ELECTRICAL CONDUCTION IN

SINGLE CRYSTALS OF ALUMINA

by

H. Mümtaz KIZILYALLI, M.E.E.

A thesis submitted for the
Degree of Doctor of Philosophy
in the University of London

Department of Electrical Engineering,
Imperial College of Science and Technology,
London

May 1974

Dedicated

to

my wife and children

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to Mr. P.R. Mason for invaluable supervision, criticism and encouragement during this project.

Also I would like to thank Professor J.C. Anderson for valuable guidance and for the facilities of the Materials Section of the Electrical Engineering Department of Imperial College, and Dr. E.A.D. White for providing sapphire platelets.

I should like to thank Dr. B.C.H. Steele and Dr. R.J. Brook for helpful discussions.

I am indebted and most grateful to The Ford Foundation, UNESCO and the Atomic Energy Authority (U.K.) for providing grants; to the Middle East Technical Univeristy of Ankara for granting leave of absence throughout my studies, and to my programme organizers Miss E. Czaykowska and Mrs. Jane Jones in the British Council for their help in different matters.

Thanks are due to Miss Simone Ruddy for typing the thesis.

Finally, I should like to thank my wife for her patience and understanding.

ABSTRACT

A study of the phenomenology of electrical conduction in single crystal alumina platelet and boule slices is described. Crystals of 8-260 microns thickness, flux and Czochralski grown, and contacted with sputtered platinum electrodes were investigated.

A series of electrical properties of these crystals was investigated over the temperature range 600°-$1500^{\circ}C$ and an oxygen partial pressure range from 1 to $2x10^{-2}$ atm. with d.c. and a.c. methods in order to understand the conduction mechanisms. The results are mainly exhibited as plots of log $\sigma_{d.c.}$ versus $^{1}/T$ at p_{O_2} = const., log $\sigma_{d.c.}$ versus log p_{O_2} at T = const., and log $\sigma(f)$ and log $\epsilon(f)$ versus log f with T and p_{O_2} constant.

The a.c. characteristics are interpreted in terms of an electrode effect and a bulk effect. The electrode effects are analysed in terms of a two layer model, and, with the help of the Kramers-Kronig relations, approximate values of the surface layer thickness and conductivity are determined. The bulk effect is analysed by a new theory, which considers the motion of a charged defect in a periodic potential due to the lattice, on which is modulated a longer period potential due to interaction with the oppositely charged defects, considered as largely immobile.

The d.c. conductivity dependence on oxygen partial pressure and temperature is compared in detail to previous results and interpretations.

CONTENTS

CHAPTER 1 - INTRODUCTION

Alumina crystals in the form of ruby have important applications in masers and lasers and as substrates in microelectronic devices and also as microwave windows in microwave devices.

They have extremely good electrical properties. However, there is no general agreement in the literature as to some of the values of their physical parameters, or which are the most important conduction processes.

One is also interested in determining the conduction mechanisms involved in single crystal alumina as a basis to understanding polycrystalline alumina. This is another very important material technologically being used for high temperature and high voltage work. Polycrystalline material will show the processes present in single crystals and also others due to grain boundaries and dislocations.

Conduction processes may be classified as intrinsic or extrinsic; electronic or ionic; if electronic, band or hopping type.

The published literature still does not allow unambiguous decisions as to which of these processes are predominant at different temperature ranges in single crystal Al_2O_3, and a main objective of the work described here is to contribute to resolving these questions.

Effects at the contacts must also be carefully considered in deriving bulk parameters from measurements. It does not seem possible, for various reasons, to eliminate them with certainty by d.c. experimental techniques (e.g. 4- terminal measurements). Two main features of this work have, therefore, been to exploit a.c. measurements of conductivity and permittivity to try to separate bulk and electrode effects, and to make a more thorough theoretical analysis

of the relationship of a.c. properties to conduction processes.

A necessary preliminary to the study of electrical conduction phenomena in insulating materials like sapphire is a discussion of point defects. In a defect free insulator (i.e. a pure, stoichiometric single crystal with a perfect lattice), the electrons may be considered as either bound to the individual atoms so having localized wave function or else as located in a filled electron band, having distributed or Bloch type wave function. There is no mechanism for ionic species or electrons to diffuse and hence no current can flow. For ionic conduction there must be defects and defect migration. Point defects are responsible for lattice diffusion which is also often termed volume or bulk diffusion.

In chapter 2, theoretical discussion shows that the ionic conductivity of Al_2O_3 should depend on the ambient atmosphere and temperature.

The principles of calculating the energy levels of the charge carriers and the dependence of their densities on atmosphere and temperature are discussed. Later macroscopic relations between the polarization and the permittivity and conductivity in a sinusoidal field, and the idea of complex permittivity are discussed and then the relationship between the time response to a step electric field and the variation with frequency of the response to a sinusoidal field is developed. Finally the well known simple "double potential well" model and "simply periodic multiple potential well" model are reviewed and a new "multi-pair potential well" model (Mason, to be published) is presented.

In chapter 3, the design of the apparatus and the experimental procedure are discussed. The resistivity of Al_2O_3 is very high and ambient dependent as just mentioned. These are the factors to be considered in the design of the sample holder in which conductivity will be measured as a function of temperature and oxygen partial

pressure at d.c. and varying frequency. It must also avoid erroneous electrical measurements due to leakage paths and other effects.

As this was a new field of research in the Electrical Engineering Department at Imperial College, all the required equipment had to be either constructed or bought before any conductivity measurements could be made. A great deal of time was spent on this preparatory work, which is described in detail in the following chapters.

In Chapter 4 the results are presented. Measurements in the c-direction only of conductivity σ and permittivity ϵ were made on boule sapphire slices and also on flux grown platelets over a temperature range of 700°C to 1500°C, an oxygen partial pressure range from 1 to 2×10^{-2} atm., at d.c. and at a.c. over a frequency range from 70 Hz to 5 MHz.

The results are chiefly exhibited as plots of log $\sigma_{d.c.}$ vs $1/T$ at p_{O_2} = const., log $\sigma_{d.c.}$ vs log p_{O_2} at T = const., and log $\sigma(f)$ and log $\epsilon(f)$ versus log f with T and p_{O_2} constant.

In Chapter 5 the results are discussed in comparison to previous work and interpretations and in the light of the new theory described in Chapter 2.

Final conclusions are drawn in Chapter 6.

In the remainder of this chapter a brief outline of previous work will be given, but critical discussion of its significance will be left until Chapter 5.

A review of the work up to 1959 was given by Cohen (1959) who commented then upon the wide disparities between the results of different workers.

Pappis and Kingery (1961) measured conductivity from 1300°C to 1750°C at p_{O_2} values from 1 to 10^{-10} atm., mostly at 10^4 Hz, but also with two and four-terminal d.c. methods. They claimed little difference

between the results of the different methods at the same temperature and pressure. They measured the thermoelectric effect at 1500°C.

Champion (1964) investigated the conductivity of sapphire with varying chromium content below 850°C in air employing the two-terminal d.c. method which is especially susceptible to surface effects. The same remarks apply to the four-probe method of Dasgupta and Hart (1965). Harrop and Creamer (1963) used a two terminal method to measure d.c. conductivity of pure and chromium doped single crystal alumina between 800°C and 1500°C, in an unspecified atmosphere, presumably air.

Matsumura (1966) studied the conductivity and the ionic transport number at p_{O_2} = 0.2 atm. (air) from 1000°K to 1750°K. He used a three-terminal method for conductivity, but without a physical gas phase guard, and galvanic cell e.m.f. measurements with a p_{O_2} difference across the crystal for transport numbers.

Peters et al (1964) demonstrated that for high resistance materials at temperatures above 1100°C the conductivity of the gas phase around the sample can be comparable to or greater than that of the sample. Moulson and Popper (1968) have also found gas phase conduction important at relatively low temperatures.

Özkan and Moulson (1970) devised a three-terminal apparatus incorporating a physical gas guard which makes the electrical leakage path in the gas phase pass through the cool region of the sample holder envelope. All work mentioned here subsequent to that date, including our own, has used a similar arrangement. They made d.c. measurements in dry air in the temperature range 600-1573°K.

Brook et al (1971) made d.c. measurements on undoped and doped single crystal alumina applying a voltage less than the decomposition voltage in the range 1000°C to 1600°C in air.

Yee and Kröger (1973) showed that the ionic transference number

of single crystals of Al_2O_3 is unity up to 1450°C and at higher temperatures t_i decreases slightly, reaching a value of 0.8 at 1650°C. They measured the e.m.f. of a concentration cell with pure O_2 and air on the two sides and elimination of thermal e.m.f.'s, surface and gas conduction effects.

The values of conductivity and activation energies reported by these workers vary widely. The values of the activation energies are perhaps the more important in interpretation. They range from 0.5 eV to 5.8 eV. High values are most likely attributable to gas phase conduction. The values of 4 to 5 eV reported by some authors cannot be due to intrinsic conduction as sapphire has been found completely ionic at least up to 1450°C (Yee and Kröger, 1973).

It would appear that the wide spread in values for conductivity and activation energies reported are most likely due to the lack of a standardized procedure in the measurement and in the wide variation in impurity levels. However, it also appears that such a method is slowly evolving. It is highly probable, on the other hand, that lack of samples of a sufficient degree of purity may continue to give widely varying results from different sources.

A great deal less work aimed specifically at the variation of conductivity with frequency as well as temperature and oxygen partial pressure has been reported. None of it has been done with a physical gas phase guard.

Tallan and Graham (1965) made measurements on pure and doped single crystals of Alumina between 10^2 and 10^5 Hz and between 700° and 1200°C, using a 3-terminal technique with no physical gas phase guard. They explained their results by an interfacial polarization mechanism, with a surface layer up to 40μ thick, having a resistivity 5 orders of magnitude greater than the bulk.

As previously mentioned, diffusion in ionic crystals also depends upon defect migration, and some diffusion coefficient measurements are of interest to the present work.

Paladino and Kingery (1962) measured the tracer diffusion coefficient of Al in polycrystalline Al_2O_3 over the range 1670° to 1905°C, using Al-26 as a tracer. They found, by comparison of the diffusion data, that aluminium ion mobility is greater than oxygen ion mobility, as would be expected from ionic size considerations.

Jones et al (1969) determined a defect diffusion coefficient in single crystal aluminum oxide from 1400° to 1850°C by measuring the movement of the colour boundary in a Ti^{3+} doped single crystal.

CHAPTER 2 - THEORY OF ELECTRICAL CONDUCTION IN IONIC SOLIDS

Electrical conduction in any material consists of the transport of electrical charge by the motion of charged carriers under the influence of an electric field. In ionic solids the charge carriers are point defects of the crystal lattice. The defect chemistry of the solid state deals with the relative densities of differenct possible types of defects in a crystal as a function of the ambient atmosphere and temperature, and their distribution in the allowed energy levels, and this is discussed first in section 2.1. Principles of calculating the actual energy levels are then discussed in terms of the Born-Mayer theory in section 2.1.3.

Defect motion also takes place under the influence of a concentration gradient, i.e. diffusion occurs in solids. The microscopic processes involved are closely related to these in conduction, and diffusion and d.c. conductivity are therefore discussed together in section 2.2.

In a time varying field, some modes of charge carrier transport are active that would cease in the steady state (i.e. in d.c. conduction). For example, some carrier motions stop after a finite distance in a d.c. field due to the carrier meeting energy barriers so high as to be virtually insurmountable in ordinary experimental times. The relative displacement of oppositely charged carriers in this way results in the material acquiring a dipole moment per unit volume, or polarisation, which will be time varying in a time varying field. This time varying polarisation necessitates that a varying current flow in the external leads applying the potential across the crystal, and hence it leads to an a.c. conductivity. In section 2.3. the macroscopic relationships

between the polarisation and the permittivity and conductivity in a sinusoidal field, and the idea of complex permittivity are first discussed, and finally the relationship between the time response to a step electric field and the variation with frequency of the response to sinusoidal fields is discussed. This last topic has important application to the calculation of the frequency response of models of carrier motion, since it is often easier to analyse the model for the step field time response and obtain the frequency response from these general relationships.

The results of sections 2.2 and 2.3. are then finally applied in section 2.4 to the analysis of microscopic models of charge transpost for their d.c. conductivity dependence on ambient atmosphere and temperature and their a.c. conductivity dependence on the same factors and additionally as a function of frequency. The well known simple "double potential well" model is presented, but in addition a new "multi-pair potential well" model (Mason, to be published) is presented. This predicts much smaller maximum slopes on the log σ - logω plot than the value of 2 given by the simple double well model, and these values are more in accord with general observations and the results presented later here.

2.1. Defect Chemistry of Solid State

2.1.1. Types of defects and electrical conductivity

The electrical conductivity, σ, is given by

$$\sigma = \sum_i (c_i e_i \mu_i)$$

where c_i is the concentration of charge carrier type i, e_i is its charge and μ_i its mobility. The charge carriers will be point defects. In a pure crystal the point defects are: vacancies, interstitials and misplaced atoms, while in an impure crystal there are also impurity atoms. In addition to these atomic defects, there are also electronic defects: the electron and the hole.

These are the basic defects. There are further point defects that result from interaction between these basic types, called associates.

If the crystal is ionic in character, these defects may ionize by losing or collecting electrical charges. Consider, for example, an oxygen vacancy in an ionic oxide. On removing the oxygen atom, two electrons are left in the region of the vacancy. The vacancy can then ionize with one or both electrons leaving the vacancy. See Figure 2.1.

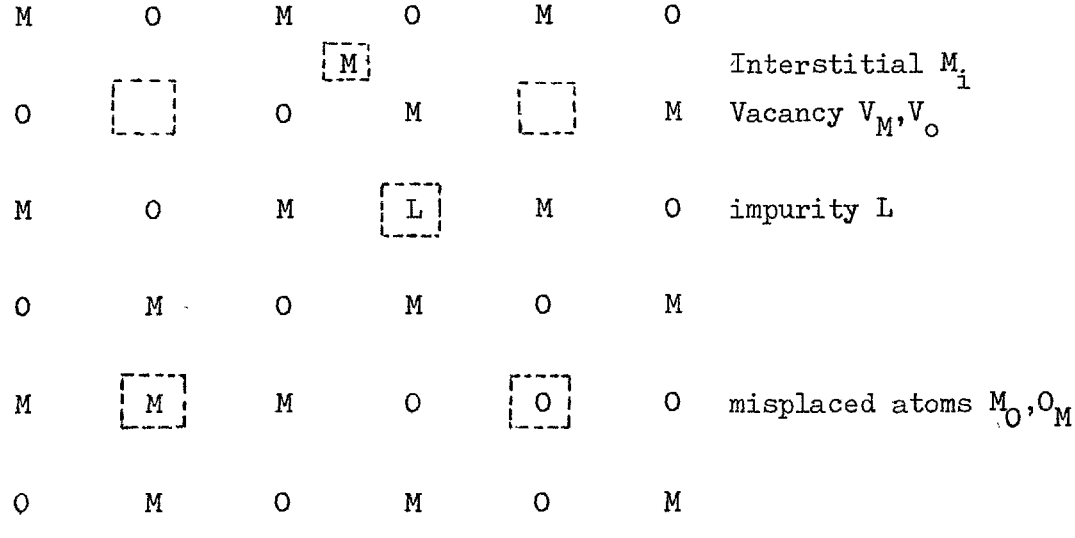


Figure 2.1. Point defects in an oxide, MO

Ionization of defects can occur by movement of electrons to either higher or lower energy levels and the ionized state may be the more stable. Another situation occurs when electrons are thermally excited from a lower energy level to a higher one. Point defects can occur in thermal equilibrium within the material and hence can be treated as chemical species in terms of equilibrium constants and the mass action law.

2.1.2. Formation of Defects and Equilibrium Concentration

In principle, all types of defects may be present in an oxide, but in general, one will predominate. The defect formation may either occur internally in the solid or through reactions with the environment. There are various rules which must be followed and obeyed in writing the correct defect reactions.

i) The ratio of the number of regular cations and anion sites in a crystalline ionic oxide is constant. Thus in a compound MO the ratio of regular M and O sites is 1:1 regardless of whether the actual composition is stoichiometric or nonstoichiometric.

ii) The total number of regular sites may change in a defect reaction, and therefore, the defect equation may include the creation or annihilation of sites. No sites are created in the formation of electronic defects.

iii) The defect equation must balance with respect to the mass, namely the number of atoms involved in the defect reaction must be the same before and after the defect formation. Vacancies have zero mass. Electronic defects do not affect the mass balance.

iv) The compounds are electrically neutral. The total effective charge is the same before and after the formation of defects.

The symbols and system used by Kröger and Vink will be employed in this thesis. In this system the type of imperfection is indicated

by a major symbol and the site occupied by a subscript. Effective charges are shown by superscript, where a cross, x, means a normal atom with zero effective charge, a dash, / , shows a negative effective charge and a dot, . , shows a positive effective charge. Negative - and positive + designate real charges. The neutral indication is sometimes omitted. For electrons and holes the effective and actual charges are the same.

Schottky disorder: This involves the presence of equivalent amounts of cation and anion vacancies. If one starts with a pair of cations and anions on regular sites within the crystal M_M and O_O, in writing the reaction one must also take into account that the formation of a Schottky pair results in the formation of two new lattice sites; hence

$$M_M + O_O \rightleftharpoons V_M'' + V_O^{\bullet\bullet} + M_M + O_O \qquad (2.1)$$

in this equation M_M and O_O may be cancelled and the net reaction becomes

$$O \rightleftharpoons V_M'' + V_O^{\bullet\bullet} \qquad (2.2)$$

Frenkel disorder: In the formation of a Frenkel defect pair, a cation on a normal site is transferred to an interstitial position and no new lattice sites are created.

$$M_M \rightleftharpoons M_i^{\bullet\bullet} + V_M'' \qquad (2.3)$$

Under conditions where a crystal is in equilibrium with the partial pressure of one of its components in an atmosphere above it, adjustments of the partial pressure will alter the composition of the crystal. Taking

as an example a binary oxide, MO, under equilibrium conditions so that the chemical potential of oxygen in the gas phase and in crystal are equal, we are saying the variations in oxygen pressure p_{O_2}, will cause variations in the ratio M:O in the solid. To see this, first suppose p_{O_2} to be fixed at the value p'_{O_2} where the oxide is stoichiometric so that $[M]_{Total} = [O]_{Total}$. $[M]_{Total}$ is the overall M concentration to which there may be several contributing terms such as $[M_M^x]$, $[M_i^x]$, and $[M_O^x]$. If p_{O_2} is increased $[O]_{Total}$ will be increased relative to $[M]_{Total}$. The excess oxygen is incorporated by defect reactions. Two of the many possible mechanisms are the incorporation of O atoms in interstitial positions according to

$$V_i^x + \frac{1}{2}\, O_2(g) \longrightarrow O_i^x \qquad (2.4)$$

or incorporation onto O sites with creation of M vacancies according to

$$\frac{1}{2}\, O_2(g) \longrightarrow O_O^x + V_M^x \qquad (2.5)$$

O_O^x has real charge $O^=$ and V_M^x has real charge V_M^{++}. The mass action law (see appendix 1) may be applied to these reactions and it follows that

$$[O_i^x] = K_1\, p_{O_2}^{1/2} \qquad (2.6)$$

or

$$[V_M^x] = K_2\, p_{O_2}^{1/2} \qquad (2.7)$$

showing that the defect concentrations depend on the environmental oxygen pressure. In writing these equations it has been assumed that the defect

concentrations $[O_i^x]$ and $[V_M^x]$ are small and hence the concentrations $[V_i^x]$ and $[O_O^x]$ of the normal lattice constituents are essentially constant.

If p_{O_2} falls below p'_{O_2}, an oxygen deficit occurs in the crystal which may, for example, be achieved as follows;

$$O_O^x \rightarrow \frac{1}{2} O_2(g) + V_O^x \qquad (2.8)$$

V_O^x is neutral because two electrons are left bound at V_O, maintaining the lattice nuetrality.

In this case the defect concentration dependence on p_{O_2} is

$$[V_O^x] = K_3 \, p_{O_2}^{-\frac{1}{2}} \qquad (2.9)$$

The existence of an impurity, L, in solution in a crystal, MO, gives rise to point defects which may be substitutional (L_M^x or L_O^x) or interstitial (L_i^x) or both. These defects may also occur in the ionised form. The impurity may occur in quasi-equilibrium or complete equilibrium. In the former, the impurity is a species having low volatility and the crystal contains a fixed total quantity of the impurity which may be divided into a number of different defect types, for example

$$[L]_{Total} = \text{Constant} = [L_i^x] + [L_M^x] \qquad (2.10)$$

In the complete equilibrium situation, a partial pressure of the impurity is maintained in the atmosphere in equilibrium with the concentration of impurity in the crystal. Some simple possible reactions for the incorporation of the impurity are as follows:

$$L(g) \rightarrow L_O^x + V_M^x \tag{2.11}$$

$$L(g) \rightarrow L_M^x + V_O^x \tag{2.12}$$

$$L(g) \rightarrow L_i^x \tag{2.13}$$

The mass action law may be applied to these equations to give the concentration dependence of the impurity defect in the crystal on the partial pressure of L in the atmosphere.

In an ionic oxide, the effectively neutral oxygen ion, O_O^x, carries a real charge, $O^=$, so that its movement would apparently imply charge transport and hence conduction. However, as an O_O^x moves one way, a vacancy moves the other. If the vacancy is effectively neutral, V_O^x it carries a real charge of -2 (from two localized electrons) so that exchange with the oxygen ion has $O^=$ moving one way and two electrons localized at a vacancy moving the other, with no resultant charge transfer. In order to produce any charge transport by their motion, defects must carry an effective charge relative to the lattice.

Consistent results are achieved either by working with real charge or with effective charges. The latter is mostly adopted here because it shows readily the effect of defects on the overall charge neutrality of the crystal, as well as indicating where coulombic interactions are possible between defects.

All defects formed may ionize. For example, oxygen vacancies act as donors and become singly or doubly charged. Consider the defect reactions

$$O_O \rightleftharpoons V_O^x + \frac{1}{2} O_2 \tag{2.14}$$

followed by

$$V_O^x \rightleftharpoons V_O^{\bullet} + e' \tag{2.15}$$

If $V_O^{\bullet}$ ionized once more

$$V_O^{\bullet} = V_O^{\bullet\bullet} + e' \tag{2.16}$$

Or combining the equations (2.14), (2.15) and (2.16) yields

$$O_O = V_O^{\bullet\bullet} + \frac{1}{2} O_2 + 2e' \tag{2.17}$$

Next consider the following defect reaction in a metal deficient oxide. A metal vacancy may be formed through the reaction of oxygen with the oxide

$$\frac{1}{2} O_2 = V_M^x + O_O \tag{2.18}$$

If V_M^x ionizes one or both of the holes then this is represented as

$$V_M^x = V_M' + h^{\bullet} \tag{2.19}$$

and

$$V_M' = V_M'' + h^{\bullet} \tag{2.20}$$

Combining the equations (2.18), (2.19) and (2.20) one gets

$$\frac{1}{2} O_2 = V_M'' + O_O + 2\,h^{\bullet} \qquad (2.21)$$

The above examples of ionization reactions lead to charged point defects and (charged) electronic defects, both of which can act as carriers, and so in general would lead to mixed, i.e. electronic plus ionic, conductivity.

However in many materials at ordinary temperatures the electronic defect mobility is much higher than the ionic defect mobility and these reactions still lead to the observation of almost pure electronic conduction.

Pure ionic conduction may be observed as the temperature is raised if the electronic mobility is very low and the ionic mobility increases more rapidly with temperature.

It may also be observed if the charged defects are formed by reactions which do not produce electronic defects. There are many possibilities for this to be achieved by the incorporation of impurities of valency different from the pure crystal constituents. It could also be achieved in a pure crystal by the formation of Schottky or Frenkel defects, as discussed for MO in section 2.1.2.

In principle also, a mixture of Schottky and Frenkel defects could occur, plus the possibility of interchange of ions between sites (Anti-structure) (Kröger (1964) p. 411, Van Gool (1966) p.57) but in practice one type will usually predominate in a given situation, due to the differences in free energy of formation (see section 2.1.3).

This situation in which the ionic defects are produced in pairs without impurities being involved may be referred to as intrinsic ionic conductivity, while when the presence of impurities is involved (to give the necessary charge balance without electronic defects) the conductivity produced may be called extrinsic ionic. Again, the situation may change from extrinsic to intrinsic with rising temperature if the concentration of defect pairs produced by the intrinsic reaction becomes overwhelmingly

greater than the impurity concentration.

2.1.3. Energy band structure and defect energy levels

In a single atom, the possible states for electrons occur at distinct and separate energies; as the atoms are brought together to form the solid, those electron states which do not overlap with the neighbouring atoms remain as narrow local levels so that electrons in these states are localized around particular atoms in the solid. The outer electron states of the individual atoms which overlap with their neighbours, are transformed into a set of new states in the solid which taken together cover a band of energy values. The extent of overlap of the electron states on the individual atoms, determines the width of the energy band; this width is the range of energies covered by the electron states in the solid contributed by the one particular level of the free atoms. The overlap also influences the size of the energy gaps, which are ranges of energy for which there are no electron states in the solid and which are as a consequence inaccessible to electrons. The outer occupied-level in the isolated atom forms the valence band in the solid; the lowest unoccupied level in the isolated atom forms the conduction band. The gap between the two is the band gap (Fig. 2.2).

In a band, the electrons are not localized and move through the crystal as free electrons subjected to a small perturbing potential provided by atoms on their sites. Electrons may also exist in states localized around particular atoms and move by jumping from the potential well around one atom to that around another. The electron mobilities for the two types of motion are very different. When a defect is added to the crystal, it destroys the periodicity locally and introduces such local electron states, which are different from the perfect crystal states and may lie within the band gap. Whether these states are occupied by electrons or not depends on the particular

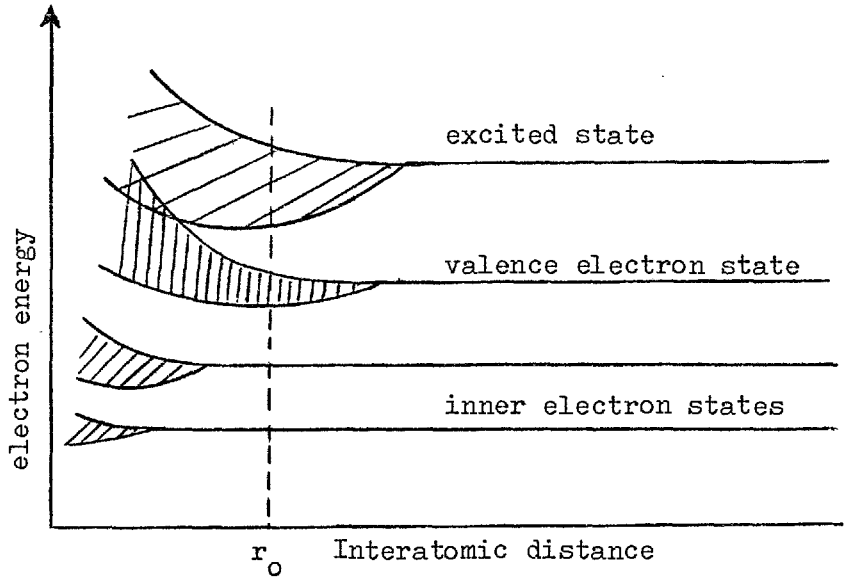


Fig. 2.2(a) Electron energy state in an elemental solid as a function of interatomic spacing.

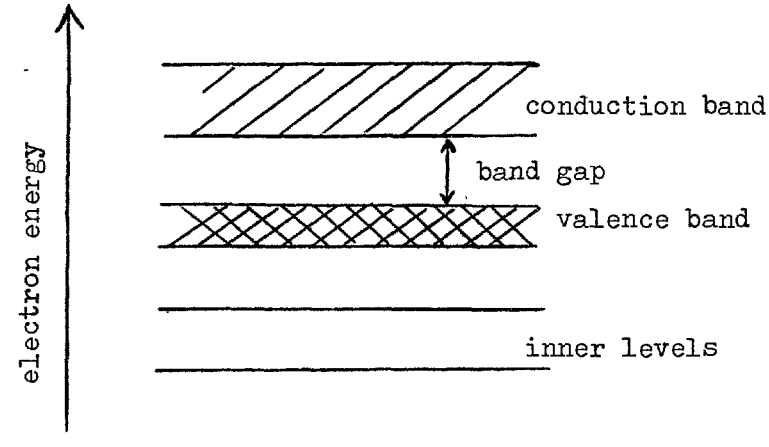


Fig. 2.2(b) Electron energy states at equilibrium interatomic spacing r_o, are plotted against a general space co-ordinate in the solid.

defect. If the neutral defect provides an occupied level within the band gap, then it is termed donor-like since the electron occupying the level can be donated to a state in the conduction band; if the neutral defect provides an unoccupied level in the gap then it is termed acceptor-like, since an electron may be excited from a state in the valence band of the host into the defect level.

Let us illustrate these considerations by an example, for which it is convenient to choose an ionic compound; since the electrons are predominantly associated with particular ions, the electron band that an electron occupies can be described in terms of the ion to which the electron is attached. Let us consider an ionic oxide, MO (Brook (1974)).

For the oxygen ions, $O^{=}$, the inner electrons occupy distinct local levels in the solid, whereas the outer electrons, by partially interacting with other $O^{=}$ ions give rise to an electron band. The energy of the highest occupied band arising from the oxygen ion level is given by the energy of the reaction

$$O^{-} \text{ (lattice)} + e^{-} \text{ (free space)} \rightarrow O^{=} \text{ (lattice)} \qquad (2.22)$$

which is the energy needed to attach the outermost electron to the oxygen ion in the solid. The energy for this equation (2.22) contains three main terms:

(1) the electron affinity of the O^{-} ion which is E_A, the energy of the reaction

$$O^{-} \text{ (free space)} + e^{-} \text{ (free space)} \rightarrow O^{=} \text{ (free space)} \qquad (2.23)$$

(2) the electrostatic energy, E_M, gained by adding an electron into the positive potential region in the unrelaxed crystal

$$e^- \text{ (free space)} \rightarrow e^- \text{ (vacant anion site)} \qquad (2.24)$$

(3) local ion movements following reaction (2.24), which reduce the electrostatic energy contribution by an amount E_P.

In sum, the energy of the band measured with respect to an electron in free space is

$$E_O = E_A - E_M + E_P \qquad (2.25)$$

The convention is to label a band or level with the species that is formed when the level is occupied; $E_{O^=}$ thus gives the approximate energy of the $O^=$ band. Similarly the outer electrons of the cation M^{++} will give rise to bands. The energy of the highest occupied band arising from the cation level is given by

$$M^{+++} \text{ (lattice)} + e^- \text{ (free space)} \rightarrow M^{++} \text{ (lattice)} \qquad (2.26)$$

which is the energy needed to attach the outermost electron to the cation in the solid. For a non-transition metal ion, this is

$$E_{M^{++}} = -E_I + E_M + E_P \qquad (2.27)$$

where E_I is the third ionization potential of the metal M, and where E_M enters with changed sign since the electron is now added to a region of negative potential within the crystal.

The highest occupied band at absolute zero temperature is fully occupied in MO and is the valence band. Since the $O^=$ and M^{++} bands are the highest occupied bands contributed by the constituents of the oxide,

the nature of the valence band depends on the relative values of $E_{O^=}$ and $E_{M^{++}}$.

The conduction band in MO is the lowest empty band and could in principle be either the next higher oxygen band or the next higher cation band. These have energies given by

$$O^= \text{ (lattice)} + e^- \text{ (free space)} \rightarrow O^{\equiv} \text{ (lattice)} \qquad (2.28)$$

and

$$M^{++} \text{ (lattice)} + e^- \text{ (free space)} \rightarrow M^+ \text{ (lattice)} \qquad (2.29)$$

The electron affinity of $O^=$ is large and positive, so $E_{M^+} < E_{O^{\equiv}}$, and M^+ band forms the conduction band.

A similar technique can be applied to the location of a defect level within the band gap provided the electron is strongly localized around the defect.

So far we have considered the perfect lattice; turning to defects V_O^x and M_i^x act as a donor; V_M^x and O_i^x act as acceptors. Substitutional impurities generally act as donors or acceptors depending on whether they have more or less valence electrons than the species they replace.

So far we have discussed the consequences of the periodic structure of the perfect crystal and the presence of the point defects in producing a set of energy levels in which the electronic defects may be distributed. The same question must now be considered for the point defects themselves, since they act as the charge carriers when the conductivity is ionic.

In section 2.1.2. it was pointed out that the Mass Action Law may be applied to relate the equilibrium ratios of the concentrations of a set of defects which may participate in a defect reaction to the reaction constant K. However, in order to calculate the absolute concentration of the various defects and the variation of K with temperature

it is necessary to calculate the energy of formation of each type of defect relative to the perfect lattice. These may be thought of as a set of energy levels for the ionic defects, analogously to the energy levels for the electronic defects.

Further, the energy of formation of a particular defect will vary with its proximity to other defects, particularly if they carry a charge of opposite sign. In principle, therefore, a defect, for example a cation vacancy, will have a whole series of formation energies depending on how many lattice sites it is away from, for example, oppositely charged anion vacancies.

Usually, however, the energy of formation is taken to refer to the energy needed to form the defect in the perfect lattice with no other defects present. This is of course the limiting value of the series just discussed where the charged defect is considered associated with the charge compensating oppositely charged defect.

The general principle of calculating the energy of defect formation is to find it as the difference between the complete energy of formation of the perfect lattice and the energy of formation of the lattice containing one of the required defects.

In order to illustrate the method we follow the presentation given by Swalin (1962), of the calculation for a vacancy.

The lattice energy of the crystal, W_ℓ, is given by the reaction

$$M^+ (g, 1 \text{ atm}) + X^- (g, 1 \text{ atm}) = MX (s, 1 \text{ atm})$$

and may be found as follows:

In an ionic crystal, the forces holding the ions together in the crystal are coulombic in nature. Thus the energy of two univalent ions separated by a distance r is

$$W_a = \mp \frac{e^2}{r} \tag{2.30}$$

where e is the electronic charge. A given ion is not affected by only one ion but by many. The net coulombic energy of this ion in the crystal is thus the sum of the repulsive and attractive interactions with all the other ions which may be written

$$W_a = - A \frac{e^2}{r} \tag{2.31}$$

where A is the Madelung constant. When the ions approach each other closely a strong repulsive interaction W_r occurs. This repulsive term is not as easy to calculate as the attractive term. In order to evaluate this term Born assumed an exponential relation of the type

$$W_r = \frac{B}{r^n}$$

where B and n are constants for a given system. The total lattice energy for N_o positive ions and N_o negative ions is

$$W_\ell = N_o (W_a + W_r) = N_o \left(- \frac{A e^2}{r} + \frac{B}{r^n}\right) \tag{2.32}$$

The equilibrium atom spacing will be r_o, where $\frac{dW_\ell}{dr} = 0$. Therefore, differentiating equation (2.32) with respect to r, setting the result equal to zero, and solving for B, we find

$$B = \frac{A e^2 r_o^{n-1}}{n} \tag{2.33}$$

Substituting equation (2.33) into (2.32), we find

$$(W_\ell)_{r=r_o} = - \frac{N_o A e^2}{r_o} \left(1 - \frac{1}{n}\right) \tag{2.34}$$

Since there are N_o atoms in the crystal, the binding energy per atom is

$$W = - \frac{A e^2}{r_o} \left(1 - \frac{1}{n}\right) \tag{2.35}$$

The process of vacancy formation on the cation sublattice may be divided into three steps:

(i) M^+ (interior of crystal) = $M^+(g)$ W_1

(ii) $M^+(g) = M^+$ (surface) W_2

(iii) Relaxation of atoms around vacancy W_3

The energy for the process (i) is the negative of equation (2.35). Thus

$$W_1 = \frac{A e^2}{r_o} \left(1 - \frac{1}{n}\right) \tag{2.36}$$

Upon completion of process (ii), one-half of the energy expended in process (i) will be regained. Thus

$$W_2 = - \frac{A e^2}{2 r_o} \left(1 - \frac{1}{n}\right) \tag{2.37}$$

The energy for vacancy formation, excluding relaxation of atoms around the vacancy, will be $(W_1 + W_2)$ or

$$W_1 + W_2 = \frac{A e^2}{2 r_o} \left(1 - \frac{1}{n}\right) \tag{2.38}$$

Let us now examine what happens in the region of the crystal where the

vacancy has been created. This is shown schematically in Fig. 2.3. In most ionic crystals, the nearest neighbours to a cation will be anions. Removal of the positive ion will have the same effect on the neighbours as the substitution of a negative charge. The interatomic distances in the crystal are governed by the balance of positive and negative charges. Removal of a positive ion will allow positive ions surrounding the vacancy to relax inward because of repulsion from neighbours surrounding the ions. Conversely the negative ions will move outward. This will result in polarization of the ions in the vicinity of the vacancy since the field about them is now asymmetric. The vacancy behaves as if it has negative charge. The cation vacancy in the crystal is treated as a spherical cavity of charge -e inside a homogeneous material of dielectric constant . The radius of the spherical cavity will be the radius of the positive ion R_+. The existence of the charge -e in the cavity will result in the polarization of the material around the cavity. Jost calculates the polarization energy, using classical dielectric theory, to be given by

$$W_{p_+} = - \frac{e^2}{2 R_+} \left(1 - \frac{1}{\epsilon}\right) \qquad (2.39)$$

The energy of vacancy formation will thus be reduced by W_{p_+} upon polarization. The energy of vacancy formation is therefore

$$W_{V^+} = \frac{A e^2}{2(R_+ + R_-)} \left(1 - \frac{1}{n}\right) - \frac{e^2}{2 R_+} \left(1 - \frac{1}{\epsilon}\right) \qquad (2.40)$$

In this equation the interatomic distance r_o has been replaced by the sum of the radii $R_+ + R_-$. Similarly, the energy of formation of an anion vacancy is

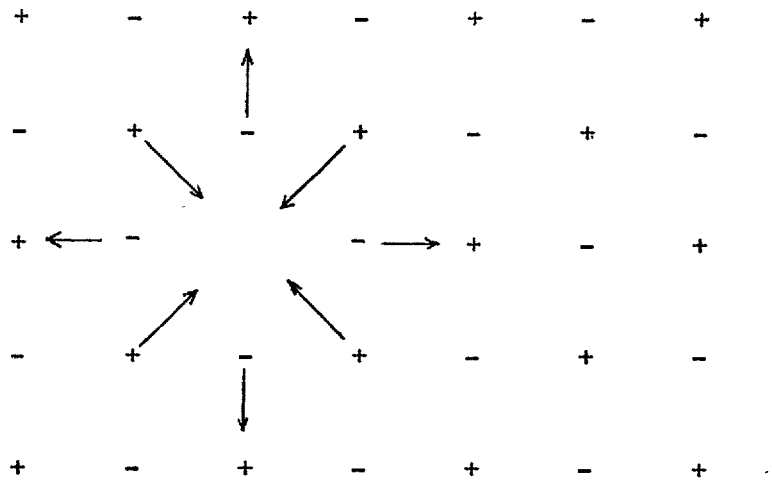


Figure 2.3. Relaxation of atoms around vacancy in an ionic crystal.

$$W_V = \frac{A e^2}{2(R_+ + R_-)} \left(1 - \frac{1}{n}\right) - \frac{e^2}{2 R_-} \left(1 - \frac{1}{\epsilon}\right) \qquad (2.41)$$

The polarization term may be very important and, in fact, may reduce the defect formation energy by several electron volts.

The method described above is in essence the Born-Mayer theory, except that this theory uses a repulsive potential given by

$$W_r = Be^{-r/\rho}$$

which is known from wave mechanical considerations to be more appropriate than the original suggestion of Born described here.

This approach has been used with considerable success in alkali halides, and even extended successfully to the calculation of association energies between a vacancy and a substitutional cation impurity (Bassani and Fumi, 1954).

It has so far however been less successful in the case of metal oxides (Boswarva, 1967). This is believed (Boswarva, 1967) to be due to the fact that the short range oxygen-oxygen ion interaction energy cannot be described accurately by either Br^{-n} or $Be^{-r/\rho}$, rather than to any basic error of the method. This difference in the form needed for W_r arises because the less ionic nature of the bond in metal oxides means that the outer electron shell of the oxygen ion approximates less closely to being closed.

2.1.4. Population of bands and defect states

At temperatures above $0^{o}K$, it is possible for electrons occupying low levels to be thermally excited into higher empty levels. Each excitation from the valence to the conduction band for example will give rise to a free electron e', in the conduction band, and free hole, $h^{\bullet}$, in the valence band. Using chemical representation for defect reactions,

and since the number of electrons in the valence band and of unoccupied states in the conduction band are virtually unchanged, we may for small defect concentrations write this process as

$$e^{x}_{\substack{\text{valence}\\\text{band}}} + h^{x}_{\substack{\text{conduction}\\\text{band}}} \rightleftharpoons e'_{\substack{\text{conduction}\\\text{band}}} + h^{\bullet}_{\substack{\text{valence}\\\text{band}}} \qquad (2.42)$$

and by the mass action law

$$[e'][h^{\bullet}] = n\,p = K_i \qquad (2.43)$$

where

$$K_i = C \exp\left(^{-\Delta E_i}/kT\right) \qquad (2.44)$$

and ΔE_i is the band gap.

The electronic species are assumed to have negligible mass, so that e^x and h^x in equation (2.42) and others can be dropped. Then (2.42) can be rewritten

$$0 \longrightarrow e' + h^{\bullet} \qquad (2.45)$$

If the process (2.42) dominates other electron and hole formation processes, and the concentration of charged atomic defects is small, then charge neutrality may be achieved by equal concentration of electrons and holes

$$n = p \qquad (2.46)$$

which is the condition for intrinsic electronic disorder.

If electrons and holes are generated predominantly by the ionisation of defects, then the electronic carriers are extrinsic. Electrons arise by excitation from donor defects, e.g.

$$M_i^x \longrightarrow e' + M_i^\bullet \qquad (2.47)\ ; \qquad \frac{[M_i^\bullet]\, n}{[M_i^x]} = K_{M_i}^{\ x} \qquad (2.48)$$

and holes arise by excitation of electrons from the valence band into acceptor levels, e.g.

$$O_i^x \rightarrow h^\bullet + O_i' \qquad (2.49)$$

by the mass action law

$$\frac{[O_i']\, p}{[O_i^x]} = K_{O_i}^{\ x} \qquad (2.50)$$

The various equilibrium constants are given by expressions of the type

$$K_x = K_x^{\,o} \exp\left(-\frac{\Delta E_x}{kT}\right) \qquad (2.51)$$

where ΔE_x is the energy separation between the defect level and the energy band involved in the excitation. According to semiconductor type statistics (Kofstad 1972) the number of electrons in the conduction band

$$n = \left(\frac{8\pi\, m_{e'}^x\, kT}{h^2}\right)^{3/2} \exp -(E_C - E_F)/kT \qquad (2.52)$$

and the number of holes in the valence band

$$p = \left(\frac{8\pi\, m_h^x \cdot kT}{h^2}\right)^{3/2} \exp - \left(\frac{E_F - E_V}{kT}\right) \qquad (2.53)$$

where $m_{e'}^x$ and m_h^x are the effective mass of the electron and hole respectively. Then the equilibrium constant

$$K_i = np = \left(\frac{8\pi\, m_{e'}^x\, kT}{h^2}\right)^{3/2} \left(\frac{8\pi\, m_h^x \cdot kT}{h^2}\right)^{3/2} \exp - \left(\frac{E_C - E_V}{kT}\right) \qquad (2.54)$$

The product np is constant at a given temperature.

Whether a given level is occupied or not depends not only on its location in the band diagram but also on the Fermi level, E_F, in the solid. From Fermi statistics, the probability that a given level, energy E, is occupied is given by

$$P(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \qquad (2.55)$$

For $E = E_F$, $p(E) = \frac{1}{2}$.

The position of the Fermi level on the band diagram depends on how many electrons are available for filling the levels. In general, reduction processes, which are removal of oxygen atoms from the crystal, leaving electrons behind, increase the electron concentration and raise the Fermi level; and oxidation processes lower the Fermi level.

We can assume small concentrations of defects so that activities may be replaced by concentrations in the mass action equations.

