

Invited Review

ELECTRICAL CONDUCTION IN DISCONTINUOUS METAL FILMS: A DISCUSSION

J. E. MORRIS

Physics Department, Victoria University of Wellington, Private Bag, Wellington (New Zealand)

T. J. COUTTS

Environmental Sciences Research Unit, Cranfield Institute of Technology, Cranfield, Bedford MK43 0AL (Gr. Britain)

(Received May 20, 1977)

The conduction mechanism in discontinuous metal films, which consist of discrete islands on an insulating substrate, has not yet been fully and satisfactorily explained.

The most popular theory to appear in the literature is that of tunnelling between islands with an electrostatic activation energy associated with single island charging. This review examines various aspects of the theory critically and with respect to the many varied experimental results. A prime objective of the review is to provoke new approaches to the problem, and several rather speculative suggestions are included to this end. Weaknesses in accepted theory and experimental progress are identified, and positive proposals are made for future research. In particular the roles of contact structures and percolation processes have been neglected for too long.

I. INTRODUCTION

1.1. Basic properties

Swann¹ drew the analogy between rain falling on a pavement and the growth of a metal film on a supporting layer. It is generally an appreciable time from the start of rainfall before a continuous film of water on concrete is formed, and so it is also with the formation of a metal film. Of course, it was not until the coming of the electron microscope that researchers were able to verify this analogy or to provide quantitative information regarding the size of the droplets (or islands or nuclei) of metal on the substrate.

It was somewhat of a surprise to researchers in thin films to find that discontinuous metal structures exhibited a significant d.c. conductivity even though no metallic filaments existed between the electrodes. In addition, these films were also found to exhibit a negative temperature coefficient of resistance (TCR) the value of which appeared to be approximately inversely proportional to the size of

Editors' Note

This paper was invited for the specific purpose of opening up discussion on what remains a very difficult scientific problem which, when resolved, will lead to important industrial application. The authors have indicated a number of directions in which research is needed, and the Editors will be pleased to receive future papers in this field.

the islands, as assessed by transmission microscopy. Furthermore, the current-voltage ($I-V$) characteristics were often found to be non-linear. These properties can be regarded as being inherent to discontinuous metal films, and although they resemble the type of feature observed in semiconductors, in that a thermally activated process is required to explain the negative TCR, it is well known that the mechanism of charge transfer is quite different in the two systems. Much theoretical and experimental effort has been directed towards attempting to explain the charge transfer process, and it is our hope that we shall be able to convince the reader that there is still a long way to go before the topic is fully understood.

1.2. Reasons for study

Initially the investigations of the electrical properties of discontinuous metal films stemmed from scientific curiosity rather than from a desire to exploit their unusual properties in devices. However, it has for some time been appreciated that these properties can be exploited, and in particular the high sheet resistance is potentially attractive in the production of high value resistors on substrates that have a very limited available area. In practice the use of island films is ruled out because of their large negative TCR and their instability when exposed to normal ambient conditions. Therefore the cermet film, consisting of a dispersion of metal particles and capillaries in a dielectric matrix, has been developed with a view to overcoming these shortcomings. It is possible to produce a mixture either of chemical compounds or of islands and capillaries and to balance the opposite TCRs of the individual components so as to obtain an overall TCR of nearly zero. Since the presence of the dielectric affords a considerable degree of passivation against atmospheric effects, the cermet seems to have notable advantages over conventional metal or alloy systems. The reason often given for studying discontinuous films is that they represent a simpler version of a cermet and would therefore assist our understanding of the potentially more usable system. This can be regarded with polite scepticism, however, as the majority of the investigations appear to us to stem from straightforward scientific curiosity.

Nevertheless, there are several other potential applications for discontinuous films which may be practicable once the problems of stability have been overcome. Electrical applications include high sensitivity strain gauges for biomedical uses^{2,3}, displacement transducers⁴, temperature sensors and IR detectors⁵, gas detectors⁶, non-linear resistors⁷, electron or light emitters⁸ and electrodes for high resolution vidicons⁹. Each of these will be discussed in the following sections. Other potential non-electronic applications include the use of island Permalloy films to enhance hysteresis effects¹⁰, the use of the optical absorption property in solar collectors and, perhaps the most unusual of all, their use as a means of producing positronium¹¹.

1.3. Aims of the review

The aim of this review is not simply to cover the standard background research in the topic (for a comprehensive bibliography see refs. 12-16 which provide an excellent introduction) but to emphasize the very recent developments in the subject which have taken place owing to a worldwide resurgence of interest. Despite a great deal of theoretical modelling and experimentation the nature of the precise charge

transfer mechanism(s) in these films is not at all clear. Therefore we shall only point out the conceptual differences between the various models. Most of these account satisfactorily for the general qualitative features of current-voltage and current-temperature characteristics, but it is commonplace to observe large discrepancies between the calculated and the actual current densities. Perhaps one of the many factors which have not yet been taken into account could satisfactorily explain these discrepancies. For example, there is still no positive evidence regarding the role of traps in the dielectric between the islands, the effect of the relaxation time of the polarized dielectric and its influence on the activation energy, the influence of island shape and field enhancement at the island edges, the influence of the contacts both on the structure and on the charge injection process and the effect of the statistical distributions of island sizes and separations. These are only a few examples of parameters which have not yet been incorporated quantitatively in any model proposed to date, and it is hoped that by bringing them to the attention of the reader we shall provoke thought which will lead to an improved understanding of the physics involved.

In addition, we shall speculate about new models and, although experimental evidence to support these models is lacking, we hope that by this means we shall persuade the reader to discover whether the ideas are valid and so shall stimulate further progress. To this end, new theoretical and experimental approaches are suggested, and it is hoped that in the near future these will be put to the test.

2. THEORY OF CONDUCTION

2.1. Electrostatically activated tunnelling

During the 1950s and into the early 1960s the electrical properties of discontinuous films were widely interpreted in terms of thermionic emission. Following Gorter¹⁷ and Darmais¹⁸, Neugebauer and Webb¹⁹ proposed a tunnelling model which included the electrostatic energy of charged islands. Most work published since then has been based to some extent on this fundamental concept. The model's potential for quantitative development makes it much more attractive than other models, particularly to experimentalists. While there are too many unanswered questions to claim that even the basic concepts are proved, there are no irreconcilable experimental discrepancies. Theoretical development is still proceeding, and we hope to indicate some of the directions this will follow. Although we shall mention other models briefly, we shall concentrate on those aspects which we believe should be incorporated into the electrostatic theory, and all experimental results will be interpreted in terms of that theory.

The basic premise of all variations of the basic electrostatic activation model is that the transfer of an electron between two initially neutral islands results in an increase in the system's electrostatic energy. Neugebauer and Webb¹⁹ concluded that the equilibrium density of charged islands (which they assumed to be the carrier density) is related to the total density N of islands by the Boltzmann distribution $n = N \exp(-\delta E/kT)$; the activation energy δE is given either by

$$\delta E = \frac{e^2}{4\pi\epsilon r} \quad (1)$$

which is the electrostatic energy of an isolated charged island of radius r (i.e. the work required to remove an electron to infinity) or by

$$\delta E = \frac{e^2}{4\pi\epsilon} \left(\frac{1}{r} - \frac{1}{2r+s} \right) \quad (2)$$

when considering the energy required to take the electron to the next island a distance s away. The positive and negative islands are regarded as charge carriers for which a net down-field tunnelling transition probability may be derived and expressed as an equivalent average mobility in order to evaluate the sheet conductivity.

Abeles²⁰ and Dryer²¹ followed Neugebauer and Webb's¹⁹ treatment but stated more explicitly that there are two types of tunnelling transition. Neutral-neutral transfers correspond to charge carrier generation, but subsequent neutral-charged tunnelling leads to a non-activated mobility term. (Abeles also introduced the equivalent concept of a finite lifetime of the charges before recombination by a charged-charged transition.) The hitherto neglected implication is that, along a given percolation path through the film, activated charge separation clearly requires a greater gap potential than the non-activated charge transfer to pass the same average current.

Swanson *et al.*²² modified Neugebauer and Webb's¹⁹ result in two ways. For n charged pairs in a finite array of N islands, they adopted the expression

$$\frac{2n}{N-2n} = \exp\left(-\frac{\delta E}{kT}\right)$$

to obtain the apparent Fermi distribution $2n = N\{1 + \exp(\delta E/kT)\}^{-1}$. They then showed that, for practical temperatures and activation energies, multiple charging of the islands may be ignored. The total current is taken to be the sum of the three types of tunnelling transition (neutral-neutral (charge separation), positive-negative (recombination) and positive-neutral or negative-neutral) weighted by the product of the appropriate population densities. Although the different transitions were identified by Swanson *et al.*, the mobility concept (which links them in a sequence) was not added.

Leaver²³ also separated the conductivity expression into charge carrier density and mobility but regarded both as thermally activated. His argument involved discussion of the dielectric constant and will be covered later in the review.

Hill^{16, 24}, Kiernan and Stops²⁵ and later Barr and Finney²⁶, in a fundamentally different approach, assumed two initially neutral islands and then calculated the tunnelling current between them in the conventional manner but with the requirement that the tunnelling electron must possess sufficient excess to provide the electrostatic energy. This is satisfied by an equivalent downward shift δE in the source island's Fermi level. The Barr and Finney²⁶ result

$$J = \frac{4\pi me}{h^3 B^2} \frac{\pi B k T}{\sin(\pi B k T)} \times \exp(-A\bar{\phi}^{1/2}) \left[\frac{\exp(-B\delta E) - \exp(-BeV)}{1 - \exp\{(\delta E - eV)/kT\}} + \frac{1 - \exp\{-B(\delta E + eV)\}}{1 - \exp\{(\delta E + eV)/kT\}} \right] \quad (3)$$

is regarded as the most exact of these (it has been derived in Appendix A) since it reduces to the standard tunnelling equation between parallel electrodes when $\delta E = 0$ and to the approximation

$$J(V, T) \approx \frac{8\pi me}{h^3 B^2} \frac{\pi B k T}{\sin(\pi B k T)} \times \exp(-A\bar{\phi}^{1/2}) \{1 - \exp(-B\delta E)\} \sinh\left(\frac{eV}{kT}\right) \exp\left(-\frac{\delta E}{kT}\right) \quad (4)$$

when $\delta E > kT$ and $\delta E \gg eV$. In this approach, the current density is evaluated at one point along its path (which is also taken to be the point of charge carrier creation) and the mobility concept is unnecessary. The main points which must be verified experimentally are whether or not the conductivity can be separated into carrier density and mobility terms and, if so, which of these is thermally activated. Conventional Hall techniques have not yet yielded this information. Observation of the drift of injected charge (e.g. by electron irradiation) might help to achieve this more easily.

1.70

2.2. Electrostatic activation energy

Disagreements about the form of the electrostatic activation energy δE arise in roughly three categories: (a) the use of an appropriate relative dielectric constant ϵ_r , (b) the assumption of particular island-gap geometries and (c) fundamental electrostatic arguments.

The choice of ϵ_r raises two questions. Is the low or high frequency value appropriate, and how does one approximate the composite medium? Early workers¹⁹ tended to use low frequency values and to average these for the substrate and vacuum (air). Since the islands are surrounded more by vacuum than by substrate, results that were lower than expected for ϵ_r could be justified. However, when low values also seemed necessary to reconcile²⁷ the measured δE value for cermets with the theoretical value, it was suggested that high frequency values were appropriate to tunnelling as a fast process^{21, 23, 27-29}. This can only be disputed if it is assumed that δE is associated with a static equilibrium charge distribution rather than with the generation process itself.

The island geometry usually assumed in the early days (e.g. by Neugebauer and Webb¹⁹) was that of a hemisphere, but the mathematics always entailed the assumption of spherical islands. Later, the more generalized island shape (Fig. 1) of an oblate spheroid was proposed by Hill¹⁶, by Dryer²¹ and by Morris^{27, 28, 30} (with limiting spherical and thin disc forms), and even an extension for finite contact angles was attempted³¹. The modifications mentioned all yield higher δE values than the simple spherical island model does for a given observable island dimension parallel to the substrate. This is because the island volume is less for both an oblate spheroid or a hemisphere than that of a sphere of equal radius, and in either case the calculation with a spherical model should yield values of δE that are smaller than the measured values.

There are more fundamental problems, however, which should be resolved first. Neugebauer and Webb's¹⁹ expression (eqn. (2)) for spherical islands is determined on the basis that an electron removed from an island to the edge of the

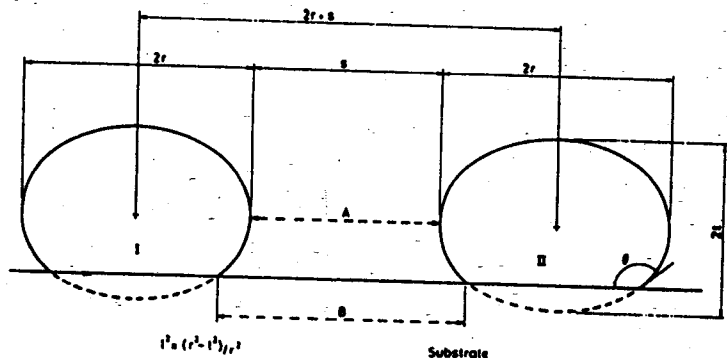


Fig. 1. The generalized oblate spheroid model of island geometry (the eccentricity is f).

next island, a distance s away, yields two adjacent charged spheres. (The geometrical modifications cited above also use this approach.) Hill¹⁶ used the same approach as Darmonis¹⁸ that the second island is a perfect conductor and that the energy required s that which would be necessary to reach the centre of the acceptor island if it were not there, i.e.

$$\delta E = \frac{e^2}{4\pi\epsilon} \left(\frac{1}{r} - \frac{1}{2r+s} \right) \quad (5)$$

Swanson *et al.*²² calculated the capacitance of two isolated spheres and, including image effects, showed that the amount of work which must be done in transferring an electron from one to the other is

$$\delta E = \frac{e^2}{4\pi\epsilon r} \left\{ 1 + \frac{r}{s} + \left(\frac{r}{s}\right)^2 + \left(\frac{r}{s}\right)^3 + \left(\frac{r}{s}\right)^4 + 3\left(\frac{r}{s}\right)^5 + \dots \right\}^{-1} \quad (6)$$

The use of image effects in a two-island model is probably not valid for an infinite array, where multiple image charges modify the elementary polarization effects¹⁶. Dryer, however, used image effects to contest the validity of Neugebauer and Vebb's expression for the field around a single charged island²¹. It is not clear whether his objection is valid if δE arises from the original charge separation.

The model developed by Abeles *et al.*²² for cermet probably holds considerable promise if it can be extended to the two-dimensional case of a discontinuous film. These workers represented the spherical island as surrounded by an infinite spherical conductor of internal diameter $r+s$ (Fig. 2(a)). The energy E_w required to remove an electron from the island to the enclosing conductor is $\frac{1}{2}(e^2/4\pi\epsilon) \times (r+s)^{-1}$. Charge equality, however, requires that two oppositely charged islands be formed, needing a total energy $E_c^0 = 2E_w$. The law of mass action was then invoked to yield a carrier density proportional to $\exp(-E_c^0/2kT)$, i.e. an activation energy δE equal to half Neugebauer and Webb's¹⁹ value and directly applicable to their concept of an equilibrium density of well-separated carriers. It could be noted that the initial charge separation (which provides the basis of

alternative approaches) may well be followed by immediate recombination. The extension to the discontinuous film has not been attempted, but a similar model such as that shown in Fig. 2(a) and (b) or (c) might be employed. The approximate model of Fig. 2(c) may ease some asymmetry problems; or alternatively the "thickness" t might be set to zero. Even if no account can be taken of the two distinct dielectric constants, the single island-conducting sheet model is undoubtedly more realistic than others for either the equilibrium or the generation concepts.

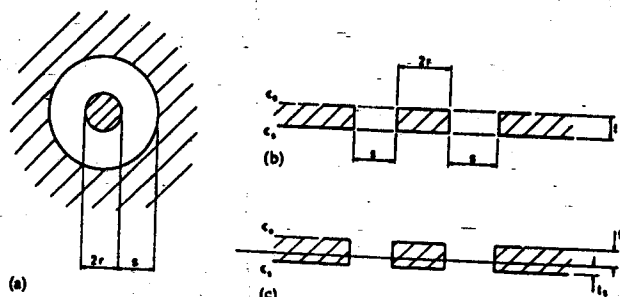


Fig. 2. (a) The cermet model of Abeles *et al.*²² which is also applicable to discontinuous films. (b) A plane view of a similar model of a discontinuous film. (c) A modified plane view where $t_0 + t_1 = t$ and $t_0 = t_1 / (\epsilon_0 + \epsilon_1)$.

Dryer *et al.*²⁹ assumed a different point of view again and pointed out that, while two charged islands must be created in the bulk of the film, only one is involved in charge injection at the electrodes. A lower activation energy is therefore required for electrode injection (if we assume that δE is linked to charge carrier creation) which, it was concluded, must be the dominant source of carriers.

Until now we have not commented upon Leaver's²³ treatment of the activation energy because it contributes to other aspects of the wider problem.

(I) The activation energy δE for the creation of charge carrier pairs is stated to be $e^2/8\pi\epsilon r$; this combines the arguments of Neugebauer and Webb (that $e^2/4\pi\epsilon r$ is the energy required to remove an electron to infinity) and of Abeles (that δE is only half of this energy).

