1}{2}$ hours) of heating could be maintained at low pressure. Some short higher-current pulses were also applied, but these caused momentary pressure bursts up to 10^{-5} torr.

For the tungsten film deposition in U.H.V., an unused uncharged filament of the same type was used, and outgassed in U.H.V. for $3\frac{1}{2}$ hours over 6 days before evaporation.

4.4 Substrate Preparation.

The substrates were Chance float-glass microscope slides, 3" x 1" x 0.04", whose constituency was given by the manufacturers as:

SiO_2	71.5%
Na_2O	12.5%
CaO	8.1%
MgO	3.5%
Al_2O_3	1.3%
TiO_2	1.8%
Sb_2O_3	1.0%

Immediately before loading into the U.H.V. chamber, the substrates were given an ultrasonically agitated wash in a hot solution of Decon 75 or Decon 90, diluted in distilled water, followed, as the manufacturers recommended, by a thorough flushing with distilled water. They were left to dry by evaporation, under lens tissues (out of contact), and before and after loading into the holder were blown free of atmospheric dust with "White Spot" nitrogen. As mentioned earlier, the ion-bombardment and bake-out which occured during the vacuum pump-cycle were expected further to ensure cleanliness. In addition, for those runs where a substrate heater was included, the substrate temperature was raised to over 100°C, and allowed to cool again to ambient, shortly before the film deposition.

4.5 Contact Preparation.

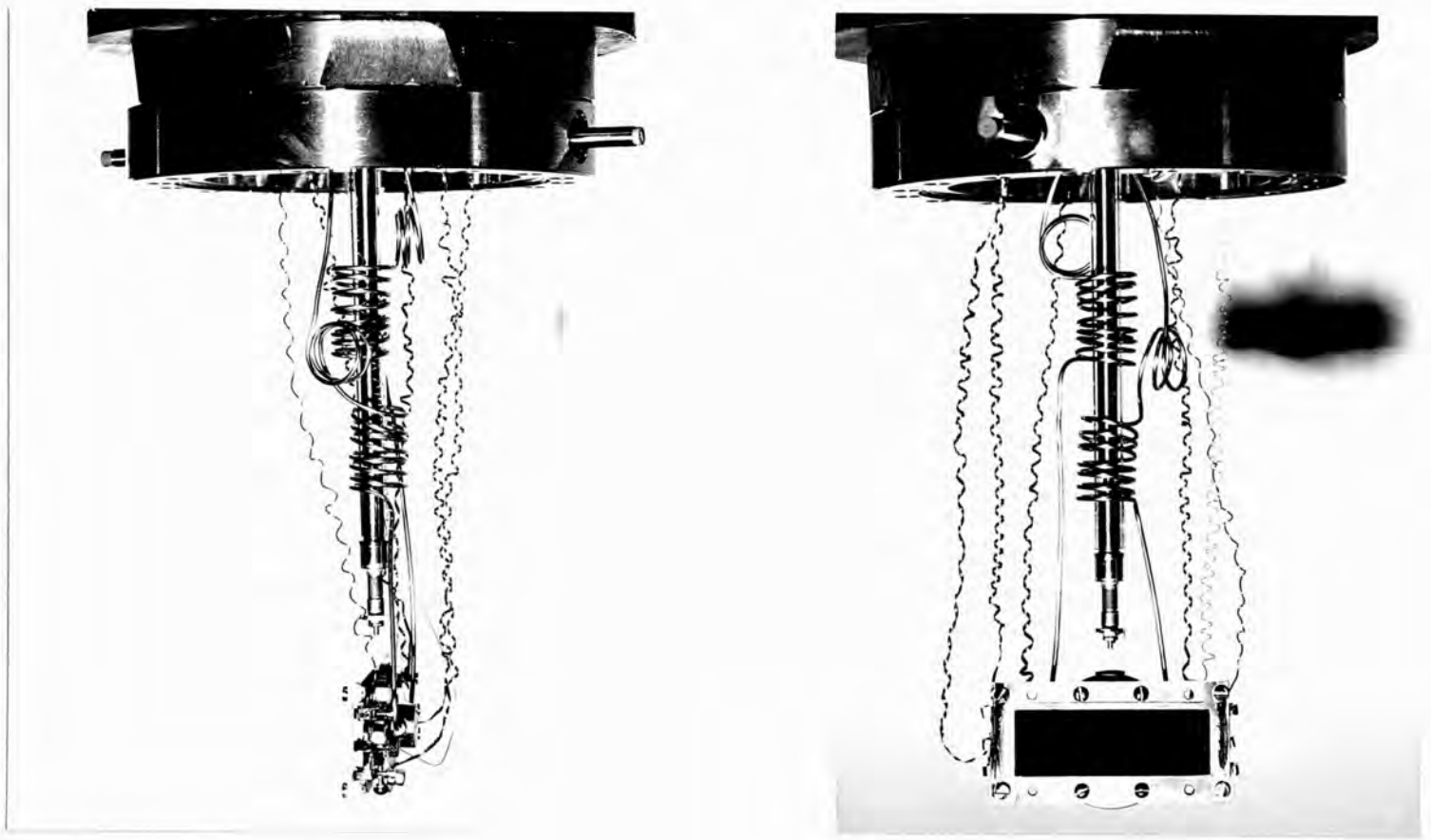
Electrical contact to the U.H.V. deposited films was made via vapour-deposited nichrome electrodes, for in situ measurements. They were deposited prior to loading into the U.H.V. system, to avoid the need for a movable shutter to define their shape and an additional evaporation source, and to avoid the contamination of the U.H.V. system with vapourised nichrome. The choice of nichrome had originally been made because of its low temperature coefficient of resistance, but since the resistance measurements were eventually made by the potential probe method, this property was not critical. Nichrome was retained, however, because it is physically tough and in particular because it easily survived the rigorous cleaning given to the substrate before its use in U.H.V. The vapour source for the deposition has been described in paragraph 4.3. The shape of the 4 electrodes

was defined by a mask made from stainless-steel shim which was held out-of-contact by 1/32", to soften the profile of the deposited film and allow good electrical contact between the electrode and the rare earth film deposited on top. At the end of each electrode remote from the rare earth film, pads were painted using Acheson Colloids "Dag 1950", which is a suspension of colloidal silver in distilled water. Small G-clamps, made of O.F.H.C. copper and stainless-steel, were fixed to the substrate; the soft, relatively thick silver pads ensured a large area intimate connection and protected the nichrome against scratching. The gold wires were then wound tightly around the screws of the G-clamps and connected at the other ends to the electrical feedthroughs integrated into the substrate holder feedthrough flange. The substrate and G-clamps are shown in Figure 3.5(b).

4.6 The Substrate Holder.

The stainless steel substrate holder assembly was designed and constructed specifically for this work, see Figure 3.5(a). The main block was welded to the feedthrough assembly, by Vacuum Generators Ltd., so that the small-bore plumbing and small reservoir were in intimate contact for maximum thermal conductivity. The plumbing was to be used for cooling the substrate by passing cold liquid through it. The substrate, and behind it the substrate heater, were held in place against the block by a mask holder. The shim mask was again out-of-contact, this time to prevent electrical connection between the rare-earth film and the earthed holder. An exploded view of the substrate assembly, including the film-heater (see Chapter V) is given in Figure 3.5(b). Figures 3.5(a), 4.2(a) and 4.2(b) show the assembled holder, with the cooling pipes and the wires for electrical connections to the film, the film-heater and the chromel-alumel thermocouple. The latter was clamped to the substrate holder block by a small screw, at the top right-hand side in Figure 3.5(a).

The movement of the holder, for vapour-deposition and for alignment parallel to the oscillating tungsten block, has been described



(a) (b)

Figure 4.2 Overall View of the Substrate Assembly and Connections (a) side view (b) front view

in paragraph 3.4.

4.7 Film Deposition in O.H.V.

The equipment, filaments and substrates used for the nichrome-film growth in O.H.V. have already been described. The sources were always carefully melted and outgassed before the substrates were exposed for deposition by moving a shutter. The vapour sources for the gold films were also F12 3x .025W, but the gold wire, 0.5mm diameter, was simply folded and fitted into the inside of the helix instead of being wound around it. The aluminium source was a strip of tungsten ribbon, $\frac{1}{2}$" wide, bent to make a "V" to hold the aluminium both before and after melting. The aluminium was a 1" length of wire, 0.1" diameter, folded into a broad "U" resting in the "V" with one arm hanging down either side, as in Figure 4.3. The distance from source to substrate was approximately 5", with the substrate always being directly above the source. The nichrome wire was 80% nickel, 20% chromium, supplied by Vactite Co., the gold was 99.999% pure, and the aluminium 99.995%, both from Johnson-Matthey Metals.

4.8 Film Deposition and Purity in U.H.V.

The rare earth wires used for the depositions were supplied by either Kochlight or Goodfellow Metals. The Gd and Dy were 99.9% pure and the Tb 99.95%. Typical analyses are given in Table 4.1. Using the filaments prepared and outgassed as described in paragraph 4.3, the depositions of the films were commenced after the practicable base pressures of $\sim 2.10^{-10}$ torr had been reached. The source-to-substrate distance was approximately 6". For gadolinium and terbium, the filament current was increased slowly to the required level, to avoid premature loss of charge, taking typically $1\frac{1}{2}$ minutes to reach 90% of the final figure. The maximum current was around 50 Amps, being slightly more, in general, for terbium deposition than for gadolinium; which is consistent with the slightly higher melting point. The maximum current was determined by experience as a compromise between deposition rate (and therefore time) and the pressure increase accompanying

the heating. For most of each deposition, the pressure was of the order of 1.10^{-8} torr or less, and the deposition was ceased when the pressure began to rise above that. The beginning and end of each deposition was defined by moving the shutter. The pressure was observed to fall rapidly as soon as the filament current was turned off, falling to $\sim 10^{-9}$ torr in seconds and to the base pressure after between 2 and 20 minutes. The growth time was typically between 1 and 6 minutes. The dysprosium deposition was achieved by sublimation. After outgassing for $1\frac{1}{2}$ hours at up to 20 Amps, with brief pulses up to 40 Amps, rapid film growth was effected by sequential pulsing of the current up to 60 Amps, each pulse lasting less than a second.

Table 4.1 Typical Metallic Impurity Analyses of Rare Earth Wires (in p.p.m.) (Goodfellow Metals)

Gd 99.9%:	Eu 171, Tb 170, Nd 166, Fe 80, Al 40, Si 20, Ca 10, Cu 1, Mg < 1
Tb 99.95%:	Y 146, Ho 84, Sm 84, Dy 82, Ca 3, Mg 3
Dy 99.9%:	Y 365, Er 85, Ho 85, Si 20, Fe 5, Mg 1, Mn < 1

Figure 4.4 shows the time required for a monolayer of nitrogen to be formed on a freshly-deposited surface, assuming unit sticking probability, vs. the pressure of nitrogen, (from a chart devised by L. de Chernatony, GEC-AEI Vacuum Consultants Laboratory). Given that the pressure at the end of each deposition was of the order of 3.10^{-8} torr, and reduced to 1.10^{-9} torr within seconds and to 3.10^{-10} torr within 15 minutes, then in the first hour after deposition, the total contamination of the surface due to gas bombardment would be $\lesssim 50\%$ of a monolayer, and thereafter at a rate of a monolayer equivalent every $2\frac{1}{2}$ hours. This is a worst case approximation, since it assumes

that the total contamination rate is independent of the residual gas constituents, that the sticking probability is unity for all molecules; and is based on an estimation of the actual curve of pressure vs. time. In order to distinguish between the effects of gas contamination and elapsed time, two films (numbers 33 and 34) of gadolinium were deposited from base pressures as high as $3 \cdot 10^{-8}$ torr. The work functions of these two films were observed to increase for tens of hours after deposition (see Chapter VI); whereas of the gadolinium films deposited from U.H.V. base pressures, only one showed a comparable change in the first few hours after deposition, and that was a decrease.

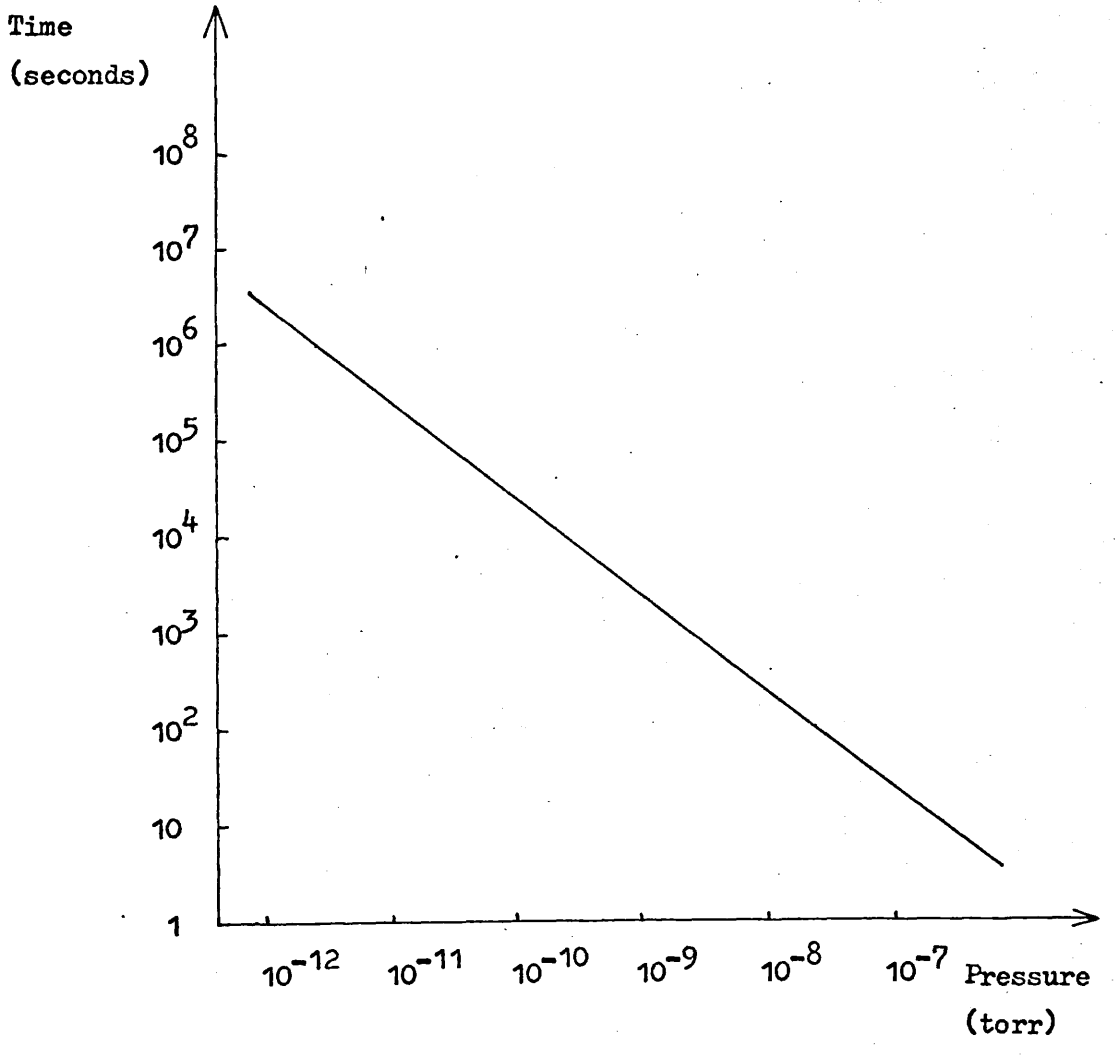


Figure 4.4 Monolayer Formation Time for Nitrogen (assuming sticking probability = 1)

After analysis according to the published "cracking patterns" [42], the Residual Gas Analyser (R.G.A.) attached to the chamber showed the general presence of water vapour and methane (CH_4), as is usual, and ammonia. The R.G.A. spectra fell into three groups. Firstly, for films up to and including number 27, water vapour provided the major peak, followed by ammonia, hydrogen, methane and ethylene (C_2H_4), with traces which could be attributable to such gases as argon, carbon dioxide, acetylene (C_2H_2), propene (C_3H_6), propane (C_3H_8), and N-butane (C_4H_{10}). Significant peaks were also seen at masses 23 to 26, which were not explainable by the standard "cracking patterns" of the more common simple gases.

On enquiring from the manufacturers of the Dag 1950 silver "paint", (Acheson Colloids) it was discovered that a small amount of proprietary binder was present therein, which was described as a "chain-link molecule, including hydrocarbons and a sodium sulphate component". Sodium having a mass of 23 a.m.u., it was assumed that this was the explanation for the peaks in the low 20's, and may also have given rise to the peaks initially attributed to the hydrocarbon gases. Unfortunately the "cracking pattern" for the binder was not known.

The second group of R.G.A. spectra was for films 28 to 30, and showed ethane (C_2H_6) and ammonia as the main constituents, followed by water vapour, methane, hydrogen and ethylene, and, as in the first group, traces for other gases and in the low 20's.

The third group of R.G.A. spectra was for films after number 30, and showed methane as most abundant, followed by water vapour, nitrogen, ammonia and hydrogen, and with other peaks as before.

The hydrocarbon gases were not thought to be contamination from rotary pump oil, as the characteristic peaks in the range 50 to 60 a.m.u.[42] were not observed under U.H.V. conditions. Typical mass spectra for the three groups are shown in Figure 4.5.

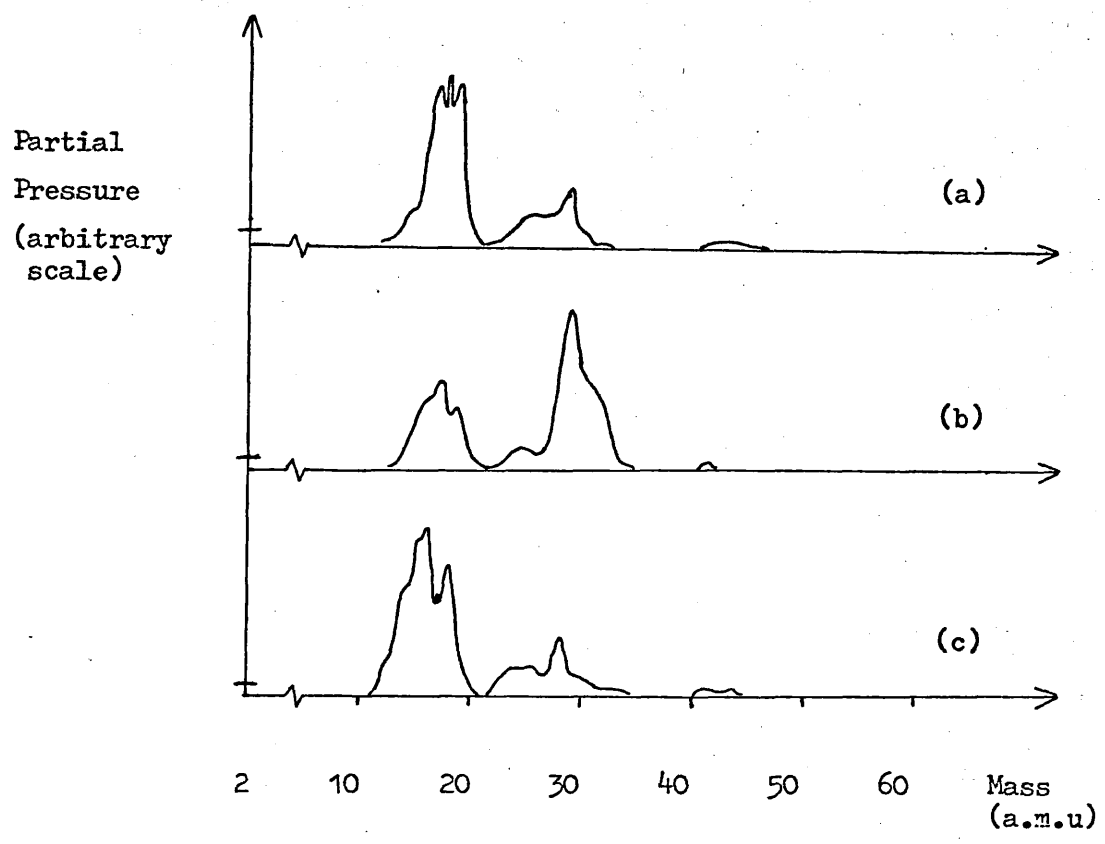


Figure 4.5 Typical Mass Spectra of Residual Gases
(a) Films up to no. 27 (b) Films 28 to 30 (c) Films 31 to 35

Surprisingly, there were very few observable changes in the mass spectra from any of the experimental variations, such as cryo-pumping, filament outgassing and evaporation, or substrate heating. For the third group of films where nitrogen was a significant residual component, it was seen to be reduced by outgassing and evaporation and increased by heating the coated substrate, suggesting inclusion in the rare earth film. Similar behaviour of the ammonia constituent was sometimes seen. The water vapour constituent sometimes varied, but not repeatably, sometimes increasing and sometimes decreasing. The sensitivity of the R.G.A. to water vapour is known not to be well behaved[43]

Since the relative sensitivity of the R.G.A. to all the various constituents was unknown, exact comparisons of partial pressures and relative bombardment rates of different gases at the sample were not calculable.

The electrical properties of the films were not separable into groups corresponding to the different mass spectra.

The tungsten film (number 35) used to check the work function of the reference plate was evaporated in 21 separate runs, the power being turned off to keep the pressure and filament feedthrough temperature within reasonable limits. The total evaporation time was 24 minutes, at pressures in the 10^{-8} torr range.

After the electrical measurements, the tungsten film was taken to the Fulmer Research Institute for analysis by X-ray fluorescence. It was to be expected that some atmospheric contamination would occur, so the analysis could only realistically indicate the non-gaseous impurities. A minimum of 97% weight tungsten was found, with the impurity being mainly Al_2O_3. Silicon and selenium were just detectable at 10keV beam voltage but not at 4keV, indicating that they were not present at the surface of the film and possibly only in the substrate, the film having been penetrated by the higher energy beam. It was thought possible that the selenium was remnant from the prior use of the analyser on a different sample, which did contain selenium. The alumina and silicon may have been due to the glass substrate, since the tungsten film was only around 1000Å thick. Therefore the purity of the film appears to have been at least 97%, possibly much closer to 100%.

4.9 Film Thicknesses

At the end of each substrate, outside the area used for resistance or work function measurements, a thin cover-slide was clamped, to provide a straight, well-defined edge for thickness measurement. The step was vacuum-coated with a layer of highly reflective silver and assembled with a plain flat glass similarly coated, to form an air-gap wedge. The principle of multiple-beam interferometry for thickness measurements is well known[46]. The multiplicity of beams generates a very small line-space ratio fringe pattern, so that a step in the wedge-thickness can readily be measured with an accuracy of better than 1% of the wavelength of the light used. The thickness of the

step is calculated as the ratio of fringe displacement to fringe separation multiplied by half the wavelength of the monochromatic light. In this work, the thickness was not required with critical accuracy, so a modestly flat reference plate was used, consisting of a piece of float-glass microscope slide. The surface finish of such slides is not specified, although the parallelism of the two working surfaces is specified to be within 5' of arc. Assuming that the worst-case surface flatness corresponded to 5' of arc over the observed area of approximate diameter 0.01mm, then the surface would be flat to within approximately 150Å, or about $\lambda/40$. Table 4.2 shows the measured thicknesses of the films, at the ends of the substrates, each being the mean of several adjacent fringe-step measurements. The spread of these measurements for each film was significantly smaller than the worst-case surface flatness estimated above, indicating that the reference slide was in fact smoother than that.

Using the formulae for thickness distribution from an extended wire source given by Holland[45] and including the physical dimensions of the source and substrate, the thickness at the very centre of the film was calculated to be 5.7% greater than at the measured step.

The errors in the measured values, estimated from the spread of fringe displacements over several adjacent fringes, were between 1 and 6%, except for the tungsten film which was 8%.

Using the same formulae, the maximum variation of thickness over the film area used for work function measurements (1.5" x 0.75"), was calculated to be 3%.

Exact measurement of the deposition rate was not possible, and it probably varied during each film. The average deposition rate for each rare-earth film was between 85 and 2500Å/minute, although almost all fell between 200 and 800Å/minute. The average deposition rate for the tungsten film was approximately 40Å/minute.

Table 4.2 Film Thicknesses

Film no.	Material	Mean thickness at end of substrate (Å)	Estimated thickness at centre of film (Å)
16	Gd	817 (±8)	864
17	Gd	997 (±18)	1054
18	Gd	1124 (±75)	1188
19	Tb	2405 (±80)	2542
21	Tb	1337 (±60)	1413
22	Tb	1124 (±70)	1188
23	Tb	1107 (±20)	1170
24	Tb	1221 (±36)	1290
25	Dy	688 (±42)	727
27	Gd	755 (±14)	798
28	Gd	1014 (±24)	1072
29	Gd	936 (±80)	989
30	Gd	1314 (±30)	1389
33	Gd	790 (±45)	835
34	Gd	657 (±30)	694
35	W	950 (±80)	1004

Chapter V The Development of a Film Heater for U.H.V. Use

5.1 Introduction.

Several techniques are available for heating a sample in vacuo, including conduction from hot fluids or heated solids, radiation from luminous or non-luminous bodies, and direct Joule heating of the sample, if conductive. The latter method was immediately rejected for this work since it would necessitate a potential gradient along the sample. Luminous radiant heating, from an incandescent wire, either naked or in the form of a light bulb, was rejected because of possible outgassing problems at the high temperatures involved and because the light would interfere with optical experiments. A non-luminous radiant heater might also be expected to cause more outgassing than an in-contact heater because it would need higher temperature or surface area than the latter. Also, with an in-contact heater it would generally be easier to avoid mechanical and electrical design difficulties inside the vacuum chamber. The plumbing to the substrate holder which was available for most of the work could have been used to carry hot fluid, e.g. kerosene, but the thermal capacity of the external temperature bath holding the kerosene would have caused a long delay. The alternative of switching to a second fluid was expected to cause problems due to the smallness of the pipe bore: the possibility of any residue from the heating fluid being frozen by the kerosene temperature, and blocking the system. Joule heating in O.H.V. systems is usually achieved with a refractory metal helix of some form, either resting in pyrofilite or wound round mica. In order to avoid contamination from these materials and to provide relatively intimate contact between the heater and substrate, it was decided to develop a heater in the form of a film on a similar substrate. Such a heater would readily be usable in any vacuum system without special fittings, such as plumbing, would be very compact, would not interfere with optical measurements, would be U.H.V. compatible, and should be capable of more uniform heat distribution than a bent wire.

It was decided to use a metal of high resistivity, in order that a relatively thick film be used. This allowed a high power dissipation, while the sheet resistance would still be sufficiently high that a simple rectangular geometry could be used rather than a meandering pattern, which would cause non-uniformity and possibly failure due to localised heating. Nichrome was chosen because of its ease of preparation and durability.

5.2 Preparation in O.H.V.

The nichrome was prepared by evaporation from wire of 80% nickel: 20% chromium as described for the contacts on the work function specimen substrate in Chapter IV. The evaporation distance was increased to 8" for better uniformity of film thickness, and hence of power dissipation and temperature when used as a heater. To promote adhesion, the glass microscope slides onto which the films were deposited were heated to $\sim 250^{\circ}C$. In the first instance, the heater took the form of a flattened helix of fine nichrome wire wound round a rectangular mica former slightly smaller than the slide. This film was used as the heater for a second film, and so on, till five films had been deposited, so that the non-uniformity of temperature distribution caused by the original wire heater was reduced. For good electrical contact each film heater was provided with end-contacts of Dag 915 silver "paint", leaving an area $2\frac{1}{2}$" long by 1" wide. The thickness of the deposited films was arbitrarily determined by the opacity, viewing the evaporation filament through the film. The resulting sheet resistivities were between 12 and 30 ohms/square, corresponding to thicknesses between 1300 and 300 $\mathring{A}$.

5.3 Initial Tests of Heater Performance.

Measurements of heater temperature vs. power dissipation were initially made using a chromel-alumel thermocouple, to see if the film was capable of producing and withstanding reasonable temperatures. Figure 5.1 shows the results, with temperatures over $280^{\circ}C$ being achievable without damage.

To assess temperature uniformity, separate microscope slides were given temperature-sensitive coatings. Manufactured by Faber-Castell under the name of "Thermocolor"[46], the coatings change colour at a fairly well defined temperature. The manufacurers specify a tolerance of $\pm 5^{\circ}C$, with an additional tolerance depending on the time taken for the colour change to occur. Table 5.1 shows the nominal temperature values for the four types used, based on the standard change time of 30 minutes. The resolution quoted is the total error range if the change occurs between 10 and 30 minutes after commencement of heating, more rapid change signifying higher temperature.

Table 5.1 "Thermocolor" Paint Temperature Indication

Type no.	Colour change	Nominal Temp. (30 mins)	Resolution (10-30 mins)
2b	Pink to Lilac	$95^{\circ}C$	$3^{\circ}C$
5	White to Brown	$175^{\circ}C$	$2^{\circ}C$
6	Green to Brown	$220^{\circ}C$	$5^{\circ}C$
7	Red to Brown	$290^{\circ}C$	$2^{\circ}C$

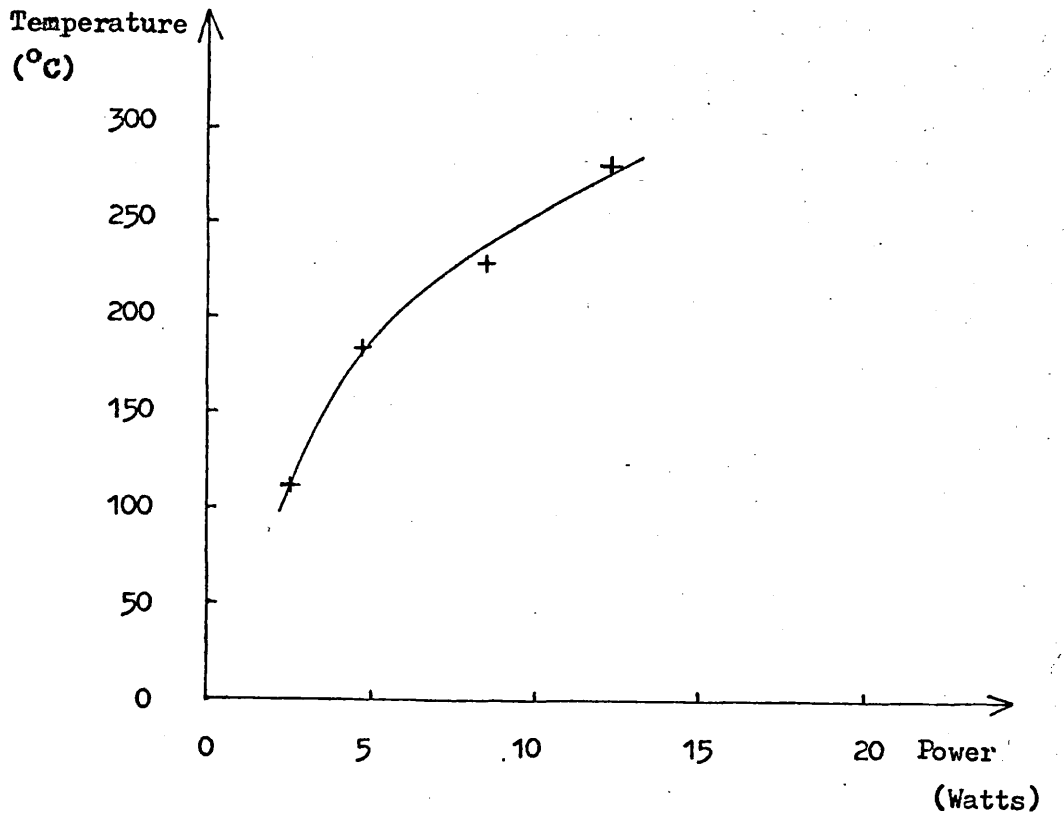


Figure 5.1 Temperature vs. Power Dissipation (initial tests)

The paints were supplied in powder form, and applied by suspending in iso-propyl alcohol and brushing onto the glass slide. After the alcohol had evaporated, the slide was placed next to a heater substrate, with the two thicknesses of glass between the heater and paint, and assembled into a holder similar to that used in the U.H.V. chamber. The tests were conducted in O.H.V. The power dissipation in the heater was gradually raised till a colour change just started, then the power was kept constant while the spread of the colour change was observed. After approximately a quarter hour the change had spread from a small spot to cover over half the film, and to almost all the film within an hour. Two tests showed similar assymmetry of heating, being eccentric on the substrate. A third attempt with the "Thermocolor" applied directly to the obverse side of the heater substrate was also asymmetrical, indicating that the assymmetry was not due to imperfect thermal contact between the two glass slides. A second batch of heater films was therefore prepared, as before except

that particular care was taken to ensure symmetrical source-to-substrate geometry.

A film was also deposited onto an alumina substrate, having a higher softening temperature than glass, to see how high a temperature could be reached with a film heater. Using a thermocouple a temperature of over 500°C was observed, after which the substrate broke, probably encouraged by mechanical stress from the clamping arrangement.

The second batch of films on glass slides was used for a 3rd Year undergraduate project[47], to continue the evaluation as described below.

5.4 Further Tests and Calibration.

Using the nichrome wire-on-mica and nichrome film heaters, measurements were repeated for temperature vs. power dissipation. Again, the film heaters were more efficient than the wire-on-mica heater, and in spite of differences in sheet resistivity were very repeatable and reproducible, see Figure 5.2.

Temperature uniformity tests were carried out slightly differently than before. Using the results in Figure 5.2, the powers required to obtain the temperatures at which the "Thermocolor" paints would change colour were deduced, and applied for 20 minutes. Three tests were conducted for each grade of "Thermocolor" on each heater, to establish reproducibility. The area and colour density indicated the uniformity of temperature distribution. Figure 5.3 shows the results from a typical film and from the original wire-on-mica heater, and that the heater should be larger than the specimen for optimum uniformity of specimen temperature. From the resolutions quoted in Table 5.1, it can be seen that the central portions of the substrates heated by the film heaters were uniform to around 5°C or better.