Similarly to section 2.1.3 where we firstly discussed the energy level structure for the electronic defects and secondly for the ionic defects, we now go on to discuss the distribution of the ionic defects among the available levels.

Since the ionic defects are massive particles, their behaviour is described by Maxwell-Boltzmann statistics.

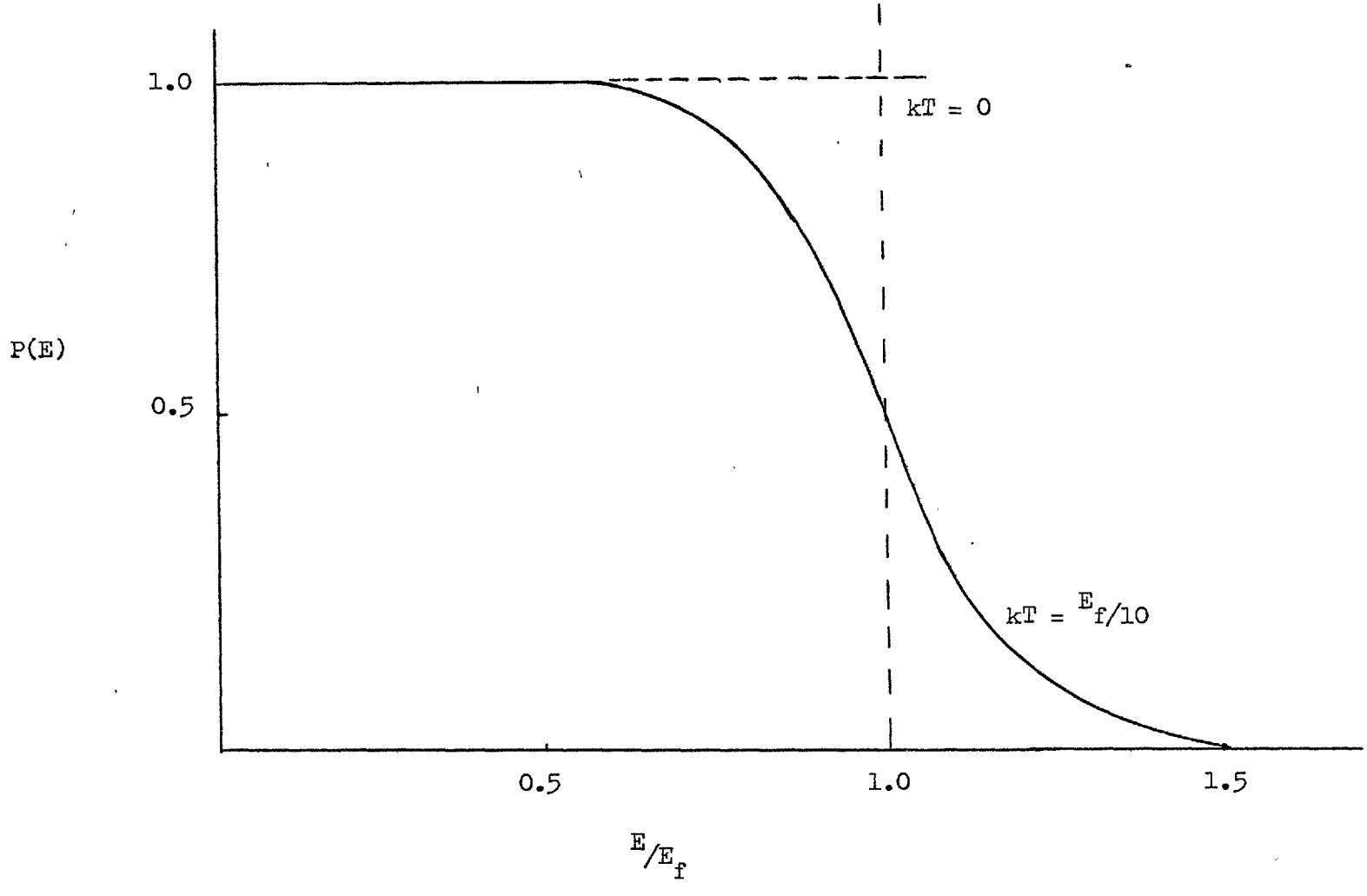


Fig. 2.4 The Fermi-Dirac distribution function; the probability, P(E) that an electron state of energy, E, will be occupied is shown for two temperatures.

For the i^{th} type of defect there will be a set of k energy levels W_{ij}, $j = 1 \cdots k$. By this we mean that there are k different types of site which the defect of type i may occupy, each with a different energy. There will in general be different number densities, N_j, of each type of site.

The probability f_j of an ion occupying a site of energy W_{ij} is given by

$$f_j = A \exp (- W_{ij}/kT) \tag{2.56}$$

where A is a constant to be evaluated, as explained below.

The number density of ions in the j^{th} type of site will therefore be

$$n_j = N_j f_j = N_i A \exp (- W_{ij}/kT) \tag{2.57}$$

If the factors N_j and W_{ij} are known, then the values of A and hence n_j may be determined from the two conditions

$$\sum_{j=1}^{k} n_j = c_i \tag{2.58}$$

and

$$\sum_{j=1}^{k} N_j = N_s \tag{2.59}$$

where c_i is as before the total concentration of defects of type i and N_s is the total number of sites available to them.

2.2. Diffusion and d.c. Conductivity

2.2.1. Macroscopic description of diffusion by Fick's Laws and Their Microscopic Basis

Experiment has shown (Kofstad, 1972) that diffusion in solids obeys Fick's Laws, the first of which is

$$j = - D \left(\frac{\partial c}{\partial x}\right)_t \tag{2.60}$$

where j represents particle flow rate per unit area of the diffusing species across a plane, c is the concentration at the plane, $(\partial c/\partial x)_t$ is the concentration gradient normal to the plane, and D is called the diffusion coefficient. The minus sign implies that the flow takes place from regions of high to low concentration.

The second law is a direct mathematical consequence of the first and of the conservation of the diffusing particles. It describes the situation when c is a function of time also, and the equation is (Kofstad, 1972)

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} \tag{2.61}$$

However, this law will not be much used in this thesis.

Fick's first law may be derived as follows (Kofstad, 1972) let us consider a series of parallel planes separated by a distance a. Let two neighbouring planes lie 1 and 2, and the volume concentrations of particles in the plane be c_1 and c_2 respectively, with $c_1 < c_2$ and a corresponding concentration gradient $\frac{dc}{dx}$, so that $c_2 = c_1 + a\left(\frac{dc}{dx}\right)$. The number of particles per unit area in the plane 1 and 2 are given by ac_1 and ac_2 respectively, and the particles in planes 1 and 2 may jump from one plane to another at a jump frequency p. The particles in

plane 1 have an equal probability of jumping to plane 2 or to the neighbouring plane in the opposite direction. The total number of particles jumping out of plane 1 per unit time is pac_1. As the particles may jump in one of two opposite directions, the number of particles jumping from plane 1 to 2 is given by $\frac{1}{2}\ pac_1$. Similarly, the number of particles jumping from plane 2 to 1 is given by

$$\frac{1}{2}\ pc_2a = \frac{1}{2}\ p\ \left[c_1a + a^2\ (\frac{dc}{dx}) \right]$$

The difference in jump rates is equal to the net flow of particles

$$j = \frac{1}{2}\ p\ ac_1 - \frac{1}{2}\ p\ (c_1a + a^2\ \frac{dc}{dx}) = -\ \frac{1}{2}\ p\ a^2\ \frac{dc}{dx} \qquad (2.62)$$

Comparing equations (2.60) and (2.62) we get

$$D = \frac{1}{2}\ p\ a^2 \qquad (2.63)$$

for one-dimensional flow. D has units cm^2/sec. If one considers a large number of jumps, n, between neighbouring planes which occurs during time t, then

$$p = \frac{n}{t} \qquad (2.64)$$

and substituting this value in equation (2.63)

$$2Dt = n\ a^2 \qquad (2.65)$$

2.2.2. Diffusion Mechanisms and Models

Lattice diffusion takes place through the movement of point defects. The presence of different types of defects gives rise to different mechanisms of diffusion. The important ones are illustrated schematically here.

Vacancy Mechanism: The diffusion is said to take place by the vacancy mechanism if an atom on a normal site jumps into an adjacent unoccupied lattice site. This is illustrated schematically in Fig. 2.5. It should be noted that the atoms move in the opposite direction to the vacancies.

Interstitial Mechanism: If an atom on an interstitial site moves to one of the neighbouring interstitial sites, the diffusion occurs by an interstitial mechanism. This is shown schematically in Fig. 2.6. Such a movement of the interstitial atom involves a considerable distortion of the lattice, and this mechanism is probable only when the interstitial atom is smaller than the atoms on the normal lattice positions.

2.2.3. Motion of Charged Defects in an Electric Field

When an electric field, E, is applied across a crystal, a force is exerted on all the charged particles. If an ion or a defect has a charge q_i, the force F_i is given by

$$F_i = q_i E \tag{2.66}$$

This force causes a directional transport of the charged particles in addition to their random thermal motion. For low enough fields, the resulting current density is given by

$$J_i = \sigma_i E \tag{2.67}$$

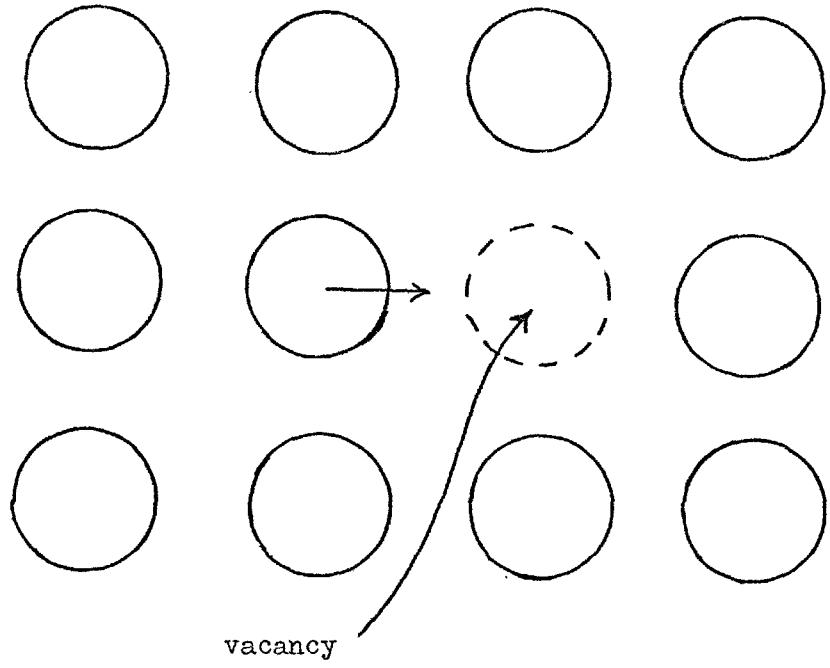


Fig. 2.5 Vacancy diffusion in solids

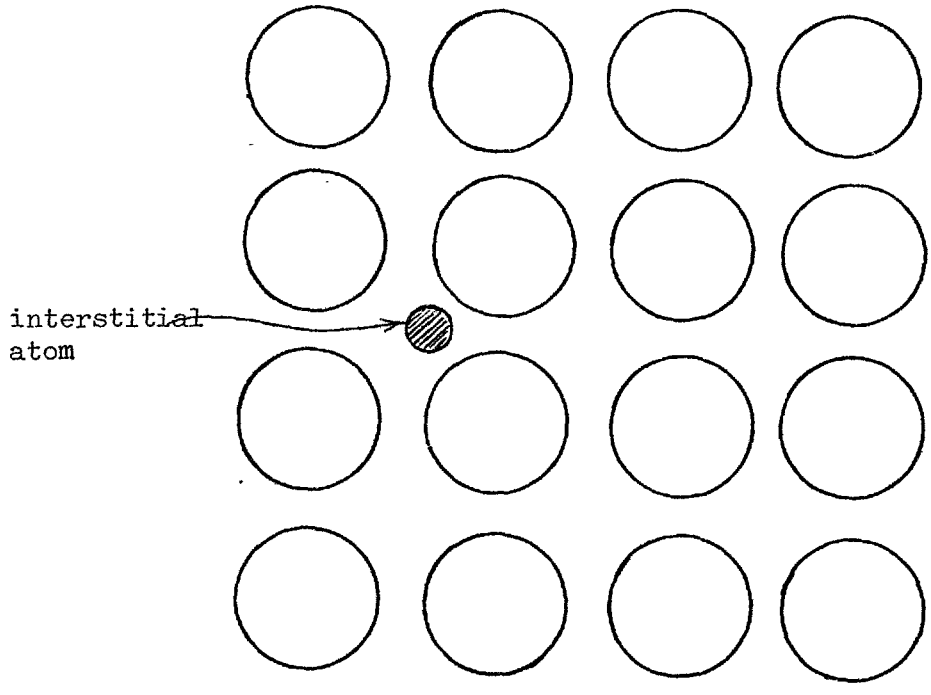


Fig. 2.6 Schematic illustration of interstitial diffusion in solids

where σ_i is the conductivity of particles of type i. If a crystal contains different types of charged carriers, the total current density J is given by

$$J = \sigma E \tag{2.68}$$

where σ represents the total electrical conductivity, and is given by

$$\sigma = \sum_i \sigma_i$$

Also σ_i is related to σ through

$$\sigma_i = t_i \sigma \tag{2.69}$$

where t_i is called the transference or transport number of the species i. In an inorganic compound the total electrical conductivity is given by the sum of the electronic and ionic conductivities,

$$\sigma = \sigma_{ion} + \sigma_{e\ell} = \sigma (t_{ion} + t_{el}) \tag{2.70}$$

where t_{ion} and $t_{e\ell}$ represent the transference numbers of ions and electrons respectively. It may be noted that $t_{ion} + t_{e\ell} = 1$. The current density J_i of the particles of type i, is related to their drift velocity, v_i, through

$$J_i = c_i q_i v_i = c_i z_i e v_i \tag{2.71}$$

where c_i is the concentration of the particles and z_i is the valence. The charge mobility μ_i is defined as the velocity in unit electric

field

$$\mu_i = \frac{v_i}{E} \tag{2.72}$$

so that the conductivity is also given by

$$\sigma_i = z_i \, e \, c_i \, \mu_i \tag{2.73}$$

2.2.4. A Simple Model for Ionic Mobility and Conductivity

A simple model, which relates the mobility of the defect to the height of the energy barrier it must surmount in moving between equivalent neighbouring sites, has long been used in the theory of ionic conductivity (Kofstad, 1972).

Once again as in section 2.2.1. a series of atomic planes at spacing a is considered. In jumping from a position in one plane to an equivalent site in the next, a defect has to surmount an energy barrier of height ΔW_o, as shown in Fig. 2.7. In the absence of any applied external perturbation, a carrier in any plane has a probability per unit time of jumping, or an average jump frequency, of p, which is given by

$$p = A \exp (- \Delta W_o / kT) \tag{2.74}$$

where A is a constant related to the frequency of the highest active lattice vibration mode of the crystal at temperature T. The number of defects jumping from plane 1 to 2 and from 2 to 1 per unit time per unit volume are equal and are each given by $\frac{1}{2} p \, c_i \, a$. In a homogeneous system there will thus be no net transport of particles.

If an electric field E is applied normal to the planes, a force

z_i e E is exerted on the defects, and accordingly the jump frequencies of the charged defects are changed. The jump frequency in the positive direction will be enhanced by the factor exp (z_i e aE/2 kT) and in the negative direction it will be decreased by the factor exp (z_i e aE/2 kT). As illustrated in Fig. 2.7 the electric field may be considered to cause a decrease in activation energy for migration in the positive direction by z_i e a E/2 and an increase by the same amount in the negative direction. The net flow of particles is given by

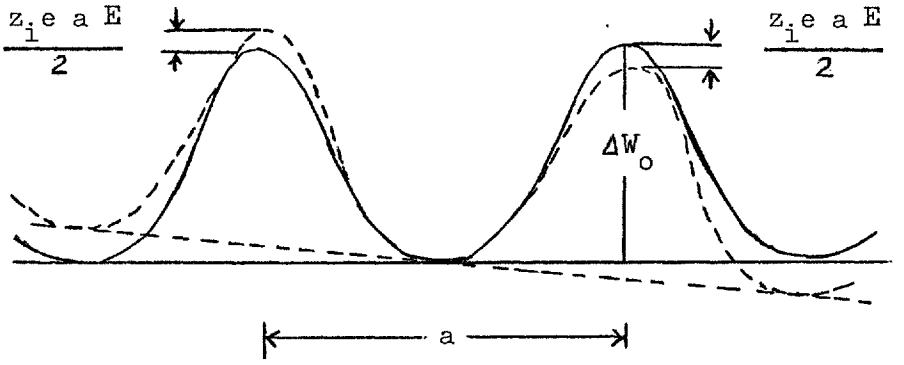


Figure 2.7. Effect of electric field in a homogeneous crystal

$$j = \frac{1}{2} p c_i a^2 \left\{ \exp\left(\frac{z_i e a E}{2 kT}\right) - \exp\left(-\frac{z_i e a E}{2 kT}\right) \right\} \qquad (2.75)$$

and for weak fields z_i a E $\ll$ 2 kT

$$j = p a^2 c_i \frac{z_i e E}{2 kT} \qquad (2.76)$$

The net flow of particles may also be expressed by

$$j = c_i v_i = c_i \mu_i E \qquad (2.77)$$

comparing equations (2.76) and (2.77) one sees that

$$\mu_i = \frac{z_i e\, a^2\, p}{2\, kT} \tag{2.78}$$

and hence by substituting (2.78) in (2.73) that

$$\sigma_i = c_i\, z_i e\, \mu_i = \frac{c_i\, z_i^2\, e^2\, a^2}{2\, KT} p \tag{2.79}$$

In view of equation (2.74) p and hence $\mu_i T$ and $\sigma_i T$ are all activated with energy ΔW_o, the inter-site barrier height.

In addition, however, one must remember that the defect concentration c_i also varies with temperature and the activation energy of c_i is therefore added to that of p when considering σ.

2.2.5. The Nernst-Einstein Relationship between μ and D

The above simple derivation for μ_i may be combined with the considerations of section 2.2.1. to yield the Nernst-Einstein relationship. Combining eqns. (2.78) and (2.63) one sees that

$$\frac{\mu_i}{D_i} = \frac{z_i e}{kT} \tag{2.80}$$

which is the required relationship.

One may note that all the parameters of the particular microscopic model which was used have been eliminated in this relationship, and indeed by general thermodynamic arguments it may be shown to be universally applicable regardless of the model used for calculating μ and D (Shewmon, 1963).

Experimentally determined diffusion coefficient are usually expressed as

$$D = D_o \exp\left(-Q/RT\right) \tag{2.81}$$

where Q is called the activation energy.

From (2.80) one sees that the activation energies of μT and D should be the same, while from (2.79) the activation energy of σT is increased above their common value by the activation energy of c_i.

2.3. Macroscopic Description of Time Varying Electrical Properties

It was mentioned in the introductory paragraphs of this chapter that the occurrence of polarization due to the movement of charge carriers over limited distances is significant for the frequency variation of conductivity and the time variation of currents under constant applied potential. Hence in the next section we revise the formal macroscopic ideas of polarization, susceptibility and permittivity, and go on in subsequent sections to review the ideas of complex permittivity and conductivity, their relationships to ordinary permittivity and conductivity, and to the time response of the material to a step function field. Finally the relationship between the real and imaginary parts of complex parameters is discussed.

2.3.1. Polarization, Susceptibility and Permittivity

These basic ideas will be described by referring to a parallel plate capacitor, as shown in Fig. 2.8, consisting of a pair of plates area A spaced distance d apart, across which a potential V is applied, resulting in total charges $\pm Q$ being stored on the plates. Hence the capacitor has capacitance $C = Q/V$. For this discussion of principle , we shall neglect fringing effects and assume that a uniform field E is set up between the plates.

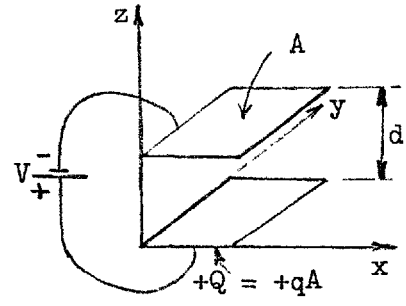


Figure 2.8

It is well known that if a dielectric medium is placed between the plates the capacitance will be increased to $C' = \epsilon' C$, where ϵ' is known as the (relative) permittivity of the medium. It will be shown in this section that this result may be explained on a macroscopic basis by the simple assumption that every elemental volume of the material takes up a dipole moment per unit volume, or polarization, P, which is given by

$P = \epsilon_o \chi' E$ where $\epsilon_o \chi'$ is a constant called the absolute susceptibility and χ' is the relative susceptibility. ϵ_o is a constant occurring in the M.K.S. system of units which is called the permittivity of free space and has the value 8.854×10^{-12} F/m.

Considering the empty capacitor, the charge densities on the inside surfaces of the plates are

$$\pm q = \pm Q/A \tag{2.82}$$

By Gauss Theorem of Electrostatics (i.e. eventually as a consequence of Coulomb's Law) the field due to a surface charge density q on the surface of a conductor is perpendicular to the surface and of magnitude $q/2\epsilon_o$, and hence in the space between the plates, adding the contributions from the two plates, it follows that

$$E = q/\epsilon_o = Q/A\epsilon_o \tag{2.83}$$

By the definition of V in terms of E it also follows that

$$E = V/d \tag{2.84}$$

Combining these equations it follows that

$$\frac{Q}{V} = \frac{\epsilon_o A}{d} \tag{2.85}$$

and since Q/V is the definition of capacitance C then

$$C = \frac{\epsilon_o A}{d} \tag{2.86}$$

for the empty parallel plate capacitor.

Now consider the capacitor filled with a medium. Consider a volume element dxdydz of the medium, as shown in Fig. 2.9, which takes up a dipole moment Pdxdydz under the influence of a field E where we assert, as mentioned above, that

$$P = \epsilon_o \chi' E \tag{2.87}$$

If the dipole moment is due to the separation of "bound" charges through a limited distance then if $\pm q_b$ are the bound charges per unit area appearing on the upper and lower elemental faces of area dxdy, then the dipole moment of the volume element is also given by $(q_b dxdy)dz$ and it follows that

$$q_b = P \tag{2.88}$$

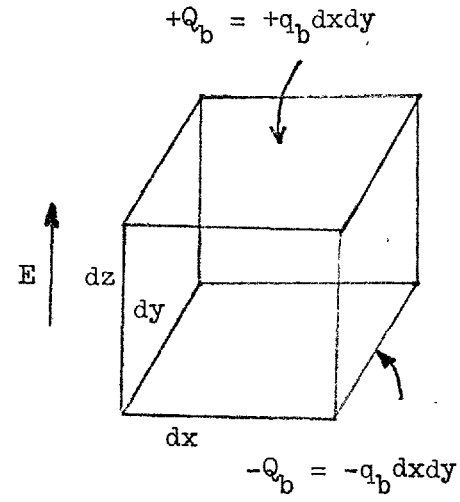


Figure 2.9

If now we consider the piling up of the elemental blocks until they fill the capacitor as shown in Fig. 2.10, then clearly there will be a charge density q_b over the whole face of the dielectric facing the metal electrode. (The charges $\pm q_b$ on the neighbouring faces of elements piled in the z-direction add to give zero volume charge density.)

If the potential difference between the plates is maintained at the same value as for the empty case, then the field E = V/d also still remains the same. However, this field must still, according to Gauss' Theorem, be given by

$$E = 2(q/2\epsilon_o) \qquad (2.89)$$

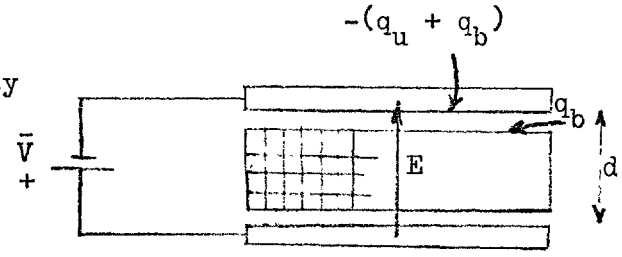


Figure 2.10

where q is the total sheet charge density at the interface of the metal electrode and the dielectric. Hence it is clear that extra charge numerically equal but of opposite sign to q_b is stored on the metal surface in addition to the previous charge stored with no medium, which is now referred to as q_u, the unbound charge, so that the total charge surface density on the plate is

$$q_t = q_u + q_b \qquad (2.90)$$

Hence from an external point of view, on applying potential V the capacitor stores charges $\pm Q'$ in the plates where

$$Q' = A\,(q_u + q_b) = Q_u + Q_b \qquad (2.91)$$

But $$Q_u = CV = \frac{A\epsilon_o V}{d} \qquad (2.92)$$

and $$Q_b = Aq_b = AP = A\,\epsilon_o \chi' E = \frac{A\epsilon_o \chi' V}{d}$$

Hence $$\frac{Q'}{V} = \frac{A\epsilon_o}{d}\,(\chi' + 1) = (\chi' + 1)C$$

and since $C' = Q'/V$ it has been shown as stated at the beginning that

$$C' = \epsilon' C$$

where $$\epsilon' = \chi' + 1 \qquad (2.93)$$

2.3.2. Complex Permittivity and Conductivity

The concepts of complex permittivity and conductivity are a special case of the idea of a complex admittance or impedance which arises from the use of the complex notation for the harmonic response of a linear system. In the case of a linear electrical system, harmonic currents and voltages are represented by complexors of the form

$$v = \mathcal{R}V_m e^{j(\omega t+\phi_1)} = \mathcal{R}V_m \left[\cos(\omega t+\phi_1) + j \sin(\omega t+\phi_1)\right] = V_m \cos(\omega t+\phi_1) \qquad (2.94)$$

and

$$i = \mathcal{R}I_m e^{j(\omega t+\phi_2)} = \mathcal{R}I_m \left[\cos(\omega t+\phi_2) + j \sin(\omega t+\phi_2)\right] = I_m \cos(\omega t+\phi_2) \qquad (2.95)$$

where $\mathcal{R}$ means "the real part of", V_m and I_m are the maximum values of V and I respectively, $\omega = 2\pi f$ where f is the frequency and ϕ_1 and ϕ_2 are phase angles referred to some arbitrary zero. The use of this complex notation is explained in many text-books on electrical engineering (e.g. Cassell, 1964). It is customary to drop the symbol $\mathcal{R}$ from the equations, and also to drop the qualifying term "complex" from voltages and currents when it is clear that harmonic quantities are being dealt with by the complex notation.

If a (complex) voltage v applied across a circuit element, as in Fig. 2.11, causes a (complex) current i to flow through it then the element is said to have (complex) impedance Z or equivalently (complex) admittance Y given respectively by

$$Z = \frac{v}{i} = \frac{V_m e^{j(\omega t+\phi_1)}}{I_m e^{j(\omega t+\phi_2)}} = \frac{V_m}{I_m} e^{j(\omega t+\phi_1-\phi_2)} \qquad (2.96)$$

and

$$Y = \frac{i}{v} = \frac{I_m}{V_m} e^{j(\omega t + \phi_2 - \phi_1)} \qquad (2.97)$$

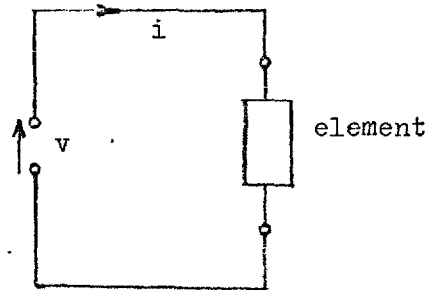


Figure 2.11

Equivalent forms of writing these results are

$$Z = |Z| \angle \phi_1 - \phi_2 = R + jX \qquad (2.98)$$

$$Y = |Y| \angle \phi_2 - \phi_1 = G + jB \qquad (2.99)$$

where $R = |Z| \cos(\phi_1 - \phi_2)$ is called the resistance

$X = |Z| \sin(\phi_1 - \phi_2)$ is called the reactance

$G = |Y| \cos(\phi_2 - \phi_1)$ is called the conductance

and $B = |Y| \sin(\phi_2 - \phi_1)$ is called the susceptance.

In electrical engineering (and in linear systems theory more generally) ideal elements are defined which impose a certain relationship between stimulus and response. An ideal capacitor is one for which the relationship

$$Q = CV$$

applies instantaneously. Since the definition of current is $i = dQ/dt$ it follows that for an ideal capacitor

$$i = C dv/dt \qquad (2.100)$$

(lower case letters imply variation with time).

The translation of this relationship into complex notation shows immediately that for an ideal capacitor defined in this way the impedance and admittance are in fact a reactance and susceptance (i.e. the real parts are zero) and in fact

$$Z_c = -jX_c \quad \text{where} \quad X_c = {}^{1}/_{\omega C}$$

or $$Y_c = jB_c \quad \text{where} \quad B_c = \omega C$$

Similarly the ideal resistor is defined to have v-i relationship

$$v = iR \tag{2.101}$$

from which it follows that

$$Z_R = R$$
$$Y_R = {}^{1}/_{R} = G$$

i.e. the impedance and admittance are entirely real.

It also follows by the simple application of Kirchoff's Laws that an ideal resistor and capacitor connected in series and considered as a single element has an impedance $Z = R - j/\omega C$, while if connected in parallel they behave as a single element of admittance $Y = {}^{1}/_{R} + j\omega C$.

The admittance of a parallel plate capacitor filled with a medium of permittivity $\epsilon'(\omega)$ and having a conductivity $\sigma'(\omega)$ will now be used to develop the concepts of complex permittivity and conductivity. The notation $\epsilon'(\omega)$ and $\sigma'(\omega)$ emphasises that these quantities themselves may vary with frequency.

The capacitor will have a capacitance $C = \epsilon'(\omega)\,\epsilon_o\,A/d$ and a conductance $G = \sigma'(\omega)\,A/d$. The geometry of the situation ensures that the

current density throughout the slab is uniform, like the field, and externally the total current observed will therefore be the sum of the currents which C and G would pass separately alone, i.e. the admittance of the capacitor is the parallel combination of its capacitive susceptance and its conductance;

$$Y = G + j\omega C = \left\{ \frac{\sigma'(\omega)}{\epsilon_o} + j\,\omega\,\epsilon'(\omega) \right\} \epsilon_o \frac{A}{d} \tag{2.102}$$

If, in a purely formal manner we were to define a material to have a complex (relative) permittivity ϵ^* with real and imaginary parts ϵ' and $-\epsilon''$ so that

$$\epsilon^* = \epsilon' - j\,\epsilon'' \tag{2.103}$$

then the admittance of a parallel plate capacitor filled with this material, proceeding formally, would be

$$Y = j\omega\,\epsilon^*\,\epsilon_o \frac{A}{d} = \left\{ j\,\omega\,\epsilon' + \omega\epsilon'' \right\} \frac{\epsilon_o A}{d} \tag{2.104}$$

Comparing real and imaginary parts of these expressions it follows that

$$\sigma'(\omega) = \omega\,\epsilon_o\,\epsilon'' \tag{2.105}$$

and ϵ' has the ordinary meaning of relative permittivity previously attributed to it.

Similarly, defining a complex conductivity $\sigma^*(\omega)$ with real and imaginary parts σ' and σ'' so that

$$\sigma^* = \sigma' + j\sigma'' \tag{2.106}$$

then the admittance of a slab sample between parallel plate electrodes would be

$$Y = \frac{\sigma^* A}{d} = (\sigma' + j\sigma'') \frac{A}{d} \tag{2.107}$$

Comparing this to Eqn. 2.104 we see that the real part is the conductivity as previously defined while

$$\sigma'' = \omega \epsilon_0 \epsilon' \tag{2.108}$$

Briefly, we may lump together the conductivity and permittivity into either a complex permittivity or a complex conductivity. These are alternative descriptions of the material, neither containing more or less information than the other, and in fact from the above it is clear that

$$\sigma^*(\omega) = j\omega \epsilon_0 \epsilon^*(\omega) \tag{2.109}$$

The real point of this notation, as we shall show in the next section, is that it is straightforward to derive either ϵ^* or σ^* from the response of the slab sample to a time step function of electric field, and in turn it is often simpler to analyse microscopic models for their time step function response than to analyse them directly for their harmonic response.

The ideas of complex permittivity and conductivity have been developed above by considering the properties of a macroscopic slab sample in which the field and current density are uniform throughout. In other words, the current density J and field E at each point are connected by the relations

$$J = j\omega\epsilon^* \epsilon_0 E \qquad (2.110)$$

or equivalently $J = \sigma^* E$ (2.111)

As a practical point, it must be mentioned that if the conditions are not uniform throughout the sample (e.g. because of effects at the electrodes) then it may not be justifiable to derive ϵ^* or σ^* from a measurement of Y for a slab sample and inversion of Eqn. 2.104. One may have, for example, to consider that ϵ^* or σ^* varies with position in the slab, and try to deduce the overall dependence of Y on ϵ^* and σ^*. These aspects will be considered further in the discussion of the results.

2.3.3. Current and Polarisation Step Time Function Responses and Their Relationship to ϵ^* and σ^*.

Again, the behaviour we are considering here is a particular example of linear system behaviour as applied to charge transport in materials. We shall therefore proceed by discussing the response of a linear system to a time varying stimulus in general terms and finally specialising to the electrical case. If F(t) is an arbitrarily time varying stimulus we shall show that r(t), the response to F(t), may be derived from A(t), the response to a special stimulus function u(t), the unit step function, provided the system is linear.

The unit step function is defined by

$$u(t) = 0 \text{ for } t < 0$$

and $$u(t) = 1 \text{ for } t \geqslant 0$$

and is represented graphically in Fig. 2.12, with a typical step response function A(t).

The system is said to be linear or have a linear region for a certain range of F(t) if for that range it is found that

(a) The response to a u(t) is a A(t).

(b) The response to a $u(t-t_1)$ + $b\,u(t-t_2)$ is a $A(t-t_1)$ + $b\,A(t-t_2)$.

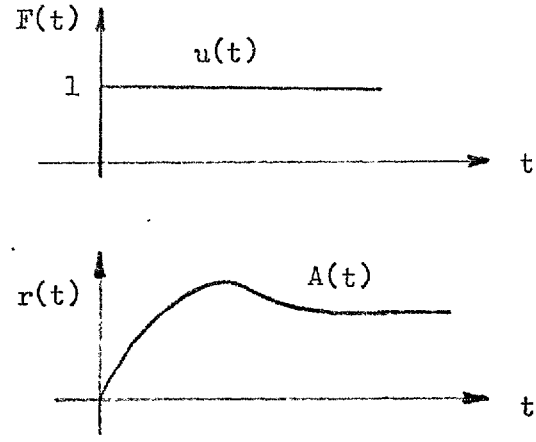


Figure 2.12

The compounding of a response r(t) as a result of applying two step functions at successive times is shown in Fig. 2.13.

An arbitrary function F(t) can be approximated by adding incremental step functions starting at intervals dt along the time axis, as shown in Fig. 2.14. When dt approaches zero, the increment in F(t) tends to the product of the slope of F(t) in the given interval and dt.

$$dF(t) = \left[\frac{dF(t)}{dt}\right] dt$$

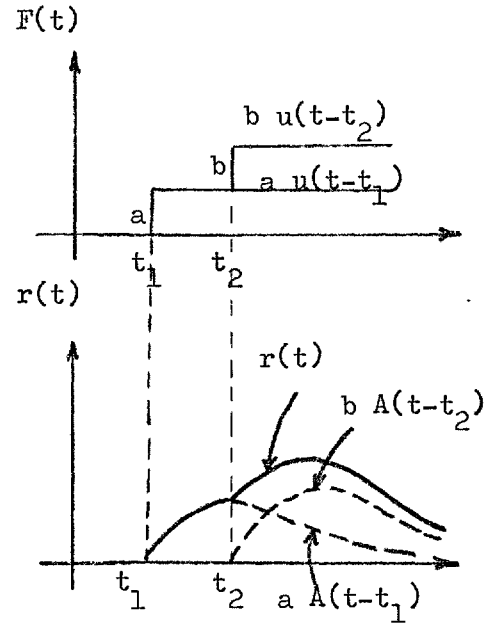


Figure 2.13

The response r(t′) at a time t′, due to the step dF(t) at time t, will be proportional to the step height and to the value of the unit step response A(t′-t) at the elapsed time t′-t later, as shown in magnified detail in Fig. 2.14.

Hence $dr(t') = \frac{dF(t)}{dt}\, A(t'-t)\, dt$

and $$r(t') = \int_{-\infty}^{t'+} \frac{dF(t)}{dt}\, A(t'-t)\, dt$$

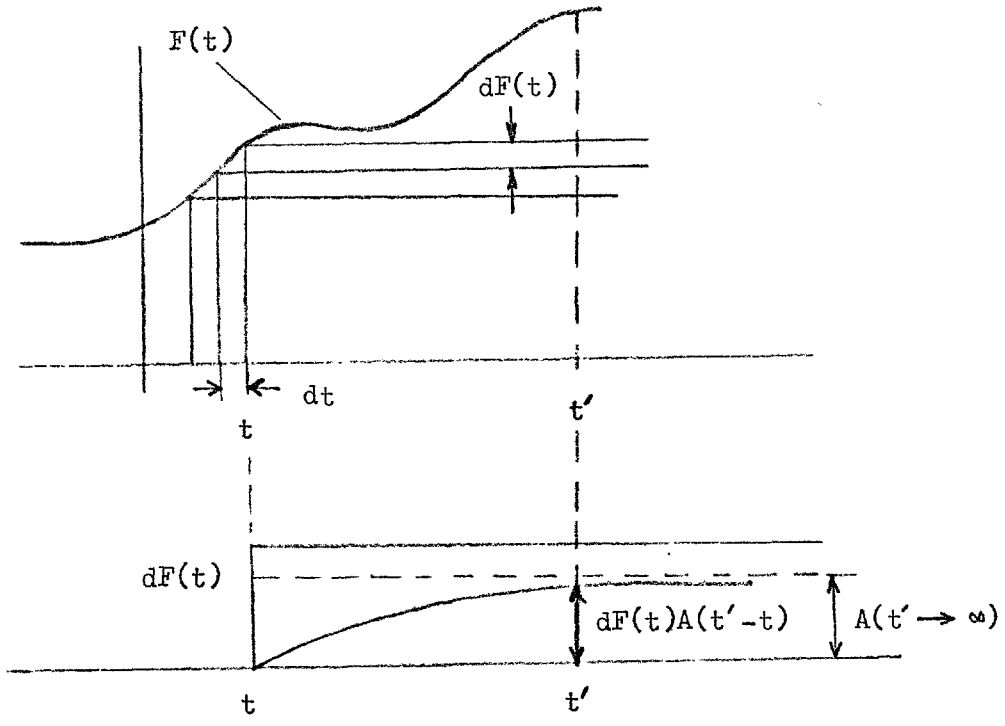


Figure 2.14

Let us substitute $\tau = t' - t$

$$r(t') = \int_{\infty}^{0^-} \frac{dF(t'-\tau)}{d\tau} A(\tau)\, d\tau = - \int_{0^-}^{\infty} \frac{dF(t'-\tau)}{d\tau} A(\tau) d\tau \qquad (2.112)$$

if $F(t) = F_m e^{st}$ where $s = p + j\omega$ then

$$r(t) = - \int_{0^-}^{\infty} F_m \frac{d}{d\tau} (e^{st} e^{-s\tau}) A(\tau)\, d\tau$$

$$= sF_m e^{st} \int_{0^-}^{\infty} e^{-s\tau} A(\tau)\, d\tau = sF_m e^{st} \mathcal{L}[A(\tau)] \qquad (2.113)$$

since the integral is the Laplace transform of $A(\tau)$ by definition.
If $p \longrightarrow 0$, $F(t)$ becomes $F_m e^{j\omega t}$ and we get $r(t)$ for a sinusoidal excitation, and it is of the form $r_m^* e^{j\omega t}$. Hence

$$r_m^* e^{j\omega t} = F_m e^{j\omega t} \lim_{p \to 0} \left\{ s \mathcal{L}[A(\tau)] \right\} \qquad (2.114)$$

Defining $A^*(\omega)$ as r_m^*/F_m, i.e. the complex parameter related to $A(t)$, then

$$A^*(\omega) = \frac{r_m^*}{F_m} = \lim_{p \to 0} [s\, A(s)] \qquad (2.115)$$

In particular, if $A(t)$ is the step response of the polarization A_p then

$$A^*(\omega) = \frac{P_m^*}{E_m} = \epsilon_o X^* \qquad (2.116)$$

the absolute complex susceptibility. On the other hand if $A(t)$ is the step response of the current $A_j(t)$, then

$$A^*(\omega) = \frac{J_m^*}{E_m} = \sigma^* \qquad (2.117)$$

the complex conductivity.

In the above development the limits t+ and 0- were used rather than simply t and 0 to take account of the possibility of A(t) containing a delta function at t = 0. The necessity for this will become apparent in what follows.

In applications of the above theory we shall be interested in two sorts of step response: (i) The polarization step response, to be denoted $A_p(t)$ hereafter, corresponding to the ϵ^* description of a material, and (ii) The current density step response denoted $A_j(t)$, corresponding to the σ^* description. These step responses tend to be used in association with different idealised microscopic models, for which they are particularly convenient, but some practical difficulties arise with real materials, and we now discuss these aspects.

(i) Polarisation step responses characteristic of models with limited charge displacement only.

Generally since P is due to limited charge displacement, which takes some time to occur after the application of E, then as shown in Fig. 2.15, $A_p(t)$ will approach a limit χ_s as t goes to ∞, where χ_s is the static susceptibility. Charge displacement may be monotonic as in (1) or may overshoot the final value giving curve (2). There may also be a very fast initial displacement which we may be able to call instantaneous for the range of speeds we are interested in, giving an initial step rise to χ_∞.

(ii) Current density step responses characteristic of models with unlimited charge displacement. Typical responses are illustrated in Fig. 2.16. By definition $\lim_{t\to\infty} A_j(t) = \sigma_s$, the static or d.c. conductivity. Usually $A_j(t) > \sigma_s$ because as time goes on space charge builds up and reduces E in part of the sample thickness, so that J reduces there also. Since as

$t \to \infty$, J becomes continuous across the sample thickness then if J reduces where E reduces it has finally to reduce everywhere. Another way of saying this is that there is a contribution to J proportional to dP/dt, but as we have just seen $dP/dt \to 0$, as $t \to \infty$.

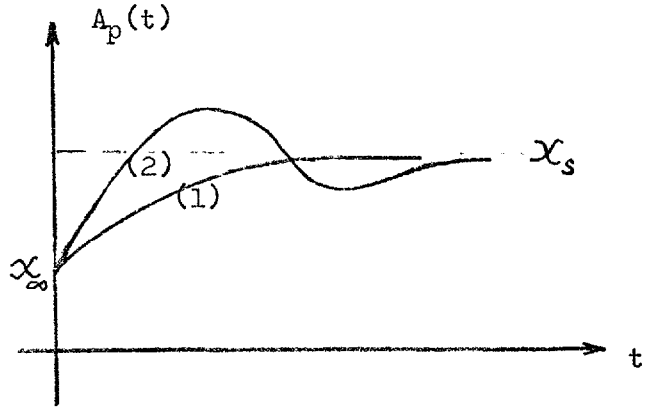


Figure 2.15

In actual observation of step responses we tend to apply a step voltage across a slab sample and observe I(t), as shown in Fig. 2.17. Then we assume that

$$A_j(t) = \frac{I(t)}{A} \Big/ \frac{V}{d} = \frac{I}{V} \frac{d}{A}$$

and we deduce $A_p(t)$ as $A_p(t) = \frac{d}{VA} \int_0^t I(t')\, dt'$

In practical materials P(t) will contribute to $A_j(t)$ observed as follows and illustrated in Fig. 2.18. A current density $\frac{dP}{dt}$ has to flow on to the plates to supply the compensating charges for P. At t = 0, $\frac{dP}{dt} \to \infty$ because of χ_∞ and the current necessary

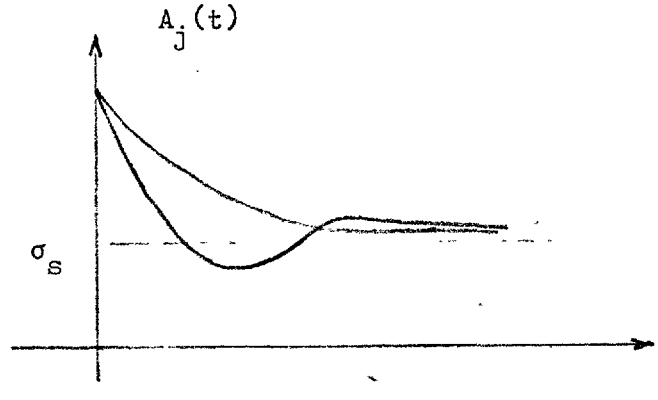


Figure 2.16

to supply the unbound charges, that is, to charge the vacuum capacitance. This means a very short, high pulse of current, whose integral with time is $\chi_\infty + 1$, must flow just after t = 0. Mathematically, this can be represented by a current

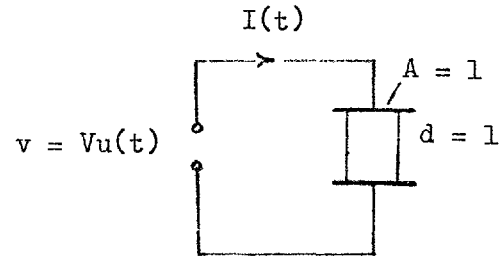


Fig. 2.17

$$(\chi_\infty + 1)\,\delta(t)$$

where $\delta(t)$ is the delta function.

