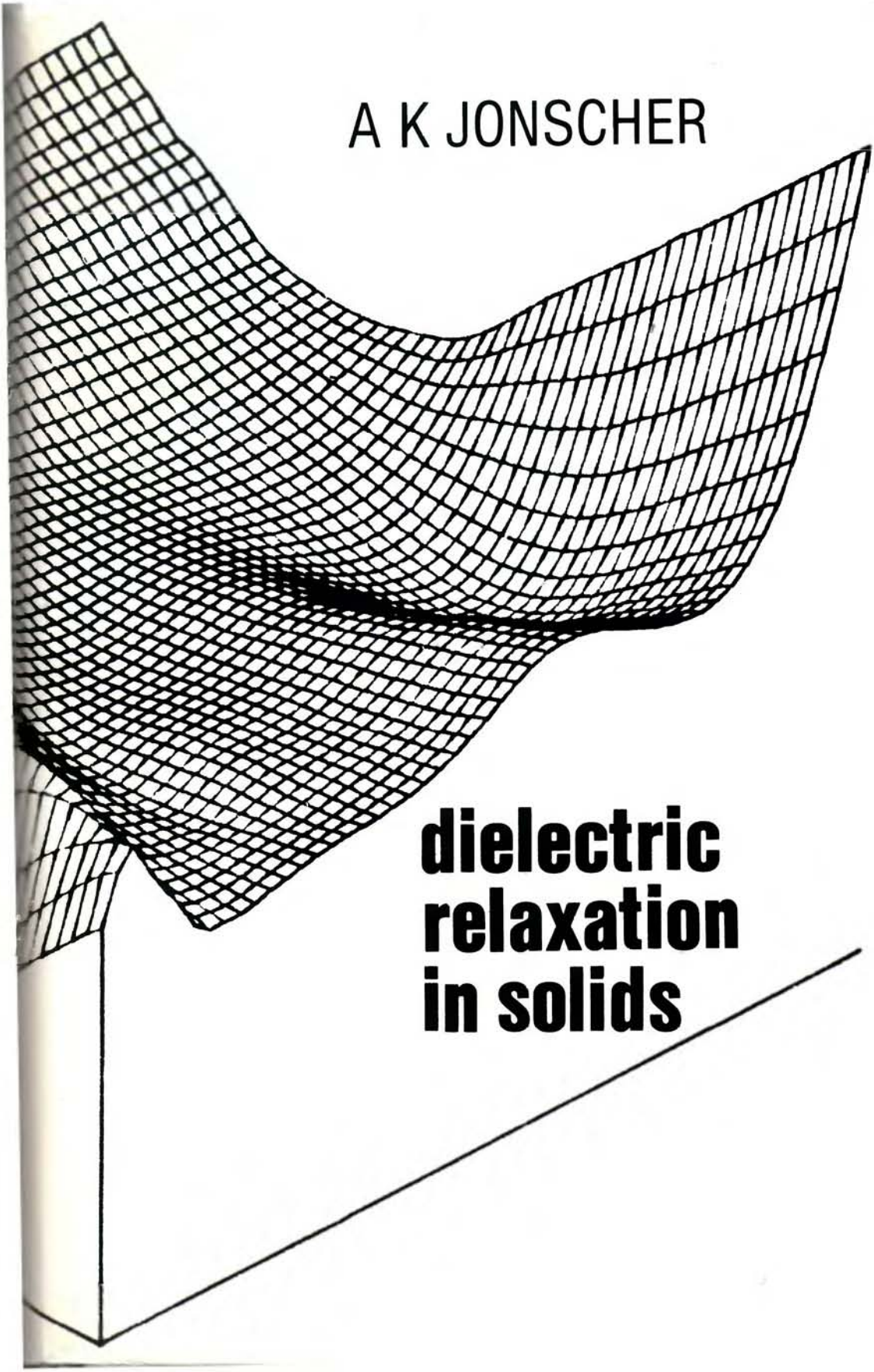
A K JONSCHER



dielectric relaxation in solids

dielectric relaxation in solids

A K Jonscher

dielectric relaxation in solids

Chelsea Dielectrics Press
London 1983

Chelsea Dielectrics Press Ltd
33 Lynwood Road London W5 1JQ

ISBN 0 9508711 0 9

Layout and cover design
by Juliusz L Englert
incorporating a diagram from the same source as Figure 3.29 by kind permission of F Meca.

Computer typeset and printed by
Page Bros (Norwich) Ltd

Contents

Chapter 3
PRESENTATION OF DIELECTRIC FUNCTIONS

Chapter 4
THE DYNAMIC RESPONSE OF IDEALISED PHYSICAL MODELS

Chapter 5
EXPERIMENTAL EVIDENCE ON THE FREQUENCY RESPONSE

Chapter 6
EXPERIMENTAL EVIDENCE ON THE TIME RESPONSE

Preface

Felix qui potuit rerum cognoscere causas†
Virgil

The quest for understanding the laws of Nature constitutes one of man's most basic urges and lies at the origin of all discovery. Another is the urge to harness the discoveries to practical ends which is the mainspring of technology. These two attributes are clearly visible in the development of the science and technology of dielectrics over the last century. In the present monograph I am primarily concerned with the first role – the understanding of the laws of dielectric relaxation – while being hopeful that beneficial consequences of this better understanding for the technology of dielectrics will follow.

A growing understanding of Nature leads to the introduction of a greater order into a previously rather confused picture and to the establishment of an organic unity in the complex range of empirically known phenomena. This unifying order produces a strong aesthetic appeal – an ordered and unified system is beautiful and the happiness to which Virgil alludes in the motto quoted above is a powerful motivation for scientific work.

This monograph proposes such a unifying and ordering treatment of the richly varied subject of dielectric relaxation. It will be left to the reader to judge the extent to which this objective is achieved.

Although I have to accept the sole responsibility for the opinions expressed in this monograph, it is my pleasure to acknowledge invaluable help from many individuals, especially from my immediate colleagues in the Chelsea Dielectrics Group. Among them I would wish to mention Professor Robert Hill whose contribution can be measured by the number of references and acknowledgements in these pages. Both he and Dr Len Dissado have developed the new many-body theory of dielectric relaxation in continuing

† Happy is he to whom it was given to understand the causes of things.

contact with the experimentalists working in our laboratory. The names of many Research Students and Visitors appear in acknowledgements, and I am particularly indebted to Mr John Pugh whose help with the development and the running of the Frequency Response Analysers in our laboratory is much appreciated.

This work would not have been possible without continuing support and understanding from successive Principals of Chelsea College – Dr Malcom Gavin, Dr David Ingram and currently Dr Charles Phelps; from the Head of the Physics Department Professor E J Burge and from my many Colleagues in the Department. Many aspects of this work were developed in close association with Professor Robert Lacoste of the Laboratoire de Génie Electrique at the Université Paul Sabatier in Toulouse. Essential material support was provided by Chelsea College, by the Science and Engineering Research Council as well as by other bodies. More detailed acknowledgements of this support, together with the account of current work in progress may be found in successive Progress Reports of the Chelsea Dielectrics Group which are available on request.

The loving understanding shown by my wife for the innumerable hours of painstaking work is also acknowledged – without it this monograph would never have seen the light of day.

Andrew K Jonscher

Chelsea Dielectrics Group

April 1983

Useful Physical Constants

magnitude of electronic charge	e	$= 1.602 \times 10^{-19}$	C
speed of light in free space	c	$= 2.998 \times 10^{8}$	ms^{-1}
permittivity of free space	ε_0	$= 10^7/4\pi c^2$	
		$= 8.854 \times 10^{-12}$	Fm^{-1}
mass of electron	m	$= 9.109 \times 10^{-31}$	kg
mass of hydrogen atom	m_H	$= 1.673 \times 10^{-27}$	kg
Boltzmann's constant	k	$= 1.381 \times 10^{-25}$	JK^{-1}
		$= 0.862 \times 10^{-4}$	eVK^{-1}
thermal energy at 300 K	kT	$\cong 1/40$	eV
Avogadro's number	L	$= 6.023 \times 10^{23}$	mol^{-1}

atomic excitation energy $1\ eV = 2.305 \times 10^4$ calorie mol^{-1}

Debye unit of dipole moment $1D = (1/3) \times 10^{-29}$ Cm·
approximately equal to charges $\pm$ e at a distance of $0.2\ Å = 2 \times 10^{-11}$m.

CHAPTER 1

INTRODUCTION

1.1 DIELECTRICS AND INSULATORS

The use of electrical insulation is as old as the science and technology of electrical phenomena; it goes back at least a century and a half, while the recognition of specifically electrostatic manifestations of electrification goes back to antiquity. Systematic investigations of dielectric properties may be traced back to the 1870's.

The accumulated experimental and theoretical material is vast and from an early stage on it was possible to discern two essentially complementary approaches to this wide-ranging subject – the study and development of insulators and of dielectrics. In this classification, insulators are materials used to prevent the flow of current where it is not desired, especially in the context of electrical and electronic engineering, and the principal interest in them lies in achieving the lowest possible electrical conduction coupled with the maximum resistance to destructive breakdown in high electric fields. Other factors such as long life, low cost, chemical inertness and the ability to withstand elevated temperatures may be added to the long list of technical specifications which must be met by modern insulating materials working sometimes under extreme external stresses. It is understandable that engineers and materials scientists searching for insulating materials suitable for specific applications were less concerned with the detailed physical mechanisms governing the behaviour of these materials, provided that their characterisation in terms of clearly defined parameters could be achieved reliably and simply. This order of priorities remains true to this day and the chief emphasis in electrical insulation science falls on the synthesis of materials and their characterisation.

By contrast with the insulation aspect, dielectric phenomena are at once more general and more fundamental – after all, insulators *are* dielectrics – and are concerned more intimately with the microscopic mechanisms of dielectric polarisation and include, especially, the transient behaviour under time-varying electric fields. Much of the available experimental evidence and a good deal of the

theoretical treatments of the subject would be of little interest to the electrical engineer searching for insulating materials.

The question of why insulators should possess their peculiarly desirable properties distinguishing them from both metals and semiconductors had to await the development of the band theory of solids in the 1930's when the distinction between metals, semiconductors and insulators was formulated in terms of the relative disposition of the conduction and valence bands and of the position of the Fermi level between them (Kittel 1976, Smith 1978). Insulators were considered to be materials with negligible numbers of thermally generated carriers of electronic nature. We now know that a small number of charge carriers, while important, is not the only criterion of acceptable insulating behaviour – the other requirement is their mobility which has to be low in good insulators. This, in turn, had to await the development of the modern theory of low-mobility hopping conduction in the last twenty years, or so (Mott and Davis 1979).

Two of the most important dielectric concepts to be developed concerned the *induced* and *orientational* polarisation in external electric fields. The first treatment of the static response by Clausius and Mossotti in the 1870's represents one of the earliest attempts to approach a many-body situation in an elementary way and it may be said to have been a century ahead of its time!

The other decisive development came in the early years of this century with Debye's theory of dynamic dielectric response of freely floating dipoles, such as might be found in dipolar liquids. This opened up a new era of investigations of the frequency- and time-dependence of the dielectric response of materials where two separate tendencies became evident as time went on. The principal interest of chemists and some physicists lay in the study of the dynamics of molecular motion in polar materials and the dielectric response represented an excellent tool for this work. The behaviour of insulators in alternating electric fields or under impulse conditions was of interest, although for different reasons, to electrical engineers. At the level of transient phenomena, the distinction between insulators and dielectrics becomes rather blurred, but it is significant that engineers were for a long time using the notion of the *anomalous* charging current which is observed when a step-function field is applied to a dielectric – as we shall see later, the very essence of dielectric behaviour was called anomalous as if it ought not to be there!

This division of interest in dielectrics between electrical engineers and chemists is clearly discernible to this day, with the addition of electrochemists whose principal interest lies in ionic conductors and, more recently, a strong group of solid state physicists working on the alternating current conduction in amorphous electronic semiconductors. The result of this state of affairs is a far reaching fragmentation of dielectric studies between groups which do not collaborate with one another, often do not see any community of interest among themselves and which use separate and different methods of evaluation, analysis and measurement. An example of this may be seen in the preference for the use of dielectric loss by chemists, loss tangent by electrical engineers and alternating current conductivity by physicists. Many examples of these differences will be found in the present treatment which seeks to integrate the various separate approaches.

One very general feature of the science of dielectrics, especially at "low" frequencies below the microwave range, is the comparative scarcity of new physical concepts during the last half-century – and this despite a very rapid growth of the understanding of many other aspects of solid state physics. We shall be able to see in the course of the development of the theory of dielectric processes that this topic throws important fresh light on the understanding of hitherto little known aspects of the theory of solids.

We shall be able to show that the time- and frequency-dependence of the dielectric response of solids is uniquely sensitive to certain relatively weak many-body interactions which largely dominate this response in wide ranges of frequency and time and also in relatively wide ranges of temperature. This influence of many-body interactions is much stronger in transient dielectric processes than in any other form of static behaviour and this probably accounts for the prevailing lack of appreciation of these phenomena in solid state physics. The only other branch of relatively modern physical theory where similar processes have been studied for some while is the theory of amorphous solids and spin glasses and many of the ideas advanced in the present work derive from that subject.

1.2 THE NATURE OF DIELECTRIC RESPONSE

We have already mentioned the fact that the dielectric response of many materials has been of immediate interest to scientists and technologists for a long time and this accounts for a great wealth

of experimental data which are available for interpretation. Furthermore, advanced forms of modern measuring equipment are on the market and are being extensively used in many centres for the study of dielectric materials. What is more problematical is the precise physical and chemical characterisation of these materials and this represents one of the principal difficulties in assessing the value of the available experimental data and in drawing any generalised conclusions from them.

The awareness of this problem is the more acute among material scientists since the development of the technology of semiconductors and also of many other materials has drawn attention to the often critical influence of even minute quantities of impurities and of structural defects on the physical properties of these materials. In relation to these, dielectric materials represent very poorly characterised systems with a very large content of impurities and imperfections. The question is legitimately raised, therefore, to what extent is it possible to assess the value of any experimental information obtained on such unreliable materials?

The present work nevertheless undertakes precisely such a generalised assessment and draws far-reaching conclusions from experimental data, justifying this action by certain aspects of this experimental information which were not perhaps sufficiently well appreciated until now. One of these aspects is the apparent "universality" of the form of the dielectric response which has been known for three-quarters of a century under the name of the Curie-von Schweidler law (von Schweidler 1907). This law states that the discharge or depolarisation currents of a wide range of dielectric materials follow the *power* law of time dependence

$$i(t) \propto t^{-n} \tag{1.1}$$

instead of the exponential relation which would correspond to the simplest first-order differential equation decay characterising the Debye mechanism.

We shall show, on the basis of many examples, that this law is indeed a "universal" one and it is only weakly influenced by the detailed physical and chemical characterisation of the materials to which it applies.

Another unique advantage of the experimental information on dielectric relaxation is the unrivalled range of the dynamic variables of time and frequency to which it applies. This wide range is further enhanced by the principle of time-temperature or frequency-tem-

perature super-position which enables data taken at different temperatures to be related to one another, thereby increasing the range and reliability of these data.

The breadth of the dynamic range is illustrated with reference to Figure 1.1 which shows the entire conceivable time and frequency

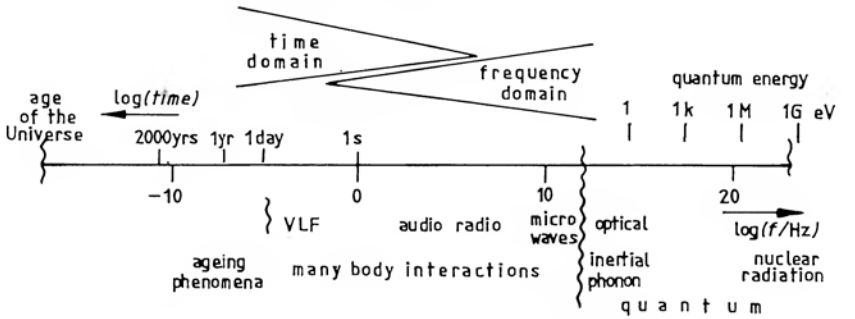


Figure 1.1 A schematic representation of the entire physically accessible range of time and of frequency on a logarithmic scale extending from frequencies corresponding to nuclear energies of the order of GeV, to the age of the universe, 10^{17} s. The energy scale is given in electron–volts. The ranges of inertial, phonon and quantum processes are indicated, and the lower frequency region is marked as "many-body interactions" in anticipation of arguments to be presented later in this work. The diverging lines indicate the increasing importance of time and frequency domain measurements. VLF stands for Very Low Frequency. The distinction between many-body relaxation processes and irreversible ageing phenomena is not meant to be rigid.

range, from frequencies corresponding to nuclear energies of 10 GeV, down to the reciprocal age of the Universe – a total of 40 decades. If we now note that the behaviour at frequencies in excess of some THz is dominated by inertial and by phonon and quantum processes which fall outside the scope of the present treatment, and if we note that the practical lower end of frequency measurements is 10^{-5} Hz, we have a total of 17 decades of "low" frequency in which to study the dielectric behaviour. This range may further be extended to twenty or more decades by the time-temperature superposition mentioned above. We are therefore in the uniquely favourable position of being able to cover experimentally up to two-thirds of the total logarithmic range of times and frequencies corresponding to the sub-quantum range.

The width of the dynamic range has as its consequence the possibility and the need to interpret the data in terms of sufficiently general theoretical models – while almost any model would do over

one or two decades, it becomes much more difficult to "fudge" the interpretation when one has ten or more decades at one's disposal.

1.3 THE PURPOSE AND SCOPE OF THE PRESENT TREATMENT

The subject of dielectrics is so rich in content and so varied in scope, ranging from theory through experimental studies to technological applications, and the resulting literature is so voluminous, that it is essential to define clearly the purpose and scope of the present work, so as not to repeat material which is well covered in other texts.

At the most elementary level it is possible to distinguish between the *static* or *steady state* response to a steady electric field and the *dynamic* response to time-varying electric fields. One particular aspect of time-dependent behaviour relates to the decay of polarisation from an initial steady state to zero after the sudden removal of an initial polarising field. In the case of inertia-less systems in the "low-frequency" sub-quantum limit defined in Figure 1.1, this decay is referred to as *dielectric relaxation* and this will later be shown to be intimately connected with the response of dielectrics under sinusoidally varying electric fields.

With regard to the static response, it is possible to state that the quantitative relationship between the polarising field and the resulting polarisation can only be calculated for the simplest cases, and is usually restricted to relatively dilute dipolar systems in which dipole-dipole interactions are not significant. In the great majority of practically important dielectric materials, it is not possible to test the relationship between static field and polarisation in terms of any specific model since we lack the essential knowledge of such numerical parameter as the density and magnitude of the dipolar species responsible for the dielectric behaviour. To that extent, therefore, the analytical treatment of the static response is of limited practical significance.

The situation is quite different in the case of the dynamic response or relaxation phenomena, since there it is possible to discern the applicability of a remarkable "universality" of the dynamic behaviour in virtually all dielectric materials, regardless of their physical and chemical properties and regardless of the nature of the polarising species, be they dipoles, hopping electrons or ions. The principle of universality of the dynamic response of dielectric relax-

ation, whether in the "time-domain" or in the "frequency domain", was first formulated on the basis of extensive experimental evidence and has resulted in a new classification of all types of dielectric response, including not only typical dielectric materials but also many others which are not normally regarded as such (Jonscher 1977, 1980). This, in turn, has led to a new theoretical approach which forms the subject of the last Chapter of the present Monograph and which provides, for the first time, a unified theoretical interpretation of dielectric data which were hitherto considered to require separate and very different treatment according to the type of material concerned. This unified approach has also led to the development of fresh insights into such related topics as the "1/f noise" and the physical basis of dielectric loss. It has also provided a consistent foundation for some of the widely used techniques in the interpretation and presentation of dielectric data, such as the "normalisation" of results obtained at different temperatures.

With the deepening understanding of the physical processes governing the frequency-domain *spectral shape* of dielectric relaxation it became possible to provide more reliable interpretation of the different forms of these spectra than was possible up to now, and we are devoting some attention to this aspect of dielectric studies. In particular, we show that some of the widely accepted "interpretations" or conclusions drawn from dielectric spectra are not justified in the light of the present analysis.

We begin in Chapter 2 with a brief discussion of some fundamental concepts relevant to the dynamic behaviour of polarisation, including the mathematical formulation of time- and frequency-domain analysis. Chapter 3 gives a detailed discussion of the often neglected topic of *presentation* of dielectric data, which provides a key to the proper evaluation of experimental results. In particular, we show the importance of using the log–log representation of the real and imaginary components of dielectric susceptibility against frequency, as opposed to the prevailing method of plotting the *linear* susceptibility against logarithm of frequency. The reason for this choice becomes evident in the light of the predominance of *power-law relationships,* which become straight lines in the log–log presentation, while forming non-descript curvilinear contours in the semi-logarithmic representation.

Chapter 4 discusses the dynamic response of idealised physical models, partly to familiarise the reader with the methods of approach and partly to show him the extent to which such idealised

models can represent the observed behaviour of dielectric solids. Chapter 5 then presents a wide selection of experimental information relating to the frequency domain, culminating in our general classification of the various types of dielectric response. This synoptic presentation covers not only traditionally recognised dielectric materials, but also many examples of dielectric response observed in systems such as semiconductor p–n junctions, amorphous electronic semiconductors, fast ionic conductors and so on, the avowed objective being to show the full breadth of the spectrum of materials which obey the universal principle. Some of the data given in our review were originally presented in different forms, e.g. as complex impedance plots or as complex permittivity plots and we had to adapt them to our purposes, as far as was possible, to maintain a unified form of presentation.

Chapter 6 gives the complementary information on the time-domain response of dielectric materials – this is much less comprehensive than the frequency domain data because the measurement techniques are less well developed, but the data included provide some specific illustrations for example, for the effects of non-linear behaviour arising from injection of charge carriers.

Up to this point we have very deliberately avoided any form of interpretation of the experimental data – we wished to let the facts speak for themselves without thrusting premature conclusions on the reader's mind. This is part of our endeavour in the present work:

to see dielectrics as they really are, and not as we'd like them to be.

Chapter 7 then gives a very brief discussion of the existing treatments of the interpretation of dielectric data – the express object here is not to enter into details which are available in many excellent texts, but rather to point out the inadequacy of the accepted interpretations in the face of the accumulated experimental evidence.

The presentation of the new many-body interpretation of the dielectric behaviour is then given in Chapter 8. We have left this discussion to this late stage because we wished it to be based firmly on a wide range of experimental evidence and also on the conclusion regarding the inadequacy of other treatments – one does not lightly advance a completely new approach to a well established subject! However, our entire approach is based on the conviction – and we assert that there is scope for conviction in science – that where

Mother Nature presents us with an overriding uniformity of behaviour in a wide range of physical situations, a fundamental unifying principle is very likely to be found behind an apparent multiplicity of different behaviours.

Throughout the treatment we limit ourselves to the region below the microwave range, since this is where the "low-frequency" universal response is seen clearly without interference from quantum and inertial phenomena. Apart from this restriction, however, we have not consciously tried to exclude any experimental evidence as being "inconvenient" or, in some way, contradicting the universality we strive to establish.

The exclusion of inertial and quantum effects does not imply that we think that many-body processes do not have a role to play there – they most certainly do, but the range of experimental information is not sufficiently wide and the influence of many-body processes not sufficiently strong to make it worth including these in the present treatment.

The present Monograph inevitably builds on the achievements of past generations of workers in this field. We rely heavily on the understanding of the microscopic processes of dipolar orientations in solids and in liquids contained in many excellent classics such as Böttcher and Bordewijk (1978), Fröhlich (1955), von Hippel (1954), Smyth (1955), Daniels (1967), McCrum et al (1967), to mention but a few.

In order not to make our treatment unduly long, we limit severely any detailed discussion of the nature of the various chemical aspects of the subject. We dwell only briefly on the subject of static behaviour, mainly on the basis of the fact that understanding of this aspect has not advanced in the last twenty years, or so, and there is no fresh evidence which we can bring to reader's attention.

We do not cover specifically ferroelectric phenomena, although we do quote some frequency response data for ferroelectric materials. There is ample specialised literature on this subject and we are only concerned with it to the extent of the dynamic behaviour. Likewise we leave aside the important subject of anisotropic dielectrics with their tensor properties, which is covered in specialised textbooks e.g. Zheludev (1971), again mainly on the basis of the fact that the specifically *dynamic* aspect of response is either not very sensitive to anisotropy, or is not sufficiently well documented. We do make some references to such cases as have come to our notice

but these are often concerned with strongly anisotropic materials, e.g. layer compounds. We are not dealing with the specialist topic of hyperpolarisation below optical frequencies, once again because there is little evidence of specifically dynamic behaviour under conditions of electric fields which are sufficiently high to give rise to significant non-linearity of dielectric response.

While hopping electronic and ionic charges enter very prominently into our discussion of the dynamics of dielectric response – we emphasize this aspect more than do typical dielectric texts – the subject of direct current flow and, more specifically of current flow under high fields where Poole-Frenkel, Poole and Schottky processes become dominant is left out of consideration. Adequate references to this field may be found in Mott and Davis (1979) and also in Jonscher and Hill (1975).

We say nothing about electrets (Sessler, 1980) and the related topic of thermally stimulated depolarisation (van Turnhout 1975).

The role of contacts and surfaces on dielectrics is certainly very important, not least because it is known to give rise to injection of charge carriers and may lead to unexpected processes taking place. Unfortunately, it must be said that our systematic knowledge of these phenomena is only relatively recent (Davies 1969, Lederer et al 1980, Barnes et al 1981) and most of it derives from the extensive studies of semiconductor surfaces and interfaces from which we argue by analogy.

An important aspect of the technology of dielectrics is ageing, especially under thermal and electrical stress in the context of electrical insulation. We indicate in Figure 1.1 that the realm of ageing phenomena is not clearly separated from that of dielectric relaxation which is of principal interest to us in the present work. A distinction which we propose to introduce, however, is that relaxation is reversible while ageing is not and we shall leave this matter at that.

A comment is called for regarding the restriction of the present treatment to the response of solid dielectrics, to the exclusion of liquids and of gases. The latter form a class in their own right in view of their low density and consequently the absence of significant inter-particle interactions. Liquids, on the other hand, share with solids the close proximity of particles and the importance of chemical bonds, and their dielectric properties are also to some extent influenced by the same many-body interactions as in the case of solids.

On the other hand, for a variety of reasons, liquids approximate more closely to the classical Debye model than do solids and our arguments concerning the universality of non-Debye responses would have been somewhat clouded – though not invalidated – by the inclusion of liquids at this stage. Solids represent in this respect a more favourable medium in which to establish the principle of many-body interactions, although later it will be possible to see that the same phenomena are also applicable to some extent in liquids.

Throughout the present work we are trying to place the main stress on conveying the physical meaning of the phenomena taking place in solid dielectrics. We believe that this form of exposition is of most direct benefit to the reader, who is interested not only in knowing *how* dielectrics behave, but principally in the reason *why* they behave in this particular manner. In this exposition, we regard mathematics as an indispensable tool for the derivation of theoretical principles and for the quantification of relationships, but we do not wish to elevate the mathematical arguments to the level of ends in themselves. Wherever possible, therefore, we employ the simplest mathematical treatment, rather than the most rigorous one, while striving to preserve the correctness of the results obtained. In this manner, the interested reader may turn to more specialised works for a complete and rigorous mathematical exposition, while the reader concerned mainly with the physical argument will be satisfied with the knowledge that a mathematical derivation is possible.

For the same reason we avoid also the derivation of formulae that are more complicated than what we can actually exploit in the course of the arguments developed in the present text. We do not wish to burden the reader with a voluminous compendium of formulae which have been derived at one time or another, but whose usefulness from our point of view is very limited. This applies particularly to the many rather formal expressions which are being used in the literature and whose physical significance is very slight.

While giving the principles on which the measurements on the time- and frequency-domain behaviour are based, we do not provide any detailed description of the techniques of dielectric measurements themselves, which constitute a highly specialised branch of knowledge. The equipment available is now at a very advanced level of development and it would be futile to try to give justice to this complex field in the present treatment. The relevant information

may be obtained from specialist literature but at present increasingly from manufacturers' instruction manuals.

On the other hand, as already stated, we do provide an extensive treatment of the methods of dealing with the data once these have been collected. This part of the analysis is much less familiar to many workers in the field of dielectric measurements and there are very few texts dealing with this aspect of analysis (Jonscher 1978, Hill 1981, Hill and Dissado 1982).

Very up-to-date reviews of several branches of the field of dielectric studies may be found in a recent volume of conference proceedings (Goodman 1980).

REFERENCES TO CHAPTER 1

Barnes C, Lederer P G, Lewis T J, Toomer R 1981, J. Electro. Stat. **10**, 107
Böttcher C J F and Bordewijk P 1978, *Theory of Electric Polarisation* Vols. I and II, Amsterdam, Elsevier
Daniels V V 1967, *Dielectric Relaxation*, Academic Press
Davies D K 1969, *Static Electrification*, Institute of Physics Conference Series No. 4, p. 29
Fröhlich H 1955, *Theory of Dielectrics*, Oxford University Press
Goodman C H L (Ed) 1980, *Physics of Dielectric Solids*, Institute of Physics Conference Series No. 58, London
Hill R M 1981, phys. stat. sol. (b) **103**, 319–28
Hill R M and Dissado L A 1982, J Phys. C: Solid State Physics **15**, 5171
Jonscher A K 1977, Nature **267**, 673–679
—— 1978, Thin Solid Films **50**, 187–204
—— 1980, *Physics of Thin Films* Vol 11, 205, M H Francombe (Ed), Academic Press
Jonscher A K and Hill R M 1975, *Physics of Thin Films* Vol 8, 169–249
Kittel C 1976, *Introduction to Solid State Physics*, John Wiley & Sons
Lederer P G, Lewis T J, Toomer R 1980, Ann. Report of Conference on Electrical Insulation and Dielectric Phenomena, National Academy of Sciences, 58
McCrum N G, Read B E and Williams G 1967, *Anelastic and Dielectric Effects in Polymeric Solids*, J Wiley, New York
Mott N F and Davis E A 1979, *Electronic Processes in Non-crystalline Materials*, Oxford University Press
Sessler G M 1980, *Electrets*, Berlin, Springer
Smith R A 1978, *Semiconductors*, Cambridge University Press
Smyth C P 1955, *Dielectric Behaviour and Structure*, McGraw-Hill
van Turnhout J 1975, *Thermally Stimulated Discharge of Polymer Electrets*, Amsterdam, Elsevier
von Hippel A R 1954, *Dielectric Materials and Applications*, J Wiley, New York
von Schweidler E 1907, Ann d Physik **27**, 711
Zheludev I S 1971, *Physics of Crystalline Dielectrics* Vols. I and II, Plenum Press, New York

CHAPTER 2

The Physical and Mathematical Basis of Dielectric Polarisation

2.1 CHARGES, DIPOLES AND CHEMICAL BONDS

At the atomic level, all matter consists ultimately of positive and negative charges balancing each other microscopically or macroscopically: the former corresponding to *overall charge neutrality*, the latter giving rise to *local space charge*, but even in that condition one expects to find the *net charge* over a given volume or sample to be zero. As examples we may note that a neutral gas consists of atoms or molecules which, while neutral, contain electrons in the outer orbits which may be detached completely if sufficient energy is supplied to the system, e.g. in the form of thermal excitation or through the absorption of a light quantum, thus *ionising* the atom or molecule and leaving a positively charged ion and the detached electron. The positive ion and electron may *recombine* again to yield the neutral molecule and the processes of ionisation and recombination may come to an equilibrium with a certain density of the charged species being maintained – this is the case of a *plasma* which is neutral overall but which contains these mobile charges in addition to any neutral particles that may be present.

A situation closely analogous to the gas plasmas arises in the case of many solids (Platzmann and Wolff 1973), where the valence electrons of a metal may be considered to constitute a cloud of negative charge which is neutralised by the *fixed* positive charge of the atomic cores constituting the lattice itself. Similarly, in intrinsic semiconductors the electrons and holes are present in equal densities and constitute a free carrier plasma, while in extrinsic semiconductors the donors or acceptors provide the fixed neutralising charge distribution for the free electron or hole charge densities, respectively.

Even though the positive and negative charges do not separate completely in space to give two separate free particles of opposite sign, the charge distribution *within* a neutral molecule may be

distorted to create a local molecular charge imbalance in which the "centres of gravity" or the first moments of the positive and negative charges, $\pm q$, become separated by a finite distance l, thus creating a *dipole* with a *dipole moment* $\mu = ql$. This concept of a dipole differs fundamentally from that of the plasma, in that the charges are not capable of being separated beyond the distance l which may be fixed or may related to the magnitude of the locally prevailing electric field, E.

The case of fixed l corresponds to situations in which the charge separation is caused by chemical interactions between *dissimilar* atoms constituting the molecule, e.g. HCl, where the different electron affinities of the atoms give rise to the appearance of net charges on each atom, in this instance $H^+ - Cl^-$. The situation leads to the creation of a *permanent dipole* whose dipole moment is dominated by the strong intramolecular interactions and is not seriously affected by external factors such as electric field and temperature. The dimensions of the dipole moment are $[\mu] = \text{C.m}$. A convenient non-SI unit frequently used in the literature to give numerical values of dipole moments is one Debye (D)

$$1\ \text{D} = (1/3) \cdot 10^{-29}\ \text{C.m}$$

which is equivalent to two charges of $\pm e$ of magnitude equal to the electronic charge separated by a distance of $2.08 \cdot 10^{-11}$ m.

The other type of dipole arises in atoms or molecules in which no significant charge separation occurs as a result of chemical interactions and in which the separation takes place between the outer valence electrons and the atomic cores. In these cases we assume that the dipole moment is proportional to the electric field according to the relation $\mu = \alpha E$, where α is the *polarisability* of the molecule.

With this definition, the dimension of polarisability is $[\alpha] = \text{m}^3$, sometimes quoted in more convenient units of $\text{Å}^3 = 10^{-30}\ \text{m}^3$. This induced dipole moment which is proportional to the electric field adds *vectorially* to any permanent dipole moment that a molecule may possess.

The formation of permanent dipoles and the polarisability of molecules are very intimately linked with the nature of the chemical bonds between atoms in molecules. In non-metallic solids these bonds may be *covalent* or *ionic*, corresponding, respectively, to the limit of negligible charge transfer and of complete charge transfer between atoms. Thus, a diatomic molecule consisting of identical

atoms, e.g. H_2 or N_2 must necessarily be covalently bonded since symmetry precludes any charge transfer. The corresponding situation arises in non-metallic elemental crystals, e.g. silicon or diamond, where the bonds are completely covalent and no charge transfer takes place between atoms. These bonds are strongly directional and their polarisability is naturally different according to the orientation of the polarising field with respect to the direction of the bonds – (see Section 2.3*c*).

Classic examples of completely ionic bonds are to be found in solids, such as the alkali halides NaCl, KCl, etc. where the charge transfer between the positive cations Na^+ and the negative anions Cl^- is complete, leading to the formation of completed valence shells in both species. The bonding forces in such solids are predominantly the Coulombic attraction between oppositely charged nearest neighbours, balanced by the strong repulsion as ions approach more closely than is allowed by their effective *ionic radii*

All compound solids consisting of dissimilar atoms, e.g. ZnS, SiO_2, BN, have partially covalent and partially ionic bonds, in which the charge transfer is finite but not sufficient to produce completed outer valence shells in the constituent ions. In addition to the induced polarisability of the individual ions, ionic solids are also polarisable by virtue of the relative shifts of the positive and negative sub-lattices from their equilibrium positions.

In addition to the electron-sharing covalent bonds and the electron-transferring ionic bonds, or some intermediate forms, non-metallic solids may exhibit a third form, the van der Waals bonds in which neither sharing not transfer of electrons takes place. These bonds are formed between atoms or molecules which are essentially chemically inert, such as the noble gas solids, e.g. solid argon, or molecular gas solids, e.g. solid hydrogen or nitrogen. In these solids a relatively weak interaction arises from the closely approaching electronic clouds. The resulting polarisability is essentially that of the constituent molecules themselves.

A very important class of solids exhibiting both covalent and van der Waals bonds are the *molecular solids* such as polymers and many organic solids in which large molecules, sometimes very large *macro-molecules*, are strongly covalently bonded internally and are then linked by the van der Waals forces between themselves. This strong disparity between the covalent and van der Waals bonds in molecular solids means that molecular motions are relatively easy, compared for example with wholly covalently bonded materials.

The process of cross-linking polymers amounts to replacing some of the weak van der Waals intermolecular bonds by strong covalent bonds. This evidently increases the mechanical rigidity of the polymer.

2.2 DIELECTRIC POLARISATION

The relative shift of the positive and negative charges in matter is referred to as *dielectric polarisation*. There exist many examples where finite polarisation is present even in the absence of an external field – for example the *permanent polarisation* in *ferroelectrics* and in *electrets*, but we shall not be concerned with those in detail in the present treatment. Our interest will be devoted entirely to the phenomenon of polarisation produced by an electric field, both through the orientation of any permanent dipoles and through the induced polarisation of individual atoms or ions.

We may write the general relation between the polarisation P and the field E as:

$$P = \varepsilon_0 \chi E + \text{higher terms in } E \tag{2.1}$$

The first term proportional to the field is of dominant importance in most systems, ε_0 is the *permittivity of free space* and χ is the *susceptibility*. The higher terms give rise to phenomena known as *hyperpolarisation* but we shall not be discussing them in detail, except in a few cases where non-linearities of the dielectric response arise in a particularly strong manner.

The polarisation P is directly related to the surface charge which appears on a polarised medium. Consider now a planar capacitor consisting of two parallel metallic electrodes at a distance w apart, as in Figure 2.1 and let the space between the electrodes be empty initially. If a voltage V is applied between the electrodes, this gives rise to a uniform field $E = V/w$, and charges $\pm Q_0 = \pm \varepsilon_0 E$ appear on the plates as a consequence of Gauss' law. Let now the space between the electrodes be filled with a material medium of susceptibility χ which becomes polarised in its own right, contributing a surface charge $Q_1 = P$, so that the total charge on the electrodes becomes the sum of the contribution of free space and of the contribution due to the material medium polarisation:

$$Q = Q_0 + Q_1 = \varepsilon_0 (1 + \chi) E \equiv \varepsilon(0) E$$

This equation defines the static permittivity $\varepsilon(0)$ of the material,

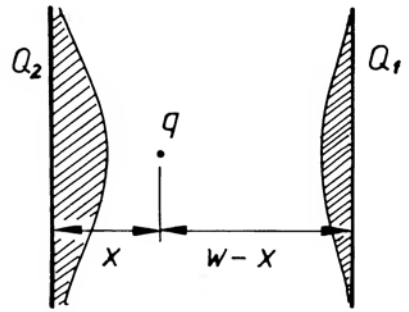


Figure 2.1 Schematic representation of the distributions of surface charges induced on two metallic electrodes at a distance w, due to the presence of a single point charge q at a distance x from one of the electrodes. The electrodes are assumed to extend to infinity in the plane normal to the plane of the paper.

in terms of the contributions of the free space and of the material medium itself. Later in the present work we shall extend the concept to the frequency dependent permittivity which will be related in an identical manner to the frequency dependent susceptibility of the materials in question.

Some authors favour the use of a different parameter, defined as the *dielectric modulus*

$$M = 1/\varepsilon = E/Q$$

which is valid also in the case of frequency-dependent behaviour. While it remains essentially a matter of taste which of these parameters one prefers, since they uniquely define the properties of the materials, we find that the permittivity is the more directly physically meaningful quantity, since in virtually all situations it is the field that constitutes the independent variable and the charge is the dependent variable, and not the other way round.

Although it is possible to devise experimental measurements which would give an indication of the value of P in the interior of a dielectric, this is not at all an easy and certainly not an accurate method of determining the dielectric polarisation. By far the most important experimental procedure for determining the polarisation below the microwave range of frequencies consists in the determination of the *electric current* resulting from a rate of change of polarisation with time,

$$i(t) = dP/dt \qquad (2.2)$$

If it is then desired to know the total polarisation, it is necessary to integrate the current over the entire range of time. This method of measurement rests on the unique relationship between the polarisation in the volume of the material and the charge Q induced on the surface. The principle is best explained with reference to Figure 2.1 which shows a pair of planar parallel electrodes at the same potential containing, initially, no matter in the space between them. Assuming now that a charge q is introduced at some point between these plates and also assuming that the area of the electrodes is large in comparison with the square of their separation w, it may be shown that the charges induced on the two electrodes are:

$$Q_1 = qx/w \qquad Q_2 = q(w-x)/w \tag{2.3}$$

so that the sum of these charges is equal to q. The proof of this may be obtained in a number of ways, one of which runs as follows (Scaife 1973). Assume that a planar *sheet* of charge of area density Q is introduced in the plane x between parallel plates extending to infinity in the plane y–z normal to x. We postulate that the charge densities induced uniformly over the entire area of the plates by the sheet at x are

$$Q_1 = bq \qquad Q_2 = (1-b)q$$

where the parameter b is as yet undetermined. This amounts to no more than the assumption of linear dependence between q and Q_1, Q_2, together with the statement that the sum of induced charges cannot exceed the primary charge. Since the field between the sheet of charge and the plates is given by Q_1/ε_0 to the right and by $-Q_2/\varepsilon_0$ to the left of the plane x, and since the potential of the two plates is the same, it follows that the parameter b has the value x/w. The next step of the argument rests on the *principle of superposition* which states that the charges induced at some boundary electrode under the action of a given distribution of charges in space represent the sum of the effects of point charges from which the distribution is composed. Since the sheet of charge corresponds to a uniform density of point charges, the induced charges corresponding to every single point charge are in the same ratio and this proves the expressions (2.3).

If we now assume an arbitrary *volume distribution* of point charges q_i at coordinates x_i, the aggregate effect of these on the plates is

$$\begin{aligned} Q_1 &= \sum_i q_i x_i/w = M_1/w \\ Q_2 &= \sum_i q_i(1 - x_i/w) = M_2/w \end{aligned} \tag{2.4}$$

This summation, which in the limit of a continuous distribution of charge $\rho(x)$ may be changed into an integration over the coordinate x, amounts to the statement that the total charges induced on the plates by a volume distribution of charge are equal to the respective *first moments* M of that distribution taken with respect to the opposite plate. We note that, since we are here asking for the total charges and not for their distribution on the plates in the plane y–z normal to x, the positions of the volume charges in the y–z plane are not relevant to this calculation.

If now the two plates are not at the same potential but have a potential difference V between them, creating in the absence of the volume charge a uniform field $E = V/w$, this field gives rise to induced charges $\pm\varepsilon_0 V/w$ and these are superimposed on any charges induced by the presence of a volume charge distribution.

It follows, therefore, that any change in the first moment of charge distribution with time results inevitably in a corresponding change of the induced charges and this change is immediately detectable as a current in the circuit connecting the two plates. This measurement provides the basis of the experimental determination of the first moment of the distribution.

Consider now the case of an overall neutral medium containing a distribution $+\rho^+(x)$ of positive space charge and a distribution $-\rho^-(x)$ of negative space charge, each amounting to the same total magnitude of charge Q:

$$Q = \int_0^w \rho^{\pm}(x)\,\mathrm{d}x$$

Each of these distributions has its own first moment $M^{\pm}$ defining its "centre of gravity":

$$s^{\pm} = M^{\pm}/Q = (1/Q)\int_0^w x\rho^{\pm}(x)\,\mathrm{d}x \qquad (2.5)$$

as shown in Figure 2.2. So long as the two distributions are exactly equal, there is no net charge on the plates, since the induced charges arising from the opposite charge distributions cancel one another exactly. If, however, due to an external field giving rise to a displacement of the positive and negative charge distributions in opposite directions, the centres of gravity s^+ and s^- should become displaced with respect to one another, then the charges induced on the plates will be

$$Q_1 = -Q_2 = (M^+ - M^-)/w = Q(s^+ - s^-)/w = Q\Delta s/w \qquad (2.6)$$

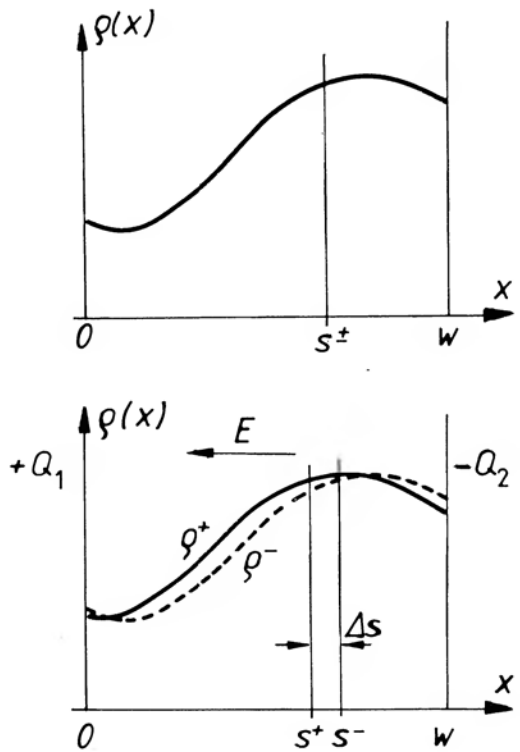


Figure 2.2 Diagram a) represents two arbitrary but equal distributions of positive and negative space charge, $\rho^+(x) = \rho^-(x)$, over a planar insulating region of thickness w. $s^{\pm}$ denotes the common centre of gravity of the distributions. In diagram b) the two distributions have been slightly displaced with respect to one another, as might be the case under the influence of an applied electric field E indicated by the arrow. Δs is the resulting displacement of the respective centres of gravity $s^{\pm}$. The resulting induced charges of opposite signs on the electrodes, $\pm Q$, are of equal magnitude.

where Δs is the extent of the displacement of the centres of gravity of the positive and negative charge distributions.

Now Q being the total charge contained in the respective positive and negative distributions, we note that $Q\Delta s$ is a *dipole moment*, which is exactly equal to the moment that would arise if two *point charges* $\pm Q$ were displaced by a distance Δs. We now assume that the plates have unit area in the y–z plane, so that Q_1 and Q_2 become charge densities per unit area, and we call these charge densities *the polarisation*. It then follows from eqn (2.6) that

polarisation is the dipole moment per unit volume and is equal to the charge per unit area on a boundary normal to the polarisation.

We conclude that a finite relative displacement of positive and negative charges produces a dipole moment throughout a given volume of space in which these charges are present and the effect of this dipole moment is to produce a charge density at the external boundaries, assumed to be normal to the displacement, which is equal to the dipole moment per unit volume. The dipole moment is a vector in the direction of the relative displacement of the two charges, and so is therefore the polarisation, although for our purposes it will be sufficient in the majority of cases to regard it as a scalar quantity, since the treatment will be restricted to one-dimensional problems. The convention is to represent the dipole moment as a vector directed from the negative to the positive charge, as shown in Figure 2.3.

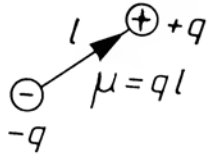


Figure 2.3 Schematic representation of a permanent molecular dipole consisting of two charges $\pm q$ at a distance l, giving the dipole moment $\mu = ql$. It is a convention that the dipole moment vector points from the negative to the positive charge.

In anisotropic crystalline media, or partially crystalline media such as oriented polymers, the dipole moment induced in the system is not necessarily oriented in the same direction as the electric field vector – the medium has in this case to be described by a tensor susceptibility (Nye 1957, Zheludev 1971).

Instead of thinking in terms of the dipole moment arising from the relative bodily displacements of two initially balancing charge distributions of opposite signs, we shall achieve the same end result if we consider the local rotations of a number of permanent dipoles or the creation of the corresponding density of induced dipoles; the dipole moment per unit volume is the summation of the individual dipole moments.

So long as there is no net volume space charge in the system, the two induced charge densities at the electrodes are the same in magnitude. This condition would not apply, for example, in an insulator into which a net space charge was injected from one electrode, as in space charge limited flow (Lampert 1965).

2.3 POLARISATION IN STATIC ELECTRIC FIELDS

We now propose to obtain the relationship between the polarisation and the electric field implied by eqn (2.1) and we shall do so in the first instance for the case of a static, i.e. time-independent field, leaving the practically more important case of time-varying fields to Section 2.5.

The basic physical fact to be borne in mind when discussing polarisation is that

> *polarisation arises from a finite displacement of charges in a steady electric field*

and this is to be contrasted with the complementary physical phenomenon of *electrical conduction* which is characterised by the fact that

> *conduction arises from a finite average velocity of motion of charges in a steady electric field.*

Our discussion in Section 2.1 clearly shows that dipoles of both the permanent and the induced type are examples of polarising species which are incapable of leading to a continuing conduction current in a static electric field, since the charges in question cannot be completely separated or dissociated under normal conditions. A dipole could only become dissociated in a field that would be sufficiently large to break the strong bonding forces of the neutral dipolar molecule and this is not normally possible in solids and liquids where various forms of electrical breakdown take place at much lower fields, 10^8–10^9 V/m.

As an example consider the fact that typical molecular bonding energies are of the order 1–10 eV, with typical spacing in molecular dipoles of the order of a few times 10^{-10} m. The field required to break the dipole would therefore amount to 10^{10}–10^{11} V/m.

We now give the elementary derivation of the expressions for the polarisation in specific cases of dipolar systems.

a) Orientational polarisation – freely floating dipoles

This is the classical type of polarisation originally treated by Debye (1945) and it refers to polar molecules which are freely floating in a dielectrically inert non-polar fluid. There are therefore no restoring forces tending to impose a preferred direction, only the randomising influence of thermal agitation.

Consider a set of N *non-interacting* dipoles, each of dipole moment μ, at a temperature T and in the presence of an external field E. The total energy of the dipole is made up of the thermal energy plus the electrostatic interaction energy:

$$W = kT - \mu E \cos\theta \tag{2.7}$$

where k is Boltzmann's constant and θ is the angle between the dipole and the field. The resulting torque is:

$$\Xi(\theta) = -\partial W/\partial\theta = -\mu E \sin\theta \tag{2.8}$$

The distribution function of dipoles throughout the spherical angle is isotropic in the absence of the field, while in the presence of the field we introduce the Boltzmann factor $\exp(-W/kT)$, using eqn (2.7):

$$g(\theta) = (N/4\pi)\exp(\mu E\cos\theta/kT) \tag{2.9}$$

where the factor 4π is introduced for normalisation. We now make use of the *low field approximation*, $\mu E \ll kT$, which will be justified later on. In this limit we may expand the exponential term and obtain

$$g(\theta) = (N/4\pi)\{1 + (\mu E/kT)\cos\theta\} \tag{2.10}$$

in which the first term corresponds to the isotropic distribution in the absence of a field and the second term represents the small perturbation favouring small angles θ.

In order to obtain the average dipole moment we take the average of $\mu\cos\theta$ over the entire solid angle:

$$\bar{\mu} = \mu\frac{\int g(\theta)\sin\theta\cos\theta\,\mathrm{d}\theta}{\int g(\theta)\sin\theta\,\mathrm{d}\theta} \tag{2.11}$$

the integration extending from 0 to 2π. Substituting $a = \mu E/kT$ and changing the variable $\cos\theta = u$, we find using the exact expression (2.9):

$$\bar{\mu} = \mu\frac{\int_{-1}^{1} u\exp(au)\,\mathrm{d}u}{\int_{-1}^{1}\exp(au)\,\mathrm{d}u} \tag{2.12}$$

which can be evaluated as follows:

$$\bar{\mu} = \mu(\coth a - 1/a) = \mu L(a) \tag{2.13}$$

where $L(a)$ is the *Langevin function* which has the limiting approximate forms:

$$L(a) = a/3 \quad \text{for } a \ll 1 \tag{2.14}$$

and

$$L(a) = 1 - 1/a \quad \text{for } a \gg 1 \tag{2.15}$$

This function is sketched in Figure 2.4 and shows the linear dependence on a and therefore on the field E in the limit of "low" fields, when the average dipole moment becomes

$$\bar{\mu} = \mu^2 E/3kT \qquad \text{for } \mu E \ll kT \tag{2.16}$$

This means that the dielectric susceptibility defined by eqn (2.1) is proportional to the inverse temperature, which may be explained physically as arising from the increasing randomising influence of thermal vibrations at higher temperatures opposing the alignment of the dipoles by the field. This prediction should be very easily verifiable experimentally and should provide a means of distinguishing between this and other mechanisms of polarisation.

Figure 2.4 shows clearly the tendency to saturation arising at sufficiently high fields and illustrates the principle that any dipolar mechanism must, in this limit, reach a complete alignment of dipoles. The corresponding higher terms in eqn (2.1) would therefore be negative, although physical situations can arise where they are positive. The most important example of the situation where complete alignment of dipoles takes place is in ferroelectrics, in which this is due to the intermolecular forces and not to the field (Burfoot 1967).

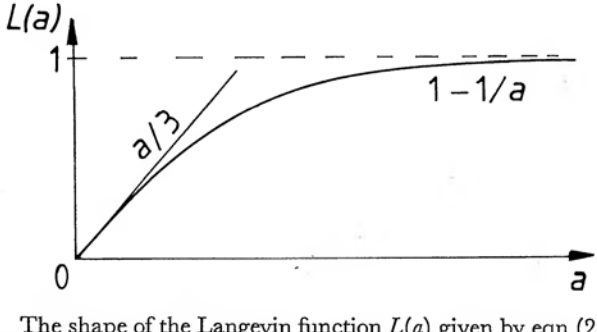


Figure 2.4 The shape of the Langevin function $L(a)$ given by eqn (2.13), showing the initial slope equal to $a/3$ and saturation for large values of the argument a.

In all other situations, where the equilibrium distribution of dipole orientations is random because there is no ordering force in the absence of an electric field, the effect of a field is to cause partial orientation according to eqn (2.13), but the magnitude of the argument of the Langevin function, which is given by the ratio of the electrical energy of the dipole to the thermal energy, is very small. A dipole moment of 1 Debye placed in a field of 10^8 V/m, which approaches the magnitude of a typical breakdown field in solids, gives $a = 0.083$ at room temperature. This shows that even in this extremely high field the alignment of ordinary molecular dipoles is only small. However, much larger values of a may be attained with certain molecular dipoles, which have very high dipole moments and correspondingly high moments of inertia, cf. Figure 6.4.

The dielectric susceptibility of our model of N *non-interacting* dipoles becomes therefore, in the limit of weak orientation,

$$\chi(0) = \frac{N\mu^2}{3\varepsilon_0 kT} \tag{2.17}$$

where the notation $\chi(0)$ signifies the zero-frequency limit, i.e. the static conditions. We note that in our convention the susceptibility χ is a dimensionless parameter and it is also a scalar in the case of isotropic media.

The practically more important case of interacting dipoles will be discussed in Section 2.4.

b) Molecular polarisability – induced dipole moment

We have already described the principle of induced electronic polarisation which essentially reflects the tendency for the outer electronic clouds to be displaced with respect to the internal atomic cores by external fields. Alternatively, in the case of ionic polarisation, we are envisaging the mutual displacements of the positive and negative sublattices, resulting in a net induced dipole moment. The characteristic feature of both these types of polarisation is their relative independence of temperature, since the forces opposing the displacements are not themselves strongly temperature dependent.

The polarisation arising from the distortion of the outer electronic shells by an external electric field is always present in atoms and molecules, whether permanent dipoles exist or not, and whether they are being oriented as freely floating or remain "rigid". Thus induced molecular polarisability may be assumed, in the first

approximation, to be simply additive to the orientational polarisation and we may therefore write:

$$\chi(0) = \frac{N_d \mu^2}{3\varepsilon_0 kT} + N_m \alpha \qquad (2.18)$$

where N_d and N_m denote, respectively, the number densities of permanent dipoles and of polarisable molecules which may, of course, be the same.

In ordered systems, such as crystals, the vectors of molecular and orientational polarisation may not coincide and they may have to be added vectorially.

c) Orders of magnitude of dipole moments and polarisabilities

It is instructive to look at the orders of magnitude of the dielectric parameters and to relate these to the structure of the molecules in question. We begin with the polarisabilities of isolated atoms which are shown in Figure 2.5 on the basis of numerical data from a collection by Teachout and Pack (1971), where we present a log–log plot of the polarisability against the atomic number. In this somewhat unfamiliar representation we see the remarkable regularity of the distribution of points. The lowest polarisabilities – of the order of 0.3 to 3 Å^3 – correspond to the noble gases with their completely filled electronic outer shells effectively screening the nuclei from the influence of external fields. At the other extreme, the Group I elements, alkali halides, have the highest polarisabilities of the order of 30 Å^3, presumably due to the highly polarisable single electron in the outermost orbit. The intermediate elements follow in regular sequences between these extremes, with relatively few irregularities, such as Al and Si.

Elemental ions have polarisabilities corresponding closely to the isoelectronic elements, i.e. Na^+ with $\alpha = 0.2$ (all values will now be quoted in Å^3) resembles Ne with 0.3, K^+ with 1 is close to Argon, and so on. Li^+ with 0.04 is almost a factor of 100 down on Li and 5 down on He. Negative ions have correspondingly higher polarisabilities.

These polarisabilities should correspond in some measure to the actual polarisabilities of ions in solids.

By contrast with isolated ions or atoms, the polarisabilities of molecules are much larger, in view of the much larger dimensions of their electronic clouds which are therefore easily "deflected" by external fields. Typical bond polarisabilities are shown in Table

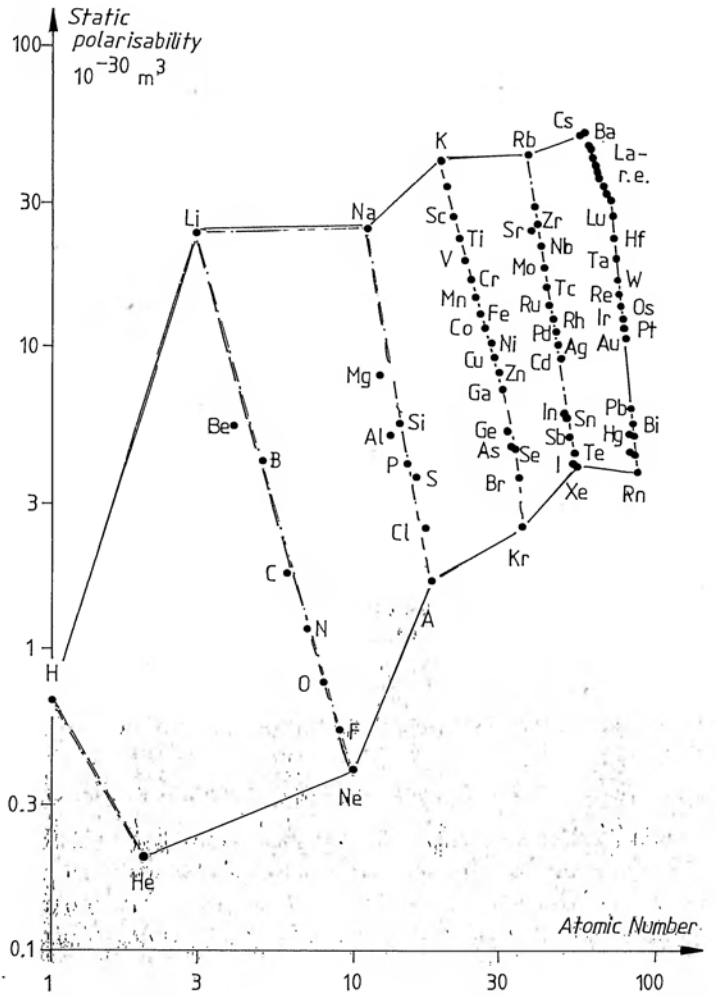


Figure 2.5 The logarithm of the static atomic polarisabilities plotted against the logarithm of the atomic number, showing the periodicity and regularity of the relationship. Data from Teachout and Pack (1971).

2.1, which gives separate data for longitudinal and transverse polarisabilities and also the average value.

Taking next the polarisabilities of complete molecules, these reflect closely the geometrical shape of the molecules and we have to specify three values corresponding to the three principal axes of

TABLE 2.1
Polarisabilities of chemical bonds parallel and normal to the bond axis and also the mean value for all three directions in space, $\alpha_m = (\alpha_\parallel + 2\alpha_\perp)/3$, in units of Å^3. From Denbigh (1940).

Bond	$\alpha_\parallel$	$\alpha_\perp$	α_m	Comments
H—H	0.93	0.72	0.79	
N—H	0.58	0.84	0.75	NH_3
C—H	0.79	0.58	0.65	aliphatic
C—Cl	3.67	2.08	2.61	
C—Br	5.04	2.88	3.60	
C—C	1.88	0.02	0.64	aliphatic
C—C	2.25	0.48	1.07	aromatic
C=C	2.86	1.06	1.66	
C=O	2.00	0.75	1.20	carbonyl

TABLE 2.2
Polarisabilities of molecules along three principal axes of symmetry in units of Å^3.

Molecule	α_1	α_2	α_3
H_2	0.934	0.718	0.718
O_2	2.32	1.21	1.21
N_2O	4.86	2.07	2.07
CCl_4	10.5	10.5	10.5
$CHCl_3$	6.68	9.01	9.01
HCl	3.13	2.39	2.39

TABLE 2.3
Dipole moments of chemical bonds and of molecules expressed in Debye units.

Bond	$\mu.3.10^{29}$ (Cm)
C—H	0.4
O—H	1.51
C—C	0
C—O	0.74
C=O	2.3
Molecule	
H_2O	1.84
CO_2	0
HCl	1.08

symmetry. Table 2.2 gives a selection of data from which we note the clear indication of molecular symmetry.

Passing now to the permanent dipole moment of bonds and molecules, we list in Table 2.3 some typical values expressed in units of Debye, and we note that all symmetric molecules, such as O=C=O or CCl_4, or symmetric bonds between identical atoms, must have zero dipole moments.

It is instructive to make a comparison of the relative contributions of the induced and orientational polarisabilities. Take the HCl molecule with a mean induced polarisability $\alpha_m = 2.64\ 10^{-30}\ m^3$ and a permanent dipole moment of $1.08\ D = 3.60\ 10^{-30}\ Cm$. In an external electric field of $10^7\ V/m$, which is a fairly high field, the resulting induced dipole moment would be $\mu_{ind} = \varepsilon_0 \alpha E = 2.3 \cdot 10^{-34}\ Cm$, which is four orders of magnitude smaller than the permanent moment of a single molecule.

However, the mean induced moment is not affected by the rotational motions of the molecules, because the electronic response is sufficiently rapid to follow the direction of the field regardless of the molecular rotations. The *net* orientational dipole moment per molecule is $\mu^2 E/(3kT)$, so that orientational polarisability is

$$\alpha_{or} = \mu^2/kT\varepsilon_0 = 1.2\ 10^{-28}\ m^3$$

which is 50 times larger than the induced polarisability (at 300 K).

Very similar considerations would apply to a polar bond, such as C = 0, the induced dipole moment would be much smaller than the permanent moment, but once again the ability of the bond moments to follow the field is severely restricted in a solid, making the effect of the induced polarisation correspondingly stronger.

Having made this comparison between the induced and permanent dipole moments we now propose to calculate the ionic polarisability resulting from the relative displacements of the sub-lattice of the positive and negative ions in the crystal. The electric force per ion in a field E is eE. With an interatomic spacing a in an assumed cubic lattice, the force per unit area is $F = eE/a^2$. If the displacement resulting from the force is Δa, the dipole moment becomes $\mu = e\Delta a$. The relation between strain and stress is $\Delta a = asF$, where s is the elastic compliance. We thus obtain for the dipole moment in a field E, $\mu = se^2E/a$, which gives a polarisability

$$\alpha = sq^2/a\varepsilon_0$$

Taking numerical values for the electronic charge, the interatomic spacing of $2\ 10^{-10}$ m, and a value of compliance appropriate to NaCl $s = 2\ 10^{-11}$ m^2/N, we obtain a polarisability for Na Cl $\alpha = 3\ 10^{-28}$ m^3 which is some two orders of magnitude higher than typical bond polarisabilities in Table 2.1.

d) Polarisation by hopping charge carriers

Induced and permanent dipoles represent one extreme form of polarisable species in which the two component charge systems are very intimately linked together, even to the point of partial interpenetration, and cannot normally be dissociated into separate charges without the application of extreme constraints. At the other extreme we may place completely free charges such as electrons in metals, or electrons and holes in crystalline semiconductors, where these charges are free to move without constraint throughout a given volume of the solid. We shall return later to a more detailed discussion of these processes which give rise to direct current conduction but which may, under certain boundary conditions, also lead to polarisation. At this moment, however, we shall introduce the concept of a class of charged species intermediate between dipoles and free carriers.

These are *hopping* charge carriers which are characterised by the fact that they spend most of the time in *localised* sites where they are subject only to relatively very small thermal vibrations, but occasionally they make a big jump or hopping transition to some neighbouring localised sites which may be one or many atomic spacings away. The concept of hopping movement has been very familiar for a long time in connection with ionic conduction, since ions move essentially only by hopping, whether by the interstitial or vacancy mechanism. A relatively more recent development is the extension of this concept to electronic charges which has found particular application in amorphous and disordered non-metallic solids, such as glasses and amorphous semiconductors (Mott and Davis 1979, Jonscher and Hill 1975). In such strongly disordered solids the normal concept of band conduction by free charge carriers do not apply and we find instead that electrons become localised and can only move by hopping between localised sites. If these sites form a continuous connected network the charges may be capable of traversing the entire physical dimensions of the sample and therefore give rise to direct current conduction that would be indistinguishable from free carrier conduction were it not for a much lower mobility – typically by many orders of magnitude – than for the corresponding free band conduction. It is inevitable,

however, that there exist "easier" and "more difficult" hops and that charges execute many reciprocating transitions between pairs of sites linked by easy transitions before making "forward" jumps to other sites involving more difficult transitions.

The probability of a hopping transition may be determined by the combined effect of the distance between the two sites and the potential barrier that has to be overcome – the transition may be a thermally assisted hop over the potential barrier or a tunnelling transition through the barrier – the latter requiring negligible activation energy. An elementary two-site situation is represented in Figure 2.6 in the form of a potential double well separated by an internal barrier and with infinitely high walls outside, indicating negligible probability of escape for the particle. We assume that the double well accommodates one charge which may occupy either of the two wells, the compensating charge of opposite sign is situated in the neighbourhood, possibly between the two wells and is

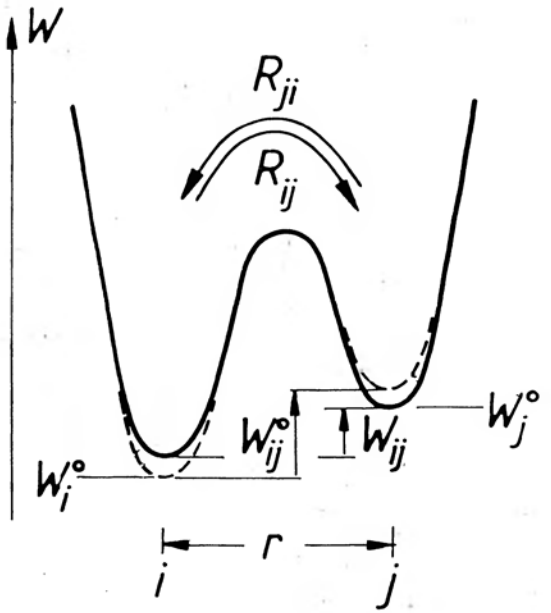


Figure 2.6 A double potential well representing the potential energy of a particle which may take up one of two "preferred" positions i and j, with the respective energies W_i^0 and W_j^0. The distance between the two sites projected in the direction of the electric field is r. In the presence of an external electric field the equilibrium energies are perturbed to W_i and W_j and the equilibrium energy difference W_{ij}^0 becomes W_{ij}. R_{ij} and R_{ji} are the transition rates between the two wells.

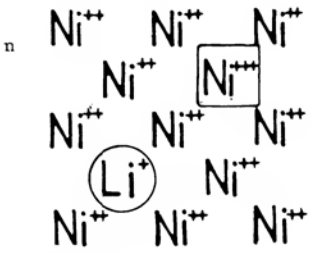


Figure 2.7 The effect of introducing a monovalent Li^+ ion in place of a divalent Ni^{++} ion in a NiO lattice. The Li^+ ion looks like a *negatively* charged site and charge balance is restored by removing an extra electron from one of the neighbouring Ni ions, thereby making it look like a singly *positively* charged site which may move in the neighbourhood of the immobile Li ion. The oxygen ions are not shown.

assumed to be rigidly fixed. A physical model corresponding to this situation may be given by a substitutional monovalent cation in a divalent cationic lattice, e.g. Li^+ in a NiO lattice, which causes one of the neighbouring Ni^{++} ions to recharge to Ni^{+++} for neutrality, Figure 2.7. However, the extra positive charge may occupy any one of several neighbouring sites and may be considered to execute hopping transitions between them. A different example is given by a non-bridging oxygen ion in a glass network which may occupy one of two or more energetically preferred positions in the lattice.

Returning to the double well diagram of Figure 2.6, we may define the time-averaged probabilities of occupation in equilibrium, f_i^0 and f_j^0 in terms of the energies W_i^0 and W_j^0 of the two wells:

$$f_j^0 / f_i^0 = \exp(W_{ij}^0 / kT) \tag{2.19}$$

where we have defined $W_{ij} = W_j + W_i$ and the superscripts ^{0} denote the equilibrium values in the absence of an externally applied electric field. The condition that the particle with a charge q should certainly be in one of the two wells is

$$f_i + f_j = 1 \tag{2.20}$$

under all conditions, i.e. even in the presence of an electric field.

In the presence of an external electric field E, the separation of the energy wells becomes

$$W_{ij} = W_{ij}^0 + qEr \tag{2.21}$$

where r is the projection of the vector connecting the two wells in space along the direction of the electric field.

This gives rise to a redistribution of the occupation probabilities between the two wells and we write:

$$f_i = f_i^0 - f'; \qquad f_j = f_j^0 + f' \tag{2.22}$$

and we make the assumption that in the steady state the actual probabilities f_i and f_j are governed by the same Boltzmann equation (2.19) with the new energy W_{ij} given by (2.21).

Writing for brevity, by analogy with (2.12), $qEr/kT = a$, we obtain after simple calculations

$$f' = f_i^0 f_j^0 \frac{e^a - 1}{1 + f_j(e^a - 1)} \simeq f_i^0 f_j^0 a \tag{2.23}$$

the last approximate equality being valid for small values of $a \ll 1$. In the presence of N identical *non-interacting* wells the total polarisation is given by $P = Nqrf'$ which leads to the static susceptibility

$$\chi_0 = \frac{q^2 r^2}{3kT\varepsilon_0} \overline{f_i^0 f_j^0} \tag{2.24}$$

where the factor 3 in the denominator comes from averaging the projections of the randonly oriented vectors r on the direction of the field and the bar denotes averaging over the factors $f_i^0 f_j^0$ if these are different for the different wells.

Bearing in mind the fact that qr may be regarded as the dipole moment of the hopping charge, we note the similarity of this expression to the dipole orientational polarisation given by eqn (2.17), the physical difference being that in the latter the dipole changes its orientation smoothly by free rotation in the fluid medium, while the orientation of the hopping charges is determined by the spatial dispositions of the allowed localised sites favoured by the structural properties of the medium and represented by the potential wells. The distribution of the hopping charges between the allowed sites is influenced by the applied field. We note that the factor $f_i^0 f_j^0$ would become 1/4 if all the initial levels were at the same energy.

It is important to note that the transition of a charge q from site i to site j is physically and mathematically indistinguishable from the corresponding rotation of a dipole through the angle π. We shall also find their time-dependent behaviour to be identical.

2.4 EFFECT OF PARTICLE INTERACTIONS

The treatment described up to the present moment specifically assumed that the polarisable particles and the dipoles present in the system may be treated as being entirely non-interacting, where the effect of N particles per unit volume may be equated to the sum over all particles of their individual contributions that would have been obtained had each of them been present in isolation from all others. It is well known, however, that this approximation is completely inadequate in describing the behaviour of assemblies, especially in the case of condensed matter where inter-particle distances are very small and interactions between them cannot be neglected. We are dealing, in effect, with *many-body interactions*, or with *interactive many-body systems*, and this requires very special techniques which were not well-developed until comparatively recently. There have arisen, therefore, over the years a number of approximate approaches to the treatment of the static polarisability which attempted, with varying degrees of success, to take account of many-body interactions.

Historically the first and, to this day fundamentally the only existing one, was the approach adopted independently by Mossotti in 1850 and by Clausius in 1879 which leads to the so-called Clausius–Mossotti formula. In essence, this approach and most of the subsequent ones, tries to represent the "reaction" of the neighbouring dipoles on the given dipole by a suitable "internal field". The result of this procedure is to modify the expression (2.18) for the static susceptibility $\chi(0)$ to the following form:

$$\frac{\chi(0)}{\chi(0)+3}=\frac{\varepsilon(0)-\varepsilon_0}{\varepsilon(0)+2\varepsilon_0}=\frac{N_d\mu^2}{9\varepsilon_0 kT}+\frac{N_m\alpha}{3} \tag{2.25}$$

We note first that in the limit of very dilute systems in which the susceptibility is much smaller than unity, this expression becomes identical with (2.18). Secondly, however, we note that as $\chi_0 \rightarrow \infty$ the left-hand side becomes equal to 3, so that the value of the parameter $(N\mu^2/3\varepsilon_0 kT)=3$ corresponds to a "dielectric catastrophe", where the cooperative interaction causes the susceptibility to go to infinity. Clearly this is not a physically plausible result, it resembles in a certain way the ferroelectric behaviour and it suggests that the choice of the internal field has been physically incorrect. Although this expression was never intended to be applied to the condensed state of matter, with its much higher particle densities than in a gaseous state, its application was nevertheless extended

to the limit of highly compressed gases and eventually to liquids, especially with certain modifications introduced by later workers. Debye noted in 1912 that the *induced* component of the polarisability corresponding to the second term in eqn (2.18) satisfies the Lorentz–Lorenz condition known since 1880 for the refractive index, in view of the rapid response time of this polarisability, which implies that the response at optical frequencies is dominated by induced polarisation and the contribution of the orientational processes is negligible. Under those conditions, Maxwell's equations state that the relative dielectric permittivity may be equated with the square of the refractive index, so that the Clausius–Mossotti relation becomes:

$$\frac{n^2 - 1}{n^2 + 2} = \frac{N_m \alpha}{3} \qquad \text{Lorentz–Lorenz}$$

so that by setting $n^2 = \varepsilon_\infty / \varepsilon_0$, the original Clausius–Mossotti equation may be written in the following form, with the contribution of the orientational polarisation of the permanent dipoles set out separately:

$$\frac{\varepsilon(0) - \varepsilon_0}{\varepsilon(0) + 2\varepsilon_0} - \frac{\varepsilon_\infty - \varepsilon_0}{\varepsilon_\infty + 2\varepsilon_0} = \frac{N_d \mu^2}{9\varepsilon_0 kT} \qquad \text{Debye}$$

This was further modified by Onsager (1936) who refined the local field argument and obtained the relation:

$$\frac{\varepsilon(0) - \varepsilon_0}{\varepsilon(0) + 2\varepsilon_0} - \frac{\varepsilon_\infty - \varepsilon_0}{\varepsilon_\infty + 2\varepsilon_0} = \frac{3\varepsilon(0)\,(\varepsilon_\infty + 2\varepsilon_0)}{[2\varepsilon(0) + \varepsilon_\infty][\varepsilon(0) + 2\varepsilon_0]} \frac{N_d \mu^2}{9\varepsilon_0 kT}$$

Onsager

Yet further refinements were introduced by Kirkwood (1939).

We do not propose to enter into any detailed discussions of these various formalisms which are treated in considerable depth in standard texts on dielectric theory referred to in Chapter 1. While they may be applicable to highly purified liquids in which charge carriers make a negligible contribution to dielectric polarisation, it is an established fact that most solid dielectrics do not obey any such expressions at sufficiently low frequencies on account of an entirely different phenomenon which will be discussed later in this book under the heading of low-frequency dispersion. This means

that as the temperature is increased – and this may mean room temperature in many cases – the low-frequency permittivity does not obey the $1/T$ law demanded by the orientational component of the Debye and other expressions, but instead shows a steady *rise* with temperature which must be ascribable to the charge carriers present in these materials – mostly ions, but possibly also electrons. This renders any meaningful measurement of the low-frequency dielectric permittivity very difficult so that any comparisons with theory become equally doubtful. An impression of the extent of this effect may be obtained by looking at the extensive tables of permittivity data compiled by Westphal and Sils (1972).

It is too early to reach firm conclusions with regard to the extent of validity of the local field approach. This approach may be regarded as a means of taking into account the interactions of other dipoles by modifying the local field acting on a given dipole in question, so that, using the polarisability of an *isolated* atom, one obtains the correct value of susceptibility for the ensemble. One might suspect that this is not the best approach and that one should regard the external field as producing the same *additional* contribution to the instantaneous local field configuration, but that the appropriate polarisability or dipole moment should be that for the particle *as part of an ensemble* which would naturally be different from that of an isolated particle.

From the point of view of the present treatment this argument does not figure very prominently for the reasons mentioned above.

2.5 TIME-DEPENDENT DIELECTRIC RESPONSE

The static response of dielectric systems to a steady electric field represents only one facet of the complete problem and, for practical purposes, a relatively insignificant one. Much more important experimentally, technologically and theoretically, is the time-dependent response to time-varying electric fields. The experimental significance stems from the availability of a vastly larger volume of data – in addition to the usual variables of temperature, composition, pressure, etc. we now have the time, or the corresponding frequency range which in many cases covers the range of between four and ten orders of magnitude. This, as we shall see, offers an excellent opportunity to compare the data with theoretical expressions and to see much more clearly than was the case with the static response, if there is agreement between them. The principal object

of the present monograph is the discussion of this time- and frequency-dependent response and we shall see later that the exploitation of the available experimental data has led us to a significant theoretical advance on the whole subject of many-body interactions in solids.

The technical significance of the time-dependent response is evident if we envisage the fact that most electrical applications of dielectrics involve the use of step-funtion, delta-function or sinusoidally variable electric fields, in comparison with which the purely static conditions represent only a limiting situation.

The most obvious physical reason for the time-dependence of the dielectric response is the inevitable "inertia" of all physical processes – no material system is capable of following arbitrarily rapidly varying driving forces. Whether this inertia is due to the mechanical inertia of the masses that have to be transferred from one position or orientation to another, or whether it is the outcome of more general rate processes, the net result is that the time dependent polarisation $P(t)$ is not the same function as the time-dependent driving field $E(t)$.

By contrast with *material* systems which necessarily exhibit delayed responses, the response of free space is instantaneous and therefore the induced charge $\varepsilon_0 E$ arising from the response of free space does follow the field instantaneously. If we take a system consisting of two plane parallel electrodes at a distance w, with the space between them filled with a dielectric material, and if we apply a time-dependent voltage $V(t)$ to the electrodes, giving rise to a *spatially uniform* field $E(t) = V(t)/w$, then the charges induced at the electrodes will be given by the sum of an instantaneous or "prompt" free space contribution and the delayed material polarisation:

$$D(t) = \varepsilon_0 E + P(t) \tag{2.26}$$

Here $D(t)$ denotes the *dielectric induction*, sometimes called *dielectric displacement*, and it gives the total charge density induced at the electrodes.

In order to put the analysis of the time dependence of the dielectric response on a proper mathematical basis we have to define the *dielectric response function* $f(t)$ which characterises the response of the dielectric medium to specified electric excitations. Quite generally, we may define three fundamental time-dependences of the exciting electric field:

the delta function, $\delta(t)$
the step function, $1(t)$, which is the integral of the former, and
the harmonic function, $\sin \omega t$ or $\cos \omega t$.

We may note here that whereas the unit step function and the harmonic functions are *dimensionless*, the delta function has to have the dimension of time, since it represents the mathematical limit of the product of a physical constraint – in this case the electric field, acting over a period of time Δt, such that the product $E\Delta t = \text{const}$ as $\Delta t \to 0$.

Any arbitrary time-dependence may be synthesised from delta-functions, step-functions, or by way of Fourier integral transformation from harmonic functions.

We now define the response function in terms of the time dependence of polarisation under delta-function excitation of strength $(E\Delta t)$

$$P(t) = \varepsilon_0 (E\Delta t) f(t) \tag{2.27}$$

In view of the physical significance of the response function, we conclude immediately that *causality* demands that there should be no reaction before action and therefore

$$f(t) \equiv 0 \quad \text{for } t < 0 \tag{2.28}$$

Equally, since we are only considering systems in which the polarisation is induced by the applied electric field, i.e. there is no *permanent* or *persistent* polarisation, we require that

$$\lim_{t \to \infty} f(t) = 0 \tag{2.29}$$

We shall later find one other condition, namely that the integral $\int_0^\infty f(t)\, \mathrm{d}t$ should be finite.

In order to proceed with the synthesis of the total time-dependent polarisation from a series of delta-function excitations corresponding to an arbitrary time-dependent field, $E(t)$, we must make the fundamental assumption of the validity of

> *the principle of superposition – that the response to consecutive elementary excitations is the sum of the responses to the individual excitations.*

This principle presupposes that the response of the dielectric system is *linear* in the exciting amplitude, which is implicit in eqn (2.27) and that this is extended to the time-dependent response.

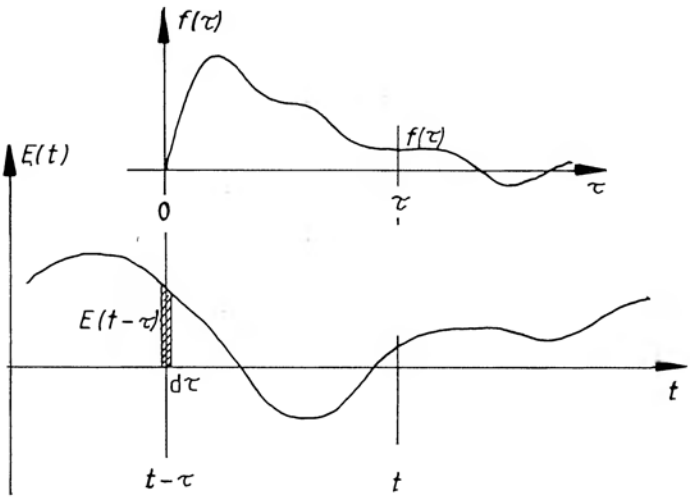


Figure 2.8 The principle of the derivation of the convolution integral (2.30), in terms of the summation of contributions of "delta-function-like" impulses $E(t-\tau)\mathrm{d}\tau$ at times τ before the instant t at which the observation is made. $f(\tau)$ is an arbitrary response function and $E(t)$ an arbitrary time-dependent electric field.

The derivation of the time-dependent polarisation proceeds now as follows. With reference to Figure 2.8, let $f(t)$ be a general response function and let the time-dependent field $E(t)$ be considered as a series of "delta functions" of strengths $E(t)\mathrm{d}t$, then response $P(t)$ at time t is given as the summation of the responses $f(\tau)\ E(t-\tau)\mathrm{d}\tau$ integrated in the time variable τ "backwards" to infinity or as far as the exciting signal extends. We therefore obtain the fundamental relationship:

$$P(t) = \varepsilon_0 \int_0^{\infty} f(\tau)E(t-\tau)\mathrm{f}\tau \qquad (2.30)$$

Now, in view of the property (2.28) we may extend the lower limit of integration down to $-\infty$ and this brings the integral into the form of *convolution or Faltung integral* of the functions $f(t)$ and $E(t)$. The physical sense of this integral is that the dielectric system retains the "memory" of its past history and this may extend in practice to times as long as hours, days or even longer.

The significance of the convolution integral (2.30) will be made clearer by considering the effect of specific waveforms of the applied field $E(t)$, particularly when they represent once-off transients. The

most elementary of them is the step function, $E_0\,1(t)$, defined by:

$$E(t) = 0 \quad \text{for} \quad t < 0$$
$$E(t) = E_0 \quad \text{for} \quad t > 0$$

In this case the integral of eqn (2.30) becomes:

$$P(t) = \varepsilon_0 E_0 \int_0^t f(\tau)\mathrm{d}\tau \tag{2.31}$$

The dielectric induction is given by

$$D(t) = \varepsilon_0 E_0 \left\{ 1(t) + \int_0^t f(\tau)\mathrm{d}\tau \right\} \tag{2.32}$$

and the resulting current flowing in the system is:

$$\begin{aligned} i_c(t) &= \mathrm{d}D(t)/\mathrm{d}t + \sigma_0 E_0 = \varepsilon_0 \mathrm{d}E(t)/\mathrm{d}t + \mathrm{d}P(t)/\mathrm{d}t + \sigma_0 E_0 \\ &= \varepsilon_0 E_0 \{\delta(t) + f(t)\} + \sigma_0 E_0 \end{aligned} \tag{2.33}$$

The delta function is due to the instantaneous response of the part of the total charge which is due to the step-function electric field itself, in other words it represents the response of "free space" contained in the volume of the dielectric material. The polarisation $P(t)$, by contrast, arises from the response of the material medium and it cannot follow the step-function field.

The term $\sigma_0 E_0$ on the right is included to take care of any direct current conductivity σ_0 of the system, which would give rise to a steady current at infinite time. The time response given by eqn (2.33) is normally dominated by the function $f(t)$ which is seen as the time-dependence of the polarising current.

This provides an important means of determining the function $f(t)$ experimentally as the response of the dielectric to the step-function charging field. This approach is much preferred to the possible alternative of measuring the polarisation response to a delta-function exciting field.

It is important to have a clear idea of the nature of this polarising current – it arises from the tendency of the polarising species in the material to respond *in a delayed manner* to the exciting field, and in this sense this current characterises the most important property of the dielectric system. On the other hand, this current has nothing to do with the mechanism of dc conductivity which is characterised by the parameter σ_0 and it is completely wrong to regard the term

in { } in eqn (2.33) as a time-dependent conductivity or, as is sometimes done in engineering textbooks, as the reciprocal of a time-dependent resistivity – to do so confuses two completely different processes:

> *the polarisation current* $\mathrm{d}P(t)/\mathrm{d}t$ *characterises the adjustment of the polarising species to a step function field and it must go to zero at infinitely long times – no charges may leave the dielectric system or enter it from the outside as a result of this process;*
>
> *the steady conduction current, or direct current (d.c) arises from continuous movement of "free" charges across the dielectric material from one electrode to the other and this current does not change in any way the "centre of gravity" of the charge distribution in the system.*

An immediate consequence of eqn (2.31) is that the polarisation after an infinitely long charging time, i.e. under a steady electric field E_0 is given by

$$P(\infty) = P_0 = \varepsilon_0 E_0 \int_0^\infty f(t)\,\mathrm{d}t = \varepsilon_0 \chi(0) E_0 \qquad (2.34)$$

according to the definition of the steady state susceptibility. This result shows that the integral of the function $f(t)$ must be finite since the steady state polarisation must remain finite.

Having reached the steady state polarisation P_0 after an infinitely long (in practice "sufficiently long") time, we may now reduce the field abruptly to zero and observe the *depolarisation current* $i_d(t)$ resulting from the return of the partially oriented polarising species back to the random arrangement which they have in equilibrium in the absence of field. Applying the principle of superposition, we may regard the *downward going* step function as the superposition of a steady value E_0 extending from $-\infty$ to $+\infty$ in time and of a negative going step function starting at time $t = 0$. Since the charging current at infinite time consists of the dc value given in eqn (2.33), the discharging or depolarisation current is given by the negative polarisation current, including the negative going delta function, if this can be resolved experimentally.

The depolarisation current does not include the direct current and it represents therefore a more convenient form of measurement of the function $f(t)$ than does the charging current. The principle of superposition implies, however, that if the polarisation current had

not been allowed to reach zero because of an insufficiently long charging time then the discharge current will show the effect of the continuing polarisation superimposed on the depolarisation current. This has serious practical implications and is dealt with as an example in Section 6.1 where we discuss the dielectric response to a square-wave field excitation.

2.6 FREQUENCY–DOMAIN RESPONSE

We shall see later that many aspects of dielectric response, especially in the presence of non-linear processes, are more easily understood in terms of the response to time-dependent signals. However, there exists a very powerful alternative approach which offers very considerable theoretical and practical advantages, provided that we are dealing with linear systems, and this is the determination of the response to harmonic excitation, i.e. sinusoidal waves. The theoretical advantage of this approach will be seen from the analysis that follows. The practical advantage lies in the fact that most engineering applications of dielectrics involve response to harmonic signals of definite frequencies and also, most importantly, very powerful measuring techniques exist which enable us to make measurements with such signals, using frequency as a parameter, with very high accuracy. The principal reason for this is that the measurement of the response at a given frequency can be made under tuned circuit conditions with a very narrow range of frequencies being taken into account, thus minimising the noise in the system. All alternating current bridges operate on this principle and the resulting technical refinements are such that very high precision measurements can be made with equipment that is relatively standard and is commercially available. By contrast, time-domain measuring equipment has to be specially built since the demand for it is not sufficient to enable commercial developments to be made at reasonable prices, and the noise limitations inherent in the wide-band measurement prove severe.

The mathematical basis for the treatment of the *frequency domain response*, as we call the study of the response as a function of frequency, rests on the Fourier transformation of a given function of time $G(t)$, defined by the Fourier transform:

$$\mathcal{F}[G(t)] = \mathcal{G}(\omega) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} G(t) \exp(-\mathrm{i}\omega t)\mathrm{d}t \qquad (2.35)$$

The Fourier transform gives the *frequency spectrum* $\mathcal{G}(\omega)$ of the

time-dependent function $G(t)$ – it tells us the amplitudes and phases and frequencies of the sinusoidal waves which make up the given time signal.

The crucial role of the Fourier transformation appears in the context of the convolution integral (2.30) defining the time-domain response of polarisation to an arbitrarily time-varying exciting field. It can be shown, see Appendix 2.1, that the Fourier transform of the convolution integral is given by the product of the Fourier transforms of the two functions under the integral. In the specific example of eqn (2.30) we may write:

$$\mathcal{P}(\omega) = \varepsilon_0 \chi(\omega) \mathcal{E}(\omega) \tag{2.36}$$

where $\mathcal{P}(\omega)$ and $\mathcal{E}(\omega)$ are the Fourier transforms of the time-dependent polarisation and field, respectively, as defined by eqn (2.35), while the *frequency-dependent susceptibility* is defined as the transform of the response function $f(t)$:

$$\chi(\omega) = \chi'(\omega) - \mathrm{i}\chi''(\omega) = \int_0^\infty f(t) \exp(-\mathrm{i}\omega t) \mathrm{d}t \tag{2.37}$$

The susceptibility is a *complex* function of the frequency, reflecting the fact that it gives information not only about the amplitude but also about the phase angle of the components of the polarisation. The real part $\chi'(\omega)$ gives the amplitude of polarisation in phase with the harmonic driving field, the imaginary part $\chi''(\omega)$ gives the component in quadrature with the field. We may regard the real and imaginary components as the cosine and sine transforms of $f(t)$, respectively:

$$\chi'(\omega) = \int_0^\infty f(t) \cos(\omega t) \mathrm{d}t \tag{2.38}$$

$$\chi''(\omega) = \int_0^\infty f(t) \sin(\omega t) \mathrm{d}t \tag{2.39}$$

It follows, therefore, that $\chi'(\omega)$ is an *even function* of frequency, while $\chi''(\omega)$ is an *odd function* in the double-sided frequency spectrum from minus to plus infinity. This mathematical concept of the extended frequency spectrum has no physical significance and we will normally only use the positive half of this interval, but the use of the extended interval has certain advantages in that the various formulae may take a more symmetric appearance, as shown below.

We therefore have the properties:

$$\chi'(\omega) = \chi'(-\omega), \quad \chi''(\omega) = -\chi''(-\omega) \tag{2.40}$$

and therefore this means that the real part can have a finite value at the zero of frequency. In fact, from eqn (2.38) and (2.34) we obtain:

$$\chi'(0) = \int_0^\infty f(t)\mathrm{d}t \tag{2.41}$$

which is the static value of the susceptibility. Similarly, the odd character of the imaginary part implies that it must vanish in the limit of zero frequency:

$$\chi''(0) = 0 \tag{2.42}$$

Equations (2.38) and (2.39) may be inverted by carrying out the reverse transformation which gives $f(t)$ in terms of either $\chi'(\omega)$ or $\chi''(\omega)$:

$$\begin{aligned} f(t) &= (2/\pi)\int_0^\infty \chi'(\omega)\cos(\omega t)\mathrm{d}\omega \\ &= (2/\pi)\int_0^\infty \chi''(\omega)\sin(\omega t)\mathrm{d}\omega \end{aligned} \tag{2.43}$$

This completes the relationships between the characteristic response function in the time domain and the frequency-domain response of the dielectric susceptibility and it shows that, in principle, the knowledge of either of these enables the other to be calculated using the Fourier transform technique. Given a set of experimental data points in the frequency or time domains, it is possible to obtain the corresponding other function by numerical integration.

We now return to the physical significance of the equation (2.36), replacing the convolution integral in the time domain by a simple product of two functions in the frequency domain. The frequency dependent susceptibility $\chi(\omega)$ defines the response of the dielectric material to a harmonic excitation at the frequency ω, and it can be readily measured by exciting the system at the particular frequency and determining the response. In this manner, one may sweep the desired frequency range and obtain the complete functional relationship. In view of the principle of superposition, the response to more complicated waveforms can be obtained by a simple summation of the appropriate frequency components from the spectrum of the incoming signal.

In our treatment of the frequency-domain response we have introduced the distinctive script symbols $\mathscr{D}$, $\mathscr{E}$, $\mathscr{I}$ to denote the Fourier transforms of the time–domain and therefore physically immediately meaningful field quantities D, E, I, in order to stress the fact that $\mathscr{D}(\omega)$ is an entirely different function of its argument than is $D(t)$. This distinction was initially introduced for purely didactic reasons, with a view to facilitating the discussion and to stressing the physical significance of what was being performed at the analytical level. In practice, one seldom distinguishes the time– and frequency–domain responses by separate symbols and from now onwards we propose to write $D(t)$ and $D(\omega)$ etc., since no confusion is likely to arise once the reader has been forewarned.

2.7 PERMITTIVITY, CONDUCTIVITY AND LOSS

Consider now Maxwell's equation defining the current in terms of the direct current conductivity and of the displacement current:

$$I = \sigma_0 E + \partial D / \partial t \tag{2.44}$$

Taking the Fourier transform of both sides of this equation we get for the corresponding frequency-domain response

$$I(\omega) = \sigma_0 E(\omega) + \mathrm{i}\omega D(\omega) \tag{2.45}$$

where we have made use of the property that the Fourier transform $\mathscr{F}[\partial D / \partial t] = \mathrm{i}\omega D$. We now use eqn (2.26) for $D(t)$, noting that eqn (2.36) gives the transform of $P(t)$, and we obtain:

$$\begin{aligned} I(\omega) &= \{\sigma_0 + \mathrm{i}\omega\varepsilon_0[1 + \chi'(\omega) - \mathrm{i}\chi''(\omega)]\}E(\omega) \\ &= \underbrace{\{\sigma_0 + \varepsilon_0\omega\chi''(\omega)}_{\textit{in phase}} + \underbrace{\mathrm{i}\omega\varepsilon_0[1 + \chi'(\omega)]\}}_{\textit{quadrature}}E(\omega) \end{aligned} \tag{2.46}$$

We now see the significance of the real and imaginary components of the complex susceptibility: the real part gives the component of displacement current which is *in quadrature with the driving field*, and therefore does not contribute to power loss, while the imaginary part gives the component of current *in phase with the driving field*, and therefore contributes to power loss. For this reason we refer to $\chi''(\omega)$ as the *dielectric loss*.

We note that the frequency–domain response of the dielectric medium may be written in terms of the *dielectric permittivity* $\varepsilon(\omega)$

which is defined by the expression:

$$D(\omega) = \varepsilon(\omega)E(\omega) = \varepsilon_0[1 + \chi'(\omega) - i\chi''(\omega)]E(\omega) \quad (2.47)$$

The real part of permittivity consists of the contributions of free space – which is necessarily real, since there can be no loss associated with free space – and of the real part of the susceptibility of the material medium itself. The imaginary component of the permittivity is entirely due to the material medium. In the case, which is very typical, where the material medium has several polarisation mechanisms coexisting and not significantly interacting between themselves, we may express the permittivity as the sum of the contributions of the individual mechanisms:

$$\varepsilon(\omega) = \varepsilon_0\left\{1 + \sum_l \chi_l'(\omega) - i\sum_l \chi_l''(\omega)\right\} \equiv \varepsilon'(\omega) - i\varepsilon''(\omega) \quad (2.48)$$

where the summation extends over all the separate polarisation mechanisms labelled with the index l.

The physical interpretation of the frequency–domain dielectric induction $D(\omega)$ is that its real part gives that component of the *total induced charge at the plates* which is in phase with the driving field $E(\omega)$ at the frequency ω, and this component includes the *free space contribution.* The imaginary component of $D(\omega)$ gives the component of charge that is in quadrature with the driving field $E(\omega)$. In the same way, we may regard the complex dielectric permittivity as the complex ratio of the induction and field at a given frequency, with the free space contribution ε_∞ appearing in the real part only.

Since dielectric measurements are concerned for the most part with the movement of charge, i.e. with the electric current, it follows from eqn (2.46) that the dc contribution σ_0 must appear in the result, since the measuring instrument cannot discriminate between *true* dielectric response which does not contain σ_0 and the *effective* which does.

If we write eqn (2.46) in the form

$$I(\omega) = i\omega\tilde{\varepsilon}(\omega)E(\omega) \quad (2.49)$$

where $\tilde{\varepsilon}(\omega)$ denotes the effective permittivity as measured by the instrument, we see that

$$\begin{aligned}\tilde{\varepsilon}(\omega) &= \varepsilon'(\omega) - i\{\varepsilon''(\omega) + \sigma_0/\omega\} \\ &= \varepsilon_0\{1 + \chi'(\omega) - i[\chi''(\omega) + \sigma_0/\varepsilon_0\omega]\}\end{aligned} \quad (2.50)$$

The significance of the last term in eqn (2.50) is that the dc conductivity makes a contribution to the apparent dielectric loss measured by a bridge or other instrument, which diverges towards zero frequency. This is not a *true dielectric* response, since it is not accompanied by any contribution to the real part of the permittivity and it arises because no instrument can distinguish between true dielectric and dc transport processes. This distinction is vital, however, for the interpretation of the physical processes involved, since as we shall see later, *direct current conduction and dielectric phenomena are essentially separate and, for the most part, independent processes.* A behaviour in which $\varepsilon''(\omega) \propto 1/\omega$ while $\varepsilon'(\omega) \rightarrow$ const as $\omega \rightarrow 0$ is therefore, conclusive evidence that the dominant process is direct current conduction in the material in question in the relevant frequency range.

2.8 KRAMERS–KRONIG RELATIONS

We conclude the present Chapter with a brief derivation of a very important relation connecting the real and imaginary parts of the complex susceptibility and known as the Kramers–Kronig relations. We do not propose to give a fully rigorous derivation, which relies on the properties of the functions of complex variables and uses contour integration in the complex plane of frequency, taking into account certain fundamental properties of the susceptibility function. Instead we will only give a simplified "proof", leaving it to the reader to follow up more specialised literature for the full treatment (Landau and Lifshitz 1962).

The Kramers–Kronig relations are ultimately a consequence of the principle of causality – the fact that the dielectric response function satisfies eqn (2.28). Now one look at equations (2.38) and (2.39) is sufficient to convince us that both $\chi'(\omega)$ and $\chi''(\omega)$ are derived from the same generating function $f(t)$ and that it should be possible in principle to "eliminate" this function and to express $\chi'(\omega)$ in terms of $\chi''(\omega)$.