The heaters used in the U.H.V. chamber were slightly shorter

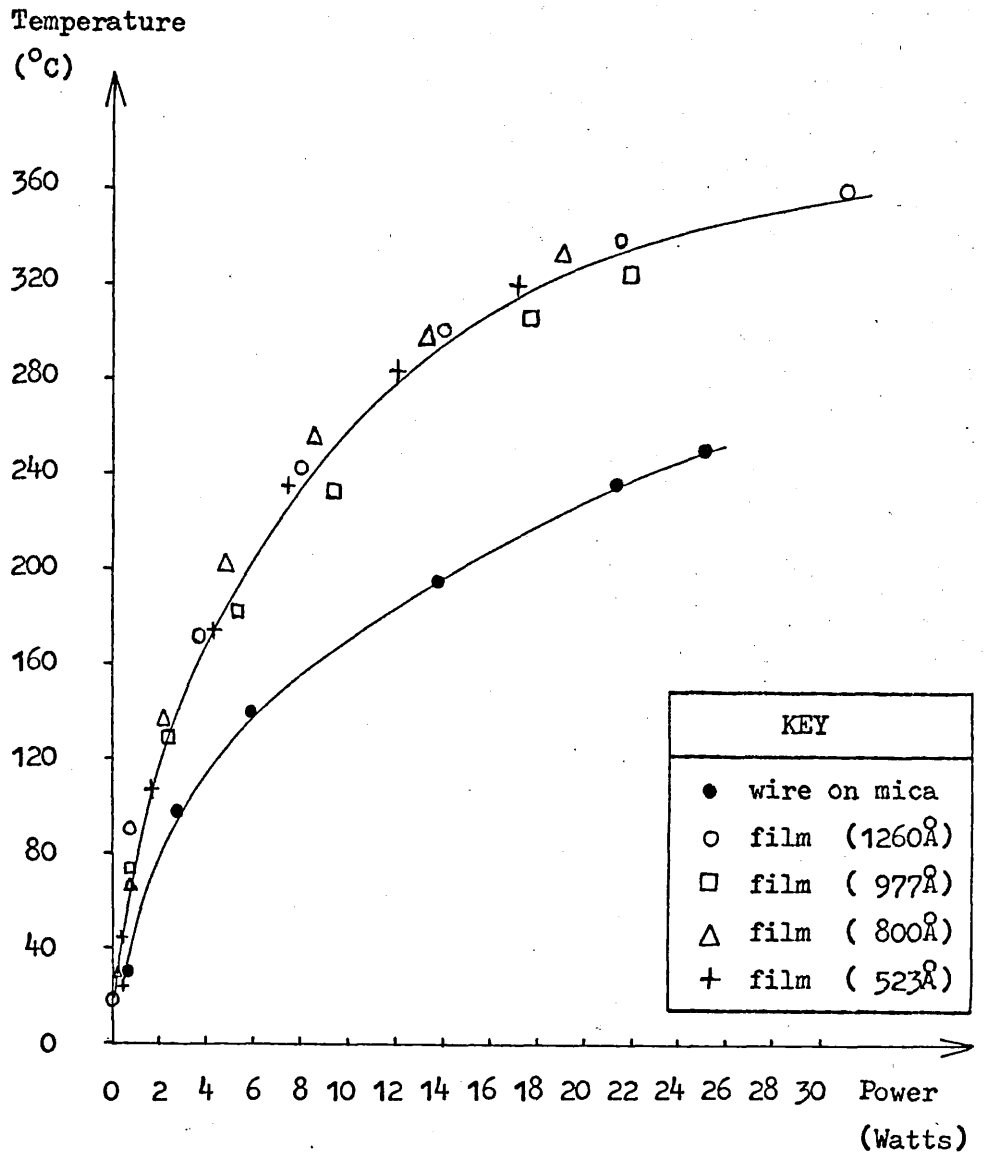


Figure 5.2 Temperature vs. Power Dissipation in Heaters (after reference 47)

than those described above, at 2" long, see Figure 3.5(b). For the work function measurements the relevant area of the specimens was $1\frac{1}{2}$" by $\frac{3}{4}$", so that temperature uniformity over that area may reasonably be assumed to be within 5°C.

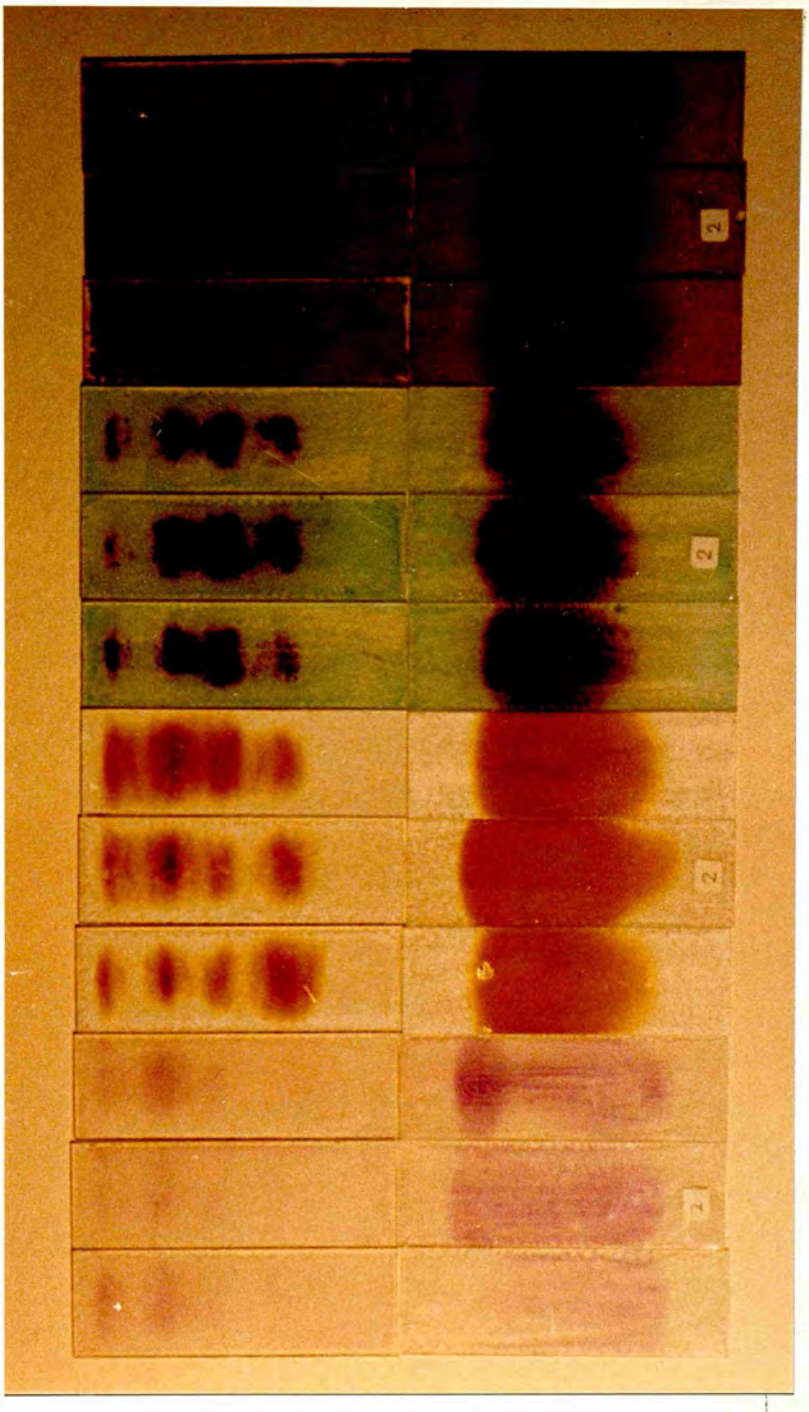


Figure 5.3. Temperature Distribution of Wire Heaters and Film Heaters using "Thermocolor" paints (after reference 47).

Chapter VI The Experimental Results for Rare Earth Metals

6.1 Introduction.

The method of work function measurement and the practical details of its execution have been described in Chapters II and III, including the ways in which the various sources of error were evaluated. The resistivities of the films were also measured, using the potential probe technique for the resistance and multiple-beam interferometry for the thickness (see paragraph 4.9). Cryogenic facilities to attain the low temperatures required for the determination of the residual resistivity of the films were not available, so the resistivities simply serve to compare the film when fresh and aged, and approximately with the values published elsewhere.

The results from various conditions and histories are presented and analysed, including the effects of aging, annealing and atmospheric contamination. Some theoretical values are derived which partly explain the spread in results. Finally, the results are summarised and compared with results published elsewhere.

6.2 Measurements on Gadolinium Films in U.H.V.

The gadolinium films were measured under various conditions and with various histories. The later films were subjected to temperatures slightly above and slightly below the Curie point, and were also annealed and the measurements repeated. The spread of values observed was larger than expected, and the c.p.d.'s were seen to vary with time, albeit slowly. To observe whether these variations were caused by gas contamination or were a function only of time, two films (no.'s 33 and 34) were prepared and measured in a slightly higher residual pressure. The results are presented in several forms in the following paragraphs.

6.2.1 Specimen Reproducibility of Gd Films

The results of measurements on the deposited films are presented in the form in which they were obtained i.e. contact potential difference between the film and the reference surface. Table 6.1 shows the c.p.d. values both at initial measurement and after aging in U.H.V. In the case of films which were heated in vacuo, the latter figure is that immediately prior to the heating. The figures given in brackets are the errors, and are the range of bucking potential over which the signal from the capacitor was at a null or minimum, plus the range of bucking potential in the measurement of the W film (no. 35) used to check the reference surface, which was $\pm$0.0015 eV. The error in the change of c.p.d. is the square root of the sum of the squares of the errors in (a) and (b).

It is evident that there was a large spread, not only in the results on fresh films but also on the aged films. There appeared to be no simple correlation between the c.p.d. and film thickness, (see paragraph 4.9) nor residual gas analysis (see paragraph 4.8) nor deposition time. There was, however, a tendency for the c.p.d. to be smaller for the higher values of the product of deposition time and pressure (see paragraph 6.6). Prior to the growth of films 27 and 30, the substrates were heated, in U.H.V., to over 370^{o}K by means of the nichrome-film heater, although deposition always occurred with the substrate initially at room-temperature and unheated except by the deposition process itself. The pre-deposition heating acted as a final substrate-cleaning operation, and may account for the generally higher c.p.d.'s (and correspondingly lower work function values) of these later films by comparison with no.'s 14, 16 and 18. However, films no.'s 33 and 34 were intentionally deposited and measured at slightly higher pressures but did not reproduce the lower c.p.d. values and therefore do not corroborate the hypothesis that the lower c.p.d.'s of the earliest films were due to contamination within the films.

Table 6.1 Contact Potential Difference Results: Gd-W

Film No.	c.p.d. (eV)		Change in c.p.d. (eV)	Age at (b) (hours)
	(a) fresh	(b) aged		
14	- (-)	-1.48 (±0.011)	- (-)	212
16	-0.927 (±0.008)	-0.998 (±0.021)	-0.051 (±0.022)	30
18	-0.628 (±0.003)	-0.670 (±0.004)	-0.052 (±0.005)	55
27	-1.366 (±0.021)	-1.417 (±0.011)	-0.051 (±0.023)	97
28	-1.005 (±0.011)	-1.005 (±0.011)	-0.000 (±0.015)	28
29	-1.270 (±0.011)	-1.325 (±0.011)	-0.055 (±0.015)	74
30	-1.060 (±0.011)	-1.095 (±0.011)	-0.035 (±0.015)	23
33	-1.030 (±0.014)	-0.910 (±0.011)	+0.120 (±0.018)	31
34	-1.617 (±0.011)	-1.305 (±0.011)	+0.312 (±0.015)	48

6.2.2 Measurements of Gd Films above and below the Curie Temperature

The Curie Temperature of bulk Gd is generally taken to be around $293^{\circ}K$[48]. The Curie Temperature of a U.H.V.-deposited film of Gd, inferred from the change in slope of a graph of resistance vs. temperature, has been measured as $290^{\circ}K$[49] and as $298^{\circ}K$[50]. As described earlier, the substrate holder was capable of being heated, by means of the nichrome-film heater, or cooled, by passing a cooling liquid through the small-bore pipework. Early attempts at using liquid nitrogen were abortive, since it was difficult to avoid condensation of water onto the liquid nitrogen and thence into the small-bore pipework where it caused a blockage. For reliable cooling, although to a lesser degree, it was decided to use a less cold liquid, and odourless kerosene was chosen because it flows readily at temperatures well below the Curie point of Gd. (It was considered that the magnetic ordering temperatures of Tb and Dy were outside the capability of the equipment in view of the difficulty with liquid nitrogen). Initial attempts to cool the kerosene were made by passing it through a helical copper tube immersed in iced water. The lowest temperature observed via the thermocouple was $289^{\circ}K$. By replacing the iced water with a solution of $CaCl_2$ cooled either by solid carbon dioxide or by liquid nitrogen, a temperature of $284^{\circ}K$ was seen. A refrigerated cooling tank was then used to cool the whole body of kerosene directly, and the lowest temperature recorded by the substrate holder thermocouple was $276.5^{\circ}K$. Since this was located at the edge of the holder and the cold reservoir extracting heat from the holder was relatively small, it was surmised that the temperature of the film was at least as distant from the ambient as the thermocouple. The temperature observed when the film was heated was also lower than expected from the calibration in Chapter V, probably for a similar reason, although the heater was larger than the film area used for work function measurements. The calibration of the thermocouple and the difference between the expected and observed (thermocouple) temperatures when the film was heated are described in Appendix III.

The results of c.p.d. measurements at various temperatures are shown in Figure 6.1. For each film two separate curves are given, for before and after annealing. In no case was there any significant change in c.p.d. in the region of the Curie point, T_c. The effects of anneal-

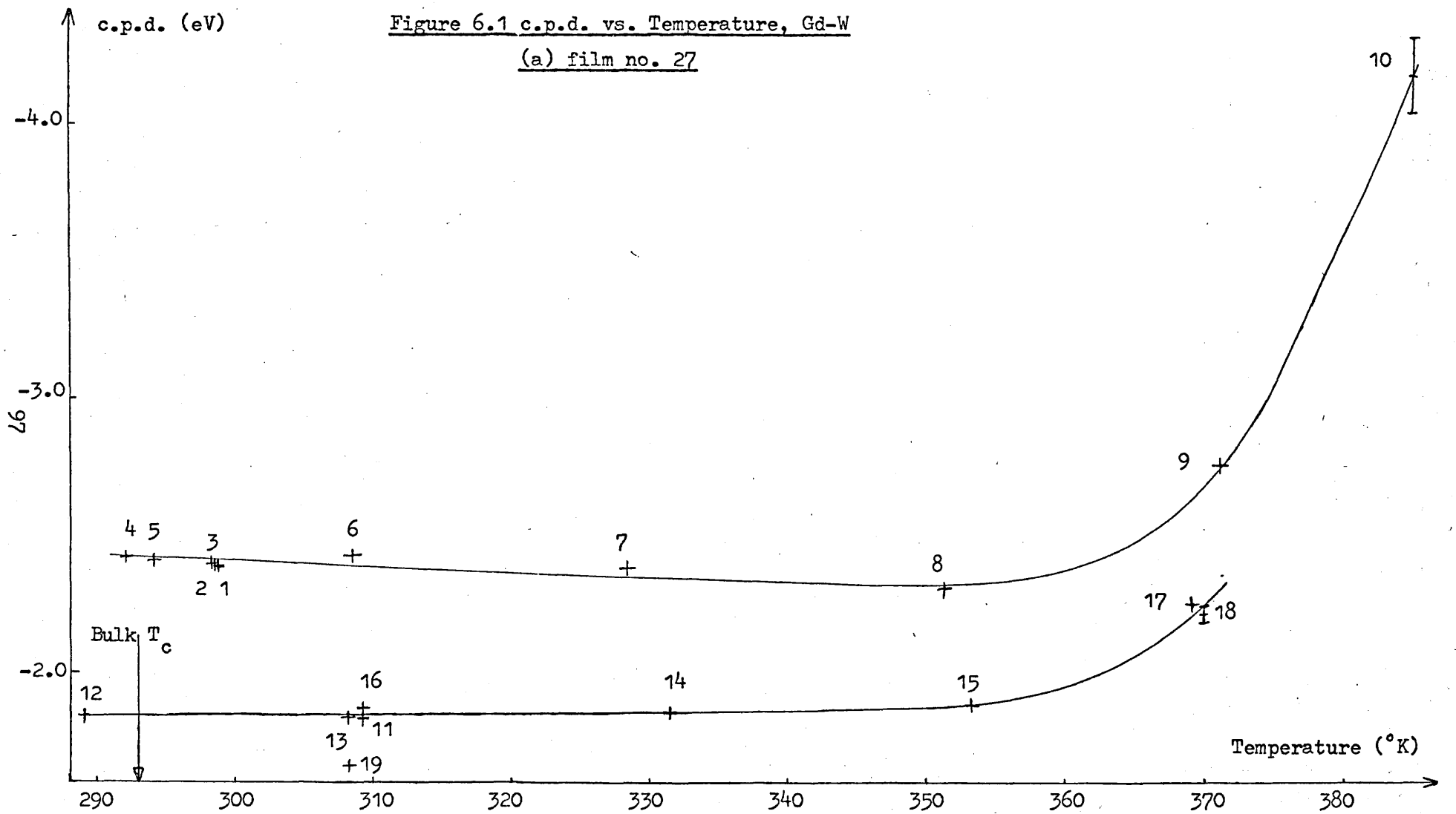


Figure 6.1 c.p.d. vs. Temperature, Gd-W

(a) film no. 27

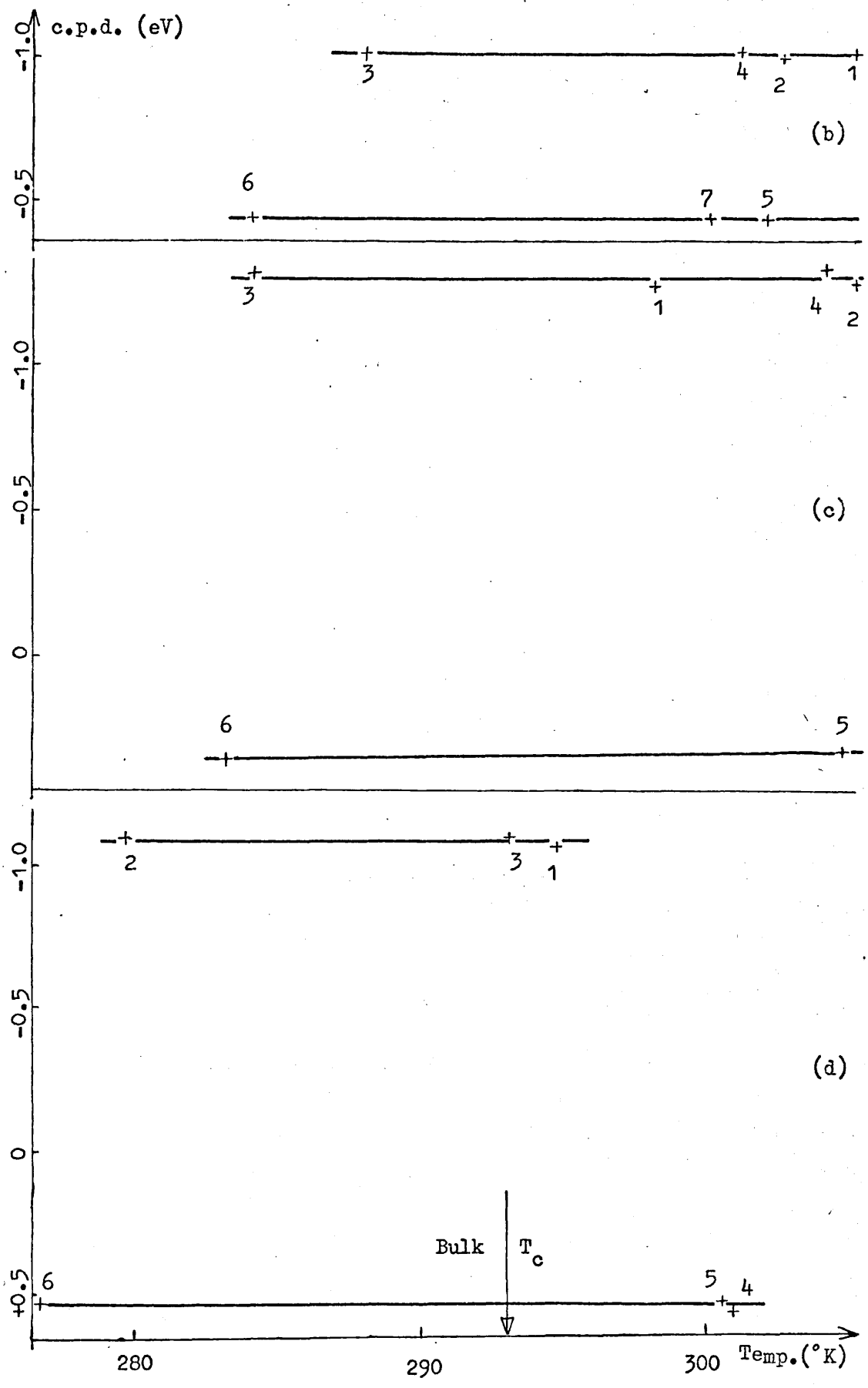


Figure 6.1 (b) film no. 28 (c) film no. 29 (d) film no. 30

ing are discussed in paragraph 6.2.4. The numbers adjacent to the experimental points indicate the sequence of measurement for each film. The errors in the c.p.d. measurements were smaller than the crosses marked on the figure. The exact values of the errors in the temperature measurements were unknown, but were likely to be quite large (see Appendix III). However, since the interest lay mainly in the relative temperatures for an individual specimen rather than the absolute values, no absolute error bars have been marked since they would be so large as to make the curves appear meaningless. For correctness the temperature values for each specimen should be replaced by values which are progressively more distant from ambient ($\sim 295^{\circ}K$) than those indicated, the relationship being different for each specimen but constant for each.

6.2.3 Effects of Aging on Gd Films

The changes in the c.p.d. between the Gd films and the tungsten reference are shown in Table 6.1, together with the ages of the films, in hours. It is perhaps remarkable that the changes observed in several were virtually identical (viz films no.'s 16, 18, 27 and 29) but no common factor was observed. What was obvious, however, was that the c.p.d.'s for the films evaporated under U.H.V. all increased on aging while the c.p.d.'s for the two films deposited under O.H.V., (no.'s 33 and 34) changed in the opposite sense and by far greater extents. The variations of the c.p.d.'s with time are shown in Figure 6.2. Again there was a significant difference between the U.H.V. and O.H.V. films, in that the former suffered far less rapid variations (with the exception of no. 16). This was, of course, to be expected since the bombardment rate of the residual atmosphere onto the surfaces of the films is proportional to the partial pressures of the gas species[40]. It seems likely, therefore, that the changes observed for the O.H.V. films resulted from contamination, although the mechanism(s) by which the changes occurred are less obvious. Also, since the c.p.d. variations in the U.H.V. films were in the opposite direction, it is likely that they were due to some effect other than gas contamination, such as crystal restructuring.

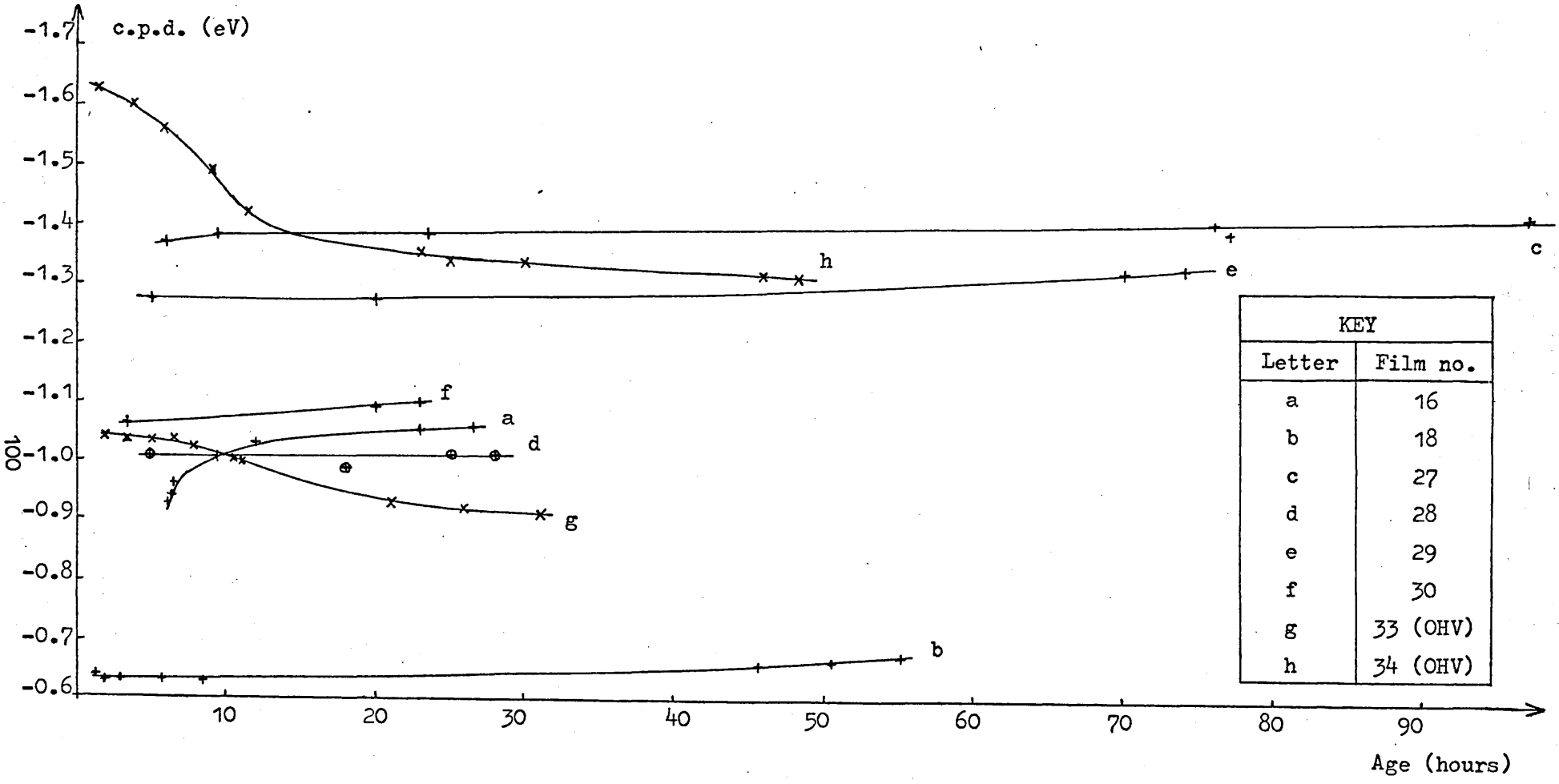


Figure 6.2 The Effect of Aging on c.p.d.:Gd-W

6.2.4 Effects of Annealing on Gd Films

After the aging period described above, some of the Gd films were heated, using the nichrome-film heater, and the c.p.d.'s remeasured. The results of measurements at or near room temperature are shown in Table 6.2, with the annealing temperature, the time to warm up and the time at the highest temperature, and the age of each film. The heater power was increased gradually, to avoid thermal shock to both the heater and the specimen. The two figures in the "Annealing Time" column show the time to reach the maximum temperature and the time for which it was sustained before the heater power was removed. The time taken to fall back to room temperature was around 3 hours.

The "Annealing Temperature" column shows both the temperatures recorded by the thermocouple situated at the corner of the substrate holder and the expected temperatures derived from the calibration curve of substrate temperature vs. heater power dissipation, Figure 5.5 (see Appendix III for discussion of this point). The variations of the c.p.d.'s as the temperature was varied are shown graphically in Figure 6.1

6.2.5 Resistivity and Electron Diffraction Measurements on Gd Films

Resistivity and electron diffraction measurements are potentially good indicators of the structure of thin films. In the rare earth metals there is considerable anisotropy in resistivity, as shown in Table 6.3 for bulk Gd, Tb and Dy[48]

Table 6.3 Bulk Resistivity of some Rare Earth Metals ($\mu\Omega$cm)

Metal	a-axis	c-axis
Gd	141	121
Tb	122	102
Dy	111	78

Table 6.2 The Effect of Annealing on c.p.d.: Gd - W.

Film I.D. no.	c.p.d. (eV) (a) Before annealing	c.p.d. (eV) (b) After annealing	Change in c.p.d. (eV)	Annealing Time (hours) warm-up	Annealing Time (hours) hold	Annealing Temp. (°K)	Age at (b) (hours)
27*	-1.417(±.011)	-0.830(±.011)	+0.587(±.015) } +0.737 (±.015)	8	2	385/450	120
	-0.840(±.011)	-0.870(±.011)	-0.030(±.015)	3½	1½	353/390	142
	-0.870(±.011)	-0.650(±.011)	+0.180(±.015)	6	1½	382/450	166
28	-1.005(±.011)	-0.432(±.011)	+0.573(±.015)	1½	½	403/500	47
29	-1.325(±.011)	+0.365(±.011)	+1.690(±.015)	2	1½	379/520	95
30	-1.095(±.011)	+0.575(±.011)	+1.670(±.015)	3	1	386/530	45
33	-0.910(±.011)	-0.230(±.011)	+0.680(±.015)	1	½	405/495	46

* Film no. 27 was heated three times.

Apart from the effects of surface roughness and of the film thickness being small compared to the electron mean-free-path, the resistivity of a film may differ from the bulk value due to impurities and defects. A further complication in the rare earths is the possibility of some or all of the film being in the f.c.c. form rather than the h.c.p. of the bulk. Investigations by other workers have shown that the f.c.c. phase can exist in films whose thickness is below 500Å [51] and that the resistivity is significantly lower. Experimentally derived resistivity ratios for the two phases are given in Table 6.4 for Gd, Tb, Dy and La. Reale's films were deposited at 10^{-9} torr [52] and Bist and Srivastava's at 10^{-6} torr [51]. The former consisted of a large number of randomly oriented grains and were measured in vacuo. The latter were measured in air, no orientation information was given.

Table 6.4 Resistivity of f.c.c./Resistivity of h.c.p.

	Reale [52]	Bist and Srivastava [51]
Gd	0.811	0.56
Tb	0.795	-
Dy	0.780	-
La	0.860	~0.33

In the present work, the resistances of the films were measured by the well-known potential probe method [53]. Four contacts were used, the two outermost for supply of current and the two innermost for measurement of the consequent potential across the specimen. Since the input impedance of the voltmeter was very much higher than the resistance of the sample, the resistances of the two potential contacts were not critical. Their only effect was to reduce the sensitivity of the measurement, and that very slightly since they also were small compared to the input impedance of the meter. The current contacts were also non-critical, since the potential measurement did not encompass them. The temperature coefficients of the contact resistances

were, therefore, also unimportant. The contact geometry used in this work is shown in Figure 3.5(b). The length and width of the samples were measured by travelling microscope, the length being taken as the distance between the centres of the two potential contacts. The potential value used to calculate the resistance was the average of measurements taken for both current directions. This was to eliminate as far as possible errors due to thermoelectric effects or rectification that may arise either from within the bulk of the sample due to inhomogeneities, from rectifying boundaries or from other defects. To eliminate inaccuracies caused by electrical pick-up in the voltmeter, the potentials were also recorded with the current source disconnected and the resulting values were subtracted from the current-driven measurements. The current source was a Keithley 610C Electrometer. The instrument may be used as an ohmmeter, for which purpose it has an accurate current source, the value being selected by the "Range" switch. The current used for this work was 10μAmps, chosen as a compromise between measurement sensitivity (limited by the voltmeter sensitivity) and the risk of electromigration in the specimen, and was the highest current available from the Keithley. The resulting current density was in the region of 1 Amp/cm^2. The accuracy and consistency of the current value was verified by voltage measurement at an internal calibration point, and was set to within $\pm$0.1%. The accuracy of the voltmeter, a $4\frac{1}{2}$ digit Solartron LM1440.2, was $\pm$(0.0033% of full scale + 0.01% of reading), and its input impedance was $2.10^{10}\Omega$. Since the specimen resistance values were between 6Ω and 52Ω and the current 10^{-5} Amps, the measured potentials were between 60 and 520μV. The sensitivity of the voltmeter was 10μV, which resulted in an accuracy of resistance measurement of between 2% and 16%, typically 3%. The accuracy of the resistivity calculations was worse, being subject to an additional error of between 1% and 9% in the film thicknesses (at the end of the specimen) and a theoretical variation of around 6% from the centre to the end (see paragraph 4.9). It was not attempted to measure the residual resistivity, caused by impurities and defects in the specimens, since the facilities were not available for the low temperatures required to do so. Hence it was not possible to measure changes in residual resistivity, due to crystal restructuring, gas absorption

etc., nor subsequently to calculate changes in the "true" resistivity due to a change of crystal orientation or structural phase. Accurate measurement of the total resistance was therefore inappropriate, since it would have been impossible either to interpret the reasons for small variations or accurately to compare the resistivities of the films with bulk values. The calculated resistivities are shown in Table 6.5.

Table 6.5 Resistivity of Gd Films

Film I.D. No.	Resistivity ($\mu\Omega$cm)	
	Fresh	Annealed
16	138 (±10)	-
18	160 (±20)	-
27	185 (±13)	178 (±13)
28	168 (±15)	155 (±14)
29	136 (±20)	-
33	74 (±10)	58 (±8)
34	140 (±14)	-

Measurements of resistance of some of the films were prevented by accidental shorting out of two or more of the naked gold wires by specimen manipulation in the course of optical measurements by other workers. Comparison with Table 6.3 shows that some figures lay within the range of bulk values, some were higher and one lower. The higher resistivities were possibly due to impurities and defects. Indeed, in those which were annealed the resistivity was decreased. In all cases the resistivity was lower than the 200$\mu\Omega$cm quoted by Bist and Srivastava[51] for thick Gd films grown at 10^{-6} torr. (This high pressure may have been the reason why earlier authors[50, 51] found difficulty in evaporating Gd.) The one sample with lower resistivity than the bulk was one of the two evaporated under O.H.V., in fact similar conditions to those of Bist and Srivastava. The resistivity ratio attri-

buted to them in Table 6.3 is a conservative one, being the ratio of their thinner films to the bulk material and ignoring both the size-effect and the higher-than-bulk resistivity of their thicker films. With both these effects included, the ratio would be below 0.36. The ratio of the measured resistivity of film no. 33 to the a- and c-axes of the bulk material were 0.525 and 0.61 respectively before annealing, and 0.41 and 0.48 respectively after annealing. These figures are consistent with the film being partly or wholly in the f.c.c. phase. This may have been due either to wide variation in film thickness so that some areas were less than 500Å thick, or to the f.c.c. phase existing in the relatively thick (835Å) film.

Except for the changes due to annealing, the resistivities were not seen to change with aging, being consistent, at the same temperatures, over several days while the c.p.d.'s were changing. The temperature coefficients of resistance were calculated over various ranges of temperature between 293° and 384°K. The values were not particularly accurate due to uncertainty in temperature and to the resolution of the resistance measurements. The results ranged from 0.05%/°K to 0.12%/°K, which bracket the published value of 0.067%/°K over the temperature range 300 to 375°K[54] for bulk Gd.

Electron diffraction observations were conducted in reflection due to the thicknesses of the films. In many cases only a small number of quite broad diffuse rings were seen, indicating a very small grain size, but of insufficient quantity and clarity to determine the crystal structure. In a few cases, however, interpretable patterns were obtained. These were probably produced by transmission through localized promontories and generally indicated an h.c.p. structure with the c-axis roughly perpendicular to the substrate surface. In one case (film no. 16) a single crystallite was observed, of considerable size judging by the smallness of the spots. At another site of the same film the pattern showed several crystallites at slightly different orientations. From another film (no. 30) a pattern was obtained which consisted of spots in rings, showing polycrystallinity. The grouping of the rings clearly indicated f.c.c. structure, although the thickness

of the film (1389Å) suggested that h.c.p. should be expected. Due to the difficulty in obtaining reflection diffraction patterns and interpreting them as certainly as transmission patterns, the results are less conclusive. It was shown, however, that the films were polycrystalline, generally of h.c.p. structure but with the possibility of some crystals in the f.c.c. phase.

6.3 Measurements on Terbium Films in U.H.V.

The measurements on the Tb films were all made at ambient temperature, the thermocouple recording temperatures between 19 and 26°C. The work functions were measured over a period of time after deposition. As with Gd, the spread of results was larger than expected and the c.p.d.'s varied slowly with time.

6.3.1 Specimen Reproducibility of Tb Films

Table 6.6 shows the c.p.d. values obtained for the various Tb films at initial measurement and after a period of time in the U.H.V. environment. Again there was no simple correlation between the c.p.d. and either film thickness, residual gas analysis or deposition time, but there was a tendency for the c.p.d. to be larger for the higher values of the product of deposition time and pressure. This is the opposite trend to that in the Gd films, and can therefore be attributed to the rare earth metals rather than the tungsten reference surface.