In addition any dA_p/dt due to electronic or lattice polarisability contributes a charging current to A_j in addition to space-charge build up or dynamic properties of conduction processes. On the other hand, the final steady current in a real material would make the polarization step response deduced by integrating the current tend to a linearly rising function of time as shown in Fig. 2.19.

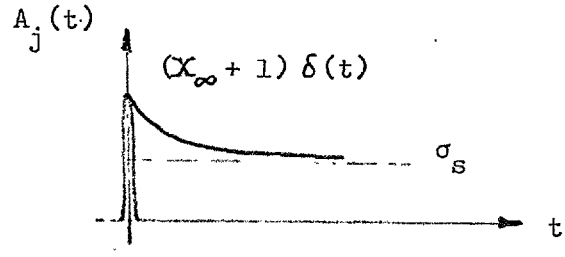


Fig. 2.18

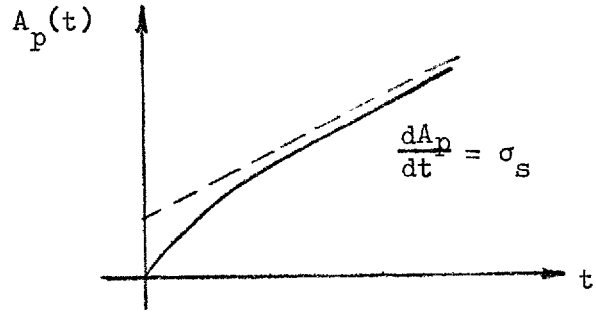


Fig. 2.19

In analysing a microscopic model for $A_j(t)$ or $A_p(t)$ in order to find $\epsilon^*(\omega)$ or $\sigma^*(\omega)$ by applying the Laplace Transform relation just deduced, due account must be taken of these complications. The actual procedures adopted in this work will be explained where they are applied.

2.3.4. The Inter-relationship of ϵ' and ϵ'' - The Kramers-Kronig Relations

In the preceding section we showed that the absolute complex susceptibility $\epsilon_o \chi^*(\omega)$ is given by

$$\epsilon_o \chi^*(\omega) = \lim_{p \to 0} s \mathcal{L}\{A_p(t)\}$$

$$= \lim_{p \to 0} s \int_o^\infty e^{-st} A_p(t)\, dt \qquad (2.118)$$

This complex susceptibility has real and imaginary parts, i.e.

$$\epsilon_o \chi^* = \epsilon_o \chi' - j \epsilon_o \chi'' \qquad (2.119)$$

Since χ' and χ'' are each related by an integral transform to the same function $A_p(t)$ there must in turn be a direct relationship between them, which again turns out to be an integral transform.

To demonstrate this relationship Eqn. 2.118 can be considered, before taking the limit, as stating that

$$\epsilon_o \chi^*(s) = s \int_o^\infty e^{-st} A_p(t)\, dt \qquad (2.120)$$

where $\chi^*(s)$ is a generalised complex susceptibility, which is a function of $s = p + j\omega$ and reduces to $\chi^*(\omega)$ as $p \to 0$.

This equation may be written

$$\epsilon_o \chi^*(s) = \int_o^\infty e^{-st} \left\{ \frac{dA_p(t)}{dt} \right\} dt = \int_o^\infty e^{-st} B_p(t)\, dt \qquad (2.121)$$

using the Laplace transform theorem relating the Laplace transform of a

function and its derivative (Cassell, 1964, p. 58) which is

$$\mathcal{L}\left\{\frac{dA_p}{dt}\right\} = s\mathcal{L}(A_p) - A_p(0) \qquad (2.122)$$

In the above, $B(t) = \frac{dA_p(t)}{dt}$ is actually the impulse function response of the polarisation, but this need not concern us further.

For the applications we shall be considering, if we exclude lattice and electronic polarizability from $A_p(t)$, then $A_p(0) = 0$, and there is no step or delta function at $t = 0$. The relationships we shall derive between χ' and χ'' will then strictly apply to the contributions to χ^* from defect motion only, which is sufficient for our purposes. In fact, the contribution to χ' from lattice and electronic polarisation can be added into the equations after. A more sophisticated approach (see Section 2.3.3.) would be to leave a step function in $A_p(t)$ at $t = 0$, and a delta function in $B_p(t)$ at $t = 0$, and define the Laplace transform with lower limit 0^-.

For the class of $B_p(t)$ corresponding to $A_p(t)$ with no step or delta functions, the limit of $p \rightarrow 0$ may be taken inside the integration in Eqn. (2.121) giving

$$\epsilon_o\chi^*(\omega) = \epsilon_o\chi' - j\epsilon_o\chi'' = \int_o^\infty e^{-j\omega t} B_p(t)\, dt \qquad (2.123)$$

or, separating real and imaginary parts

$$\epsilon_o\chi'(\omega) = \int_o^\infty B_p(t) \operatorname{Cos}\omega t\, dt = \int_o^\infty B_p(x) \operatorname{Cos}(\omega x)\, dx \qquad (2.124)$$

and

$$\epsilon_o\chi''(\omega) = \int_o^\infty B_p(t) \operatorname{Sin}\omega t\, dt = \int_o^\infty B_p(x) \operatorname{Sin}(\omega x)\, dx \qquad (2.125)$$

where the second forms may be written since t is only a "dummy" variable

which disappears after integration*.

Eqns. (2.123) and (2.124) are one-sided cosine and sine Fourier transforms respectively, and according to the Fourier inversion theorem may be inverted to give

$$B_p(x) = \frac{2}{\pi} \int_0^\infty \epsilon_0 \chi'(\omega) \operatorname{Cos}\omega x \, d\omega = \frac{2}{\pi} \int_0^\infty \epsilon_0 \chi'(\mu) \operatorname{Cos} \mu x \, d\mu \quad (2.126)$$

and $$B_p(x) = \frac{2}{\pi} \int_0^\infty \epsilon_0 \chi''(\omega) \operatorname{Sin}\omega x \, d\omega = \frac{2}{\pi} \int_0^\infty \epsilon_0 \chi''(\mu) \operatorname{Sin} \mu x \, d\mu \quad (2.127)$$

where again the dummy variable has been changed in the second form*.

Substituting Eqn. (2.127) into (2.124) one gets

$$\chi'(\omega) = \frac{2}{\pi} \int_0^\infty dx \left\{ \operatorname{Cos}\omega x \int_0^\infty \chi''(\mu) \operatorname{Sin} \mu x \right\} d\mu$$

$$= \frac{2}{\pi} \int_0^\infty d\mu \left\{ \chi''(\mu) \int_0^\infty \operatorname{Cos}\omega x \operatorname{Sin} \mu x \right\} dx \quad (2.128)$$

where firstly $\operatorname{Cos}\omega x$ has been taken inside the integration w.r.t. μ, and then the order of integration interchanged.

The inner integral is improper, since its value oscillates continuously as the upper limit leads to infinity. However, Eqn. (2.127) may be written as

$$\chi'(\omega) = \frac{2}{\pi} \lim_{R \to \infty} \int_0^\infty d\mu \left\{ \chi''(\mu) \int_0^R \operatorname{Cos}\omega x \operatorname{Sin} \mu x \right\} dx$$

$$= \frac{2}{\pi} \lim_{R \to \infty} \int_0^\infty \chi''(\mu) \frac{1}{2} \left(\frac{1 - \operatorname{Cos}(\mu + \omega)R}{\mu + \omega} + \frac{1 - \operatorname{Cos}(\mu - \omega)R}{\mu - \omega} \right) d\mu \quad (2.129)$$

Since $\chi''(\mu)$ is a slowly varying function, then the product of $\chi''(\mu)$ and

---

* The reason for making these variable changes is to avoid confusion between variables in Eqn. (2.128), as will become clear on reaching it.

either of the cosine terms becomes, for large R, a rapidly alternating function with slowly decreasing envelope as $\mu \to \infty$, and the integral from $\mu = 0$ to ∞ is zero. Hence Eqn. (2.129) becomes

$$\chi'(\omega) = \frac{2}{\pi} \int_0^\infty \chi''(\mu) \frac{\mu}{\mu^2 - \omega^2} \, d\mu \tag{2.130}$$

A similar manipulation of Eqns. (2.126) and (2.125) leads to

$$\chi''(\omega) = \frac{2}{\pi} \int_0^\infty \chi'(\mu) \frac{\omega}{\omega^2 - \mu^2} \, d\mu \tag{2.131}$$

2.4. Microscopic Models for A.C. Electrical Properties of Ionic Conductivity

In this section the analysis of some models for ionic charge carrier motion in solids is discussed. Two well-known much used models are reviewed, and a more detailed analysis given of a new model (Mason, to be published).

2.4.1. Motion of Ionic Defects

The concept of charged ionic defects moving around randomly in a crystalline lattice by thermal activation over energy barriers between equivalent sites has already been introduced in section 2.2. The consequences for diffusion in a concentration gradient and drift in a d.c. electric field were also analysed.

It will be shown in section 2.4.3. that this model predicts no variation of σ with ω except a monotonic decrease setting in at extremely high frequencies ($\sim 10^{13}$ Hz). This is quite contrary to general experimental experience, in which $\sigma(\omega)$ shows an increasing trend with ω, at least up to the MHz range. The model must therefore neglect at least one feature of ionic defect motion essential to the occurrence of frequency variation in σ.

The essential shortcoming seems to be the over-simplifying assumption that the depth of the potential well is the same at every site available to the defect. There are several reasons why this is not so.

Since the crystal is electrically neutral (at any rate with no field applied at electrodes to inject space charge) then there must be equal numbers of oppositely charged defects. If those of one sign and type are the predominantly mobile ones, then the sites available to them as nearest neighbours of the oppositely charged ones constitute a set of sites of lower potential energy. Similarly the next nearest neighbours form another set, of energy intermediate between that of nearest

neighbours and of still further removed sites. Again, uncharged defects may also provide sites of different energy.

Once the sites are no longer all considered to be equivalent, the one-dimensional nature of the treatment in section 2.2 is no longer justified, since drift in the field direction occurring as a result of jumps both along and across the field may differ from that in which jumps are constrained along the field direction. In the former case some mobile defects may avoid ever falling into lower energy wells as they traverse the crystal, but in the latter all mobile defects will in turn be compelled to pass through low energy sites in their path.

However, the treatments given in this section will be essentially one-dimensional as a first approach to the problem. They show that once a variable well depth is assumed then a $\sigma(\omega)$ which is finite at $\omega = 0$ and increases with ω is predicted, even for a one-dimensional treatment. The one-dimensionality may well be, however, a factor in the rather poor numerical agreement with the experimental results, although there are various other possible causes which will be examined in Chapter 5.

The basic physical reason why assuming a variable well depth leads to frequency variation of σ is as follows. With no field applied, the depths of energy wells for mobile defects on opposite sides of an attractive centre (oppositely charged immobile defect) will be the same, although lower than for further removed sites. Their average occupation numbers will therefore be equal.

If a step field is applied, the average occupation numbers for the up-field and down-field neighbouring sites will change with time until they reach a new steady state consistent with the jump probabilities and occupation densities at each site giving the product necessary to provide the final steady state current. In short, a polarization builds up with a time transient, and the arguments of section 2.3 show this will lead

to frequency variation in σ.

We now go on to consider the three models mentioned above.

2.4.2. The Isolated Double Potential Well Model

This model has long been used as a simple picture of relaxation processes in dielectrics. For our purposes it has the defect that it predicts no d.c. conductivity. It is assumed that a particle of charge

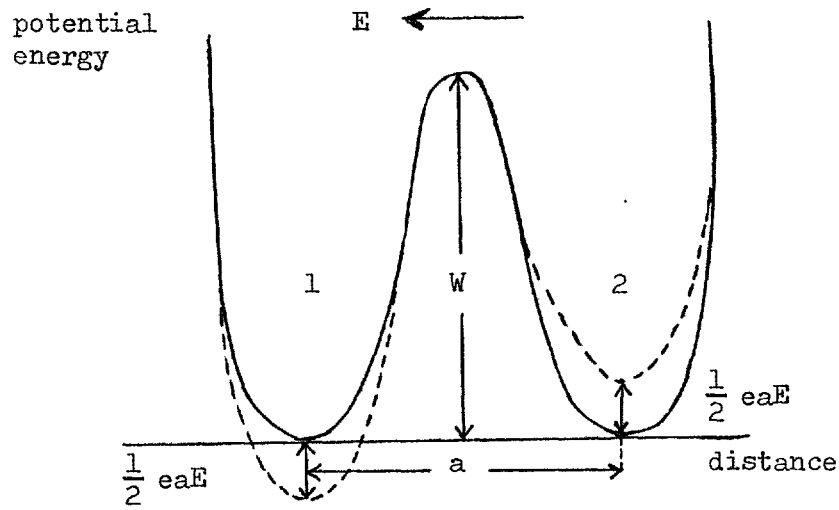


Fig. 2.20 Double potential well: the two wells contain one charged particle which may occupy either well.

e may be in one or other of two sites 1 and 2, located at distance a apart in the field direction. These sites are defined as minima of the potential energy as shown in Fig. 2.20 and the potential rises on each side to a very high value, in effect, ∞. An electric field E acting on the system causes a difference in the potential energy of the two sites and the figure is drawn with solid lines for the condition without field, and with dotted lines for the presence of a field.

The potential difference due to the field E is

$$V_1 - V_2 = eaE \qquad (2.132)$$

This model is equivalent to a dipole in that a movement of the charge from 1 to 2 or vice versa is equivalent to a turn by 180° of the angle of a dipole of moment

$$\mu = \frac{1}{2} ea \tag{2.133}$$

which one might imagine to be hinged about the centre point between 1 and 2. (There will, of course, be a compensating charge of opposite sign somewhere, providing the other half of the dipole.) There will be N_d bistable dipoles, where N_d is the number of defects of one sign per unit volume. We also assume equal potential energy for the sites 1 and 2 in the absence of an electric field. The model described would have no dynamic properties if it were on a macroscopic scale because the charged particles would not have the energy to jump the potential hill between the two potential wells 1 and 2. However, in a microscopic assembly the N_d bistable dipoles must be imagined as being located in a heat reservoir which consists of spontaneously active particles which exchange energy with each other and the dipoles. Hence the directions of the dipoles fluctuate. A charge carrier situated in a well 1 occasionally acquires an energy sufficient to lift it over the potential barrier, and drops into the well 2 associated with it. On arrival in 2 the energy of the carrier is returned to the heat reservoir, and the carrier then stays in 2 until such time as it acquires enough energy from the reservoir to return over the hill to 1. The probability of a jump by a charge in a double potential well can be derived from statistical thermodynamics. The number of dipoles jumping per unit time from 1 to 2 is given in terms of the difference of potential energy between the two wells as

$$p_{12} = A\, e^{-\frac{W - \mu E}{kT}} \tag{2.134}$$

where T is the absolute temperature, k is the Boltzmann constant and A a factor which may or may not depend on temperature. The equation presupposes that $W \gg kT$; the energy W is usually described as an activation energy. On the other hand, in general

$$\frac{\mu E}{kT} \ll 1 \tag{2.135}$$

and the expression for the frequency of jumps can be simplified to

$$p_{12} = p_o \left(1 - \frac{\mu E}{kT}\right) \tag{2.136}$$

where

$$p_o = A\, e^{-W/kT} \tag{2.137}$$

is the frequency of jumps in the absence of an applied field. The jump frequency in the opposite direction from 2 to 1, is given in the case sketched by

$$p_{21} = p_o \left(1 + \frac{\mu E}{kT}\right) \tag{2.138}$$

The average population of charges in wells 1 and 2 of the N_d bistable dipoles will not change with time if the number of charges jumping per unit time, from left to right, equals that jumping from right to left; namely, if

$$N_1 p_{12} = N_2 p_{21} \tag{2.139}$$

where N_1 is the number of occupied wells 1 and N_2 that of occupied wells 2. Since the total number of occupied wells per unit volume, that is

the total number of bistable dipoles, is constant, we have

$$N_1 + N_2 = N_d \tag{2.140}$$

Equations (2.139) and (2.140) permit the calculation of N_1 and N_2 in equilibrium, and hence the polarization in equilibrium. In general, the polarization per unit volume is given by that number of dipoles in one direction which is not compensated by dipoles in the opposite direction, namely

$$P = (N_1 - N_2)\,\mu \tag{2.141}$$

In the steady state $N_1 - N_2$ and hence $P = P_s$ can be calculated from equations (2.139) and (2.140), using equations (2.136) and (2.138), we get

$$N_1\, p_o\, (1 - \tfrac{\mu E}{kT}) = N_2\, p_o\, (1 + \tfrac{\mu E}{kT}) \tag{2.142}$$

which yields

$$N_1 - N_2 = (N_1 + N_2)\,\tfrac{\mu E}{kT} \tag{2.143}$$

Substituting (2.143) and (2.140) in (2.141), gives

$$P_s = N_d\,\tfrac{\mu^2 E}{kT} \tag{2.144}$$

The time-dependent properties of the model follow from the fact that the change in the number of dipoles in 1 is equal to the outflow to 2 less the inflow from 2, thus

$$\frac{dN_1}{dt} = - N_1 p_{12} + N_2 p_{21} \tag{2.145}$$

On the other hand, differentiating equation (2.140) with respect to time, yields

$$- \frac{dN_2}{dt} = \frac{dN_1}{dt} \tag{2.146}$$

$$\text{or} \quad \frac{d(N_1 - N_2)}{dt} = 2 \frac{dN_1}{dt} \tag{2.147}$$

and substituting (2.147) in (2.145) we have

$$\frac{1}{2} \frac{d(N_1 - N_2)}{dt} = - N_1 p_{12} + N_2 p_{21} \tag{2.148}$$

using equations (2.136) and (2.138), then

$$\frac{1}{2} \frac{d(N_1 - N_2)}{dt} = - N_1 p_o \left(1 - \frac{\mu E}{kT}\right) + N_2 p_o \left(1 + \frac{\mu E}{kT}\right) \tag{2.149}$$

$$= - p_o (N_1 - N_2) + p_o (N_1 + N_2) \frac{\mu E}{kT} \tag{2.150}$$

Combining the last equation with equation (2.141) yields

$$\frac{1}{2p_o} \frac{dP}{dt} + P = \frac{N_d \mu^2 E}{kT} = \frac{N_d e^2 E a^2}{4\, kT} \tag{2.151}$$

Rewriting equation (2.151) for $E = E_1(t)$

$$\frac{dP}{dt} + (2p_o)P = \frac{N_d e^2 a^2 p_o E_1}{2\, kT} \tag{2.152}$$

The general solution of this equation is

$$P = L\, e^{-2p_o t} + \frac{N_d e^2 a^2 E_1}{4\,kT} \tag{2.153}$$

Using the initial condition, P = 0 when t = 0, yields

$$L = -\frac{N_d e^2 a^2 E_1}{4\,kT}$$

Hence

$$\frac{P}{E_1} = A_p = \frac{N_d e^2 a^2}{4\,kT}\left[1 - e^{-2p_o t}\right] \tag{2.154}$$

On the other hand,

$$\epsilon_o \chi^* = \lim_{p \to o} \left\{ s\mathcal{L}\left[A_p(t)\right] \right\}$$

and remembering that

$$\mathcal{L}\left[e^{at}\right] = \frac{1}{s - a} \quad \text{and} \quad \mathcal{L}\left[u(t)\right] = \frac{1}{s}$$

therefore

$$\epsilon_o \chi^* = \lim_{p \to o} \left\{ s\, \frac{N_d e^2 a^2}{4\,kT} \left(\frac{1}{s} - \frac{1}{s + 2p_o}\right) \right\} = \frac{N_d e^2 a^2}{4\,kT}\, \frac{2p_o}{j\omega + 2p_o} \tag{2.155}$$

This equation may also be written

$$\epsilon_o \chi^* = K_1 \frac{1}{1 + j\Omega/2} \tag{2.156}$$

where $\Omega = \omega / p_o$ is the applied electric field <u>angular</u> frequency normalised to the spontaneous jump frequency of the carriers over the barrier in the absence of any field and $K_1 = \frac{N_d e^2 a^2}{4\,kT}$. (Using this "mixed" normalisation of an angular to ordinary frequency yields neater equations in later

developments.)

The real and imaginary parts of this equation are

$$\chi'\epsilon_o = K_1 \frac{1}{1 + \Omega^2/4} \tag{2.157}$$

and

$$\chi''\epsilon_o = K_1 \frac{\Omega/2}{1 + \Omega^2/4} \tag{2.158}$$

where

$$\chi^* = \chi' - j\chi''$$

The complex conductivity σ^* is given by $j\omega\epsilon_o\epsilon^*$ (see section 2.3.2) and hence

$$-j\omega\epsilon_o\sigma^* = jK_2 \frac{\Omega/2}{1 + j\Omega/2} = \sigma' + j\sigma'' \tag{2.159}$$

with real and imaginary parts

$$\sigma' = K_2 \frac{\Omega^2/4}{1 + \Omega^2/4} \tag{2.160}$$

and

$$\sigma'' = K_2\left[\frac{\Omega/2}{1 + \Omega^2/4} + \frac{\epsilon_o\Omega}{2}\right] \tag{2.161}$$

where

$$K_2 = \frac{2N_d e^2 a^2 p_o}{4kT} = 2p_o K_1$$

Measuring instruments are generally calibrated directly in terms of conductivity σ' and permittivity ϵ'. The characteristics of equations (2.160) and (2.157) for σ' and ϵ' may be most conveniently displayed on log-log plots. In Fig. 2.21 are shown plots of $\log(\epsilon_o\chi''/K_1)$ vs $\log\Omega$ and

log (σ'/K_2) vs logΩ.

This model predicts zero d.c. conductivity (on setting $\Omega = 0$ in eqn. (2.160)) and the log-log plot rises from $-\infty$ with a constant slope of 2 until $\Omega \sim \frac{1}{2}$, when it flattens out rapidly and remains constant as $\Omega \rightarrow \infty$. The log-log plot of ϵ' shows an inverse type of behaviour.

In general, real materials also show a d.c. conductivity. If we suppose for the moment that this has the arbitrary value K_2/K, then this value of σ_{dc} plots as a horizontal line at - log K on the log(σ/K_2) plot, as shown by the dotted line for K = 50 in Fig. 2.21. The total conductivity is the arithmetic sum of σ_{dc} and σ', and becomes as shown by the chain dotted line. The combined plot may also be considered as a plot of $(\sigma' + \sigma_{dc})/\sigma_\infty$ where σ_∞ is the limiting total conductivity as $\Omega \rightarrow \infty$. Similarly, a plot of $(\sigma' + \sigma_{dc})/\sigma_{dc}$ would have exactly the same shape but be shifted up the axis by log K. Note also that $K = (\sigma_\infty - \sigma_{dc})/\sigma_{dc} \sim \sigma_\infty/\sigma_{dc})$.

2.4.3. Multiple Identical Wells

The steady state properties of this model have been discussed in section 2.2.4. It can be shown simply that this model does not predict a conductivity rising with frequency, so a full mathematical analysis will not be given.

Consider the transient response to a step electric field. Fig. 2.22(a) shows the energy diagram before the field is applied. The probabilities of a defect occupying any one of the wells are all equal, with common value, say f. The probability per unit time of a particle in well n jumping to well n + 1 or n - 1 is also the same for all wells, say with value p. Hence the forward and reverse defect fluxes are the same over each barrier, and the same for all barriers with value fp.

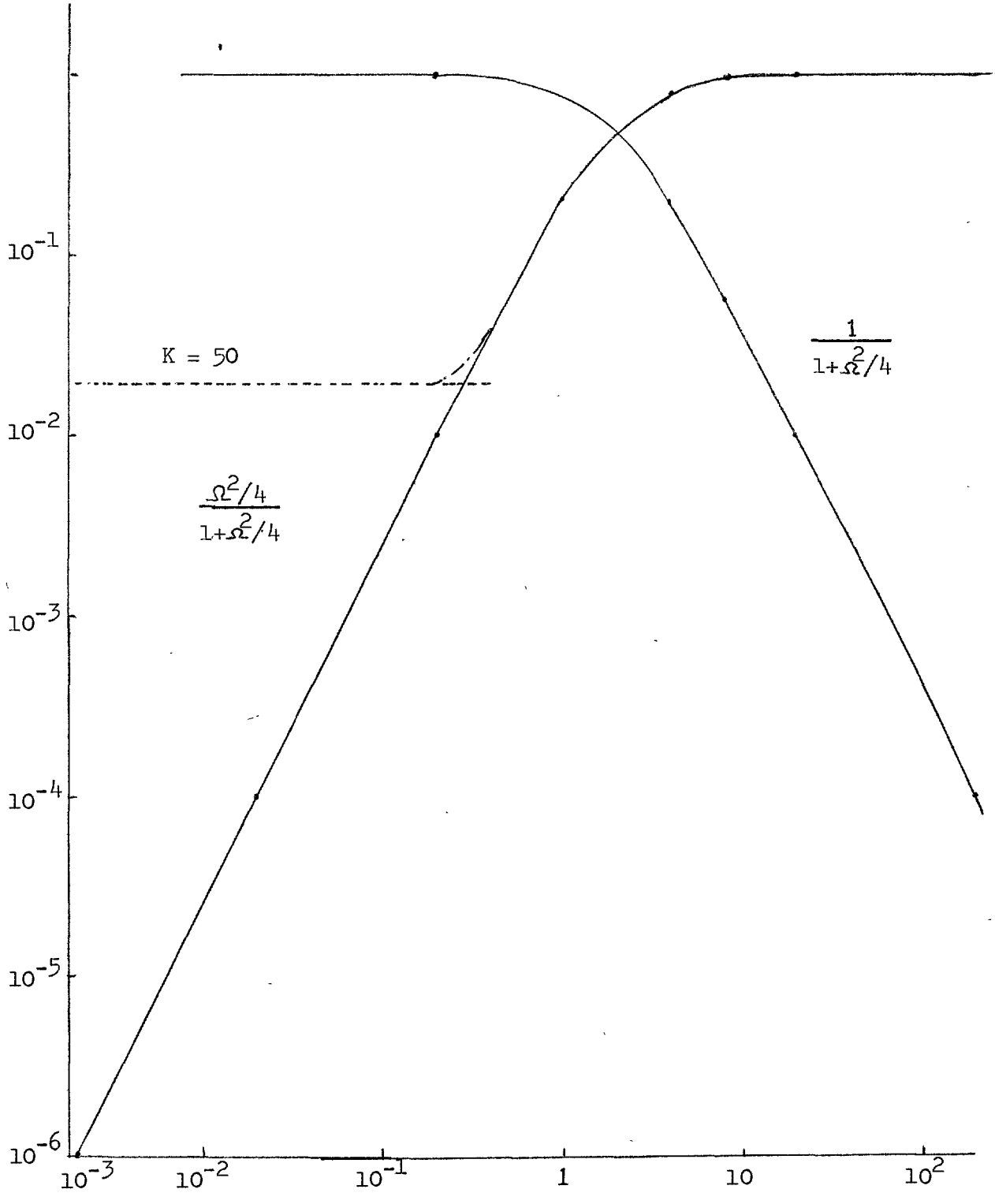


Figure 2.21 Isolated double well

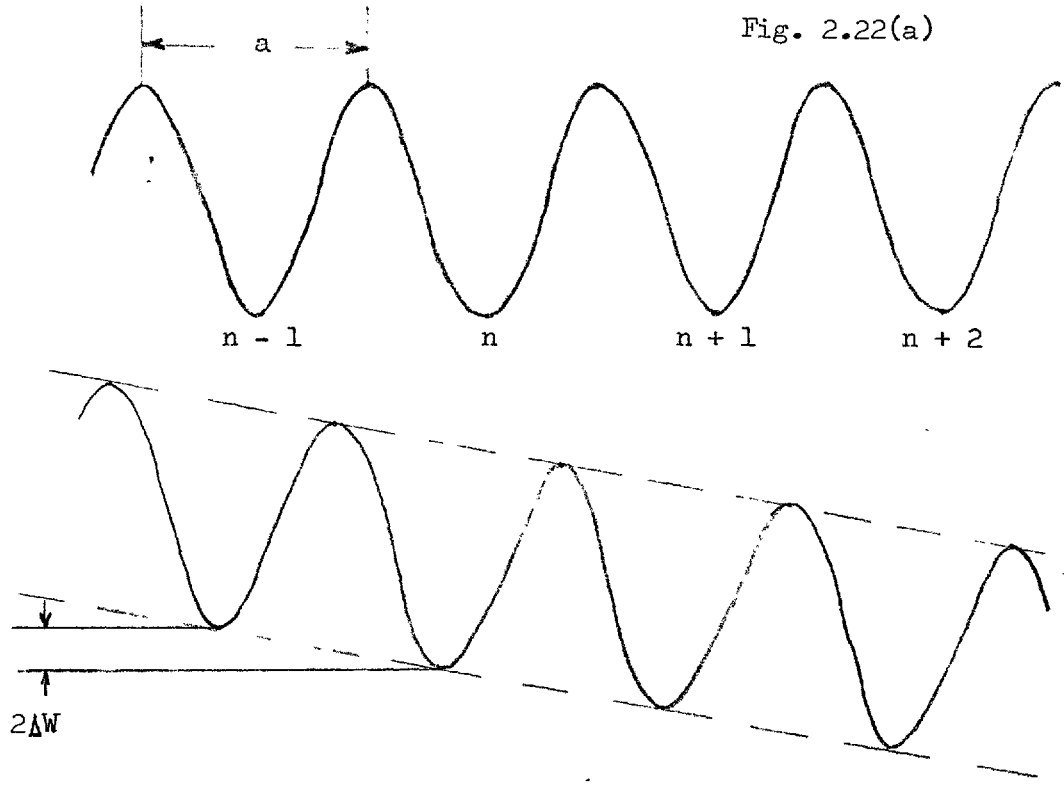


Fig. 2.22(a)

Fig. 2.22(b)

On applying the step electric field, the energy diagram takes on an average slope, as shown in Fig. 2.22(b). The heights of the barrier for downfield and upfield jumps are decreased and increased respectively by $\Delta W = zeaE$, where z is the effective charge of the defect, and hence the jump probabilities become $p(1 \pm \delta)$, for the downfield and upfield jumps, where $\delta = \exp(\Delta W/kT)$.

The change in the energy diagram slope and hence the variation of p follows the applied field instantaneously. Since f does not change immediately, at first sight there is an instantaneous increase of $2f\delta$ in the downfield particle flux above the upfield particle flux over energy barrier.

In fact, one must consider that a defect cannot be considered to have joined the flux until it has actually crossed the barrier. The

defects which are enabled or prevented to cross the barrier by the change in height must be considered as vibrating within the potential well. Hence, those which are just approaching the barrier as it is lowered will be immediately able to cross, whereas those which have just turned will not cross until one period of the vibration later. Hence the differential flux and hence the electric current will build

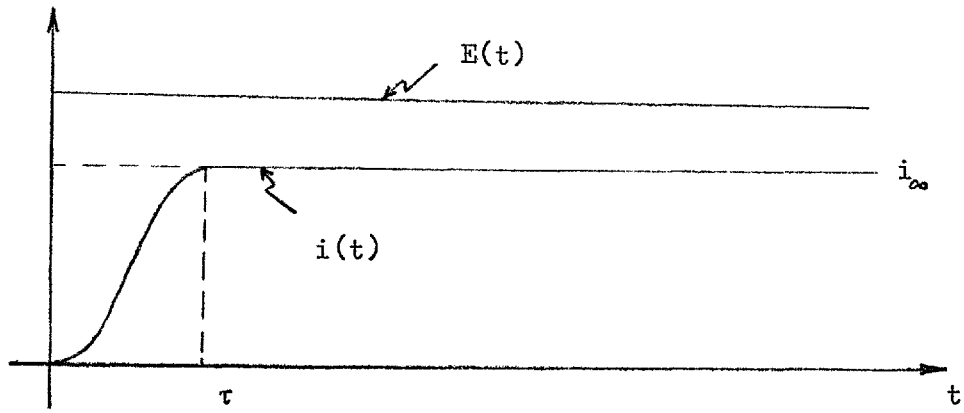


Fig. 2.23 τ = period of defect vibration in well.

up to its final value in one period of the vibrational frequency of the defects in the wells, as shown in Fig. 2.23.

Since the differential flux is the same over all the barriers, it follows that f cannot change in any well, and hence there is no changing polarization. The current also remains constant after the initial build up.

Neglecting the initial rise, it is clear that the current will follow a harmonic field effectively instantaneously, and maintain the same value for a given field, up to frequencies approaching that of the defect vibration. Since this is of the order 10^{12} to 10^{13} Hz, this

model predicts constant conductivity up to such frequencies.

If a harmonic field of frequency higher than this is applied, one must imagine the potential diagram of Fig. 2.22(b) tipped back and forth from the position shown to one with equal and opposite slope at this rate. Clearly few, and eventually none of the defects which would be affected by the barrier height change will have time to cross the lowered barrier, and the conductivity will fall off to zero.

In fact, this effect will operate at high enough frequencies whatever distribution of barrier heights one postulates, and one must always expect conductivity to decrease at high enough frequencies. However, at this stage the treatment would have to be done in terms of quantised lattice vibrations (phonons) interacting with the defects.

2.4.4. A New Varying Depth Multi-Well-Pair Model

In this section a new multi-well-pair model is described and analysed. The well depths are given a variation to account for the interaction between (mobile and immobile) defects mentioned at the opening of this chapter, and the analysis yields an a.c. and a d.c. conductivity which are both related to the numbers of defects, and the jump probabilities over the various barrier heights between the wells.

2.4.4.1. The General Model and Particular Cases

The general model is that mobile charged defects move in a potential distribution of the type shown in Fig. 2.24.

In this figure the short periodicity represents the barriers between equivalent lattice sites for the mobile defect, spaced at distance a, whereas the superimposed slower longer range variation represents the attractive potential of a relatively immobile defect situated between the wells -1 and 1. The immobile defect is here pictured as being on a lattice site belonging to the opposite constituent

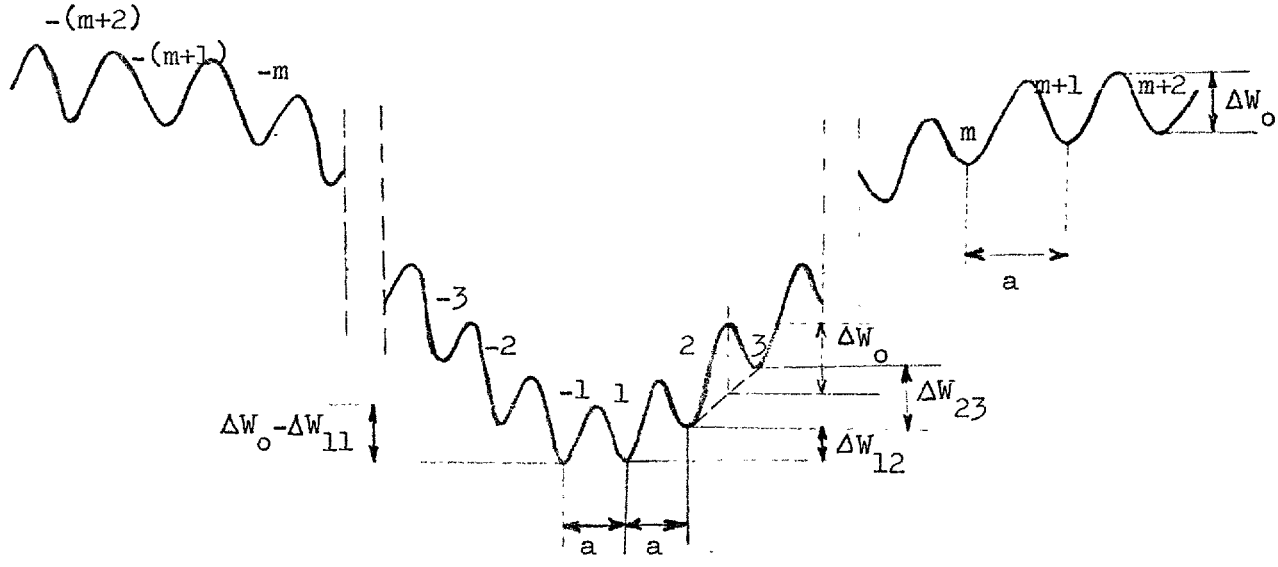


Fig. 2.24

of the crystal from that to which the mobile defect belongs, e.g. at an O site if we consider mobile V_M in MO.

The barrier height far from the immobile defect is taken as ΔW_o, as is also the average of the barrier heights for jumps to right and left over barriers closer in, except for barrier -1 to +1. This barrier is taken to be of height $\Delta W_o - \Delta W_{11}$ so that reasonable assumptions may be made according to what the immobile defect is considered to be. The potential difference between the bottoms of the i^{th} and $i+1^{th}$ wells is labelled $\Delta W_{i,i+1}$. The barrier height for jumps away from the centre is therefore $(\Delta W_o + \frac{1}{2}\Delta W_{i,i+1}) = (\Delta W_o + \frac{1}{2}\Delta W_{-i,-(i+1)})$ and for jumps towards the centre it is $(\Delta W_o - \frac{1}{2}\Delta W_{i,i-1}) = (\Delta W_o - \frac{1}{2}\Delta W_{-i,-(i-1)})$.

The natural expectation is that $\Delta W_{i,i+1}$ will become smaller and smaller as i increases, becoming eventually negligible for some value of $i \geqslant m+1$. It will be seen that the mathematical analysis results in an

m x m matrix which has to be inverted to give $\epsilon^*(\omega)$ or $\sigma^*(\omega)$.

For real attractive potentials $\Delta W_{i,i+1} \rightarrow 0$ only as $m \rightarrow \infty$, but some arbitrary decision has to be made to "cut-off" the variation at some value of m and take $\Delta W_{i,i+1} = 0$ thereafter.

In the following sections the general analysis for any value of m is given, and the results for ϵ^* and σ^* expressed in terms of the solution to the m x m matrix inversion in section 2.4.4.5. The evaluation of $\epsilon^*(\omega)$ or $\sigma^*(\omega)$ for arbitrary m and $\Delta W_{i,i+1}$ variation would then best be done by numerical matrix inversion on a computer. This has not been undertaken yet.

The particular cases m = 1 and m = 2 with arbitrary $\Delta W_{i,i+1}$ variation have been solved algebraically and the results with some numerical evaluations are given in sections 2.4.4.6. and 2.4.4.7.

It is also possible to solve algebraically to arbitrary order in m the case $\Delta W_{11} = \frac{1}{2}\Delta W_{i,i+1} = \text{Const}$, with the spacing between wells -1 and 1 increased to 2a. The results are given in section 2.4.4.8. with a numerical evaluation from m = 6. These parameters allow the solutions for the m^{th} order to be built up as recurrence relations with the m-1 and m-2 order solutions, and are chosen solely for this mathematical convenience. However, this case could be considered to perhaps have some relevance to an immobile defect sitting in a site usually available to a mobile defect, e.g. an immobile M^{3+}_M in MO with mobile V_M.

2.4.4.2. Zero Field Conditions

The transition probabilities without field are as follows:

$$p_{-11} = p_{11} = A \exp\left(-\frac{\Delta W_o - \Delta W_{11}}{kT}\right) = p_{\infty} \exp W_{11}$$

where $p_{00} = A \exp\left(-\frac{\Delta W_o}{kT}\right)$ and $W_{11} = \frac{\Delta W_{11}}{kT}$

Similarly

$$p_{-1-2} = p_{12} = A \exp\left(-\frac{\Delta W_o + \frac{1}{2}\Delta W_{12}}{kT}\right) = p_{00} \exp\left(-\frac{1}{2} W_{12}\right)$$

$$p_{-2-1} = p_{21} = A \exp\left(-\frac{\Delta W_o - \frac{1}{2}\Delta W_{12}}{kT}\right) = p_{00} \exp\left(\frac{1}{2} W_{12}\right) \qquad (2.162)$$

$$p_{-2-3} = p_{23} = p_{00} \exp\left(-\frac{1}{2} W_{23}\right)$$

$$p_{-3-2} = p_{32} = p_{00} \exp\left(\frac{1}{2} W_{23}\right) \quad \text{etc.}$$

On the other hand, the relative probabilities of occupancy of the sites are given by Maxwell-Boltzmann statistics as

$$f_{-1} = f_1 = f_m \exp\left(W_{12} + W_{23} + \cdots\cdots + W_{m-1,m}\right)$$

$$f_{-2} = f_2 = f_m \exp\left(W_{23} + \cdots\cdots + W_{m-1,m}\right) \qquad (2.163)$$

$$f_{-(m-1)} = f_{m-1} = f_m \exp W_{m-1,m}$$

and one notes also that

$$\sum_{i=1}^{m} f_i N_i = (N_d)^{1/3}$$

where N_i is the number of sites of type i and N_d is, as before, the volume density* of mobile (and hence also immobile) defects.

---

*Since the model is one-dimensional the linear density of defects along the "lattice" is $(N_d)^{1/3}$

Hence

$$N_i f_m \exp (W_{12} + \ldots + W_{m-1,m}) + N_2 f_m \exp (W_{23} + \ldots + W_{m-1,m}) +$$

$$\ldots + N_{m-1} f_m \exp W_{m-1,m} + N_m f_m = (N_d)^{1/3} \qquad (2.164)$$

and $$N_1 = N_2 = \ldots = N_{m-1} = 2(N_d)^{1/3} \qquad (2.165)$$

and therefore also

$$N_m = (N_S)^{1/3} - 2\,(m-1)\,(N_d)^{1/3} \qquad (2.166)$$

where N_S is the volume density of lattice sites available to the mobile defect and N_m is the linear density of sites with "normal" potential.

Solving eqn. (2.164) for f_m and using equations (2.165) and (2.166) then

$$f_m = \frac{(N_d/N_S)^{1/3}}{1 - 2(m-1)\left(\frac{N_d}{N_S}\right)^{1/3} + 2\left(\frac{N_d}{N_S}\right)^{1/3} \left\{ \exp (W_{12} + \ldots + W_{m-1,m}) + \ldots + \exp W_{m-1,m} \right\}} \qquad (2.167)$$

Then f_1 and f_2, etc. may be found by substituting eqn. (2.167) in equations (2.163).

The relationships

$$f_i\, p_{ij} = f_j\, p_{ji} \qquad (2.168)$$

may also be noted for future use. They follow directly from (2.162) and

(2.163). Their physical meaning is that the flux from the i^{th} to j^{th} well is equal to that from the j^{th} to i^{th} as must be the case in equilibrium, or in fact in a steady state.