(II) In (I), ϵ' is the effective permittivity of the medium. For cermet, the insulator dielectric constant is modified by the concentration f of metal spheroidal islands to $\epsilon' = \epsilon(1 + 2\gamma f)/(1 - \gamma f)$ where γ is determined by island eccentricities. The problem is considerably more complex in discontinuous films, and Leaver is the only worker (to our knowledge) to attempt it. He does not claim to have achieved an exact solution, but by the combination of argument and evaluation of the total potential at a charged sphere (including the induced dipoles on the spheroidal islands of the surrounding two-dimensional array) he has shown that an expression equivalent to the cermet one may be derived. The problem is to evaluate an effective replacement for $f = \frac{1}{2} \sum_i \{ r_i^3 R_i^{-2} (R_i - r)^{-2} \}$ where R_i is the centre separation of the charged and polarized islands (of volume $(4/3)\pi r_i^3$). Further argument and approximation lead to the estimation of an approximate expression in terms of r

11.45 - 3:00
11.15, last 5 min

1.50

and s . Although there is clearly room for improvement, Leaver's method of attack here is the most realistic to date.

(III) The charge density is activated electrostatically, *i.e.* the low frequency value of ϵ' is appropriate. Leaver has pointed out that the simple charged-neutral transfer is also activated because tunnelling is accomplished too fast for relaxation mechanisms to follow. If we use the approach of Dryer *et al.*²⁹, the activation energy is therefore the difference between a high frequency charging energy and the static component already present in the charged island. The final expression

$$\delta E_{\text{mobility}} = \frac{e^2}{4\pi} \left(\frac{1}{\epsilon'_{\text{optical}}} - \frac{1}{\epsilon'_{\text{static}}} \right) \left(\frac{1}{r} - \frac{1}{2r+s} \right)$$

includes both the usual proximity term $r^{-1} - (2r+s)^{-1}$, since the charge moves across only one gap, and the polarization energy of the dielectric.

Naturally, if tunnelling were to proceed sufficiently slowly for relaxation mechanisms to follow, the mobility would not be activated. Electrons which strike the barrier surface at glancing incidence (at less than about 10^{-3} rad according to Leaver's calculation) satisfy this condition and provide a significant component of the current at room temperature if $\delta E_{\text{mobility}}$ is sufficiently high (50% for 0.35 eV). For lower values of the mobility activation energy, the transition occurs at lower temperatures and is accompanied by a drop in the total δE to the charging contribution alone.

(IV) At small activation energies it is conceivable that sufficient islands are charged for the usual Boltzmann relation to fail, and Leaver's expression for n pairs of charged islands, $n = N \{2 + \exp(\delta E/kT)\}^{-1}$, is similar to that of Swanson (see Section 2.1). In this circumstance the charge separation can no longer be regarded as infinite, and the charging energy δE is modified to $(e^2/4\pi\epsilon') (r^{-1} - \alpha R^{-1})$ where R is the effective charge separation and α a type of Madelung constant which may be calculated for a given film geometry (*i.e.* average values of r and s).

(V) Finally Leaver has recognized the problem of the distributed nature of the film and has developed an elementary form of percolation theory to handle it. In a two-dimensional array, the datum particle (from which tunnelling is about to take place) has approximately five nearest neighbours. One of these has just supplied the charge in question and does not enter the argument further. The charge will actually choose the smallest gap to one of the other four, and so the average tunnelling distance is taken to be the mean of the lowest quartile of the distribution of island separations. The datum particle is then taken (a little more arbitrarily) to have a radius equal to the mean of the uppermost quartile of the size distribution. Estimates for f (see (II)) and hence ϵ' are obtained by considering two of the surrounding islands to be separated from the datum by the mean of the lowest quartile and the other three separations to be the means of the other three quartiles.

Obviously we can argue about the details of this treatment, but its value lies in indicating the way for more sophisticated work. As a logical application of common sense to estimate a result it deserves commendation. It should be noted that, unlike the other percolation approaches based on variable-range hopping/tunnelling, this is a fixed-range model which does not recognize that percolation paths will change with temperature.

It will be clear by now that the basic concept of the origin of δE is uncertain. Apparently one electrostatic energy is required to separate two charges initially (adjacent islands and high frequency ϵ_r), while another value is necessary to maintain an equilibrium charged island density (well-separated charges and static ϵ_r). The problem is how to choose the appropriate model or how to reconcile the two concepts in a coordinated theory and to establish which effect determines δE . The choice of the low or high frequency ϵ_r will depend upon the result.

2.3. Tunnelling parameters ✓

Equation (4) may be rewritten for low fields as $\sigma = \sigma_0 \exp(-\delta E/kT)$ where σ_0 is usually taken as constant to a first approximation for a given film. According to eqn. (4)

$$\sigma_0 = K \frac{8\pi m e^2}{h^3 B^3 k T} \frac{\pi B k T}{\sin(\pi B k T)} \{1 - \exp(-B \delta E)\} \exp(-A \bar{\phi}^{1/2}) \quad (7)$$

where K includes a geometrical term to account for coplanar tunnelling^{16, 27}.

It is doubtful whether this expression is valid if one of the mobility viewpoints described in Section 2.1 is adopted. The non-activated mobility will be related to the tunnelling current with $\delta E = 0$, and from eqn. (3) (provided BeV is small)

$$\sigma_0 = K' \frac{4\pi m e^2}{h^3 B} \frac{\pi B k T}{\sin(\pi B k T)} \exp(-A \bar{\phi}^{1/2}) \quad (8)$$

where K' is now related to the recombination length²⁰. In the rest of this section we can ignore these points since most aspects of the tunnelling problem are common to either viewpoint.

It will be noted in the derivation in Appendix A that the electron energy is separated into two components parallel and perpendicular to the island surface. It has been claimed by Abeles²⁰ that this approach is valid only for tunnelling between plane parallel electrodes and that a curved surface emits more readily than a planar one does³³. The comment has some validity and raises the equivalent problem of determining an effective tunnelling area. A satisfactory approach would be to adopt the division of energy technique and to integrate the contribution of each elemental surface area. The effective tunnelling distance must be determined for each of these by establishing equipotential contours between islands and by using the condition that the tunnelling path must be perpendicular to these.

Any such calculation will require a conformal mapping technique with assumptions of island shape, contact angle with the substrate etc. It will be easier to consider the cermet case first.

The potential barrier height between islands is usually assumed to be reduced from the height of the substrate conduction band or metal work function (depending on the assumed tunnelling path) by image effects. (Hill¹⁶ has stated that the plane parallel electrode case is a reasonable approximation, but Dryer²¹ has objected to this and has claimed that barrier lowering between islands is significantly less.) Any such assumption is, in fact, highly questionable because of the speed of the tunnelling process³⁴ which is completed before charge relaxation within the islands can occur.

Hill¹⁶ has shown that a Schottky barrier can form at the island-substrate interface and that island Fermi levels may shift as a result of charge exchange, but there has been no satisfactory quantitative development of these points.

It is generally accepted (and certainly the experimental evidence so far supports the view) that tunnelling between islands takes place through the substrate. Schmidlin³⁵ has pointed out that the presence of a positive ion at the centre of a tunnelling barrier considerably enhances the transmission probability, although Gadzuk³⁶ has argued that this would be unlikely to occur in practice since the position of the ions would not in general be at the centre of the gaps and so the enhancement would not be as great. (As far as we are aware, the literature does not cover the case of negative ions.) A single ion acts as an electron trap and reduces both the barrier height and the barrier width, and tunnelling takes place in two steps. The width reduction is usually expressed as a reduction in effective electron mass m^* to permit the continued use of the observed gap width. It will be appreciated that the two-step tunnelling process is much more than twice as likely as the single-step process owing to the inverse exponential relationships. In many cases, gap widths of the order of 100 Å have been observed, where measured currents are much larger than expected from simple tunnelling theory, and so it would seem that multiple ion-trap "hopping" or tunnelling can occur between islands. The model thus becomes a great deal more complicated since there must be a finite interval between each hop corresponding to the dwell time of an electron within the trap. Thus the charge mobility will be modified.

It will be apparent that this model is similar to Hill's early model (Section 2.4) for films of wide gaps and large islands^{16,37}. Clearly, not all of the ions will be situated at the same energy level; hence some thermal activation will be necessary for trap-to-trap transitions within the barrier, so that the process resembles hopping conduction in amorphous materials.

Clearly, whether a given transition occurs depends upon both the spatial and the energy separation of the traps, which is a similar concept to the hopping model of conduction. The thermal dependence of this process is swamped by the island electrostatic energy unless the latter is very small (*i.e.* large islands).

If tunnelling is the conduction mechanism, it may well be a multiple transition process since single-step tunnelling theory predicts currents that are much smaller than those observed. Furthermore, the effective tunnelling area calculation for a single-step process will yield values even smaller than those usually calculated on the basis of a planar model. Comparisons of theoretical thermionic and tunnel current densities incorrectly assume equal electrode areas and hence unduly favour tunnel currents. Even so, thermionic emission is expected to dominate at large gap widths¹⁶ if only single steps are involved.

2.4. Other theories

Since Swann¹ first attacked the conduction problem, a great variety of possible mechanisms have been proposed. Some lead immediately to a negative conclusion. Neugebauer³⁸, for example, has shown that substrate-island expansion mismatch cannot lead to an activated temperature dependence. Although space-charge-limited conduction with traps (which has been suggested as a mechanism in parallel with thermionic emission³⁹) seems to yield the correct temperature and field

dependences, the parameters evaluated from both measurements are not self-consistent²⁸.

For most other models a conclusion is less easily reached. A variety of Kronig-Penney models have been considered. A regular array of finite potential wells separated by identical potential barriers is obviously an unrealistic model but it does lead to a band structure²⁸. Ball *et al.*⁴⁰ used a random array but they achieved an activation energy equal to the difference between the barrier height and the Fermi level, which is subject to all the objections levelled at the thermionic model below.

Prior to the introduction of the electrostatic activation concept, thermionic emission was widely accepted as the mechanism responsible for charge transfer⁴¹⁻⁴³. Measured activation energies, however, are much less than the metal-insulator work functions, and image reduction of the barriers is usually invoked to explain the discrepancy. The measured range of δE requires very close islands for direct thermionic transfer but is not unreasonable for emission into the conduction band of the substrate¹⁶. A comparison of tunnel and thermionic currents shows that both tunnelling and thermionic substrate currents are possible for typical gap dimensions. Certain experimental results, however, cannot be explained on a thermionic model, even if it is extended. These results include strain measurements, the variation of resistance with time upon exposure to air, the magnetoresistance and the Hall effect which are described in Sections 3.7, 3.8 and 3.12. We have therefore totally discounted thermionic emission as a mechanism, even in wide gap films, since these also show the properties mentioned.

The existence of finite permitted energy bands in small particles is well known. Mostovetch and Vodar⁴⁴ included these in a Kronig-Penney model to establish a qualitative picture of semiconducting islands separated by insulating gaps. Hartman⁴⁵ considered an electron in the island potential well to determine the permitted energy levels. Electrons may then tunnel between the permitted levels in adjacent islands, except for the zero level which is uncrossed even at very low fields. (Zero-level transitions may also be inhibited by electrostatic considerations.) Predicted activation energies are reasonable, but Abeles²⁰ has objected to the one-dimensional treatment and has claimed that these values are too high by one or two orders of magnitude and that the existence of finite bands can be effectively ignored. The activation energy is taken to be the difference between the ground state and the first excited levels, *i.e.* $n = 0$ and $n = 1$. However, since transfer can only occur from near the Fermi level, where for small islands $n \approx 10^4$, the activation energy must be very small indeed. The model cannot therefore be regarded as realistic for this reason alone, apart from the usual objection that the application of finite fields should uncross permitted bands in adjacent particles and should cause the conductance to decrease and not to increase as is observed.

This point is not necessarily valid since it is unreasonable to expect adjacent islands to be identical. The model has never been extended to include distributed island dimensions (where we would expect only the wider high order levels to be crossed) or to predict the average behaviour as the field is increased.

One of the most recent theoretical propositions was made by Licznarski⁴⁶ based on the island-substrate charge transfer discussed by Hill¹⁶. One of the basic postulates of Licznarski is that the substrate work function is greater than that of the

island and that electrons are injected into the substrate by the island, thus raising the dielectric potential barrier. The equilibrium injection density, and hence the barrier height, are functions of temperature, leading to a complicated "activation energy" of the correct order that is dependent on the island size provided the metal work function and the substrate electron affinity are similar. This qualitative agreement does not explain the properties of those discontinuous films (e.g. platinum) in which positive charge is transferred to the substrate¹⁶. Essentially the Licznarski model proposes tunnelling through an inter-island barrier which decreases with increasing temperature or island size. Kadlec⁴⁷ has shown that the height of a trapezoidal barrier decreases with temperature but that the predicted variation is unlikely to lead to an activation energy.

A variety of other models have been proposed and are given names such as "trap-assisted tunnelling", "impurity conduction" etc. These range from that developed by Herman and Rhodin⁴⁸ after Wei⁴⁹ of an electron moving within the substrate to a point outside the island's electrostatic field (with δE similar to that of Neugebauer and Webb¹⁹) to those classified as "impurity band" and "hopping" conduction used by Tick and Fehner⁵⁰ to describe movement through a surface adsorbate monolayer. The latter picture has really only been developed qualitatively^{51, 52}. Hill has introduced a model based on tunnelling between substrate traps distributed in energy and space^{16, 37}. The experiment used will be discussed later (Section 3.9), but there are distinct similarities to the later concept of variable-range hopping, which may be a better view of transport across the gap than a single tunnelling step*. The essential difference between the models appears to hinge on the density and degree of localization of the sites via which transfer takes place.

2.5. Percolation approaches

Over the course of the last two or three years percolation theory has become widely used in solid state physics to explain a variety of phenomena such as conduction in amorphous materials. Latterly, there have been several attempts to apply it to conduction in discontinuous metal and cermet films, and the new theories represent a considerable improvement upon the original models although they are based upon them. The percolation approach allows us to extend our arguments beyond the idealized models of an array of equally sized and regularly spaced particles to the point where the random nature of the structures observed in practice can actually be taken into account. Before discussing the models which have invoked percolation theory it is appropriate to make a few general remarks first.

Percolation can best be described by resorting to a definition used by Pike and Seager⁵³ who stated that "... percolation models are composed of sites and bonds between sites.... In general their relation to physical problems is made by identifying the site with sources of interaction and the bonds with interactions of some minimum strength or greater.... The essence of percolation theory is to determine how a given set of sites, regularly or randomly positioned in some space, is interconnected."

* Variable-range hopping is discussed in this review in two separate contexts and it is important that these are not confused: (1) between traps within the inter-island gaps as an alternative to single-step tunnelling between islands (the present topic); (2) between islands as island sizes and gap widths dictate different optimum tunnelling paths at different temperatures (introduced in Section 2.5).

For charge transfer between metal islands in a two- or three-dimensional matrix, the islands can be taken as the sites and the extent of the overlap of the electronic wavefunctions of the islands can be regarded as the strength of the interaction. Therefore we can define the critical condition such that the contribution to the overall film conductivity of any pair of islands can be neglected unless the resistance to activated tunnelling between them is equal to or less than that to bulk or surface charge transfer through or on the dielectric substrate. The usual objection levelled at the application of percolation theory in this context is the abruptness of the transition which it predicts between a totally insulating and a conducting system. By defining the critical threshold for the existence of a conducting path between two sites in terms of the properties of the substrate, this objection can be overcome, and we can say that the array of islands (*i.e.* the film) will first conduct when an infinite cluster of critical "resistors" is formed; the magnitude of each "resistor" (adjacent pair of islands) is equal to or less than that of the substrate between them. This cannot be regarded as a strictly accurate definition, however, but it is perhaps helpful in overcoming the conceptual difficulties. It should also be pointed out at this stage that there is a second critical threshold in this context which is the transition from the discontinuous to the semicontinuous state. This aspect has been more extensively studied than has the variation of conductivity with temperature, and it will be discussed first. By so doing it is intended to indicate the possibly surprising similarity between the behaviour of this class of films and that of other, apparently quite different, physical systems.

2.5.1. The discontinuous-semicontinuous transition

The topic of percolation conduction in a variety of mixed insulator-conductor systems has recently been discussed by Coutts⁵⁴ who has pointed out that the difficulty in making quantitative predictions about their conductivity is due to the fact that the ratio of the conductivity of the two species is of the order of 10^{22} so that very small changes in composition can lead to very large changes in conductivity. However, predictions of the variation in conductivity of mixed systems with the fraction of the conducting component can be made, and even if this only gives an average value and uses a wide distribution it is still of interest. The critical condition for the formation of an infinite cluster can be expressed mathematically in the following way. If the probability of any bond between sites being present is p , an infinite cluster will not be formed unless $p \geq p_c$ where p_c is the critical probability. Frisch *et al.*⁵⁵ have shown that p_c can be determined from the empirical law

$$Zp_c = \frac{D}{D-1} \quad (9)$$

where Z is the coordination number and D the dimensionality of the lattice. For most of the regular crystallographic lattices this law is found to give an accurate estimate of p_c , although Sykes and Essam⁵⁶ have shown that small differences exist when these values are compared with the precise values determined by series expansions. To relate the mathematical concept of criticality to real systems we only need to multiply p_c by f , the packing fraction for the particular lattice under consideration, in order to obtain the critical area or volume fraction. For a simple cubic lattice with $Z = 6$ and $D = 3$ we have $p_c = \frac{1}{4}$. The maximum fraction f of the volume of this lattice which can be occupied by a random close-packed distribution

of equal-sized spheres is $\pi/6$. Thus we would expect the formation of infinite clusters at a critical volume fraction of

$$V_c = fp_c \approx 13\%$$

In the general context of conduction in mixed systems we would thus expect metallic conduction when the proportion of conducting spheres is equal to or greater than this value. Unfortunately, no account has been taken of any size distribution of the particles, their shape and the manner of preparation of the composite as pointed out by Pike and Seager⁵³. This comment is of course equally applicable to discontinuous films, but nevertheless the technique is attractive as a means of predicting the variation in conductivity after the transition from conduction by activated tunnelling to normal metallic conduction.