6.3.2 Effects of Aging on Tb films

The overall effects of aging on the c.p.d. results are shown in Table 6.6 and the variations as a function of time are shown graphically in Figure 6.3. The changes in c.p.d. were predominantly in the same sense as those on Gd, and therefore gas contamination was not thought to be a primary cause for these films (see paragraph 6.2.3). Two films underwent c.p.d. changes in the opposite direction. Film no. 19 was the first Tb film to be deposited and the evaporation source was outgassed at a lower current than all the others. Since the unchar-

Table 6.6 Contact Potential Difference Results: Tb-W

Film I.D. No.	c.p.d. (eV)		Change in c.p.d. (eV)	Age at (b) (hours)
	(a) fresh	(b) aged		
19	-1.171 (±0.003)	-1.115 (±0.011)	+0.056 (±0.012)	44
20	-1.055 (±0.021)	-1.078 (±0.021)	-0.023 (±0.029)	21
21	-1.375 (±0.016)	-1.382 (±0-016)	-0-007 (±0.022)	18
22	-1.570 (±0.021)	-1.600 (±0.021)	-0.030 (±0.029)	26
23	-1.437 (±0.026)	-1.430 (±0.011)	+0.007 (±0.028)	19
24	-1.727 (±0.031)	-1.845 (±0.011)	-0.118 (±0.033)	13

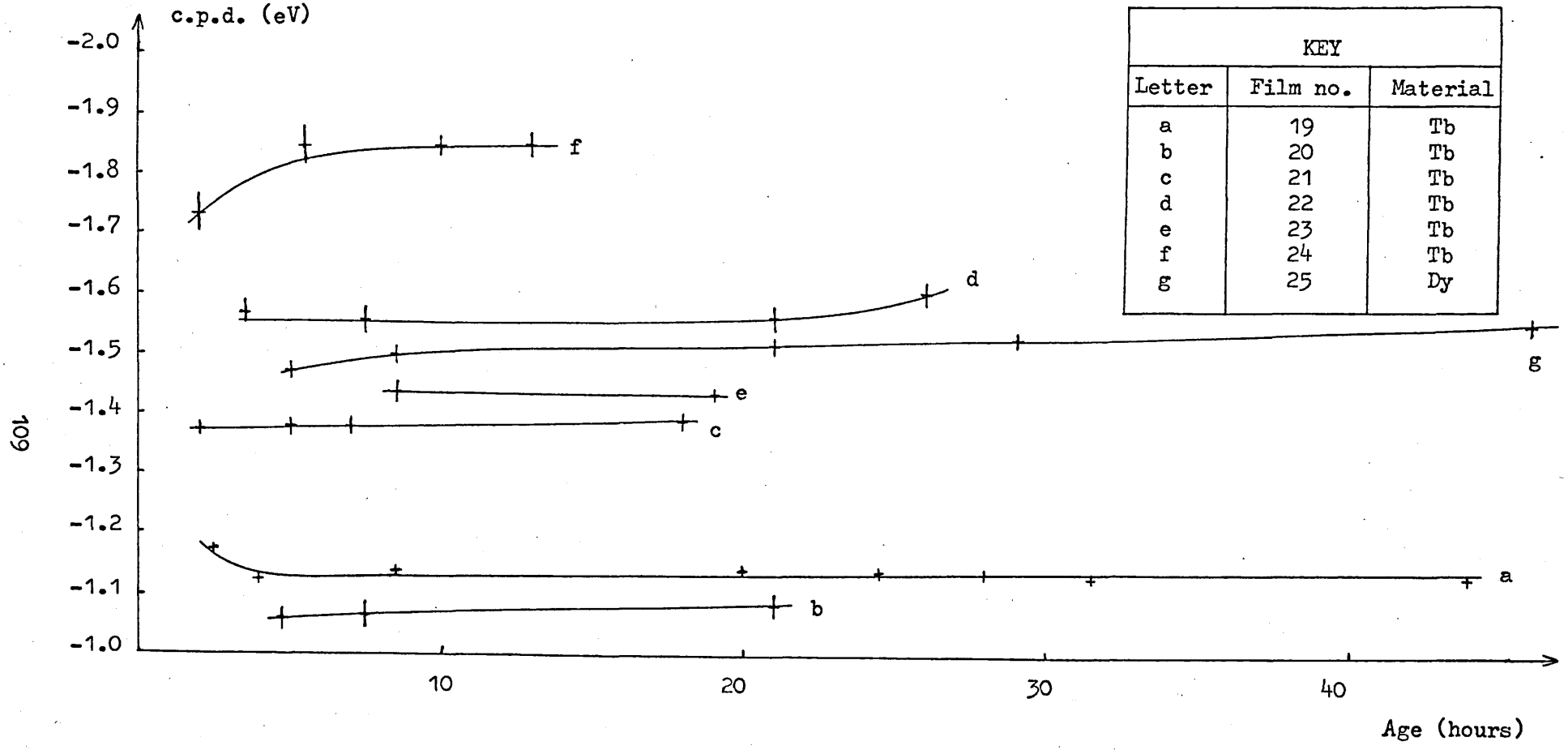


Figure 6.3 The Effect of Aging on c.p.d.:Tb-W and Dy-W

acteristically rapid early change in c.p.d. was in the same sense as those observed in the O.H.V.-deposited Gd films, and since the resistivity was abnormally low, as for one of the O.H.V. Gd films (see Tables 6.5 and 6.7), it is possible that the outgassing was inadequate and the film consequently contaminated. The other Tb film for which a reduction in c.p.d. was observed was no. 23. It is perhaps a coincidence that this was the first film deposited after the system had accidentally been let up to atmospheric pressure by air ingress at a leak, rather than by nitrogen. This may have had a destabilising effect on the tungsten reference surface, though the system was baked and otherwise run as normal. The residual gas analyser was, unfortunately, not as sensitive as usual, and the results were somewhat inconclusive, but a slightly higher proportion than usual of oxygen was detected.

6.3.3 Resistivity and Electron Diffraction Measurements on Tb Films

The results of the resistivity measurements and calculations on the Tb films are shown in Table 6.7.

Table 6.7 Resistivity of Tb Films

Film I.D. No.	Resistivity ($\mu\Omega$cm)
19	71 ($\pm$16)
20	-
21	-
22	116 ($\pm$17)
23	136 ($\pm$12)
24	111 ($\pm$13)

The values for films 22 to 24 compared well with the bulk values for Tb in Table 6.3, allowing for a little residual resistivity. The value for film no. 19 was very low, similar to that for the Gd film no. 33. As explained above, there was a possible common factor of contamination of

the film due to inadequate outgassing (low current in the case of no. 19 and low vacuum in the case of no. 33). The ratios of the measured resistivity to the bulk resistivities were 0.58 and 0.7 for the a- and c-axes respectively. Although these are lower than Reale's figure, they are compatible with Bist and Srivastava's figures for Gd and La.

The electron diffraction observations of the Tb films were also conducted in reflection. As with the Gd films some patterns were obtained which consisted of broad diffuse semi-circles, indicating small-grained polycrystallinity. No clear spot patterns were found. The poor quality of the patterns and the high background intensity hampered interpretation, but from the regularity of the ring-spacings the structures were more likely to have been h.c.p. than any other.

6.4 Measurements on Dysprosium Film in U.H.V.

The measurements on Dy were also made at ambient temperature. Only one film was deposited, so specimen reproducibility was untested. The effect of aging in U.H.V. has been included with the Tb results in Figure 6.3 (film no. 25). The c.p.d. value at first measurement was -1.465 ($\pm$0.016)eV and after aging -1.545 ($\pm$0.016)eV, a change of -0.080 ($\pm$0.022)eV, and therefore in the same direction as the changes in both the Gd and Tb films. The resistivity of the film was 105 ($\pm$13) $\mu\Omega$cm which falls between the a-axis and c-axis values given in Table 6.3. Attempts at obtaining reflection electron diffraction information on this film were unsuccessful.

6.5 Measurements in O.H.V.

After all the measurements in U.H.V. had been completed, the reference surface was set up in an O.H.V. plant and the c.p.d.'s re-measured. All the substrates had therefore been exposed to atmospheric contamination, albeit mostly in dessicators, for varying amounts of time. The earliest film, no. 14, had been so exposed for 204 days and the latest, no. 35, for 15 days. Measurements were nevertheless

Table 6.8 Comparison of c.p.d Results in U.H.V. and O.H.V.

Film type and I.D.no.		last value in U.H.V. (eV)	c.p.d. value in O.H.V. (eV)	change (eV)
Gd	14	-1.48 (±.011)	-1.33 (±.010)	+0.15 (±.015)
	16	-0.998 (±.021)	-1.42 (±.006)	-0.422 (±.022)
	18	-0.670 (±.004)	-0.847 (±.009)	-0.177 (±.011)
	27	-0.650 (±.011)	-0.977 (±.009)	-0.327 (±.014)
	28	-0.432 (±.011)	-1.205 (±.006)	-0.773 (±.013)
	30	+0.537 (±.011)	-1.015 (±.006)	-1.552 (±.013)
	33	-0.230 (±.011)	-0.930 (±.006)	-0.700 (±.013)
	34	-1.305 (±.011)	-1.302 (±.009)	+0.003 (±.014)
Tb	19	-1.115 (±.011)	-1.627 (±.009)	-0.512 (±.014)
	20	-1.078 (±.021)	-1.572 (±.009)	-0.494 (±.023)
	21	-1.382 (±.016)	-1.330 (±.011)	+0.052 (±.019)
	22	-1.600 (±.021)	-1.282 (±.014)	+0.318 (±.025)
	23	-1.430 (±.011)	-1.347 (±.009)	+0.083 (±.014)
	24	-1.845 (±.011)	-1.280 (±.011)	+0.565 (±.015)
Dy	25	-1.545 (±.016)	-1.482 (±.009)	+0.063 (±.018)
W	35	-0.003 (±.0015)	+0.010 (±.005)	+0.013 (±.006)

conducted under O.H.V. conditions to eliminate the effects of gas in the capacitor space (see paragraph 3.2.7). The results are shown in Table 6.8 with the last U.H.V. values for comparison. The changes in c.p.d. were seen to vary considerably, indicating that they were not solely due to a change in the work function of the reference surface.

The resistances of some of the films were also remeasured under O.H.V. and at atmospheric pressure. The results are shown in Table 6.9. Note that errors are expressed as a percentage, not in ohms.

Table 6.9 Resistances of Rare Earth Films

Film Type and No.		Resistance (Ω)			Change in Resistance (%)	
		Last U.H.V.	O.H.V.	760 torr	U.H.V. to O.H.V.	U.H.V. to 760 torr
Gd	18	29 (±3%)	38 (±3%)	-	+31 (±4)	-
	27	48 (±2%)	52 (±2%)	52 (±2%)	+8.5 (±3)	+8.5 (±3)
	28	31.25 (±3%)	32 (±3%)	32 (±3%)	+2.4 (±4)	+2.4 (±4)
	29	29.5 (±3%)	27.5 (±3%)	27.25 (±3%)	-6.8 (±4)	-7.6 (±4)
	30	-	27.5 (±3%)	27.25 (±3%)	-	-
	33	19 (±5%)	18.25 (±5%)	19 (±5%)	-4 (±7)	0 (±7)
	34	44 (±2.5%)	41.75 (±2.5%)	42 (±2.5%)	-5.1 (±3.5)	-4.5 (±3.5)
Tb	19	6 (±16%)	9 (±10%)	-	+50 (±19)	-
	22	21 (±5%)	20 (±5%)	-	-4.75 (±7)	-
	23	25 (±4%)	20.5 (±5%)	-	-18 (±6.5)	-
	24	18.5 (±5%)	17.5 (±5%)	18 (±5%)	-5.4 (±7)	-2.7 (±7)
Dy	25	31 (±3%)	32.5 (±3%)	-	+4.8 (±4)	-

In view of the extreme reactivity of the rare earths, the small extent of these changes was surprising. Some of the films exhibited a decrease in resistance after atmospheric exposure, which indicated the

presence of effects other than conversion of the surface(s) from metallic to dielectric properties.

6.6 Discussion.

In order to study the effects of various parameters or conditions on the work functions or resistivities of the specimens, it was convenient to arrange the identification numbers of the films in sequences for each material, each sequence starting with the film which had the lowest value of any given parameter and ending with that having the highest (or, where both positive and negative values existed, starting with most negative and ending with most positive, or vice-versa). In this way it was possible to see whether any correlation existed between a property of the material and an experimental variable. In an attempt to account for the large spreads in values of the c.p.d.'s and resistivities, the following parameters have been compared:

c.p.d. at first measurement,
pressure x time during deposition,
age at first measurement,
film thickness,
c.p.d., aged,
age at the above,
change in c.p.d. due to aging,
change in c.p.d. due to annealing,
annealing temperature,
annealing time,
annealing temperature x time,
c.p.d. after atmospheric exposure,
change in c.p.d. due to atmospheric exposure,
resistivity in U.H.V.,
resistivity after atmospheric exposure,
change in resistivity due to atmospheric exposure.

As has been stated earlier, it was expected that the tungsten

reference surface would be repeatable from run to run, since it underwent nominally identical procedures each time. Also, within each run the tungsten was exposed to the residual atmosphere for a long time prior to the deposition of the rare earth film, so it might be expected to be as much contaminated as possible within the U.H.V. conditions. To check that the variations in c.p.d. were not due to changes in the work function of the tungsten, it is appropriate to examine the dependence of the c.p.d. on the amount of contamination generated during the deposition process, i.e. pressure x time (P x t). For the Gd films the c.p.d. had a predominantly inverse dependence on (P x t) whereas for the Tb films a direct relationship was observed. Hence the spread of results was not due to contamination of the tungsten surface either before or during the deposition of the films. It has also already been pointed out that the c.p.d.'s for the two Gd films deposited and observed at the higher pressures were seen to decrease relatively rapidly, whereas for the Gd, Tb and Dy films deposited and measured under the best U.H.V. conditions the c.p.d.'s increased slowly. Post-deposition contamination was not the dominant factor in the aging behaviour. Since the c.p.d.'s were negative, the work functions of the films were lower than that of the tungsten, therefore the work functions of the films were decreasing with age, and the older the film the greater the effect.

The effects of annealing on c.p.d. were inversely in sequence with the annealing temperatures (as recorded by the thermocouple). The range of temperatures was small, however, so this may have been a spurious correlation. The work functions of the films increased on annealing, the magnitude of the increase being smaller for the lower initial work function values. The perturbations in the sequences (work function, annealing temperature and annealing effect) were films no.'s 27 and 33. The latter was deposited at high pressure and the former was heated three times and was significantly older than the other films at the time of these measurements. On the first and third heating cycles the work function was increased, but over the second it decreased slightly, similar to unheated aging. This may have been because the

highest temperature in the second cycle was lower than that in the first, and it is reasonable to assume that no further crystal reorganisation occurred. This corroborates the interpretation that the changes observed during heating were largely due to annealing and not to external causes such as gas contamination or change in the work function of the reference surface. There was no correlation between annealing time and its effect. It is interesting to note that the annealing changes were in the opposite direction to those attributed to aging (except for film no. 33 and the 2nd cycle of no. 27). The effect of annealing on resistivities was in direct sequence with the annealing temperature, though the same limitation as to range of temperatures applies. Increases of work function on annealing have been observed by other workers in a wide range of metals[55, 56], including Gd[57]. The resistivities before and after atmospheric contamination were in identical sequence order except for one Gd film (no. 18) and one Tb (no. 19). However, some were seen to increase and some to decrease. The change in resistance of metal films on exposure to gas contamination has been considered to have three possible causes[58]:

(a) decrease in effective thickness of the film, due to the surface losing its metallic properties,

(b) decrease in mobility of conduction electrons, due to an increase of diffuse scattering at the top surface or at grain boundaries,

(c) decrease in concentration of conduction electrons throughout the film, due to sharing with the adsorbed gas atoms.

All three mechanisms would produce an increase in resistance and therefore cannot by themselves explain the present results.

The effects of gas contamination on the work function of simple metals have been considered for both physical adsorption[55] and chemical adsorption[59]. Experimental results from group VIII metals and for gold[60] showed that the change in work function due to

physical adsorption had a linear relationship with the work function of the clean surface, being greater for lower initial value. The present results on Gd and Tb also showed linear trends on exposure to the atmosphere, see Figure 6.4. The difference in the polarity of c.p.d.'s can be explained by assuming that the tungsten reference surface also suffered a change in work function on exposure to the atmosphere, which was constant, i.e. the origin may be moved by that amount. A further similarity existed between the two sets of results, that is the change in slope of c.p.d. with temperature, In the simple metals[55] this change occurred at between 370 and 410°K and in the case of Gd at 360°K, see Figure 6.1(a). There are three possible causes. Firstly, a reaction with the glass substrate, for example an exchange of ions between the films and the glass, which was rich in the oxides of alkali metals such as Na, Ca and Mg. Secondly, desorption of one or more common gaseous contaminants. Thirdly, the onset of effects of strain imposed by differential thermal expansion either between the film and the substrate or between the substrate and its holder. Such a strain might be expected to alter the Fermi Level via the lattice volume, hence the work function. Effects of this type have been observed by other workers[61, 62]. The third effect does seem least likely, however, since it would not be expected to occur within such a narrow range of temperatures for a range of dissimilar metals. The basic postulate for the physisorption variation was anisotropy[55]. A model to explain the effects of chemisorption has also been based on site heterogeneity, or anisotropy[59]. The model is too limited to apply to this work since it treats only two types of surface site, or patch, and, by implication, only one type of adsorbed species. It also requires a knowledge of the extent of coverage of both types of site by the contaminant. However, research on single crystal specimens, at different orientations, and with individual contaminants would benefit from this approach.

6.7 Theoretical Values.

The work function of Gd has been calculated as 3.07eV by Ramey and Katzberg[63]. The value was given by[64]:

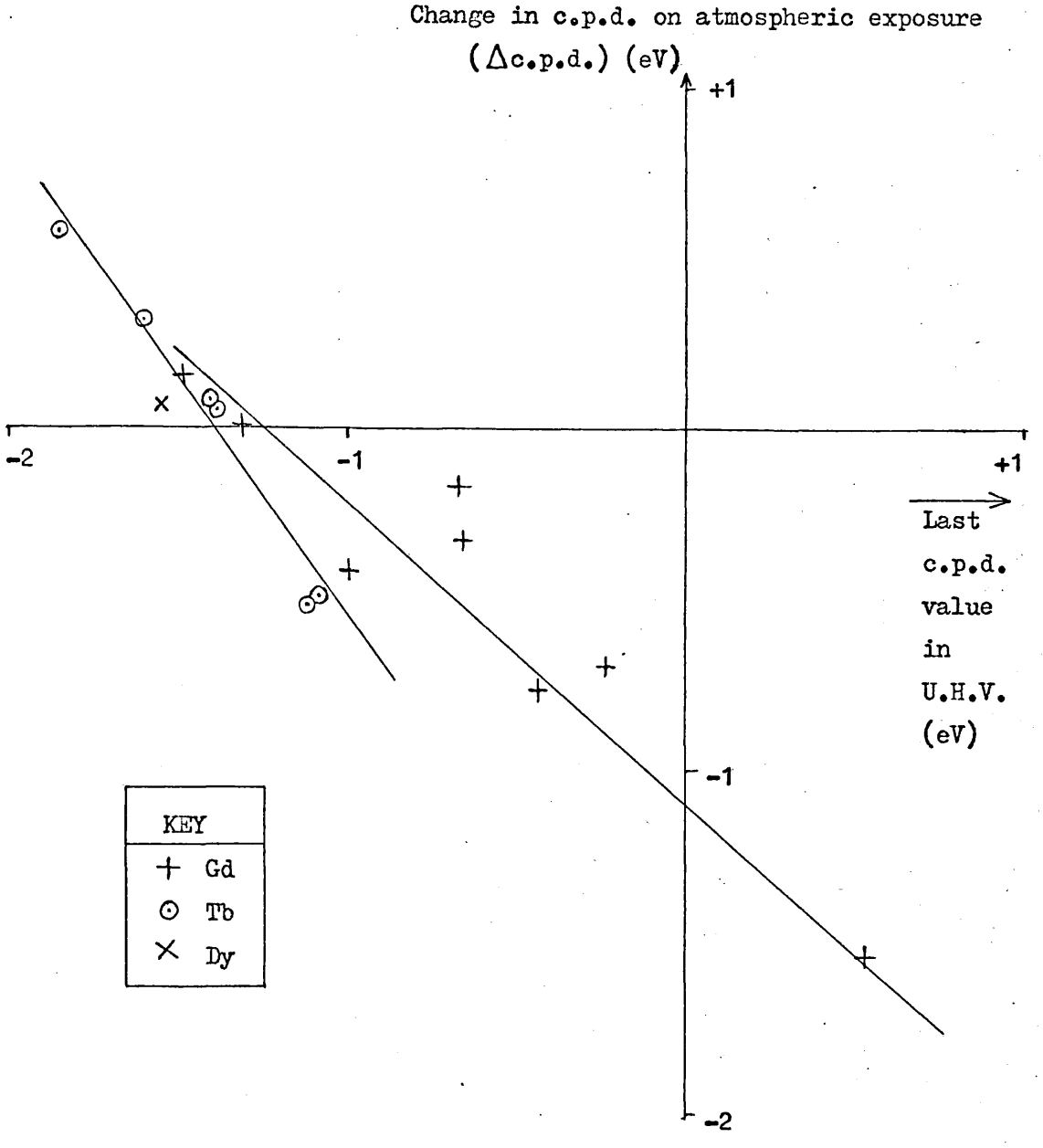


Figure 6.4 The Effect of Atmospheric Exposure on c.p.d.

$$\phi = \frac{(x + 0.15)}{0.44} \qquad \text{equation (6.1)}$$

where $$x = 0.31\left(\frac{n+1}{r}\right) + 0.50$$

n is the no. of electrons in the valence shell, taken to be 3

and r is the radius of the atom, taken to be 1.623Å[65].

Using more recent values for the atomic radii of Gd, Tb and Dy[66], i.e.1.801, 1.783 and 1.775Å respectively, the work functions are 3.042, 3.058 and 3.065eV respectively. These results must be treated with caution, however, since they assume simple atomic and lattice structures, which are not strictly valid for the rare earths, and the coefficients in the equations have been derived empirically from measurements of the work functions of other rare earths (La, Ce, Pr, Nd, Pm and Sm) by other workers, using thermionic emission[67] and photoelectric[68] techniques.

An alternative procedure is to calculate the work function from the surface dipole barrier, the Fermi Level and the chemical potential. This type of calculation has been reported by Sahni, Krieger and Gruenebaum[69] (SKG) and by Lang and Kohn[70] (LK) for a "jellium" model, in which only the valence electrons were considered to contribute to the electron density and the ion cores (the positive nuclei and the core electrons) were replaced by a positive background charge of density equal to that of the electron gas. The surface region of the metal is shown schematically in Figure 6.5. The hatched area represents the uniform positive-charge background, which has charge density $n_+ = \frac{k_F^3}{3\pi^2}$

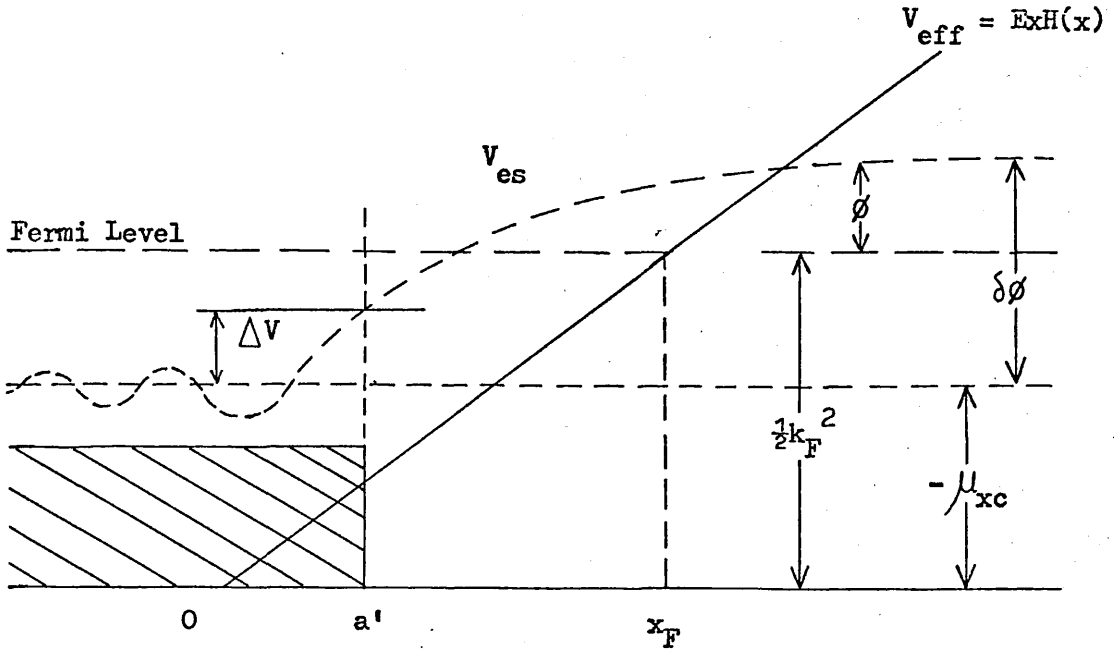


Figure 6.5 Schematic representation of the metal surface

The effective potential, V_{eff}, at the metal surface is given by

$$V_{eff} = ExH(x) \qquad \text{equation (6.3)}$$

where E = Field Strength,

x = Linear dimension perpendicular to the surface,

H(x) = step function = 0, $x \leqslant 0$

= 1, $x > 0$

V_{es} is the electrostatic potential,

ΔV is the difference in the electrostatic potential between the surface and the bulk of the metal,

$\delta\phi$ is the surface dipole barrier,

μ_{xc} is the exchange-correlation contribution to the chemical potential,

$\frac{1}{2}k_F^2$ is the Fermi Energy

and a' is the position of the metal surface, fixed by the requirement of overall charge neutrality.

Because of the shape of the potential, this scheme is called the Linear Potential model.

The work function was calculated from

$$\phi = \delta\phi - \tfrac{1}{2}k_F^2 - \mu_{xc} \qquad \text{equation (6.4)}$$

$$\delta\phi = 4\pi \int_{-\infty}^{+\infty} x n_t(x)\,dx \qquad \text{equation (6.5)}$$

where n_t = total charge density

and $$\mu_{xc} = \frac{d(n_e \varepsilon_{xc})}{dn_e} \qquad \text{equation (6.6)}$$

where n_e = electron density

and ε_{xc} = exchange-correlation energy per particle for the uniform electron gas.

The latter is related to the Wigner-Seitz radius, r_s, which is the radius of a sphere of volume equal to the primitive lattice cell, and is given by[70]

$$\varepsilon_{xc} = \frac{-0.458}{r_s} - \frac{0.44}{(r_s + 7.8)} \qquad \text{equation (6.7)}$$

The results of these calculations for a range of values of r_s (in atomic units. The atomic unit of distance is equal to 0.529Å.) are given in Table 6.10. Alternative formulae for the correlation function give slightly lower results[69], but with similar range.

The Wigner-Seitz radius is determined from the mean electron density $\bar{n}_e$ by

$$\frac{4\pi}{3} r_s^3 = \frac{1}{\bar{n}_e} \qquad \text{equation (6.8)}$$

Table 6.10 Work Functions Calculated from Jellium Linear Potential Model

r_s (atomic units)	ø (eV)	
	SKG	LK
2.0	4.12	3.89
2.5	3.68	3.72
3.0	3.31	3.50
3.5	3.01	3.26
4.0	2.74	3.06
4.5	2.59	2.87
5.0	2.63	2.73

For bulk Gd the Wigner-Seitz radius is 3.78 atomic units[71] (a.u.) which, by interpolation from Table 6.10, indicates a work function of 2.85eV by SKG and of 3.15eV by LK. The density of f.c.c. Gd is lower than that of h.c.p. Gd by 6.8%[52]. The Wigner-Seitz radius is therefore 2.3% higher at 3.87 a.u., indicating a work function for the f.c.c. phase of approximately 2.80eV by SKG and 3.11eV by LK.

The corresponding radii for bulk h.c.p. Tb and Dy may be estimated from their atomic numbers and atomic volumes, compared to those of Gd, using equation (6.8) and the formula for electron density:

$$\overline{n_e} \propto \frac{Z}{\Omega} \qquad \text{equation (6.9)}$$

where Z = Atomic no. (no. of electrons/atom),

Ω = Atomic volume (cm^3/mole).

The results are shown in Table 6.11. The indicated work functions for Tb and Dy are 2.88eV and 2.96eV respectively by SKG and 3.17eV and 3.19eV respectively by LK.

Table 6.11 Some Atomic Parameters of Gd, Tb and Dy

Element	Z	$\Omega^{(66)}$	r_s
Gd	64	19.91	3.78
Tb	65	19.30	3.72
Dy	66	19.03	3.68

The above calculations have been based on a one-dimensional model. A first-order approximation to the anisotropy of the work function may be made by calculating the anisotropy of the electron density. For each crystallographic orientation, the anisotropy parallel to the surface will be ignored so that one-dimensional calculations may be performed, normal to the surface. The anisotropy of electron densities in different crystal directions may be estimated from the band structure[72]. In each region of the Brillouin Zone (B.Z.), there are an integral number of bands below the Fermi Level. Each point in this region (i.e. each set of momenta values allowed by the periodic boundary conditions) can accomodate 2 electrons per band, one of each spin. Since the number of points per unit area is fixed, equal to $^1/N$ where N is the number of unit cells in the repeating segment of the region, then the number of electrons accomodated in each of the regions of the B.Z. is given by $^2/N$ x (area of the region) x (number of bands below the Fermi Level in the region). The number of bands can be determined directly from the band structure diagrams[71], which were determined[73] by relativistic augmented plane wave calculations based on the atomic configurations Gd ($4f^7\ 5d^1\ 6s^2$) and Dy ($4f^{10}\ 6s^2$). The area of each region in k-space for which each band lay below the Fermi energy was determined from diagrams of intersections of

the Fermi surface with the symmetry planes of the B.Z.[73]. The lower half of the primitive B.Z. is reproduced in Figure 6.6, with lattice vectors $\underline{b}_1$, $\underline{b}_2$ and $\underline{b}_3$. The points of high symmetry are labelled and show the part of the B.Z. on which the band calculations were based and which was repeated 24 times to produce the whole B.Z.

The intersections with the Fermi Surface are shown in Figure 6.7 (a) and (b) for Gd and Dy respectively. The number of bands below the Fermi Level in each position are shown, and the results of manual integration of (area x number of bands) for each symmetry plane are tabulated.

The calculations strictly represent the density of states, but since ambient temperatures were being considered, it has been assumed that they approximate to the actual electron densities in the various planes of k-space.

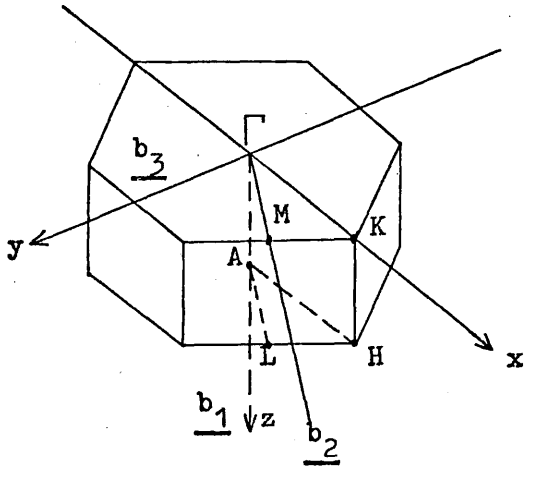
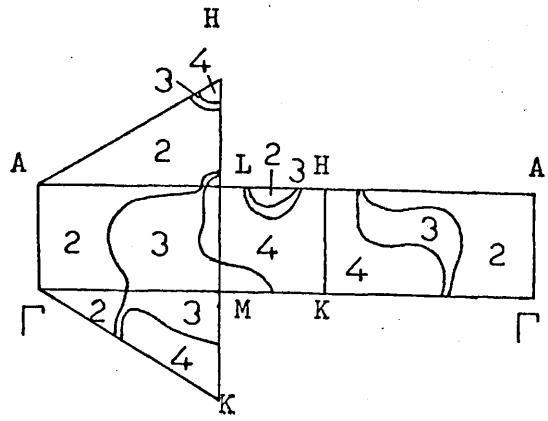


Figure 6.6 The Lower Half of the Primitive Brillouin Zone for h.c.p. Crystal

Transforming to real space, the planes in the B.Z correspond to directions in the crystal lattice.

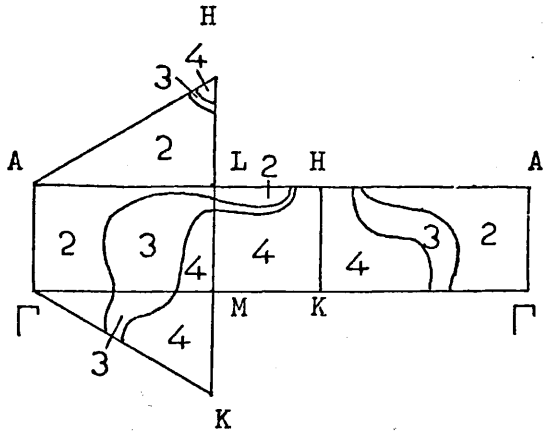
The calculations have been performed for a limited number of planes of intersection, so the anisotropies may be greater than shown in the tables in Figure 6.7. The variations from the median values are $\pm$50% for both Gd and Dy, and by interpolation[73] we may assume a similar anisotropy for Tb.

The anisotropy of the work function may be calculated from the anisotropy of the electron density by the relation between r_s and $\bar{n}_e$ in equation(6.8).Although the Wigner-Seitz radius does not, by definition,take account of anisotropy, the values of r_s in Table 6.10 may be replaced by values of $\bar{n}_e$, and the unidimensionality of the model allows the replacement of $\bar{n}_e$ (the average bulk value) by the unidirectional electron density. Assuming that the average value of electron density corresponds to the average bulk value $\bar{n}_e$, and consequently to an r_s value for Gd of 3.78 a.u., then the extreme values of electron density correspond to r_s values of 3.30 a.u. and 4.76 a.u. These give work function values of 3.15eV and 2.60eV respectively by SKG and 3.35eV and 2.80eV respectively by LK. This is an anisotropy range of almost 20% in both cases. Similar ranges are obtained for Tb and Dy. However, as mentioned above, larger ranges may exist, since these calculations included only a limited selection of crystal directions. Closer agreement to the experimental values is obtained with the LK calculations. However, since the films were polycrystalline, it cannot be assumed that the full range of anisotropy would be apparent, since the method measured the area-weighted average.In the case of the c-axis being predominantly perpendicular to the surface, the measured work function would be expected to be towards the higher value, since the plane HAΓK which transforms to the c-axis has the highest value of (Band density x area)(see Figure 6.7) hence of electron density, n_e. Lang and Kohn[70] have shown that the crystal face with highest work function is that with lowest surface energy, and which therefore would be expected to be preferentially present at the surface of a vacuum-deposited film. After the work described above had been completed on each specimen, optical measurements were made on most of the films from



(a) Gd

Plane	Electron band density x area (arbitrary units)	
A H L	369	
A L M Γ	1006	Median
Γ M K	543	723
L H K M	726	spread
H A Γ K	1077	± 354



(b) Dy

Plane	Electron band density x area (arbitrary units)	
A H L	354	
A L M Γ	865	Median
Γ M K	580	721
L H K M	715	spread
H A Γ K	1088	± 367

Figure 6.7 Intersections of Fermi Surface and Brillouin Zone (a) Gd (b) Dy

which the optical conductivities were calculated[74]. Since this parameter is also dependent on the electron density, the results of the work function and optical conductivity experiments were compared, but no correlation was found. The range of results was not inconsiderable, as shown in Table 6.12.