2.4.4.3. Steady State Condition with Constant Applied Field

The total probability per unit time that a defect is in any one of $(N_S/N_d)^{1/3} = n_S$ wells and jumps to the right is

$$\sum_{i=1}^{n_s} f_i \, p_{ij} \qquad \text{where } j = i + 1$$

Similarly the total probability per unit time of a jump to the left is

$$\sum_{i=1}^{n_s} f_i \, p_{ik} \qquad \text{where } k = i - 1$$

Hence the total differential probability of a jump to the right is

$$\sum_{i=1}^{n_s} f_i \, p_{ij} - \sum_{i=1}^{n_s} f_i \, p_{ik} = p_t$$

when the field is zero, then $p_t = 0$. The applied field will increase p_{ij} to $p'_{ij} = p_{ij}\,(1 + \delta)$ and decrease p_{ik} to $p'_{ik} = p_{ik}\,(1 - \delta)$ where $\delta = \frac{zea\,E}{2kT}$. The differential probability

$$p_t = \cdots + f'_{-2}\,p'_{-2,-1} + f'_{-1}\,p'_{-1,1} + f'_{1}\,p'_{1,2} + f'_{2}\,p'_{23} + \cdots\cdots$$

$$\cdots - f'_{-1}\,p'_{-1,-2} - f'_{1}\,p'_{1,-1} - f'_{2}\,p'_{21} - f'_{3}\,p'_{31} - \cdots\cdots$$

or

$$p_t = \ldots + \left[f'_{-2}\, p_{-2,-1}\,(1+\delta) - f'_{-1}\, p_{-1,-2}\,(1-\delta)\right] + \left[f'_{-1}\, p_{-1,1}\,(1+\delta) - f'_{1}\, p_{1,-1}\,(1-\delta)\right] + \left[f'_{1}\, p_{12}\,(1+\delta) - f'_{2}\, p_{21}\,(1-\delta)\right] + \ldots + \left[f'_{q}\, p_{q,q+1}\,(1+\delta) - f'_{q+1}\, p_{q+1,q}\,(1-\delta)\right] + \ldots\ldots$$

Each term in square brackets is the differential flux rate of defects over a barrier, and in the steady state these are all equal, otherwise the well populations would still be changing. Also far from the defect, i.e. $i = m + 1$ onwards, then $f'_{m+1} = f'_{m+2} = \ldots = f_{m+1}(0)$, the probability before applying E and $p_{m+1,m+2} = p_{m+2,m+1} = p_{00}$ so that

$$f'_{m+1}\, p_{m+1,m+2}\,(1+\delta) - f'_{m+2}\, p_{m+2,m+1}\,(1-\delta) = 2\delta f(0)_{m+1}\, p_{00} \qquad (2.169)$$

and since all terms equal this at the steady state and there are $(N_S/N_d)^{1/3}$ terms then

$$p_t = \left(\frac{N_S}{N_d}\right)^{1/3} 2\delta\, f(0)_{m+1}\, p_{00} \qquad (2.170)$$

Also mobility

$$\mu = \frac{a\, p_t}{E} = \frac{zea^2}{kT}\left(\frac{N_S}{N_d}\right)^{1/3} f(0)_{m+1}\, p_{00} \qquad (2.171)$$

then d.c. conductivity*

$$\sigma_{d.c.} = N_d ze\, \mu = \frac{N_d z^2 e^2 a^2}{kT}\left(\frac{N_S}{N_d}\right)^{1/3} f(0)_{m+1}\, p_{00} \qquad (2.172)$$

---

*The bulk volume density is used here to give a bulk conductivity.

For m = 0 which is the elementary multi-well model with no differences in well depths then

$$f(0)_{m+1} = \left(\frac{N_d}{N_S}\right)^{1/3}$$

Hence

$$\mu = \frac{ze a^2}{kT} p_{00} \tag{2.173}$$

and

$$\sigma_{d.c.} = N_d \frac{z^2 e^2 a^2}{kT} p_{00} \tag{2.174}$$

and the elementary result of section 2.2.4. is recovered.

2.4.4.4. Equations for Transient Conditions

The time rate of change of the number of carriers in sites type i is

$$\frac{dn_i}{dt} = p_{ki}\, n_k - (p_{ik} + p_{ij})\, n_i + p_{ji} n_j \tag{2.175}$$

where k = i - 1 and j = i + 1. Since

$$n_i = N_i f_i \text{ and } N_i = 2N_d \tag{2.176}$$

then substituting these into the above equation gives,

$$\frac{df_i}{dt} = p_{ki} f_k - (p_{ik} + p_{ij})\, f_i + p_{ji} f_j \tag{2.177}$$

Without field, from Eqn. (2.168)

$$p_{ki}f_k = p_{ik}f_i \text{ and } p_{ij}f_i = p_{ji}f_j$$

and therefore $\frac{df_i}{dt} = 0$ as expected.

The probabilities of occupancy with field f_i will not be written with primes since $f_i(0)$ will be written for f_i when $E = 0$ for $t < 0$. Now in particular

$$\frac{df_1}{dt} = p'_{-1,1}f_{-1} - p'_{1,-1}f_1 - p'_{12}f_1 + p'_{21}f_2$$

$$= p_{-1,1}(1+\delta)f_{-1} - p_{1,-1}(1-\delta)f_1 - p_{12}(1+\delta)f_1 + p_{21}(1-\delta)f_2$$

$$= p_{-1,1}f_{-1} - p_{1,-1}f_1 - p_{12}f_2 + p_{21}f_2 + \delta\left[p_{-1,1}f_{-1} + p_{1,-1}f_1 - \right.$$

$$\left. -p_{12}f_1 - p_{21}f_2\right] \qquad (2.178)$$

and

$$\frac{df_{-1}}{dt} = p_{1,-1}(1-\delta)f_1 - p_{-1,1}(1+\delta)f_{-1} - p_{-1,-2}(1-\delta)f_{-1} + p_{-2,-1}(1+\delta)f_{-2}$$

$$= p_{1,-1}f_1 - p_{-1,1}f_{-1} - p_{-1,-2}f_{-1} + p_{-2,-1}f_{-2} + \delta\left[- p_{1,-1}f_1 - \right.$$

$$\left. - p_{-1,1}f_{-1} + p_{-1,-2}f_{-1} + p_{-2,-1}f_{-2}\right] \qquad (2.179)$$

Due to the symmetry of the potential pattern, f_i will be greater than $f_i(0)$ by the same amount as f_{-i} is smaller, so that in particular

$$f_1 + f_{-1} = 2f_1(0) \tag{2.180}$$

Defining also

$$g_i = f_i - f_{-i} \tag{2.181}$$

and using $p_{ij} = p_{-i-j}$, then Eqns. (2.178) and (2.179) lead to

$$\frac{dg_1}{dt} = p_{-1,1}g_{-1} - (p_{1,-1} + p_{12})g_1 + p_{21}g_2 + 4\delta(p_{1,-1} - p_{12})f_1(0) \tag{2.182}$$

Since $g_{-1} = -g_1$, this can also be written

$$\frac{dg_1}{dt} = -(2p_{1,-1} + p_{12})g_1 + p_{21}g_2 + 4\delta(p_{1,-1} - p_{12})f_1(0) \tag{2.183}$$

but this simplification is not available for $\frac{dg_i}{dt}$ with $i > 1$. Proceeding similarly

$$\frac{dg_2}{dt} = p_{12}g_1 - (p_{21} + p_{23})g_2 + p_{32}g_3 + 4\delta(p_{21} - p_{23})f_2(0) \tag{2.184}$$

and in general with $k = i - 1$ and $j = i + 1$ then

$$\frac{dg_i}{dt} = p_{ki}g_k - (p_{ik} + p_{ij})g_i + p_{ji}g_j + 4\delta(p_{ik} - p_{ij})f_i(0) \tag{2.185}$$

For an m-well-pair model we assume $g_m \equiv 0$ and hence $\frac{dg_m}{dt} \equiv 0$, so that

$$\frac{dg_{m-1}}{dt} = p_{m-2,m-1}g_{m-2} - (p_{m-1,m-2} + p_{m-1,m})\, g_{m-1} +$$

$$4\delta(p_{m-1,m-2} - p_{m-1,m})f_{m-1}(0) \qquad (2.186)$$

Hence the $g_i(t)$ are the solutions to a set of simultaneous first order linear differential equations. Each $g_i(t)$ results in a time dependent dipole moment given by

$$\mu_i(t) = (2i - 1)\frac{aze}{2}\, g_i(t) \qquad (2.187)$$

and hence the dipole moment associated with one immobile defect will be

$$\mu(t) = \frac{aze}{2} \sum_{i=1}^{m} (2i - 1)\, g_i(t)$$

and for N_d defects per unit volume the polarisation will be

$$P(t) = N_d \frac{zae}{2} \sum_{i=1}^{m} (2i - 1)g_i(t) \qquad (2.188)$$

If $g_i(t)$ and hence $P(t)$ were the response to a unit step function applied field, then $P(t)$ would be in fact $A_p(t)$ and the theory of section 2.3.3. could be applied to find

$$A_p(s) = \mathcal{L}\left\{A_p(t)\right\} \qquad (2.189)$$

and hence

$$\epsilon_0 \chi^*(s) = s\, Ap(s) \qquad (2.190)$$

and eventually

$$\underset{o}{\in} \chi^*(j\omega) = \lim_{p \to 0} \left\{ s\, Ap(s) \right\} \tag{2.191}$$

However, assuming $E = E_1(t)$ in Eqn. (2.188) and Laplace transforming shows that

$$A_p(s) = N_d \frac{zae}{2} \sum_{i=1}^{m} (2i - 1) g_i(s) \tag{2.192}$$

In other words, we can find $\underset{o}{\in} \chi^*(j\omega)$ if we can contrive to find the $g_i(s)$ without necessarily finding $g_i(t)$. In fact, this is done in the next section.

2.4.4.5. General Solution in the Frequency Domain by Laplace Transform

The simultaneous equations of the type of Eqn. (2.185), with particular cases Eqns. (2.183) and (2.186) may all be Laplace Transformed bearing in mind that $\mathcal{L}\left[\frac{dg_i}{dt}\right] = s\, g_i(s) - g_i(0)$ and that all $g_i(0) = 0$. The equations may then be re-arranged by collecting the coefficients of g_i on to the right and the constants on to the left to yield, in matrix form

$$[f] = [p]\,[g] \tag{2.193}$$

where

$$f = \frac{4\delta}{s} \begin{bmatrix} (p_{1,-1} - p_{12})\, f_1(0) \\ (p_{21} - p_{23})\, f_2(0) \\ \vdots \\ (p_{ik} - p_{ij})\, f_i(0) \\ \vdots \\ (p_{m-1,m-2} - p_{m-1,m})\, f_{m-1}(0) \end{bmatrix} \tag{2.194}$$

and $[g]$ is a column vector with elements $[g_1 \ldots g_{m-1}]$ and

$$[p] = \begin{bmatrix} (2p_{11}+p_{12}+s) & -p_{21} & 0 & \cdots & 0 & 0 \\ -p_{12} & (p_{21}+p_{23}+s) & -p_{32} & 0 & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & & & p_{m-2,m-1} & (p_{m-1,m-2}+p_{m-1,m}+s) \end{bmatrix} \tag{2.195}$$

The $g_i(s)$ may thus be found by matrix inversion of Eqn. (2.193). Once the $g_i(s)$ are found then $A_p(s)$ is given by Eqn. (2.192), and consequently $\epsilon_o \chi^*(\omega)$ and $\sigma^*(\omega)$ may be found as previously noted.

All the $g_i(s)$ contain the factor $4\delta/s = 2z\,eaE/kTs$ and may also be manipulated by the use of the $f_i p_{ij} = f_j p_{ji}$ relationships to contain the factor $p_{00} f_1(0)$, i.e. each $g_i(s)$ may be written in the form $[2zea\,E\,p_{00}\,f_1(0)/kTs]\;G_i(s)$ where the actual form of $G_i(s)$ depends on the particular f_j contained in g_i, and we shall not write out particular cases.

Hence one finds

$$\epsilon_o \chi^*(s) = s\,A_p(s) = \frac{N_d\, z^2 e^2 a^2\, P_{00}\, f_1(0)}{kT}\; G(s) \tag{2.196}$$

where

$$G(s) = \sum_{i=1}^{m} (2i - 1)\, g_i(s) \tag{2.197}$$

and therefore

$$\epsilon_o \chi^*(\omega) = \frac{N_d\, z^2 e^2 a^2\, P_{00}\, f_1(0)}{kT}\; G(j\omega) \tag{2.198}$$

and

$$(\sigma^* - j\omega\epsilon_o) = j\omega\epsilon_o\chi^*(\omega) = \frac{N_d z^2 e^2 a^2 p_{00}}{kT} f_1(0)\left\{j\omega G(j\omega)\right\} \qquad (2.199)$$

The absolute susceptibility and the a.c. conductivity, σ', are each given by the real parts of Eqns. (2.198) and (2.199) respectively.

Using Eqn. (2.172) of section 2.4.4.3. the ratio $\sigma'/\sigma_{d.c.}$ is seen to be

$$\frac{\sigma'}{\sigma_{d.c.}} = \left(\frac{N_d}{N_s}\right)^{1/3} \frac{f_1(0)}{f_{m+1}(0)p_{00}} \mathcal{R}\left\{j\omega\, G(j\omega)\right\}$$

which may be written

$$\frac{\sigma'}{\sigma_{d.c.}} = K'\mathcal{R}\left\{j\Omega\, G(j\Omega)\right\} \qquad (2.200)$$

where

$$K' = \left(\frac{N_d}{N_s}\right)^{1/3} \exp\left(\Delta W_{12} + \Delta W_{23} + \ldots + \Delta W_{m-1,m}\right) \qquad (2.201)$$

and $\Omega = \omega/p_{00}$ as previously defined.

In addition the ratio of total conductivity to d.c. conductivity is given by

$$\frac{\sigma}{\sigma_{d.c.}} = \frac{\sigma_{d.c.} + \sigma'}{\sigma_{d.c.}} = 1 + \frac{\sigma'}{\sigma_{d.c.}} \qquad (2.202)$$

2.4.4.6. The One Depressed Well-Pair Case

In terms of the general treatment just given we have m = 1 and $g_2 \equiv 0$. Hence the d.c. conductivity is given by Eqn. (2.172) as

$$\sigma_{d.c.} = \frac{N_d z^2 e^2 a^2}{kT} \left(\frac{N_s}{N_d}\right)^{1/3} p_{00}\, f_2(0) \tag{2.203}$$

There is only one equation for dg_i/dt, which is Eqn. 2.182, i.e.

$$\frac{dg_1}{dt} = -(2p_{11} + p_{12})g_1 + 4\delta\,(p_{11} - p_{12})f_1(0)$$

and the solution of its Laplace transform is

$$g_1(s) = \frac{4\delta}{s} \frac{(p_{11} - p_{12})}{(2p_{11} + p_{12} + s)}\, f_1(0)$$

Hence

$$\epsilon_o \chi^*(\omega) = \frac{N_d z^2 e^2 a^2 f_1(0) p_{00}}{kT} \cdot G(j\omega) \tag{2.204}$$

where

$$G(j\omega) = \frac{(p_{11} - p_{12})}{p_{00}(2p_{11} + p_{12} + j\omega)}$$

Using $p_{11} = p_{00} \exp W_{11}$ and $p_{12} = p_{00} \exp(-W_{12}/2)$ this last equation may be re-written

$$G(j\omega) = \frac{\exp W_{11} - \exp(-W_{12}/2)}{p_{00}(2 \exp W_{11} + \exp(-W_{12}/2) + j\Omega)} \tag{2.205}$$

For comparison to the isolated double well case we may write the combined result for $\epsilon_o \chi^*(\omega)$ as

$$\epsilon_o \chi^*(\omega) = K_1' \frac{1}{1 + j\Omega/[2 \exp W_{11} + \exp(-W_{12}/2)]} \tag{2.206}$$

where

$$K_1' = \frac{N_d z^2 e^2 a^2 f_1(0)}{kT} \cdot \frac{\exp W_{11} - \exp(-W_{12}/2)}{2\exp W_{11} + \exp(-W_{12}/2)} \qquad (2.207)$$

and Eqns. (2.206) and (2.207) may be compared to Eqn. (2.156), whence it is seen that K_1' has replaced K_1 and $\Omega_{1/2}$, the value of Ω at which $\epsilon_o \chi'$ reaches half its final value (see Eqn. 2.159), is shifted from 2 to $2 \exp W_{11} + \exp(-W_{12}/2)$. Thus the relaxation frequency has been shifted from $2p_o$ by two factors; a) the different barrier height attributed to the central barrier and b) the influence of the jumps into the neighbouring wells shown by the appearance of W_{12} in the expression for $\Omega_{1/2}$.

From Eqns. (2.203) and (2.205) one may write the ratio

$$\frac{\sigma'}{\sigma_{d.c.}} = \left(\frac{N_d}{N_s}\right)^{1/3} \exp W_{12}\, \Re\left\{\frac{\left[\exp W_{11} - \exp(-W_{12}/2)\right] j\Omega}{2\exp W_{11} + \exp(-W_{12}/2) + j\Omega}\right\}$$

$$= K \frac{\left[\Omega/2 \exp W_{11} + \exp(-W_{12}/2)\right]^2}{1 + \left[\Omega/2 \exp W_{11} + \exp(-W_{12}/2)\right]^2} \qquad (2.208)$$

where

$$K = \left(\frac{N_d}{N_s}\right)^{1/3} \exp W_{12}\left[\exp W_{11} - \exp(-W_{12}/2)\right] \qquad (2.209)$$

which is of the form for the isolated well pair with an added d.c. conductivity as discussed in section 2.4.1. Curves of σ'/σ_∞' and σ/σ_∞' were already plotted in Fig. 2.21 for the case K = 50. In this previous case however, K was assumed to provide an arbitrary d.c. conductivity not related to the model for a.c. conductivity, whereas with the present model the value of K also depends on the parameters

of the model giving the a.c. conductivity. The value of $\Omega_{1/2}$ for σ' is shifted in just the same way as discussed above for $\epsilon_o \chi'$.

For $W_{11} = 0$, i.e. the central barrier equal in height to the isolated well pair barrier, the relationships between the models are seen very clearly, the modifications consisting entirely of extra terms in exp $(- W_{12}/2)$.

As an illustrative example, a calculated curve of Eqn. (2.209) with $W_{11} = 0$, $K = 20$ and $N_d/N_s = 10^{-3}$ is shown in Fig. 2.25. These parameters determine exp W_{12} as 213.3 which corresponds to $\Delta W_{12} =$ 0.82 eV at $T = 1500^{o}C$.

2.4.4.7. The Two Depressed Well-Pair Case

For this case we have $m = 2$ and $g_3 \equiv 0$ and the d.c. conductivity as given by Eqn. (2.172) as

$$\sigma_{d.c.} = \frac{N_d z^2 e^2 a^2}{kT} \left(\frac{N_s}{N_d}\right)^{1/3} p_{00} f_3(0) \tag{2.210}$$

The two by two matrix equation for $g_1(s)$ and $g_2(s)$ may be solved by Kramer's rule to give

$$g_1(s) = \frac{4\delta}{s}\left[(p_{11}-p_{12})(p_{21}+p_{23}+s)f_1(0) + p_{21}(p_{12}-p_{23})f_2(0)\right]/\Delta \tag{2.211}$$

$$g_2(s) = \frac{4\delta}{s}\left[(p_{12}-p_{23})(2p_{11}+p_{12}+s)f_2(0) + p_{12}(p_{11}-p_{12})f_1(0)\right]/\Delta \tag{2.212}$$

where Δ is the determinant of the matrix given by

$$\Delta = (p_{21}+p_{23}+s)(2p_{11}+p_{12}+s) - p_{21}p_{12} \tag{2.213}$$

Using $f_i p_{ij} = f_j p_{ji}$ and removing the factor $4\delta f_1(0)/s$ from both $g_1(s)$

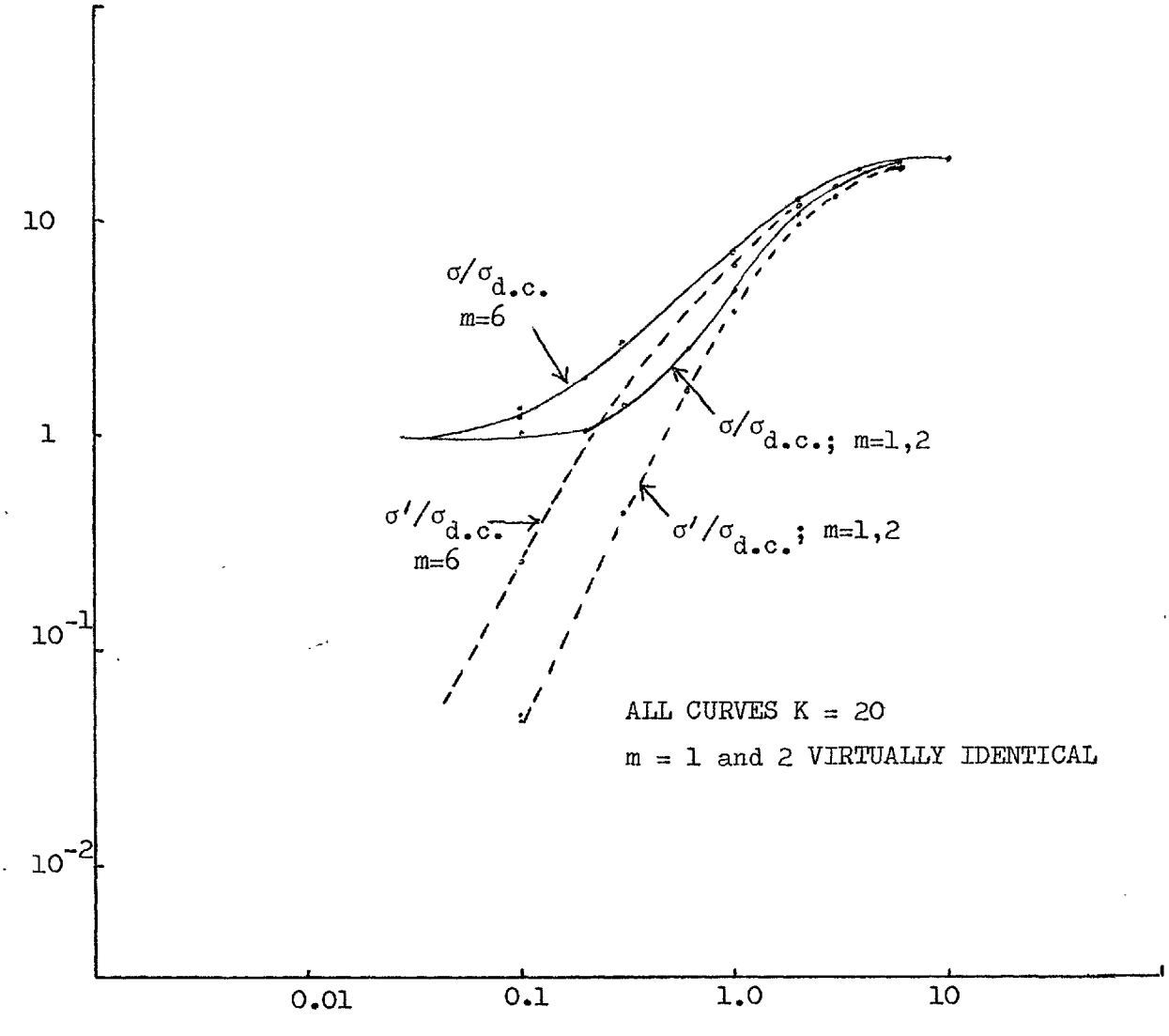

10
1
10^{-1}
10^{-2}
0.01
0.1
1.0
10
$\sigma/\sigma_{d.c.}$
m=6
$\sigma/\sigma_{d.c.}$; m=1,2
$\sigma'/\sigma_{d.c.}$
m=6
$\sigma'/\sigma_{d.c.}$; m=1,2
ALL CURVES K = 20
m = 1 and 2 VIRTUALLY IDENTICAL

Figure 2.25

and $g_2(s)$ leaves as remainders

$$G_1(s) = \left[(p_{11} - p_{12})(p_{21} + p_{23} + s) + p_{12}(p_{12} - p_{23}) \right] / \Delta \quad (2.214)$$

and

$$G_2(s) = \left[\frac{p_{12}}{p_{21}} (p_{12} - p_{23}) (2p_{11} + p_{12} + s) + p_{12}(p_{11} - p_{12}) \right] / \Delta \quad (2.215)$$

Using Eqn. (2.197) yields

$$G(s) = G_1(s) + 3G_2(s) \quad (2.216)$$

and hence

$$\epsilon_o \chi^*(\omega) = \frac{N_d z^2 e^2 a^2}{kT} f_1(0) G(j\omega) \quad (2.217)$$

By substituting for p_{ij} in terms of W_{ij} and with some algebraic manipulation one finds that

$$\Delta = s^2 + Xs + Y \quad (2.218)$$

where $X = p_{00}\left[2\exp W_{11} + \exp(- W_{12}/2) + \exp(W_{12}/2) + \exp(- W_{23}/2)\right]$ (2.219)

and

$$Y = p_{00}^2 \left[2\exp W_{11} \exp \frac{W_{12}}{2} + 2\exp W_{11} \exp(- W_{23}/2) + \right.$$

$$\left. + \exp(- W_{12}/2) \exp(- W_{23}/2) \right] \quad (2.220)$$

and that

$$G_1(s) = As + B \tag{2.221}$$

where

$$A = p_{00}\left[\exp W_{11} - \exp(-W_{12}/2)\right] \tag{2.222}$$

and

$$B = p_{00}^2\left\{\left[\exp W_{11} - \exp(-W_{12}/2)\right]\left[\exp(W_{12}/2) + \exp(-W_{23}/2)\right] + \exp(-W_{12}/2)\left[\exp(-W_{12}/2) - \exp(-W_{23}/2)\right]\right\} \tag{2.223}$$

and that

$$G_2(s) = Cs + D \tag{2.224}$$

where

$$C = p_{00}\exp(-W_{12})\left[\exp(-W_{12}/2) - \exp(-W_{23}/2)\right] \tag{2.225}$$

and

$$D = p_{00}^2\left\{\exp(-W_{12})\left[\exp(-W_{12}/2) - \exp(-W_{23}/2)\right]\left[2\exp W_{11} + \exp(-W_{12}/2)\right] + \exp(-W_{12}/2)\left[\exp(W_{11}/2) - \exp(-W_{12}/2)\right]\right\} \tag{2.226}$$

Hence we may write G(s) in the form

$$G(s) = \frac{(A + 3C)s + B + 3D}{s^2 + Xs + Y} \tag{2.227}$$

We shall only consider the case where $G_2(s) \ll G_1(s)$, when

$$G(j\omega) = \frac{BY + (XA - B)\omega^2 + j\left[(AY - BX)\omega - A\omega^3\right]}{Y^2 + (X^2 - 2Y)\omega^2 + \omega^4} \tag{2.228}$$

Hence

$$\epsilon_o \chi'(\omega) = \frac{N_d z^2 e^2 a^2 f_1(0)}{kT} \cdot \frac{BY + (XA - B)\omega^2}{Y^2 + (X^2 - 2Y)\omega^2 + \omega^4} \tag{2.229}$$

and

$$\sigma'(\omega) = \frac{N_d z^2 e^2 a^2 f_1(0)}{kT} \cdot \frac{(BX - AY)\omega^2 + A\omega^4}{Y^2 + (X^2 - 2Y)\omega^2 + \omega^4} \tag{2.230}$$

Evidently, the parameters W_{11}, W_{12} and W_{13} enter into Eqns. (2.229) and (2.230) in such a complex way that a general discussion would be very long and complex, and hardly worthwhile in view of the extreme simplifications which have been made to arrive at even these expressions. However, once values of W_{11} etc. have been chosen and the numerical values of A, B, X and Y determined, it is possible to calculate out $\epsilon_o \chi'(\omega)$ or $\sigma'(\omega)$ on a desk calculator without too excessive a labour. In fact, calculations of the ratio $\sigma'(\omega)/\sigma_{d.c.}$ were made. Using Eqns. (2.229) and (2.210) this is given by

$$\frac{\sigma'}{\sigma_{d.c.}} = \left(\frac{N_d}{N_s}\right)^{1/3} \exp(W_{12} + W_{13}) \cdot \frac{A}{p_{00}} \mathcal{R}\left\{\frac{\frac{1}{A}(BX - AY) + \omega^4}{Y^2 + (X^2 - 2Y)\omega^2 + \omega^4}\right\}$$

$$= K\mathcal{R}\left\{\frac{\frac{1}{A'}(B'X' - A'Y') + \Omega^4}{Y'^2 + (X'^2 - 2Y')\Omega^2 + \Omega^4}\right\} \tag{2.231}$$

where A', B' etc. are the same quantities as A, B, etc., but with the multiplying factors in p_{00}, p_{00}^2 etc. omitted, $\Omega = \omega/p_{00}$ as before and

$$K = \left(\frac{N_d}{N_s}\right)^{1/3} A' \exp(W_{12} + W_{13})$$

$$= \left(\frac{N_d}{N_s}\right)^{1/3} \exp(W_{12} + W_{13})\left[\exp W_{11} - \exp(-W_{12}/2)\right] \tag{2.232}$$

which may be compared to Eqns. (2.208) and (2.209) for the single depressed well-pair case. Since the numerical value of the ratio of polynominals in Ω in Eqn. (2.230) rises from zero to 1 as Ω rises from 0 to ∞, then K has the same significance as previously, and gives the ratio $\sigma'_{\infty}/\sigma_{d.c.}$.

As a particular case the same values as for the single well-pair were taken, i.e. K = 20 and $N_d/N_s = 10^{-3}$; so that exp $(W_{12} + W_{23})$ must have the value 224.8, i.e. almost the same numerical value as exp W_{12} in the previous case. At T = 1500°C this corresponds to $W_{12} + W_{23} = 0.828$ eV. The total depression $\Delta W_{12} + \Delta W_{23}$ was divided in the ratio 5:1, so that $\Delta W_{12} = 0.662$ eV and $\Delta W_{23} = 0.166$ eV. This ratio corresponds to the potential differences for a $1/r$ potential originating at a point mid-way between sites -1 and +1. The resulting curves of log $\sigma'/\sigma_{d.c.}$ and log $(1 + \sigma'/\sigma_{d.c.})$ against log f are plotted on the same axes as for the one depressed well-pair case in Fig. 2.25.

2.4.4.8. Any Number of Equal Potential Steps

An examination of the general theory in sections 2.4.4.4. and 2.4.4.5. will show that a considerable algebraic simplification of the equations occurs for the particular case

$$\Delta W_{12} = \Delta W_{23} = \Delta W_{34} = \ldots = \Delta W_{m,m+1};$$

$$\Delta W_{ij} = 0,\ i,\ j \geqslant m + 1;$$

$$\exp W_{11} = \frac{1}{2} \exp W_{12}$$

provided also that the spacing between sites -1 and +1 is increased to 2a.

Eqn. (2.172), repeated here, is immediately valid for $\sigma_{d.c.}$;

$$\sigma_{d.c.} = \frac{N_d z^2 e^2 a^2}{kT} \left(\frac{N_s}{N_d}\right)^{1/3} f_{m+1}(0)\, p_{00} \tag{2.233}$$

The values of $g_i(s)$ are to be found by matrix inversion of Eqn. (2.193), which was

$$[f] = [p][g]$$

For the present model however, writing

$$\alpha = \exp\,(W_{i,i-1}/2); \quad i \geqslant 1 \tag{2.234}$$

then $$p_{i,i-1} = p_{00}\,\alpha \tag{2.235}$$

and $$p_{i,i+1} = p_{00}/\alpha \tag{2.236}$$

and the expressions for $[f]$, $[p]$ and $[g]$ become

$$f = \frac{4\delta}{s} p_{00} \begin{bmatrix} (\alpha - \frac{1}{\alpha})\, f_1(0) \\ (\alpha - \frac{1}{\alpha})\, f_2(0) \\ \vdots \\ (\alpha - \frac{1}{\alpha})\, f_m(0) \end{bmatrix} \tag{2.237}$$

$$p = p_{00}^{m} \begin{bmatrix} (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}}) & -\alpha & 0 & \cdots & 0 \\ -\frac{1}{\alpha} & (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}}) & -\alpha & & \\ \vdots & & & & 0 \\ & & & & -\alpha \\ 0 & \cdots & -\frac{1}{\alpha} & & (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}}) \end{bmatrix} \qquad (2.238)$$

and $[g]$ is again the column vector of the g_i with i = 1,2 ...,m-1. The $g_i(s)$ for the m well-pair case are thus given by

$$g_i^m(s) = \frac{M_m^i}{\Delta_m} \qquad (2.239)$$

where Δ_m is the determinant of the matrix $[p]$ for the m pair case, and M_m^i is the determinant of this matrix modified by replacing the i th column by $[f]$; e.g. for i = 1

$$sM_m^1 = \frac{4\delta}{s} p_{00}^{m} \begin{vmatrix} (\alpha - \frac{1}{\alpha})f_1(0) & -\alpha & 0 & 0 & 0 \\ (\alpha - \frac{1}{\alpha})f_2(0) & (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}}) & -\alpha & 0 & \vdots \\ \vdots & & & & \\ & & & & -\alpha \\ (\alpha - \frac{1}{\alpha})f_m(0) & \cdots & -\alpha & & (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}}) \end{vmatrix} \qquad (2.240)$$

The following recurrence relationships may be proven by elementary expansion of the determinants concerned;

$$\Delta_1 = (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}})$$

$$\Delta_2 = (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}})^2 - 1$$

$$\Delta_3 = (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}})\,\Delta_2 - \Delta_1 \qquad (2.241)$$

$$\vdots$$

$$\Delta_m = (\alpha + \frac{1}{\alpha} + \frac{s}{p_{00}})\Delta_{m-1} - \Delta_{m-2}$$

etc.

and

$$s\,M_1^1 = (\alpha - \frac{1}{\alpha})\,f_1(0)$$

$$s\,M_2^1 = (\alpha - \frac{1}{\alpha})\,f_1(0)\left\{\Delta_1 + \frac{1}{\alpha}\right\}$$

$$s\,M_3^1 = (\alpha - \frac{1}{\alpha})\,f_1(0)\left\{\Delta_2 + \frac{\Delta_1}{\alpha} + \frac{1}{\alpha^2}\right\} \qquad (2.242)$$

$$\vdots$$

$$s\,M_m^1 = (\alpha - \frac{1}{\alpha})\,f_1(0)\left\{\Delta_m + \frac{\Delta_{m-1}}{\alpha} + \cdots \frac{\Delta_1}{\alpha^{m-2}} + \frac{1}{\alpha^{m-1}}\right\}$$

etc.

Hence Eqns. (2.241) and (2.242) allow the evaluation of $s\,g_1^m(s)$. The analysis for $g_i^m(s)$ with $i > 1$ has not been carried out. Assuming that $g_1^m(s)$ is the most important term, it follows that

$$\epsilon_o \chi^*(s) = \frac{N_d z e a}{2} \quad \frac{4\delta}{s} \quad \frac{s M_m^1(s)}{\Delta_m(s)}$$

$$= \frac{N_d z^2 e^2 a^2}{kT} \quad \frac{M_m^1(s)}{\Delta_m(s)}$$

$$= \frac{N_d z^2 e^2 a^2}{kT} \quad (\alpha - \tfrac{1}{\alpha})\, f_1(0) \quad \frac{L_m'(s)}{\Delta_m(s)} \tag{2.243}$$

where $L_m(s) = M_m^1(s)/(\alpha - \frac{1}{\alpha})\, f_1(0)$ (2.244)

To find an expression for $\sigma'/\sigma_{d.c.}$ we may use

$$\sigma' = \mathcal{R}\{j\omega \epsilon_o \chi^*(j\omega)\}$$

$$= \frac{N_d z^2 e^2 a^2}{kT} \quad (\alpha - \tfrac{1}{\alpha})\, f_1(0) \quad \mathcal{R}\left\{\frac{j\omega L_m'(j\omega)}{\Delta_m(j\omega)}\right\} \tag{2.245}$$

and hence using Eqn. (2.233) then

$$\frac{\sigma'}{\sigma_{d.c.}} = \left(\frac{N_d}{N_s}\right)^{1/3} \quad \frac{(\alpha - \frac{1}{\alpha})}{p_{00}} \cdot \frac{f_1(0)}{f_{m+1}(0)} \quad \mathcal{R}\left\{\frac{j\omega L_m'(j\omega)}{\Delta_m(j\omega)}\right\}$$

$$= \left(\frac{N_d}{N_s}\right)^{1/3} (\alpha - \tfrac{1}{\alpha}) \cdot \alpha^{2m} \quad \mathcal{R}\left\{\frac{j\Omega L_m'(j\Omega)}{\Delta_m(j\Omega)}\right\} \tag{2.246}$$

where the factor

$$K = \left(\frac{N_d}{N_s}\right)^{1/3} \quad (\alpha - \tfrac{1}{\alpha})\, \alpha^{2m} \tag{2.247}$$

has the significance of giving the ratio $\sigma'/\sigma_{d.c.}$ as before.

A calculation of $\sigma'/\sigma_{d.c.}$ for m = 6 with K = 20 and $N_d/N_s = 10^{-3}$ has been made. This requires $\alpha = 1.644$ corresponding to $\Delta W_{12} = \Delta W_{23} = \ldots = \Delta W_{m-1,m} = 0.152$ eV at a temperature T = 1500°C. The results are plotted in Fig. 2.25 with these for the single and two well-pair cases.

CHAPTER 3 - APPARATUS AND EXPERIMENTAL PROCEDURE

The theoretical discussion of chapter 2 showed that the conductivity of Al_2O_3 depends on the ambient atmosphere and temperature. Also as mentioned in chapter 1, the resistivity of Al_2O_3 is very high. These are the important factors to consider in the design of the apparatus and method of measurement.

In order to be able to measure conductivity σ as a function of oxygen partial pressure p_{O_2} and temperature T at d.c. and varying frequency, we need a sample holder designed so that the p_{O_2} and T can be varied over a range but held constant at particular values for fairly long periods (up to some days) to allow equilibriation of the sample and ambient atmosphere. It must also be designed to avoid erroneous electrical measurements due to parallel leakage paths and other effects. To show how these requirements have been met, the principles involved in measuring $\sigma_{d.c.}$ and $\sigma_{a.c.}$ for high resistivity materials and their effect on the sample holder design will be discussed.

The construction of the sample holder will then be described, and finally the ambient atmosphere and temperature control arrangements described.

3.1. Principles of Conductivity Measurements

Electrical conductivity was measured using a three terminal method with a circuit similar to that described by Moulson and Popper (1968) for both the d.c. and a.c. cases.

3.1.1. d.c. measurements

The circuit diagram for d.c. method is shown in Fig. 3.1. A voltage of 1 volt (less than decomposition voltage) is applied across the crystal from a stabilized d.c. power supply (Fluke, Model 415B)

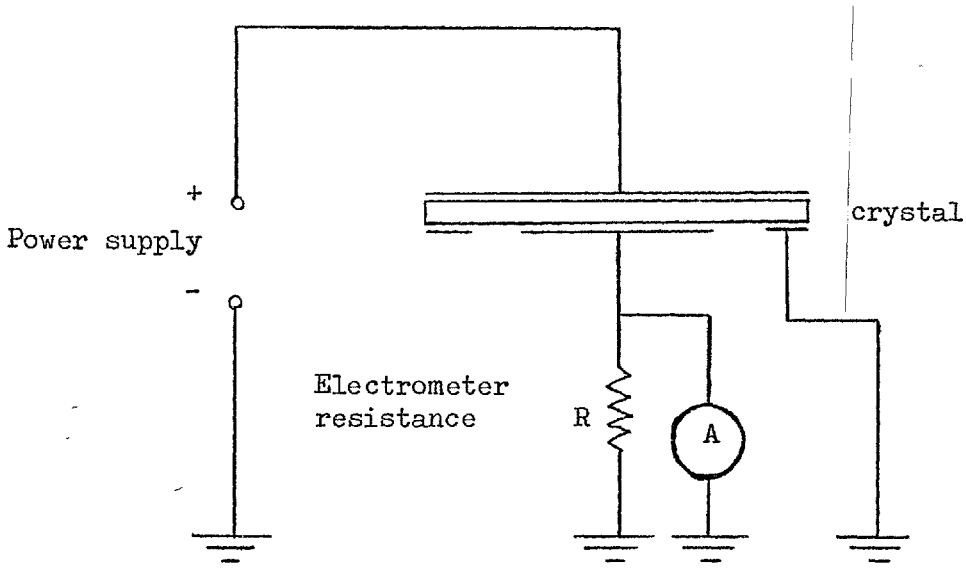


Figure 3.1 Circuit diagram

and the corresponding current measured by an electrometer type instrument (Keithley Electrometer Type 602 Solid State).

In measuring the current through the specimen at a given voltage, the current must be checked in different combinations of amplifier gain and input resistor, giving the same nominal sensitivity (Mills, 1968). If the same result is obtained on more than one range, then this value is assumed to be correct. The furnace heating current was turned off while making current measurements to avoid interference from it.

3.1.2. a.c. measurements

Two sets of instruments were used for this purpose. A Wayne-Kerr Universal Bridge Type B221 was used to measure conductance and capacitance between frequencies 70-20,000 Hz. A Wayne-Kerr Radio Frequency Bridge Type B601 and Source and Detector SR 268 were used to obtain conductance and capacitance values between frequencies 100 KHz to 5 MHz. The r.m.s. value of the voltage applied was always 0.5 volt to avoid applying a voltage that might have a peak value greater than the decomposition voltage. The furnace was again turned off as while making d.c. measurements.

An advantage of the Bridges for the present application is their ability to make accurate measurements even when long connecting cables are used. The design of the Bridges is based on the transformer ratio-arm principle. The unknown impedance is balanced against standards of conductance and capacitance in parallel. Tappings on the two bridge transformers, connected to decade controls, permit measurements to be made accurately on a wide range of impedance in any quadrant of the complex plane.

A simplified diagram of the circuit arrangement is shown in Fig. 3.2, where Z_u and Z_s are the unknown and standard impedances

respectively. Balance will be indicated by the detector, connected to the secondary of the current transformer T2, when equal currents flow from either end of the centre-tapped primary of T2. In this condition, the potential across the primary will be zero and, therefore, the right-hand terminals of Z_u and Z_s will be at a neutral potential. Thus the same voltage is applied to both impedances and, for equal currents to flow in the two halves of T2 primary, the resistive and reactive components of the unknown impedance must be equal to those of the standard.

The transformers are so designed that very heavy shunting is possible without seriously affecting the accuracy of measurements. Components can in most instance, therefore, be measured in situ.

3.2. Construction of the Sample Holder

The sample holder employs a fully guarded 3-terminal electrode system and is similar to that of Moulson et al (1968). The sample holder is shown schematically in Fig. 3.3. and consists of a stainless steel body A, which receives the alumina furnace tube B at one end where it is sealed by an O-ring C compressed by a threaded ring nut D, while it is closed by an end plate E at the other end. The electrodes, and connecting wires are made from platinum, 10% rhodium-platinum and 13% rhodium-platinum.