Essam *et al.*⁵⁷ have shown that the conductivity of a lattice with $p \geq p_c$ is given by

$$\sigma = \sigma_0(p - p_c)^\beta \quad (10)$$

where σ_0 and β are constants whose values are governed by the particular lattice. If we replace p and p_c by A and A_c or V and V_c , then this law can be applied to real systems. Coutts⁵⁴ has plotted $\log \sigma/\sigma_0$ against $\log(V - V_c)$ for a variety of systems including simulated lattices, thick silver films, hot-pressed carbon black and thin film cermets and has shown that all of these have a value of β in the range 1.3–1.9, thus indicating that the behaviour of microscopic and macroscopic systems is very similar. Of particular interest in the present context is the fact that Abeles *et al.*⁵⁸ have successfully applied the technique to cermet films, whilst more recently Liang *et al.*⁵⁹ have shown that it is equally applicable to two-dimensional bismuth films. These workers plotted $\log \sigma$ against $\log(A - A_c)$ to obtain the critical exponent β , having first obtained A_c from a plot of σ/σ_0 against A . It is interesting to note that their value of A_c (0.67) corresponds to an equivalent film thickness of about 9 nm. Andersson⁶⁰ studied the variation in the critical thickness with the temperature of the substrate and found that this was nearly constant at about 5 nm until the temperature exceeded approximately 100°C, after which it increased very rapidly to more than 20 nm at about 150°C. In view of the above comments regarding a critical area and the findings of Liang *et al.*⁵⁹ it would have been most interesting to have examined whether A_c was influenced by deposition temperature. On the basis of percolation theory A_c is expected to be fixed for a given substrate-metal combination, and a unique curve for the variation of conductivity with area coverage should be found. The deposition parameters such as temperature, rate, applied voltage etc. which are always found to influence plots of conductivity against thickness should have little effect if conductivity is plotted against fractional area.

Having shown that a percolation approach is capable of predicting the onset of semicontinuous conduction and the variation in conductivity with the metallic volume or area fraction thereafter, we shall now discuss its application to totally discontinuous conduction and the variation in conductivity with temperature.

2.5.2. Variation of conductivity with temperature

A particularly relevant, but apparently little recognized, piece of work was carried out by Zeller and Giaever⁶¹ who were interested in the effect of a layer of

small metal particles embedded in the central region of a metal-insulator-metal tunnel junction. This system was used to study the effect of the distribution of the activation energies of the particles on the tunnel junction. Interestingly it was shown that even for very small-sized particles with energy bands split into individual levels the mismatch of these could not give rise to the activation energy. This is identical with the criticism of the theory of Hartmann⁴⁵, about which we expressed doubt earlier, although it is not this point which is of interest here. The system was extended by Zeller⁶² who deposited many layers of particles within the junction, adjacent layers being separated by oxide barriers. This structure was considered to be analogous to that of pseudo-one-dimensional metals in which defects in the metal-like strands, such as vacancies or impurity atoms, were akin to the oxide barriers between islands. In both cases it was argued that the conductivity in the metallic component was unimportant and that it was the gaps which dominated the resistivity. Furthermore, and most important, it was considered that only nearest neighbour tunnelling occurred. The variation of the conductivity of both the one-dimensional metals and the small particle analogue with temperature was found to be of the form $\log \sigma \propto T^{-1/2}$. This is a result of particular significance to our discussion since the small particle analogue is precisely the system of interest to us here.

It is of particular interest that in deriving the well-known $T^{-1/4}$ law of charge carrier hopping between localized sites in amorphous materials Mott and Davis⁶³ simply optimized the tunnelling distance between the sites and the energy separation of the sites. If the same technique is applied to a one-dimensional conductor in which the hops are not confined to nearest neighbours but are allowed to be of variable range, then a $T^{-1/2}$ law follows immediately. Similarly, if the hops can take place in two dimensions then a $T^{-1/3}$ law is predicted. Knottek *et al.*⁶⁴ have actually examined this aspect of the problem by evaluating the power of T for a range of thicknesses. They have shown that when the films exceed a critical thickness there is indeed a change from a $T^{-1/2}$ to a $T^{-1/3}$ dependence, and this evidence tends to support the conclusion that it is the variable-range nature of the hopping process which gives rise to the particular temperature dependence.

However, Bernasconi⁶⁵ has analysed disordered systems in terms of percolation along critical paths in which only hopping between nearest neighbours is permitted. He has shown mathematically that even under this restriction a $T^{-1/2}$ dependence is expected provided that there is a distribution of activation energies. For the particular distribution assumed (*i.e.* $\rho(E) = 1/E_m$ where E_m is regarded as the maximum activation energy) it was shown that a plot of $\log \sigma$ against $1/T$ exhibited curvature over a central temperature range but approached straight lines asymptotically at high and low temperature, the slopes of which could be used to determine the coordination number Z and the maximum activation energy E_m of the system. Bernasconi processed Zeller's data and ingeniously calculated a value for the relative dielectric constant of the oxide which he claimed was in good agreement with experimental observations. Although a $T^{-1/2}$ dependence was predicted for the particular distribution assumed, Bernasconi pointed out that other forms could result if different distributions were assumed. This remark will be seen to be significant later. In view of the fact that the distribution of particle sizes in discontinuous films is now known to be of log-normal form, as has recently

been demonstrated by Grandqvist and Buhrmann⁶⁶ and by Andersson and Grandqvist⁶⁷, it may be expected that the distribution of activation energies is not as simple as that suggested by Bernasconi.

The similarity between amorphous semiconductors, one-dimensional metals and discontinuous films appears to have passed quite unnoticed, and this is really the purpose of giving so much space to the first two topics. Even the very recent theories which invoke percolation models, and which will be discussed next, bear a strong general resemblance to the ideas above, although differing in detail.

Abeles *et al.*³² have argued that, because of the mechanism of growth of two- and three-dimensional discontinuous films, which involves the diffusion of adatoms, large islands are to be associated with large inter-island separations so that within relatively macroscopic volumes (or, presumably, areas) of a film the ratio of separation to size will be approximately constant. By making this assumption and by accepting the standard expression for the activation energy it is a fairly straightforward exercise to show that the product of the activation energy and the separation is a constant whose value is calculable solely from the volume (or area) concentration of the metal. Abeles *et al.* have developed an expression for the conductivity of a given pair of islands, and, as for Hill's much earlier model, this involves two exponential terms, one of which ($\exp(-A\bar{\phi}^{1/2}s)$) is proportional to the tunnelling probability and the other ($\exp(-\delta E/kT)$) to the activation energy term. Since it was assumed that $s\delta E = \text{constant}$ it is reasonable to maximize the product of these terms to obtain the value of the separation of smallest resistance. This procedure is very similar to that originally used by Mott, which was mentioned earlier, and it leads to the now well-known $T^{-1/2}$ dependence for the conductivity.

In deriving their expression for the low field conductivity, Abeles *et al.* have argued that the most probable tunnelling path for an electron is along the chain of particles of constant activation energy and therefore constant spacing. They have then obtained the total film conductivity by integrating over all possible spacings. This assumes that there is no cross-linking between the chains, which seems a rather unlikely possibility.

Abeles *et al.* obtained good straight-line relationships for a number of metal-insulator combinations when they plotted $\log \sigma$ against $T^{-1/2}$, and they naturally accepted this as confirmation of their theory. However, Hill and Coutts⁶⁸ later pointed out that the very fact of making such a plot almost presupposes its validity, and they re-plotted the data in the form of $\log \delta E$ against $\log T$, as had been suggested earlier by Hill⁶⁹. This procedure indicated that the $T^{-1/2}$ dependence of the conductivity was not nearly as well established as may have been believed earlier. In a reply to this criticism Sheng and Abeles⁷⁰ suggested that the deviations may have been caused by the fact that Hill *et al.* processed all the published data including those measured at very low temperatures. However, it is at just such temperatures that a change of the power dependence might be expected, as we shall see later.

Heinrichs *et al.*⁷¹ also made the assumption that s/r was constant and applied the type of critical path analysis discussed earlier in this section. The starting point of their analysis was similar to that used by Abeles *et al.*³² in that they assumed that charge transfer was by a hopping mechanism and that the conductivity could be represented in the form

$$\sigma = \sigma_0 \exp\left(-2\alpha s_{ij} - \frac{E_{ij}}{2kT}\right) \quad (11)$$

where the meaning of the symbols will now be apparent. The critical-percolation technique of Ambegaokar *et al.*⁷², which was originally applied to amorphous materials, expresses the conductivity in the form

$$\sigma(T) = \sigma_0 \exp(-s_p) \quad (12)$$

where s_p is a critical value of the exponent such that

$$2\alpha s_{ij} + \frac{\beta E_{ij}}{2} < s_p$$

This can be rationalized to give a quadratic equation in s , the solutions of which are

$$s_{\pm} = \frac{1}{4\alpha} s_p \pm (s_p^2 - 4\alpha\beta f)^{1/2} \quad (13)$$

under the restriction that $s_p^2 > 4\alpha\beta f$ where f is a constant arising from the assumption that

$$sE = \text{constant} = f$$

These solutions tell us that there is a range of values of s available to an electron making a transition from an island centred at $s = 0$. Heinrichs *et al.* then proceeded to calculate the total available volume, within a film, into which tunnelling could occur. They applied the condition that the film would not conduct unless the critical available volume was equal to or greater than the critical volume fraction (as discussed earlier for the transition from tunnelling to metallic conduction). This, of course, presupposes that tunnelling occurs only via nearest spatial neighbours and that conduction first occurs when there is an infinite cluster formed by the overlap of the shells of available volume. This definition is very similar to the one we used earlier although it excludes variable-range tunnelling. Proceeding from this condition it was then possible to show that, over precisely definable ranges of temperature, the conductivity followed the $T^{-1/2}$ law predicted by Abeles *et al.* The ranges could be evaluated on the basis of structural data and were found to be in good agreement with those observed in practice. Perhaps surprisingly Heinrichs *et al.* claimed that, whereas in the case of hopping conduction in amorphous materials there is a clear change of the temperature dependence with dimensionality⁶⁴, the corresponding exponent for discontinuous films is expected to be independent of temperature. However, they pointed out that the range of temperatures over which the law is expected to apply will certainly change with film thickness.

The principal criticisms of this work are the assumption that $s/r = \text{constant}$ and the failure to recognize the defect in plotting $\log \sigma$ against $1/T^{1/2}$. As we pointed out earlier, observation of a straight line is not necessarily all the evidence required.

To some extent these criticisms have been overcome by Hill and Coutts⁷³. They assumed that the form of the particle size distributions $f(r)$ was any one of three

simple functions of the island radii and they claimed that, as far as the problem of charge transfer was concerned, the range of sizes was defined by finite upper and lower limits. The probability distribution of the centre-to-centre spacing $f(R)$ was assumed to be random since no alternative evidence was available.

These definitions were then used to derive another function $g(E_s)$ which represented the probability distribution of activation energies: because of the finite range of particle sizes, the range of $g(E_s)$ was also, of necessity, finite. The next step in the analysis was to obtain the distribution of edge-to-edge separations $g(s)$, since this, rather than the centre-to-centre spacings, is the principal interest in tunnelling. The probable number of particles available to an electron leaving a reference particle was then written as

$$N = \int g(s) \int g(E_s) ds dE_s \quad (14)$$

which was evaluated for all three assumed distributions of particle sizes in both two and three dimensions.

The remainder of the analysis followed the general critical percolation approach used by Ambegaokar *et al.*⁷² in deriving the formula for hopping conduction in amorphous materials. Since charge transfer between particles involves both spatially and energetically different sites, the probability of transfer can be written as

$$\exp\left(-\frac{E_s}{kT} + 2\alpha s\right) \quad (15)$$

where α is the characteristic decay distance of the electronic wavefunction. The array of particles was divided into two classes; firstly, a submatrix which has a probability of being connected (*i.e.* conducting) that is greater than some minimum critical value, and secondly a disconnected part which takes no part in the conduction process. If the critical probability is $\exp(-h_c)$, then hopping between sites would not be expected to occur unless

$$\exp\left(-\frac{E_s}{kT} + 2\alpha s\right) > \exp(-h_c) \quad (16)$$

(This can readily be seen to be a similar approach to that used by Heinrichs *et al.*⁷¹) Furthermore, the array as a whole will not conduct unless the average number of neighbours to which hopping can take place is greater than a critical value given by

$$N_c = \frac{D}{D-1} \quad (17)$$

Thus there will be no conduction unless

$$N > N_c$$

Figure 3 shows a diagram of s against E_s as described by eqn. (16). The quantities E_u and E_l are the upper and lower limits of E_s as dictated by the finite size range and are the limits over which E_s was integrated in eqn. (14). From the diagram it can be seen that there are two areas to consider: firstly, the rectangle bounded by the abscissae $s = 0$ and $s = (1/2\alpha)(h_c - E_u/kT)$ and the ordinates $E_s = E_u$ and $E_s = E_l$; secondly,

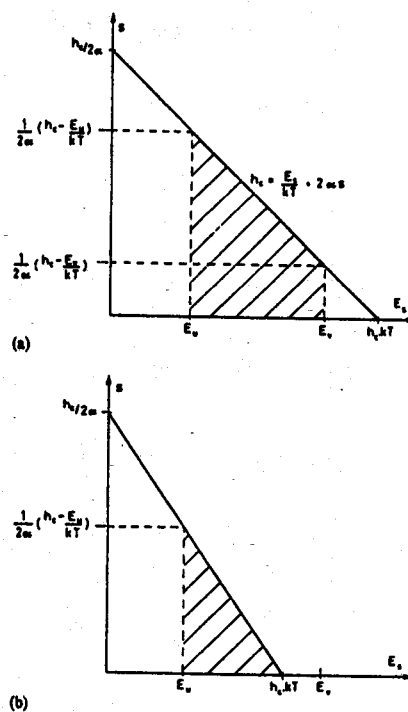


Fig. 3. A plot of s against E_s for the critical path model.

the triangular region bounded by the abscissa $s = (1/2\alpha)(h_c - E_u/kT)$, the ordinates $E_s = E_u$ and $E_s = E_l$ and the line $s = (1/2\alpha)(h_c - E_s/kT)$.

Integration of eqn. (14) over the rectangular area led to the result that $h_c \approx 1/T$ which corresponds to simply activated conduction. From the diagram it can be appreciated that this area will be dominant at high temperatures or if E_u is only slightly less than E_c . Thus, simply activated conduction can be expected at high temperatures or for very uniform structures when the range of E_s is small.

Integration over the triangular area gives an expression for the critical coordination number which under certain simplifying assumptions yields either $h_c \approx 1/T^{1/2}$ or $h_c \approx 1/T^{1/3}$ (the latter applies at low temperature). Thus this theory also predicts the temperature dependence found by Abeles *et al.*³² and by Heinrichs *et al.*⁷¹ but shows that this is a transitional form between the high and low temperature limits. Hill and Coutts⁷³ pointed out that, although these functional forms result from assumed distributions, it is apparent that similar forms will result from other distributions but that the ranges over which the various laws apply are likely to change.

Clearly, further work is necessary before the validity of this theory can be checked, and the authors have recently learned that very low temperature studies are soon to be made by Andersson⁷⁴.

It must also be emphasized that all models in this section are fundamentally dependent upon the assumption of single-step tunnelling between islands. For any trap-hopping alternative within the gaps it would seem that traps should be roughly evenly distributed between islands, and the exponential dependence of resistance on gap width is replaced by a simple linear dependence.

3. EXPERIMENTAL RESULTS

3.1. Experimental techniques

There are too many variations of the basic procedures employed in the preparation of discontinuous films to permit a comprehensive review, and we must therefore be content to make a few general observations. Experimental results from different laboratories are often totally contradictory. In all probability the explanation for this lies in different preparation techniques, but it is only when these details are provided that the origin of a discrepancy can be pinpointed. The problem is, of course, that there are so many significant influences on film properties that it is not practicable either to monitor every variable during preparation or to describe each step in publication. It is safest to rely only on the measurement of A as B is varied with all other parameters constant, and even then only within a single sequence of experiments from one laboratory. Attempts to compare absolute values from different groups rarely yield useful conclusions.

Discontinuous films have been prepared in various laboratories by sputtering, thermal evaporation and electron beam deposition in vacuum pressures ranging from less than 10^{-9} to 10^{-4} Torr. Of these, sputtering seems to provide the greatest reproducibility⁵².