The dependence of the work functionsof nickel and of gadolinium on magnetic phase has been observed experimentally and considered theoretically by several authors[75 to 83]. Some experiments observed a small change in the work function[76, 83] or in its temperature dependence[77] at the Curie point, whereas others observed no change[75, 78, 79]. Similarly some theoretical papers have predicted such a change[80, 81], while others have pointed out inconsistencies in the theory[82] or a modification to the theory[75] to allow agreement with either case. Since the available experimental and theoretical values for the behaviour of the relatively simple nickel are so contradictory, it would be premature to draw a firm conclusion as to whether the present experimental results confirm one view or another. As shown in Figure 6.1, neither the work function nor its temperature dependence were seen to change at the Curie point. It can also be inferred that the value of the temperature coefficient was lower than the resolution of the measurements, i.e. less than 0.5meV/°K. (The temperature coefficient of the work function of the reference surface was probably less than 6μeV/°K[32] and would not affect the c.p.d. significantly.)

It should be recognized that the above models for the work function, its anisotropy and temperature dependence are extremely simple. Even those parts which specifically relate to the rare earths, i.e. the band structures and intersection diagrams, are far from being completely accurate, since they are based on free-atom models of the elements. They therefore do not completely describe either the bulk lattice or the surface.

6.8 Summary of Results on Rare Earth Films.

The principal qualitative and quantitative results are re-stated

here in the interests of clarity, and the values of other workers (where available) are added for comparison.

No change was observed in either the work function of Gd or its temperature-derivative in the region of the Curie point. The temperature coefficient was less than 0.5meV/^{o}K in that region.

The work functions of Gd, Tb and Dy films decreased during aging in U.H.V. Residual atmospheric contamination was not thought to be the dominant cause, since films aged in higher pressures exhibited dissimilar behaviour. No changes in resistivity were observed during this period.

The effects of annealing on Gd were to increase the work function and decrease the resistivity. The change in the work function was less, the lower the initial value. Changes of a similar magnitude (33%) have been observed on annealing Gd deposited onto tungsten[57], although at a higher annealing temperature (1100^{o}C) and complicated by the difference in patch-averaging due to the field-emission method employed. This result is consistent with a reduction in surface energy caused by the annealing.

On exposure to atmosphere (760 torr of nitrogen then air), the changes in c.p.d. between the films and the tungsten reference showed a linear trend in dependence on the initial value. Since the films deposited and aged at relatively high pressure (numbers 33 and 34) were seen to increase in work function, it seems reasonable to suppose that the U.H.V.-deposited films also underwent an increase on exposure to the atmosphere, and therefore to make sense of Figure 6.4 the tungsten must also have increased in value. Hence the lower initial value of the rare earth work function the larger the increase on atmospheric exposure. This is a similar result to those obtained for Xe adsorption on Au, Pt, Pd, Ir, Rh, Ni and Fe[55]. The changes in resistance due to atmospheric exposure were less well behaved, being positive in some cases and negative in others. Contamination was obviously not the sole cause; this may explain the deviation from linearity in

the work function effect (Figure 6.4).

Table 6.12 summarises the numerical results from the present and other published work, both experimental and theoretical. The c.p.d. results have been converted to work function values by adding 4.542eV, the average value of the tungsten reference (see paragraph 3.3). The two films (no.'s 33 and 34) deposited and aged under lower vacua have been omitted from this table.

The c.p.d. measurements of Ramey and Katzberg[63] employed an oscillating dielectric rather than oscillating the electrodes, and may have been subject to errors due to charge accumulation[28]. The measurements were described as being made "at room pressure". If this is to be interpreted strictly then further error would accrue due to further charge collection on collision with the gas and to surface contamination. The sample was described as being polycrystalline with a [0001] texture perpendicular to the surface, which was mechanically polished from the bulk using powders, ethanol and water. The same "cleaning method" was used for the reference surfaces, Ag and Nb.

The photoemission studies of Blodgett, Spicer and Yu[79] estimated ϕ from the Fowler plot of $(\text{quantum yield})^{\frac{1}{2}}$ vs. photon energy. This is not the most accurate method of estimating work function from photoemission data, and in any case has a complex dependence on the surface patch distribution (see paragraph 2.3). Measurements above and below the Curie temperature indicated that shifts in Gd band structures from one state to the other were less than 0.1eV.

Similar comments apply to the results of Eastman[84], with the additional complication that the films were deposited onto quartz which had previously been coated with chromium. The crystallites of the Gd were between 100 and 200Å in size and randomly orientated. Busch et al[85] observed photoemission from a Gd film deposited in U.H.V. at liquid helium temperatures at photon energies as low as ~2.8eV.

The Gd samples used in the field-emission experiments of Eckstein et al[57] were measured before and after annealing. A considerable

Table 6.12 Summary of Rare Earth Data

(a) Work Function (eV)

Experimental	Gd		Tb		Dy
	Median	Range	Median	Range	
Fresh*	3.545	0.74	3.151	0.67	3.077
Aged (U.H.V.)*	3.467	0.81	3.081	0.77	2.997
Aged + Annealed*	4.501	1.22	-	-	-
After atmospheric exposure* (unannealed)	3.409+	0.57	3.089+	0.35	3.060+
After atmospheric exposure* (annealed)	3.451+	0.23	-	-	-
Kelvin method	2.9 (±0.006)(63)				
Photoemission	3.1 (±0.1)(79)				
Photoemission	3.1 (±0.15)(84)				
Photoemission	~2.8(85)				
Field Emission Fresh	2.4 (±0.3)(57)				
Field Emission Annealed	3.2 (±0.3)				

Theoretical		Mean	Range	Mean	Range	Mean	Range
Gordy-Thomas		3.042	-	3.058	-	3.065	-
Linear Potential Model (LK)*	h.c.p.	3.15	≠ 0.55	3.17	≠ 0.55	3.19	≠ 0.55
	f.c.c.	3.11	-	-	-	-	-

* Denotes Present Work

\+ These values are in error by the extent to which the W reference changed due to atmospheric exposure.

Table 6.12 (Continued)

(b) Resistivity ($\mu\Omega$cm)

Experimental	Gd		Tb		Dy
	Median	Range	Median	Range	
Fresh*	160	49	104	65	105
Aged (U.H.V.)*	160	49	104	65	105
Aged + Annealed*	167	23	-	-	-
After atmospheric exposure* (unannealed)	210	-	109	6	110
After atmospheric exposure* (annealed)	160	66	-	-	-
Bulk, a-axis	141 [48]		122 [48]		111 [48]
Bulk, c-axis	121		102		78
Film (h.c.p.)	~ 200 [51]				
Film (f.c.c.)	~ 73 [51] ~ 106 [52]		~ 89 [52]		~ 74 [52]

(c) Optical Conductivity ($sec^{-1} \times 10^{-14}$) [74]**

Experimental	Gd		Tb		Dy
	Median	Range	Median	Range	
Aged (U.H.V.)*	[27 17.5]	[16 15]	[19.5 8]	[9 6]	[24 13]
Aged + Annealed *	[29 16]	-	-	-	-

* Denotes Present Work

** The two values quoted are those measured at photon energies 1.8eV and 3.0eV respectively

increase was observed, even though the Gd was less than randomly orientated initially due to epitaxy on the fine W tip. The quoted accuracy of the absolute values was poor, probably because the technique required the comparison of the emission curves from the naked and Gd-coated W tip, which may have been slightly dissimilar in shape. Since epitaxy occurred, and since the technique does not measure the area-weighted average work function, some discrepancy in values was to be expected between these and c.p.d. results. The median values for the present Gd films were higher than those of other workers, by approximately 0.4eV for fresh U.H.V.-deposited surfaces. It is to be expected that photoemission results, for example, would be lower than the c.p.d.-derived values, due to the former's bias towards patches of low work function as well as to the Schottky effect. The lowest unannealed film work function values from the present work are in very good agreement with most of the previously published results on clean Gd.

Chapter VII The Work Functions of Silicon, Gold, Aluminium and Tungsten, and their Use in Far Infra-red Diode Detectors

7.1 Introduction.

Measurements have been made of the work functions of materials other than rare earths, whose use in the construction of point-contact diodes for the detection of far infra-red radiation (at 337μm wavelength) has been studied both experimentally and theoretically[86]. These studies are described in this Chapter, and the relevance of work functions to the operation of the devices is shown. A model is presented to explain the performance of the diodes, which is based on the calculations of both tunnel and thermionic current flow between a hemispherical whisker tip and a plane electrode.

Except for the preparation of the metal plane electrodes and the work function measurements, the experimental work described in this Chapter was performed by another postgraduate student, Martin W. Booton. The development of the theoretical model resulted from continual joint discussion, starting with the tunneling theory mentioned briefly in paragraph 2.6.8.

7.2 The Use of Whisker-antenna Diodes as Far Infra-red Detectors.

Point-contact diodes have been used to detect far infra-red radiation for a number of years[87, 88]. The diodes are formed by bringing into contact a finely-pointed whisker and the surface of a metal or semiconductor. The whisker acts as an antenna at the submillimetre wavelength[89], and the diode, consisting of the junction, rectifies the received signal. Due to the compactness of the junction, the time-constant of such devices is very small, giving them a rapid response time. By comparison, the photoconductive detectors, such as bulk InSb[90] are almost as fast but require cryogenic temperatures, while thermal detectors such as the Golay cell are very slow, having a response time of the order of milliseconds and a recovery time of over a second. The disadvantages of point-contact diodes are lower responsi-

vity than the more conventional detectors[87, 88, 89, 91, 92], and frailty and susceptibility to mechanical vibration.

During the earlier investigations[91, 92, 93] of point-contact diode performance, several interesting phenomena had been noticed. Firstly, two polarity reversal features were observed[93], one with mechanical adjustment and one with incident power. These were interpreted respectively as being due to variations of the difference in work function between the whisker and the plane surface, and to a voltage-dependent reversal in the rectification of the diode, based on calculations only of tunnel currents. The very rapid reversal of signal with increasing incident power is significant. Secondly, the highest responsivity was obtained as the whisker was withdrawn just prior to its losing contact with the surface of the substrate[93, 94]. Thirdly, the performance was predicted and observed to be improved with reduction in contact area[91]. The experimentally observed superior responsivities of Ge over GaAs diodes[92] were attributed to smaller contact areas on the harder Ge, whereas it had been expected that the GaAs devices would prove better, especially at high frequencies, due to the latter's higher mobility. Finally, the responsivities of the GaAs diodes were optimized by applying a slight forward bias[92]. Similar characteristics have also been observed in the behaviour of point-contact diodes at near infra-red[95, 96] and microwave[97] frequencies.

7.3 Diode Construction.

The diodes were constructed by the common practice of bringing a finely-pointed wire into contact with a plane substrate of metal or semiconductor. The whiskers were formed by etching tungsten wires, of 20μm diameter, in an electrolyte of potassium hydroxide, using a few volts a.c., resulting in a fine conical tip, see Figure 7.1 (a). Neither the electrolyte strength nor the voltage were found to be critical. Base electrodes of gold and aluminium were prepared by vacuum evaporation onto glass and the tungsten sample was simply a piece of ribbon of the type used to make evaporation boats. The silicon substra-

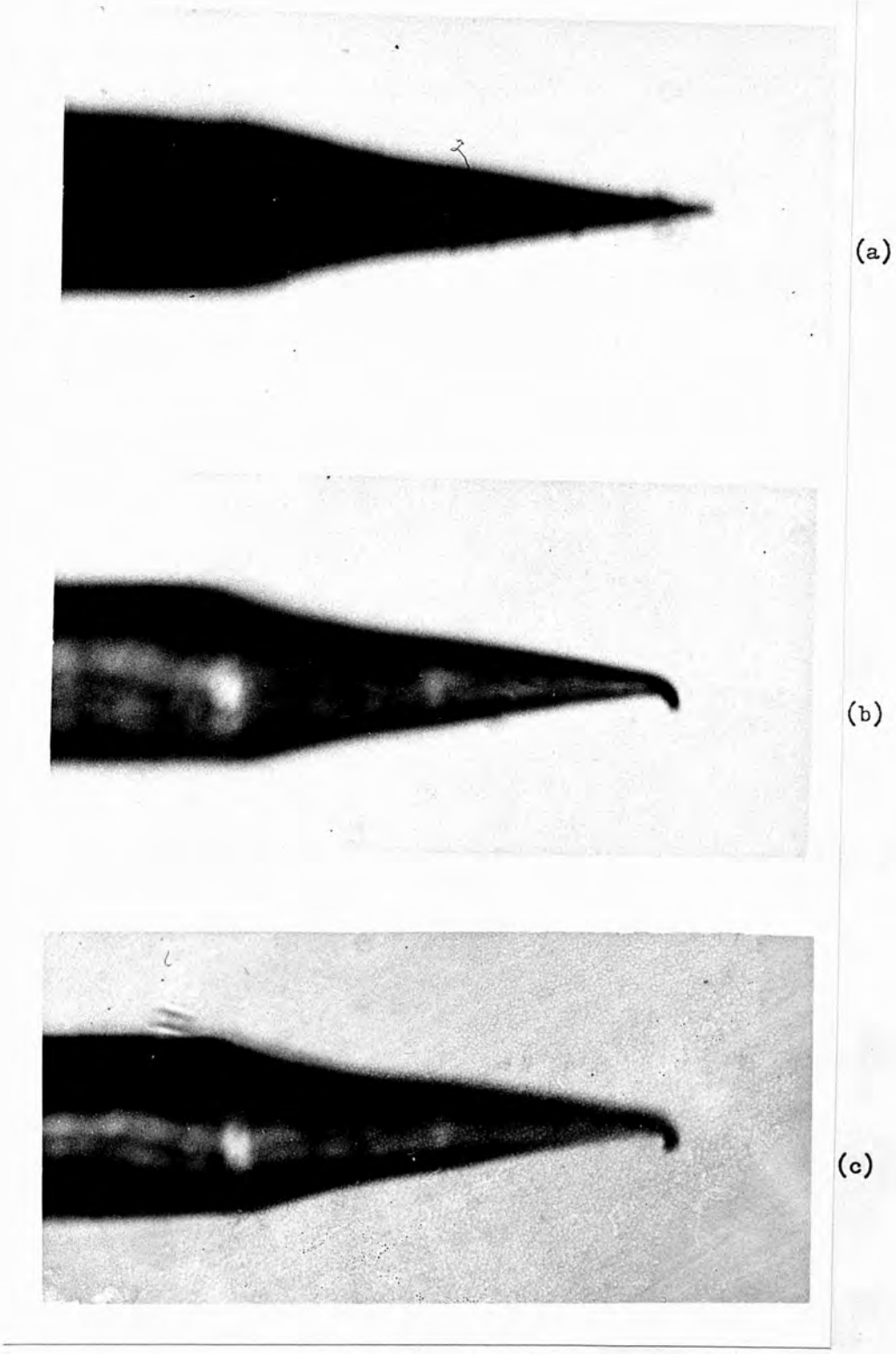


Figure 7.1 Whisker Antenna Tip at 1300X Magnification
(a) before use (b) after one contact (c) after four contacts

tes were in the usual disc form, and were n-type, doped with antimony with a resistivity between 0.005 and 0.020Ωcm. The germanium and gallium arsenide were in the form of small chips. In each case, the base electrode was held rigidly in an adjustable holder, see Figure 7.2 and electrically earthed. The antenna, held by a folded copper wire, was brought to close proximity and the final adjustments and subsequent pressure variations were effected by moving the base electrode, via a micrometer with hydraulically reduced movement. It is commonly assumed[91, 93] that the base electrodes have an insulating surface layer, caused by atmospheric contamination. The devices were therefore Metal-Insulator-Metal (M.I.M.) or Metal-Insulator-Semiconductor (M.I.S.) structures.

7.4 Experimental Investigations.

7.4.1 Diode Measurements

The experiments fell into two categories: firstly, voltage-responsivities of the diodes, including measurements of output impedance by shunting the diodes to reduce the signal by 6dB; secondly, measurements of the current-voltage characteristics, using short pulses to limit the power dissipation. The first step was mechanically to adjust each diode and obtain a response to the few milliwatts of 337μm radiation from a HCN laser, which was chopped at 11Hz for ease of observation on the oscilloscope to which the diode antenna was directly connected. For a M.I.M. diode, in particular on gold, only very light whisker pressure was required for detection to occur, because of the thinness of the insulating layer on the surface. The output signal was highest at lowest pressures, but instability sometimes occurred, due to vibration of the diode fixture. Similarly for the M.I.S. devices, the highest responsivities were observed with the lowest pressure, or greatest insulator thickness, but since the responsivities were superior to those of the M.I.M. diodes, the decrease of signal due to stabilizing pressure was less unacceptable. Further pressure changed the signal through zero to opposite polarity, generally of less amplitude than the initial signal. The reduction of

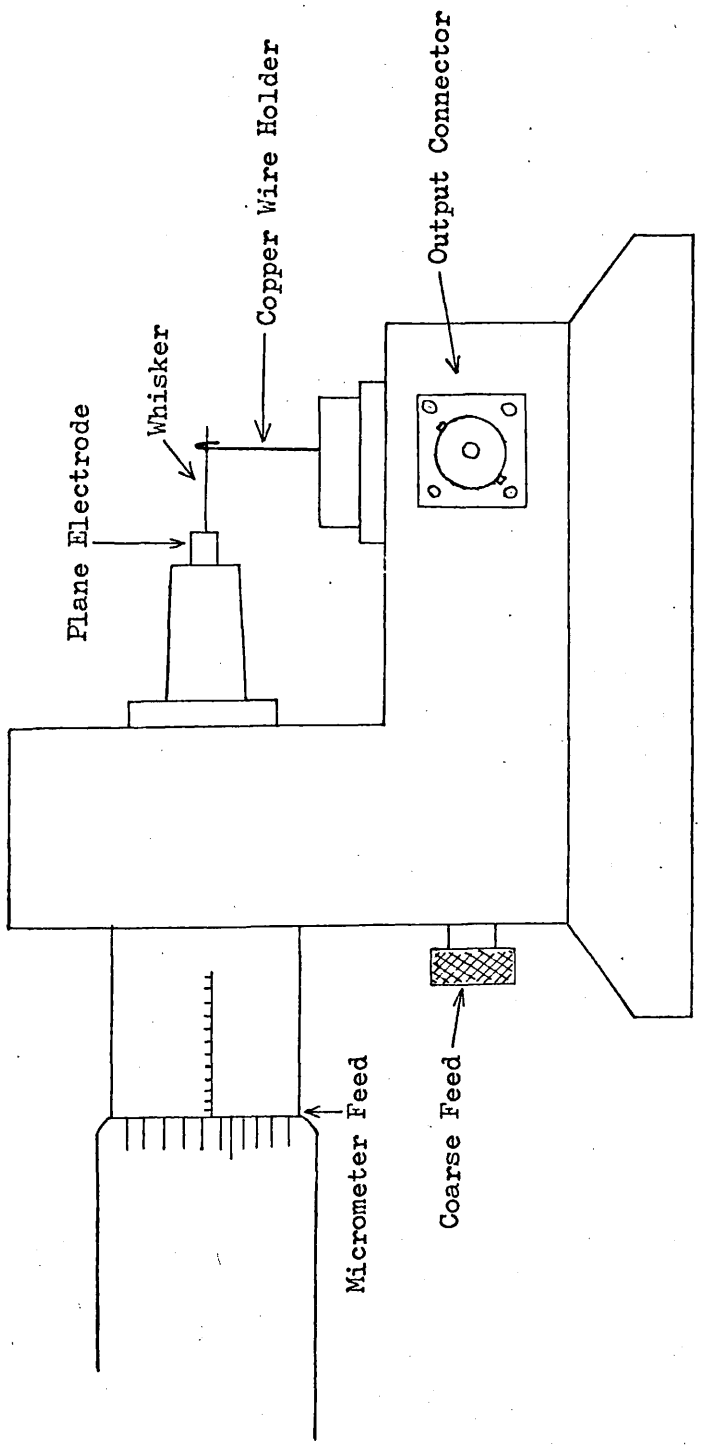


Figure 7.2 The Diode Holder

pressure at any time during this procedure produced a reproducible change in signal, indeed the whole cycle could be repeated several times with repeatable variation and polarity reversal, with only slight reductions in maximum signal due to blunting of the whisker tip. Figure 7.1 shows a whisker (a) before use, (b) after one contact and (c) after four contacts.

The original responsivity was regained on re-sharpening the whisker. The blunting effect was particularly noticeable in M.I.M. diodes, as it was usually necessary to resharpen the whisker after each application. The following measurements were made with the whisker pressure adjusted for near-optimum performance, the responsivities being calculated by dividing the d.c. voltage obtained from the diode by the c.w. power incident on the antenna, measured using a calibrated Golay cell. The responsivities were measured as a function of applied d.c. bias. No dependence was observed for the M.I.M. diodes, but the results for the Si, Ge and GaAs were as shown in Figure 7.3. The most notable was the tungsten-GaAs, which produced no output when unbiased but was by far the most reponsive with a small forward bias.

This was also the only diode of which the noise on the signal was sufficiently large to be observed on our equipment. The noise equivalent power (n.e.p.) of a detector is defined as the input signal power required to produce unity signal-to-noise ratio[92]. The amplitudes of the signal and noise voltages were measured simultaneously on an oscilloscope. The noise power was obtained by dividing the r.m.s. noise voltage by the responsivity, determined from the signal voltage. For comparison purposes, the n.e.p. is usually quoted for a 1Hz bandwidth. The measurement bandwidth, B_w, in this work was 10^6Hz. Due to the measurement being of a c.w. laser (chopped at 11Hz) the lower frequency limit was d.c. and the upper was calculated, from the time constant of the combination of diode, coaxial cable and oscilloscope input impedance, to be $\sim 10^6$Hz. The noise would be predominantly Johnson noise and shot noise, so the noise voltage was proportional to $\sqrt{B_w}$[98].

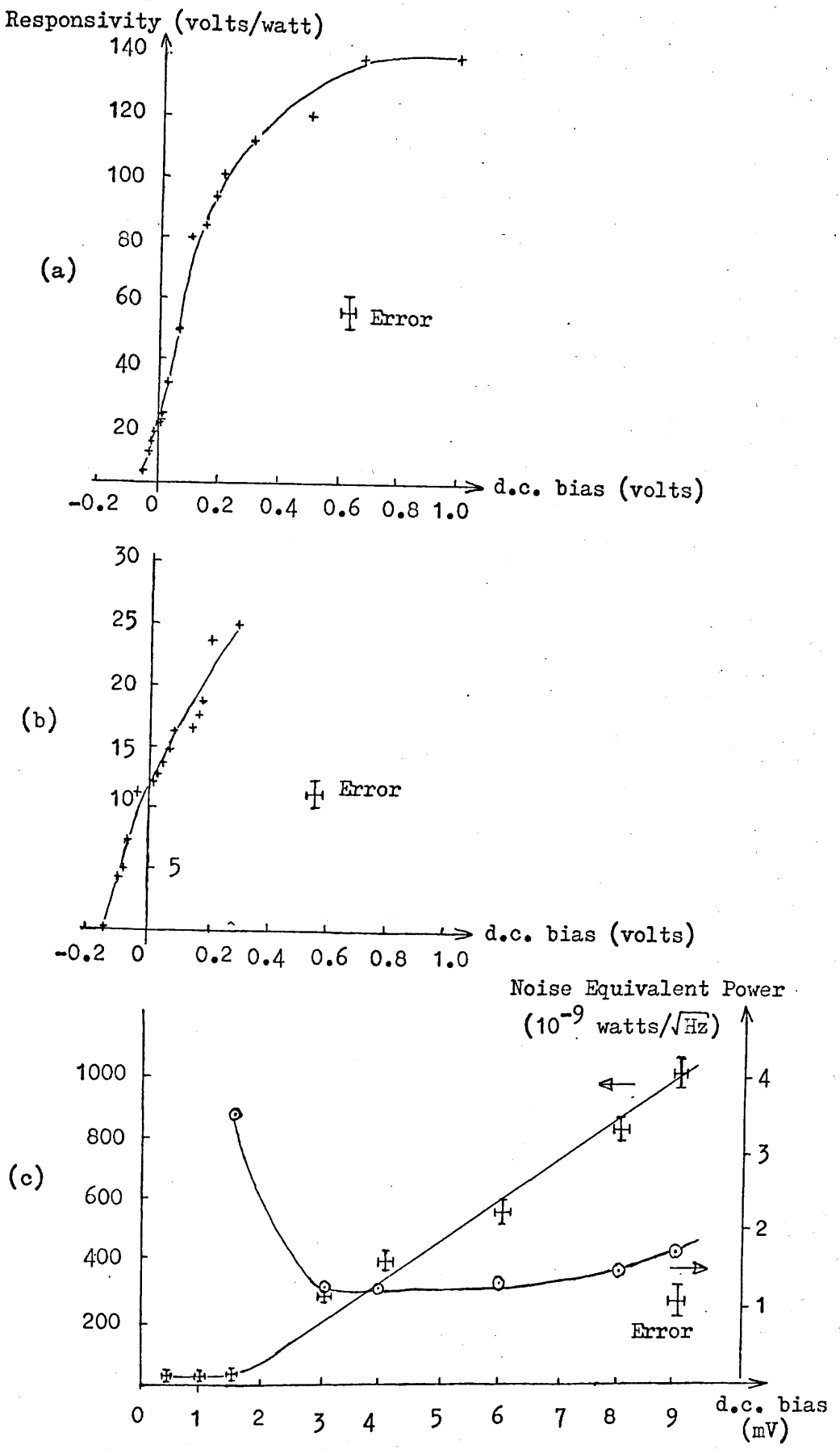


Figure 7.3 Responsivity vs. d.c. bias (a) W-Ge (b) W-Si (c) W-GaAs

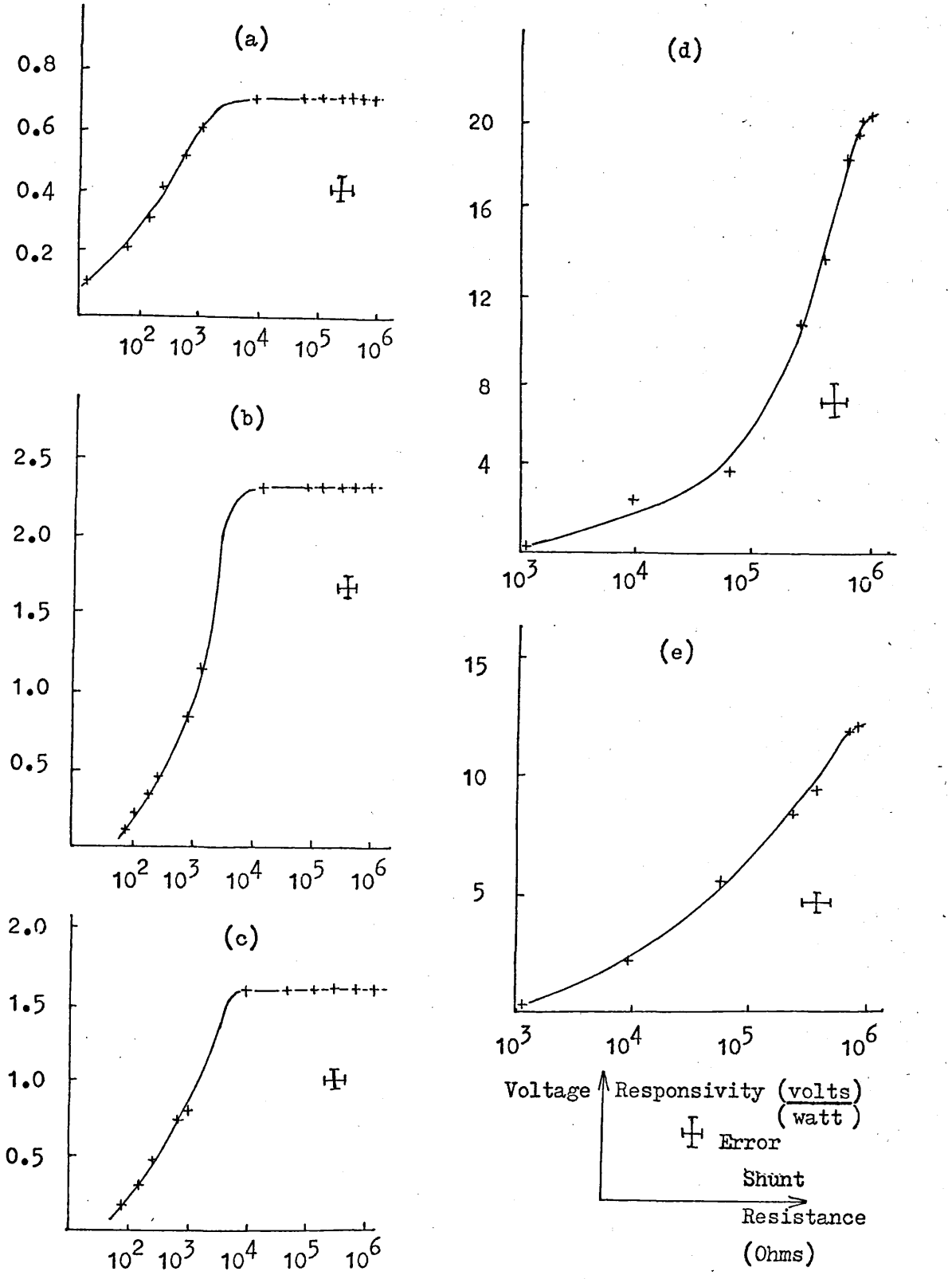


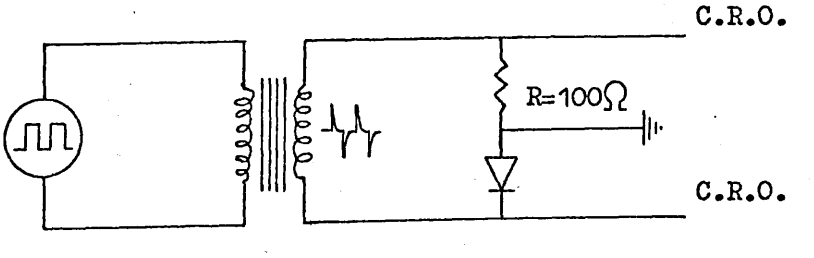
Figure 7.4 Responsivity vs. Shunt Resistance
(a) W-Au (b) W-Al (c) W-W (d) W-Ge (e) W-Si

The Johnson noise arose from the resistive components, R_B and R_S in the equivalent circuit, and the shot noise from the same sources and possibly from trapping states in the dielectric layer. It was not possible to eliminate frequency-dependent noise by, for example, inserting a high-pass filter, since the laser was operating in the c.w. mode. The n.e.p. values for a range of bias voltages shown in Figure 7.3 (c) were referred to a 1Hz bandwidth by dividing by $\sqrt{B_w}$. It is not easy to compare them with other published results, since various bandwidths are given, with no indication of the noise spectral distribution. However, these n.e.p. results appear better than previous figures for GaAs point-contact diodes[92] although not as good as the cryogenically operated bolometers[90].

By observing the dependence of responsivities on the value of a shunt load, as shown in Figure 7.4, the output impedance of each diode was taken to be equal to the shunt value at which the signal fell by 6dB from the open-circuit level (i.e. peak-voltage/2). In applications where the detector response time is important, a load in shunt with the detecting circuit is used to reduce the time constant. The output impedance of the diode is important since it determines the extent to which the shunt may be reduced without significant loss of signal voltage.

The current-voltage characteristics of each diode were measured using positive and negative-going spikes, derived by transforming an audio-frequency (900Hz) square wave, see Figure 7.5, after Green et al[93]. The voltages across the series resistor and the diode were

Figure 7.5 Circuit used for I vs.V measurements



displayed on an oscilloscope, and their values observed and plotted for successive values of the driving amplitude. The resulting I vs. V curves are shown in Figure 7.6. The asymmetries in the M.I.M. diodes were very small, whereas for the M.I.S. diodes the reverse currents were smaller than the forward currents, but noticeably different from the high-impedance values expected of conventional diodes. (See also Becklake et al[92].) The zero-bias resistances, estimated from the slopes of the I-V curves at the origin, were in good agreement with the output resistances, as shown in Table 7.1.

Table 7.1 Measured Zero-Bias Resistances and Output Resistances

Diode Type	Zero-Bias Resistance(Ω)	Output Resistance(Ω)
Tungsten-Silicon	$2.3\ (\pm 0.3).\ 10^4$	$7.8\ (\pm 0.1).\ 10^5$
Tungsten-Germanium	$4.5\ (\pm 0.3).\ 10^4$	$3.2\ (\pm 0.1).\ 10^5$
Tungsten-Gold	220 (±25)	200 (±15)
Tungsten-Aluminium	930 (±65)	1100 (±100)
Tungsten-Tungsten	770 (±30)	900 (±30)

7.4.2 Work Function Measurements

The work functions of the evaporated metal electrodes and of the silicon discs were measured in O.H.V., using the vibrating capacitor technique. The other samples were too small for this method, and were not measured. The reference electrode was platinum foil, washed in Decon 75, rinsed with distilled water and cleaned in situ in the vacuum chamber by gas-discharge bombardment. The value of the platinum work function was taken to be 5.32eV[99], giving values for the plane electrodes which were in good agreement with published values, see Table 7.2. A variation of work function with crystal orientationcould be expected to occur both in the polycrystalline plane electrodes and in the tungsten whisker.For example, Table 7.3 shows the variation of work function in tungsten[1]. The difference in work function could

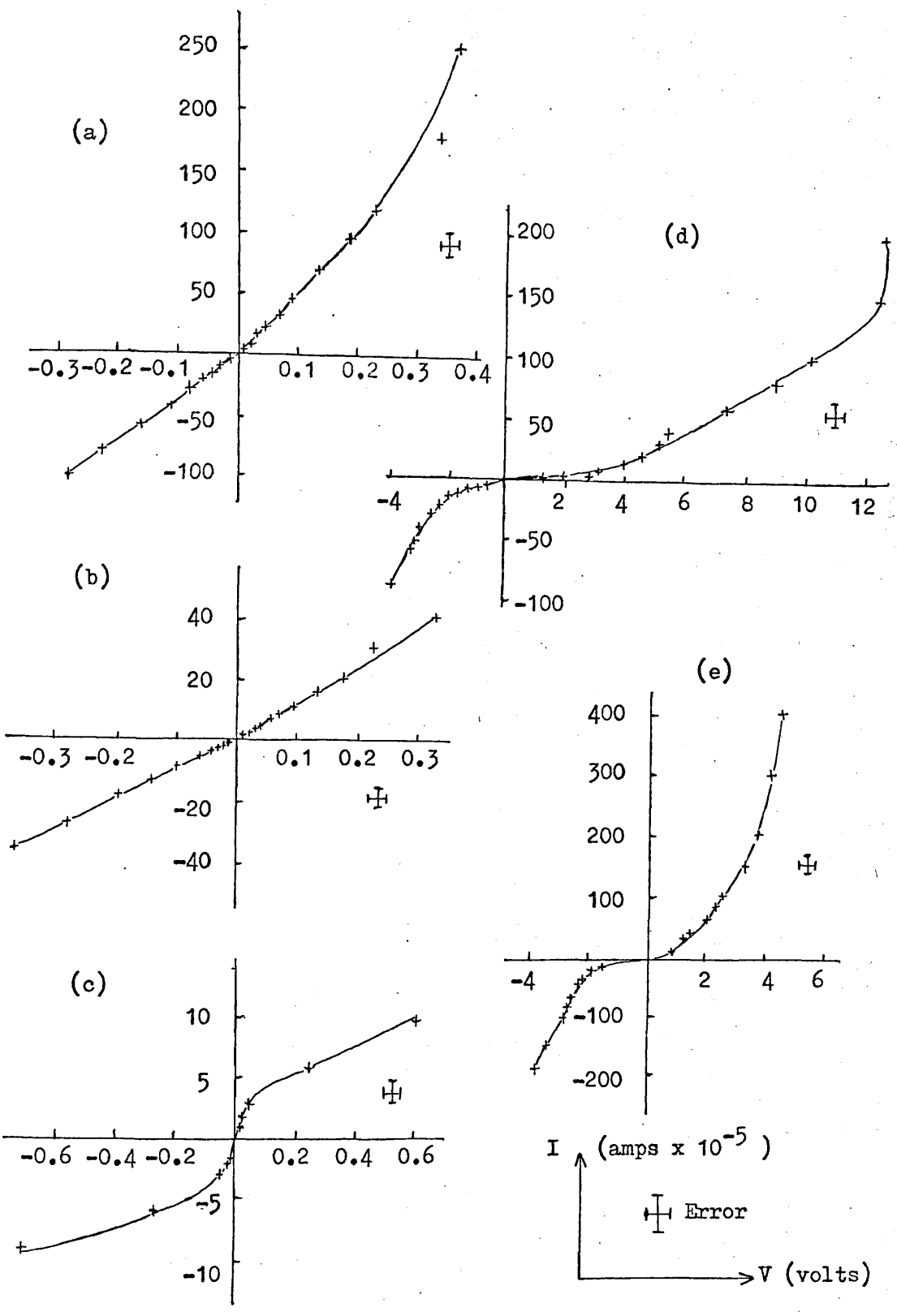


Figure 7.6 Current vs. Voltage Characteristics
(a) W-Au (b) W-Al (c) W-W (d) W-Ge (e) W-Si

Table 7.2 Average Work Functions of the Plane Electrodes

Material	Present Work	Published Results	Reference
Aluminium	4.26 ±0.02eV	4.24 ±0.03eV	1
Gold	5.39 ±0.03eV	5.32 ±0.1eV	1
Silicon	4.85 ±0.02eV	4.85 ±0.05eV	1, 100

therefore vary significantly, perhaps even as much as to reverse in polarity, due to movement of the tip from place to place on the electrode and by either mechanical deformation of the tip or on exposure of a new crystal face when re-etched. This was not thought solely responsible for the observed behaviour of the diodes, however, because of the reproducibility and repeatability. An alternative explanation is offered below.