We proceed as follows. First we prove the following transformation:

$$\int_{-\infty}^{\infty} \frac{\sin xt}{x-\omega}\,dx = \cos \omega t \int_{-\infty}^{\infty} \frac{\sin(x-\omega)t}{(x-\omega)t}\,d[(x-\omega)t] + \sin \omega t \int_{-\infty}^{\infty} \frac{\cos(x-\omega)t}{(x-\omega)t}\,d[(x-\omega)t]$$

where the integrals denote the Cauchy principal value, i.e. they ignore the imaginary contributions arising from integration through the pole at $x = \omega$. The first integral is equal to π, the second vanishes so that we obtain;

$$(1/\pi) \fint_{-\infty}^{\infty} \frac{\sin xt}{x - \omega} \mathrm{d}x = \cos \omega t \tag{2.51}$$

The integral on the left is known as the *Hilbert transform* of the function $\sin xt$. We now substitute this into (2.38):

$$\chi'(\omega) = (1/\pi) \int_{-\infty}^{\infty} f(t) \fint_{-\infty}^{\infty} \frac{\sin xt}{x - \omega} \mathrm{d}x \, \mathrm{d}t$$

$$= (1/\pi) \fint_{-\infty}^{\infty} \frac{1}{x - \omega} \int_{-\infty}^{\infty} f(t) \sin xt \, \mathrm{d}t \, \mathrm{d}x$$

In these manipulations we have extended the integration of eqn (2.38) to $-\infty$ which is permissible in view of the causality principle. The second integral in this expression is equal to $\chi''(\omega)$ in view of eqn (2.39), so that we may finally write:

$$\chi'(\omega) = \frac{1}{\pi} \fint_{-\infty}^{\infty} \frac{\chi''(x)}{x - \omega} \mathrm{d}x \tag{2.52}$$

and similarly:

$$\chi''(\omega) = -\frac{1}{\pi} \fint_{-\infty}^{\infty} \frac{\chi'(x)}{x - \omega} \mathrm{d}x \tag{2.53}$$

These are the Kramers–Kronig relations which express the value of either $\chi''(\omega)$ or $\chi'(\omega)$ at a particular value of the frequency ω in terms of the integral transform of the other throughout the entire frequency range $(-\infty, \infty)$. In view of what was said above about the even and odd character of these functions, we may change the range of integration to $(0, \infty)$ and thus obtain the one-sided Kramers–Kronig integrals:

$$\chi'(\omega) = \frac{2}{\pi} \fint_{0}^{\infty} \frac{x\chi''(x)}{x^2 - \omega^2} \mathrm{d}x \tag{2.54}$$

$$\chi''(\omega) = -\frac{2\omega}{\pi} \fint_{0}^{\infty} \frac{\chi'(\omega)}{x^2 - \omega^2} \mathrm{d}x \tag{2.55}$$

Comprehensive tables of Hilbert transforms may be found in *Tables of Integral Transforms* (Erdelyi 1954).

We shall find that the Kramers–Kronig relations have a fundamental significance for our approach to the interpretation of the dielectric behaviour of solids. They are also very useful in certain experimental situations where they may enable the values of one of the functions to be obtained from those of the other when for some reason the measurements of them are impeded. This will be discussed later.

One very immediate consequence of the Kramers–Kronig relations is their evaluation for the case of zero frequency:

$$\chi'(0) = \frac{2}{\pi}\int_0^\infty \frac{\chi''(x)}{x}\,dx = \frac{2}{\pi}\int_{-\infty}^{\infty} \chi''(x)\,d(\ln x) \qquad (2.56)$$

This relates the *polarisation increment* for a given polarisation mechanism or combination of mechanisms, to the area of the loss curve plotted against the logarithm to the base e of the frequency. This immediately shows that a mechanism leading to a strong polarisation must inevitably give rise to correspondingly high losses somewhere in the frequency spectrum, Figure 2.9.

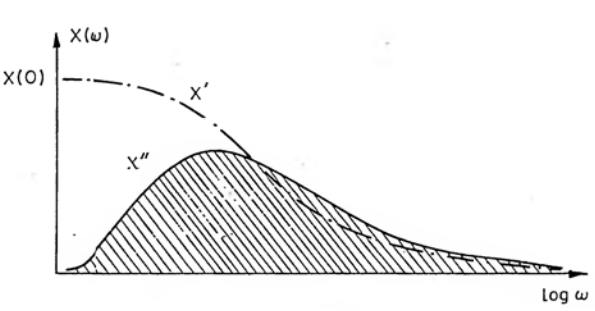


Figure 2.9 The significance of the relation (2.56) between the polarisation increment $\chi(0)$ and the area under the curve of $\chi''(\omega)$, in linear scale, against the logarithm of frequency.

The same conclusion may be expressed in a different way: it is impossible to have a loss-free dielectric material of finite susceptibility. It is possible to extend this even further and to state that it is impossible to have a dispersion-free dielectric material, i.e. one which shows frequency-independent real and imaginary parts. This follows from the fact that the Hilbert transform of a constant is equal to zero, as is clear from the inspection of the integrals in question.

We conclude therefore that a variation of the dielectric parameters with frequency – known as dispersion – is an essential property of all dielectric materials.

On the whole, most dielectric responses are characterised by the fact that their loss is significant only in certain restricted ranges of frequency, sometimes two or three decades, sometimes rather more, as indicated schematically in Figure 2.10. One of the consequences of the Kramers–Kronig relations is that the regions of negligible loss make a negligible contribution to the integral (2.52), in other words $\chi'(\omega)$ is almost independent of frequency outside the regions of significant loss, remaining constant *below* the lossy region and rapidly decreasing to zero *above it*. This means that, in the presence of several *substantially non-overlapping* loss processes, the regions between the successive loss peaks are characterised by an almost constant value of the dielectric permittivity which is determined by

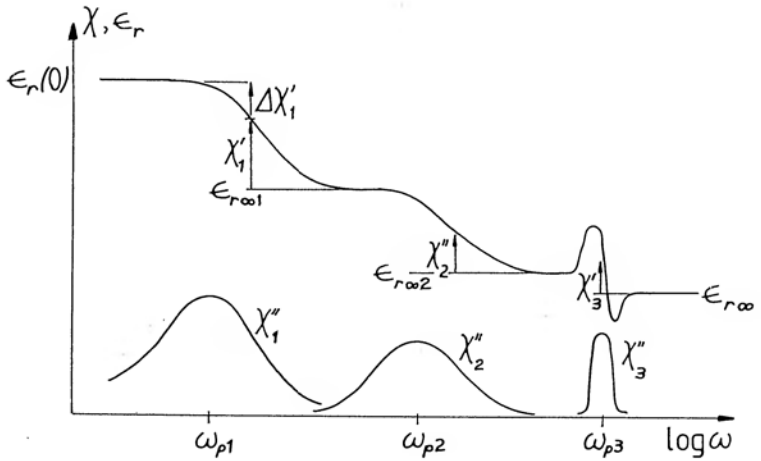


Figure 2.10 An illustration of the manner in which the real part of the total dielectric permittivity, shown here as the *relative* permittivity $\varepsilon_r' = \varepsilon'(\omega)/\varepsilon_0$ is made up of contributions of all loss processes at frequencies higher than the frequency in question. Three processes are shown, denoted by the subscripts $\alpha = 1, 2, 3$, the last one being a resonance process, and they are assumed to be reasonably well separated in frequency. The real part of the susceptibility $\chi'_\alpha(\omega)$ of every process adds to the sum of the contributions of all higher processes which define the "high frequency permittivity" $\varepsilon_{\infty\alpha}$ for that process. The dielectric decrement $\Delta\chi'_\alpha(\omega)$, as defined by eqn (3.40) is also indicated for the first process. Note that the real part of the permittivity remains flat in regions where losses are negligible. The ordinate scale is shown to be *linear* in this case to bring out more clearly the additive properties of the contributions to ε_r'.

the total area under *all the higher-frequency loss peaks*, as derived in the eqn (2.56) and shown in Figure 2.10.

This being the situation, it is often possible to define, for the purpose of the study of a limited region of frequency, a "high frequency permittivity" ε_∞ consisting of the free space contribution ε_0 and of the contributions of all the other polarising processes at higher frequencies, so that eqn (2.48) may be written for any such chosen polarisation mechanism α:

$$\varepsilon(\omega) = \varepsilon_{\infty\alpha} + \varepsilon_0\{\chi'_\alpha(\omega) - i\chi''_\alpha(\omega)\} \tag{2.57}$$

To the extent to which $\varepsilon_{\infty\alpha}$ can, in fact, be sensibly defined in any given situation, i.e. to the extent to which higher-lying loss peaks are sufficiently far removed in frequency from the region under investigation, it is possible to obtain the susceptibility $\chi_\alpha(\omega)$ by subtracting ε_∞ from the measured values of the real part $\varepsilon'(\omega)$.

The constant value of ε_∞ is not included in the result of Kramers–Kronig's transformation from $\chi''(\omega) = \varepsilon''(\omega)/\varepsilon_0$, since the Hilbert transform of a constant is zero, and neither is the singularity $\sigma_0(\omega)$ arising from the presence of dc conductivity recovered from the Kramers–Kronig transformation of the real part $\varepsilon'(\omega)$ or $\chi'(\omega)$, because dc conductivity does not contribute to the real part of the permittivity.

These considerations highlight the significance of Kramers–Kronig transformations in the processing and interpretation of experimental data, which are frequently obscured by the presence of the dc conductivity in the loss results at low frequencies and of the ε_∞ term in the $\varepsilon'(\omega)$ data at high frequencies. Both these perturbing influences may be removed by using the Kramers–Kronig transformation from the other data which are not affected (Lovell 1974).

A program for numerical computation of Kramers–Kronig transforms has been developed by Lovell in BASIC and is given in Appendix 2.2. This program converts the directly measured data on the real part of the capacitance, $C'(\omega)$, into the coresponding frequency-dependent conductance $G(\omega) = \omega C''(\omega)$, and C' and C'' are related to the permittivity components ε' and ε'' by the same geometrical factor. In order to make the fullest possible use of the finite range of measured frequencies, the program extrapolates the measured results by one decade at either end, thereby giving more reliable data within the measuring window. It should be appreciated that this procedure may lead to errors near the ends of the range

if the true response should show some unexpected change of trend just outside the measurement window, but long experience in the Chelsea Dielectrics Group has shown no perceptible complications arising from this source. It should be noted that the program accepts data and outputs the computed values at constant intervals on the logarithmic frequency range, as is customary to record these data in practice.

APPENDIX 2.1

FOURIER TRANSFORM OF THE CONVOLUTION INTEGRAL

Define the Fourier transforms

$$\mathcal{F}f(t) = \chi(\omega)$$

and

$$\mathcal{F}E(t) = \mathcal{E}(\omega)$$

Then

$$\begin{aligned}
&\mathcal{F}\left\{\int_{-\infty}^{\infty} f(\tau)E(t-\tau)\mathrm{d}\tau\right\} \\
&= \int_{-\infty}^{\infty} \exp(-\mathrm{i}\omega t)\left\{\int_{-\infty}^{\infty} f(\tau)E(t-\tau)\mathrm{d}\tau\right\}\mathrm{d}t \\
&= \int_{-\infty}^{\infty} f(\tau)\left\{\int_{-\infty}^{\infty} \exp(-\mathrm{i}\omega t)E(t-\tau)\mathrm{d}t\right\}\mathrm{d}\tau \\
&= \int_{-\infty}^{\infty} f(\tau)\exp(-\mathrm{i}\omega\tau)\left\{\int_{-\infty}^{\infty} \exp[-\mathrm{i}\omega(t-\tau)E(t-\tau)\mathrm{d}t\right\}\mathrm{d}\tau \\
&= \int_{-\infty}^{\infty} f(\tau)\exp(-\mathrm{i}\omega\tau)\mathrm{d}\tau \int_{-\infty}^{\infty} E(s)\exp(-\mathrm{i}\omega s)\mathrm{d}s \\
&= \chi(\omega)\mathcal{E}(\omega)
\end{aligned}$$

APPENDIX 2.2

COMPUTER PROGRAMS FOR KRAMERS–KRONIG TRANSFORMATION C → G AND G → C

```
DATA FORMAT

    ALL DATA FROM DIELECTRIC MEASUREMENTS ARE RECORDED ON 8-HOLE PAPER
TAPE USING THE STANDARD ASCI CHARACTER SET.  INFORMATION THUS RECORDED
INCLUDES A HEADER CONCERNING THE SAMPLE TYPE, APPLIED AC AND DC FIELDS
AND FREQUENCY RANGE INSTRUCTIONS. THE DATA FOLLOWS, EACH POINT BEING
RECORDED IN TERMS OF FREQUENCY (HZ), CAPACITANCE (PF), CONDUCTANCE (MHO)
AND TAN DELTA. A TRAILER IS USED TO INDICATE DATA TERMINATION.
    THE PROGRAMME IS PRIMARILY FOR USE WHERE THE FREQUENCY IS INCREMENTED
LOGARITHMICALLY.

INPUT ROUTINE
    LINES 10, 11 AND 12 INPUT THE HEADER.
    LINES 16 AND 17 PERMIT THE OPTION OF EXCLUDING UNDESIRED DATA.
    LINES 30 TO 55 INPUT THE DATA.
    LINE 33 DETECTS THE TRAILER AND TRANSFERS CONTROL TO THE ANALYSIS
COMMENCING AT LINE 60.

OUTPUT ROUTINE (TO PAPER TAPE PUNCH)
    LINE 12 OUTPUTS THE HEADER.
    LINE 26 OUTPUTS A "DATA" STATEMENT.
    LINES 1794 TO 1798 OUTPUT THE TRANSFORMED DATA.
    LINE 412 APPENDS THE TRAILER.

    A SIMPLIFIED INPUT ROUTINE MAY BE SUBSTITUTED WHICH NEEDS ONLY
FREQUENCY AND CAPACITANCE OR CONDUCTANCE INFORMATION AS FOLLOWS:-
    DELETE LINES 10,11,12,16,17,30 TO 60.
    FOR THE C TO G TRANSFORM:-
    INSERT 30 N=0
           32 N=N+1
           34 INPUT #1:X
           36 IF X<0 THEN 60
           38 INPUT #1:Y
           40 REM
           42 GOTO 32
           60 REM
    IN THE TRANSFORM FOR G TO C LINE 40 SHOULD BE CHANGED:-
           40 Y=Y/2/PI/X
    INPUT DATA MAY BE IN EITHER ASCENDING OR DESCENDING FREQUENCY ORDER.
    THE PARAMETER X DEFINES FREQUENCY (HZ).
                  Y DEFINES EITHER C (F) OR G (MHO). Y>0 IS NECESSARY.
    INPUT IS TERMINATED WHEN A NEGATIVE FREQUENCY IS ENCOUNTERED.

    A COMPATIBLE PUNCH OUTPUT ROUTINE MAY ALSO BE SUBSTITUTED:-
    DELETE LINES 12,26,412,1794,1796 AND 1798.
    FOR THE C TO G TRANSFORM:-
    INSERT 26 PRINT NLN(2);"DATA";NLN(2); \ PRINT #2:RPT(100,0);
           412 PRINT #2:FMT("A")-1 \ PRINT #2:RPT(50,0);
           1794 PRINT #2:FMT("A")U
           1796 PRINT #2:H2
    IN THE TRANSFORM FOR G TO C LINE 1796 SHOULD BE CHANGED:-
           1796 PRINT #2:H2/1E12
```

```
1 PRINT "KRAMERS KRONIG TRANSFORM  C  TO  G"
3 DIM D5(200)
4 DIM F(100),Y1(100),Y2(100)
5 DIM Q(50),V(50),A(50),B(50),C(50)

5 DIM Q(50),V(50),A(50),B(50),C(50)
6 DIM D(50),E(50)
9 I=1
10 INPUT #1:Z$
11 IF Z$="DATA" GO TO 16
12 PRINT Z$ \ PRINT #2:Z$ \ GO TO 10
16 PRINT "INPUT DATA SETS TO BE OMITTED ? (ENTER 0 TO TERMINATE)" \ N9=0
17 N9=N9+1 \ INPUT D5(N9) \ IF D5(N9)>0 THEN 17 \ D5(N9)=1000 \ N9=1
18 PRINT
20 PRINT "FREQUENCY RANGE OF OUTPUT DATA"
21 PRINT "      MIN        "; \ INPUT F3
22 PRINT "      MAX        "; \ INPUT F2
23 PRINT "POINTS/DECADE  "; \ INPUT T1
26 PRINT NLN(2);"DATA";NLN(2); \ PRINT #2:"DATA";NLN(2);RPT(40,0);
27 PRINT " F (HZ)          C (F)          G/W (F)"
28 PRINT "                 GIVEN          COMPUTED";NLN(2);
30 FOR N=1 TO 200
32 INPUT #1:Z$
33 IF SEG$(Z$,1,3)="FIN" GO TO 60
34 IF Z$<>"OK" GO TO 32
40 INPUT #1:Z$
42 IF D5(N9)=I GO TO 53
44 X=VAL(SEG$(Z$,1,13))
45 Y=VAL(SEG$(Z$,16,27))
46 Y=Y*1.00000E-12
48 IF Y<=0 THEN N=N-1 \ GO TO 55
50 Q(N)=LOG(X) \ V(N)=LOG(Y)
51 GO TO 54
53 N9=N9+1 \ N=N-1
54 I=I+1
55 NEXT N
60 N=N-1
62 IF Q(N)>Q(1) THEN 95
65 FOR I=1 TO N
70 D(I)=Q(N+1-I)
75 E(I)=V(N+1-I)
80 NEXT I
90 FOR I=1 TO N \ Q(I)=D(I) \ V(I)=E(I) \ NEXT I
95 I=0
96 H1=0 \ H2=0 .
100 IF INT(N/2)*2=N GO TO 120
110 GO TO 130
120 Q(N+1)=Q(N-2)
125 V(N+1)=V(N-2)
130 N1=INT((N+2)/2)
140 FOR J=2 TO N1
150 A2=0 \ B2=0 \ C2=0 \ D2=0
160 J1=2*J-3
165 K=2*J-1
180 FOR I=J1 TO K
190 L=I+1
200 M=I+2
210 IF (L-J1)>2 THEN L=L-3
220 IF ((M-J1))>2 THEN M=M-3
230 A2=A2+V(I)*(Q(L)-Q(M))
240 B2=B2-V(I)*(Q(L)^2-Q(M)^2)
250 C2=C2+V(I)*Q(L)*Q(M)*(Q(L)-Q(M))
260 D2=D2+Q(I)^2*(Q(L)-Q(M))
265 NEXT I
270 A(J)=A2/D2
280 B(J)=B2/D2
290 C(J)=C2/D2
295 NEXT J
300 M1=5
```

```
305 N5=1
310 K=1
315 S1=0 \ S2=0 \ S3=0 \ S4=0
320 L=K+M1-1
325 FOR I=K TO L
330 S1=S1+Q(I) \ S2=S2+V(I) \ S3=S3+Q(I)^2 \ S4=S4+Q(I)*V(I)
335 NEXT I
340 R=M1
345 A(N5)=0
350 B(N5)=(R*S4-S1*S2)/(R*S3-S1^2)
355 C(N5)=(S2-B(N5)*S1)/R
360 IF N5>1 GO TO 385
365 N5=N1+1
370 K=N-M1+1
375 GO TO 315
385 Y1=1
395 FOR I5=1 TO T1
400 S=I5-1
405 F=F3*Y1*10^(S/T1)
410 IF F<=F2 GO TO 420
412 PRINT #2;NLN(2);"FIN KK C TO G";NLN(15);RPT(100,0);
414 PRINT RPT(10,7);NLN(1);"FIN KK C TO G";NLN(15);
416 END
420 IF F<.5*EXP(Q(1)) THEN STOP
425 GOSUB 1000
430 NEXT I5
435 Y1=10*Y1
440 GO TO 395
1000 REM
1010 U=F
1020 S1=.8*U
1030 T=1.25*U
1040 X=LOG(S1)
1050 Y=LOG(T)
1060 V=LOG(U)
1070 N1=INT((N+2)/2)
1080 I=0 \ J=0 \ M=0
1090 FOR K=1 TO N1
1100 R=Q(2*K-1)
1110 IF K=N1 THEN R=Q(N)
1120 IF I=0 THEN IF X<R THEN I=K
1130 IF J=0 THEN IF Y<R THEN J=K
1140 IF M=0 THEN IF V<R THEN M=K
1150 IF J>0 GO TO 1200
1160 NEXT K
1170 IF I=0 THEN I=N1+1
1180 IF J=0 THEN J=N1+1
1190 IF M=0 THEN M=N1+1
1200 Q1=Q(1)-2.303
1210 N3=N1+1
1220 IF I>1 THEN Q1=Q(2*I-3)
1230 IF I=N3 THEN Q1=Q(N)
1240 H2=0
1250 U1=Q1 \ U2=X \ P2=I \ GOSUB 2000
1260 Q2=Q(N)+2.303
1270 IF J<N3 THEN Q2=Q(2*J-1)
1280 IF J=N1 THEN Q2=Q(N)
1290 U1=Y \ U2=Q2 \ P2=J \ GOSUB 2000
1300 DEFFNZ(I,X)=EXP(A(I)*X^2+B(I)*X+C(I))
1320 C1=FNZ(I,X)
1330 C2=FNZ(J,Y)
1350 F1=(C2-C1)/(T-S1)
1360 G=(C1*T-C2*S1)/(T-S1)
1370 P=(T-U)/(U-S1)
1380 R=(T+U)/(U+S1)
```

```
1380 R=(T+U)/(U+S1)
1390 H2=H2-U*F1*LOG(P*R)/2-G*LOG(P/R)/2
1400 J=J+1
1410 IF J>N3 GO TO 1520
1420 FOR K=J TO N3
1430 Q2=Q(N)+2.303
1440 IF K<N3 THEN Q2=Q(2*K-1)
1450 IF K=N1 THEN Q2=Q(N)
1460 Q1=Q(2*K-3)
1470 IF K=N3 THEN Q1=Q(N)
1480 U1=Q1 \ U2=Q2 \ P2=K \ GOSUB 2000
1490 IF K<0 GO TO 1520
1500 NEXT K
1510 REM
1520 I=I-1
1530 IF I<1 GO TO 1630
1540 FOR K=1 TO I
1550 L=I-K+1
1560 Q1=Q(1)-2.303
1570 IF K<I THEN Q1=Q(2*L-3)
1580 Q2=Q(2*L-1)
1590 IF L=N1 THEN Q2=Q(N)
1600 U1=Q1 \ U2=Q2 \ P2=L \ GOSUB 2000
1610 IF K<0 GO TO 1630
1620 NEXT K
1630 FOR K=1 TO 2
1640 IF K=1 GO TO 1690
1650 Q3=Q(N)+2.303
1660 L=N3
1670 T=.5
1680 GO TO 1720
1690 Q3=Q(1)-2.303
1700 L=1
1710 T=-.5
1720 G=FNZ(L,Q3)
1730 F0=EXP(Q3)
1740 H=LOG(ABS((F0-U)/(F0+U)))
1750 H2=H2+H*G*T
1755 NEXT K
1760 G=FNZ(M,V)
1780 H2=H2*.63662
1790 PRINT FMT("E",10)U,G,H2
1792 G=G*1.00000E+12
1793 H2=H2*2*PI*U
1794 PRINT #2:"OK"
1796 PRINT #2:FMT("E",13)U;" ";G;" ";H2;" ";0;NLN(2);
1798 PRINT #2:RPT(20,0);
1800 RETURN
2000 REM
2010 DEFFNW(X,I)=EXP(A(I)*X^2+B(I)*X+C(I))/(U/EXP(X)-EXP(X)/U)
2020 L2=INT((U2-U1)*100/2.303)
2030 I2=INT(L2/4)*2
2040 IF I2<4 THEN I2=4
2050 W2=I2
2060 K2=I2-1
2070 H3=(U2-U1)/W2
2080 T2=U1
2090 E2=0
2100 O2=0
2110 FOR J8=1 TO K2 STEP 2
2120 O2=O2+FNW(T2,P2)
2130 E2=E2+FNW(T2+H3,P2)
2135 T2=T2+2*H3
2140 NEXT J8
2149 H6=0 \ H8=0
```

```
2149 H6=0 \ H8=0
2150 H1=(4*E2+2*D2-FNW(U1,P2)+FNW(U2,P2))*H3/3
2160 H2=H2+H1
2165 IF ABS(H1)<1.00000E-04*ABS(H2) THEN K=-1
2170 REM
2190 RETURN



1 PRINT  \ PRINT "KRAMERS KRONIG TRANSFORM   G TO C" \ PRINT
3 DIM D5(200)
5 DIM Q(50),V(50),A(50),B(50),C(50)
6 DIM D(50),E(50)
8 PRINT #2:RPT(200,0);
9 I=1
10 INPUT #1:Z$
11 IF Z$="DATA" GO TO 16
12 PRINT Z$ \ PRINT #2:Z$ \ GO TO 10
16 PRINT "INPUT DATA SETS TO BE OMITTED ? (ENTER 0 TO TERMINATE)" \ N9=0
17 N9=N9+1 \ INPUT D5(N9) \ IF D5(N9)>0 GO TO 17 \ D5(N9)=1000
18 PRINT
20 PRINT "FREQUENCY RANGE OF OUTPUT DATA"
21 PRINT "      MIN        "; \ INPUT F3
22 PRINT "      MAX        "; \ INPUT F2
23 PRINT "POINTS/DECADE  "; \ INPUT T1
24 PRINT  \ PRINT "VALUE OF C INF, TO BE ADDED  (PFS)     "; \ INPUT C9
26 PRINT NLN(2);"DATA";NLN(2); \ PRINT #2:"DATA";NLN(2);RPT(40,0);
27 PRINT " F (HZ)          C (F)          G/W (F)"
28 PRINT "                 COMPUTED       GIVEN";NLN(2);
30 FOR N=1 TO 200
32 INPUT #1:Z$
33 IF SEG$(Z$,1,3)="FIN" GO TO 60
34 IF Z$<>"OK" GO TO 32
40 INPUT #1:Z$
42 IF D5(N9)=I GO TO 53
44 X=VAL(SEG$(Z$,1,13))
45 Y=VAL(SEG$(Z$,29,41))
46 Y=Y/(2*PI*X)
48 IF Y<=0 THEN N=N-1 \ GO TO 55
50 Q(N)=LOG(X) \ V(N)=LOG(Y)
51 GO TO 54
53 N9=N9+1 \ N=N-1
54 I=I+1
55 NEXT N
60 N=N-1
62 IF Q(N)>Q(1) THEN 95
65 FOR I=1 TO N
70 D(I)=Q(N+1-I)
75 E(I)=V(N+1-I)
80 NEXT I
90 FOR I=1 TO N \ Q(I)=D(I) \ V(I)=E(I) \ NEXT I
95 I=0
96 H1=0 \ H2=0
100 IF INT(N/2)*2=N GO TO 120
110 GO TO 130
120 Q(N+1)=Q(N-2)
125 V(N+1)=V(N-2)
130 N1=INT((N+2)/2)
140 FOR J=2 TO N1
150 A2=0 \ B2=0 \ C2=0 \ D2=0
160 J1=2*J-3
165 K=2*J-1
180 FOR I=J1 TO K
190 L=I+1
```

```
190 L=I+1
200 M=I+2
210 IF (L-J1)>2 THEN L=L-3
220 IF ((M-J1))>2 THEN M=M-3
230 A2=A2+V(I)*(Q(L)-Q(M))
240 B2=B2-V(I)*(Q(L)^2-Q(M)^2)
250 C2=C2+V(I)*Q(L)*Q(M)*(Q(L)-Q(M))
260 D2=D2+Q(I)^2*(Q(L)-Q(M))
265 NEXT I
270 A(J)=A2/D2
280 B(J)=B2/D2
290 C(J)=C2/D2
295 NEXT J
300 M1=5
305 N5=1
310 K=1
315 S1=0 \ S2=0 \ S3=0 \ S4=0
320 L=K+M1-1
325 FOR I=K TO L
330 S1=S1+Q(I) \ S2=S2+V(I) \ S3=S3+Q(I)^2 \ S4=S4+Q(I)*V(I)
335 NEXT I
340 R=M1
345 A(N5)=0
350 B(N5)=(R*S4-S1*S2)/(R*S3-S1^2)
355 C(N5)=(S2-B(N5)*S1)/R
360 IF N5>1 GO TO 385
365 N5=N1+1
370 K=N-M1+1
375 GO TO 315
385 Y1=1
395 FOR I5=1 TO T1
400 S=I5-1
405 F=F3*Y1*10^(S/T1)
410 IF F<=F2 GO TO 420
412 PRINT #2;NLN(2);"FIN KK G TO C";NLN(15);RPT(100,0);
414 PRINT RPT(10,7);NLN(1);"FIN KK G TO C";NLN(15);
416 END
420 IF F<.5*EXP(Q(1)) THEN STOP
425 GOSUB 1000
430 NEXT I5
435 Y1=10*Y1
440 GO TO 395
1000 REM
1010 U=F
1020 S1=.8*U
1030 T=1.25*U
1040 X=LOG(S1)
1050 Y=LOG(T)
1060 V=LOG(U)
1070 N1=INT((N+2)/2)
1080 I=0 \ J=0 \ M=0
1090 FOR K=1 TO N1
1100 R=Q(2*K-1)
1110 IF K=N1 THEN R=Q(N)
1120 IF I=0 THEN IF X<R THEN I=K
1130 IF J=0 THEN IF Y<R THEN J=K
1140 IF M=0 THEN IF V<R THEN M=K
1150 IF J>0 GO TO 1200
1160 NEXT K
1170 IF I=0 THEN I=N1+1
1180 IF J=0 THEN J=N1+1
1190 IF M=0 THEN M=N1+1
1200 Q1=Q(1)-2.303
1210 N3=N1+1
1220 IF I>1 THEN Q1=Q(2*I-3)
```

```
1220 IF I>1 THEN Q1=Q(2*I-3)
1230 IF I=N3 THEN Q1=Q(N)
1240 H2=0
1250 U1=Q1 \ U2=X \ P2=I \ GOSUB 2000
1260 Q2=Q(N)+2.303
1270 IF J<N3 THEN Q2=Q(2*J-1)
1280 IF J=N1 THEN Q2=Q(N)
1290 U1=Y \ U2=Q2 \ P2=J \ GOSUB 2000
1300 DEFFNZ(I,X)=EXP(A(I)*X^2+B(I)*X+C(I))
1320 C1=FNZ(I,X)
1330 C2=FNZ(J,Y)
1350 F1=(C2-C1)/(T-S1)
1360 G=(C1*T-C2*S1)/(T-S1)
1370 P=(T-U)/(U-S1)
1380 R=(T+U)/(U+S1)
1390 H2=H2+F1*(T-S1)+U*F1*LOG(P/R)/2+G*LOG(P*R)/2
1400 J=J+1
1410 IF J>N3 GO TO 1520
1420 FOR K=J TO N3
1430 Q2=Q(N)+2.303
1440 IF K<N3 THEN Q2=Q(2*K-1)
1450 IF K=N1 THEN Q2=Q(N)
1460 Q1=Q(2*K-3)
1470 IF K=N3 THEN Q1=Q(N)
1480 U1=Q1 \ U2=Q2 \ P2=K \ GOSUB 2000
1490 IF K<0 GO TO 1520
1500 NEXT K
1510 REM
1520 I=I-1
1530 IF I<1 GO TO 1760
1540 FOR K=1 TO I
1550 L=I-K+1
1560 Q1=Q(1)-2.303
1570 IF K<I THEN Q1=Q(2*L-3)
1580 Q2=Q(2*L-1)
1590 IF L=N1 THEN Q2=Q(N)
1600 U1=Q1 \ U2=Q2 \ P2=L \ GOSUB 2000
1610 IF K<0 GO TO 1760
1620 NEXT K
1760 G=FNZ(M,V)
1780 H2=H2*.63662
1790 PRINT FMT("E",10)U,H2,G
1792 H2=H2*1.00000E+12
1793 G=G*2*PI*U
1794 PRINT #2;"OK"
1796 PRINT #2;FMT("E",13)U;" ";H2;" ";G;" ";0;NLN(2);
1798 PRINT #2;RPT(20,0);
1800 RETURN
2000 REM
2010 DEFFNW(X,I)=EXP(A(I)*X^2+B(I)*X+C(I))/(1-U^2/EXP(2*X))
2020 L2=INT((U2-U1)*100/2.303)
2030 I2=INT(L2/4)*2
2040 IF I2<4 THEN I2=4
2050 W2=I2
2060 K2=I2-1
2070 H3=(U2-U1)/W2
2080 T2=U1
2090 E2=0
2100 O2=0
2110 FOR J8=1 TO K2 STEP 2
2120 O2=O2+FNW(T2,P2)
2130 E2=E2+FNW(T2+H3,P2)
2135 T2=T2+2*H3
2140 NEXT J8
2149 H6=0 \ H8=0
2149 H6=0 \ H8=0
2150 H1=(4*E2+2*O2-FNW(U1,P2)+FNW(U2,P2))*H3/3
2160 H2=H2+H1
2165 IF ABS(H1)<1.00000E-04*ABS(H2) THEN K=-1
2170 REM
2190 RETURN
```

REFERENCES TO CHAPTER 2

Burfoot J C 1967, *Ferroelectrics, An Introduction to the Physical Principles*, D van Nostrand

Debye P 1945, *Polar Molecules*, Dover, New York

Denbigh G K 1940, Trans. Faraday Soc. **36**, 936

Erdelyi A 1954, *Tables of Integral Transforms* Vol. II, McGraw–Hill, New York

Jonscher A K and Hill R M 1975, *Physics of Thin Films* Vol. 8, 169–249

Kirkwood J G 1939, J. Chem. Phys. **7**, 911

Lampert M A 1965, *Injection Currents in Solids*, Academic Press, New York

Landau L D and Lifshitz E M 1962, *Electrodynamics of Continuous Media*, Pergamon, Oxford

Lovell R 1974, J. Phys. C: Solid State Physics **7**, 4378–4384

Mott N F and Davis E A 1979, *Electronic Processes in Non-crystalline Materials*, Oxford University Press

Nye J F 1957, *Physical Properties of Crystals*, Clarendon Press, Oxford

Onsager I 1936, J. Am. Chem. Soc. **58**, 1486

Platzmann P M and Wolff P A 1973, *Waves and Interactions in Solid State Plasmas*, Academic Press, New York

Scaife B K P 1973, in *Problems in Physical Electronics*, R L Ferrari and A K Jonscher (Eds), Pion, p. 287

Teachout R E and Pack R T 1971, Atomic Data **3**, 195

Westphal W B and Sils A 1972, *Dielectric Constant and Loss Data*, Massachusetts Institute of Technology Technical Report AFML TR 72 39

Zheludev I S 1971, *Physics of Crystalline Dielectrics* Vol. I and II, Plenum Press, New York

CHAPTER 3

Presentation of Dielectric Functions

3.1 INTRODUCTION

The dielectric functions, $f(t)$ in the time domain and the complex frequency domain function $\chi(\omega)$, represent the most basic form of information about the dielectric properties of a substance and they form the basis of most interpretational analysis on experimental data. However, even at this fundamental level there are various options open to the presenter of data, depending upon the nature of the particular response: for example it may be advantageous to plot the frequency dependent conductivity $\sigma(\omega)$ instead of dielectric loss and it may be desirable to plot $\varepsilon'(\omega)$ instead of $\chi'(\omega)$. What is even more important, however, is the often complicated process of arriving at the basic information implied by these functions, which presupposes the uniformity of the measured sample, i.e. the absence of separate regions characterised by different dielectric properties. The presence of inhomogeneities may complicate very considerably the appearance of the raw data, as they are obtained from the measuring system used in their determination and failure to recognise properly the origin of these complications may lead to erroneous and highly misleading conclusions from experimental data. It is therefore necessary to discuss in detail some of the problems involved here.

There exists, however, the further difficulty that through a process of slow and often erratic evolution over the decades, there have arisen a number of separate, sometimes not very logical, conventions regarding the presentation of information acceptable in a particular experimental context. This has the disadvantage, that the users of a particular convention may be perfectly familiar with the significance of certain features, for example of a complex impedance plot, but someone not so familiar may find it very difficult to understand the meaning and the interpretation of data in these particular forms. What is more, a critical analysis of particular time-honoured methods of presentation often reveals serious drawbacks which actually

hamper proper interpretation and we shall attempt to point those out in the course of our discussion.

Dielectric data have one distinctive feature compared, for example with most other physical data relating to steady state situations – they involve the variables of time or frequency, in addition to any variables such a temperature, pressure, composition etc which apply in all material measurements. This means that one is dealing with a very large volume of data and their proper handling and presentation is particularly important. In addition, the availability of Fourier and Kramers–Kronig transformations presents further valuable opportunities for checking internal consistency of the data and for obtaining the maximum amount of information but, once again, they entail the ability to handle and store the required information.

The aim of the present Chapter is to give a critical introduction into these various methods of presentation and of handling of data and to acquaint the reader with the most important difficulties and pitfalls. It is not our object, at the present stage, to enter into the details of the actual physical interpretation of the data which will have to await the final chapters of this monograph – at this stage we intend to equip the reader with the optimal tools for the proper presentation of the information that may be at his disposal.

3.2 ADMITTANCE, IMPEDANCE, PERMITTIVITY

In the frequency range below 1 GHz which we have chosen for the present treatment for reasons explained in Chapter 1, measurements are carried out almost exclusively on samples in the form of two terminal devices consisting of two metallic or otherwise conducting *plates* with the dielectric medium filling the space between these plates. Figure 3.1 a) shows a rather arbitrarily shaped block of dielectric material with two conducting areas of equally arbitrary shape and area. We may now define the *capacitance* of this device in terms of the ratio of the *charge* Q induced on the plates to the *voltage* V applied between the plates:

$$C = \mathrm{d}Q/\mathrm{d}V \tag{3.1}$$

This definition is completely general, regardless of the geometry of the device used and is independent of the homogeneity and linearity of the material under investigation. It is valid, for example, for a

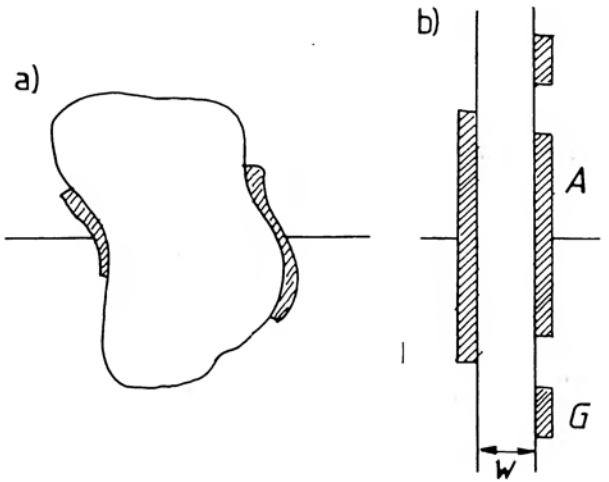


Figure 3.1 Capacitor structures. Diagram a) represents a general configuration of an arbitrarily shaped dielectric body with two metallic electrodes partially covering the surface at opposite faces of the dielectric block. Although the capacitance may be determined exactly and the frequency dependence of the dielectric parameters may be measured, it would be difficult to ascribe exact values to the dielectric permittivity because of the complicated shape. Diagram b) shows a more regular planar geometry of thickness w and with a guard electrode G surrounding the measuring electrode of surface area A, to avoid leakage currents along the surface and to eliminate fringing fields which would distort the uniform field pattern and would lead to errors in the correct assessment of the geometrical factor A/w. This would give an incorrect value of the dielectric permittivity, but it would not affect its spectral behaviour as function of frequency.

dielectric system of such completely indeterminate geometry as an electrolytic capacitor in which one metallic plate has been deeply etched to form intricate tunnels or alternatively consists of a sintered body of metallic powder in order to enhance the effective area of the capacitor. The other "plate" consists of either solid or liquid electrolyte filling the intricate voids in the base electrode. The dielectric medium in this case consists of a thin film of anodically produced oxide whose exact thickness is unknown. We are thus faced with a largely unknown geometry, which does not in any way preclude the usefulness of the device as a capacitor. Another example of an unusual geometry is represented by a p–n junction in a semiconducting material (cf Section 4.8) — here the junction area defines a space charge region from which free carriers have been swept away leaving behind a dielectric medium to which contacts are made by "plates" represented by the neutral semiconducting material on the n- and p-sides. In this particular instance, the

charge is not induced "on" the electrodes but appears as an extension of the space charge region and the system is strongly nonlinear in the applied bias. It is also true to say that there is no direct way of determining the thickness of the dielectric region, although indirect methods may be used.

Although irregularly shaped samples such as those shown in Figure 3.1 a) may sometimes be used, it is far more convenient and advisable to use samples of well defined geometry, such as the parallel-plate structure shown in diagram b), with two plates of area A adjacent to a planar slab of thickness w, usually with the additional stipulation that the lateral dimension of the electrode is much larger than the distance w, so that *fringing effects* may be neglected. These arise because the field is only well defined in the area well removed from the edges and therefore quantitative deductions about the value of the permittivity and loss can only be made on samples in which these fringing fields are negligible.

Subject to the stipulation that $A \gg w^2$, assuming that the dielectric material between the plates is homogeneous and that the electrodes make an intimate contact with this material, and also that the material has a linear response with respect to the amplitude of the applied voltage, we may write the expression for the electric field $E = V/w$, the total charge is given by $Q = A\mathfrak{Q}$, where $\mathfrak{Q}$ is the charge density per unit area of the plate. Making use of the fact that the dielectric induction D represents the total charge density induced on the electrodes by a field E, we may write

$$Q = A\mathfrak{Q} = AD = A\varepsilon E = \varepsilon AV/w \tag{3.2}$$

which with eqn (3.1) gives the expression for the capacitance:

$$C = \varepsilon A/w \tag{3.3}$$

In this manner, a measurement of the capacitance which is readily performed in terms of the measurement of the physical quantities of charge and of voltage, leads directly to the determination of the dielectric permittivity. If the applied voltage is a steady voltage V_0, the capacitance so obtained is the steady state capacitance C_s. If, on the other hand, the applied voltage is an alternating signal at a frequency ω, then the measured capacitance is a complex quantity whose real and imaginary parts correspond directly to the real and imaginary components of the complex permittivity:

$$C(\omega) = C'(\omega) - iC''(\omega) = (A/w)\{\varepsilon'(\omega) - i\varepsilon''(\omega)\} \tag{3.4}$$

$C'(\omega)$ corresponds to the ordinary capacitance, while the imaginary

component $C''(\omega)$ represents the dielectric loss component. In many engineering applications it is customary to define the *loss angle* δ by which the phase of the inducation $D(\omega)$ lags behind the driving voltage $E(\omega)$. The tangent of this loss angle is given by the relation:

$$\tan \delta = C''(\omega)/C'(\omega) = \varepsilon''(\omega)/\varepsilon'(\omega) \tag{3.5}$$

The practical significance of tan δ is that it represents the ratio of the energy dissipated *per radian* in the dielectric to the energy stored at the peak of the polarisation. It also gives a direct indication of the sensitivity required of the measuring system for the assessment of any particular material. However, for the purpose of the critical discussion of the interpretation of dielectric data this quantity is not very significant, since it represents the ratio of two frequency-dependent components which does not convey as much information as the separate plots of the component parts. We shall not therefore use tan δ in our discussion unless we are forced to do so by the absence of other information. In low-loss materials, where the dispersion of the real part of the permittivity is negligible, in view of eqn (2.56), it is evident that the *frequency dependence* of tan δ is the same as that of $\varepsilon''(\omega)$, although the absolute values are different.

One practical advantage of tan δ as a figure of merit of a dielectric material consists in its independence of the geometry of the sample – it is the ratio of two parameters which both contain the same geometrical factor. In situations in which the geometry of the sample is not known, for example in a semiconductor p–n junction or an electrolytic capacitor, so that $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ cannot be determined separately, tan δ remains the only characteristic which can be reliably obtained.

It is convenient for some applications to define the *geometrical capacitance* of a capacitor in terms of the value of capacitance that would be obtained with the same geometry but with the dielectric medium being replaced by free space (vacuum):

$$C_0 = \varepsilon_0 A/w \tag{3.6}$$

so that the ratio of the measured capacitance to the geometrical capacitance gives the ratio of the true complex permittivity to the permittivity of free space, known as the *relative dielectric permittivity*:

$$\varepsilon_r(\omega) = C(\omega)/C_0 = \varepsilon(\omega)/\varepsilon_0 \tag{3.7}$$

the real part of which is known as the *dielectric constant*.

It is customary to describe alternating field and current response in terms of *phasors* or *rotating vectors* in the complex plane – see Appendix 3.1 – and we propose to do so in the following discussion. In those terms the ratio of the current to the voltage is known as the *admittance*:

$$Y(\omega) = I(\omega)/V(\omega) \tag{3.8}$$

and this is, in general a complex quantity, indicating a phase shift ϕ between the two phasors. Physically this ratio is best understood in terms of a fixed reference voltage driving a current through a system, which may be represented by a component I' *in phase* with V and another component I'' *in quadrature* with V, as shown in Figure 3.2. This may be represented by the equivalent circuit of Figure

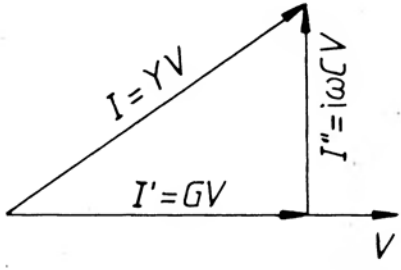


Figure 3.2 The phasor diagram of a current I which is leading the voltage V by an angle ϕ. Y is the admittance of the system and it may be represented by the equivalent circuit of Figure 3.3.

3.3 showing the parallel connection of a conductance G_0 and a capacitance C_0, with the same voltage V driving the current

$$I(\omega) = \{G(\omega) + i\omega C(\omega)\}V(\omega) \equiv Y(\omega)\,V(\omega) \tag{3.9}$$

where the conductance and capacitance are assumed to be *ideal* or *frequency-independent* or *purely real*. In the case of C_0 this means that

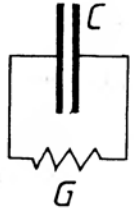


Figure 3.3 A parallel combination of an ideal, frequency-independent capacitance C and conductance G which may represent the system in which the current is as shown in Figure 3.2.

there is no dielectric loss associated with it, since any loss that may be present in a real capacitor would be represented by the conductance G. We see, therefore that an ideal capacitance has an admittance vector pointing at right angles, the current is in full quadrature with the driving field or voltage. For a slightly imperfect capacitor the condition is that the loss current is much smaller than the displacement current, $G_0 \ll \omega C_0$ and we note that $\delta = \pi/2 - \phi$.

It is equally important to define for other purposes the reciprocal of the admittance which is known as the *impedance*

$$Z(\omega) = 1/Y(\omega) = V(\omega)/I(\omega) \tag{3.10}$$

which represents, therefore the *voltage* resulting from the flow in a system of a reference *current*, Figure 3.4,, and corresponds to a series

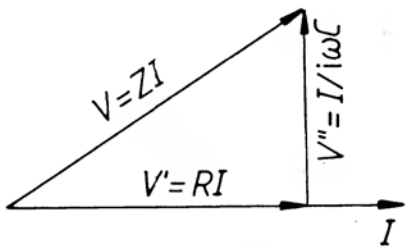


Figure 3.4 The phasor diagram of a voltage V which is leading the current I by an angle ϕ. Z is the impedance of the system and may be represented by the equivalent circuit of Figure 3.5.

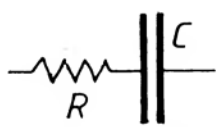


Figure 3.5 A series combination of an ideal capacitor C and a resistor R which may represent the system in which the voltage is as shown in Figure 3.4.

R–C circuit, as shown in Figure 3.5, for which we may write:

$$V(\omega) = \{R_0 + 1/i\omega C_0\} I(\omega) \equiv Z(\omega)\, I(\omega) \tag{3.11}$$

It is clear, therefore, that the admittance representation is the natural way of describing physical phenomena in which two mechanisms exist *in parallel*, so that the same voltage – or field – drives two components of current through the system, while the impedance representation is particularly suitable for those situations in which two physically different regions are *in series* with one another, so

that the same reference current causes two separate voltage drops across separate regions. A typical example of the former situation is the presence of a finite dc conduction mechanism in parallel with the dielectric polarisation which is inevitably present in all materials. The second, series combination is found in all systems where a barrier region is present adjacent to a bulk conducting or semiconducting material – the barrier is depleted of charge carriers and appears as a capacitance, while the bulk looks like a series resistance.

Since $Z(\omega)$ is the inverse of $Y(\omega)$, it follows that the transformation of one into the other may be achieved either analytically or graphically, following the rules of inversion, cf Appendix 3.2. Thus eqn (3.9) is represented as a straight line in the complex Y diagram, and its inversion is therefore a semicircle of diameter $1/G_0$ in the complex Z plane, Figure 3.6. Similarly, the series circuit represented

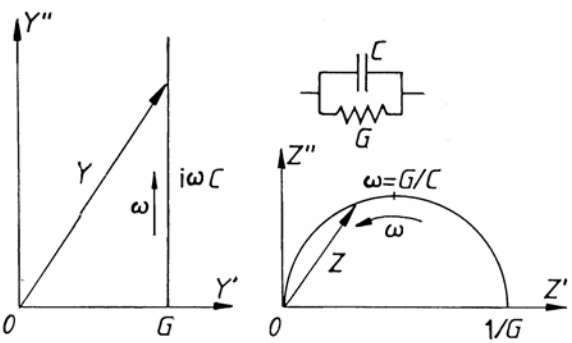


Figure 3.6 The locus of the tip of the admittance vector Y for the parallel circuit shown in the inset is represented by a vertical straight line going through the real axis at $Y' = G$. Its inversion into the impedance plane is a circular arc centred on the real axis and cutting this axis at the origin and at $Z' = 1/G$. The arrows indicate the direction of increasing frequency.

by eqn (3.10) corresponds to a straight line in the complex Z plane and it inverts into a semicircle in the Y plane, Figure 3.7. The corresponding analytical expressions are given below:

$$Y = i\omega C_0 \frac{1 - i\omega\tau}{1 + \omega^2\tau^2}; \quad \tau = R_0 C_0 \quad \begin{matrix} \textit{series } R\text{–}C \\ \textit{circuit} \end{matrix} \tag{3.12}$$

$$Z = (1/G_0) \frac{1 - i\omega\tau}{1 + \omega^2\tau^2}; \quad \tau = C_0/G_0 \quad \begin{matrix} \textit{parallel } G\text{–}C \\ \textit{circuit} \end{matrix} \tag{3.13}$$

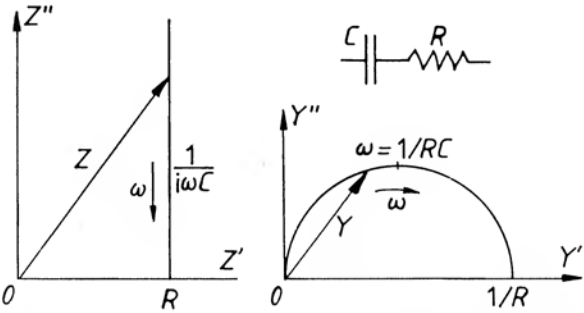


Figure 3.7 The locus of the tip of the impedance vector Z for the series circuit shown in the inset is represented by a straight line, while its inversion into the complex admittance plane represents a circular arc centred on the real axis and cutting it at $Y' = 0$ and $Y' = 1/R$. The arrows indicate the sense of increasing frequency.

It is immediately evident that the "wrong" representations are much more complex than the "natural" ones and this is one of the reasons for trying to find what is the most appropriate representation for any particular situation. We shall return to the treatment of more complicated situations later but for the present we propose to make one further transformation.

We may always *define* a complex capacitance $C^*(\omega) = C'(\omega) - iC''(\omega)$ such that the current is given by:

$$I(\omega) = i\omega C^*(\omega)\, V(\omega) \tag{3.14}$$

from which we infer that with the definition (3.9):

$$C^*(\omega) = Y(\omega)/i\omega = (A/w)\{\varepsilon'(\omega) - i\varepsilon''(\omega)\} \tag{3.15}$$

thus the effective capacitance is related to the admittance by eqn (3.15) and this constitutes the basis for the determination of the effective permittivity of a dielectric system from the measurement of admittance. We must bear in mind at all times, however, that the dielectric permittivity can only be reasonably defined for a *homogeneous* material, so that if there is reasonable presumption that the sample in question is not homogeneous, say that it consists of two or more regions in series, then there is no point in transforming the admittance data directly into permittivity.

With reference to eqn (3.12) we note that the effective capacitance or a series R–C circuit is:

$$C^* = C_0 \frac{1}{1 + i\omega\tau} = C_0 \frac{1 - i\omega\tau}{1 + \omega^2\tau^2}; \quad (\tau = R_0 C_0) \tag{3.16}$$

The significance of this result will become clear in the context of the discussion of the physical models giving the Debye response in Chapter 4.

It is well known that the expression (3.16) represents in the complex plane the semicircular contour as shown in Figure 3.8, which gives the plot of the real against the imaginary part of C^*.

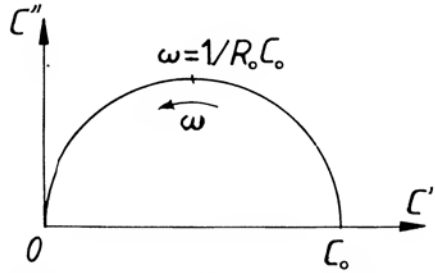


Figure 3.8 The complex capacitance plot, as defined by eqn (3.16) for a series R–C circuit. The imaginary component $C''(\omega)$ has its maximum value at a frequency $\omega = 1/RC$.

3.3 MORE COMPLICATED EQUIVALENT CIRCUITS

Having acquired some acquaintance with the handling of the equivalent impedance, admittance and complex capacitance or permittivity of a few very simple ideal circuits consisting of elements that did not show any inherent frequency dispersion in their own right, we are now in the position to look at a few more complicated circuit combinations. It should be very clearly understood that this exercise is not undertaken in the spirit of trying to interpret every observed dielectric response of *real* materials, as has been only too often the case in the literature on dielectrics. The sole object here is to familiarise the reader with the significance of the existence of series and parallel physical mechanisms and with the effect which these are likely to have on the observed response of dielectric systems, since it is only possible to interpret the physical response of *individual* processes and it therefore becomes imperative to disengage them from one another by suitable manipulation of the respective diagrams.

Thus, if presented with a suspected *series* combination of physical mechanisms, it is advisable to plot the complex Z diagram to see if a clear division into two separate regions is evident. This will always be the case if the series element is a resistance and it may be possible for a series capacitance.

The following examples will also help to illustrate the respective advantages of plotting the frequency-dependent capacitance and conductance, depending upon the presence or absence of conductance in the limit of zero frequency. They will also stress the importance of using the correct representation in order to obtain the simplest form of results – this is exactly equivalent to the use of the correct coordinate system, e.g. cartesian, spherical polar, cylindrical etc in solving analytical problems involving specific geometries.

i) Series R–C *in parallel with* C_∞

It will be shown in Section 4.5 that the response of a series R–C circuit is identical with that of an ideal Debye mechanism. The parallel frequency-independent C_∞ here corresponds to any physical process present together with the Debye process, which has no dispersion in the frequency range of interest – this may be the free space capacitance, or it may correspond to fast responding processes with dispersion at much higher frequencies, e.g. the induced polarisation. The complete response is shown in Figure 3.13 d). The natural form of representation here is the complex Y diagram which consists of a semicircle corresponding to the R–C case and of a vertical "spur" representing the capacitive contribution $i\omega C_\infty$. The addition of these two *must be carried out vectorially* in the complex Argand diagram, taking into account the frequency dependences of the two components. This depends in the present instance upon the ratio C/C_∞, if this is large, then the semicircle is almost completely traversed with increasing frequency before the vertical spur makes a significant contribution, if the ratio is close to unity, the semicircle does not become well developed.

The Z diagram corresponding to this circuit is rather complicated and we do not give it here as it is not the natural form of representation. On the other hand the complex C^* or ε is very simple, since it represents a semicircle corresponding to the R–C circuit, offset by C_∞. The $C''(\omega)$ diagram is exactly as in the case of the series R–C circuit, the reason being that the presence of the parallel free space capacitance does not in any way affect the loss arising from the Debye combination. On the other hand the $C'(\omega)$ diagram shows the limiting capacitance C_∞.

ii) Resistance in series with parallel G–C *combination*

This is a very common situation in which a barrier with some conductance across it is placed in series with a bulk conducting region of resistance R_0. The natural form of representation is the Z diagram, while the Y diagram would be rather complicated. Since

dc conduction path is now evidently present down to zero frequency, the convenient parameter is the effective conductance, $G(\omega)$, together with the real part of the effective capacitance, $C'(\omega)$. The full expressions for these are easily shown, in the limit of a *small series resistance*, to be:

$$G(\omega) = G_0 \frac{1 + \omega^2\tau^2/b}{1 + \omega^2\tau^2} \qquad (3.17)$$

$$b = R_0 G_0 \ll 1$$
$$\tau = R_0 C_0$$

$$C'(\omega) = C_0 \frac{1}{1 + \omega^2\tau^2} \qquad (3.18)$$

and the frequency dependence is shown schematically in Figure 3.9 together with the significant points in terms of the basic circuit parameters.

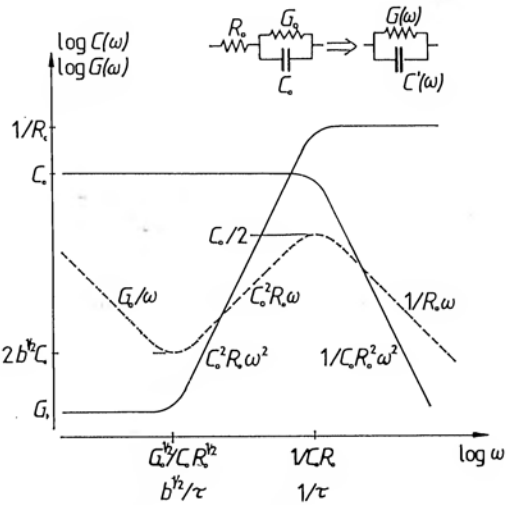


Figure 3.9 The admittance components $G(\omega)$ and $C'(\omega)$ corresponding to the series – parallel combination of frequency-independent components R_0, C_0 and G_0 shown in the inset. The assumption $R_0 C_0 = b \ll 1$ implies that the series resistance has a much smaller value than the reciprocal "loss" conductance G_0. The relaxation time is defined as $\tau = R_0 C_0$. The dotted line gives the equivalent loss component $C'' = G(\omega)/\omega$ which is lowered with respect to its proper level for clarity.

As an illustration we also give the plot of the imaginary part of the effective capacitance, which in the high-frequency region represents an ordinary loss peak:

$$C''(\omega) \simeq C_0 \omega\tau/(1 + \omega^2\tau^2)$$

while at low frequencies it becomes dominated by the conductance component G_0/ω.

At this point it is advisable to note a few general rules regarding the relationship between the imaginary part of the capacitance and the effective conductance:

$$C''(\omega) = G(\omega)/\omega \tag{3.19}$$

and the same rules relate to the corresponding bulk parameters

$$\varepsilon''(\omega) = \sigma(\omega)/\omega \tag{3.20}$$

Rule 1:
A loss peak in $C''(\omega)$ or $\varepsilon''(\omega)$ corresponds to a region of $G(\omega) \propto \omega^2$ followed by a saturated value of conductance.

Rule 2:
A strong dispersion of loss at "low" frequencies, with $C''(\omega)$ or $\varepsilon''(\omega)$ proportional to $1/\omega$ while C' or ε' remains independent of frequency, corresponds to dc conduction.

iii) Capacitance in series with parallel G–C *combination*
This represents an important representation of the physically often occurring situation in which the bulk region characterised by the parallel dc conductance and capacitance is bordered by a "barrier" region in which the dominant element is the capacitance C_s. Since the bulk capacitance C_0 corresponds to the physical dimensions of the entire sample, while the barrier may be presumed to be much thinner than the sample as a whole, it is reasonable to assume that in most physically realisable situations $C_s \gg C_0$. The "natural" representation now is the impedance diagram, shown in Figure 3.13 f) which corresponds exactly to the admittance diagram in c). As the barrier capacitance becomes dominant, so the circular arc is traversed more fully before the "spur" region is entered with decreasing frequency.

The complex admittance now becomes, in the limit of dominant barrier capacitance:

$$Y(\omega) \simeq G_0(1 + i\omega\tau) \quad \begin{matrix} \tau = C_0/G_0 \\ C_s \gg C_0 \end{matrix} \tag{3.21}$$

which is shown in Figure 3.13 f) and which is identical in form to the Z diagram and also to the Y diagram for the combination shown in diagram d). The expressions for the real and imaginary components of the effective permittivity, in the limit of high barrier capacitance, become:

$$\varepsilon(\omega) \simeq C_s \frac{1 + i\omega\tau}{1 + i\omega\tau r} = C_s \frac{1 + r\omega^2\tau^2 - i\omega\tau r}{1 - r^2\omega^2\tau^2} \qquad (3.22)$$

$$\text{with } r = C_s/C_0 \gg 1$$

The loss peak occurs at $\omega_p = 1/r\tau = G_0/C_s$.

The remarkable similarity between the dielectric response of the circuit combinations represented by the diagrams d) and f) in Figure 3.13 should be noted, in the limit of sufficiently separated frequency responses of the two component parts.

iv) Two parallel circuits in series

This is a generalisation of the circuit to cover the possible existence of two different regions, each characterised by a dc conductance and a capacitance and is shown in Figure 3.10. Quite clearly, the cases shown in Figure 3.13 b) to f) represent limiting cases of this more general configuration.

The case of arbitrary values of the two sets of parameters, C_{12} and G_{12} is complicated and we shall only be concerned with the limiting case corresponding to

$$\begin{aligned} &C_1 \ll C_2 \qquad && G_1 \gg G_2 \\ &C_2/C_1 = r \gg 1; \quad && \tau_{12} = C_{12}/G_{12} \end{aligned} \qquad (3.23)$$

The physical significance of this choice of parameters is that region 1 corresponds to the relatively conducting volume of the sample and region 2 represents a highly capacitive but weakly conducting barrier – in fact we are adding a finite conductance to the previous example of a non-conducting barrier.

The impedance may be written *exactly* in the form:

$$Z(\omega) = \frac{1/G_1}{1 + i\omega\tau_1} + \frac{1/G_2}{1 + i\omega\tau_2} \qquad (3.24)$$

which represents the *vectorial* summation of two semi-circles in the complex plane. In the limit represented by the conditions (3.23) the two semicircles are well separated and are as shown in Figure 3.10, with the smaller arc corresponding to the presumed *volume* response, the larger arc describing the *barrier*.

Our assumptions involve the following inequalities:

$$\tau_1 \ll r\tau_1 \ll \tau_2$$

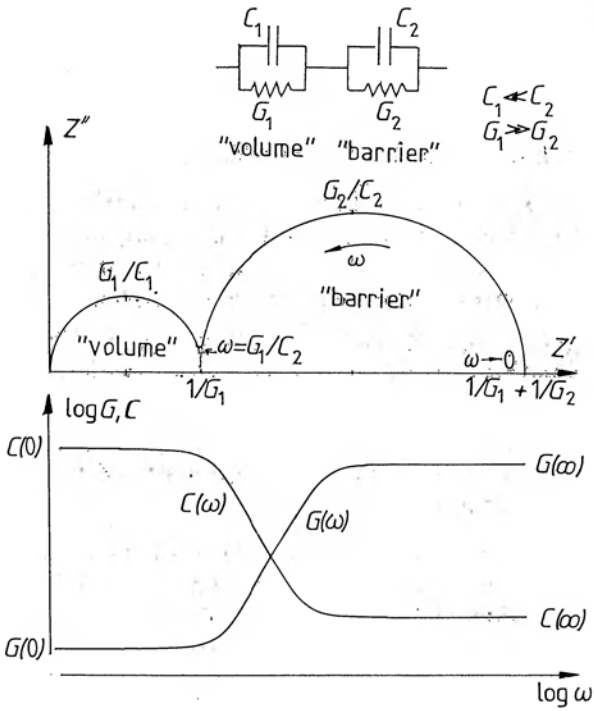


Figure 3.10 The complex impedance representation and the frequency dependence of the effective parallel capacitance and conductance of the series – parallel combination of resistors and capacitors shown at the top. The impedance diagram is drawn for the case of a clear separation of the time constants, $G_1/C_1 \gg G_2/C_2$, which produces the clear separation into two semicircles. The descriptions "volume" and "barrier" denote a physically plausible assignment of the two regions if $C_1 \ll C_2$ and $G_2 \ll G_1$, in which case the two regions may be presumed to be physically connected in series. The exact expressions for the limiting values at zero and at infinitely high frequencies are as follows:

$$G(0) = \frac{G_1 G_2}{G_1 + G_2} \qquad G(\infty) = \frac{C_1^2 G_2 + C_2^2 G_1}{(C_1 + C_2)^2}$$

$$C(0) = \frac{G_1^2 C_2 + G_2^2 C_1}{(G_1 + G_2)^2} \qquad C(\infty) = \frac{C_1 C_2}{C_1 + C_2}$$

The shape of the frequency response graphs should be taken to be schematic only.

and the positions of the frequencies corresponding to the reciprocals of these relaxation times are indicated in the impedance diagram. The inversion of this expression into the admittance plane gives:

$$Y(\omega) = G_2 \frac{(1 + i\omega\tau_1)(1 + i\omega\tau_2)}{1 + i\omega\tau_1 r}$$

but the exact shape of this expression is not easily envisaged. It is more significant that the high frequency region corresponding to the smaller arc in the impedance diagram corresponds to the vertical line

$$Y(\omega) \simeq G_1 + i\omega C_1; \quad \omega \gg 1/r\tau_1$$

while at very low frequencies we are inverting the arc in the impedance plot near $1/G_2$:

$$Y(\omega) \simeq G_2 + i\omega C_2; \quad \omega \ll 1/\tau_2$$

The $C(\omega)$ and $G(\omega)$ plots appear approximately as shown in Figure 3.10. The important point, however, is that the investigation of the high-frequency arc, corresponding to the "volume" response is easily achieved by inversion into the admittance plane and hence by direct determination of the parallel conductance and capacitance. On the other hand, if we wish to look in detail at the behaviour of the barrier, it is advisable to invert the impedance with respect to the point $1/G_1$ on the real axis, i.e. to invert the quantity

$$Z_2(\omega) \simeq Z(\omega) - 1/G_1$$

since the inverse of this is again a straight line in the admittance plane.

This is an important general rule which should be applied to all measurements where there is the suspicion that the behaviour is governed by the interaction of two series elements – the first step should be to represent the original experimental data as the impedance diagram and check if there is evidence of a clear separation into two series regions. If this is the case, then these regions should be inverted separately into the admittance plane, each with respect to the point at which the high-frequency data extrapolate to the real axis, Figure 3.11 a).

Having used the impedance diagram to resolve any unwanted series elements, whether resistive or otherwise, which hamper proper analysis of the remaining data, it may be advisable to use the admittance plot to eliminate any *parallel dc conductance* G_0 that may be present, as shown in Figure 3.11 b). The remaining element may then be considered to be a generalised capacitance $C(\omega)$ *consisting of both real and imaginary components*. The reason for removing the dc conductance is that the latter has nothing to do with the dielectric properties properly speaking and the interpretation of the results would be hampered by the resulting singularity near the origin of the frequency scale.

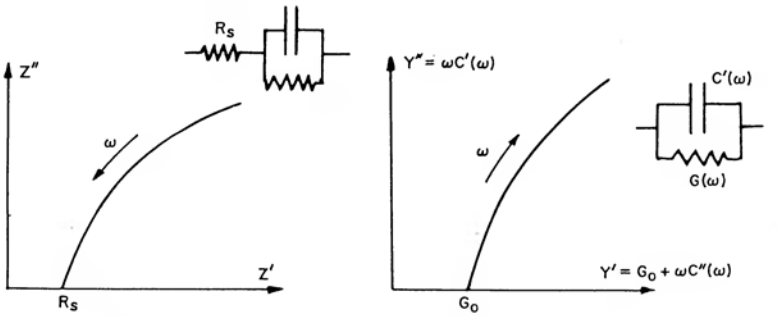


Figure 3.11 a) The use of the complex impedance diagram to determine a series resistance R_s that is independent of frequency and is placed in series with a circuit containing conductance and capacitance, which may themselves be frequency dependent.

b) The use of the complex admittance diagram to separate out the direct current conductance G_0 from the complex ac properties represented by an equivalent parallel circuit.

v) Distributed R–C *line*

Our discussion so far has been concerned with lumped-component networks which are well suited to represent a wide range of physical situations and their understanding is basic to the interpretation of most dielectric data. We shall leave until a later Section the discussion of even more complicated lumped component networks, leading to the contentious issue of continuous or semi-continuous distributions of simple networks. At this point we propose to introduce one special form of distributed network – the uniform distributed *R–C* line, Figure 3.12. This may correspond, for example, to the case of a capacitor in which one or both electrodes have a high sheet resistance. In order to conform to the one-dimensional geometry assumed in the present derivation it is necessary to have a *stripe geometry*, the other likely form of geometry with a circular electrode of high sheet resistance and a central contact to this electrode would require more complicated solutions involving Bessel functions, but the principle of the analysis remains the same.

One reason for the treatment of the linear case is that this gives an exact analogy of the physical process of diffusion, either of heat (Carslaw and Jaeger 1959) or of injected charge carriers, cf Section 4.9.

With reference to Figure 3.12, assume that the resistance per unit length in the x direction which measures the distance from the external contact is r and is constant everywhere, while the capaci-

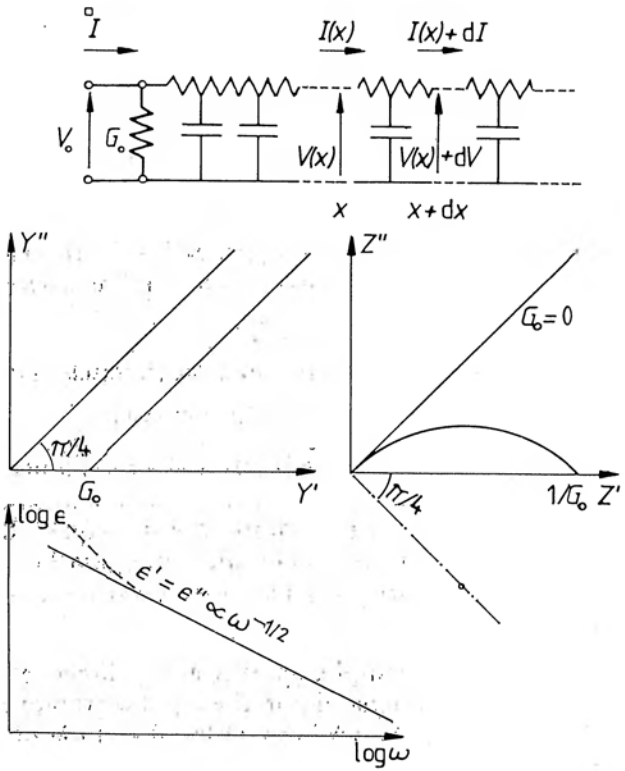


Figure 3.12 A distributed *R–C* line of infinite length, with a parallel conductance G_0 showing the currents and voltages at a distance x from the origin. The effective impedance and admittance diagrams are shown, with zero and finite values of G_0. The real and imaginary components of the effective permittivity are equal and both depend on frequency as $\omega^{-1/2}$.

The dotted contour corresponds to the presence of a finite dc conductance G_0 which causes a rapid rise of $\varepsilon'' \propto 1/\omega$. The distributed *R–C* equivalent circuit shown is an exact representation of the physical processes of diffusion of either heat or of particle densities.

tance per unit length is c. We then have the following differential equations for the voltage and current increments at the coordinate x:

$$\mathrm{d}V = -I(x) r \, \mathrm{d}x$$

$$\mathrm{d}I = -V(x) \, \mathrm{i}\omega c \, \mathrm{d}x$$

which lead to the differential equations for voltage and for current:

$$\mathrm{d}^2 V/\mathrm{d}x^2 = A^2 V(x) \quad \text{with } A = (1 + \mathrm{i})(rc/2)^{1/2}\omega^{1/2}$$

giving

$$V(x) = V_0(\cosh Ax - B \sinh Ax) \tag{3.25a}$$

and

$$I(x) = -(1/r)\,\mathrm{d}V/\mathrm{d}x = -(V_0A/r)(\sinh Ax - B \cosh Ax) \tag{3.25b}$$

Here V_0 is the input voltage to the line and the constant B depends on the boundary conditions at the far end – in the case of an infinitely long line $B = 1$. The current into the line is therefore:

$$I_0 = V_0AB/r$$

and this gives the input admittance into an infinitely long line:

$$Y(\omega) = (1 + \mathrm{i})(c/2r)^{1/2}\omega^{1/2} \propto (\mathrm{i}\omega)^{1/2} \tag{3.26}$$

The complex admittance diagram is also shown in Figure 3.12 and is represented by a straight line at an angle $\pi/4$, implying that the real and imaginary parts of the admittance are equal. If there is a parallel dc conductance, the admittance diagram becomes displaced and the impedance diagram then becomes a semicircle inclined at the angle $\pi/4$.

The importance of this example, apart from its direct relevance to a class of physical phenomena, lies in the fact that it represents the first "breakaway" from the uniform Debye-like type of response offered by lumped-component R–C networks. The complex permittivity is obtained from eqn (3.26):

$$\varepsilon(\omega) \propto (\mathrm{i}\omega)^{-1/2} \tag{3.27}$$

which is also shown in Figure 3.12. This is remarkable in two respects: the real and the imaginary parts are equal and there is no sign of any loss peak – the real and the imaginary components increase indefinitely towards low frequencies.

This type of behaviour will be found highly relevant in the context of our subsequent discussion of dielectric properties of a wide range of materials.

3.4 SUMMARY OF SIMPLE CIRCUIT RESPONSES

We now bring together in Figure 3.13 the dielectric responses of simple networks *consisting entirely of frequency-independent* components, resistance, capacitance and, in one case, inductance. We begin with the series R–L–C circuit, anticipating the result of a more detailed analysis of the harmonic oscillator, which represents its mechanical

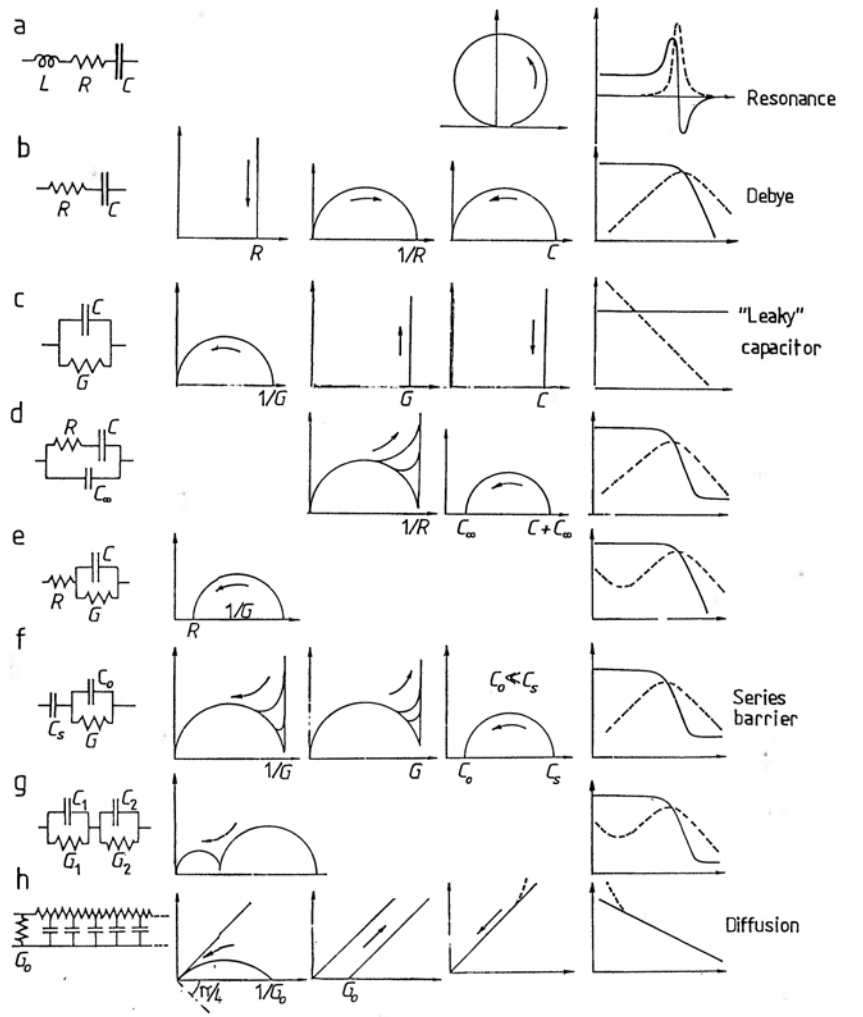


Figure 3.13 Schematic representations of the properties of simple circuit combinations of ideal, frequency-independent elements of capacitance C, conductance G or resistance R, and inductance L, as shown in the first column. The second column gives the complex impedance plot Z, the third the complex admittance plot Y and the fourth the complex capacitance plot C^*. The fifth column gives the corresponding plots of $\log C'(\omega)$ (solid lines) and $\log C''(\omega)$ (dotted lines), against $\log \omega$. Comments refer, where appropriate, to simple physical significance of the various models. Arrows indicate the sense of increasing frequency. Where more than one contour is shown, these refer to varying ratios of the respective components or time constants.

equivalent, to be found later in Section 4.2. The importance of this circuit and of its mechanical equivalent lies in the fact that it is the only circuit which shows the phenomenon of *resonance*, as distinct from the *relaxation behaviour* shown by all the others.

The physical difference between relaxation and resonance lies in the fact that the latter corresponds to an *oscillatory* time domain response, where the discharge current changes sign periodically, and this is associated with the presence in the system of two complementary forms of energy storage. In the *R–L–C* circuit the capacitance stores electrostatic energy while the inductance stores the magnetic energy, in its mechanical counterpart, the compliance stores potential energy, the inertia stores kinetic energy. In the case of an electromagnetic wave we have electrostatic and magnetic energies in the components of the electric and magnetic fields of the wave. These two energies may interchange, producing periodic oscillations which are damped to the extent to which the *dissipative* processes, represented by the resistance R gradually reduce the total energy of the system.

By contrast, the other types of circuit consisting only of capacitances and conductances, cannot transfer energy to another form and they therefore give in the time domain a continuous decay without overswing into the other sense of current flow. Correspondingly, there is no change of the real part of the susceptibility which would imply that the system is swinging in antiphase with the driving signal.

We note that all the relaxation networks, except that corresponding to the distributed *R–C* line in diagram h) and the trivial case of parallel *C–G* in diagram c), give one form or another of a Debye response with some direct current conduction in some cases.

These schematic diagrams show clearly that to every equivalent circuit corresponds a "natural" form of representation giving a very simple description of the dielectric parameters, while other, less "natural" representations lead to considerable complications and the resulting difficulty of interpretation. This should be clearly borne in mind when approaching any dielectric data which do not appear to convey a reasonably simple message – one should strive to find if their appearance can be significantly simplified by adopting a different form of representation.

We shall see later that the great majority of dielectric materials do not show any of the simple responses of the Debye or diffusive types described above. Nevertheless, an examination of their impedance

and admittance graphs before analysing the data in further detail provides a valuable means of eliminating unnecessary confusion and should always be recommended, since the presence of series combinations of dissimilar regions is always a distinct possibility, whatever the detailed nature of the response of the individual regions may be.

One very important consideration should be mentioned with regard to the form of presentation of experimental data. The object of any graphic presentation may be said to be two-fold. At the simplest level, the object is to convey the general *visual impression* of the type of response obtained, especially where a rapid appraisal is needed to check if the response falls into one of well recognised characteristic forms. This is, therefore, almost a "fingerprinting" approach which conveys all the required information to the trained observer looking for a given pattern. Because this is a rapid and familiar procedure, it is often used and it has its proper place, especially at the level of technological assessment of materials.

There is, however, a second, deeper level of assessment, where one is concerned with a serious physical interpretation of the phenomena observed experimentally and one goes beyond the "fingerprinting" approach outlined above. At this level, one is no longer satisfied merely with the determination of a simple equivalent network, or with some empirical mathematical expression, but one is seeking the significance of the observed behaviour at the level of physical mechanisms. To do this effectively, however, one has to go through the process of proper representation of the experimental information in order to obtain the most out of it.

As an example, take the impedance representation of the parallel circuit shown in Figure 3.13 c). Experimental data presented in this form may appear to fit a semicircular contour rather closely, but the eye is not very sensitive to small departures from a circular shape and a much more critical test would be to represent the data in the admittance plot, where any departures from a straight line would immediately become evident. We shall be showing examples of this type of analysis in the following Chapters.

3.5 LOGARITHMIC IMPEDANCE AND ADMITTANCE PLOTS

The availability of very wide ranges of frequency and of time variables in dielectric measurements has long been accepted as

necessitating the use of logarithmic representation, at least for the abscissa of the various plots, even if the ordinates were still plotted linearly, as is generally done in the literature. This was the only meaningful way of representing the whole available ranges of data on a single graph. In addition, we have stressed the value of representing the ordinates in the logarithmic scales as well, for reasons of clarity and conformity with the prevailing power-law dependences of dielectric parameters on frequency. As an example, we note that even the classical Debye response is much more sensitively assessed when presented in the log–log scales, in which the loss consists of two straight lines with slopes ± 1 and a transition region, rather than a curve in the semi-logarithmic representation which is difficult to appraise visually.

All these arguments become even more strongly applicable to the representation of the complex admittance and impedance, both of which may span as many decades of absolute values as there are decades of frequency available. For instance, in the case of two parallel circuits in series, as in Figure 3.13 g), the diameters of the two circular arcs may be in the ratio of several powers of ten and it becomes completely impracticable to represent them both meaningfully on the same linear plot. For this reason, the Chelsea Dielectrics Group have introduced the method of logarithmic plotting of the complex Z and Y diagrams which offers significant advantages in several respects. Since, however, the procedure is unfamiliar to most people, it is advisable to dwell briefly on the logarithmic representations of the simple circuits and to explain some of the implications. Once the reader has become familiar with this approach, other more complicated problems will not represent serious difficulties.

The most important contour occurring in Figure 3.13 is a circular arc with its centre on the real axis and passing through the origin. Denoting the real part by x and the imaginary part by y, we may write the equation of an arc of diameter a

$$x^2 - 2ax + y^2 = 0$$

while the region near the origin may be expanded

$$y = (2ax)^{1/2}$$

which gives

$$\log y = \tfrac{1}{2} \log x + \text{const} \tag{3.28}$$

We note also that the point (0, 0) has the logarithmic coordinates $(-\infty, -\infty)$.

This means that the part of the circular arc next to the origin is represented in the logarithmic plot by a straight line of slope $\frac{1}{2}$ extending to infinity downwards to the left. Now bearing in mind the nature of the logarithmic scale, the part of the circle beyond the peak becomes strongly compressed in the horizontal direction and the point $(2a, 0)$ recedes to minus infinity. The shape of the circular arc in logarithmic coordinates is shown schematically in Figure 3.14.

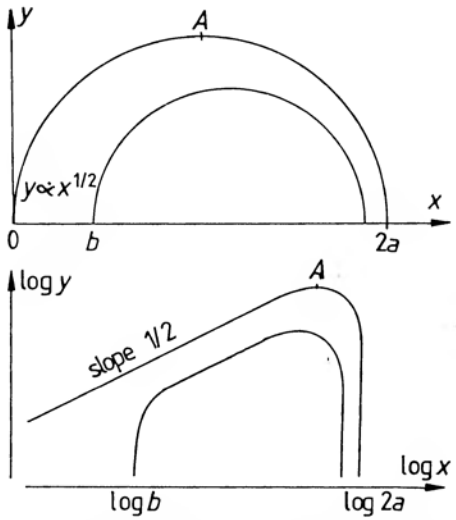


Figure 3.14 The linear representation of a circular arc passing through the origin and of a displaced arc, such as may correspond to an impedance or an admittance diagram. The lower diagram gives the corresponding logarithmic representation, showing the slope $\frac{1}{2}$ in the first case and rapidly dropping characteristic wherever the linear plot cuts the x-axis at a finite value.

However, if the circular arc does not pass through the origin, then the shape of the logarithmic graph is changed as shown in Figure 3.14 with the left part of the arc becoming a vertical line at the appropriate abscissa. It is evident, therefore, that two non-overlapping circular arcs, as in Figure 3.13 g), will be represented by a combination of the former two shapes. An actual example of this type of response is found in Figure 3.15 showing the effect of the amplitude of the applied signal on the dielectric response of a sample which clearly has a barrier region in series with the bulk. At low signal amplitudes the barrier impedance dominates the behaviour and we see the ratio of the diameters of the barrier and volume regions as almost 1000. Despite this very large ratio, the

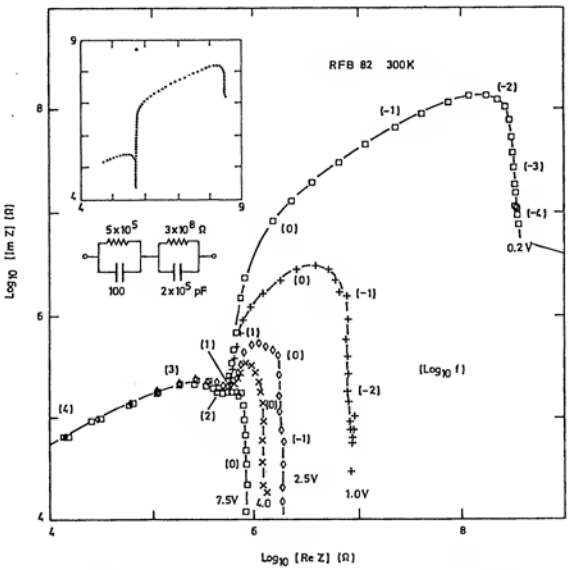


Figure 3.15 The logarithmic representation of the complex impedance plot facilitating the display of a large range of values of Z on one diagram. Data refer to a sputtered film of STAG glass with Al electrodes. The series barrier impedance "collapses" as the signal amplitude increases. RMS signal shown in volts, frequencies in Hz. Inset shows the representation of a matching equivalent circuit with the indicated parameters. From Jonscher and Frost (1976).

small volume response is perfectly resolved on the logarithmic scale, although the linear representation would give this region larger than the diameter of the dot. As the signal amplitude increases, so the barrier becomes less important, because of its strongly non-linear response, and the successive impedance plots show the dramatic collapse of the barrier impedance, while volume impedance remains unchanged within experimental resolution.

It should be noted that with modern digital equipment, this type of plotting of experimental data is made as easy as the linear presentation.

If the circular arc is not straight but is inclined to the horizontal, as in Figure 3.12, then the initial part of the arc is well approximated by a straight line, so that the slope of the logarithmic plot becomes $+1$ before it reaches the peak.

Since

$$\frac{\mathrm{d}\log y}{\mathrm{d}\log x} = \frac{x}{y}\frac{\mathrm{d}y}{\mathrm{d}x}$$

it follows that the apex of the semicircle in linear coordinates corresponds directly to the apex in the logarithmic presentation. This may help to locate the point at which $\omega = 1/\tau$.

3.6 THE RESPONSE OF A "UNIVERSAL" CAPACITOR

We now anticipate the results of our survey of experimental data which will be presented in Chapter 5, by introducing the concept of a power-law frequency response of dielectric relaxation found in a very wide range of dielectric materials, at least in the frequency range above any loss peaks that may be discernible. This can be expressed by the relationship for the complex capacitance $C_n(\omega)$ and the corresponding susceptibility:

$$\chi(\omega) \propto C_n(\omega) = B(\mathrm{i}\omega)^{n-1}$$

$$= B\{\sin(n\pi/2) - \mathrm{i}\cos(n\pi/2)\}\omega^{n-1} \qquad (3.29)$$

in which the exponent n defines the frequency dependence. This power law relation is unique in that the real and imaginary components of the complex susceptibility are *the same functions of frequency* and therefore remain in a constant frequency-independent ratio. This implies that the power law of eqn (3.29) is the only function remaining invariant under the Kramers–Kronig transformation, except for a multiplicative constant. The condition for this is that the exponent should fall in the range

$$0 < n < 1 \qquad (3.30)$$

Eqn (3.29) neglects the contribution C_∞ which should normally be added to the real part of this expression in order to make it correspond to the complete dielectric permittivity of the material in question.

We shall find in Chapter 5 that the response of real materials covers the entire range of the exponent n given by eqn (3.30), with the values in the upper half of this range corresponding to typical *high* frequency behaviour, while values of n below $\frac{1}{2}$ may be observed at *low* frequencies under certain conditions. In the present discussion we are not concerned with the physical significance of the universal relationship, our sole purpose is to describe the type of response to be expected under a range of situations likely to be encountered in experimental conditions, so that the reader may be prepared when he finds them.

One immediate consequence of eqn (3.29) is that the ratio

$$\frac{C''(\omega)}{C'(\omega) - C_\infty} = \chi''(\omega)/\chi'(\omega) = \cot(n\pi/2) \qquad (3.31)$$

is independent of frequency, in complete contrast with, for example, the series R–C circuit response, eqn (3.16) which is the same as the ideal Debye response and for which the ratio is given by $\omega\tau$.

This means that the plots of $\log \chi'$ and $\log \chi''$ vs $\log \omega$ are two *parallel* straight lines with the appropriate logarithmic separation corresponding to eqn (3.31), Figure 3.16.

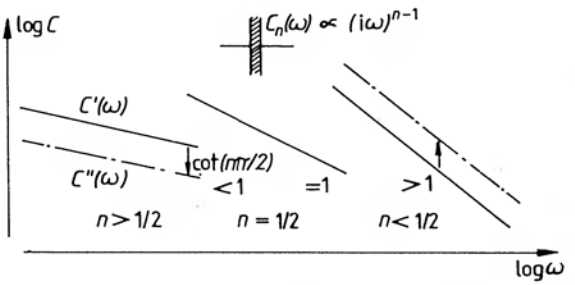


Figure 3.16 Schematic representation of the frequency dependence of the real and imaginary components of the capacitance of a "universal capacitor", in the double-logarithmic scales. Three cases are shown, with the exponent less than, equal to and greater than half. The real part of the capacitance, less any value of C_∞ that may be appropriate, is shown by the solid line, the imaginary part by the chain-dotted line. Negligible dc conductance is assumed.

We note that if $\frac{1}{2} < n < 1$, the real part of susceptibility is higher than the imaginary part, while if $0 < n < \frac{1}{2}$ the converse is true. In the special case of $n = \frac{1}{2}$ the two components are equal and this is the case already discussed in Section 3.3 *v*) corresponding to a distributed R–C line and also corresponding to the diffusive boundary condition to be discussed in Section 4.9. We also note that the limiting case of $n \to 1$ gives *frequency-independent* dielectric loss and real part, with the further implication of a *vanishingly small* loss. We shall see later that this is a special condition which is found experimentally in very low loss materials. Finally, the other limiting condition, $n \to 0$ corresponds to *both* $\chi''(\omega)$ *and* $\chi'(\omega)$ falling with frequency approximately as $1/\omega$, in complete contrast with the parallel dc conductance shown in Figure 3.13 c), where the real part is constant with frequency. This limiting condition will also be found to correspond to a very well defined physical situation, even though it has not until recently been recognised as such.

It should be clearly understood that the two limiting conditions for the exponent n being equal to unity and zero, respectively, correspond to mathematical singularities and are not physically admissible, although they may be approached experimentally within the margin of error.

The ac conductivity corresponding to the universal relation is given by the power law, following eqn (3.20):

$$\sigma(\omega) = a\omega^n + \sigma_0 \tag{3.32}$$

where we have added the dc conductivity which may be present in any given material. We shall see later that the ω^n law is very frequently observed in a wide range of materials, in keeping with its "universal" character.

The sine and cosine Fourier transforms of the universal law are (Erdelyi 1954):

$$\begin{aligned} \mathcal{F}_c(\omega^{n-1}) &= \Gamma(n)\cos(n\pi/2)t^{-n} \\ \mathcal{F}_s(\omega^{n-1}) &= \Gamma(n)\sin(n\pi/2)t^{-n} \end{aligned} \tag{3.33}$$

which shows that the time-dependence of the relaxation current after the sudden removal of a polarising field follows the power law t^{-n} which is widely observed in practice and is known as the Curie–von Schweidler law (von Schweidler 1907) this should be contrasted with the purely exponential law which is the Fourier transform of the Debye law eqn (3.16).

Returning to eqn (3.29) we note that the complex susceptiblity plot corresponding to the universal response represents a straight line inclined at an angle $(1-n)\pi/2$ to the horizontal, Figure 3.17, at least in the high-frequency region in which this universal law is applicable.

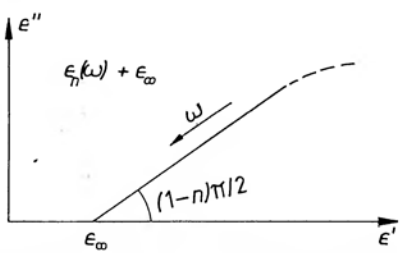


Figure 3.17 A schematic representation of the high-frequency end of the complex permittivity (Cole–Cole) plot for a universal material behaviour, given by $A(i\omega)^{n-1} + \varepsilon_\infty$. The dotted contour indicates the effect of the presence of a loss peak at lower frequencies.

The complex admittance becomes, from eqn (3.15):

$$Y(\omega) = B(\mathrm{i}\omega)^n + \mathrm{i}\omega C_\infty + G_0 \qquad (3.34)$$
$$= G_0 + B\{\cos(n\pi/2) + \mathrm{i}\sin(n\pi/2)\} + \mathrm{i}\omega C_\infty$$

which is shown schematically in Figure 3.18 both in linear and in logarithmic representations. The expression in { } represents a straight line in the linear representation with an angle $n\pi/2$ to the horizontal, while the contribution of the C_∞ term becomes progressively more significant as the frequency increases. A dc conductance G_0 is usually present, especially in the more conducting materials such as ionic conductors.

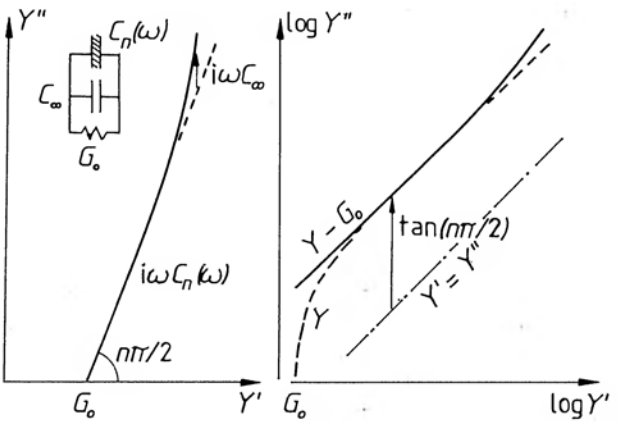


Figure 3.18 The complex admittance diagrams in linear and logarithmic coordinates for a universal capacitor with a constant capacitance C_∞ and a conductance G_0 in parallel. The linear diagram corresponds to a straight line inclined at an angle $n\pi/2$ to the horizontal and shows a departure at high frequencies due to C_∞. The logarithmic diagram is a line at unit slope and at a distance corresponding to the ratio $\tan(n\pi/2)$ from the line $Y' = Y''$. In order to obtain a straight line it is necessary to subtract G_0 from the real part, otherwise the dotted contour is obtained.

We now consider the complex impedance representation. The inversion of the { } term in eqn (3.34) is a straight line with the same slope, remembering that we are conventionally plotting the complex conjugate of the impedance, i.e. we plot Z in the upper half plane, instead of the lower. If we take the dc conductance into consideration, but ignore the C_∞ term, we find that the complex impedance locus becomes a semicircle inclined to the horizontal at the angle $(1 - n)\pi/2$ as shown in the Appendix. We may write:

$$Y = G_0 + B(\mathrm{i}\omega)^n = G_0\{1 + (\mathrm{i}s)^n\} \qquad (3.35)$$

where the variable $s = \omega(B/G_0)^{1/n}$. The complex impedance is therefore given by the semicircle shown in Figure 3.19. We shall see many examples of this type of inclined circular impedance diagram especially for ionic conductors.

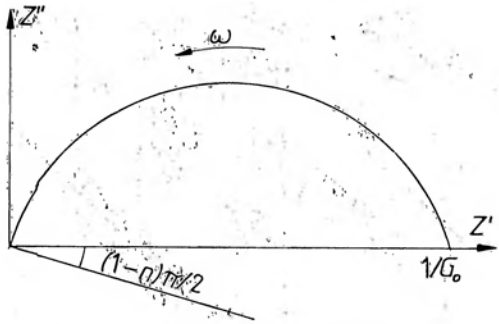


Figure 3.19 The complex impedance diagram of the system shown as admittance in Figure 3.18. To the extent to which the admittance may be represented by a straight line, the impedance diagram is a circular arc inclined to the real axis. An example of this type of response is shown in Figure 5.36.

As an illustration of the various types of dielectric response, we plot in Figure 3.20 the logarithmic impedance diagrams for a series R–C circuit, a semicircular diagram corresponding to the parallel G–C circuit and the universal response for the limiting cases of $Z'' \gg Z'$ or very low loss capacitance and $Z'' \ll Z'$ or almost pure conductance, with the intermediate diffusive case for which $Z'' = Z'$. The completely different character of the universal response should be noted. We also include the inclined circular arc diagram corresponding to the parallel combination of a universal capacitor and a dc conductance.

Consider next the series combination of a universal capacitor, represented by the expression

$$C_n(\omega) = B(\mathrm{i}\omega)^{n-1}$$

with an ordinary resistance R. This is the direct counterpart of the "Debye element" of Figure 3.13 b). It is easily shown that the effective complex capacitance of this combination is given by:

$$1/C^* = (1/B)\omega^{1-n}\sin(n\pi/2) + \mathrm{i}[\omega R + (1/B)\omega^{1-n}\cos(n\pi/2)]$$

Despite its apparent complexity, this expression gives simple results in the limit of low and high frequencies, in relation to $(RB)^{-1/n}$. At low frequencies we find that C^* becomes equal to C_n, in other

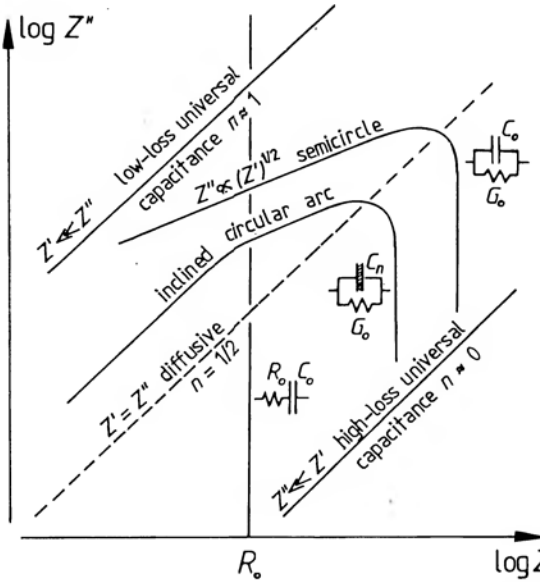


Figure 3.20 A schematic representation in the logarithmic coordinates of the impedance plots for various combinations of resistive elements with ideal or universal capacitances. The ideal components are denoted by the subscript "0". The dotted line $Z' = Z''$ represents a reference – above it the loss component is less than the capacitive component, while the reverse is true below this line. The difference between the semicircle centred on the real axis and the inclined circular arc should be noted. The values of the exponent n corresponding to the various characteristics shown are indicated.

words the series resistance does not have any effect, as might have been expected from elementary considerations. At high frequencies, the expansion of the result gives the approximate relations:

$$C^* = \frac{\sin(n\pi/2)}{BR^2\omega^{1+n}} - \mathrm{i}/(\omega R)$$

which shows that the imaginary part goes as $1/\omega$, just as for the ideal capacitor in series with the resistor, but the real part goes as ω^{-1-n} which becomes equivalent to the Debye combination Fig 3.13 b) when the exponent becomes equal to unity.