The inner tube F (5 mm i.d. and 8 mm o.d.) carried the lead that makes contact with the centre electrode. The alumina guard cylinder G (10 mm i.d., 14 mm o.d.) is centred around the inner tube by using two alumina inserts; the outer surface of this cylinder is covered with 10% rhodium-platinum foil (about 40 cm long) which makes contact with the guard ring on the sample. Thus the guard ring on the sample has been effectively extended to a region outside the hot zone of the

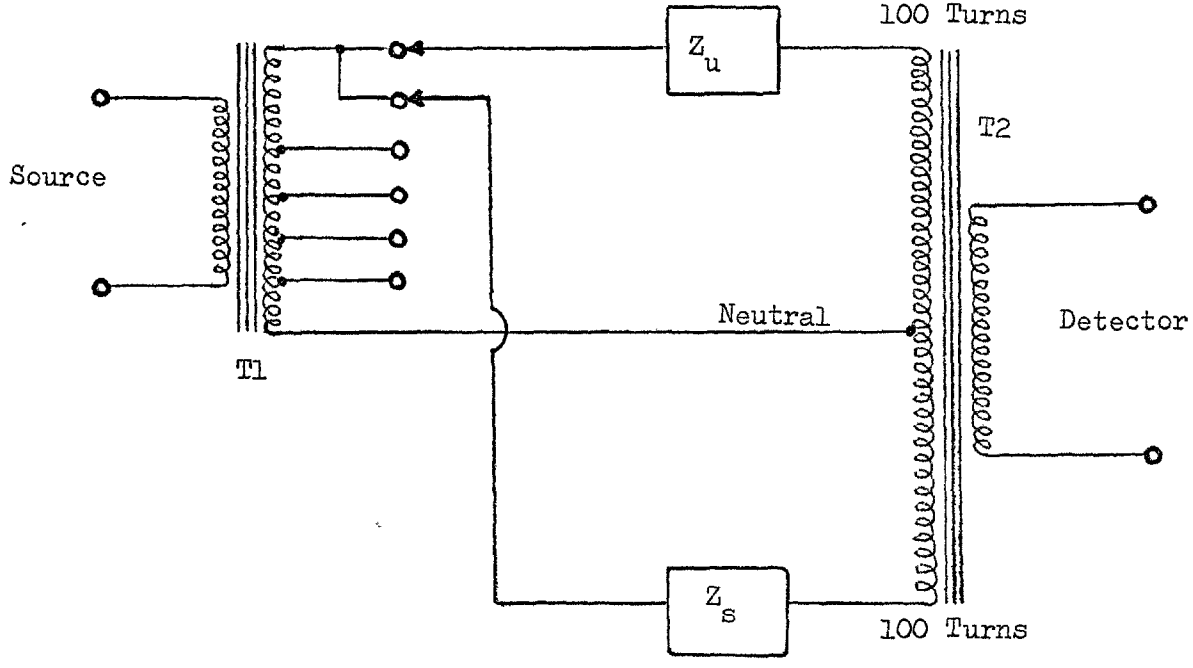


Figure 3.2. Principle of operation

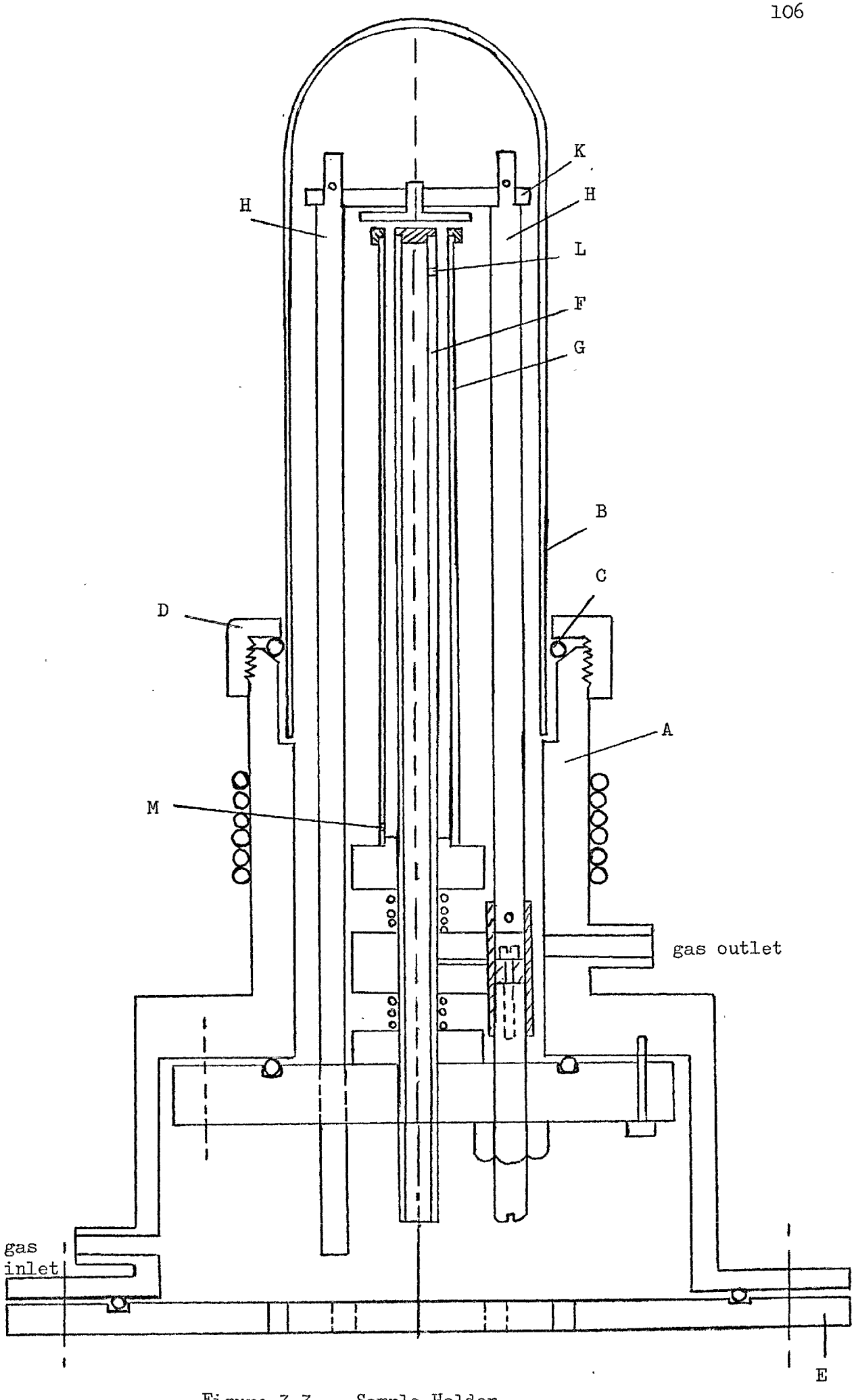


Figure 3.3 Sample Holder

apparatus. The two stainless steel springs maintained the inner and the guard cylinder in definite contact with the specimen. The supporting rods H and circular alumina piece K were attached by alumina pins. The whole system has been designed to be vacuum tight. All leads have been brought out through the stainless steel base-plate via vacuum lead throughs (type 7D, Edwards Vacuum Components, Crawley, Sussex). The hot junction of a platinum : 13% rhodium-platinum thermocouple has been located outside of the guard cylinder, close to the specimen.

Coaxial connections to the sample holder were made through BNC connectors which are attached to the base-plate.

To maintain a proper gas flow through the sample holder, gas comes in through the "gas inlet", goes through tube F and then through the holes L and M into the space between tube F and G and finally goes out through the "gas outlet".

3.3. Sample Ambient Control

3.3.1. Gas control system

This is shown diagrammatically in Fig. 3.4. Argon and oxygen are purified as follows:

(i) Soda lime removes carbon dioxide. Since the reaction produces moisture, it must be placed first in the system.

(ii) Silica gel absorbs moisture.

(iii) Magnesium Perchlorate acts as a fine abstracter of moisture.

(iv) The gas then flows through titanium granules that are heated to about 700-800°C to remove oxygen.

Soda lime will not be used while CO/CO_2 mixtures flow through the system. All connections are made by using glass or stainless steel tubing since plastic tubing is porous to moisture and oxygen. The stainless steel tubes were cleaned before assembly with a special treatment to get rid of lubricating oils remaining inside from the extrusion process used in

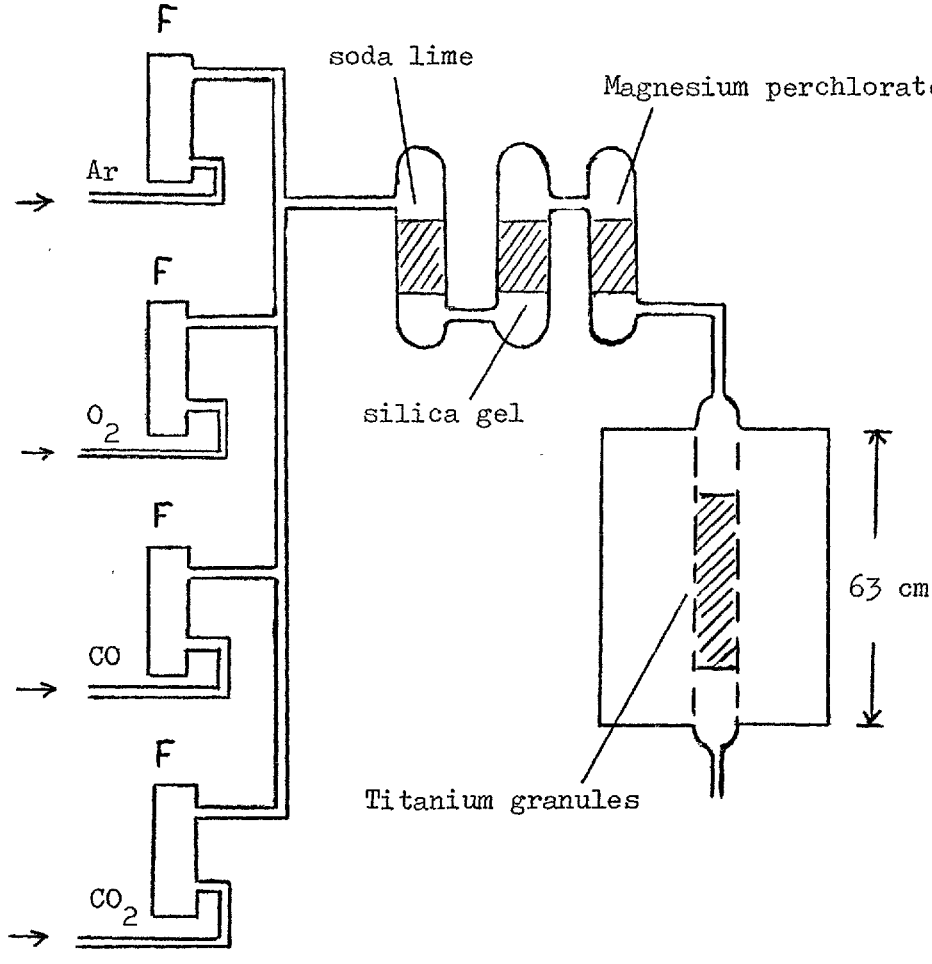


Figure 3.4 Gas Purification System

manufacture. The purified gas goes through the oxygen monitor to check the partial pressure of oxygen. The oxygen monitor uses a solid state oxide electrolyte which conducts by oxide ion movement. The principle of such a device consists in setting up a cell of type (Steele, 1968)

$$p'_{O_2} \mid \text{Solid Oxide Electrolyte} \mid p''_{O_2}$$

in which the cell reaction at high temperature involves the transfer of oxygen ions from the side of the high oxygen partial pressure p''_{O_2} towards the lower p'_{O_2} through the movement of doubly charged vacancies. The oxygen partial pressure p'_{O_2} and p''_{O_2} fix the chemical potential given by

$$\mu_{O_2} = \mu^{\circ}_{O_2} + RT \ln p_{O_2} / p^{\circ}_{O_2} \tag{3.1}$$

where $p^{\circ}_{O_2}$ is the pressure of oxygen in the standard state. The EMF of the cell will therefore result from the difference in chemical potentials μ''_{O_2} and μ'_{O_2} which are related by

$$\mu_{O_2} = \Delta\mu''_{O_2} - \Delta\mu'_{O_2} = 4\,EF \tag{3.2}$$

where E is the electromotive force and F is Faraday's constant. The potential difference set up between two electrodes is given as

$$E = \frac{RT}{4F} \ln p''_{O_2} / p'_{O_2} \tag{3.3}$$

Operating with a known oxygen partial pressure on one side, the unknown oxygen partial pressure can be calculated from the measured potential difference.

Construction of the cell.

The essential components of the cell are two electronically conducting electrodes in physical contact with a solid electrolyte, as shown in Fig. 3.5. Platinum is used as the electrodes because of its superior chemical resistance. The use of such cells is limited by the onset of electronic conduction in the solid electrolyte, which reduces the measured potential below theoretical value. The point at which electronic conduction occurs is a function of both temperature and partial pressure of oxygen.

For this reason $Zr_{0.85}Ca_{0.15}O_{1.85}$ is used for which the electronic conduction within temperature range 600°C to 1500°C and oxygen partial pressure $1 < p < 10^{-25}$ is so small that for most practical purposes equation (3.3) can be safely applied to provide accurate values for the approximate oxygen partial pressure change.

A standard electrode of fixed oxygen activity, consisting of equal amounts of Fe and its oxide FeO, has been used because it exhibits a high degree of reversibility.

There are four flowmeters, type 1100, manufactured by Rotameter Manufacturing Co. Three of the flowmeters have 15 cm long tubes and one has a 30 cm long tube. The partial pressure of oxygen can be adjusted by changing the flow rate of argon and oxygen to the calculated values. The accuracy of the readings depends on the length of the tube and is better for the longer tube: $\pm 3\%$ of indicated flow for the 30 cm long tube and $\pm 4\%$ for the 15 cm long tube.

The oxygen monitor furnace and the titanium furnace are horizontally wound with 24 and 16 gauge Kanthal wire respectively.

3.3.2. Sample Holder Furnace

The experiments were performed in a horizontal molybdenum wound

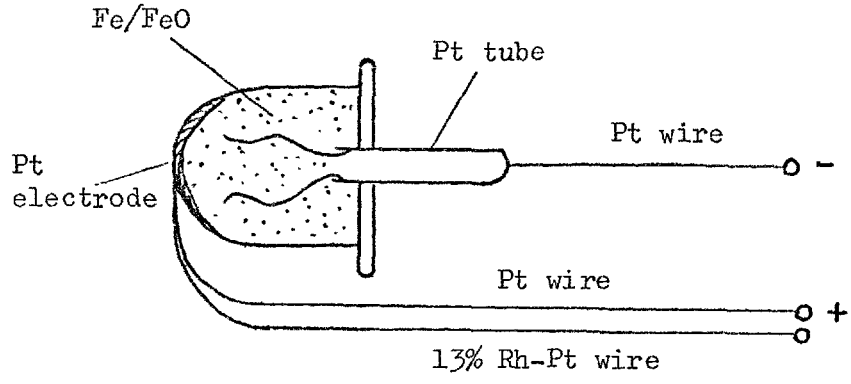


Figure 3.5 Oxygen monitor

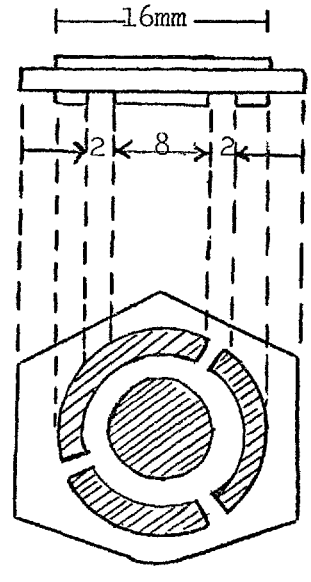


Figure 3.6 Crystal with electrodes on

furnace. The heating element is a molybdenum tape 0.5 mm thick and 3 mm wide, wound on to a recrystallized alumina tube which is about 95 cm long and 65 mm in diameter, at approximately six turns per inch, giving a total resistance of just five ohms when cold. The winding tube is sealed into a metal container filled with alumina powder. A continuous flow of forming gas must be kept running all the time when the furnace is in use to avoid oxidation of the molybdenum wire. The temperature is controlled by a sheathed thermocouple (platinum : 13% rhodium-platinum) placed in the centre, in conjunction with an automatic controller (Eurotherm, PID/SCR series). The cold junction, which is connected in series, is kept at $0^{o}C$ in a water-ice mixture.

3.4. Samples

The samples were of two kinds. The first kind was platelets grown by vapour transport from a lead fluoride flux melt in our own Crystal Growth Laboratory. These crystals have their c-axis perpendicular to the plane of the platelets and measurements were made on specimens of thickness 13 and 17μ.

The second kind of crystal was obtained from Crystals Systems Inc., Salem, Mass., U.S.A., and were 260μ thick. All samples were electroded by sputtering platinum for 1 hour at 100 w power and 3μ argon pressure in an M.R.C. sputtering unit. Under these conditions 3000 Å thick platinum layers were obtained. Pure platinum was used as the electrode material and has been sputtered onto the sample in the three-terminal electrode configuration. The electrode dimensions are given in Fig. 3.6. Aluminium and molybdenum masks were used to produce the electrode patterns.

The way in which the measurements were organized: In the first and second runs $\sigma_{ac}(f)$ was measured at two temperatures first, but after this

a standard procedure was adopted, i.e. the p_{O_2} was set and runs of rising and falling temperature made. At each set temperature we waited for constant $\sigma_{d.c.}$ then measured $\sigma_{a.c.}(f)$.

CHAPTER 4 - RESULTS

Measurements were made on a total of 5 single crystal samples as detailed in Table 4.1.

Samples 1 to 4 were flux-grown platelets which proved to be very fragile to temperature cycling, and this restricted the amount of information which could be obtained from them before they failed by cracking. The table shows the highest temperature reached and the range of a.c. measurements made, and if d.c. measurements were completed at the highest temperature.

Mass spectrographical analysis of both flux and Czochralski grown crystals shows that they contain a larger amount of silicon (368 and 267 ppm respectively) and in addition the flux grown crystal contains iron (1238 ppm). See tables 4.2 and 4.3.

Samples numbers 1, 2 and 4 were not heated above 1050°C, which will be called low temperatures, and samples numbers 3 and 5 were taken up to 1350°C and 1500°C respectively which will be termed high temperatures.

The a.c. conductivity versus frequency plot of Sample 1 in Fig. 4.1 shows that 950°C and 850°C curves approach each other if the frequency is above 5000Hz. The measurements at 850°C were made as temperature reduced from 950°C which was the highest temperature for this sample and the highest frequency was 20 KHz. No d.c. measurements had been made when the crystal failed.

Sample 2 was heated up to 1050°C, but measurements were made while going up for the first time and as coming down. One observes, on first heating to 750°C, a three orders of magnitude change in a.c. conductivity over the frequency range from 70 Hz to 1 MHz, but this variation reduces with temperature to less than one order at 1050°C. When the same measurements were made on coming down from 1050°C, the a.c. conductivity changes

Table 4.1.

Sample no.	Thickness	Type of growth	Highest Temp.	Maximum frequency and d.c.	p_{O_2} at.
1	17μ	flux-grown	950°C	20 KHz	1
2	13μ	" "	1050°C	1 MHz and d.c.	1
3	13μ	" "	1350°C	1 MHz and d.c.	1 and 10^{-5}
4	7.83μ	" "	850°C	1 MHz and d.c.	1
5	260μ	Czochralski	1500°C	1 MHz and d.c.	1, 10^{-1} and $2x10^{-2}$

Table 4.2 Mass Spectrographic Analysis of Sapphire Sample 3

Element	Impurity concentration (ppm atomic)	Element	Impurity concentration (ppm atomic)	Element	Impurity concentration (ppm atomic)
U	0.1	Ce	0.1	Cu	17
Th	0.1	La	0.1	Ni	10
Bi	0.1	Ba	1	Co	11
Pb	16	Cs	3	Fe	1,238
Tl	0.4	I	0.1	Mn	12
Hg	0.4	Te	0.3	Cr	34
Au	0.1	Sb	3	V	3
Pt	10,088	Sn	1	Ti	46(3)
Ir	0.3	In	1	Sc	3
Os	0.4	Cd	1	Ca	340
Re	0.2	Ag	0.2	K	32
W	0.4	Pd	1	Cl	150
Ta	34(1)	Rh	3,408	S	102
Hf	0.4	Ru	0.7	P	30
Lu	0.1	Mo	0.7	Si	368
Yb	0.4	Nb	0.1	Al	Matrix
Tm	3	Zr	0.7	Mg	(4)
Er	0.3	Y	0.3	Na	35
Ho	0.1	Sr	0.1	F	33
Dy	0.4	Rb	16	O	
Tb	0.1	Br	22	N	
Gd	0.5	Se	0.2	C	
Eu	0.2	As	11	B	423
Sm	0.4	Ge	28	Be	(5)
Nd	1	Ga	19(2)	Li	3
Pr	0.1	Zn	7		

(1) Possible contamination by source.

(2) Possible contamination from previous sample.

(3) High detection limit due to interference by C ions.

(4) Cannot be determined due to interference by C ions.

(5) Cannot be determined in the presence of Al.

Table 4.3 Mass Spectrographic Analysis of Sapphire Sample 5

Element	Impurity concentration (ppm atomic)	Element	Impurity concentration (ppm atomic)	Element	Impurity concentration (ppm atomic)
U	0.01	Ce	0.01	Cu	0.7
Th	0.01	La	0.01	Ni	2
Bi	0.5	Ba	0.01	Co	0.1
Pb	0.02	Cs	0.01	Fe	11
Tl	0.01	I	0.01	Mn	1
Hg	0.03	Te	0.03	Cr	0.4
Au	(Counter electrode	Sb	0.02	V	3
Pt	0.03	Sn	0.5	Ti	23
Ir	0.02	In	11(2)	Sc	0.1
Os	0.03	Cd	0.08	Ca	10
Re	0.02	Ag	0.3	K	11
W	0.03	Pd	0.04	Cl	227
Ta	3(1)	Rh	0.3(3)	S	5
Hf	0.03	Ru	0.03	P	10
Lu	0.01	Mo	21	Si	267
Yb	0.03	Nb	0.01	Al	Matrix
Tm	0.01	Zr	0.02	Mg	170
Er	0.03	Y	0.2(4)	Na	3
Ho	0.01	Sr	0.01	F	1
Dy	0.04	Rb	0.1	O	
Tb	0.01	Br	20	N	
Gd	0.04	Se	0.02	C	
Eu	0.02	As	0.1	B	0.3
Sm	0.04	Ge	0.1	Be	(5)
Nd	1	Ga	0.9	Li	
Pr	0.01	Zn	0.7		

(1) Contamination from source.

(2) Possible contamination from previous sample.

(3) Possible contamination from source.

(4) Possible contamination by a previous sample.

(5) Cannot be determined in the presence of Al.

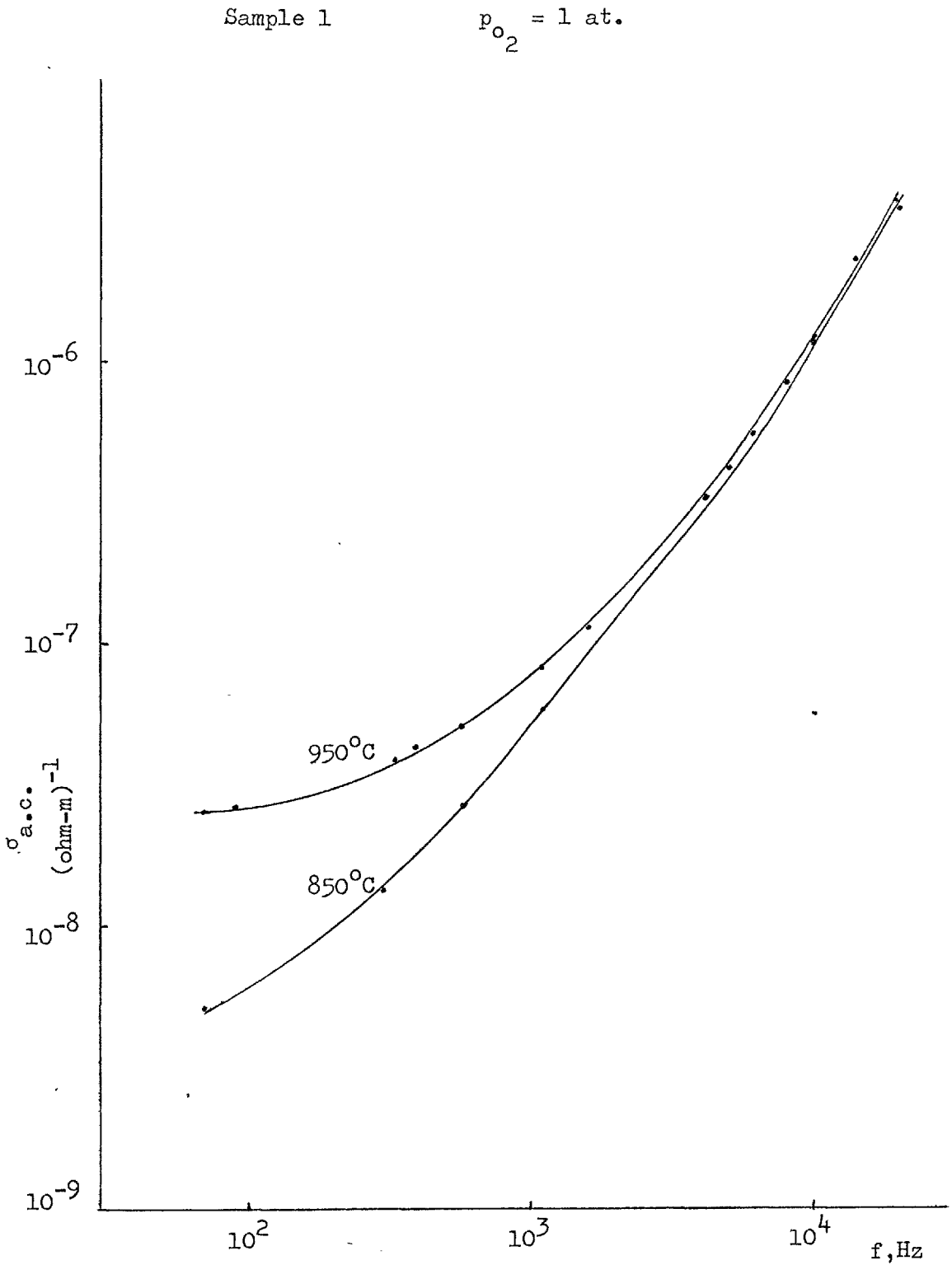


Figure 4.1 Conductivity against frequency

within the same frequency range, became less than one order of magnitude at 750°C and 850°C. See Fig. 4.2 and 4.3. In conclusion one observes that at low temperatures, i.e. below 1050°C, one gets different a.c. conductivity values while making measurements with temperature increasing or decreasing. Not all previous work described in the literature seems to have taken such effects into account, as will be discussed in more detail in the discussion. The a.c. conductivity versus frequency curves, at 1050°C, 850°C, 750°C and 500°C with descending temperature in pure oxygen, are shown in Fig. 4.4. Corresponding d.c. measurements of sample 2 again in pure oxygen and the log σ vs $^{1}/T$ plot in Fig. 4.5, show an activation energy of 0.8 eV in the low temperature region.

The a.c. conductivity versus frequency curves at 750°C and 1050°C at different oxygen partial pressures are plotted in Fig. 4.6. As one observes from these curves a.c. conductivity reduces as the partial pressure of oxygen reduces. The d.c. conductivity against partial pressure of oxygen curves are plotted at 1050°C and 750°C in Fig. 4.7. Sample no. 2 fractured after this amount of information had been obtained, before further d.c. measurements were taken.

Sample 3 was heated up to 1350°C, and a.c. conductivity versus frequency curves are plotted for p_{O_2} = 1 at and 10^{-5} at. in Fig. 4.8 and 4.9. respectively. At high temperatures and low frequencies, a.c. conductivity does not change much as the frequency increases; but it has a frequency dependence at higher frequencies of the form $\omega^{0.25-0.8}$.

The d.c. conductivity versus temperature curves in Fig. 4.10 have two regions. A low temperature region which has an activation energy of 0.95 eV for p_{O_2} = 1 at. and 0.69 eV for p_{O_2} = 10^{-5} at. and a high temperature region with activation energies of 3.1 eV for p_{O_2} = 1 at and 2.98 eV for p_{O_2} = 10^{-5} at., these values compare with the activation energy of 2.9 eV of Brook et al (1971) of flux grown

Sample 2 750°C p_{O_2} = 1 at

(1) As temperature going up for the first time

(2) As temperature going down from 1050°C

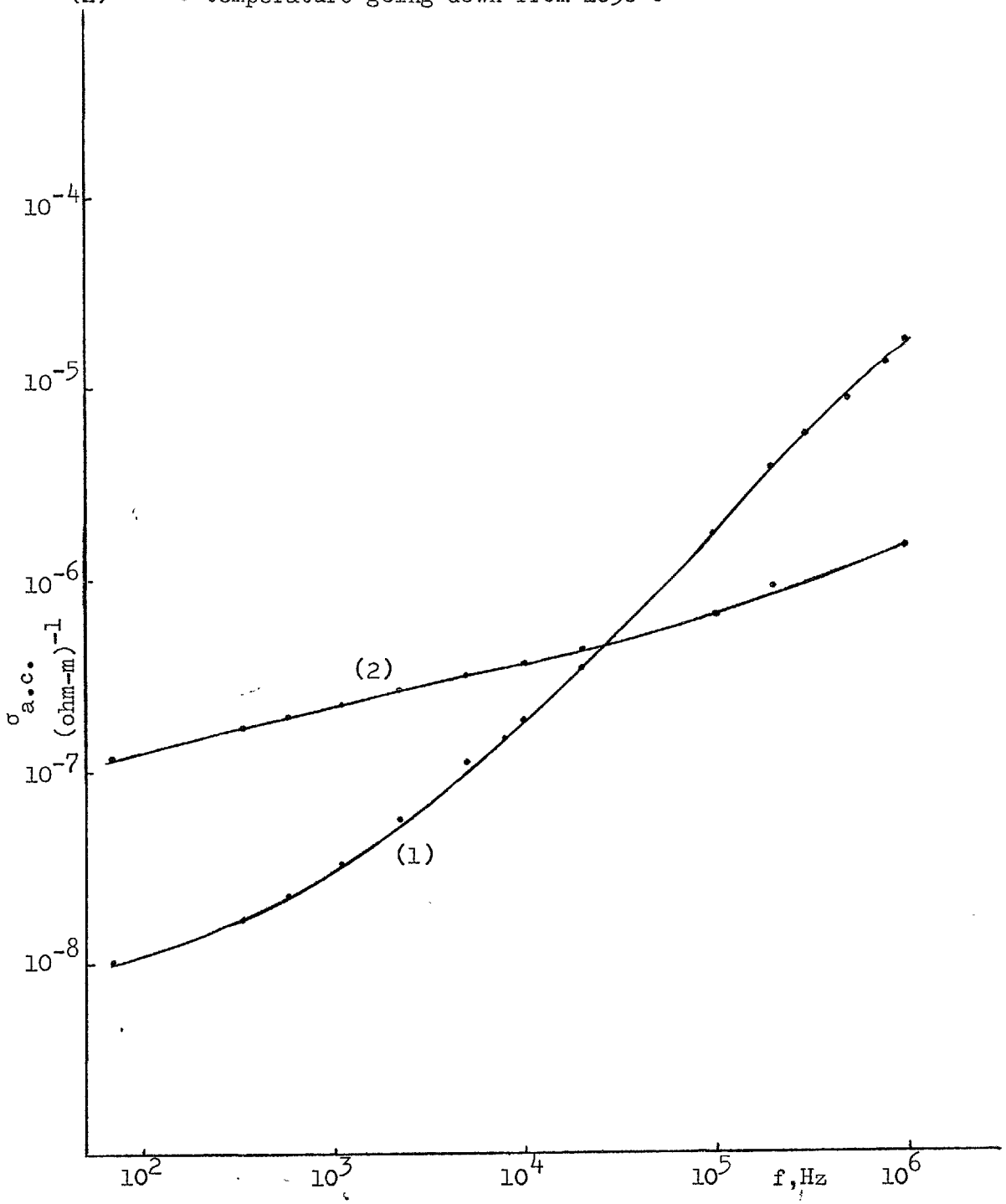


Figure 4.2. Conductivity versus frequency as temperature going up and down

Sample 2 850°C p_{O_2} = 1 at

(1) As temperature going up for the first time

(2) As temperature going down from 1050°C

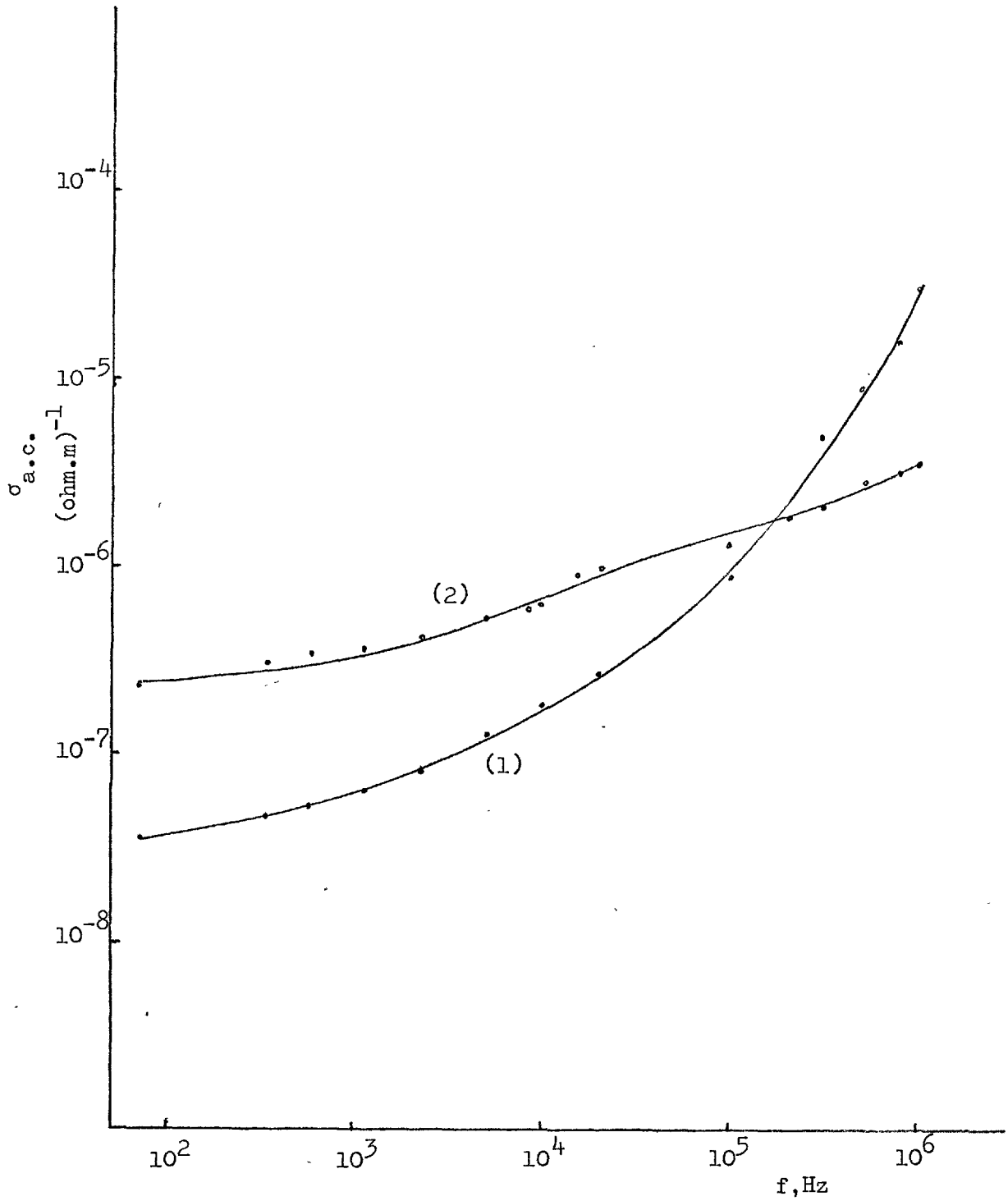


Figure 4.3. Conductivity as function of frequency as temperature going up and down

Sample 2 p_{O_2} = 1 at.

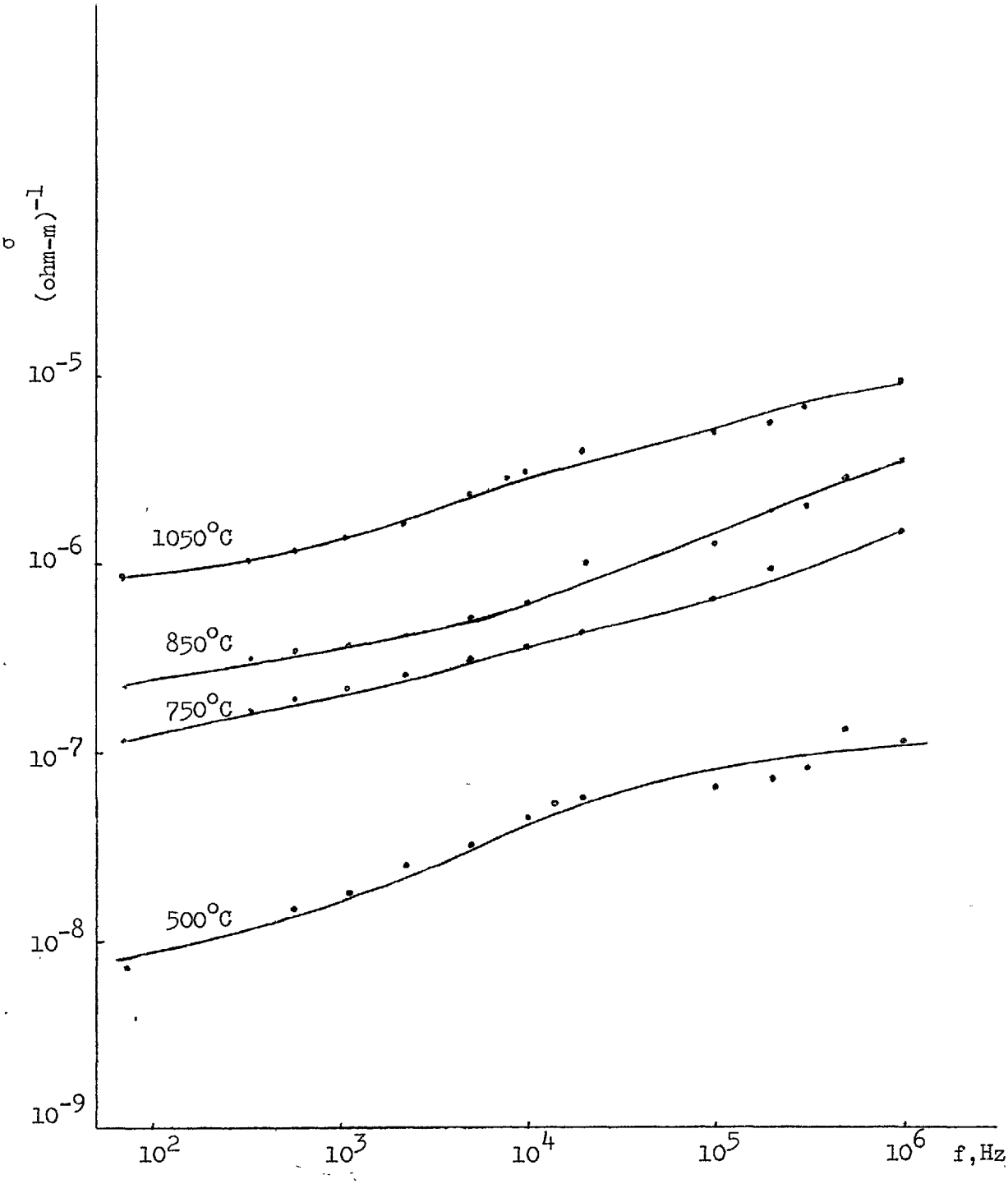


Figure 4.4. Conductivity versus frequency at various temperatures

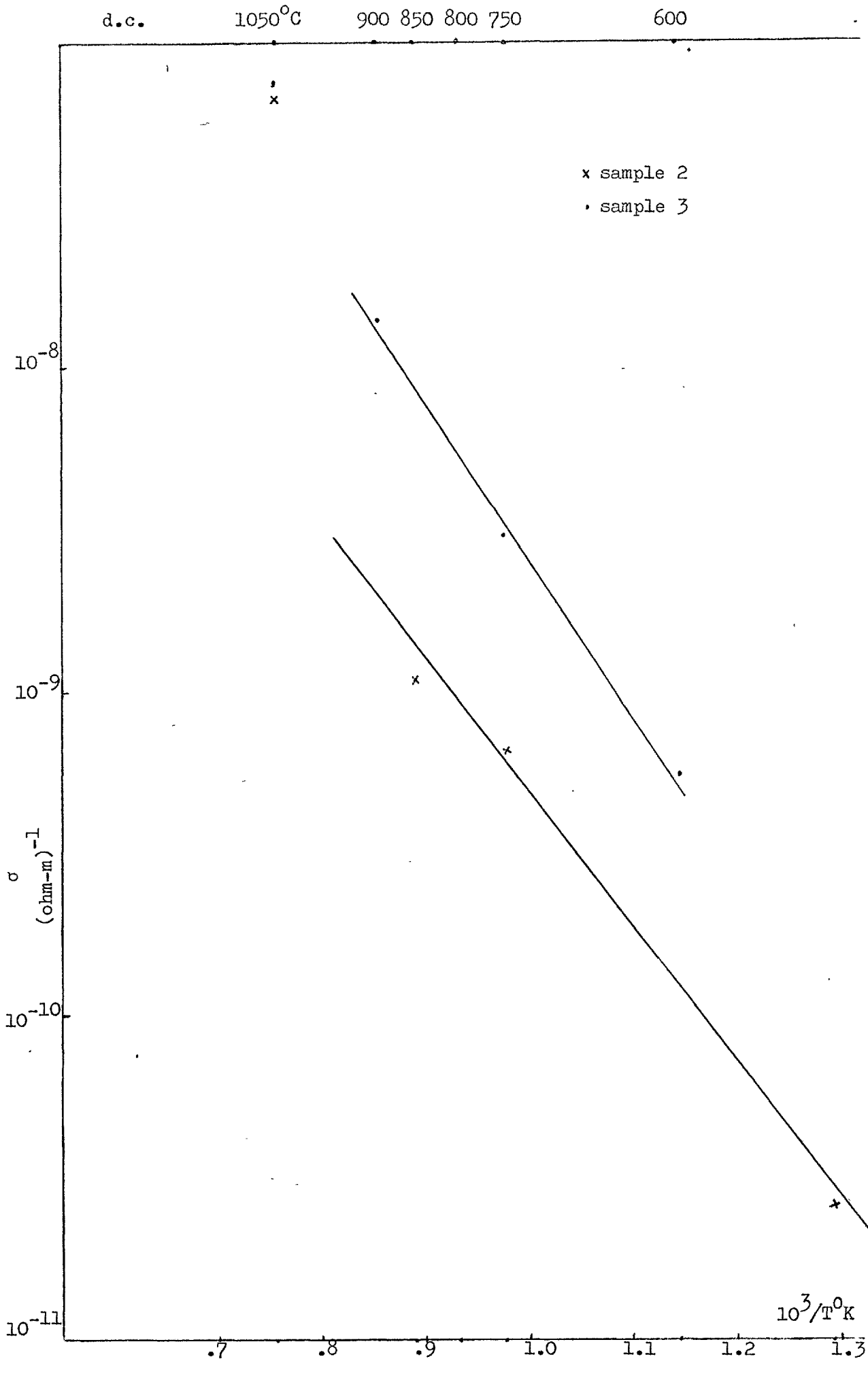


Figure 4.5. Temperature dependence of d.c. electrical conductivity.