Table 7.3 Work Functions of Tungsten at Different Crystal Orientations

Crystal Orientation	Work Function (eV)
110	5.3 ±0.12
111	4.4 ±0.03
116	4.3 ±0.03
100	4.6 ±0.08

7.5 Theoretical Considerations

The conduction in a M.I.M. structure can be explained by a combination of thermionic and tunnel contributions. Expressions for the currents flowing between plane parallel electrodes separated by a thin insulating layer have been derived by Simmons for both

tunnel[26, 101] and thermionic[25] mechanisms. The currents flowing between a plane electrode and a small hemispherical one representing the whisker tip were calculated by numerical integration of a series of concentric rings, to describe the performance of the diode. The theoretical responsivity and its dependence on various parameters were calculated by substituting these computed values in the expressions given by Green et al[93]. The response times of both M.I.M. and M.I.S. diodes were estimated from a simple equivalent circuit[92] with values derived from both calculation and experiment.

7.5.1 Tunnel and Thermionic Conduction Characteristics

The potential barrier between two dissimilar plane parallel electrodes separated by a thin insulating layer is shown in Figure 7.7, with the effect of image forces ignored (a) or (b) (after Simmons[26]). The notation is the same as in Chapter II, $E_{F1,2}$ are the Fermi Energies of the two metals and $\psi(x)$ is the x-dependent barrier height, equal to ψ_1 and ψ_2 at the two metal-insulator interfaces.

Simmons' derivations of the tunnel[26, 101] and thermionic[25] current flowing in such a structure at moderate or low temperatures ($< 300^{\circ}K$) are given in Appendix I. Using the same notation, that J_1 is the net tunnel current density flowing with the electrode of lowest work function biased negative to the other, by a low potential, $V, < \frac{\Delta\phi}{e}$

$$J_1 = J_0 \left[\psi_1 \exp\left(-B\bar{\psi}_1^{\frac{1}{2}}\right) - (\bar{\psi}_1 + eV) \exp\left(-B(\bar{\psi}_1 + eV)^{\frac{1}{2}}\right) \right]$$

equation (7.1)

in which the mean barrier height

$$\bar{\psi}_1 = \psi_1 + \left(\frac{s_1+s_2}{2s}\right)(\Delta\phi - eV) - \left(\frac{1.15\lambda s}{\Delta s}\right) \ln\left(\frac{s_2(s-s_1)}{s_1(s-s_2)}\right)$$

equation (7.2)

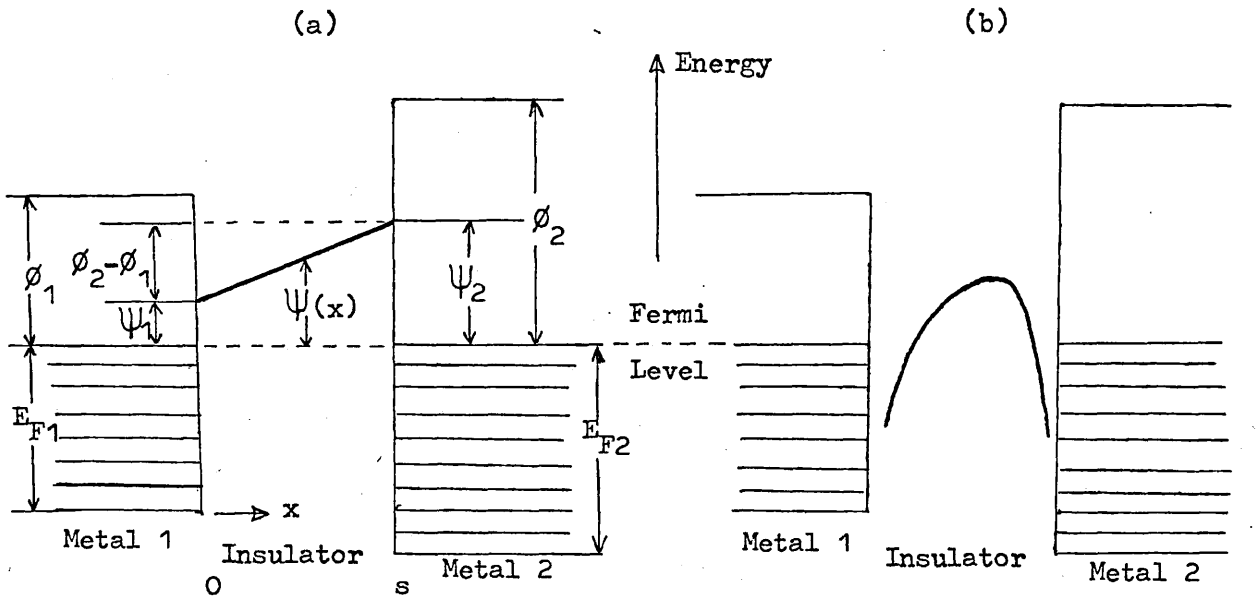


Figure 7.7 Energy Diagram of Potential Barrier between Dissimilar Electrodes
(a) Image forces ignored (b) Image forces included

The limits of the barrier at the Fermi level are (see Appendix I)

$$s_1 = \frac{9.2\lambda s}{3\psi_1 + 4\lambda - (eV - \Delta\phi)} - \frac{1.2\lambda s}{\psi_2 - eV}$$

$$s_2 = s - \frac{1.2\lambda s}{\psi_2 - eV}$$

equations (7.3)

With opposite polarity of bias, $V < \psi_1/e$ the net tunnel current density is

$$J_2 = J_0\left[\bar{\psi}_2 \exp\left(-B\bar{\psi}_2^{\frac{1}{2}}\right) - (\bar{\psi}_2 + eV)\exp\left(-B(\bar{\psi}_2 + eV)^{\frac{1}{2}}\right)\right]$$

equation (7.4)

in which the mean barrier height,

$$\bar{\psi}_2 = \psi_2 - \left(\frac{s_1+s_2}{2s}\right)(eV + \Delta\phi) - \left(\frac{1.15\lambda s}{\Delta s}\right)\ln\left(\frac{s_2(s-s_1)}{s_1(s-s_2)}\right)$$

equation (7.5)

and the limits of the barrier at the Fermi level are

$$s_1 = \frac{1.2\lambda s}{\psi_2}$$

$$s_2 = s - \frac{9.2\lambda s}{3\psi_2 + 4\lambda - 2(eV+\Delta\phi)} + s_1$$

equations (7.6)

In the above equations, the saturated zero-field current density $J_0 = e/2\pi h(\Delta s)^2$,

the constant $B = \frac{4\pi\Delta s}{h}(2m)^{\frac{1}{2}}$

$$\Delta s = s_2 - s_1 \quad ,$$

$$\Delta \phi = \phi_2 - \phi_1 = \psi_2 - \psi_1$$

$$\lambda = \frac{e^2 \ln 2}{8\pi \varepsilon_r \varepsilon_o s} \quad ,$$

e and m are the electronic charge and mass, respectively, and h is the Planck constant.

For thermionic conduction, the Richardson-Dushman equation was used to derive the net current densities J_1' and J_2' , where the subscripts denote the bias polarity as before:

$$J_1' = AT^2 \exp\left(\frac{eV - \psi_2}{kT}\right) \exp\left[\frac{\left(14.4(7 + \varepsilon_r s(\Delta\phi - eV))\right)^{\frac{1}{2}}}{\varepsilon_r skT}\right]$$

equation (7.7)

and $$J_2' = AT^2 \exp\left(\frac{-\psi_2}{kT}\right) \exp\left(\frac{14.4(eV + \Delta\phi)}{\varepsilon_r s(kT)^2}\right)^{\frac{1}{2}}$$ equation (7.8)

where the Richardson constant, $A = 4\pi mek^2/h^3$.

Low voltages have been assumed, $0 \leqslant V \leqslant \frac{\Delta\phi}{e}$ but $\frac{eV}{kT} > 2$.

Simmons' numerical calculations for plane parallel electrodes showed that the tunnel currents would dominate the behaviour of a M.I.M. sandwich at smaller thickness of the insulating layer, and thermionic at larger. The thickness at which the dominance changed depended on temperature, dielectric constant, and the difference between the work functions of the electrodes. Another particularly interesting feature

was that at low voltages the forward direction for tunnel current was when the electrode of lower work function was positivelybiased, but that for thermionic current was with the same electrode negatively biased. Furthermore, as the bias voltage was increased, the dominance of one mechanism over the other could change. Hence, the rectification ratio of the structure, defined as

$$\eta = \left|\frac{I(V)}{I(-V)}\right| = \left|\frac{J(V)}{J(-V)}\right| \qquad \text{equation (7.9)}$$

could vary, even so much as to change from >1 to <1, i.e. change diode polarity, depending on the voltage, temperature, barrier thickness and the work functions of the electrodes.

For this work, numerical calculations were based on a physically more appropriate geometry, which is shown in Figure 7.8. Instead of two plane electrodes, one was part of a small sphere, representing the tip of the antenna. The insulating layer was considered to be homogeneous, to have a clearly defined interface with the plane electrode, and to be of uniform thickness, d. The whisker tip, radius A_w, was inserted into the insulator layer such that the minimum distance to the base electrode was s.

In order to calculate the currents flowing in such a device, the spherical surface was considered to consist of rings, of radius r and constant width

$$\delta r = \frac{R}{1000} .$$

Assuming typical values for A_w of 0.1 to 1.0μm [93, 94, 95] and for d of the order 100Å [102], the curvature of the tip was small compared to insertion depth, so that each annulus was considered to be plane parallel with the base electrode. The distance, t, of a ring from the plane metal electrode was given by

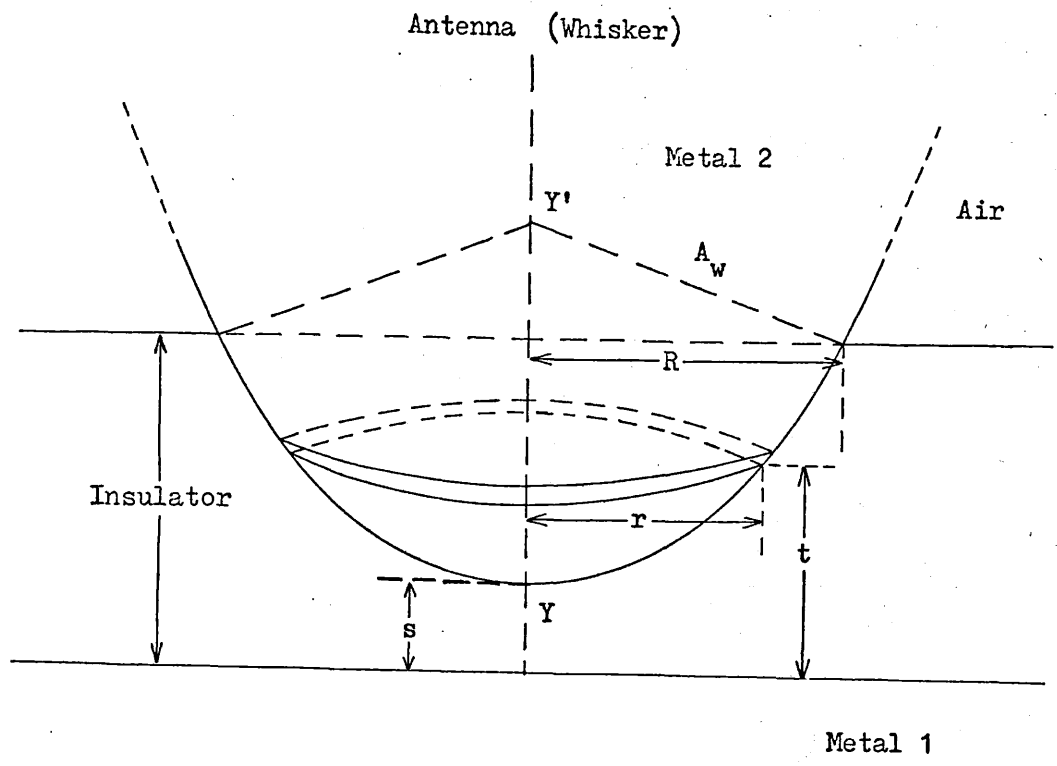


Figure 7.8 Schematic Representation of a Hemispherical Whisker Tip Partially Embedded in the Insulating Layer of a Plane Electrode

$$t = s + A_w - (A_w^2 - r^2)^{\frac{1}{2}}$$ equation (7.10)

The contribution from each annulus was then calculated from equations (7.1), (7.4), (7.7) and (7.8) for current density, multiplied by the area of the annulus, $2\pi r\,\delta r$, and using equation (7.10) for the effective insulator thickness to each ring. The upper limit of r was $r = R = \left(2A_w(d-s) - (d-s)^2\right)^{\frac{1}{2}}$, and the tip was divided into 1000 annuli of radii $\left(\frac{R}{1000}\right)$ to R, so that equation (7.9) became

$$\eta = \left| \frac{\sum_{r=0}^{R} 2\pi \left(J_2(r) + J_2'(r)\right) r\,\delta r}{\sum_{r=0}^{R} 2\pi \left(J_1(r) + J_1'(r)\right) r\,\delta r} \right|$$ equation (7.11)

The computer program, which included the calculations of responsivity from the formulae in Paragraph 7.5.2, is described in Appendix II.

As an example, the results of the I vs. s computations are shown in Figure 7.9 for the parameter values given in the key. The scale of the figure is too small to show differences between J_1 and J_2 and between J' and J_2'. The heights of the potential barriers, ψ_1 and ψ_2, were chosen to be between zero and the values of the work functions, ϕ_1 and ϕ_2. For physically reasonable cases, values of ψ_1 between 0.5 and 4.0 were taken, with $\Delta\phi = \Delta\psi$ and $\varepsilon_r = 8$[102]. The result for η (from equation (7.11)) with the same parameter values as in Figure 7.9 is shown in Figure 7.10.

These results were used to calculate the responsivities of M.I.M. diodes as described in 7.5.2 below.

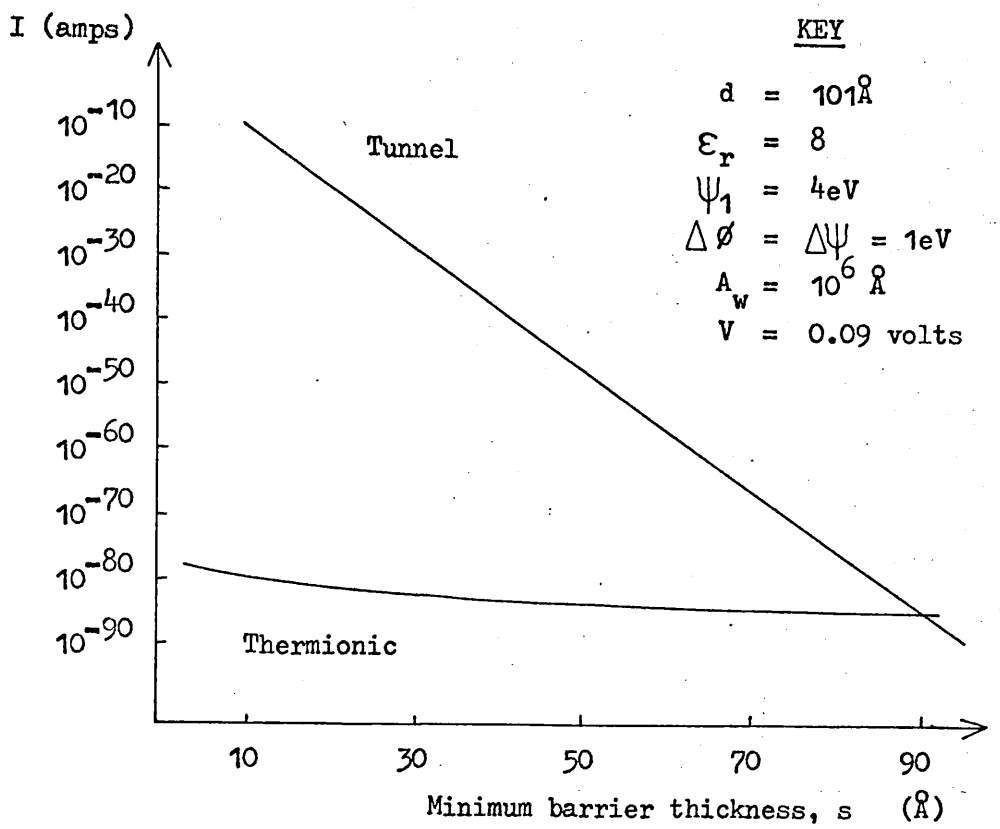


Figure 7.9 Computed I vs s Characteristics of Tunnel and Thermionic Conduction in a Point-contact M.I.M. Diode

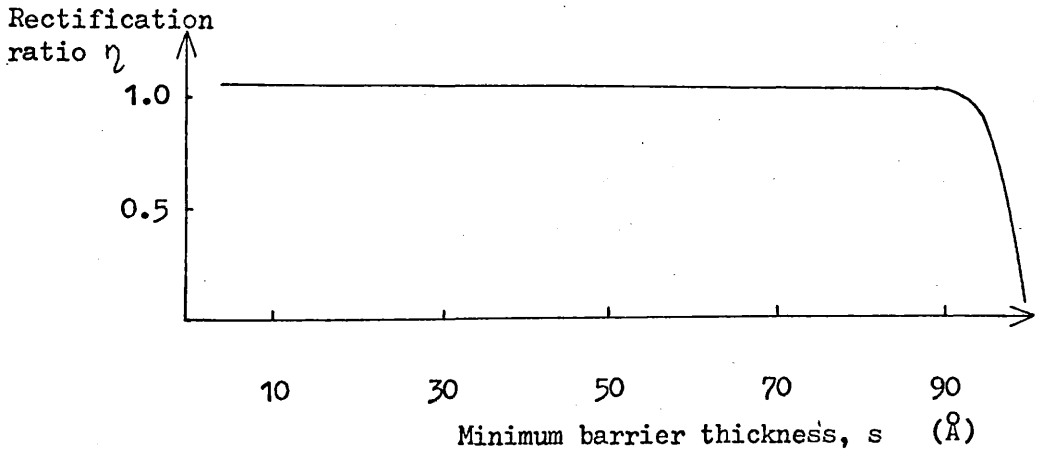


Figure 7.10 Computed Rectification Ratio vs. s in a Point-contact M.I.M. Diode

7.5.2 Responsivity: Theory and Calculations

The current-voltage characteristic of a nearly-linear diode may be described by[91]:

$$I = aV + bV^2 + cV^3 \ldots \qquad \text{equation (7.12)}$$

where $a, b, c, \ldots$ are constants.

If a small power level P is coupled into the diode, at angular frequency ω, then the resistance of the diode, R_B, may be approximated to $1/a$. The voltage appearing across the diode is approximately:

$$V = (2PR)^{\frac{1}{2}} \sin\omega t = \left(\frac{2P}{a}\right)^{\frac{1}{2}} \sin\omega t. \qquad \text{equation (7.13)}$$

Substituting (7.13) into (7.12) and ignoring terms higher than quadratic since V is small,

$$I = a\left(\frac{2P}{a}\right)^{\frac{1}{2}} \sin\omega t + \left(\frac{bP}{a}\right)(1 - \cos 2\omega t) \qquad \text{equation (7.14)}$$

which has a d.c. component, $I_D = \frac{bP}{a}$

Therefore the detected voltage $V_D = I_D R_B = \frac{bP}{a^2}$

Hence, the voltage responsivity, $\gamma = \frac{V_D}{P} = \frac{b}{a^2}$ volts/watt

equation (7.15)

The rectification ratio calculated from equation (7.12) is

$$\eta(V) = \left|\frac{I(V)}{I(-V)}\right|$$

$$= \frac{aV + bV^2 + cV^3}{aV - bV^2 + cV^3} \qquad \text{equation (7.16)}$$

$$\simeq 1 + \frac{2bV}{a}$$

$$\eta(V) \simeq 1 + \frac{2\gamma V}{R_B}$$

Hence $$\gamma \simeq \frac{\left(\eta(V) - 1\right).R_B}{2V} \qquad \text{equation (7.17)}$$

The theoretical responsivity of a point-contact diode could therefore be calculated directly from a knowledge of the I-V characteristics. For this purpose, the diode barrier resistance R_B was taken as the average of the forward and reverse resistances, calculated from the computed I. vs. V values. This was reasonable since it had already been assumed, for low V, that $R_B \simeq 1/a$, i.e. the diode was nearly linear in the area of origin. Some of the computed results of γ vs. s are given in Figure 7.11(a) and (b), for various values of A_w, V and ψ_1. The solid curves indicate a positive signal from the diode; the dotted curves a negative signal. The value of ε_r for these curves was 8. For $\varepsilon_r = 3$, the curves were substantially the same. For example, in the lower family of curves, a change of ε_r from 8 to 3 was equivalent to a change in the tip radius, A_w, from 10^{20}Å to 10^{40}Å, within half an order of magnitude or better, over the whole range of s. A change in the barrier heights produced a much larger effect. The two families of curves in Figure 7.11(a) are for ψ_1 = 4eV and ψ_1 = 1eV. For clarity, the curves for ψ_1 = 0.5eV are shown separately in Figure 7.11(b). The higher values of V_D have been omitted to comply with the theoretical assumption that $eV \leq \psi_1$. For comparison with the curves of Figure

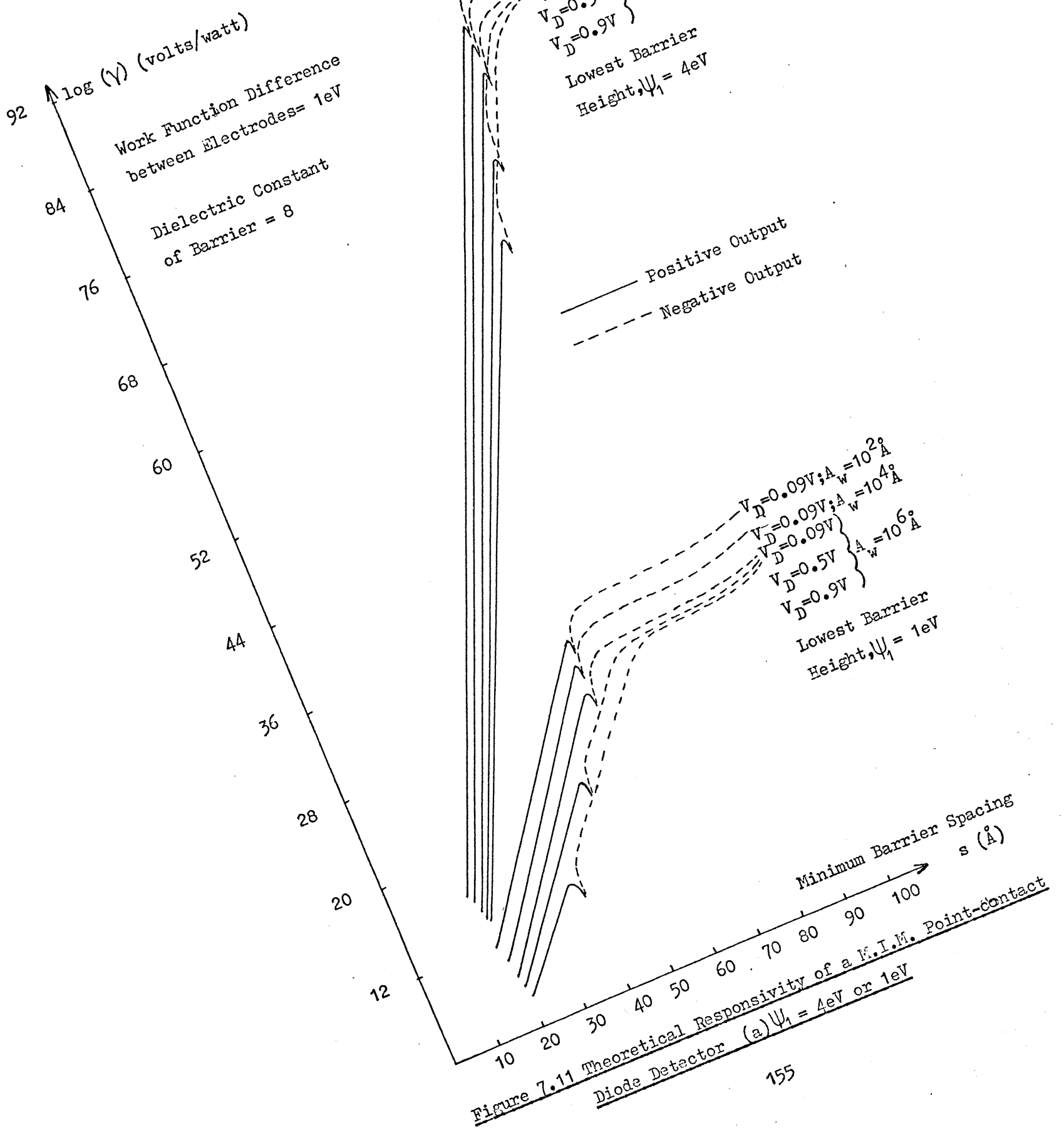


Figure 7.11 Theoretical Responsivity of a M.I.M. Point-contact Diode Detector (a) ψ_1 = 4eV or 1eV

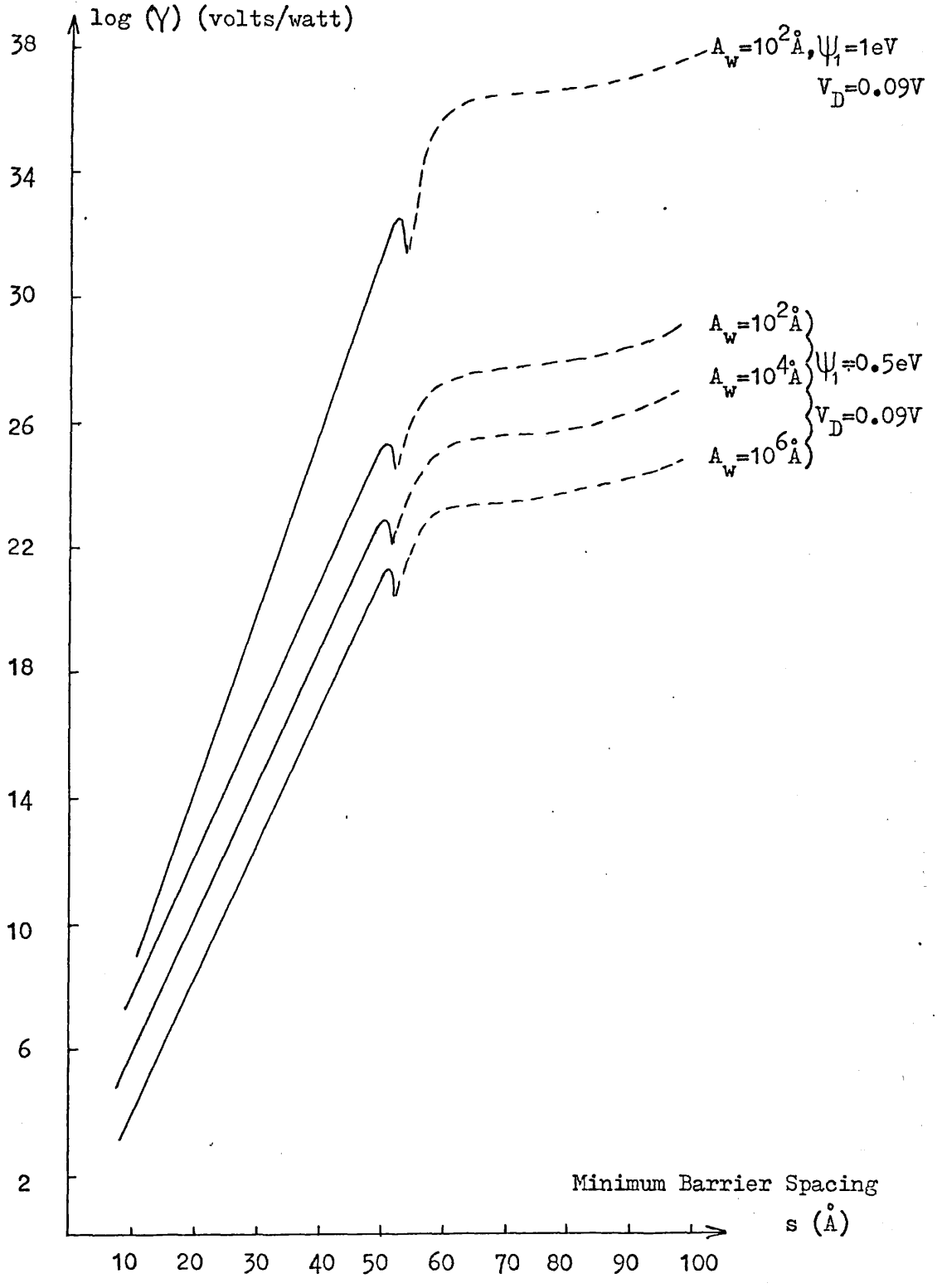


Figure 7.11 (b) ψ_1 = 1eV or 0.5eV

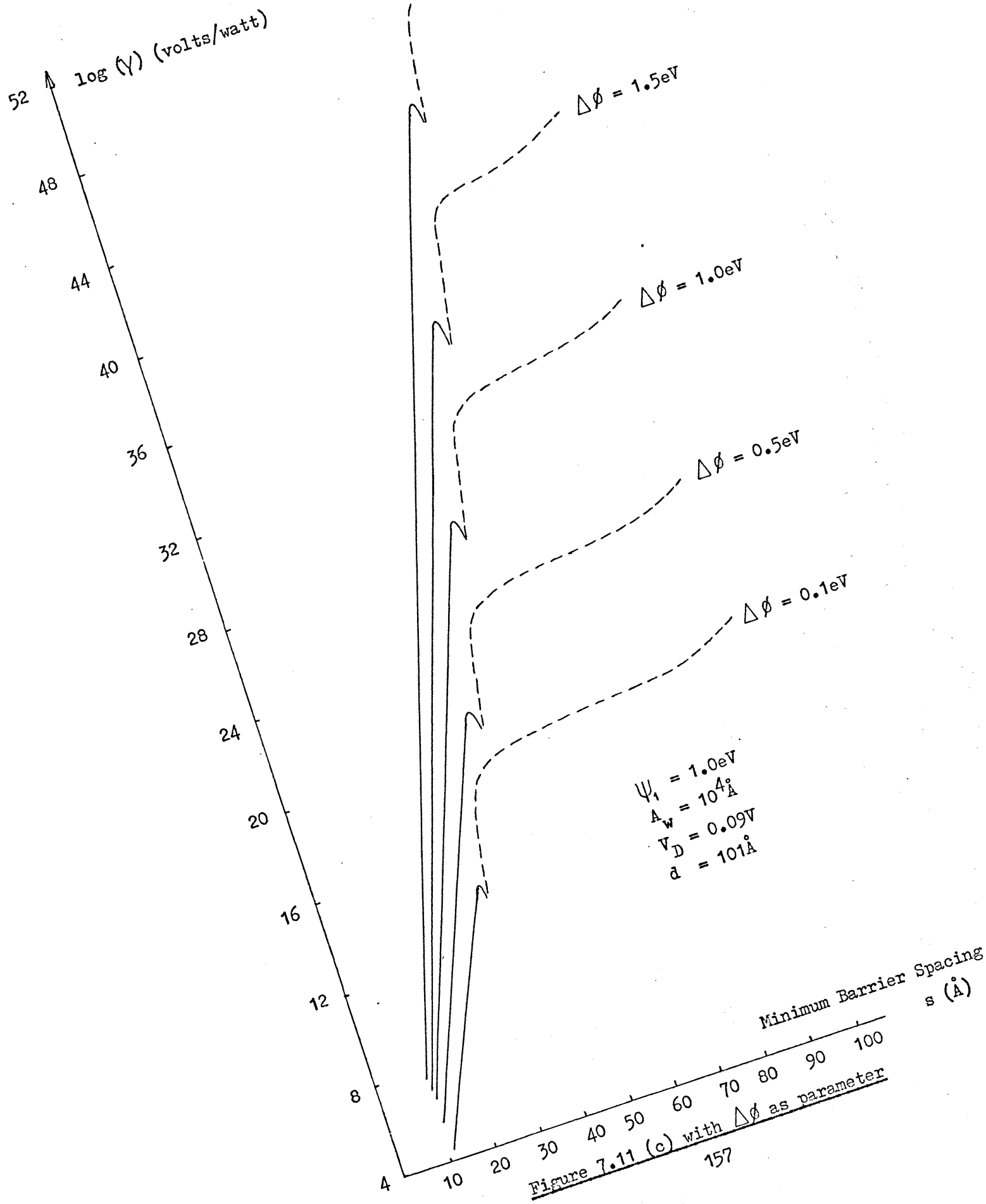


Figure 7.11 (c) with $\Delta\phi$ as parameter

7.11(a), the curve for ψ_1 = 1eV has been reproduced, with all other parameters the same as the uppermost of the group for ψ_1 = 0.5eV. The value $\Delta\psi = \Delta\phi$ = 1eV in all three cases. The exact relationship between ψ and ϕ is not clear, although it is always assumed that $\psi < \phi$ and $\Delta\psi = \Delta\phi$.

To examine the effect of the work functions on the diode performance, γ was calculated as a function of s, as before, for a wide range of values of $\Delta\phi$, from 0.1eV to 2eV. The results are shown in Figure 7.11(c), with the values of the other parameters in the key. It is apparent that, apart from the value of s, the barrier heights and work functions are by far the most critical parameters in the model.

7.5.3 Response Times

The response times of isolated M.I.M. and metal-semiconductor junctions are extremely small, being in the region of 10^{-16} seconds[91] and 10^{-12} seconds[92] respectively. The barrier capacitance and the series or spreading resistance, especially in the M.I.S. diodes, might be expected to affect the performances at some high frequencies. To estimate their effect, a simple equivalent circuit was used[92], as shown in Figure 7.12, where C_B is the barrier capacitance and R_B, R_S are the barrier and spreading resistances.