The complete frequency response is shown in Figure 3.21, where the diagram a) gives the schematic presentation while diagram b) gives the calculated response for a series of different values of the exponent n. We note that as n approaches unity, in the limit of a loss-less capacitor, the response approximates to that given in

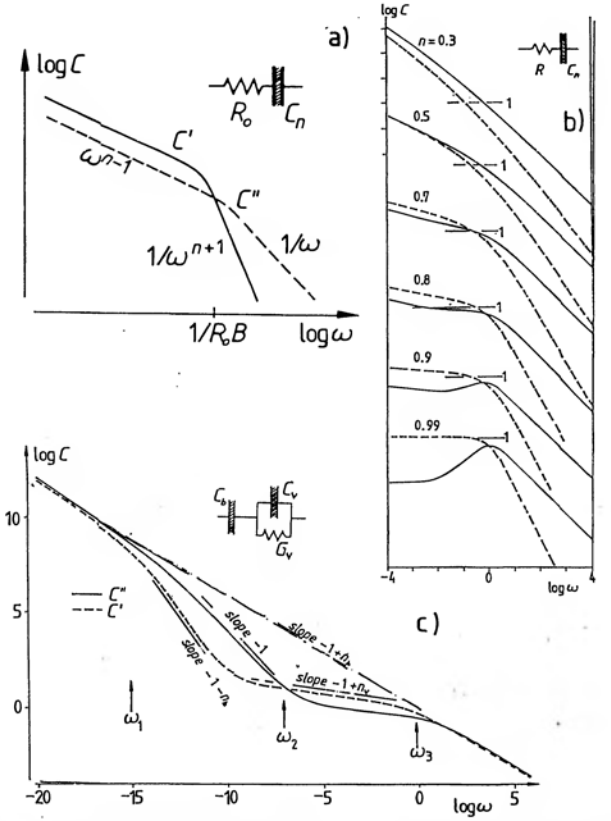


Figure 3.21 The frequency response of circuits involving universal lossy capacitors of the type $C_n = B(i\omega)^{n-1}$. Diagram a) gives the response of a series combination of C_n with a resistor R, diagram b) gives a number of calculated frequency dependences for a range of values of the exponent n. Diagram c) gives the response of a series–parallel circuit shown, with two universal capacitors, one of which corresponds to a series barrier region, while the other forms the equivalent of a volume region with its parallel conductance G_v. The values of parameters assumed in the calculation are as follows: $B_v = 1$, $G_v = 10^{-6}$, $n_v = 0.85$, $B_s = 1$, $n_s = 0.4$. At very low frequencies the volume behaviour is dominated by the conductance and the response is that of a series combination of C_b and G_v, which is therefore closely similar to that seen in Figure 3.21 a). At intermediate frequencies the series resistance gives rise to a slope of $-1 - n_b$ for C' and a slope of -1 for C''. At still higher frequencies the volume capacitance becomes important against the volume conductance and the slope becomes that corresponding to the volume parameter n_v. Finally, at very high frequencies the more steeply falling barrier capacitance becomes once again dominant against the less dispersive volume capacitance. Thus the very low and very high frequency behaviour have a common slope $-1 + n_b$. While the values of frequency corresponding to the various regions are purely arbitrary, the example illustrates the need to have a very wide frequency range if the true behaviour of complicated equivalent circuits is to be properly recognised.

Figure 3.13 b). The individual diagrams are displaced vertically for clarity and the position of the ordinate value 1, corresponding to a loss-less capacitor value, is indicated on each diagram.

Diagram c) represents the behaviour of a series combination of one universal capacitor with a value of exponent n_b, which might correspond to a barrier region, and a parallel combination of another universal capacitor with exponent n_v and with a parallel conductance G_v, representing the volume region of a sample, corresponding to a generalisation of the circuit of Figure 3.13 f). We note that at frequencies $\omega \ll \omega_1 = (G_v/B_b)^{1/n_b}$ the response is dominated by the barrier capacitance C_b; in the interval $\omega_1 \ll \omega \ll \omega_2 = (G_v/B_v)^{1/n_v}$ the series resistance $1/G_v$ with C_b dominates; in the interval $\omega_2 \ll \omega \ll \omega_3 = (B_b/B_v)^{1/(n_v - n_b)}$ the volume capacitance predominates, while at high frequencies $\omega \gg \omega_3$ the barrier capacitance becomes smaller than the volume capacitance and therefore again dominates the response. The series resistance interval is completely analogous to that shown in Figure 3.21 a).

The complementary combination with a series resistance, corresponding to Figure 3.13e) is shown in Figure 3.22 which also shows the corresponding logarithmic impedance plot. Two cases are shown – corresponding to a finite volume conductance and to zero volume conductance for direct current. The volume capacitor is, of course lossy, and there is always a finite alternating current conductivity.

It is relevant to point out at this stage the difference in behaviour between the *series* connection of a dispersive capacitance, as in Figure 3.21 c), and the presence of a *parallel dispersive mechanism* within the same bulk material which may be characterised by its own parallel conductance and high-frequency capacitance. This is especially important in those situations where the capacitance in question is strongly dispersive, with the exponent n close to zero, and therefore has rapidly rising C' and C'' towards low frequencies. Specific examples of this type of behaviour will be found in Section 5.6.

It is clear from an inspection of Figure 3.21 that the presence of a series dispersive capacitor with a more weakly dispersive bulk results in the appearance of a bulge in the plot of $C'(\omega)$ which departs significantly from parallelism with $C''(\omega)$ with its slope close to -1 and only becomes parallel at lower frequencies. By contrast, the presence of a parallel bulk dispersive mechanism manifests itself by a completely parallel run of $C'(\omega)$ and $C''(\omega)$, as seen for example in Figure 5.45. A good contrast between the

two types of behaviour is found in Figure 5.43 where the humid sand shows a typical parallel behaviour, while sand soaked in water shows series behaviour.

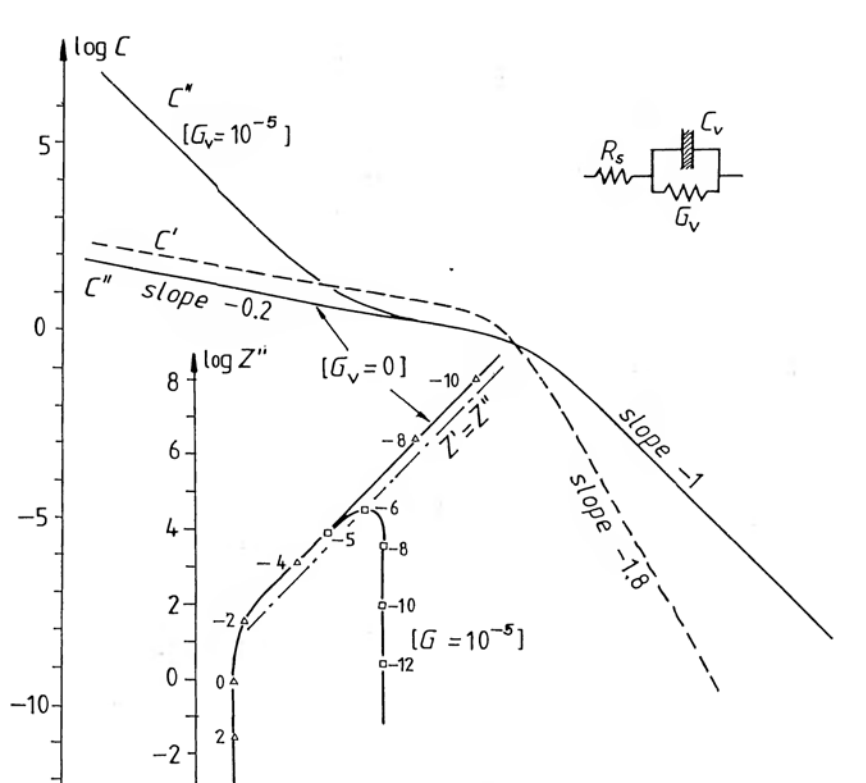


Figure 3.22 The response of a circuit consisting of a series resistance R_s, corresponding to some external or other resistance, and a "volume" capacitance $C_v = B_v(i\omega)^{1-n_v}$, with a parallel conductance G_v. The main plot gives the frequency dependence of the capacitances $C'(\omega)$ and $C''(\omega)$ for the circuit, the inset shows the log–log plot of the complex impedance. The figures on the impedance plot give values of the logarithms of the frequency ω for the particular choice of parameters $B_v = 1$, $R_s = 10$, $G_v = 0$ and 10^{-5}, $n_v = 0.8$. The slopes of the various components of the plots give the power laws of the frequency dependence. The complex impedance plot shows that the ratio Z''/Z' is greater than unity, corresponding to the assumed value of $n_v > \frac{1}{2}$.

3.7 REPRESENTATION IN THE COMPLEX PERMITTIVITY PLANE

We have already mentioned the complex capacitance plot in conjunction with Figure 3.8 for a series combination of an ideal capacitor and a resistor which, as we shall see later, correspond to the classical Debye type of response. The corresponding complex permittivity plot, differing from the complex capacitance by a geometrical factor, has been widely used since its introduction by R H and K S Cole (1941) and is known under the name of Cole–Cole representation. The complete equivalent circuit corresponding to a physically realistic dielectric of Debye characteristics would contain a further ideal parallel capacitor C_∞ to represent the free-space capacitance or the value of permittivity at sufficiently high frequencies where the losses are negligible. This is shown in Figure 3.23 and we note that by shifting the origin of the complex plane to the point ε_∞ we are able to express the complex susceptibility χ^* by the same diagram.

In those cases where the material in question behaves in a manner resembling closely the Debye behaviour, the complex permittivity plot approximates a semicircle and this gives a relatively rapid means of characterising the response, which is given by the complex relation:

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{\varepsilon(0) - \varepsilon_\infty}{1 + i\omega\tau} \tag{3.36}$$

where $\varepsilon(0)$ is the static permittivity and ε_∞ the permittivity at "infinitely high" frequencies, the latter being due to all processes with much higher response rates than the mechanism under consideration at the moment. τ is the relaxation time of the Debye process.

From the presentation of the experimental data in Chapter 5 it will be evident, however, that the behaviour of most dielectric materials departs in varying degrees from the Debye response and it became necessary to modify the empirical expression representing the Cole–Cole plot. One such modification was proposed by Cole and Cole (1941) and it is given by

$$\varepsilon(\omega) - \varepsilon_\infty \propto \frac{1}{1 + (i\omega\tau)^{1-\alpha}} \tag{3.37}$$

where the parameter α denotes the angle of tilt of the circular arc from the real axis, – eqn (3.37) denotes the mathematical operation

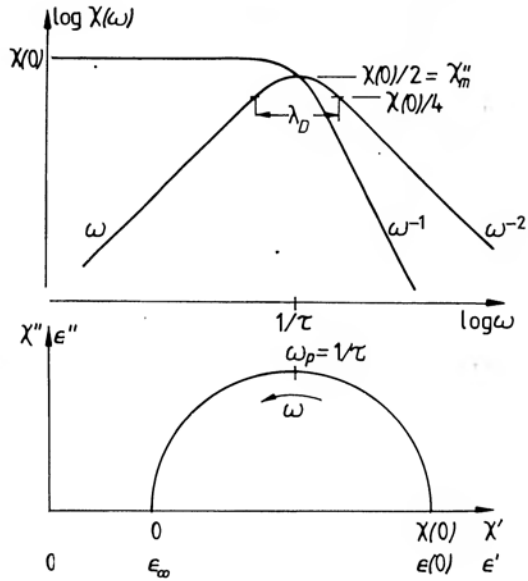


Figure 3.23 The frequency dependence of the real and imaginary components of the susceptibility χ of an ideal Debye system corresponding to eqn (3.36). The loss peak is symmetric in the log ω representation and has a characteristic width at half-height $\lambda_D = 1.144$ decades. The lower diagram shows the corresponding complex χ diagram, as well as the change of coordinates to convert it to the complex permittivity ε diagram.

of tilting the complex plane graph by $\alpha\pi/2$. It is clear therefore, that this parameter has no physical significance whatever, it is a purely empirical factor to describe the appearance of the experimentally obtained Cole–Cole plot.

While the Cole–Cole expression correctly maps small departures from the ideal Debye response, it is not adequate to represent more severe forms of non-Debye behaviour and this has led Davidson and Cole (1951) to propose the following modification:

$$\varepsilon(\omega) - \varepsilon_\infty \propto \frac{1}{(1 + i\omega\tau)^{1-\beta}} \tag{3.38}$$

which gives a pear-shaped form of diagram which has an angle of tilt of the tangent to the plot given by $(1 - \beta)\pi/2$, Figure 3.25. This formula corresponds reasonably well to a number of actually observed dielectric responses. Once again, the parameter β has no direct physical significance.

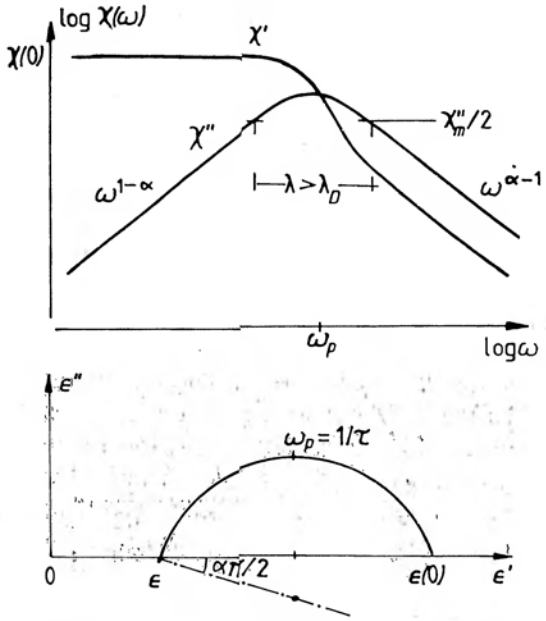


Figure 3.24 The frequency dependence of the real and imaginary components of the susceptibility χ corresponding to the Cole–Cole expression (3.37), with the lower diagram showing the complex plane representation. Note the parallelism of the real and imaginary parts at frequencies above the peak, implying a constant ratio $\chi''(\omega)/\chi'(\omega)$. The loss peak is symmetric but is broader than the Debye peak.

There are many examples of dielectric behaviour which cannot be mapped by either of these last two expressions, both of which contain only one adjustable parameter to describe the shape of the complex plot. For this reason, a further generalisation has been introduced by Havriliak and Negami (1966), consisting in a combination of both Cole–Cole and Cole–Davidson expressions:

$$\varepsilon(\omega) - \varepsilon_\infty \propto \frac{1}{\{1 + (i\omega\tau)^{1-\alpha}\}^{1-\beta}} \tag{3.39}$$

This two-parameter formula is capable of fitting many of the observed results, but its enhanced flexibility cannot be usefully exploited by the complex plane plot as such, since this is in itself an inherently insensitive way of presenting dielectric information, especially in regions of the frequency spectrum away from the loss peak itself. Moreover, the very fact that this plot represents the dielectric information in a form in which the frequency is an implicit variable, restricts the usefulness of the procedure as a complete

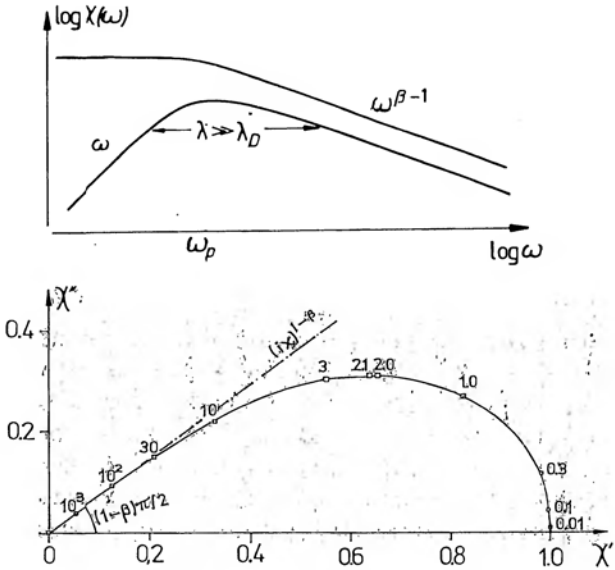


Figure 3.25 The frequency dependence of the real and imaginary components of the susceptibility corresponding to the Cole–Davidson expression (3.38), with the calculated complex plane representation in the lower diagram with $\beta = 0.6$. The loss peak is asymmetric, its low-frequency branch is proportional to ω. The slope of the high-frequency branch depends on β and the relation $\chi''(\omega)/\chi'(\omega) = \text{const}$ applies. The figures refer to the values of the dimensionless frequency parameter $x = \omega\tau$.

means of characterising the data. Far more important, from our point of view, is the frequency–domain representation and it is worth noting here that the original Cole–Cole expression (3.37) corresponds to a broadened but symmetric loss peak in the log-frequency representation, since we may write the following approximate relations:

$$\Delta\chi'(\omega) \equiv \chi(0) - \chi'(\omega) \propto \sin(\alpha\pi/2)(\omega\tau)^{1-\alpha}$$
$$\chi''(\omega) \propto \cos(\alpha\pi/2)(\omega\tau)^{1-\alpha} \qquad \text{for } \omega\tau \ll 1 \qquad (3.40)$$

$$\chi'(\omega) = \tan(\alpha\pi/2)\chi''(\omega) \propto (\omega\tau)^{\alpha-1} \qquad \text{for } \omega\tau \gg 1 \qquad (3.41)$$

The first equation (3.40) defines the *dielectric decrement* $\Delta\chi'(\omega)$, which is always proportional to the dielectric loss in all fractional power laws. The Havriliak–Negami expression gives the respective limiting behaviours:

$$\chi'(\omega) \propto 1 - (1-\beta)\sin(\alpha\pi/2)(\omega\tau)^{1-\alpha}$$
$$\chi''(\omega) \propto (1-\beta)\cos(\alpha\pi/2)(\omega\tau)^{1-\alpha} \qquad \omega\tau \ll 1 \qquad (3.42)$$

TABLE 3.1
The summary of various spectral functions and their power-law exponents* (from Hill and Jonscher (1983)).

Process	Susceptibility function	exponent for $\omega \ll \omega_p$		exponent for $\omega \gg \omega_p$	
		$\Delta\chi'(\omega)$	$\chi''(\omega)$	$\chi'(\omega)$	$\chi''(\omega)$
One parameter					
Debye	$(1 + i\omega/\omega_p)^{-1}$	2.0	1.0	-2.0	-1.0
Cole–Cole (1941)	$(1 + \{i\omega/\omega_p\}^{1-\alpha})^{-1}$	$1-\alpha$	$1-\alpha$	$\alpha-1$	$\alpha-1$
Fuoss–Kirkwood (1941)†	$2(\omega/\omega_p)^{\gamma}(1 + \{\omega/\omega_p\}^{2\gamma})^{-1}$	γ	γ	$-\gamma$	$-\gamma$
Davidson–Cole (1951)	$(1 + i\omega/\omega_p)^{-\beta}$	2.0	1.0	$-\beta$	$-\beta$
Williams–Watts (1970)	$\sum_{s=1}^{\infty} \frac{\Gamma(\Delta s)}{(s-1)!} \left\lvert \frac{\exp(-i\Delta\pi/2)}{\omega^{\Delta}\omega_p^{-\Delta}} \right\rvert^{s}$	2.0	1.0	$-\Delta$	$-\Delta$
Two parameter					
Havriliak–Negami (1966)	$\{1 + (i\omega/\omega_p)^{(1-\alpha)}\}^{-\beta}$	$1-\alpha$	$1-\alpha$	$-\beta(1-\alpha)$	$-\beta(1-\alpha)$
Jonscher (1975)	$\{(\omega/\omega_1)^{-m} + (\omega/\omega_2)^{1-n}\}^{-1}$	m	m	$n-1$	$n-1$
Hill (1978)	$\omega^m(\omega_p^{2s} + \omega^{2s})^{-(m+1-n)/2s}$	m	m	$n-1$	$n-1$
Dissado–Hill (1979)	$(1 + i\omega/\omega_p)^{n-1} \times {}_2F_1\left(1-n, 1-m; 2-n; \frac{\omega_p}{\omega_p + i\omega}\right)$	m	m	$n-1$	$n-1$

* All the spectral parameters α, β, Δ, γ, m and n are fractional and positive.
† The Fuoss–Kirkwood relationship gives only the imaginary component.

and

$$\chi'(\omega) = \tan(n\pi/2)\chi''(\omega) \propto (\omega\tau)^{n-1}$$

where $\omega\tau \gg 1$ (3.43)

$$n = \alpha + \beta - \alpha\beta$$

This shows two different logarithmic slopes at low and high frequencies. The Cole–Davidson expression corresponds to $\alpha = 0$ in these formulae, which means that the low-frequency slope is unity and the high-frequency slope is $\beta - 1$. The Cole–Cole formula corresponds to $\beta = 0$.

We have mentioned already the principal limitations of the complex plane representation as the loss of explicit frequency information and the loss of accuracy in the assessment of the response at longer distances from the loss peak frequency. The most important limitation, however, is the absence of any particular physical model which is capable of explaining the postulated forms of dependence, so that the exercise of fitting the various parameters amounts to purely formal "fingerprinting" of the response and cannot lead to an elucidation of the underlying physical reality since the parameters α and β are not based on the physics of dielectric interactions.

Several other empirical expressions have been proposed to represent the dielectric relaxation function, either in its entirety or the imaginary component only. Among these may be mentioned the Fuoss–Kirkwood (1941), Williams–Watts (1970), Jonscher (1975) which is given by eqn (5.3), and Hill (1978). We shall see later in Chapter 5 that only those having two parameters, α and β, m and n in the exponents can represent the majority of experimentally determined relaxation functions.

Table 3.1 gives a summary of these various functions and the exponents in the corresponding power law relations for $\Delta\chi'(\omega)$ and $\chi''(\omega)$ in the limits of the "high" and "low" frequency approximations. The Table also gives the analytical expression derived by Dissado and Hill (1979) from their many-body model of dielectric relaxation, which is discussed in Chapter 8.

3.8 REPRESENTATION OF THE TEMPERATURE DEPENDENCE

The dielectric permittivity is a complex function of at least two variables – frequency and temperature, although pressure may be

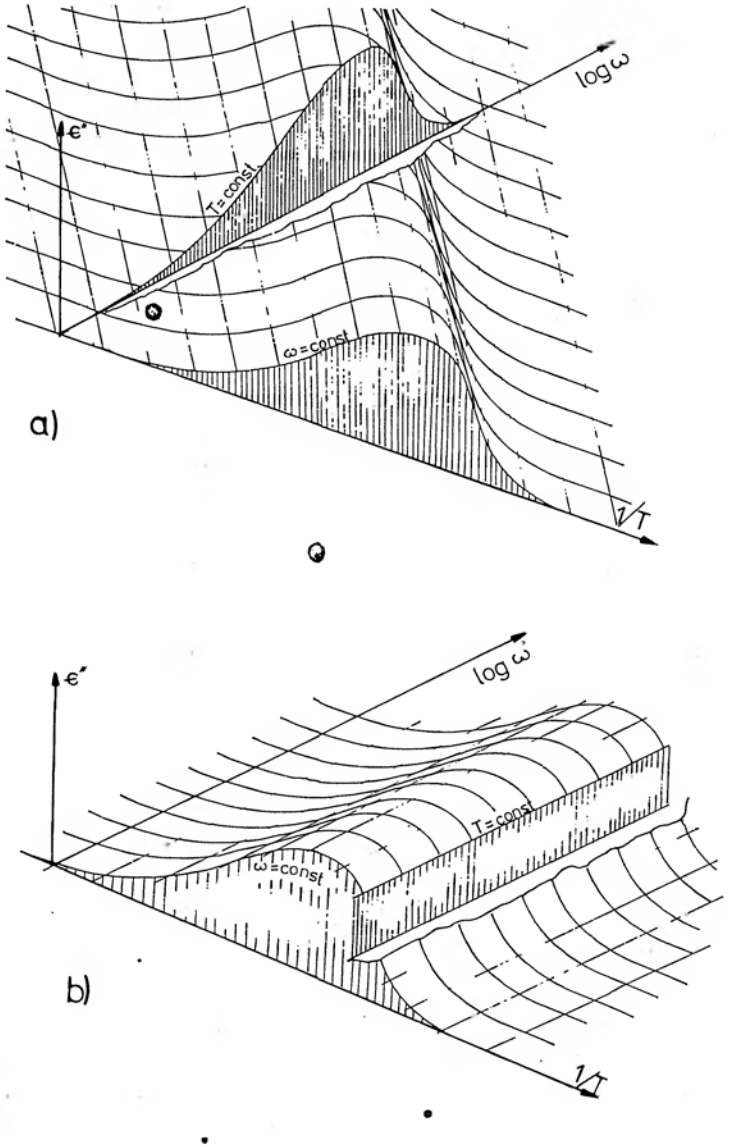


Figure 3.26 A schematic diagram of the three-dimensional relation $\varepsilon''(\omega, 1/T)$ for the limiting case of an ideal Debye response (a) and for the temperature–dependent but frequency–independent "non-Debye" response – diagram b). The measurement of the temperature dependence at a constant frequency would give very similar results in both cases, from which it could be inferred mistakenly that the material is Debye-like in each case. From Jonscher (1978).

another physical variable. A complete representation should therefore comprise two "three-dimensional" plots of $\chi'(\omega, T)$ and $\chi''(\omega, T)$ but these are cumbersome and are therefore seldom employed, although modern computer graphics enable one to plot third-angle projections in two dimensions.

The prevailing method of representation consists therefore in plotting the frequency dependence with temperature as parameter or vice versa. In the special case of an ideal Debye response characterised by an exponentially temperature dependent relaxation time the eqn (3.16) may be written in the form:

$$\chi(\omega) \propto \frac{1}{1 + i\omega\tau_\infty \exp(W/kT)} \tag{3.44}$$

where τ_∞ is a suitable pre-exponential factor and W is the activation energy. From this expression it is clear that the susceptibility is the same function of frequency as it is of the exponential of W/kT, or in another representation, the same function of $\log_e\omega$ as of W/kT. It makes no difference, therefore, whether one is plotting the frequency response or the temperature response in the appropriate coordinates, one obtains an identical loss peak which moves to higher temperatures with increasing frequency, or conversely, to higher frequencies with increasing temperature. This is shown in the schematic representation of Figure 3.26.

The problem becomes rather more serious if the response is not of the Debye type, since there is then no *a priori* way of relating the temperature dependence to the frequency dependence in any unambiguous manner and in these circumstances it is *far more meaningful to plot the frequency dependence* with temperature as a parameter, than the other way round.

It is found quite generally that the frequency dependence does not change very drastically with temperature, at least over temperature ranges over which the material does not alter its structure in any significant way. This means that it is often possible to *normalise* the data for different temperatures by shifting the frequency spectra laterally into coincidence obtaining sometimes a single "master curve" which gives a complete description of the behaviour when accompanied by the locus of the translation point (Hill and Dissado 1982).

The technique consists in taking a family of loss spectra in log–log representation, with temperature as parameter, Figure 3.27 a), and placing a tracing paper over it. An arbitrary point "A" is marked

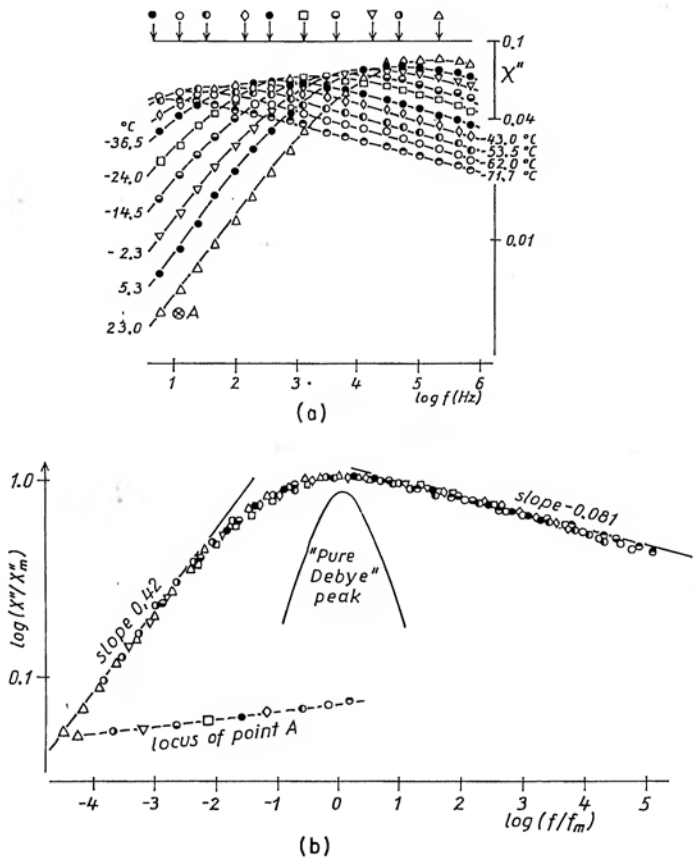


Figure 3.27
a) Dielectric loss peaks for polyethylene terephthalate (5% crystalline) in the lower temperature region (after Ishida et al 1962).
b) The normalised loss curve obtained by lateral displacement of the loss curves in diagram a). The dotted line represents the locus of the marker point 'A' in diagram a). The theoretical shape of a pure *Debye* loss peak is also indicated. From Jonscher (1975).

on the original sheet, although it may be advantageous to record its coordinates of frequency and loss amplitude for future reference. The $\log f$ and $\log \chi''$ axes are then marked on the tracing paper and the corresponding scales are traced. Taking the set of data corresponding, say, to the highest temperature, the experimental points are traced on the tracing paper, together with the position of the reference point "A". The tracing sheet is next translated laterally and vertically, *parallel to itself*, until the data for the next temperature run overlap as closely as possible with the first set of data, they are

then recorded on the tracing sheet and the new position of the reference point is marked. This procedure is repeated as many times as there are temperature runs, until a complete *master curve* is obtained. When the spectral shape of the loss characteristic does not change significantly with changing temperature, the master curve is uniquely determined. Its spectral range is larger than the measurement range by the amount of the shift of the reference point along the frequency axis, so that the resulting curve contains much more information than any single temperature run. Moreover, the density of data points on the master curve is significantly higher than on the original temperature runs, thus increasing the reliability of the information contained in the master curve, as may be seen in Figure 3.27 b).

The locus of the reference point "A" gives information about the frequency dependence – the horizontal translation – and the amplitude dependence – the vertical translation, of the loss spectra, while the knowledge of the scaling of the axes provide complete information about the absolute values of the loss spectra for any temperature. The information provided in a master curve and in the reference point locus enables, therefore, the complete reconstruction of the original data.

Very many sets of experimental data may be satisfactorily normalised into reasonably single–valued master curves, thereby *proving* that the spectral shape of the loss characteristics *remains invariant* with temperature. This property implies that the complete frequency- and temperature-dependence of loss may be represented by the *product*

$$\chi''(\omega, T) = \chi(0) F[\omega/\omega_p(T)] \tag{3.45}$$

where $\chi(0)$ denotes the temperature-dependent *amplitude factor* and $F[\omega/\omega_p(T)]$ is the *spectral shape function* of the normalised frequency ω/ω_p, where $\omega_p(T)$ is the temperature-dependent loss peak frequency which is usually strongly temperature dependent and may show the characteristic activated or Arrhenius behaviour implied by eqn (3.44):

$$\omega_p(T) = \nu_0 \exp(-W/kT) \tag{3.46}$$

with an energy of activation W. By plotting the lateral translation of the reference point, which is already on a logarithmic scale, against $1/T$ it is possible to determine whether this simple behaviour is, in fact, obeyed or whether a more complicated temperature dependence prevails, and there are many instances of this situation, as will be seen in Chapter 5.

It is common to find the normalisation procedure carried out in semi-logarithmic representation, in which case it is not generally possible to obtain more than a reasonable coincidence of the immediate neighbourhood of the loss peak itself, since the shape of the curved contours of loss does not admit of satisfactory normalisation of the "wings" of the peak. One is therefore in a position to determine the activation energy of the loss peak frequency and also the amplitude dependence on temperature, but little can be said about the spectral shape $F(\omega/\omega_p)$.

We conclude, therefore, that the normalisation in the log–log representation, leading to the derivation of the product rule, eqn (3.45), corresponds to the natural behaviour of dielectric systems and this procedure is strongly recommended.

We have concentrated so far on the loss component, $\chi''(\omega)$, of the complete dielectric response since this shows more pronounced features than the corresponding real part of the susceptibility, $\chi'(\omega)$, being odd and even functions of frequency, respectively. It is clear, however, that the real part contains in principle the same information as the imaginary part and parallel normalisation of the real data may considerably strengthen the process of data presentation and interpretation, especially if there are complicating features in the acquisition of such data. It may happen, for instance, that the dielectric behaviour is overshadowed by a strong direct current conduction which masks the loss behaviour by the G_0/ω term. In this situation the processing of the real part is the only way forward towards a sensible interpretation of data. Examples of this will be seen in Chapter 5.

It is important to note that the real part of susceptibility must be obtained from the measured permittivity – or capacitance – data, by subtracting a suitable value of ε_∞ or C_∞ and this requires some judgement. If, as is often the case, the loss characteristic follows a well defined power law, then it is a simple matter to determine the Kramers–Kronig-compatible position of the real part at the *highest* frequency in any one loss curve. It is then possible to determine the value of ε_∞ required to bring this point to its proper position, and to use this value in subtracting from all other $\varepsilon'(\omega)$ points at lower frequencies for *the same temperature* run. The reason for restricting this operation to one temperature is that the value of ε_∞ may be temperature dependent. One repeats this operation for the other temperatures, in each case aiming to place the corresponding highest-frequency point in the Kramers–Kronig-com-

patible position with respect to the loss curve. Having obtained the set of ε_∞ values for the various temperatures it is then possible to normalise the resulting $\chi'(\omega)$ curves, using *the same displacement* of the reference point as for the loss normalisation.

Just as ε_∞ may obscure the behaviour of the real part of susceptibility so the presence of a dc conductivity may mask the low-frequency behaviour of dielectric loss. The term σ_0/ω is usually unmistakable and it may be relatively simple to subtract the corresponding values from the measured loss, in order to obtain a "true" loss response. However, there are sometimes problems in determining exactly what value to subtract, since the resulting shape of the rising part of the loss peak is very sensitive to that. By subtracting too low a value of σ_0/ω one obtains a very shallow loss peak which may turn up again towards low frequencies, by subtracting too high value one obtains a steeper peak which may become negative. The evaluation of the dielectric decrement $\Delta\chi'(\omega)$ may help here by defining the slope of the rising part of the loss curve.

One may attempt to normalise the loss data even without subtracting the dc conductivity where this is present, but it should be clearly understood that the *lateral* shift of the characteristic with the -1 slope does not correspond to any physical change of frequency, since no such frequency shift can arise from the temperature dependence of $\sigma_0(T)$. It is fortuitous that with the slope of -1 in log–log representation, the lateral shift is exactly the same as the corresponding vertical shift, which should have been effected. The activation energy determined from the lateral shift does correspond, therefore, to the proper activation energy of the dc conductivity.

The use of the real part of the susceptibility in the processing of experimental data becomes particularly important in cases where the loss appears to show a -1 slope that might be interpreted as dc conductivity, but the real part likewise shows signs of rising steeply towards low frequencies. This is the case of strong low-frequency disperison, which will be described in detail in Section 5.6 and which should be clearly distinguished from the genuine dc conduction, for which the real part does not show any dispersion.

If more than one loss process is in evidence in a given set of experimental data, normalisation procedures may still be applicable, especially if the range of temperatures and frequencies available is sufficiently large, but one should clearly bear in mind the fact that two separate processes are likely to have different values of the corresponding activation energies or, more generally, have

different temperature dependence if they are not simply activated. In this case, the separate parts, e.g. two loss peaks, or a loss peak and a dc process, will "move" at different rates in the lateral direction with varying temperature, so that normalisation is only possible over restricted frequency ranges corresponding to the dominance of one of these processes at a time. The resulting normalisation curve becomes therefore diffuse in the regions corresponding to transitions from one process to the other and it is impossible to be very precise in determining the exact frequency shape in these situations. Examples of this will be seen in Chapter 5.

In cases where the shape function itself changes either continuously or discontinuously with temperature, normalisation can at best determine the variation of the loss peak frequency and amplitude with temperature, no other conclusions can be drawn. The physical significance of this type of behaviour will be discussed in Chapter 8.

Properly used, the normalisation technique represents one of the most powerful analytical tools at the disposal of the students of dielectric behaviour. It can greatly strengthen the reliability of experimental data and may extend very significantly the range of frequencies available for interpretation of data. It should also be mentioned that the same normalisation procedures may be used in cases where the external variable is not temperature but, for example, pressure, humidity, composition or other variable which affects the rate processes in the system under study. Examples of this will be shown in Chapter 5.

The evaluation of activation energies is facilitated by the use of the expression for W in terms of the change $\Delta(10^3/T)$ per one decade of the frequency shift:

$$W = 0.198/\Delta\,(10^3/T) \quad \text{eV} \tag{3.47}$$

if, as is customary, the reciprocal temperature scale is expressed in terms of $10^3/T\,\mathrm{K}^{-1}$.

Temperature is not the only external variable which influences the relaxation spectra in the manner described above. Exactly similar effects may be seen with pressure, the increase of which slows down the rate processes, without affecting the shape of the spectra. An example of this is shown in Figure 3.28 and relates to data obtained by Williams et al (1972) on a blended copolymer Acrilonitrile–Butadiene. Two sets of measurements were carried out at different temperatures, each at four different pressures. All data normalise

very satisfactorily into a single master set of curves for the real and the imaginary components of the susceptibility which form a Kramers–Kronig-compatible set. These data show, in particular, that the shape of the response remains invariant under the effects of both temperature and pressure, and that the effects of rising pressure may be nullified, as far as the loss peak frequency is concerned, by an increase of temperature.

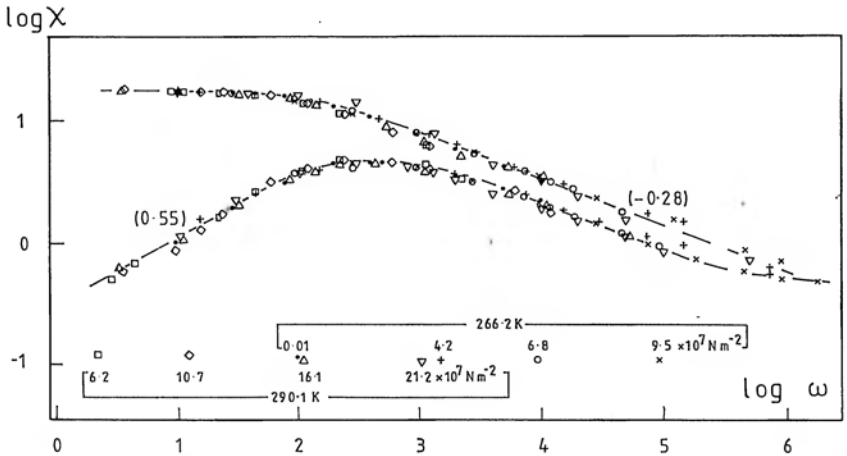


Figure 3.28 Normalised susceptibility plots for 40% Acrilonitrile–Butadiene polymer from data by Williams et al (1972), taken at two temperatures of 266.2 and 290.1 K, each over a range of pressures. The normalisation shows that temperature and pressure have exactly equivalent effects, the influence of an increased pressure may be cancelled by a rise in temperature. In all cases the shape of the loss and real part of susceptibility remains invariant. Plotting of data and normalisation by R M Hill from the original measurements.

Another external variable which may influence the response of certain dielectric systems is humidity, the effects of which are illustrated, together with a normalisation procedure, in Figure 5.43.

The application of normalisation procedures to the alternating current conductivity in the frequency domain is illustrated in Figure 5.41, while the complementary process of normalising time–domain data on relaxation current is shown in examples in Chapter 6. A special case of non-linear voltage dependence of current is shown in Figure 6.13, where the normalisation results in the derivation of a single master curve. In this particular instance the variables are the time for the "kink" and the amplitude of the current.

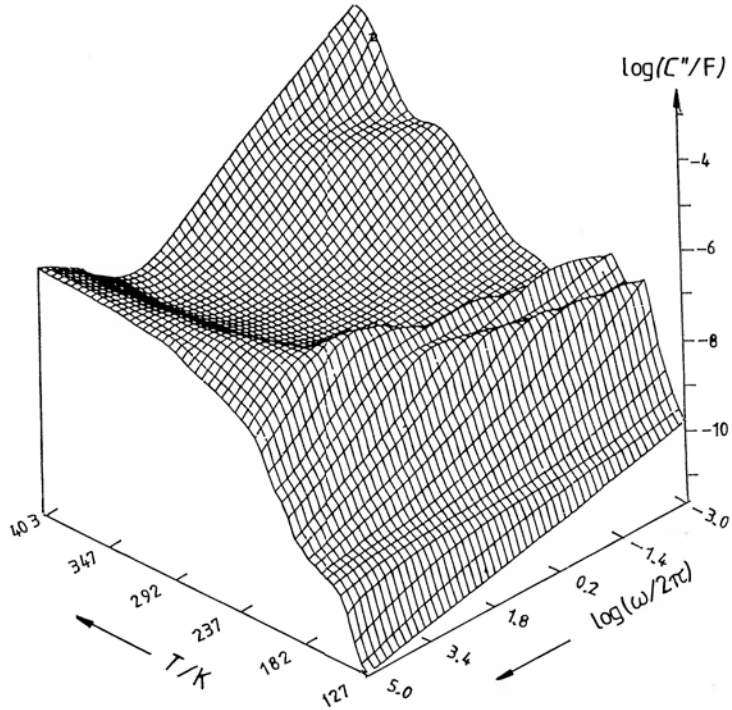


Figure 3.29 A computer graphic representation of the dependence on frequency and on temperature of dielectric loss in a tantalum foil electrolytic capacitor. From Meca (1981).

A powerful and visually very effective way of presenting experimental information on dielectric loss and polarisation consists in three-dimensional computer graphics in third angle projection, which can be generated automatically from numerical input data with frequency and, say, temperature, as variables. An example of this method of presentation is shown in Figure 3.29 which refers to loss in a tantalum foil capacitor. This method enables a rapid appraisal of a complex situation to be made more easily than can usually be done with a family of two-dimensional curves.

APPENDIX 3.1

TIME DOMAIN, ROTATING VECTORS AND FREQUENCY DOMAIN

We should always remember that

Mother Nature works with real quantities in the time domain

It may be convenient mathematically to define complex quantities such as the dielectric susceptibility and to speak of rotating vectors in the complex plane, but these are only mental images of real physical processes proceeding in real time. It is therefore instructive to work through the time domain formulation and arrive at the frequency domain formalism in order to better appreciate the significance of the various steps involved.

Let the applied field be given by the harmonic relation:

$$E = E_0 \cos \omega t \tag{1}$$

where E_0 denotes the magnitude of the field at its peak. Taking eqn (2.26) and adding the free space contribution which is always exactly in phase with the applied field, we get:

$$\begin{aligned} D(t) &= \varepsilon_0 E(t) + P(t) \\ &= \varepsilon_0 E_0 \left[\cos \omega t + \int_0^\infty f(\tau) \cos \omega (t - \tau) \mathrm{d}\tau \right] \end{aligned}$$

Expanding the cosine terms and using the definitions (2.38), (2.39) and (2.48), we get:

$$D(t) = [\varepsilon'(\omega) \cos \omega t + \varepsilon''(\omega) \sin \omega t] E_0 \tag{2}$$

which is a purely real quantity, as it should be because it denotes a physical entity in real time. The corresponding current density is obtained from eqn (2.44) in the form:

$$I(t) = \{[\sigma_0 + \omega \varepsilon''(\omega)] \cos \omega t - \omega \varepsilon'(\omega) \sin \omega t\} E_0 \tag{3}$$

Note that sin ωt *lags* behind cos ωt by an angle $\pi/2$, $-\sin \omega t$ *leads*. Hence $D(t)$ has a component lagging behind $E(t)$, the current leads $E(t)$.

We now note with reference to Figure 3.30:

$$\cos \omega t = \mathcal{R}\, e^{i\omega t}$$

$$-\sin \omega t = \mathcal{R}\, ie^{i\omega t}$$

where $\mathcal{R}$ denotes "the real part of".

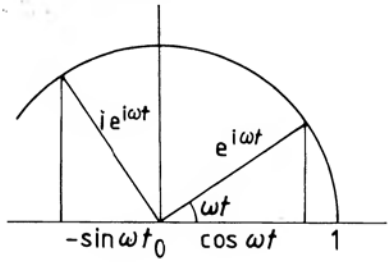


Figure 3.30

We may therefore write in our convention:

$$I(t) = \mathcal{R}[\sigma_0 + \omega\varepsilon''(\omega) + i\omega\varepsilon'(\omega)]\, e^{i\omega t} E_0 \tag{4}$$

The quantity in the square braces is complex and we may define a complex quantity:

$$I_0 = [\sigma_0 + \omega\varepsilon''(\omega) + i\omega\varepsilon'(\omega)]E_0 \tag{5}$$

where E_0 is a purely real quantity. The ratio I_0/E_0 is shown in Figure 3.31. This corresponds exactly to the notation in eqn (2.46) which referred to the components of a Fourier spectrum. The angle ϕ denotes the angle by which the phase of the current leads that of the applied field.

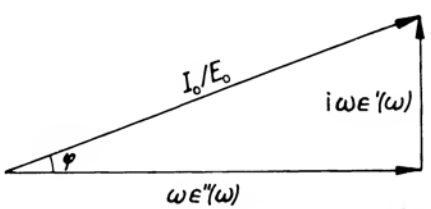


Figure 3.31

The current density may therefore be represented by the symbolic expression:

$$I(t) = \mathcal{R}[I_0\, e^{i\omega t}] \tag{6}$$

The widely accepted rotating vector or "phasor" notation represents real time-dependent quantities as real parts of vectors in the complex plane rotating with an angular speed ω (radians per second).

The vector in the bracket of (6) is shown in Figure 3.32. By dropping the $\mathcal{R}$ symbol and the term $\exp(i\omega t)$ we find that we may regard eqn (5) as expressing the current-field dependence in the phasor notation.

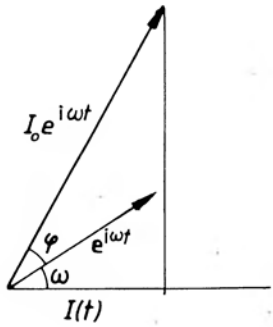


Figure 3.32

In this notation the time-dependence is implicit in the equations but the term $\exp(i\omega t)$ has been dropped. This amounts to fixing the rotating vector diagram in space, eqn (5), and the corresponding diagram gives all the necessary information. The current is referred to the phase of the driving field.

In the frequency domain we are concentrating attention on the square bracket in (5), since the frequency dependent parameters $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ are uniquely related to the real-time dependence of the dielectric system under step-function excitation, $f(t)$.

APPENDIX 3.2

INVERSION IN THE COMPLEX PLANE

Consider the locus of the complex number:

$$y = 1 + (is)^{n\pi/2} = |y| \exp(i\varphi)$$

represented by a straight line A inclined at an angle $\alpha = n\pi/2$ and passing through the point 1 on the real axis, Figure 3.33. Drawing a circle centred on the normal to the line A and passing through

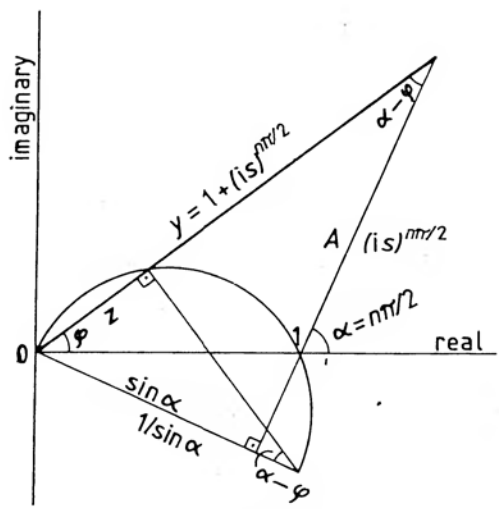


Figure 3.33

the coordinates 0 and 1 on the real axis, the point of intersection of the line y with the arc of the circle defines a vector z whose magnitude is given by the similarity of triangles as:

$$|z| = 1/|y|$$

Since z has the same angle as y, it follows that z is the complex conjugate of the reciprocal of y. In our convention, we are plotting in any case the complex conjugate, since it is more convenient to plot impedances in the upper half plane.

REFERENCES TO CHAPTER 3

Carslaw H S and Jaeger J C 1959, *Conduction of Heat in Solids*, Clarendon Press, Oxford
Cole K S and Cole R H 1941, J. Chem. Phys. **9**, 341
Davidson D W and Cole R H 1951, J. Chem. Phys. **19**, 1484
Dissado L A and Hill R M 1979, Nature Lond. **279**, 685
Erdelyi A 1954, *Tables of Integral Transforms* Vol. II, McGraw-Hill, New York
Fuoss R M and Kirkwood J G 1941, J. Amer. Chem. Soc. **63**, 385
Havriliak S and Negami S 1966, J. Polymer Sci, Pt C, **14**, 99
Hill R M 1978, Nature Lond **275**, 96
—— 1981, J. Mat. Sci. **16**, 118
Hill R M and Dissado L A 1982, J. Phys. C: Solid State Physics **15**, 5171
Hill R M and Jonscher A K 1983, Contemporary Physics **24**, 75
Ishida Y, Matsuo M and Yamafuji K 1962, Kolloid Z. u. Z. f. Polymere **180**, 108
Jonscher A K 1975, Colloid Polymer Sci. **253**, 231
—— 1978, Thin Solid Films **50**, 187
Jonscher A K and Frost M S 1976, Thin Solid Films **37**, 267
Meca F 1981, PhD Thesis (University of London)
von Schweidler E 1907, Ann. d. Physik (Leipzig), Series 4, **24**, 711
Williams G and Watts D C 1970, Trans. Far. Soc. **66**, 80
Williams G, Watts D C and Nottin J P 1972, J. Chem. Soc. Far. Trans. II **68**, 16

CHAPTER 4

The Dynamic Response of Idealised Physical Models

4.1 INTRODUCTION

It will be instructive to present now the mathematical treatment of the frequency response of some simple physical models of dielectric polarisation – partly as an illustration of the method of treatment and as an exposition of the physical significance of specific features of the dielectric functions. We begin with the treatment of the harmonic oscillator and we show the adaptations of the basic model to cover various physical situations, including the case of free charge carriers, and we follow this with a treatment of the floating dipoles – the original Debye model – and of hopping charge carriers such as ions or localised electrons in solids. We also include an analysis of certain semiconductor phenomena which are directly relevant to dielectric processes, namely Schottky barriers, p–n junctions and generation-recombination processes. We conclude with a discussion of the diffusive model which represents a departure from the one-particle models considered hitherto and may be regarded as a simple example of many-particle cooperative systems.

4.2 THE HARMONIC OSCILLATOR

The simplest model of a harmonic oscillator is given by a charge $-e$ with a mass m bound elastically to an equilibrium position with a force constant λ and subjected to damping constant ms. This mechanical model represents rather well the situation in a free atom, where the outer electron shell is readily displaced with respect to the positively charged centre consisting of the atomic nucleus proper with the relatively strongly bound inner shell electrons. Denoting by y the displacement from the central equilibrium position and postulating a harmonic electric field of amplitude E and frequency ω we may write the equation of motion:

$$my'' + msy' + \lambda y = -e\,E\exp(\mathrm{i}\omega t) \tag{4.1}$$

where the primes denote differentiation with respect to time. We may solve this equation for the displacement by standard methods and noting that the polarisation of N identical non-interacting oscillators is given by $P = -eNy$, so that dividing by the amplitude E of the field we obtain the following expressions for the real and imaginary components of the complex dielectric susceptibility

$$\chi'(\omega) = \Omega_p^2 \frac{\Omega^2 - \omega^2}{(\Omega^2 - \omega^2)^2 + s^2\omega^2} = B^2 \frac{1 - x^2}{(1 - x^2)^2 + k^2x^2} \tag{4.2}$$

$$\chi''(\omega) = \Omega_p^2 \frac{s\omega}{(\Omega^2 - \omega^2)^2 + s^2\omega^2} = B^2 \frac{kx}{(1 - x^2)^2 + k^2x^2} \tag{4.3}$$

where

$$x = \omega/\Omega \quad k = s/\Omega \tag{4.4}$$

and

$$\Omega_p = (e^2N/\varepsilon_0 m)^{1/2}$$

is known in a different context as the *plasma frequency*, i.e. it is the natural frequency of oscillation in a free gas plasma, while

$$\Omega = (\lambda/m)^{1/2} \tag{4.5}$$

is the natural frequency of oscillation of the harmonic oscillator in the absence of damping. A brief discussion of the plasma oscillations is given in Appendix 4.2.

Figure 4.1 shows the plots of $\chi'(\omega)$ and of $\chi''(\omega)$ for three values of the damping constant $k = s/\Omega$.

The general shape of these graphs confirms the even and odd character of the real and imaginary components, respectively. The real part shows an initial *rise* with frequency, before reaching a peak and going through zero to *negative values* at frequencies in excess of the natural frequency Ω. The physical significance of this behaviour is that for $\omega > \Omega$ the displacement is *in antiphase* with the driving field, due to the effect of the inertia of the oscillator. The region of strongest dispersion is necessarily accompanied by a peak in the imaginary component – a very general type of behaviour resulting directly from the Kramers–Kronig relations. The initial rise in $\chi'(\omega)$ is referred to as *anomalous dispersion*, since it corresponds to a trend that is exactly opposite to that found in all other systems in which inertia is not the dominant influence, i.e. in which *relaxation* rather than *resonance* is the basis of the behaviour.

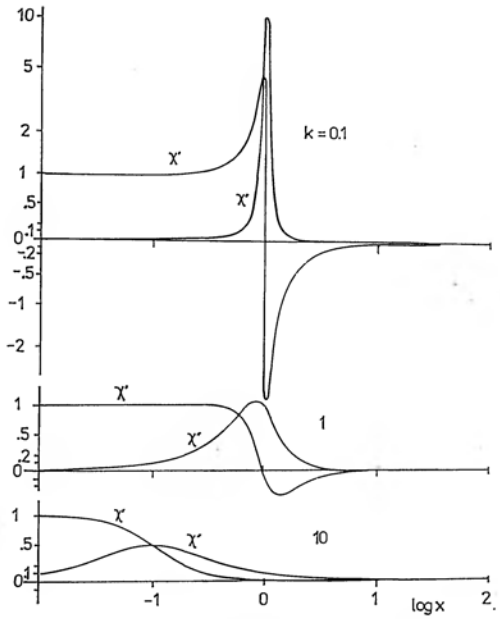


Figure 4.1 The graphical representation of the real and imaginary components of the complex susceptibility given by eqns (4.2) and (4.3), respectively, for three values of the damping coefficient k, plotted against the logarithm of the normalised frequency x. The ordinate scale is quasi-logarithmic, it is expressed as $\sinh^{-1}\chi$ which becomes logarithmic for large values of the argument but is linear for small values, changing sign in the normal way and is therefore preferred to the simple logarithmic representation.

The effect of the damping constant is clearly seen – for $k = 10$ there is no anomalous dispersion and no negative undershoot of the real part at high frequencies. The maximum value of loss is given by B/k. As with all resonance phenomena, the width of the loss peak increases with the increase of damping.

Instead of solving the equation of motion (4.1) with a harmonic driving force, we may apply a delta-function transient, in which case the time–dependent solution is found to be:

$$f(t) = (\omega_p^2/\alpha)\exp(-st/2)\sin(\alpha t) \tag{4.6}$$

where

$$\alpha^2 = (\lambda/m) - s^2/4 \tag{4.7}$$

We note that this is, in fact, the dielectric response function, being the response of the displacement, i.e. polarisation to a delta-function excitation. Figure 4.2 shows the functional form of this response, which is a damped harmonic oscillation, as might be expected from the nature of the problem.

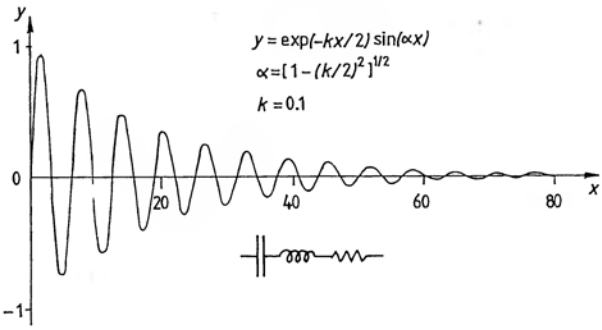


Figure 4.2 The time–domain response of a damped harmonic oscillator under delta-function excitation, given by eqn (4.6), together with the corresponding equivalent circuit.

It may be easily checked that the time– and frequency–domain responses calculated in this manner do, in fact, represent Fourier transforms of one another, while the real and imaginary components are a Hilbert pair.

One characteristic feature of the function $f(t)$ may be mentioned here. The initial rise is linear with t and this means that the behaviour is dominated by inertial effects – only later on do the restoring forces and the damping process come into action.

It is worth pointing out that the response of the damped harmonic oscillator is exactly analogous to that of a series circuit containing capacitance, conductance and resistance, as shown in the inset of Figure 4.2. Denoting the charge on the condenser by q, we may write down the equation of the time dependence of charge:

$$Lq'' + Rq' + Cq = V(t) \tag{4.8}$$

where $V(t)$ is the driving voltage. We note by comparison with eqn (4.1) that the electromechanical analogues are

mass – inductance
damping – resistance or conductance
restoring force – capacitance
mechanical displacement – charge

More detailed discussion of the application of these models to the properties of solids may be found, for example, in Ward (1971).

Excellent examples of these resonances are found in solids. An analysis of the dynamics of a perfect lattice of atoms of ions held

together by elastic forces shows that the motions of individual atoms or ions can be resolved into simple harmonic component oscillations, with damping arising from thermal effects (Elliott and Gibson 1974, Smith 1978, Kittel 1964).

Particularly important from our point of view in the present context are those vibrations in which the neighbouring atoms or ions in a unit cell move in antiphase with respect to one another. These *optic modes* of vibration have a fairly well defined frequency which is of the order 10^{13} Hz and they are of two basic types – those where the motion is along the line joining the equilibrium positions of the particles, called *longitudinal optic* (LO) modes, and those where the oscillations are in the plane normal to the line of centres, called the *transverse optic* (TO) modes. The latter couple particularly strongly to the transverse electromagnetic waves in materials which are wholly or partially ionic, since these motions create oscillating dipoles which are driven by the electric field of the electromagnetic wave. The resulting strong absorption at the frequency of the TO mode can be observed as a very narrow resonance when the perturbing thermal vibrations have been reduced by lowering the

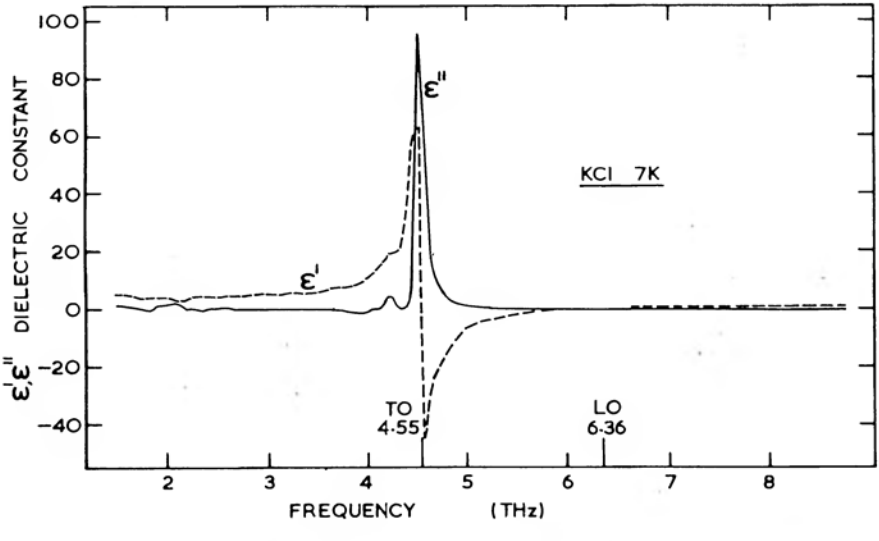


Figure 4.3 The real and imaginary parts of the dielectric permittivity of KCl at 7 K, shown *linearly* against frequency. The measurements were taken using the technique of Dispersive Fourier Transform Spectroscopy in which the sample forms part of a Michelson interferometer and the measurement determines the complex refractive index $\tilde{n} = \varepsilon^{1/2}$. TO and LO denote the transverse and longitudinal optic modes, respectively, as explained in the text (Parker et al 1979).

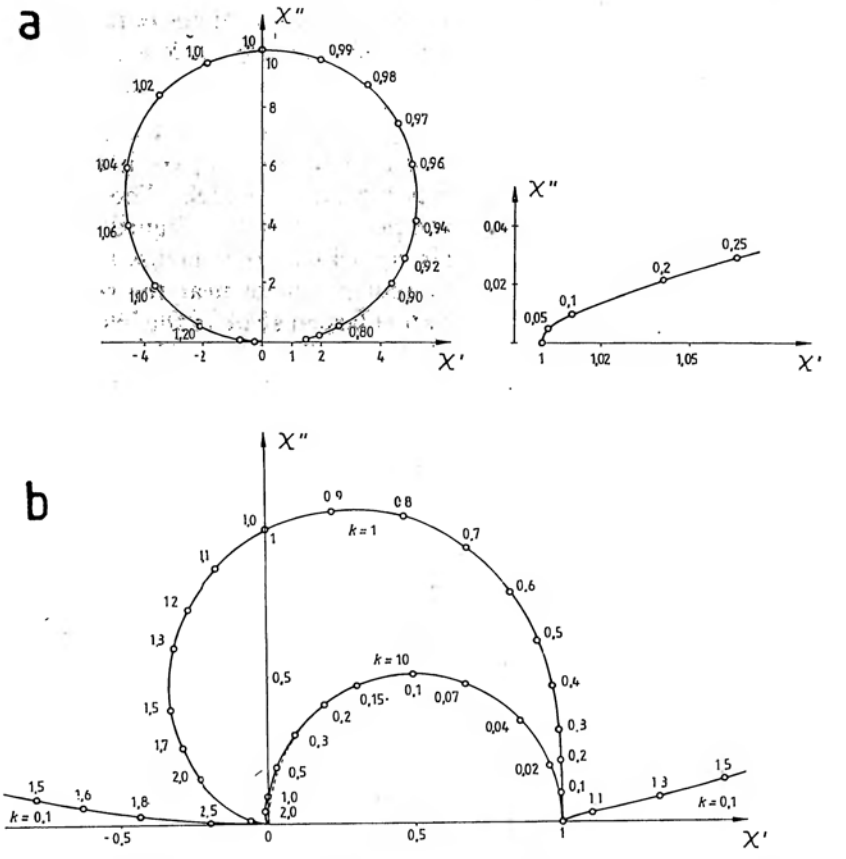


Figure 4.4 The complex plane susceptibility plots for a damped harmonic oscillator, following the equation

$$\chi = 1/(1 - x^2 + \mathrm{i}\, kx)$$

where k is the damping coefficient. Diagram a) corresponds to a weakly damped oscillator with $k = 0.1$, such as the one shown on the top of Figure 4.1. The inset shows in greater detail the part corresponding to small values of the frequency parameter x. Diagram b) gives the extremities of the characteristic corresponding to the same oscillator, together with the complete plots for more strongly damped oscillators, with $k = 1$ and 10, which gradually approach the ideal Debye behaviour corresponding to infinitely strong damping. The latter is shown by the dotted line.

temperature. Figure 4.3 shows the response of the ionic crystal KCl at 7 K in the form of frequency-dependence of $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ in the far infrared region of the spectrum. The sharpness of the resonance may be compared with the theoretical curve of

Figure 4.1, but the different modes of presentation of both vertical and horizontal scales should be noted – Figure 4.1 is logarithmic, Figure 4.3 linear.

It may be instructive to see the complex plane Cole–Cole plot for an inertial system with variable damping, to compare it with the corresponding plots of the various relaxation models. The analysis is given in Appendix 4.1 and the plot is shown in Figure 4.4. For small damping, $k = 0.1$, diagram a), which corresponds to Figure 4.1 and 4.2, the plot is almost circular, except near the real axis, where the response is shown on an enlarged scale in the inset. Plots for increasing values of the damping coefficient are shown in diagram b), and it is clear that the response goes over into the Debye response which corresponds to infinitely strong damping.

4.3 AN INTERTIALESS SYSTEM WITH A RESTORING FORCE

The omission of inertia is always justified in the limit when damping or viscous or resistive forces dominate. This is equivalent to the omission of inductance from our electrical circuit and corresponds physically to the linear analogy of Debye's "floating" dipoles which will be discussed in Section 4.5.

To obtain this limit, we set $m = 0$ while choosing the damping constant in such a way that

$$ms/\lambda = \tau \tag{4.8}$$

which is a time constant and remains finite. We also retain a finite value of the product ms, so that s tends to infinity. We now re-write eqn (4.7) in the form:

$$\alpha^2 = -(s^2/4)\{1 - 4(\lambda/s^2 m)\}$$

where the second term in { } tends to zero, so that we may expand the bracket giving:

$$\alpha = \mathrm{i}\left(\frac{s}{2} - \frac{1}{\tau}\right) \tag{4.9}$$

Using this expression in eqn (4.6), expanding $\sin(\alpha t)$ in terms of exponentials of an imaginary argument and making use of eqn (4.8) it is easily shown that the response function becomes a pure decaying exponential:

$$f(t) = \frac{e^2 N}{\varepsilon_0 \lambda \tau} \exp(-t/\tau) \tag{4.10}$$

An analogous operation on eqn (4.2) and eqn (4.3) gives in the present limit the following expressions for the complex susceptibility:

$$\chi(\omega) = \frac{\chi(0)}{1 + i\omega\tau} \tag{4.11}$$

where the static value of the susceptibility is given by:

$$\chi(0) = e^2 N/\varepsilon_0 \lambda \tag{4.12}$$

while the real and imaginary components are

$$\chi'(\omega) = \chi(0)\frac{1}{1 + \omega^2\tau^2} \tag{4.13}$$

$$\chi''(\omega) = \chi(0)\frac{\omega\tau}{1 + \omega^2\tau^2} \tag{4.14}$$

These response functions correspond exactly to eqn (3.36) and they are shown in Figure 3.33 in the logarithmic representation. We note that the width of the loss peak at half maximum height is equal to 1.144 decades and the loss peak is symmetric in the logarithmic frequency representation.

We note further that the real and imaginary parts of the susceptibility may be represented in parametric form by eliminating $\omega\tau$ between equations (4.13) and (4.14):

$$\{\chi' - \tfrac{1}{2}\chi(0)\}^2 + \chi''^2 = \{\chi(0)/2\}^2 \tag{4.15}$$

which represents a semicircle in the complex χ-plane of radius $\frac{1}{2}\chi(0)$ centred on the real axis and going through the origin, as shown in Figure 3.33. This is a characteristic plot for the exponentially decaying response function. The frequency corresponding to the loss peak, i.e. to the highest point on the semicircle is the *loss peak frequency*:

$$\omega_p = 1/\tau \tag{4.16}$$

We note in the case of a viscous system with negligible inertial effects that the real part of the susceptibility is a monotonically decreasing function of frequency – there is no anomalous dispersion as in the inertial case.

The purely exponential decay of polarisation given by eqn (4.10) shows that the polarisation obeys the first-order differential equation

$$\mathrm{d}P/\mathrm{d}t = -P(t)/\tau \tag{4.17}$$

which implies that the rate of decay of polarisation after the removal of the electric field is proportional to the instantaneously remaining polarisation in the system. It follows that the polarisation current under step-function charging or discharging is given by the relation:

$$i(t) = \varepsilon_0 E_0 \frac{e^2 N}{\varepsilon_0 \lambda \tau} \exp(-t/\tau) \tag{4.18}$$

and the integral of this gives the initial polarisation

$$P(0) = (e^2 N/\lambda) E_0$$

in agreement with the expression (4.12) for steady state susceptibility.

We note that the response function given by eqn (4.10) implies a step rise at $t = 0$ under delta-function excitation, which is the expected response of an inertialess system.

4.4 FREE CHARGE CARRIERS WITH COLLISIONS

As the next special case consider an oscillator with inertia and damping but without any restoring force. The case corresponds to a free charge carrier for which there is no definite "equilibrium" position. We now set $\lambda = 0$, so that from eqn (4.7) $\alpha = is/2$, and the application of eqn (4.6) would give

$$f(t) = (\sigma_0/\varepsilon_0)\{1 - \exp(-st)\}$$

where

$$\sigma_0 = \omega_p^2/s = e^2 N/ms \tag{4.19}$$

However, we note that this form of the response function is inadmissible since it corresponds to a finite value as $t \to \infty$ which is contrary to the basic property of the response function that it should vanish for infinitely long times since there can be no permanent polarisation as a result of the excitation by a delta function. Since, however, a free carrier with a finite mass receiving a delta-function impulse must show a finite displacement because there is no restoring force, we "correct" the response function by subtracting the constant value from the expression in { } thus giving

$$f(t) = -(\sigma_0/\varepsilon_0) \exp(-st) \tag{4.20}$$

The remarkable feature of this result is the negative sign of the response function which is otherwise identical in form to the expo-

nential dependence given by eqn (4.10). It follows therefore that the dielectric susceptibility has the same frequency dependence as (4.11) but the signs of the real and imaginary components are *negative*. Thus free carriers make a negative contribution to the real and imaginary components of the dielectric permittivity. However, although negative loss and negative conductivity imply *power gain*, there is no net gain since the conductivity must include the direct current contribution, so that the total alternating current conductivity is:

$$\sigma(\omega) = \sigma_0 - \frac{\sigma_0}{\omega^2 + s^2} = \frac{\sigma_0}{1 + \omega^2/s^2} \tag{4.21}$$

This gives the well-known result which may be determined more directly by considering the dynamics of free charge carriers and which shows that the ac conductivity is always positive but decreases rapidly above the collision frequency s, as shown in Figure 4.5.

The negative contribution to the real part of the dielectric permittivity is

$$\chi'(\omega) = -\omega_p^2/(\omega^2 + s^2) \tag{4.22}$$

but this value should be understood to subtract from the appropriate positive contribution ε_∞ to the permittivity arising from other polarising mechanisms present in the material, as well as from the permittivity of the free space. It follows that free carrier systems may have a negative permittivity at sufficiently low frequencies, if

$$\varepsilon_0 \omega_p^2/s^2 > \varepsilon_\infty \tag{4.23}$$

The total permittivity is

$$\varepsilon'(\omega) = \varepsilon_\infty - \varepsilon_0 \omega_p^2/(\omega^2 - s^2)$$

and if condition (4.23) is satisfied then this becomes zero at a frequency given by

$$\omega_1 = \{\varepsilon_0 \omega_p^2/\varepsilon_\infty - s^2\}^{1/2} \simeq \omega_p \tag{4.24}$$

the second approximate equality being applicable when the principal polarisation mechanism is free charge polarisation and when $\omega_p^2 \gg s^2$. Both these conditions are normally met in metals and the negative permittivity at low frequencies implies that electromagnetic waves cannot propagate below the plasma frequency, since the refractive index, $n = \varepsilon^{1/2}$, becomes purely imaginary.

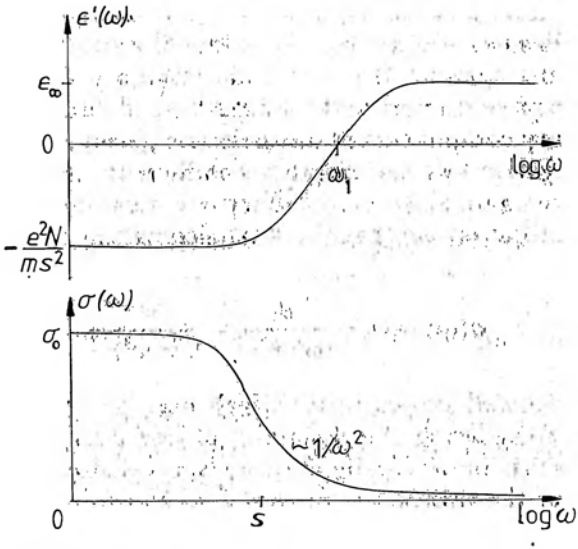


Figure 4.5 The dielectric response of a free-carrier system with a collision frequency s, plasma frequency ω_p and dc conductivity σ_0. ε_∞ is the permittivity due to fast-responding processes, including the free space contribution ε_0. The upper diagram shows the real part of the permittivity, which becomes negative below a frequency ω_1 given by eqn (4.24). Electromagnetic waves cannot propagate below ω_1. The lower diagram shows the corresponding ac conductivity $\sigma(\omega)$.

4.5 DIPOLES FLOATING IN A VISCOUS FLUID

We now return to the model of dipoles floating freely in a viscous medium as proposed originally by Debye in 1912, to account for the observed behaviour of many dipolar liquids and described briefly in Section 2.3 a) in the limit of steady state response. To treat the transient case, we apply a method of analysis developed for this purpose by Debye (1945). The orientations of the individual, non-interacting dipoles in the assembly are given by the positions of "representative" points formed by the intersections of dipole vectors with the surface of a sphere, the dipoles being assumed to occupy the centre of this sphere, Figure 4.6. Under equilibrium conditions the representative points cover the sphere uniformly, giving a uniform distribution over the spherical angle, $g_0(\theta) = N/4\pi$, from eqn 2.9. Under the action of a perturbing electric field the representative points move on the surface of the sphere and have to satisfy a continuity equation which may be written as follows:

$$\mathrm{d}g/\mathrm{d}t = \partial g/\partial t + \nabla \cdot \Phi = -g_1/\tau \qquad (4.25)$$

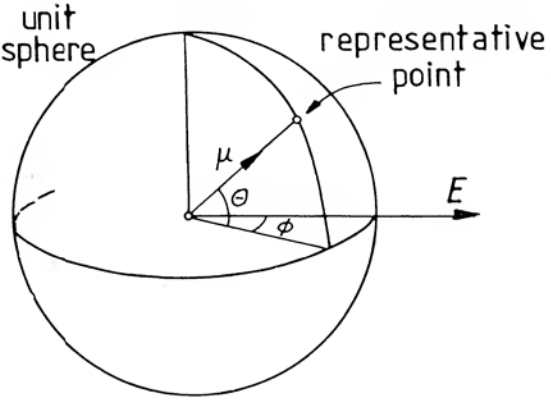


Figure 4.6 The coordinates θ and ϕ of the representative point on a unit sphere defining the orientation of a dipole moment μ situated at the centre of the sphere. E is the direction of an externally applied electric field.

where Φ is the "flux" of the representative points, g_1 is the perturbation of the distribution function from its equilibrium value g_0 and τ is the relaxation time of the perturbed distribution.

The flux Φ has only one component in the spherical polar coordinates defined by (r, θ, ϕ):

$$\Phi = \{0, g(\theta)v_\theta, 0\}$$

where v_θ is the velocity determined by the torque Ξ and the viscosity coefficient η:

$$v_\theta = \mathrm{d}\theta/\mathrm{d}t = \eta\Xi(\theta) = -\eta\mu E \sin\theta$$

where we have used eqn (2.8).

The divergence in spherical polar coordinates appropriate to our present situation is given by:

$$\begin{aligned}\nabla \cdot \Phi &= \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\{g(\theta)\sin\theta\,\mathrm{d}\theta/\mathrm{d}t\} \\ &= -\frac{\eta\mu E}{\sin\theta}\frac{\partial}{\partial\theta}\{g(\theta)\sin^2\theta\}\end{aligned} \tag{4.26}$$

Taking the limit of small perturbations, $g_1 \ll g_0 = N/4\pi$, we obtain from (4.25) the following equation for the small perturbation:

$$\frac{\partial g_1}{\partial t} - \frac{N}{2\pi}\eta\mu E\cos\theta = -\frac{g_1}{\tau}$$

which has the following small-signal solution in an alternating electric field:

$$g_1(\theta) = \frac{N}{4\pi}\frac{\mu E}{kT}\frac{1}{1+i\omega\tau}\cos\theta \tag{4.27}$$

where we have made the following substitution in order to make the result accord with the steady state solution (2.10):

$$2\eta\tau = 1/kT \tag{4.28}$$

This condition relates immediately the relaxation time τ to the viscosity η of the fluid in which the dipoles are "floating" and since η is normally thermally activated with an energy W, we may also write:

$$1/\tau = \omega_p = \nu\exp(-W/kT) \tag{4.29}$$

where ν is a suitable frequency and we have equated $1/\tau$ with the loss peak frequency.

We are now able to carry out the averaging over all dipolar orientations exactly as in eqn (2.11) and obtain by analogy with eqn (2.15) in the limit of low fields the following expression for the complex susceptibility:

$$\chi(\omega) = \frac{N\mu^2}{3\varepsilon_0 kT}\frac{1}{1+i\omega\tau} \tag{4.30}$$

The steady state value of this in the limit of zero frequency is the same as in (2.17), while the frequency dependence is the same as for the inertialess harmonic oscillator.

Equation (4.30) represents the classical Debye equation for the frequency response of freely-floating non-interacting dipoles which are being aligned by the field and also suffer randomising collisions by the medium in which they float.

Since the maximum amplitude of the frequency factor in eqn (4.30) is equal to $\frac{1}{2}$, the loss peak amplitude is given by

$$\chi''(\omega_p) = \tfrac{1}{2}\chi(0) = (\mu^2 N/3\varepsilon_0 k)(1/T) \tag{4.31}$$

Taking now into account eqn (4.29) which applies to thermally activated loss peak frequency with the activation energy W, we may derive the relation between the loss peak amplitude and the loss peak frequency:

$$\chi''(\omega_p) = \tfrac{1}{2}\chi(0) = -\text{const}\,.\log[\omega_p(T)] \tag{4.32}$$

Hence in a Debye system a plot of the linear loss peak amplitude against the logarithm of the loss peak frequency should yield a straight line, as shown schematically in Figure 4.7 a).

However, even if the loss peak frequency were not thermally activated with a unique energy W, as may be the case with some so-called α peaks in polymers and glasses, cf. Section 5.3 and Figure 5.19, eqn (4.31) is still valid within the approximations used and we would expect a linear plot of the logarithm of the loss peak against the logarithm of the absolute temperature, Figure 4.7 b).

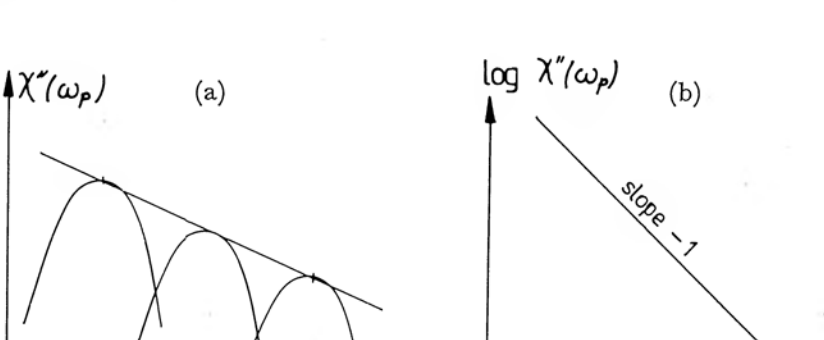


Figure 4.7 The relationship between the loss peak frequency and the loss peak amplitude expected for a Debye system with a simply activated loss peak frequency, diagram a). Even if the loss peak frequency is not simply activated, e.g. in an α loss process, the temperature dependence of the loss peak amplitude should be as shown in diagram b).

4.6 CHARGE HOPPING BETWEEN TWO POTENTIAL WELLS

The dynamic response of a charge hopping between two potential wells can be obtained by reference to the static analysis in Section 2.3 e) and to Figure 2.6. While the static case could be solved exclusively with reference to the Boltzmann equilibrium between two energy levels and was not in any way affected by the actual *transition probabilities* R_{ij} R_{ji} between these wells, the dynamic case has to take these into account.

Under equilibrium conditions the actual transition rate is given by

$$\nu = f_i^0 R_{ij}^0 = f_j^0 R_{ji}^0 \tag{4.33}$$

With the application of a field the barrier for jumping out of well i changes by $-qrE/2$ and that for jumping out of well j by $+qrE/2$ so that we may write the approximate expressions, expanding the exponentials as in (2.23):

$$R_{ij} = R_{ij}^0(1 + a/2); \quad R_{ji} = R_{ji}^0(1 - a/2)$$

while the occupation probabilities change by $\pm f'$ as in (2.22). The net rate of change of occupancy is therefore given by:

$$-df'/dt = R_{ji}^0(f_j^0 + f')(1 - a/2) - R_{ij}^0(f_i^0 - f')(1 + a/2)$$

Bearing in mind eqn (4.33), defining

$$R_{ij}^0 + R_{ji}^0 = \frac{1}{f_i f_j} = \frac{1}{\tau} \tag{4.34}$$

and neglecting higher order terms in small quantities we obtain the following equation for the rate of change of the occupancy:

$$\frac{df'}{dt} = -\frac{f'}{\tau} + \mathcal{N} a \tag{4.35}$$

which is solved in the small-signal approximation at a frequency ω:

$$(i\omega + 1/\tau) f' = \frac{q\mathcal{N}}{kT} E$$

from which we obtain the susceptibility for N double wells

$$\chi(\omega) = \frac{N\mu^2}{3kT\varepsilon_0} \mathcal{N}\tau \frac{1}{1 + i\omega\tau} \tag{4.36}$$

This gives the same static result as eqn (2.24) and, once again, shows the same Debye frequency dependence.

In view of what had been said about the equivalence of charges hopping between two wells and dipoles jumping between preferred orientations, one expects an identical frequency dependence from a set of jumping dipoles if their re-orientations occur between two preferred directions.

4.7 DIELECTRIC PHENOMENA IN SEMICONDUCTORS

i) Semiconductor materials

Electronic and ionic semiconductors do not normally show significant dielectric behaviour below the THz frequency range which we have set as a limit for our considerations. The reason for this lies in the effect of their dc conductivity σ_0 on the dielectric loss, the σ_0/ω term dominating at low frequencies, thus making dielectric measurements difficult if not impossible. For example, a "transistor quality" silicon sample with $\sigma_0 = 10^2\,\Omega^{-1}\,m^{-1}$ would give a loss σ_0/ω equal to the real part of the permittivity at 200 GHz, while even "high-resistivity" semiconductors with $\sigma_0 = 10^{-4}\,\Omega^{-1}\,m^{-1}$ would still be difficult to measure below a MHz.

Dielectric phenomena appropriate to free carriers – the fall of conductivity according to eqn (4.21) or, in suitable conditions the change of sign of the dielectric permittivity according to eqn (4.24) leading to total reflection of incident radiation – are well known but they fall in the infra-red region of the spectrum, in view of the high values of the collision frequency s.

While this is true of the bulk behaviour of semiconductors, important dielectric phenomena arise at interfaces between semiconductors and metals, where predominantly insulating Schottky barriers may appear, or even more importantly at p–n junctions. In these regions the equilibrium carrier densities which are present in the bulk are removed due to favourable potential energy configurations and the remaining semiconductor lattice has essentially dielectric properties. These effects have been known since the beginning of the science and technology of metal-semiconductor rectifiers in the 1930's and they form an important branch of present-day semiconductor science.

To a very good approximation these semiconducting barrier regions may be considered as *ideal*, i.e. loss-free capacitors, except for the dc contribution of any leakage currents that may be present. This is how they are being treated in most textbooks and scientific papers dealing with their dielectric properties, for example as blocking capacitors in semiconductor contacts, as voltage-dependent capacitances in p–n junctions and as spurious blocking barriers in ionic conductors. On closer examination, however, it becomes clear that these barrier regions may exhibit important dielectric properties in their own right, showing a dependence of loss and polarisation on frequency which departs considerably from that expected on the

basis of simple considerations. Several examples of this will be shown in Chapter 5 and is for this reason that we are giving in the present Section a brief treatment of the fundamentals of semiconductors designed to provide a background to their dielectric properties. There can be no question of giving a comprehensive treatment of this wide subject within the scope of the present monograph, so that the interested reader who is not familiar with semiconductor physics should turn to specialist texts, for example Henisch (1957), Smith (1978), Grove (1967), Sze (1969), to mention but a few. Our discussion will be concerned predominantly with electronic semiconductors – the case of ionic conductors may be treated analogously, while more detailed considerations will be found in specialist literature, especially on the so-called fast ion conductors (Vashishta et al 1979) which show dielectric behaviour discussed in Chapter 5. It should be noted that all ionic conductors have *hopping localised* charge carriers, not free carriers as in electronic semiconductors.

A barrier region depleted of electronic charge carriers responds primarily as a dielectric material with a purely real permittivity ε_∞ which is determined by the lattice itself. However, there may be present in this lattice other, more slowly responding processes, which therefore contribute to the dielectric loss in the system. Of these precesses, the removal and replenishment of the equilibrium charge densities take times of the order of the *space charge relaxation time* $\tau_s = \varepsilon_0/\sigma$, cf. Appendix 4.2, and this is typically in the MHz to GHz region, so once again it may be considered to be rather fast. The two really slow processes are charge *generation and recombination* and the complementary process of charge *trapping and detrapping*. These processes will be considered in some detail because of their importance in p–n junctions. In the case of ionic lattices, there may also be present slowly hopping charge carriers of the type considered in Section 4.6.

In order to provide a minimal amount of background information for the non-specialist reader, we begin with the concept of electronic semiconductors. These are non-metallic solids in which the chemical valence bonds are fully saturated, in the sense that no more electrons can be accommodated in the bonding process. Typical bonds are either covalent or ionic, or a mixture of the two. An important conclusion of the band theory of solids is that energetically the valence electrons occupy all available levels within a band of energies which is known as the *valence band*, and is shown as the shaded region in Figure 4.8. A full valence band does not contribute to the

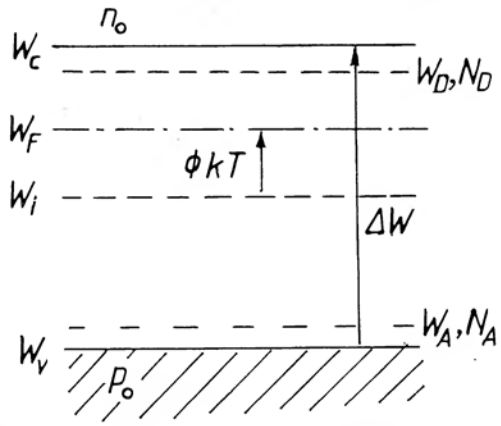


Figure 4.8 The energy band diagram of a homogeneous semiconductor with an energy gap ΔW, donor density N_D at a level W_D, acceptor density N_A at W_A, intrinsic level W_i and a Fermi level W_F at an energy difference ϕkT above W_i. W_c and W_v denote the edges of the conduction and valence bands, respectively. The free carrier densities of electrons and holes are n_0 and p_0, respectively. The position of the Fermi level indicates an n-type semiconductor.

electrical conductivity of the solid, since the electrons occupy all available levels and an external electric field cannot change their energies and cannot therefore impart any net momentum to them. This is the reason why materials with completely filled valence bands and no other electrons in the system are insulators.

Above the top of the valence band there is a region in which no free electron can exist – this is known as the *forbidden gap* – of width $\Delta W = W_c - W_v$, i.e. the difference between the top of the valence band and the bottom of the next-higher allowed *conduction band*, also shown in Figure 4.8. The width of the forbidden gap is related to the strength of the chemical bonds in the solid – it is larger in diamond, 5.5 eV, than in silicon, 1.1 eV and in germanium, 0.7 eV. In compound materials it is large in ionic solids with high difference of electronegativities, such as alkali halides, and it decreases with increasing atomic number, for example NaCl has a larger gap than KBr.

It is a fundamental property of semiconductors and insulators that their conduction band is normally completely empty at the absolute zero of temperature. This means that there are no electrons available to move in the conduction band and the material is therefore an insulator, in complete contrast with metals in which conduction is possible because no excitation is required to bring carriers into the

conducting levels there. In a perfect semiconductor free of lattice defects and of *extrinsic* impurities, the only way in which an electron may be removed from the valence band and placed in the conduction band is by excitation across the forbidden gap. This may be achieved by thermal excitation or it may take place through the absorption of a light quantum of sufficient energy. In either case, there will now be conduction by electrons in the conduction band, and by the missing electrons in the valence band which behave as *positive* charges and are referred to as *holes*. We will not be concerned further with the photoconductive process which results when the excitation is by photons and will discuss only the thermal excitation which produces equal densities of *intrinsic carriers*

$$n_i = (N_c N_v)^{1/2} \exp(-\Delta W/2kT) \tag{4.37}$$

where N_c and N_v are known as the effective densities of states in the conduction and valence bands, respectively, and are of the order $10^{24} - 10^{25}\,\mathrm{m}^{-3}$. At room temperature the exponential factor amounts to 10^{-9} with $\Delta W = 1$ eV, e.g. in silicon, and 10^{-26} with 3 eV, so that materials with forbidden gaps above that range contain negligible amounts of intrinsic carriers and are therefore insulators, at least in the absence of impurities.

However, the intrinsic densities are only of importance in materials containing correspondingly low densities of extrinsic impurities and defects and most semiconductors contain such impurities in much larger quantities. Such impurities may perform several different functions. Some impurities contain one electron more in their valence shell than the regular lattice atom for which they substitute – for instance P in Si, or Te substituting for As in GaAs. This extra electron does not take part in bonding of the atom in question and it is easily split off and becomes free to move in the conduction band. Once there, it behaves as a virtually free electron, suffering occasional collisions with the lattice but being able to follow externally applied electric fields. The atom from which the electron became detached remains as a positively charged centre in the lattice. The energy required to ionise the non-isoelectronic centre, known as the *donor*, is much smaller than the energy ΔW required to break the bond in the regular lattice and to produce the intrinsic excitation. These shallow donors are therefore shown in the proximity of the conduction band in Figure 4.8. Typical energies may be $0.01 - 0.05$ eV and they are comparable with thermal energies at normal temperatures. This means that in the presence of a density N_D of such donors, the electron density in the conduction

band increases to an equilibrium value n_0 which may reach N_D at sufficiently high temperatures and which certainly exceeds n_i by many orders of magnitude.

A different type of impurity may contain one electron less than is required for normal bonding, e.g. B in Si, and this leaves the valence band short of one electron. This missing electron may not be localised at the impurity atom which gave rise to it, but may become excited into the valence band, by allowing one of the valence band electrons to drop into the missing bond attached to the impurity, thereby rendering the impurity atom negative. The resulting hole is then free to contribute to conduction in the valence band in an exactly analogous way to an electron in the conduction band. The impurity giving rise to the appearance of a hole is called an *acceptor*, since it accepts an electron from the valence band, thus creating a hole. A density N_A of acceptors increases the hole density p_0 beyond the intrinsic value n_i to the limit of N_A at sufficiently high temperatures.

The equilibrium densities n_0 of electrons and p_0 of holes are governed by *Fermi–Dirac statistics* which defines the occupation probabilities of all levels, those corresponding to conducting or propagating states as well as those of localised states such as donors and acceptors. The occupation probability of a level at energy W is defined by the Fermi–Dirac distribution function

$$f(W) = \frac{1}{1 + \exp[(W - W_F)/kT)]} \tag{4.38}$$

where W_F is defined as the *Fermi level* and corresponds to that energy at which the probability of occupation by an electron is $\frac{1}{2}$. The occupation probability falls off exponentially above the Fermi level, since the probability of an electron having sufficient energy to be there is decreasing rapidly. Likewise, the probability of finding a hole, i.e. an unoccupied level, $1 - f$, falls off exponentially as the energy W decreases below W_F, indicating that electronic states are almost fully occupied there.

In an intrinsic semiconductor the Fermi level falls almost mid-way in the forbidden gap, to provide equal occupation probabilities for holes in the valence band and for electrons in the conduction band – there is a slight departure from mid-gap position because of an asymmetry of the two bands. Denoting the energy level taken by the Fermi level in an intrinsic semiconductor by W_i, we may define the position of the Fermi level under all conditions in terms of a

dimensionless energy parameter $\phi = (W_F - W_i)/kT$, as indicated in Figure 4.8, and the equilibrium densities of electrons and holes are then obtained in the form:

$$n_0 = n_i \exp \phi \qquad p_0 = n_i \exp - \phi \tag{4.39}$$

which leads to the relation:

$$n_0 p_0 = n_i^2 = (N_c N_v) \exp(-\Delta W/kT) \tag{4.40}$$

Since the density of one carrier species, either electrons or holes, depending upon the prevailing density of donors or acceptors, is liable to be much higher than the intrinsic value n_i, it follows that the other density is correspondingly smaller, although it always remains finite. Thus we refer to the *majority* and *minority* densities of carriers in the free bands, and we denote semiconductor materials as *n-type* if the majority carriers are electrons arising from the preponderance of donor impurities, and as *p-type* if the majority carriers are holes arising from dominant acceptor impurities. It is evident that the Fermi level is situated below the intrinsic level in p-type materials and above the intrinsic level in n-type materials.

Electrons in the valence band and holes in the conduction band are referred to as "free" carriers in the sense that they are able to propagate in these bands with relatively little hindrance from "collisions" with impurities and with vibrational states of the lattice, known as phonons. These collisions determine the effective *mobilities* μ_e and μ_h of electrons and holes, defined by the expression for the net drift velocity v for electrons and holes in a field E:

$$v_e = -\mu_e E \qquad v_h = \mu_h E$$

Mobilities are considered as positive numbers and are measured in units m^2/Vs. Typical values in crystalline semiconductors are 0.1 and 0.03 for electrons and holes in silicon, respectively, 0.76 for electrons in gallium arsenide and 0.36 and 0.18 for electrons and holes in germanium, respectively. The resulting electron and hole conductivities are additive, since electrons move in the opposite sense than holes, as is implied by the velocity relations, but they carry the opposite sign, so that the conductivity does not depend upon the sign of the relevant carriers. We may write therefore the expression for the electrical conductivity in static electric fields:

$$\sigma_0 = e(\mu_e n + \mu_h p) \tag{4.41}$$

To this dc conductivity may be added any ac contribution arising from dielectric loss processes which may be present in the material,

for instance dipolar or hopping charge contributions. The value given by eqn (4.41) corresponds directly to that given by (4.19) if we note that the mobility is given by the relation $\mu = e/ms$ for each carrier species separately. The frequency dependence of this conductivity is given by eqn (4.21).

Many non-metallic solids, whether they be characterised as semiconductors or semi-insulators according to the value of the conductivity, do not conduct by free carriers but by discontinuous hopping movements of charge carriers between preferred localised sites. One form of such transport was considered in Section 4.6 where the charge was assumed to be confined to two sites only, so that it could not contribute to dc conductivity and behaved exactly like a dipole. However, there is no reason to suppose that most charges are so confined and they may execute consecutive hops between sites arranged in a three-dimensional array, with either equal or different probabilities of individual hops. Such hopping charges can contribute to dc conductivity as well as giving a finite ac effect. Their mobility in static fields is much lower than the free band mobilities, generally less or much less than $10^{-5}\,m^2/Vs$, and with values as low as 10^{-10} being known. A general treatment of this widely developed subject may be found in Mott and Davis (1979). Here we wish to stress the fact that hopping motion is *the only possible* form of transport for all forms of ionic conductors, since ions cannot propagate in "free bands" – their masses are much too high for this to happen. The band structure as such is therefore not applicable to ionic conduction, but the concept of conductivity is and ionic conductivity is indistinguishable from electronic in simple experimental situations, except for their contrasting response under alternating field.

ii) Schottky barriers and p–n junctions

Having given a very brief outline of the bulk properties of semiconducting materials, we are now in a position to describe some of the interfacial phenomena which are of considerable interest in the context of dielectric properties. An interface between a metal and a semiconductor represents a discontinuity in the structural sense – one crystal lattice being replaced by another with completely different electronic properties – but it also represents a discontinuity in the potential distribution compared with the bulk semiconductor. These potential discontinuities have a fundamental influence on the electrical behaviour of the interface.

There are several physical causes why the Fermi level at the interface between a semiconductor and a metal, or a semiconductor and a

dielectric, should not lie at the same position with respect to the conduction and valence band edges as in the bulk, where its position is dictated by the impurity content of the material. Among these reasons may be quoted the difference of work function or contact potential difference, the presence of interfacial states which play the role analogous to bulk states in the material, or accumulation of ionic charges to which the interface represents an impenetrable barrier. Detailed discussion of these phenomena may be found in the references quoted earlier, in particular in Henisch (1957), Grove (1967) and Sze (1969). For our present purpose it is sufficient to describe the result of two basically different types of potential mismatch which can be found in semiconductor – metal contacts, and they apply with suitable modifications also to dielectric-metal interfaces as well.

The interface may be such that the Fermi level tends to come closer to the band edge than in the bulk material, as shown in Figure 4.9 for both n-type and p-type material. In this case there is an *increase* of the respective majority density near the interface, favoured by the potential profile in which the negative charges seek positions of lowest energy, positive charges seek positions of highest energy. It should be intuitively evident that this type of contact would assure good continuity of charge flow from the metal to the semiconductor and vice versa – it would act as an *ohmic contact*. The

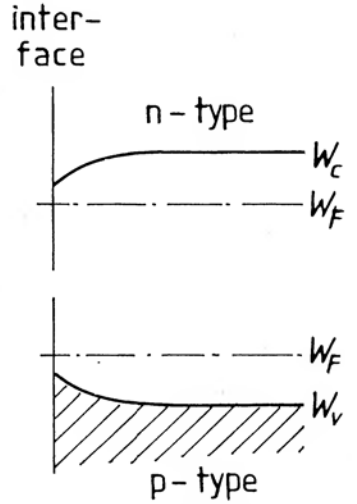


Figure 4.9 The band structure of an ohmic contact in n-type (top) and p-type (bottom) semiconductors. The band configurations refer to the proximity of an interface and they correspond to *separate* situations.

conductance of a system comprising a block of semiconductor and two such contacts would be determined mainly by the bulk semiconductor, since the two contacts would not constitute any significant additional resistance to the flow of charges.

An opposite type of mismatch may arise and produce the situation shown for an n-type semiconductor in Figure 4.10. The material

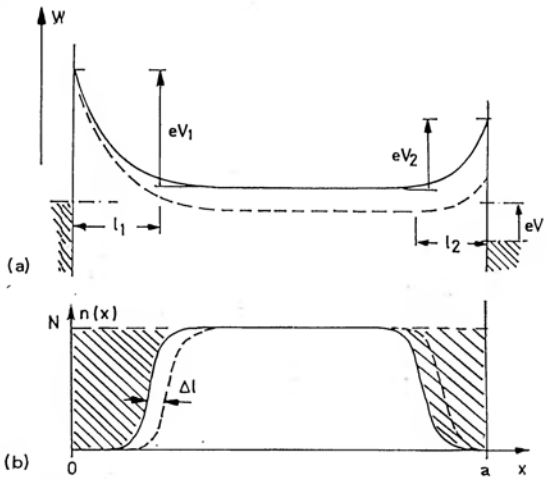


Figure 4.10 The potential energy diagram (a) and the distribution of the free carrier density (b) in a metal-semiconductor-metal sandwich with blocking contacts due to Schottky barriers of heights V_1 and V_2 and widths l_1 and l_2, respectively. The application of an external bias V changes the potential and density distributions as shown by the dotted lines. The density $n(x)$ is shifted laterally by distance Δl. Shading in b) denotes space charge.