Normally, glass substrates are used—preferably a high resistance ion-free smooth surface variety. Alkali halide crystals are also used but must be cleaved *in vacuo*, and outgassing presents problems³⁹. The aim is to achieve reproducibility so that the film structure can be reliably varied by changing a single parameter. Hill⁷⁵ has identified one source of the problem as surface mobility. The refractory metals³⁹ and platinum which have a small critical nucleus, a high nucleation density and a low surface mobility are therefore preferred to gold as the deposited metal, although the latter has been the most popular choice because it is easier to deposit (and more inert than nickel, silver or chromium which are also widely employed). The most important factor, however, is probably surface cleanliness and structure. A variety of cleaning procedures using detergents, organic solvents, vapour degreasing and rinsing in distilled or deionized water has been used since the more aggressive agents such as hydrofluoric acid may pit the glass surface. The essential step for reproducibility is (after conventional chemical degreasing) to bake²⁴ the substrate (for more than 4 h at 300°C in 10^{-5} Torr) to drive off the surface layer of water. Instead of perfecting the existing surface in this way, other workers have chosen to create a new surface in order to eliminate the need for substrate outgassing⁷⁶, e.g. by using⁷⁷ a film of 100–200 nm of sputtered Al_2O_3 or by using⁵⁰ SiO which also provides a suitable support for electron microscopy. Once a clean surface has been

prepared it can be contaminated by the residual atmosphere of the vacuum system, and to minimize this effect deposition should commence immediately and should be as rapid as is compatible with reasonable control over the process.

Substrate temperature also affects film structure and may be controlled in a number of ways: by radiation heating of the substrate face, by contact with a grounded heat sink or by direct electrical heating of the substrate rear. Electrical charges also affect film structure, and although substrate charge may be removed by suitable heating⁷⁸ the charge content of the vapour stream (consisting of both electrons and positive ions⁷⁹) remains. This may be reduced by biasing the source⁶⁰ or by applying a potential across the vapour stream⁷⁹. Monitoring the film resistance during deposition yields lower resistance films at a given stage of deposition than otherwise. (Hill²⁴ has commented that monitored film resistances continue to fall to unusable values after deposition ceases, possibly owing to a field-induced coalescence.) If we consider substrate temperature again, deposition at low temperature tends to produce an initially uniform distribution (because of reduced surface mobility). Agglomeration at higher temperatures then yields a more reproducible structure than is possible by direct high temperature deposition⁸⁰. Deposition rate also influences structure by a direct effect on island nucleation and a secondary effect due to the variation in the radiation heating of the substrate by the source. Rate and "thickness" are usually monitored by a quartz crystal microbalance. Probably the most often neglected influence on film structure is that of illumination intensity⁸¹ which appears to operate by electromagnetic interaction with the island structure⁸². While little can be done about the source illumination of the substrate, deposition in a darkened chamber will at least exclude external variations. There appears to be considerable variation in the nature of the contacts employed (e.g. contacts of the same material as the film, contacts of a different metal, silver paste as contacts) and when they are deposited. We are not aware of any particular advantages of varying the contact once the ohmic requirement is satisfied, and we shall make no further comment. However, the film structure near contacts is generally different to that in the central regions, and its effects are discussed in Section 3.6.

High resistances require the use of a suitable electrometer, and low frequency a.c. measurement²⁴ is favoured to avoid polarization effects. It is surprising, in view of the uncertainty of the role of film contacts, that four-terminal resistance measurement is not more common and that no attempt has been made at contactless measurement. The latter would require very low frequency inductive or capacitive coupling to avoid a.c. effects (Section 3.11). *RC* discharge techniques to measure high resistances⁸³ would probably encounter non-linearities due to non-ohmic effects (Section 3.5) although the technique has been employed by Mitchinson and Pringle⁸⁴.

Other more specialized aspects of experimental techniques will be discussed in the appropriate sections below.

3.2. Film structure

A typical sequence of film structures with increasing thickness is shown in Fig. 4. Most theoretical treatments assume a fairly uniform distribution of small islands (e.g. Fig. 4(a), (b)), but in practice non-uniform agglomerated films are often found

(e.g. Fig. 4(c), (d)). The double-peaked island distribution is due to secondary nucleation within the gaps and is obtained by depositing a mobile metal on a hot substrate^{30, 30, 79, 85}. Discontinuous properties may persist even in nearly continuous structures. Films even thicker than those shown but still consisting of small islands between either massive polycrystalline agglomerates or long winding filaments can exhibit discontinuous film properties. Hill²⁴ has proposed four structural categories: (I) small particles, small gaps; (II) small particles, large gaps; (III) large particles, small gaps; (IV) large particles, large gaps. He has concluded that tunnelling processes are limited to small gaps, while thermionic emission is necessary for large gap conduction. We have already rejected thermionic emission and have suggested that multiple trap tunnelling/hopping may occur within larger gaps. All experimental results are examined in terms of this type of mechanism.

Film structure is normally viewed directly by transmission electron microscopy, but first the film must be removed from the substrate. To achieve this the film is usually coated with carbon and stripped from the substrate. If a soluble substrate is used (e.g. rocksalt), it can be readily dissolved^{19, 86}. The carbon layer may be floated away from a glass substrate by immersion in water, the process being greatly facilitated by suspending the substrate over water for a few days beforehand²⁴.

The problem with this technique is that it is never certain whether the structure

observed is the same as that on the original substrate. In fact it is probably not, since the technique requires the metal islands to adhere better to the carbon than to the glass. The contact angle is a function of the interfacial binding energy, so island shape changes are expected and have been observed⁸⁷ under similar conditions. The purpose of the pre-suspension over water is to loosen the island-substrate adhesion by using the effects of water vapour. Increased surface mobility and some coalescence is therefore possible, but unless this preliminary procedure is undertaken it is particularly difficult to remove the contact regions. Recently it has been suggested²¹ that contact effects may play a major role in determining film properties, and it is imperative that this region is inspected on a routine basis for differences from the general area.

Clearly, if the film structure on the substrate could be determined, the problem of island shape changes would not arise. The most promising technique by which this may be achieved is the optical absorption/transmission of the film as a function of wavelength. In principle it seems probable that island sizes, separations and shapes and the distributions of these parameters may be determined by use of the appropriate assumptions⁸⁸⁻⁹⁵, but the theory is still inadequately established for confident use. It is also inappropriate for use with aggregated films²⁸. The advantage that the results are representative of a large film area is shared by the use of small angle electron scattering^{96, 97} which unfortunately also requires film removal from the substrate. Nevertheless, it is surprising that the latter technique has not found wider use, even though it has only been proved for small island structures.

The usual result published is a distribution of island sizes^{6, 98} although the distribution of gap widths is occasionally reported^{86, 98} (and this tends to be broader¹⁹).

The condition required by Abeles³² percolation approach that the ratio of island-to-gap dimensions is nearly constant is definitely not fulfilled in discontinuous films²³ where the two appear to be uncorrelated. One of the perennial problems is the assignment of appropriate island and gap dimensions in agglomerated film structures for comparison with measured values of δE . In general, agreement with theory requires minimum island and typical gap dimensions²⁸.

At low substrate temperatures islands tend to develop as discs and at high temperatures as spheres. Many structural investigations conclude that the best approximation to island shape is the oblate spheroid^{28, 48}, which is also the minimum energy configuration for a charged island^{28, 78}. Where this conclusion is based on the assumption of a known deposited average thickness caution should be exercised since it is almost certain that the sticking coefficient is substantially less than unity in the very low thickness range of interest^{78, 99}. Furthermore it probably decreases at high temperature⁷⁶, and at low deposition rates it is essentially zero¹⁰⁰ except where nucleation occurs at surface imperfections.

Recent work has dealt with the question of island shape in some detail^{60, 66}, and the contrary conclusion has been reached that the islands are prolate spheroids. (Leaver²³ has also indicated that tests in his laboratory show islands to be best represented in this form, with axial ratios of around 2:1.) This study begins with the assumption that the distribution of island volumes is given by the log-normal distribution function

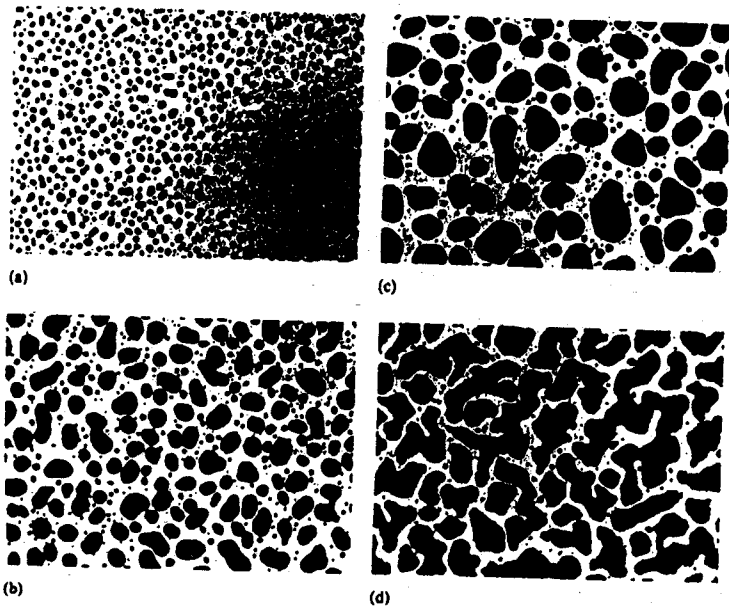


Fig. 4. Electron micrographs of ultrathin gold films deposited at 0.05 nm s^{-1} . The thicknesses are (a) 1 nm, (b) 4 nm, (c) 6 nm and (d) 15 nm. (From Kazmerski and Racine⁶.)

$$\frac{dN(v)}{dV} = \frac{1}{(2\pi)^{1/2} \ln \sigma_v} \exp \left[-\frac{1}{2} \left\{ \frac{\ln(V/\bar{V})}{\ln \sigma_v} \right\}^2 \right]$$

where \bar{V} is the median island volume and σ_v the geometric standard deviation. Each of the minor and major axes (a , b) and an "equivalent" diameter $(ab)^{1/2}$ independently fit the log-normal distribution also. Eccentricity increases with film thickness and coalescence as the films begin to show filamentary characteristics. The analysis is pursued further still to compare deposited "thicknesses" with total prolate spheroidal island volumes. As a final step, island eccentricities are plotted against film thickness for deposition with and without an applied field. In accordance with the results of Ahilea and Hirsch¹⁰¹ the field leads to increased eccentricity and elongation. Most of the experimental results leading to the proposition of the oblate spheroid are equally valid in support of the prolate spheroid, and the prolate spheroid which has been largely neglected in theoretical work will have to be included in any future arguments.

There are several structural unknowns which have never been satisfactorily investigated. The island-substrate contact angle has never been directly established for combinations of interest. A recent note has suggested¹⁰² that its variation with temperature should be measured directly. Nor has it been established whether the amorphous "carpet" between islands postulated by Tick and Fehlner⁵⁰ actually exists and if so whether it is due to surface contamination, a physisorbed water layer or single metal atoms. In addition the melting point depression exhibited by small metal islands has been well established¹⁰³⁻¹⁰⁶ experimentally down to a diameter of about 2.5 nm, but extrapolation of these results suggests that islands of diameter less than about 1 nm melt at around room temperature. This is typical of the island dimensions encountered and used to explain the activation energy. If the islands are liquid or the outer skin making up most of the island volume is liquid, what are the implications for the tunnelling theory? How is the Fermi energy affected at the surface? Many of the fundamental concepts would break down, and a complete revision of the theory would be needed.

Undoubtedly, the most useful advance in the structural area would be the production of regular structured films. Instead of attempting to produce a constant island size with random gaps and positions, it may be possible to produce a square array of islands. This might be achieved by setting up perpendicular standing wave patterns on the substrate surface, but it is clearly easier in theory than in practice. It appears that electromagnetic waves¹⁰⁷ offer the only possibility of achieving appropriate wavelengths. Alternatively, lines of islands might be achieved by sweeping the substrate with a sufficiently low energy beam of metal ions—provided that adequate focusing were possible.

3.3. Activation energy

The wide acceptance of an electrostatic activation energy is quite remarkable considering the continuing lack of agreement between calculation and measurement. It probably stems to some degree from dissatisfaction with all alternative propositions. However, most experiments yield values of the correct order for observed structures and, while many workers have found that prediction exceeds

measurement, an equal number have claimed the opposite to be true. There is still doubt as to the appropriate dielectric constant or island-gap geometry to be used.

The question of whether the high or low frequency dielectric constant should be used cannot be completely resolved by published results. Comparisons of measured δE with structure put¹³ ϵ between 14 and 0.4; this must be compared with an expected low frequency ϵ , of less than 6, say, or high frequency values between 1 and 2. (It should be noted that different reports use different theoretical formulae in reaching these figures.) Leaver's²³ recent arguments relating to dielectric constant and activation energy were summarized in Section 2.2. According to these, there are two separate activation energies (charge density and mobility), and the appropriate dielectric constants are different for each. The query posed above—whether the high frequency or low frequency value should be used—is clearly oversimplified, and the discrepancies which the traditional comparison would produce vary strongly with film structure.

One area of uncertainty lies in the choice of appropriate structural parameters to use in the theoretical expressions for δE . Very few workers have taken into account the contact area or have allowed for any structural differences. Within the bulk of the film there is clearly a distribution of r and s . Hill²⁴ selected minimum gap widths (in an attempt to find the preferred current percolation path) and the most popular island dimension; however, Morris^{28,31}, when working with aggregated films (where the density of small islands is low but still appears to dominate the resistive properties), took typical gaps and minimum island dimensions. Recent developments in the application of percolation theory may help to clarify this issue, but it is still open to question whether or not contact effects dominate and whether examination of the interior film is relevant to δE at all^{21,29,108}.

Van Steensel¹⁰⁹ has attempted to establish whether the high frequency value for ϵ should be used by using a ferroelectric substrate. Dingle¹¹⁰ has recently confirmed his result that δE does not change at the substrate Curie point. While this suggests strongly that the low frequency value (which should change at the Curie point) is not involved, it is not entirely conclusive. Ferroelectricity is essentially a bulk property, and it is by no means certain that it exists sufficiently close to the surface to be relevant to electrostatic activation. Clearly, bulk crystal properties never extend right to the surface, and this uncertainty brings the result into question¹¹⁰.

Far too few authors have given details of how they have selected r and s (or from where on the film) and the expression used for δE often varies. Few authors have made quantitative comparison between theoretical and experimental δE values. Nevertheless, an indication of the typical variations of δE with structure is given in Fig. 5.

The most remarkable aspect of discontinuous film conductance is the wide range of temperature and conductance over which δE remains constant (e.g. sometimes as low as 77 K). In fact the theoretical requirement²⁴⁻²⁶ that $\delta E \gg kT$ in order for the simple expression $\sigma = \sigma_0 \exp(-\delta E/kT)$ to hold is one argument against this approach, since the linear $\log \sigma$ versus T^{-1} relationship seems to hold even when $\delta E \approx kT$. While it is true that a curvature of the plot is sometimes observed^{21,24,39}, the usual linearity over wide ranges with the observation of non-uniform structures must be a constant source of amazement. However, the

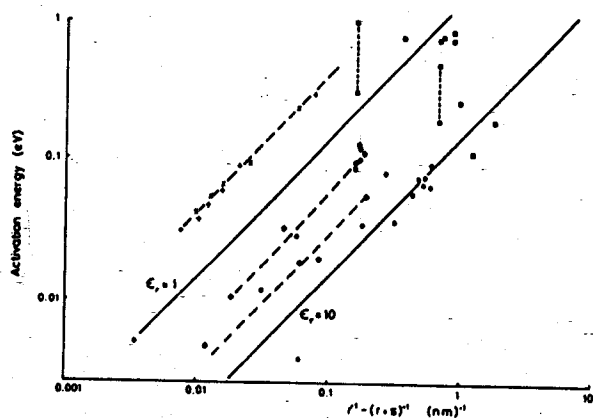


Fig. 5. The measured activation energy as a function of the observed geometry: —, theoretical results for $r = 1$ and $\epsilon = 10$; ●, experimental results of Tick and Fehlner²⁰; ○, Hill's²⁴ experimental results using a soda glass substrate; □, Hill's²⁴ experimental results using a 7059 substrate; ■, Morris's experimental results; ◇, Herman and Rhodin's⁴⁸ experimental results using a deposition temperature of 300 K; ◆, Herman and Rhodin's⁴⁸ experimental results using a deposition temperature of 100 K; ×, Kazmerski and Racine's experimental results using a deposition rate of 0.05 nm s^{-1} ; +, Kazmerski and Racine's⁶ experimental results using a deposition rate of 0.5 nm s^{-1} .

percolation approach of Hill and Coutts⁷³ may be able to reconcile some of the difficulties.

As indicated in Section 2.5, recent developments in the theory based on a percolation model suggest that a linear Arrhenius plot should only be expected at relatively high temperatures and that in this regime agreement between calculated and observed activation energies is usual; average structural parameters are used to obtain the calculated activation energy⁷³. At lower temperatures Abeles *et al.*³² and Heinrichs *et al.*⁷¹ have predicted a $T^{-1/2}$ variation, but Hill and Coutts⁷³ have suggested that this is merely a transitional region between the high temperature T^{-1} behaviour and the low temperature $T^{-1/3}$ variation.

Few published results cover as wide a temperature range as might be desired. However, when the deposition temperature is exceeded, irreversible structural changes are produced and below about 77 K problems of high resistance measurement are encountered. Over the widest published range the empirical formula $\sigma = \text{const. } T^n \exp(-\delta E/kT)$ has been derived⁴⁴.

The fact that δE remains constant with distributed geometries may, however, support models based on any type of distributed picture where the substrate conductivity is modified by the overall effect of the islands, e.g. the band model of Palatnik *et al.*⁵¹ and Boiko *et al.*⁵²

Herman and Rhodin⁴⁸ have found that $\delta E \propto r^{-1}$ to first order and furthermore that $\delta E \propto (\text{island mass})^{-1/2}$ indicating that the islands are nearly disc-like¹¹. As the deposition temperature was increased the relationships still held but a new

value of z , was necessary, suggesting that the platelets changed to a more spherical form.