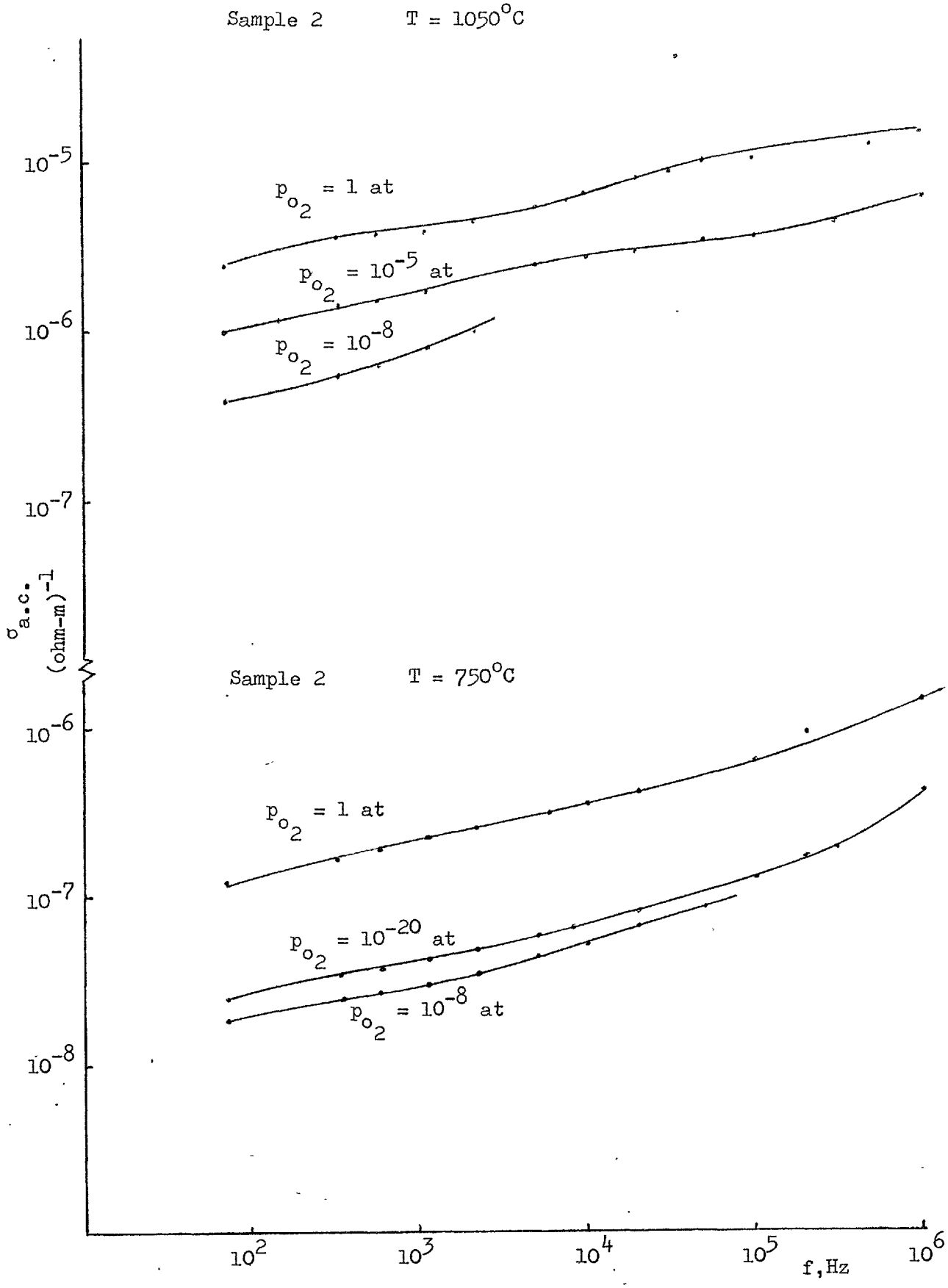


Figure 4.6. Conductivity versus frequency at various oxygen partial pressures

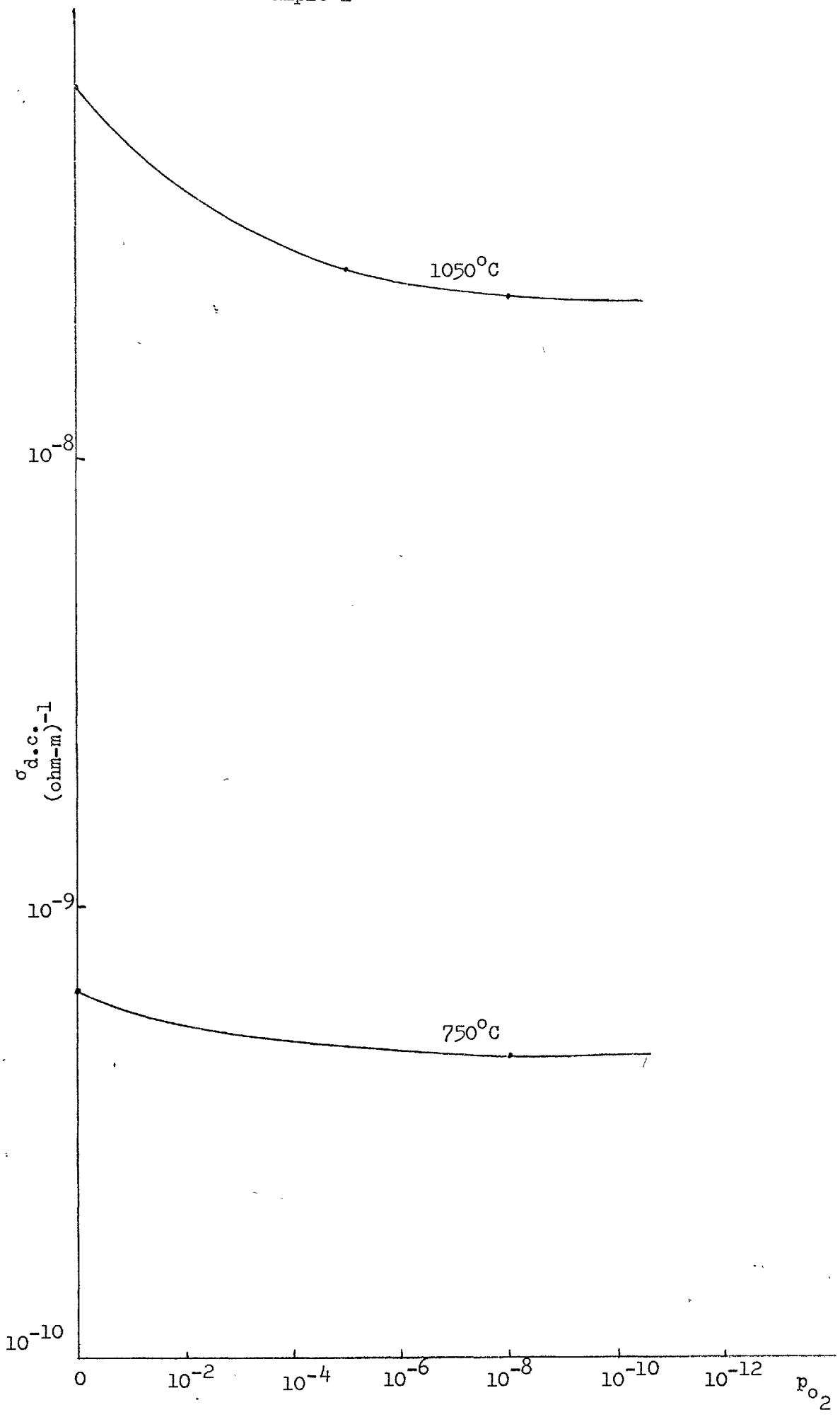


Figure 4.7. d.c. conductivity versus partial pressure of oxygen at various temperatures

Sample 3 p_{O_2} = 1 at

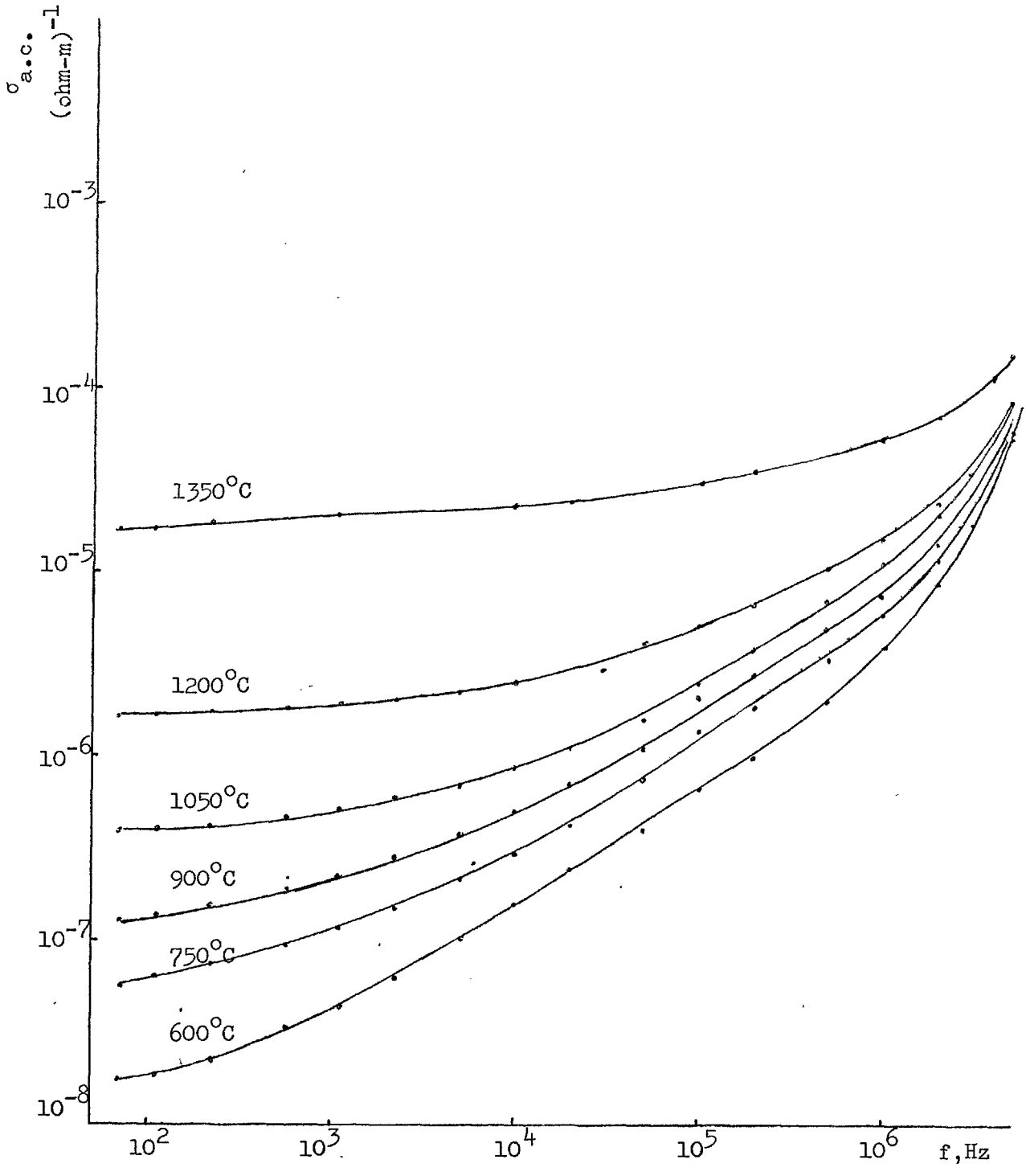


Figure 4.8. Conductivity versus frequency at various temperatures

Sample 3 $p_{O_2} = 4 \times 10^{-5}$ at

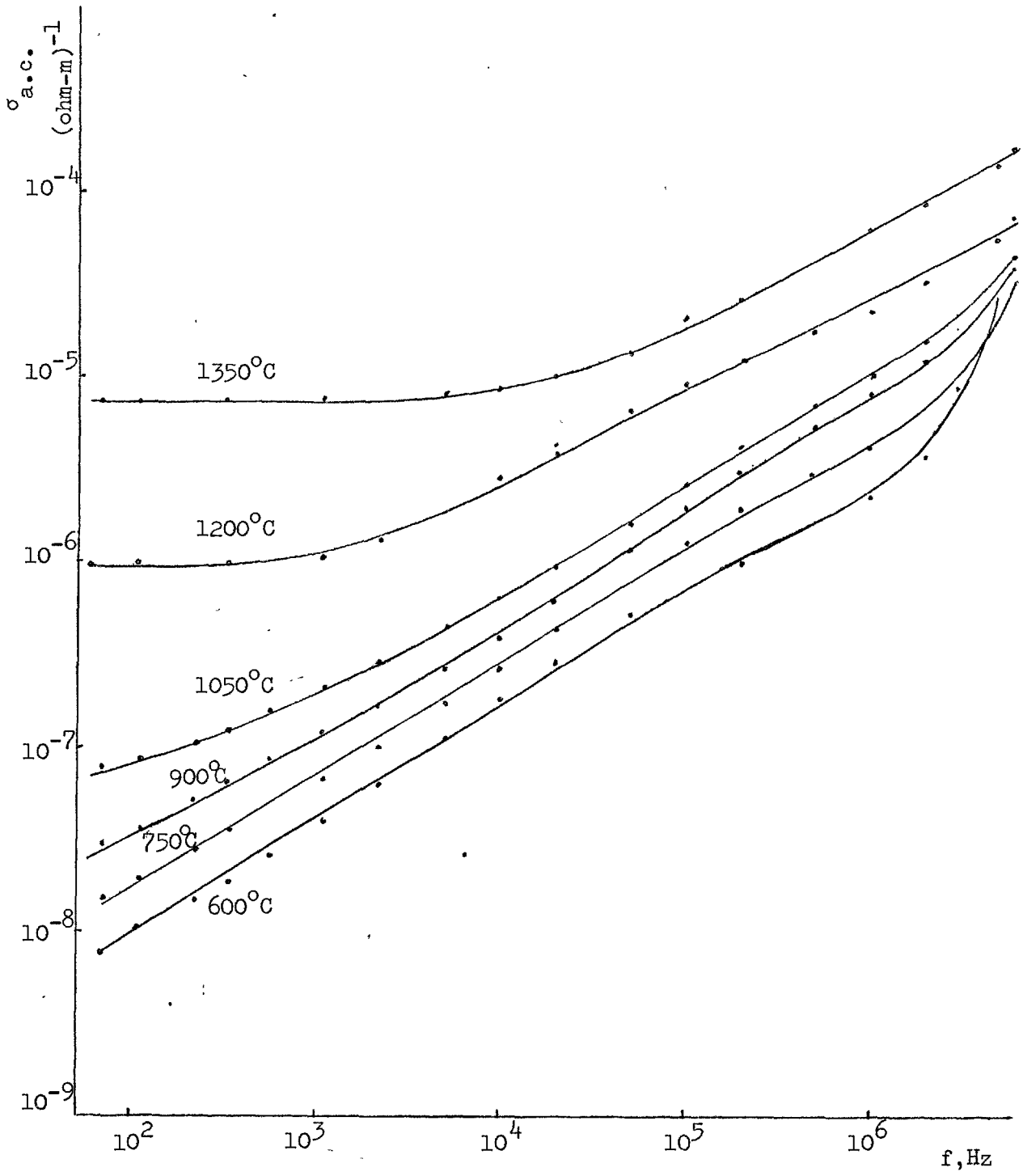


Figure 4.9. Conductivity against frequency at different temperatures.

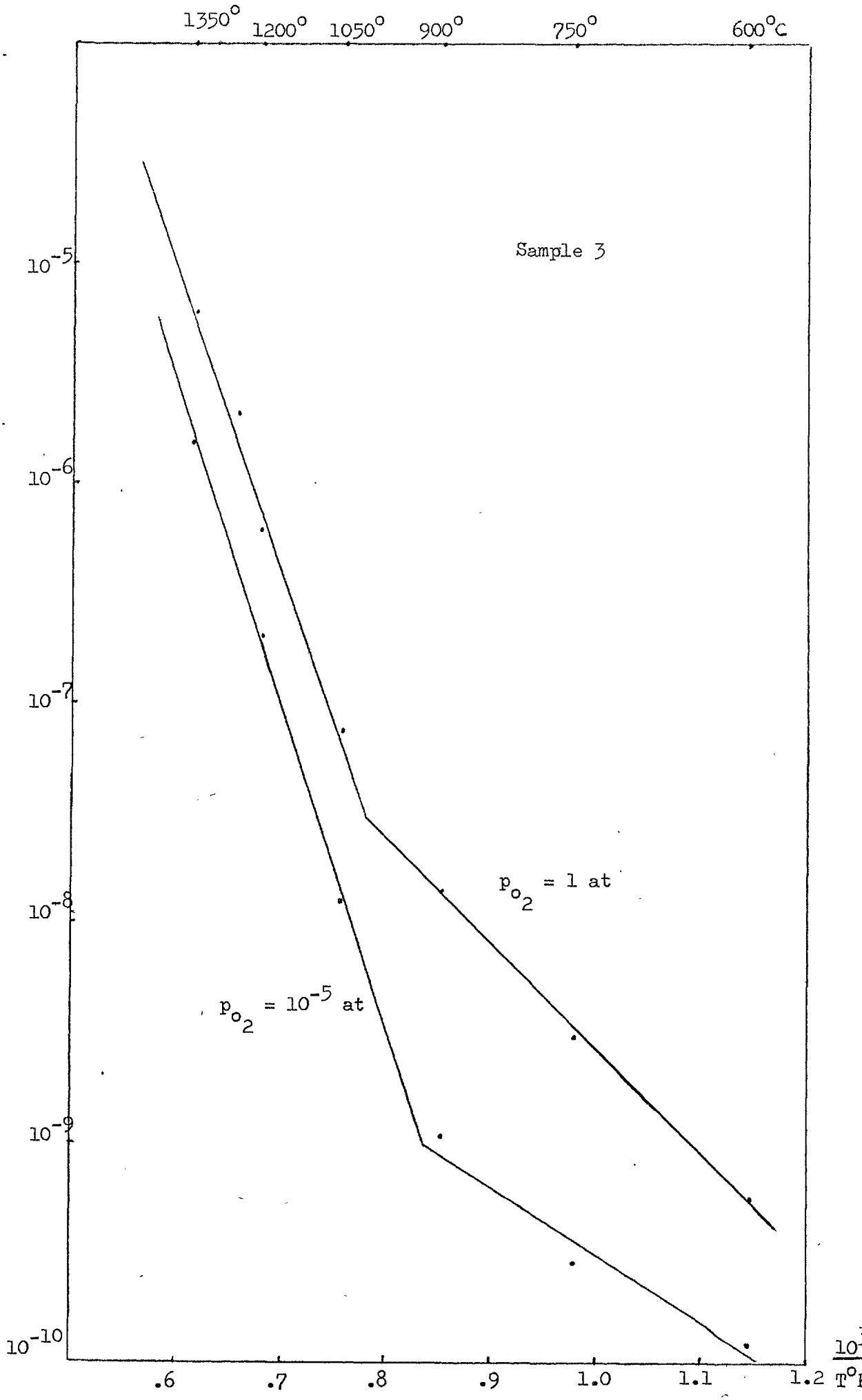


Figure 4.10. d.c. conductivity as a function of tempera ure at various p_{O_2}

sapphire at high temperatures. The d.c. conductivity values at $p_{O_2} = 4x10^{-5}$ at. are lower than those at $p_{O_2} = 1$ at. From Fig. 4.11 it is clear that activation plots of $\sigma_{a.c.}$ versus $^1/T$ where $\sigma_{a.c.}$ is measured at a constant frequency as used by some researchers, will give different slopes from plots of $\sigma_{d.c.}$ versus $^1/T$. In this figure as a representative example, d.c. and a.c. conductivities measured at f = 10 KHz at different temperatures are plotted for comparison. It is apparent that both the conductivity and activation energies are different in the two cases and one gets higher a.c. than d.c. conductivity values. The difference between the two becomes larger as the temperature goes down. For this sample significant changes with temperature are observed in the capacitance versus frequency characteristic and apparent permittivity against frequency curves at different temperatures and 1 at. oxygen partial pressure are drawn in Fig. 4.12 and at $4x10^{-5}$ at. in Fig. 4.13.

More than two orders of magnitude capacitance variation are observed at 1350°C, but this difference decreases as the temperature goes down. The d.c. conductivity against p_{O_2} curves at various temperatures are plotted in Fig. 4.14.

Sample 4, which was 7.83μ thick, was only heated up to 850°C, in pure oxygen before failure, hence the a.c. conductivity versus frequency curves are intermingled after about 3 KHz, as shown in Fig. 4.15. The d.c. conductivity values compare to previous samples and are shown in Fig. 4.16.

Sample 5 was a Czochralski grown crystal, obtained from Crystal Systems Inc., Salem, Mass., U.S.A., and was 260μ thick. Experiments were made at three partial pressures of oxygen to be able to investigate the effect of oxygen partial pressure on sapphire crystals. Analysis of these results shows that d.c. conductivity is oxygen partial pressure

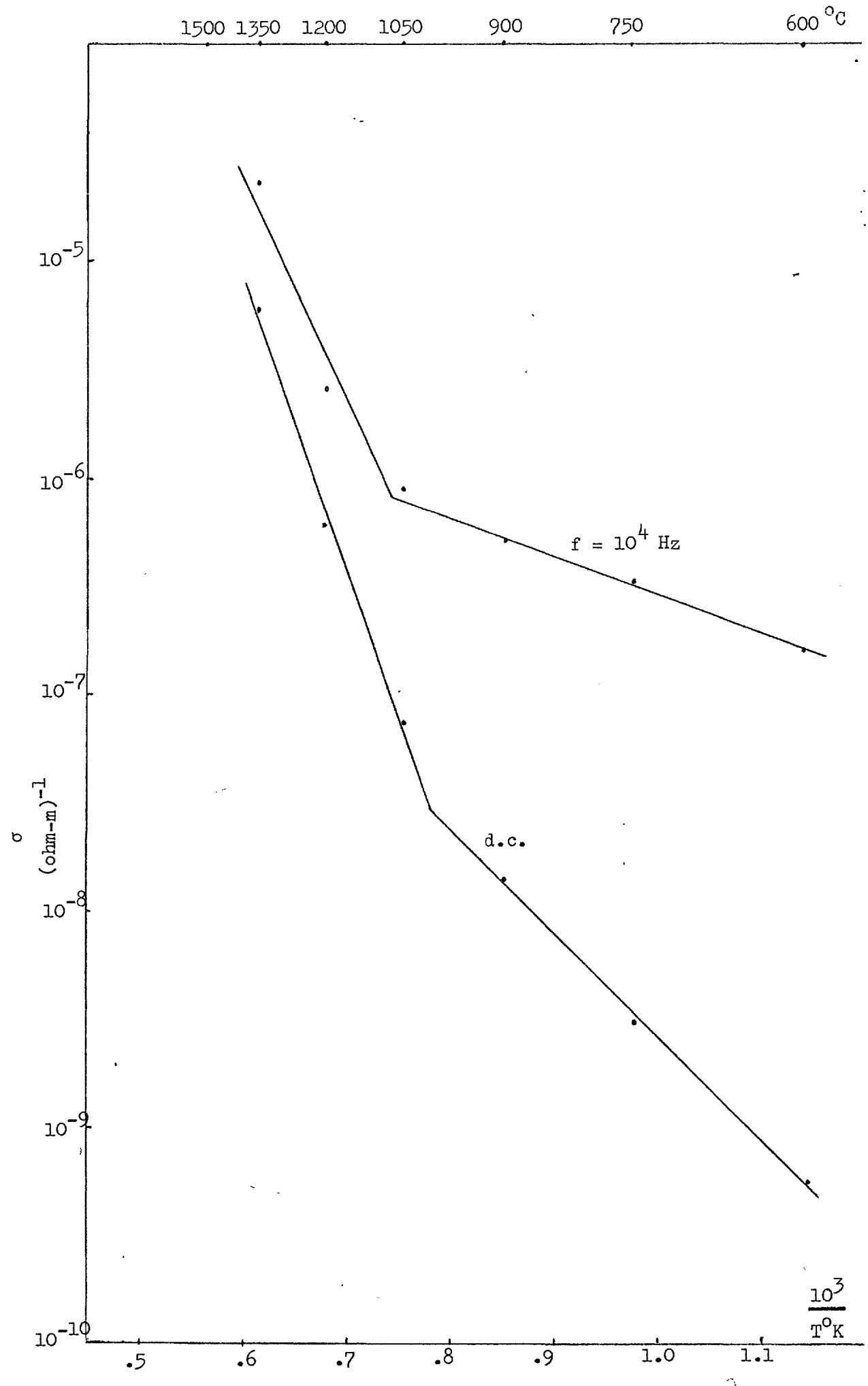


Figure 4.11 Comparison of d.c. and a.c. conductivity as f(T)

Sample 3 p_{O_2} = 1 at

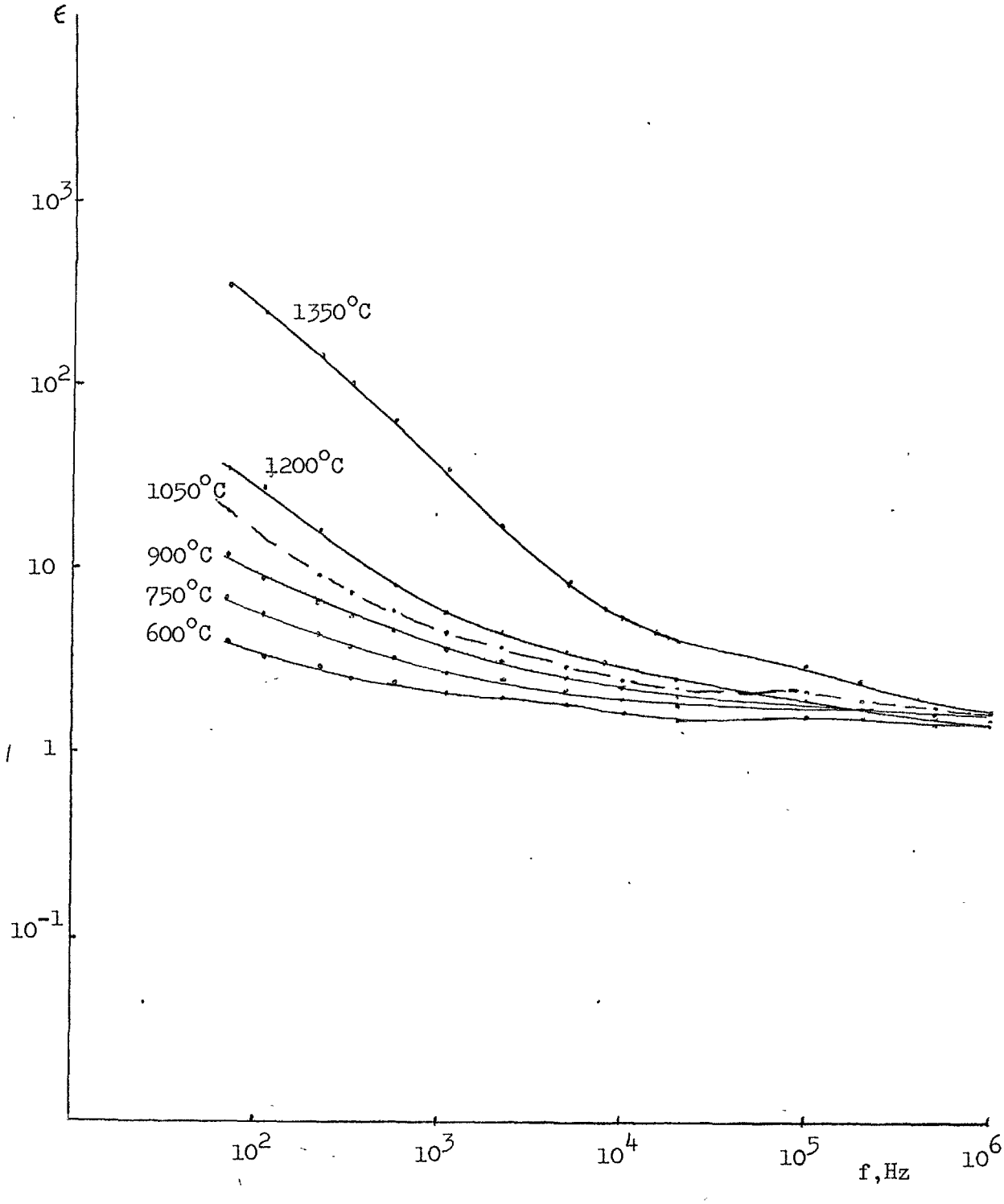


Figure 4.12 Permittivity versus frequency at various temperatures

Sample 3 $p_{O_2} = 4 \times 10^{-5}$ atm

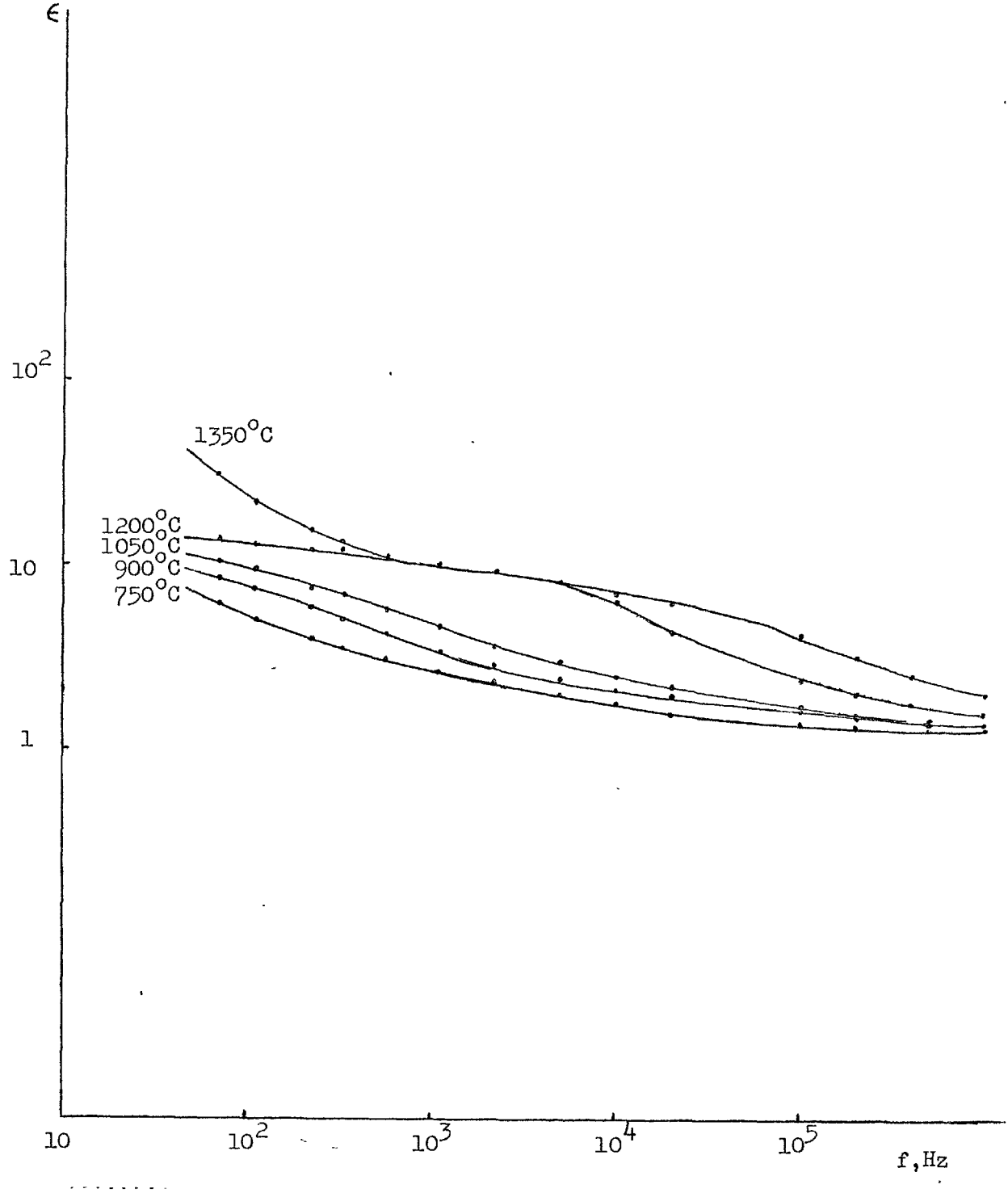
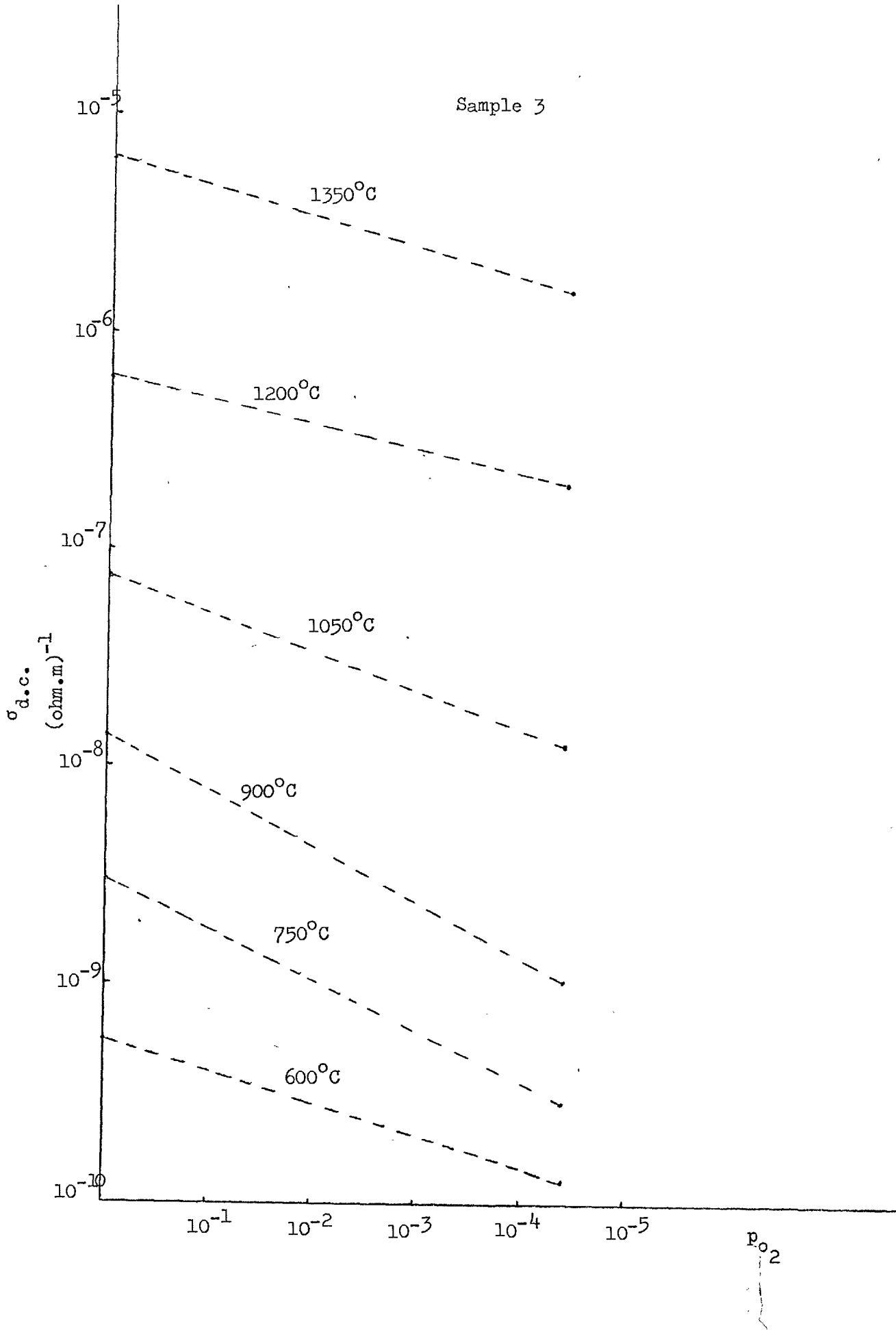


Figure 4.13 Permittivity against frequency at different temperatures

Figure 4.14 D.C. conductivity against oxygen partial pressure at various temperatures



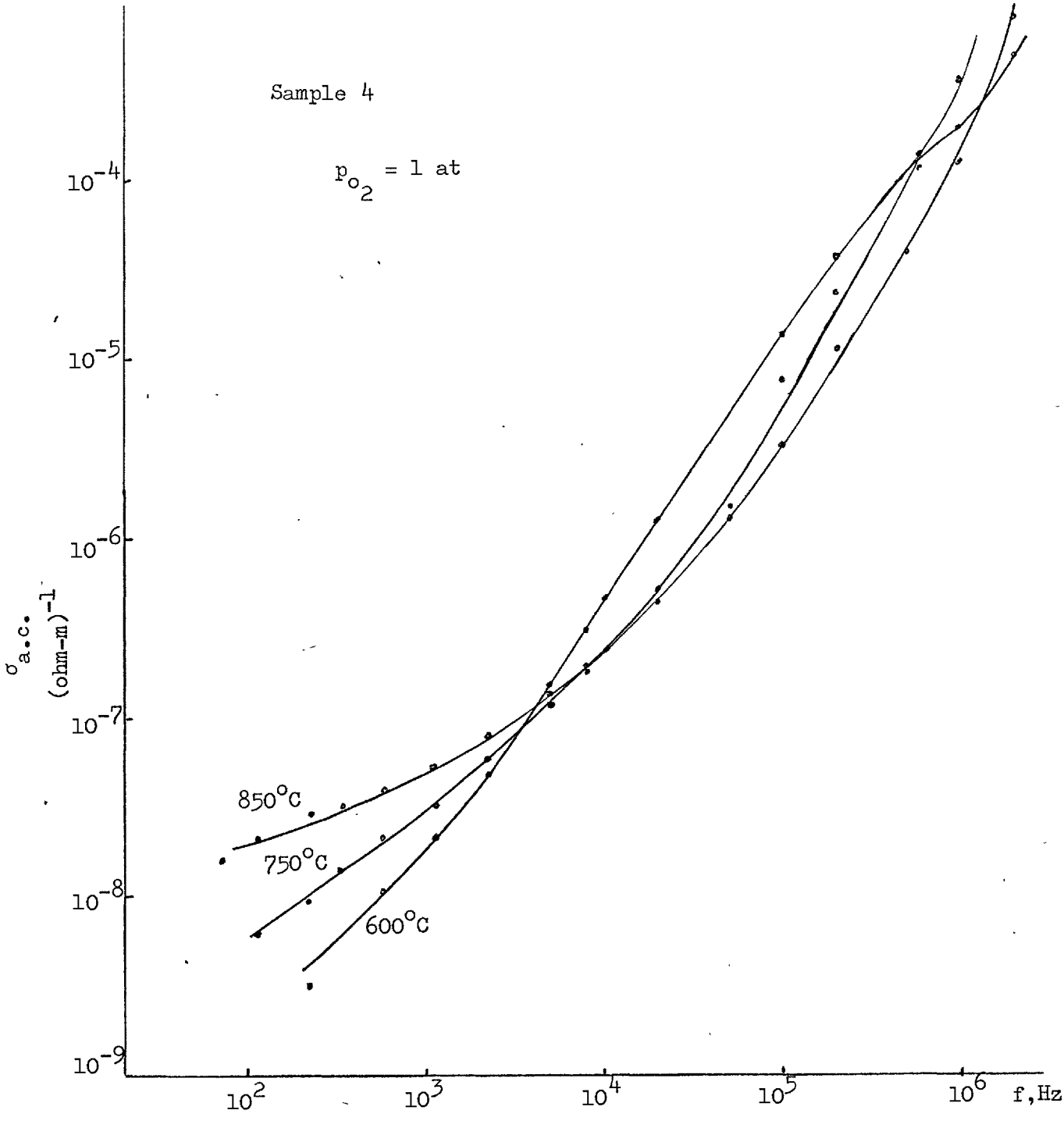


Figure 4.15 Conductivity as a function of frequency at various temperatures.

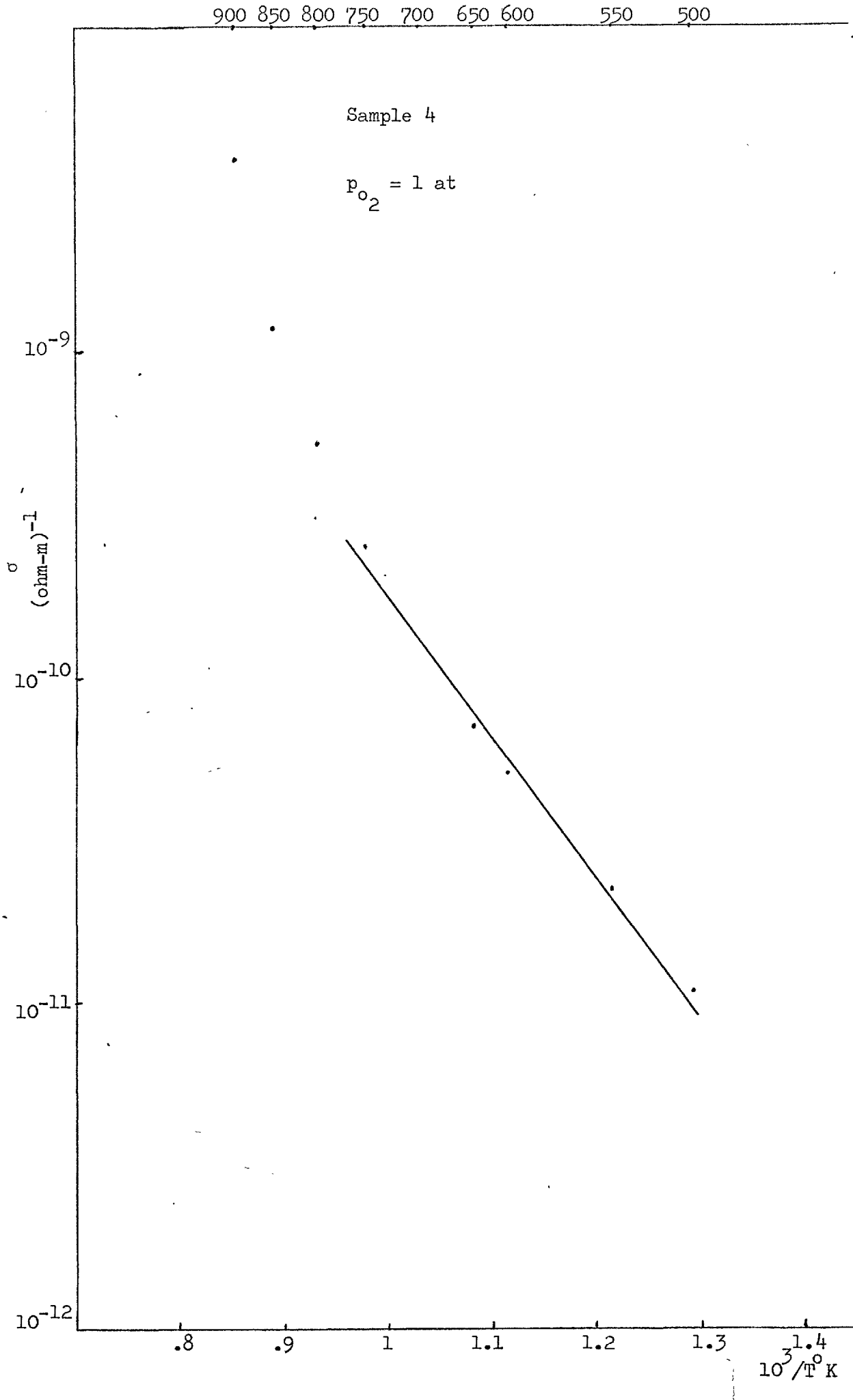


Figure 4.16 D.C. conductivity versus temperature

dependent, i.e. $\sigma = \sigma_o p_{O_2}^k$. In Figures 4.17, 4.18 and 4.19, plots of log σ vs $^1/T$ give activation energies, at both low and high temperatures, which are higher than the flux grown crystals. They change slightly at different oxygen partial pressures, the low temperature average is about 1.5 eV and the high temperature average is about 4.5 eV. The a.c. conductivity versus frequency curves are plotted at several temperatures and three oxygen partial pressures in Figures 4.20, 4.21 and 4.22. One again observes at high temperatures, namely 1500°C and 1350°C, a flat region at low frequencies followed by a slow increase on the log σ vs log f plot. This occurs at all partial pressures, but with small changes in overall level. As the temperature goes down to 1200°C and below one cannot observe the above mentioned flat region. Although the lowest temperature that measurements were made at was 900°C in pure oxygen, as the oxygen partial pressure reduces it was not possible to make measurements with the Radio Frequency Bridge at this temperature because the conductivity becomes too low. The d.c. conductivity versus p_{O_2} curves at various temperatures are plotted in Fig. 4.23.