From Jonscher (1976)

has a uniform density N of shallow donors extending up to the interface and this determines the bulk conductivity, but the rising potential in the interfacial barrier region causes an exponential fall-off of the density of free electrons, following eqn (4.39) which is still applicable to this situation. This leaves behind the exposed positive charge of the donor levels, which are no longer neutralised by the electrons emitted from them, as in the bulk material. The barrier heights V_1 and V_2 at either contact are determined under equilibrium conditions by the interfacial factors mentioned above and they need not be the same, if the structure is not symmetric. The thicknesses l_1 and l_2 of the two barrier regions are determined from the solution of Poisson's equation for the space charge regions,

which under certain simplifying assumptions become:

$$l_{1,2} = (2\varepsilon V_{1,2}/eN)^{1/2} \tag{4.42}$$

Taking a dielectric constant $\varepsilon_r = 10$, corresponding to silicon, and assuming a barrier height of 1 eV, which would be typical of many interfacial situations, we obtain the barrier thicknesses corresponding to typical impurity densities as shown in Table 4.1. The lowest impurity density quoted corresponds to highest purity materials

TABLE 4.1

Impurity density	10^{19}	10^{21}	10^{23}	10^{25}	$10^{27}\ m^{-3}$
Barrier width at 1 eV	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9} m

normally available in semiconductors, and hence also in typical dielectrics. One is therefore unlikely ever to find barrier thicknesses in excess of 10 μm and values ten times smaller are more typical of high-purity materials. At the other extreme, $10^{27}\ m^{-3}$ approaches 10% of the total atomic density in a solid and it may be considered typical of the conditions prevailing in ionic solids near a blocking electrode, with the resulting barrier width of 1 nm which amounts to no more than three atomic spacings. Such a barrier would have a capacitance of 0.1 F/m^2.

Depletion barriers are known as *Schottky barriers* after the German physicist who first explained the operation of metal-semiconductor rectifiers in the 1930's.

If an external potential difference V is applied to the system shown in Figure 4.10, there is negligible current flow so long as $V < V_{1,2}$ in view of the blocking nature of both electrodes to the transport of electrons. There is therefore a negligible potential drop in the bulk material and the entire potential V is accommodated by making one of the barriers slightly higher, the other slightly lower. This results in a shift of the entire charge density distribution by an amount Δl, so that the centre of gravity of this distribution moves by the same amount. Application of eqn (2.6) gives the change of the induced surface charge at the metal electrodes *per unit area* as:

$$\Delta Q = \pm eN\Delta l\,(a - l_1 - l_2)/a \simeq \pm eN\Delta l \tag{4.43}$$

where the approximate equality applies when the thickness a of the semiconductor slab is much larger than the combined widths of the two barrier regions, which is usually the case. In this approximation the result does not depend upon the thickness of the neutral semiconductor and is determined solely by the change of the barrier

widths, which are related directly to the changes of the barrier heights $\Delta V_1 = -\Delta V_2$ and are given by:

$$V = \Delta V_1 + \Delta V_2 = eN(l_1 + l_2)\Delta l/\varepsilon \tag{4.44}$$

There is no significant potential drop in the bulk material in view of the absence of current flow. Since the effective capacitance is given by the ratio $\Delta Q/V$, we obtain the following expression for the capacitance of the system per unit area:

$$C = \frac{\varepsilon}{l_1 + l_2} \tag{4.45}$$

This capacitance is, of course, exactly that of the two barriers in series and may be regarded as the capacitance of a parallel-plate condenser with separation between metallic plates equal to $l_1 + l_2$ and with the permittivity equal to that of the semiconductor.

Our derivation of this capacitance presupposed a negligible voltage drop in the interior of the sample, which would be true in the case of steady state voltage and with negligible carrier transport through the barriers. A different situation arises with sinusoidally varying voltage signal, when the displacement current required to charge the barrier capacitances becomes significant and requires a driving field in the bulk of the material. We now have the ordinary series combination of the capacitance given by eqn (4.45) with the resistance of the bulk of the semiconductor. The complex capacitance of this series $R - C$ combination has a frequency dependence identical to that of an ideal Debye system.

It should be noted that the barrier capacitance described above corresponds to a *perfect* dielectric region of thickness $l_1 + l_2$ and permittivity ε placed in series with a conducting material. The only departures from this ideal model which are implied in the system considered here are the effects of the dc conductance G_0 through the barriers, which would contribute a term of the form G_0/ω as the loss component. There are so far no other loss processes in the system, since the dielectric behaviour of the barrier itself would be expected to be dominated by the lattice response of the semiconductor which only becomes significant at very high frequencies.

The concept of a Schottky barrier is easily extended to a semiconductor structure in which the complete transition from n-type to p-type material occurs within a single-crystal block of semiconductor, which has been doped with predominantly donor type impurities on one side of a plane and with predominantly acceptor

type impurities on the other side. The resulting situation is shown in Figure 4.11 in which the energy bands make a gradual transition from their equilibrium position in the p-type material to their equilibrium position in the n-type material. The result is exactly equivalent to the juxtaposition of two Schottky barriers of opposite curvatures and the junction plane in which the impurity profile is changing suddenly from donor to acceptor doping is the point at which the two barriers are fitted to give continuous potential and field contours.

The distributions of charge carriers in the neighbourhood of the *p–n junction* formed in this way is also shown in Figure 4.11. A region depleted of both carrier species is created, its total width l

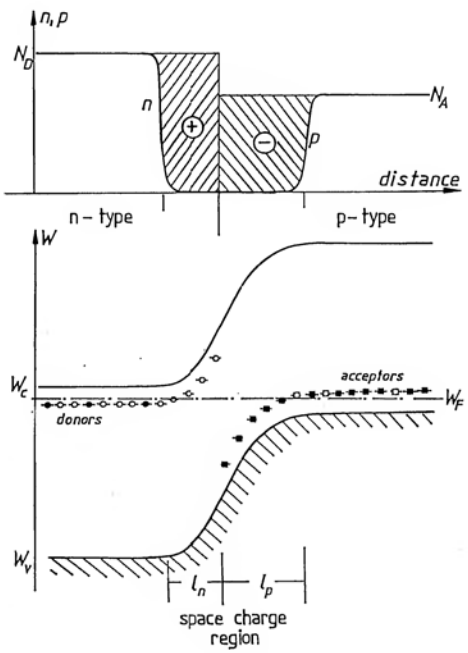


Figure 4.11 Equilibrium relations in a p–n junction. The upper diagram shows the spatial distribution of donor and acceptor densities changing abruptly in the junction plane. It also shows the distributions of the majority electrons, $n(x)$ on the n-side and of the majority holes, $p(x)$ on the p-side. The shaded areas denote the net space charge arising from the un-neutralised donors and acceptors where the free charge carriers have been swept away. The lower diagram shows the corresponding energy band configuration with the widths l_n and l_p of the space charge region on the two sides. Circles denote donors, squares acceptors, black symbols denote occupation by electrons. The Fermi level W_F is constant throughout the system in equilibrium.

being the sum of the two widths l_n and l_p on both sides of the junction. The same considerations apply as in the case of the Schottky barrier regarding the resulting capacitance, with the important difference that a p–n junction normally shows very much lower leakage currents than a metal-semiconductor contact, primarily because of the absence of a discrete interface. This makes true dielectric loss measurements on p–n junctions much easier than on Schottky barriers.

It should be noted that both p–n junctions and Schottky barriers represent capacitive elements whose widths, and therefore capacitances, are dependent on the magnitude of any applied steady bias – both are also *asymmetric* structures, which are sensitive to the polarity of the applied bias. One such polarity – that making the p-type region positive and/or the n-type region negative causes a reduction of the barrier height and a rapidly increasing current with bias – this is the *forward* bias, while the opposite polarity results in the flow of very small "leakage" current and is called the *reverse* polarity.

The diffusion potential V_D of p–n junctions and Schottky barriers tends to decrease with rising temperature due to the shift of the Fermi level towards the middle of the forbidden gap. Typical values in silicon p–n junctions are 0.5 V at 300 K and almost 1.0 V at 77 K – the latter being nearly the value of the energy gap.

iii) Charge generation/recombination processes

Our discussion of the barrier and junction phenomena was so far concerned exclusively with the presence of shallow donor and acceptor levels supplying free carriers in the bulk regions and constituting the fixed space charge in areas from which free charge carriers have been swept away by potential barriers. The responses of these shallow levels to externally applied potentials may be considered to be rapid on the time scale of signals of interest to us in the present context – they are certainly shorter than 1 ns. Likewise, the speed of adjustment of the free charges themselves is dictated by the space charge relaxation time discussed in Appendix 4.2.

We now have to consider the effect of *deep localised levels* within the forbidden gap. These are always present in all materials since they arise from lattice defects which do not act as shallow donors or acceptors and they may also be due to certain impurities, such as transition metals in silicon or germanium. These deep levels are shown schematically in Figure 4.12 at a single energy W_t – the

subscript "t" stands for *trap* or *trapping level* – and they have a density N_t which may be appreciable in comparison with the donor and acceptor densities. An electron falling into one of these deep levels from the conduction band may become *trapped* or immobilised in it for a significant length of time, and likewise an electron falling from one of these levels into a hole in the valence band may effectively trap that hole until it becomes released by re-excitation of an electron from the valence band into the trap state.

Figure 4.12 shows four types of transitions between the valence and conduction bands and the deep levels. R_{ct} corresponds to electron transitions into the traps, R_{tc} is the opposite process of excitation from the traps into the conduction band. Likewise, R_{tv} corresponds to electron transitions from the traps into the valence band holes,

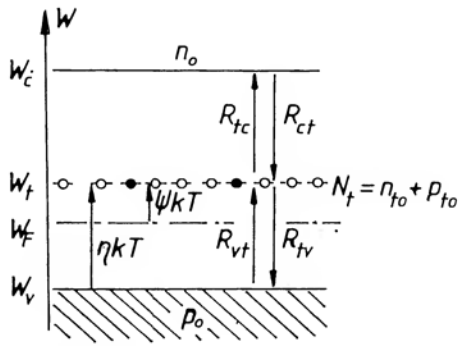


Figure 4.12 The band diagram of a semiconductor with a single level of deep traps of density N_t which communicate with the conduction and valence bands through the rates R_{ct} etc. shown by arrows. The occupancy of these levels is determined by the positions of the Fermi level in the system. The density of empty states is p_{to}, that of full states n_{to}.

while R_{vt} is the opposite process of excitation of electrons from the valence band into the localised traps. We note that R_{vt} followed by R_{tc} gives rise to the *generation* of an electron–hole pair, while R_{ct} followed by R_{tv} causes the *recombination* of an electron and a hole, i.e. their annihilation. Although generation and recombination may occur directly between the conduction and valence bands, the presence of deep levels enhances this process very considerably – by many orders of magnitude, because it makes it easier for the charges to obtain the smaller amounts of energy that are required for an indirect transition via a localised level.

The equilibrium densities of electrons in the traps, n_{to} are given by the Fermi–Dirac statistics, eqn (4.38)

$$n_{to} = N_t/(1 + e^{\psi}), \qquad p_{to} = N_t - n_{to} = N_t/(1 + e^{-\psi}) \tag{4.46}$$

Define the density

$$p_1 = N_v e^{-\eta} \tag{4.47}$$

corresponding to the density of holes that would be present in the valence band if the Fermi level coincided with the trapping level. The equilibrium hole density then becomes:

$$p_0 = N_v e^{-\eta+\psi} = p_1 e^{\psi} \tag{4.48}$$

The two equilibrium transition rates to and from the valence band may now be written in terms of certain rate constants r_v and r_v' and of the products of the numbers of carriers which are available for making the transitions and of numbers of empty spaces into which these transitions can be made:

$$R_{tv}^0 = r_v n_{to} p_o \qquad R_{vt}^0 = r_v' N_v p_{to} \tag{4.49}$$

The two rates must be equal in equilibrium, so that the constant r_v' may be eliminated:

$$r_v' N_v = r_v n_{to} p_o / p_{to} = r_v p_o \frac{1 + e^{-\psi}}{1 + e^{\psi}} = r_v p_o e^{-\psi} \tag{4.50}$$

Similar considerations apply to the other two rates so that we may write finally:

$$\begin{aligned} R_{tv}^0 &= r_v n_{to} p_o \qquad & R_{vt}^0 &= r_v p_o e^{-\psi} p_{to} \\ R_{ct}^0 &= r_c p_{to} n_o \qquad & R_{tc}^0 &= r_c n_o e^{\psi} n_{to} \end{aligned} \tag{4.51}$$

The equilibrium situation described by these relations may be perturbed in one of two different ways, both of which will be relevant to our considerations. The first of these consists in *injection* of excess carrier densities, for example by photo-excitation or by the application of a forward bias to the p–n junction – both processes increase the densities of minority and majority carriers by the same amount, $\Delta n = \Delta p$, so that neutrality is maintained (Jonscher 1960). The presence of these excess densities causes an increase in both *downward rates* R_{ct} and R_{tv} without, to a first approximation, altering the upward rates R_{vt} and R_{tc}, at least in the case which is known as the Schockley–Read model where the occupancy of the trap level is assumed to remain unchanged in the presence of injection. In this limit, which reflects quite closely the actual behaviour of many

semiconductors, we obtain the following differential equations for the rates of change of carrier densities:

$$\frac{dn}{dt} = \frac{d\Delta n}{dt} = R_{tc} - R_{ct} = r_c[n_{to}n_o\, e^{\psi} - p_{to}(n_o + \Delta n)] \quad (4.52)$$

which after some simplifications, taking into account eqns (4.51), gives

$$d\Delta n/dt = -r_c p_{to}\Delta n = -\Delta n/\tau \quad (4.53)$$

which defines a time constant τ known as the *lifetime of injected carriers*.

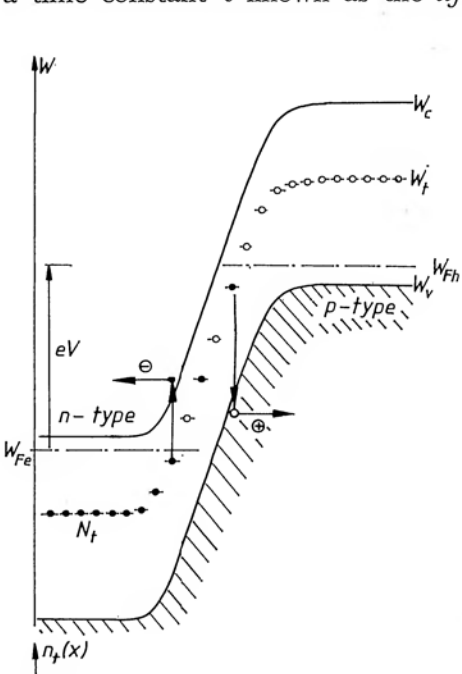


Figure 4.13 The energy band diagram of a p–n junction containing one deep trap level of density N_t and energy W_t, with an applied reverse bias V. The perturbation of thermal equilibrium gives rise to removal of electrons from traps near the edge of the n-type material, and to their addition to the traps near the edge of the p-type material. Trap occupancy is complete below the electron Fermi level W_{Fe} and is zero above the hole Fermi level W_{Fh}, with intermediate filling in the range between these two levels, as shown in the lower diagram. The effect of a slight further increase of bias is shown by the shaded regions which indicate the removal and addition of electrons to traps. The arrows indicate the excitations and subsequent transport of an electron and a hole.

The physical significance of this result is that, in the approximations used in this model, the rate of decay of excess carrier densities is proportional to these excess densities themselves. This means that the recombination process would give an *exponential* time decay if the source of injection, e.g. the light, were to be suddenly turned off, and the rate constant is $1/\tau$.

Typical values of lifetimes vary between tens or hundreds of microseconds in transistor quality silicon and germanium, and some nanoseconds in materials such as GaAs.

If instead of injecting carriers one were to remove them from some region, one would induce the opposite process of carrier *generation* which would tend to restore equilibrium disturbed by the removal process. This situation occurs typically in a p–n junction that has been biased in the reverse direction, as shown schematically in Figure 4.13. Here carrier generation takes place in the space adjacent to the space charge region and also, with a different time constant, in the space charge region itself (Sah, Noyce and Shockley 1957, Jonscher 1960). Under steady state bias this gives rise to the reverse leakage current of the p–n junction which appears like a constant conductance G_0 across the junction capacitance. If a small alternating signal is superimposed on the steady bias, with a frequency ω, then the generation process is modulated by that signal, since the edges of the space charge region move in step with the applied signal as the width of the space charge region is modulated. The resulting alternating generation/recombination process follows a time and frequency dependence which is obtained by solving the recombination equation (4.53) with a suitable harmonic driving term at the frequency ω. Using the phasor convention we obtain for the generation/recombination rate resulting from the modulation of the width of the space charge region the equation:

$$\mathcal{R}(\omega) = \mathcal{R}_0/(1 + \mathrm{i}\omega\tau) \tag{4.54}$$

where $\mathcal{R}_0$ is a suitable rate constant depending upon the magnitude of the space charge modulation, i.e. on the small signal voltage applied to the system, and also on the absolute magnitudes of the recombination rates and the configuration of the barrier profile.

Now consider the effect of a single generation process within the space charge region of the p–n junction on the external circuit. Figure 4.14 illustrates the physical situation. A generation process occurring at a certain time t_1 creates an electron–hole pair which is separated by the prevailing electric field in a time which might

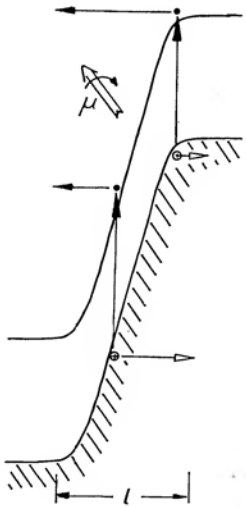


Figure 4.14 An illustration of the effect of generation of an electron–hole pair in the space charge region of a reverse biased p–n junction. The rapid transport of the two carriers gives rise to an effect equivalent to a sudden switch of a dipole between opposite orientations. The dipole moment is *el* regardless of where in the space charge region the generation process takes place.

be estimated as being the width of the space charge region divided by a suitable velocity. The latter we shall assume to be of the order of 10^5 m/s, i.e. the maximum attainable drift velocity of carriers, so that with a typical width of 1 μm the separation time is of the order 10^{-11} s.

This individual generation process, which is delayed with respect to the agency producing it, i.e. the voltage applied to the junction, by the mean time τ produces an effect exactly equivalent to a sudden switching of a dipole in a system with jumping dipoles, as described in Section 4.6.

The response of the capacitance of the space charge region of the p–n junction is therefore determined by the following processes:

i) the adjustment of the free carrier densities at the edges of the space charge region – this takes place with a time constant given by the space charge relaxation time τ_s, Appendix 4.2, generally falling in the nanosecond to picosecond range.
ii) the dielectric response proper of the lattice itself – this is dominated by the atomic and ionic relaxations in the infra-red range of the spectrum, i.e. tera-hertz frequencies.

iii) the response of the generation-recombination processes, determined by their lifetime which falls in the microsecond to nanosecond region, corresponding to frequencies in the megahertz to gigahertz range.

Only the last mechanism is likely to be of interest in the frequency range with which we are concerned here and the resulting frequency dependent complex capacitance $\tilde{C}(\omega)$ of a p–n junction contains therefore a frequency-independent component C_∞ arising from all the fast processes i) and ii), and an additional frequency-dependent component which obeys exactly the same frequency law as the classical Debye process in the physically entirely different context of dipolar phenomena:

$$\tilde{C}(\omega) = C_\infty + \Delta C/(1 + i\omega\tau) \qquad (4.55)$$

This is the conclusion of our analysis of generation/recombination processes in semiconductor p–n junctions, on the assumption of a simple recombination process obeying an exponential dependence on time. This conclusion is not influenced by the precise choice of the recombination process – it would be valid for any linear recombination process, whether it involved a single intermediate deep centre or several of them, or even if the recombination occurred directly between the conduction and valence bands in a bimolecular manner – the latter provided that the excitation level was sufficiently low. The sole requirement in the present situation is that the individual generation/recombination events should be *completely independent* of one another, so that the processes in question are *non-interactive*, the result of a large number of events occurring randomly is equivalent to the simple summation of their individual effects.

We note that a steady generation rate, such as would arise from the presence of a constant bias on a junction, results in a steady current representing the summation of an infinite series of impulses shown in Figure 4.14. This therefore gives rise to a constant conductance G_0 in parallel with the capacitance of the junction. Even if the signal is not constant but varies in time at a rate which is slow in comparison with $1/\tau$, this still gives an alternating current exactly in phase with the signal, i.e. a purely real conductance. A rise in conductance corresponding to the proper "dielectric" effect of the generation rate would only be expected at frequencies approaching $1/\tau$ and higher.

This means that a p–n junction in which the recombination time was of the order of 1 microsecond would only show "dielectric"

effects at Megahertz frequencies. Experience shows that junctions show such effects at much lower frequencies – down to Hertz region or below, and another mechanism has to be invoked to explain these phenomena.

iv) Trapping phenomena

This second mechanism is likely to be found in a different type of action that may be found in deep levels shown in Figure 4.12, if these levels show substantial transition rates to *one band* only. If, for example, the only interaction is with the conduction band, then the centre becomes an *electron trap*, since once an electron enters it from the conduction band it can only become re-excited back again, remaining in a trapped condition in the meantime. Since re-excitation requires a supply of energy, while recombination downwards disposes of energy, the mean trapping times are much longer than average recombination times, sometimes by many orders of magnitude. Such slow electron traps, and their counterparts in communication with the valence band, the hole traps, would therefore provide very slow responses to external stimulation and we shall discuss their role in a little detail.

The process may be seen in relation to a reverse biased p–n junction in Figure 4.13. Here a single deep trapping level is shown which is completely occupied on the n-side, since it lies deep below the Fermi level, and is completely empty in the region of p-type material where it is high above the Fermi level. In the presence of a reverse bias, the Fermi level becomes split into two "quasi-Fermi levels" which have the property that between them the occupancy of the centre is not dictated by their relative positions with respect to these Fermi levels, but is somewhere intermediate between full and empty, as shown in the lower diagram of Figure 4.13.

At any given reverse bias, including zero bias if required, there is a definite occupancy of these deep levels throughout the space charge region. If the bias is now increased, by the addition of a small-signal voltage, the space charge width increases slightly, thereby causing an adjustment of the occupancy of some of the levels at the planes in which the quasi-Fermi levels cut through the deep levels. The lower diagram in Figure 4.13 shows that some of these levels on the n-side have their electrons removed from them by excitation into the conduction band, where their effect is exactly the same as that of generated electrons in Figure 4.14. Likewise, if there were some hole traps on the p-side, they would need to have some holes removed, or electrons added to them in the new situation.

In order to obtain a quantitative estimate of the rate of removal of electrons from deep traps as the Fermi level changes position with respect to them, we shall assume that the situation is spatially uniform, instead of being confined to a small volume near the plane where the Fermi level crosses the trap level – this will not affect the *time dependence* of the processes in question, only their absolute magnitude.

Writing in the general case of transitions to and from the conduction band only, and making use of equations (4.47)–(4.51) we have the following rates:

$$R_{ct} = r_c n p_t \tag{4.56}$$

$$R_{tc} = r_c n_o (p_{to}/n_{to}) n_t = r_c n_o n_t \, e^{\psi_0} \tag{4.57}$$

where ψ_0 is an equilibrium position of the Fermi level around which there are small fluctuations, $\psi = \psi_0 + \psi_1$. We assume that the free carrier density in the conduction band follows very rapidly any changes of the Fermi level, so that $n = n_o \, e^{-\psi_1} \simeq n_o(1 - \psi_1)$. The rate of change of trapped electron density is found to be:

$$dn_t/dt = [(N_t - n_t)(1 - \psi_1) - n_t \, e^{\psi_0}] r_c n_o$$

and writing for the trapped density $n_t = n_{to} + n_t'$, we obtain after simple transformation, taking into account the cancellation of the equilibrium rates, the following equation for the small-signal trapped charge density as a function of the small-signal variation of the position of the Fermi level:

$$dn_t'/dt = -n_t'/\tau_t - \psi_1/K \tag{4.58}$$

where

$$1/\tau_t = r_c n_o (1 + e^{\psi_0}) \tag{4.59}$$

and

$$1/K = r_c n_o p_{to} \tag{4.60}$$

The rate equation (4.58) is of the same form as the recombination equation (4.53), except that it includes the driving term, while eqn (4.53) was obtained for the time-decay of density after the removal of injection. The general solution of this equation is of the same form as (4.54), except for the different magnitude of the trapping time constant τ_t compared with the recombination time τ. We are interested in values of ψ close to zero, i.e. the Fermi level is crossing the trap level. Assuming r_c to be of the same order of

magnitude in both cases, we note that with the Fermi level at the trapping states, the free carrier density n_0 is much smaller than the density p_{t0} in the recombination case in p-type material. This shows that the trapping time τ_t is much longer than the recombination time τ, as was indicated on the basis of qualitative arguments.

This analysis proves that the effect of trapping states obeying the simple relations assumed in the present case is to give the same frequency dependence as the ideal Debye process, but the time constant may now be very long, i.e. the phenomena in question would be expected to be observable at low and very low frequencies.

It may be noted that the recombination times and trapping times described above are likely to be strongly temperature-dependent, in view of the thermal nature of the recombination and trapping processes in question.

We shall see in Chapter 5 examples of junction response which confirm the existence of loss processes and which show strong temperature activation of the trapping rates, but which also depart strongly from the ideal Debye response, indicating clearly that the simple exponential time dependence is not consistent with experimental observations.

4.8 DIFFUSIVE TRANSPORT

We conclude this review of the principal simple models of dielectric response with an example of a behaviour that is not normally associated with the subject of dielectrics, but which illustrates several interesting physical principles. While the models discussed hitherto were essentially one-particle models in which the effect of an assembly of particles was the summation of the effects of the individual particles in the assembly, the next model to be considered is in a sense a "cooperative" one, in that it requires the presence of a definite distribution of particles in space and in time. This constitutes an essential novely of this model in comparison with the other one-particle models which all gave a Debye type of response, with the exception of the free electron model.

There are many situations in physics and in physical chemistry in which charge carriers are introduced into a volume of material under conditions of complete neutrality, i.e. there exist two types of charge carriers of opposite signs and one is always available to neutralise the effect of the other. A well known example of this type of behaviour is found in a forward-biased p–n junction which injects

minority carriers into extrinsic regions of opposite types on both sides of the junction. Since the *majority* carriers are always present in significantly larger numbers than the injected minority carriers, their supply to the injection region is unrestricted and neutrality is always maintained. This means that the transport of the injected densities occurs by a nearly-field-free diffusive mechanism in which a gradient of the density of both carrier species, $n(x, t)$ has to be established in order to produce the flow of the particle current $j(x, t)$. The diffusion current in one dimensional case is:

$$j(x, t) = -D\,\partial n(x, t)/\partial x \tag{4.61}$$

while the continuity equation for charge flow in the absence of recombination may be written in the form

$$\partial n/\partial t = -\partial j/\partial x = D\partial^2 n/\partial x^2 \tag{4.62}$$

where D is the diffusion coefficient for the charge carriers. As boundary conditions we take a slab of semiconductor of thickness w with the injecting contact at $x = 0$ and an ohmic contact maintaining zero excess density at $x = w$. The density at the injecting contact is assumed to consist of a steady value due to a static bias and a small-signal component of frequency ω:

$$n(0, t) = n(0) + n' \exp(i\omega t) \tag{4.63}$$

The steady state solution due to the density $n(0)$ is given by the condition of continuity $\mathrm{div}\, j = 0$, which means that

$$n_0(x) = n(0)(1 - x/w) \tag{4.64}$$

as shown in Figure 4.15. This density gradient gives rise to the

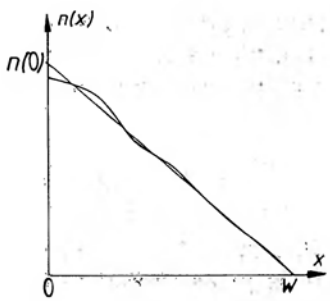


Figure 4.15 The distribution of density of injected excess charge carriers, e.g. electrons in a forward-biased p–n junction, or ions in an electrolyte next to an electrode, showing the steady state distribution by the faint line and the distribution under an alternating signal by the thick line.

steady state (dc) current which in the present instance is given by:

$$j_0 = Dn(0)/w \tag{4.65}$$

which is independent of both space and time variables, since we assume that there is no recombination.

The time-dependent component is obtained as a solution of eqn (4.62) in view of the linearity of the system which admits additivity of solutions. Because of the harmonic time dependence we replace the time-derivative with $i\omega$ and obtain the small-signal equation:

$$\partial^2 n'/\partial x^2 = (i\omega/D)n' = n'/L^2 \tag{4.66}$$

where we have defined the complex "ac diffusion length"

$$L = (D/i\omega)^{1/2} = (2D/\omega)^{1/2}/(1 + i) \tag{4.67}$$

The solution for n' is of the general form:

$$n' = n'(0)\{\cosh(x/L) + A \sinh(x/L)\} \tag{4.68}$$

where the constant A has to be determined to obtain $n'(w) = 0$. Differentiating this expression with respect to distance we obtain the small-signal current at the injecting contact:

$$j'(0) = n'(0)A/L = A(\omega/2D)^{1/2}(1 + i)n'(0) \tag{4.69}$$

The important conclusions for the present are that the current may be expressed in the form:

$$j'(\omega) = \text{const } \omega^{1/2}(1 + i)V' \tag{4.70}$$

where we have postulated the existence of a linear relationship between the small-signal applied voltage V' and the density n' – a very reasonable assumption under small-signal conditions. We may therefore finally obtain the effective susceptibility of the system using eqn (3.15):

$$\chi(\omega) \propto \frac{Y(\omega)}{i\omega} \propto (1 - i)\omega^{-1/2} \propto (i\omega)^{-1/2} \tag{4.71}$$

This result is remarkable in that the frequency dependence of the dielectric susceptibility is proportional to $\omega^{-1/2}$ and also the real and imaginary components of the susceptibility – and of the admittance – are equal. This is a complete departure from the Debye response and the full significance of this will become apparent later in the context of a more general type of frequency dependence.

For the present we shall mention that the peculiar frequency dependence of diffusing charge carriers derives from the very specific "inertia" of the diffusion process – the fact that a density gradient has to be established before the current can flow and this gives rise to very slow responses. A schematic representation of the instantaneous density distribution in the case of an alternating small-signal boundary condition superimposed on a steady gradient is shown in Figure 4.15. Very good illustrations of the importance of attenuation of alternating diffusive flow are to be found in heat flow problems which are governed by identical relations. It is therefore interesting to note that the diurnal heat cycles do not penetrate more than a few centimetres into the ground, while even the very slow and severe summer–winter temperature cycles hardly go beyond a metre. Further discussion of this may be found in Carslaw and Jaeger (1949).

Although we have carried out the analysis on the specific example of a forward biased p–n junction, very similar considerations apply to interfacial electrochemical processes, where ionic species dissociate at the electrodes and diffuse into the neutral bulk of the electrolyte – a process recognised by Warburg as early as 1890, with the result that this type of admittance is known in electrochemical circles as the Warburg impedance. Furthermore, similar behaviour may also be found in solid electrolytes – materials with exceptionally high ionic mobilities which have attracted a good deal of attention recently, cf. Section 5.5 c).

4.9 CONCLUDING COMMENTS

Our discussion of the frequency response of idealised physical systems has revealed a number of features which it would be instructive to highlight at the present moment in the development of the argument. Any system which can be described by a first-order linear differential equation of the type

$$dy/dt = -y/\tau + \text{a force term}$$

must necessarily lead to a Debye-type response in the frequency domain and to an exponential relaxation on sudden removal of a polarising force. In these sytems the rate process is proportional to the perturbation itself and there is a restoring force present in the system tending to re-establish the equilibrium. Our inertialess system with restoring force, floating dipoles, charge between potential wells and all forms of recombination and trapping processes

in which there is linear rate of return to equilibrium, all show the same type of response.

Departures from the classical relaxation law arise when there are inertial forces at play, since they imply a second-order differential equation, and where there are diffusive – i.e. collective – effects, since a second-order equation is also involved.

The case of free electrons also departs from the standard relaxation because there is no physical restoring force in that situation.

It is worth stressing here that, while a number of departures from the classical relaxation law has been described, these represent very well defined frequency- and time-dependences, so that there is no scope for "fitting" any arbitrary functions in terms of these laws, for example when the experimental data take the form of broadened and asymmetric loss peaks. The only way to "fit" the data in terms of these elementary laws would be to assume superposition of several such laws, and this is being done in the context of Debye-like relations in the approach which uses a Distribution of Relaxation Times, Section 7.2.

While one may attempt to model a broadened loss peak in terms of summations of ideal Debye responses with a range of relaxation times, it would be very difficult to represent a loss process with a fractional power law of exponent different from $\frac{1}{2}$ by a superposition of diffusive processes which necessarily have exponents $\frac{1}{2}$ exactly.

APPENDIX 4.1

THE COMPLEX SUSCEPTIBILITY OF AN INERTIAL SYSTEM WITH RESTORING FORCE.

It is interesting to obtain approximate expressions for the complex plane plots of the susceptibility of an inertial system with restoring force, as given by eqns (4.2) and (4.3), so as to be able to compare them with the more familiar inertialess (Debye) and similar plots. Taking

$$\chi(\omega) = \frac{1}{1 - x^2 + ikx}$$

with $k \ll 1$ signifying small damping, and taking x close to unity, we set $z = 1 - x \ll 1$

$$\chi(z) \simeq (2z + ik - ikz)^{-1} \simeq (2z + ik)^{-1}$$

This corresponds to the nearly circular part of the plot in Figure 4.4 which crosses the imaginary axis at $\chi'' = 1/k$. For small values of $x < k \ll 1$ we may write:

$$\chi(\omega) = (1 - x^2 + ikx)^{-1} \simeq 1 + x^2 - ikx$$

which corresponds to the part shown in the inset. Finally, for very large values of x we find:

$$\chi(\omega) \simeq (-x^2 + ikx)^{-1} \simeq -(1/x^2)(1 + ik/x)$$

which gives the negative branch close to the origin.

Corresponding plots for larger values of the damping coefficient k are also shown and it is clear that the Debye response corresponds to the limit $k \to \infty$.

APPENDIX 4.2

RELAXATION OF "FREE" CHARGE

The presence of excess free charge in a material gives rise to a local high electric field which tends to "blow up" that charge, thus giving a rapid dispersion of space charge. An approximate analytical treatment of this situation may be obtained as follows.

Let the space charge be ρ, arising, for example, from injection of majority electrons into the solid, or a thermal fluctuation locally, or any other means. The continuity of particle current is obtained from Maxwell's equation

$$j + \partial D/\partial t = \text{curl}\, H \tag{A4.1}$$

and from Poisson's equation

$$\text{div}\, D = \rho \tag{A4.2}$$

By taking divergence of (A4.1) and noting that div curl $\equiv 0$ we obtain

$$\text{div} j = -\partial\rho/\partial t \tag{A4.3}$$

Now set

$$j = \sigma E \simeq \sigma_0 E \tag{A4.4}$$

where σ_0 is the local equilibrium electrical conductivity. We may then write

$$\text{div} j = \text{div}(\sigma E) \simeq \sigma_0\, \text{div}\, E = (\sigma_0/\varepsilon)\rho \tag{A4.5}$$

The approximation used here assumes that the injection of excess charge does not alter significantly the conductivity σ_0. We then obtain from (A4.3) and (A4.4) the relation:

$$\partial\rho/\partial t \simeq -(\sigma_0/\varepsilon)\rho = -\rho/\tau_s$$

which implies an exponential decay of any space charge

$$\rho = \rho_0 \exp(-t/\tau_s) \tag{A4.6}$$

where τ_s is the space charge relaxation time.

Up to the present moment the analysis involved the decay of a disturbance of density in a *strongly damped* or *collision-dominated* electron gas. This condition is implied by the use of the conductivity equation (A4.4) which states, in effect, that the velocity of charges is proportional to the force acting on them – a typical viscosity-limited behaviour.

Consider next the case of a *free carrier* gas in which collisions may be neglected on the time-scale under discussion. This may correspond to a truly free-carrier gas, e.g. a rarefied plasma, or to a solid state electron gas in a metal or a semiconductor, where collisions occur at a rate of $10^{14}\,s^{-1}$, but in which we are interested in the response to electric fields at optical frequencies of $10^{16}\,s^{-1}$. Under these conditions metals behave as almost free carrier systems (Platzman and Wolff 1973).

In this case we put $j = eNv$ in the continuity equation (A4.3) and write the equation of motion in the absence of collisions:

$$\partial v/\partial t = eE/m$$

After simple manipulations we obtain

$$\partial^2\rho/\partial t^2 = -\Omega_p^2\rho \qquad \text{(A4.7)}$$

which implies that any disturbance of space charge produces an undamped oscillatory response

$$\rho = \rho_0 \exp(i\Omega_p t) \qquad \text{(A4.8)}$$

where Ω_p is the plasma frequency referred to in eqn (4.4). The physical significance of this result is that a density fluctuation producing a local space charge causes opposing electric fields to build up which in due course repel the excess charge, but in doing so cause it to "overshoot", i.e. to produce a space charge elsewhere. The absence of collisions means that this phenomenon is not damped, at least not in the first approximation.

REFERENCES TO CHAPTER 4

Abramovitz M and Stegun I A 1965, *Handbook of Mathematical Functions*, Dover Publications, New York

Carslaw H S and Jaeger J C 1949, *Operational Methods in Applied Mathematics*, Oxford University Press

Debye P 1945, *Polar Molecules*, Dover Publications, New York

Elliott R J and Gibson A F 1974, *Solid State Physics*, Macmillan, New York

Erdelyi A 1954, *Tables of Integral Transforms* Vol. II, 249, McGraw-Hill, New York

Grove A S 1967, *Physics and Technology of Semiconductor Devices*, John Wiley, New York
Henisch H K 1957, *Rectifying Semiconductor Contacts*, Oxford, Clarendon Press
Jonscher A K 1960, *Principles of Semiconductor Device Operation*, G Bell and Sons London and John Wiley, New York
—— 1976, Thin Solid Films **36**, 1
Kittel C 1964, *Quantum Theory of Solids*, John Wiley, New York
Mott N F and Davis E A 1979, *Electronic Processes in Non-crystalline Materials*, Oxford University Press, London and New York
Parker T J, Mok C L and Chambers W G 1979, IEE *Conference Publication* 177, 207
Platzman P M and Wolff P A 1973, *Waves and Interactions in Solid State Plasmas*, Academic Press, New York
Sah C T, Noyce R N and Shockley W 1957, Proc. IRE 46, 1076
Smith R A 1978, *Semiconductors*, Cambridge University Press, 2nd Edition
Sze S M 1969, *Physics of Semiconductor Devices*, John Wiley, New York
Vashishta P, Mundy N J and Shenoy G K (Eds) 1979, *Fast Ion Transport in Solids*, Elsevier, Amsterdam
Ward I M 1971, *Mechanical Properties of Solid Polymers*, John Wiley, New York

CHAPTER 5

Experimental Evidence on the Frequency Response

5.1 INTRODUCTION

The reader is by now familiar with the background of the theory of dielectric relaxation, at least in its more idealised manifestations, to be able to appreciate the nature of the dielectric properties of *real materials*. The present Chapter gives a broad review of the dielectric response of a wide range of materials with the object of giving the reader an unbiased impression of the true behaviour, as distinct from the idealised model responses, because only in this way will he be able to appreciate the limitations of the latter. Accordingly, this chapter will be uncommitted with regard to interpretation of data – no attempt will be made to pass value judgements on the significance of the results – the emphasis will be on a clear presentation of the results in a unified manner. This should enable comparisons to be made between different sets of data, many of which may not have appeared in the same context previously. This point is very important, because tradition and custom have contributed to often strangely ineffective ways of presenting dielectric information which then admit of many different interpretations, thereby compounding certain misconceptions regarding the significance of these results.

Thus, for example, the common way of plotting the frequency dependence of the real and imaginary components of permittivity consists in using the *linear* ε'' and ε' against the logarithm of frequency, while the much more informative plot is the log–log representation, since as we shall see, these functions often give the power law dependence which is immediately recognisable in the logarithmic presentation while giving a rather meaningless curvature in the semi-logarithmic plot. Similarly, the widely used complex permittivity diagram, known as the Cole–Cole plot, is very insensitive to the detailed frequency dependence of loss, especially in the vital regions further away on either side of the loss peak frequency.

Wherever possible, we shall attempt to give the real and imaginary components of the *susceptibility* for one particular mechanism, since this gives the best chance of having a meaningful picture.

The presentation of the experimental information is arranged according to the following plan. We begin with examples of dielectric response approximating as closely as possible to the ideal Debye behaviour and from there we proceed to examples of increasingly strong departures from this ideal behaviour, while still remaining with materials in which there exist recognisable *loss peaks*. The next class of response types will be characterised by a complete absence of dielectric loss peaks and will show instead a region of strong dispersion of both the real and the imaginary parts of the susceptibility. This type of behaviour will be identified with a well defined mechanism which is completely different from the commonly accepted Maxwell-Wagner interfacial mechanism that is being invoked in such cases.

We shall also have the occasion to present examples of dielectric losses which are virtually independent of frequency over extended ranges. These losses will again be identified with certain physical conditions.

In our presentation we shall depart in one more respect from traditionally accepted conventions: our selection of examples will not be made with reference to any particular material or class of materials but will use examples of a given type of response drawn from the entire spectrum of materials. The intention here is very definite: to convince the reader that the dielectric response of very different solids shows certain similarities which will help us decisively at a later stage in determining the true nature of the physical mechanisms involved.

Indeed, it may be said that to the extent to which our entire approach to the dielectric response of solids has contributed to a better understanding of the physical principles underlying this response, this was possible because of the synoptic approach we have taken, unfettered by the accepted models or by preconceived ideas as to *how dielectrics should respond, instead of how they do respond.*

An important point to be noted here concerns the extent to which it is necessary to understand in detail the physical and chemical structure and general behaviour of these often very different materials in order to be able to appreciate their dielectric response.

The view taken in the present work is that this is not necessary insofar as these various materials *follow a common pattern of dielectric behaviour*. For this reason we refrain from a detailed presentation of the general properties of ferroelectric materials, liquid crystals, fast ionic conductors or, for that matter, of water, each of which would form the subject of a major review in its own right. The examples of their dielectric behaviour which are quoted here do not necessarily lend themselves to a clear interpretation within the classical systems of dielectric theory and we shall propose to give some interpretation in the light of the new theory in Chapter 8.

5.2 NEAR-DEBYE RESPONSES

There exist very few examples indeed of the ideal Debye response in condensed matter generally, let alone in solids. One of the more sensitive tests of conformity to Debye consists in a detailed analysis of the frequency dependence of the *real* part of the susceptibility at frequencies in excess of the loss peak frequency ω_p. The test relies on the fact that Debye behaviour shown in Figure 3.23 gives a clear divergence of the real and imaginary components, whose logarithmic slopes are -2 for $\chi'(\omega)$ and -1 for $\chi''(\omega)$. This Kramers–Kronig-compatible result represents a *singularity* of dielectric behaviour, and departure, however slight, towards power-laws of the type ω^{n-1} with $n \neq 0$, demands as a consequence of Kramers–Kronig relations that both should follow *the same* frequency dependence, giving a frequency-independent ratio, Figure 3.24

$$\chi''(\omega)/\chi'(\omega) = \cot(n\pi/2) \qquad 0 < n < 1 \tag{5.1}$$

This result has nothing whatever to do with any particular physical mechanism or interpretation – it is the consequence of the principle of causality, on which Kramers–Kronig is based and its applicability testifies to the self-consistency of the results obtained experimentally.

The requirement here is the availability of very accurate data for $\varepsilon'(\omega)$ from which a suitable value of ε_∞ may then be subtracted to give $\chi'(\omega)$. This procedure is not as arbitrary as may appear at first sight, since it would not be possible to obtain a spurious "fit" to $1/\omega^2$ instead of $1/\omega$, or vice versa, over any significant frequency range.

The nearest approximation to the Debye response which we have found in the considerable range of experimental material investi-

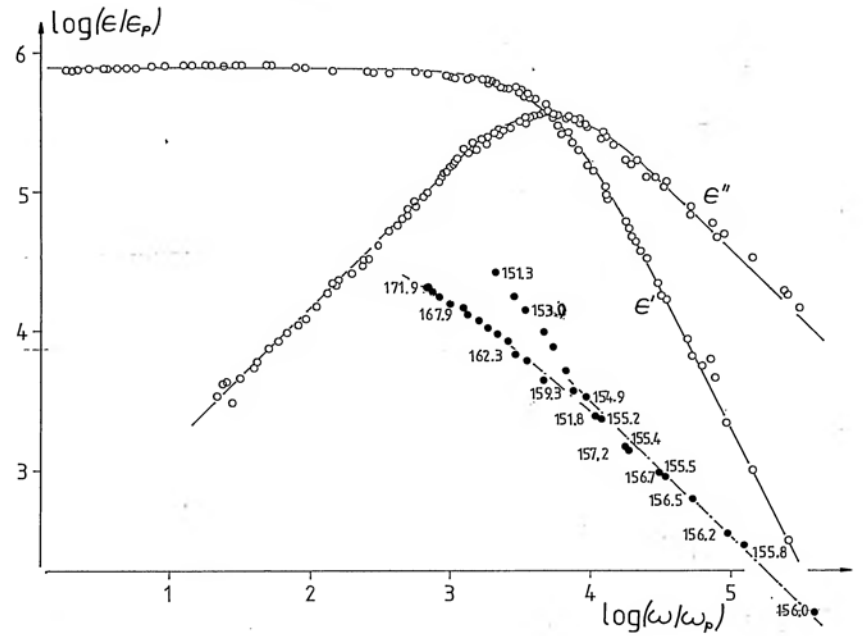


Figure 5.1 An example of an "ideal" Debye response in a ferroelectric single crystal of Cesium dihydrogen phosphate (CsH_2PO_4) which undergoes a ferroelectric transition at around 156 K. The data are normalised from measurements at five spot frequencies of 1.0, 12.0, 72.4, 251.2 and 1,000.0 MHz and in a temperature range between 151 and 172 K. In the measuring range the values of the permittivity ε were so high that it was not necessary to subtract the ε_∞. The solid points give the locus of the reference point (*cf.* Section 3.8) and the slope of the chain-dotted line at temperatures around the Curie temperature $T_C = 156.0$ K is drawn parallel to the loss slope of -1. The significance of this result will be discussed in Chapter 8, but for the present it is sufficient to point out the agreement between the normalised points (open circles) and the theoretical Debye response drawn as solid lines. It should be noted that there does not appear to be any departure from the -2 slope of $\varepsilon'(\omega)$. The logarithmic scales are purely arbitrary, the only significant information is the magnitude of the decades. From experimental data by Deguchi et al (1982), normalisation by R M Hill, reproduced by permission.

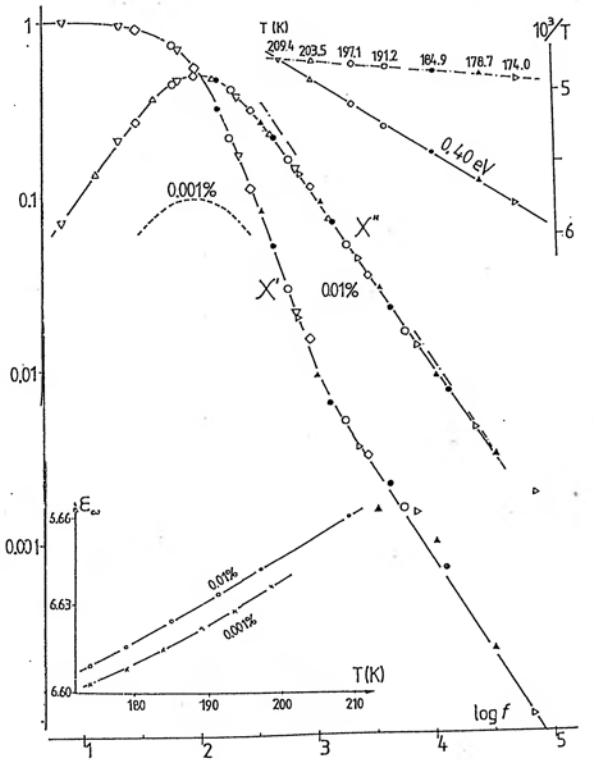
gated both in-house and from published literature data is shown in Figure 5.1 which relates to the response of a ferroelectric single crystal, Cesium dihydrogen phosphate and the data have been normalised from measurements taken at five spot frequencies and over a range of temperatures on either side of the Curie temperature. The more detailed significance of the shape of the normalisation trace and the reasons for this particular behaviour of ferroelectrics will be discussed in Chapter 8, for the present we point out the

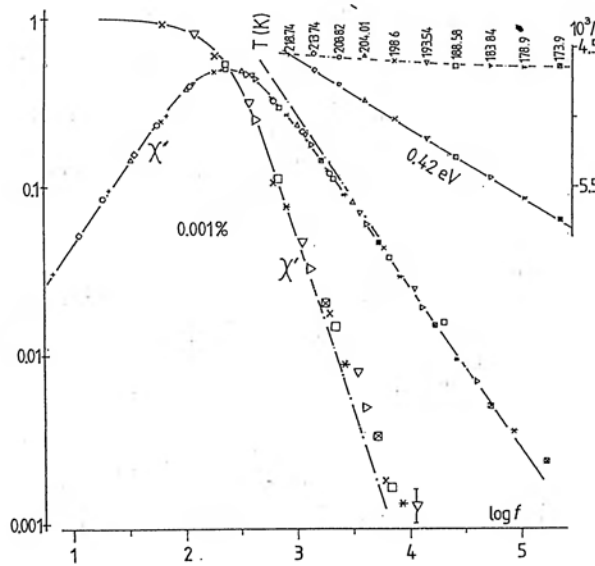
close agreement with the Debye response, suggesting that to the extent of the accuracy of the available data there is no departure from it.

A good example of the situation in which the experimental data reveal a clear departure from Debye response is shown in Figure 5.2 giving the data for single crystal CaF_2 doped with Er to the extent of 1 part in 10^4 (0.01%) (Fontanella et al 1978). Each trivalent Er ion replacing for the divalent Ca ion requires the introduction of an extra interstitial F^- ion to preserve charge neutrality and the $Er^{+++} - F^-$ pairs represent dipoles whose orientation changes abruptly as the interstitial F^- ions take up any one of the equivalent nearest-neighbour positions surrounding the Er^{+++} ions. The experimental data covering a range of temperatures were fitted by choosing the best value of ε_∞ to subtract from the real part of the permittivity. This fit shows unmistakably that the relation (5.1) is well maintained beyond the loss peak and separation between the real and imaginary parts gives a good indication of the departure from unit slope.

It is worth noting that the doping density of 1 part in 10^4 corresponds to a mean distance between the dipoles of 22 lattice spacings which would suggest a weakly interacting system – almost independent dipoles. Measurements on crystals with an even lower concentration of 1 part in 10^5 gave a lower loss, indicated by the dotted line in Figure 5.2 but the resulting accuracy of the real part was insufficient to enable one to derive a reliable fit for the high-frequency limit. This result suggests that attempts to produce true Debye-like responses by "diluting" the dipolar species in order to remove interactions do not give conclusive results since the magnitude of the loss decreases more rapidly than the strength of the interactions.

A similar analysis of data for water at 20°C is shown in Figure 5.3 relating to data from Mason et al (1974). Measurements on water are available with a considerable accuracy and the paper quoted gives carefully statistically weighted data to obtain the most reliable pattern of behaviour. The best values of the exponents in the presumed expressions ω^m below the loss peak and ω^{n-1} above the peak are given, showing clearly departures from the true Debye relation. We note that water is considered to be one of the classic Debye dielectrics so that the demonstration of this departure from true Debye behaviour, however slight, constitutes a significant result.


T (K)
209.4 203.5 197.1 191.2 184.9 178.7 174.0
10³/T
5
6
0.40 eV
0.001%
χ''
χ'
0.01%
1
0.1
0.01
0.001
ε_ω
5.66
6.63
6.60
0.01%
0.001%
T(K)
180 190 200 210
log f
1 2 3 4 5


T (K)
218.74 213.74 208.62 204.01 198.6 193.54 188.58 183.84 178.9 173.9
10³/T
4.5
5.5
0.42 eV
χ''
0.001%
χ'
1
0.1
0.01
0.001
log f
1 2 3 4 5

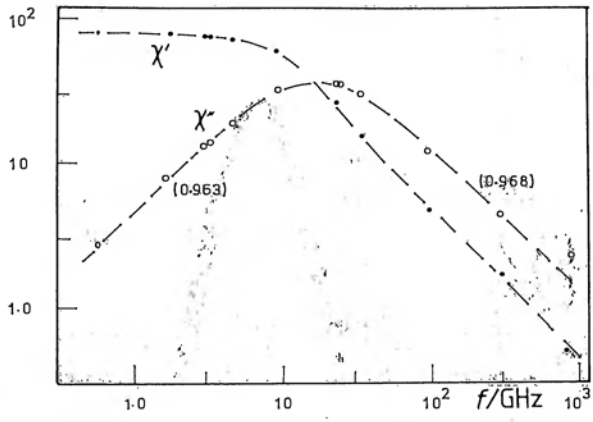


Figure 5.3 The "best fit" to the values of the dielectric permittivity of water at 20°C plotted as the real and imaginary parts of the susceptibility, subtracting the value $\varepsilon_\infty = 3.8\ \varepsilon_0$ from the real part $\varepsilon'(\omega)$. Data from Mason et al (1974), plotting by R M Hill. Note the frequency scale is in GHz. The slopes shown are consistent with the values of the real and imaginary components, particularly beyond the loss peak where the self-consistency of the two curves provides an additional fitting criterion.

The loss spectrum of ice at several temperatures normalised to a common peak is shown in Figure 5.4. The comparison with the theoretical Debye shape shows the very slight deviation towards broader peak response, although this mode of presentation does not easily admit of a quantitative evaluation of the departure. It is clear, however, that the peak tends to progressive broadening with decreasing temperature, in common with many other materials.

A different mode of presentation of similar data is shown in Figure 5.5 where the Cole–Cole plots are shown for a range of temperatures and for the orientation of the field parallel and perpendicular to

Figure 5.2 The dielectric response of CaF_2 represented as the real and imaginary components of the susceptibility against frequency. The data obtained at various temperatures are normalised and the loci of the displacement points are shown, together with the corresponding activation diagrams plotted against $1/T$. The values of ε_∞ used in the normalisation procedure are shown in the upper diagram as functions of temperature, indicating an almost linear rise with T, very similar for both samples. The upper diagram refers to a sample doped with Er to 1 part in 10^4, the lower 1 part in 10^5. The dotted peak in the upper diagram represents the relative position of the loss peak in the less heavily doped sample. The more heavily doped sample shows clear departure from the Debye relation in the sense of Figure 5.1, the data for the purer sample cannot be judged with certainty because of the limitations of sensitivity.

Data from Fontanella et al (1978) presentation by Jonscher (1980a).

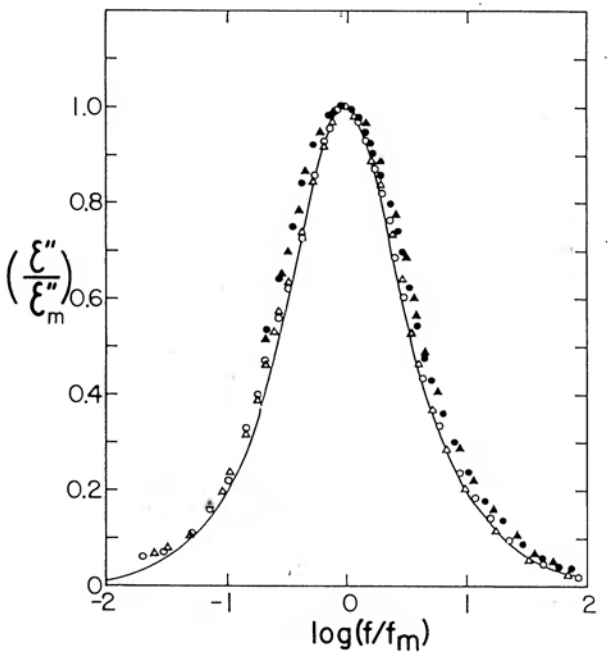


Figure 5.4 Normalised plots of the dielectric loss of ice against frequency at several temperatures. The solid line is the theoretical curve for a Debye relaxation. ○, parallel to the c-axis at 238.1 K, ● do. at 201.5 K, △ perpendicular to the c-axis at 235.9 K, ▲ do. at 202.6 K. f_m signifies the loss peak frequency, ε''_m the loss peak amplitude.

From Johari and Jones (1978). Reproduced from the *Journal of Glaciology* by permission of the International Glaciological Society.

the c-axis. The plots become progressively more circular as the temperature increases and at the same time a low-frequency "spur" appears which is due to some form of charge carrier transport – probably of the Warburg diffusive type.

Many examples of slightly broadened loss peaks, manifesting themselves as slightly tilted Cole–Cole plots shown in Figure 5.6 for the case of triglycine sulphate, are found in ferro- and anti-ferro-electric materials in the region of "giant dispersion", usually at GHz frequencies. The behaviour is very similar on both sides of the Curie temperature and the loss peak frequency is seen to increase as the temperature moves away from the Curie temperature.

This type of response is closely similar to that shown in Figure 5.1 but in that case the semicircle would be centred on the real axis — the Cesium dihydrogen phosphate behaves like a more perfect material than the TGS shown in Figure 5.6.

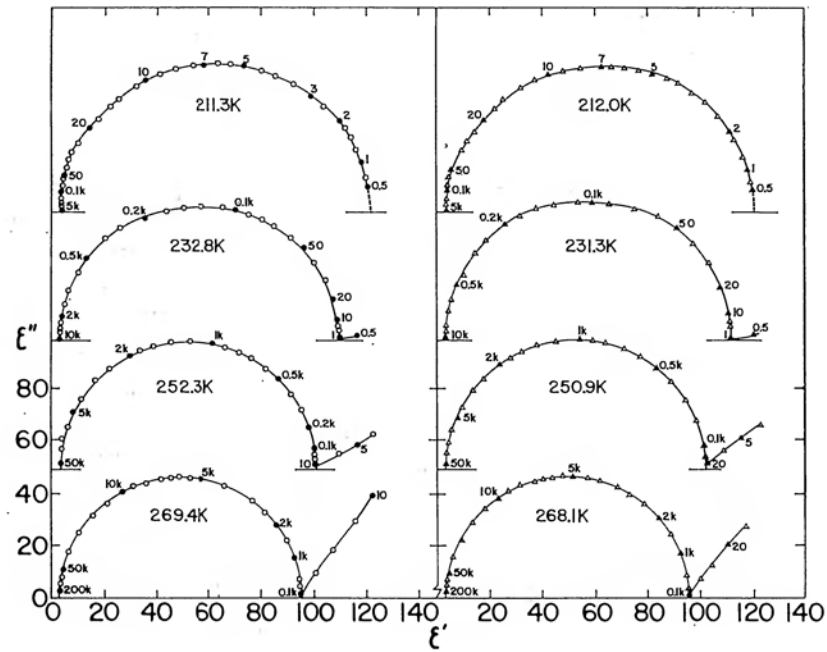


Figure 5.5 Complex plane plots of the relative permittivity and loss of ice at several temperatures. The left half is for measurements with the c-axis oriented parallel to the electric field, and the right half for that oriented perpendicular to the field. Numbers besides the filled data points are the frequencies in Hertz. The plots have been progressively shifted upwards as indicated by the axis.

Johari and Jones (1978). Reproduced from the *Journal of Glaciology* by permission of the International Glaciological Society.

A remarkable example of near-Debye behaviour is the dielectric response of a p–n junction in silicon from which deep-level-forming impurities had been carefully removed. Figure 5.7 shows the data for $C'(\omega)$ plotted *linearly* and $C''(\omega)$ against frequency, for a range of temperatures. The normalisation of the loss data is also shown and so is the plot of the activation energy which is a straight line, indicating that the activation energy does not change in the temperature range under study. The rise of loss at high frequencies is virtually independent of temperature and it corresponds to the effect of the series resistance of the diode. The loss peak amplitude increases slightly with increasing temperature, which is consistent with the fact that the width of the space charge region decreases due to the fall of the diffusion potential V_D. The reduction of the dispersion of capacitance with falling temperature is due to the gradual movement of the loss peak out of the frequency window.

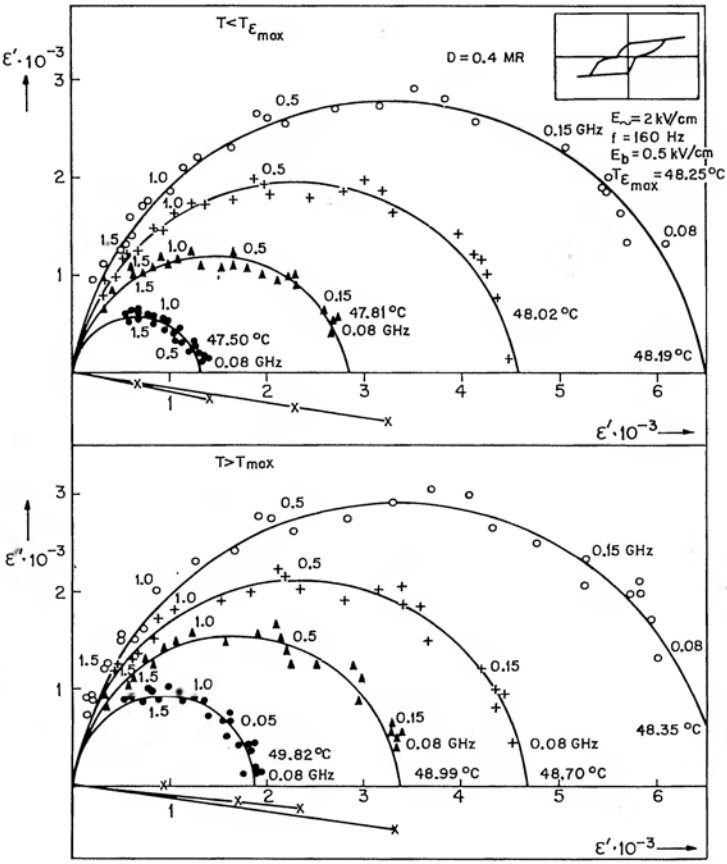


Figure 5.6 Complex permittivity diagrams for Triglycine Sulphate (TGS) below and above the Curie temperature. The frequencies indicated are in GHz. The inclination of the lines of centres of the circular arcs gives a measure of the departures from the ideal Debye behaviour.

From Pawlaczyk (1977).

The high-frequency capacitance is strongly temperature dependent due to the variation of the diffusion potential causing a narrowing of the space charge region with increasing temperature, as outlined in Section 4.7. At the lowest temperatures there appears a relatively frequency-independent loss and this corresponds to the remarkably low value of the loss angle of 10^{-4} radians – approaching that of the best insulators. The relaxation time at 265 K is of the order of 10 μs, which is eminently compatible with minority carrier recombination, without the need to invoke deep trapping levels.

We note that in the case of p–n junctions it is more appropriate to give the dielectric data in terms of the complex capacitance rather than as the complex permittivity, since the width of the space

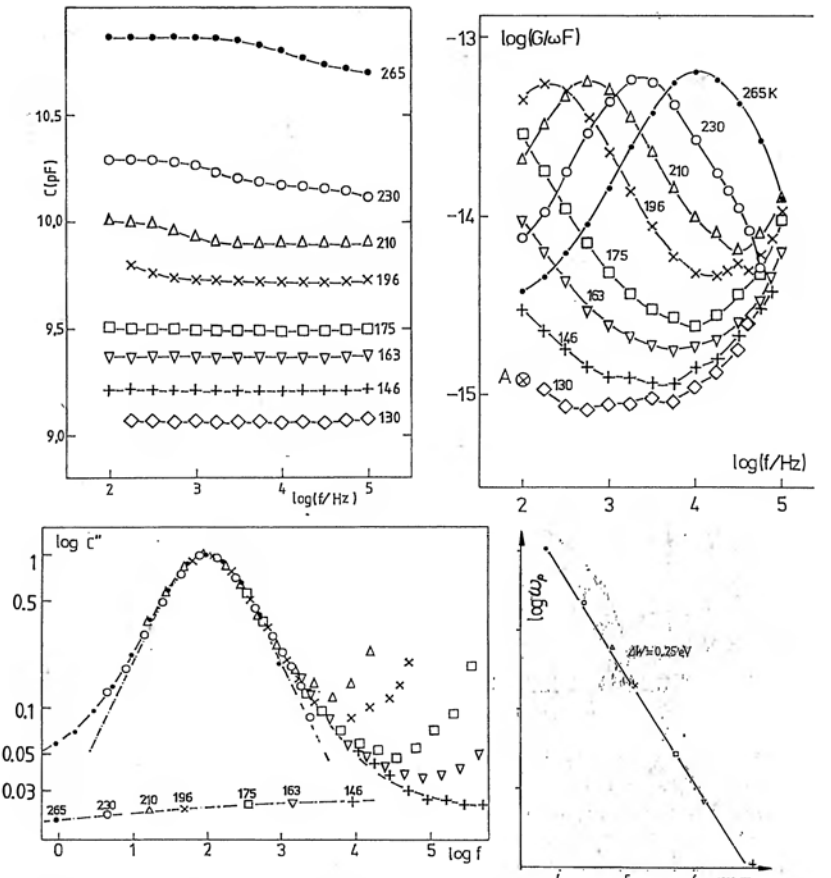


Figure 5.7 The frequency dependence of the complex capacitance of a silicon p–n junction believed to contain few deep trapping levels. Diagram a) shows the real part $C'(\omega)$ plotted linearly against $\log f$, diagram b) gives the corresponding loss $G(\omega)/\omega$, both for a range of temperatures. Diagram c) gives the normalisation of loss, with the locus of the translation of the representative point "A" in b), with the Debye contour shown by the chain-dotted line. Diagram d) gives the plot of the logarithm of the peak frequency against reciprocal temperature, indicating the presence of a single well defined activation energy of 0.27 eV. Loss data at high frequencies show clear evidence of series resistance of approximately 50 Ω which is consistent with the construction of the device.

From C K Loh (1978).

charge region is not known independently, quite apart from the fact that many devices are encapsulated and it is not possible to obtain their area.

A remarkable example of anisotropy of the dielectric response is shown in Figure 5.8 giving the complex ε plot for a liquid crystal, i.e. a molecular liquid in which the molecules have a considerable degree of coordination. With the electric field parallel to the molecules the response resembles closely that of a ferroelectric, with the important difference that the frequencies for the "giant dispersion" are much lower in the liquid crystal, the plot is indistinguishable from a Debye semicircle. The perpendicular orientation, on the other hand, gives much lower absolute values of the permittivity and a very distorted complex plane plot. The frequency range of the loss peak is also quite different in the two cases. This suggests that different dipoles, or different configurations of dipoles, may be responsible for the two situations and that the interactions between these dipoles must be different.

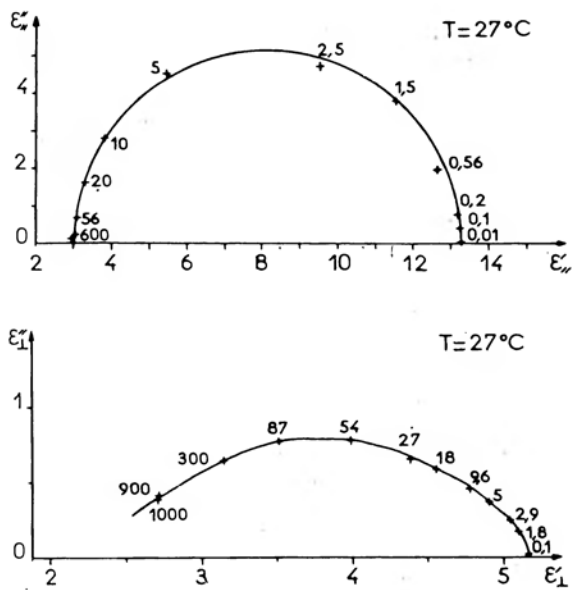


Figure 5.8 The complex plane diagram of the permittivity of the smectic phase of a liquid crystal, with frequencies given in MHz. The permittivity parallel to the orientation of the molecules gives a classical Debye plot, that normal to the molecules gives a broad and asymmetric plot.

From Druon and Wacrenier (1977).

5.3 BROADENED AND ASYMMETRIC DIPOLAR LOSS PEAKS

We have started our presentation of the experimental evidence in the frequency domain with a description of some of the examples resembling rather closely the Debye response, although we have noted that the ideal Debye response is hardly ever found in solids. Moving now towards the more generally observed types of behaviour we find a continuous variation of the breadth of the loss peaks, which may be expressed in terms of the slopes of the two branches in the logarithmic presentation. We shall not be concerned very much with the classification of these in terms of the Cole–Cole, Cole–Davidson and Havriliak–Negami representations, since these correspond to no more than a purely formal classification.

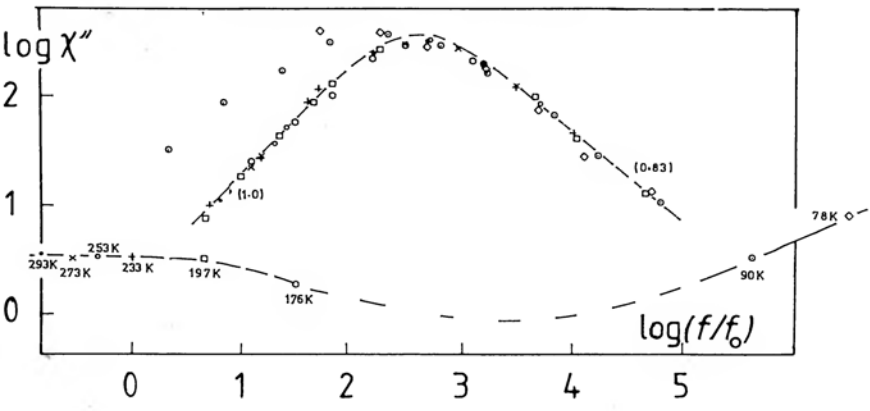


Figure 5.9 Dielectric loss data for tricyclohexyl carbinol, normalised for a wide range of temperatures. The data above 176 K were measured in the range 10^9–10^{11} Hz (Meakins 1956), those for 90 and 78 K were measured at frequencies below 10^6 Hz (Johari and Dannhauser 1969). The leading slope m of the high-temperature, high-frequency data is 1.0. A very definite change of slope occurs on going to lower frequencies and lower temperatures.

Plot by courtesy of R M Hill.

We begin our presentation with two examples of peaks in which the leading low-frequency edge has a slope of unity, within experimental error, since this case corresponds to an important limiting case in our theoretical interpretation. Figure 5.9 shows the normalised data for Tricyclohexyl carbinol, taken over a wide range of temperatures and in the relatively high frequency range 10^9–10^{11} Hz. The two lowest temperatures, which also correspond

to lower measuring frequencies, show a marked broadening of the leading edge, but the high-frequency, high-temperature behaviour has a slope of exactly unity. Our second example refers to butyl stearate and covers a range of almost ten decades of frequency, Figure 5.10, with two clearly discernible loss peaks. The high-frequency peak gives a sudden change of the leading slope to unity at around 10^9 Hz, while little can be said with any assurance regarding the trailing edge. The significance of these results will be discussed in Chapter 8.

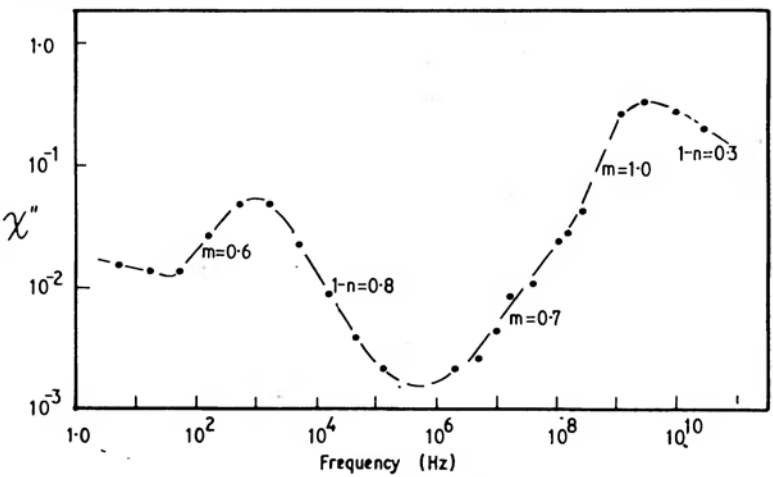


Figure 5.10 The dielectric loss spectrum for butyl stearate at 20°C for a very wide range of frequencies (Dryden 1957). Two distinct loss peaks are seen, both broader than Debye, but the slope $m = 1.0$ should be noted on the leading side of the high-frequency peak, at 10^9 Hz.

Plot by courtesy of R M Hill.

a) Polymeric materials

Many of our examples will refer to polymeric materials which are characterised by the presence of a phase transition at the so-called glass transition temperature, T_g. At temperatures above T_g the material behaves in a liquid-like manner, with a very high viscosity, but showing a measure of quasi-crystalline order. Below T_g, on the other hand, the system becomes "frozen in" and the behaviour is more solid-like. At T_g we find a sudden change in many properties, e.g. heat capacity and thermal expansion and the dielectric behaviour changes likewise, with the so-called α loss peaks above T_g being relatively narrower, those below T_g known as the β peaks being relatively much broader. The two types differ also in their temperature dependence, as will be discussed below.

The dielectric loss spectrum of the α peak of polyvinyl acetate in a range of temperatures is shown in Figure 5.11 where the real and imaginary components are plotted on a *linear* scale against the logarithm of frequency. Apart from the strong rise of loss at low

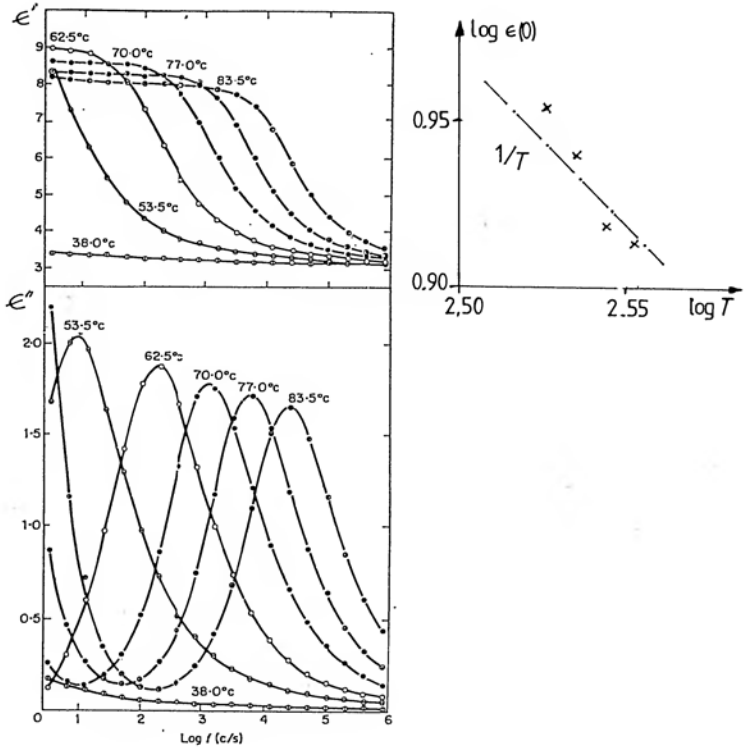


Figure 5.11 The frequency dependence of the real and imaginary components of the relative permittivity of polyvinyl acetate covering a range of temperatures in the α relaxation region above T_g. Note semilogarithmic plotting. From Ishida et al (1962). The diagram on the right gives the variation of the low-frequency limit of the real part with temperature in logarithmic representation, indicating that the data are consistent with a $1/T$ dependence. *cf.* Figure 4.7.

By permission of Dietrich Steinkopff Verlag, Darmstadt.

frequencies and high temperatures due to the onset of ionic dc conduction the loss peaks show a slightly broadened shape, a strong shift of the peak frequency ω_p with temperature and a relatively much weaker decrease of amplitude with increasing temperature.

The latter is also confirmed by a slight decrease of the low-frequency permittivity $\varepsilon(0)$, which is plotted as a function of T in a logarithmic representation to show that, within the limited accuracy of our data, the response is consistent with the $1/T$ behaviour expected of Debye-like, i.e. thermally disordered systems. This treatment is equivalent to that described in Figure 4.7, where loss peak amplitude was plotted.

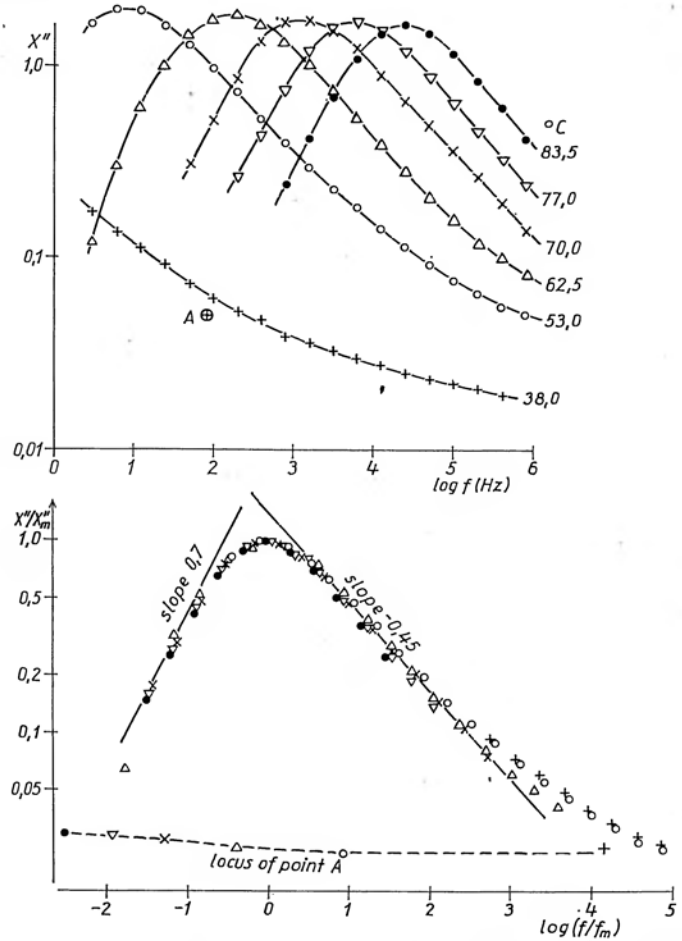


Figure 5.12 Loss data from Figure 5.11 plotted logarithmically, upper diagram, and normalised with respect to temperature, lower diagram. The locus of normalisation shows a very slight drop, consistent with the temperature dependence observed in Figure 5.11, but the logarithmic presentation reduces the sensitivity of this amplitude shift. The normalised plot suggests the onset of a second process at the high-frequency low-temperature end of the spectrum.

From Jonscher (1975). By permission of Dietrich Steinkopff Verlag, Darmstadt.

We now turn to the presentation of the same data in the logarithmic representation, which we prefer, shown in Figure 5.12 as loss data and the corresponding normalisation. In this presentation the temperature dependence of the amplitude cannot be assessed with sufficient precision, but the normalisation of the shape is much easier and so is the proper assessment of the nature of the frequency response away from the peak. We note that the peaks become slightly broader and a high-frequency "tail" develops as T decreases.

The corresponding data for polydian carbonate are shown in Figure 5.13 where we note that a very good normalisation is possible giving a single master curve, with an almost temperature-independent loss peak amplitude, witnessed by the horizontal locus of the displacement points.

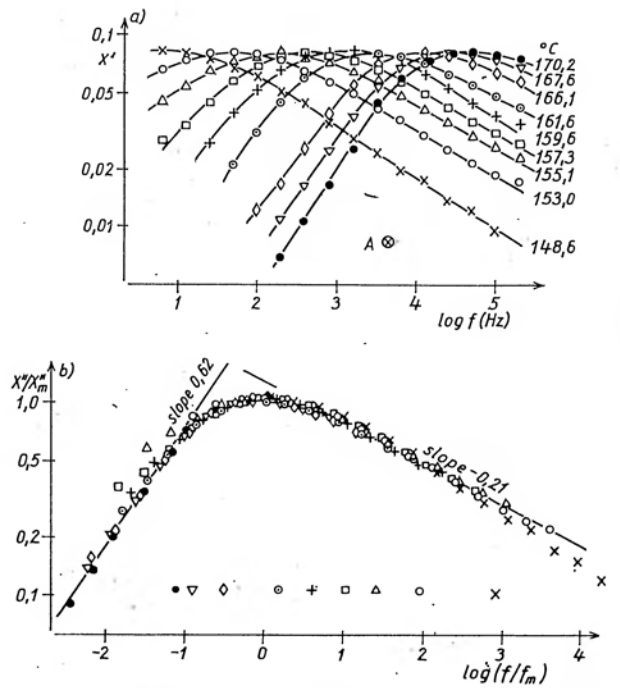


Figure 5.13 Dielectric loss peaks for the α relaxation in polydian carbonate for a range of temperatures, based on original data from Ishida and Matsuoka (1965). The normalisation shown in the lower diagram proves that all the points, except possibly those for the lowest temperature, fall on a single normalising curve.

Figure reproduced from Jonscher (1975). By permission of Dietrich Steinkopff Verlag, Darmstadt.

Not all peaks normalise to a single master cruve, showing a shape which is independent of temperature. Examples of varying shape are shown in Figure 5.14 and 5.15 in which both slopes below and above the loss peaks become smaller with decreasing temperature. Both these show a clear increase of amplitude with temperature.

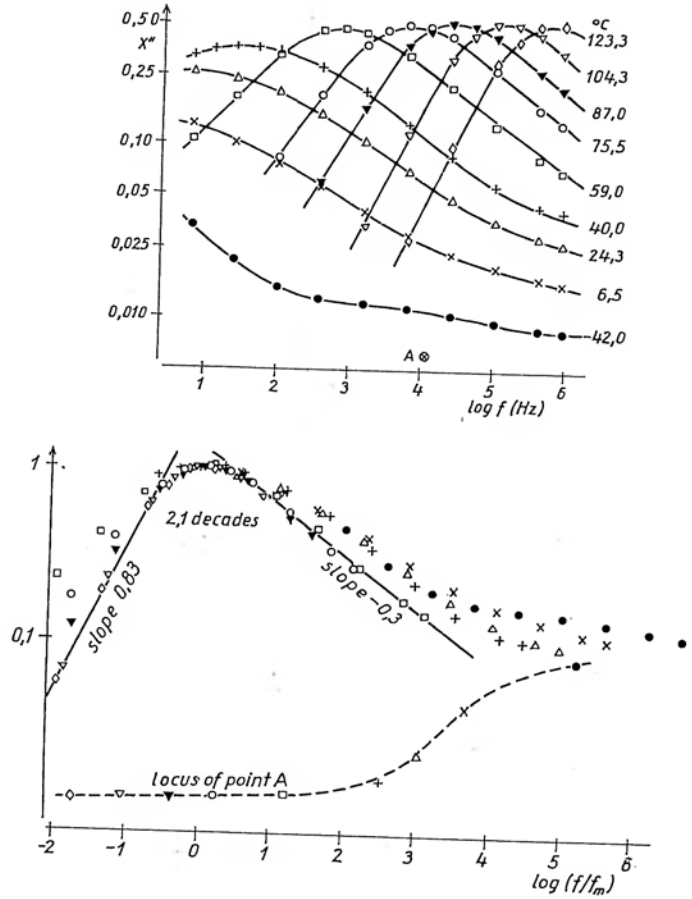


Figure 5.14 The dielectric loss peaks for the α relaxation of poly n-butyl methacrylate after Ishida (1969). The normalisation shows a very large secondary loss process at the lowest temperatures, with a corresponding strong decrease of amplitude, manifesting itself as a sudden rise of the locus of the representative points. The secondary process does not normalise with the same shift as the main peak, indicating that its activation energy is different.

From Jonscher (1975). By permission of Dietrich Steinkopff Verlag, Darmstadt.

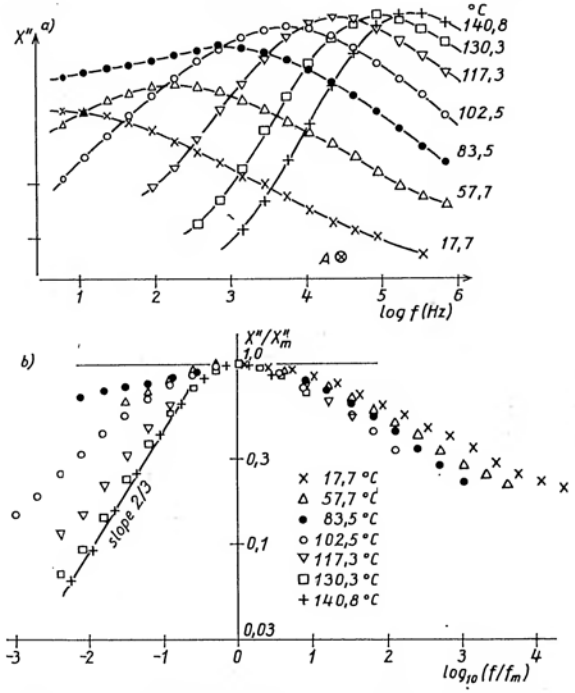


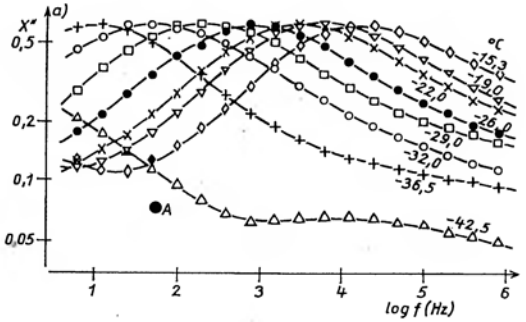
Figure 5.15 Dielectric loss data for polyethyl methacrylate taken from Ishida and Yamafuji (1961). The normalisation shows a very strong variation of the low-frequency slope – the exponent m – and a less pronounced variation of the high-frequency slope – exponent $1 - n$, with a correspondingly strong variation of the loss peak amplitude. The consistent increase of loss peak amplitude with rising temperature should be noted – a trend in complete contrast to that shown in Figure 5.11.

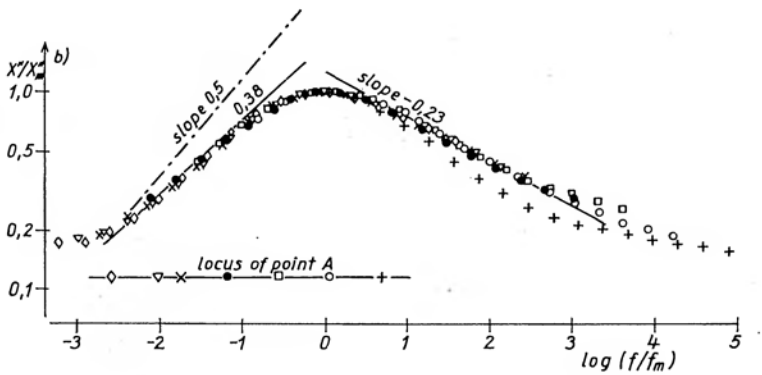
From Jonscher (1975). By permission of Dietrich Steinkopff Verlag, Darmstadt.

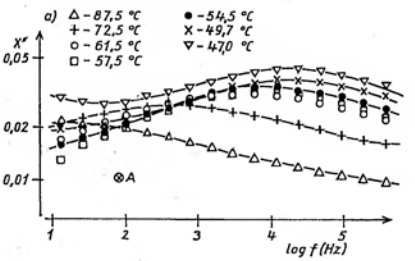
An example of the contrast between the α and β peaks in the same material is shown in Figure 5.16 relating to polychloroprene where the much broader nature of the β peaks is apparent. Again the amplitude of the α peak is almost independent of temperature, while the β peak is increasing.

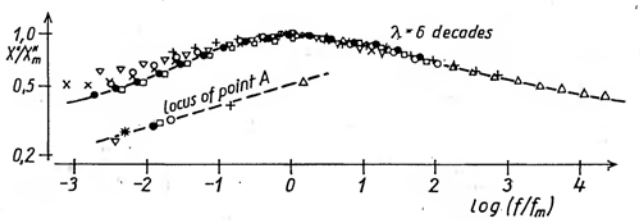
An example of an almost perfect normalisation to a single master curve is shown in Figure 3.37 relating to the low-temperature peak in polyethylene terephthalate, where the loss peak amplitude again shows a definite increase with temperature.

The following discussion of the broadened loss peaks will be facilitated by a reference forward to eqn (5.3) which defines the power–law relations below the loss peak frequency, $\chi''(\omega) \propto \omega^m$ and


a)
χ''
0,5
0,2
0,1
0,05
°C
-15,3
-19,0
-22,0
-26,0
-29,0
-32,0
-36,5
-42,5
A
1
2
3
4
5
6
log f (Hz)


b)
χ'/χ''m
1,0
0,5
0,2
0,1
slope 0,5
0,38
slope -0,23
locus of point A
-3
-2
-1
0
1
2
3
4
5
log (f/fm)


a) △ - 87,5 °C
+ - 72,5 °C
○ - 61,5 °C
□ - 57,5 °C
● - 54,5 °C
× - 49,7 °C
▽ - 47,0 °C
χ''
0,05
0,02
0,01
⊗A
1
2
3
4
5
log f (Hz)


1,0
χ'/χ''m
0,5
0,2
λ ≈ 6 decades
locus of point A
-3
-2
-1
0
1
2
3
4
log (f/fm)

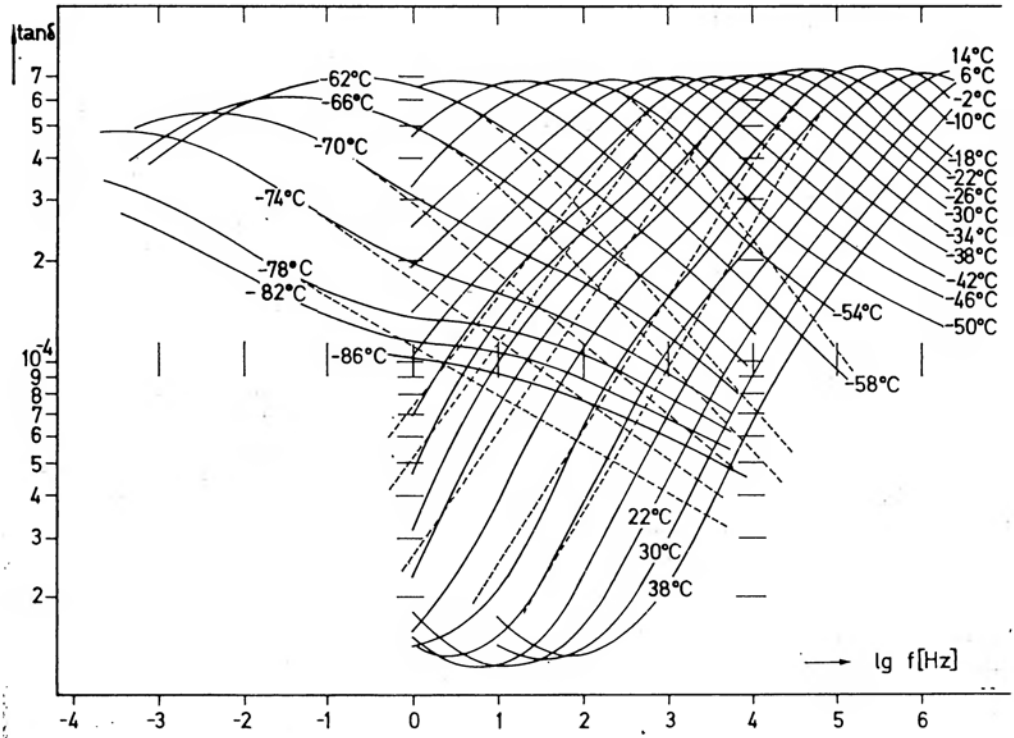


Figure 5.17 The dielectric loss spectrum of poly-isobutylene for a range of temperatures and a wide range of frequencies. Note that because the tan $\delta < 10^{-3}$, the dispersion of $\varepsilon'(\omega)$ is negligible and the tan δ plot is effectively the same as the $\varepsilon''(\omega)$ plot, except for a constant factor ε'.

From Stoll et al (1972). By permission of Dietrich Steinkopff Verlag, Darmstadt.

Figure 5.16 Dielectric loss data for polychloroprene in the high- and low-temperature regions, corresponding to α and β relaxations, respectively. Data from Matsuo et al (1965).

The high-temperature data in the upper diagrams normalise to an almost perfect master curve with a marked asymmetry. The lowest temperature data suggest the presence of a secondary loss peak descending from higher frequencies.

The lower-temperature data in the lower diagram reveal a very broad and asymmetric peak, whose width at half-height is 6 decades, whose amplitude increases strongly with rising temperature. The leading slope decreases markedly at the lowest temperatures.

From Jonscher (1975). By permission of Dietrich Steinkopff Verlag, Darmstadt.

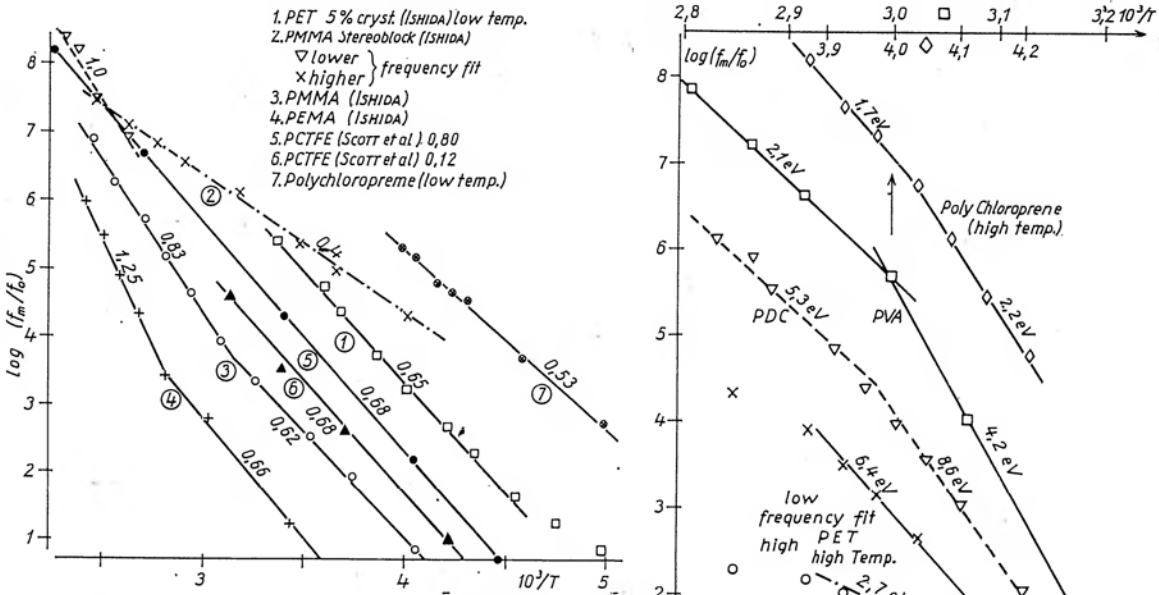


Figure 5.18 Activation energy diagrams, obtained by plotting the logarithm of the loss peak frequency f_m, or the amount of the horizontal shift of the normalisation locus, against the reciprocal temperature, $1/T$. The lower temperature peaks for a range of polymeric materials are shown in the diagram on the left, which also gives lower activation energies whose values are given on the diagram in electron-volts. The plots corresponding to higher-temperature peaks are shown in the right-hand diagram and the activation energies are also indicated. The plots are displaced horizontally and vertically for clarity, the diagrams on the right also have different scales of the abscissa.

From Jonscher (1975). By permission of Dietrich Steinkopff Verlag, Darmstadt.

above the loss peak frequency, $\chi''(\omega) \propto \omega^{n-1}$, where m and n are constants for any particular characteristic and they fall in the range $(0, 1)$.

We conclude our presentation of the dielectric response of dipolar polymeric materials at temperatures other than in the cryogenic range, which will be discussed separately, with an example of a rather complex behaviour of polyisobutylene, Figure 5.17. The amplitude of the main loss peak is almost independent of temperature while the shape narrows very significantly with increasing temperature, making normalisation rather uncertain. A further complication is the presence of a secondary loss mechanism which produces a peak with a different activation energy from the main one, so that the secondary peak is on the high-frequency slope of the main peak at low temperatures and appears on the low-frequency slope at high temperatures.

The temperature dependence of the α and β peak frequencies for a range of polymeric materials is shown in Figure 5.18 in the form of the usual activation plot of log ω_p against $1/T$, the slope of which

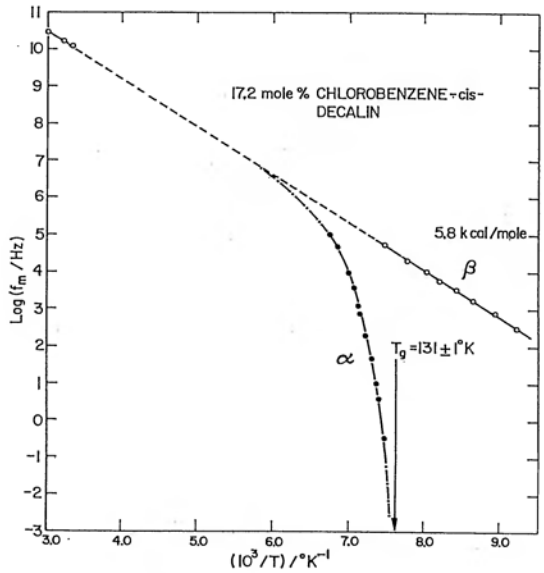


Figure 5.19 The frequency of the loss peak for the chlorobenzene–cis-decalin system showing a glass transition and demonstrating the different types of activation plots for the α and β peaks.

From Johari (1976).

may be interpreted as the activation energy. The β peaks give, on the whole, well defined straight line plots, indicating a single activation energy the value of which falls in the range 0.5–0.7 eV. In the case of the α peaks there is, in general, a tendency to curved plots, with the slope increasing with decreasing temperature. The values of the activation energy obtained by drawing the tangent to the plot appear unreasonably high in some cases and an alternative viewpoint makes physically much more sense. Figure 5.19 shows the complete plot of the loss peak frequency against temperature for Chlorobenzene–cis-Decalin mixture which is a glass-forming substance with $T_g = 131$ K. It is clear from this plot that the β peaks follow the classical Arrhenius relation, eqn (4.29) with a well defined activation energy, but the critical temperature T_g implies a catastrophe in which the frequency ω_p tends to zero, giving an artificially high slope which cannot, however, be interpreted as an activation energy in the conventional sense. It is interesting to note that at very high temperatures the response falls on the projection of the β relaxation from lower temperatures.

It may be intriguing to note the relatively narrow range of activation energies of the β processes, between 0.5 and 0.8 eV, approximately. An explanation of this was proposed by Hill (private communication) in terms of the physical processes which determine the conditions of measurements. Taking the pre-exponential factor ν in the expression defining the loss peak frequency

$$\omega_p = \nu \exp(-W/kT) \tag{5.2}$$

as 10^{12} Hz, corresponding to typical lattice vibration frequencies, we obtain the numerical expression for the logarithm

$$\log \omega_p = 12 - 5.04\, W(10^3/T)$$

where W is the activation energy expressed in electron-volts. Taking a typical frequency "window" available to most experiments between 10 Hz and 1 MHz, and a typical temperature range between −50 and 200°C we obtain the shaded area in Figure 5.20 as the experimentally accessible region. With these data, we see that loss peaks with energies in excess of 1 eV will not appear except at much higher temperatures or lower frequencies, while those with energies less than 0.4 eV will not be reached until much lower temperatures and higher frequencies. To that extent we may say, therefore, that the observed values of activation energies are self-selecting through the nature of the typical experimental conditions.