It is interesting to note the different experimental approaches to the problem here. Most authors have estimated mean values of r and s , have calculated a theoretical δE value and have compared this with experiment²⁴, whilst others have plotted experimental δE values against an assumed function of r and s . (This is a superior approach since it obviates the need for any assumption about the dielectric constant.) Monchaud *et al.*¹¹² found the values of r and s required to match the thickness and island density to the observed values of δE and concluded that these were similar to the observed values of r and s .

3.4. Tunnelling parameters

We can say little about the agreement between experimental and theoretical values of σ_0 since little relevant information is provided in the literature. Clearly σ_0 and δE must be considered separately, but usually only δE is tabulated. To extract σ_0 from δE and the film resistance requires the latter to be normalized for a square film.

Furthermore, the actual film dimension must be known if possible contact effects are to be taken into account. Usually, none of this information is provided. It can be assumed, however, that film dimensions are constant for a given publication, and some results from δE and the resistance at a known temperature or from activation plots can be deduced. To a first approximation and on the assumption of single-step tunnelling between islands, it is expected that $\log(\sigma_0)^{-1}$ will be directly proportional to the island separation within a single set of results. Furthermore, we also expect s to decrease as the film thickness increases, and therefore σ_0 to increase as δE decreases. However, experimentally σ_0 does not usually increase as δE decreases.

Kazmerski and Racine⁶ found a very slight steady decrease in σ_0 (less than 30%) with increasing thickness over a range where gap width, though observed to be decreasing, was nearly constant. Neugebauer and Webb¹⁹ obtained a similar result, but the reduction in σ_0 as δE decreased from 0.64 to 0.18 eV exceeded four orders of magnitude. It can only be concluded that either the theory is in error or that the dominant effect of thickness increase is island coalescence. Both these results were obtained for platinum films. Hill²⁴ did not present sufficient results for a given system to establish the trends, but at least his results for platinum on soda-lime glass showed that the variation of σ_0 with gap was in the correct direction. Owing to ionic drift, however, results on soda glass are of dubious value. For other systems (Pt-7059, Au-soda-lime) σ_0 varied by 1.5–2 orders for essentially identical gap widths. The results of Herman and Rhodin⁴⁸ and of Morris^{28, 31} show no consistent trends whatsoever.

Possibly it is expecting too much of the theory to predict observed values of σ_0 when σ_0 contains so many possible variables. The principal points being made in this section are as follows.

(1) Together with comparisons of experimental and theoretical δE values, we might expect to find some reference to σ_0 —such data are not readily found in the literature. Tabulations of room temperature resistance are not particularly instructive.

(2) Where such data can be deduced there is usually no obvious consistency^{28, 31, 48}; this is probably the reason that they are so seldom found.

(3) Where there is a consistent trend, it is not as expected^{6, 19}.

(4) It is highly dubious to calculate the absolute resistance of a specific film and to claim agreement with theory if no consistent trend can be established over a selection of films.

We shall add the general observation (only occasionally made in the literature) that σ_0 generally tends to be higher than expected¹³. The example quoted by Kiernan and Stops¹¹³, where σ_0 is over an order greater than predicted, is quite conservative. Large discrepancies are often accounted for by adjusting m^* , $\bar{\phi}$ etc.

As a separate exercise, individual tunnelling parameters have been evaluated by Hill²⁴ who noted a curvature of the Arrhenius plot at high temperatures. He attributed this curvature to the term $\pi BkT/\sin(\pi BkT)$ which will increase from unity as πBkT increases. The dominant dependence on tunnelling parameters is the exponential term $\exp\{-4\pi d(m^*\bar{\phi})^{1/2}/h\}$. $m^*\bar{\phi}$ can therefore be evaluated from the absolute magnitude of the conductance evaluated in the linear range where πBkT is small. The assumption made was that the absolute error in $m^*\bar{\phi}$ will be small even though first-order approximations must be made to m^* , $\bar{\phi}$ and gap width in the calculation. More serious problems arise with the possible existence of preferred percolation paths and in the estimation of the effective tunnelling area (Hill has used $0.25 \times$ average particle size \times film width) and the length for coplanar (substrate) tunnelling with a very uncertain island geometry. The main objection, however, is to the use of a single film conductance value at all. It would be better to determine $m^*\bar{\phi}$ from a plot of $\log \sigma_0$ versus s . Such a plot, however, does not show the expected variation, as already pointed out.

Nevertheless, when $\bar{\phi}/m^*$ is evaluated from the curved portion of the Arrhenius plot, the results for $\bar{\phi}$ and m^* are quite reasonable: $\bar{\phi} \approx 0.9$ eV, $m^* \approx \frac{1}{2} \times$ rest mass for 7059 glass. Hill then proceeded to modify $\bar{\phi}$ to allow for image reduction of the potential barrier and he compared the result with the bulk activation energy of conduction in the substrate† to estimate the effect of band bending due to interface charge transfer. The final value is very close to the difference between the substrate electron affinity and the metal work function.

In fact the agreement with expectation is quite astonishing and is a powerful argument for single-step tunnelling between islands. In addition, credence is lent to the statement that, while silver, gold and chromium are electron injecting, electrons flow from the glass substrate to the platinum island in accordance with expectation. However, Morris³⁰ has pursued a similar line of argument based on the fact that no curvature was noted on his activation plots and has concluded that the effective mass must be less than one-thousandth of the rest mass. (A similar value has been obtained for cermets¹¹⁴.) This result must point to multiple-trap-assisted tunnelling (hopping) but may be due to greater surface contamination. If this is the case, a distinction must be drawn between two different types of surface impurity since Kazmerski and Racine⁶ found increased curvature (*i.e.* increased B) in films

† It should be noted that Dryer²¹ has claimed that this will be dominated by ion currents and therefore bears no relation to the substrate's electronic properties. Hill has not mentioned whether ionic effects were taken into account.

deposited under poor vacuum, which is probably indicative of an increased barrier height (see Section 3.8).

We note here that Boiko *et al.*¹¹⁵ have evaluated $\bar{\phi}$ and m^* from strain data (for gold on mica) as 3 eV and 2.8 times the rest mass respectively.

3.5. Non-ohmic effect

As for many discontinuous film phenomena, observations of non-ohmic effects vary widely. The first positive statement that can be made is that the film conductance universally increases with applied field E_a . Beyond this, results (or rather authors' interpretations of their results) fall into two main categories: (a) $\log \sigma$ increases as $E_a^{1/2}$ or (b) σ is of the form $A(T) + B(T)E_a^{1/n}$ where $n > 1$.

We shall consider case (a) first, since it is usually associated with variations in the electrostatic activation energy. The usual approach is to consider the activation energy δE within the context of an applied field. The electrostatic energy which must be thermally supplied is decreased by an amount which is supplied by the field itself. Thus (on a model^{19, 28, 116, 117} corresponding to eqn. (2))

$$\delta E = \frac{e^2}{4\pi\epsilon} \left(\frac{1}{r} - \frac{1}{r+s} \right) - seE_a \quad (18)$$

where E_a , the field within the gap, is related to E_a by $sE_a = (2r+s)E_a$. In developing this equation, the activation energy was regarded as a "barrier" to charge separation, and it can be seen that the expression is only valid up to a minimum field

$$E_{a\min} = \frac{e}{4\pi\epsilon}(r+s)^{-2} \quad (19)$$

beyond which the "barrier" develops a maximum and

$$\delta E = \frac{e^2}{4\pi\epsilon r} + reE_a - 2 \left(\frac{e^3 E_a}{4\pi\epsilon} \right)^{1/2} \quad (20)$$

until the field reaches $E_{a\max} = (e/4\pi\epsilon)r^{-2}$ when δE vanishes. (Similar treatment has covered the cases of ellipsoidal and disc islands³¹.) This approach predicts a linear field reduction of δE at low fields and the $E_a^{1/2}$ reduction at high fields. The theoretical variation by Morris¹¹⁶ (Fig. 2), is remarkably similar to the observations of Weitzkamp and Bashara¹¹⁷ (Fig. 4) but to few others. The general approach, however, does predict that non-ohmic effects, as they are due to a reduction of δE , will be more prevalent at low temperatures, in agreement with observation^{19, 44, 118, 119}.

These effects are also more noticeable in thicker films¹¹², presumably because of field enhancement by island growth. The most general discrepancy is an apparent absence of the linear region^{19, 21, 29, 39, 44, 60, 120}. In these cases it is suggested that the low field dependence is in fact linear with E_a .

When the non-ohmic effect is so widely attributed to δE reduction it is surprising that more investigators have not attempted to measure δE as a function of field. Hill²⁴ has shown that δE reduces linearly with $E_a^{1/2}$ for one field, and for a very small reduction (Fig. 6) Morris³¹ has plotted δE as a function of $E_a^{1/2}$ for assumptions of both spherical and disc island geometries. Quantitative agreement with theory was not particularly good and required the assumption of considerable

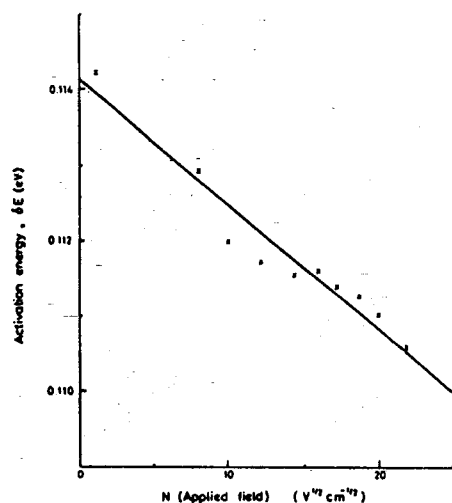


Fig. 6. The activation energy of conduction for a thin film of platinum plotted against the square root of the applied field. (From Hill²⁴.)

field enhancement due to the aggregated film structures. The feature of special interest in these results is the high field saturation without δE approaching zero. (Others^{44, 120} have also observed high field saturation with no mention of δE reaching zero.) Morris³¹ has attempted to explain the effect by including all the field effect contributions, and this appears to account satisfactorily for the high field saturation, but the necessary consistency between theory and experiment for both δE and $\log \sigma$ could only be obtained for structural parameters that are very different from those actually observed. A different approach is at present being pursued and has yielded promising results. Details will be published shortly¹²¹, but the basic model involves a distribution of island-gap dimensions along the current path. Most of the applied voltage is dropped over the high resistance combinations, and hence the initial value of δE will drop rapidly with field until the voltage is distributed more uniformly. One early result of the study suggests that the type of field effect described above requires the ratio r/s to be approximately constant along the path. This condition is, of course, the basic assumption of the percolation model of Abeles *et al.*³² and is unfortunately somewhat dubious.

Van Itterbeek *et al.*¹¹⁹ have found a linear relationship between $\log \sigma$ and $E_s^{1/n}$, where n appears to be 3 for nickel and iron and possibly higher for cobalt. Hirsch and Friedman¹²² have attempted an explanation based on an absorbed gas layer, but we suspect that some separate mechanism might be involved with ferromagnetic islands. Hök *et al.*⁷⁷ have noticed a cube root effect with aluminium contacts to a gold film. The origins of the different results are still not clear and deserve further investigation.

Boiko *et al.*⁵² noted a further complication. At very low fields σ increased

rapidly with field. At moderate fields the conduction was ohmic, and in the high field region $\log \sigma$ increased as $E_s^{1/2}$. The low field result needs verification but could possibly be caused by a distributed geometry as described above.

Milgram and Lu¹²³, Joglekar *et al.*¹²⁴ and Kollarits and Speiser¹²⁵ have analysed the non-ohmic effect in terms of a two-process model including a space-charge-limited current. While their raw data seem very similar to those obtained by others it must be noted that the relationship $I = AV + BV^2$ which they developed appears to provide an adequate fit. A and B are functions of temperature, but no simple relationships in terms of structural parameters were obtained.

It is known that the field effect is not caused by Joule heating in the film nor is there any redistribution of field along the film, so it is not purely a contact effect⁴⁴.

While field enhancement appears to be reasonable in aggregated films, Weitzenkamp and Bashara¹¹⁷ found that a quantitative fit required lower fields than those observed. Their interpretation of this result led them to the conclusion that the field (*i.e.* the activation "barrier") extends over several islands. This view would be in agreement with earlier comments that charges must achieve more than a single gap separation before they can be regarded as "free".

3.6. Contact effects

During the preparation of this review we have concluded that the role of the contact is one of the most neglected aspects of the conduction problem in discontinuous films. Since little attention has been paid to the contact area until recently, few publications describe film structure in the contact region. The discussion in this section therefore tends to be rather speculative.

We begin with Dryer's proposition that, since charge carrier creation in the bulk of the film involves the complementary charging of two islands, the charging of a single island adjacent to an electrode requires only half the charging energy^{21, 29, 126}. This idea leads to a more fundamental modification of the theory than a mere division of the film into end and middle regions with differing δE values. If we assume first that the film structure is uniform from contact to contact, the substantially lower energy required may mean that this is the only carrier generation process of any significance, especially where $\delta E \gg kT$. The model then develops along the following lines. An electron is injected at the negative end, singly charging an adjacent island (originally neutral); the subsequent passage of the electron through the film is by tunnelling from charged islands to neutral islands with no associated energy change (on the assumption that all islands are identical and that there is a non-activated mobility). Similarly, electron removal from an island adjacent to the positive terminal may be regarded as hole injection. Clearly the model is very similar to the charge density and mobility approach of Abeles³² (and recombination effects must be included) but with the new concept of charge injection. A simple experiment to confirm the model might involve localized heating of the centre of the film to determine whether the conductance variation is strong (owing to δE) or weak (owing to tunnelling mobility).

The idea has certain attractive features. The widespread conflict between theoretical and experimental δE values (Fig. 5) is neatly explained as far as available data permit. First, Hill's experimental values for δE are roughly half those expected using a reasonable low frequency ϵ_s for the substrate. Furthermore, Hill²⁴ has

commented that the structures described were the same adjacent to the electrodes. The excellent agreement of the tunnelling part of Hill's conductance (Section 3.4) is also consistent with the model. Hill's¹⁶ evaluation of the current density implies a uniform distribution of charge carrier generation across the film. This model is more relevant to the proposed electrode injection than to the bulk of the film (as intended) where experiment shows that current is largely confined to percolation paths¹⁷. The results of Morris²⁸ quoted in Section 3.3 were obtained from nearly continuous highly aggregated structures. Island dimensions were taken from those observed between large agglomerates or meandering filaments. In at least some of these films, a more transparent region was visible near the electrodes where a more discontinuous structure is now believed to exist. Up to now the existence of this region had been ignored, not only because it is difficult to strip the film near the electrode from the substrate but also because the field distribution was uniform along the film. It would appear that carrier density may be determined by this region, whilst the mobility is governed by the intervening region. As pointed out earlier, Leaver²³ has suggested that the mobility may also be activated because of polarization and relaxation effects. No information is available on the end structures of other films in Fig. 5, but we tentatively assume that these may also differ from those observed.

Different end structures have been noted by several authors, and a variety of explanations have been advanced. Goer *et al.*¹²⁶ have claimed that the different structure does not appear unless a field is applied during deposition and that the increased coalescence (*i.e.* larger islands, wider gaps) is initiated by field enhancement in the electrode region and is then further increased by the larger potential drop maintained by the different structure. (It should be noted that, in Fig. 5, Morris's films were deposited with an applied field and Hill's films without.) This proposition disagrees with an alternative explanation of filamentary structures in the bulk of the film; these are also attributed to field effects while the contact regions similarly have a large island, large gap structure^{101, 128}. Ahilea and Hirsch^{101, 128} have suggested that the electrodes trap radiation from the deposition source and that the agglomeration results from high temperatures near the contacts, although intuitively the metallic electrodes would be assumed to have a much higher IR reflectivity than the very tenuous film itself (see also p. 56). Possibly both effects play an important role.

Roussow *et al.*¹²⁹ showed that vapour diffusion around the mask led to a significant gradation of the contact into the film area, and Borziak *et al.*¹³⁰ have also related different structures to geometric masking effects. Borziak *et al.* have differentiated between (i) homogeneous films (no structural difference at the contacts) obtained by ensuring a sharp electrode edge, (ii) symmetric inhomogeneous films where the end structures are identical but different from the middle area and (iii) asymmetric inhomogeneous films where two distinct end structures are formed by offsetting the source position relative to the substrate.

Clearly, the origins of end structures are not well understood and would be a profitable area for future study.

Borziak *et al.*¹³⁰ have described some intriguing observations with their inhomogeneous films. The asymmetrical type produce asymmetric resistance properties, yielding diode-type behaviour (Fig. 7) and possibly the most immediately practicable application of the films. The implication here is that the

charge injection processes for electrons and holes are quantitatively dissimilar. Only Dryer²¹ has really tackled the problem quantitatively to date, and he did not recognize this possibility. Borziak *et al.*¹³⁰ also noted that the potential drops were greatest near the contacts (consistent with the model since gap widths are greatest here), that they varied with polarity (consistent with differences in electron and hole injection) and that they were always greatest at the positive end, as shown in Fig. 8. (There is also a suggestion that this produces a thermal gradient along the film.) In contrast, Dittmer¹²⁰ obtained maximum potential drop at the negative end, but with films "formed" by the electrical breakdown of continuous films.