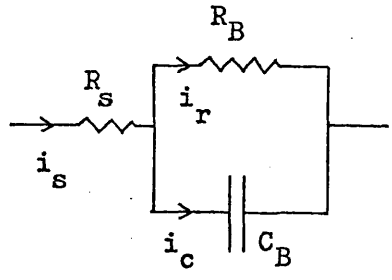


Figure 7.12 Equivalent Circuit of a Point-contact Diode

The power absorbed in the barrier resistance, $P_B = i_r^2 R_B$

in the barrier capacitance, $P_C = 0$

and in the spreading resistance, $P_s = i_s^2 R_s$

Now $i_s = i_c + i_r$; $\dfrac{i_c}{j\omega C_B} = i_r R_B$

Total coupled power, $P = P_B + P_s + P_C = i_r^2 R_B + \left|(i_c+i_r)^2 R_s\right|$

Therefore $$\frac{P_B}{P} = \frac{i_r^2 R_B}{i_r^2 R_B + \left|1+2j\omega C_B R_B - \omega^2 C_B^2 R_B^2\right| i_r R_s}$$

$$= \frac{R_B}{R_B + (1+\omega^4 C_B^4 R_B^4 - 2\omega^2 C_B^2 R_B^2 + 4\omega^2 C_B^2 R_B^2)^{\frac{1}{2}} R_s}$$

$$= \frac{R_B}{R_B + R_s(1+\omega^2 C_B^2 R_B^2)}$$

$$= \frac{1}{\left(1 + \dfrac{R_s}{R_B}\right)\left(1 + \dfrac{\omega}{\omega_c}\right)^2} \qquad \text{equation (7.18)}$$

where $$\omega_c = \frac{\left(1 + \dfrac{R_s}{R_B}\right)^{\frac{1}{2}}}{C_B (R_B R_s)^{\frac{1}{2}}}$$

The high frequency performance will therefore be limited, ω_c being the -3dB frequency, or corner frequency, since the ratio $\dfrac{P_B}{P}$ at $\omega = \omega_c$ will be exactly half that at $\omega = 0$.
For an order of magnitude calculation, let us assume a junction diameter of the order of 0.1μm [91,994], insulating layer of thickness 50Å and relative permit tivity of 8[102], then $C_B \simeq 10^{-16}$ Farads. Assuming that the barrier resistances of the M.I.M. diodes

could be approximated to the measured resistance, since the spreading resistance would be low, $\sim 1\Omega$, then equation (7.18) indicates values for ω_c between 45 and 100THz. For the M.I.S. diodes, let us assume a barrier resistance similar to the M.I.M. devices, permittivity of 4.5[103] and spreading resistances equal to the measured resistances, since now the barrier resistances are small compared to the overall values, then ω_c is in the region of 15THz. Hence all the diodes would be expected to be fully sensitive at $337\mu m$ (891GHz).

7.6 Comparison of Experimental and Theoretical Results

Comparison of experimental and theoretical responsivities of point-contact diodes reveals the following similarities:

1. The polarity of the signal reversed with increasing pressure/decreasing insulator thickness, s, proceeding through a gradual change on either side of the rapid reversal.

2. The responsivity was greater at initial contact/greatest barrier thickness than after the polarity reversal had occurred. Similar experimental results to both items 1 and 2 had been observed earlier with both M.I.M.[93] and M.I.S.[94] diodes. The behaviour was repeatable from diode to diode and reproducible as the whisker pressure was increased or decreased, corresponding to decreasing and increasing barrier thickness. This indicates that the variation was mechanistic, rather than a random variation of, say, work function difference. There was, however, some evidence of change in $\Delta\phi$, since the actual polarity on initial contact was not always the same. This would be the result if the relative magnitudes of the two work functions were reversed, see 7.4.2. since the forward and reverse directions of current would be exchanged, (from the definitions of J_1, J_2, J_2', and J_2') resulting in the reciprocation of η and a polarity reversal of $(\eta - 1)$

and γ. Nevertheless, the qualitative behaviour, items 1 and 2, was always observed.

3. The polarity of the signal could be reversed by varying the incident power[93], Figure 7.11(a) shows that at a fixed value of s, the responsivity will vary with the value of V, and at some values of s the responsivity polarity may change. For example, in the lower family of curves, at $s = 50\text{Å}$, an increase of V from 0.09V to 0.9V will cause a change in γ from $+10^{26}$ to -10^{22} volts/watt.

4. The responsivity was improved by minimising the area of the contact. Figure 7.11(a) shows an increase of γ directly proportional to a decrease in whisker tip radius.

5. The best response was obtained with a high-mobility semiconductor. This result had been expected, but not always obtained in earlier work[92], since higher mobility provides a lower spreading resistance, R_s. Equation(7.18)shows the dependence of $\frac{P_B}{P}$ (i.e. the ratio of power absorbed in the barrier to total absorbed power) on the circuit values. In the simple case where $\omega \ll \omega_c$, it is apparent that the lower the value of R_s, the more power will be absorbed in the detecting junction.

From these comparisons, the theoretical model can be seen to be compatible with all the experimental observations, as mentioned in paragraph 7.2. Indeed, the model could be used to predict all but the improved responsivity with bias, observed with the M.I.S. devices. Although the theory has been derived for a M.I.M. barrier, it is interesting to note that the behavioural characteristics of the M.I.S. diodes with variation of whisker pressure were qualitatively similar to those of the M.I.M. A possible explanation is that the mechanism described by the model contributes to a d.c. voltage appearing across the deple-

tion layer near the semiconductor surface. The polarity of this component will vary accordingly, and the rectification ratio of the semiconductor device will be modified, becoming greater with forward bias and smaller with reverse. The M.I.S. devices may therefore be expected to behave in a qualitatively similar manner to the M.I.M. diodes, but to exhibit higher rectification ratios, and therefore higher responsivities, as experimentally observed. The application of appropriate external bias certainly improved the responsivities of the M.I.S. diodes, as illustrated in Figure 7.3, yet had little or no effect on the M.I.M. diodes, possibly because the latter were so nearly linear that the rectification ratios were not significantly affected. There was, however, a discrepancy of scale between the experimental and theoretical results, in that the theoretical responsivities were enormously greater than those achieved in the laboratory. It has been assumed that the insulator barrier was of uniform thickness, homogeneous, and had clearly defined interfaces with the electrodes. It has been shown theoretically[104] that local variations in the insulator thickness can cause disagreement between theoretical and experimental values of current density by several orders of magnitude. There is therefore ample scope to affect the rectification ratio, η , and the barrier resistance, R_B, and hence the responsivity. Earlier experimental results[105] have also shown appreciable variations in current values due to impurities in the insulator layer, caused by producing the insulator in less than perfectly clean conditions, whereas it has been implicitly assumed that the layer was a perfect dielectric. Since it is possible to produce rectification with nominally symmetrical devices[102] (i.e. both electrodes of the same work function), it may be that such observations result from an asymmetry, caused by, for example, a poorly defined transition from metal to dielectric at one or both interfaces. Figure 7.11 shows how important are the values of the barrier heights.

A further assumption implicit in the model was that the work functions and barrier heights at each metal-insulator interface were single-valued, i.e. that at each electrode the grain size was larger than the junction area. This is unlikely to be met in practice, so that

the effective values of the work functions and barrier heights would be complex functions of the actual grain distributions in the neighbourhood of the whisker tip. The possible variation in $\Delta\phi$ as the whisker tip was moved from site to site on the planar electrode would be reduced, since neither extremely low nor high values would occur individually. Also, it seems likely that this would generally result in a smaller difference between the effective work functions at the interfaces, hence a lower responsivity. For example, at the parameter values shown in Figure 7.11(c), an order of magnitude reduction in $\Delta\phi$ (from 1.0 to 0.1eV) would cause a reduction in responsivity of around 14 orders. Reflection electron micrographs of the aluminium films showed a range of grain size from $< 40\text{Å}$ to $> 400\text{Å}$. The grain size in tungsten ribbon has been reported as $\sim 0.01\text{mm}$ [34], and it seems reasonable to assume that the similarly manufactured wire is similar. Hence the films had grain sizes smaller than the junction area, so that the effect just described would obtain, whether the whisker tip was single- or multi-grained.

No attempt has been made to adjust the various parameters in order to achieve quantitative agreement between the experimental and theoretical values. So many were not actually measured due to the inherent practical difficulties that there were too many unknown to facilitate such a fit. If the assumed values of ε_r, s, d, A_w, and V were substantially accurate, then such agreement would necessitate great reduction in the values either of the barrier heights at the interfaces, ψ_1 and ψ_2, or of the work function difference ($\Delta\phi = \phi_1 - \phi_2 = \psi_1 - \psi_2$) or both. It is obvious also that for the high values of responsivity given in Figure 7.11, even a very low power level coupled into a detector would cause a very high voltage to be developed across the barrier, which would not only render the theoretical assumptions $V < \frac{\Delta\phi}{e}$ and $V \leq \psi_1/e$ invalid, but would exceed the dielectric strength of the insulator.

Direct comparison was further frustrated by the inequality of the

measured power incident on the diode assembly and the power absorbed by the device.

For the derivation of the tunnel current density equations, see Appendix I, it was assumed that the contribution from electrons above the Fermi level was negligible. Figure 7.13[106] illustrates the possible mechanisms contributing to the overall current flow:

(i) Thermionic emission
(ii) Thermally assisted tunneling
(iii) Field assisted tunneling
(iv) Tunneling from the Fermi level or below

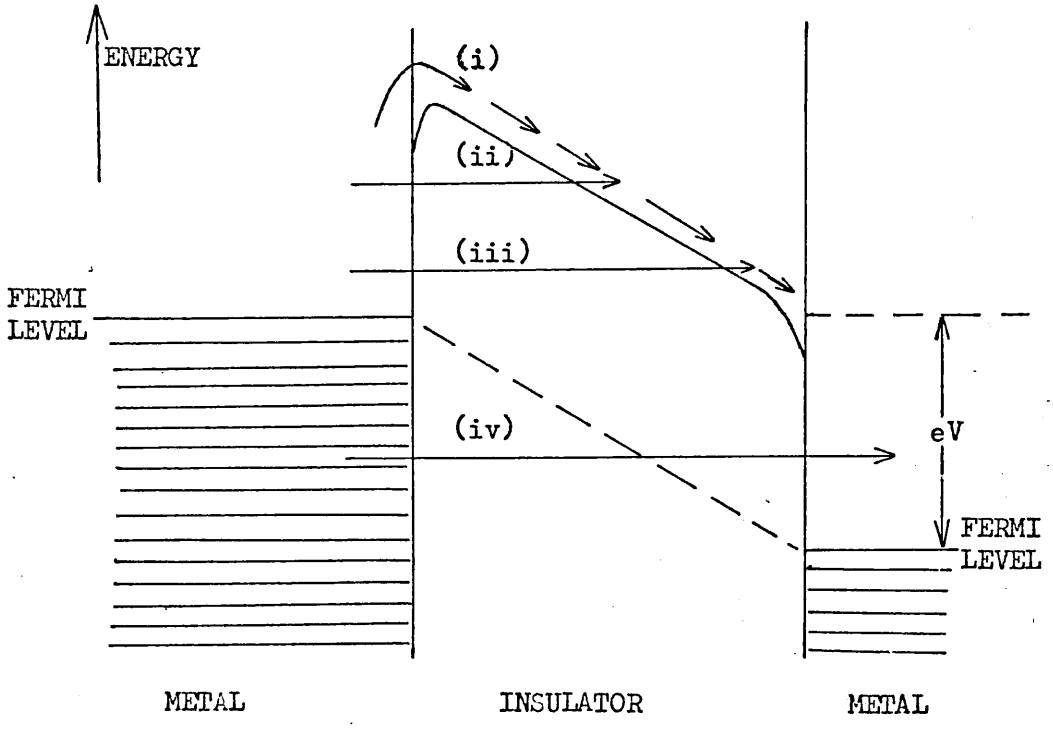


Figure 7.13 Current Mechanisms in a Biased M.I.M. Structure (after Pitt[106])

In the present model, mechanisms (ii) and (iii) have been neglected. Whether this was physically reasonable in the particular devices described here has not been quantitatively assessed. However, it seems reasonable to assume that the qualitative features of the model will be

retained whether or not these mechanisms were significant compared to mechanism (iv). If the forward directions were the same for all three tunneling contributions, then the γ vs. s curves would change by a scaling factor; both in the γ axis since the barrier resistance R_B and the rectification ratio η would be affected, and in the s axis since the thickness at which the dominant mechanism changed from tunnel to thermionic would be affected. If the forward directions were not the same, then a polarity reversal of γ would still occur, but where one tunneling mechanism was succeeded by another rather than by thermionic conduction. In any case the barrier resistance R_B would increase with increasing thickness s and would be likely to dominate the $|\gamma|$ vs. s characteristic, by being more strongly dependent on s than is $(\eta - 1)$. Therefore the shape of the curves in Figure 7.11 would not be changed.

One further check was made on the calculations. Since the computations were limited by the outermost surface of the insulator, the current flow from metal to metal via the air above the insulator has been neglected, for simplicity. For low values of s, the distance, t, to the uppermost rings was so much larger and the incremental currents so much smaller that they did have negligible effect. For larger s, however, this effect was not so easily ignored. Calculations showed that for s values greater than 60Å there was a small difference between the calculated responsivities for insulator thickness, d, of 101Å and 10101Å, see Figure 7.14. This was attributed to the inclusion in the 10101Å case of a greater active area (see Figure 7.8), and hence lower barrier resistance, R_B. The factor $(\eta - 1)$ was virtually unchanged at the highest values of s since η was by then very much less than unity. Although the theoretical consideration of the dielectric - air interface would make the model more complete, changes of the magnitude indicated in Figure 7.14 do not significantly alter the good qualitative comparison between the model and the experimental observations described above.

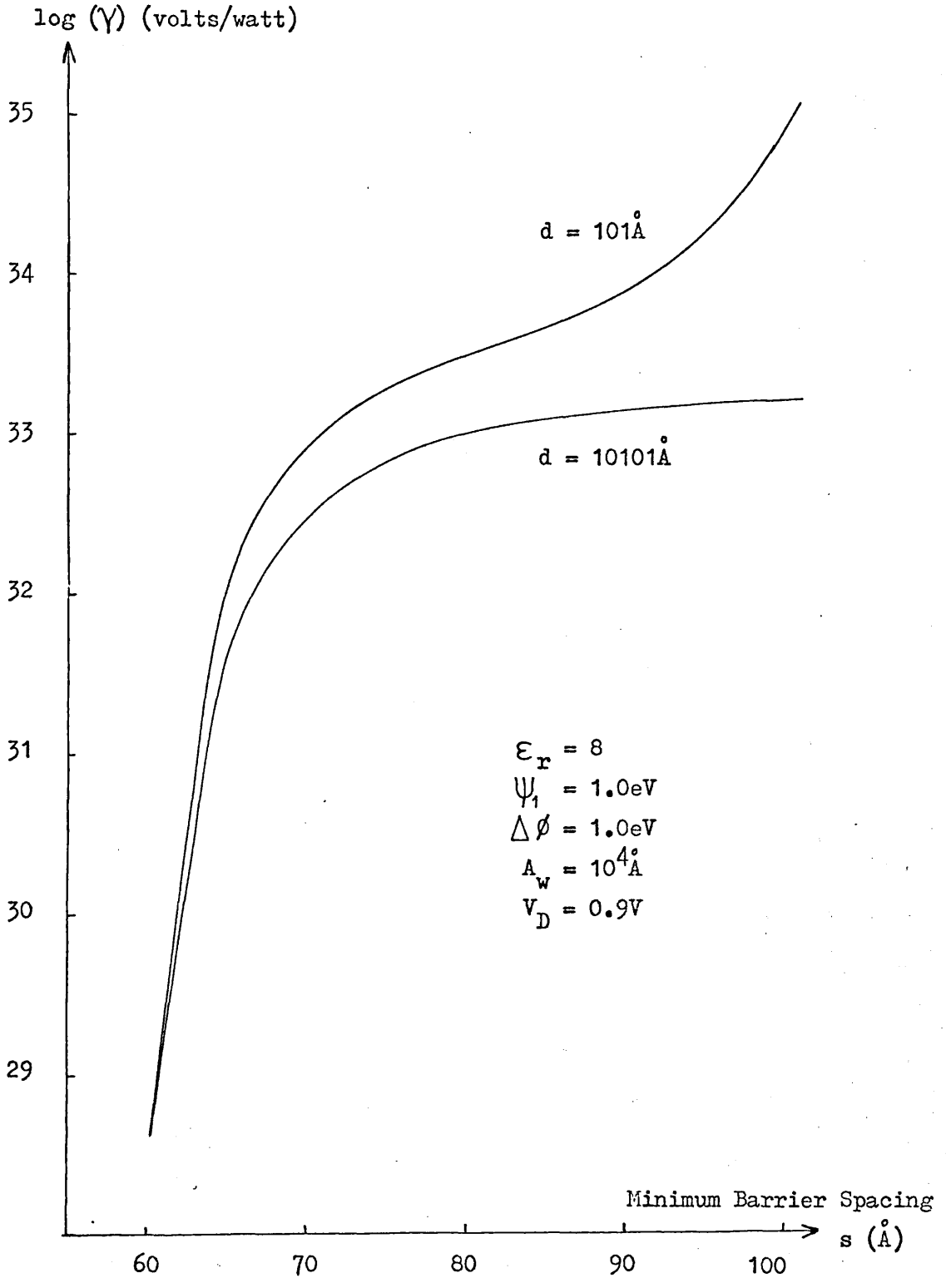


Figure 7.14 The Effect of Insulator Layer Thickness on Responsivity

Chapter VIII Conclusions and Some Suggestions for Further Work

8.1 Introduction

The role of this final chapter is to collate the developments, achievements and failures of the work described earlier and to suggest some avenues for further work which will make good the shortcomings or build on the successes. The main body of the work has been concerned with accurate measurement of the work functions of some rare earth specimens, while the development of a model for fast, sensitive, far infra-red detectors has been a specific application to which the work functions of the materials involved have been of great importance. These two sectors will be discussed separately except where there is some common factor.

8.2 Conclusions

8.2.1 Work Function Measurements

The experimental method used in this work was chosen to provide the first measurements of the true work functions of high-purity samples of the rare earths Gd, Tb and Dy. Earlier measurements had been made either on heavily contaminated specimens [63] or by photoemission [79,84,85] or field-emission [57] techniques subject to the errors described in Chapter II. The vibrating capacitor technique also avoided magnetic fields and, being a c.p.d. measurement, gave the area-weighted average work function of the polycrystalline films. This has been analytically verified. Because the method is not biased to patches of particular work function values (e.g. high or low) and therefore observes the whole area of the surface, it is likely to be most sensitive to contamination-induced changes. For exactly the same reason it is likely to result in a larger spread of measured values from polycrystalline samples than any of the non-c.p.d. methods which are biased. For example, photoemission measurements will favour patches of lowest

work function whatever fraction of the sample is so formed.

To facilitate measurements on high-purity specimens, the vibrator mechanism was designed specifically for this work to enable the measurements to be made in situ on U.H.V. evaporated films, and to allow other measurements, resistive and optical, to be made in situ on the same specimens. Some of the possible sources of error were shown empirically or analytically to be negligible. The remaining errors in c.p.d. measurement amounted typically to 11meV, although this figure is not indicative of the sensitivity of the method. It arose from the polycrystallinity of the samples which caused the minimum signal from the capacitor to be non-zero and to spread over a range of bucking potential, V_E. The sensitivity of the method was limited by the p.s.d. used to observe the signal. In this work it was of the order of 1meV, determined by the smallest change in V_E which it was possible to detect outside the region of the signal minimum. With more sophisticated electronics and smooth, single-crystal capacitor plates the sensitivity of the method should be reducible to a very much lower figure.

The U.H.V. deposition and storage of the rare earth films renders them among the purest and cleanest observed to date [57,79,84,85]. No general correlation was found between the variation of measured work function and the residual atmosphere observed within the U.H.V. chamber. The films were sufficiently thick to avoid quantum size effects and to be of the usual bulk h.c.p. structure[51]. There was some evidence from electron diffraction measurements showing the c-axis to be roughly perpendicular to the substrate surface. This is consistent with the resistivity values measured in the plane of the surface being nearer the higher a-axes values than those of the c-axes (see table 6.3) (although the high values could also have been caused by a significant residual resistivity). However, there was electron diffraction and resistivity evidence to indicate the presence of f.c.c. phase in a few cases.

To facilitate changes in the temperatures of the specimens,

either for experimental variation or for annealing, a compact heater has been developed which is U.H.V.-compatible, easily fitted and does not interfere with optical measurements. It provides more uniform distribution of temperature than a wire heater and is more efficient. Temperatures of over 550°K have been achieved on glass substrates and over 770°K on alumina.

The results obtained from the work function measurements on the rare earths Gd, Tb and Dy have been described and discussed in detail in Chapter VI. These are the first measurements on clean (U.H.V.) Gd by the c.p.d. technique and on Tb and Dy by any technique. The median values obtained for the fresh films of Gd, Tb and Dy were 3.545eV, 3.151eV and 3.077eV respectively. The range of values for Gd and Tb were 0.74eV and 0.67eV respectively. The lowest value for Gd was in very good agreement with most of the previous measurements on clean Gd by other workers. As stated earlier, the relatively large spread of results may be due to the c.p.d. technique's measurement of the area-weighted average, whereas other methods are less evenly weighted or specifically biased towards one type of patch, i.e. low work function in the cases of photoemission and of thermionic emission with a high collecting field, higher work function for thermionic emission at weak field[1]. The work functions were found to decrease with age in U.H.V. but to increase with annealing (in U.H.V.) and with exposure to nitrogen then air. The latter changes were to be expected due to energetically favoured crystal reorganisation and to gross non-metallic surface contamination respectively. The reason for the decrease during U.H.V. storage is not known and is therefore a candidate for further study. It is possible that the reduction was caused by a (partial) monolayer of positive ions reducing the dipole layer at the surface of the specimen, but the variations of the work functions of the two films prepared and stored at slightly lower vacua were monotonically increasing, which suggests otherwise, since gross contamination caused an increase.

Perhaps the most significant result was the absence of a change in either the work function or its temperature dependence at

the Curie temperature. The theoretical postulates for such a change at T_c were either (a) a change in the chemical potential[80] (the calculation has since been shown to be invalid[82]) or (b) a change in the average magnetisation[81], which determined both the gradient with which the electron density decayed from n_+ (the uniform positive charge background) to zero (at a distance from the surface) and the gradient with which the magnetisation decreased near the surface. If it is assumed that the magnetisation varies linearly from its bulk value at a point within the material to zero at a point outside, then for specimen temperatures close to the Curie point, the surface magnetisation is linearly dependent on that temperature difference, whereas the bulk magnetisation varies as the square root[107]. The linear temperature dependence provides no discontinuity at the transition from ferromagnetic to paramagnetic phase, in either the work function or its temperature coefficient. An exponential spatial decrease in surface magnetisation has been used to predict a small (0.135eV) change in the work function of nickel at its Curie point[81], which has been observed by some workers[66,76] but not by others[75,77]. The present experimental result for Gd therefore supports the view that the magnetisation at the surface is linearly dependent both spatially and on temperature. It is also consistent with the photoemission results of Blodgett, Spicer and Yu[79] who found no change in the band structure at the Curie temperature.

New calculations have been made of the mean work functions of Gd, Tb and Dy and of their anisotropies based on the electronic band structure, yielding 3.15eV, 3.17eV and 3.19eV respectively for the work functions and 0.55eV for their anisotropies. The extent of the anisotropy estimate is conservative since a necessarily limited number of crystal directions has been used. Both values are lower than the experimental results on Gd, although one would only expect to observe the full range of anisotropy by measuring single crystals (either patches or bulk). Therefore, either the real anisotropy is very much larger than the limited calculations suggest or another factor was responsible for the

spread of results. The difference between the mean work functions of the h.c.p. and f.c.c. phases of Gd has been calculated as 0.04eV and was clearly not the cause. The calculations were based on a positive background charge model with a linear potential gradient near the surface, but with a variation with crystal direction. The addition of pseudopotentials to represent more realistically the crystal lattice has yet to be made for these complex metals. It is interesting to note that for the simpler h.c.p. metals Zn and Mg, the first-order pseudopotential corrections increased the theoretical work functions in the c-axis by 0.38eV [70]. The pseudopotentials would also contribute to the anisotropy [70], whether to increase or decrease it remains to be seen.

8.2.2 Far Infra-red Detector Diode Model

The experimental characteristics of point-contact metal-insulator-metal (M.I.M.) and metal-insulator-semiconductor (M.I.S.) diodes used as detectors or mixers for very high frequency radiation have been reliably and repeatedly observed by many workers, including those diodes fabricated in this department for the detection of far infra-red radiation (337μm wavelength) from an H.C.N. laser [86]. The reasons for their behaviour have been imperfectly understood. A new model has been developed [86] in which both tunnel and thermionic currents are included and account is taken of the physical and electrical parameters encountered in such diodes in practice, including the work functions of the electrodes used in their construction. All the experimentally observed features are qualitatively described by the model, including the effects of variations in the thickness of the barrier between the two electrodes, the radius of the tip of the point-contact whisker and the voltage (corresponding to input power) appearing across the diode. Although the model is strictly applicable only to M.I.M. devices, the M.I.S. diodes exhibited similar qualitative behaviour [86]. Quantitative agreement was not so good. Several possibilities have been suggested to account for this, including microscopic geometrical variations, impurities in the insulator layer, polycrystallinity of the electrodes, and inefficiency of the whisker as an

antenna and contributions to the tunneling current by electrons from above the Fermi level. The model predicts improved responsivity from diodes whose electrodes have high work functions, leading to large barrier heights, and a large difference between their respective work functions. It is not easy to verify this experimentally, however, because the factors described above for the poor quantitative agreement are not repeatable from material to material. Also the two requirements of large work functions and a large difference between them are counteractive.

Apart from the thickness of the insulator between the tip of the whisker and the plane electrode, the most critical parameters in the model are the work functions, and the difference between them, which determine the heights of the barriers at the metal-insulator interfaces.

8.3 Suggestions for Further Work

8.3.1 Work Function Measurements

The least well understood features of this work were the spread of the results and the decrease of the work function during aging in U.H.V. To determine whether anisotropy was responsible for the spread it would be appropriate to measure the work functions at different crystal directions of a monocrystalline sample. One way to achieve clean monocrystalline surfaces would be to deposit fresh layers under U.H.V. onto the surfaces of one or more bulk crystals, which are now more readily available. A more sophisticated, and hopefully more realistic, model of the work function and its anisotropy should be attempted, possibly using pseudopotential perturbations from the uniform background model as has already been performed for some of the simpler elements [70]. It would be interesting to discover whether the work functions of the various faces of rare earths are monotonically dependent on the packing density of atoms within them, as the simpler theory predicts, and whether the anisotropy is sufficienty extensive to account for the present results. The use of monocrystalline

surfaces would also remove probably the principal objection to the field emission retarding potential method[18], which does not require a surface of known work function for reference. However, the technique is not strictly accurate for all surfaces, only those for which the Sommerfeld free-electron model is satisfied[19]. It would be useful to study the crystal structure of the specimens directly, for example by Low Energy Electron Diffraction, and to monitor the surface for contamination, using an energy discriminator in conjunction with the L.E.E.D. or a specific technique such as Auger Electron or Secondary Ion Mass Spectroscopy.

Although the present results suggest that the observed decrease in work function on aging in U.H.V. was not due to a small amount of surface contamination, it would perhaps be more conclusive if the measurements were able to be taken more rapidly after the deposition. The manner of electrical connection to the specimen, and to the heater and thermocouple, could be redesigned in such a way as to enable the substrate to be moved rapidly from the deposition position to the measurement position while the vibrator assembly was at least approximately aligned and ready to be operated. On the present equipment this would entail only minor changes to the vibrator, if any, mainly to allow the reference surface to be chielded during the deposition of the specimen, but major changes to the electrical feedthroughs. The specimen holder would ideally be able to be raised by at least 25mm and rotated through at least 180° without entangling the electrical connections. The most obvious solution would be a set of rotating and sliding copper contacts (not carbon brushes), with one of each type in series in each interconnection to allow both movements simultaneously. A rather complicated mechanical design would be required to provide these together with electrical isolation between them. An alternative solution has been suggested by Hölzl and Schrammen[108] which avoids sliding contacts. The oscillating electrode moves in a plane parallel to the surfaces of the electrode pair, rather than perpendicular as in the present work. The specimen holder is only required to be moved very slightly, to align the two surfaces

comprising the capacitor, while the evaporation proceeds from "behind" the oscillating electrode. Indeed it is possible to measure the c.p.d. during the deposition either of the specimen itself or of a contaminating material, since the lateral oscillation of the reference electrode repeatedly exposes the substrate to the source. The parallelism and reproducibility of the capacitor gap may be difficult to control accurately, but the authors claim an accuracy of 1meV.

The behaviour of the work function at magnetic transition temperatures is still unclear both experimentally and theoretically, although it is known to be dependent on the surface magnetisation, how it differs from the bulk magnetisation and its spatial distribution. A more reliable and complete knowledge of the characteristics of the work function at both the Néel and Curie points should help to advance the understanding of magnetic behaviour, and should be pursued for the rare earths and for the other magnetic materials.

Having produced such specimens and measured their work functions, other, preferably related, parameters should be measured. The optical conductivity values which were obtained from the present films were calculated from reflectance measurements (74) which penetrate quite deeply into the material. It is to be expected that better correlation would be found if the optical, or other, parameters which were considered were obtained from measurements which are as surface-sensitive as the work function, such as ellipsometry.

8.3.2 Far Infra-red Detector Diodes

Although point-contact whisker diodes are useful as rapid-response detectors of extremely high frequency radiation operating at convenient temperatures, their general application is hampered by a relatively poor responsivity, particularly in the case of M.I.M. devices, and by their physical frailty. Checking the predictions of the model, that the responsivity would be enhanced

by large barrier heights at the conductor-insulator interfaces and by a large work function difference, would be facilitated by experimental redesign to obviate some of the practical difficulties mentioned earlier. The use of U.H.V. as the insulating barrier would preclude impurity conduction, although the constancy of the electrode separation would be difficult to ensure. The use of single crystals for the electrodes would more clearly identify the respective work functions and possibly minimise the irregularity of the spacing between them. Surfaces with work functions lower than the rare earths are available, but are probably not suitable for this purpose. The low work function elements are difficult to use (e.g. caesium melts at 28°C) and the conventional coated cathodes are not described by this model. It therefore seems sensible to extend the use of the monocrystalline rare earth and other specimens proposed above to construct diodes in U.H.V. using various whisker materials to alter the work function difference. Further theoretical studies should be made, to quantify the effect of electrons originating from above the Fermi level and to extend the model more exactly to describe the M.I.S.diodes. The improved understanding of their behaviour should enable better detectors to be made.

Some attempts have already been made to improve the robustness of the devices, using electrochemical etching to produce submicron-dimensioned electrode-tips in metal films supported by insulating substrates[109]. The more difficult task of producing such a tip which just contacts another metal or semiconductor electrode does not appear yet to have been accomplished. One of the major difficulties, that of obtaining a low capacitance in such a rugged whisker diode, might be achieved as follows (See Figure 8.1). The base electrode is coated with an insulator sufficiently thick to minimise the capacitance between the two electrodes. A window in the insulator is produced by an appropriate chemical etchant. The top layer of metal from which the whisker electrode is to be fabricated is then deposited overall, and defined by suitable photolithographic techniques into a narrow strip which protrudes over the window and contacts the underlying electrode , but with the necessary thin

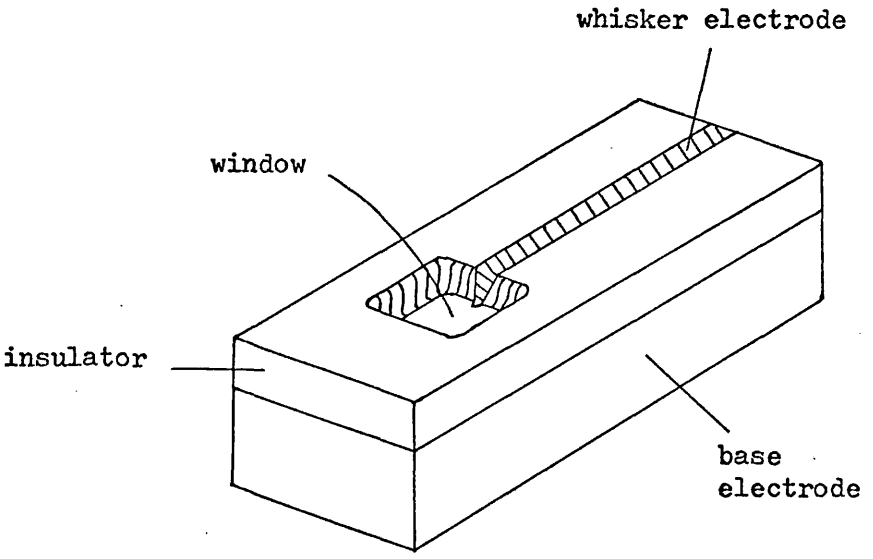


Figure 8.1 Suggested Construction of a Robust Whisker Diode

insulating barrier between them. The electrochemical etching produces a tapered end to the whisker electrode as it is withdrawn from the solution, the difficulty lies in completing the withdrawal as the whisker tip is still just in contact with the base electrode. Another possibility would be to produce a narrow track of metal film on the surface of an intrinsic semiconductor or an insulator which can be converted at a localised area into a high mobility semiconductor, for example by ion implantation, at the site where the submicron tapered end of the whisker has been previously defined. The second method probably entails more difficult fundamental materials problems, but once these were solved would be more amenable to production.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the encouragement of his supervisor, Dr. R.F. Miller, and useful discussions particularly on the interpretation of electron diffraction patterns. He is also indebted to Dr. M.M. Moore for the X-ray photographs of the tungsten samples and to Mr. R. LeSadd, of G.E.C. Medical, Wembley, for supplying the disc from which the tungsten reference block was cut.

Thanks are also due to the technicians for their aid, particularly Messrs. R.F. Elton, G. Hayward and J. Taylor for making the extra hardware required, Mr. M.E. Thyer for microphotography and multiple-beam interferometry, Mr. R. May for the electron diffraction patterns, Mr. R. Burton for the photographs of the equipment and Mr. R. Mason for his assistance in preparing the tungsten reference and the flame-polished borosilicate substrate.

The cooperation of Dr. L.S. Julien and Mr. R.B. Inwood in sharing the U.H.V. equipment is also acknowledged. Additional thanks are due to R.B.I. for allowing the inclusion of his optical conductivity results prior to publication. Mr. M. Belgrave's efforts in testing and calibrating the film heaters are also recognised.

The work on the point-contact whisker diodes arose from a need for a fast detector for studies of an HCN laser conducted by Mr. M.W. Booton, who made all the measurements on the diodes. The development of the model to describe their performance was conducted jointly through repeated discussion and refinement.

The author is also obliged to the University of London for providing the facilities and to the Science Research Council for a Research Studentship.

Finally, it is a pleasure to thank Miss C. Chandler for her patience and hard work in typing the manuscript.

Appendix I Derivation of Current Equations for M.I.M. Diodes

The equations (7.1) to (7.8) from which the current-densities in the point-contact diodes were calculated were taken directly from three papers by Simmons. His derivations are described here, but with the notation changed where necessary to be compatible with this work.