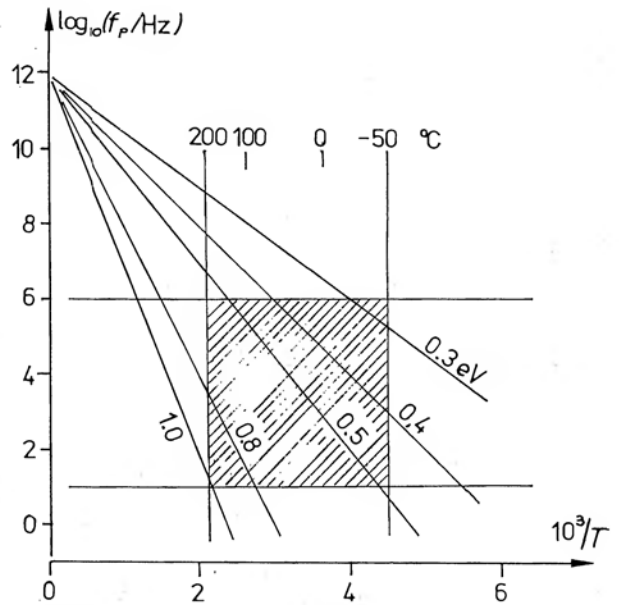


Figure 5.20 An illustration of the reasons for the apparently narrow range of activation energies found in the low-temperature peaks of dipolar materials, including polymers. If the activation energy can be expressed by a relation of the form of eqn (5.2), with ν corresponding to a typical lattice frequency of 10^{12} Hz, and if the available temperature and frequency "window" is as indicated by the shaded region, the observable activation energies fall in the range 0.3–1.0 eV.

Courtesy R M Hill.

b) Other dipolar systems

There are very many dipolar systems, with either molecular dipoles or with hopping ionic species which give the effect of dipolar behaviour by being confined to hopping between nearest neighbour positions around a fixed site. These materials are often more thermally stable than polymers, in that they do not show the phenomena of glass transition, so that their dielectric loss peaks are more easily normalised into a single master curve, sometimes covering a wide range of temperatures.

Figure 5.21 shows a compilation of loss peaks of a wide range of materials, the broader lower-temperature ones on the left, the narrower higher-temperature ones on the right. Some of these are glass-forming systems undergoing the glass transition process similarly as in polymers, e.g. 3-methyl-3-heptanol, others are mixtures of organic liquids and included in the collection is also the low-temperature response of silica glass–Suprasil in the range 4.2 K

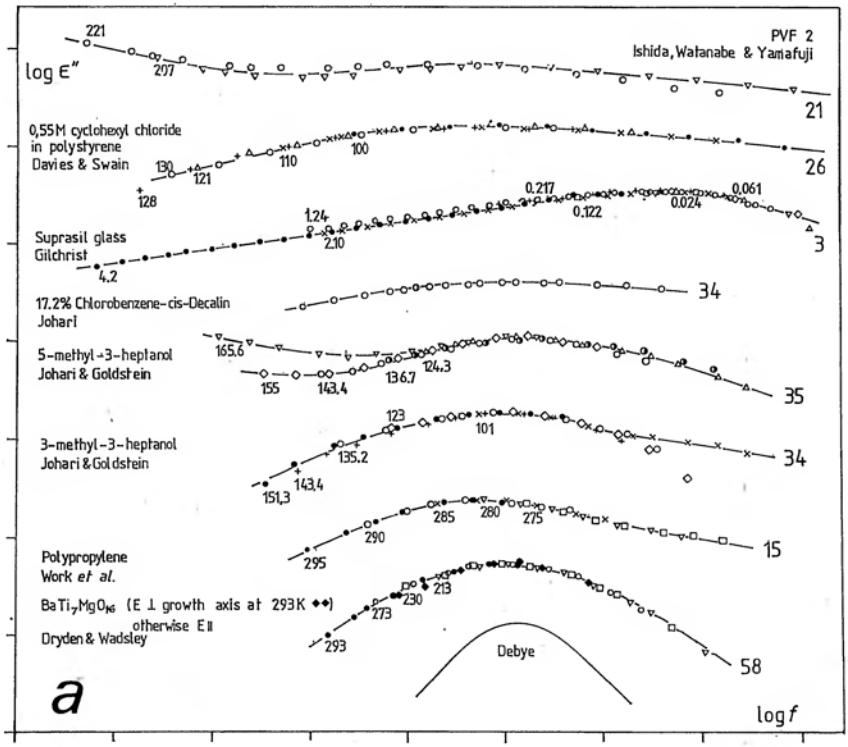
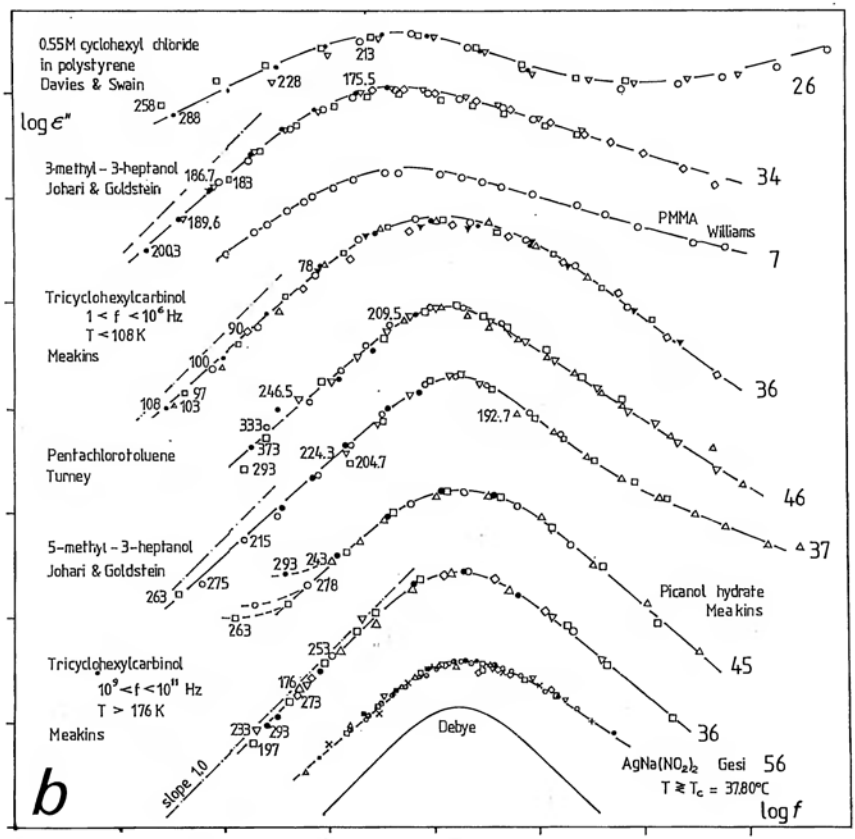


Figure 5.21 Two compilations of dielectric loss data for a range of dipolar materials, plotted logarithmically and normalised for the temperatures indicated as numbers (in K) at the corresponding lowest frequency points. The individual plots are displaced vertically and horizontally for clarity. The set in a) gives the broader, lower-temperature peaks, that in b) the narrower higher temperature peaks. The chain-dotted lines indicate the slopes corresponding to $m = 1$. The shapes of ideal Debye responses are shown. The numbers on the right of each diagram refer to literature data in Table 5.1.

The logarithmic plots and normalisations by courtesy of R M Hill.

down to 24 mK. The principal object of showing these loss peaks is to emphasize their power–law character on either side of the loss peak, manifesting itself through the linear regions, where there is no overlap of more than one mechanism. The very broad peaks in a) tend to an almost frequency-independent behaviour, while in other cases there is evidence of either a second low-frequency process or another higher-frequency process taking over. The low-temperature Suprasil response illustrates the very wide range of power–law below the loss peak – this type of result can only be



obtained at very low temperatures, since at higher temperatures the onset of direct current conduction masks the low-frequency dielectric behaviour. A wide range of frequencies can be covered easily if the exponent m becomes small because of the small variation of loss.

It is noteworthy that none of the responses shown in these diagrams give a low-frequency slope steeper than $+1.0$ and a high-frequency slope steeper than -1.0. The significance of these values will become clear in the context of the theoretical discussion of Chapter 8.

A special class of dielectric systems are ferroelectrics, whose special properties include the persistence of polarisation in the absence of external electric field and the existence of a critical temperature, the Curie temperature T_c, above which the ferroelectric ordered

phase goes over into the disordered paraelectric phase. We saw the "giant" dispersions in these materials in Figures 5.1 and 5.6 where the behaviour approximated closely to the Debye response. We now show another example of ferroelectric behaviour in the vicinity of T_c represented as a normalised plot of the real and imaginary components of $\chi(\omega)$ for $AgNa(NO_2)_2$ in Figure 5.22. The remarkable feature of this response, as also in Figure 5.1, is the shape of the locus of the displacement point which follows the not hitherto seen, let alone understood trajectory, consisting of two almost parallel lines which are also parallel to the high-frequency slope of the loss peak. Starting below the Curie temperature, where the material is ferroelectric, the amplitude increases and the frequency decreases as we approach T_c and the opposite trend sets in above T_c. Many more examples of this behaviour have been found and the full significance of these results was only recently brought out by the work of Dissado and Hill (1979, 1980) and will be discussed in Chapter 8.

c) Dipolar response at cryogenic temperatures

One of the most striking features of the dielectric response distinguishing it sharply from most forms of steady state response, such as transport of charges, is the persistence of dielectric loss down to the lowest temperatures in the milli-Kelvin range. This behaviour is somewhat unexpected and it might not have been noticed, let alone studied in detail, were it not for the technological interest aroused some years ago in the development of cryogenic electrical machinery and cables. At the same time, the full theoretical significance of these phenomena only became clear in the context of the new "universal" theory which will be discussed in Chapter 8.

The behaviour of many dipolar materials at cryogenic temperatures resembles that at higher temperatures, except that the relationship between the loss peak frequency and amplitude and the temperature does not follow the same pattern (Phillips 1970, Carson 1973). Yet other materials show a completely frequency-independent loss which will be discussed more conveniently under Section 5.7.

Figure 5.23 shows the data for two samples of slightly oxidised polyethylene, the first in the range 1.24–17.1 K which shows a simple normalisation, the second in the lower range 0.029–4.2 K in which there is clear evidence of a critical temperature around 0.2 K. In the first case the peak amplitude increases steadily with decreasing temperature while the peak frequency decreases. In the

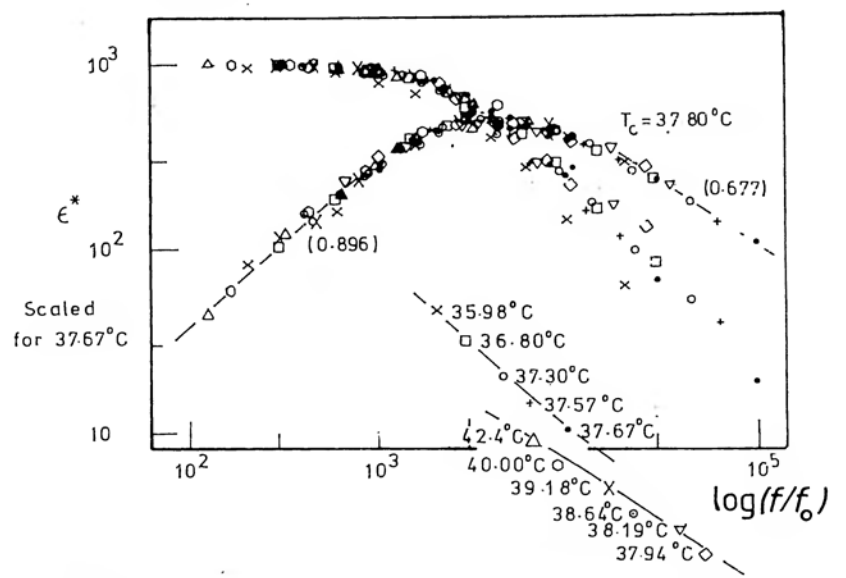


Figure 5.22 Normalisation of the real and imaginary components of the permittivity of $AgNa(NO_2)_2$ below and above its Curie temperature, showing that the locus of the normalisation points follows two almost parallel lines with slopes close to the high-frequency slope of the loss curve.

Data from Gesi (1972), normalisation and plotting by courtesy of R M Hill.

second case the behaviour is complicated by the existence of the critical temperature. Both peaks are slightly broader than Debye, their width at half-height is $\lambda = 1.5$–1.7 times the Debye width of 1.144 decades. They are also slightly asymmetric.

Figure 5.24 shows a remarkable example of the response at 4.2 K for a "pure" sample without antioxidant, which gives a completely "flat" loss at a very low level, and also the response of two samples with different types of antioxidant which both give very distinct loss peaks of almost symmetric nature and approximately twice the Debye width. The presence of antioxidant in this context may be associated with the introduction of some dipolar impurities into the normally non-polar polyethylene and it may also give rise to some cross-linking of polymer chains.

Very interesting features of low-temperature response appear in Figure 5.25 relating to 2,4,6, tri-tert-butyl phenol (Isnard and Gilchrist 1980) which are normalised and scaled at 51.7 K by R M Hill. The high-frequency slope $1 - n$ is almost independent of temperature, but the low frequency slope m shows a sharp dip, see

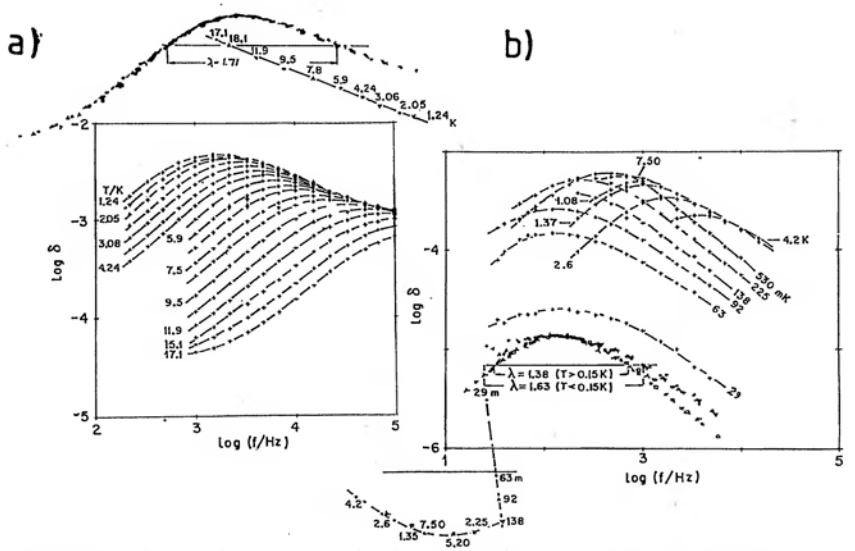


Figure 5.23 Dielectric loss peaks in the cryogenic range of temperatures for two slightly oxidised samples of polyethylene, based on numerical data by Gilchrist (1977) for sample a) and Gilchrist (1979) for sample b). Normalisation is given with the corresponding loci of the displacement points. Sample a) shows a steady trend in the entire temperature range 1.24–17.1 K, with the amplitude of the peak increasing and the frequency decreasing with decreasing temperature. A single mechanism appears to be present. Sample b) in the temperature range 4.8 − 0.029 K shows a clear break in the trend at around 150 mK, with the loss peak frequency becoming constant at the lowest temperatures and the amplitude falling rapidly.

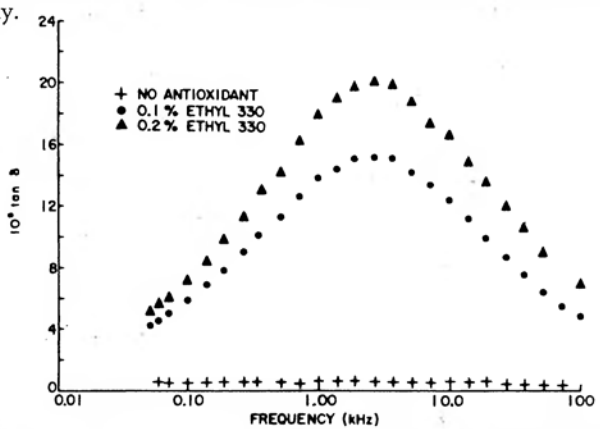


Figure 5.24 The dielectric loss of polyethylene with and without antioxidant at 4.2 K. The "pure" sample has a flat loss in the frequency range investigated, the two samples with antioxidant give approximately symmetric peaks with width $\lambda = 2.1$ decades. For the present purpose, the effect of the antioxidant is to introduce additional dipoles into the material.

From Thomas and King (1975).

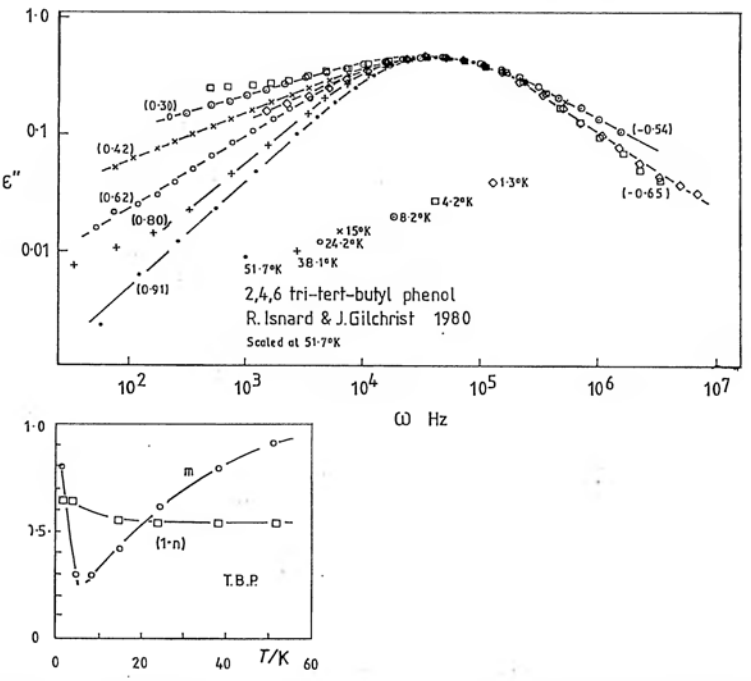


Figure 5.25 Low-temperature data for 2,4,6 tri-ter-butyl phenol, with the slopes $1 - n$ and m plotted in the inset as functions of temperature.

Data from Isnard and Gilchrist (1980). Plotting by R M Hill.

inset, indicating a discontinuity of behaviour. The dielectric response reveals itself as a very sensitive probe into structural details of the material under conditions where no other measuring technique might be able to reveal any significant changes.

Our last example of the low-temperature dipolar responses concerns the class of materials known as clathrates, in which molecular "cages" consisting of H_2O molecules arranged in pentagonal dodecahedra and similar structures, contain individual "guest" molecules which may be polar and which are confined in an almost spherically symmetric potential. In principle, one would therefore expect these dipoles to be almost free to rotate and to behave like ideal Debye systems, except for the interactions between the neighbouring dipoles. In fact, the behaviour departs very strongly from the Debye pattern. Figure 5.26 shows three examples of response at low temperatures, drawn in the form of complex ε plots which indicate a region at high frequencies with the universal frequency dependence, resulting in a straight line tail to the complex plane plot.

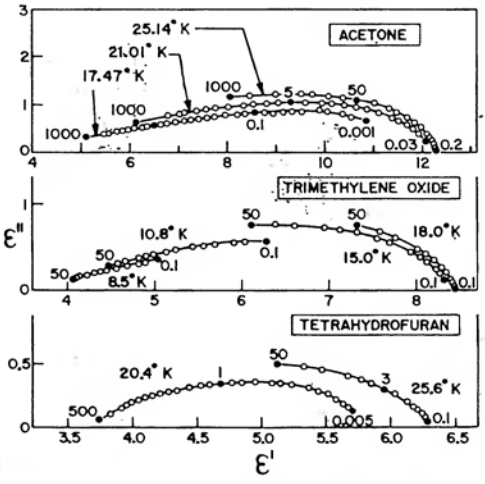


Figure 5.26 Low-temperature complex permittivity plots for clathrates containing different guest molecules as indicated. Temperatures in K, frequencies in kHz. The strongly flattened and asymmetric shape of these plots should be noted. From Gough et al (1973).

A very recent survey of the low-temperature properties of dielectrics may be found in Phillips (1981).

d) Characterisation of dielectric loss peaks

Our survey has shown that the prevailing form of frequency dependence of the dielectric loss peaks may be represented by the empirical law combining two power-laws, respectively below and above the peak frequency ω_p (Jonscher 1975a)

$$\chi''(\omega) \propto \frac{1}{(\omega/\omega_p)^{-m} + (\omega/\omega_p)^{1-n}} \tag{5.3}$$

in which the exponents m and $1 - n$ fall in the range (0, 1) and the peak frequency is generally temperature dependent with either a simple activation law, eqn (5.2) or with some more complicated relationship. The applicability of this empirical relation is limited at low frequencies by the onset of direct current conduction, while at the higher frequencies perturbing processes are often visible arising from the presence of some series resistance in the measuring system, or from the overlap of some other loss processes.

The question of overlap of different dielectric processes raises some very important questions of principle. It is often argued that since the dielectric response of many materials is complicated by the

presence of multiple peaks and other features, there is no point in trying to find a "universal law" which would represent the response in a more satisfactory manner than the simple and arbitrary superposition of Debye mechanisms, to be discussed in Chapter 7. We have to leave it to the reader's judgement whether the empirical law (5.2) does or does not represent a useful approximation to the behaviour of most materials and whether it may be used as a "basic element" for the construction of more complicated spectra by a superposition of two, or possibly three separate such mechanisms.

Our point is that this law is capable of representing the observed behaviour of many dielectrics in a remarkably wide range of frequencies – spanning as many as eight to ten decades in some cases, allowing for the normalisation process. The strong presumption arises, therefore, that this power law does accurately reflect a very fundamental process and its physical implications should be further examined since, if it could be shown to correspond to a well defined physical mechanism, it would represent a significant step forward to the meaningful interpretation of the processes of dielectric polarisation. This does not, in any sense, preclude the possibility that in any given material there may be more than one single "universal" process in operation – indeed, it is most surprising that so many different materials appear to be dominated by *one single* "universal" law (5.3) in the entire available frequency range.

In order to advance our understanding of the physical implications of the universal law it is instructive to study the incidence of the values of the exponents m and n throughout the spectrum of materials. Figure 5.27 shows a plot of these exponents for 100 dipolar materials (Hill 1981a) using the two exponents as coordinates in a square array. Where the exponents vary with temperature, the characteristics of the most well-defined spectrum were taken.

The top right-hand corner with $m = 1$ and $1 - n = 1$ corresponds to the ideal Debye characteristic. The diagonal corresponds to symmetric peaks, for example those given by the empirical Cole–Cole expression, eqn (3.37) and the Fuoss-Kirkwood (1941) expression:

$$\varepsilon'' = \frac{\varepsilon''_m}{(\omega\tau_0)^{\alpha} + (\omega\tau_0)^{-\alpha}} = \varepsilon''_m \operatorname{sech}(\alpha \ln \omega\tau_0) \tag{5.4}$$

which corresponds to our universal law with equal exponents.

TABLE 5.1

The characteristic exponents m and $1 - n$ shown in Figure 5.27, together with experimental data and reference numbers relating to the list at the end of the Table.

Number in Fig.	Material	m	$(1 - n)$	Ref. No.	Comments
1	Polyacetaldehyde	0.64	0.54	1	Amorphous form
2	Polyacronitrile	0.18	0.33	2	
3	Polyethylene	0.82	0.35	3	$1.3 < T < 17.1$ K Sample PE1
4	Polyethylene	0.66	0.53	3	$29 < T < 4200$ mK Sample PE 10Ma
5	Polyethylene	0.68	0.63	3	$20 < T < 4200$ mK Sample PE 14
6	Polyethylene terephthalate	0.77	0.28	4	α-peak
7	Polyethylene terephthalate	0.42	0.08	4	β-peak
8	Methyl methacrylate	0.25	0.33	5	
9	Polymethyl acrylate	0.74	0.25	6	
10	Polymethyl methacrylate	0.66	0.28	7	$100 < P < 2950$ atm, $283 < T < 363$ K
11	Polymethyl methacrylate	0.45	0.17	8	
12	Poly-*n*-butyl methacrylate	0.71	0.29	9	
13	Polychloroprene	0.38	0.23	10	α-peak
14	Polychloroprene	0.17	0.16	10	β-peak
15	Polydian carbonate	0.62	0.21	11	α-peak
16	Polyethyl methacrylate	0.49	0.18	12	
17	Poly-*n*-hexyl methacrylate	0.90	0.35	13	
18	Poly-*n*-octyl methacrylate	0.84	0.38	13	
19	Poly-cyclohexyl methacrylate	0.068	0.25	14	
20	Poly-2-chlorocyclohexyl methacrylate	0.32	0.44	14	
21	Poly-4-chlorocyclohexyl methacrylate	0.23	0.31	14	
22	Poly-4-chlorocyclohexyl acrylate	0.42	0.25	14	
23	Polypropylene	0.37	0.24	15	
24	Polypropylene carbonate	0.90	0.45	16	α-peak
25	Polypropylene carbonate	0.85	0.71	16	β-peak
26	Polypropylene oxide	0.39	0.31	17	

27	Polyvinyl acetate	0.83	0.43	6	
28	Polyvinyl acetate	0.76	0.38	18	In solution in Galva 15
29	Polyvinyl acetate	0.56	0.42	18	In solution in Galva 60
30	Polyvinyl acetate	0.50	0.29	19	
31	Polyvinyl fluoride	0.28	0.11	20	
32	Polyvinylidene fluoride	0.07	0.49	21	α-peak
33	Polyvinylidene fluoride	0.53	0.13	21	β-peak
34	Polyvinylidene fluoride	0.05	0.12	21	γ-peak
35	Polyvinyl formal	0.25	0.24	19	
36	Polyvinyl chloroacetate	0.57	0.40	6	
37	Polyvinyl chloroacetate	0.60	0.43	18	
38	Polyvinyl chlorodiphenyl	0.24	0.22	25	
39	Poly-γ-benzyl-L-glutamate	0.42	0.81	23	Solid form
40	Poly-γ-benzyl-L-glutamate	0.55	0.51	24	In solution in E.D.C.
41	Poly-γ-benzyl-L-glutamate	0.79	0.55	24	In solution in dioxane
42	Butyl stearate	0.56	0.77	25	$\omega_p < 10^9$ Hz
43	Butyl stearate	1.0	0.33	25	$\omega_p > 10^9$ Hz
44	Cyclohexyl chloride in polystyrene	0.47	0.30	26	α-peak
45	Cyclohexyl chloride in polystyrene	0.21	0.08	26	β-peak
46	Phthalic anhydride	0.87	0.45	27	3.72 wt%
47	Nylon 610	0.60	0.57	28	$T = 483$ K ($T_m = 497.8$ K)
48	Nylon 610	0.30	0.51	28	$T = 413$ K
49	Nylon 610	0.18	0.41	28	$T = 373$ K
50	P-methoxyphenylazoxy-p-butylbenzene	0.75	0.77	29	Nematic phase
51	P-methoxyphenylazoxy-p-butylbenzene	0.60	0.81	29	Isotropic phase
52	*N*-heptyl-cyanobiphenyl	0.97	0.68	30	Both isotropic and nematic phases
53	Acetic acid	1.0	0.35	31	$\omega_p > 10^9$ Hz
54	Bromobenzene in decalin	0.71	0.28	32	
55	Chlorobenzene-pyridene	0.93	0.35	33	α-peak 43.4 conc.
56	Chlorobenzene-pyridene	0.15	0.10	33	β-peak 43.4 conc.
57	Chlorobenzene-cis-decalin	0.89	0.19	34	α-peak
59	Chlorobenzene-cis-decalin	0.06	0.15	34	β-peak
59	2-Methyl-3-heptanol	0.78	0.27	35	$P = 2.5$ kb
60	2-Methyl-3-heptanol	1.0	0.38	35	$P = 4.4$ kb

TABLE 5.1 continued

Number in Fig.	Material	m	$(1-n)$	Ref. No.	Comments
61	3-Methyl-3-heptanol	0.86	0.33	34	α-peak
62	3-Methyl-3-heptanol	0.49	0.34	34	β-peak
63	5-Methyl-3-heptanol	0.90	0.67	34	α-peak
64	5-Methyl-3-heptanol	0.23	0.25	34	β-peak
65	5-Methyl-3-heptanol	0.65	0.68	35	$P = 3.55$ kb
66	5-Methyl-3-heptanol	0.98	0.57	35	$P = 1.57$ kb
67	Tricyclohexyl carbinol	1.0	0.83	36	$10^9 < \omega < 10^{11}$ Hz; $T > 176$ K
68	Tricyclohexyl carbinol	1.0	0.74	37	$\omega < 10^6$; $T < 108$ K
69	Neo-hexanol	0.72	0.95	37	β-peak
70	Glycerol	0.55	0.46	38	$P = 3.1$ kb
71	Glycerol	0.50	0.69	38	$P = 4.4$ kb
72	Menthol	0.99	0.22	39	
73	2:4:6-tri-t-butylphenol	1.0	0.66	36	$10^9 < \omega < 10^{10}$ Hz; $176 < T < 293$ K
74	*n*-docosyl bromide	0.47	0.63	40	
75	Picric acid	0.85	0.79	41	
76	Polyethylene adipate	0.06	0.09	42	
77	M.B.B.A.	0.85	0.73	43	
78	Methyl stearate	0.57	0.47	44	$P < 2$ kb; $213 < T < 243$ K
79	Polyvinyl acetyl	0.45	0.24	19	$343 < T < 393$ K
80	Pinacol hydrate	0.89	0.86	45	
81	Pentachlorotoluene	0.87	0.64	46	
82	Trichloroethane–cyclohexane	0.59	0.53	47	Conc. range 1:1 to 0.8:1
83	Trichloroethane–cyclohexane	0.83	0.74	47	Conc. range 0.4:1 to 0.2:1
84	Suprasil glass	0.15	0.31	3	$20 < T < 4200$ mK
85	Impurity cations in quartz	0.69	0.85	48	
86	Silicon monoxide/platinum cermet film	0.54	0.71	49	
87	Silicon p–n junction	0.83	0.78	50	
88	Silicon p n junction	0.16	0.60	50	

89	Anodised tantalum	0.42	0.36	51	
90	Anodised aluminium	0.58	0.92	51	
91	Rochelle salt	1.0	0.96	52	
92	KD_2PO_4	1.0	0.93	53	
93	Triglycene sulphate	0.97	0.99	54	
94	M.A.S.D.	0.81	0.75	56	$138 < T < 200$ K
95	$AgNa(NO_2)_2$	0.90	0.68	56	
96	$AgNaNO_2$	1.0	0.83	57	
97	$BaTi_7MgO_{16}$	0.68	0.55	58	
98	H_2O	0.96(3)	0.96(8)	59	293 K selected published values
99	H_2O	0.78	0.76	60	181 K
100	D_2O	0.94	0.95	61	$T < 262.3$ K

LITERATURE REFERENCES TO FIGURE 5.27 AND TABLE 5.1

1 G. WILLIAMS, *Trans. Faraday Soc.* **59** (1963) 1397.
2 Y. ISHIDA, O. AMANO and M. TAKAYANAGI, *Kolloid-Z.u.Z.f.P.* **172** (1960) 129.
3 J. le G. GILCHRIST, private communication (1977).
4 Y. ISHIDA, M. MATSUO and K. YAMAFUJI *Kolloid-Z.u.Z.f.P.* **180** (1962) 108.
5 Y. KAMAWURA, S. NAGAI, J. HIROSE and Y. WADA, *J. Polymer Sci.* **A-2** (1969) 1559.
6 J. D. FERRY, M. L. WILLIAMS and E. R. FITZGERALD, *J. Phys. Chem. Soc.* **59** (1955) 403.
7 G. WILLIAMS, *Trans, Faraday Soc.* **60** (1964) 1556.
8 Y. ISHIDA, *J. Polymer Sci.* **A-2**(7) (1969) 1835.
9 S. STRELLA and R. ZAND, *J. Polymer Sci* **25** (1957) 105.
10 M. MATSUO, Y. ISHIDA, K. YAMAFUJI, M. TAKAJANAGI and F. IRIE, *Kolloid-Z.u.Z.f.P.* **201** (1965) 89.
11 Y. ISHIDA and S. MATSUOKA, *Amer. Chem. Soc. Polymer Preprints* **6** (1965) 795.
12 S. STRELLA and R. ZAND, *J. Polymer Sci.* **25** (1957) 97.
13 *Idem, J. Polymer Sci.* **31** (1958) 45.
14 J. HEIJBOER, T.N.O. Central Laboratory Publication Number 435, Holland, 1972.
15 R. N. WORK, R. D. McCAMMON and R. G. SABA, *J. Chem. Phys.* **41** (1964) 2950.
16 R. PAYNE and I. E. THEODOROU, *J. Phys. Chem.* **76** (1972) 2892.
17 S. YANO, R. R. RAHALKAR, S. P. HUNTER, C. H. WANG and R. H. BOYD, *J. Polymer Sci.* **14** (1976) 1877.
18 D. J. MEAD and R. M. FUOSS, *J. Amer. Chem. Soc.* **63** (1941) 2382.
19 B. L. FUNT and T. H. SUTHERLAND, *Can. J. Chem.* **30** (1952) 940.
20 Y. ISHIDA and K. YAMAFUJI, *Kolloid-Z.u.Z.f.P* **200** (1964) 50.
21 Y. ISHIDA, M. WATANABE and K. YAMAFUJI, *ibid.* **200** (1964) 48.
22 R. M. FUOSS, *J. Amer. Chem. Soc.* **63** (1941) 378.
23 K. HIKICHI, K. SAITO, M. KANEKO and J. FURUICHI, *J. Phys. Soc. Jap.* **19** (1964) 527.
24 L. A. DISSADO and R. M. HILL, unpublished work (1980).
25 J. S. DRYDEN, *J. Chem. Phys.* **26** (1957) 604.
26 M. DAVIES and J. SWAIN, *Trans. Faraday Soc.* **67** (1971) 1637.
27 G. WILLIAMS and P. J. HAINS, *Faraday Symp. Chem. Soc.* **6** (1972) 14.
28 R. H. BOYD and C. H. PORTER, *J. Polymer Sci.* **A-2(10)** (1972) 647.
29 J. P. PARNIEUX, A. CHAPOTON and E. CONSTANT, *J. de Phys.* **36** (1975) 1143.
30 A. BAKA, P. G. OWEN and A. H. PRICE, *Mol Cryst. Liq. Cryst.* **51** (1979) 273.
31 E. CONSTANT and A. LEBRUN, *J. de Chemie Physique* **61** (1964) 166.
32 G. P. JOHARI and C. P. SMYTH, *J. Chem. Phys.* **56** (1972) 4411.
33 G. P. JOHARI and M. GOLDSTEIN, *ibid.* **53** (1970) 2372.
34 *Idem. ibid.*, **55** (1971) 4245.
35 G. P. JOHARI and W. DANNHAUSER, *ibid.* **50** (1969) 1862.
36 R. J. MEAKINS, *Trans. Faraday Soc.* **52** (1956) 320.
37 G. P. JOHARI, *Ann. N. Y. Acad. Sci.* **279** (1976) 117.
38 G. P. JOHARI and E. WHALLEY, *Faraday Symp. Chem. Soc.* **6** (1972) 23.
39 H. CACHET and J. C. LESTRADE, *Acad. Sci. Paris* **259** (1964) 541.
40 J. S. DRYDEN and S. DASGUPTA, *Trans. Faraday Soc.* **51** (1955) 1661.
41 R. J. MEAKINS, *ibid.* **51** (1955) 371.
42 Y. ISHIDA, K. YAMAFUJI and K. SHIMADU, *Kolloid-Z.u.Z.f.P.* **201** (1965) 49.
43 V. K. AGARWAL and V. P. ARORA, *J. Chem. Phys.* **66** (1977) 2817.
44 M. BROADHURST, in "Dielectric Properties of Polymers" edited by F. E. Krans (Plenum Press, New York, 1972) p. 129.
45 R. J. MEAKINS, 1961, *Prog. in Dielectrics* **3** (1961) 170.
46 A. TURNEY, *Proc. I.E.E.* **100** (1953) 46.
47 R. FINSEY and R. van LOON, *J. Chem. Phys.* **63** (1975) 4831.
48 E. K. SNOW and P. GIBBS, *J. Appl. Phys.* **35** (1964) 2368.
49 D. MYO, Ph.D. Thesis, Imperial College, London, 1979.
50 A. K. JONSCHER, private communication (1980).

51 F. MECA, PhD. 1981. Thesis, Chelsea College, London.
52 F. SANDY and R. V. JONES, *Phys. Rev.* **168** (1968) 1181.
53 R. M. HILL and S. K. ICHIKI, *ibid.* **130** (1963) 150.
54 *Idem, ibid.* **128** (1952) 1140.
55 Y. MAKITA and M. SUMITA, *J. Phys. Soc. Jap.* **31** (1971) 792.
56 K. GESI, *Jap. J. Appl. Phys.* **11** (1972) 1745.
57 I. HATTI, *J. Phys. Soc. Jap.* **24** (1968) 1043.
58 J. S. DRYDEN and A. D. WADSLEY, *Trans. Faraday Soc.* **54** (1958) 1574.
59 P. R. MASON, J. B. HASTED and L. MOORE, *Adv. Mol. Rel. Proc.* **6** (1974) 217.
60 S. R. GOUGH and D. W. DAVIDSON, *J. Chem. Phys.* **52** (1970) 5442.
61 G. P. JOHARI and S. J. JONES, *Proc. Roy. Soc. A.* **349** (1976) 467.

The top side gives points compatible with the empirical Cole–Davidson formula (3.38) and also with the frequency-domain transform of the Williams-Watts (1970) law:

$$f(t) \propto t^{\beta-1} \exp(-t/\tau) \tag{5.5}$$

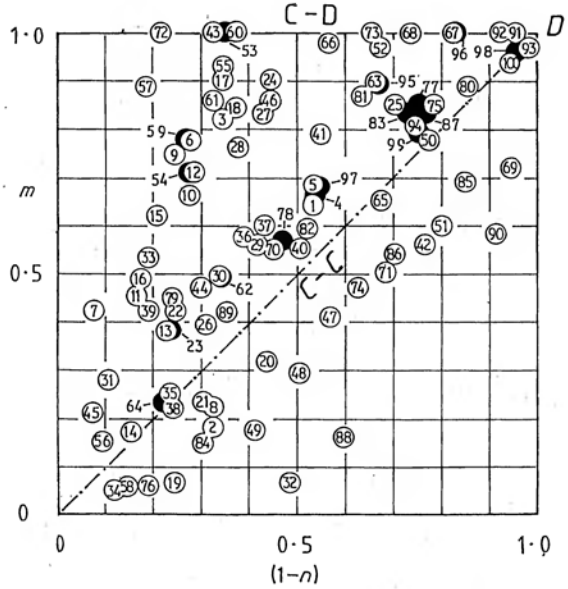


Figure 5.27 A plot of the characteristic exponents m and $1-n$ defined by eqn (5.3), for one hundred dipolar materials. Each material is represented by a point with coordinates m and $1-n$, numbers refer to the Table 5.1 which gives literature references. Where it has not been possible to label the points internally, an external label is given relating to a solid contour. The Debye process corresponds to the top right-hand corner marked D, the Cole–Davidson expression corresponds to $m = 1$ on the top side marked C–D, while the Cole–Cole symmetric characteristic corresponds to the diagonal C–C. Where the exponents m and n show a variation with temperature, the "best" characterised values are shown.

From Hill (1981a).

It is evident that the behaviour of a great majority of the materials shown in Figure 5.27, which are generally representative of all dielectrics with dipolar loss peaks, do not conform to any of the "single parameter" functions and that recourse must be had to two-parameter representations, with two independent parameters, such as the exponents α and β in the Havriliak–Negami law (3.39), or our exponents m and n.

It should be noted that most of the data falling on the top side of the square, i.e. showing the value $m = 1$ for the low-frequency branch of the spectrum, correspond to measurements taken at very high frequencies in excess of 10^8 Hz, such as those shown in Figures 5.9 and 5.10. This point will be taken up in Chapter 8.

We conclude, therefore, that the behaviour of real dielectric materials indicates unmistakably the presence of at least *two separate and independent processes*, characterised by the exponents m and n, both falling strictly within the ranges (0, 1). A satisfactory theory of dielectric response must be capable of explaining this general feature in terms of recognisable physical processes and must be able to do so largely independently of the detailed physical and chemical nature of the materials involved, as is clear from the list in Table 5.1 explaining the types of materials for which coordinates are given.

5.4 THE DIELECTRIC BEHAVIOUR OF p–n JUNCTIONS

We have outlined in Section 4.8 the principles of the frequency-response arising from generation-recombination processes in semiconductors, with particular reference to the behaviour of semiconductor p–n junctions. In these rather unusual "dielectric" systems the inherent delay in the response with respect to the applied signal arises not from the presence of identifiable dipoles in the space charge region – in fact, it is very difficult to see how dipoles having the desired low-frequency response might arise – but from the delay in the generation or de-trapping of charge carriers at the edges of the space charge region, as shown in Figures 4.13 and 4.14.

Despite the great technological importance of p–n junctions and despite the fact that their capacitive properties have been well

recognised for many decades, there is relatively little experimental information on the dielectric spectrum of their loss and polarisation as functions of frequency. Recent work shows, however, that p–n junctions have very characteristic responses and their study has contributed significantly to an improvement of our theoretical understanding of the nature of the trapping and recombination processes in semiconductors.

We have already shown the results of measurements on a relatively "pure" p–n junction, which gave an almost Debye-like spectrum, Figure 5.7. In the present Section we shall augment this picture with several spectra of different p–n junctions which depart very drastically from the Debye behaviour but which nevertheless give the same general pattern of "universal" response found in dipolar materials, eqn (5.3). This remarkable similarity of behaviour with a very different basic physical process underlying it encourages a critical reappraisal of the relevant theory.

Figure 5.28 gives the complex susceptibility spectra of a silicon rectifier diode. The loss data are as measured, with strong evidence of dc conduction at low frequencies and higher temperatures. The data have been normalised with respect to temperature, and the locus of the displacement points is shown on the same diagram, together with the relevant activation plot which gives an activation energy of 0.48 eV in the higher temperature region.

Once outside the influence of the dc conduction processes, the spectrum shows a well-defined power law with $m = 0.18$ and this is followed beyond the peak by a relatively steep power law with $n = 0.42$ extending for at least three decades. The corresponding activation energy at the lower temperatures is 0.35 eV.

The real part of the susceptibility was obtained from the $C'(\omega)$ data by requiring that the high-frequency part should be compatible with the very well defined power law for the loss, eqn (3.31). This gave values of C_∞ for the different temperatures – we note that the capacitance of a p–n junction depends on temperature not only through the dependence of the permittivity, but also, and to a large extent, through the temperature dependence of the width of the space charge region, so that we cannot determine the absolute values of $\varepsilon(\omega)$ or $\chi(\omega)$. It is clear from the examination of the real part of $\chi'(\omega)$ that it saturates very slowly towards low frequencies, which is characteristic of the behaviour of systems with small values of the exponent m, in accordance with the known properties of the relevant Kramers–Kronig transforms, eqn (3.40) (Hill 1981b). It

is also clear that there is no evidence of another lower-frequency process setting in the region dominated by the dc conduction – this is important, because it proves that the dielectric response of this particular junction follows a single law given by eqn (5.3) over a range of *ten decades* of frequency. This is the more remarkable since the normalisation procedure has shown the presence of a variation

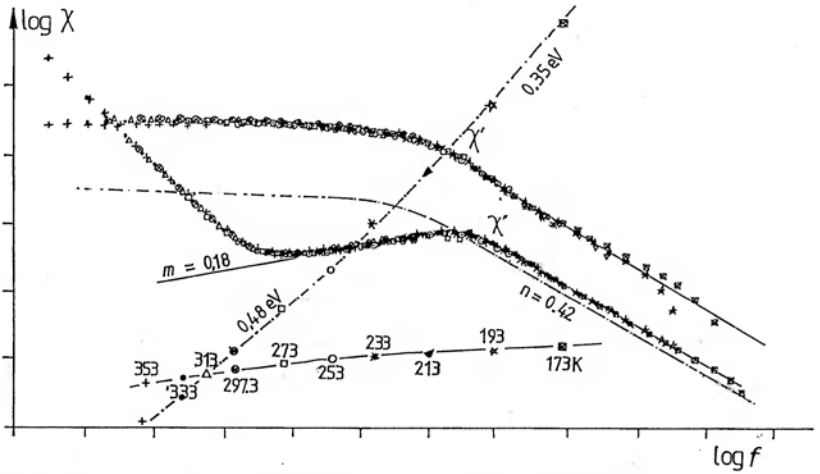


Figure 5.28 The dielectric spectrum of the real and imaginary components of the susceptibility of a silicon p–n junction diode, taken over a range of temperatures. Since the thickness of the space charge region is not known precisely, the data correspond to $\chi'(\omega) \propto C'(\omega) - C_\infty$ and $\chi''(\omega) \propto G''(\omega)$, where C_∞ is a suitably chosen value which gives a Kramers–Kronig compatible frequency dependence for the real and imaginary parts at frequencies in excess of the loss peak frequency. The data for the real part $\chi'(\omega)$ are displaced vertically for clarity – their proper position with respect to the imaginary part is shown by the chain-dotted line. The locus of the normalisation points shows that the amplitude of the loss peak increases with increasing temperature. The chain-dotted line superimposed on the diagram is the activation energy plot of the horizontal shift of the normalisation points against a $1/T$ scale (not shown), giving two extreme values of activation energies. The steeply rising loss at low frequencies corresponds to the dominance of direct current conduction.

From Charoensiriwatana (1982).

of the activation energy, as witnessed by the rising slope of the activation plot in Figure 5.28. These measurements were supplemented by a detailed study of the dependence of the dielectric response on the applied steady bias, which has the effect of changing the width of the space charge region and also altering the distribution of localised charges in the space charge region. Figure 5.29

gives the results, processed in a similar manner as above, for 0, 2, 5 and 10 volts reverse bias. Measurements with forward bias are not practicable because of the dominant direct current conduction in this mode.

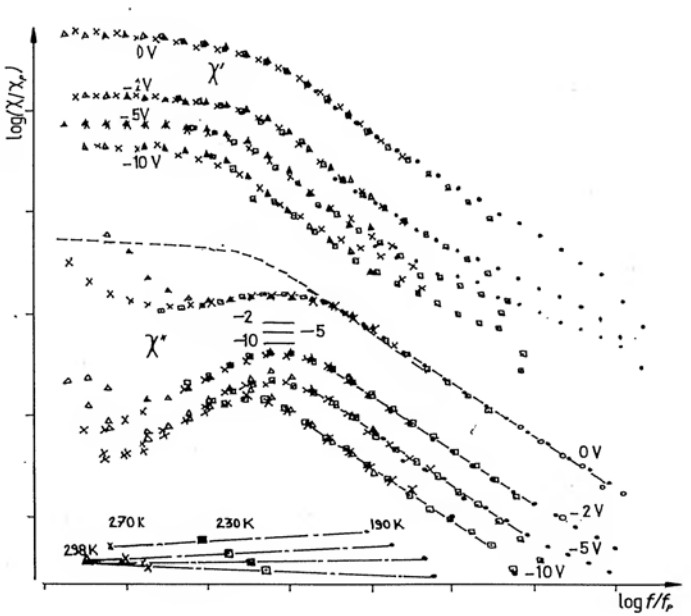


Figure 5.29 The effect of a steady reverse bias on the loss spectrum of the diode shown without bias in Figure 5.28. The data for individual temperatures have been normalised similarly as in Figure 5.28. The plots corresponding to the real and imaginary components are displaced with respect to one another for clarity, but the individual plots for different biases are in their correct relative positions within each set. The positions of the loss peaks in the frequency scale are correctly given, indicating a slight decrease of the loss peak frequency with increasing bias. The normalisation points are indicated at the bottom of the diagram, they are the same for the real and imaginary components and they are referred to a common point for all four biases at 298 K. The regular shift of the temperature points with increasing reverse bias indicates that the activation energy is increasing. The dotted line in the loss plot shows the correct position of the real part for zero bias. The general conclusion is that the spectrum of relaxation does not change with reverse bias, apart from a slight shift in frequency and a change of amplitude, the latter due to changing width of the space charge region.

The short horizonal lines on the loss plots indicate the corresponding positions of C_∞ which provide a measure for the effective width of the space charge region at each bias. There is little variation of $C_\infty(0)/C_\infty(-V)$ for different temperatures. The implication is that χ''_m/C_∞ decreases with increasing bias, while χ''_m/χ' does not, since they must remain Kramers–Kronig compatible.

From Charoensiriwatana (1982).

The very interesting result of these measurements is that the *shape* of the loss *spectrum* does not depend on bias – the normalised data show a parallel shift with respect to the zero-bias response. What is significant, however, is the reduction of the *amplitude* of loss and that by a larger factor than the corresponding reduction of C_∞ which may be presumed to depend on the thickness of the space charge region. The absolute reduction of the ratio of loss amplitude to C_∞ is consistent with the interpretation of the process as a series combination of a lossy region near the edge of the space charge region and a relatively loss-free but voltage-dependent capacitance of the main space charge region itself.

A completely different semiconductor system is represented by a GaAs/GaAlAs double heterostructure laser diode junction whose loss spectrum for a wide range of temperatures is shown in Figure 5.30 a). Here the high-temperature low-frequency response is clearly dominated by the direct current conduction processes which one would expect in these junctions with their high density of deep levels and very short life-time of the order of 10^{-9} s. As the temperature decreases towards 50 K the low-frequency part of the spectrum collapses until the loss becomes nearly flat and at 35 K there appears a rising part of a new loss peak which becomes even more pronounced at 4.6 K.

These data were further processed together with the real part of the capacitance using a computer and a Visual Display Unit to obtain the best possible self-consistent set of data by subtracting a dc conductance from the low-frequency loss and a C_∞ from the high-frequency $C'(\omega)$ data. The result is shown in Figure 5.30 b)

Figure 5.30

a) The dielectric loss data for a GaAs/GaAlAs double heterostructure laser diode, taken over a wide range of temperatures and presented without subtraction of the dc component which dominates the high-temperature behaviour.

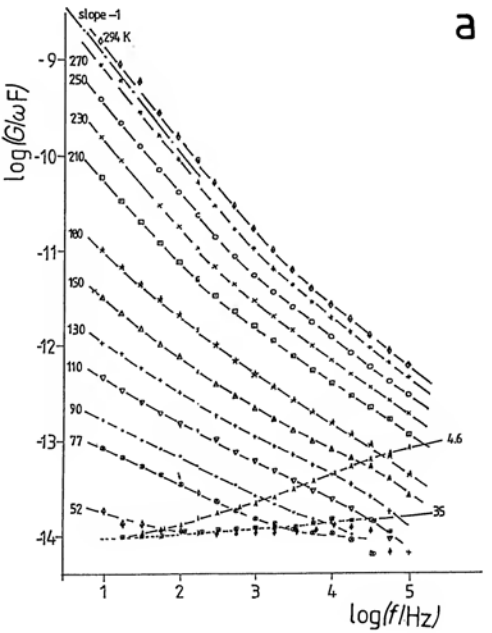
Note the flattening of the response at 50 K and the appearance of a loss peak at 4.6 K.

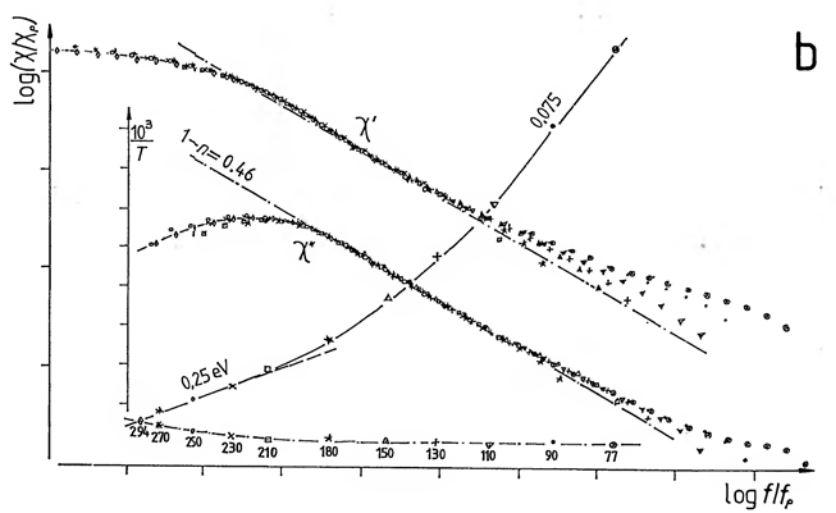
b) The same data after a self-consistent adjustment of G_0 and C_∞ to obtain a Kramers–Kronig compatible set of normalised curves. The real and imaginary components are displaced vertically for clarity. The locus of the displacement points and the activation diagram are also plotted.

The tailing off at high frequencies is a reflection of the flattening of loss at low temperatures in a).

Note the continuous variation of the activation energy.

Charoensiriwatana (1982).


a
slope −1
294 K
270
250
230
210
180
150
130
110
90
77
52
4.6
35
log(G/ωF)
log(f/Hz)


b
log(χ/χ_p)
χ'
χ''
0.075
1−n=0.46
0,25 eV
10³/T
294 270 250 230 210 180 150 130 110 90 77
log f/f_p

in the form of the normalised response for the very wide temperature range 77–294 K. The first impression is the success of the self-consistent treatment in uncovering the loss peak from underneath the "mountain" of dc conduction loss. The second surprise is the similarity of the overall pattern of the response to the silicon junction – the steeper high-frequency slope than the low-frequency slope, confirmed here again by the slow saturation of the real part. The presence of a secondary high-frequency process is evident here and this is consistent with the evident tendency to a flat response at lower temperatures. One further interesting feature is the large range of activation energies which fall from 0.25 eV at the higher temperatures to 0.075 eV below 100 K.

Our survey of the dielectric properties of p–n junctions concludes with the data of a silicon $n^+ - p$ junction bombarded with 0.7 MeV electrons. This treatment is known to create a considerable amount of lattice disorder, resulting in the appearance of deep levels with well defined activation energies. Figure 5.31 shows the loss data normalised for temperatures in the range 80–293 K showing three distinct loss peaks in addition to the dc conduction. The activation plot superimposed on the main diagram gives the activation energies.

This diagram illustrates very clearly the problem arising with normalisation in the presence of several partially overlapping processes – it is impossible to obtain a completely consistent normalisation across the entire spectrum and one has to concentrate on the individual principal features, in the present instance the loss peaks, leaving the valleys between them rather less well defined. The loss peaks are not sufficiently well separated to enable their slopes to be determined with assurance, only the n exponents of the outer two peaks could be obtained and they are in the region of 0.4. This shows that the peaks are significantly broader than the Debye shape, but they are not as wide as those seen in Figure 5.28 and 5.30. Their activation energies compare well with data obtained independently on material treated in the same manner, for example using the technique of Deep Level Transient Spectroscopy (DLTS).

It is appropriate to comment here on the relative merits of the "dielectric" spectroscopy technique described above, alternatively known as admittance spectroscopy, and the other techniques of DLTS and Thermally Stimulated Capacitance (TSCap).

Both these latter techniques are much more rapid and require less sophisticated equipment than the admittance spectroscopy, but

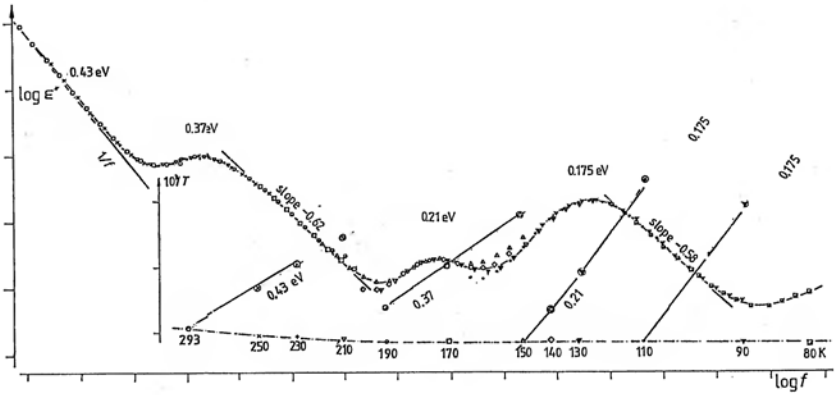


Figure 5.31 The dielectric loss spectrum after normalisation for a silicon $n^+ - p$ junction bombarded with 0.7 MeV electrons showing, in addition to the low-frequency direct current conductivity manifesting itself by the $1/f$ slope, three distinct loss peaks whose activation energies are indicated. The diagram contains two separate plots which are superimposed on one another – the normalised loss *versus* logarithm of frequency plot with its associated locus of the diaplacement, and the plot of the same displacement *versus* $10^3/T$ which gives the activation energies. In a plot covering four differently activated processes in 15 decades of normalised frequency, it is impossible to obtain a uniform normalisation throughout, due to the overlap of mechanisms. The normalisation has therefore centred on the loss peaks, leaving the valleys less well defined. The high-frequency slopes of two of the peaks are indicated, both corresponding to the exponent $n \cong 0.4$. The peaks are not sufficiently well separated to enable the exponents m to be determined with any accuracy, neither can the middle peak be characterised with certainty.

From Favaron 1982.

they are capable essentially of giving the activation energies and the capture cross sections of the deep centres, while not revealing anything about their spectral properties. As a rapid means of characterising deep centres these methods are therefore adequate, their limitations have been well presented by Favaron (1982). On the other hand, the dielectric admittance spectroscopy reveals information about the frequency or time dependence of the dynamic behaviour of deep traps which, as we have seen, departs very significantly from the classically accepted exponential or Debye model. This information could not be obtained normally from the other techniques, because they inherently lack the necessary bandwidth and yet the lack of this type of information has contributed to the wide acceptance of the simple exponential model which is implicit in most theoretical treatments of recombination and trapping processes. Studies of the kind shown above open the way, therefore, to a fresh appreciation of the need to develop a more

realistic dynamic model of recombination and trapping processes in semiconductors, a model which would account for the fractional power law relaxation which resembles in many respects the dipolar relaxation seen in all dielectrics.

Anticipating the discussion of strong low-frequency dispersion in Section 5.6 it is worth noting that none of the many types of p–n junctions studied in the Chelsea Dielectrics Group has shown any sign of dispersion of the real part $C'(\omega)$ even at the lowest frequencies and highest temperatures investigated. The significance of this will become clear in the discussion of Chapter 8.

5.5 DIELECTRIC RESPONSE WITHOUT LOSS PEAKS

In our discussion so far we have covered the various types of dielectric response involving loss peaks, with the implication that the loss becomes arbitrarily small towards very low frequencies, in agreement with the general requirement that the loss should go through zero at zero frequency as required by eqn (2.42). We now come to an essentially different type of response which is found in many dielectric systems and which we shall come to associate with the presence in the material of partially mobile charge carriers. Here no loss peaks are observable down to the lowest measurable frequencies and instead the response goes over into a steeply *rising* branch towards low frequencies. This might be expected in the presence of a dc conductivity, which would give a contribution of the form σ_0/ω, eqn (2.50), while leaving the real part independent of frequency. We shall see, however, that a further, hitherto not well understood phenomenon is often apparent at low frequencies in charge-carrier-dominated systems in which both the real and the imaginary part of the complex susceptibility rise steeply towards very low frequencies.

When dealing with charge carrier systems it is often preferable to plot the ac conductivity instead of the dielectric loss, as explained in Chapter 3 and we shall use both loss and conductivity plots, according to circumstances. The principal consideration will be the criterion of "true" dielectric response – if the phenomenon in question is ordinary direct current conduction not accompanied by a strong dispersion of the real part of the permittivity, then the conductivity plot is preferred and corresponds more closely to the physical situation, but if there is evidence of simultaneous strong low-frequency dispersion of the real as well as the imaginary part

then the process in question is a genuine dielectric pheomenon and the loss plot is preferred to give more direct comparision with the real part. We mention here that the concept of charge carriers making a contribution in their own right to the dielectric polarisation is unfamiliar to some people who regard charge carriers as only being capable of producing direct current conduction and to be specifically excluded from all consideration in the context of dielectric behaviour. We shall argue later in some detail that this attitude is necessarily restrictive and runs plainly counter to good experimental evidence as well as to good theoretical arguments, but for the present we merely wish to ask the reader to keep an open mind and let the facts speak for themselves, remembering the discussion in Chapter 4 where *hopping charge carriers* were shown to possess Debye-like dielectric properties.

a) Charge carriers in dielectric materials

It is desirable at the outset of this discussion to say a few words about the physical properties of charge carriers in dielectric materials. We are familiar already with the concept of *free* charge carriers – electrons and holes – in semiconductors which move in conduction and valence bands, respectively, with mean free paths between collisions spanning many interatomic spaces. These free charge carriers have large thermal velocities even in the absence of an external field and their drift velocity even in the presence of the highest attainable fields is at most comparable and usually very much smaller than the thermal velocity. The dielectric response of free charge carriers was described in Section 4.4 and we concluded that their dielectric effects only become significant at frequencies of the order of the reciprocal collision time, which fall in the 10–100 GHz range. This means that the dielectric effects of free charge carriers are negligible in the frequency range which we are predominantly concerned with in the present context.

When discussing semiconducting materials, therefore, we shall not be concerned with any free charge carriers that may be present in them, except insofar as these may give rise to such a high dc conductivity that dielectric effects may not be easily measurable. On the other hand, low-frequency dielectric responses are to be expected from any *localised* charge carriers that may be present in a semiconductor, especially in conditions where the effects of the free charge carriers are not dominant, for example at sufficiently low temperatures or in the space charge regions of p–n junctions.

Localised charge carriers may contribute to dielectric relaxation in

two fundamentally different ways which both result in a *delayed response* of current to the applied field. On the one hand, these carriers may suffer a delayed release from the localised levels into the free band where they take part in the ordinary conduction process – this is the generation–recombination process described in Section 4.8. On the other hand, these localised charges may be displaced from their original positions by the action of an external field by means of *hopping* transitions between localised levels, not involving excitations into the respective free bands. This hopping conductivity corresponds to an extension of the two-potential-well model described in Section 4.6 to many interconnecting wells forming an extended network of possible paths that the charge may follow. The main point to bear in mind is that these hopping transitions have very different probabilities according to the relative distances between the localised sites and to their separation in energy, so that certain *easy* transitions will be executed many times in both directions, as in a two-well system, while the more difficult ones are only traversed relatively less frequently. This shows clearly that a hopping charge carrier shows both *dielectric* characteristics, insofar as it behaves like a jumping *dipole* in its reciprocating motions, and simultaneously *conducting* characteristics resulting from its extended hopping over many sites. The point to note here is that the dielectric properties are determined by the *easiest* transitions, while the conducting properties are determined by the *most difficult* transitions which limit the free percolation of charges from one electrode to the other.

The appearance of localisation in a solid is very intimately connected with disorder, such as structural or compositional defects. In crystalline semiconductors any such point defects give rise to localised states which, if their energy separation from the edges of the conduction and valence bands is sufficiently large, act as effective *traps* for electrons or holes, cf. Section 4.7. If the density of these traps is sufficiently large for tunnelling between them to be possible, transport may occur in these states by thermally assisted tunnelling, i.e. hopping.

The density of these traps increases with increasing disorder and a limiting situation is attained in the case of completely disordered or amorphous materials, in particular in amorphous semiconductors which have received a considerable amount of attention in recent years.

Localisation may also take place under conditions of apparent structural order if the medium is sufficiently strongly polarisable

to respond to the presence of a charge by a significant local distortion which leads to a lowering of the potential energy of the carrier with respect to the undistorted medium. This is referred to as the "carrier digging it own potential well" and the resulting quasiparticle consisting of the charge carrier and its surrounding distortion is known as a *polaron*. Polarons are necessarily localised and may move by hopping when sufficient energy becomes available for their excitation out of the local potential well.

We may mention one other contributory factor to the establishment of localisation and this derives from the nature of the chemical bonding in many dielectric materials of the organic type. In these molecular solids described briefly in Section 2.1 the weak van der Waals bonds between neighbouring molecules give rise to very narrow allowed bands, with correspondingly high effective masses of the resulting charge carriers and this, together with the presence of disorder, contributes further to the process of localisation.

Whatever the detailed physical reasons, the fact remains that most commonly used dielectric materials, including virtually all polymers and the majority of glasses and ceramics, do not show any evidence of the presence of free charge carriers, which would give rise to much higher levels of dc conductivity than are actually observed. The inference is that the only carriers that may be present are necessarily very low mobility carriers, such as hopping electronic charges inevitably are. Needless to say, ionic charges are by their very nature hopping carriers, since the concept of a free band for ions has no physical sense, and their mobilities are correspondingly extremely low, even in materials with very high ionic conductivities such as the so-called "fast ion conductors".

b) Alternating current conductivity of hopping charges

Interest in the ac conductivity of hopping charge carriers first arose through the studies by Pollak and Geballe (1961) on compensated silicon – that is material containing both donor and acceptor impurities – at temperatures close to the absolute zero. Under these conditions the only possible mechanism of conduction is by hopping between neighbouring donors or acceptors and this provides one of the most thoroughly studied materials for this type of conduction. Pollak and Geballe found that the ac conductivity obeyed the empirical law of frequency dependence given by the power law of eqn (3.32), with the exponent n in the region 0.7–0.8. For several years this result remained essentially a curiosity until the rise of

interest in the properties of amorphous semiconductors has led to the realisation that the same power law relation applied to many other materials thought to conduct by hopping electrons.

A collection of data relating to the ac conductivity of a number of materials is shown in Figure 5.32, where $\log \sigma(\omega)$ is plotted against $\log \omega$ with individual sets of data denoted by a letter of the alphabet being displaced with respect to one another for clarity, while retaining a common frequency scale. It is immediately clear that the power law of eqn (3.32) is generally obeyed with the exponents n falling in the range $0.6 < n < 1$, with three exceptions – the two materials under f and the very high frequency region of m, these will be discussed later. We note that at the lower end of the frequency range many of the materials show a flattening of the characteristic towards what might be mistaken for the dc conductivity, although closer examination will reveal the absence of complete saturation in most of the cases under consideration. This incompletely saturated low-frequency conductivity will be discussed in more detail in the context of Section 5.6.

Some further general features may be mentioned. The temperature dependence of the lower frequency response is much stronger than that of the high-frequency end. The exponent n either remains constant or decreases slightly with increasing temperature. While some data extend only over a limited range of frequency, say, 3–4 decades, others are much more extended and there is no doubt that, on the whole, the power law relation is very well obeyed, and certainly no other simple law could fit the observed facts nearly as well.

We now consider the significance of the slopes n exceeding unity, as in cases f and m. The implication of this is that the dielectric loss is *rising* with increasing frequency, while for $n < 1$ it is falling. Thus we conclude that the response of anthracene and β carotene under f corresponds to a very shallow loss peak *superimposed on an almost constant background* extending over many decades of frequency. Similarly, the high-frequency part of the response of As_2S_3 under m corresponds to a sharp, nearly Debye-like rising loss above 10 GHz – a feature already familiar from several other high-frequency data for polymeric and other materials.

As regards the range of materials shown in Figure 5.32, they represent a wide selection of structural, chemical and physical properties. The include classic examples of electronic hopping systems, e.g. the single crystal silicon at very low temperatures a), or

the chalcogenide glasses l), the classic example of an ionic conductor, b), as well as many materials that are not evidently either one or the other. Their common feature is that, unlike the materials characterised by the presence of loss peaks, they invariably show a gradual transition towards dc or very slowly varying conductivity at low frequencies.

While it is clear that all the materials shown in Figure 5.32 follow the frequency dependence given by the power law (3.32), and this is remarkable in itself, we next present the same information on common $\log \sigma(\omega)$ as well as $\log \omega$ scales, Figure 5.33. The most surprising conclusion from this is that the ac component of the conductivity for all these very different materials has an absolute value within at most four decades, while the majority of materials fall within two decades. The chain-dotted lines drawn correspond to frequency-independent losses ($n = 1$) of $\chi'' = 10^{-3}$ and 10, respectively. Compared with this relatively very narrow spread of absolute values of the ac conductivity, the dc conductivity is seen to span a range of well upwards of 10 decades – as might be expected for this very varied selection of materials.

It is also clear that the temperature dependence of the dc component of conductivity is much stronger than that of the ac component, the latter being relatively insensitive to temperature.

Having noted the "giant dispersion" found in ferroelectric materials at sufficiently high frequencies, Figures 5.1, 5.6 and 5.22, where a nearly-Debye-like response is found, it is interesting to note that the same materials show a completely different behaviour at lower frequencies. Figure 5.34 gives the data for the frequency dependence of the real and imaginary components of the susceptibility for the same triglycine sulphate as in Figure 5.6 but in the frequency range between 10 and 10^5 Hz and in the temperature interval -10 to $+70$°C which includes the Curie temperature, $T_c = 50$°C. The corresponding temperature dependence of the real and imaginary parts of the permittivity is shown in Figure 5.35 and shows the familiar rapid rise of ε' and ε'' as T_c is approached from either below or above. While this temperature behaviour is completely classical, the frequency response shown in Figure 5.34 is quite unfamiliar and follows the universal law *both* above and below T_c.

c) Fast ionic conductors

We now look at the dielectric behaviour of fast ionic conductors (Vashishta et al 1979) which owe their properties to the existence

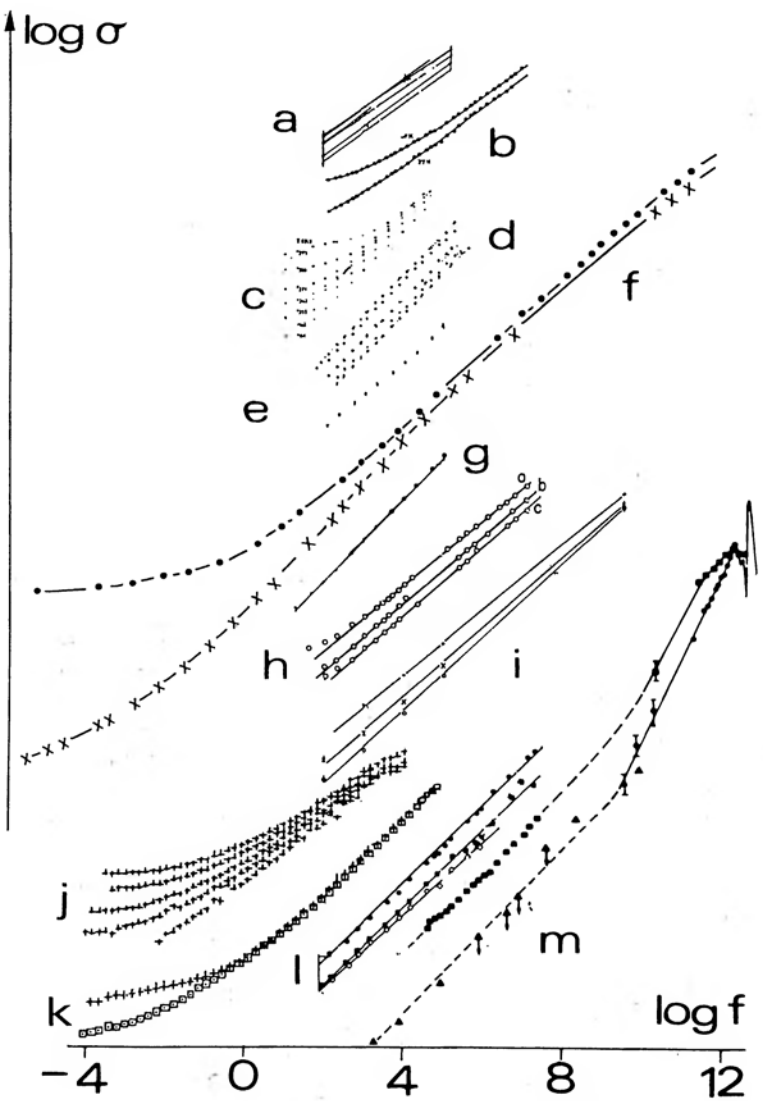


Figure 5.32 A complilation of ac conductivity data for a range of materials presented on a common log f basis in Hertz, but displaced vertically for clarity. Data sets denoted by one letter are on a common log σ scale.

a) Single crystal silicon in the impurity hopping range of electrons, 3.0 K (bottom), 4.2, 8.0 and 12.0 K (Pollak and Geballe 1961).

b) Single crystal β alumina at 77 and 87 K – a classical fast ion conductor by Na^+ ions (Grant et al 1977).

c) Glow-discharge deposited amorphous silicon, 84–295 K (Abkovitz et al 1976)

d) A range of chalcogenide glasses at 293 K (Roberts and Polanco 1972).

e) Single crystal anthracene with 1 M saline solution as contact, 294 K (Abkovitz et al 1975).
f) Single crystal anthracene (crosses) and β carotene (dots) at 294 K (Hughes and Pethig 1975).
g) Trinitrofluorincne (TNF) – Polyvinylcarbazole (PVK) – a molecular glass with relatively "free" hopping motion by electrons (Abkovitz et al 1974).
h) A range of P_2O_5–FeO–CaO glasses at 300 K (Murawski and Gzowski 1974).
i) A glass of the ccmposition V_2O_5–P_2O_5 at three temperatures (Sayer et al 1971).
j) Evaporated amorphous silicon monoxide at 211–297 K (Frost and Jonscher 1975).
k) A 9-layer stearic acid film between an Al and a Au electrode, 300 K, in the dark (squares) and with ultra-violet light (crosses) (Careem et al 1977).
l) Three samples of amorphous films at 300 K. Top to bottom As_2Se_3, Se and As_2S_3. Measurements believed to correspond to bulk properties without interference from electrode processes (Lakatos and Abkovitz 1971).
m) Two samples of Se_2Se_3 at 300 K with the frequency range extending to far infra-red and showing a steeply rising region above 10 GHz (Strom and Taylor 1974).

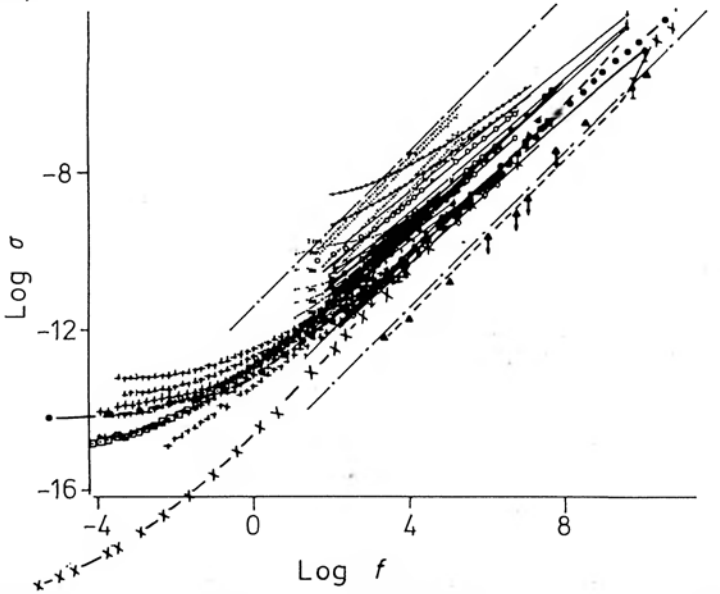


Figure 5.33 The ac conductivity data from Figure 5.32 placed on a common log $\sigma/(\Omega\ \mathrm{cm})^{-1}$ scale and showing the remarkably narrow range of absolute values of ac conductivity for very different materials, all obeying the universal law (3.32). The upper and lower chain-dotted lines correspond to frequency-independent loss, respectively, $\chi'' = 10$ and 10^{-3}. This is a corrected version of a diagram originally published by Jonscher (1977).

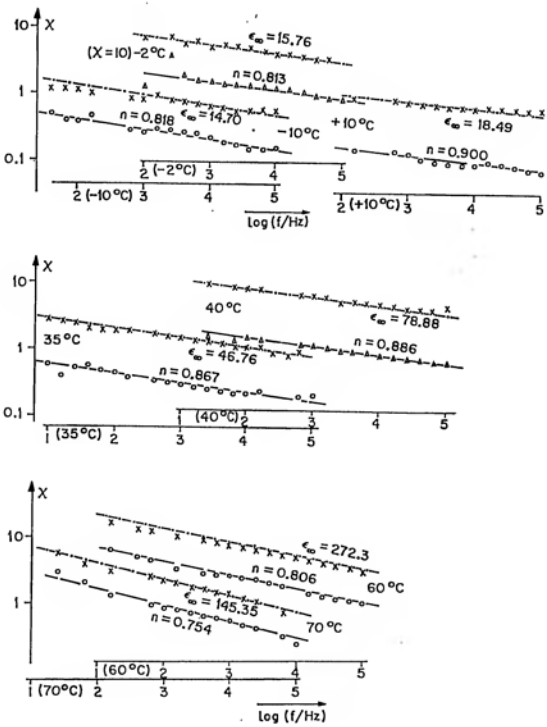


Figure 5.34 The real and imaginary components of the susceptibility of Triglycine Sulphate in the "low" frequency region in a range of temperatures below and above the Curie temperature, $T_c = 50°C$. The continuous lines represent the best fit through the experimental loss points, the chain-dotted lines are in Kramers–Kronig-compatible positions for $\chi'(\omega)$. The values of ε_∞ have been determined to place the points of the real part $\varepsilon'(\omega) - \varepsilon_\infty$. The various sets of data are displaced in both the vertical and the horizontal directions for clarity, the frequency scales are drawn accordingly. Note that the universal power law relation applies both above and below T_c.

From Jonscher and Dube (1978).

of "redundant" lattice vacancies arising not from disorder, as would be the case in ordinary ionic conductors, but are an intrinsic property of their structure. This greatly facilitates ionic movements, since it is not necessary to create vacancies by thermally displacing ions from their normal positions. We shall begin by reference to Figure 3.19, showing an inclined circular arc impedance diagram often found in ionic conductors. This was shown to be a consequence of the universal response of these materials, (Jonscher 1975b/1978a), although it was not normally so interpreted. Figure 5.36 gives experimental data for β PbF_2 with good circular arcs inclined

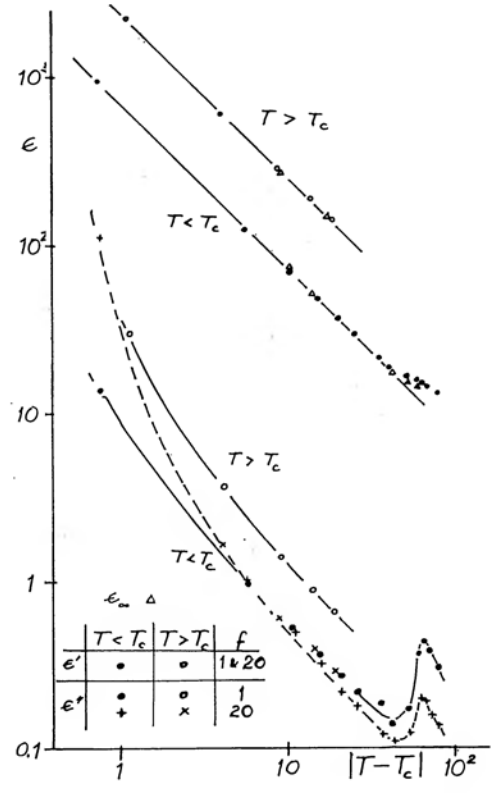


Figure 5.35 The temperature dependence of the real and imaginary components of the permittivity for Triglycine Sulphate, plotted against the logarithm of $T - T_c$. The significance of the symbols is explained in the inset. The triangles denote the values of ε_∞ obtained from Figure 5.34, the implication being that the loss process shown in Figure 5.34 makes only a small contribution to the total permittivity.

From Jonscher and Dube (1978).

at an angle of 0.33 rad, giving $\alpha = 0.21$. In addition to these circular arcs, however, we also note the presence of very definite inclined "spurs" corresponding to the lowest frequencies – these are the beginnings of a second set of inclined circles which, very fortunately for the interpretation of data, are well separated in frequency from the former. The equivalent circuit of this is shown as a series combination of two parallel circuits representing the volume and barrier regions – the latter being identified with the low-frequency "spurs" in the impedance diagrams. The conductances G_v and G_b represent the dc mechanisms and the fact that the barrier region

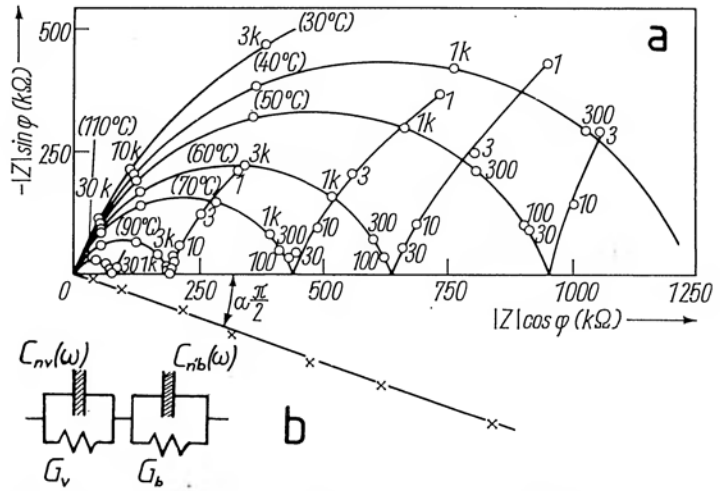


Figure 5.36 The complex impedance plot for β PbF_2 taken from the original paper by Reau et al (1975) – diagram a). The equivalent circuit of this system is represented in diagram b) and consists of a series combination of volume and barrier admittances with their universal capacitors $C_n(\omega)$ and conductances G. Numbers indicate frequencies in Hz.

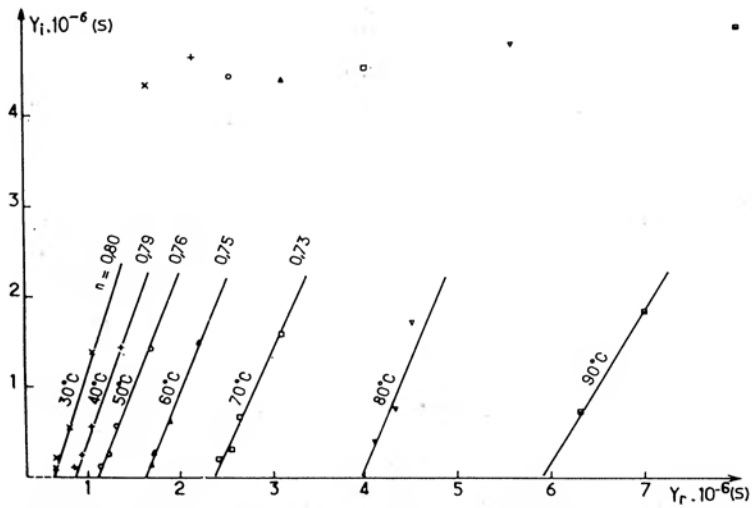


Figure 5.37 The complex admittance plots for the circular arcs corresponding to the high-frequency parts of the impedance characteristics of Figure 5.36. The straight lines are drawn through the lower-frequency points, the deviations at higher frequencies are due to the effect of the frequency-independent component C_∞ as explained in Figure 3.18. The numerical values of the exponent n are determined from the slopes of these lines.

From Jonscher and Reau (1978).

has a non-zero conductance is proved by the evident curvature of the spurs. It is essential to note that the equivalent circuits can only correspond to the observed data if the two capacitors are taken to be "universal" in the sense of the definition (3.29), an ideal frequency-independent capacitor could not under any circumstances represent the data. In accordance with the procedures outlined in Chapter 3 we now invert the high-frequency circular arcs into the admittance plane, which may be done without any further complications since the impedance arcs go through the origin and may therefore be expected to invert into inclined straight lines, as is seen in Figure 5.37. This figure shows clearly the advantages of the admittance representation, compared with the impedance plots: the eye is more sensitive to small departures from a straight line than from a circle. It is now clear that the inclination of these admittance plots does vary with temperature, which was not clear from the apparent position of all the centres of arcs on one line in figure 5.36. It is also clear that the highest frequency points do not fall on the extrapolation of the lower straight lines – they lie distinctly higher and this is due to the effect of C_∞. It is also clear that the parallel dc conductance increases rapidly with rising temperature, as indicated by the intercepts on the real axis.

We may now take the analysis of these data one stage further by plotting the frequency dependence of the dielectric parameters – the most meaningful representation in the final analysis. With reference to eqn (3.34) we note that the determination of the dielectric loss, which corresponds to the *real* part of $Y(\omega)$ is complicated by the presence of the large value of G_0 which makes the exact determination of the true dielectric loss component rather

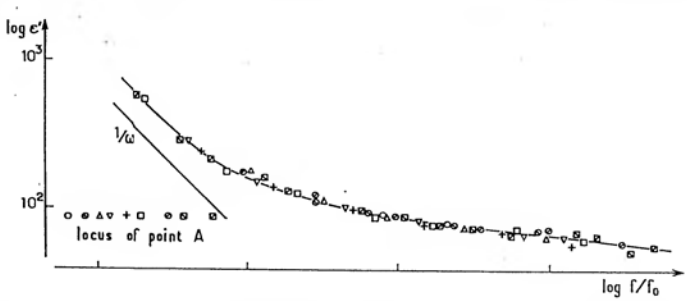


Figure 5.38 The real part of the relative dielectric permittivity of the volume region of a sample of β PbF_2 determined from the imaginary part of the admittance in Figure 5.37. The locus of the point "A" gives an activation energy of 0.25 eV, which is almost half of the activation energy of the dc conductance $G_0(T)$.

From Jonscher and Reau (1978).

uncertain. By comparison, the determination of ε' is completely unambiguous from the *imaginary* part of $Y(\omega)$ and only the subsequent determination of the susceptibility may be affected by the finite contribution at high frequencies from the ε_∞. The results of this determination, and of the subsequent normalisation of $\varepsilon'(\omega)$ is given in Figure 5.38 which clearly falls into two parts – one corresponding to the universal response with a value of the exponent $n = 0.9$, and the low-frequency part which shows a rapid dispersion of the type to be described in the following Section. There is a certain discrepancy between the value of n determined in this manner and the values indicated in Figure 5.37 and it may be that the reason for this is the inevitable complicating influence of ε_∞ which may not be apparent in Figure 5.38 but which would cause a distortion of the plot in the sense of smaller slope, or larger values of n.

In the determination of the absolute values of $\varepsilon'(\omega)$ we have used the geometrical capacitance of the sample and we note the relatively high values of ε_r which is rather typical of many ionic conductors.

The same experimental data offer a rather unique opportunity of analysing the low-frequency "spur" region of the impedance diagram, corresponding to some form of *barrier* region which is in series with the volume of the sample. A detailed presentation of all available data, not all of which were presented in Figure 5.36 for clarity, is shown in the form of the admittance diagrams in Figure 5.39 a) in *linear* representation. The consecutive plots are displaced with respect to one another along the real axis, but they all go through the origin for the temperatures 40–80°C, indicating that there is no measurable parallel barrier conductance and once again showing the greater precision of the admittance plot, compared with the impedance which seemed to indicate the presence of some curvature. The exponent has consistently the value $n = 0.62$ which is very definitely smaller than the corresponding bulk values. The data for the highest temperatures give a larger scope since there are more points, even though they are not all shown in Figure 5.36. A *logarithmic* admittance plot giving all points is shown in Figure 5.39 b) after the subtraction of the now measurable parallel conductance of the barrier region. The offset from the line $Y_i = Y_r$ is clearly visible, although an exact determination of the exponent n is not possible from this graph.

Having shown that the barrier region itself also corresponds to a universal capacitance, with a different value of n' from the bulk

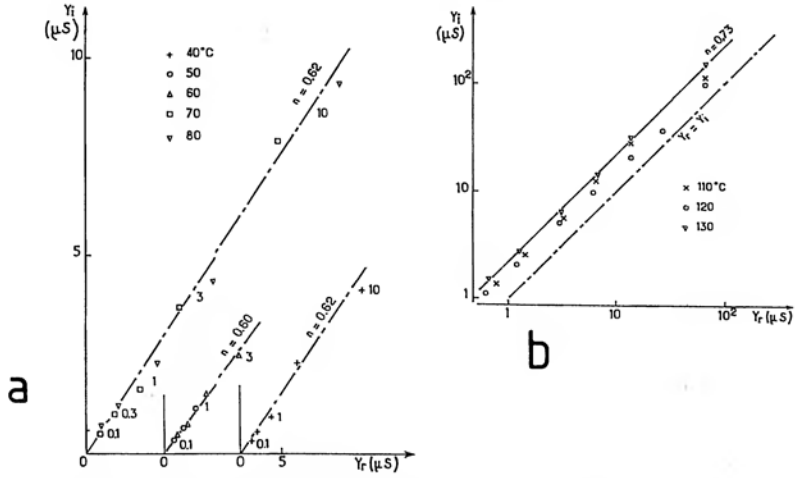


Figure 5.39 The inversion of the "spur" data into the admittance plane to reveal the "universal" character of the barrier region capacitance. Diagram a) corresponds to the lower temperature data and is plotted in linear coordinates, its slope defining the value of the exponent n through the relation slope $= \tan(n\pi/2)$. There is no evidence of dc conductance in these data. Diagram b) is in logarithmic coordinates and corresponds to the higher temperatures, after the subtraction of the dc conductance. The numbers in a) denote frequencies in Hz.

From Jonscher and Reau (1978).

material, it is possible to go one step further and evaluate the absolute value of that capacitance for various temperatures. Further details of this type of analysis and of the physical conclusions which may be drawn from it may be found in the paper by Jonscher and Reau (1978) from which these data were taken.

We have taken the example of β-PbF_2 to show what can be done with experimental data covering a sufficiently large range of frequencies and temperatures. The important conclusions from this analysis are as follows:

i) the response of typical ionic conductors follows the "universal" pattern, exactly as for electronic conductors,
ii) both the volume and the barrier responses may be described in terms of universal properties, although the values of the exponents n_v and n_b are not normally the same.

This presentation of the dielectric response of ionic conductors will be concluded with an entirely different result relating to sodium β-alumina. This is one of the favourite ionic conductors which has a layer structure of chemical composition $Na_2O \cdot 11\ Al_2O_3$ with the

Na^+ ions relatively free to move in the planes defined by the Al_2O_3 spinel blocks. The conductivity within the planes is therefore high and is due entirely to Na^+ ions, while normal to these planes there is virtually no conductivity, either electronic or ionic.

When measured with metallic electrodes normal to the planes of high conduction, the samples present typical impedance diagrams characteristic of a series R–C combination, where the resistance is that of the bulk material and the capacitance is due to the contact barrier which is not completely transparent for sodium (Hooper 1977).

While these measurements in the planes of high conductivity are important for assessing the bulk conductivity of the material, it is interesting to look at the behaviour *normal* to the planes, since it is in this direction that genuinely dielectric processes may be observed. Figure 5.40 shows the data in the form of a normalised plot of the real and imaginary components of the susceptibility, the former having been obtained from the real part $\varepsilon'(\omega)$ in the usual manner

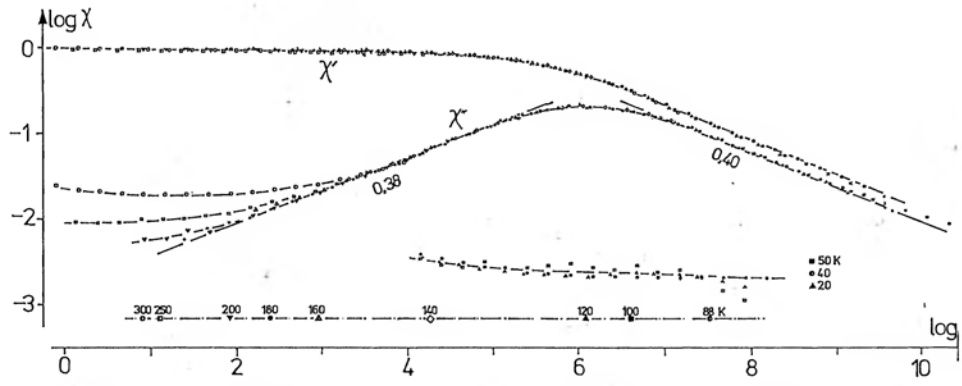


Figure 5.40 The dielectric response of single crystal sodium betha alumina, measured normal to the planes of easy conduction, where there is no detectable dc conductivity, in the temperature range 20–300 K. Normalisation was carried out for all temperatures for which the loss peak is visible, the lowest temperature data corresponding to frequency- and temperature-independent loss are reproduced without normalisation. There is no visible change of the loss peak amplitude between 88 and 300 K. The real part was obtained by subtracting for each temperature a suitable value of $\varepsilon_\infty(T)$ to obtain Kramers–Kronig compatible behaviour with the imaginary part. The continuous line through the points corresponding to $\chi'(\omega)$ – where it can be discerned between the data points – corresponds to the numerically computed Kramers–Kronig transform from the line drawn through the $\chi''(\omega)$.

From Deori et al (1983).

by subtracting at each temperature a suitable value of $\varepsilon_\infty(T)$ to make the response compatible with the loss component. A very wide range of temperatures has been covered and this gives rise to a total frequency range of eleven powers of ten. There is no trace of dc conductivity even at the highest temperatures used in the present study. The normalised loss peak is almost exactly symmetric and very broad and there is evidence of a low-frequency process which gives a plateau with some temperature dependence. It is remarkable that at temperatures below 50 K the loss peak disappears and the loss becomes almost independent of frequency and also independent of temperature between 20 and 50 K. The mean activation energy obtained by plotting the horizontal translation of the normalisation point against the reciprocal temperature is 0.17 eV. For comparision, it may be noted that the dc conductivity activation energy is 0.13 eV for a similar material (Hooper 1977), but the two processes are completely different. It is believed that this loss peak arises from some finite moisture content in the material – its amplitude becomes smaller as the sample is dried (Deori et al 1983). The implication is that an ion such as OH^- or H^+ can take up one of two equivalent positions, giving rise to a dipole-like peak which finally disappears when the thermal energy becomes insufficient to excite the motion.

It would be easy to quote many more examples of the universal response without loss peak, i.e. going over into what appears to be direct current conduction at sufficiently low temperatures, but we will change the emphasis of presentation before proceeding further. The power-law frequency dependence given by eqn (3.32) which we repeat below

$$\sigma(\omega) = \sigma_0(T) + a(T)\omega^{n(T)} \tag{5.6}$$

with an emphasis on the temperature dependence of the various parameters is very familiar when the exponent lies in the range $0.6 < n < 1$ and is often identified with the presence of hopping conduction by electrons. This view is frequently found in studies of a very wide range of compounds, some of them organic, where the authors measure the ac conductivity and uncritically conclude that the dominant mechanism is electronic hopping, while there is no other evidence to support such a statement and often there is ample evidence that this cannot be the case.

The following section will take up the detailed discussion of a limiting case of the value of the exponent n.

5.6 STRONG LOW-FREQUENCY DISPERSION

There exist many examples of a very slowly varying conductivity at low frequencies, of the type already mentioned in connection with the data shown in Figure 5.32. Although this is a very widely observed behaviour, it had not been recognised as a specific type until very recently (Jonscher 1978b). One example of this type of

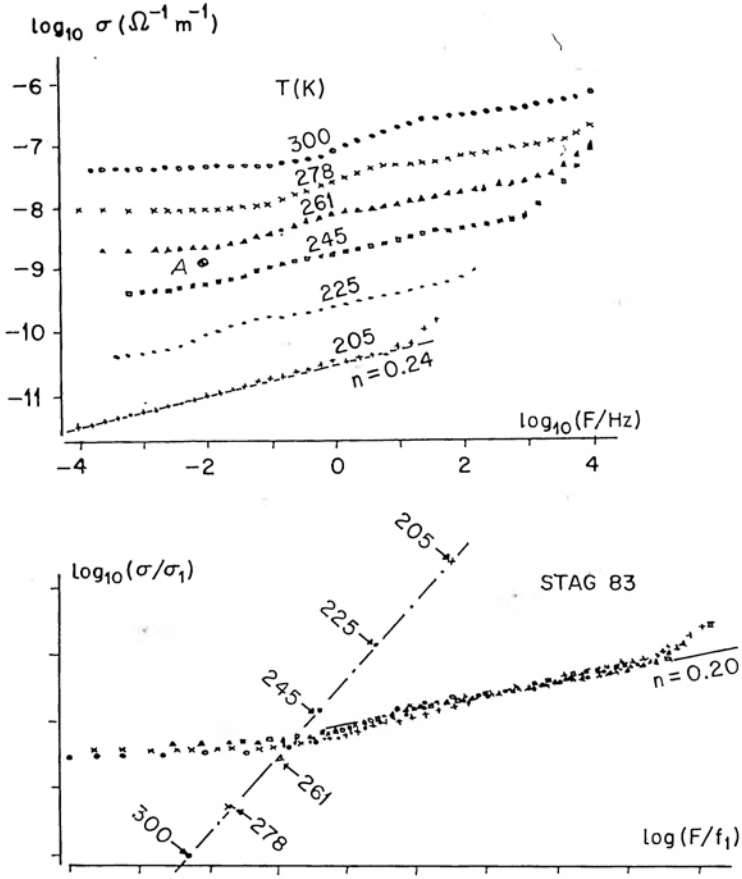


Figure 5.41 The upper diagram shows the frequency dependence of the conductivity of a STAG chalcogenide glass for a range of temperatures. The slight kinks at the higher temperatures are the result of the overlap of the response of a barrier, as shown in Figure 3.15, but the main body of the data relate to bulk response of the glass. The lower diagram shows the result of normalisation with the locus of the representative point "A" now having unit slope. The arrows indicate the positions of the points on the assumption of a simple activated process.

From Jonscher and Frost (1976).

response is shown in Figure 5.41 relating to a sample of chalcogenide glass composed of Silicon, Tellurium, Arsenic and Germanium (STAG), over a range of temperatures. Particularly at the lowest temperature, the power-law relation is clearly discernible and the value of the exponent is 0.24 – much lower than the classically accepted range described above.

This example of the normalisation of *ac conductivity* rather than loss and susceptibility shows the different sense of translation of the data between different temperatures. The locus of the representative point now has a unit slope in the logarithmic plot, where it would have been horizontal in the loss normalisation. This comes from the relation between conductivity and loss $\sigma(\omega) = \omega\varepsilon''(\omega)$.

Another good example of strong dispersion may be seen in Figure 5.42 giving the frequency dependence of the conductivity and of the real part of the permittivity for single crystal alumina – Al_2O_3.