One measures G and C with the bridges but it is possible to calculate $R = \frac{G}{G^2 + (\omega C)^2}$ and $X = \frac{C}{G^2 + (\omega C)^2}$ and then plot resistance versus reactance. A typical result is shown in Fig. 4.24. These curves show a distribution of relaxation time since the shape of the curves are broadened rather than being semicircular. If these curves are extrapolated to find the intersection with the resistance axis on the low frequency side, one obtains extrapolated R values. The approximate activation energy, corresponding to these extrapolated resistance values, is smaller (3.38 eV) than the d.c. value at high temperatures. Assuming these curves belong to the bulk and there exists another curve between the d.c. value and the extrapolated resistance value which belongs to electrode processes, then the

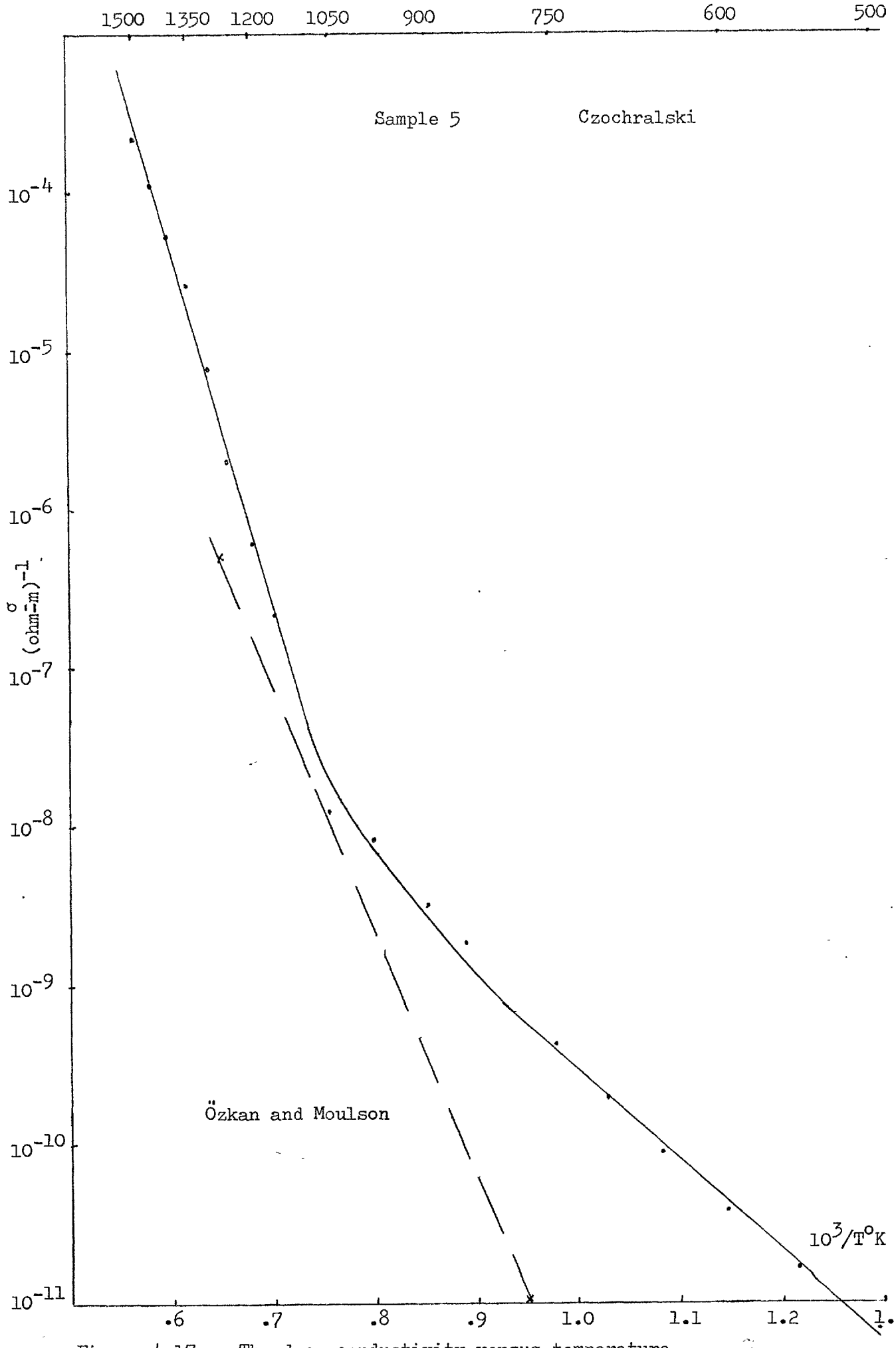


Figure 4.17 The d.c. conductivity versus temperature

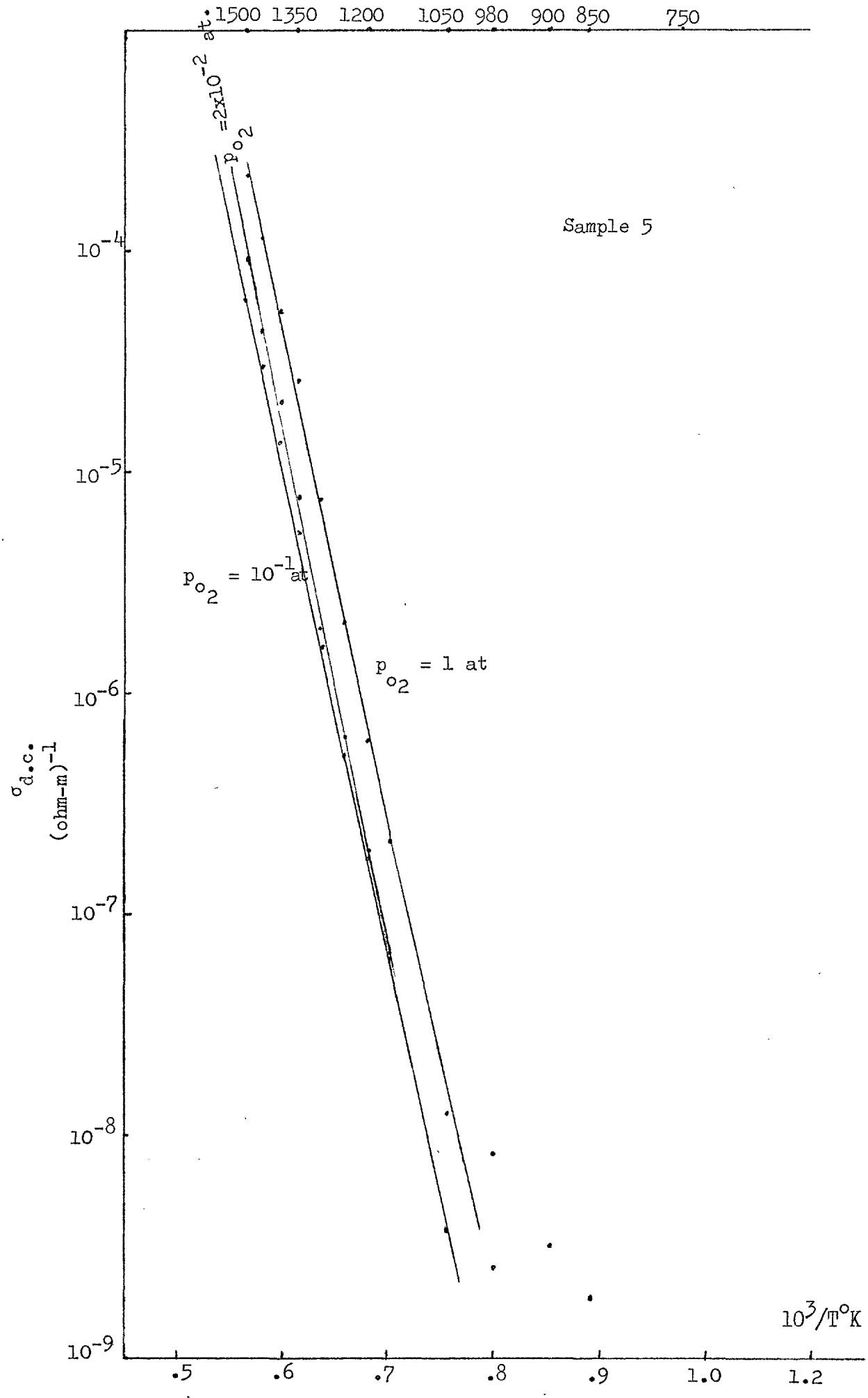


Figure 4.18 D.C. conductivity as f(T) at various p_{O_2}

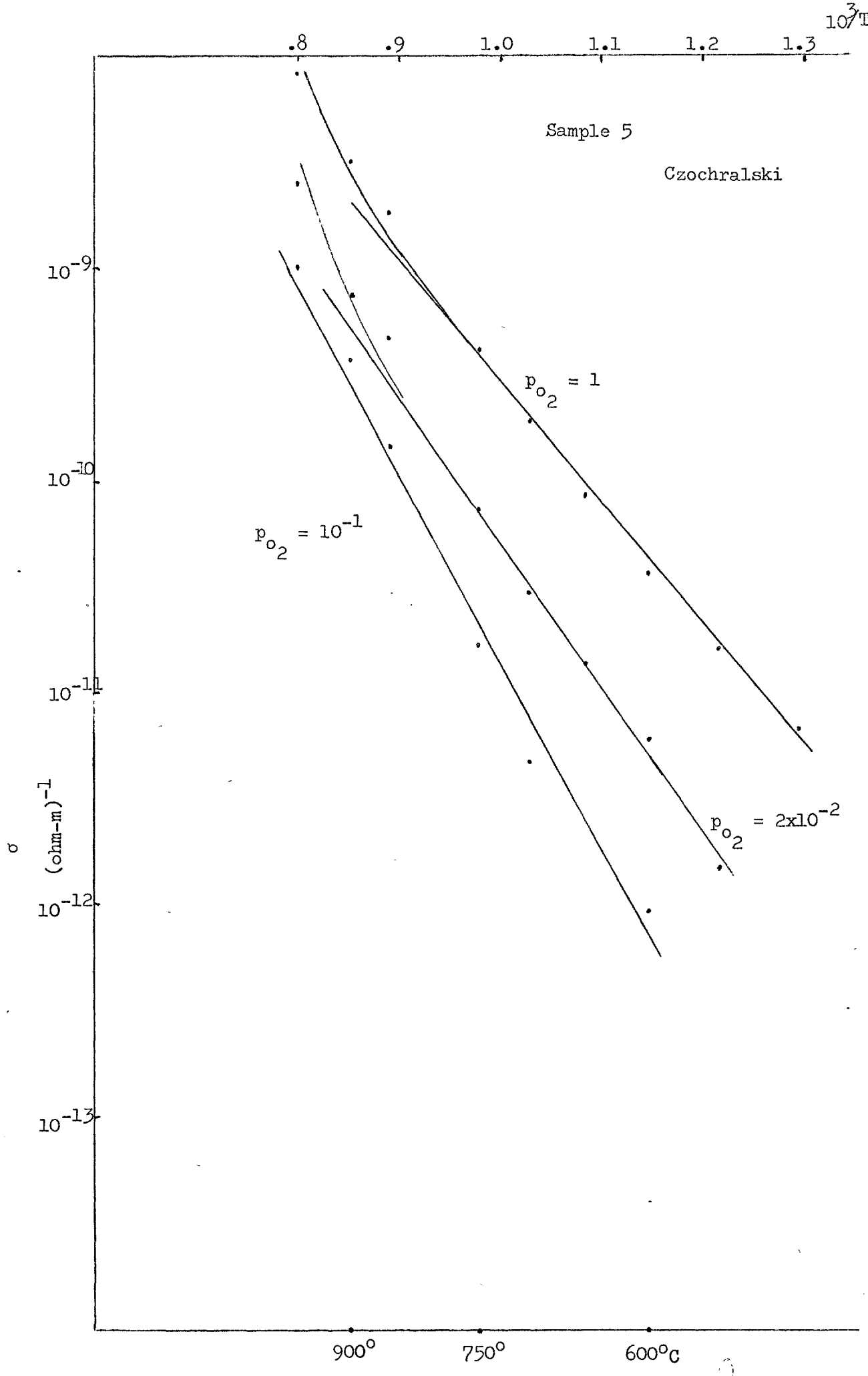


Figure 4.19 D.C. conductivity versus temperature at various p_{O_2}

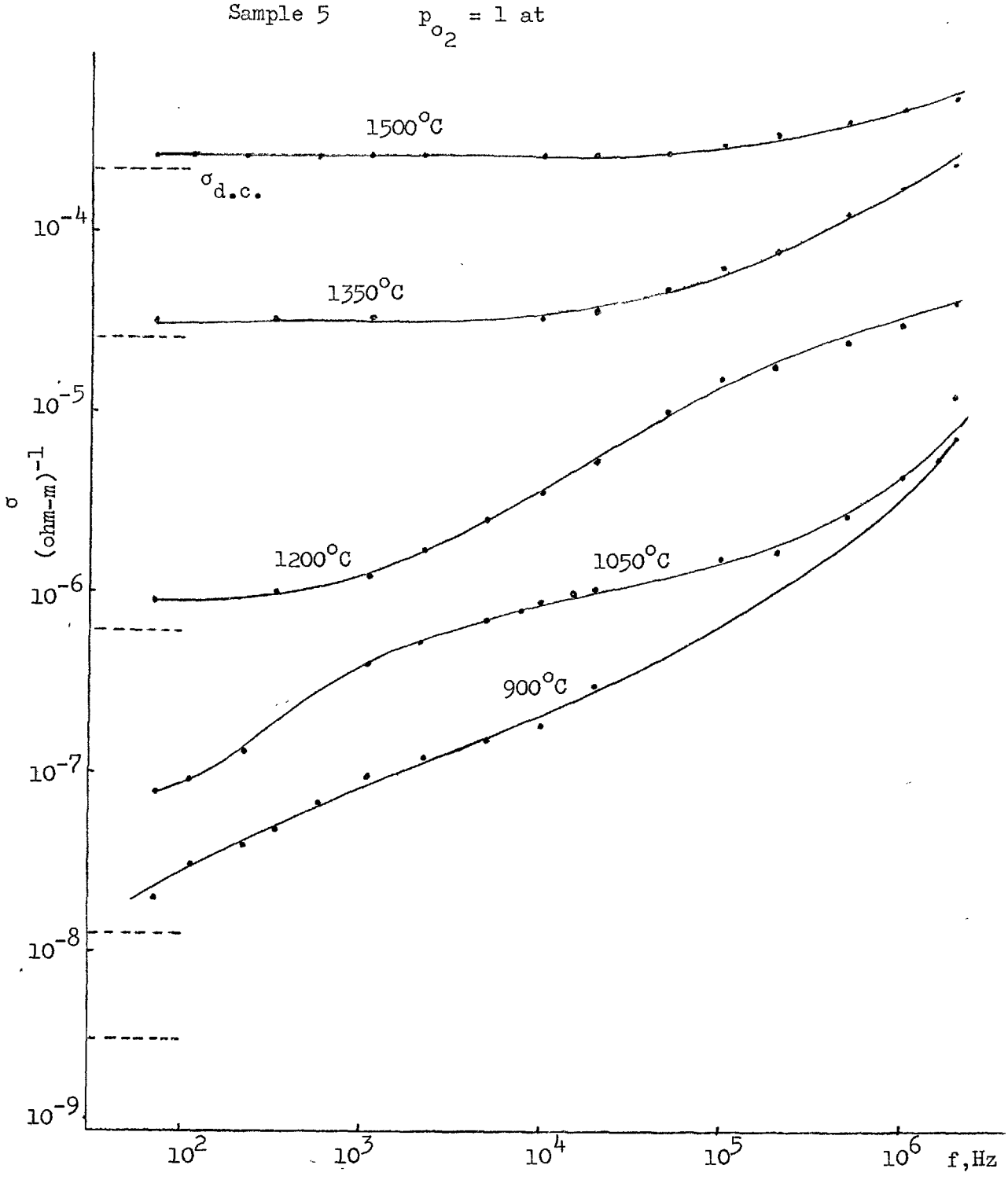


Figure 4.20 Conductivity against frequency at different temperatures

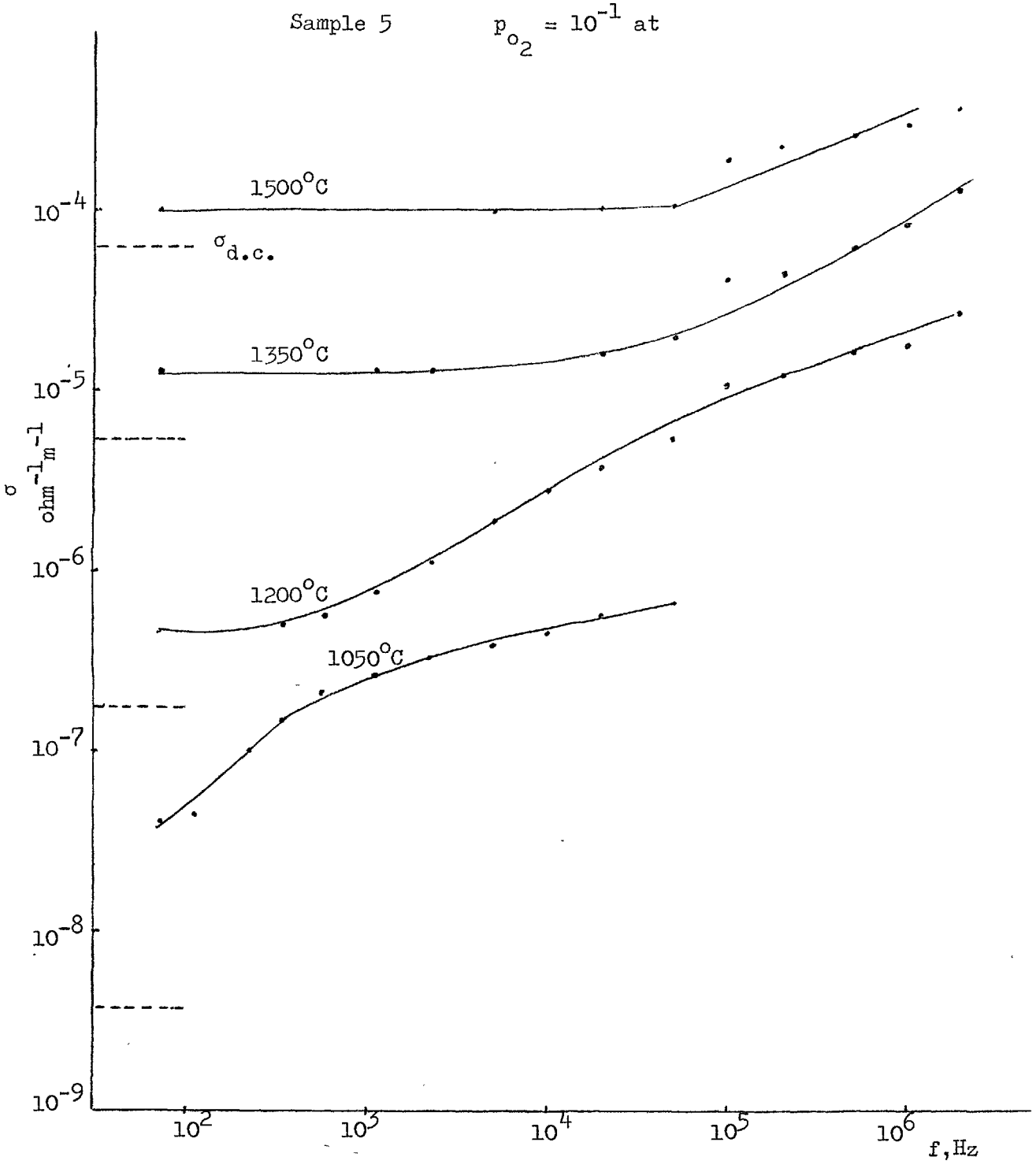


Figure 4.21 Conductivity against frequency at various temperatures

Sample 5 $p_{O_2} = 2 \times 10^{-2}$ at

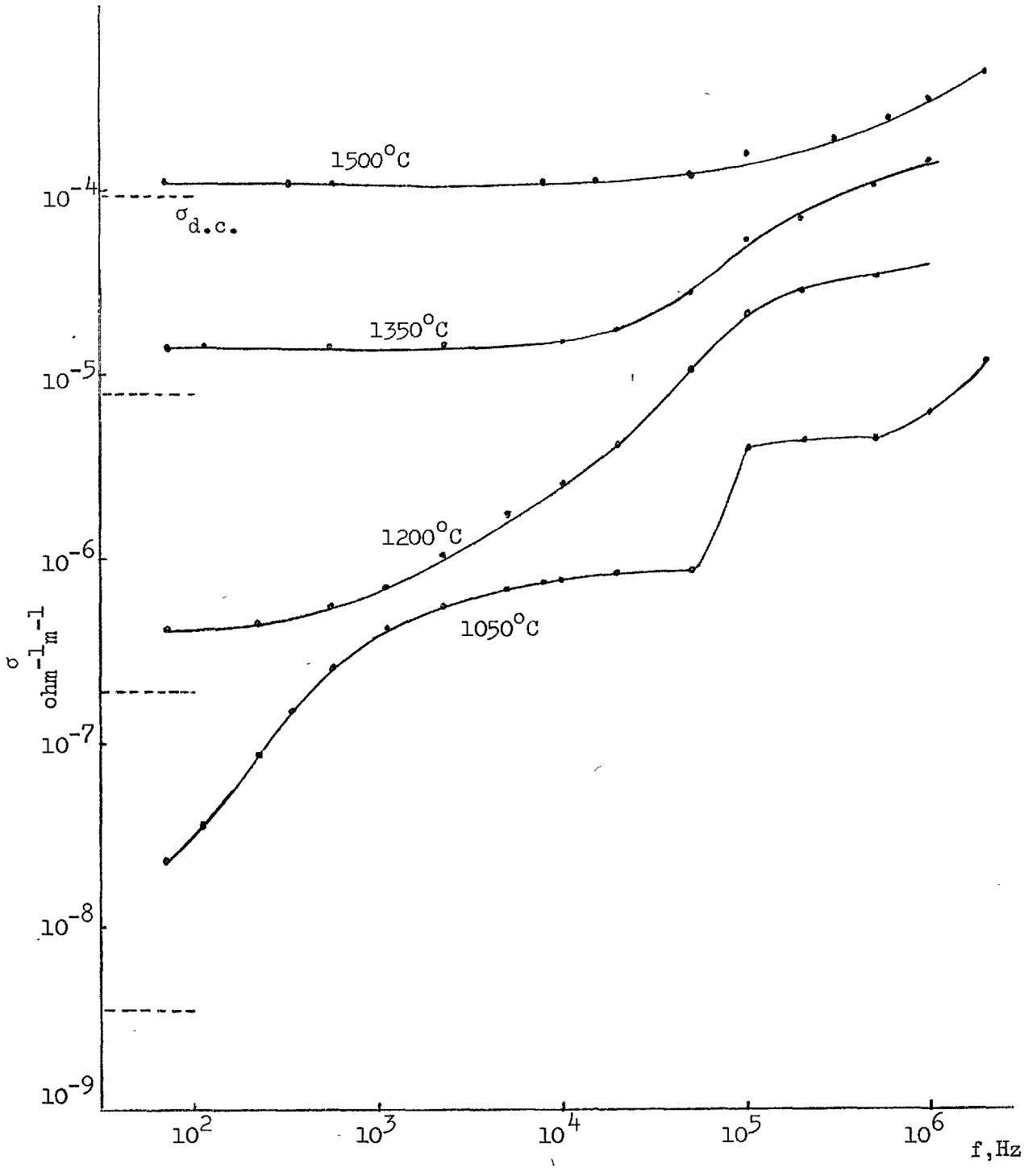


Figure 4.22 Conductivity versus frequency at various temperatures

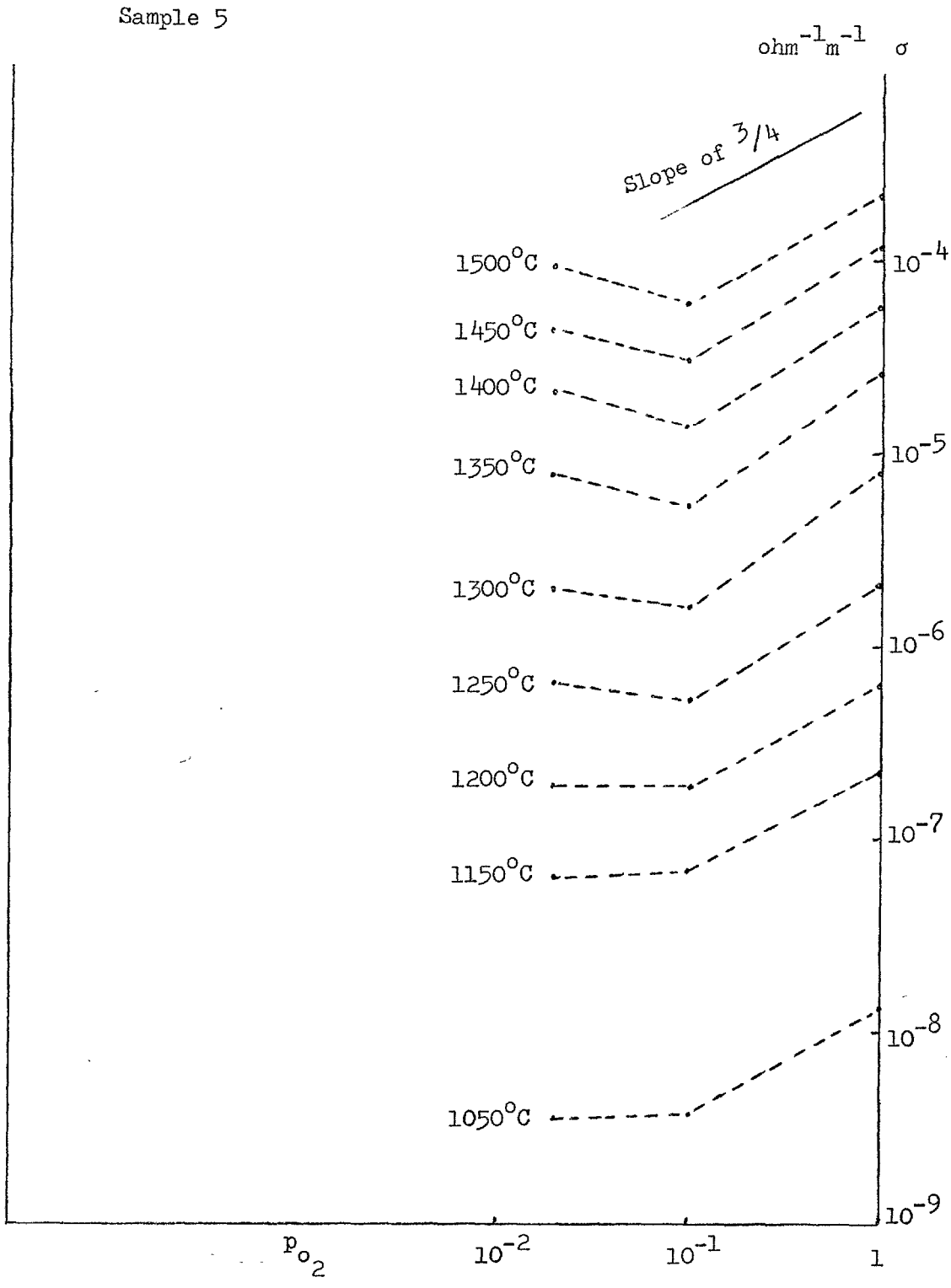


Figure 4.23 Effect of oxygen partial pressure on electrical conductivity of single crystal sapphire

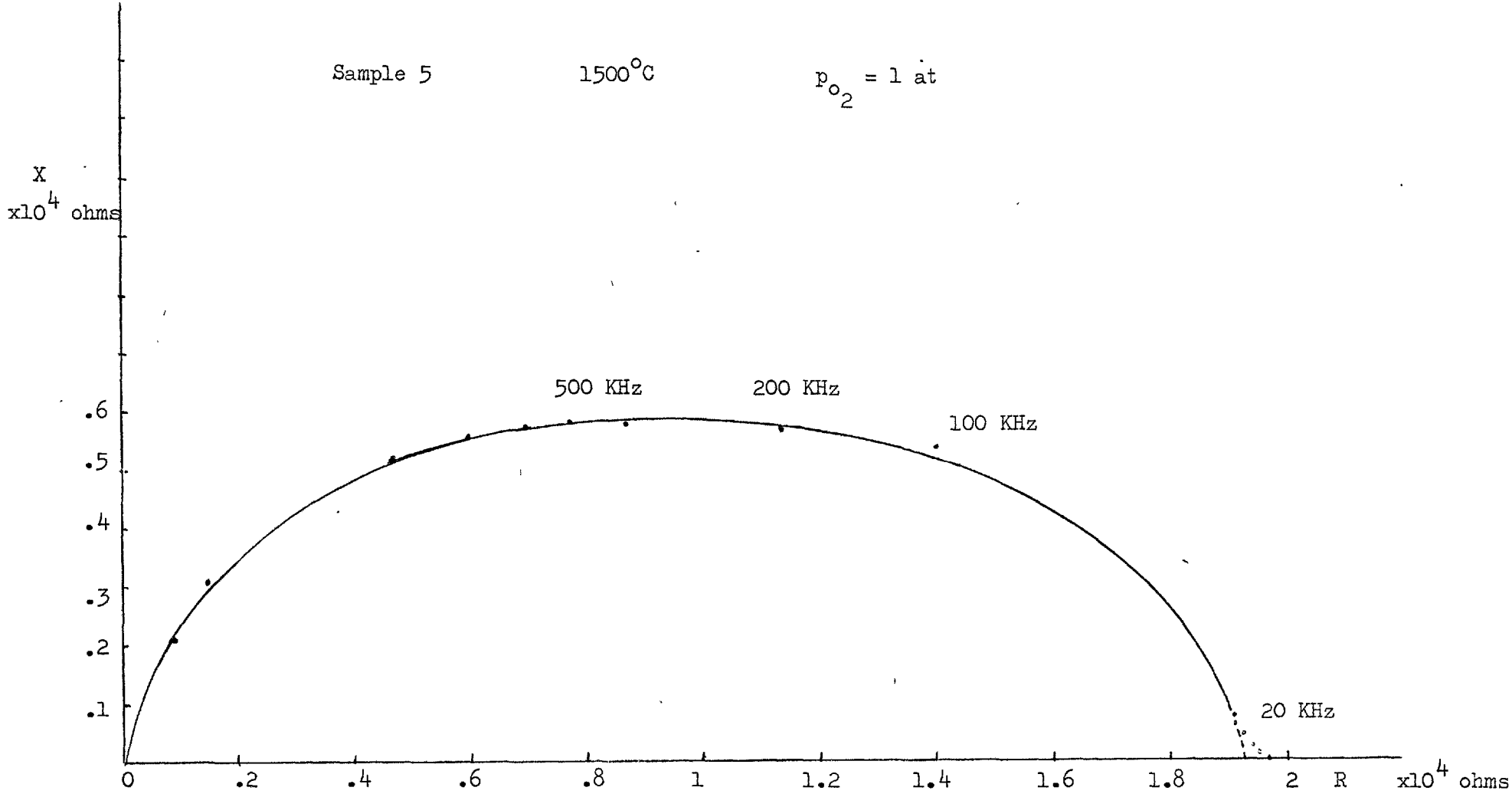


Figure 4.24 Reactance against resistance curve

extrapolated resistance value is the true value of the bulk. One has to make low frequency measurements to be able to prove this assumption.

Permittivity versus frequency curves have been plotted at various temperatures and oxygen partial pressure, as shown in Figs. 4.25, 4.26 and 4.27. Dispersion is larger at 1500°C and p_{O_2} = 1 at. and gets smaller as both temperature and oxygen partial pressure go down.

d.c. measurements generally took a long time (about 15 hours), since the current decayed from a high value to its steady state value upon application of 1 volt d.c. at a particular temperature.

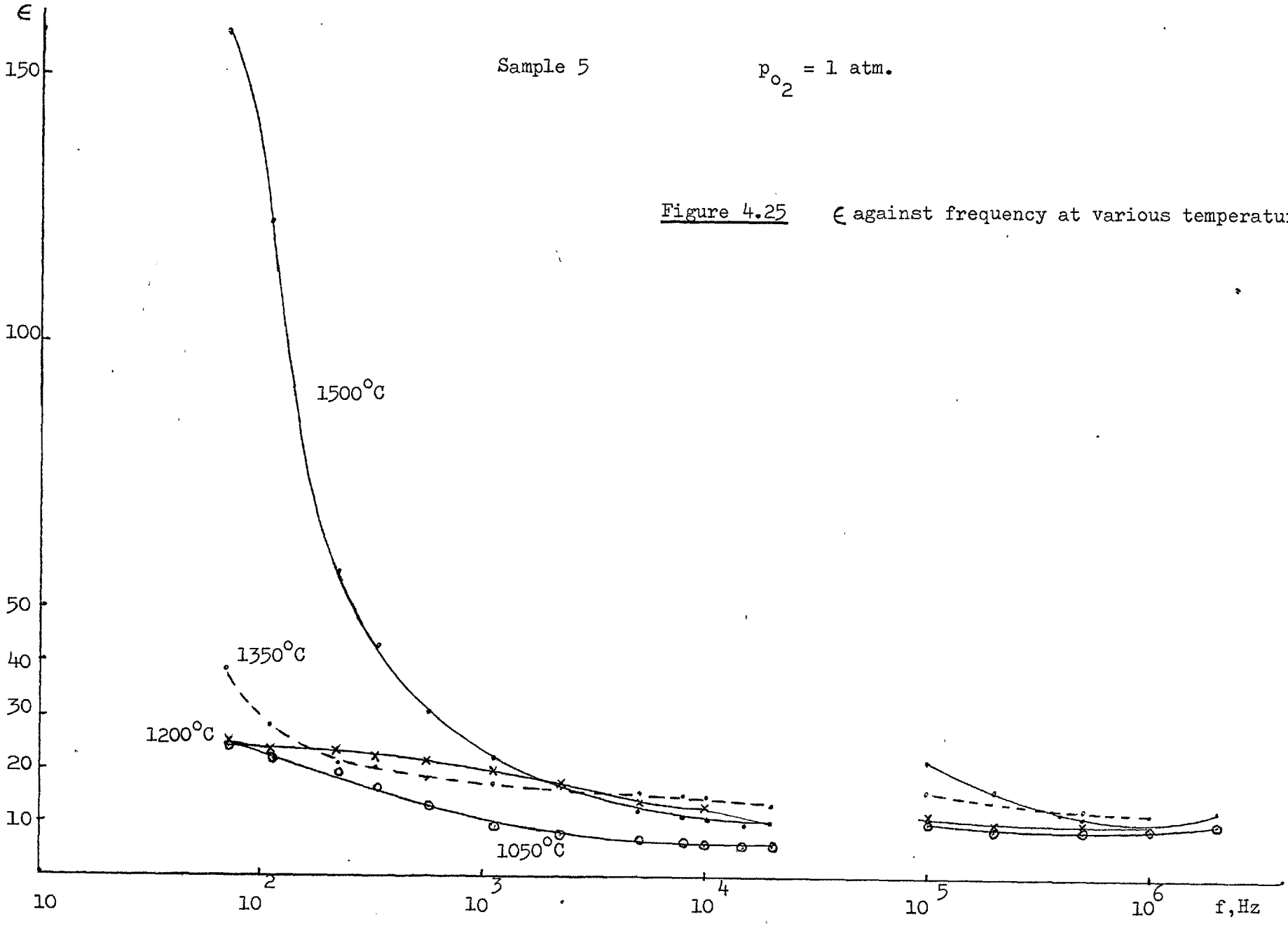


Figure 4.25 ϵ against frequency at various temperatures

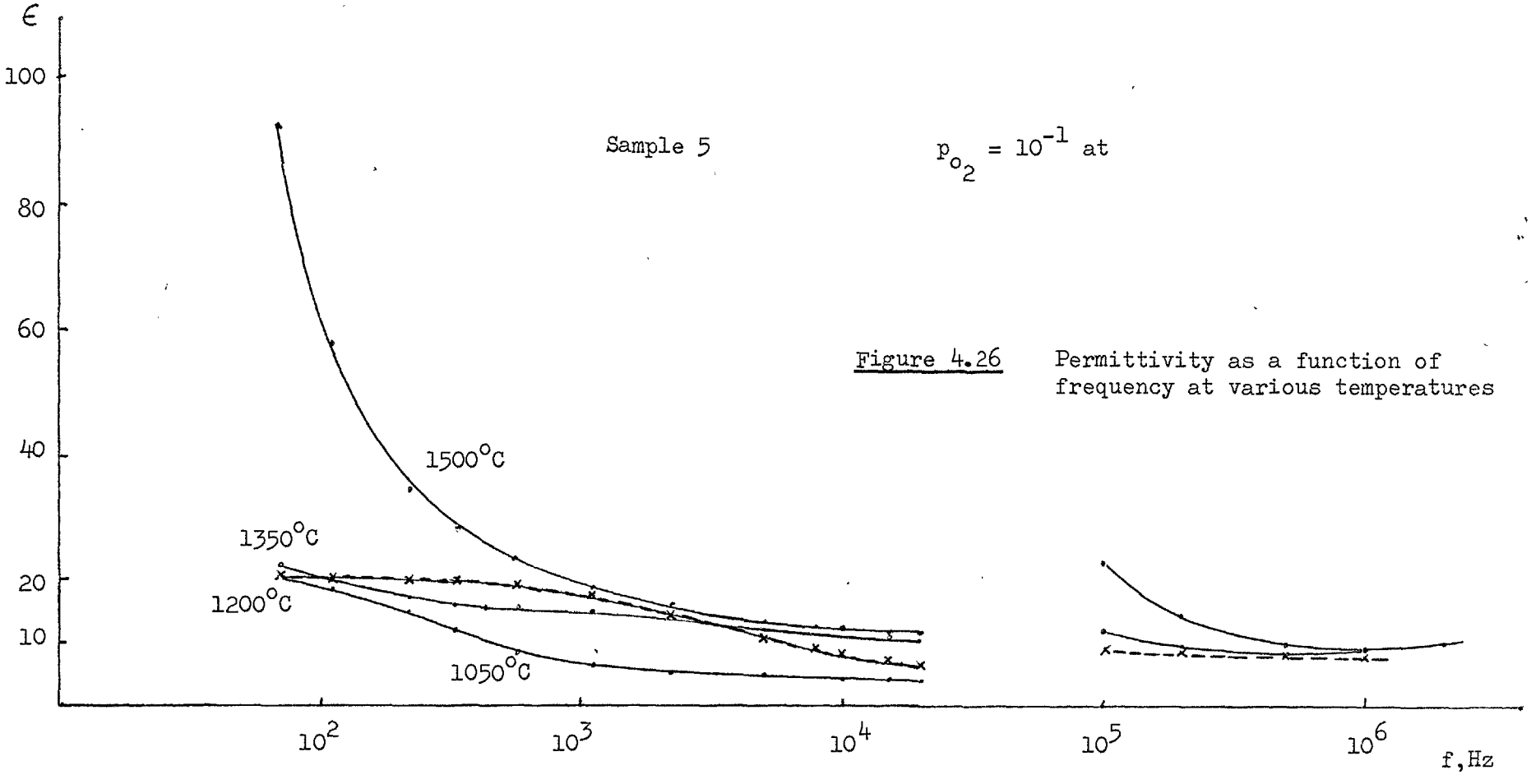


Figure 4.26 Permittivity as a function of frequency at various temperatures

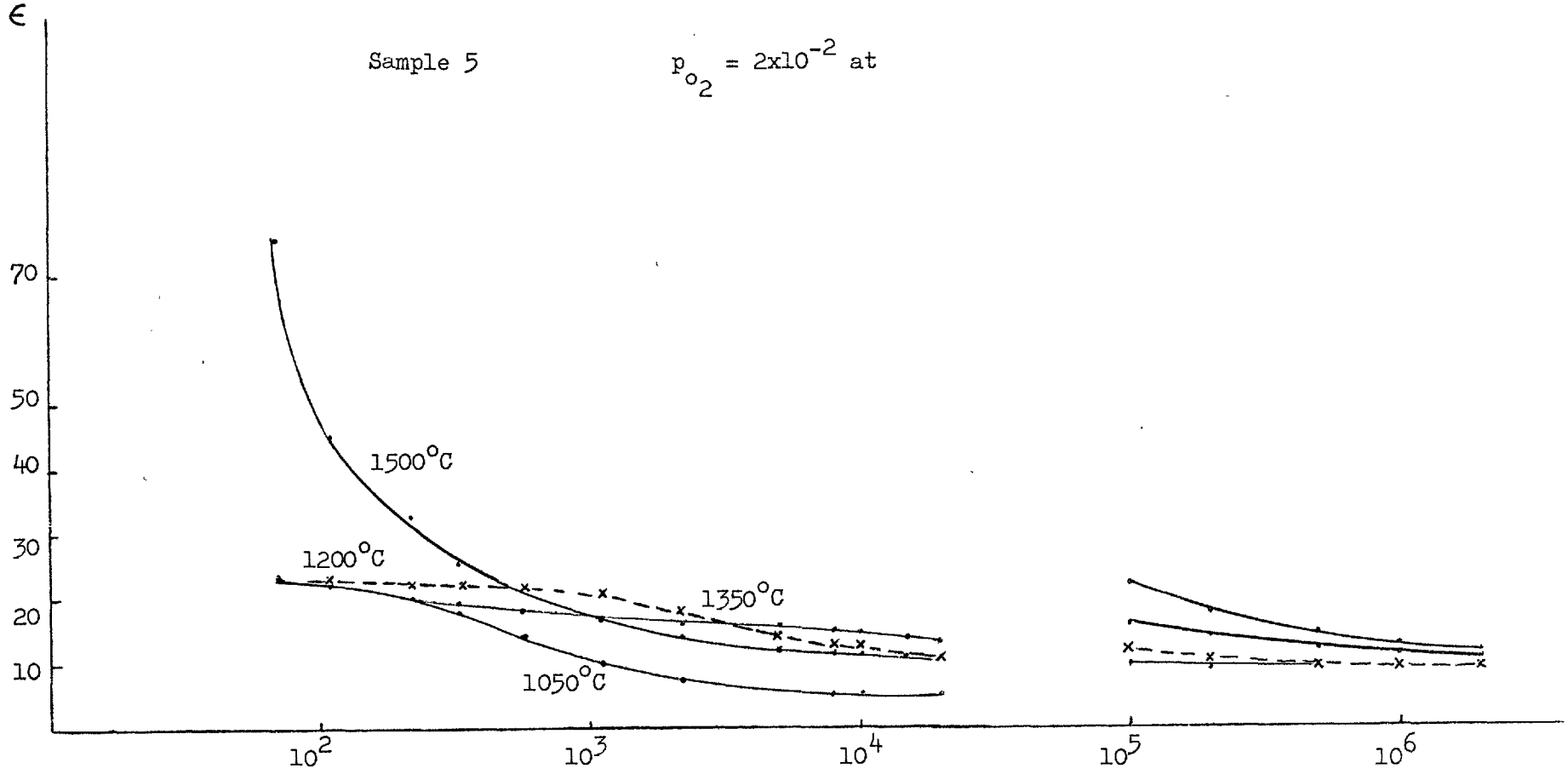


Figure 4.27 Permittivity against frequency at various temperatures

CHAPTER 5 - DISCUSSION

Previous work on Alumina was only barely outlined in the introduction, Chapter 1, since a discussion of the critical factors determining procedure and equipment design was not given until Chapter 3.

Hence in the first section of this chapter some general comments on previous work are made in the light of the discussion of Chapter 3 and the results of the present work presented in Chapter 4.

In the second section the d.c. conductivity dependence on oxygen pressure and temperature, which was observed, is compared in detail to previous results and interpretations.

In the third section the a.c. characteristics are discussed and an interpretation made in terms of an electrode effect and a bulk effect. The electrode effects are analysed in terms of a two layer model (Volger, 1960) and with the help of the Kramers-Kronig relations approximate values of the surface layer thickness and conductivity are determined and compared with previous results (Tallan and Graham, 1965) obtained at lower temperatures.

The bulk effect is analysed by the general theory described in Chapter 2 using particular parameters appropriate to Alumina.

5.1. Comparison to Previous Work

Pappis and Kingery (1961) carried out their conductivity measurements on single crystal Al_2O_3 over a wide range of temperature and oxygen partial pressures using two- and four-terminal d.c. and a.c. ($f = 10^4$ Hz only) techniques, which however do not prevent surface and gaseous conduction. Peters et al (1964) demonstrated that for high resistance materials

at high temperatures above 1100°C the conductivity of the gas phase around the sample can be comparable to or greater than that of sample. Moulson and Popper (1968) have also found gas phase conduction important at relatively low temperatures. Pappis and Kingery (1961) plotted log conductivity versus 1/T at various p_{O_2} and log conductivity versus log oxygen partial pressure at several temperatures. They also made thermoelectric electromotive force (Seebeck effect) measurements to determine the polarity of the charge carriers over a range of p_{O_2} but these again are affected by surface and gaseous carriers. Champion (1964) has investigated the conductivity of sapphire below 1000°C and employed two-terminal method, which is especially susceptible to surface effects. The same applies to the four-probe method of Dasgupta and Hart (1965). Harrop and Creamer (1963) attributed the conduction to electrons and holes; they have used the two terminal method.