AI.1 Tunnel Currents (26, 101)

The potential barrier between two similar metal electrodes separated by an insulating film of thickness s and with a potential difference, V, is shown in Figure AI.1

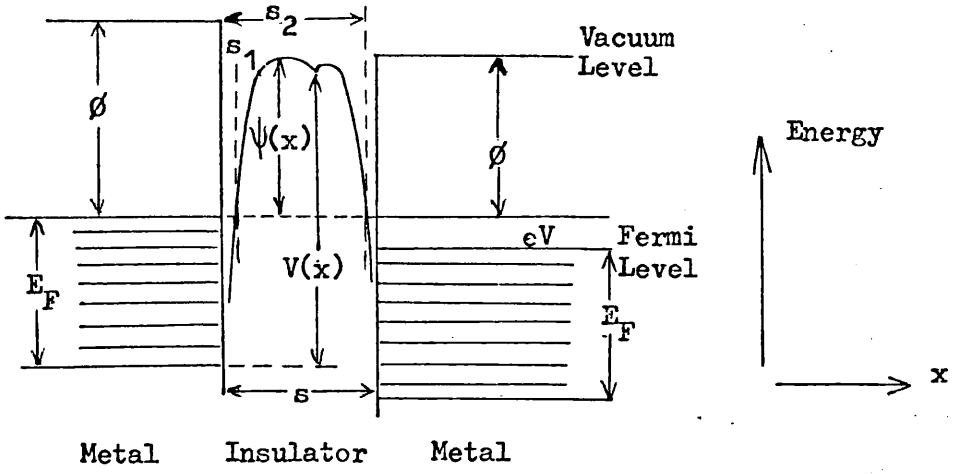


Figure AI.1 Potential Barrier in an Insulating Film between two Electrodes

The probability $D(E_x)$ that an electron can penetrate a barrier of height V(x) is given by the W.K.B. approximation:

$$D(E_x) = \exp\left(\frac{-4\pi}{h}\int_{s_1}^{s_2}\left(2m(V(x) - E_x)\right)^{\frac{1}{2}}dx\right) \quad \text{equation(AI.1)}$$

where $E_x = \frac{m v_x^2}{2}$ which is the energy component of the electron in the x direction

The net tunnel current density is the difference between the currents flowing from electrode 1 to electrode 2 and vice-versa, which are calculated from the probability function and the Fermi-Dirac distribution function, f(E), to give

$$J = \int_0^{E_m} D(E_x)\,\zeta\, dE_x \qquad \text{equation (AI.2)}$$

where E_m is the maximum energy of the electrons in the electrodes, and ζ is a function of f(E), E_F and V.

Writing $D(E_x) \simeq \exp\left(-B(E_F + \overline{\psi} - E_x)\right)^{\frac{1}{2}}$

where $\overline{\psi} = \dfrac{1}{s_2 - s_1}\int_{s_1}^{s_2} \psi(x)dx,$

which is the mean height of the barrier above the Fermi level of the negatively biased electrode,

and $B = 4\pi (s_2 - s_1)(2_m)^{\frac{1}{2}}/h$ at low voltages,

evaluation of equation(AI.2) gives

$$J = J_o\left[\overline{\psi}\exp\left(-B\overline{\psi}^{\frac{1}{2}}\right) - (\overline{\psi} + eV)\exp\left(-B(\overline{\psi} + eV)^{\frac{1}{2}}\right)\right] \qquad \text{equation (AI.3)}$$

where $J_o = e/2\pi h(s_2 - s_1)^2$.

It had been assumed that the temperature is low enough that $E_x \simeq E_F$, and that the barrier width $\Delta s = s_2 - s_1$ is constant for $E_x > V_{(x)} > E_F$.

Equation (AI.3) can be applied to any shape of potential barrier. For a structure with dissimilar electrodes, the barrier is asymmetric, and the net current densities are defined as J_1 when the electrode of lower work function is biased negative to the other and J_2 when the opposite bias is applied.

The mean barrier height may be obtained from the expression

$$\overline{\psi} = \frac{1}{\Delta s}\int_{s_1}^{s_2} \psi(x)dx.$$

Using a parabolic approximation for the shape of the potential barrier when account is taken of the image forces arising from the metal-dielectric interfaces, the mean barrier height with the electrode of lower work function negatively biased is given by

$$\bar{\psi} = \frac{1}{\Delta s}\int_{s_1}^{s_2}\left(\psi_1 + (\Delta\phi + eV)(x/s) - 1.15\lambda s^2/x(s-x)\right)dx$$

equation (AI.4)

where $\lambda = e^2 \ln 2/8\pi\varepsilon_r\varepsilon_o s$

and $\Delta\phi = \phi_2 - \phi_1$

The limits s_1 and s_2 are the roots of the equation $(x) = 0$, and are given by

$$\left.\begin{aligned} s_1 &= \frac{9.2\lambda s}{3\psi_1 + 4\lambda - (eV - \Delta\phi)} - \frac{1.2\lambda s}{(\psi_2 - eV)} \\ s_2 &= s - \frac{1.2\lambda s}{(\psi_2 - eV)} \end{aligned}\right\} \quad 0 < V < \frac{\Delta\phi}{e}$$

Integration yields

$$\bar{\psi} = \psi_1 + \left(\frac{s_1 + s_2}{2}\right)(\Delta\phi - eV) - \left(\frac{1.15\lambda s}{\Delta s}\right)\ln\left(\frac{s_2(s-s_1)}{s_1(s-s_2)}\right) = \bar{\psi}_1$$

equation (AI.5)

Equation (7.1) is simply obtained by substituting equation (AI.5) for $\bar{\psi}_1$ into equation (AI.3) for J_1. Equations (7.2) and (7.3) are the same as those above for $\bar{\psi}_1$, s_1 and s_2.

Similarly, when the electrode of higher work function is negatively biased, the mean barrier height is given by

$$\bar{\psi} = \frac{1}{\Delta s}\int_{s_1}^{s_2}\left(\psi_2 - (\Delta\phi + eV)(x/s) - 1.15\lambda s^2/x(s-x)\right)dx$$

In this case the integration limits are given by

$$\left.\begin{aligned} s_1 &= 1.2\lambda s/\psi_2 \\ s_2 &= s - \frac{9.2\lambda s}{3\psi_2 + 4\lambda - 2(eV + \Delta\phi)} + s_1 \end{aligned}\right\} V < \psi_2/e$$

and the integration yields:

$$\overline{\psi} = \psi_2 - \left(\frac{s_1+s_2}{2s}\right)(eV+\Delta\phi) - \left(\frac{1.15\lambda s}{\Delta s}\right) \ln\left(\frac{s_2(s-s_1)}{s_1(s-s_2)}\right) = \overline{\psi}_2$$

equation (AI.6)

The current equation is obtained by substitution of equation (AI.6) for $\overline{\psi}_2$ into equation (AI.3) for J_2, while equations (7.5) and (7.6) are again copied from the above expressions for $\overline{\psi}_2$, s_1 and s_2.

AI.2 Thermionic Currents [25]

In considering the thermionic current flow through the structure, the Richardson-Dushman equation gives the current for one metal-dielectric interface as

$$J' = AT^2 \exp(-\psi/kT)$$

where $A = 4\pi mek^2/h^3$

For the two-electrode structure, the net current is the difference between the currents flowing in opposite directions and is given by

$$J = AT^2 \left(\exp(-\psi_1'/kT) - \exp(-\psi_2'/kT)\right)$$

where ψ_1' and ψ_2' are the maximum barrier heights above the Fermi level of the negatively and positively biased electrodes, respectively. When a potential V exists between the electrodes,

$$\psi_2' = \psi_1' + eV$$

so that $J' = AT^2 \left(\exp(-\psi_1'/kT)\,(1 - \exp(eV/kT))\right)$

equation (AI.7)

The barrier heights are obtained by the substitution of approximations to the barrier shape into the integrand of equation (AI.4). The maximum height is then obtained by determining the first turning point, by differentiation. In the absence of applied potential, the resulting barrier height is

$$(\psi)_{max} = \psi_2 - 14.4(7+\varepsilon_r s\Delta\phi)^{\frac{1}{2}}/\varepsilon_r s$$

measured with respect to the electrode of lower work function. With the electrode of lower work function negatively biased by a potential $V \leqslant \Delta\phi/e$

$$(\psi)_{max} = \psi_2 - eV - \left(\left(14.4(7+\varepsilon_r s\ \overline{\Delta\phi - eV})\right)^{\frac{1}{2}}/\varepsilon_r s\right)$$

equation (AI.8)

measured with respect to the electrode of lower work function.

With the electrode of higher work function negatively biased, for any value V,

$$(\psi)_{max} = \psi_2 - \left(\left(14.4(7+\ \varepsilon_r s\ \overline{\Delta\phi + eV})\right)^{\frac{1}{2}}/\varepsilon_r s\right)$$

equation (AI.9)

measured with respect to the electrode of higher work function.

The net currents flowing with the electrode of lower work function biased negatively and positively are J_1' and J_2' respectively. By substituting equations (AI.8) and (AI.9) into equation (AI.7) in turn, and assuming $eV/kT > 2$, the currents are given by

$$J_1' = AT^2 \exp\left(\frac{eV - \psi_2}{kT}\right) \exp\left(\frac{\left(14.4\left(7+\varepsilon_r s(\Delta\phi - eV)\right)\right)^{\frac{1}{2}}}{\varepsilon_r skT}\right)$$

equation (AI.10)

and

$$J_2' = AT^2 \exp\left(\frac{-\psi_2}{kT}\right) \exp\left(\frac{14.4(eV+\Delta\phi)}{\varepsilon_r s(kT)^2}\right)^{\frac{1}{2}}$$

equation (AI.11)

Appendix II Computer Program for M.I.M. Model

The current equations (7.1) to (7.8) were evaluated and summed for the geometry of Figure 7.8, and the responsivities were calculated from equation (7.17), using a computer program written in FORTRAN IV and run on a Control Data Corporation 6600 machine. The symbols in the equations were changed where necessary to a form which the computer would accept, as follows:

PHI	=	$\Delta\phi$
AT2	=	AT^2
Z1	=	ψ_1
Z2	=	ψ_2
P	=	ε_r
D	=	d
A	=	A_w
S	=	s
RADIUS	=	R/1000
DR	=	δr
R	=	r
SS	=	t

} Figure 7.8

S11	=	s_1
S12	=	s_2
DS1	=	Δs

} for tunnel currents in equation (7.1)

S21	=	s_1
S22	=	s_2
DS2	=	Δs

} for tunnel currents in equation (7.4)

PI1	=	$\overline{\psi}_1$
PI2	=	$\overline{\psi}_2$
G1, G2		intermediate functions of the S parameters, to simplify some of the statements.
CON1	=	Argument of first exponential term in equation (7.1)
CON2	=	" " " " " " " (7.4)
CCC1	=	" " second " " " " (7.1)
CCC2	=	" " " " " " " (7.4)
ZERO1	=	J_o in equation (7.1)
ZERO2	=	J_o " " (7.4)

ARG1 = Argument of first exponential term in equation (7.8)

ARG2 = " " second " " " " " (7.8)

ARG3 = " " first " " " " " (7.7)

ARG4 = " " second " " " " " (7.7)

CUR1 = current contribution to J_1 from one annulus

CUR2 = " " " J_2 " " " "

SUM1 = Summation of CUR1

SUM2 = " " CUR2

THERM1, THERM2, TOTAL1, TOTAL2 = intermediate terms to simplify the statements for CURTH1 and CURTH2

CURTH1 = J_1' for the summation of all the annuli

CURTH2 = J_2' " " " " " " "

TOT1 = $J_1 + J_1'$ for the summation of all the annuli

TOT2 = $J_2 + J_2'$ " " " " " " "

RATIO1 = η for thermionic current contributions (= J_2'/J_1')

RATIO2 = η " tunnel " " (= SUM2/SUM1)

RATIO3 = η " total currents (= TOT2/TOT1)

R1 = R_B for thermionic currents

R2 = R_B " tunnel "

R3 = R_B " total "

GAMA1 = γ for thermionic currents

GAMA2 = γ " tunnel "

GAMA3 = γ " total "

The currents were expressed in amps.cm^{-2}; work functions and barrier heights in electron-volts; voltages in S.I. volts and the various barrier geometrical parameters in Ångstroms. The expressions for the barrier limits became

$$s_1 = \left(\frac{46}{\varepsilon_r}\right) \Big/\!\Big/ \left(3\psi_1 + \frac{20}{\varepsilon_r s} - (V - \Delta\phi)\right) - \frac{6}{\varepsilon_r(\psi_2 - V)}$$

$$s_2 = s - \frac{6}{\varepsilon_r(\psi_2 - V)}$$

} For J_1

and

$$s_1 = \frac{6}{\varepsilon_r \psi_2}$$

$$s_2 = s - \left(\frac{46}{\varepsilon_r}\right) \Big/\!\Big/ \left(3\psi_2 + \frac{20}{\varepsilon_r s} - 2(V + \Delta\phi)\right) + s_1$$

} For J_2

Also $J_o = 6.2 \cdot 10^{10}/(\Delta s)^2$,

$B = -1.025\Delta s$,

$kT = 0.025$, for $T = 300^\circ K$

and $\lambda \doteq \frac{5}{\varepsilon_r s}$

Using $m = 9.11 \cdot 10^{-8}$ gm,

$e = 1.6 \cdot 10^{-19}$ coulomb

$k = 1.38 \cdot 10^{-16}$ erg/deg,

$h = 6.62 \cdot 10^{-27}$ erg-sec.

then $AT^2 = 1.08 \cdot 10^7$ amps/cm^2 at $T = 300^\circ K$

The calculations were executed by nested DO loops: the innermost summed the annular contributions by varying r (R in the program), another varied the minimum thickness s (S in the program), and the next controlled the value of the applied voltage, V (also V in the program). Programs were variously run with either fixed values of A_w, ε_r, ψ_1 and $\Delta\phi$ or with further loops to repeat the central computations with additional values of these parameters.

The following program listing shows loops for all the parameters except ε_r and ψ_1. The values of the parameters used in the loops were specified in the DATA statements at the top of the listing, using dummy variables in the loop statements. Starting with the innermost, the DO number and its parameter were

1	r, ring radius
33	d, thickness of insulator layer
2	s , minimum thickness of insulator layer
5	V , voltage
4	$\Delta\phi$, work function difference
90	A_w, whisker tip radius.

In the example shown, the fixed values for ε_r and ψ_1 are given by P = 8.0 and Z1 = 1.0 (eV) respectively.

```
DIMENSION U(3),W(5),S(100),R(1000),RR(3),DD(2)
DATA U/0.9,0.5,0.09/
DATA W/0.1,0.25,0.5,0.75,1.0/
DATA RR/1000000.,10000.,100./
DATA  DD/101.,10101./
AT2=10.8E+06
Z1=1.0
P=8.0
DO 90 II=1,3
A=RR(II)
WRITE(3,85)A
WRITE(3,77)
DO 4 L=1,5
PHI= W(L)
WRITE(3,80)PHI
Z2=Z1+PHI
ARG1=-Z2*40.
DO 5 M=1,3
V=U(M)
CONST=-V*40.
ARG3=(V-Z2)*40.
WRITE(3,70)V
WRITE(3,77)
WRITE(3,20)
WRITE(3,77)
DO 2 MM=1,10
J=MM*10
S(J)=FLOAT(J)
DO 33 KK=1,2
```

```
      D=DD(KK)
      RADIUS=(SQRT(2.0*A*(D-S(J))-(D-S(J))**2))/1000.0
      DR=2.0*3.1416*RADIUS
      SUM1=0.0
      SUM2=0.0
      THERM1=0.0
      THERM2=0.0
      DO 1 I=1,1000
      R(I)=RADIUS*(FLOAT(I))
      SS=A+S(J)-SQRT(A*A-R(I)*R(I))
      S21=6./(P*Z2)
      S22=SS-(46./P)/(3.*Z2+20./(P*SS)-2.0*(V+PHI))+S21
      DS2=S22-S21
      G2=S22*(SS-S21)/(S21*(SS-S22))
      PI2=Z2-((S21+S22)/(2.*SS))*(V+PHI)-(5.75/(P*DS2))*(ALOG(G2))
      CCC2=-1.025*(S22-S21)*(SQRT(PI2+V))
      CON2=-1.025*(S22-S21)*(SQRT(PI2))
      ZERO2=(6.2E+10)/(DS2*DS2)
      CUR2=(PI2*(EXP(CON2))-(PI2+V)*(EXP(CCC2)))*DR*R(I)
      CUR2=CUR2*ZERO2
      S11=(46./P)/(3.*Z1+(20./(P*SS)-(V-PHI)))-(6./P)/(Z2-PHI)
      S12=SS-(6./P)/(Z2-V)
      DS1=S12-S11
      G1=S12*(SS-S11)/(S11*(SS-S12))
      PI1=Z1+(S11+S12)*(PHI-V)/(2.*SS)-(5.75/(P*DS1))*(ALOG(G1))
      CON1=-1.025*(S12-S11)*(SQRT(PI1))
      CCC1=-1.025*(S12-S11)*(SQRT(PI1+V))
      ZERO1=(6.2E+10)/(DS1*DS1)
      CUR1=(PI1*(EXP(CON1))-(PI1+V)*(EXP(CCC1)))*DR*R(I)
      CUR1=CUR1*ZERO1
      SUM1=SUM1+CUR1
      SUM2=SUM2+CUR2
      ARG2=SQRT(14.4*(V+PHI)*1600.0/(P*SS))
      ARG4=(SQRT(14.4*(7.+P*SS*(PHI-V))))*40./(P*SS)
      TOTAL1=EXP(ARG4)*DR*R(I)
      TOTAL2=EXP(ARG2)*DR*R(I)
      THERM1=THERM1+TOTAL1
      THERM2=THERM2+TOTAL2
    1 CONTINUE
      RATIO1=(EXP(CONST))*THERM2/THERM1
      RATIO2=SUM2/SUM1
      CURTH1=AT2*(EXP(ARG3))*THERM1
      CURTH2=AT2*(EXP(ARG1))*THERM2
      SUM2=SUM2/1.E+16
      SUM1=SUM1/1.E+16
      CURTH2=CURTH2/1.E+16
      CURTH1=CURTH1/1.E+16
      TOT2=CURTH2+SUM2
      TOT1=CURTH1+SUM1
      RATIO3=TOT2/TOT1
      WRITE(3,100)J,RATIO2,RATIO1,RATIO3,SUM2,SUM1,CURTH2,CURTH1,TOT2,
     CTOT1
```

```
      R1=V*(1.0/CURTH2+1.0/CURTH1)/2.
      R2=V*(1.0/SUM2+1.0/SUM1)/2.
      R3=V*(1.0/TOT2+1.0/TOT1)/2.
      GAMA1=((RATIO1-1.)/(2.*V))*R1
      GAMA2=((RATIO2-1.)/(2.*V))*R2
      GAMA3=((RATIO3-1.)/(2.*V))*R3
      WRITE(3,105)GAMA2,GAMA1,GAMA3
      WRITE(3,77)
   33 CONTINUE
    2 CONTINUE
    5 CONTINUE
    4 CONTINUE
      WRITE(3,88)
   90 CONTINUE
      STOP
  100 FORMAT(1X,I2,1X,9(E11.4,1X))
   70 FORMAT(1X,2HV=,E12.5)
   80 FORMAT(1X,4HPHI=,E12.5)
   20 FORMAT(2X,1HS,6X,6HRATTUN,6X,6HRATTHM,6X,6HRATCOM,5X,7HTUNNEL2,5X,
     C7HTUNNEL1,4X,8HTHERMAL2,4X,8HTHERMAL1,4X,8HCOMBINE2,4X,8HCOMBINE1)
   77 FORMAT(1X)
  105 FORMAT(4X,3(E11.4,1X))
   85 FORMAT(1X,2HA=,F9.1)
   88 FORMAT(1H1)
      END
```

Appendix III Calibration of Thermocouple and Comparison with Film-Heater

The pair of chromel-alumel ($T_1.T_2$) thermocouples used throughout the U.H.V. measurements was calibrated by means of standard temperature values, comparison being made to the same mercury-in-glass thermometer as used above, at ambient temperature and enclosed with the reference junction in a bag to exclude draughts. The standard temperatures were [110]:

(a) Melting point of (clean) mercury : 234.3°K

(b) Ice point (distilled water) : 273°K

(c) Transition of sodium sulphate decahydrate : 305.4°K

(d) Boiling point of (distilled) water : 373°K

(e) Triple-point of benzoic acid : 395.4°K

Points (b) and (d) were simply the normal steady-state conditions. Point (a) was achieved by immersing the "active" junction in mercury, freezing with liquid nitrogen and allowing to thaw. The couple was under tension and pulled out when the Hg melted. Re-immersion produced an identical potential. Points (c) and (e) were achieved by raising the materials above their melting points, immersing the couple and observing the potentials as the materials cooled. At the melting points the rate of cooling became very slow, due to the latent heat, so that the point of interest could be determined from the pronounced inflexions in the respective cooling curves.

In order to compare the results with the standard curves for this type of thermocouple [111] for which one junction is held at 0°C, the results have been plotted in figure AIII.1 with the ambient temperatures (in °C) subtracted from the standard temperatures. To determine the temperature during an experiment the ambient temperature, recorded on the same thermometer, was added to the temperature taken from the calibration curve. The agreement between the standard curve and the calibration curve was quite good, being only a few degrees different.

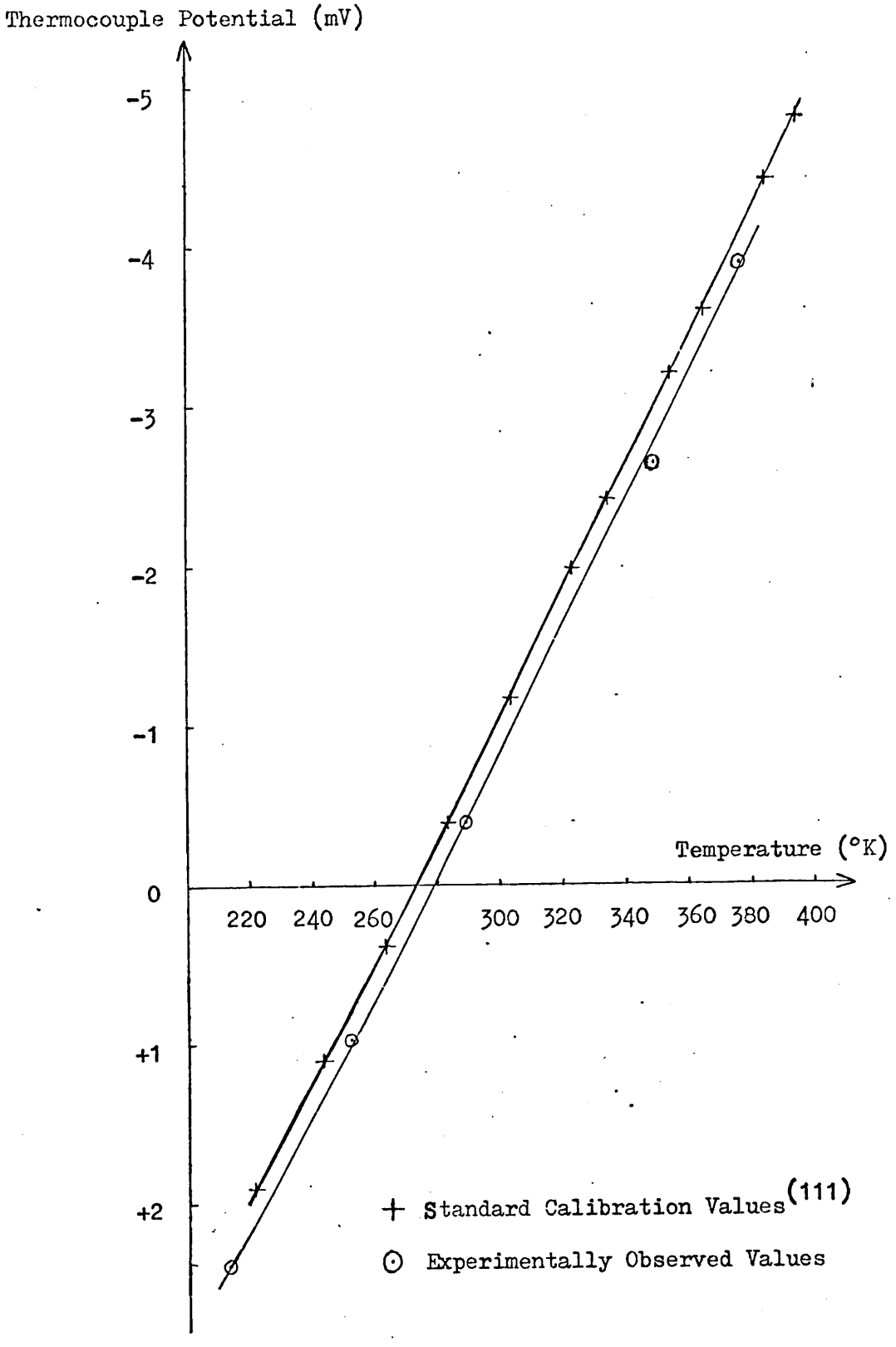


Figure AIII.1 Thermocouple Calibration Curve

The placement of the thermocouple at a small distance from the specimen was expected to cause an error in the indication of the temperature at the specimen surface. The holder used in the evaluation of the heaters was slightly different from that used in the U.H.V. and did not have provision for a thermocouple. Figure AIII.2 compares the calibration curve of the heater (from figure 5.2) with the temperature measured by the thermocouple (in the U.H.V. holder) as a function of heater power dissipation. Although the discrepancy was large, it was not critical since the experimentation which required temperature measurement was close to and encompassed the ambient temperature. Since the thermocouple would also measure the low temperature as being closer to ambient than the surface of the specimen (both above and below ambient) it would have given reliable indication that the Gd was above and below the Curie point in the respective cases. For the measurements of the higher annealing temperatures, both the thermocouple and heater-calibration temperatures have been quoted (table 6.2). Since the reproducibility of the heaters has been demonstrated, it was considered that the temperatures at the surfaces of the specimens were more likely to be close to those indicated by the heater calibration curve than by the curves of the out-of-contact thermocouple.

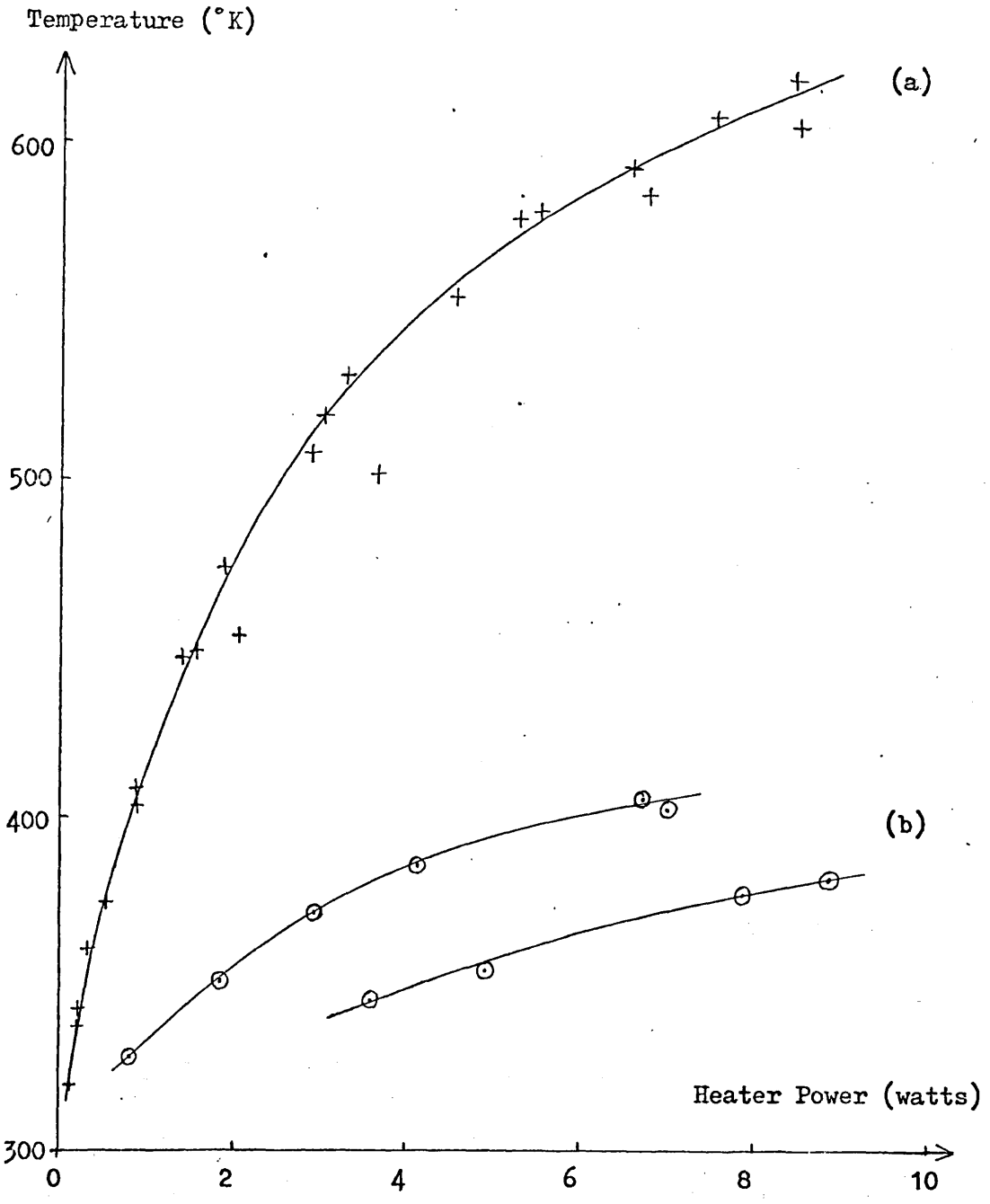


Figure AIII.2 Specimen Temperature vs. Heater Power Dissipation

(a) from heater calibration (fig. 5.2) (b) as measured by thermocouple in U.H.V. position, for various films

REFERENCES

1 RIVIERE JC in "Solid State Surface Science" Vol 1. Ed. M. Green. Ch.4. (Dekker 1969)

2 HERRING C & NICHOLS MH Rev. Mod. Phys. 21 185 (April 1949)

3 RHODIN TN & ADAMS DL "Treatise on Solid State Chemistry" vol 6A, Ed. N.B. Hannay,(Plenum Press 1976)

4 SMOLUCHOWSKI R Phys. Rev. 60 661 (1941)

5 DUBOIS E Ann. Phys. 14 627 (1930)

6 UHLIG HH J. Appl. Phys. 22 1399 (1951)

7 VAN DER ZIEL A "Solid State Physical Electronics" (Prentice-Hall 1968)

8 SEITZ F "Modern Theory of Solids" (McGraw Hill 1940)

9 CUSACK N "The Electrical and Magnetic Properties of Solids" (Longmans Green 1958) Chapters 2 to 4

10 KNAPP AG Surface Science 34 289 (1973)

11 GREINER RA "Semiconductor Devices and Applications" (McGraw Hill 1961)

12 FOWLER RH Physical Review 38 45 (1931)

13 CUSACK N "The Electrical and Magnetic Properties of Solids" (Longmans Green 1958) Chapter 5

14 FOWLER RH & NORDHEIM LW . Proc. Roy. Soc. (London) A119 173 (1928)

15 NORDHEIM LW Proc. Roy. Soc. (London) A121 626 (1928)

16 GOOD RH & MULLER EW Handbuch der Physik Vol 21 Ed. S. Flügge (Springer, 1956)

17 YELINSON MI, DOBRYAKOVA FF, KRAPIVIN VF, MALINA ZA & YASNOPOL'SKAYA AA Radio Eng. Electron (USSR) (English Translation) 6 1191 (1961)

18 HOLSCHER AA Surface Science 4 189 (1966)

19 STRAYER RH, MACKIE W, & SWANSON LW Surface Science 34 225 (1973)

20 ZISMAN WA. Rev. Sci. Instrum. 3 367 (1932)

21 DELCHAR T, EBERHAGEN A & TOMPKINS FC J. Sci.Instrum. 40 105 (1963)

22 OATLEY CW Proc. Roy. Soc. (London) A155 218 (1936)

23 SHELTON H Phys. Rev. 107 1553 (1957)

24 ANDERSON PA Phys. Rev. 47 958 (1935)

25 SIMMONS JG J. Appl. Phys. 35 2472 (1964)

26 SIMMONS JG J. Appl. Phys. 34 2581 (1963)

27 HANDY RM Phys. Rev. 126 1968 (1957)

28 KELLNER K Research 5 341 (1952)

29 MACDONALD JR & EDMONDSON DE Proc. I.R.E. 49 453 (1961)

30 DE BOER JSW, KRUSEMEYER HJ & BURHOVEN JASPERS NC Rev. Sci. Instrum. 44 1003 (1973)

31 D'ARCY RJ & SURPLICE NA J. Phys.D. Appl. Phys. 3 482 (1970)

32 HOPKINS BJ & RIVIERE JC Proc. Phys. Soc. 81 590 (1963)

33 HOPKINS BJ & RIVIERE JC Brit. J. Appl. Phys. 15 941 (1964)

34 DOBSON RJ & HOPKINS BJ Brit. J. Appl. Phys. (J. Phys.D.) Series 2 1 1241 (1968)

35 KOLM HH & FRATT GW Phys. Rev. 99 1644 (1955)

36 BLOTT BH & LEE TJ J. Sci. Instrum. (J.Phys.E) Series 2 2 785 (1969)

37 ANTES LL & HACKERMAN N J. Appl. Phys. 22 1395 (1951)

38 ANON Technical Literature no. 395/1/1971 "Tungsten" (Metalwerk Plansee A.G. & Co. Reutte, Austria)

39 TAYLOR AJ "Some Optical and Magneto-optical Studies on Thin Absorbing Films" PhD Thesis Univ. London 1972

40 ROBINSON NW "The Physical Principles of Ultra-High Vacuum Systems and Equipment" (Chapman and Hall, London 1968)

41 TEGART WJM°G "The Electrolytic & Chemical Polishing of Metals in Research and Industry" (Pergamon 1956)

42 CRAIG RD & HARDEN EH "The Interpretation of Mass Spectra in Vacuum Measurement" Vacuum 16 67 (1965)

43 Private communication with A.E.I. Mass Spectrometer engineers

44 TOLANSKY S "An Introduction to Interferometry" (Longmans 1973) Chapter 10

45 HOLLAND L "Vacuum Deposition of Thin Films" (Chapman and Hall, London 1966) Chapter 5

46 Supplied by Headland Engineering Developments Ltd., Melon Road, London S.E.15, from whom technical literature is available

47 BELGRAVE M 3rd Year undergraduate project. Royal Holloway College (1974) (Unpublished)

48 LEGVOLD S in "Magnetic Properties of Rare Earth Metals" Ed. R.J. Elliott (Plenum Press London 1972)

49 SCHULER C Zeitschrift für angewandte Physik 15 218 (1963)

50 LODGE FMK PhD Thesis Univ.Durham (1969)

51 BIST BMS & SRIVASTAVA ON Thin Solid Films 24 137 (1974)

52 REALE C Vacuum 25 403 (1975)

53 GERRITSEN AN Handbuch der Physik Ed. S. Flügge vol.19 (Springer Berlin 1956)

54 COLVIN RV & ARAJS S Phys. Stat. Sol. 4 37 (1964)

55 MULLER J Surface Science 42 525 (1974) and Surface Science 49 681 (1975)

56 SUHRMANN R & WEDLER G Z. für angew. Phys. 2 70 (1962)

57 ECKSTEIN W, GEORG KF, HEILAND W, KIRSCHNER J & MULLER N Z. Naturforsch 25a 1981 (1970)

58 SURPLICE NA, MULLER J & SINGH B Thin Solid Films 28 179 (1975)

59 CROWELL AD Surface Science 49 649 (1975)

60 NIEUWENHUYS BE, BOUWMAN R & SACHTLER WMH Thin Solid Films 21 51 (1974)

61 CRAIG PP & RADEKA V Rev. Sci. Instrum. 41 258 (1970)

62 MINTS RI, MELEKHIN VP & PARTENSKII MB Sov. Phys. Solid State 16 2330 (1975)

63 RAMEY RL & KATZBERG SJ J. Chem. Phys. 53 1347 (1970)

64 GORDY W & THOMAS WJ J. Chem. Phys. 24 439 (1956)

65 GORDY W J. Chem. Phys. 22 1470 (1954)

66 ELLIOTT RJ in reference 48

67 SCHUMAKER EE & HARRIS JE J. Am. Chem. Soc. 48 3108 (1926)

68 RENTSCHLER HC, HENRY DE & SMITH KO Rev. Sci. Instrum. 3 794 (1932)

69 SAHNI V, KRIEGER JB & GRUENEBAUM J Phys. Rev. B. 155 1941 (1977)

70 LANG ND & KOHN W Phys. Rev. B. 3 1215 (1971)

71 FREEMAN AJ in reference 48

72 SLATER JC "Quantum Theory of Molecules and Solids" Vol.2 "Symmetry and Energy Bands in Crystals" (McGraw Hill 1965) pages 171-3