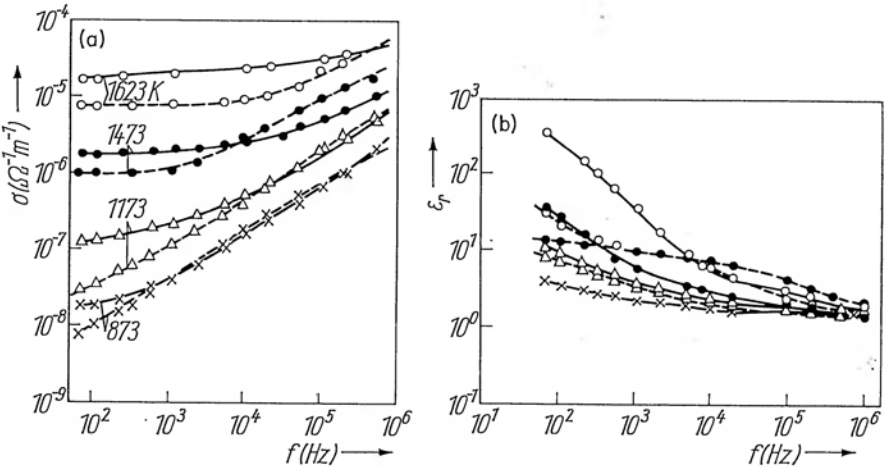


Figure 5.42 The frequency dependence of the electrical conductivity and of the relative permittivity over a range of temperatures for single crystal alumina. The strong dispersion of permittivity and the slowly variable ac conductivity at low frequencies should be noted.

From Kizilyalli and Mason (1976).

Here the "low" temperature is 873 K, "high" 1632 K, so we are dealing with ionic conduction and this is further supported by the fact that the low-frequency conductivity is significantly enhanced by the pressure of oxygen in the ambient. While the "low" temperature data show a relatively "conventional" power law relation

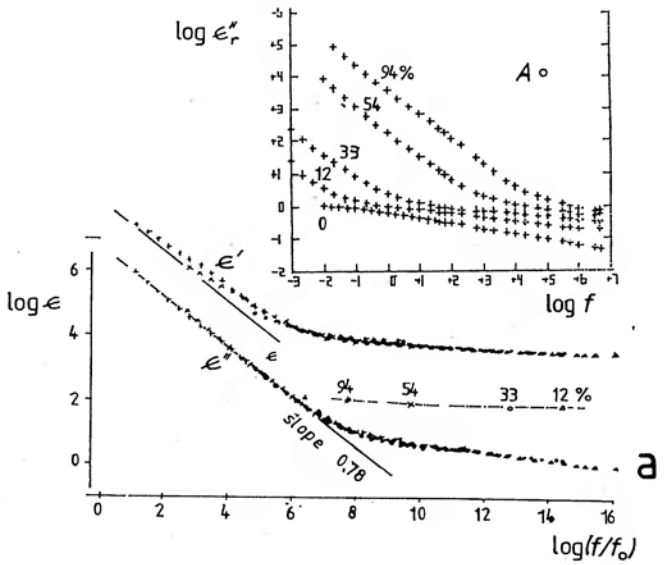
with $n = 0.63$, especially at low oxygen pressures, the tendency to "almost dc", i.e. nearly frequency–independent conductivity becomes evident with rising temperature.

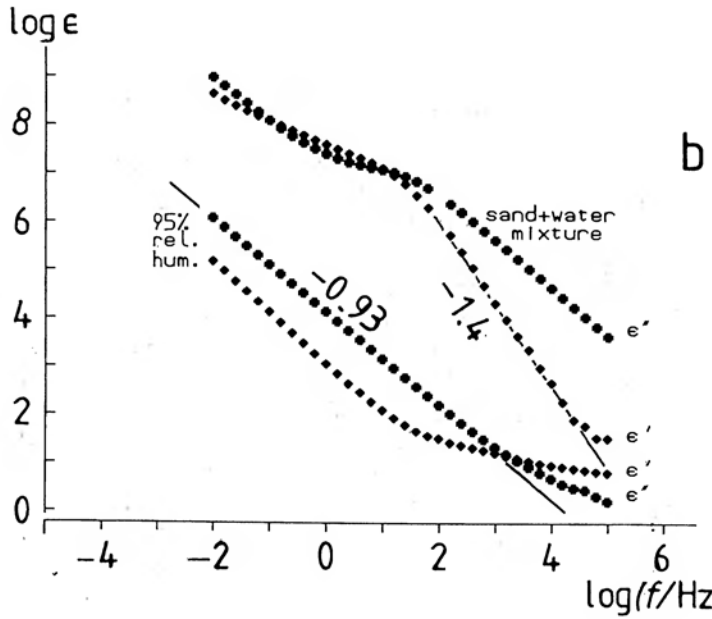
In view of eqns (3.20) and (3.31), we note that the frequency dependence of conductivity of the form ω^n necessarily entails the "universal" form of frequency dependence of *both* $\chi'(\omega)$ *and* $\chi''(\omega)$ on ω^{n-1}, whatever the value of n in the range $(0, 1)$. We conclude, therefore, that *small* values of n entail *strongly* frequency–dependent susceptibility, i.e. strong dispersion of the dielectric parameters. This is clearly borne out in the case of Figure 5.42 by the accompanying plot of the real part of the permittivity which shows a strong dispersion by up to two orders of magnitude at low frequencies. This proves conclusively that the phenomenon in question is not simply dc conductivity.

An unusual dielectric result is shown in Figure 5.43 a) giving the results of measurmeents on loose sand with varying levels of humidity of the air surrounding it. The inset shows a family of characteristics of $\varepsilon''(\omega)$ which clearly consist of two regions of frequency dependence, one being the almost flat universal law with n close to unity and the other showing a very strong dispersion with small values of the exponent. The lower diagram shows that it is possible to normalise the data very satisfactorily and the dotted line gives the locus of the representative point "A" as the humidity varies. The corresponding diagram for the real part is also shown, but it has been displaced because of the confusing overlap that

Figure 5.43 Dielectric measurements on loose sand with varying humidity content. Diagram a) refers to the same sample of sand compressed between planar metallic electrodes with air of variable humidity circulating in the container. The inset shows the family of plots of $\varepsilon''(\omega)$ for a range of humidites indicated as % relative humidity, while the main diagram shows the normalisation of these results with humidity as parameter, together with the locus of the displacement of the reference point "A". The same diagram also shows the corresponding data for the real part $\varepsilon'(\omega)$ normalised with the same displacements as $\varepsilon''(\omega)$ but translated vertically for clarity, only the two slopes being drawn in Kramers–Kronig compatible positions. The high-frequency part shows $n = 0.8$, the low-frequency part $n = 0.22$. Diagram b) gives similar results for a sample at the highest humidity of 95% and the same sample soaked in liquid water. The two diagrams give a very clear contrast, with the humid sand corresponding to a superposition of two parallel mechanisms with the respective values of $n = 0.07$ and a much higher value at high frequencies. The sand–water mixture shows typical series behaviour as illustrated in Figure 3.21 c), with $n \cong 0.5$ at low frequencies.

Original data Shahidi et al (1975), Shahidi (1977), presentation Jonscher (1978b).


log ϵ''_r
94%
54
33
12
0
A °
log f
log ϵ
ϵ'
ϵ''
ϵ
94
54
33
12 %
slope 0.78
log(f/f$_o$)
a


log ε
95% rel. hum.
-0.93
-1.4
sand+water mixture
ε'
b
log(f/Hz)

would result from the fact that the real part is higher than the imaginary part for $n > \frac{1}{2}$ while the opposite is true for $n < \frac{1}{2}$. This example shows very clearly, on the one hand, that the low frequency dispersion forms a natural extension of the high-frequency universal regime and, on the other hand, that normalisation is possible with humidity as well as temperature as the variable parameter. This means that the *rate processes* in both cases are influenced in a similar manner. Diagram b) shows similar data for the same sample of sand placed, respectively, in 95% humidity air and soaked in water. The former is similar to diagram a), with the strongly dispersive region characterised by a value of $n = 0.07$ and an almost perfect parallellism of $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ at low frequencies, giving way to a less dispersive behaviour at higher frequencies for which insufficient data exist to enable one to determine the value of n. The sample soaked in water shows the behaviour characteristic of series combination of a strongly dispersive capacitor with a less dispersive "bulk" behaviour, as illustrated schematically in the low-frequency part of Figure 3.21 c) – our experimental range is insufficient to reveal the presence of a final strongly dispersive region.

The phenomenon of strong low-frequency dispersion is particularly noticeable in ionic conductors and it is shown very clearly in Figure 5.44 giving the dielectric loss for a one-dimensional tunnel structure of the Hollandite type in which potassium ions are relatively freely mobile. At the lowest temperatures shown in this diagram, 77 K, the behaviour is a classic example of the universal type, with a value of the exponent $n = 0.85$, while at the higher temperatures there is clear evidence of a loss peak being superimposed on the general universal trend. At temperatures in excess of 223 K we find the onset of a strong dispersion which is shown in more detail in Figure 5.45 where we see clearly that *both* $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ follow the same trend, thus excluding the possibility that this is simply the effect of the onset of dc conduction. The diagram of Figure 5.45 a) represents the result of normalisation of data for several different temperatures, both the real and the imaginary data being normalised with the same frequency shifts, with the resulting activation energy being closely similar to the dc activation energy. It should be noted that the real and imaginary components have been displaced with respect to one another vertically for clarity. It may also be noted that the real part of $\varepsilon(\omega)$ does not normalise well throughout the entire frequency range, indicating that the two processes are physically quite distinct – the normalisation here has been carried out with particular attention to the low-frequency response.

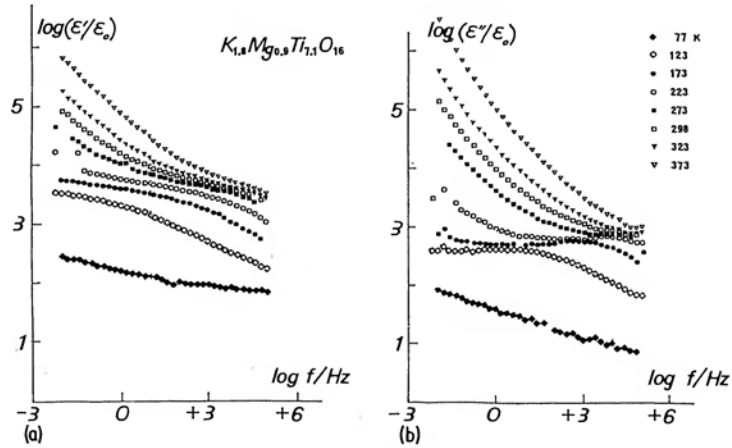


Figure 5.44 The frequency dependence of the dielectric permittivity and loss of the ionic conductor Hollandite of the composition $K_{1.8}Mg_{0.9}Ti_{7.1}O_{16}$ in a range of low temperatures. The loss for 77 K follows a single universal law with $n = 0.85$ over seven decades of frequency, the corresponding ε' data deviate slightly at high frequencies due to the influence of ε_∞. At 123 K a clear loss peak is superimposed on the universal trend, moving with increasing temperature to higher frequencies. At the highest temperatures strong low-frequency dispersion sets in which is shown in more detail in Figure 5.45.

From Jonscher et al (1979a).

Figure 5.46 shows the dispersion in a semiconductor-doped glass which in its pure state has a relatively flat loss. The high temperature of measurement means that electronic and ionic conduction is well developed and the dispersion is seen to extend over seven decades of frequency and of loss. The loss has a constant slope corresponding to the value of $n = 0.050$, within experimental error, over the entire range of measurement and the capacitance behaves consistently with Kramers–Kronig relations.

The unique character of this strong low-frequency dispersion should be clearly borne in mind – it does not represent any simple effect of a *series* barrier, such as has been shown in Figure 3.13 f), which would have given a slope of -2 for the log C'. This unique character is clearly seen in the effect which this type of dispersion has on the complex impedance diagram (Jonscher 1978a) which is not that of a spur that could be attributed to a series capacitance, as in Figure 5.36. The complex impedance diagram for the same sample is shown in Figure 5.45b) where we note the very gradual transition from a strongly inclined arc *which is not circular*, into a very diffuse and strongly inclined spur, without any clear-cut transition between the two.

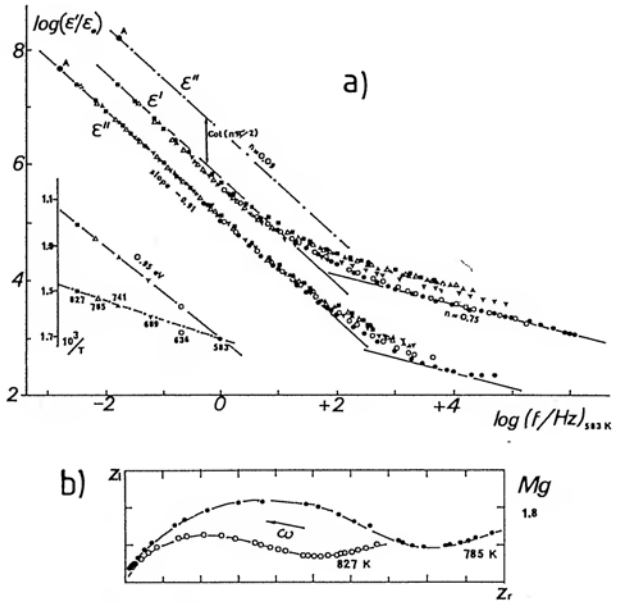


Figure 5.45 The dielectric response of the same Hollandite sample as in Figure 5.44 at high temperatures. Diagram a) shows the normalised data for both $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$, displaced with respect to one another in order to avoid the confusion from crossing-over. The correct position of the loss data with respect to the real part is shown by the chain-dotted line with the reference point "A" indicated in its correct position. There is some discrepancy in the normalisation of the high-frequency data of loss. The inset shows the temperature dependence of the translation frequency, with a simple activation energy of 0.95 eV. Diagram b) shows the complex impedance plots for two temperatures in the strongly dispersive range, showing the effect of strong dispersion on the appearance of the complex impedance. The usual inclined circular arc becomes completely distorted into an almost horizontal line. No "fitting" of a circular arc could credibly be performed in this type of situation. Neither is it possible to assign any dc conductivity to this diagram.

From Jonscher et al (1979a).

We now refer to the point first raised in Section 3.6 and alluded to in the present discussion – the distinction between a strongly dispersive barrier region placed physically *in series* with a less dispersive bulk region, and a homogeneous region in which two physical processes with different values of the exponent n are present *in parallel*. This is shown schematically in Figure 5.47 as two equivalent circuits with the corresponding frequency dependences of their effective complex permittivities. The presence of a distinct "bulge" in the $\varepsilon'(\omega)$ plot is evidence of the series combination, as

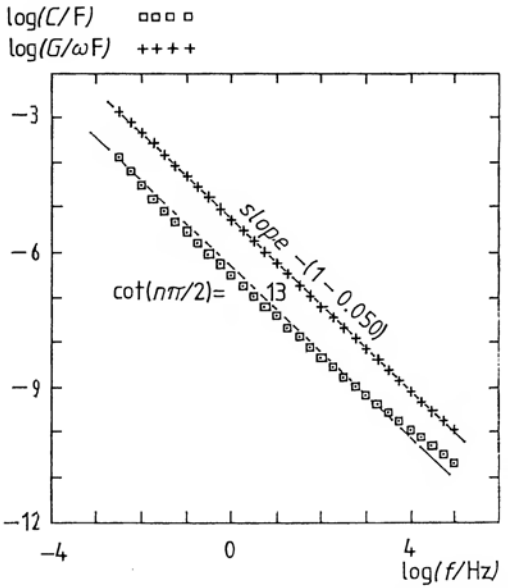


Figure 5.46 Low-frequency dispersion in a semiconducting glass, SiO_2 doped with SnO_2 and Sb_2O_5 intended for use as a conducting glaze. Measurements taken at 800 K. The straight line drawn through the loss data G/ω gives $n = 0.050$ over seven decades of frequency. The line drawn through the C' corresponds to the Kramers–Kronig compatible position with respect to the loss.

From Doyle (1981).

shown in Figure 3.21 c), the parallel plots of $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ being evidence of parallel volume processes. These series and parallel combination are not easily picked out on the impedance and admittance plots, since the shapes of their respective components are strongly distorted, as may be seen in Figure 5.45 b).

It should be remembered that the separate processes with different values of the exponents n have in general different activation energies. This is evident from their different physical character, the strongly dispersive region being closely akin to dc conduction, with substantial charge movements in the material, the high-frequency regime being due to more "dipole-like" localised hopping of charges, if not to molecular dipoles as such. This difference of activation energies may pose problems in normalisation procedures, since the only topographically distinctive point in the spectrum of Figure 5.47 b) is the transition between the two dispersive regions. It is easy to bring consecutive spectra corresponding to different temperatures into coincidence by horizontal translation, but there is no *a priori* reason to assume that this is the correct procedure.

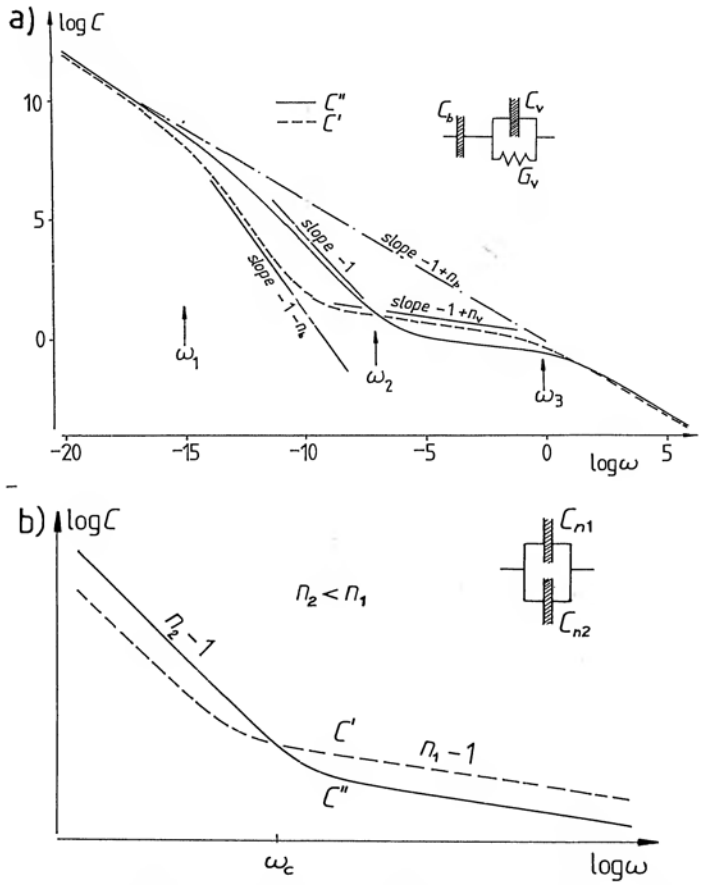


Figure 5.47 The distinction between the series and parallel combination of a strongly dispersive region and a less dispersive process. The series combination a) corresponds to the presence of a physically separate barrier of highly dispersive properties in series with a bulk region with less dispersive character. The parallel combination b) corresponds to the subsistence within a single volume of the material of two differently dispersive mechanisms, of which the more strongly dispersive one dominates at low frequencies, while the less strongly dispersive becomes important at higher frequencies. The transition frequency is denoted by ω_c, in conformity with the analysis.

Moreover, horizontal or approximately horizontal translation yields a much more definite normalisation to steeply rising or falling spectra than to nearly frequency-independent spectra. This may explain the apparent success of a single translational normalisation of such complex spectra as in Figure 5.43 – it does not prove that the two components of the dispersion have the same activation energies or humidity dependence.

Two further points should be made about the strongly dispersive and near-dc behaviour. The first is that normalisation by the shift in *frequency* is manifestly inappropriate when dealing with the dc limit, where the entire change of the spectrum is due to *amplitude* variation of the dc conductivity. However, since the slope of the corresponding loss component G_0/ω is rigorously -1, horizontal translation gives exactly the same logarithmic shift as does the vertical translation, so that the determination of activation energies remains valid, and that is the only physically significant parameter that can be derived from this procedure. It follows, however, that the case of strong low frequency dispersion should be treated similarly as predominantly an amplitude shift, although further research is needed to determine this point with more certainty.

The second point concerns the simultaneous presence of strong low-frequency dispersion and direct current conductivity. The difference between the respective frequency dependences is too small to be detectable in most cases and the only indication that dc conductivity may be present is an abnormally high ratio of loss to real part of permittivity, since dc conductivity does not contribute to $\varepsilon'(\omega)$.

We shall leave a detailed discussion of the interpretation of the physical significance of the strong dispersion until a later chapter, at the present time suffice it to say that the behaviour is very widely observed in many materials, especially at elevated temperatures, examples may be found in Wesphal and Sils (1972). The fact that this behaviour forms a natural limit of the universal behaviour suggests strongly that it should have a similar physical interpretation and this point will be taken up in Chapter 8. In the meantime, it might be observed that the existence of this type of behaviour has been recognised for many decades under the name Maxwell-Wagner phenomena, relating to interfacial polarisations between a bulk sample and the electrodes, or within grains – real or imagined – within the bulk of the material. Such processes are essentially combinations of resistor-capacitor networks (Volger 1960, Haberey and Wijn 1968) and it is difficult to see how they could explain the good power-law relations actually seen in these situations, in terms of accidental combinations of barrier parameters. This point will be taken up in Chapter 7.

What is most remarkable in these results is the very wide range of values of the effective permittivity that is being measured, without any sign of tailing off towards the lowest frequencies to give a

limiting value. This poses some fundamental questions regarding the nature of the prevailing dielectric response. One particular problem is the possible role of interfacial electrochemical reactions – a subject that is far from well understood at present.

One important test should be carried out, wherever possible, with regard to the *linearity* of the dielectric behaviour in terms of the applied electric field. Briefly, most bulk dielectric processes are to a good approximation linear, so that doubling the signal amplitude doubles the current and therefore leaves the capacitance and conductance readings *unchanged*. However, barrier phenomena, of which Figure 3.15 was a good example, tend to be very non-linear and become less important as the applied voltage *per barrier* becomes significant compared with the thermal energy kT/e. This test is easily carried out if there are only two, or very few barriers in series, as in Figure 4.10, but the position becomes difficult when there are large numbers of "grains" with possible interfacial barriers in series, since then it may become difficult to apply a sufficiently high overall voltage to satisfy the thermal energy criterion. One might argue that a material consisting of very many very small grains is macroscopically "homogeneous" and we cannot, therefore, detect its microscopic inhomogeneity by this method.

An alternative test which is often more satisfactory is the repetition of the measurement on another sample with a different nominal thickness. If the processes in question are bulk phenomena, the impedance will rise in proportion to the thickness, if there are dominant interfacial processes, there will be no change in the impedance, and therefore in the capacitance parameters. The difficulty lies in the fact that samples of different thicknesses may not be easily available and this test is therefore difficult to apply in many experimental situations.

A certain way of testing the linearity of the response is to check the applicability of the Kramers–Kronig relations – these would not be expected to apply to non-linear systems. By this test, most of the examples of strong dispersion quoted above were linear to a fair extent. However, there is always the unresolved problem of the *quantitative* consequences of nonlinearity on the applicability of Kramers–Kronig and Fourier transformations, while they should not apply in the *strict* mathematical sense, it is established experimentally that a measure of "tolerance" exists which it is difficult to quantify exactly. As always in the case of interfacial phenomena, extreme caution is called for in the interpretation of the results.

5.7 FREQUENCY-INDEPENDENT LOSS

The opposite extreme of the dielectric response compared with the strong low-frequency dispersion described in Section 5.6 is the limiting case of the exponent n tending toward unity, ie. a virtually frequency-independent loss. While theoretically this limit is forbidden, just as is the case of $n = 0$, there are many experimental situations in which the loss is so weakly dependent on frequency as to be virtually frequency-independent. Just as the strong low-frequency dispersion had not been recognised until recently as an identifiable case in its own right, so neither is the frequency-independent loss recognised as a definite type of dielectric response possessing its own physical significance. This is fully understandable within the framework of the "old" dielectric philosophy, based heavily on the Debye approach, in which the notion of the power law as a universally applicable principle was at best an embarrassing fact, without any theoretical standing in its own right. However, once we accept that the universal power law is not only an *experimental fact*, but also has a plausible *theoretical justification*, the acceptance of the limiting conditions becomes much easier and, indeed, inescapable and it is in this spirit that we propose to treat the following discussion.

We have already shown in Figure 5.24 an example of a very "flat" or frequency-independent loss observed in a high-purity polyethylene at very low temperatures. There are many examples of similar behaviour, at all temperatures but, in general, below the range where significant dc conduction and other charge movements take place. Figure 5.48 shows the loss data for a range of hot pressed ceramics, covering a particularly wide range of frequencies extend-

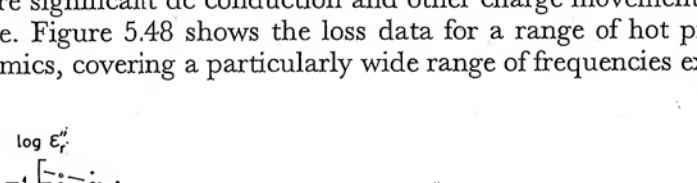


Figure 5.48 The dielectric loss of hot pressed ceramics in a wide range of frequency ● ● ● ● Si_3N_4, ○ ○ ○ ○ ○ 5 wt % MgO/Si_3N_4 and × × × × Sialon $Si_2Al_4O_4N_4$, all at room temperature. A clear tendency to low-frequency dispersion is visible, superimposed on a virtually frequency-independent loss at higher frequencies, including microwave data at 10 GHz.

Based on data of Thorp and Sharif (1977), presentation Jonscher (1980b).

ing up to 10 GHz. Apart from a slight dispersion at lower frequencies which is not sufficiently pronounced to enable one to make a firm interpretation, the general trend is evidently towards a flat response.

A very interesting class of dielectric systems is represented by *Langmuir films* of long-chain molecules, typically fatty acids, which may be transferred as coherent monolayers from a water surface on to solid substrates. A recent series of reviews may be found in a special issue on Langmuir Films of Thin Solid Films (Barlow, 1980).

Each monolayer is approximately 2.5 nm thick – for the case of stearic acid which has a carbon chain of 18 carbon atoms – and films may be built up from monolayer thickness to 40, and more layers, thus ranging from thicknesses in which electron tunnelling is the dominant mechanism of transport, to virtually "bulk" samples. Langmuir films have good dielectric properties and they are particularly suitable for the study of the effect of high electric fields, since small voltages are sufficient to produce such fields in relatively very thin samples.

Detailed dielectric measurements exist on Langmuir films covering a wide range of film thicknesses, frequencies and temperatures. By way of example, Figure 5.49 gives the dielectric loss data for a 9-layer stearic acid film – chemical formula $CH_3(CH_2)_{16}COOH$ – between aluminium electrodes. Two distinct regions of response are visible – at the lower frequencies the response is of the "universal" type with values of the exponent equal to 0.6–0.7, while at the higher frequencies the response becomes virtually flat. It is also significant that the temperature response in these two regions is quite different, as may be seen by the crossing over of the various curves.

A more detailed impression of the temperature dependence of loss at two fixed frequencies is shown in Figure 5.49 b) where three temperature peaks are discernible and these may be correlated with known phase transformations in the stearic acid lattice. It is also interesting to note that at least one of these temperature peaks is virtually independent of frequency. In order to investigate this phenomenon further, a detailed study was made of the loss in a different sample over a more restricted frequency range but using a more sensitive measuring instrument. The results of this study are also shown in Figure 5.49 c) where the frequency plot shows an intricate pattern of intersecting curves without apparent order.

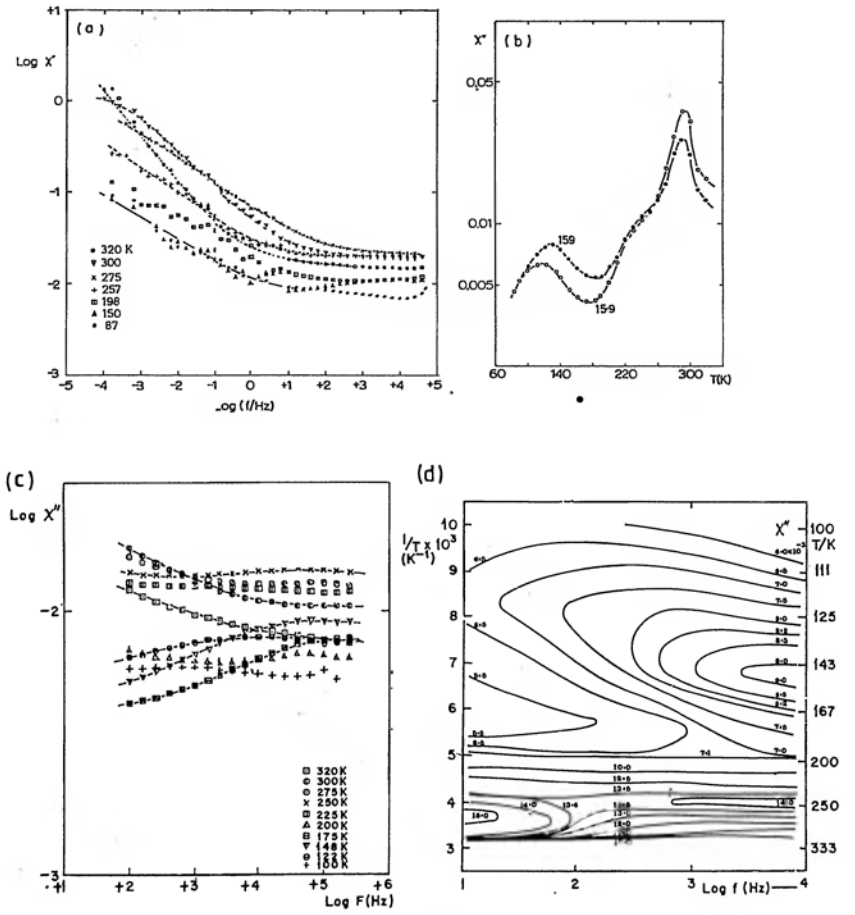


Figure 5.49 Dielectric loss data for multilayer stearic acid films between Al electrodes.

a) Loss data for a 9-layer film over a range of temperatures showing two distinct ranges of frequency- and temperature-dependence.
b) The temperature dependence of this loss at two constant frequencies of 15.9 and 159 Hz, both in the frequency-independent region of diagram a), showing three loss peaks, one of which at least is independent of frequency.
c) More detailed loss measurements on a 13-layer sample between 100 and 320 K. Note the extended loss scale, showing that the loss is almost independent of frequency.
d) The same data represented as contours of constant loss against log $\omega/2\pi$ and $1/kT$, showing the presence of two loss peaks, one at the higher temperatures which is independent of frequency and a lower-temperature one moving in a manner similar to the activated peak in Figure 3.36 a).

From Careem and Jonscher (1977).

However, a much more meaningful picture emerges when a contour map is drawn of constant loss against log ω and $1/T$, where it is evident that there are two loss peaks, one virtually independent of frequency – it runs substantially parallel to the frequency axis, while the other crosses the frequency-temperature plane diagonally.

This question of temperature dependence is of considerable significance for the evaluation of experimental data. In the classical Debye system the relaxation time τ is given by the relation (4.29), and the susceptibility is a function only of the product $\omega\tau$, cf. eqn (4.30):

$$\chi(\omega) = \chi(\omega\tau) = \chi\big(\exp(\log \omega + W/kT)\big) \qquad (5.7)$$

so that contours of constant loss correspond to $\log \omega + W/kT =$ const. The three-dimensional plot of loss against temperature and frequency is therefore as shown schematically in Fig. 3.26 a) where it is clear that the contour along the line of constant W/kT with log ω as variable is *identical with* a contour along a line of constant frequency with W/kT as variable.

Provided, therefore, that the dielectric material in question obeys the Debye characteristics, it is equally valid to take the data as functions of frequency with temperature as a constant parameter, or with variable temperature at a constant frequency. Since measurements at constant frequency are often easier than measurements at constant temperature, requiring much less specialised electronic equipment, it has become customary among workers in the field of dielectrics to take only the contours along constant frequency, with temperature as a variable. If, as is often the case, these variable temperature data showed a well developed loss peak, the inference was made, explicitly or implicitly, that the material follows the Debye characteristics.

The results of Figure 5.49 show conclusively that this is a completely unjustified assumption, in that many, perhaps the majority of dielectric systems obey completely different temperature/frequency relations which may be represented schematically by the diagram of Figure 3.36 b) which corresponds directly to the lower part of the contour map in Figure 5.49 where the entire temperature peak runs parallel to the frequency axis. We may note at this juncture that the behaviour of ferro-electrics shown in Figures 5.34 and 5.35 represents a very similar aspect of frequency-temperature dependence – the temperature dependence is now in the form of a singularity at the Curie temperature, while the loss is almost flat in frequency, corresponding to the universal law.

Returning now to the data shown in Figure 5.49 a) we note that the two separate regions of frequency dependence characterise two different physical mechanisms, of which the low-frequency one may be regarded as being primarily associated with charge carriers present in the material, giving values of the exponent in the typical range $0.6 < n < 0.8$, while the high-frequency regime is more closely linked with the properties of the *dielectric lattice* and is typified by a nearly frequency-independent loss. We shall have further occasions to note this type of behaviour.

The frequency-independent loss corresponding to values of the exponent n close to unity and hence, through eqn (3.31), to *low values* of tan δ. For this reason one expects to find this behaviour especially in low-loss materials and an example of this is shown in Figure 5.50 giving data for several samples of low-loss polyethylene

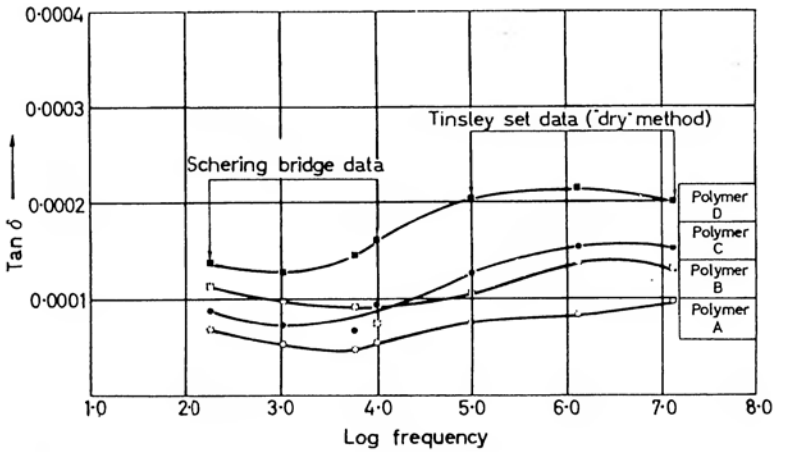


Figure 5.50 Examples of almost frequency-independent losses in a range of polyethylenes at room temperature, covering five decades of frequency in the audio and radio ranges. The loss scale is linear and the variation over the frequency range does not exceed a factor of two.

From Reddish (1962).

over five decades of frequency. Polyethylene is a non-polar material and loss peaks in it are due to additives and impurities. These data are plotted on a linear scale, so that we conclude that loss changes by less than a factor of two in five decades of frequency, which is very flat indeed.

As against this, we find that the temperature dependence of loss is relatively much stronger as may be seen in Figure 5.51, also for

polyethylene, down to 10 K. Here the data cover two decades of frequency and in this range loss changes by up to a factor of three, while the temperature dependence covers a factor of ten.

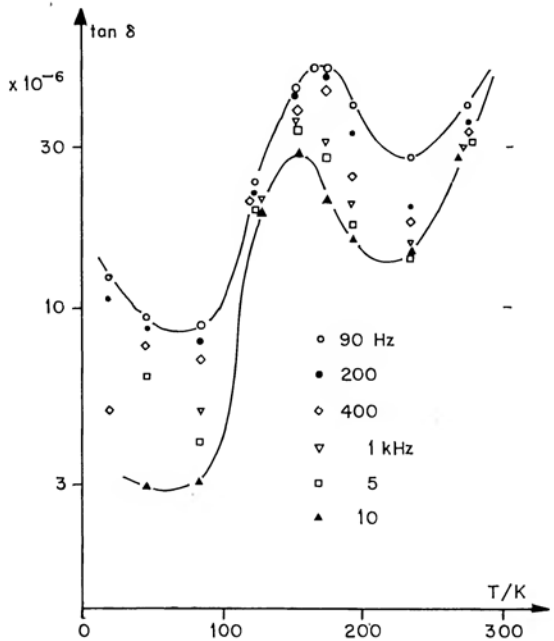


Figure 5.51 The temperature dependence of dielectric loss between 10 and 300 K for various frequencies of polyethylene with antioxidant, showing the relatively much stronger variation with temperature than with frequency at a constant temperature.

Unpublished data, by courtesy of the Director, Laboratoire de Genie Electrique de Toulouse, CNRS, Universite Paul Sabatier Toulouse.

A general feature of the "flat" dielectric loss in frequency is that it is also independent of temperature to a large degree, except where the temperature changes the structure of the material, as in the case shown in Figures 5.49 and 5.51. This remains true even when the loss is not completely flat, with n around 0.9. One rather extreme example of this type of behaviour is shown by the ionic conductor hollandite already mentioned earlier. Figure 5.52 shows the data taken in the temperature range extending down to 5.4 K, showing that the response becomes virtually temperature independent below 77 K and this remains true even for the strongly dispersive region at very low frequencies.

It is relevant to point out the experimentally observed fact that all the available evidence regarding the lowest loss materials, of the order of a few microradians loss angle, points invariably to the existence of an essentially flat loss in frequency, rather than any

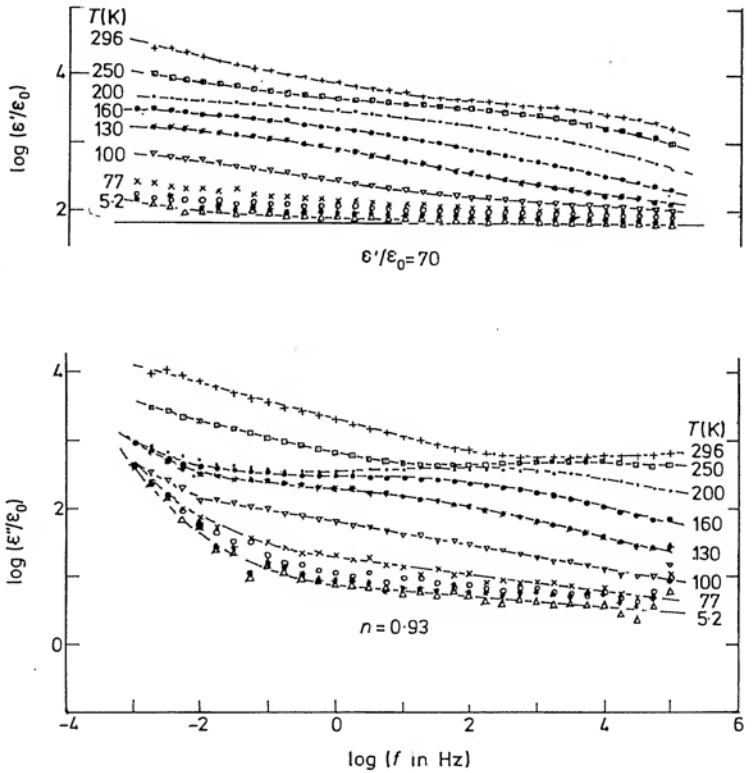


Figure 5.52 The frequency dependence of the complex dielectric permittivity of the ionic conductor of the hollandite family $K_{1.6}Mg_{0.8}Ti_{7.1}O_{16}$ in the log–log representation, with temperature as parameter. The temperature range overlaps that in Figure 5.44 but extends down to 5.2 K. The symbols for the lowest temperatures are × 77 K, ○ 60 K, * 30 K, △ 5.2 K. The upper diagram represents the real part, with the horizontal line corresponding to the value 70, the lower diagram represents the imaginary part.

From Deori and Jonscher (1979). © The Institute of Physics.

particular loss peaks. Any loss peaks due to the presence of what we might call "extrinsic" impurities are invariably superimposed on an irreducible flat spectrum of loss which may be presumed to be due to the "intrinsic" lattice dipoles or charges.

5.8 SUPERPOSITION OF DIFFERENT MECHANISMS

We conclude our presentation of the experimental evidence on frequency domain behaviour of dielectric materials with a brief discussion of the effects of superposition of different dielectric mechanisms arising from the presence in one sample of more than one process with overlapping activity in frequency. Many examples of such overlap may be quoted and we propose to mention here three cases – the superposition of the "pure lattice" response with, respectively, dipolar species added to the material, and ionic or electronic charges injected into the system. The dipolar admixtures are most readily introduced by suitable diffusion, since dipoles cannot be drifted in an electric field, and the most common impurity here is moisture which can also dissociate into positive and negative ions.

The effect of the addition and subsequent removal of water in polyethylene is shown in Figure 5.53 – water is seen to introduce a strong loss peak whose width only slightly exceeds the Debye width. The underlying trend is an almost frequency-independent loss which is characteristic of pure polyethylene.

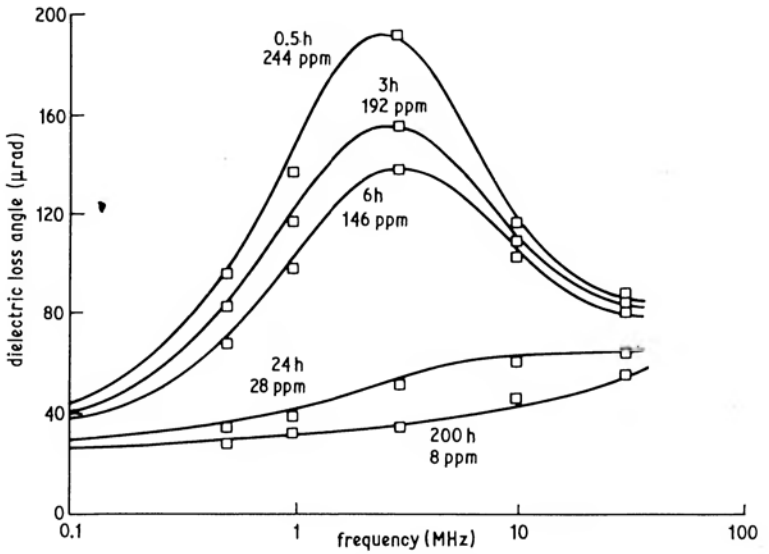


Figure 5.53 The effect on dielectric loss of successive drying of a sample of polyethylene quenched in water at 95°C, with drying times and water content indicated. The large initial peak due to water molecules disappears, leaving behind a very flat loss at the level of 20–40 μ radians.

From Ayers (1979).

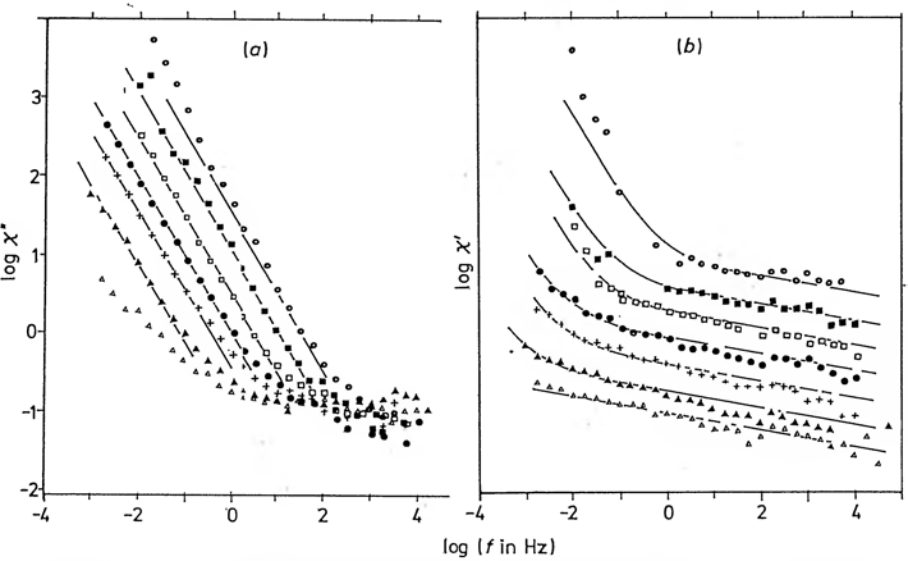


Figure 5.54 The frequency dependence of the imaginary part (*a*) and the real part (*b*) of the dielectric loss of a 15-layer film of stearic acid sandwiched between two aluminium electrodes, measured at 343 K, for applied signals of amplitude 2.0×10^8 (○), 1.3×10^8 (■), 1.2×10^8 (□), 8×10^7 (●), 6×10^7 (+), 1.6×10^7 (▲) and 2.6×10^6 (△) V m^{-1}. The straight lines in (*a*) and the ultimate low-frequency curve in (*b*) have the slope -1. The loss data were corrected for series resistance effects while the $\chi'(\omega)$ data were obtained by subtracting the same value of ε_∞ from all the data measured experimentally. The $\chi'(\omega)$ curves in (*b*) are displaced with respect to one another by $\frac{1}{3}$ of a decade to avoid overcrowding and the straight lines are drawn in corresponding positions, to show that there is no significant effect of electron injection on the "high-frequency" response.

From Jonscher et al (1979b). © The Institute of Physics.

Ionic and electronic charge carriers are very easily introduced into dielectrics by means of injection from electrodes or, in the case of electrons, by direct electron bombardment. The current understanding of contact charging of insulators is still relatively rudimentary (Rose–Innes 1981), but it is evident that contact potential differences between the electrodes and the bulk dielectric inevitably lead to electron injection Davies (1969) much as in the case of metal-semiconductor contacts. The important difference in comparison with the latter is the absence of high densities of donor and acceptor impurities in typical dielectrics.

Ionic carriers, on the other hand, are more readily removed from solids than introduced into them in view of the difficulty with which they traverse interfacial barriers.

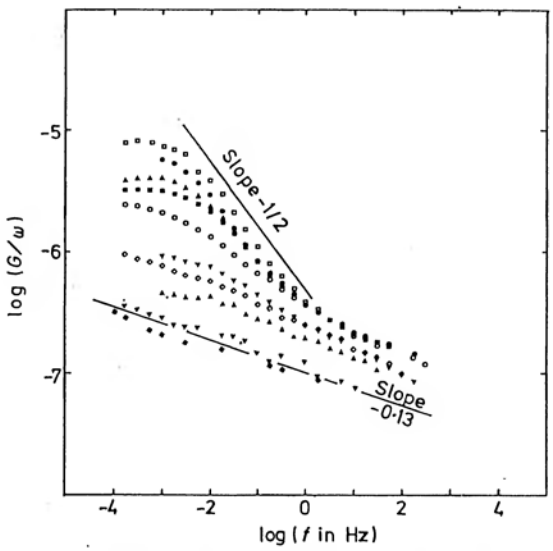


Figure 5.55 The frequency dependence of the dielectric loss, expressed as $G/2\pi f$, of a sintered tantalum electrolytic capacitor measured under an alternating voltage signal of 0.2 V and with a steady bias of 0 (□), 0.10 (●), 0.50 (△), 0.75 (■), 1.0 (○), 1.5 (▼), 2.0 (◇), 3.0 (▲), 5.0 (▽) and 9.5 (◆) V. As the steady bias increases, so the dielectric loss at low frequencies decreases rapidly to reach a steady value following the universal law with the exponent $n = 0.87$.

From Jonscher et al (1979b). © The Institute of Physics.

Examples of the opposite effects on dielectric loss and polarisation of the injection of electrons and of the removal of ions are shown in Figures 5.54 and 5.55. The former relates to stearic acid multilayers and shows the rapid increase of the low-frequency dispersion as the signal amplitude on the film is increased, while the latter shows the rapid drop of loss in a tantalum oxide electrolytic capacitor as the bias is increased. In both cases the ultimate trend is to a simple power law frequency dependence with a relatively small slope – exponent n not very different from unity.

Our last example shows how different superimposed polarisation processes may be separated by subtracting data taken over a range of frequencies at different temperatures. This technique can only work, of course, if some of the overlapping mechanisms have a weaker dependence of temperature than others, since only then is it possible to separate out the temperature-dependent element. Figure 5.56 gives the results of such measurements on a multilayer film of stearic acid, taken at 91, 293 and 325 K. Apart from a series resistance limitation at high frequencies, there is some evidence of

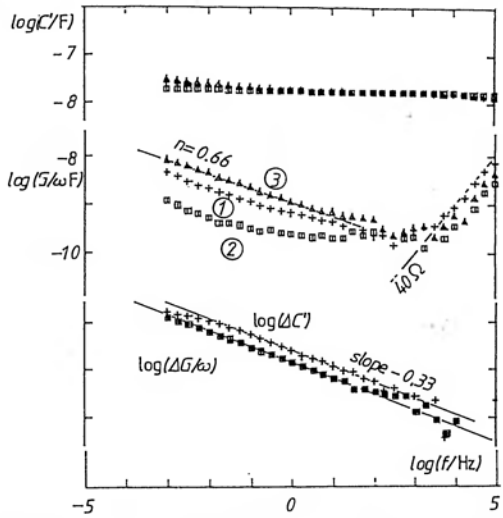


Figure 5.56 The dielectric response of a three-layer stearic acid film between aluminium electrodes at three temperatures: 1–293 K, 2–91 K and 3–325 K. The upper diagram shows the capacitance $C'(\omega)$, the middle diagram the loss $G(\omega)/\omega$ and the lower diagram gives the incremental complex capacitance $\Delta C'(\omega)$ and $\Delta G(\omega)/\omega$ obtained by subtracting the 91 K data from the 325 K data. The two straight lines are drawn in the Kramers–Kronig compatible ratio, showing agreement over seven decades of frequency, even though the original data show some complicating features.

From Millany and Jonscher (1980).

a low-frequency dispersion which manifests itself by a power-law dependence with $n = 0.66$, but on the whole the trends are not clearly defined. However, by subtracting the data taken at 91 K from those taken at 325 K one obtains a complete elimination of all other trends which do not apparently depend on temperature, i.e. the series resistance and the "flat" low-temperature loss, and we are left with the pure power-law with $n = 0.66$ for both the real and the imaginary components of incremental susceptibility, extending now over seven decades of frequency and with the correct Kramers–Kronig ratio.

This technique promises to be very effective, especially where measurements are being collected in an automated fashion and are available on tape or other digital output.

We complete this review of the frequency–domain information with an example of the dielectric relaxation spectrum of Poly-vinylidene difluoride (PVDF) taken over the temperature range 213–373 K and shown in Figure 5.57 as a normalised graph of loss. We have

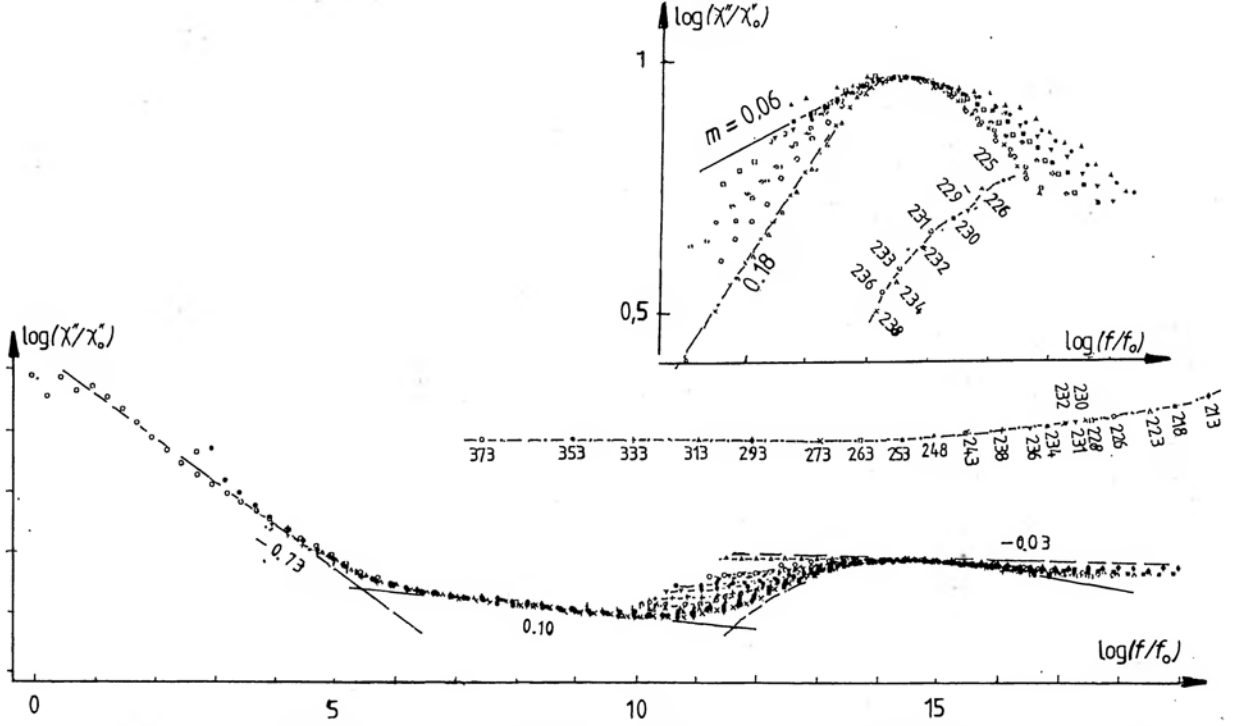


Figure 5.57 An extended relaxation spectrum of loss in polyvinylidene difluoride, measured in a wide range of temperatures and normalised over an effective frequency range of 20 decades. The spectrum shows all the relaxation types discussed in the present Chapter, from strong low-frequency dispersion, through loss peaks with variable slopes to virtually flat loss at low temperatures and high frequencies. The detailed structure of the peak region is shown in the inset on an expanded ordinate scale. Temperatures in Kelvin are indicated on the locus of the representative point. Note the very narrow temperature range over which the peak shape is changing rapidly.

From measurements by Bodakian (1982).

chosen this example because it shows in one material the entire range of differen: forms of spectral response reviewed so far. At the highest temperatures and lowest frequencies there is clear evidence of strong low-frequency dispersion, with the value of $n = 0.27$ over more than four decades of frequency. Our data are inconclusive with regard to the presence of a loss peak at the lowest frequencies, although other similar measurements failed to show the presence of a loss peak. There then follows a region of very low dispersion with $n = 0.90$ which goes over gradually into a distinctive loss peak whose leading slope varies between almost zero and 0.18, as may be seen more clearly in the inset where the scale of ordinates has been considerably expanded. There is likewise a definite variation of the high-frequency slope, but the overall value of $n = 0.97$ appears to be the limiting low-temperature feature. The rise of the displacement locus at lowest temperatures indicates that the loss peak amplitude has been decreasing in this range. In this case it appears that the loss peak becomes so shallow that it ultimately gives way to a virtually flat loss. It should be noted that the entire frequency spectrum comprises twenty decades of frequency.

5.9 SURVEY OF FREQUENCY RESPONSE INFORMATION

The extensive review of the frequency response data for a very wide range of materials given in the present Chapter was intended to give the reader a good grasp of the *actual performance of real dielectrics*, as opposed to the response expected from idealised models of dielectric systems which was presented in Chapter 4. It must be clear by now that the actually observed response bears very little relation to the idealised models which lead one to expect little else than the ideal Debye or the diffusive law which looks like the universal law for the special case of $n = \frac{1}{2}$.

With such extensive departures from some idealised model response, the question may naturally arise whether the model is at all applicable, or whether it is only applicable to certain limiting situations, while the great majority of real-life situations are governed by some entirely different models. The main message of the present work is that this is indeed the case, that a completely fresh approach should be made to the interpretation of the dielectric response of solids, because the old-established approaches are incapable of explaining any but the very few cases which approximate to the ideal response.

Our approach to the theoretical interpretation of the observed data will be made much easier if we can form an overall picture of the real response of materials and, in particular, if we can deduce from this picture certain general rules of behaviour, because any such general rules would strengthen our conviction that we had a valid model. We therefore propose to summarise the results of our survey of dielectric properties with a view to deriving a general pattern of behaviour.

Our summary is presented in Figure 5.58 which gives a schematic represenation of the various response types, both as plots of $\log \chi'$ and $\log \chi''$ against $\log \omega$ and also as plots of the complex χ. On the right is the – practically non-existent – ideal Debye response, and then moving towards the left are the symmetric, the slightly asymmetric and the strongly asymmetric broad peaks, all relating to the dipolar materials. We list some examples of systems which exhibit the particular types of response. Still further to the left we have the response of hopping charge carrier systems which is characterised

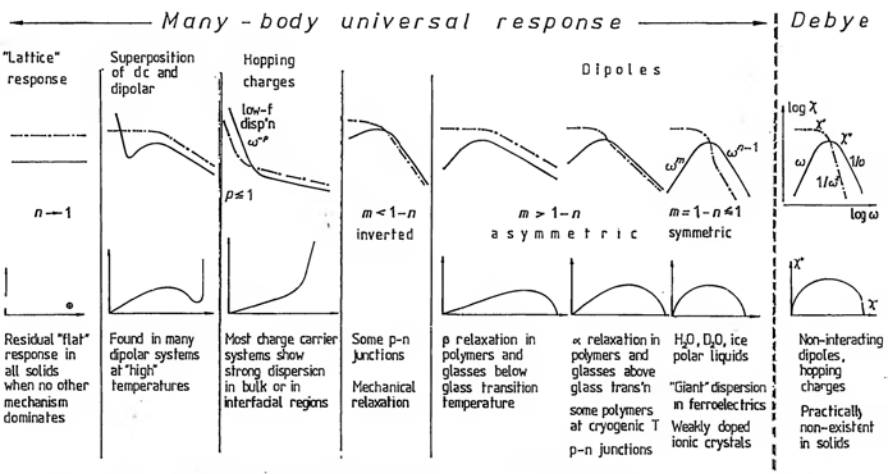


Figure 5.58 The general classification of all types of dielectric responses found in solids. The upper row gives the diagrammatic representation of the $\log \chi'(\omega)$ (chain-dotted line) and $\log \chi''(\omega)$ (solid line) against $\log \omega$, the lower row gives the corresponding complex susceptibility plots. Typical materials giving the various types of response are indicated. The extreme right gives the practically non-existent case of the Debye response, moving to the left we find increasingly broader loss peaks for dipolar systems, further to the left the charge carrier responses corresponding to the strong low-frequency dispersion and to dc conductivity. On the extreme left is the limiting case of "flat" frequency- and temperature-independent loss.

From Jonscher (1980b).

by the absence of loss peaks and which may show a very strong dispersion at low frequencies. This low-frequency dispersion follows the same universal law as the higher-frequency response, but with a much smaller value of the exponent n. There is, finally, the other extreme of response in the form of a practically frequency independent χ'' and χ' which we have identified with the residual response of dielectric lattices.

We stress the fact that this is the first *complete* classification of *all experimentally observed* types of dielectric response in the frequency range below the microwave region of the spectrum and extending down to the lowest attainable frequencies of 10^{-4}–10^{-6} Hz. It takes into account not only the generally accepted dipolar response, but also the equally important charge carrier responses of both hopping electronic and ionic nature. This extension of the range of the polarising species might not be acceptable to some dielectrics specialists who tend to regard with suspicion any charge carriers as allegedly contributing only to dc conduction. We have proved experimentally that this view is unduly restrictive and that hopping charge carriers do give rise to genuinely dielectric responses, as we have defined them in Chapter 2. However, the final justification for the inclusion of charge carriers among the polarising species will be found in the experimental observation that they follow *the same law of frequency dependence* as the classical dipolar species above the loss peak frequency.

We are now able to state that, above the loss peak frequency, where such exists, *all dielectric systems obey the universal law*;

$$\chi'(\omega) \propto \chi''(\omega) \propto \omega^{n-1} \tag{5.8}$$

with the exponent in the range

$$0 < n < 1 \tag{5.9}$$

and with the consequence that the real and imaginary parts of the susceptibility are in a frequency-independent ratio:

$$\chi''(\omega)/\chi'(\omega) = \cot(n\pi/2) \tag{5.10}$$

which has the very simple physical significance that the ratio of

$$\frac{\text{energy lost per radian}}{\text{energy stored}} = \cot(n\pi/2) \tag{5.11}$$

The universal law might otherwise appear completely arbitrary and it would probably not have been selected as a very obvious choice, except that, as we have seen in many instances, it is *extremely well obeyed* experimentally – so well, in fact, that there can be no doubt that it represents the best empirical approximation to the true response. This universal law also has two unique features, not possessed by any other functional relation. The first of these is the constancy of the ratio (5.10) which is the consequence of the fact that the universal law is the only functions that remains invariant, except for a constant, under Kramers–Kronig transformation. The second will be discussed in Chapter 8 in the context of the time-domain response of many-body systems when it will be made clear that this particular relation is, in fact, *the only possible relation to be obeyed by interactive many-body systems.*

The range of validity of the universal relation extends from frequencies at which the quantum and phonon effects become dominant, which means in practice below the infra-red region of the spectrum, down to either the loss peak frequency in dipolar systems or to the onset of the strong low-frequency dispersion in charge-carrier-dominated systems. Subject to this limitation of its range, the universal relation is obeyed under extremely general conditions, *viz.*

> *in all physical structures: single crystal, polycrystalline, amorphous and glassy;*
> *all types of chemical bonding: covalent, ionic, molecular;*
> *all types of systems: inorganic, organic, biological*
> *all possible polarising species: dipoles, hopping electrons, polarons, ions;*
> *all geometrical configurations: from bulk to narrow interfacial regions, molecular thickness films, planar and intricate geometries, continuous and discontinuous media.*

The universal law also covers the entire range of the exponent n which is mathematically allowed for this type of relation to be valid as a Kramers–Kronig-compatible system. The discovery of the two extreme values of n corresponding, respectively, to the strong low-frequency dispersion and to the frequency-independent loss has thus completed the classification to coincide with what will be shown in Chapter 8 to be the theoretically admissible range.

Dipolar systems obey the same universal law at frequencies above the loss peak frequency ω_p and their response below this frequency is governed by a second power law, corresponding to the relation

(5.3):

$$\left.\begin{aligned} \chi''(\omega) &\propto \omega^m \\ \chi'(\omega) &= \chi(0) - \frac{1-n}{m}(\omega/\omega_p)^m \cos(n\pi/2) \\ &= \chi(0) - \text{const.}\ \chi''(\omega) \end{aligned}\right\} 0 < m < 1 \qquad (5.12)$$

The discovery of this second power law valid below the loss peak frequency will be found later to have been very important in the development of the new theory of the dielectric response based on the many-body interactions.

We thus have a complete picture of the experimental situation relating to the dielectric response of the complete range of solids and this picture is subsumed by two power laws, with either positive or negative exponents, and by a thermally activated transition frequency between the respective regions of applicability of these power laws.

We have also seen that the effect of temperature, to a first approximation, is to translate the entire frequency response bodily in the direction of increasing frequency with rising temperature – the shape of the frequency response being relatively less strongly affected. The behaviour at very low temperatures represents a special case, the most surprising feature of which is that there remains a remarkably high level of loss, or of dielectric "activity" down to the lowest temperatures at which any dc conductivity would have been completely frozen out.

REFERENCES TO CHAPTER 5

Abkovitz M, Le Comber P G and Spear W E 1976, *Communications in Physics* **1**, 175

Abkovitz M, Blossey D F and Lakatos A I 1975, Phys. Rev. **B12**, 3400

Abkovitz M, Lakatos A I and Scher H 1974, Phys. Rev. **B9**, 1813

Ayers S 1979, Proc. IEE 126, 711

Bodakian V 1982, Ph.D. Thesis, University of London, Chelsea College

Barlow W A (Guest Ed) 1980, Thin Solid Films Vol 68 *'Langmuir — Blodgett Films'*

Careem M A and Jonscher A K 1977, Phil. Mag. **35**, 1489

Careem M A, Jonscher A K and Taiedy F 1977, Phil. Mag. **35**, No. 6, 1503–1508

Carson R A 1973, Proc. Roy. Soc. London **A332**, 255–268

Charoensiriwatana V 1982, Ph.D. Thesis, University of London, Chelsea College

Deguchi K, Okaue E and Nakamura E 1982, J Phys Soc Japan **51**, 3569

Davies D K 1969, Brit. J. Appl. Phys. **2**, 1533

Deori K L and Jonscher A K 1979, J. Phys. C: Solid State Physics **12**, L289

Deori K L, Leech A G and Jonscher A K 1983, to be published

Dissado L A and Hill R M 1978, Nature **279**, 685

—— 1980, Phil Mag **B41**, 625

Doyle B 1981, Ph.D. Thesis, University of London, Chelsea College
Druon C and Wacrenier J M 1977, J. de Physique **38**, 47
Dryden J S 1957, J. Chem. Phys. **26**, 604
Favaron J 1982, Ph.D. Thesis, University of London, Chelsea College
Fontanella J, Jones D L and Andeen C 1978, Phys. Rev. **B18**, 4454
Frost M S and Jonscher A K 1975, Thin Solid Films, **29**, No. 1, 7–18
Fuoss R M and Kirkwood J G 1941, J. Amer. Chem. Soc. **63**, 385
Gesi K 1972, Jap. J. Appl. Phys. **11**, 1745
Gilchrist J le G 1977, J. Phys. Chem. Solids, **38**, 509–516
—— 1979, private communication
Gough S R, Hawkins R E, Morris B and Davidson D W 1973, J. Phys. Chem. **77**, 2969
Grant R J, Hodge I M, Ingram M D and West A R 1977, Nature **266**, 42
Haberey F and Wijn H P J 1968, phys. stat. sol. **26**, 231
Hill R M 1981a, J. Materials. Science **16**, 118–124
—— 1981b, phys. stat. sol. (b) **103**, 319
Hooper A 1977, J. Phys. D: Applied Physics **10**, 1487
Hughes D and Pethig R 1975, *Dielectric Materials, Measurements and Applications*, IEE Publication 129, 52, IEE London
Ishida Y 1969, J. Polymer Sci. A-2, **7**, 1835
Ishida Y, Matsuo M and Yamafuji K 1962, Kolloid Z.u.Z.f.Polymere **180**, 108
Ishida Y and Matsuoka S 1965, American Chemical Soc., Polymer Preprints 6, No. 2, 795–798
Ishida Y and Yamafuji K 1961, Kolloid Z.u.Z.f.Polymere **177**, 97–116
Isnard R and Gilchrist J le G 1980, Chemical Physics **52**, 405
Johari G P 1976, Annals. N.Y. Acad. Sci. **279**, 117
Johari G P and Dannhauser W 1969, J. Chem. Phys. **50**, 1862
Johari G P and Jones S J 1978, J. Glaciology, **21**, No. 85
Jonscher A K 1975a, Colloid and Polymer Sci. **253**, 231
—— 1975b, phys. stat. sol. (a) **32**, 665
—— 1977, Nature **267**, No. 5613, 673–79
—— 1978a, J. Materials Science **13**, 553–562
—— 1978b, Phil. Mag. B 38, 587
—— 1980a, J. Phys. D. **13**, L89
—— 1980b, *Physics of Thin Films*, M. Francombe (Ed.), Vol. 11, Academic Press, London and New York, p. 205
—— 1981, J. materials Sci. **16**, 2037
Jonscher A K, Deori K L, Reau J M and Moali J 1979a, J. Materials Science **14**, 1308
Jonscher A K and Dube D C 1978, Ferroelectrics **17**, 533
Jonscher A K and Frost M S 1976, Thin Solid Films **37**, 267–273
Jonscher A K, Meca F and Millany M H 1979b, J. Phys. C: Solid State Physics **12**, L293
Jonscher A K and Reau J M 1978, J. Materials Science **13**, 563
Kizilyalli H M and Mason P R 1976, phys. stat. sol. (a) **36**, 499
Lakatos A I and Abkovitz M 1971 Phys. Rev. B. **3**, 1791
Loh C K 1978, private communication
Mason P R, Hasted J B and Moore L 1974, Adv. Mol. Rel. Proc. **6**, 217
Matsuo M, Ishida Y, Yamafuji K, Takayanagi M and Irie F 1965, Kolloid Z.u.Z.f. Polymere **201**, 89–93
Meakins R J 1956, Trans. Far. Soc. **52**, 320
Millany M H and Jonscher A K 1980, Thin Solid Films **68**, 257–273

Murawski L and Gzowski O 1974, phys. stat. sol. (a) **24**, K115
Pawlaczyk C 1977, Fizyka Diel. i Radiospektr. IX, 53 (in Polish), PWN Poznan
Phillips W A 1970, Proc. Roy. Soc. London A319, 565–581
—— 1981 *Physics of Dielectric Solids*, C H L Goodman (Ed.), Institute of Physics Conference Series No. 58, p. 64
Pollak M and Geballe T H 1961, Phys. Rev. **122**, 1745
Réau J M, Claverie J, Campet G, Deportes C, Ravaine D, Souquet J L and Hammou A 1975, C.R. Acad. Sci. Paris **280**, 325
Reddish W 1962, Pure and Applied Chemistry **5**, 723
Roberts G G and Polanco J I 1972, Solid State Commun. **10**, 709
Rose Innes C 1981, *Physics of Dielectric Solids*, C H L Goodman (Ed.), Institute of Physics Conference Series No. 58, p. 122
Sayer M, Mansingh M, Reeves J M and Rosenblatt C J 1971, J. Appl. Phys. **42**, 2857
Shahidi M, Hasted J B and Jonscher A K 1975, nature **258**, 595
Shahidi M 1977, private communication
Stoll B, Pechhold W and Blasenbrey S 1972, Kolloid Z.u.Z.f. Polymere **250**, 1111
Strom U and Taylor P C 1974, *Amorphous and Liquid Semiconductors*, J Stuke and W Brenig (Eds.), 375, Taylor and Francis, London
Thomas R A and King C N 1975, Appl. Phys. Letters **26**, 406
Thorp J S and Sharif R I 1977, J. Materials Science **12**, 2274
Vashishta P, Mundy J N and Shenoy G K (Eds.) 1979, *Fast Ion Transport in Solids*, Elsevier, N. Holland
Volger J 1960, *Progress in Semiconductors*, A F Gibson (Ed.) **4**, 205
Westphal W B and Sils A 1972, *Dielectric Constant and Loss Data*, MIT Technical Report, AFML-TR-72-39 April
Williams G and Watts D C 1970, Trans. Far. Soc. **66**, 80

CHAPTER 6

Experimental Evidence on the Time Response

6.1 THE ROLE OF TIME–DOMAIN MEASUREMENTS

Chapter 2 introduced the notion of the essential equivalence of time– and frequency–domain responses in *linear systems*, where it was shown that the time-dependence of current in response to a step–function field and the frequency dependence of the dielectric susceptibility in response to sinusoidal excitation were connected by the Fourier integral transformation. To that extent, therefore, there is no fundamental difference between the two types of measurements so far as linear dielectric systems are concerned and the choice between them must rest on two types of considerations:

i) the convenience of the measuring process, and
ii) the study of departures from linearity in the dielectric response of the materials in question.

On the first point, it should be noted that, in principle, time–domain measurements have the advantage of requiring *one single sweep* to determine the entire corresponding frequency range which might, for its part, require a large number of spot measurements in the frequency spectrum. The argument runs therefore, that it is more "economical" on time to measure in the time domain and to transform into the frequency domain. There are, however, some very serious drawbacks which to a large extent invalidate this simple reasoning.

The first of these is the availability of equipment which is well developed commercially for the frequency domain measurements in the form of an extensive range of alternating current (ac) bridges covering a wide span of frequencies from the sub-audio, of the order of 1 Hz, to the microwave, 10–30 GHz, and extending even further into the terahertz region in the form of open resonators. The fundamental advantage of these frequency domain techniques lies in the fact that they all work on a suitably narrow band of frequency

at any single measurement, by using tuned circuits or otherwise synchronous detection over many cycles, with the effect that the signal-to-noise-ratio is very substantially improved in comparison with wide band detection. The result of this is that measurements of dielectric loss angle – the most significant criterion of measurement – can be made easily at the level of 100 μ radians, they are possible with advanced methods down to 10 μ rad and can be made under extreme conditions down to very few μ rad.

This level of sensitivty is not achievable in the time domain, for the simple reason that the measurement is made under wide band conditions, since in the very nature of the experiment one has to determine the response of a dielectric system over a very wide range of time. The practical implication of this is that one is severely limited in the available signal-to-noise ratio of the system. A possible remedy is to go back to the principle of repeated scans, by accumulating data over a number of "shots" and averaging, which can be done easily with modern electronic signal-processing equipment, but that immediately vitiates the alleged advantage of "speed" of measurement. An even more precise but much more time-wasteful method would be to employ some "box-car-detector" principle, where a large number of step-functions are applied and the response is scanned by passing a "strobe" which measures only in a selected time interval. This latter method is most useful in systems with short response times – microseconds or less, but it becomes unworkable in the type of dielectric measurements with which we are concerned in the present monograph.

The significant point to be noted here is that we are confronted with the principle of information collection – in order to extract a certain amount of information out of the inevitable background of noise one requires a minimum amount of time to enable repeated sampling to be carried out, whether this be in the form of sinusoidal signals or repeated step functions.

To this should be added the further point that the technical perfection of commercially available ac bridges is very high in view of the many decades of development work that has been done on them. Compared with that time–domain equipment has to be "tailor-made" for the specific purpose in hand and it is difficult to obtain ready-made systems. Furthermore, the limitation of time–domain measurements by the required risetimes and response times of the signal and of the detection equipment, respectively, sets an upper limit on the effective frequency range that can be employed.

The same considerations apply in reverse for the frequency–domain operation – the difficulty of generation and detection of sinusoidal signals increase very rapidly below the 1 Hz limit down to which the ordinary bridge techniques are applicable. Although equipments exist on the market which are capable of measuring in the frequency domain down to 10^{-4}–10^{-5} Hz, the potential advantage of time–domain techniques increases as one moves toward these very low frequencies.

One limitation of the speed with which time–domain measurements can be carried out arises from the "memory" retained by all dielectric systems and arising from the fact that, in principle, the dielectric response function $f(t)$ defined by eqn (2.27) extends to infinity and it is only a matter of practical considerations how far should one wish to extend the experiment in any given situation.

To illustrate this point, take the case of the "universal" response defined by eqn (3.33) and assume that we apply to the system in question a square wave of duration T, which represents the practical limit of what should have been an infinitely long step function.

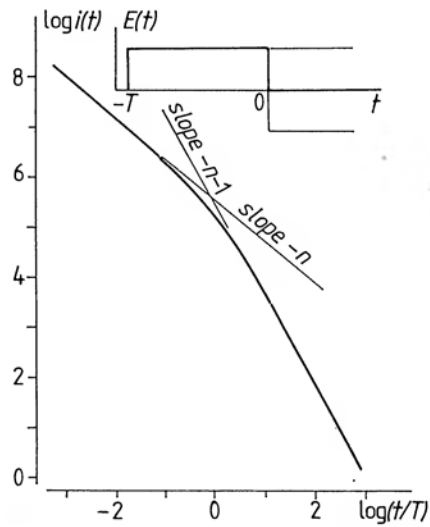


Figure 6.1 An illustration of the effect of a finite charging time T, simulated by two superimposed step functions starting at $-T$ in the positive sense and at 0 in the negative sense. The resulting discharge current which would be obtained for a material showing a pure power law $t^{-0.8}$ consists of one branch with the correct exponent, for $t \ll T$, and this is followed by a second branch with the slope -1–0.8 at times much longer than T.

After this "charging time" T, the sample is short-circuited, i.e. the field in the sample is zero again, assuming all the time that the system is linear and therefore free from space charge and other effects which might tend to distort the field in the interior. We now note with reference to the diagram in Figure 6.1 that the square wave of finite duration T may be represented by two step functions, one positive starting at $t = -T$ and the other negative of the same amplitude starting at $t = 0$ the overall response being determined by the *principle of superposition* according to which the first step function is assumed to act for an infinitely long time and the effect of the second one is superimposed on that of the first.

The current is given therefore as the sum of two terms:

$$i(t) \propto (T + t)^{-n} - t^{-n} \tag{6.1}$$

and in the limit of short times, $t \ll T$

$$i(t) \propto t^{-n} \tag{6.2}$$

since the first term is negligible in comparison with the second in (6.1). On the other hand, in the limit of long times, $t \gg T$, we have

$$i(t) \propto t^{-n}\{(1 + T/t)^{-n} - 1\} \simeq nTt^{-n-1} \tag{6.3}$$

We conclude, therefore that the discharge current at times that are short compared with the preceding charging time follow the genuine characteristics expected of the material in question, but at times long compared with the charging time the slope of the logarithmic plot of the discharge current becomes much steeper than the genuine characteristics.

The conclusion is that in order to study the discharge characteristics of a material with a view to determining its dielectric response, it is necessary to charge the material prior to the beginning of the discharge process for at least *ten times longer* than the maximum desired time of the discharge measurement. Apart from serious errors that may arise if this condition is not fulfilled, these requirements mean that the time taken for an experiment is much longer than the actual measuring time itself. The same conditions apply to the opposite process of charging current measurement: the material must be thoroughly discharged for at least ten times as long as it is intended to measure the charging current for. A more detailed discussion of this will be found in Appendix 6.1.

We now return to the problem of non-linearity which is very important in relation to the study of the dielectric properties of real materials. Although it is very convenient to make the assumption of linearity in order to simplify the analysis of the results of measurements, it remains an experimental fact that very few systems are truly linear either in the presence of high electric fields or at high temperatures, where charge movements may take place on a significant scale. Under these conditions the measurement of the dielectric response by any synchronous or narrow band method invariably rejects all information relating to the non-linearity, since the detected signal consists only of the fundamental frequency of the driving signal and leaves out all higher harmonics which might give an impression of the extent of departure from linearity.

It would be possible, in principle, to study non-linear behaviour by looking directly at the second or third harmonic – depending on the nature of the system, but this becomes possible where the dominant, i.e. the *real part* of the susceptibility becomes significantly non-linear, whereas the non-linearity of the much smaller *loss component* of the characteristic has to become very strong in order to show any significant effect on the production of higher harmonics (Le Sueur and Jonscher 1972, Jonscher 1973).

The most direct manner of investigating the non-linearities of the dielectric response consists therefore in the measurement of the time–domain response, since this method probes directly the response according to the principle, already mentioned in Section 2.6, that

"Mother Nature works in the time domain"

Any non-linearities that may be observed are therefore far easier to interpret, since they represent the natural response of the system to specific driving conditions and we shall see examples of this later in this Chapter.

We wish to mention here the fact that the non-linearity of the real part of the dielectric permittivity can normally only be observed at very high electric fields, so that the observation of any significant phenomena is limited to very high frequencies, in the optical range, where breakdown does not intervene, or alternatively to very short pulses in, for example, liquids. We shall not be concerned with this class of phenomena in the present treatment.

An extreme limit of non-linear behaviour is found in the case of ferro-electrics where domain switching changes the state of polarisation in a discontinuous manner, but once again this class of response falls outside the scope of our work.

One special aspect of dielectric non-linearity is represented by the process of charge injection and transport in insulating materials and we shall have occasion to study this in some detail, especially as this represents a rather important aspect of the behaviour of many materials.

It is worth mentioning that the time–domain response offers a particularly easy way of testing for non-linearities by simply varying the amplitude of the applied step field. The results can be very conspicuous while the corresponding frequency–domain measurements with variable amplitude of the signal generally tend to be much less sensitive. We shall show several examples of this type of behaviour.

Non-linearity of dielectric response is best tested experimentally in frequency–domain measurements by taking a reading at a given signal amplitude and repeating this reading at a higher value of the amplitude – if the reading remains the same, the system may be said to be linear, if not, non-linearity is evident.

It is relevant to note here that the applicability of Kramers–Kronig relations is restricted to linear systems and yet experience shows that these transformations appear to be relatively "tolerant" to finite amounts of non-linearity in the systems under study. One contributory factor to this apparent insensitivity is that the frequency–domain measurement tends to "linearise" the system by taking only the fundamental frequency result and rejecting higher harmonics. Data for the real and imaginary components of the susceptibility evaluated in this manner may then show a remarkable degree of agreement with their Kramers–Kronig transforms – neither are faithful representations of the response of a non-linear system, but they remain self-consistent expressions of the linearised response. On the other hand, the Fourier transformation between the frequency– and time–domain responses may not be equally satisfactory, since the time–domain behaviour contains the non-linearity which may have been eliminated from the frequency–domain data.

It will be clear to the reader that the available experimental data relating to the time–domain response of the dielectrics are much

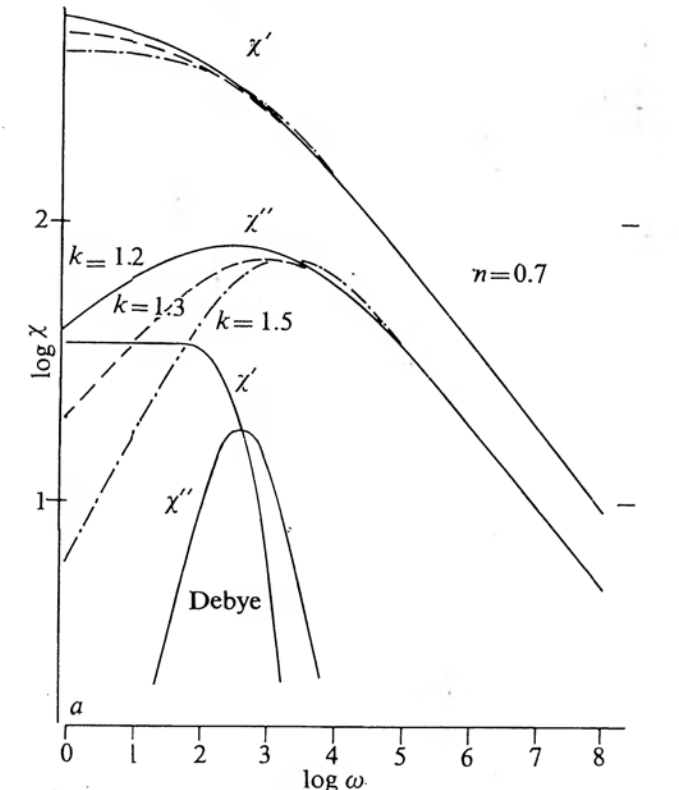
less numerous than the corresponding frequency–domain information. This is due to the fact, already stressed earlier, that equipment for the measurements in time domain is far less standardised and less well developed than the frequency domain equipment.

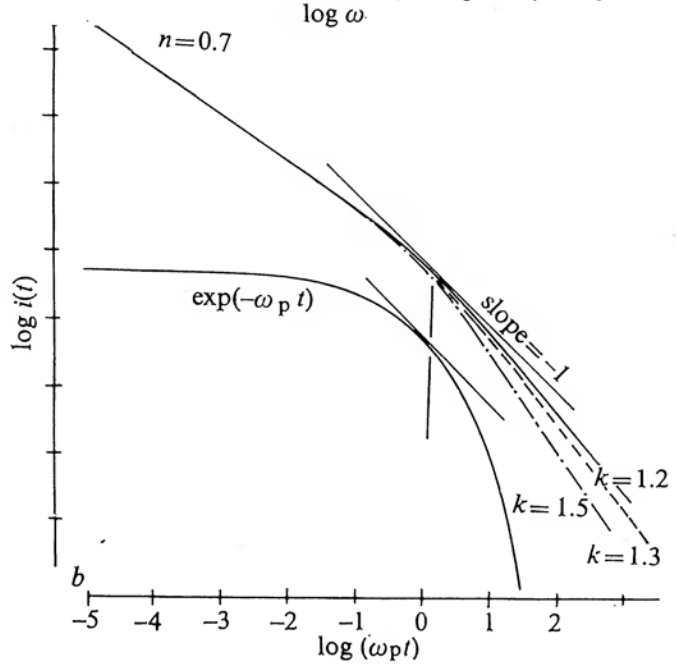
6.2 THE SIGNIFICANCE OF LOSS PEAKS IN THE TIME–DOMAIN

The universal relation in the time domain, $f(t) \propto t^{-n}$, which represents the Fourier transform of the frequency–domain response in which both the real and the imaginary parts of the susceptibility are proportional to ω^{n-1}, cannot be valid in the entire time range from zero to infinity. Its limitation at very short times is evidently set by the inevitable inertial processes which require that $f(t) \propto t$. Likewise, the universal relation with $n < 1$ would correspond to an infinitely large total charge in the system at infinitely long times, and this is physically inadmissible. It is therefore physically necessary that the function $f(t)$ should vary more rapidly than with a logarithmic slope $-n$, where $n < 1$. Now consider the empirical expression derived from our survey of the experimental data for a very wide range of dipolar materials and given by eqn (5.3). At frequencies much in excess of the loss peak frequency ω_p this becomes the universal law, ω^{n-1}, while at frequencies much below the loss peak frequency we have the other power law ω^m. Each of these laws separately transforms into t^{-n} and t^{-m-1}, respectively, according to eqn (3.33), and it is possible to show that the combined equation gives, in fact a combination of these two laws in such a manner that as *short times* $t \ll 1/\omega_p$ the former law applies which corresponds to the *high frequency* response, while the low-frequency response transforms at *long times* $t \gg 1/\omega_p$ into t^{-m-1}. Thus the

Figure 6.2 Equivalent time– and frequency–domain representations of dielectric behaviour giving a loss peak. Diagram b) shows the time–domain dependence for the case of a universal response with the exponents in eqn (6.4), $m = k - 1 = 0.2$, 0.3 and 0.5, $n = 0.7$. An exponential dependence corresponding to a Debye behaviour is also shown on the same time scale. Diagram a) gives the numerically Fourier-transformed time–dependent relation, showing the corresponding loss peaks.

The vertical scales are arbitrarily displaced.


χ′
χ″
k = 1.2
k = 1.3
k = 1.5
n = 0.7
log χ
2
1
χ′
χ″
Debye
a
0 1 2 3 4 5 6 7 8
log ω


n = 0.7
log i(t)
exp(−ω p t)
slope = −1
k = 1.2
k = 1.5
k = 1.3
b
−5 −4 −3 −2 −1 0 1 2
log (ωpt)

complete time–domain response may be represented by the approximate relation:

$$f(t) \propto i(t) \propto \frac{1}{(\omega_p t)^n + (\omega_p t)^{1+m}} \tag{6.4}$$

which in the logarithmic representation consists of two straight lines, one at short times with a slope $-n$ and another, at long times, with a slope $-1-m$, as shown in Figure 6.2. It becomes evident, therefore, that the physical significance of the loss peak in the frequency–domain lies in the existence in the time domain of two *physically separate and consecutive* processes. The loss peak frequency represents the reciprocal time at which a transition from one process to the other takes place in the time domain (Jonscher 1975).

We stress this interpretation because it was, to our knowledge, the first attempt to interpret the loss peak as the outcome of an interplay of two separate physical processes taking place sequentially in the time domain, and not in terms of some single or combined Debye-like loss peaks in the frequency domain. We shall show later that the new interpretation of the dielectric behaviour shows rigorously that this is precisely what happens and that there are two entirely different classes of relaxation processes which dominate in the respective time ranges.

We now note the close similarity between Figures 6.1 and 6.2 b) despite the fact that they are due to entirely different physical causes and this shows the importance of allowing sufficiently long charging and discharging times before the measurements in order to avoid possible mistakes.

The fact that there exists a direct relationship between the presence of a loss peak in the frequency domain and the corresponding transition in the time–domain from a region of slopes less than minus one to a steeper slope, acquires a special significance in the case of the classical Debye relaxation process. There, the exponential characteristic function is represented in the logarithmic plot as a continuously curving line the slope of which passes through minus one at $t = 1/\omega_p$, as may be seen in Figure 6.2. Now the exponential law is the result of a single, "homogeneous" decay process, eqn (4.17), so that in this particular case the presence of a loss peak is the necessary consequence of the single exponential law. In all other cases of non-exponential decay we have to invoke the presence of two separate processes in order to interpret a loss peak.

6.3 THE HAMON APPROXIMATION

The experimentally established existence of the universal relation in the frequency domain, and more generally of the two regions of power law in eqn (5.3) makes it possible to develop an approximate procedure for the evaluation of the Fourier transform from the frequency to the time domain (Hamon 1952). We know from eqn (3.33) that the time-domain response corresponding to the universal law is proportional to t^{-s}, where the exponent s is introduced for generality in place of n or $m+1$. Taking a sample of area A, thickness w and hence the geometrical capacitance $C_0 = A\varepsilon_0/w$, let the measured current response to a *voltage* step of amplitude V_0 be given by the relation

$$I(t) = Kt^{-s} \tag{6.5}$$

where the constant K is determined from the experiment. Taking now the definition of the current response given by eqn (2.33) and ignoring the delta-function and the dc contributions as irrelevant for the present purpose, we may express the current *density* $i(t)$ in response to a field step E_0 in terms of the current $I(t)$ in response to a voltage V_0 and obtain from eqn (6.4):

$$f(t) = \frac{K}{C_0 V_0} t^{-s} \tag{6.6}$$

Using the definition of dielectric loss given by eqn (2.39) and its inverse relation (2.43), we may write:

$$\chi''(\omega) = F_s(t^{-s}) = \frac{K}{C_0 V_0} \Gamma(1-s)\cos(s\pi/2)\omega^{s-1} \tag{6.7}$$

where we have made use of the Fourier transforms (3.33) in which we have replaced ω by t and $n-1$ by $-s$ which is permissible, since the Fourier transformation works both ways.

Define now a time t in terms of the reciprocal frequency ω, $\omega = 2\pi f = 1/t$, so that eqn (6.6) may be written in the form:

$$\chi''(1/t)_{rad} = \chi''(1/2\pi t)_{Hz} = \frac{\Gamma(1-s)\cos(s\pi/2)}{C_0 V_0} tI(t) \tag{6.8}$$

This expression relates the dielectric loss at a radian frequency $1/t$ or a frequency $1/2\pi t$ in Hertz, to the charging or discharging current $I(t)$ at the time t.

It should be noted that this relation is valid not only in the usual range of the exponent $0 < s < 1$, corresponding to the loss at frequencies above ω_p, but remains valid also in the range of positive slope of the loss curve below ω_p, where we may put formally $m = s - 1$, which means that the range of the exponent covering this part of the loss curve is $1 < s < 2$. Defining the parameter

$$a = \Gamma(1 - s)\cos(s\pi/2) \tag{6.9}$$

we find the numerical dependence of a on s as shown in Figure 6.3, so that for $s = 0$ corresponding to the strong low-frequency dispersion $a = 1$, for $s = 1$ corresponding to a flat loss $a = \pi/2$, while for positive slopes corresponding to the rising part of the loss curve a increases rapidly beyond $m = \frac{1}{2}$. Figure 6.3 shows also a sketch of the corresponding loss characteristic with the slopes drawn at appropriate points of the abscissa.

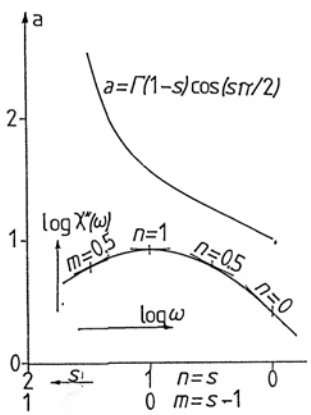


Figure 6.3 The dependence of the parameter a in the Hamon approximation, eqn (6.9), on the exponents s, $n = s$ and $m = s - 1$, given on the abscissa. The inset represents a logarithmic plot of the loss-frequency dependence which has the property that the slopes m and $1 - n$ drawn at the respective frequencies have values corresponding to the values of the exponent s at the same abscissae. This enables an estimate to be obtained of the parameter a depending upon the local slope of the logarithmic loss-frequency plot.

We note that from approximately $m = \frac{1}{2}$ on the positive slope to $s = -1$ on the negative slope, i.e. over the entire range of negative slopes and up to $m = \frac{1}{2}$ on the positive slope, the parameter a changes by a little over a factor of 2. Furthermore, eqn (6.8) remains valid even though the exponent s is not a constant but changes

slowly with frequency giving a curved loss-frequency plot in the logarithmic representation, where we may establish local relations between loss and polarisation current of the form:

$$\chi''(f)_{Hz} = \frac{a}{C_0 V_0} \frac{1}{\pi f} I(1/2\pi f) \tag{6.10}$$

or the equivalent expression:

$$\chi''(1/2\pi t)_{Hz} = \frac{a}{C_0 V_0} tI(t) \tag{6.11}$$

Provided that the exponents m and n fall in the range indicated, the parameter a may be set equal to a constant value of the order of unity with a relatively small error. This means that most loss characteristics may be obtained from the corresponding polarisation currents $I(t)$, by plotting $tI(t)$ *logarithmically against* $1/2\pi t$, and this then becomes a plot of $\chi''(f)$ with the appropriate scaling factor a/C_0V_0 with the parameter a set to a value between 1 and 1.5, according to the nature of the loss curve. This procedure is known as the Hamon approximation and is very useful for a rapid appraisal of loss data from the polarisation currents, or vice versa.

Equation (6.11) shows that the loss peak occurs at the frequency $\omega_p = 1/t_1$ where t_1 is the time at which the logarithmic slope goes through the value $\mathrm{d}(\log i)/\mathrm{d}(\log t) = -1$, as shown in Figure 6.2.

6.4 EVIDENCE FOR INERTIAL EFFECTS

We have already mentioned in the context of eqn (4.6) that the effect of inertial phenomena on the time–domain response is the presence of a linear rise with time at short times, to be followed later by a steady decay towards zero. It is virtually impossible to obtain experimental evidence fo these processes in the case of solids, since the inertial regime corresponds to frequencies of the order of the optical and far-infrared region, i.e. times in the pico-second range and below. We propose instead to give a very elegant example of this type of response obtained with biological macromolecules with correspondingly high inertia, floating freely in a viscous fluid medium. The experiment consists in the application of a delta-function force and in the monitoring of the alignment of the molecules at later times, in accordance with the definition of the characteristic function $f(t)$. However, the experiment is not carried out in the conventional manner by the application of a short pulse

of a sufficiently strong electric field, since this would inevitably lead to the dielectric breakdown of the sample, but instead a giant laser pulse of very short duration is applied and this causes an alignment of the dipoles not through the direct action of the electric field, since this is oscillating very rapidly and the permanent dipoles simply do not have a chance to follow, but through the generation of *induced* dipoles which then tend to align the molecules in the general direction of the oscillating field because of the inherent anisotropy of these large molecules. Since the induced dipoles respond rapidly even on the scale of the frequencies corresponding to the visible light of the laser, the energy of alignment is proportional to the square of the field and there is only an alignment and no oscillations result from this cause. The dipole moment of the resulting polarisation is then probed with a steady beam of low-intensity laser light which simply measures the resulting birefringence and provides a direct measure of the instantaneous polarisation of the sample.

Figure 6.4 shows the results of such measurements on Tobbaco Rattle Virus and it is clear that the rising part of the inertial response persists long after the disappearance of the exciting giant pulse.

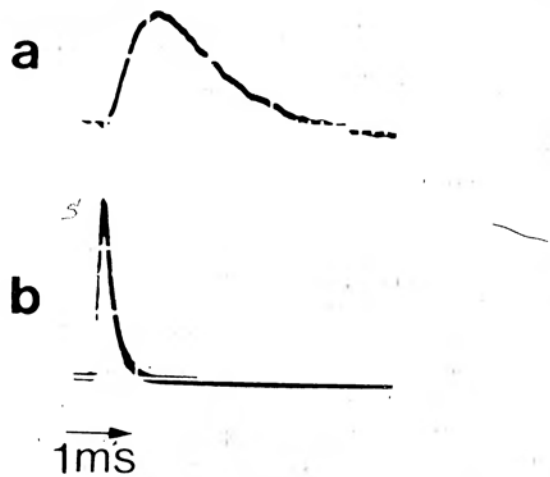


Figure 6.4 A demonstration of inertial effects in large biological molecules by means of laser-induced birefringence of a solution of Tobacco Rattle Virus (TRV) at a wavelength of 488 nm. The upper trace a) shows the birefringence of TRV, the lower trace b) shows the fixed — Q YAG laser pulse of approximately 8 kW maximum power. Both traces on a time scale shown by the arrow.

From Jennings and Coles (1974). Reprinted by permission from *Nature*, Vol. 252, p. 33.

6.5 LONG-TIME BEHAVIOUR IN LOW-LOSS POLYMERS

The time–domain technique was perfected by W Reddish and his group in the 1950's and they used it extensively for the determination of the dielectric loss in a wide variety of low-loss polymers at frequencies down to 10^{-4} Hz, by Fourier transforming from the time data. The technique enabled the frequency spectrum to be extended by several decades, compared to what was technically feasible by bridge and other techniques at that time (Hyde 1970).

The results of some early measurements on oxidised high-density polyethylene are shown in figure 6.5 a), where the current response is plotted in the manner familiar in electrical engineering circles, as $\log(E/i)$ since it is customary to refer to "time–dependent resistivity", even though this concept has very little physical significance. The data were measured at several temperatures and the normalisation of these results is possible by lateral shifting in the normal manner. However, since the normalisation of loss peaks corresponds to nearly horizontal displacements, the normalisation of current plots must correspond to displacements along the slope unity in the logarithmic representation, since the loss is related to the product of t and $i(t)$ from eqn (6.11). The frequency or time activation energy is then obtained from the *horizontal component* of the logarithmic displacement plotted against $1/T$. If the locus of the displacement point has a slope -1 then the corresponding loss peak amplitude is constant in temperature.

The use of the time–domain measurements as an extension of the range of bridge measurements in the frequency–domain is illustrated in Figure 6.6 giving the extended dielectric loss spectrum of PMMA in a range of ten decades. The lowest two decades are transformed from the time–domain, the upper $4\frac{1}{2}$ decades are measured directly, while the middle range of $3\frac{1}{2}$ decades shown dotted in the figure is interpolated. This type of contour map of loss with temperature and frequency as variables corresponds directly to the three-dimensional representation of Figure 3.29 and the loss contour shows a peak changing with frequency and temperature in a manner consistent with the presence of an activation energy.

The implication of the transient response of the charging currents for the determination of direct current conductivity is discussed in Appendix 6.2.

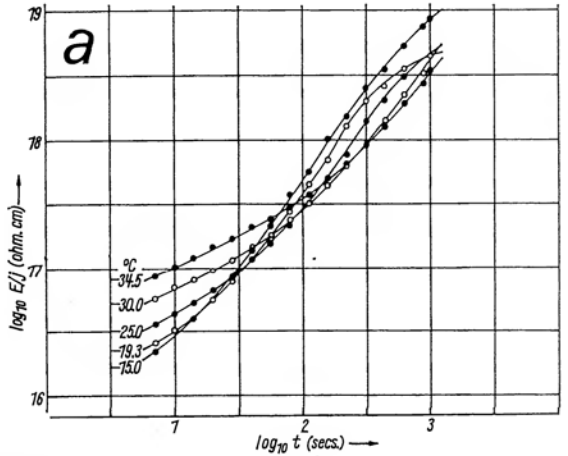
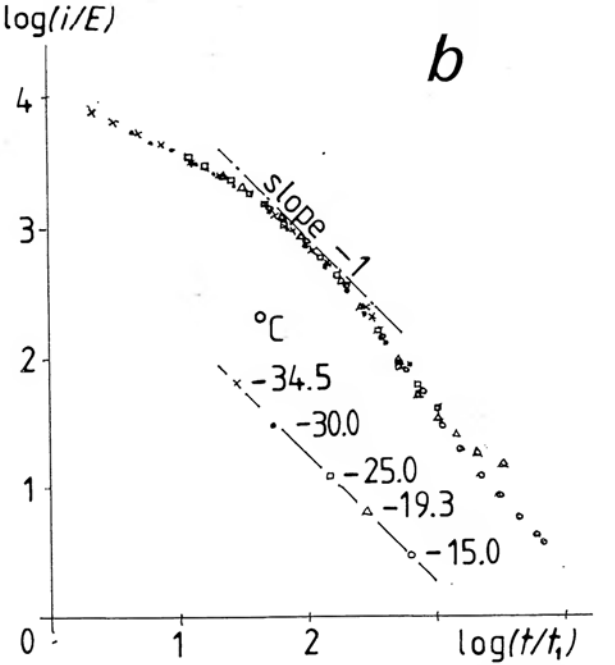


Figure 6.5
a) Time–domain data of a brand of polyethylene, Marlex 50, slightly oxidised. The data refer to a range of temperatures and they are plotted in the customary engineering representation as $\log(E/j)$, where j is the current density, against $\log t$. The solid line indicates unit slope and "A" represents a reference point for the normalisation in diagram b).