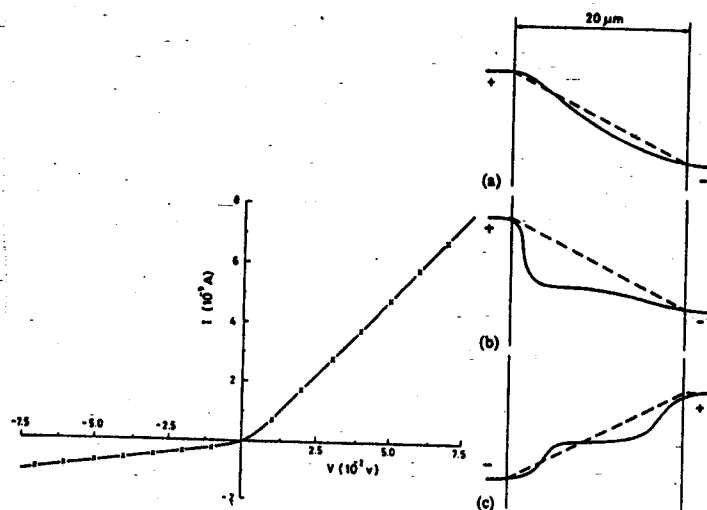


Fig. 7. The dependence of film current on voltage and polarity for an inhomogeneous structure. (From Borziak *et al.*¹³⁰)

Fig. 8. The potential distribution across discontinuous gold films: (a) homogeneous; (b) inhomogeneous; (c) inhomogeneous (reverse field). (From Borziak *et al.*¹³⁰)

Films "formed" by Morris¹³¹ in a similar way showed the asymmetric and switching characteristics obtained by Borziak *et al.* Ahilea and Hirsch^{101, 128} noted potential drops close to the electrodes with silver films but a uniform field distribution with platinum. Where the distribution was non-uniform, coating the region close to the electrodes destroyed the discontinuous properties. Until shown otherwise, we are tempted to assume that where field distribution is concentrated at the electrodes the intervening region is actually continuous.

Dryer *et al.*²⁹ have pointed out that a field enhancement factor of $2(l/R_e)^{1/2}$ operates at the electrode, where l is the film length and R_e is the radius of curvature of the electrode edge. Certainly on the basis of this model we would expect the field reduction of δE to be operative only in the contact region, but no one has yet analysed data along these lines. The very high field enhancements necessary to force

a fit between theory and Morris's³¹ data, however, seem to support the idea. (Nevertheless, high field saturation of the non-ohmic effect without the expected zero δE value suggests that there may be a secondary contribution to δE from within the film¹³².) It is worth noting that Ahilea and Hirsch^{10, 101} and Hök *et al.*⁷⁷ have reported $V^{1/3}$ non-ohmic effects which vanished when the contact region was overcoated but that Mostovetch and Vodar⁴⁴ and Milgram and Lu¹²³ have claimed to show that non-ohmic effects are not related to the electrodes. This confusion should be resolved by further careful structural and electrical measurements.

The role of the contact region proposed here is also supported by a.c. measurements (Section 3.2) and switching effects (Section 3.13).

3.7. Strain

There has been considerable interest in the application of discontinuous films as strain gauges. It is not surprising therefore that much of the literature describing the variation of film resistance with applied strain has an applied bias, *e.g.* it describes how to maximize the sensitivity or the stability^{2, 133}. We shall concentrate on results which are relevant to the conduction process itself.

The traditional treatment of strain sensitivity^{77, 134-136} uses only the width of the tunnelling barrier; *i.e.* the film resistance is expressed as

$$R = K \exp\left(\frac{\delta E}{kT}\right) \exp\left\{\frac{4\pi}{h}(2m^*\bar{\phi})^{1/2}s\right\} \quad (21)$$

and

$$\gamma = \frac{s}{R} \frac{dR}{ds} = \frac{4\pi}{h}(2m^*\bar{\phi})^{1/2}s \quad (22)$$

If K varies with s in the model adopted, the expression is modified slightly^{30, 136}.

Nishiura *et al.*¹³⁶ have pointed out that the known increase in γ with s is in agreement with this model but directly contradicts expectation for a thermionic conduction mechanism where an approximately inverse relationship is predicted. The thermionic model is therefore totally incompatible with strain experiments.

Data analysis on the basis of eqn. (22) provides reasonable support for a tunnelling mechanism. Parker and Krinsky¹³⁴ have obtained quantitative agreement and have concluded that tunnelling must occur by successive hopping between traps. Nishiura *et al.*¹³⁶ found that γ was proportional to s with reasonable accuracy for different films, but their absolute value of $m^*\bar{\phi}$ seems a little high. We tend to agree with their own comments on the accuracy of the determination of s by observation of simultaneously deposited films on different substrates. This is never particularly reliable.

The simplified treatment above is based upon the assumption that δE is independent of strain. Witt¹³⁷ has considered the point in more detail and has concluded that δE variations can be neglected for the range of structural parameters of interest. The experimental facts remain, however, that δE decreases with strain^{30, 115} and that this effect may yet prove to be significant. Morris³⁰ has treated the subject theoretically and has shown that the δE reduction is consistent with a model of flattened disc-like islands which also extend with the substrate. The observation of the δE decrease therefore supports the flat island model (for which

there is other evidence) and suggests that island-substrate adhesion is strong (which is rather unexpected). Quantitative agreement between theory and experiment was poor. Boiko *et al.*¹³⁸ have recorded the simple variation (γ proportional to s) for small gaps, but at large gap widths γ decreased as s increased until the gauge factor vanished, as shown in Fig. 9 (and presumably would eventually go negative). Without actually evaluating the change in δE , m^* and $\bar{\phi}$ may be derived from the results by including the second-order effect of $\pi BkT/\sin(\pi BkT)$ in the analysis.

The temperature coefficient β of the gauge factor is given by

$$\beta = \gamma^{-1} \left(\frac{\partial \gamma}{\partial T} \right)_\varepsilon = -\alpha + \frac{\partial}{\partial T} \left(\frac{\partial R}{\partial s} \right)_T \left/ \left(\frac{\partial R}{\partial \varepsilon} \right)_T \right. \quad (23)$$

where ε is the strain and where $\alpha = -\delta E/kT^2$ is the film TCR. Witt and Coutts¹³⁹ have suggested that $\beta = -\alpha$ and that the second term is universally zero, but while this appears to be the case for continuous films¹⁴⁰ and a reasonable approximation for many discontinuous structures¹⁴¹ it is clearly not true where δE varies significantly with strain. The intriguing possibility exists of developing a strain gauge with a zero temperature coefficient of gauge factor.

Noise and stability are the principal problems with strain experiments^{30, 137, 139}. Islands probably migrate, substrate-island adhesion may fail and (in nearly continuous films) filaments may break. Results should be improved by the

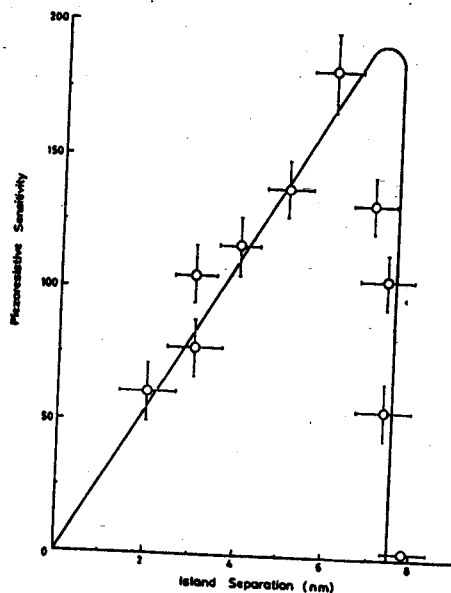


Fig. 9. The dependence of the strain sensitivity γ on the gap size s between islands in a film of gold on mica. (From Boiko *et al.*¹³⁸)

use of low amplitude cyclic straining with lock-in techniques for the recovery of small variations from the noise.

3.8. Gas absorption

It is well known that discontinuous film properties may change drastically after the film's removal from vacuum. For this reason general experimentation must be performed with the films *in vacuo* wherever possible. Any experimental data measured in air, for example, must be taken quickly to minimize the effect of change, and indeed it is debatable whether meaningful results can be obtained in air at all. We must distinguish between two types of environmental effects.

(a) The ambient atmosphere may modify a certain film behaviour which is directly caused by some other agency. Sections 3.18 and 3.19, for example, will describe the modifying effects of atmosphere on irreversible structural changes.

(b) In this section we are concerned with effects which are directly caused by the environment itself. In general these may be subdivided into two categories: (i) interaction with the metal island (e.g. oxidation) which tends to be irreversible; (ii) substrate effects (bulk or surface) which seem to be reversible.

Fehlner¹⁴² has described the first type of action with the example of oxygen on zirconium films. At a pressure of 10^{-7} Torr the oxygen formed a monolayer on the metallic island which increased the resistance sharply as the island work function (and hence the tunnelling barrier) increased. Subsequently, the resistance increased more slowly as the island oxidized and the effective gap width increased. Fehlner derived the inter-island spacing and the room temperature oxidation rate from the experiment. The former agreed with direct observation of the film, and the latter agreed with previously published data. Later work by Deshpande¹⁴³ suggests that oxidation rates determined in this way are consistently less than accepted values (although similar) possibly because of the very low pressures used.

The second type of effect has been described by Morris⁸⁵ for gold films in air. Basically, the resistance variation with time is of the form

$$\log R(T, t) = A_1 + A_2 t^{1/2} \exp\left(-\frac{p}{T}\right) \quad (24)$$

(provided the film temperature is constant) where p is characteristic of the absorption of the ambient gas by the substrate material. An equation of this form was derived on the basis of the increase in the tunnelling barrier by substrate absorption of air⁸⁵. By determining the rate of resistance variation with time when T is constant, p may be derived for the gas-substrate combination. Values determined for air or oxygen in soda-lime or Corning 7059 glass were in reasonable agreement with published data. This type of effect is reversible, and the activation energy of conduction was not affected.

Both effects are entirely consistent with the electrostatically activated tunnelling model, and the second specifically points to substrate tunnelling. The results could not be explained by a thermionic model.

A number of other workers have noted reversible decreases in resistance. Kazmerski and Racine⁹ noted that the $\pi BkT/\sin(\pi BkT)$ curvature discussed by Hill²⁴ occurred at very much lower temperatures (about 200 K) for films exposed to

oxygen or deposited under poor vacuum. Clearly B is larger for these films, i.e. $(m^*/\bar{\phi})^{1/2}s$ is larger. This could indicate that Morris's⁸⁵ argument based entirely upon increases in $\bar{\phi}$ is an oversimplification. Certainly m^* , $\bar{\phi}$ and the effective gap width are not entirely independent in a real barrier, but the interrelationships are ill defined. The reversible decreases in conductivity observed with the absorption of hydrogen, nitrogen, carbon monoxide or air are clearly related to this sine term, but it is still not easy to reconcile these results with the previous theory. Perhaps the lower temperature range employed permits an adsorbed surface layer to form, providing an additional parallel conduction path. Hirsch and Friedman¹²² have also used very low temperatures and have obtained $\log R$ versus $t^{1/2}$ linearity for a resistance decrease. Their theory is similar to that of Morris⁸⁵ (to explain the $t^{1/2}$ behaviour) but incorporates a voltage term. They immersed films in liquid air, nitrogen, hydrogen and helium, however, rather than in a gaseous environment and the resistance appears to decrease in all cases. Mostovetch and Vodar⁴⁴ have also noted resistance decreases with adsorption at very low temperatures.

A number of other workers have noted oxidation effects^{112, 125, 144}, and Kollarits *et al.*¹²⁵ have noted that these do not influence the activation energy. Others have invoked an explanation involving substrate absorption^{117, 145}. From these last papers it would appear that argon has no effect on film resistance, while the effects of nitrogen¹⁴⁶ and water vapour are still in dispute. It would be instructive to investigate the effects of these and other gases (e.g. hydrogen, helium) more fully.

Williams and Stone¹⁴⁷ have also identified a physisorbed surface layer of water as having a major effect on film conductance in the laboratory atmosphere. When this surface layer begins to boil off at 100°C, the film resistance increases drastically. Comparison of these results with those of Morris⁸⁵ indicates that laboratory humidity has a significant effect and suggests a possible application of the films as humidity sensors (see Fig. 10).

Much more complex reversible absorption/desorption effects have been studied by Barr¹⁴⁸ using hydrogen with palladium films. The consistent explanation of all observations involves a fast work function increase with surface adsorption by the island (fast resistance increase) and an increase in island lattice parameter (slow resistance decrease as the gap shrinks) as slow absorption into the island continues. At first sight it appears that substrate absorption with the H^+ ions decreasing the barrier height would be an alternative explanation of the last effect.

Only one publication describes the effect of a metal vapour (mercury) where condensation simply short circuits islands and capillaries¹⁴⁹; Lazarov and St. Manev¹⁴⁹ have also found that oxygen absorption by a discontinuous gallium film at 77 K decreases the film resistance and that water vapour leads to either a decrease or an increase of resistance depending on the measuring field. This last result is particularly interesting and needs intensive investigation. If the results previously described are all subject to the measurement field employed, the conclusions may be totally invalid.

Note that most of the workers cited above regard the gas either as absorbed by the substrate bulk to modify the bulk trap density or as adsorbed on the metal surface to modify the work function. No one has attempted to combine both effects, nor has anyone attempted to predict the relative significance of these effects.

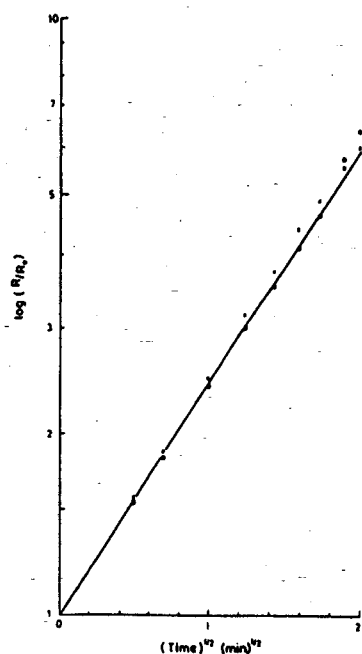


Fig. 10. The resistance increase during exposure to wet (O) and dry (x) air. (From Morris³¹.)

3.9. Substrate bias and ion drift

Hill³⁷ was the first to apply a d.c. potential to the rear of the substrate and to study the resultant modification in film properties (Fig. 11). To a first approximation it is expected that mobile Na^+ ions in the substrate will migrate towards the film with a positive applied bias, increasing the film conductance owing to an increased ion density within the gaps, and that the opposite will occur with negative bias. Modulation of the film conductance has in fact only been observed for substrates with an appreciable Na^+ content²⁴ and correlates well with typical trap densities³⁷. Physical evidence of Na^+ migration has been observed by Fehlner and Irving¹⁵⁰, and the drift time constants also suggest ion movement. The variation of electrical properties cannot, however, be unambiguously explained on the basis of this simple model.

In the first place, the activation energy must also be changed by substrate bias³⁷. Secondly, although positive bias increases the film conductance for a soda-lime substrate, this seems to be the result of activation energy reduction rather than of a change in tunnelling barrier which seems to increase. Thirdly, the use of a borosilicate substrate gives a highly non-linear effect³⁷. The explanation proposed by Hill^{24, 37} is based upon the existence of electron trap sites within the forbidden energy gap. Furthermore, the assumed energy distribution of the traps decreases

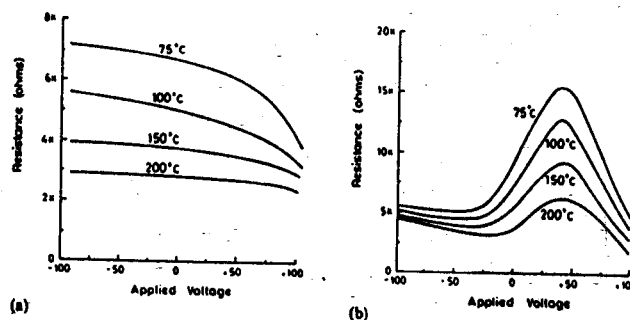


Fig. 11. The variation in resistance of gold films as a function of applied bias voltage and temperature: (a) soda glass substrates (0.05 in thick); (b) borosilicate glass substrates (0.02 in thick). (From Hill³⁷.)

away from the allowed bands to a minimum within the gap. The different effects with the two substrates (Fig. 11) can be explained in terms of the relative positioning of this minimum and the Fermi level.

The results quoted all apply to low resistance quasi-continuous films, and this leads to the possibility that the thermal activation between the barrier traps might dominate over simple electrostatic activation energy effects.

Before entirely accepting this viewpoint, the implications of yet another observation must be considered. Only platinum films show no bias effect, and only platinum is not electron injecting but instead forms a depletion layer at the island-substrate interface²⁴. Clearly the substrate bias will induce a net surface charge on the film which will modify the interface charging effects between island and substrate. The effects of this island-substrate charge exchange on the activation energy has never been adequately accounted for, although Licznerski⁴⁶ has made some progress in this direction. This point must be developed theoretically before Hill's interpretation can be fully accepted.

An essential point in Hill's model is the assumed energy distribution of the traps. Verification should be possible by spontaneous thermal release experiments.

Once again, without independent information regarding the charge density and the mobility it is difficult to make reliable predictions, but interesting results should be obtained with non-ionic substrates if the charge capacitively injected into the discontinuous film by the bias field is comparable with the carrier density. This limit was probably not reached in Hill's experiments where the bias field and δE were very low. It should also be noted that no account appears to have been taken of possible current drain from the bias supply.

Substrate ions also affect the usual tunnel current when they are mobile. In this case drift effects are to be expected and have been observed¹⁴⁶ with films on soda-lime substrates where the mobile ion density (Na^+) is sufficiently high. An initially uniform distribution of ions within the barrier will provide weak trap sites and an anomalously large tunnelling current. The application of an electric field across the gap removes these ionic traps from the centre of the barrier to the edge (leaving, incidentally, negative space charge in the gap) where they have a negligible effect.