73 KEETON SC & LOUCKS TL Phys. Rev. 168 672 (1968)

74 INWOOD RB PhD Thesis Univ. London (1978) (Unpublished)

75 PORSCH G & HOLZL J Thin Solid Films 34 235 (1976)

76 CARDWELL AB Phys. Rev. 76 125 (1949)

77 COMSA G, GELBERG A & IOSIFESCU B Phys. Rev. 122 1091 (1961)

78 CHRISTMANN K, ERTL G & SCHOBER O Surface Science 40 61(1973)

79 BLODGETT AJ, SPICER WE & YU A Y-C "Optical Properties and Electronic Structure of Metals" Ed. F. Abeles (North-Holland, Amsterdam 1966)

80 WONSOWSKI SV, SOKOLOV AV & WEXLER AS Fortschritte der Physik 4 216 (1956)

81 PANT MM & RAJAGOPAL AK Solid State Communications 10 1157 (1972)

82 BETTERMAN N Z. Naturforsch 31a 1424 (1976)

83 PIERCE DT & SPICER WE Phys. Rev. Lett. 25 581 (1970)

84 EASTMAN DE Phys. Rev. B. 2 1 (1970)

85 BUSCH G, CAMPAGNA M, GOTTI D & SIEGMANN HC Phys. Rev. Lett. 22 597 (1969)

86 FRAYNE PG, CHANDLER N & BOOTON MW J. Phys. D: Appl. Phys. 11 2391 (1978)

87 HOCKER LO, SOKOLOFF DR, DANEU V, SZOKE A & JAVAN A Appl. Phys. Letts. 12 410 (1968)

88 DANEU V, SOKOLOFF DR, SANCHEZ A & JAVAN A Appl. Phys. Letts. 15 398 (1969)

89 MATARESSE LM & EVENSON KM Appl. Phys. Letts. 17 8 (1970)

90 PUTLEY EH Proc. Phys. Soc. 76 802 (1960)

91 GREEN SI J. Appl. Phys. 42 1166 (1971)

92 BECKLAKE EJ, PAYNE CD & PREWER BE J. Phys. D: Appl. Phys. 3 473 (1970)

93 GREEN SI, COLEMAN PD & BAIRD JR Proc. Submillimetre Wave, Microwave Res. Inst. Symp. Vol.10 (Polytechnic press of Polytechnic Institute of Brooklyn (1970) p.369

94 MARHIC ME & JASSBY DL Report UCIA 34P157, PELR3, UCLA ENG 7204 prepared for U.S.Atomic Energy Commission, under contract no. AT(04-3) - 34 (1972)

95 FARIS SM, GUSTAFSON TK & WIESNER JC I.E.E.E. J. Quantum Electronics 9 737 (1973)

96 JAVAN A "Technology and Physics of Infrared and point-contact diodes." M.I.T. Technical report AFCRL TR - 75 - 0496 (1975)

97 KWOK SP, HADDAD GJ & LOBOV G J. Appl. Phys. 42 554 (1971)

98 SMITH RA, JONES FE & CHASMAR RP "The Detection and Measurement of Infra-red Radiation" 2nd Edn. (Clarendon Press, Oxford 1968) Chapter V

99 FOMENKO VS "Handbook of Thermionic Properties" Ed. G.V. Samsonov (Plenum N.Y. 1966)

100 THANAILAKIS A Conference on "Metallic Semiconductor contacts" U.M.I.S.T. April 1974

101 SIMMONS JG J. Appl. Phys. 34 1793 (1963)

102 POLLACK SR & MORRIS CE J. Appl. Phys. 35 1503 (1964)

103 KAYE GWC & LABY TH "Tables of Physical & Chemical Constants" Longmans)

104 CHOW CK J. Appl. Phys. 34 2599 (1963)

105 CHEN TT & ADLER JG Solid State Communications 8 1965 (1970)

106 PITT CW Microelectronics and Reliability, 9 239 (1970)

107 WOLFRAM R, DEWAMES RE, HALL WF & PALMBERG PW Surface Science 28 45 (1971).

108 HOLZL J & SCHRAMMEN P Applied Physics 3 353 (1974)

109 FARIS SM & GUSTAFSON TK J. Vac. Sci. Technol. 12 1356 (1975)

110 ROBERTS JK & MILLER AR "Heat and Thermodynamics" (Blackie 1960) page 29

111 Chemical Rubber Company Handbook (1977)

J. Phys. D—D6676—P G Frayne N Chandler and M W Booton—PAPER—
J. Phys. D: Appl. Phys., Vol. 11, 1978. Printed in Great Britain

Point-contact diode submillimetre detectors

P G Frayne, N Chandler† and M W Booton‡
Physics Department, Royal Holloway College, Egham, Surrey

Received 30 May 1978

Abstract. Point-contact whisker diodes have been used to detect 337 μm radiation giving responsivities up to 1000 V W^{-1}. The respective merits of metal–insulator–metal and metal–insulator–semiconductor diodes are examined. To explain the characteristics of responsivity with electrical and mechanical variations, a model is presented based on thermionic and tunnel currents flowing between a hemispherical and a plane electrode.

1. Introduction

The use of point-contact diodes as detectors in the submillimetre region has some advantages over other infrared detectors. Unlike the bulk InSb detector (Putley 1960), liquid helium is not required, and because of the low capacitance of its junction a point-contact diode has an inherently fast response, in contrast to room-temperature thermal detectors such as the Golay cell. The ability to rectify at far IR frequencies was demonstrated by Frayne (1967) with commercially produced diodes. However, these and devices made by contacting a pointed wire antenna on to a metal or semiconductor surface (Daneu *et al* 1969, Matarrese and Evenson 1970, Becklake *et al* 1970) possess lower responsivities than more conventional detectors. For this reason a programme was undertaken to produce a device with increased sensitivity, and to propose a model of the mechanism taking place at a point-contact to aid future development.

2. Experimental details

Diode detectors were produced by electrolytically etching to a point one end of a straight length of 20 μm diameter tungsten wire, and bringing this point into contact with the atmospherically contaminated surface of a metal or semiconductor. By orienting the tungsten wire antenna in the beam of an HCN laser operating at 891 GHz and by suitably applying pressure to the contact, a DC potential was generated across the junction. For both metal and semiconductor devices it was found that the greatest output from the diode was obtained on initial contact between the tungsten and the insulating layer. Increasing the pressure on the wire resulted in reduction of the responsivity, with ultimate reversal of the polarity of the output voltage. The output obtained, once reversal had taken place, was less than that produced on initial contact. The high responsivity could

† Present address: Marconi Research Laboratories, GEC-Marconi Electronics Ltd, Great Baddow.
‡ Present address: GEC Hirst Research Centre, Wembley.

0022-3727/78/0014-2000 $02.00

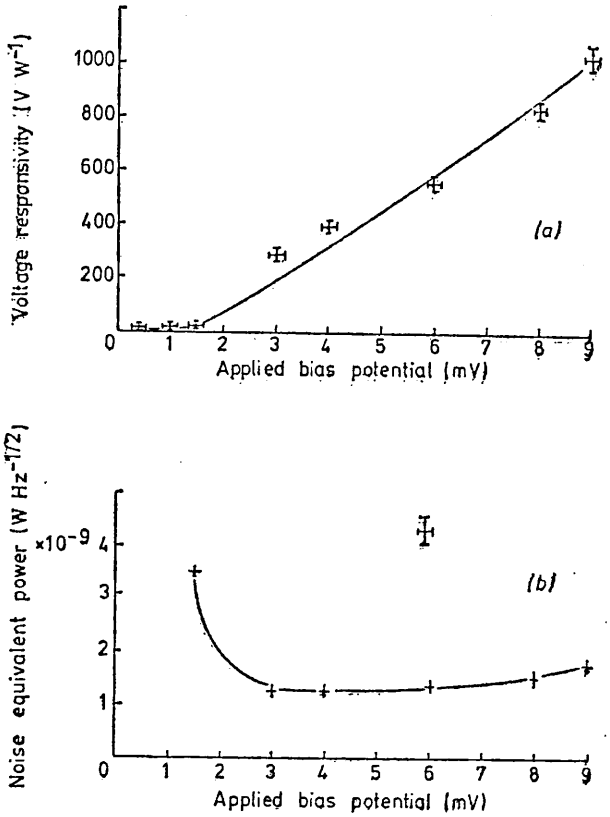


Figure 4. Tungsten–GaAs diode. (*a*) voltage responsivity versus applied bias potential, (*b*) Noise equivalent power versus applied bias potential. Crossed bars indicate uncertainty.

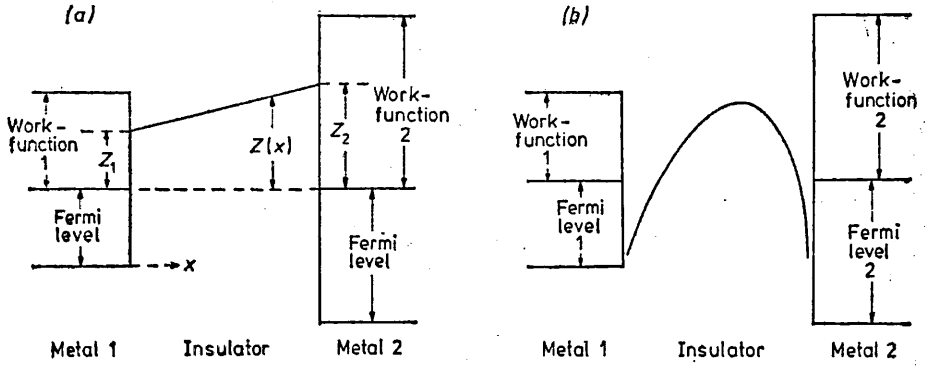


Figure 5. Energy diagram of potential barriers between dissimilar electrodes: (*a*) image forces neglected; (*b*) image forces included.

the current density for the opposite sense of applied voltage, then the rectification ratio of the diode may be expressed as

$$\eta = \frac{\int_0^R 2\pi J_2(r) r \, dr}{\int_0^R 2\pi J_1(r) r \, dr}$$

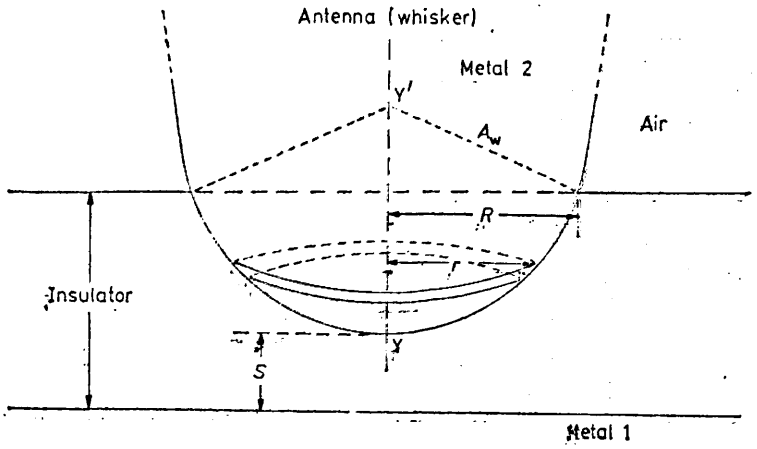


Figure 6. Diagrammatic representation of a hemispherical whisker tip partially embedded in the insulating surface layer of a plane electrode.

In this equation, r represents the perpendicular distance from the line of symmetry YY' in figure 6 to the circumference of the sphere and $r=R$ at the insulator–air interface. Using Simmons' equations for the current density for tunnel (1963b) and thermionic (1964) conduction, the above expression for η has been evaluated numerically. Similarly, the resistance of the barrier R_B has been calculated so that the voltage responsivity expressed by Green (1971) as

$$\gamma = \frac{(\eta-1)\,R_B}{2V}$$

may be plotted against the minimum electrode spacing S. Values of the current flow between the electrodes used to calculate η and R_B were the summation of the currents arising from the tunnel and thermionic effects. The responsivity curves displayed in figure 7 show a polarity reversal due to a change in dominance from tunnel currents at small electrode spacing to thermionic processes at higher S. The figure also shows the dependence of γ on the radius A_w of the whisker tip, on the detected voltage V_D and on the minimum barrier height Z_1. For clarity, the curves for the lowest chosen value of Z_1 are shown separately in figure 8, with curve F of figure 7 included to facilitate comparison. The higher values of V_D have been omitted to comply with the theoretical assumption that $eV \leqslant Z_1$. The difference between the work-functions of the two electrodes and hence between the heights Z_1 and Z_2 of the barriers at the metal–insulator interfaces has been taken as 1 eV for these examples.

4. Discussion

The theoretical model shows qualitative agreement with the experimental observation that the voltage responsivity of point-contact diodes is greatest at large electrode separation and then decreases and finally reverses polarity as the spacing is reduced.

The poor quantitative agreement may be attributed to the imperfect efficiency of the antenna and to the simplifications in the model. It has been shown (Chow 1963) that local perturbations in the insulator thickness can change the resistance by several orders of magnitude. Imperfections in the dielectric would also reduce the resistance significantly

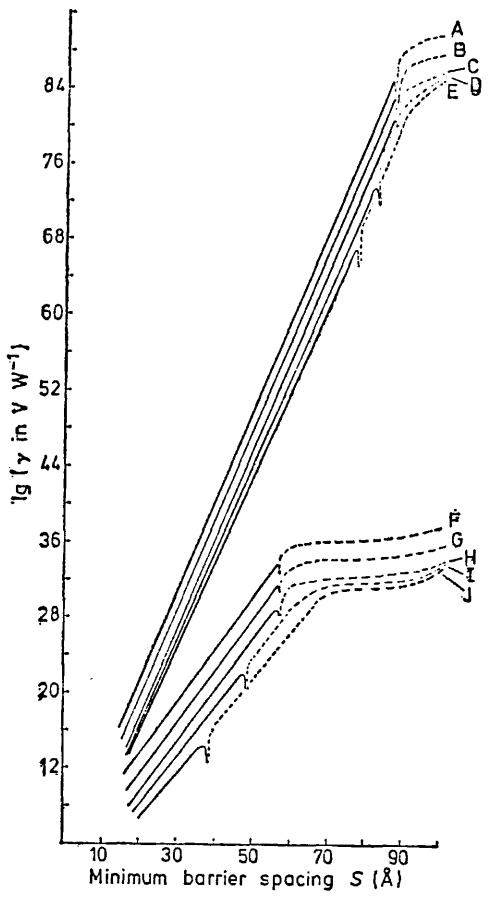


Figure 7. Theoretical voltage responsivity of a MIM point-contact diode detector. The dielectric constant of the barrier is 8; ——— positive γ, — — — negative γ.

Curve	A	B	C	D	E	F	G	H	I	J
V_D(V)	0·09	0·09	0·09	0·5	0·9	0·09	0·09	0·09	0·5	0·9
A_w (Å)	10^2	10^4	10^6	10^6	10^6	10^2	10^4	10^6	10^6	10^6
Z_1 (eV)	4	4	4	4	4	1	1	1	1	1
Z_2 (eV)	5	5	5	5	5	2	2	2	2	2

(Chen and Adler 1970). Since it is possible to produce rectification with nominally symmetrical devices (Pollack and Morris 1964) there may also be an effect due to an asymmetry caused by, for example, a poorly defined transition from metal to dielectric at one or both interfaces, resulting in reductions of barrier heights and of effective barrier thickness.

The significant factor is that the experimental characteristics are both reversible and repeatable, indicating a mechanistic rather than a random variation. Green (1971) has expressed the current–voltage characteristics of a MIM diode as

$$I=aV+bV^2+\ldots$$

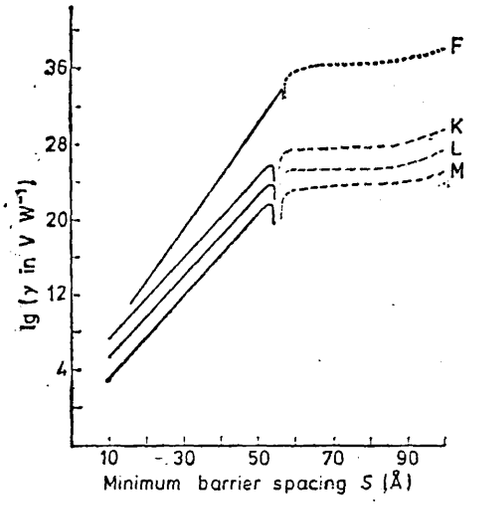


Figure 8. Theoretical voltage responsivity of a MIM point-contact diode detector. The dielectric constant of the barrier is 8; $V_D = 0{\cdot}09$ V; ——— positive γ, — · — · — negative γ.

Curve	F	K	L	M
A_w (Å)	10^2	10^2	10^4	10^6
Z_1 (eV)	1	0·5	0·5	0·5
Z_2 (eV)	2	1·5	1·5	1·5

where a and b are constants, and the instantaneous voltage across the diode is

$$V=(2PR_B)^{1/2}\sin\omega t=\left(2V_D\frac{a}{b}\right)^{1/2}\sin\omega t$$

where V_D can be seen to depend on the RF power P coupled into the device. Figure 7 also indicates that at a fixed dielectric thickness, variation of and, at some barrier spacings, reversal of the polarity of the responsivity occurs with changing V_D and hence incident power. This property has been demonstrated experimentally by Green *et al* (1970).

The high impedance of the MIS diodes when compared to the MIM structures indicates a relatively slow response. Assuming a typical device to have a junction diameter of 0·1 μm (Green 1971), an insulating layer of dielectric constant 8 (Pollack and Morris 1964) and a thickness 50 Å, then the measured impedances of the MIM diodes in table 3 indicates that the upper 3 dB output point for such devices can be between 45 and 100 THz. For a tungsten–silicon diode with a similar barrier resistance, a series resistance approximately equal to the output impedance indicated in table 3 and a dielectric constant of 4·5 (Kaye and Laby 1966), then the upper 3 dB point is in the region of 15 THz. Although the use of MIS detectors at higher frequencies is restricted by the limited bandwidth, figure 1 shows that they have a distinct advantage over MIM devices at 337 μm by possessing greater responsivities.

The theoretical model of the point-contact diode has been developed for an MIM barrier. However, it is interesting to note that the performance characteristics of the tungsten–semiconductor diodes with variation of whisker pressure are similar to that of the tungsten–metal diodes. A possible explanation is that the mechanism described by the MIM model contributes to a DC bias voltage across the depletion layer near the semi-

conductor surface. The polarity of this component will vary accordingly and the rectification ratio of the semiconductor device will be modified, becoming greater with forward bias and smaller with reverse. The MIS devices may therefore be expected to behave in a similar qualitative manner to the MIM diodes but to exhibit higher rectification ratios, and therefore higher responsivities as observed experimentally.

6. Conclusions

The experimental behaviour of MIM point-contact diode detectors can be explained in terms of thermionic and electron tunnelling processes taking into account the geometrical shape and dimensions of the insulating barrier between the antenna tip and the base electrode.

Responsivities of 1000 V W^{-1} have been obtained at 337 μm from biased tungsten–GaAs detectors. However, the inherent high resistivities of semiconductor devices indicate that for the highest frequencies MIM diodes must be used and these may operate up to 100 THz.

Acknowledgments

The authors wish to thank the Science Research Council for a support grant. The investigation was undertaken as partial fulfilment for higher degrees while two of the authors (NC and MWB) were in receipt of research studentships.

References

Beckiake E J, Payne C D and Prewer B E 1970 *J. Phys. D: Appl. Phys.* **3** 473–81
Chen T T and Adler J G 1970 *Solid St. Commun.* **8** 1965
Chow C K 1963 *J. Appl. Phys.* **34** 2599–602
Daneu V, Sokoloff D, Sanchez A and Javan A 1969 *Appl. Phys. Lett.* **15** 398
Frayne P G 1967 *Electron. Lett.* **3** 338–9
Green S J 1971 *J. Appl. Phys.* **42** 1166
Green S J, Coleman P D and Baird J R 1970 *Proc. Submillimetre Wave, Microwave Res. Inst. Symp.* Vol 10 Polytechnic press of Polytechnic Institute of Brooklyn N.Y.
Haas G A and Thomas R E 1966 *Surface Sci.* **4** 64–88
Kaye G W C and Laby T H 1966 *Tables of Physical and Chemical Constants* (London: Longmans)
Matarrese L M and Evenson K M 1970 *Appl. Phys. Lett.* **17** 8–16
Pollack S R and Morris C E 1964 *J. Appl. Phys.* **35** 1503
Putley E H 1960 *Proc. Phys. Soc.* **76** 802–5
Riviere J C 1969 *Solid State Surface Science* ed M Green (New York: Dekker)
Simmons J G 1963a *J. Appl. Phys.* **34** 1793–803
—— 1963b *J. Appl. Phys.* **34** 2581–90
—— 1964 *J. Appl. Phys.* **35** No. 8, pp 2472–81
Surplice N A and D'Arcy R J 1970 *J. Phys. E: Sci. Instrum.* **3** 477
Zisman W A 1932 *Rev. Sci. Instrum.* **3** 367

be recovered, however, as the output voltage was reproducible with cyclic variation of the contact pressure.

Although some exceptions did occur, making contact on to various samples of a particular material usually resulted in the same polarity of output on initial contact. The experimental observation of polarity reversal by Green *et al* (1970) was explained by variation of the difference between the work-functions of the two electrodes as the tip of the antenna contacted different areas of the substrate. To investigate this situation, tungsten ribbon and evaporated films of gold and aluminium were used as plane metal electrodes. By virtue of the large surface areas, the Kelvin vibrating capacitor technique (Zisman 1932, Surplice and D'Arcy 1970) was used to measure the average value of the work-function of each polycrystalline film. Measurements were also made on silicon, but the samples of the other semiconductors investigated, germanium and gallium arsenide, were not large enough for the method employed.

An indication of the possible deviation of the work-function over the surface of a material has been given by Haas and Thomas (1966). This variation was explained by the metal surface being composed of regions of different crystal orientation. The work-function of the point of the tungsten wire was unknown. However, the measured values for the plane electrodes given in table 1 are sufficiently close to the documented values for the different crystallographic orientations of tungsten in table 2 (Riviere 1969), that

Table 1. Measured average work-functions of substrate materials.

Substrate material	Work-function (eV)
Silicon	4·85±0·02
Gold	5·39±0·03
Aluminium	4·26±0·02

Table 2. Work-functions of crystal orientations in tungsten (after Riviere 1969).

Tungsten orientation	Work-function (eV)
110	5·3±0·12
111	4·4±0·03
116	4·3±0·03
100	4·6±0·08

polarity reversal of the work-function difference was experimentally possible. The situation therefore exists in which successive contacts of a whisker could yield devices with a range of work-function differences, due either to the patchy nature of the plane electrodes or to different crystal faces of the tungsten wire being exposed, either by repeated etching or by mechanical deformation. However, the repeatability and reversibility of the variation of the responsivity with applied pressure implied that, once contact had been made, the patches were large enough and the antenna drive accurate enough to eliminate random work-function changes through lateral movement of the whisker tip. The voltage responsivities of the various detectors were found after optimisation, by dividing the DC output voltage from the diode by the CW power incident on the antenna, measured using a calibrated Golay cell.

The responsivities of detectors formed by contacting tungsten antennae on to aluminium, gold, tungsten and doped silicon and germanium electrodes are shown in figure 1 as a function of the resistive load. The variation of detector performance against load was investigated by placing a shunt resistor across the diode. This is a useful method for reducing the response time of the output circuit. The output impedances derived from the curves given in table 3 are in good agreement with the zero-bias resistances estimated from the current–voltage characteristics of figure 2. These curves, which were measured by applying short pulses of current to the diode from a differentiated 900 Hz square wave, in order to limit power dissipation, show reasonable asymmetry between the forward and reverse bias conditions for the tungsten–semiconductor devices. However, only a small

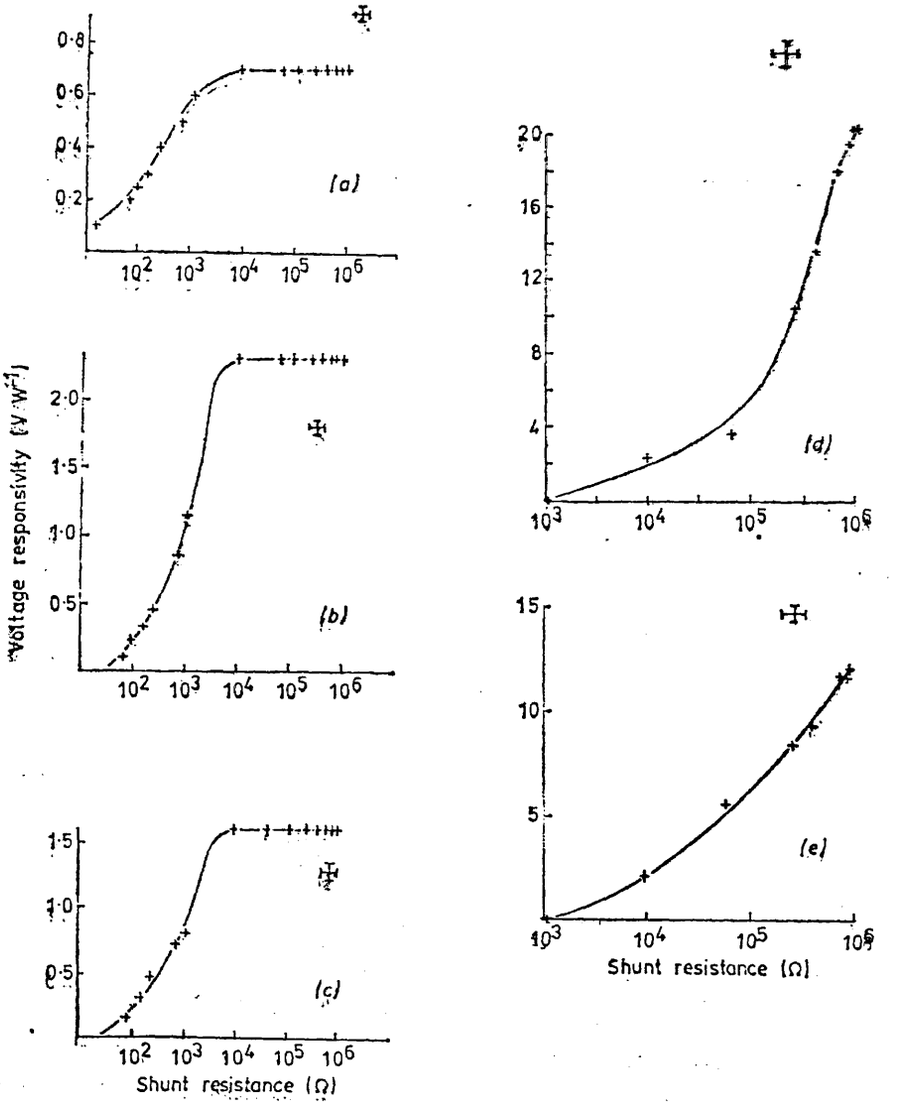


Figure 1. Experimentally observed voltage responsivities of point-contact diode detectors: tungsten to (*a*) gold, (*b*) aluminium, (*c*) tungsten, (*d*) germanium, (*e*) silicon. Crossed bars indicate uncertainty.

Measured zero-bias resistance and output impedance of point-contact diodes.

Table 3.

Diode	Zero-bias resistance (Ω)	Output impedance (Ω)
Tungsten–silicon	$(2{\cdot}3\pm0{\cdot}3)\times10^{4}$	$(7{\cdot}8\pm0{\cdot}1)\times10^{5}$
Tungsten–germanium	$(4{\cdot}5\pm0{\cdot}3)\times10^{4}$	$(3{\cdot}2\pm0{\cdot}1)\times10^{5}$
Tungsten–gold	220 ± 25	200 ± 15
Tungsten–aluminium	930 ± 65	1100 ± 100
Tungsten–tungsten	770 ± 30	900 ± 30

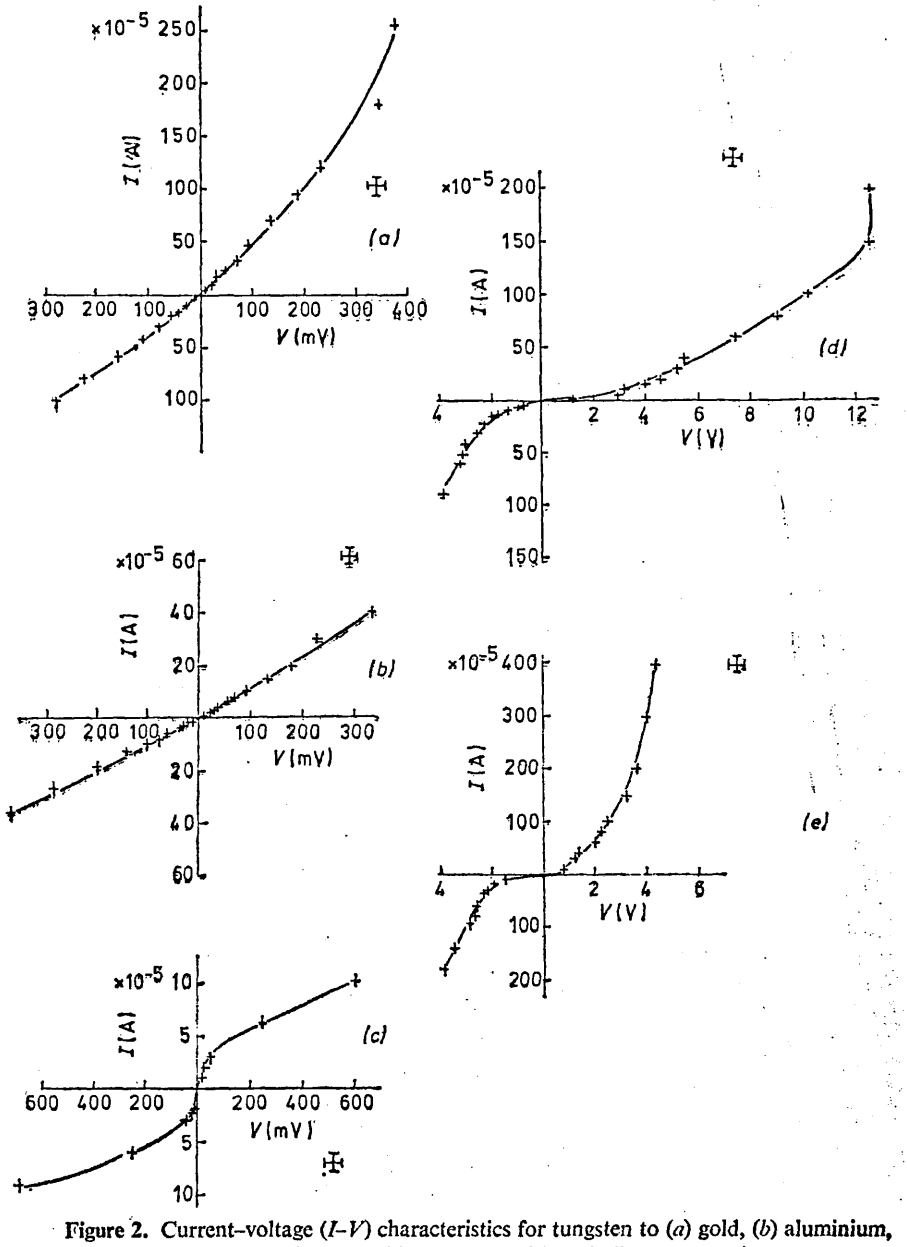


Figure 2. Current–voltage (I–V) characteristics for tungsten to (*a*) gold, (*b*) aluminium, (*c*) tungsten, (*d*) germanium, (*e*) silicon. Crossed bars indicate uncertainty.

difference is observed for the tungsten–metal diodes, hence the poorer responsivity. The nearly linear characteristic of the metal–insulator–metal (MIM) diodes also explains the observation that enhancement of the detector output was not achieved by applying DC bias to the diode.

In contrast, the metal–insulator–semiconductor (MIS) detectors showed a useful increase in responsivity with, as shown in figure 3, the application of the appropriate bias

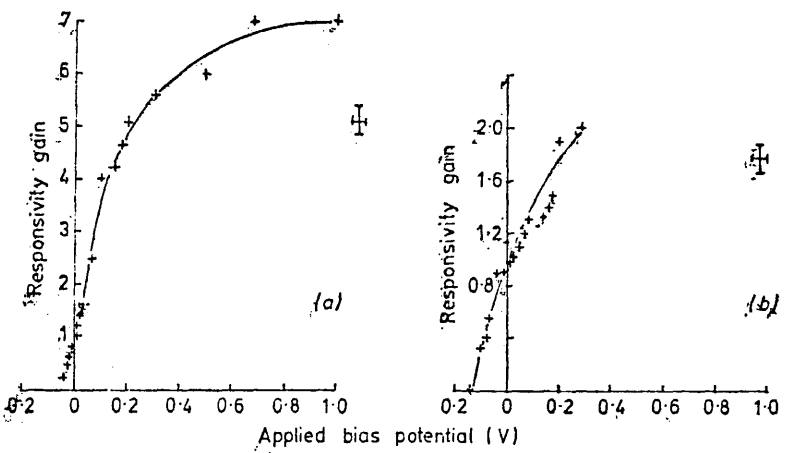


Figure 3. Responsivity gain versus applied bias for (*a*) tungsten–germanium, (*b*) tungsten–silicon. Crossed bars indicate uncertainty.

potential. The most notable result was produced by contacting a tungsten antenna on to a crystalline sample of undoped GaAs. This diode produced no rectified output in the unbiased state, but with the application of a small DC potential, a responsivity was obtained which was far greater than observed in either silicon or germanium. Figure 4 shows the response of this detector as a function of applied bias, together with the noise equivalent power. It was only with this tungsten–GaAs diode that the noise was sufficiently large to be measured without additional preamplification.

3. Theory

The simple potential distribution existing between two plane parallel electrodes separated by an insulating material and connected by an external circuit is represented diagrammatically by figure 5(*a*). Simmons (1963a,b, 1964) has explained current flow through such a structure in terms of thermionic and electronic tunnelling mechanisms, while taking into account distortion of the field by image forces which produce a potential barrier of the type illustrated in figure 5(*b*). The operation of a point-contact diode cannot adequately be explained using the analysis derived for two plane parallel electrodes, for Chow (1963) has shown that a non-uniform spacing can significantly alter the current–voltage characteristic. The diode will be represented by a planar electrode coated with a uniform dielectric layer, into which is partially inserted the sharpened hemispherical tip of the antenna. This geometry is shown in figure 6 where the interface between the plane metal and the insulator is clearly defined. Using the convention of Simmons (1963b) that J_1 represents the current density when the electrode of lower work-function is negatively biased and J_2