From Reddish and Barrie (1958).



b) The same data as in a) normalised by lateral shifting and inverted to give a normalised plot of $\log(j/E)$ which corresponds to our manner of presentation. The locus of the normalisation points runs parallel to the slope −1, which is also drawn to show the position of the normalised loss peak. The fact that the normalisation runs parallel to this slope of −1 implies that the loss peak does not change its amplitude with frequency.

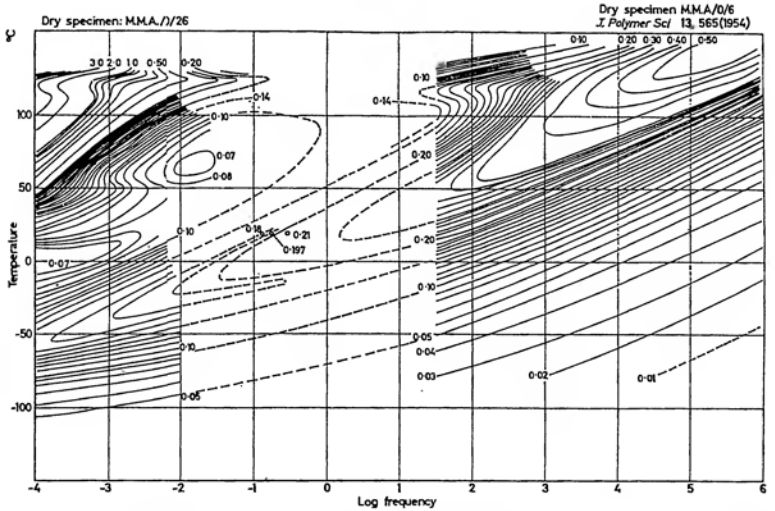


Figure 6.6 Contour map of dielectric loss in Polymethyl methacrylate (PMMA) plotted with frequency and temperature as variables, covering ten decades of frequency and a range of 250 K. The two lowest decades were obtained by Fourier transformation from the time–domain measurements, the upper 4½ decades are direct bridge measurements. The dotted contours denote interpolated values.

From Reddish (1962).

6.6 DETECTION OF NONLINEARITIES BY TIME–DOMAIN MEASUREMENTS

We showed in Section 2.5 that the time–domain response of a linear system to rising and falling step functions is symmetric and that the difference between the charging and discharging currents should be equal to the time–independent direct current, if this is measurable:

$$i_c(t) - i_d(t) = i_0 \tag{6.12}$$

It is a measure of nonlinearity of most dielectric systems that this expression is hardly ever satisfied and instead the difference between the charging and the discharging currents depends on time and may even become negative.

An example of this type of behaviour is shown in Figure 6.7 where results of time–domain measurements of charging and discharging currents are shown for a sample of stearic acid under various conditions of the amplitude of the applied step and of temperature.

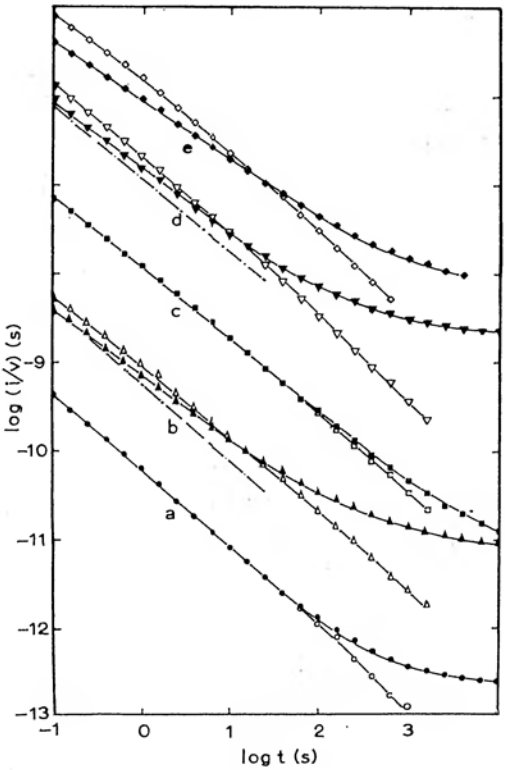


Figure 6.7 The logarithmic representation of the charging currents, $i_c(t)$ – full symbols – and discharging currents, $i_d(t)$ – open symbols – for Al – Stearic Acid – Al sandwich structures. Curves a)–d) relate to a sample of 11 monolayers thickness, curves e) to 13 monolayers. The temperature and average field were as follows.

Curves	a	b	c	d	e	
Average field	6.8	6.8	1.1	9.1	4.8	$\times 10^7$ V/m
T	250	300	300	300	320	K

The currents are scaled by the applied voltage V and successive sets of curves are displaced vertically by one decade with respect to one another for clarity. The chain dotted lines on curves b) and d) represent the correct relative positions of the lower-stress characteristics given by lines a) and c), respectively. The conductance scale refers to curve a).

Jonscher and M A Careem (1975).

While at the lowest temperatures and fields the charging and discharging currents cannot be resolved to any satisfactory degree at short times and it is therefore impossible to verify the applicability of eqn (6.12), at least there is no evidence of it not being applicable. The situation is changed completely, however, if either the temperature or the amplitude of the field is raised sufficiently when we observe complete *crossing over* of the charging and discharging currents at short times, which is completely incompatible with the response of a linear system. It was suggested in the context of that study that the phenomenon of crossing over is due to the injection of space charge into the dielectric during the charging process, leading to gradual accumulation of excess charge density near one or both electrodes, according to the nature of the interface. When the field is removed abruptly at the commencement of the discharge process, the initial discharging current consists of the true depolarisation current with the current due to the withdrawal of the previously injected charges superimposed on it. While the polarisation current is not expected to be non-linear within the range of practically attainable field amplitudes, in view of the essential linearity of the dielectric response already mentioned earlier, the process of injection is evidently strongly non-linear and responds easily to such factors as temperature and field.

Very good examples of non-linear dielectric systems are found among ionic conductors or materials in which the element of ionic conduction is at least appreciable. One of the most interesting examples of this type of behaviour is found in electrolytic capacitors which consist of a very thin anodic oxide layer on a suitable metal substrate such as aluminium or tantalum, with the top electrode consisting of either a solid or a liquid electrolyte. Here ion transfer between the electrolyte and the anodic oxide layer may take place relatively easily, leading to strong transient phenomena, while in the long term the steady current is negligibly small. Figure 6.8 shows a collection of charging and discharging currents in a tantalum electrolytic capacitor over a range of temperatures between 230 and 393 K. Each pair is displaced vertically with respect to the preceding one to avoid the confusion of many overlapping curves and to emphasize the relative positions of the charging and discharging currents. It is noticeable that the phenomenon of crossing over is most pronounced at intermediate temperatures. At the lowest temperature the charging and discharging currents are almost indistinguishable except at the longest times, while the charging current is significantly higher through most of the range of time at the highest temperature.

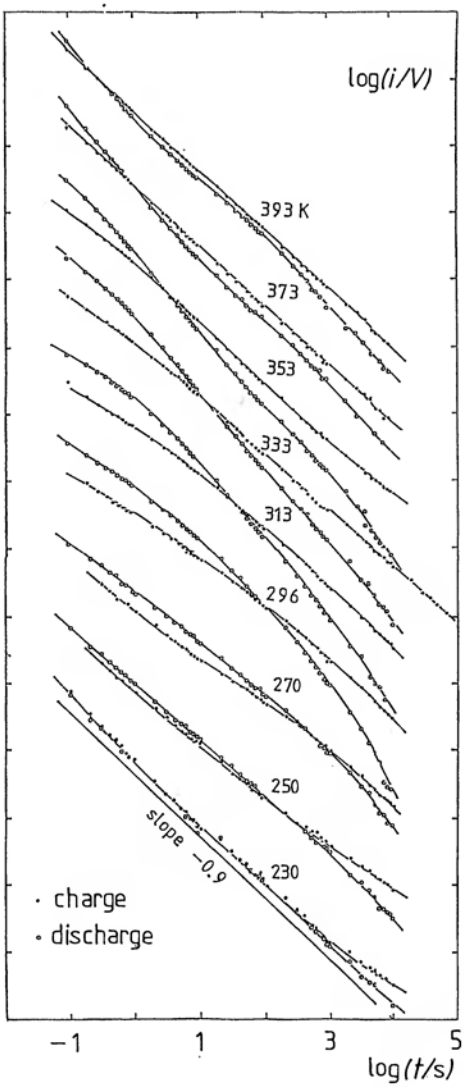


Figure 6.8 Charging and discharging currents in a tantalum electrolytic capacitor (●●●● and ○○○○○, respectively) drawn as sets of pairs corresponding to different temperatures at a constant voltage step amplitude. Each set is displaced vertically by one decade with respect to the previous one for clarity.

The crossing over of $i_c(t)$ and $i_d(t)$ should be noted, especially in the intermediate range of temperatures.

From Meca and Jonscher (1979).

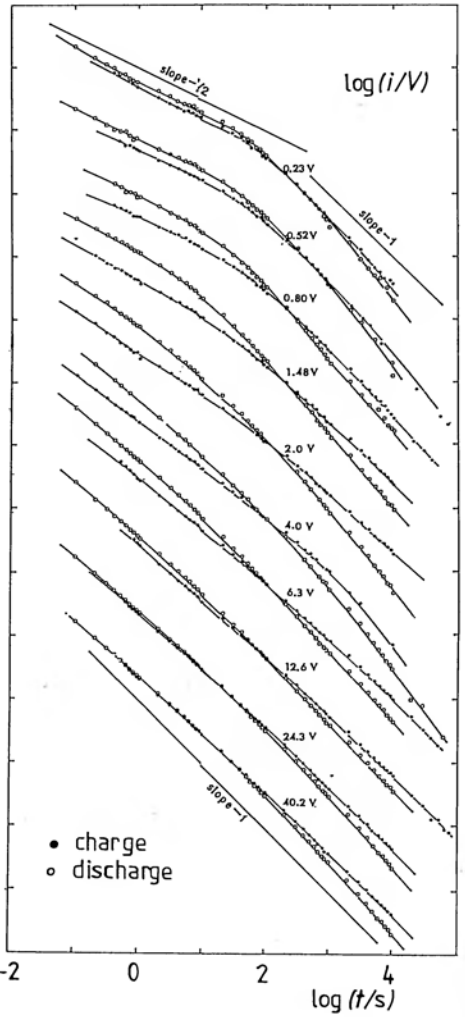


Figure 6.9 Charging and discharging currents in a tantalum electrolytic capacitor (●●●●● and ○○○○○, respectively), drawn as pairs corresponding to variable step amplitude and at a constant temperature of 296 K. Each set is displaced with respect to the previous one by one decade for clarity. At low voltages the short-time slope is $-\frac{1}{2}$, the slope changes gradually with increasing voltage. Maximum crossing over occurs at medium voltages.

From Meca and Jonscher (1979).

Different aspects of nonlinearity appear in Figure 6.9 showing a corresponding collection of data with the amplitude of the applied step voltage as parameter, at a constant temperature. Here the voltage covers a range of 200:1 and we note the complete change of the nature of the response. At low voltages the initial part is well represented by the $t^{-1/2}$ relation which could be presumed to correspond to some diffusion process, while the later stages of the polarisation and depolarisation currents are steeper than -1. The corresponding frequency–domain response would therefore consist of a loss peak with the high-frequency behaviour being dominated by the $\omega^{-1/2}$ dependence. As the voltage amplitude increases, the crossing over becomes initially more pronounced, while the clear distinction between the slopes $-1/2$ and slopes steeper than -1 gradually becomes blurred. This might be interpreted as showing the diminishing contribution of diffusion in high driving fields. It is remarkable, however, that even at the highest voltages there is no evidence of direct current conduction at very long times.

As an illustration of the extent of agreement between the time– and frequency–domain measurements in the linear regime we show in

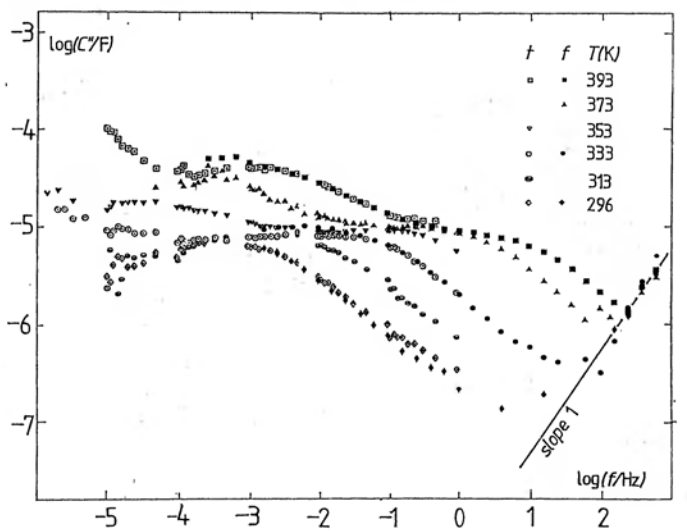


Figure 6.10 The frequency dependence of the dielectric loss G/ω for a 50 μF electrolytic capacitor based on a sintered tantalum powder body. The full symbols, marked by f in the inset, for all six temperatures were obtained directly in the frequency domain, the open symbols (t) are Fourier transformed data from time–domain measurements using the smallest available step amplitude. The rise at high frequencies is due to series resistance, probably from the electrolyte. From Meca and Jonscher (1979).

Figure 6.10 the superposition of directly measured frequency–dependent loss, in the form of $G(\omega)/\omega$ for the same capacitor and of the Fourier transformed time–domain data, both taken at the lowest amplitudes. The agreement of the data in the region in which their frequencies overlap is very good and shows both the extent of linearity and the reliability of the Fourier transformation itself, which in this case was performed numerically. No such agreement could be obtained with higher amplitude responses.

Non-linear response may also be studied by superposing a rising or falling step on a steady bias maintained for a very long time before and after the application of the step. The bias has the effect of "conditioning" the system by either injecting or removing any charge carriers and the step subsequently "samples" the conditioned system. Figure 6.11 shows the results of such measurements and it should be noted that the discharge currents corresponding to different bias voltages are here represented in their correct relative positions, i.e. they are not displaced as in the other figures but represent the genuine influence of bias on the level of the discharge current which is seen to *fall* by nearly an order of magnitude as the

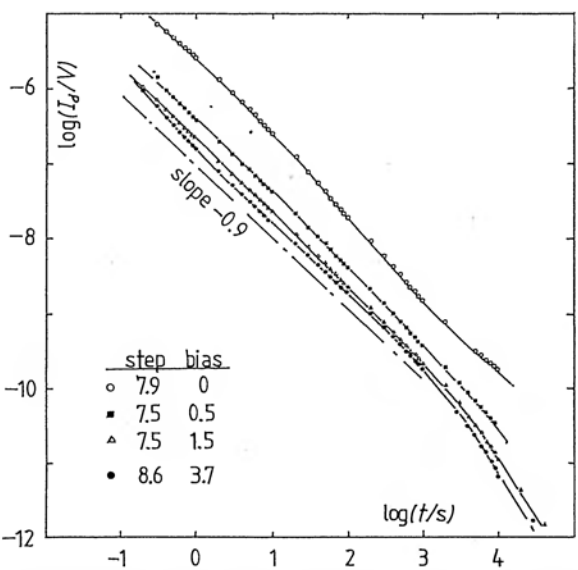


Figure 6.11 The effect of a steady bias superimposed on a step on the discharging currents of the same capacitor as shown in Figure 6.10. Measurements were taken at 296 K. The meaning of the symbols is explained in the inset, step of 7.5 V with bias of 0.5 V indicates the change from 8.0 to 0.5 V to obtain the discharge curve. From Meca and Jonscher (1979).

bias *increases*. Despite this very strong effect on the amplitude of the current, the time–dependence does not appear to change to any significant extent.

6.7 CONTRIBUTION OF CHARGE CARRIERS TO THE DIELECTRIC RESPONSE

The role of charge carriers in the determination of the dielectric response of solids was already mentioned in the context of Chapter 5, especially in relation to the low-frequency dispersive phenomena. Time–domain measurements are especially important in the study of these effects, but their influence on the frequency dependence of the dielectric loss is also significant. The reason for the strong influence of charge carriers on non-linearity lies in the fact that their injection and removal are strongly non-linear, while the response of the dipolar "lattice" of the material remains essentially linear within the attainable range of voltages.

The reflection of the phenomena shown in Figure 6.11 on the frequency domain response is shown in Figure 5.55 giving the loss measured under a small alternating signal superimposed on a steady bias applied to the same electrolytic capacitor. Increasing bias has the expected effect of lowering the loss and it is interesting to note the changing nature of the frequency dependence, from the apparently diffusion – dominated power law with $n = \frac{1}{2}$, to the lattice–dominated exponents $n = 0.87$ at a bias sufficiently high to cause the removal of ionic charges.

The removal of charge carriers, especially ionic ones, from dielectrics is a well-known phenomenon and is used to "clean" materials from charges introduced by earlier treatments. It is evident, however, that the effectiveness of this method depends critically on the properties of the contacts applied to the material, since these have to "block" the passage of the charges in question in the presence of an external bias. If this is not the case and the contact injects charges at sufficiently high voltages, then the effect on the dielectric response is to increase the loss, or the corresponding ac conductivity, and simultaneously to increase the real part of the susceptibility, in a manner typical of strong low-frequency dispersion. An example of this is shown in Figure 5.54 relating to stearic acid multilayers with aluminium electrodes. The conductance at low frequencies is visibly not saturated, while at the highest voltages the behaviour appears to be dc dominated, but the strong dispersion of $\chi'(\omega)$

shows a contribution of the low-frequency dispersive behaviour. It may be inferred that aluminium contacts to stearic acid are capable of injecting charge carriers, probably electrons which would be expected to be more readily introduced than ions across metallic–dielectric interfaces.

An interesting and as yet insufficiently elucidated aspect of the time–domain response concerns the comparative behaviour of the polarising current under conditions of dark discharge and of the

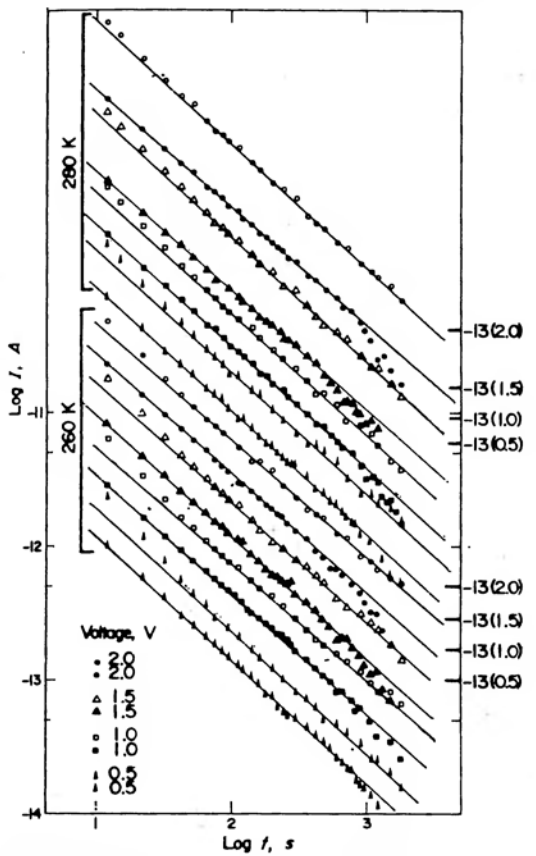


Figure 6.12 The dark discharge currents (solid symbols) and the photo–decay currents at constant voltage (open symbols) after the removal of an ultraviolet illumination. The readings were taken at 0.5, 1.0, 1.5 and 2.0 volts and at two temperatures of 260 and 280 K. The left-hand current scale refers to the lowest set of currents, all other sets are displaced vertically for clarity and the markers on the right indicate the position of 10^{-13} A. The strict parallellism between the dark and photo-discharge currents should be noted.

Jonscher and Buddhabadana (1978).

decay of photoinduced current after the cessation of illumination that had produced photocurrent under steady state conditions. Experiments carried out on stearic acid multilayers have shown that the two currents decay in an identical manner, despite the fact that the initial and boundary conditions are completely different in the two cases. Figure 6.12 gives the results of such measurements and the accompanying Table 6.1 summarises the initial and boundary conditions applicable in both situations. Since in the dark case it is reasonable to expect the depolarisation current to be dominated by the dipolar contribution, while in the photodecay case the influence of injected charges must be appreciable, it follows that there exists a strong "coupling" between the charges and the lattice which forces the two decay processes to conform to the same law.

TABLE 6.1
The boundary and initial conditions applicable to the dark discharge and to the photo-relaxation.

Parameter	Dark discharge	Photo-decay
Initial field distribution	substantially uniform	significantly influenced by injected space charge
Polarisation in steady state	mainly due to dipoles, with some contribution from any dark injected charge carriers at higher voltages and temperatures	in addition to the dipolar contribution, there is a significant polarisation due to the photo-generated charge distribution
change of potential	from finite to zero	no change
final field distribution	zero field	substantially uniform
the change of polarisation arises from	mainly dipolar relaxation	mainly mobile charge redistribution

Jonscher and Buddhabadana (1978).

The time– and frequency–dependent phenomena accompanying the injection of space charge into dielectrics are very intimately related to the corresponding steady state processes, in particular the well-known subject of space-charge-limited currents (Lampert 1965). We do not propose, however, to draw closer links between the static and transient cases, since there appears to be little information of significance which could be brought into this discussion from the experimental angle. On the whole, space-charge-limited currents correspond to an increase of the effective capacitance by, at most, a factor of two, and there must therefore be a corresponding increase of loss.

6.8 OTHER CHARGE CARRIER PHENOMENA

Our review of the time–domain response of dielectrics would be incomplete without the mention of several other phenomena which influence the movement of charge carriers in dielectrics and which become especially visible in the time–domain measurements. Some of these may be regarded as somewhat marginal to the strict topic of dielectric response of solids, others may not be sufficiently well understood at the present time to warrant a detailed discussion in the present context.

a) Charge injection and surface potential

The fact that any contact between two materials, be they both dielectrics or one dielectric and the other a metal, leads to charge transfer between them is well known and has been widely studied. The immediate evidence of charge injection is obtained from the build-up of potential at the surface which can be easily measured by suitable capacitive probes. The surface charge, once left to itself, tends to decay gradually by drift into the interior of the material under the action of its own electric field (Davies 1969, Hill 1975, Rose-Innes 1981).

This may be used to study the mobility of charges in low-mobility solids. A rather little understood part of this process is the fact that the charge collected on subsequent breaking and making of the contact tends to increase, apparently indefinitely, thus making it difficult to explain the phenomenon simply in terms of contact potential difference.

Accumulations of surface charges may give rise to fields that are sufficiently high for surface breakdown to occur in extreme conditions.

b) Energy loss arising from the movement of charges

The fact that a charge carrier moving through a lossy dielectric should lose energy in the process of changing the local polarisation of the surrounding medium is intuitively obvious and yet it does not appear to have received much detailed attention (Fröhlich and Platzman 1953, Jonscher 1980).

Yet it is apparent that this type of interaction between a moving charge carrier and the surrounding lattice must bring about a strong coupling, of which some evidence had already been seen in Figure 6.12. It is to be expected that the growing understanding of the dielectric processes will lead to the elucidation of this important problem.

c) Dispersive charge flow

This subject has acquired considerable importance in recent years in the context of charge transport in amorphous semiconductors and semi-insulators (Pfister and Scher 1977, 1978, Scher 1977). The fundamental difference in comparison with charge transport in crystalline materials lies in the fact that, while in the latter the movement of an injected pulse of charges in a steady electric field produces a steady drift of the pulse with the appropriate drift velocity, subject only to a slight broadening of the pulse by diffusion, in amorphous and other low-mobility solids the identity of the pulse becomes completely smeared out, apparently through very heavy trapping of charges and their very slow release under thermal excitation. Thus while the concept of transit time arises naturally in crystalline solids as the time when the current due to the charge motion decreases rapidly to zero, in the second case we have no obvious stage which could be identified with the transit of any particular charges. Instead, the current-time dependence shows a type of relationship shown in Figure 6.13 and this type of behaviour

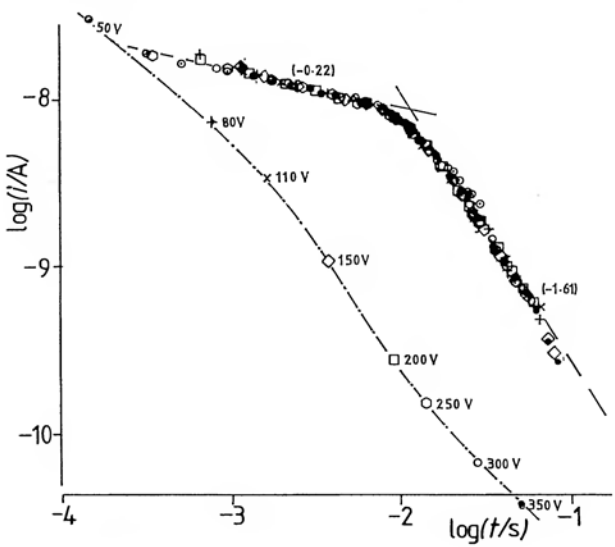


Figure 6.13 Time response of the decay current in a sample of poly-vinyl karbazole (PVK) after perturbation by an electron beam pulse, with variable voltage across the sample. The results for different voltages have been normalised in the usual manner, with the symbols connected by the chain-dotted line corresponding to the displacement of a representative point, with the respective values of the voltage indicated. The two exponents of the power law are shown as -0.22 and -1.61.

From measurements by Hirsch (1979) plotting and normalisation by courtesy of R M Hill.

is known under the name of *dispersive transport*. The current-time dependence is of the power–law type, consisting of two branches separated by a distinct kink which tends to be identified with a "transit time" in the system although the dependence of this time on the amplitude of the electric field causing the drift is not simply inverse, as would be the case in ordinary drift. What is very interesting and intriguing is that the power–law dependence observed in dispersive flow resembles strongly the time–domain response of dielectric polarisation in the absence of drift, Figure 6.2 b). The full significance of this similarity is not clear and many workers do not believe that any relationship exists, but further studies may reveal that the process is another manifestation of charge-lattice coupling.

d) Charge carrier systems with strong dispersion

It is very interesting to observe the time–domain response of systems with strong low-frequency dispersion of the type discussed in Section 5.6, especially at times long compared with the reciprocal frequency at which the transition occurs from the strongly dispersive regime to the less rapidly varying power law, cf. Figures 5.43 and 5.45. Since this regime obeys the same universal power law of frequency dependence, ω^{n-1}, but with an exponent n close to zero, the Fourier transform into the time domain, eqn (3.33) is valid and indicates a very slowly decaying power law for the current, $i(t) \propto t^{-n}$. Time–domain measurements would immediately indicate the presence of any non-linearities in this behaviour, as well as the eventual behaviour of current at very long times.

There are very few experimental data available for this regime, but Figure 6.14 shows a set of charging and discharging currents in a fast ionic conductor of the hollandite type, for a range of amplitudes. At short times the response appears like the tail end of a normal power law with values of the exponent in excess of $\frac{1}{2}$, and with the amplitude approximately linear in the applied potential. This is followed by a range of some three decades of time in which the current appears to be independent of time – a "quasi–dc" regime – once again linear in voltage, until finally the current begins to drop more rapidly.

It is highly significant that both the charging and the discharging currents follow identical paths in the first two regions, indicating that the processes in question are reversible, including the low-frequency dispersive region. At 1 V amplitude, the two currents remain identical even in early stages of the final drop, the charging

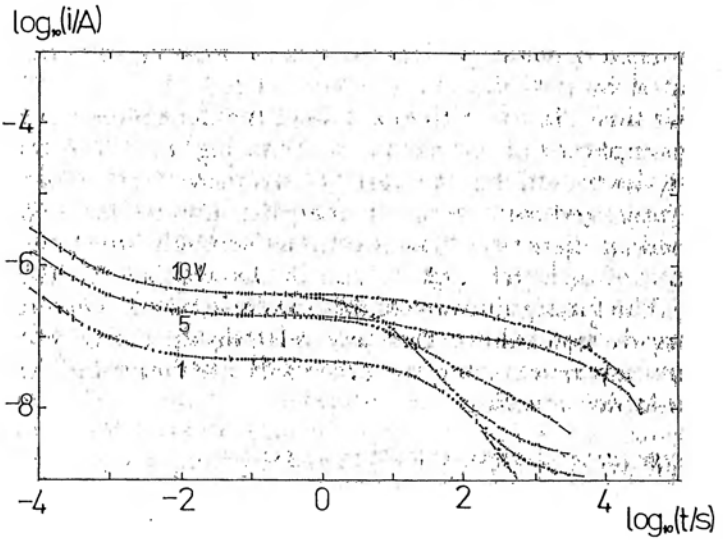


Figure 6.14 Direct plot of the charging and discharging currents in a sample of a fast ionic conductor of the hollandite type. The graph was obtained on-line as the experiment proceeded and the three sets of characteristics correspond to the stated values of the step voltage applied to the sample. The charging and discharging currents are indistinguishable at shorter times and tend to separate at progressively longer times as the amplitude decreases. The behaviour is linear at short times and becomes strongly non-linear at long times.

Unpublished data from Chelsea Dielectrics Group.

current becoming apparently constant only at the longest times. At the two larger amplitudes the discharge current becomes distorted and the charging current shows a complex structure. At this stage the system is strongly non-linear in the applied voltage.

This kind of time–domain experiment is most instructive in the study of strong low-frequency dispersion and the advantages of time–domain measurement are particularly clear.

6.9 CONCLUSIONS REGARDING TIME–DOMAIN EVIDENCE

When compared with the frequency–domain information, time–domain data are evidently much less abundant and less reliable in terms of accuracy, mainly because of the less well developed equipment which is available for this type of work. Where the dielectric systems under study may be said to be linear in their response with

respect to the amplitude of the exciting electric field, so that the ratio $i(t)/V(t)$ remains constant, the validity of the Fourier transformation is assured and the time–domain evidence should represent a fully complementary set with respect to the frequency–domain data. To that extent, it is merely a matter of experimental convenience which method is being used. Our experimental evidence shows, however, that the application of even modest voltages to a dielectric capacitor system may result in perceptible departures from linearity, such as the "crossing over" of the charging and discharging currents which clearly point to the presence of some injection processes that would be very likely to remain completely undetected by frequency–domain measurements. Likewise, at very long times, there is evidence of processes setting in which may not be simply the direct current conduction limit in the case of charging currents and which represent the equivalent of the strong low frequency dispersion showing as a slowly falling current at long times.

We have already pointed out the value of the time–domain representation in the understanding of the true nature of the dielectric loss peaks in the frequency domain – as the transition between two sequential processes of different physical nature. This is an important element in our understanding of the new approach to the dielectric response of solids – not as an accidental superposition of Debye-like processes, but as a well defined sequence of physical processes which are characterised by a very specific form of time dependence (Jonscher 1975).

The Fourier transform of the universal law (5.8) obeyed at frequencies in excess of the loss peak frequency ω_p is

$$i(t) \propto t^{-n} \qquad 0 < n < 1, \; t \ll 1/\omega_p \tag{6.13}$$

and this is the well-known empirically established Curie–von Schweidler law formulated over seventy years ago (von Schweidler 1907).

On the other hand, the Fourier transform of eqn (5.12) valid below ω_p is:

$$i(t) \propto t^{-1-m} \qquad 0 < m < 1; \qquad t \gg 1/\omega_p \tag{6.14}$$

so that these two power laws determine the time–domain response of dipolar systems in which a loss peak is seen in the frequency domain.

Furthermore, the behaviour of carrier-dominated systems with their strong low-frequency dispersion below a frequency ω_c may be described by the same type of power law as eqn (6.13) but with a very small value of n. It may be advantageous for the purpose of subsequent discussion to introduce the notation $n = 1 - p$ with p close to unity, to denote the low-frequency dispersion regime, in which case the long-time response of charge carrier systems will be denoted by

$$i(t) \propto t^{-1+p} \qquad p \simeq 1; \qquad t \gg 1/\omega_c \tag{6.15}$$

corresponding to a very slowly time-varying current, as shown in Figure 6.14.

The complete representation of the universal dielectric response in the time–domain, covering both dipolar loss peaks and the strong low-frequency dispersion associated with the charge carrier dominated systems may be represented in the unified form of the depolarisation current as a function of time:

$$i(t) \propto t^{-s}; \qquad 0 < s < 2 \tag{6.16}$$

with the exponent s taking values in different ranges at "long" and "short" times, respectively, as shown in Table 6.2.

TABLE 6.2
The values of the exponent s in the universal relation $i(t) \propto t^{-s}$ corresponding to various physical polarising systems and to "short" and "long" times, respectively.

System	"short" times $t \ll 1/\omega_{c,p}$	"long" times $t \gg 1/\omega_{c,p}$
Dipoles	$s = n$	$s = 1 + m$
Charges		$s = 1 - p \qquad p \lesssim 1$
Near-Debye	$s \to 0$	$s \to 2$
Debye	$\exp(-\omega_p t)$	

These universal dielectric relationships are shown schematically in Figure 6.15 which distinguishes between charge carrier and dipolar systems and also between "short" and "long" times. Several points need stressing in this connection.

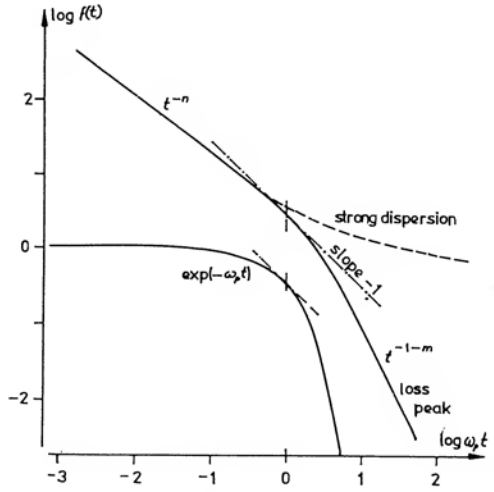


Figure 6.15 The two alternative forms of the universal dielectric response of solids in the time domain, corresponding to the two types of polarising species which may be present. Both follow the same power law at "short" times with a logarithmic slope smaller than −1, while at "long" times dipolar systems give a steeper power law and charge carrier systems obey a much slower time dependence, as shown by the dotted line. The exponential Debye response is shown for comparison.

a) The presence of two power laws

We have shown on the basis of convincing experimental evidence both in the frequency and in the time–domain, that the universally observed form of dielectric response which is followed in *all* dielectric systems, regardless of their detailed physical and chemical properties, consists of *two power laws* which follow sequentially in the two regions of time. There is an intermediate law joining these two regions, the precise form of which cannot be easily obtained from the available experimental data and which will be described in detail in the theoretical discussion in Chapter 8.

We note that it is not possible to give an adequate representation of the dielectric response on the basis of a single mechanism only, corresponding to the use of a single parameter, e.g. in the Cole–Cole, Cole–Davidson and Williams–Watts expressions.

b) The temperature dependence of the universal law

All experimental evidence clearly points to the fact that the critical transition time $1/\omega_p$ in dipolar systems or $1/\omega_c$ in charge carrier systems are strongly temperature dependent, showing for the most part some form of activated behaviour with a well defined activation

energy, although more complicated forms of dependence are seen in the case of the α relaxation processes in polymers and also in the response at very low temperatures.

By contrast with this strong temperature dependence of the transition time, the exponents m and n or p determining the power law behaviour are relatively very weakly temperature dependent. The only strong dependence may occur in the vicinity of any transition temperatures such as between the glassy and supercooled liquid states in polymers.

c) Limiting forms of response at "zero" and "infinite" times

The power law t^{-n} diverges at zero time and it is essential to note the physical limitation of its validity at sufficiently short times which ensures that the mathematical singularity does not arise. This limitation is provided by several processes which are not covered by the existing experimental evidence, on which our empirical laws are based. One of them is the presence of quantum processes, involving either phonon or photon interactions which become apparent in the frequency range in excess of THz and which would manifest themselves in the frequency–domain by strong loss peaks and dispersions of the real part, while their time–domain counterpart would be an oscillatory dependence with some damping. The second limitation, which would be expected to be observable in a similar range of frequencies and times is the presence of inertial effects which are again experimentally inaccessible. There is a third limitation, based on the strict theoretical treatment of the model explaining the power laws, and this will be discussed in Chapter 8.

The important conclusion at the present moment is that the presence of a singularity at zero time in the universal relation (6.16) is not a bar to its validity at longer times, since the power law is not intended to apply at arbitrarily short times. The power law describes extremely well the behaviour of dielectrics in a very wide range of times and there is no need to insist that *the same law* should necessarily be valid at arbitrarily short times, since the theoretical treatment to be given in Chapter 8 takes care of the short-time effects.

A more serious difficulty arises with respect to the carrier response corresponding to the dotted line in Figure 6.15. This clearly leads to a divergence of the total charge stored in the system and this is physically impossible. Experimental evidence shown in Figure 6.14 indicates that the nearly flat behaviour is followed by a steep drop which eliminates this problem of infinite charge, and we do not at

present understand sufficiently well the full implications of this behaviour. With regard to the corresponding frequency–domain response, all the examples shown in Chapter 5, and most examples known to us, show a continuing rise of ε' and ε'' down to the lowest attainable frequencies so that at the present time we have to accept that insufficient experimental basis exists for pronouncing definitively on the long-time behaviour in carrier systems. However, once again, this does not invalidate the *experimentally well established fact* that the behaviour described does exist and is seen over wide ranges of both time and frequency.

d) The Debye "singularity"

We stress again the singular nature of the classical Debye response, which does not correspond to any mathematical limit of the universal law but represents a completely different form of law relating to completely different physical circumstances. The Debye law was derived on the assumption of *non-interacting* dipolar or charge carrier species. Experimentally we have shown beyond reasonable doubt that this form of behaviour does not exist in real systems and we therefore conclude that this basic assumption is not satisfied in condensed matter. Our theoretical treatment in Chapter 8 will propose a completely different interpretation in terms of many-body processes, which are not only plausible but manifestly inevitable in condensed matter. Furthermore, we note that attempts to move away from interaction, for example by diluting the polarising species to the point where their interactions might become negligible, does not lead to experimentally verifiable Debye behaviour, since the magnitude of the signal becomes too weak to measure with precision, Figure 5.2.

e) Time–domain response of the polarisation

Our discussion of the time–domain response of dielectrics has concentrated so far on the step-function response of the depolarisation *current* which is given by the response function $f(t)$. This choice is dictated by the practical consideration of convenience of measurement – it is far easier to measure the current than the polarisation. Nevertheless, it is important to have a clear idea of the behaviour of the polarisation itself, since this is the primary physical variable, rather than its time derivative which is the polarisation or depolarisation current.

Assuming that the universal law given by eqn (6.13) is valid in the entire physically accessible time range, we may write the following

expressions for the polarisation in the charging mode:

$$P(t) = \varepsilon_0 E_0 \int_0^t f(t)\,\mathrm{d}t = \frac{A}{1-n} t^{1-n} \tag{6.17}$$

and in discharge:

$$P(t) = \varepsilon_0 E_0 \int_t^\infty f(t)\,\mathrm{d}t = \text{const} - \frac{A}{1-n} t^{1-n} \tag{6.18}$$

We note that this formulation involves a divergence of the polarisation at infinitely long times but this is not a serious problem for the present discussion.

The functional dependence of t^{1-n} is shown schematically in Figure 6.16 for a range of values of the exponent n. The limit $n \to 0$

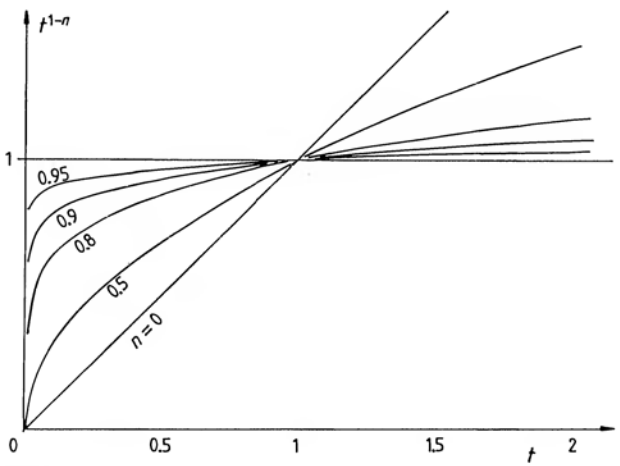


Figure 6.16 The function t^{1-n} for a range of values of the exponent n. The approximation $t^{1-n} \cong 1 + (1-n)\ln t$ is valid for values of n close to unity, over a wide range of values of t.

corresponds to strong low-frequency dispersion, with its tendency for the polarisation to grow indefinitely. The other interesting limit is that of the "flat" loss, $n \to 1$, in which the polarisation under step–function charging builds up rapidly to an almost constant value and changes little thereafter.

The limit $n = 1$ represents a singular solution for which the polarisation should show a logarithmic time dependence, since

$$\int_a^t \mathrm{d}t/t = \ln t - \ln a \tag{6.19}$$

and this form of relaxation law is often seen in different contexts, especially in disordered systems.

While eqn (6.19) may represent a mathematical singularity within the family of solutions (6.17) for all values of $n \neq 1$, it is interesting to investigate in more detail the transition from the power law (6.17) to the logarithmic law (6.19) as the exponent approaches the critical value 1. To this end write

$$1 - n = \alpha \ll 1$$

Equation (6.17) now becomes:

$$P(t) = At^{\alpha}/\alpha \tag{6.20}$$

But

$$\ln t^{\alpha} = \alpha \ln t = \ln[1 + (t^{\alpha} - 1)] \cong t^{\alpha} - 1 \tag{6.21}$$

the last approximation being valid for $t^{\alpha} \cong 1$, or $\ln t \ll 1/\alpha$.

From (6.20) and (6.21) we obtain the relations for charging

$$P(t) \propto \frac{1}{\alpha} + \ln t \propto 1 + \alpha \ln t \tag{6.22}$$

and for discharging

$$P(t) \propto \text{const} - \alpha \ln t \tag{6.23}$$

We have thus derived the required relationship between the power law and the limiting logarithmic time dependence, showing that within the states range of time in which eqn (6.21) is valid, the logarithmic law may be indistinguishable from the power law solution.

APPENDIX 6.1

THE MINIMUM DURATION OF CHARGING AND DISCHARGING

In order to estimate the effect of finite charging times on subsequent discharge, and vice versa, let us assume that the material under study has a simple universal characteristic $f(t) \propto t^{-n}$ throughout the experimentally accessible range of time. We take the exact eqn (6.1) and we calculate the time t_e at which the logarithmic slope S reaches a given value.

$$S = \frac{\mathrm{d} \log i(t)}{\mathrm{d} \log t} = -n \frac{1 - z^{n+1}}{1 - z^n}$$

where

$$z = t/(T + t)$$

Setting this equal to -1 we solve numerically and obtain the values of t_e given in Table 6.3. The significance of this time is that, had the measurements been continued beyond t_e, the presence of a loss peak in the frequency–domain would have been inferred from the change of slope.

We may impose an even more stringent condition, that the slope should not change from the correct value $-n$ by more than, say, 10 or 20%. These times are also given in the Table.

All values in the Table are given in terms of the ratio T/t_e etc, i.e. the ratio of the *minimum* charging time required to the desired duration of discharge measurement. The same considerations apply to the measurement of charging currents.

TABLE 6.3

n	T/t_e	$T/t_{10\%}$	$T/t_{20\%}$
0.95	21.4		
0.9	10.8	12.3	
0.8	5.3	17.8	7.2
0.7	3.3	28.4	10.6
0.6	2.6	52	17.3
0.5	1.6	118	33

APPENDIX 6.2

TIME–DOMAIN RELAXATION AND DC CONDUCTIVITY

The time–domain relaxation according to the power law $i(t) \propto t^{-n}$ may interfere with the determination of dc conductivity σ_0 of a material, since a considerable delay may have to be allowed after the application of a bias voltage before the relaxation current falls below the steady state value of the current appropriate to the dc conductivity. To obtain a quantitative estimate of this effect we use equation (2.33), omitting the delta-function which is normally lost in the measuring process:

$$i_c(t)/E_0 = \varepsilon_0 f(t) + \sigma_0$$

The present analysis applies predominantly to low-loss materials, for which $n \cong 1, \chi''(\omega) = \text{const}, \tan\delta = \varepsilon_0\chi''/\varepsilon_\infty \ll 1$, with the result that $\varepsilon \cong \varepsilon_\infty$.

Equation (6.8) gives in these circumstances:

$$\chi''(\omega) \cong (1.5/\varepsilon_0)\, ti(t)/E_0 = \text{const}$$

where the current here is the depolarisation current, not including the dc component, i.e. it is equal to $\varepsilon_0 f(t)$.

The implication of this is that

$$i(t)/E_0 = \varepsilon_0 f(t) \cong \varepsilon_0\chi''(\omega)/t = \varepsilon_\infty \tan\delta/t$$

so that finally

$$i_c(t)/E_0 \cong \varepsilon_\infty \tan\delta/t + \sigma_0$$

Assuming the value $\varepsilon_\infty = 10^{-11}$ F/m, we obtain the diagram of Figure 6.17 which shows that, according to the particular combination of $\tan\delta$ and σ_0 it may be necessary to wait for a significant length of time before a reliable measurement of σ_0 can be made.

The above analysis applies to the limiting case of $n = 1$, corresponding to a frequency-independent loss. In the situation where this is no longer the case, the same analysis would still be applicable

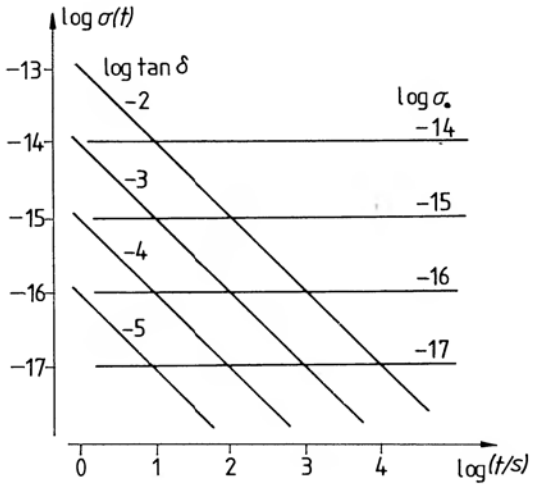


Figure 6.17 The "time-dependent conductivity", $\sigma(t) = i(t)/E$, in $(\Omega m)^{-1}$ for low-loss materials with a dc conductivity σ_0, assuming that the exponent n of the universal law is equal to unity and that the real part of the permittivity is independent of frequency and is $\varepsilon' = 10^{-11}$ F/m. In order to attain the dc conductivity indicated it is necessary to wait for a time that is sufficiently long to enable the corresponding discharge current to fall below the dc value. The parameter is tan δ which for the low-loss materials in question is equal to the angle δ itself (in radians).

if one took for tan δ the value appropriate to the frequency corresponding to the longest time at which the measurement is being made.

REFERENCES TO CHAPTER 6

Davies D K 1969, Brit. J. Appl. Phys. **2**, 1533
Fröhlich H and Platzman R L 1953, Phys. Rev. **92**, 1152
Hamon B V 1952, Proc. IEE 94, 151
Hill R M 1975, J. Phys. C: Solid State Physics **8**, 2488–2501
Hirsch J 1979, private communication
Hyde F J 1970, Proc. IEE 117, 1891
Jennings B R and Coles H J 1974, Nature **252**, No. 5478, 33–34
Jonscher A K 1973, *Electrets, Charge Storage and Transport in Dielectrics*, M M Perlman (Ed.), The Electrochemical Society, 269–284
—— 1975, Nature **256**, 566–568
—— 1980, J. Phys. D: Appl. Phys. **13**, L137
Jonscher A K and Buddhabadana S 1978, Solid State Electronics **21**, 991
Jonscher A K and Careem M A 1975, Phys. Letters **A55**, 257
Lampert M A 1965, *Injection Currents in Solids*, Academic Press, New York
Le Sueur E J and Jonscher A K 1972, Thin Solid Films **12**, S9

Meca F and Jonscher A K 1979, Thin Solid Films **59**, 201–219
Millany H M and Jonscher A K 1980, Thin Solid Films **68**, 257–273
Pfister G and Scher H 1977, *Proceedings of the International Conference on Liquid and Amorphous Semiconductors*, Edinburgh, W E Spear (Ed.), p. 197
—— 1978, Advances in Physics **27**, 747–798
Reddish W 1962, Pure and Applied Chemistry **5**, 723
Reddish W and Barrie I T 1958, IUPAC Symposium, Wiesbaden, p. 1 A1
Rose-Innes C 1981, in *Physics of Dielectric Solids*, C H L Goodman (Ed.), Institute of Physics Conf. Series No. 58, p. 122
Scher H 1977, *Proceedings of the International Conference on Liquid and Amorphous Semiconductors*, Edinburgh, W E Spear (Ed.), p. 209
von Schweidler E 1907, Ann. d. Physik **24**, 711

CHAPTER 7

Previously Accepted Interpretations

7.1 INTRODUCTION

The review of the experimental evidence relating to a wide range of dielectric materials, summarised in Figure 5.58, shows clearly the wide variation of types of dielectric response, whether involving loss peaks or strong low-frequency dispersion, whose only common feature is the presence of power–law relations between the loss and the frequency. The corresponding time–domain representation likewise involves power–law relations for the charging and discharging currents and it has been recognised as the generally observed type of behaviour since the beginning of the present century.

We saw in Chapter 4, on the other hand, that the response of most ideal physical systems involved a Debye-type exponential dependence on time and the corresponding Fourier transform into the frequency domain, with the exception of the resonant response and of the diffusive system. This concept of the Debye relaxation has somehow dominated the dielectric thinking for a very long time, despite the evident discrepancies in relation to the experimentally observed behaviour. It is surprising, perhaps, that in view of these discrepancies so much effort should have gone in the past into attempts to adapt the Debye philosophy to suit physical reality, rather than into devising essentially new methods of approach.

The purpose of the present Chapter is to give a brief review of the most important currently accepted interpretations of the dielectric response of materials, as a preparatory step to the introduction of our own alternative approach. It is not proposed to dwell in detail on the various theories, since they have been adequately presented elsewhere – a very recent and comprehensive review may be found in Böttcher and Bordewijk (1978), several other classic texts are listed in Chapter 1. Instead, our aim will be to present a critical appraisal of these various theories, pointing out in what respects they fall short of the requirements imposed by the inescapable experimental evidence.

A newcomer to the subject of dielectric relaxation might be excused the feeling of a certain surprise at the multiplicity of different models, all trying to explain what might strike him as the fairly common patterns shown in Figure 5.58. This feeling of surprise arises, however, only after a thorough review of *all experimental evidence* covering a *wide range of materials*. In fact, however, most dielectric studies in the past have been carried out in comparative isolation between the different branches, usually centred on groups of materials such as polar liquids, ceramics, glasses, low-loss polymers, ionic conductors, amorphous electronic conductors etc. It is not difficult to see how, within a narrow sector of materials and of types of response, a method of presentation and of interpretation becomes accepted which looks plausible enough. Reference across these self-imposed subject boundaries was rendered the more difficult for the acceptance in these various groups of different and often incompatible methods of presentation of data. A typical example is the acceptance of the complex permittivity plots in the context of polar liquids and the equally exclusive use of complex impedance plots by fast ionic conductor workers. Departures from straight circular arc plots are then explained by both groups in terms of distributions of relaxation times, while our analysis shows that these two modes of response rely on physically entirely different mechanisms. Furthermore, experimental data were often confined to relatively narrow ranges of frequency, making the "fitting" of more or less arbitrary theories fairly easy. Only by looking at a frequency range of more than, say, five or six decades, consisting of the actual measuring range extended by normalisation of data from different temperatures, can the extent of agreement or its lack be properly assessed.

Not the least cause of a lack of common appreciation of the nature of the dielectric response was the prevailing habit of plotting the data semi-logarithmically, which meant that the all-pervasive power–law relations were being missed, all loss peaks looking as somewhat broadened versions of the Debye shape, but without any sense of order appearing. Only by plotting the data in a log–log representation, as we are doing, can the true universality of behaviour be seen without difficulty.

Even the acknowledged validity of the Curie–von Schweidler time-domain law for many dielectric materials was not sufficient to prompt a serious enquiry into the possible alternative explanation of the power–law relation, eqn (6.13).

For the purpose of our discussion we shall divide the various theories into the following broad categories:

a) distributions of relaxation times (DRT's);
b) distributions of hopping probabilities;
c) correlation function approaches;
d) local field theories;
e) diffusive boundary conditions – Warburg impedance;
f) interfacial phenomena and the Maxwell-Wagner effect;
g) transport limitations at the boundaries.

These will be briefly discussed in the following sections.

7.2 DISTRIBUTIONS OF RELAXATION TIMES (DRT's)

These theories (Garton 1946, Frohlich 1955, Macdonald 1962a, b, 1963) are among the earliest and, in some ways, the least sophisticated attempts to interpret evident departures from the ideal Debye behaviour, and although they are not regarded as a panaceum by the more critical workers, their hold on the more popular front is very strong, to the point where for many people the DRT is a self-evident explanation not requiring closer justification. The attraction of DRT's lies in the eminent plausibility of the fact that there should exist a distribution of relaxation times in any solid material, since there is likely to be some non-uniformity of local situations of the individual dipoles or charges. Furthermore, the assertion of a DRT is impossible to prove as being either right or wrong – there is usually no independent way of showing whether a particular form of DRT is applicable in a given case. This places the proponents of this approach in a favourable position and any argument usually does not get very far.

Mathematically the summation of Debye responses corresponding to a distribution of relaxation times is carried out in terms of integration of a function $g(\tau)$ defining the distribution:

$$\chi(\omega) = \int_0^\infty \frac{g(\tau)}{1 + i\omega\tau} d\tau \tag{7.1}$$

This may be regarded as a special form of integral transform, analogous to Fourier, Laplace or Hilbert transforms, through which one function, in this instance the Debye function given by eqn (4.11), is transformed into a different function, the desired empirically observed response function $\chi(\omega)$, with the aid of a suitable kernel which in this instance represents the physical concept of the distribution of relaxation times.

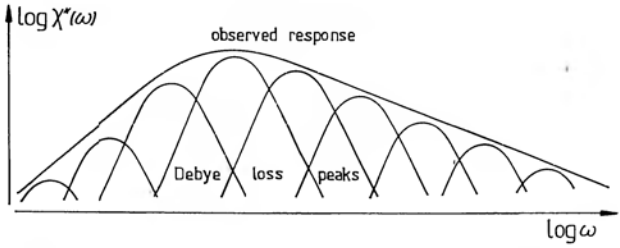


Figure 7.1 A schematic representation of the significance of the existence of a distribution of relaxation times in the application to the interpretation of a "non-Debye" loss peak in a dielectric material. The distribution could be discrete or quasi-continuous.

This operation may be represented graphically by the construction shown in Figure 7.1 relating to the frequency domain, or by the corresponding time-domain representation shown in Figure 7.2. It is evident that within very wide limits, any sufficiently slowly varying function of frequency $\chi(\omega)$ may be represented by the integral in eqn (7.1), so that virtually all forms of relaxation behaviour can be "explained" by this means. On the other hand, it would be impossible to represent in this way resonance phenomena such as that shown in Figure 4.1.

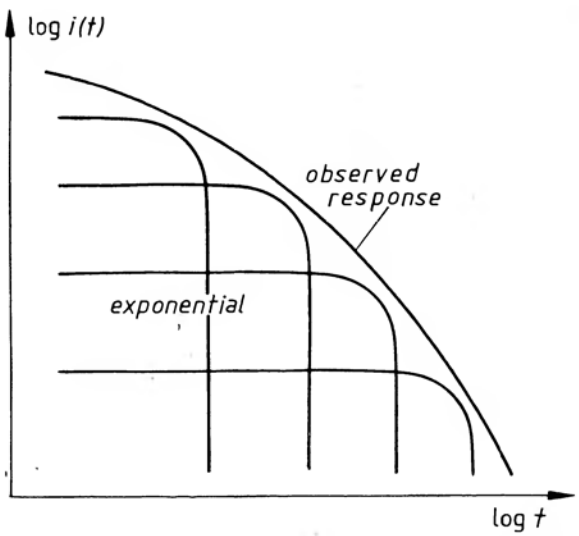


Figure 7.2 A schematic illustration of the distribution of relaxation times in the time–domain corresponding to a fitting of exponentials of the form $\exp(-\omega_p t)$ with a range of loss peak frequencies ω_p to simulate a given response.

In the limit of sufficiently wide loss peaks, the functional form of $g(\tau)$ follows reasonably close that of $\chi(1/\omega)$, as may be seen from an examination of Figure 7.1. Considerable efforts went in the past into the mathematical inversion of eqn (7.1), which would yield the analytical form of $g(\tau)$ for a given input function $\chi(\omega)$. In the case of analytical expressions for $\chi(\omega)$, no exact analytical procedure exists analogous to the inversion of Fourier or Laplace integrals, while approximate procedures for treating empirically given susceptibility functions are available. It is important to bear in mind, however, that any effort spent on this operation is only justified to the extent to which it is thought that the information obtained in this manner has specific physical significance and justification. Merely to obtain a distribution of relaxation times as a means of characterising a given susceptibility function does not advance the true understanding of the situation, since this DRT function cannot, by the nature of its derivation, contain any more information than does the experimentally determined susceptibility function from which the analysis started.

A complete *physical interpretation* of the situation, as distinct from the purely mathematical formulation of the problem in terms of eqn (7.1) would now require a discussion of the reasons for this particular form of distributions of relaxation times in any given material and it is here that difficulties start. There are many good reasons why distributions of relaxation times should arise in solids, the most obvious being the presence of inhomogeneities both on the macroscopic and on the microscopic or atomic levels. Thus, considering a particular polymeric material, it is evident that not all dipoles present in it are likely to find themselves in exactly the same environment, with some being more free to rotate than others. Similarly, the dipoles present even in single crystal and well ordered materials may find certain orientations more favoured than others and certain transitions between these orientations may be easier than others. These variations of local transition probabilities would reflect themselves not only in the variation of the absolute values of the relaxation times τ, but also in the corresponding variations of the activation energies W and of the pre-exponential factors ν in eqn (4.29). The importance of the variation of activation energies arises in connection with the temperature dependence of the dielectric susceptibility, as may be seen from the following considerations.

With reference to Figure 7.1, we note that as the temperature changes, the variation of the individual relaxation times causes a

shift of the respective loss peaks on the logarithmic frequency axis with the relevant activation energies. We may now consider two extreme situations. Either all activation energies are the same and the entire distribution of relaxation times is caused by a variation of ν, in which case the response at different temperatures is shifted bodily in frequency, without any change of shape, implying that normalisation is complete, as for example in Figures 3.27 and 5.21. Alternatively, the variation of τ is caused entirely by the variation of the activation energy W, in which case the resulting response of dielectric loss would suffer a change of shape with varying temperature. Of these two, the later hypothesis is more plausible, since there does not appear to exist any very good physical reason why the "attempt-to-jump" frequency ν should vary, in some cases over many decades with the same activation energy.

A much more likely physical behaviour would consist in an interdependence between the values of ν and of the energy W, as shown in Figure 7.3 which illustrates the so-called *compensation rule* which applies to many activated processes and which requires that the smallest values of the activated quantity have the highest activation energies. The consequence of this type of behaviour would be a narrowing of the loss peaks with increasing temperature, which is in qualitative agreement with some of the observed responses (Jonscher 1977).

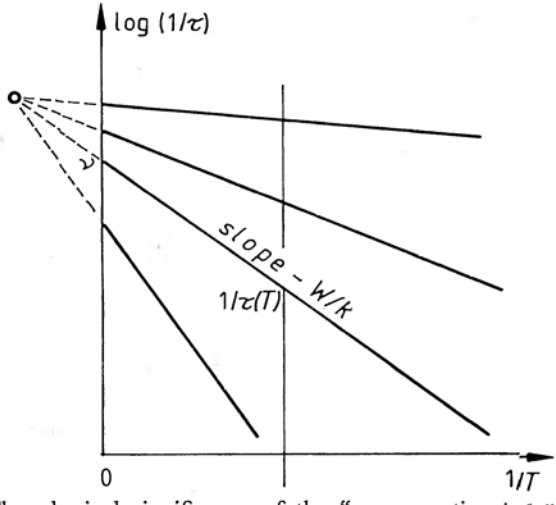


Figure 7.3 The physical significance of the "compensation rule", linking the pre-exponential factors ν and the corresponding activation energies W in eqn (4.29). To obtain a range of relaxation times $\tau(T)$ at any given T one has to take different and *corresponding* values of energies and pre-exponential factors.

Attempts have been made in a few specific cases where the crystal structure was known with sufficient detail to enable the calculation of transition probabilities to be made, to obtain the shape of the spectrum from first principles (Lauritzen 1958). A measure of agreement was obtained with experimental data, with one or two peaks and in a limited range of frequencies in the 10^8–10^9 Hz range. In the light of the discussion to be given in Section 8.4, this type of calculation may be said to refer to the principal thermally-activated transitions, while saying nothing about the many-body relaxations which determine the spectral shape of the loss peak outside the region of the loss peak frequency. To that extent, therefore, these calculations may be said to leave the central problem of large deviations from the Debye shape out of consideration. Certainly no attempt to calculate from first principles the shape of a really broad peak has come to our notice.

With reference to the time-domain representation of Figure 7.2, we may write down an analogous expression to (7.1) in terms of a distribution of *loss peak frequencies*, $h(\Omega)$:

$$f(t) = \int_0^\infty \exp(-\Omega t)\, h(\Omega)\, \mathrm{d}\Omega \tag{7.2}$$

which is immediately recognisable as the Laplace transform of $h(\Omega)$. The inversion of this function to give the required distribution $h(\Omega)$ for a given response function $f(t)$ is relatively simple, since tables of inverse Laplace transforms are readily available (Erdelyi 1954) and it is certainly much simpler than the corresponding inversion of eqn (7.1) which involves manipulation of Stietjes integrals (Böttcher and Bordewijk 1978, p. 523).

Comparing the definitions (7.1) and (7.2) and using the Fourier transform relation (2.37) it is easy to show that

$$h(\omega_p) = \omega_p^{-1} g(1/\omega_p) \tag{7.3}$$

which is dimensionally correct since $g(\tau)$ has the dimensions of reciprocal time, while $h(\omega_p)$ is dimensionless.

To illustrate the principle, consider first the trivial case of a Debye response for which the function $f(t) = \exp(-\omega_p t)$. Inspection of eqn (7.2) shows that the distribution function is a delta function centred on the required frequency ω_p, $h(t) = \delta(\Omega - \omega_p)$, which is physically self-evident. Taking next the limiting form of a "flat" response in frequency, the required response function is $f(t) = 1/t$, the inverse Laplace transform of which is $h(\omega_p) = \text{const}$, corresponding to a "flat" distribution in frequencies ω_p.

Next note that in our representation on a logarithmic time axis, the contributions to the current $f(\log t)$, with equal logarithmic intervals in loss peak frequency ω_p, must have the "strength" or amplitude given by

$$h(\omega_p)\,\mathrm{d}\omega_p = \text{const}\,\mathrm{d}\omega_p \propto \omega_p\,\mathrm{d}\omega_p/\omega_p = \omega_p\,\mathrm{d}\log\omega_p$$

so that the amplitude should go as ω_p or as $1/t$. Comparable relations would be obtained for other forms of distributions. In the general case of the universal fractional power law time dependence t^{-n}, the inverse Laplace transform of this is $\omega^{n-1}/\Gamma(n) \propto h(\omega_p)$. The corresponding distribution in logarithmic intervals is therefore:

$$h(\log\omega_p) \propto \omega_p^n \propto t^{-n}$$

Thus the logarithmic frequency distribution follows the time–dependence of the relaxation current, as shown schematically in Figure 7.2.

One of the principal difficulties in assessing the validity of the DRT approach consists in the impossibility of *proving* that a given form of $g(\tau)$ is actually *wrong*, while equally, the protagonists of this approach cannot prove that their distributions are in fact correct. The most that can be said is that a given experimental result is *compatible with* the presence of a certain DRT, which does not imply in the least that this is the correct solution.

Our main argument against the DRT philosoply rests essentially on the universality of the dielectric response of so many different materials, all giving the same type of lower–law dependence on frequency and on time. Any serious proof of the validity of DRT's should, therefore, give a more general justification of the universality in terms of these distributions and should explain the independence of the response from the detailed physical properties of these very different media. No such generalisation has ever been proposed, to our knowledge, and for this reason we regard this approach as not objectively acceptable in the face of the experimental evidence at our disposal. This does not alter the fact that many people in the applied area of dielectrics science, especially in Engineering, are very attached to this type of interpretation.

7.3 DISTRIBUTIONS OF HOPPING PROBABILITIES

Just as a localised electronic or ionic charge hopping between two or more preferred sites represents an exact counterpart of a dipole jumping between preferred spatial orientations, so the dipolar DRT

approach has an exact counterpart in the equivalent theories which consider the observed results as arising from corresponding distributions of hopping probabilities. These theories became especially popular in the context of amorphous and glassy electronic semiconductors (Pollak 1971, Austin and Mott 1969, Scher and Lax 1973, Butcher and Morys 1973, Butcher and Ries 1981) but they are also being invoked in the context of ionic semiconductors.

While the level of mathematical sophistication applied to these hopping theories is higher than that found in the dipolar DRT's, the same fundamental objections apply to both – the approach is essentially a one-particle treatment and there are no means of determining the expected distributions from first principles or from independent experimental evidence. Once again, if these distributions were really responsible for the observed behaviour shown, for example in Figure 5.33 for very different materials, it would be necessary to carry out a generalisation of the theory to explain why the same form of behaviour is seen in all.

The uncritical acceptance by many workers of the distributions of hopping probabilities goes to the point where it is confidently stated that the presence of a power–law frequency dependence of conductivity following eqn (3.32) with the exponent in the range $0.6 < n < 0.8$ constitutes in itself a *proof* of the applicability of the hopping mechanism. This simplistic view ignores the fact that many other materials exist which are not hopping electronic conductors and which show exactly the same type of frequency dependence.

One class of hopping materials which obey the same type of frequency dependence are the ionic conductors, especially the so-called fast ion conductors, examples of which were given in Chapter 5. Once again, the generally accepted "theories" relating to the behaviour of these materials invoke more or less well defined "distributions" of some parameters, even though the interpretations in question are mainly concerned with the interpretation of inclined circular arcs in the complex impedance plane, as shown in Figure 5.36. We have shown that this type of behaviour is exactly equivalent to our "universal" dielectric response and it would therefore be very surprising if the same type of distribution of relaxation times were found in these as in the electronic hopping conductors.

One important point relating to all forms of distribution functions of either dipolar or particle relaxation times is the fact that, with the entire body of experimental evidence in mind, we never see any indication of a *finite range* of relaxation times. The reason for this

is that the summation of Debye-like response functions covering a finite range of relaxation frequencies between, say, $\Omega_1 < \omega_p < \Omega_2$, should produce a dependence of loss as $\chi''(\omega) \propto \omega$ for $\omega \ll \Omega_1$ and $\chi''(\omega) \propto 1/\omega$ for $\omega \gg \Omega_2$. This type of behaviour is conspicuously absent even in dielectric measurements covering many decades of frequency.

Under the general heading of distributions of probabilities and relaxation times one should mention the fact that the purely empirical expressions due to Cole and Cole, eqn (3.37), and to Cole and Davidson, eqn (3.38), are often "explained" in terms of not clearly specified distributions of some parameters, without any particular attempts being made to specify the physical nature of these distributions. The numerical values of the exponents α and β are sometimes taken to be "measures" of the width of the distributions involved. Needless to say, these arguments cannot be regarded as serious interpretations in any physically rigorous sense – at best they may be considered as means of qualitatively characterising the type of response.

7.4 CORRELATION FUNCTION APPROACHES

The correlation function is a mathematical expression of the time-evolution of the perturbation arising from an initial disturbance of a dipole from its equilibrium orientation, $\mu(0)$. If $\mu(t)$ is the orientation at a time t, then the correlation function is defined as the average over all dipoles of the product:

$$C(t) = \langle \mu(0) \, . \, \mu(t) \rangle / \mu^2 \qquad (7.4)$$

The analysis of correlation functions has been carried out extensively, often at a high level of mathematical refinement (Glarum 1960, Cole 1965, Nee and Zwanzig 1970, Williams 1972, Cole 1973, Fulton 1975). An excellent recent review may be found in Cole (1980).

Correlation function approaches may be divided into two basic types, one being purely descriptive, i.e. describing the nature of the time-dependent function in eqn (7.4) by means of some mathematical approximation, without at the same time producing any specific justification of this on physical grounds. This aspect is not very helpful in the understanding of what determines the physical nature of dielectric relaxation.

A second type of approach begins from first principles for a given physical model and it may be said to represent a rigorous analysis, within the limitations of the necessarily simplified model which is postulated at the outset. Some recent studies are concerned with numerical solutions of systems not admitting of closed-form analytical solution. It turns out, however, that those models that are tractable analytically or even numerically, relate only to one-particle processes, i.e. to auto-correlations, where the dipole in question is assumed to interact only with itself. Once any form of many-body interactions is introduced through cross-correlations between different dipoles, the mathematical difficulties become insurmountable.

One method of treating many-body interactions within the scope of correlation function approach consists in the application of the Ising model to linear-chain systems (Bozdemir 1981a, b). This method leads to finite distortions from the Debye shape and to considerable modification of the single-particle relaxation frequencies.

Once again, however, the general conclusion from these various approaches is that, while they may successfully describe slight departures from the Debye ideal, they do not give any proper insight into the more serious departures and they are unable to account for the observed universality of response.

7.5 LOCAL FIELD THEORIES

A very considerable amount of theoretical effort has gone over the years into the proper formulation of the *local* situation arising at the site of a particular dipole or charge under the influence of an external electric field and also taking into account the interactions of the neighbouring dipoles or charges. Most of this work is concerned with the *static polarisation*, where the various theories described in Section 2.4 attempt to represent the physical situation in terms of closed-form mathematical expressions.

For all their possible usefulness in the treatment of static behaviour, the application of local field theories to polarisation response in time-varying electric fields has not led to any striking agreement with the empirically determined behaviour. While slight departures from the Debye model could be explained on this basis – as also on the basis of several other theories, it is difficult to see how more serious departures could be treated in this way. Accordingly, we

do not propose to give any more detailed treatment of this subject, especially since it is difficult to find such treatments in most standard dielectric textbooks.

7.6 DIFFUSIVE BOUNDARY CONDITIONS

Section 4.8 gave the description of the diffusive transport in the particular example of a forward-biased p–n junction. Very similar considerations apply to the case of an interface between an electrolyte and an electrode, provided that there exist two mobile carrier species of opposite signs, so that neutrality may be preserved in the presence of a density gradient (Jonscher 1975, 1978).

The great advantage of the diffusive model (Warburg 1899, 1901, Mitoff and Charles 1972) lies in the fact that, uniquely amongst all the other physical models, it provides a power–law type of dependence of frequency and on time, breaking away drastically from the ubiquitous Debye relationship predicted by all the other models. The half-power law so obtained is at least in the same general category as the universally observed fractional power–law and this accounts for some of the attraction this model has, especially in the context of fast ion conductors.

These advantages notwithstanding, it has to be stated that the usefulness of this model is severely limited by the fact that in practice the particular value of the exponent $n = \frac{1}{2}$ is the exception rather than the rule and that the model itself is quite rigid in not admitting any significant departures from this value. For example, there is no way in which one can envisage any form of superposition of several diffusive mechanisms to give an overall effect of an exponent departing from $\frac{1}{2}$. We therefore conclude, that while this model may well represent a plausible means of interpreting the dielectric properties of a narrow class of materials which both satisfy its physical assumptions and give the particular power–law with $n = \frac{1}{2}$, it has to be said that it cannot provide a generally applicable model to explain the behaviour of the great majority of carrier-dominated dielectrics.

Under the heading of diffusive models we mention another attempt to combine the diffusive process with dipolar phenomena, proposed by Glarum (1960) who postulated the existence of a "triggering" mechanism for the dipolar reorientations. According to this model, the dipole could not undergo a transition from one orientation to another without the intervention of a "defect" executing a random

diffusive motion. It is not clear what the nature of this defect might be, nor what is the precise form of triggering, but for our present purposes it is sufficient to state that the only result of this process is to introduce the half-power law into the dipolar processes, so that this theory cannot be considered to constitute a significant addition to the catalogue of generally applicable processes.

7.7 INTERFACIAL PHENOMENA AND THE MAXWELL–WAGNER EFFECT

Strongly dispersive behaviour at low frequencies had been noted for a long time in many dielectric materials and it has attracted a considerable amount of attention, especially on the part of electrical engineers. The simplest form of interfacial effect is represented by a capacitive layer arising near an electrode as a result of the formation of a Schottky barrier which is less conducting than the bulk material of the sample. The effect of this on the dielectric behaviour of the structure as a whole has already been described in Chapter 3, where strong dispersions were shown to be possible if the ratio of the bulk capacitance to the capacitance of the barrier is sufficiently small. However, the frequency dependence of this series combination is seldom in agreement with the observed behaviour of strong dispersion, which may be typically of a power–law form with an exponent of the order of -1, as shown in Chapter 5. The often used interpretation consists then in assuming a whole series of such barrier phenomena (Volger 1960, Haberey and Wijn 1968), or even a complete series-parallel array of barrier-volume effects, such as might arise if the material consisted of grains separated by more insulating inter-grain barriers.

This kind of inhomogeneous medium effect is known under the general name of Maxwell–Wagner effect (Maxwell 1954, Wagner 1913) and it can be developed into a whole distribution of interfacial effects, providing a direct counterpart to the distributions of Debye-like relaxation times in the volume of a dipolar material. This model therefore, has all the superficial attractions, as well as all the drawbacks of the distribution theories – an apparent plausibility coupled with the virtual impossibility of proving conclusively that it is either right or wrong. We believe that the experimental evidence for the power–law dispersion with exponents approaching the value -1 is sufficiently well established to require here again a more general form of interpretation explaining why a power–law should be found other than by a purely accidental combination of barrier parameters.

7.8 TRANSPORT LIMITATION AT THE BOUNDARIES

This approach is concerned with a different aspect of interfacial processes, this time with the partial limitation of charge movements across interfaces between electrodes and the electrolyte or between different parts of an inhomogeneous material (Macdonald 1974, 1976). It therefore represents an extension and a refinement of the simple Maxwell–Wagner model, allowing for partial transmission of charge species, typically ions rather than postulating completely blocking barriers. These models were studied in conjunction with ionic conductors which resemble closely very dense liquid electrolytes. Apart from considerable mathematical difficulties arising from the need to handle several simultaneous differential equations with boundary conditions, describing the motions of the various species, there is the inherent difficulty of knowing independently the various transmission parameters which effectively constitute disposable parameters.

It would appear that these models are capable of explaining the response of some strongly interface-dominated systems in which the bulk material possesses a sufficiently high electrical conductivity to appear effectively like a pure resistance in series with a barrier. On the other hand, it would not be possible by this means to explain the fractional power–law behaviour of bulk ionic conductors, such as those shown in Figures 5.36 and 5.45 which are typical of many other cases.

7.9 THE NEED FOR AN ALTERNATIVE APPROACH

We have given this brief account of the various hitherto accepted interpretations of the form of frequency dependence of the dielectric response in order to demonstrate to the reader that many theoretical approaches have been developed over the years in attempts to account for the large variety of observed types of dielectric response. They suffer, however, either from lack of plausibility or from a very restricted range of applicability to specific systems, or finally they are incapable of accounting for the more serious departures from the Debye behaviour observed experimentally.

Of the reasons for these shortcomings we have already mentioned the far-reaching fragmentation of dielectric research into specialised areas not communicating with one another and also a rather uninformative way of presenting dielectric data in a manner which fails to bring out the important features of the behaviour. These twin

reasons account for the lack of any overall integrating approach to the treatment of dielectric relaxation, since there did not appear to exist any strong reason for attempting such an approach. We believe that the classification given in Chapter 5 and based on a far-reaching analysis of a wide range of materials is fundamental in focussing attention for the first time on the *universality* of the dielectric response which not only does not follow Debye, but instead follows a very well defined power–law relationship. Once this universality became apparent, the quest for an all-embracing theoretical interpretation became imperative, since it was too much to expect that the various sectional models which might even be plausible in their respective contexts should all give the same universal law.

In addition to the universality of the frequency dependence of the dielectric susceptibility there was also the evidence of the persistence of strong dielectric activity down to the lowest temperatures, which clearly calls for the existence of dominant non-activated processes, i.e. some form of tunnelling phenomena which would be applicable to heavy ions and dipoles as well as to electrons. This alone would have called for a serious re-examination of the currently accepted models, since they all involve thermally activated processes or, at most, electronic or protonic tunnelling.

Our search for an alternative approach to the interpretation of the dielectric relaxation phenomena in solids was based, therefore, on the conviction that the observed universality of response, expressed in terms of power-law relations in frequency and in time, required a more generally applicable and physically uniform theoretical framework than the largely fragmented approaches described above. This conviction went deeper than the mere concept of "elegance" of a generalised theory replacing a multitude of specialised ones – it was rooted in the concept that a simple experimental result demands an equally simple physical model which is not based on a range of more or less arbitrary assumptions.

We shall be showing in Chapter 8 that the new many-body model of dielectric relaxation has all the advantages of simplicity and generality demanded by the situation and that it is capable of explaining within a single framework the totality of experimental results.

REFERENCES TO CHAPTER 7

Austin I G and Mott N F 1969, Adv. Phys. **18**, 41
Böttcher C J F and Bordewijk P 1978, *Theory of Electric Polarisation*, Second Edition, Vol. II, Elsevier
Bozdemir S 1981a, phys. stat. sol. (b) **103**, 459
—— 1981b, phys. stat. sol. (b) **104**, 37
Butcher P N and Morys P 1973, J. Phys. C **6**, 2147
Butcher P N and Ries B 1980, Phil. Mag. **B44**, 179
Cole R H 1965, J. Chem. Phys. **42**, 637
—— 1973, Mol. Phys. **26**, 969
—— 1980, *Physics of Dielectric Solids*, C H L Goodman (Ed.), IoP Conference Series No. 58, p. 1
Erdelyi A 1954, *Tables of Integral Transforms*, Vol. I, McGraw-Hill, New York
Fröhlich H 1955, *Theory of Dielectrics*, Oxford University Press
Fulton R L 1975, J. Chem. Phys. **62**, 4355
Garton C G 1946, Trans. Far. Soc. XLII A, 56
Glarum S H 1960, J. Chem. Phys. **33**, 1371
Haberey F and Wijn H P J 1968, phys. stat. sol. **26**, 231
Jonscher A K 1975, phys. stat. sol. **32**(a), 665
—— 1977, J. of Electrostatics **3**, 53
—— 1978, J. Materials Science **13**, 553
Lauritzen J I 1958, J. Chem. Phys. **28**, 118
Macdonald J R 1962a, Physica **28**, 485
—— 1962b, J. Chem. Phys. **36**, 345
—— 1963, J. Appl. Phys. **34**, 538
—— 1974, J. Chem. Phys. **61**, 3977
—— 1976, *Electrode Processes in Solid State Ionics*, M Kleitz and J Dupuy (Eds.), Reidel, Dordrecht
Maxwell J C 1954, *Treatise on Electricity and Magnetism*, 3rd Edition, Dover Publications, New York
Mitoff S P and Charles R J 1972, J. Appl. Phys. **43**, 927
Moore E J 1974, J. Phys. C. **7**, 339, 1840
Nee T W and Zwanzig R 1970, J. Chem. Phys. **52**, 6353
Pollak M 1971, Phil. Mag. **23**, 519
Scher H and Lax M 1973, Phys. Rev. **B7**, 4491, 4502
Volger J 1960, *Progress in Semiconductors*, A F Gibson (Ed.), Vol. 4, 205, Heywood & Co, London
Wagner R J 1913, Ann. d. Physik **40**, 817
Warburg E 1899, Ann. Phys. Chem (Neue Serie) **67**, 493
—— 1901, Ann. Phys. (Germany) **4**, 6, 125
Williams G 1972, Chem. Rev. **72**, 55

CHAPTER 8

The Many-body Universal Model of Dielectric Relaxation

8.1 THE CONDITIONS FOR THE OCCURRENCE OF THE UNIVERSAL RESPONSE

Chapter 5 established the applicability of the universal power–law of the frequency dependence of the complex susceptibility, eqns (5.8)–(5.12), quite regardless of the physical and chemical properties of the very different systems under examination. Chapter 6 gave the experimental evidence for the corresponding time domain response, eqn (6.16).

The universal law was found to be valid over large ranges of frequency and time and also over remarkably wide ranges of temperature, including cryogenic temperatures. Since the universal power law is so evidently insensitive to the many material properties which one would imagine should determine the dielectric response, it is pertinent to ask which features of the polarisation processes are common to all materials exhibiting the universal behaviour. This would provide an important indication as to the direction in which one has to look for a physical interpretation of the dielectric response in condensed matter.

A search for a common explanation is all the more imperative since the ubiquitous fractional power law is not an evident solution of a differential equation corresponding to some familiar mechanism, and yet its very universality suggests strongly that a generally applicable mechanism is at play.

The first common property which is evident in all systems under consideration is the inevitable presence of interactions arising from the close proximity of atoms and molecules in condensed matter. It has been realised that this close proximity of interacting species gives rise to far-reaching consequences, hence the efforts over the years to develop local field theories, correlation functions approaches and so on. Strangely, however, the full logic of this

realisation has never been pursued and this is the reason why all these theories are in effect approximations based on adjustments to one-particle treatments. Consequently, they could never account for any but the relatively slight departures from the ideal Debye responses.

While it is true that almost any departures from Debye behaviour can be explained in terms of distributions, we have pointed out elsewhere the lack of physical plausibility of these approaches. At the same time the essential difference has to be stressed between *superpositions* on one-particle relaxations and *interactive* relaxation processes. The former give ordinary summation of behaviours in linear systems, the latter modify in a very profound way the one-particle responses in a cooperative system.

The importance of many-body interactions as the *dominant* and not merely a modifying influence in dielectric relaxation was recognised in the early stages of the present investigations (Jonscher 1975) when it became clear that a common mechanism was required for a wide range of apparently disjointed situations. This has resulted eventually in the development of the rigorous quantum-mechanical theory by Dissado and Hill (1979, 1980) an outline of which will be presented in the following sections.

The second common feature of all solid dielectrics is the *abrupt* or *discontinuous* nature of the dipolar or charge carrier transitions between their preferred "stationary" orientations or positions. This is in complete contrast with the smooth transitions of the classical Debye dipoles which were presumed to be "floating" in a continuous viscous medium. It is also very different from the motion of "free" charge carriers in crystalline conductors which move with a constant velocity over mean free paths covering many interatomic spacings and undergo abrupt changes of velocity in random collisions, while their positions change continuously. Abrupt transitions occur even in liquids, so that some of the considerations presented below apply there as well.

Since the particles undergoing sudden transitions are charged, these transitions involve equally sudden changes of potential which may be "seen" by neighbouring particles which in their turn respond and so on. The interactive nature of the system implies, therefore, that any sudden individual transition of a charge or a dipole brings about a series of "chain" responses stretching both in time and in space well beyond the time and position of the initiating transition.

These two criteria of the universal response – the presence of many-body interactions and the sudden nature of the individual transitions – are completely general and may be presumed to be largely independent of the detailed physical and chemical nature of the materials in question. Thus, the potential variation resulting from a transition is the same whether the charges in question are dipolar, ionic or electronic. Furthermore, the short-range order in most solids is basically similar regardless of the detailed nature of the long-range order. Equally general is the property implicit in this model that there exist two very different time scales corresponding, respectively, to *microscopic* and *macroscopic* processes.

In this classification, microscopic transitions are the very rapid oscillations of individual dipoles or charges and these would be the only processes at play in a non-interactive medium. Because of particle-particle interactions, however, these microscopic oscillations slowly transmit themselves, or "leak out", to other regions of the medium which adjusts itself slowly to the rapid microscopic movements.

The interactions in question may be direct between the active particles, dipoles or charges, for example through Coulombic interactions, or alternatively they may take place through the intermediary of the "inert" matrix or lattice in which the active species are embedded. This transfer of excitations through the lattice is facilitated by the intimate interplay between dielectric polarisation and strain in the lattice–strain relaxation following very similar laws as dipolar relaxation, as will be mentioned later in Section 8.6.

The comparatively recent realisation of the importance of these interactive processes for dielectric relaxation coincided in time with the development of a growing theoretical understanding of cooperative processes in such branches of physical science as critical phenomena, magnetic interactions and the properties of disordered solids.

8.2 A DESCRIPTIVE APPROACH TO MANY-BODY INTERACTION

Many-body interactions represent one of the less well developed branches of physics and the reason for this is as much a lack of evident sense-perception experience of these phenomena which would make it easy to construct "mechanistic" or "ball-and-stick"

models, as also the considerable level of mathematical difficulty in the theoretical treatment. For example, the idea of dipolar rotation or of hopping motion of a charge carrier is intuitively self-evident, sometimes to the detriment of a deeper understanding, since a deceptively simple model appears to explain the entire situation and it requires a considerable mental effort to rise above this level. Even a *par exellence* non-classical process such as quantum mechanical tunnelling of a single particle can be readily visualised as the familiar "burrowing" of a mole underground. We are far less sure of our conceptual grip on the physical situation in the otherwise well known case of screening in a gas of free charges, such as electrons or ions where each individual charged particle is envisaged as being surrounded self-consistently by all the others avoiding that particular particle, in view of their mutual repulsion. While there is no difficulty in visualising this avoiding of one particular particle, to extend the model mentally to *all* particles simultaneously requires a considerable stretch of imagination.

The problem of applying many-body theory to dielectric relaxation is not made easier by the remarkable and unexpected fact that many-body interactions seemed not to affect, significantly, the working of such well established theoretical subjects as the band theory of solids which is expressly a one-electron approximation. We may say that most solid state phenomena can be understood in terms of this simple theory and where there is a suspicion that something is lacking, the remedy usually consists in adjusting some of the many parameters at one's disposal, so as to make the simple theory "fit" the data. Evident cases of complete breakdown of the one-electron theory are plasmon phenomena which do not fit at all into the model and have to be treated separately (Platzman and Wolff 1973).

An early indication of the need for many-body theory was provided by the analysis of heavily doped crystalline semiconductors (Fistul 1969). However, in this as in many other situations where many-body processes were thought to be important, there existed the serious difficulty of obtaining a sufficiently strong basis of experimental facts to guide the theoreticians and to provide an incentive to develop a new approach.

In this respect, the totality of dielectric phenomena in the frequency and time–domains provides a unique wealth of evidence which is not parallelled in any other branch of physics, as we saw in the earlier chapters. This means that in our case it is not possible to

"fudge" the theory by a judicious choice of disposable parameters if one wants to take a truly general viewpoint. For this reason dielectric relaxation provides a uniquely suitable vehicle for the development of many-body theory, since no other approach could be shown to be adequate to the task.

a) The screened hopping model

The lack of an intuitive understanding of many-body phenomena makes it desirable to develop various forms of approaches to the description of such systems which may advance our understanding of their dynamic response. The first such model, to our knowledge, was based on the requirement of generality of the energy criterion, eqn (5.11), which is applicable in the region of dielectric relaxation beyond the loss peak frequency (Jonscher 1975a).

From a strictly theoretical standpoint this "theory" is very simplistic, almost naive in approach and it cannot provide much quantitative information to aid the detailed interpretation of the universal behaviour. However, its advantage lies precisely in the fact that it uses relatively familiar concepts and the basic validity of these concepts makes a better understanding of more complex models correspondingly easier.

The basic argument starts from the proposition that if a model can be found satisfying the universal criterion (5.10), which is itself a direct consequence of the universal power law relationship (5.8) beyond the loss peak frequency, then there is no need to solve any particular differential equations describing the dynamic response of the system under consideration. This is because the Kramers–Kronig relations are valid quite generally regardless of the detailed physical nature of the processes governing the behaviour, and these relations uniquely satisfy the requirements for self-consistency of the solution in question. In this way the energy criterion (5.11) uniquely determines the frequency response of the system, without the frequency dependence having to be derived explicitly.

This argument may strike some people as unacceptable since they are used to solving specific differential equations to obtain definite physical laws. Some even assert that a physical law cannot be valid if it is not a solution of a differential equation describing the system in question. A little reflection should suffice, however, to convince the reader that such an argument is mistaken, since there is no obvious rule why it should be necessary to impose such a restriction on the validity of physical laws.

The basis of the screened hopping model is the well established fact that assemblies of charged particles exhibit the phenomenon of self-consistent *screening*, which is seen in gaseous plasmas and also in free electron gases in semiconductors and metals. The phenomenon itself is *par excellence* a many-body process and the generally accepted simple derivation of the Debye screening radius r_D is not in doubt. The screened Coulomb potential of any one particle of charge q in an assembly of density N in a medium of dielectric permittivity ε is given by the relation:

$$\Psi(r) = (q/4\pi\varepsilon r)\exp(-r/r_D) \tag{8.1}$$

with a Debye screening distance defined by

$$r_D = (\varepsilon kT/q^2N)^{1/2}$$

where r is the distance from the particle in question. The unscreened Coulomb potential corresponding to a bare particle in an infinite neutral medium is given by the pre-exponential factor in eqn (8.1). The presence of the exponential term implies that by virtue of Gauss' law the particle is surrounded, self-consistently, by a screening charge of magnitude $-q$, so that at distances $r \gg r_D$ the charge in question is "invisible" in the assembly.

We now propose to apply this simple concept to assemblies of *localised* charges which spend most of the time on defined localised sites and only from time to time make abrupt transitions under thermal excitation. These charges could be electrons or ions, or they could represent the individual charges of dipoles of finite length – as all *physical* dipoles necessarily are.

The situation here is similar to that obtaining in free carrier plasmas, mutual interactions tend to screen charges from one another through repulsion of like and attraction of oppositely charged species. The fundamental difference in comparison with free charges is that localised charges are not entirely free to adjust their positions, and therefore their local density, in a fully self-consistent manner with respect to the local potential. The analysis takes this into account by postulating that instead of complete screening by a charge $-q$, we now have *incomplete screening* by a charge $-(1-p)q$, where p is a screening parameter which may take values between 0 and 1.

Figure 8.1 shows the situation of a localised charge q occupying one of two allowed sites i and j and being surrounded by the opposite screening charge. If now the charge hops over to the other site in a time that is *very short* in comparison with other time

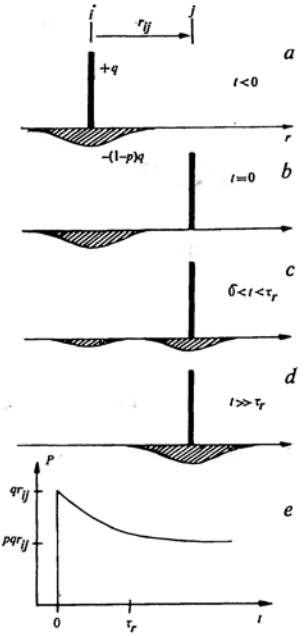


Figure 8.1 The screened hopping model in which a charge is constrained to hop between two localised sites i and j at a distance r_{ij} apart. The charge q is screened by an opposite charge $-(1-p)q$, diagram a). When the charge jumps discontinuously to the other site, the screening charge is momentarily left behind, diagram b), following relatively slowly in a time τ_r until a new steady state is established, diagrams c) and d). The time dependence of the polarisation resulting from this sequence of events is shown in diagram e).

From Jonscher (1975a). Reprinted by permission from *Nature*, Vol. 253, p. 717.

constants of the system, then the screening charge is momentarily left behind since it cannot follow such a rapid movement and this is shown in diagram b). There then follows a relatively *slow* transfer of the screening charge to the new site, diagram c), until a new equilibrium is established, d). The rate of transfer of the screening charge is determined by the natural rate of transitions for localised charges between individual sites, since the screening of any one hopping charge is effected by other similar hopping charges. We define a screening time τ_r, which is not to be confused with the Debye relaxation time with which it has nothing whatever in common. We do not even imply that the time dependence is in any sense exponential – the exact time dependence of this screening is immaterial for the purpose of the present argument, the only

important point is that the time scale of τ_r is very much longer than the *actual transit time* of the charge between the two sites. The latter is of the order of the reciprocal lattice frequency for ionic transitions, i.e. $\simeq 10^{-14}$ s, while for electrons it is shorter still.

We now consider the polarisation arising from the abrupt transfer of the charge followed by the slower movement of the hopping screening charge. Diagram 8.1 e) shows that the initial polarisation is qr_{ij}, where r_{ij} is the distance between the sites i and j, while the final polarisation is pqr_{ij}, after the complete transfer of the screening charge. Noting that this entire process takes place in an electric field E which may be considered constant on the time scale of these events, the work done by the field is Eqr_{ij}, while the energy finally stored in the system is $Epqr_{ij}$. The energy loss or dissipation arising from this process is due to the fact that the screening charge has to move *against* the applied field E under the influence of the strong local interactions with the hopping charge. Since the same energy loss is involved in each and every hopping transition, we obtain the energy ratio:

$$\frac{\text{energy lost per transition}}{\text{energy stored}} = \frac{1-p}{p} = \cot(n\pi/2) \qquad (8.2)$$

the last equality being the consequence of the energy criterion (5.10). This relation establishes a direct connection between the screening coefficient p and the exponent n in the power law (5.8). The limit $p = 0$, i.e. also $n = 0$, corresponds to complete screening, as in a free charge system, while $p = 1$ and $n = 1$ corresponds to the absence of screening, as would be the case with immobile charges which are unable to follow local changes of potential.

The functional relationship between the parameters n and p is shown in Figure 8.2 which shows how the full range of the exponents n between 0 and 1 can be obtained, in agreement with the empirical observations, by varying the screening parameter p, which here represents in a simplified way *the strength of many-body interactions* between charges and dipoles in the dielectric system under consideration.

Our screened hopping model involves three characteristic time scales, in addition to the practically instantaneous actual hopping transitions: the natural frequency ν with which the charges jump in equilibrium between the sites, the screening adjustment time τ_r and the frequency ω of the applied electric field. Since the time τ_r is determined by the individual transitions, it is reasonable to

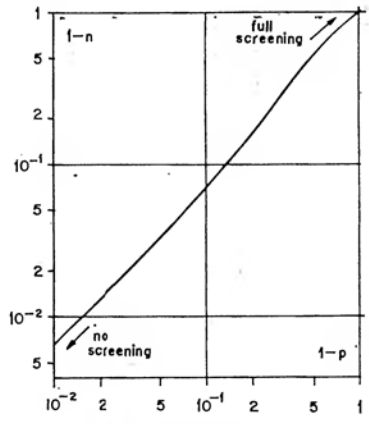


Figure 8.2 The relation between the exponent $1-n$ in the universal law and the screening parameter $1-p$ in the model, as given by eqn (8.2). Small values of n correspond to "full" screening, values of n close to unity correspond to "weak" screening.

suppose that $\tau_r > 1/\nu$. We may therefore distinguish the following three ranges of behaviour depending upon the relative value of the operating frequency ω:

a) $\omega < 1/\tau_r < \nu$ Particles and screening mobile, universal response governed by the energy criterion;

b) $1/\tau_r < \omega < \nu$ screening cannot follow the movement of particles and the system behaves in a similar manner to the Debye model below the relaxation frequency;

c) $1/\tau_r < \nu < \omega$ the system cannot follow the field, the response drops off as in the Debye model beyond the relaxation frequency.

A careful analysis of experimental data for a wide range of materials shown in the present Review fails to reveal the presence of any well developed Debye "tail" up to frequencies in excess of some 10 GHz where quantum and phonon effects take over. We conclude, therefore, that the regions b) and c) are not visible and this implies that the frequencies $1/\tau_r$ and ν fall above the 10 GHz limit. This seems, in any case, perfectly reasonable in terms of the likely orders of magnitude of atomic and molecular frequencies.

The arguments presented above should be valid equally for electronic and ionic charges and we would expect little change with the variation of the nature of the short-range order in the material.

They should be valid regardless of the nature of the chemical bonds involved. They could also be readily transferred to the dipolar situations in which the dipoles have *finite lengths* – as is the case with all physically realistic dipoles. These dipoles can screen one another in exactly the same manner as do hopping charges, except possibly that the screening may be weaker than in charge carrier systems.

We therefcre have a simple qualitative model which has the required property of satisfying the energy criterion and which is generally applicable to a wide range of physical and chemical conditions. This model explains the observed universal frequency dependence, *without any reference to distributions of parameters*. This does not imply that we deny the existence in physical systems of distributions of energies, relaxation frequencies etc., it only suggests that these distributions are not necessary *as such* to explain the universal dielectric response. The variation of the exponent n with temperature and with the type of material may be understood in terms of the variation of the screening parameter p: systems with more "mobile" carriers would have smaller values of n, i.e. stronger dispersion with frequency in view of a more effective screening, systems with weaker screening, e.g. induced as compared with permanent dipoles, would have values of n closer to unity, i.e. a flatter frequency response.

The essential features of our model are stated once more:

i) the existence of screening as an expression of many-body interactions;
ii) the existence of two distinctly separate time scales – the practically instantaneous transition time for the individual hopping movements and the relatively much longer screening adjustment time τ_r.

It turns out that the second of these features is also required in the rigorous theory and this supports the value of the screened hopping model as a qualitative guide to a better understanding of the polarisation processes.

Whatever its advantages, the screened hopping model suffers from the fundamental limitation that it is not capable of describing the appearance of the loss peak – for this we require the rigorous theory which will be described in Section 8.3. It was suggested, however, at that early stage (Jonscher 1975b) that the loss peak may be the manifestation of the existence of two separate and independent relaxation processes operating *sequentially* in the time–domain, one dominating at times shorter than $1/\omega_p$, the other at longer times.

This suggestion foreshadowed in a sense the findings of the later exact theory.

b) The role of disorder in the dielectric response

We shall be arguing that the role of many-body phenomena in determining the nature of the dielectric response of solids is intimately bound with the inevitable presence of a finite amount of disorder in all orientationally polarisable media (Jonscher et al 1980). We note that only orientational polarisation mechanisms show the relatively slow responses which are characterised by the universal behaviour, while the induced polarisation is normally dominated by inertial processes and its response times are typically in the range of 10^{-14} s.

A perfectly ordered crystalline array does not admit of any transitions by any of its constituents without the prior creation of point defects–perfect order is *rigid*. By contrast, disordered media are much more *flexible* since, by their very nature, there is no unique disordered configuration. This means, therefore, that a disordered system is capable of rearrangements of the type envisaged in Figure 8.5 b), which involves *small* adjustments of positions or orientations of *large numbers* of particles. These small adjustments requiring negligible energies for their induction are known under the name *configurational tunnelling*.

A considerable body of theoretical understanding of disorder has been accumulated in the context of studies on amorphous and glassy solids, particularly semiconductors. The concept of disorder being implicit in orientational polarisation in dielectrics is much less well established and we wish to discuss it briefly here.

Two types of disorder have to be considered in the present context – positional disorder of individual particles or the corresponding orientational disorder of dipoles, and the resulting energy disorder in the surrounding matrix. The two are intimately connected by the fact that the energy of a system is directly related to the orientational or positional arrangement of its constituent parts. One of the fundamental difficulties and limitations of the theory of disordered solids lies in the virtual impossibility of describing exactly, or even statistically, the state of the system. It turns out, however, that our understanding of the dielectric polarisation can be advanced sufficiently even on the basis of such imperfect information as we possess, using very general concepts of disorder.