The tunnelling resistance therefore increases with time. If the applied field is now removed, a residual voltage caused by the ion displacement may be measured, and this decays exponentially with time.

There are still a number of unanswered questions about the results described. The reported behaviour with reversal of the applied field¹⁴⁶ is not totally consistent with the model described. Furthermore the displacement current which must accompany the ion movement should be subtracted from the total measured current. The experiment should also be performed at a variety of temperatures in order to establish that the expected variation of the residual voltage decay time corresponds to that of the ionic mobility.

The measurement of film resistance on ionic substrates clearly requires a.c. monitoring to avoid polarization effects^{24, 27, 118}.

3.10. Noise

To our knowledge, only two groups have successfully studied noise in discontinuous metal films. Williams and coworkers^{145, 151} have considered noise power as a function of frequency and temperature. Mazzetti and Stepanescu¹⁵² have investigated the effects of both substrate absorption and (with others) substrate bias on noise power spectra.

The only theoretical treatment is based on a model of electron tunnelling between islands and has led to an approximate noise spectral density¹⁵¹:

$$\frac{e^2(f)}{\Delta f} \propto R^4 f^{-1} T^{-3} \exp\left(-\frac{4\delta E}{kT}\right) \quad (25)$$

In general, a $1/f$ spectral density is observed, the power actually varying between 0.8 and 1.2. The experimental variation with film resistance is in reasonable agreement and, while an Arrhenius plot of

$$\ln\left\{T^3 e^2(f)/\Delta f\right\}$$

against $1/T$ is not quite the straight line predicted (Fig. 12), the deviation may be accounted for by a more exact expression which includes the $\pi BkT/\sin(\pi BkT)$ term to allow for the temperature dependence of tunnelling¹⁴⁵.

Activation energies δE may be found from the Arrhenius noise plots, and these agreed well with those measured from the d.c. resistance, provided that $\delta E > kT$ as assumed in the theory. Although δE ostensibly enters eqn. (25) as part of the temperature dependence of the tunnelling process, it would be more precise to say that it is part of the temperature dependence of the film resistance. The theoretical and experimental agreement was not inconsistent, however, with a model of non-activated tunnelling within the film and activated charge injection at the electrodes, as was discussed earlier.

Mazzetti and Stepanescu¹⁵² found that, while the conductance fluctuations $\langle \Delta G^2 \rangle$ and the conductance G both decreased sharply upon exposure of the film to air, $\langle \Delta G^2 \rangle/G$ remained roughly constant. This is stated as being in agreement with their theory (which predicts $\langle \Delta G^2 \rangle$ to be nearly proportional to G) but appears to contradict that described above. Subsequent experiments to modify the noise characteristics by the application of substrate bias yielded interesting results¹⁵³. A

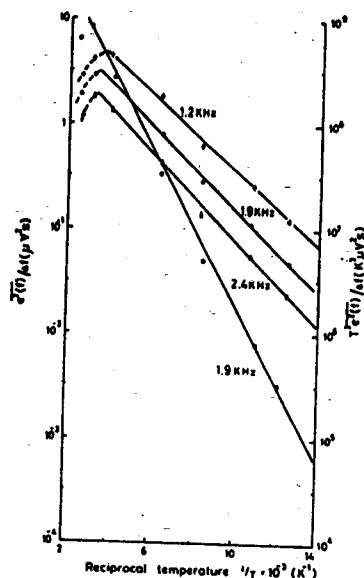


Fig. 12. The variation of current noise with the reciprocal of the absolute temperature for a platinum film with $\delta E = 0.026$ eV measured from d.c. conductivity. The corresponding value determined at a frequency of 1.9 kHz from the above figure is 0.022 eV. (From Williams and Stone¹⁴⁷.)

positive bias field injected mobile positive ions (no effect with an ion-free substrate) into the tunnelling gap, and a negative field removed them. The positive bias decreased the film resistance (as expected) and the noise level. The noise varied by a factor of 10 for a resistance variation of about 15%. The effect was reversible but could be "frozen in" by cooling with bias applied. (While the $1/f$ spectrum may be observed for both bias conditions, an interesting deviation is illustrated which varies with bias.) We are not entirely convinced that the results disprove the trap-assisted tunnelling model (as claimed). However, the tunnelling-based prediction that noise is proportional to R^4 is certainly not borne out by these results.

3.11. A.c. effects

An equivalent circuit of the island-gap combination is shown in Fig. 13: R_∞ and $R_0(T)$ represent the resistance of the island and the gap respectively. The parallel gap element $H(\omega)$ is usually assumed¹¹⁸ to be the gap capacitance C_0 . Although the experimental variation of film impedance with frequency (Fig. 13) is generally consistent with this type of model^{131, 143, 154, 155}, there is one major discrepancy which requires explanation.

Joglekar¹⁵⁶ has found that the theory of Fuchs¹⁵⁷ and Mayadas and Shatzkes¹⁵⁸ for conduction in continuous films may be extended to discontinuous films if r.f. measurement of the resistance is used so that $R_0(T)$ is short circuited by

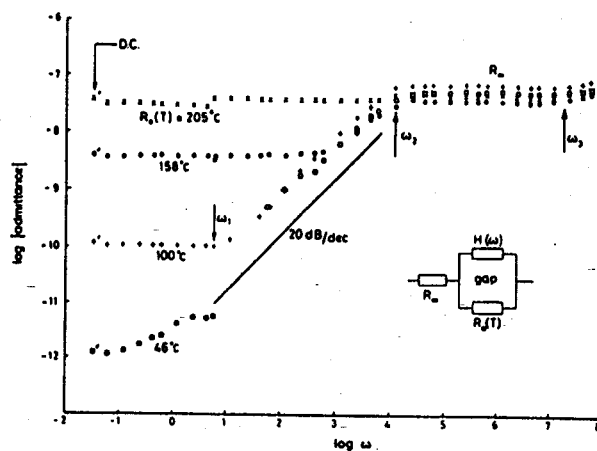


Fig. 13. The admittance of a discontinuous film as a function of frequency and temperature. (From Morris¹³¹.)

C_0 and only R_∞ is observed. By this method Harkay¹⁵⁹ has shown that R_∞ may exceed $R_0(T)$ even when the d.c. Arrhenius plot shows no curvature. This provides a warning that d.c. measurements may not be totally relied upon. R_∞ is dependent upon the island size owing to mean free path limitations. Morris¹³¹ has noted that transient effects are also consistent with the equivalent impedance network, even when the impedance-frequency Bode plot requires a more complex equivalent circuit than that above. Upon exposure of a molybdenum film to oxygen, the gap and island resistances increased and the capacitance decreased; these variations may be individually identified as functions of time¹⁶⁰. The problem arises when the gap capacitance is evaluated from the corner frequencies, yielding values between 0.2 pF (similar to the inter-crystallite capacitance in semiconductor films) and 10 pF (or even higher in some extreme cases)¹³¹. If the gap capacitance is estimated from a plane parallel electrode approximation, a much smaller value is obtained. Alternatively, the single-electron charging energy used to evaluate the electrostatic activation energy may be employed with a similar result. There is clearly something incorrect with this model.

Tick and Fehlner⁵⁰ separated a.c. resistance from the reactive component and found that the resistance was constant with frequency for palladium films; they obtained a gap capacitance of 2 pF. For their mixed Pd-Au island films, however, the a.c. resistance decreased with frequency as $\omega^{0.9}$, and this result was attributed to hopping through an "amorphous carpet" between islands. Morris¹³¹ adopted this explanation for the excessive apparent capacitance, but it seemed to be inconsistent with the phase shifts noted. We propose an alternative explanation.

This is based upon the concept that the activation energy is related only to electron injection at the electrodes (see Section 3.6). At sufficiently high frequencies the charge may be injected directly through the self-capacitance between the contact

and the main body of the discontinuous film, bypassing the activation mechanism. This self-capacitance will be much greater than that usually calculated between the two electrodes. Quantitative calculation of the proposal is not possible without information on the structure of the contact regions and the distribution of resistance between the main film and that close to the contact. This explanation is supported by the evidence that contact effects dominate film properties. In addition the results of Morris¹³¹, for example, were obtained from films with visible structure differences near the electrodes. One of these films also showed the asymmetrical effects characteristic of structure differences between the two electrode regions.

No doubt there are other plausible explanations for the apparent discrepancy in island capacitances. If the metal islands inject electrons into the substrate, the tunnelling barrier may be dominated by two narrow Schottky barriers at the island-substrate interfaces. Measured capacitances might then be those of the Schottky barriers alone. A platinum island, however, forms a depletion layer at the interface, and there should therefore be no Schottky barrier¹⁴. Since excessive C_0 values have also been determined for platinum films¹⁵⁴, we can eliminate this possible explanation of high capacitances.

Films have been observed to possess an inductance at very low frequencies for nearly continuous filamentary structures which also exhibit asymmetrical current-voltage characteristics and switching effects¹³¹. The original explanation based on filamentary inductance cannot be dismissed, nor can the concept of substrate ion drift. If further study eliminates these, however, a possible link with the switching mechanism might be pursued with profit.

3.12. Hall effect and magnetoresistance

We have already suggested that Hall effect measurements might resolve some of the fundamental questions about the conduction mechanism. In early experiments directed primarily at continuous films^{161,162} the magnitude of the Hall coefficient was found to decrease rapidly with film thickness to the limit of continuity. Initial unsuccessful attempts to observe a Hall effect were therefore made in an atmosphere of doubt as to whether any effect existed and of certainty that, if one did exist, it would be very small. Until recently, similar efforts to measure a Hall effect in cermets had also met with no success^{163,164}.

To our knowledge, only one group has successfully studied Hall and magnetoresistive effects in discontinuous films^{52,138,165}, while another has achieved significant magnetoresistance results for a cermet¹⁶⁶. These results have important implications for the conduction theory, and there is an obvious need for others to attempt similar measurements.

Measured Hall mobilities⁵² are much lower than quoted values for electron mobility in the substrate. If conduction takes place by thermionic emission of electrons from a metal island into the substrate conduction band, then the measured mobility is expected to be that of the substrate. The conclusion that thermionic emission does not play a significant part may be further generalized, i.e. any proposed mechanism which requires the conduction electrons to drift in the substrate conduction band may be discounted. It must be noted, however, that the Hall effect measurements were made on films composed of oriented filaments rather than of islands¹³⁸. The activation energies were low, and presumably both metallic

and activated tunnelling conduction mechanisms would contribute to measured properties. Exactly which led to the Hall effect is open to question.

Considerably more effort is needed in the pursuit of reliable Hall effect data, particularly where observation can be confidently related to an ideal discontinuous structure. It is not obvious when a Hall effect might arise in the ideal island array where the electrons do not "drift" in any usual sense. Provided there is no electric field within the islands the electron energy distribution should be isotropic, and any Hall effect must originate in the actual tunnelling process. It must therefore be determined whether or not maximum tunnelling probability is shifted from the electric field direction by the application of a perpendicular magnetic field. Even if this is the case, there may still be no observable Hall effect in a real film if the degree of disorder is sufficient to prohibit transport anywhere except along a small number of percolation paths.

The same point that the magnetic field does not affect an isotropic electron energy distribution within the island must be made when considering transverse magnetoresistance. Two experimental results have been reported, with totally different physical origins, each strongly indicative of the tunnelling mechanism.

Boiko *et al.*^{11,5} have stated that the effective width of the tunnelling barrier in a transverse magnetic field B_{\perp} is

$$\lambda(B_{\perp}) \approx \lambda(B_1) \sin^{-1} \left\{ \frac{\alpha(0)}{\lambda(B_1)} \right\} \quad (26)$$

where

$$\lambda(B_1) = 2(2m^* \bar{\phi})^{1/2} / \pi e B_1$$

Provided that $\alpha(0)/\lambda(B_1) \ll 1$, the increase in film resistance is given by

$$\frac{\Delta R(B_{\perp})}{R(0)} \approx \frac{4\pi}{h} (2m^* \bar{\phi})^{1/2} \{\alpha(0)\}^3 \{\lambda(B_1)\}^{-2}$$

The measured effect shown in Fig. 14 has two components. It is the positive component which interests us, and this clearly shows the predicted dependence on B^2 . (The small negative effect is due to localized spin ordering within the metallic filaments and should vanish with more discontinuous structures.) Once again a tunnelling mechanism is experimentally supported, but there are two points requiring further clarification. The increase in $\Delta R/R$ with temperature suggests that $m^* \bar{\phi}$ is temperature dependent, possibly because of phonon interactions with the tunnelling electron. If this is the case, then the temperature dependence is sufficiently marked that it must be incorporated into the basic theory. Equation (26) has also been quoted by Boiko *et al.*^{11,5} without reference, derivation or suggestion of its physical origin (presumably a classical trajectory in the magnetic field) and requires independent verification.

Earlier suggestions that the effect of the magnetic field upon the islands can be neglected are only valid for non-magnetic materials. Helman and Abeles¹⁶⁶ have analysed the temperature-dependent magnetoresistance results obtained by Gittleman *et al.*¹⁶⁷ with a Ni-SiO₂ cermet according to an activated tunnelling model. If the magnetic moments of the two islands involved in a tunnelling event are not parallel, additional or less thermal energy $2\Delta E_m$ must be supplied depending on

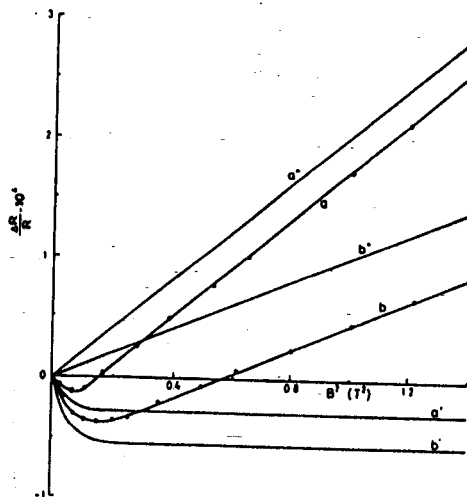


Fig. 14. The transverse magnetoresistance of a discontinuous film of gold on mica ($R = 2.9 \times 10^7 \Omega/\square$): curve a, experimental results at 300 K, resultant of the negative and positive components a' and a'' ; curve b, experimental results at 100 K, resultant of the components b' and b'' . (From Boiko *et al.*^{11,5})

whether the spin of the tunnelling electron is parallel or antiparallel respectively to that of the source grain. Consequently the approximate tunnel conductance must be rewritten as

$$\begin{aligned} \sigma(H, T) &\propto \exp(-A\bar{\phi}^{1/2}) \left\{ \frac{1}{2}(1+P) \exp\left(-\frac{\delta E + \Delta E_m}{kT}\right) + \right. \\ &\quad \left. + \frac{1}{2}(1-P) \exp\left(-\frac{\delta E - \Delta E_m}{kT}\right) \right\} \\ \sigma(H, T) &= \sigma(0, T) \left\{ \cosh\left(\frac{\Delta E_m}{kT}\right) - P \sinh\left(\frac{\Delta E_m}{kT}\right) \right\} \end{aligned}$$

where $\frac{1}{2}(1+P)$ and $\frac{1}{2}(1-P)$ are the probabilities of the two electron spin orientations. Now ΔE_m differs for the two cases of widely separated islands (superparamagnetic) and islands in close contact (ferromagnetic), leading to two separate transverse magnetoresistive effects which dominate at low and high temperatures respectively. The experimental results are clearly the sum of two distinct mechanisms.

Two totally independent studies of transverse magnetoresistance have been described.

- (1) Both strongly support tunnelling as the transfer mechanism.
- (2) The two theories approach the subject from opposite viewpoints—one considers only the tunnelling term, the other only the activation energy (and in a special case).

(3) The experiments provide strong evidence for a tunnelling mechanism and for an activation energy at least related to island energies.

(4) The theories have yet to be combined.

3.13. Switching

Switching has been noted by a number of authors^{21, 28, 52, 108, 125, 168-170} and naturally has created interest in some potential applications. Little effort has been directly applied to discovering its origins. It would appear that two distinct mechanisms have been observed. The first is S-type switching where the resistance is initially high and falls to a low value with the application of a critical field. The characteristics tend to be erratic, low to high resistance transitions also occur, results are usually not reproducible, the film is often destroyed by testing and it is believed that most observations may be explained by Joule self-heating with structural changes. In the second type (N-type), the film resistance rises from an initially low value, and repeated cycling tends to be more reproducible and non-destructive. Typical N-type characteristics are shown in Fig. 15.