Cohen (1959) performed the earliest review of the situation. Matsumura (1966) used no physical guard to eliminate gas phase conduction, and in addition he employed a voltage greater than the decompositions potential of sapphire. Özkan and Moulson (1970) made d.c. conductivity measurements by a three-terminal method with a physical gaseous conduction guard in dry air (p_{O_2} = 0.2 atm.), i.e. with an equipment on which we based our design. However they applied 85 volts across the sample and the non-linear behaviour of current voltage characteristic may be due to application of a voltage higher than the decomposition voltage. Their activation energy of 4.0 eV is comparable to our result of 4.39 eV made at $p_{O_2} = 10^{-1}$ atm. on Czochralski grown single crystals, and the conductivities were comparable as shown by Figure 4.17 where their results are plotted with ours.

Brook, Yee and Kröger (1971) made measurements on pure and doped flux and Czochralski grown crystals again using a three terminal

physically guarded d.c. method. They proposed a defect model for the oxide on the basis of a combination of their work and that of Pappis and Kingery on the oxygen pressure dependence of conduction in Al_2O_3. However, since the latter involves data taken by 2 or 4-terminal methods, which avoid contact potential but not surface conduction effects, confirmation of the detailed model would require further studies of the fully guarded type made on crystals equilibrated at different oxygen pressures. They concluded that the conduction in Al_2O_3 is substantially ionic with aluminum interstitial ions as charge carriers. A key factor in their argument is, however, the positive sign of the carriers deduced from thermoelectric measurements by Harrop and Creamer (1963) and Pappis and Kingery (1961). Both these authors seem confused about the relationship between the polarities of the hot and cold contacts and the carrier sign. The equipment used by Pappis and Kingery has already been criticised above. That of Harrop and Creamer appears to use a physical gas guard, but not an electrical guard ring.

Later Yee and Kröger (1973) conducted e.m.f. measurements on a concentration cell using a biassed guard to eliminate thermal e.m.f. effects, and with gas flow arrangements especially designed to prevent any partial pressure gradients across the crystal face, or any mixing of the gas streams leading to a false p_{O_2} value. They found that single crystal Al_2O_3 in O_2/air has an ionic transference number of 1 up to 1450°C; at higher temperatures t_i decreases slightly, reaching a value of 0.8 at 1650°C.

Tallan et al (1965) and (1963) carried out a.c. measurements over a temperature range and frequency range, and explained their results by an interfacial polarization mechanism, with a surface layer up to 50μ thick, having resistivities 5 orders of magnitude greater than bulk. T.P. Jones et al (1969) determined a defect diffusion coefficient

in single crystal aluminum oxide from 1400° to 1850°C by measuring the movement of the colour boundary in a Ti^{3+} doped single crystal.

Our experiments resulted in plots of the d.c. conductivity versus 1/T at constant oxygen partial pressure; the d.c. conductivity versus p_{O_2} at constant temperature; the conductivity and permittivity versus frequency for several combinations of temperature and partial pressure.

5.2. The Pressure and Temperature dependence of the d.c. conductivity

In general the d.c. conductivity versus 1/T plots, at constant p_{O_2}, are in broad agreement with previous work, at temperatures above 1050°C. The activation energy is insensitive to p_{O_2} over the range investigated (1 to 2×10^{-2}) suggesting that the number of mobile defects, but not the type, varies with p_{O_2} in this temperature range.

Below 1050°C activation energies become smaller, and depend on the p_{O_2}, changing between 0.8 and 1.1 eV for the flux grown crystals. Hence there exists a region with lower activation energy between 900°C and 500°C. The corresponding low temperature region activation energy for Czochralski crystals is 1.5 eV which is relatively higher than that of flux grown crystals.

The high temperature region activation energies and conductivity values of flux grown platelets, are comparable to those of Brook et al (1971) and those of Czochralski grown crystals are comparable to those of Özkan and Moulson (1970).

There is nothing in these results to positively contradict their interpretations. Brook et al proposed that the conduction in Al_2O_3 is due to aluminum interstitial ions at high oxygen pressure. Özkan and Moulson tried to interpret their results in terms of electronic conduction in bands, which now seems inappropriate in view of the work of Yee and Kröger (1973).

The $\sigma_{d.c.}$ versus p_{O_2} isotherms are rather incomplete even for Sample 5. However, the sort of behaviour observed could be explained by the following sort of model;

In terms of real charges:

$$\frac{1}{2} O_2(g) + 2\ Fe_{Al}^{++} \longrightarrow 2\ Fe_{Al}^{+++} + O_O^{=} + \frac{2}{3} V_{Al}$$

or the same equation in terms of effective charges may be written as follows,

$$\frac{1}{2} O_2(g) + 2\ Fe' \longrightarrow 2\ Fe^{x} + O_O + \frac{2}{3} V_{Al}'''$$

where the equilibrium constant is

$$K_i = \frac{[Fe^x]^2 [O_O] [V_{Al}'''] ^{2/3}}{p_{O_2}^{1/2} [Fe']^2}$$

The charge neutrality condition is

$$[Fe_{Al}'] = [Si_{Al}^{\bullet}]$$

Since $[O_O]$, $[Fe^x]$ and $[Fe']$ are constant, therefore

$$[V_{Al}'''] = K_i'\ p_{O_2}^{3/4} \quad \text{and} \quad \sigma_{V_{Al}} = q[V_{Al}''']\mu[V_{Al}'''] \propto p_{O_2}^{3/4}$$

this predicts a variation of log σ on log p_{O_2} broadly as observed in Fig. 4.23, but the absolute values depend on both the density and the

mobility of V_{Al}'''.

5.3. Oxygen partial pressure and temperature dependence of the a.c. conductivity

The general features of the results described in the previous chapter lead us to an interpretation in terms of a bulk effect and a contact effect. In general we attribute the slower rise in σ' at higher frequencies with little change in ϵ', to the bulk processes. The difference between σ' at 70 Hz at high temperatures and $\sigma_{d.c.}$ and the considerable variation in ϵ' at low frequencies we attribute to the electrode effects. At high temperatures there is a flat region of σ' as previously remarked which suggests that the effect of the surface on σ' is complete once f rises above about 70 Hz. On the other hand, as f decreases, the rise in ϵ' by 70 Hz becomes larger as temperature increases. This suggests that the characteristic frequency of the surface process rises with temperature, but its character is such that the total change it causes from $\sigma_{d.c.}$ to some intermediate value σ_m decreases with temperature. Even without a detailed model of the surface, which we have not constructed, it is possible to discuss the consistency of the variation in $\sigma'(\omega)$ and $\epsilon'(\omega)$ in this region by the Kramers-Kronig equations, and this is done in the next sub-section.

In the sub-section following that, the variations of $\sigma'(\omega)$ are discussed in terms of the general theory of Chapter 2 in relation to the particular parameters one might expect for Al_2O_3.

5.3.1. Electrode (Low Frequency) Process

Since the material has a lattice contribution χ'_∞ to $\chi^*(\omega)$ then the Kramer-Kronig relation equations (2.130) and (2.131) must be interpreted as relating the increase in χ' above χ'_∞ to χ'', and χ'' to the increase in χ' above χ'_∞, i.e. equations (2.130) and (2.131) are replaced by

$$\chi'(\omega) - \chi'_{\infty} = \frac{2}{\pi} \int_0^{\infty} \chi''(\mu) \frac{\mu}{\mu^2 - \omega^2} \, d\mu \tag{5.1}$$

and

$$\chi''(\omega) = \frac{2}{\pi} \int_0^{\infty} \frac{(\chi'(\mu) - \chi'_{\infty})\omega}{\omega^2 - \mu^2} \, d\mu \tag{5.2}$$

but

$$\sigma' = \omega \epsilon_o \chi'' = \sigma_{tot} - \sigma_{d.c.}$$

Therefore multiplying both sides of eqn. (5.2) by $\omega\epsilon_o$ yields

$$\sigma_{total} - \sigma_{d.c.} = \frac{2\epsilon_o}{\pi} \int_0^{\infty} \frac{(\chi' - \chi'_{\infty})}{\omega^2 - \mu^2} \omega^2 \, d\mu$$

The integrand is the product of $\chi' - \chi'_{\infty}$ and $\frac{\omega^2}{\omega^2 - \mu^2}$, and for a high enough value of ω, $\chi' - \chi'_{\infty}$ goes to zero before μ approaches ω so that the product is finite. Let $\omega \longrightarrow \infty$ i.e. much greater than frequency at which the low frequency dispersion in σ is completed; and call the conductivity at this frequency σ_m, then

$$\sigma_m - \sigma_{d.c.} = \frac{2\epsilon_o}{\pi} \int_0^{\infty} (\chi' - \chi'_{\infty}) \, d\mu$$

or

$$\sigma_m - \sigma_{d.c.} = \frac{2\epsilon_o}{\pi} \int_0^{\infty} (\epsilon' - \epsilon'_{\infty}) \, d\mu \tag{5.3}$$

The integral is the area under the $\epsilon' - \epsilon'_{\infty}$ versus frequency curve, hence designating the area by $A(\omega)$ we get

$$\sigma_m - \sigma_{d.c.} = \frac{2\epsilon_o}{\pi} A(\omega)$$

or since $\omega = 2\pi f$, then

$$A(f) = \frac{1}{4\epsilon_o} (\sigma_m - \sigma_{d.c.}) \qquad (5.4)$$

The area which should be found beneath the curve if measurements could be continued to lower frequencies, may be obtained by substituting the values of ϵ_o, σ_m and $\sigma_{d.c.}$ in this equation. Knowing this area, a value of ϵ_s may be estimated by linearly extrapolating the $\epsilon' - \epsilon'_\infty$ curve towards lower frequencies so as to give an area A beneath the curve by the time f = 1 Hz is reached. Assuming ϵ_s is not much greater than this value at 1 Hz, the rest of the curve will make a negligible contribution to A. In all cases the extrapolated curve had a slope smaller in magnitude than 2, which is required for physical consistency as explained below.

Equation (5.4) has been used to obtain static permittivity at 1500°C and 1350°C and at oxygen partial pressures of 1, 10^{-1} and 2×10^{-2} atm. An example of extrapolation is shown for T = 1500°C and p_{O_2} = 1 atm. in Fig. 5.1. The static permittivity is highest at p_{O_2} = 1 atm. at both temperatures and reduces with p_{O_2}.

Volger (1960) has reviewed the modelling of interfacial dispersion by layers of different permittivity and conductivity, and gives specific equations for the ratios of ϵ' and σ' above and below the resulting dispersion in terms of the ϵ'_i and σ'_i (i = 1,2) of the two layer model. As is well known, this two layer model predicts a Debye-type dispersion with σ' and ϵ' having slopes of 2 at their steepest parts which is the steepest slope which such a model can explain. It is seen from the application of the Kramer-Kronig relations just discussed that the value of $\sigma_m - \sigma_{d.c.}$ (where σ_m refers, for the low frequency dispersion at high

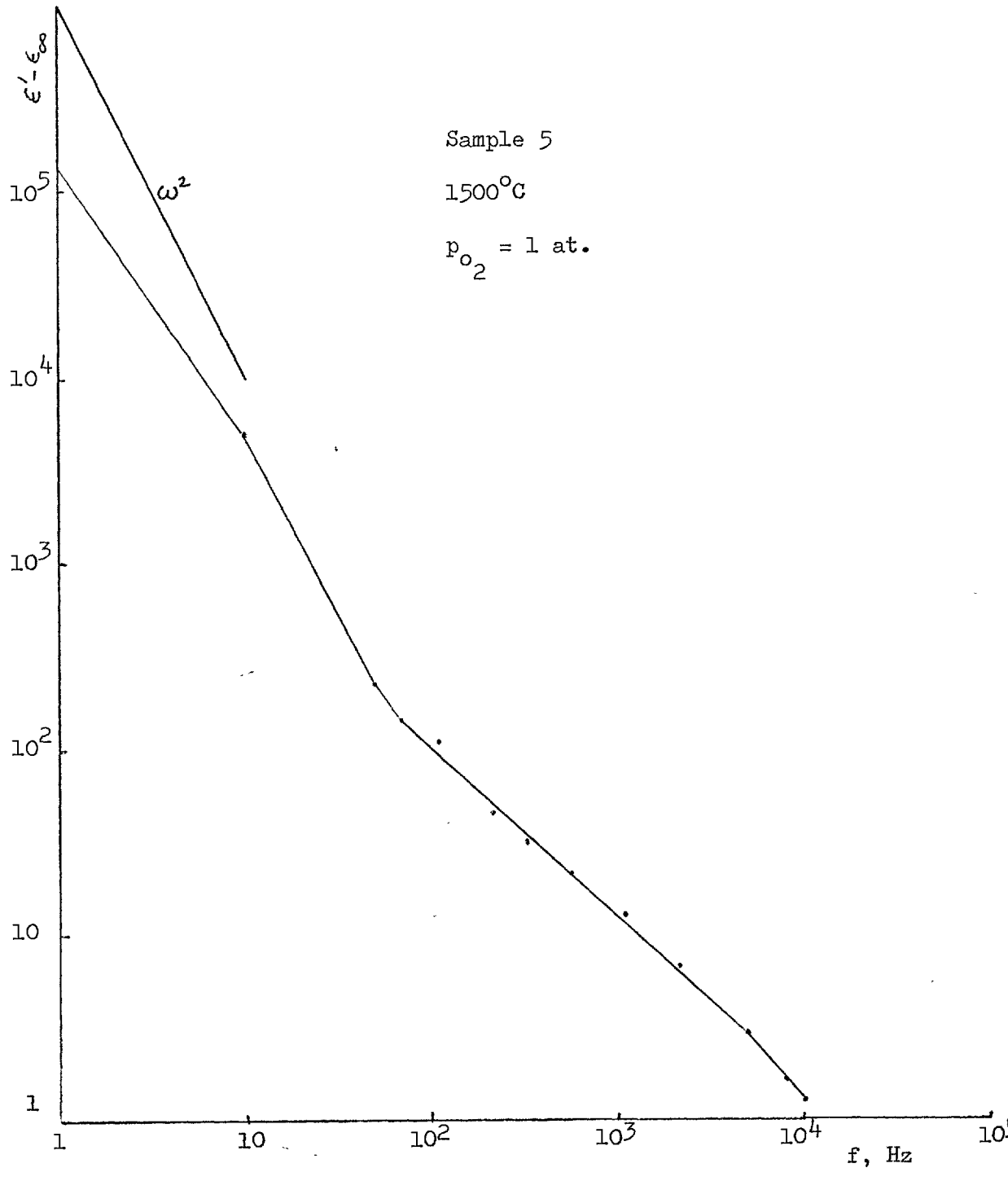


Fig. 5.1 $\epsilon' - \epsilon_\infty$ versus frequency.

temperatures to approximately the 70 Hz value) do not require as steep a slope as 2 to explain them, so that it is within the bounds of feasibility. This also means that they cannot be explained exactly in terms of the 2-layer model, but that a multiple-layer model or graded variation in ϵ' and σ' would have to be used. Assuming, however, that the two layer model applies in a qualitative fashion, it is seen from Volger's equations for the two-layer case that

$$\sigma_s = \frac{d_1 + d_2}{\frac{d_1}{\sigma_1} + \frac{d_2}{\sigma_2}} \tag{5.5}$$

and

$$\sigma_m = \frac{(d_1 + d_2)\ \frac{d_1\sigma_1}{\epsilon_1^2} + \frac{d_2\sigma_2}{\epsilon_2^2}}{\left(\frac{d_1}{\epsilon_1} + \frac{d_2}{\epsilon_2}\right)^2} \tag{5.6}$$

Assuming the blocking layer has an insignificant resistance ($d_1\rho_1 \ll d_2\rho_2$) but that its resistivity is so high that $d_1\rho_1^2 \gg d_2\rho_2^2$ then Volgers result for ϵ_s in this case reduces to

$$\frac{\epsilon_s}{\epsilon} \simeq \frac{d_1\rho_1^2}{d_2\rho_2^2} \tag{5.7}$$

when it is assumed that $\epsilon_1 = \epsilon_2 = \epsilon$.

Combination of equations (5.5), (5.6) and (5.7) and making use of the inequalities to neglect small terms yields

$$4\left(\frac{d_1}{d}\right)^2 + \frac{\epsilon_s}{\epsilon_\infty}\ \frac{d_1}{d} + \left(\frac{\sigma_m}{\sigma_s} - 1\right)^2 = 0$$

where d is $d_1 + d_2$, then, since $\epsilon_s/\epsilon_\infty \gg 1$, one finds

$$\frac{d_1}{d} \cong \frac{(\frac{\sigma_m}{\sigma_s} - 1)^2}{\epsilon_s/\epsilon_\infty} \tag{5.8}$$

Eqn. (5.7) may be rewritten in the following form

$$\frac{\sigma_2}{\sigma_1} \cong \sqrt{\frac{d}{d_1}\frac{\epsilon_s}{\epsilon}} \tag{5.9}$$

On the other hand eqn. (5.6) may be simplified as

$$\sigma_2 = \frac{\sigma_m}{1 + \frac{d_1}{d}\frac{\sigma_1}{\sigma_2}} \tag{5.10}$$

by making use of $d_2 \cong d$.

Finally eqns. (5.8), (5.9) and (5.10) have been used to calculate d_1, σ_2/σ_1, σ_1 and σ_2 at 1500°C and 1350°C at oxygen partial pressures of 1, 10^{-1} and 2×10^{-2} atm. and the results are tabulated in Table 5.1. One observe that the thickness of the thin layer d_1 increased from an average of about 100 Å at 1500°C to an average of about 5,000 Å at 1350°C, i.e. by a factor of 50. Tallan and Graham (1965) computed a thickness of surface layer of about 40μ at a temperature of 1034°C, which seems in reasonable agreement, if one considers the change in the layer thickness as the temperature goes down.

5.3.2. Bulk (High Frequency) Process

In Chapter 2 it was shown that a multiple well-pair model could predict a log σ - log f characteristic of lower slope than 2, and calculated curves of log $\sigma'/\sigma_{d.c.}$ and log$(\sigma_{tot}/\sigma_{d.c.})$ versus log Ω were plotted in Fig. 2.25 for the 6 well-pair case with $\sigma_\infty/\sigma_{d.c.} = 20$, which necessitated $\Delta W_{12} = 0.152$ eV.

Table 5.1

Temperature	1500°C			1350°C		
p_{O_2}, atm.	1	10^{-1}	2×10^{-2}	1	10^{-1}	2×10^{-2}
σ_s $(ohm\text{-}m)^{-1}$	2.21×10^{-4}	6.21×10^{-5}	9.3×10^{-5}	2.65×10^{-5}	5.43×10^{-6}	7.91×10^{-6}
σ_m $(ohm\text{-}m)^{-1}$	2.62×10^{-4}	1×10^{-4}	1.13×10^{-4}	3.33×10^{-5}	1.3×10^{-5}	1.43×10^{-5}
ϵ_s F/m	1.3×10^{5}	4×10^{4}	2×10^{4}	1.27×10^{3}	7×10^{3}	3×10^{3}
d	260 μ	260 μ	260 μ	260 μ	260 μ	260 μ
d_1	6.88 Å	242 Å	60 Å	0.135 μ	0.72 μ	0.555 μ
σ_2/σ_1	7×10^{4}	6.6×10^{3}	8.66×10^{3}	4.95×10^{2}	5.02×10^{2}	3.71×10^{2}
σ_1 $(ohm\text{-}m)^{-1}$	3.74×10^{-9}	1.5×10^{-8}	1.3×10^{-8}	6.72×10^{-8}	2.59×10^{-8}	3.85×10^{-8}
σ_2 $(ohm\text{-}m)^{-1}$	2.62×10^{-4}	10^{-4}	1.13×10^{-4}	3.33×10^{-5}	1.3×10^{-5}	1.43×10^{-5}

The experimental plots of log σ vs log f show a general similarity to Fig. 2.25, although the slope is in general even less. However, it seems probable that the slope would reduce even more for higher numbers of well-pairs.

If the results of the 6 well-pairs K = 20 model are assumed to be appropriate in order of magnitude to describe the experimental results, then we can try to test the applicability of the theory to alumina.

The model predicts (see Fig. 2.25) that $\sigma' = \sigma_{d.c.}$ (or rather σ_m in view of our identification of σ_m to the bulk conductivity in the last subsection) at $\Omega_2 = 0.22$, or that $\sigma_{tot} = 2\sigma_m$ at this value of Ω. Hence the experimental curves should pass through $\sigma_{tot} = 2\sigma_m$ at a frequency

$$f_2 = \frac{\omega_2}{2\pi} = \frac{\Omega_2 p_{oo}}{2\pi} = \frac{0.22 p_{oo}}{2\pi} \qquad (5.11)$$

Hence to find f_2 we need to know or estimate p_{oo}. One way to do this is to try to find p_{oo} from σ_m. According to Eqn. (2.233) we have for m = 6;

$$p_{oo} = \frac{\sigma_m kT}{N_d z^2 e^2 a^2 f_7(0)} \left(\frac{N_d}{N_s}\right)^{1/3} \qquad (5.12)$$

In order to find p_{oo} from the measured σ_m one has therefore to find z, a, N_d, N_s and $f_7(0)$. In fact $f_7(0)$ may be found from Eqn. (2.167) once N_d and N_s are known, using the appropriate value of $\Delta W_{12} = 0.152$ eV for the 6 well-pair case. We shall assume tentatively z = 3, the normal Al ion charge.

The values of a and N_s are determined by the crystal lattice of Al_2O_3. Fig. 5.2 shows a unit cell of Al_2O_3, with the O^{2-} ions omitted for clarity, according to Pauling and Hendricks (1925). The position

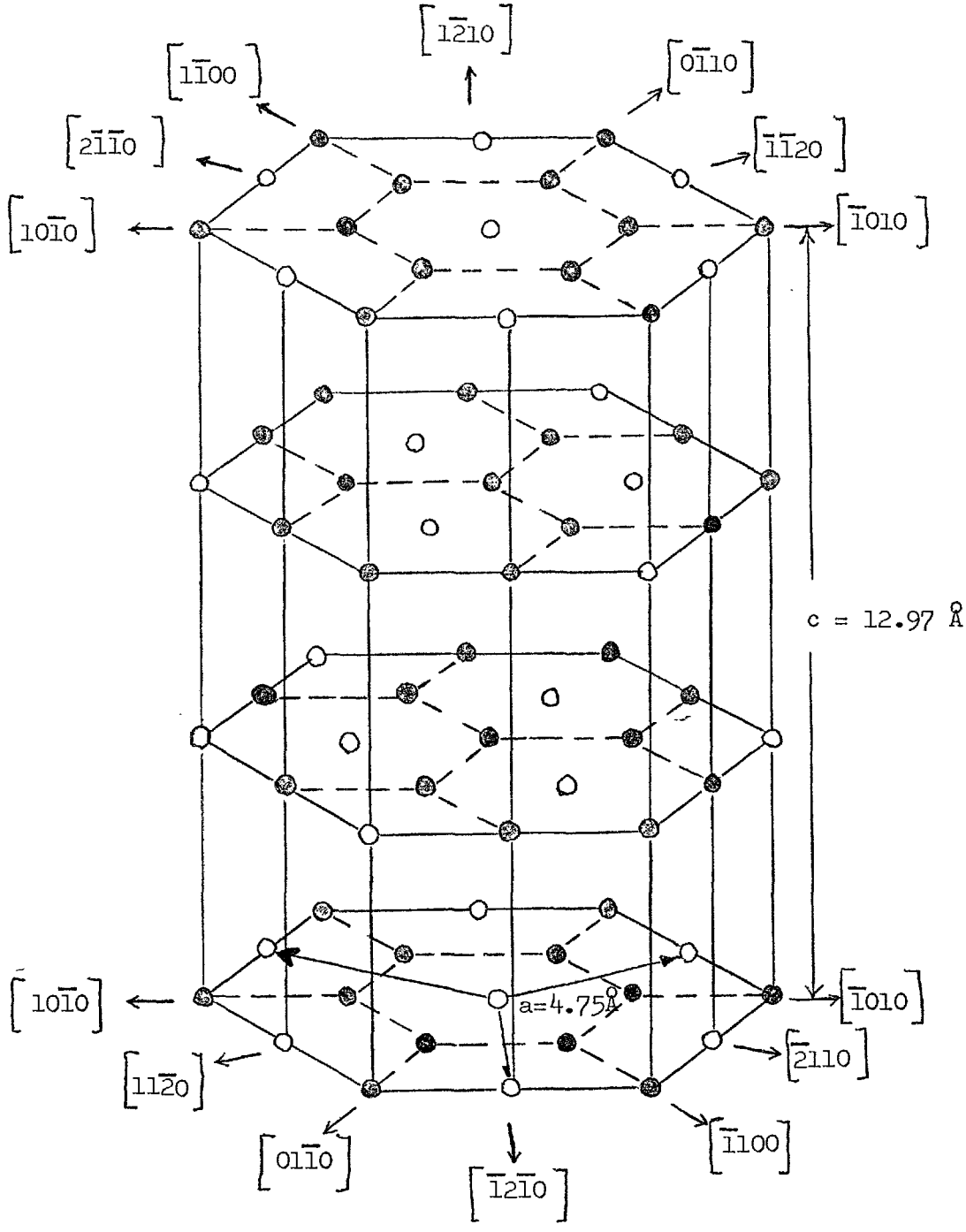


Figure 5.2. The arrangement of aluminium ions and holes in the hexagonal unit cell in sapphire

of holes or interstices which could also accommodate Al^{3+} ions are also marked. From the figure, the unit cell volume V is

$$V = 6\ a^2\ c\ \tan 30^\circ = 5.06\times10^{-28}\ m^3$$

Hence the number of unit cells per m^3 is

$$N_{cells} = {}^1/V = 1.97\times10^{27}\ .\ m^{-3}$$

Again from the figure, the unit cell contains 24 Al^{3+} ions and 12 holes. Once a defect, e.g. V'''_{Al} , has been created, it is difficult to know if one has to consider only 24 sites available to it or all 36. This would depend on a calculation of the relative free energies of the configurations of the type described in section 2.1.3. Again, for a tentative calculation, 36 sites per cell will be taken. Hence

$$N_s = 36\ N_{cells} = 7.09\times10^{28}\ m^{-3} \qquad (5.13)$$

The jump distance a should not be confused with the a-dimension of the unit cell. Since the crystals were measured along the c-axis, then from Fig. 5.2 again we take

$$a = c/3 = 2.16\ \text{Å}$$

A value of N_d may be found from σ_m and the diffusion coefficient D of the defect responsible for σ_m by means of the Nernst-Einstein relationship

$$\mu = \frac{z\ e\ D}{kT} \qquad (5.14)$$

and the conductivity - mobility relationship

$$\sigma = N_d \, ze\mu \tag{5.15}$$

which combine to give

$$N_d = \frac{\sigma kT}{D z^2 e^2} = 5.990 \times 10^{13} \frac{\sigma_m T}{D} \tag{5.16}$$

assuming σ_m is the appropriate value of σ, as explained above.

Values of D for Al in Al_2O_3 measured by tracer techniques have been given by Paladino and Kingery (1962). However, if they are used in Eqn. (5.16) in conjunction with our results for σ_m the resultant values for N_d are greater than N_s.

Jones, Coble and Mogab (1969) point out that the tracer diffusion coefficient actually measures a weighted sum of the diffusion coefficients of the individual defects contributing to the tracer diffusion. By observing the movement of the colour boundary in a Ti doped Al_2O_3 crystal they believe they have measured D directly for an individual, but unidentified defect. The values obtained are far higher than those of Paladino and Kingery, in our temperature range.

If we tentatively assume this is the defect responsible for conductivity, then Eqn. (5.16) gives $N_d = 9.28 \times 10^{24}$ at 1500°C and $N_d = 5.4 \times 10^{24}$ at 1350°C at p_{O_2} = 1 atm. and the same order of magnitude in the other p_{O_2}.

Using all the above factors in Eqn. (5.12) at T = 1500°C, p_{O_2} = 1 atm. gives the result

$$p_{oo} = 4.62 \times 10^9$$

and using Eqn. (5.11) then

$$f_2 = 161 \text{ MHz}$$

This compares with an experimental result of

$$f_2 \simeq 1 \text{ MHz}$$

which is a disappointingly poor agreement.

Apart from the various assumptions made above, finding the parameters appropriate to Alumina to use in Eqn. (5.12), the whole treatment may of course be criticised on account of its one-dimensionality. This means that every defect which travels right through the solid has to go down into the deeper potential wells at some stage, whereas in a three-dimensional crystal there will be paths through avoiding the deep wells. In addition, there will be a greater proportion of "undepressed" sites for a given N_d/N_s and value of m (no. of well-pairs) in a three-dimensional treatment than in a one-dimensional treatment.

To test this idea a very rough three-dimensional calculation has been made to relate $\sigma_{d.c.}$ to p_{oo}. It is based on the idea that the conductivity will be given by

$$\sigma_{d.c.} = \frac{n_7 z^2 e^2 a^2}{kT} p_{oo} \qquad (5.17)$$

where n_7 is the number of defects in the "undepressed" wells, i.e. those defects at a greater distance than 6 lattice sites from an immobile defect. Evidently

$$n_7 = N_7 \, f_7(0)$$

where N_7 is the density of such sites and $f_7(0)$ is the Maxwell-Boltzmann probability factor that such a site be occupied. If C is the number of nearest neighbours, then it is shown in appendix 2 by a rough calculation that

$$N_7 = N_s - C\,(1 + 2^2 + \cdots + 6^2)\, N_d \tag{5.18}$$

and

$$f_7^{-1} = C\left\{\left[\exp(6\,W_{12}) - 1\right] + 2^2\left[\exp(5\,W_{12}) - 1\right] + \cdots\cdots + \cdots + 6^2\left[\exp(W_{12}) - 1\right]\right\} + \frac{N_s}{N_d} \tag{5.19}$$

Combining Eqns. (5.17), (5.18) and (5.19) to give an equation for p_{oo}, one finds, using the same values of σ_m, N_d, N_s, z, a and ΔW_{12} as previously that the results for p_{oo} are reduced by a factor of about 100.

It seems reasonable to keep the value of $\Omega_2 = 0.22$, since this was deduced by a one-dimensional calculation for both $\sigma_{d.c.}$ and σ'. Hence values of f_2 have been calculated by means of Eqn. (5.11) using the p_{oo} values derived from the approximate three-dimensional calculation.

Table 5.2 tabulates the f_2 values thus calculated with those observed for temperatures of 1500°C and 1350°C and the three oxygen partial pressures.

Table 5.2

T (°C)	1350			1500		
p_{O_2}, atm.	1	10^{-1}	2×10^{-2}	1	10^{-1}	2×10^{-2}
Calculated f_2(Hz)	1.84×10^6	6.8×10^5	7.6×10^5	10.6×10^6	3.7×10^6	4.2×10^6
Observed f_2(Hz)	1.3×10^5	8×10^4	5×10^4	1.3×10^6	2.5×10^5	4×10^5

CHAPTER 6 - CONCLUSIONS

Our overall conclusion is that the variation of conductivity in single crystal Al_2O_3 with temperature, p_{O_2} and frequency is consistent with ionic conduction, although we have no single result which makes this indisputable.

Although the results are in general agreement with previous work, there are some specific areas of disagreement, which will be pointed out in the following as they arise in connection with reviewing the evidence for our main conclusions.

The most novel conclusion is that there is a bulk as well as a surface effect in the a.c. behaviour of Al_2O_3 crystals, and that the frequency dependence of the bulk effect suggests that the interaction between mobile charged defects and oppositely charged immobile defects is quite long range, i.e. significant to at least the 6th neighbour.

The evidence for the two processes is the variation of $\sigma_{a.c.}$ and ϵ at high and low temperatures as already discussed, and the evidence for the long range interaction is the necessity to go to a 6 well-pair model to give computed results for σ which show a slope of $\sigma(f)$ beginning to get down to the observed values at high frequencies and temperatures. Reasonable agreement has been obtained for the frequency at which σ should start to rise with approximate calculations using parameters appropriate to Al_2O_3.

However previous workers (Tallan and Graham, 1965) have concluded that only surface effects are important. They only went up to 1034°C and did not measure $\sigma_{d.c.}$ in the same rig, and did not use a physical guard. Moulson and Popper (1968) found that this is important even at relatively low temperatures. Their results differ from ours in that

they still saw quite large variation (two orders of magnitude) in ϵ' in the KHz region at 1034°C, and the rise in σ had saturated at 10 KHz. Our results at 1050°C show relatively little dispersion in (a factor of 4) and σ is still rising at 10 KHz and even at 10^6 Hz.

This suggests that our samples have a real difference in properties from theirs. We believe the surface mechanism in our samples has moved well below 10^2 Hz and we still have an order of magnitude difference between $\sigma_{d.c.}$ and σ (70 Hz) at 1050°C.

The frequency at which interfacial dispersions occur depend sensitively on the σ_2/σ_1 ratios and d_1/d_2 ratios and the large differences between our results and Tallan and Graham's (1965) may be due to the accessibility of the electrode surface to the ambient gas, or to the equilibrating procedure. Also they used Pt paste, while we used sputtered Pt electrodes. However, our extrapolated results for layer thickness show order of magnitude agreement. Bauerle (1969) found that the Pt paste electrodes were quite non-porous while sputtered Pt electrodes were reasonably porous. He made Pt paste electrodes porous by passing a heavy current through them for several minutes at 800°C.

Two points of practical significance from the σ(f) measurements are:-

i) There is not much difference between $\sigma_{d.c.}$ and σ_m at temperatures greater than 1350°C, therefore high temperature d.c. measurements in the literature are probably acceptable.

ii) The activation energies, quoted in the literature, obtained from $\sigma_{a.c.}$ measurements, at f = 10 KHz, are not significant.

The second, but very tentative main conclusion, is that the previous interpretation of Brook et al. that Al interstitials are the mobile defects, is not supported by this work. A factor in their

argument was the $\sigma_{a.c.}$ versus p_{O_2} dependence of Pappis and Kingery (1961) at high p_{O_2} and our work, as far as it goes, does not agree with theirs. Their slope is 0.2 and ours is 0.6. We have suggested an interpretation in terms of V_{Al}, but this is only very speculative. The results of Pappis and Kingery may be criticised (as mentioned above) on the grounds that many σ measurements were taken at 10^4 Hz, so that only those at 1350°C and above will be close to $\sigma_{d.c.}$ values.

The above conclusions were mainly based on the $\sigma(f)$ and $\sigma_{d.c.}(p_{O_2})$ variations. The remaining aspect of our results is the $\sigma_{d.c.}$ versus $(1/T)$ plots but these do not lead to any very striking conclusions. The experimental results are in reasonable agreement with post 1968 work (i.e. those where a physical guard was used) over the same temperature range. However, at lower temperatures the results are at variance with those of Özkan and Moulson (1970) in showing a new region of lower activation energy. Brook et al (1971) did not measure below 1000°C. This behaviour resembles that of other ionic conductors, e.g. alkali halides, and is usually attributed to association. However, we have not attempted any detailed interpretation of this.

Summing up, the results and discussion show that systematic measurements of $\sigma(f)$ and $\sigma_{d.c.}$ versus T, p_{O_2} and further refinement of the theory of $\sigma(f)$ should eventually yield valuable information about defect densities, barrier heights and jump frequencies and the variation of interaction potential with distance between oppositely charged defects, and indeed the present study has already suggested that this interaction is probably a good deal longer range than usually considered (i.e. nearest neighbour only in the double potential well).

For Al_2O_3 single crystals, in particular, the most informative measurements to make next would be conductivity as a function of oxygen partial pressure at a p_{O_2} higher than 10^{-1} atm. and at lower p_{O_2} with a

sample holder where physical guard is incorporated. Conductivity measurements at temperatures higher than 1600°C would also be interesting in the light of Yee and Kröger's (1973) work showing that conductivity starts to become mixed. The theoretical work should be extended into 3-dimensions or at least numerical methods may be used to solve the present matrix equations in more general cases.

APPENDIX I

Law of Mass Action

The rate or speed of a chemical reaction is proportional to the concentration of each of the reactants. In the reaction between A and B to produce C and D

$$A + B \rightarrow C + D$$

the speed is proportional to the number of moles per litre of A. The speed is also proportional to the number of moles per litre of B. If the number of molecules of A in one litre is doubled and the number of B molecules is trebled, the rate will be six times faster, hence

$$R \propto [A] \times [B]$$

where R is the rate of reaction, and the brackets signify "moles per litre" of A and B. 1 mole contains 6.02×10^{23} molecules.

Since the rate is also dependant upon other factors, namely, the nature of A and B and temperature, it is necessary to add a term to the proportionality so that the proportionality sign may be replaced by an equality sign

$$R = k \times [A] \times [B]$$

Once the reactants have been chosen, their nature is fixed. Consequently, the magnitude of k for a given reaction at a given temperature does not change. This quantity k is referred to as the rate

constant for the reaction. We next consider the reverse reaction

$$C + D \rightarrow A + B$$

The speed of this reaction is given by the expression

$$R_r = k' \times [C] \times [D]$$

The numerical value of k' cannot be the same as that of k because the natures of C and D differ from that of A and B.

Now let us consider the reaction

$$2A + B \rightarrow C \quad \text{or} \quad A + A + B \rightarrow C$$

In this case, the speed of reaction is proportional to the square of the concentration of A, for it is necessary that two molecules of A collide for the reaction to occur. The reaction is

$$R = k \times [A] \times [A] \times [B] = k \times [A]^2 [B]$$

At equilibrium, the speed of the forward reaction is equal to the speed of the reverse reaction

$$R_f = R_r$$

For the equilibrium $A + B \rightleftharpoons C + D$

$$R_f = k [A][B]$$
$$R_r = k' [C][D]$$

Then at equilibrium

$$k\,[A]\,[B] = k'\,[C]\,[D]$$

rearranging, we get

$$\frac{k}{k'} = \frac{[C]\,[D]}{[A]\,[B]}$$

For a given reaction at a specific temperature, the quantities k and k′ do not change. Hence we may replace k/k' by the term K

$$K = \frac{[C]\,[D]}{[A]\,[B]}$$

The quantity K is called the equilibrium constant.

APPENDIX 2

Approximate 3-Dimensional Calculation of N_7 and $f_7(0)$

The derivation of Eqns. (5.18) and (5.19) for N_7 and $f_7(0)$ on an approximate three dimensional basis is as follows.

The site density in the crystal is N_s, and it contains defect density N_d, and each site has C nearest neighbours at distance a. Let N_1, N_2 ... N_6 be the numbers of sites at distances a, 2a, ... 6a from the immobile defects, and N_7 the number of all the sites beyond distance 6a. The numbers N_i will increase approximately as the volumes of shells with radii a, 2a, ... 6a, etc. Hence we can write

$$N_1 = CN_d; \quad N_2 = 2^2\, CN_d; \quad \ldots\ldots \quad N_6 = 6^2\, CN_d \tag{A.1}$$

Hence $N_7 = N_s - (N_1 + N_2 + \ldots\ldots + N_6)$

$$\text{i.e.} \quad N_7 = N_s - CN_d\,(1 + 2^2 + \ldots + 6^2) \tag{A.2}$$

which is Eqn. (5.18).

Again

$$f_i = f_7(0)\ (\exp w_{i,7}) \tag{A.3}$$

where $w_{i,7} = \Delta W_{i,7}/kT$ and $\Delta W_{i,7}$ is the energy difference between the bottom of the i^{th} well and that of the 7th and other wells beyond. Hence

$$f_1 = f_7(0) \exp 6\, w_{12}; \quad f_2 = f_7(0) \exp 5\, w_{12};$$

$$\ldots\ldots\ f_6 = f_7(0) \exp w_{12} \tag{A.4}$$

If $n_i = N_i f_i$ is the average number in wells type i then we have

$$\sum_{i=1}^{m+1} n_i = \sum_{i=1}^{m+1} N_i f_i = N_d \tag{A.5}$$

For m = 6 and using A.1 and A.4 for N_i and f_i, one finds

$$CN_d f_7(0) \exp 6w_{12} + 2^2 CN_d f_7(0) \exp 5w_{12} + \ldots + 6^2 CN_d f_7(0) \exp w_{12}$$

$$+ N_s f_7(0) - CN_d f_7(0) (1 + 2^2 + \ldots + 6^2) = N_d$$

which reduces to

$$f_7^{-1}(0) = C\left\{\left[\exp(6w_{12}) - 1\right] + 2^2\left[\exp (5w_{12}) - 1\right] + \ldots\ldots\right.$$

$$\left. + 6^2 \left[\exp(w_{12}) - 1\right] \right\} + \frac{N_s}{N_d} \tag{A.6}$$

which is Eqn. (5.19).

REFERENCES

Bassani, F. and Fumi, F.G. (1954) Nuovo Cimento, 9, 274

Bauerle, J.E. (1969) J. Phys. Chem. Solids, 30, 2657-2670

Boswarva, I.M. and Franklin, A.D. (1967) Mass Transport in Oxides, Proceedings of a Symposium

Brook, R.J. (1974) in Electrical Conductivity in Ceramics and Glass, Part A, Edited by N.M. Tallan, Marcel Dekker

Brook, R.J., Yee, J. and Kröger, F.A. (1971) J. Amer. Ceram. Soc. 54 (9), 444-54

Cassell, W. (1964) Linear Electric Circuits, Wiley

Champion, J. (1964) Brit. J. Appl. Phys. 15, 633

Champion, J. (1968) Proc. Brit. Ceram. Soc. 10, 51

Cohen, J. (1959) J. Amer. Ceram. Soc. 38 (9) 441

Daniel, Vera V., (1967) Dielectric Relaxation, Academic Press

Dasgupta, S. and Hart, J. (1965) Brit. J. Appl. Phys. 16 (5), 725

Fröhlich, H. (1958) Theory of Dielectrics, Oxford University Press

Harrop, P.J. and Creamer, R.H. (1963) Brit. J. Appl. Phys. 14, 335

Jones, T.P., Coble, R.L. and Mogab, C.J. (1969) J. Amer. Ceram. Soc. 52 (6), 331-334

Kofstad, Per. (1972) Nonstoichiometry, Diffusion and Electrical Conductivity in Binary Metal Oxides, Wiley-Interscience

Kröger, F.A. (1964)Chemistry of Imperfect Crystals, North-Holland

Matsumura, T. (1966) Canadian J. Physics 44, 1685-1698

Mills, J.J. (1968) Research on Electrical Conductivity and Conduction Mechanisms in Alumina

Moulson, A.J. and Popper, P. (1968) Proc. Brit. Ceram. Soc., No. 10, 41-50

Özkan, O.T. and Moulson, A.J. (1970) Brit. J. Appl. Phys. 3(6), 983-86

Paladino, A.E. and Kingery, W.D. (1962) J. Chem. Phys. 37(5), 957-962

Pappis, J. and Kingery, W.D. (1961) J. Amer. Ceram. Soc. 44(9), 459-464

Pauling, L. and Hendricks, S. (1925) J. Amer. Chem. Soc. 47, 781

Peters, D.W., Feinstein, L.F. and Peltzer, C. (1964) J. Chem. Phys., 42 (7), 2345-46

Reilly, M.H. (1970) J. Phys. Chem. Solids, 31, 1041

Shewmon, P.G. (1963) Diffusion in solids, McGraw Hill

Steele, B.C.H. (1968) Electromotive force measurements in high temperature systems, p. 3-27. Institution of Mining and Metallurgy

Tallan, N.M. and Graham, H.C. (1965) J. Amer. Ceram. Soc. 48 (10) 512

Tallan, N.M. and Detwiller, D.P., (1963) J. Appl. Phys. 34 (6) 1650

Van Gool, W. (1966) Principles of Defect Chemistry of Crystalline Solids, Academic Press

Volger, J. (1960) in Progress in Semiconductors, Vol. 4, Dielectric Properties of solids in relation to imperfections, John Wiley and Sons Inc.

Yee, J. and Kröger, F.A. (1973) J. Amer. Ceram. Soc. 56 (4), 189-191