Positional and orientational disorder arises very naturally in connection with orientational polarisation – they are inseparably linked with it, since by definition this type of polarisation consists in partial ordering of a randomly oriented or positioned array of dipoles or charges. At one extreme we have a perfectly ordered array of dipoles, completely oriented in the sense of ferroelectric interactions, where the only possible reorientation is the complete reversal of the entire domain – a strongly non-linear process which we are not considering in the present work. Any linear polarisation mechanism, on the other hand, involves slight reorientations of dipoles by an ordering effect of the external electric field which is opposed by the disordering tendency usually due to thermal vibrations.

Even if we had a perfectly ordered sub-lattice of dipoles in a "matrix", these dipoles would by themselves and in the absence of an external field be randomly oriented, except in the ferroelectric case below the Curie temperature. The disorder would become even more profound if the dipoles became randomly positionally distributed in the matrix, which is by far the most common situation in real solids. We now have to consider the effect of the orientational disorder on the constituent dipoles themselves and on the surrounding matrix. Since interactions between dipoles depend on their mutual orientations, there is a range of these interactions to which any particular dipole is exposed on account of the random orientations of its neighbours. Each dipole experiences, therefore, a slightly different environment and has a correspondingly slightly different energy from the others.

This position is even further complicated by the interactions between any particular dipole and the surrounding matrix in which it is embedded. These interactions influence the energy of the dipole and also influence the matrix through the strains introduced by the presence of the randomly oriented dipole. The resulting disorder in the matrix further influences other dipoles in the neighbourhood, and so on. This shows the physical impossibility of having macroscopically well defined energy levels in a dipolar system, even though each microscopic level is well defined within its environment.

The presence of point defects, whether randomly distributed or not, leads in turn to a second level of disorder – that arising in the neighbourhood of each such defect as the result of the local strain introduced by the defect. This strain affects significant numbers of nearest neighbours, changing their energetic and spatial characteristics – we note that for a simple cubic lattice the number of

neighbours up to the third-nearest to any particular site is 26, while going up to the sixth-nearest neighbours, i.e. up to a distance of 2.45 lattice spacings, we include 72 lattice points. Thus this type of strain disorder has far-reaching consequences and involves large numbers of sites per defect. Once the presence of such disorder has been envisaged, it is easy to see that the number of possible configurations involving slight displacements of atoms from their proper lattice sites becomes very large and thus the probability of configurational changes by transitions involving very small energies is correspondingly high.

It should be clearly borne in mind that a real solid finds itself at any temperature in a state of dynamic equilibrium in which the constituent atoms of the lattice execute small oscillations around their equilibrium positions and the tunnelling transitions between different configurational states should be regarded as resulting from this dynamic situation.

We conclude from this brief discussion that a dielectric system is necessarily disordered and this, in the presence of interactions, leads to the high probability of configurational tunnelling by dipoles or charged particles, without any need for these transitions to be thermally excited. A simple illustration of these transitions will be given later in the context of Figure 8.5.

c) The correlated states

An isolated parabolic potential well has sharply defined stationary energy levels corresponding to the ground state and to the excited states of a particle confined to that well. A pair of such wells in sufficiently close proximity to give rise to interactions, shows a splitting of the unperturbed energy levels, as shown in Figure 8.3. By a direct extension of this argument, we may conclude that a large number of interacting potential wells develops quasi-continuous bands of levels, the lowest of which has a width in energy which will be denoted by 2ξ. The density of these *correlated states* is high and they may be regarded as describing the movements of the centroids of assemblies of particles, rather than of individual particles in a non-interacting system. Since these levels are the result of particle interactions, they do not fit into the conventional model of energy bands in solids, which by the assumptions made in its derivation is restricted to single-particle systems. There is therefore no way in which these bands of correlated states can be placed in the conventional picture of solids involving the valence band, the conduction band and the localised levels arising from imperfections.

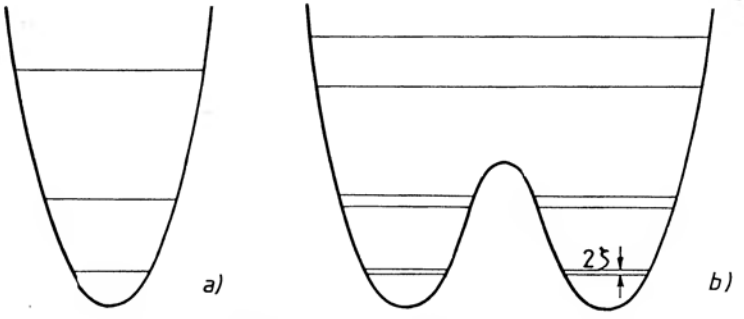


Figure 8.3 Diagram a) shows a single parabolic potential well with its stationary energy levels. Diagram b) gives the corresponding picture for a double well in which interaction splits the levels into pairs, with the ground state splitting being 2ζ and increasing rapidly for higher levels.

One characteristic difference between single-particle and many-particle excitations is the magnitude of the respective energies. Single-particle excitations are typically of the order of 1 eV and may span the range from 10 eV for the excitation of an electron from the valence into the conduction band in diamond or in covalently bonded carbon chains in polymers, down to milli-eV for the excitation of donor and acceptor levels in extrinsic semiconductors. The energies of many-body excitations are, by contrast, much smaller, of the order of milli-eV at most or smaller.

The physical origin of the narrow band of correlated states is therefore to be found in the weak but finite interactions which cause small variations in the energies of interacting particles as the result of movements of other particles – the system is "soft" or "flexible", none of the energies are uniquely defined in any absolute sense, everything depends on the movement of everything else.

While the energy span 2ζ of the band of correlated states is small, their total number and therefore their density in energy is very large, since they correspond to the order of density of dipolar particles per unit volume.

An important property of correlated states is their *inaccessibility to thermal excitations*, which means that the distribution in energy is sharply cut off, as shown in Figure 8.4, in which respect they differ completely from the familiar distributions of electrons and other particles showing thermal spreads of energy. The reasons for this inaccessibility to thermal excitations may be found in the character of these states which represent interactions *between* particles and not

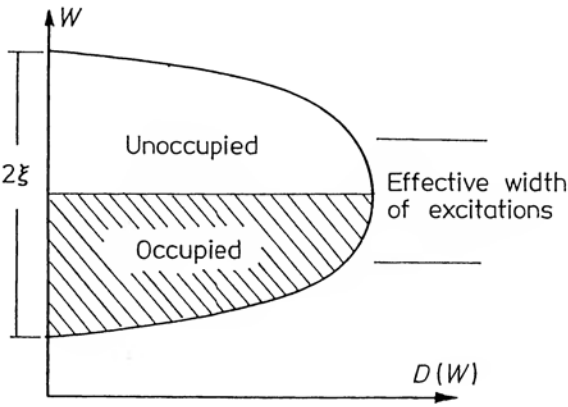


Figure 8.4 The density in energy of correlated states, showing the "occupancy" of these states in equilibrium and the effective width of excitations over which the density is approximately independent of energy.

excitations of individual particles themselves. On the one hand, this means that the excitation energies are much smaller than the typical values of the thermal energy kT, so that single phonon interactions are not possible. On the other hand, thermal motions are less effective in influencing the mutual interactions between particles, while acting primarily on assemblies of these particles.

The excitations of the correlated states occur through the sudden transitions by individual charges of dipoles in the assembly when they hop or "jump" between preferred orientations or positions. These hopping motions leave the remaining charges in instantaneous positions of non-equilibrium, just as in the screened hopping model the jump leaves the screening charge in a non-equilibrium position. In terms of the diagram of Figure 8.4 this excitation amounts to an upward shift of the population within the narrow band, giving rise to the storage of the excitation energy as will be discussed in Section 8.6.

d) "Large" and "small" transitions

As pointed out above, a perfectly ordered system is completely "rigid" and admits of no rearrangement of the constituent atoms or ions until a degree of disorder has been introduced, for example by the creation of a vacancy or an interstitial defect and this requires a considerable amount of energy. Once a vacancy has been created, further motion of atoms is relatively easy by a neighbouring atom hopping into the vacancy, which is therefore translated to another site, and so on. This hopping motion requires energies typically in

the order of 1 eV to take the particle over potential barriers hindering its motion in the lattice. The resulting steady state – direct current – conductivity for ionic transport is of this type, while hopping electronic conduction in compensated crystalline semiconductors is similar, except that transitions between full and empty levels involve very low activation energies because of the high dielectric permittivity of the medium and the low effective masses of charge carriers in shallow states.

An entirely different situation arises in disordered systems in which the constituent particles do not have firmly assigned positions in the lattice and where the energies and positions of the individual particles are affected by their neighbours. The resulting metastable situation leads to many alternative configurations of particles having very similar energies and transitions from one configuration to another may take place with very little energy input and with only very small adjustments of the positions of individual particles.

The situation reflects some of the features of a "milling crowd", as distinct from parade order in a military unit. In the crowd, there is no fixed arrangement of individuals and a fluid-like motion of the whole is possible without any single member moving a long way from a given position. The many-body interactive character of the motion is emphasised by the mutual jostling of individuals in the crowd, since no one can move without affecting the positions of his nearest neighbours.

We shall be referring to the first type of transitions in an otherwise ordered array with defects as *large* transitions, while the mutual slight rearrangements of particles in a disordered system will be called *small* transitions. Since in the latter case many particles are moving in a concerted manner, the total effect of many such small transitions may be equivalent to that of a single large transition. Where the particles carry opposite charges forming an overall neutral assembly, their mutual motions produce a changing dipole moment and they correspond therefore to dielectric relaxation. If a system has reached a state of equilibrium in the presence of a steady electric field, the removal of this field will give rise to a relaxation towards a new steady state via transitions changing the total dipole moment per unit volume.

The two types of transitions are shown schematically in Figure 8.5 in which the upper diagram corresponds to a "large" transition in an ordered array with one vacancy on a "square" site. This vacancy corresponds to a certain dipole moment calculated in terms of the

difference between the "centres of gravity" of the distributions of "square" and "round" charges. The dipole moment is indicated by the arrow at the top. A transition by a neighbouring "square" particle into the vacancy creates a change in the dipole moment as indicated by the lower arrow in diagram a).

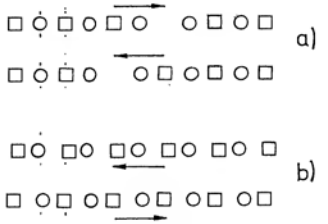


Figure 8.5 A schematic illustration of "large" transitions occurring in an ordered array with a vacancy, diagram a), and of "small" transitions occurring in a "disordered" array in which the constituent species – round and square symbols – are slightly displaced as shown by the vertical lines, with respect to the "lattice" positions. The rearrangement of these displaced particles in the sense shown in the lowest line gives the same change of dipole moment as the large transition in a).

Diagram b) shows a hypothetical "disordered" lattice in which the "square" and "round" particles occupy positions slightly displaced in opposite senses from the regular positions in a). These small displacements of oppositely charged particles in opposite senses create a net dipole moment indicated by the upper arrow in b) and equal in magnitude to that in a). If now the individual particles were to take up new "disordered" positions corresponding to displacements in opposite senses with respect to the "lattice" positions, the sum of all these small transitions would amount to the same change of dipole moment as the single large transition in a). In this way a disordered system may produce a large change of dipole moment without a single particle making a large transition.

Since the small transitions in narrow bands of energy have the fundamental property that they require negligible energies, they are referred to as *configurational tunnelling*, by analogy with quantum mechanical tunnelling of individual electrons, which also does not require any input of energy. The essential difference lies in the fact that in quantum mechanical tunnelling a light particle, such as an electron, makes a relatively large transition, while in configurational tunnelling large numbers of particles, which may be very heavy, make very small transitions.

8.3 THE INFRA-RED DIVERGENCE MODEL

a) The inapplicability of exponential relaxation in time

Our survey of experimental evidence presented in Chapters 5 and 6 has clearly shown the inapplicability of the simple exponential law of relaxation in virtually all condensed matter systems, and particularly in solids. The exponential relaxation in time and its frequency–domain equivalent of Debye relaxation apply only to non-interacting linear relaxors, i.e. those for which the potential energy is a quadratic function of displacement. The fact that this type of relaxation is virtually non-existent in nature points to the inapplicability of these assumptions in condensed matter. The physical reasons for the presence of interactions are intuitively evident and they have been outlined in the preceding sections – their origin lies in the interactions between the individual dipoles or charges and the "lattice" or matrix in which they are embedded. The implication is that any movement or relaxation of any one dipole or charge affects other neighbouring dipoles or charges, so that in an interacting system

nothing can move without everything else being affected.

Once this is the case, it becomes physically impossible to have exponential relaxation processes and one has to search for alternative mechanisms that would lead to the observed universal power law which replaces the idealised exponential process.

The impossibility of exponential relaxation may be demonstrated on a more mathematical basis by the following argument, adapted in a simplified form from Dissado and Hill (1983a). Let $a_l(t)$ be a displacement of a particle on the l-th site at time t. The relaxation of that particle after the removal of the external constraint is governed by the differential equation taking into account non-linear interactions in the system:

$$\mathrm{d}a_l(t)/\mathrm{d}t = (\mathrm{i}\omega_l - \gamma_l)\mathrm{a}_l(t) - \sum_{l'} \mathrm{a}_l(t)\mathrm{a}_{l'}(t)\, V_{ll'}$$

where $V_{ll'}$ is the coupling potential between displacements at sites l and l', ω_l is the natural frequency of oscillation at site l and γ_l is the corresponding damping coefficient. In the absence of coupling between different sites, i.e. in a non-interacting system, the solution of this equation is purely exponential with damping determined by the magnitude of ω_l/γ_l. However, with a finite amount of coupling the exponential solution is impossible.

b) Physical concepts in infra-red divergence

There exists a class of many-body interactions known in theoretical physics under the name of infra-red divergence (IRD) and manifesting themselves as X-ray edge anomaly, phonon-broadening of infra-red absorption spectra, Čerenkov radiation and Bremsstrahlung (Duke and Mahan 1965, Ferrel 1969, Yuval and Anderson 1970, Hopfield 1969, Mahan 1974). The common features of IRD phenomena are:

a) the presence of a sudden large excitation – an X-ray photon or infra-red photon being absorbed – which perturbs the potential or the Hamiltonian of the system;
b) the existence of lower-energy excitations which form a continuum over a band of energies and whose density of states in energy satisfies a condition defined later in the Appendix. They are the electron-hole pairs near the Fermi energy of the conduction band in metals, or the phonon spectrum of the semiconductor in which the IRD occurs.

For the purpose of the present treatment it is sufficient to state that the time-evolution of the relaxation of the system after the sudden excitation follows the IRD law t^{-n} where the exponent n depends on the detailed parameters of the system, but is always smaller than unity. A simple derivation of the properties of IRD is given in Appendix 8.1.

This power–law of relaxation applies to any interacting system in which the potential energy diagram contains quartic, i.e. 4th order terms, arising from the interaction, as in Figure 8.3 b), in addition to the usual second order terms, Figure 8.3 a). The higher order terms result in a Hamiltonian containing contributions of a non-linear form which are at the origin of IRD (Hohenberg and Halperin 1977, Patashinskii and Pokrovskii 1979).

The condition defined in the Appendix which the density of states should satisfy for IRD to be possible, need not strictly apply over the entire range of energies 2ζ. It would be sufficient for it to be satisfied over a limited range of the most likely transitions. For one particular case, in which this condition implies a constant density of states, this is shown in Figure 8.4 where the nearly constant density in the middle range should still be sufficient to give the power–law behaviour with sufficient approximation for excitations of small magnitude with respect to the total width 2ζ.

The application of IRD processes to dielectric relaxation was first proposed by Ngai (Ngai 1979, 1980a, b, Ngai et al 1979, Ngai and White 1979) who developed a detailed argument for the existence of low-energy excitations in the form of correlated states in most materials showing the universal dielectric behaviour.

An equivalent argument was carried further by Dissado and Hill (1980, 1981a) who took into account the experimental fact that the complete power law response of dipolar systems was characterised by *two independent* exponents m and n, as shown in Figure 5.27, corresponding evidently to two physically separate processes.

The exponent n is given by (see Appendix 8.1, eqn A8.13):

$$n = |V_f/2\xi|^2 \tag{8.3}$$

where V_f is the average excitation potential arising from a sudden hopping transition of a charge or dipole. Since only excitations which are smaller than the width of the band of states 2ξ are effective, this immediately shows why the exponent n has to fall in the range $0 < n < 1$ in agreement with experimental observations. The significance of the other exponent m will be discussed below.

The development of the Dissado & Hill (D & H) model is an example of a successful collaboration between experimental and theoretical teams, each bringing its own specific contribution. Without the theory, the empirical classification of Figure 5.57 and the corresponding time–domain information would remain an interesting but not very satisfying compilation of information. On the other hand, the development of theory was only made possible by the continuous interaction with experimentalists whose data helped to guide the theory at important cross-roads, to take the correct route. The point is that, like so many other physical processes, dielectric relaxation is so complicated that it is virtually impossible to develop a factually correct and theoretically tractable model, because too many parameters are uncertain and only a continuing comparison with experiment enables the correct choices to be made.

This was well illustrated by the initial approach to the problem of universality of the dielectric response – so long as one was restricting attention to specific groups of materials in narrow ranges of frequency, it was possible to develop otherwise plausible models fitting the particular situations, but a more general inspection of the empirical classification soon dispelled the illusion that one was dealing with physically realistic models. This has led to the formulation of our general conditions for the occurrence of the universal

behaviour, Section 8.1, and these in turn have led directly to the model of IRD. The next essential point to be considered were the deviations from the t^{-n} law or the corresponding deviations from the ω^{n-1} law at low frequencies and also the strong low-frequency dispersion.

c) The Dissado–Hill model of "large" and "small" transitions

This model is based on a combination of single-particle and many-body transitions shown in the energy diagram of Figure 8.6 which represents the potential energy of an assembly of a large number of interacting dipoles or charged particles in a two-level system. The large excitation energy Δ corresponds to the "large" transitions of single particles from the orientation or position corresponding to one of the two preferred states, to the other allowed position or orientation. The bottoms of the potential wells are split by an energy $2B_{eff}$ which is determined by the local conditions and also by the externally applied field E. Even more importantly, however, the splitting depends on the actual occupancies of the two minima, i.e. on the net dipole moment M

$$B_{eff} = B + kT_cM + QE \tag{8.4}$$

where T_c is the Curie–Weiss interaction parameter of the system and Q is the magnitude of the dipole moment corresponding to the transition of a charge from one minimum to the other. The shading of the bottoms of the two potential wells represents the correlated states of width 2ξ shown in Figure 8.4. The model envisages three types of transitions denoted, respectively, by the arrows a, a', b and c. Transitions a and a' correspond to the large single-particle transitions over the barrier height Δ, with thermal excitation in the case a and with partial tunnelling in case a'. Transitions b and c, on the other hand, are configurational tunnelling small transitions which do not involve thermal assistance. These various transitions will now be discussed separately.

d) Analysis of "Large" transitions

These transitions take one particle over the potential barrier and the dipole moment is given in terms of the difference between the occupancies of the two minima. In the case of equilibrium the dipole moment is given in the Curie–Weiss form

$$M_e = \tanh\left(\frac{B + kT_cM_e}{kT}\right) \tag{8.5}$$

With the equilibrium perturbed, for example by the removal of a polarising field, the rate of change of M is determined by the relation

(Brereton and Davies 1977):

$$\frac{dM}{dt} = -\nu_A \cosh\left(\frac{B + kT_cM}{kT}\right)\left\{M - \tanh\left(\frac{B + kT_cM}{kT}\right)\right\} \quad (8.6)$$

in which ν_A is a thermally activated frequency of transitions:

$$\nu_A = \nu_0 \exp(-\Delta/kT) \quad (8.7)$$

exactly as in the classical Debye model, ν_0 being the natural jumping frequency of the order of $10^{13}\ s^{-1}$.

The temperature parameter T_c defining the strength of the interaction through eqn (8.4) may, in general, depend on the temperature T and we are not at the present moment in a position to describe this dependence in a comprehensive manner. However, the particular case in which this parameter is independent of temperature corresponds to the situation where T_c defines a transition temperature in the system of dipoles and this *dipole alignment transition* will be shown to possess very interesting properties about which the Dissado–Hill theory can make specific predictions which are clearly verifiable by experimental data. This situation will be discussed further in Section 8.4 c).

Equation (8.4) is non-linear in M and this property will find application in the treatment of large-signal responses corresponding to very high electric fields (Dissado and Hill 1981b), but for the purpose of the present calculation of the linear response defining the dielectric susceptibility this expression may be linearised by defining the small-signal departure M' from the equilibrium value M_e, with $M = M_e + M'$

$$\frac{dM'}{dt} = -\nu_A M' \frac{d}{dM}\left\{\left(\cosh\frac{B + kT_cM}{kT}\right) \times \left(M - \tanh\frac{B + kT_cM}{kT}\right)\right\}_{M=M_e} \quad (8.8)$$

the derivative being taken at $M = M_e$. The result of this differentiation may be written in the form:

$$\frac{dM'}{dt} = -\omega_p M' \quad (8.9)$$

where the rate constant ω_p is given by

$$\omega_p = \nu_A \cosh\left(\frac{B + kT_cM_e}{kT}\right)\left(1 - \frac{T_c}{T}(1 - M_e^2)\right) \quad (8.10)$$

The solution of this equation gives the classical Debye result with the relaxation frequency given by eqn (8.10) which differs rather significantly from the classical Debye value (8.7). The resulting rate of relaxation in the time-domain is:

$$f_1(t) = \frac{dM'}{dt} = -\omega_p \exp(-\omega_p t) M'(0) \tag{8.11}$$

where $M'(0)$ is the initial deviation from equilibrium caused by the field. This value is obtained by differentiating the dipole moment with respect to the parameter B which is influenced by the external field E:

$$M'(0) = Ed(\partial M_e/\partial B)_{M=M_e} \tag{8.12}$$

and this gives the expression for the initial value of the dipole moment:

$$M'(0) = \frac{E \cdot d}{kT} \frac{1 - M_e^2}{1 - (1 - M_e^2) T_c/T} \tag{8.13}$$

Equations (8.10)–(8.13) give the complete description of the relaxation process arising from the large transitions over the potential barrier Δ. We note that this transition rate is of the Debye form with the important modification that the relaxation frequency depends on the interaction parameters B and T_c and that the initial value $M'(0)$ is given as a function of the equilibrium value M_e and of T_c.

e) The small flip transitions

These transitions are denoted by the arrow b in Figure 8.6 and they represent a tunnelling mode between different configurations giving a net change of the total dipole moment. They are therefore polarisation relaxing transitions of the type envisaged earlier by Ngai et al (1979). At times which are long in comparison with the reciprocal width ζ of the band of states, expressed in units of Planck's constant h, i.e. for $\zeta t \gg 1$, second-order perturbation theory leads to the result that the relaxation proceeds according to the law

$$f_2(t) \propto \cos(n\pi/2)(\zeta t)^{-n} \tag{8:14}$$

where the exponent n is given by eqn (8.3) and depends therefore on the magnitude of the excitation potential V in relation to the width 2ζ of the correlated states band.

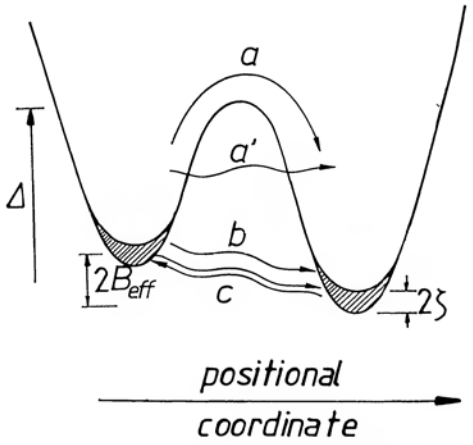


Figure 8.6 The potential energy diagram of a many-body two-level system representing the energy of a large number of interacting systems. The two wells correspond to the preferred orientations or positions. The splitting of the bottoms by $2B_{eff}$ is shown, the shaded regions represent the energy states of width 2ζ resulting from particle interactions. Arrow *a* denotes large thermally assisted transitions over the barrier Δ, arrow *a'* is the same with tunnel assistance. Arrows *b* and *c* refer to configurational tunnelling transitions of the flip and flip–flop types, respectively.

From Jonscher et al (1980).

These transitions are denoted by the arrow *b* in Figure 8.6 and *transitions*. As pointed out before, the interparticle interactions broaden the pairs of states into a band the maximum width of which is fixed by the relevant phonon frequency, i.e. for a double-minima by the phonons within an individual well. Excitations of the band can span the entire range of values from zero to 2ζ.

The flip transitions constitute the dominant relaxation mechanism at relatively short times, $1/\zeta \simeq 10^{-12}\,\mathrm{s} \ll t \ll 1/\omega_p$, since it will be shown later that this mechanism dominates until the time $1/\omega_p$, which falls typically between nanoseconds and milliseconds. The flip transitions set in, therefore, much *before* the Debye process of large transitions has time to develop and they are responsible for the Curie–von Schweidler law found in most dielectric materials. They occur entirely within the band of correlated states and they do not couple their energy to the phonon bath – they excite instead the correlated states themselves which store their energy until it can be imparted to the large transitions and only then be transmitted through them to the phonon bath.

The flip processes are independent of the thermally excited large transitions and they are competitive with the latter, so that the joint probability is the product of the rates (8.11) and (8.14) and this is the theoretically correct result, not the sum of probabilities as might be expected at first glance.

We also note that the alternative way of interpreting the physical significance of the value of the exponent n is to related it to the extent of correlation of the individual flip transitions between themselves: a fully correlated system corresponds to $n \rightarrow 1$, a completely uncorrelated system to $n \rightarrow 0$. We note therefore, that the latter case corresponds to the post-peak behaviour in the Debye system which is evidently applicable to uncorrelated dipoles. Fully correlated flip transitions, on the other hand, correspond to a very "rigid" system in which no single transition can occur without being accompanied by all the other dipoles flipping as well – this process therefore has an inherently low probability and corresponds to the "flat" low-loss response described in Section 5.7.

It should be pointed out that the t^{-n} law, if carried to zero time, would give rise to a singularity which would be physically impossible. However, the full solution of the IRD problem gives for times $\zeta t \ll 1$ a finite solution, as shown in the Appendix 8.1, and an initial decay of Gaussian form with zero gradient at zero time. It should be noted that this analysis ignores the inertial effects discussed briefly in Section 4.2. Experimental confirmation has been obtained of the existence of the oscillations at very short times, corresponding to microwave and far-infra-red frequencies (Dissado and Hill 1983a).

f) Fluctuations or flip-flop transitions

We have already pointed out that the IRD response given by eqn (8.14) and corresponding to the flip transitions cannot continue indefinitely, since this would lead to a physically inadmissible result. We therefore require a different physical mechanism that would give the response of the system at times much longer than $1/\omega_p$ and would be related to the exponent m in the empirical law of dipolar response.

Dissado and Hill have proposed that this mechanism is given by local fluctuations of the dipole moment which retain the average value of the total polarisation and which may be regarded as synchronous transitions in opposite senses at different points in the system, giving zero net change of dipole moment. These are called by them the *flip-flop* processes and they are denoted by the pair of

arrows c in Figure 8.4. Since they are cooperative processes of essentially similar type to the flip transitions, they must follow an essentially similar power law analogous to (8.14) which will be written in the form $(\eta t)^{-m}$, where η is a frequency in the range 10^8–10^{10} s^{-1} (Joffrin and Levelut 1975) and the exponent m is defined by analogy with eqn (8.3) in terms of the mean excitation potential for flip-flop transitions, V_{ff}

$$m = |V_{ff}/2\eta|^2 \tag{8.15}$$

while 2η is seen in the present context as corresponding to the width of correlated states involved in the flip-flop transitions. As in the case of flip transitions, the exponent m may be considered as characterising the degree of correlation of the flip-flop transitions among themselves, the value unity corresponding to complete correlation, the value zero to uncorrelated transitions.

The decay of these excitations is followed by restoration processes

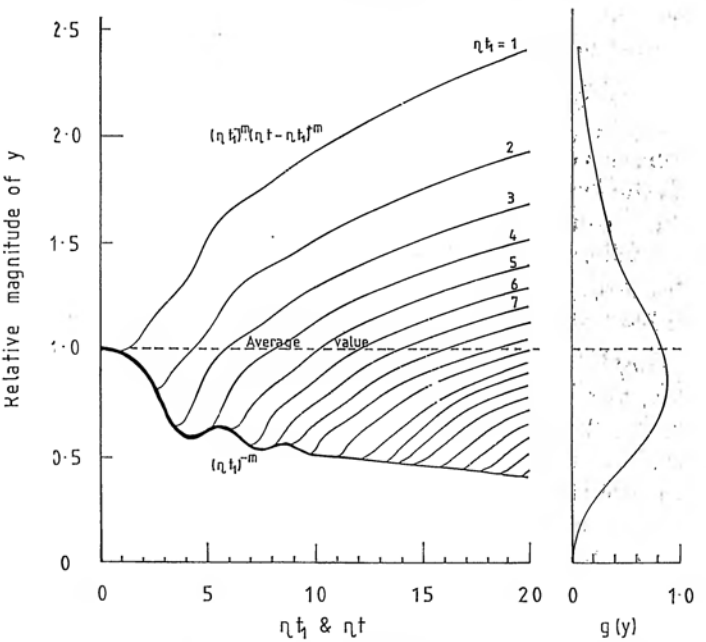


Figure 8.7 Schematic representation of the fluctuation processes represented by the exact expression of which eqn (8.16) is an approximation for $\eta t > 1$, normalised to their time-independent average value and denoted by $y(t)$. The diagram assumes for simplicity that at zero time all systems were at their average value. The diagram on the right shows the distribution function $g(y)$ of y.

From Hill et al (1981).

which follow the law $\{\eta(t-t_1)\}^m$ where t_1 is the time for the onset of the restoration process. The time development of these fluctuation processes may be simplified to the form

$$f_3(t_1) \cong t_1^{-m}(t-t_1)^m \tag{8.16}$$

A schematic representation of this time development is shown in Figure 8.7 which brings out clearly the decay and the recovery processes, the latter being spaced at equal intervals of the initiation times t_1. The fluctuating base line marked $(\eta t_1)^{-m}$ for brevity represents the complete solution of the equivalent of eqn (A8.14) exhibiting an oscillatory behaviour. We note that the time average of the fluctuation processes over any time period t is independent of time

$$(1/t)\int_0^t t_1^{-m}(t-t_1)^m \,\mathrm{d}t_1 = \Gamma(1-m)\,\Gamma(1+m) = \text{const in time}$$

and the fluctuations shown in Figure 8.7 are normalised to that average value. Also shown on the right of that figure is the distribution function of these fluctuations.

Similarly as for the flip transitions, the flip-flop processes do not couple to the thermal bath, they merely redistribute the energy within the correlated system, the final transfer to phonons occurring through the large transitions which are thermally activated. However, the flip-flops cause a redistribution of the excitation energy between different points in the system and in this manner they influence the ultimate excitation of the large momentum- and energy-dissipating transitions.

g) The complete analytical development of relaxation

The total rate of change of polarisation, which is the depolarisation current $i(t)$ flowing upon a sudden removal of a steady polarising field, and which is therefore related to the time-dependent characteristic dielectric function $f(t)$, has to be calculated by taking the average of the joint probabilities of $f_1(t)$ and $f_2(t)$ over the fluctuations $f_3(t)$. In the rigorous solutions developed by Dissado and Hill, which neglect only the inertial effects, phonon and quantum excitations, the complete solution is given by the plots shown in Figure 8.8 where the normalised current response is plotted logarithmically against time for two values of the exponent n, one in the middle range $n = 0.5$, the other for the upper reaches, $n = 0.9$. The entire time dependence falls into five regions. At times short compared with the reciprocal frequency ζ^{-1} the behaviour is Gaussian, clearly avoiding the singularity which would result if the

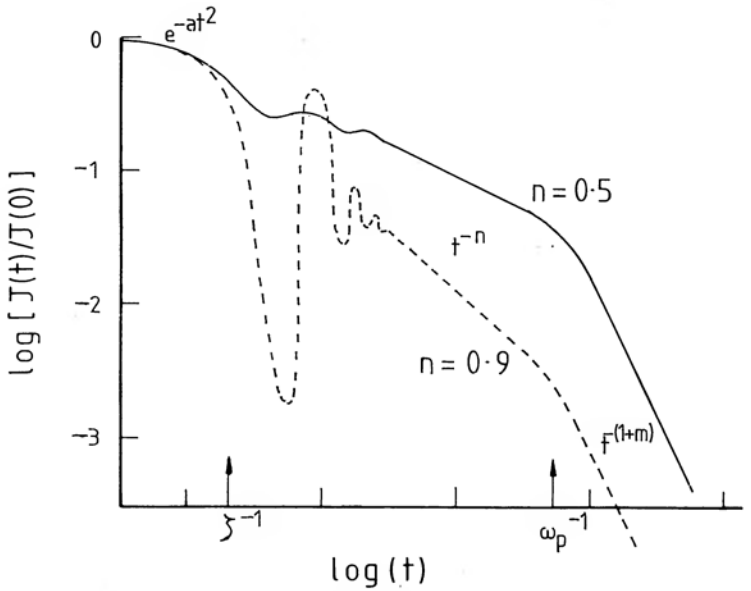


Figure 8.8 The complete time-dependence of relaxation current after step-function field excitation, for two values of the exponent n. This solution is well-behaved at zero time, although it ignores inertial effects, showing a Gaussian time dependence, and this is followed by an oscillatory region at times of the order of $1/\zeta$. There then follow two power–law regions, the Curie–von Schweidler law t^{-n} and the final t^{-1-m}, separated by a narrow exponential range.

From Hill and Jonscher (1983).

Curie–von Schweidler law t^{-n} were to continue to zero time. This is followed by an oscillatory region at times in excess of $1/\zeta \cong 10^{-12}$ s, or so, the flip processes establish themselves with some transient behaviour before the onset of the classical Curie–von Schweidler power law t^{-n}. That part of the response may be studied in more detail on the basis of approximate solutions which neglect the early stages of the time development, which is derived as follows:

$$f(t) \propto i(t) = \frac{\mathrm{d}M(t)}{\mathrm{d}t} = -\omega_p \cos(n\pi/2)\, M'(0)\, \zeta^{-n} I(t) \qquad (8.17)$$

where the entire time-dependence is subsumed in the factor $I(t)$ given by

$$I(t) = \frac{\int_0^t \exp\{-\omega_p(t-t_1)\}(t-t_1)^{m-n} t_1^{-m}\, \mathrm{d}t_1}{\int_0^t (t-t_1)^m t_1^{-m}\, \mathrm{d}t_1} =$$

$$= \exp(-\omega_p t)\, t^{-n}\, {}_1F_1(1-m;\, 2-n;\, \omega_p t) \times \frac{\Gamma(1+m-n)}{\Gamma(2-n)\,\Gamma(1+m)} \tag{8.18}$$

in which ${}_1F_1(\ ;\ ;\ )$ is the confluent hypergeometric function and $\Gamma(x)$ is the Gamma function.

Equation (8.18) gives the time-dependence of the depolarisation current after the transient stage at very short times shown in Figure 8.8 which does not concern us directly in the present work. The behaviour of the material in this approximation is specified by three parameters, ω_p, m and n which are the three empirically determined factors in the complete description of the dielectric relaxation process. Each of these has its specific physical significance and, in principle, is capable of being calculated if the physical condition of the system can be specified with sufficient accuracy, which is hardly ever the case in practice.

The general expression (8.18) may be Fourier transformed into the frequency domain, giving the complex dielectric susceptibility as the product of two factors:

$$\chi(\omega) = \chi(0)\, F(\omega/\omega_p) \tag{8.19}$$

in which the *amplitude* factor is given by

$$\chi(0) = \omega_p^n \cos(n\pi/2)\, \zeta^{-n} N \,.\, d[M'(0)/E] \times \frac{\Gamma(1+m-n)}{(1-n)\,\Gamma(1+m)} \tag{8.20}$$

while the entire frequency dependence is represented by the *shape* factor which is a function of the reduced frequency $x = \omega/\omega_p$

$$F(x) = (1+\mathrm{i}x)^{n-1}\, {}_2F_1(1-n,\, 1-m;\, 2-n;\, (1+\mathrm{i}x)^{-1}) \tag{8.21}$$

in which ${}_2F_1(\ ,\ ;\ ;\ )$ is the Gaussian hypergeometric function.

Equation (8.19) provides the theoretical justification for the frequently used normalisation procedure for the dielectric polarisation which is the basis of Chapter 5, a justification which is notably absent in earlier theories of dielectric polarisation. Provided that the parameters m and n remain independent of temperature, the normalisation results in a single master curve, as is seen in many examples in Chapter 5, while cases of temperature-dependent m and n result in normalisation plots consisting of several separate, not completely overlapping curves.

Equations (8.18) and (8.21) provide the complete description of the time- and frequency-dependence of the dielectric response and the amplitude of these responses is given by the corresponding terms in eqns (8.17) and (8.20). We now propose to describe some of the most important consequences of these general formulae, which give remarkably simple relationships in certain specific cases.

The expansion of the hypergeometric function (8.18) provides the following ranges of time-dependence:

$$i(t) \propto t^{-n} \quad \text{for} \quad 1/\zeta \ll t \ll 1/\omega_p \tag{8.22}$$
$$i(t) \propto \exp(-\omega_p t) \quad t \simeq 1/\omega_p \tag{8.23}$$
$$i(t) \propto t^{-m-1} \quad t \gg 1/\omega_p \tag{8.24}$$

and these are exactly the relations obtained on the basis of experimental data for a very wide range of dipolar materials, Figure 6.15. The corresponding expressions for the real and imaginary components of the complex dielectric susceptibility are as follows:

$$\chi'(\omega) = \chi''(\omega)\tan(n\pi/2) \propto \omega^{n-1} \quad \zeta \gg \omega \gg \omega_p \tag{8.25}$$
$$\chi(\omega) \propto (1 + i\omega/\omega_p)^{-1} \quad \omega \simeq \omega_p \tag{8.26}$$
$$\chi(0) - \chi'(\omega) \propto \chi''(\omega) \propto \omega^m \quad \omega \ll \omega_p \tag{8.27}$$

These expressions again correspond to the empirically determined dielectric behaviour of all dipolar materials.

The time-dependent relaxation process following a sudden removal of a polarising field may be considered as a succession of four consecutive stages, each of which is dominated by a particular type of transition. The very first stage covering a period of up to times of the order $t < 1/\zeta \simeq 10^{-12}$ s represents the delay necessary to establish the flip processes and corresponds to the initial stages of the time-evolution of the IRD process. This period may be dominated additionally by quantum phenomena and falls outside the range of dielectric responses which are the subject of the present study. Secondly, the t^{-n} law itself is a result of the small flip transitions whose probability of occurrence is much higher at this stage than that of the flip-flop transitions on account of the very large number of dipoles requiring to make a transition in the early stages of the relaxation. The flip transitions dominate the behaviour up to times of the order $1/\omega_p$ which represents the "natural" delay time for the onset of the Debye-like large transitions dominating the relaxation process at this third stage. Both the flip and the thermally excited large transitions cause a relaxation of the dipole moment. However, at times $t \gg 1/\omega_p$ the fourth stage sets in

which the state of polarisation approaches equilibrium and the "directional" processes of the second and third stages give way to fluctuations corresponding to the flip-flop transitions which, by themselves, do not give rise to any further change of the dipole moment but instead they redistribute the energy locally and thereby slow down the moment-altering transitions which complete the relaxation process according to the t^{-m-1} law.

It is important to bear clearly in mind that all three types of transitions take place all the time, at various relative rates, as may be inferred from eqn (8.18). The various stages of response described above correspond to the dominance of the respective processes, but not to the exclusion of the others.

8.4 THE CONSEQUENCES OF THE DISSADO–HILL THEORY

a) The significance of the loss peak

The immediate conclusion from the expressions presented in Section 8.3 is that the three fundamental processes denoted by *a*, *b* and *c* in Figure 8.6 govern three separate stages of dielectric relaxation in the time domain after the onset of the flip transitions at times of the order of $1/\zeta$, or the corresponding three ranges of frequency. It is very gratifying to find that the classical Debye process is retained in the new formalism as an integral part of the relaxation process, but a part governing only the response in the immediate neighbourhood of the loss peak frequency ω_p or the corresponding time. This means that the loss peak itself remains very closely associated with the Debye process, even though its numerical value may be modified in comparison with the classical value through the various terms in eqn (8.10). The importance of this conclusion lies in the fact that it preserves the validity of the very large body of theoretical and experimental work which associates the loss peak frequency with specific vibrational and relaxational processes in dielectric materials with predominantly dipolar polarisation mechanisms.

On the other hand, the new theory makes it clear that the validity of the "pure" Debye process is strictly limited in both frequency and in time and that the "wings" of the response on either side of the loss peak frequency are determined by other processes which are definitely many-body in nature and which give rise to the characteristic power-law relations in both time and frequency. It

is clear, moreover, that these power laws are established as specific manifestations of many-body processes and that they do not require any distributions of Debye-like processes for their justification.

This reasoning does not in any sense preclude the existence in materials of more than one Debye-like process of large transitions, each with its own characteristic activation energy and amplitude and with its associated many-body "wings". We have shown in Chapter 5 examples of dielectric responses having more than one loss peak, separated by larger or smaller frequency intervals. However, where a "monolithic" or "simple" loss peak exists in any given frequency range in a particular material, there we may assert with some assurance that there exists only one *dielectrically active* thermally exicted process, while others may be present but for some reason are not dielectrically effective.

b) The temperature dependence of the loss peak

The classical Debye–Langevin dipolar response given by eqns (4.30) and (2.13) requires that the loss peak amplitude should decrease linearly with the logarithm of the loss peak frequency. While this type of behaviour is sometimes seen in dipolar materials, for example in the case of the α peaks in polymers, there are many examples where it is clearly not applicable and where one has to invoke a different process. If the loss retains its spectral *shape* as well as its *amplitude* with changing temperature, then the integral of loss in the log-frequency space is invariant and this means that the polarisation increment $\chi(0)$ is also invariant with temperature, eqn (2.56). The implication of this is that a given field produces in the steady state a constant polarisation, regardless of temperature. This can be understood in terms of the polarisation being limited by non-thermal forces, for example by elastic constraints, instead of being limited by thermal agitation, as in the classical Debye process. At the same time, and very characteristically, the *rate processes* returning the system back to equilibrium after the removal of the field remain thermally activated, often with very simple single-energy processes, governed by large transitions over the barrier in Figure 8.6 which determine the loss peak frequency ω_p. This type of behaviour is seen in the data shown in Figures 5.13, 5.16 and 5.17 for dipolar materials, and is also clearly seen in p–n junctions in which the nature of the loss process is quite different. The temperature-independent dielectric increment in p–n junctions is due to the fact that the density of generation-recombination centres in the space charge region is constant with temperature.

There exist numerous cases in which the loss peak amplitude *increases* with increasing temperature, for instance in polymeric β peaks, as shown in Figures 5.16 and 5.21 a). A proper understanding of these phenomena would require further developments of our theoretical analysis, with special reference to the interplay between dielectric and mechanical relaxations.

c) Dipole alignment transitions

This special case of dielectric response is referred to by Dissado and Hill (1979, 1980) as the situation in which the characteristic temperature T_c is itself invariant with temperature which is by no means an obvious theoretical condition but which means physically that the material shows a *critical temperature* T_c. An examination of eqn (8.20) shows that the following relation applies between ω_p and $\chi(0)$:

$$\chi(0) = (\omega_p/\zeta)^n \cos(n\pi/2)\, N\,.\, d^2(1 - M_c^2)/(kT) \times \{1 - (T_c/T)(1 - M_c^2)\}^{-1} (1/m)\, \Gamma(1 - n) \quad (8.28)$$

The detail in this expression is a consequence of the particular Curie-Weiss mean field approximation used by Dissado and Hill.

It is possible to identify two regions of behaviour in this expression. Where the activation factor implicit in ω_p dominates the behaviour, which is the case well away from the transition temperature T_c, the relation becomes

$$\chi(0) \propto \omega_p^n \quad T \text{ away from } T_c \quad (8.29)$$

while in the region close to the transition temperature we have:

$$\chi(0) \propto \omega_p^{-1+n} \quad T \text{ close to } T_c \quad (8.30)$$

Eqns (8.29) and (8.30) constitute clear theoretical predictions which are verifiable experimentally and, by their very nature they could not result from any arbitrary combination of relaxation times, or any other specific previously proposed theory of relaxation. An illustration of this type of response is shown in Figure 5.22 and in Figure 5.23 a). Another example showing the complete range of relationships (8.29) and (8.30) as given in Figure 8.9 which relates to a ferroelectric ceramic (Brown 1981). There are many other examples of this type of behaviour. Sometimes more than one critical temperature exists, leading to the appearance of several separate branches in the $\chi(0) - \omega_p$ relationship, as for example in Rochelle salt which is the first ferroelectric material to be recognised as such and which is remarkable by the fact that it has two ferroelectric Curie temperatures at 249 and 291 K, being ferroelec-

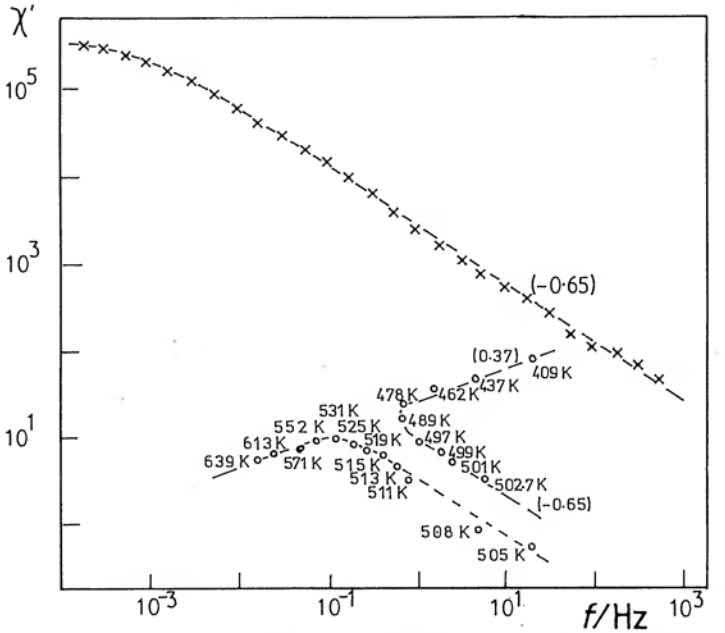


Figure 8.9 The relation between the loss peak amplitude and frequency, plotted in logarithmic scales, showing excellent agreement with eqns (8.29) and (8.30). The material is a doped ferroelectric ceramic based on lead zirconate and titanate. The sample in question showed a value of the exponent $n = 0.35$, and the slopes of the displacement locus are $+0.37$ away from the critical Curie temperature and -0.65 near T_c. The temperatures at which measurements were made are shown next to the characteristic points. The normalisation was carried out on the real part $\chi'(\omega)$ because a high dc conductivity masked the loss response.

From Brown (1981).

tric only in the narrow temperature interval between these. Figure 8.10 gives the normalisation of data originally obtained by Sandy and Jones (1968). This normalisation clearly shows four branches corresponding to eqn (8.30) and there are insufficient experimental data to delineate the other branches corresponding to temperatures further away from the critical temperature.

It is interesting to dwell briefly on the case given by eqn (8.30) for which the displacement of the characteristic point in normalisation procedure runs parallel to the high-frequency branch of the loss peak. This is the response expected of systems undergoing dipole alignment transitions, but it is probable that some other systems may also exhibit this type of behaviour, as is almost certainly the case for the low-temperature response of polyethylene shown in Figure 5.23a).

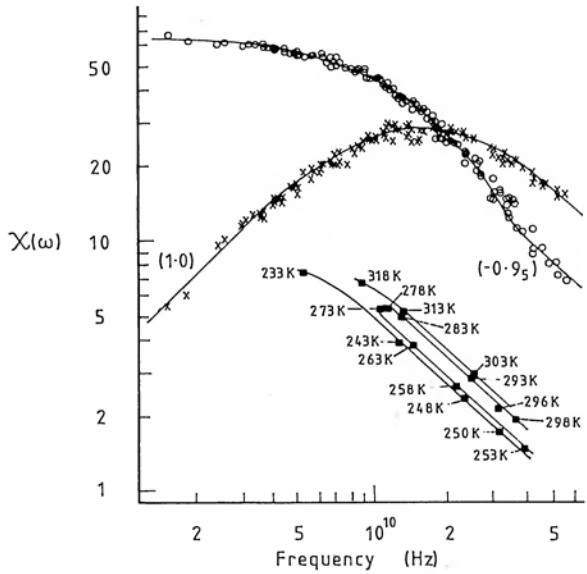


Figure 8.10 Normalised susceptibility plot for the dielectric loss in Rochelle salt which is a ferroelectric material with two Curie temperatures and with ferroelectric properties only between these temperatures. The two minima in the datum point trace correspond to 253 and 298 K giving rise to the four branches of the characteristic plot with the slope $n - 1$. The values found for this sample are $m = 1.0$ and $n = 0.05$. The plot is scaled at the temperature of 318 K.

From Hill and Jonscher (1983).

The frequency domain response is shown schematically in Figure 8.11 a) and the corresponding time–domain response in diagram b). The implication is that the short-time n-process is temperature independent in its polarisation *rate*, which is the depolarisation current, and that the duration of this stage of the depolarisation process is determined by the temperature-dependent loss peak frequency. Since the polarisation increment $\Delta\chi(0)$ is determined by the area under the (linear) loss curve on log(frequency) axis, the total initial polarisation increases with falling temperature, as the loss peak frequency decreases, as required by eqn (8.30). We do not at present understand the full implications of this simple physical conclusion, but it appears to be significant.

d) The exponents m and n

The Dissado–Hill theory associates these exponents with the degree of correlation between the flip-flop transitions for m and between the flip transitions for n. In each case the value of unity corresponds

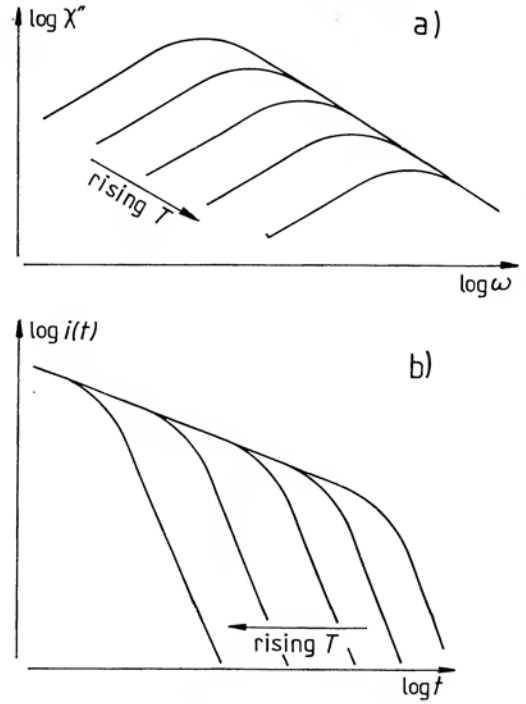


Figure 8.11 Schematic representation of the consequence of the loss amplitude and frequency response described by eqn (8.30), for the behaviour in the frequency domain, diagram a), and in the time domain, diagram b). In the time domain the initial rate of change of polarisation is independent of temperature.

to fully correlated transitions, the value zero to uncorrelated transitions. The value $m = 0$ corresponding to completely uncorrelated flip-flops resembles mathematically the case in which there are no flip-flop transitions at all, i.e. where the final stages of the relaxation process follow the exponential term in eqn (8.23). The Fourier transform of this gives the Davidson-Cole behaviour corresponding to eqn (3.38) in which the loss below the peak frequency is proportional to ω. Thus the cases $m = 0$ and $m = 1$ give the same frequency response below the peak, although the latter is very unlikely to happen on physical grounds.

One particular case of the absence of flip-flop transitions occurs when the frequency of the large transitions determining the loss peak is sufficiently high, so that the flip-flop process does not have time to develop at all. This occurs under conditions where $\omega_p > 1/\eta \cong 10^9 - 10^{10}\,\mathrm{s}^{-1}$ and it accounts for most of the points on the upper edge of the diagram in Figure 5.27.

The case $n = 0$ corresponds again to completely uncorrelated flip transitions and it represents the limiting case of a Debye system. The other extreme of $n = 1$ corresponds to a fully correlated system in which every flip transition causes some other flip transition to occur, so that the system may be regarded as very "rigid" and therefore difficult to excite. This is the situation in the experimentally determined limit of "flat" low level loss the physical cause of which is the inherently low probability of flip transitions in a highly correlated system, while mathematically the low loss is a direct consequence of the Kramers–Kronig transformation, eqn (8.2) in the limit of $n \rightarrow 1$. The existence of flat loss as the limiting form of dielectric behaviour in solids, but not in liquids (Reddish 1978, Lynch 1978) may be understood in terms of the essential rigidity of solids, favouring strong correlations between flip transitions after the removal of thermally activated large transitions. In liquids, on the other hand, the continual thermal fluctuations destroy any vestigial structures that may exist and this makes strong correlation extremely unlikely, so that high-purity liquids may show dielectric losses descending below the limit of detection of the best available equipment.

Where a frequency-independent loss is seen with values of the slope close to zero, representing the residual loss process in low-loss materials, one may have a number of possible situations. In predominantly dipolar materials one may be dealing with the case of *both* exponents m and $1 - n$ tending to zero, corresponding to strongly correlated flip transitions and, therefore, correspondingly weakly correlated flip-flop transitions. It is very likely that the generally limited accuracy of experimental measurements will make it difficult to discern the presence of a small peak superimposed on a broad flat background. Some of the data in Figure 5.21 and those in Figure 5.50 may fall into this category.

An alternative possibility is that one is dealing with only the limiting case of a flip process, in which case the loss peak with its activated processes must be presumed to lie outside the available frequency "window", for instance as in Figure 5.48.

At the present time we do not yet have any exact theoretical guide to the temperature dependence of the exponents n and m. We know from experience that they may show a temperature dependence tending to smaller values of m and $1 - n$ with decreasing temperature, as may be seen in Figures 5.15 and 5.16. A very interesting situation is seen in Figure 5.25 relating to very low-temperature

data, where $1 - n$ decreases slightly with increasing temperature, while m shows a distinct minimum.

Very marked and abrupt changes of the exponents may be seen in systems undergoing a structural change, for example the β to α transition or going through the glass transition temperature in glassy systems such as polymers. The more "solid-like" or rigid structure below T_g shows a flatter loss peak than the more "liquid-like" above T_g.

These observations point clearly to the close relationship between short-range structural order and the values of the exponents. So long as the order remains unchanged with varying temperature, the exponents remain constant, if the order changes slowly or abruptly, the exponents reflect this immediately. To that extent, therefore, there is little point in looking for any general rules regarding the temperature dependence of these exponents – what is more relevant is to ask for their dependence on the structural order.

The dependence of the exponents m and n on the structurally determined correlations of flip and flip-flop transitions is shown schematically in Figure 8.12.

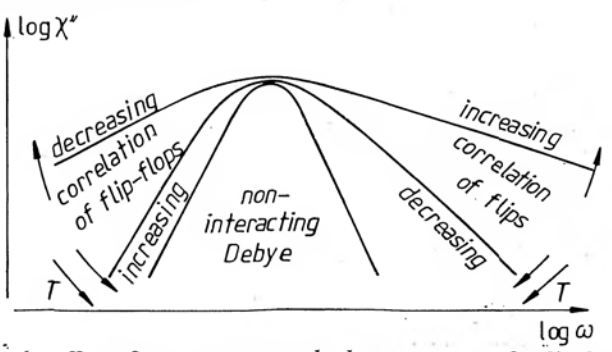


Figure 8.12 the effect of temperature on the loss spectrum of a dipolar material, as a consequence of changing correlations of flip and flip-flop transitions resulting from varying short-range order. The non-interacting Debye response is shown as a singular case. The emphasis in this diagram is on the *shape* of the spectral distribution, its *amplitude* and loss peak *frequency* have been normalised in the usual manner.
From Jonscher (1981).

In this context it is interesting to note the considerable stability of the shape of p–n junction loss characteristics over large intervals of temperature. This may be understood by the stability of semiconductor structure, with complete absence of phase transitions and other perturbing phenomena.

The very different behaviour of ferroelectrics at "high" and "low" frequencies requires comment. In the region of the "giant" dispersion shown in Figures 5.1, 5.6, 8.9, 8.10 and also to some extent in Figure 5.8 relating to a liquid crystal, the response is nearly Debye-like since in this frequency region the whole strongly coupled dipolar system behaves like a single molecule which does not suffer any interactions with other parts of the system. By contrast, at much lower frequencies, such as are shown in Figure 5.34, the loss arising from the "giant" dispersion of the ferroelectric transition has become so small that other, disorder-dominated processes dominate, giving the "universal" response. In either case, the amplitude of the loss process is governed by the co-responding high-frequency permittivity and the frequency dependence remains similar on either side of the Curie temperature.

This sensitivity of the exponents m and n to the structure of the material represents one of the principal *diagnostic* uses of dielectric measurements – the shape of the loss spectrum representing a sensitive tool for the detection of order changes.

e) The temperature dependence of "flat" loss

One of the more intriguing aspects of the temperature dependence is the case of the "flat" loss showing a clear peak in temperature, as illustrated schematically in Figure 3.36 and in the experimental data of Figures 5.49 and 5.51. One concludes that even though the flat loss represents a manifestation of the temperature-independent flip transitions, the *amplitude* of the loss, due to these transitions is sensitive to the temperature through the sensitivity of the structure to temperature.

f) The narrow range of ac conductivities

We have noted with reference to Figure 5.33 the remarkably narrow range of the absolute values of the alternating current conductivity which covers barely four orders of magnitude in a remarkably wide range of dielectric and semiconducting materials. This we have contrasted with the much wider range of the corresponding direct current conductivities which cover many powers of ten between the more highly conducting materials and the relatively more insulating ones. This behaviour of dc conductivity is not surprising in view of the wide variation of the available carrier densities and of their mobilities in the various materials. In this context, however, the narrow range of ac conductivities is rather surprising.

A possible explanation of this behaviour may be found in the concept of the limiting "flat" loss in all materials, arising from the

"lattice" contributions and corresponding typicaly to loss tangents of the order of 10^{-4}. This loss would give a freqiency dependence of conductivty proportional to ω, parallel to the chain-dotted lines in Figure 5.33 and the contribution of charge carriers may become negligible in comparison with this loss in the more insulating samples. The inference here is that the magnitudes of the lattice losses themselves do not vary greatly in most materials.

8.5 CLUSTERING AND STRONG LOW-FREQUENCY DISPERSION

The inherent disorder which is associated with all dielectric materials has already been stressed in Section 8.2 b) and this feature has far-reaching consequences for the dielectric polarisation in all materials. The point is that an apparently disordered material may, in fact, contain regions of relatively greater order, separated by regions of less order and this type of structure is widely encountered in many disordered materials. Examples of this type of structure may be found in glasses (Goodman 1975), in liquids (Hodgkinson 1976, Cohen and Jortner 1974) and in ceramics and their presence is either suspected or considered very likely in many other contexts. It is well known that the various granular regions may have different physical properties, e.g. the electrical conductivity, permittivity and also mechanical properties. It would be evident, therefore, that the presence of such non-uniformities might reflect itself on the dielectric relaxation of solids. We know already how sensitive the relaxation process is to local order and variations of the local order would be felt in the response of a material.

This subject was put on a quantitative basis by Dissado and Hill (1983a, b), Hill and Dissado (1982) who have extended the approach to the concept of a *cluster* which is defined as a region of coherence of molecular excitations within a larger body of a solid. A cluster in this sense does not entail any particular *structural* order falling within its boundaries – it is defined by the range of interactions of the type involved in polarisation relaxation. On the basis of detailed experimental information, Dissado and Hill suggest that cooperative interactions in dielectric relaxation involve a certain *correlation length* ξ_c over which the correlations of excitations are stronger than outside that length. These correlation lengths determine the scale of "granularity" of the medium, i.e. its sub-division

into clusters, with different relaxation properties of polar species within the clusters and relaxations caused by inter-cluster exchanges which are rather less strongly correlated than the former. The analysis of *intra*-cluster relaxation follows exactly the same path as that presented in Section 8.3 for the flip processes determining the short-time or high-frequency relaxation of a dielectric medium. *Inter*-cluster relaxations, characterised by a weaker interaction, are associated in the case of dipolar materials, with the flip-flop or m-processes determining the low-frequency or long-time relaxation. These processes limit the time-range of the intra-cluster relaxation and impose the t^{-1-m} regime in the time domain, after the exponential relaxation process, Figures 6.15 and 8.8.

In dipolar polarisation the interacting clusters correspond to *bound* charges, giving rise to loss peaks as a consequence of averaging over the fluctuations. By contrast, in charge carrier polarisation the interacting clusters are formed of *quasi-mobile* charges and the resulting different interaction with the field causes an effective interchange of t and t_1 in eqn (8.16), thus leading to the change from m to $-p$ in the expressions for the total current and for the complex susceptibility. This leads directly to the replacement of the loss peak by strong low-frequency dispersion.

The frequency at which the transition from the intra-cluster n process to the inter-cluster p process takes place is referred to as ω_c and the presence of strong dispersion tends to suppress the loss peaks from dipolar interactions, except in rare situations where dipolar and carrier processes are present simultaneously and the peak coincides with ω_c.

The magnitude of the exponent p is related to average energy of intercluster interactions, just as the magnitudes of m and n were related to the energies of their respective excitation bands. In general, theory suggests that the magnitude of p should be close to unity, in agreement with a large body of experimental evidence.

This development of the understanding of the relaxation processes connected with quasi-mobile charge carriers brings to its conclusion the outline of a unified theory of relaxation in solids. Charge-transport-dominated clusters provide a link between the dipole and dipole-like charge relaxations, on the one hand, and the ultimate dc conductivity, on the other hand. Strong low-frequency dispersion is an essential link in this complete picture.

8.6 ENERGY RELATIONS IN THE MANY-BODY THEORY

a) Stored energy in the static and transient regimes

When a dielectric medium is polarised to a statis polarisation P_0 in a steady electric field E_0, the energy density stored in the system is

$$W_0 = \tfrac{1}{2}E_0P_0 = \tfrac{1}{2}\chi(0)\,E_0^2 = \tfrac{1}{2}P_0^2/\chi(0) \qquad (8.31)$$

As the system is subjected to depolarisation, so that $P(t) \to 0$, the energy of the system is also reduced to zero but there is no way in which the instantaneous energy remaining in the system can be related to the instantaneous value of the polarisation, even though the entire frequency spectrum of the dielectric susceptibility be known. The reason behind this somewhat surprising result is that the instantaneous polarisation depends on the entire past history of the sample, so that it is not sufficient to know the instantaneous value of $P(t)$ but one has to specify the time "path" on which this value was reached.

Even more importantly, however, we know very little about the actual mechanisms which determine the transfer of energy stored initially in the polarisation to the heat bath and eventually to the outside heat sink. To take a simple example, a dielectric being charged "infinitely slowly" up to a polarisation P_0, suffers a vanishingly small loss of energy to the lattice, since the system is in quasi-equilibrium at all times and no loss of energy is involved – the total energy expended by the external source is equal to the stored energy W_0. If now the system should be equally slowly discharged by reducing the field to zero at a very slow rate, the entire energy W_0 would be transferred back to the source in the form of the discharge current working against the instantaneous applied field.

Going over next to the time–dependent regime in which finite rates of charging and discharging are applied, it is intuitively evident that a loss of energy is being necessarily incurred as a result of the changing polarisation. In the case of sinusoidally varying applied fields of a frequency ω, this manifests itself as a phase lag between the applied field and the resulting polarisation. In the case of a Debye dielectric, the phase lag corresponds to a *constant time lag* $\tau = 1/\omega_p$ regardless of the applied frequency, so that the phase lag increases as the frequency increases, $\phi = \tan^{-1}(\omega\tau) \simeq \omega\tau$. This constant time lag is, in fact, not so much a physical necessity, as

rather a direct consequence of the assumptions underlying the Debye process, namely the postulated existence of a constant time delay implicit in the differential equation (4.17).

In complete contrast with this classical Debye behaviour, the "universal" response beyond the loss peak frequency gives a *constant phase lag* $\phi = \tan^{-1}(1-n)\pi/2 \simeq (1-n)\pi/2$, independently of frequency. This is the basis of the statement that the energy lost per cycle is a constant fraction of the energy stored in the system at the peak of the cycle, which is such a remarkable property of the universal mechanism of polarisation. This property was used as the basis of the screened hopping model and we consider it as a fundamental feature of considerable theoretical significance.

The implication of this property is that the harmonic change of polarisation from a maximum value to zero entails a constant loss, regardless of the frequency at which this change takes place. The simplest way of thinking of this at the molecular level is to see it in terms of *every single transition incurring a constant amount of loss of energy*.

A characteristic feature of the harmonic drive is that the instantaneous polarisation and the driving field are in quasi-equilibrium, except for the, usually small, phase lag. An entirely different situation arises when the material is driven under step-function excitation by either charging or discharging – the latter corresponds to a sudden short-circuiting of a charged capacitor. Under these circumstances there is a *major instantaneous imbalance* between the prevailing polarisation and the field. This means that there is likely to be a much larger loss of energy in the process of changing the polarisation from zero to a finite value or the other way round – we would expect both amounts to be the same, in view of the "symmetry" of the charging and discharging operations.

Taking the discharge first, a system charged initially in a steady field E_0 has the stored energy which is given by eqn (8.31). If now the field is *suddenly* reduced to zero by applying an ideal short-circuit to the capacitor, the entire stored energy has to be dissipated internally, since there is no means of coupling it out in view of the assumed zero resistance in the external system. This means that the energy lost in *suddenly* reducing the field to zero is equal to the total initially stored energy. Taking now the opposite situation, where a step-function voltage is applied to an ideal capacitor without

any series inductance or resistance, the energy expended per unit volume in the charging process is given by

$$W_c = \int_0^\infty i(t)\, E\, \mathrm{d}t = \varepsilon_0 E_0^2 \int_0^\infty f(t)\, \mathrm{d}t = \varepsilon_0 E_0^2 \chi(0) \qquad (8.32)$$

which is exactly *twice* the stored energy given by eqn (8.31). This result is fully consistent with the previous one for the discharge, since the energy lost is the same in both (Jonscher 1978).

The implication of these results is that a charging process in which the charging or discharging current varies as $f(t)$ leads to the maximum energy loss – as much is lost as is stored. Any slower charging process gives rise to a lower level of loss implied by the loss angle in harmonic charging, where the ratio of energy lost per cycle to energy stored may be as low as 10^{-3} or 10^{-4} in low-loss materials.

The very high level of loss involved in rapid changes of polarisation in lossy dielectric media is highly relevant in the process of energy loss by moving charges which represents one of the less well understood aspects of charge transport in dielectrics (Fröhlich and Platzman 1953, Jonscher 1980).

b) Transfer of energy to the heat bath

We are now in a position to look in some detail at the mechanism of energy transfer between the polarising species – dipoles or hopping charges – and the heat bath of the dielectric lattice. In the case of the Debye mechanism this transfer occurs by direct exchange of energy of the charges whose movements are slightly delayed with respect to the field by the inherent delay time τ. The energy loss is a direct *consequence* of the time lag inherent in the Debye process.

With the Dissado–Hill model of many-body interactions the situation is essentially different, since the primary process is the constant energy loss per transition implicit in the universal law, while the phase lag *adjusts itself accordingly* so that the external circuit is capable of supplying the required energy. We may therefore look at the individual types of transitions shown in Figure 8.6, with a view to ascertaining their respective loss processes.

Starting with the large transitions taking place over the barrier Δ, which may or may not be partially tunnel-lowered, these transitions have a direct coupling to the phonon bath in exactly the same way as in the classical Debye transitions – they are of the same nature and their energy loss is determined solely by the actual

time delay of each individual transition with respect to the driving electric field. At frequencies close to the loss peak frequency this delay is the same as in the Debye case, but at other frequencies the delay is larger below the peak and smaller above it. The important point to note is that these transitions provide the major channel of energy transfer to the lattice.

A completely different situation arises in the case of the small configurational tunnelling transitions of the flip type. These transitions do not couple directly to the heat bath, but they excite the correlated states by creating an instantaneous imbalance resulting from the very rapid transitions in a system which cannot adjust instantaneously to a new situation. This is equivalent to the raising of the occupied states within the band of correlated states shown in Figure 8.4. Each transition causes a raising of the energy of a state by the amount $n\zeta$ and the combined effect of many such transitions occurring in the initial stages of the discharge process during the t^{-n} stage of relaxation described in Section 8.3f) is to store an appreciable amount of energy in the system of correlated states. This energy can only "leak out" to the lattice through the intermediary of the "large" transitions which do not begin to set in appreciably before the characteristic time $1/\omega_p$, but the effect of the large transitions continues until considerably longer times than in the purely Debye system, since they must eventually discharge the entire energy to the lattice but at the slower rate dictated by the m process. The energy storing action of the m and n transitions is an essential feature of the model and is at the basis of the broader loss peaks.

The role of the flip-flop transitions in the dissipation of energy lies in their action of transferring the energy of the correlated states from one position to another in the system, thereby facilitating the excitation of the energy-dissipating large transitions. The flip-flop transitions themselves dissipate neither energy nor dipole moment in the system as a whole.

Turning attention for the moment to the operation in the frequency regime at a constant frequency in excess of the loss peak frequency, we note that this regime is dominated by the flip transitions which give rise to the ω^{n-1} law. While driving the system at this frequency, we are not allowing the large transitions to take place appreciably, since their natural frequency is much lower. The only energy loss takes place therefore through the finite coupling between the excitations of the correlated states and the large transitions which are

thus induced to occur prematurely. In this regime, the actual value of the driving frequency is of little significance in determining the coupling to the large lossy transitions. A given amplitude of polarisation, corresponding to a given number of polarising transitions predominantly of the flip type, excites therefore a given number of correlated states and these can only lose their energy through the large transitions. This is the physical reason for the observed constant ratio of energy lost per cycle to energy stored at the peak, which is also the necessary cause of the "universal" ω^{n-1} frequency dependence.

The more strongly correlated a system, the larger the exponent n, the smaller the probability of exciting the highly correlated flip transitions. A strongly correlated system is also a highly "rigid" or perfect system in which there are few defects through which the thermally excited large transitions may be taking place. This is, therefore, the limiting condition of a flat loss found in low-loss systems.

c) Dielectric and mechanical loss

A relationship between the dielectric and mechanical loss peaks has long been known to exist (McCall 1981) and it has been associated with the common role of the Debye-like transitions which define the principal energy transfer processes – the difference between dielectric and mechanical losses being primarily the different form of "drive" – electrical in one case, mechanical strain in the other, while the active species is the same, namely the lattice dipoles.

An essentially different relationship has been derived by Hill and is described in Hill and Jonscher (1983). The principal line of argument is as follows. Given that the energy stored in the flip transitions is a fraction n of the maximum energy, the energy lost in the process of excitation is $1 - n$. If now the energy that is lost electrically were *entirely* converted into mechanical energy and stored as such, then the mechanical storage is equal to electrical loss. One would expect, therefore, that the mechanical relaxation process in a comparable frequency range should be governed by power law relations with exponents which are given by:

$$n_d + n_m = 1 \tag{8.33}$$

and similarly for the flip-flop processes

$$m_d + m_m = 1 \tag{8.34}$$

TABLE 8.1
Relationships between the dielectric and mechanical correlation parameters (from Hill and Jonscher 1983).

Material		Mechanical	Dielectric	Sum	Reference
Chlorocyclohexylacrylate	m	0.57	0.42	0.99	Heijboer (1972)
	n	0.28	0.75	1.03	Heijboer (1972)
Poly-n-octylmethacrylate	m	0.11	0.84	0.95	Strella and Chinai (1958)
	n	0.27	0.62	0.89	Dannhauser et al (1958)
Poly-n-hexylmethacrylate	m	0.28	0.9	1.18	Strella and Chinai (1958)
	n	0.38	0.65	1.03	Child and Ferry (1957a)
Poly-n-butylmethacrylate	m	0.25	0.71	0.96	Strella and Zand (1957)
	n	0.3	0.71	1.01	Child and Ferry (1957b)
Trio-*o*-tolylphosphate	n	0.55	0.4	0.95	Shears et al (1974)
Di-n-butylphthallate	n	0.5	0.45	0.95	Barlow and Erginsav (1972)

where the subscripts d and m denote dielectric and mechanical relaxation, respectively. This prediction is of a much more fundamental natue and covers a wider range of frequencies than the loss peak relatonship previously known and it is therefore very interesting to note the extent of agreement with experiment. There exist relatively very few data relating to both electrical and mechanical losses on specific materials and even fewer cover sufficiently wide frequency ranges to enable the exponents m and n to be determined with any accuracy.

Table 8.1 shows *all known data* compiled by R M Hill, giving the values of the respective exponents and also their sum. It is most remarkable that the agreement with the theoretical prediction should prove to be so good in all cases which could be traced. This result offers a significant support for the veracity of the theoretical treatment presented here and it also suggests that very similar considerations govern the mechanical responses as do the dielectric ones. This opens up exciting new prospects for considerable advances in the understanding of the mechanical relaxation processes which hitherto have always been considered on very similar lines to the Debye-based dielectric theory.

8.7 THE DYNAMICS OF TRAPPING AND RECOMBINATION IN SEMICONDUCTORS

The results of dielectric spectroscopy of p–n junctions in semiconductors were presented in Chapter 5 alongside other specifically dielectric data and the surprising conclusion from these was the far-reaching similarity of the results for dipolar and recombination/trapping systems, in view of the fact that the physical mechanisms responsible for the relaxation processes are completely different in both cases. The near-Debye response shown in Figure 5.7 relating to a high-purity p–n junction with very few deep levels gives a relaxation time of the order of 1 μs at 300 K, which may be presumed to represent a very likely value of the recombination time of electron hole pairs in silicon. We do not have sufficient experimental evidence to infer that recombination processes necessarily give near-Debye responses, but there is no doubt that the examples of a less pure silicon diode shown in Figure 5.28 and the very similar response of GaAs shown in Figure 5.30 reveal far-reaching departures from the Debye ideal, suggesting strongly non-exponential, power-law processes. In these two instances there is little

doubt that the relaxation times correspond to deep trapping rather than straight recombination processes. Other examples such as in Figure 5.31 show less drastic but still significant departures from Debye response in irradiated junctions.

The general conclusion from these and other measurements that the trapping processes in semiconductors do not obey exponential time developments and follow instead power laws is confirmed by a wider review of luminescence and photoconductivity in solids – all these depart very strongly from the exponential laws and very often follow power laws instead (de Polignac and Jonscher 1983).

There is little doubt that many-body interactions play some role in the transition of electrons to and from deep levels – several phonons are inevitably involved, and in addition there may be some interaction between neighbouring centres, especially in view of the effect of rapid "dipolar flip" arising from the sudden translation of the liberated charge in the conduction or valence band, as described by Figure 4.14. However we do not at the present time possess a theory capable of explaining this type of trapping dynamics and it remains to be hoped that the establishment of certain experimental facts will give rise in due course to some relevant theoretical studies, just as has been the case with the entire question of dielectric relaxation described in the present work.

These considerations may be placed on a more analytical basis by noting that the detrapping current in a p–n junction may be expressed as the *rate of detrapping* of charge carriers in Figure 4.13. This means that the number density of trapped charges is the equivalent of polarisation in a dipolar material. We may therefore write the following formal expression for the trapped charge density after a delta-function light or other form of excitation

$$\Delta n(t) = A\,(L\Delta t)\,h(t) \tag{8.35}$$

which is written by analogy with the corresponding polarisation equation (2.27) and where A is the normalising factor expressing the strength of the light coupling to charge carriers, while $L\Delta t$ is the strength of the delta function. The response to an arbitrary time-dependent illumination may then be written by analogy with eqn (2.30) as

$$\Delta n(t) = A \int_0^\infty h(\tau)\, L(t-\tau)\, \mathrm{d}\tau \tag{8.36}$$

of which the Fourier transform into the frequency domain may be

written in the form analogous to (2.36):

$$\Delta_2(\omega) = Ah(\omega)\, L(\omega) \propto H'(\omega) - iH''(\omega) \qquad (8.37)$$

where $h(\omega)$ and $L(\omega)$ are the Fourier transforms of $h(t)$ and $L(t)$, respectively, the latter being the amplitude of the driving light or other exciting signal. The complex nature of the Fourier transform (8.37) means that the response of trapped charges is delayed in phase with respect to the driving sinusoidal signal.

Under conditions of step-function excitation we obtain the transient response

$$\Delta n(t) = A\, L_0 \int_0^t h(\tau)\, \mathrm{d}\tau \quad \text{in charging} \qquad (8.38)$$

$$\Delta n(t) = A\, L_0 \int_t^\infty h(\tau)\, \mathrm{d}\tau \quad \text{in discharging} \qquad (8.39)$$

and

$$\Delta n(\infty) = A\, L_0 \int_0^\infty h(\tau)\, \mathrm{d}\tau$$

under steady state illumination (8.40)

If the response function $h(t)$ is purely exponential then the corresponding frequency–domain response is analogous to Debye, as in the case shown in Figure 5.7, while serious departures from that shape imply that the response function is of power–law type. These equations form the basis for the experimental determination of the response function of trapping processes at deep levels in semiconductors and semi-insulators.

A similar analysis may be applied to the frequency-dependent response of a photoconductor, in which case the important difference consists in the fact that photoconduction is proportional to the carrier density and not to the rate of change of density. Thus photocurrent is in this instance analogous to the dielectric polarisation.

Finally, once again it is possible to extend the analysis to the luminescent response of phosphors and similar materials, where the light output is proportional to the rate of change of carrier density which makes it therefore analogous to depolarisation current and to the current in a p–n junction.

The advantages of the frequency-domain method of measuring the photoconductive or luminescent response of deep traps are based on the same considerations as those of the determination of the

dielectric response function, as explained in Section 6.1. The information contained in, for instance, Figures 5.28 or 5.31 could never have been obtained by conventional processes of time–domain measurement in view of the limited range of amplitudes which are normally accessible to measurement.

8.8 DIELECTRIC DIAGNOSTICS OF MATERIALS

After our review of experimental data both in the time and frequency domains and after our theoretical discussion of the interpretation of the universality of dielectric behaviour in a very wide range of materials, it remains to sum up the salient points of the approach presented in the present Monograph. We ask the question: What does dielectric measurement give as tangible evidence of value to the experimenter, the theoretitian, the fundamental scientist and the applications-oriented engineer?

As regards fundamental polarisation processes, we are able to distinguish immediately between dipolar and charge-carrier-dominated processes, in terms of the division in Figure 5.58 into characteristics showing loss peaks and strong low-frequency dispersion, respectively. If the dielectric response is clearly dominated by loss peaks, with possibly a dc component at the lowest frequencies but without any strong low-frequency dispersion, then the conclusion must be that the dominant process is a dipolar polarisation, whether this be due to molecular dipoles, e.g. in polymers, or to ionic dipoles created by strongly bound ionic species, as for example in CaF_2, Figure 5.2.

The activation energy of the loss peak frequency may be analysed by the normalisation techniques and it may be interpreted in the usual manner in which loss peaks have always been interpreted in classical dielectric approaches, corresponding in the present case of the "large" transitions in Figure 8.6. The dipolar species corresponding to the activation energies and to the frequencies noted in this part of the study may then be associated with the classically developed arguments for the various types of transitions in dipolar materials – with the various molecular conformations in, for example, polymeric or ionic solids being associated with specific parameters. To that extent, the study of the loss peak frequency itself provides no more information than the classical approaches, with the possible reservation about the interpretation of the precise value of the frequency and of the amplitude, none of which is well understood in terms of classical theories either.

The next important point in the new approach to dielectric diagnostics is the *spectral shape* of the frequency dependence of loss and polarisation. The sheer quantity of information contained in this spectral shape is potentially much larger than that contained in the loss peak frequency and amplitude, provided that we are able to make use of it. Here, the presence of well-defined exponents m and n helps considerably in the interpretation of data, since if simple power laws are not discernible the conclusion must be that complicating overlaps between different mechanisms take place and discussion is difficult. However, the values of m and n provide valuable information about the structural relations in the system under study, the values corresponding to nearly Debye-like behaviour indicating a considerable degree of liquid-like disorder in the system, with the consequent lack of correlation between flip and flip-flop transitions. At the other extreme, the relatively very "flat" loss peaks testify to the presence of relatively ordered systems, in which a considerable "rigidity" of structure prevents uncorrelated transitions from taking place, with solid-like behaviour and relatively low loss.

An important aspect of interpretational analysis of data is the invariance of the exponents m and n with temperature or pressure in certain ranges of these variables. This clearly indicates that the short-range order in the relevant range remain invariant, retaining the spectral shape of the response. Where, on the other hand, this spectral shape changes with variation of temperature or of other variables, then the conclusion must be that phase transitions take place, involving a sudden or gradual variation of the short-range order, as the case may be.

We have stressed in the present approach the separation of frequency- and temperature-dependence of dielectric response – the presence of loss peak in temperature at constant frequency was shown not to prove that the material shows Debye-like response, only a frequency sweep can prove this point conclusively.

The low-temperature limit of dielectric behaviour is of particular interest in our approach, since this can only correspond either to tunnelling transitions of relatively very light protons, or to many-body configurational tunnelling in systems involving much heavier entities.

We now come to the interpretation of the "post-peak" response of dielectrics, at frequencies in excess of the loss peak frequency. The fact that this shows generally a well-defined power law of the type

given by eqn (5.8), does not, *per se* prove that one is dealing with hopping electrons, or even more generally with other charge carriers such as ions. This is the great misunderstanding which is widely found in the literature, where the presence of a power–law of this type was often taken to mean the dominance of hopping electronic conduction. Quite apart from the fact that this law does not distinguish between electronic and ionic charge carriers, in the absence of clear loss peak evidence it cannot even rule out the role of dipoles as the dominant polarising species. There are many examples in the literature where an insufficient frequency range fails to reveal the presence of a loss peak, so that only the power law is in evidence and this is then taken as proof of hopping electronic conduction.

Just as a loss peak is the ultimate proof of a dipolar process, so the charge carrier response is finally confirmed by the presence of strong low-frequency dispersion. Once again, this should be clearly distinguished from the dc conductivity – the latter is independent of the polarising dipolar species and is always present in all polar materials at sufficiently high temperatures. The former shows that charge carriers of either electronic or ionic nature are hopping in the system and their collective interactions lead to strong low-frequency dispersion. We may distinguish here between "bulk" and "barrier" responses if the data enable a sufficiently clear conclusion to be reached between Figures 5.47 a) and b). Neither of these behaviours, however, is compatible with any "classical" type of dielectric response, unless one is prepared to accept the purely accidental combinations of Maxwell–Wagner parameters.

The existence of the "flat" frequency-independent dielectric loss, with its connotation of a logarithmic time-dependence of polarisation, eqn (6.23), will continue to puzzle us for some time yet – the exact physical significance of this remarkable form of dielectric response is not easy to grasp.

One powerful tool of experimental analysis easily available with modern Frequency Response Analysers is the ability to subtract data obtained at different temperatures or in other similar situations, where the superposition of two different mechanisms may be clearly demonstrated on the basis of data which do not otherwise suggest any such simple result, as in Figure 5.56.

The adaptation of dielectric techniques to the study of the dynamic response of trapping centres in semiconductors, photoconductors and phosphors opens up one of the potentially richer fields in solid state physics. Here the theoretical interpretation is less well

advanced than n the dipolar and charge carrier polarisations, but the availability of good experimental data should stimulate the corresponding theoretical effort. In the meantime, it is already becoming clear that these measurements reveal material features which are not readily accessible to other methods of measurement and are not ever suspected to exist – e.g. in semi-insulating gallium arsenide. If, one day, a unified theory could embrace dielectric polarisation and trapping processes in semiconductors, this would represent a major advance in our understanding of the physics of the solid state. Even without such a unified theory, however, the fact that the dynamics of trapping processes can be shown to follow power laws in time and in frequency, instead of the normally assumed exponential relations, is itself of considerable significance.

The absolute values of the loss peak amplitudes or of the low-frequency dispersive processes continue to defy detailed analysis – but neither are they capable of being treated in the classical approaches, excluding the simplest Debye one which is so seldom seen in practice. This aspect, together with the closely related absolute magnitude of the steady state polarisation, $\chi(0)$, represent a continuing challenge to our theoretical understanding of many-body processes in solids.

One other area of considerable interest is the interaction between charge carriers and the dielectric matrix in which they are moving, as outlined in Section 6.7. This is not only relevant to low-field transport in dielectric materials, but becomes vital in the context of a physical approach to the causes of dielectric breakdown.

8.9 CONCLUSIONS

The purpose of the present Monograph was to establish, on the basis of wide-ranging experimental evidence, the principle of universality of dielectric behaviour, with its specific power–law dependence on frequency and on time, and to present an interpretation of this behaviour in terms of the newly formulated many-body theory of dielectric relaxation. In the course of our treatment of this subject we have tried to convince the reader that the experimentally observed behaviour is not compatible with any of the hitherto accepted theories of dielectric response. The principal encouragement to the formulation of an alternative approach stemmed from the very existence of universality of behaviour which covered all materials showing dielectric-like behaviour, whether

they be of the traditionally accepted kind or not. This universality accommodates a wide range of specific behaviours, some of which depart from the classical Debye response to the extent of being completely unrecognisable.

The behaviour of all real-life systems invariably departs to a finite extent from that of simplified or idealised models which may have been introduced as an aid to comprehension and analysis of a given situation. In some cases the departures of the actual behaviour from ideality are sufficiently small to be treated as perturbations and corresponding small modifications may be introduced into the idealised model to take account of this. These modifications may then provide a means of improving our understanding of the physical reality governing the response of the process in question. As examples of this we may quote the introduction of localised impurity levels which perturb the band structure of nearly perfect semiconductors, or small deviations from Ohm's law in the conduction of electric current.

A stage may be reached, however, where the perturbation becomes as important as, let along stronger than, the original ideal effect itself. Under these conditions, successive approximations to the idealised theoretical solution are no longer convergent and the system becomes intractable by the perturbation method. At this point, it may be necessary to develop from the beginning a completely fresh approach which attacks the essence of the non-ideal system directly and does not presuppose at all the existence of any ideal system. To follow the examples quoted earlier, the modern approach to completely disordered and amorphous solids attempts to tackle disorder itself, and the theory of high-field conduction attacks the non-linear flow as such.

This is exactly the situation in the theory of electric polarisation in near-Debye systems, where it is possible to introduce small perturbations of the ideal Debye model by means of correlation functions or by superposition of two or more simple non-interacting Debye relaxors. By the time one is trying to explain the behaviour of systems departing from the ideal by as much as many of the examples in the review of Chapter 5, the method becomes increasingly questionable in its validity and, in particular, in its physical plausibility, for example in the case of very wide distributions of relaxation times. The approach adopted in the present treatment starts, therefore, from the proposition that the ideal Debye behaviour does not exist other than as a mathematical abstraction and

that the essenceof all relaxation phenomena in condensed matter lies in the *interactve* nature of the movement of the polarising species.

Thus, instead of aking the usual approach of looking at the *departures* from the ideal Dbye response, we have taken the opposite approach of seeking a direct explanation of the power–law behaviour in its most general forn. The resulting many-body theory has proved to be fully able to account for most of the different manifestations of the universal response and has done this within a single framework whose generality is such that it can be adapted to all solid and even liquid situaions.

It is noteworthy that some of these types of responses could not reasonably be accounted for in terms of the established theories, for instance the imiting situations of flat loss with $n = 1$ and of strong low-frequency dispersion with $n \rightarrow 0$, the variety of low-temperature behaviours and the recombination and trapping laws.

Using the available very rich material of experimental data relating to dielectric behaviour of solids, it became possible to develop and to refine the new many-body theory to the point where it can not only look at the entire range of strictly dielectric responses, but is capable of moving into new areas of solid state physics. The reason for this is that many-body interactions are indeed ubiquitous in condensed matter, but their effects tend to be rather subtle and their experimental determination is not always feasible on a scale which enables a detailed theory to be developed. However, once the outline of the theory has been developed on the basis of the uniquely sensitive dielectric data, its application can proceed much more easily in other branches of physics. In this manner it became possible to interpret the ubiquitous "1/f noise" (Hill et al 1981), to look afresh at the age-old problem of dielectric breakdown as a cooperative phenomenon, to formulate a new approach to the transport of charges in semi-insulating materials, to interpret the shape of NMR spectra, the behaviour of magnetic systems (Dissado and Hill 1981c) and the non-exponential recombination laws (Jonscher et al 1982). In this way, a new and rich source of insights into the theory of condensed matter is being opened up.

As is often the case, the direct approach to the dominant mechanism, interactions in the present case, results in a greatly simplified physical model in comparison with the results of multiple approximations which do not lead to any physically plausible conclusions. Even though the mathematical treatment of the interactive system may not be very familir to the majority of people concerned with

the study of dielectrics, the physical model is essentially simple and the concepts involved become easily comprehensible once the initial obstacle of unfamiliarity has been removed.

It must be said that one can never be *absolutely sure* that a particular model, however otherwise plausible and theoretically rigorous, is *the unique true* model. The question of the ultimate acceptability of a theory lies in the creation of a consensus of opinion in the relevant professional circle that the balance of probability lies on the side of that theory. One of the essential elements in creating this consensus is the *simplicity* and the *generality* of the model in question, since the principle that

Mother Nature prefers simplicity

is a powerful guide to correct solutions.

On this note we leave the reader to form his own opinion about the plausibility and generality of the present many-body model of dielectric polarisation – let him weigh up the evidence presented here, and any other evidence of which he is aware, and let him then set it against the present theory and against the other theories that he knows of. A consensus will thus be formed without fear or favour or prior commitment.

APPENDIX 8.!

THE INFRA-RED DIVERGENCE

In the followirg we give a simplified treatment of the infra-red divergence based on a derivation by Hill (Jonscher and Hill 1982).

The vibrational Hamiltonian in the ground and excited states of a centre in an insulating crystal can be written in terms of creation and annihilation operators as (Hopfield 1969):

$$\mathrm{h}_{ground} = \sum_i E_i a_i^\dagger a_i \tag{A8.1}$$

$$\mathrm{h}_{excited} = \sum_i E_i a_i^\dagger a_i + \sum \lambda_i (a_i^\dagger + a_i) \tag{A8.2}$$

where E_i is the energy eigenvalue of the i-th state and λ_i is the first-order interaction energy between the states i.

On the application of a sudden electromagnetic pulse, such as might arise from the sudden reorientation of another dipole, each oscillator either remains in the ground state, or is excited with a probability $(\lambda_i/E_i)^2$. The mean number of oscillators is then:

$$\bar{N} = \sum_i (\lambda_i/E_i)^2 = \int_0^\zeta \frac{\lambda^2(E) \,.\, N(E)\, \mathrm{d}E}{E^2} \tag{A8.3}$$

where the summation over the levels i has been extended to an integral over energy, $N(E)$ is the density of states per unit energy and ζ is the upper limit of these states.

The mean energy is

$$\bar{E} = \sum_i (\lambda_i^2/E_i) = \int_0^\zeta \frac{\lambda^2(E) \,.\, N(E)\, \mathrm{d}E}{E} \tag{A8.4}$$

from which the distribution function of emission energy can be determined as

$$g(E)\, \mathrm{d}E = \lambda^2(E) \,.\, N(E)\, \mathrm{d}E \tag{A8.5}$$

In the particular case in which $g(E) = A \,.\, E$, i.e. it is proportional to E, we have from (A8.3) and (A8.4):

$$\bar{E} = A\zeta \tag{A8.6}$$

$$\bar{N} = A \int_0^\zeta \frac{\mathrm{d}E}{E} \tag{A8.7}$$

so that the mean energy is well defined but there is a logarithmic divergence in the number of excited states of oscillators at zero energy.

In second-order perturbation theory a logarithmically divergent number of low-energy excitations gives a time development of the response of the system to the pulse which is cooperative. The response in terms of the time-dependent current under step-function field excitation is given by the Fourier transform:

$$i(t) = \exp\left[\sum_i (\lambda_i/E_i)^2 [\exp(-iE_i t) - 1]\right] \qquad \text{(A8.8)}$$

which in the limit of small energy differences and with the particular distribution function given by (A8.5) can be re-written in the integral form:

$$\exp\left\{-A\int_0^{i\xi t} \frac{1 - e^{-x}}{x} dx\right\} \qquad \text{(A8.9)}$$

which is equivalent to

$$\exp\{-A[\gamma + \ln(i\xi t) + E_1(i\xi t)]\} \qquad \text{(A8.10)}$$

where E_1 is the exponential integral and γ is Euler's constant. The ratio of the average energy of excitation to the width of the available energy range is

$$\bar{E}/\xi = A$$

which if we set $A = n$ gives

$$\ln i(t) = -n[\gamma + \ln(i\xi t) + E_1(i\xi t)] \qquad \text{(A8.11)}$$

as obtained by Dissado and Hill (1980). In that derivation the characteristic parameter n was defined in terms of the average value of λ_i which is given by

$$\langle\lambda^2\rangle/\xi = \sum_i \lambda_i^2 = \bar{E} \qquad \text{(A8.12)}$$

and hence using (A8.6)

$$n = \langle\lambda^2\rangle/\xi^2 \qquad \text{(A8.13)}$$

The derivation presented here emphasises that the correlation index n is the average *energy* of excitation within the band and is identical to the parameter defining the density of excitation states.

Returning to eqn (A8.11) we note that the polarisation current is

given by:

$$i(t) = \mathrm{Re}\{(\mathrm{i}\zeta t)^{-n} \exp[-n\mathrm{E}_1(\mathrm{i}\zeta t)]\, \mathrm{e}^{-n\gamma}\} \qquad \text{(A8.14)}$$

denoting the real part of the expression in { }. The first factor in this expression gives the familiar power law which dominates for large values of the argument:

$$i(t) \propto (\zeta t)^{-n} \quad ; \quad \zeta t \geqslant 10 \qquad \text{(A8.15)}$$

while the second factor modulates this power law with an oscillatory function in the intermediate range of arguments, $1 < \zeta t < 10$, and for short times the solution tends to a constant and finite value

$$i(t) \propto \exp(-n\zeta^2 t^2/4) \quad ; \quad \zeta t < 1 \qquad \text{(A8.16)}$$

The exact solution at short times gives the correct physical behaviour corresponding to a finite response instead of the divergent t^{-n} law. We have already commented on the fact that this solution ignores inertial effects. It is noteworthy that the function shows strongly oscillatory behaviour for large values of the exponent n.

REFERENCES TO CHAPTER 8

Barlow A J and Erginsav A 1972, Proc. R. Soc. A **327**, 175

Brereton M G and Davies G R 1977, Polymer, **18**, 764

Brown M E 1981, J. Materials Science **16**, 1410 and Ph.D Thesis, University of London

Child W C and Ferry J D 1957a J. Colloid. Sci. **12** 389

—— 1957b J. Colloid Sci. **12** 327

Cohen M H and Jortner J 1974 J. Physique **C4**, 345

Dannhauser W, Child W C and Ferry J D 1958, J. Colloid Sci. **13**, 103

de Polignac Anne and Jonscher A K 1983, to be published

Dissado L A 1982, Physica Scripta **T1**, 110

Dissado L A and Hill R M 1979, Nature **279**, 685

—— 1980, Phil. Mag. **B41**, 625

—— 1981a, J. Materials Science **16**, 1410

—— 1981b, J. Materials Science **16**, 638

—— 1981c, J. Phys. C: Solid State Physics **14**, L649

—— 1983a, to be published

—— 1983b, to be published

Duke C B and Mahan G D 1965, Phys. Rev. **A139**, 1965

Ferrell R A 1969, Phys. Rev. **186**, 399

Frohlich H and Platzman R L 1953, Phys. Rev. **92**, 1152

Fistul V I 1969, *Heavily Doped Semiconductors*, Plenum Press, New York

Goodman C H L 1975, Nature **257**, 370

Heijboer J 1972, TNO Central Lab., Publication No. 435

Hill R M and Dissado L A 1979, Nature **281**, 286
—— 1982, J. Phys. C: Solid State Physics **15**, 5171
Hill R M, Dissado L A and Jackson R 1981, J. Phys. C: Solid State Physics **14**, 3915
Hill R M and Jonscher A K 1983, Contemporary Physics **24**, 75
Hodgkinson R J 1976, J. Phys. C: Solid State Physics **9**, 1467
Hopfield J J 1969, Comments on Solid State Physics **11**, 40
Hohenberg P C and Halperin B J 1977, Rev. Mod. Phys. **49**, 435
Joffrin J and Levelut A 1975, J. Phys. (Paris) **36**, 11
Jonscher A K 1975a, Nature **253**, 717
—— 1975b, Nature **256**, 566
—— 1978, J. Phys. C: Solid State Physics **11**, L601
—— 1980, J. Phys. D: Applied Physics **13**, L137
—— 1981, J. Materials Science **16**, 2037
Jonscher A K, Dissado L and Hill R M 1980, phys. stat. sol. (b) **102**, 351
Jonscher A K, Charoensiriwatana V, Favaron J and Loh C K 1983, to be published
Lynch A C 1978, private communication
Mahan G D 1974, Solid State Physics **29**, 175
McCall D W 1981, in *Physics of Dielectric Solids*, C H L Goodman (Ed), Institute of Physics Conference Series No. 58, p. 46
Ngai K L 1979, *Comments on Solid State Physics* **9**, 127
—— 1980a, ibid., **9**, 141
—— 1980b, Phys. Rev. **B22**, 2066
Ngai K L and White C T 1979, Phys. Rev. **B20**, 2475
Ngai K L, Jonscher A K and White C T 1979, Nature **277**, 185
Platzman P M and Wolff P A 1973, *Waves and Interactions in Solid State Plasmas*, Academic Press
Patashinskii A Z and Pokrovskii V I 1979, *Fluctuation Theory in Phase Transitions*, Pergamon Press, New York
Reddish W 1978, private communication
Sandy F and Jones R V 1968, Phys. Rev. **168**, 1181
Shears M F, Williams G, Barlow A J and Lamb J 1974 J. chem. Soc. Faraday Trans. II, **70**, 1783
Strella S and Chinai S N 1958 J. Polym. Sci. **31**, 45
Strella S and Zand R 1957, J. Polym. Sci. **25**, 105
Yuval G and Anderson P W 1970, Phys. Rev. **B1**, 1522

Author Index

Numbers indicate pages on which a publication is referred to by the name of its sole author, or both names in two-author work, or the first author's name if more than two authors. All names are listed with bold numbers relating to pages at the ends of chapters where the complete literature references are given. Pages 198 and 199 give complete references to Table 5.1.

Subject Index

ANDRZEJ K JONSCHER was born in Warsaw, graduated with 1st Class Honours in Electrical Engineering at Queen Mary College (University of London) in 1949 and obtained PhD in 1952 for work under the late Professor Harry Tropper, leader of a well-known Dielectrics Laboratory there. In 1951 he joined the Scientific Staff at the Hirst Research Laboratory of the General Electric Company in Wembley where he worked on the physics of semiconductor devices. In 1960 appeared his book "Principles of Semiconductor Device Operation".

In 1962 he became Reader and in 1965 Professor of Solid State Electronics at Chelsea College in the University of London, where interest in amorphous semiconductors gradually led him to studies of the dielectric properties of solids. He discerned a striking "universality" of the relaxation behaviour in a wide range of materials and for a time became "enfant terrible" of the dielectric establishment who found it difficult to accept his unorthodox ideas.

In recent years he has been applying dielectric measuring techniques to the study of recombination and trapping processes in semiconductors, semi-insulators and phosphors, and also became involved in a study of dielectric breakdown in collaboration with the Laboratoire de Génie Electrique de Toulouse.

He is the author and co-author of over one hundred research papers and reviews, is a Fellow of the Institute of Physics and a Fellow of the Institution of Electrical Engineers.