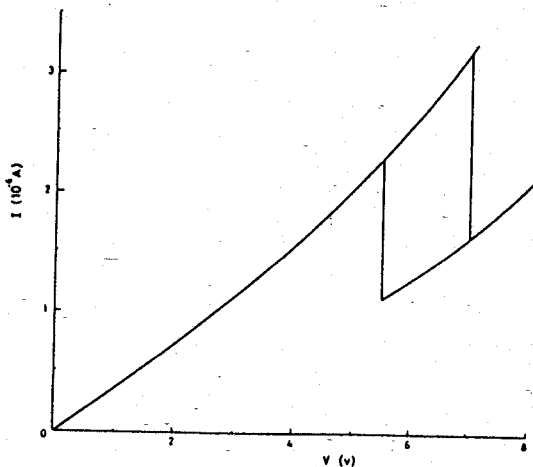


Fig. 15. The current-voltage characteristic of a discontinuous film exhibiting N-type switching. (From Borziak *et al.*¹⁷⁰)

In principle, any negative TCR material may display a thermal S-type negative resistance. Joule heating of the film decreases the resistance, increasing the current and hence Joule heating, until an equilibrium situation is reached¹³¹. Such characteristics must clearly be investigated using a constant-current source if thermal runaway is to be avoided. As far as discontinuous films are concerned, either a high field (or high field enhancement) or extreme current "bunching" within the film (or both effects) is required to obtain the necessary dissipation level. It is significant then that films where this type of negative resistance has been observed

were either highly agglomerated^{28, 168} or were deposited with a very short gap²¹. Most testing has been done with constant voltage, so the short film lifetimes and lack of reproducibility are not surprising. That the films survive at all is probably due to non-uniformity of structure and current limiting elsewhere in the film. Clearly structural changes (coalescence and filamentary breakages) give rise to second-order effects including multiple instabilities and the possibility of resistance increases from a low value, but we regard these as perturbations of the principal thermal mechanism. That there is a thermal origin has been well established by direct observation¹⁶⁸, and this type of switching has also been observed in cermets by Madjid and Anderson¹⁷¹.

The second type of switching, N-type, is of more practical interest since it appears to be more reproducible and non-destructive^{130, 169}.

Borziak *et al.*¹³⁰ have proposed a plausible explanation of this phenomenon. It has been noted that the potential distribution is non-uniform along some films, particularly those with a discontinuous region near the contact and a semi-continuous centre, as at least some of these films appear to have. It has been suggested that a potential minimum develops (as has been observed by Dittmer¹²⁰) associated with a space charge region near a contact. As the film current increases, the potential minimum decreases as more space charge builds up, leading to an accelerating decrease in conductance. This picture has yet to be fully developed. It has also been noted that increasing either the temperature or the frequency of an a.c. voltage test signal decreases any switching tendency but that the equilibrium state becomes a high resistance state¹³¹. These effects seem consistent with the space charge concept and must be kept in mind for any theoretical treatment.

3.14. Emission

Photon and electron emission characteristics from the surface of discontinuous films provide yet another field of study related to the conduction mechanism. The two effects definitely appear to be related^{130, 172, 173} and they suggest further applications of these films as efficient electron emitters or as sources of visible light^{120, 174}.

Electron emission will be considered first. Dittmer¹²⁰ interpreted his data in terms of Fowler-Nordheim field emission and the expected linearity between I_e/V^2 and V^{-1} (where I_e is the emission current). However, Kulyupin and Nepijko¹⁷⁵ have cast doubt on the result with a very neat experiment. They first pointed out that plotting I_e against $P^{-1/2}$ (where P is the power input to the film) gives equally good linearity and that the question of a field emission or hot electron process cannot be resolved by this technique. They then studied the variation of the Fowler-Nordheim constants (assuming this model) with film strain. The necessary relationship between these constants was not maintained as the substrate was extended, and it was concluded that emission is due to hot electrons rather than to field. In addition Dittmer¹²⁰ further considered I_e as a function of island size on the Fowler-Nordheim model without convincing agreement. Borziak *et al.*¹⁷⁶ used the electron temperature T_e , as given by $kT_e = \{(kT)^2 + \alpha P\}^{1/2}$ (where α is a constant), and confirmed the relationship $I_e = \text{const.} \exp\{-\phi/(\alpha P)^{1/2}\}$ where it is assumed¹⁷³ that $\alpha P \gg kT$. The work function ϕ was varied by the addition of barium oxide.

The electroluminescence intensity similarly seems to be characterized by a

relationship of the form $W \approx \exp\{-hv/(aP)^{1/2}\}$ although the spectrum tends to show more structure than the formula implies^{174, 176}. This structure definitely depends on the substrate, with an absorption edge corresponding to the band gap of the semiconducting substrates¹⁷⁴. For glass substrates the structure is understandably more complex and should yield information about the trap distribution within the tunnelling gap^{173, 176}. Electron temperatures typically range from 0.1 to 0.4 eV which seem reasonable for tunnelling electrons¹⁷⁰.

It is clear that visible photon emission only occurs at high fields and at localized areas of the film. In addition, it appears that aggregated film structures are often necessary to achieve sufficient field strength¹⁷³. At high powers some fluctuation is noticed and this is caused by structural changes. Further evidence that both effects have the same origin is provided by the onset of both electron and photon emission at similar field strengths and by the correlation of fluctuations¹⁷⁶.

We have already discussed the question of substrate tunnelling and the effect of traps to assist this process. It was shown that an electron may tunnel with thermal assistance to an energetically higher trap level in this process. Presumably the opposite may be true, and if an electron reaches a deep trap it may lose energy by radiating a photon. As electron energies increase, thermionic emission from the film will also increase.

Two other points must be mentioned. Firstly, the onset of both types of emission seems to coincide with the onset of non-ohmic effects¹⁷⁶. This result may indicate a new line to pursue with non-ohmic effects and requires wider verification. Secondly, experiments by Borziak *et al.*⁹ on cathodoluminescence have shown that the emission spectrum of this phenomenon depends upon the metal of the film—possibly upon interband transitions. We suggest that cathodoluminescence studies of discontinuous films with varying island sizes should yield information on the nature of the islands (*e.g.* energy level quantization, whether islands are liquid etc.) which may be vital to the conduction theory.

One fact seems certain—that the existence of substrate and island spectra in electroluminescence and cathodoluminescence respectively rules out any interpretation as a microplasma effect.

3.15. Silicon monoxide overcoat

A number of workers have overlaid discontinuous films with a deposited layer of SiO. The expectation is that the effect of two adjacent media of different dielectric constant will be removed and that the inter-island tunnelling area will be increased. Most publications merely state that the film resistance decreased^{144, 168, 177}, but as in other experiments it is more useful to describe the effects on δE and σ_0 independently. With regard to σ_0 , discrepancies emerge. The SiO caused a small increase in δE for gold films²⁷, a significant increase for chromium⁴⁶ and a small decrease for palladium³⁰. With gold the effect on σ_0 was indeterminate, but σ_0 increased considerably for chromium and decreased slightly with palladium.

Interpretation of these effects must obviously be highly speculative. With chromium all the observed effect occurs within the deposition of a few monolayers of SiO. Thermal annealing by radiation from the SiO source can be rejected for pre-annealed films. Explanations have been offered on the basis of an increased tunnelling area¹⁴⁴ (increases σ_0), an increased surface impurity density (increases

σ_0) and a decreased island surface tension¹⁷⁷. The last process may have led to island swelling, increased σ_0 and decreased δE . While this combination of effects has not been reported, theoretical analysis leads to the verified result that the change ΔR is proportional to $R^{3/2}$. Other explanations have been based upon charge exchange processes: one such explanation is based on the assumption, for example, that charge redistribution will lead to oblate islands becoming more spherical²⁷.

Clearly no general conclusion is yet possible in this area, and more detailed consistent data would be most welcome. The points of interest, however, are that σ_0 does not universally increase with increased tunnelling area and that δE does not universally decrease dramatically owing to an increase in the dielectric constant of the surrounding medium. Does this mean that the substrate dielectric constant dominates over that of air-vacuum (and if so why?) or is the high frequency value applicable and therefore only a small change expected? We view the data with caution, would like to see more and note that the results, as they stand, pose problems for the electrostatic model to explain.

3.16. Resistance variation during deposition

While undoubtedly many workers have monitored film resistance during deposition, few have published their results. If the conduction mechanism were well established, the resistance variation during deposition could yield valuable information on island nucleation, growth and coalescence processes. Alternatively, if the latter were known (with the additional knowledge of how an applied field affected them), then the results could assist the determination of the conduction process. Unfortunately both are still subjects for investigation. We note here the work of Andersson¹⁷⁸ and his technique of determining the onset of continuity from the peak in the curve of dR/dt versus t .

3.17. Post-deposition drift

In general the resistance of any film will vary after deposition. For low resistance films the resistance decreases with time, and for high resistance samples it increases^{2, 38, 44, 179-182}. Sometimes mixed effects have been observed, and it has generally been assumed that the two types of behaviour identify the films as continuous or discontinuous. For discontinuous films the rate of resistance variation with time (a) is greatest for the highest film resistance, (b) is less for refractory metals⁴⁴ and (c) may be decreased² by overcoating with SiO. Three types of explanation have been offered.

(a) The simplest of these is the temperature theory^{180, 181}. It is well known that the film is heated during deposition by radiation from the source. When the deposition ceases the film will cool freely to ambient. For a negative TCR film the falling temperature will lead to an increasing resistance. Certainly this thermal effect will be present, and any other proposed model must take it into account. The model is in complete qualitative agreement with observation for both discontinuous and cermet films, but it cannot be established whether or not this describes the facts fully. Excess film temperatures calculated from the model are greater for Au-SiO cermets and decrease more slowly than for discontinuous gold. These differences are expected because of the greater radiation from the hotter SiO source and a lower cooling efficiency owing to a lower surface-to-volume ratio for the cermet.

The second explanation is based on the shape changes that are undergone by fixed islands. Nishiura and Kinbara¹⁰² assumed the islands to have an oblate spheroidal shape and regarded this as variable with time after deposition. The proposition was that the inter-island gap, and hence the tunnelling resistance, would increase as the islands tended to a more spherical form. There are two problems with the model which otherwise seems to be very successful in interpreting the data. The first point is that the oblate spheroid is an equilibrium shape since it is a minimum energy configuration for a charged island. A driving mechanism for the change in eccentricity is required, and charge exchange between island and substrate is an attractive possibility. Estimates of the time taken for the process to occur vary from milliseconds to hours¹⁰³. This is not suggested in the proposal, however, which merely used the shape as a mathematical convenience for a changing radius of curvature at an island extremity. Any attempt to treat the model analytically leads to inconsistencies since surface self-diffusion will not occur with a true equilibrium shape such as the ellipsoid. The second problem lies with the concept of tunnelling between extremities of island curvature, which ignores the evidence of substrate influence. Andersson¹⁰⁴ has merely assumed that the gap width will rise exponentially towards some final value as island shapes change to minimize the total energy. Morris¹⁰² has suggested that, since atoms will be added to the island by substrate surface diffusion as well as by direct vapour impingement, the island will be growing around its line of contact with the substrate. When deposition ceases, this process will continue momentarily, but with all the atoms adherent the island will have a greater substrate contact area than equilibrium conditions demand. The system may be represented by an island (which is assumed for simplicity to be spherical) with an increasing contact angle (*e.g.* from $\pi/2$ (hemisphere) to π (sphere)).

This picture is the reverse of a spherical drop spreading to wet the substrate. The tunnelling distance between lines of contact will increase, as before, to produce a resistance variation. Regardless of how the mainly unknown parameters are chosen, the log resistance variation is essentially completed over two decades of time (as observed), and derived parameters are reasonable if inconclusive¹⁰⁸.

The final category of explanation is also the most complex. Phillips *et al.*¹⁰⁵ and Skofronick and Phillips¹⁰⁶ have studied and modelled agglomeration effects. Their predicted resistance changes are based on Hartman's conduction model and so some caution must be exercised for the reasons given in Section 2.4. One justification of the agglomeration and coalescence model is that overall δE and R_0 both increase²¹. (The observation of a $t^{1/2}$ relationship with its implication of a diffusion process is insufficiently established.) Both, however, would also increase according to the shape change models. (Consider, as an extreme example, two thin touching discs becoming two separate spheres.) A more effective argument for coalescence is based upon variations in the film behaviour with surface treatments. Oxygen seems to have a variable effect¹⁰⁷, and surface water appears to inhibit the change⁷⁷. Both of these can conceivably be explained by the contact angle model of Morris¹⁰², but this has not been considered.

One interesting observation is the variation in gauge factor. Instead of increasing monotonically with resistance, it reaches a maximum and then decreases. This is consistent with the often neglected influence of δE which increases with gap width¹⁰⁵.

3.18. Annealing

The resistance of a discontinuous film changes irreversibly when the film is heated beyond its deposition temperature or the highest temperature previously experienced¹⁹. This change is similar in many ways to post-deposition drift but it proceeds more rapidly once the critical temperature is exceeded. For any given film, it is expected that there will be an equilibrium stable structure corresponding to each temperature. Provided the temperature is held constant, the structure will tend asymptotically towards that final form (post-deposition drift). If the temperature is raised, the structure will change rapidly, the high temperature will bring about a new equilibrium state and the final structure will be stable when cooled to the original temperature.

This concept may be applied to whichever structural parameter is assumed to change (*e.g.* island eccentricity, contact angle, coalescence), and it is assumed that the same structural modifications are involved with both annealing and post-deposition drift.

With all models of structural change, R_0 is expected to increase upon annealing. Increases in δE are expected from island shape changes, while decreases are predicted by coalescence models. As with drift, a decrease in gauge factor has been noted by Neumann and Sutton².

Where R_0 increases, decreases in δE have been noted with palladium, gold, silver and platinum^{19,48,50} and increases with chromium, gold and tungsten^{21,39,46}. Hill¹¹¹ has specifically identified that the decrease in eccentricity of the oblate spheroidal island is due to annealing. (It should be noted that the quantitative detail is for a prolate spheroid island and not the oblate spheroid assumed in the discussion.) Tick and Fehlner⁵⁰ have noted that, when δE decreases, the deposition of a further small amount of material also produces a decrease in δE with a rise in R_0 . It is therefore expected that decreases in δE may be regarded as a coalescence phenomenon while a δE increase is associated with island shape changes. Coalescence or some other form of ordering (possibly leading to changes in preferred conduction channels) undoubtedly takes place in low resistance³⁹ semicontinuous or mixed TCR films⁴⁴. A secondary agglomeration process leads to a counter-effect of increasing film resistance by breaking continuous filaments¹⁷⁹.

A number of other results in this area are worthy of mention. The oxidation of nickel films has been observed¹¹², but, as with drift effects, rhenium yields anomalous data¹⁴⁴. It is well known that the application of a monitor field during film growth produces a low resistance film⁷⁹, but the structural difference between this film and another deposited simultaneously without a field vanished with annealing¹⁴⁵.

An applied voltage may also produce irreversible annealing, and once again there are two possible explanations. In semicontinuous films, current bunching and field enhancements may provide sufficient localized Joule heating to produce thermal annealing effects¹⁴⁴. In thinner films, polarization, stretching¹⁰¹ or other arguments related to field phenomena are employed.

A final comment must be made on a unique observation. Williams and Stone¹⁴⁷ found that resistance changes with platinum films occurred at 100°C and they interpreted these in terms of desorption of water from the films. In air R_0 increased and a physisorbed substrate surface layer was assumed to be lost, while in

vacuo R_0 decreased owing to the release of chemisorbed water which created more trap sites within the substrate. The irreversibilities of these desorption processes are a little surprising, and it is not entirely clear why the difference exists between behaviour in air and in vacuum. However, it is well established that surface water layers affect island mobility and coalescence kinetics, and the explanation probably lies here rather than in the trap argument above.

The difference between annealing results for discontinuous and cermet films is also significant. The thermal annealing of activated conduction cermets leads to a decrease in R_0 and an increase in δE . This effect has been explained in terms of the precipitation of gold and silicon atoms from the dielectric¹⁴. A similar effect is not expected with discontinuous films even with the existence of an "amorphous carpet" since surface mobilities are higher and inter-island atoms are more likely to join a nearby island.

Harkay and Crowell¹⁵ have introduced an interesting analytical tool which greatly assists the interpretation of annealing changes. They have used Hill's four structural types and have set these up as four quadrants of a plane. The chosen axes ($\ln\{\delta E(\text{eV})/0.025\}$, $\ln\{R_\infty/R_0(\infty)\}$) arbitrarily define the limits of large or small islands, wide or narrow gaps (where $R_0(\infty)$ is the high temperature intercept of the Arrhenius plot). The changes in electrical properties may be mapped in the plane in order to identify the structural changes readily.

4. CONCLUSIONS, SPECULATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

It will by now be apparent to the reader that the phenomenon of charge transfer in discontinuous metal films is far from well understood. In the past it has always been considered that the mechanism could be adequately accounted for qualitatively by the widely invoked model of activated tunnelling. However, in this review we have produced several pieces of evidence which cast doubt on even this, almost revered, theory. It seems likely, however, that activated tunnelling is the mechanism responsible for charge transfer, but to remove some of the anomalies which have been found when films are, for example, annealed or exposed to gases it is vital that we obtain many more quantifiable results relating to the structure and surface state of the films. By this we mean simply that, although a variety of models has been proposed to explain apparent anomalies, there is insufficient numerical evidence which can be used to test the theories.

To some extent, however, these anomalies can almost be regarded as secondary in importance to the problem of actually confirming numerically the validity of the charge transfer model. Most of the theories apply to arrays of two or three particles, and these are usually taken to be representative of the film as a whole. It is not too surprising therefore that very large discrepancies between the predicted and observed conductivities are usually obtained. There are many possible origins of such discrepancies; in this section we shall try to summarize those which we feel are particularly important and where possible we shall suggest experimentation which perhaps could shed some light on the problems.

The theories we have been discussing are concerned with explaining the electrical conductivity of films, and it is clear that this is very sensitive to the geometry of the array of islands. Several workers (e.g. Kazmerski and Racine⁶) have

carried out careful analyses of the size distributions of the islands, and this is a useful first step towards clarifying the effect of non-uniformity. In addition, though, we also need evidence of the distributions of island separations, and very few data of this sort exist. To test the more recently proposed percolation theories^{23, 22, 71, 73} it is vital that we have both sets of information. It should be a relatively straightforward matter to obtain the distributions first by producing transmission electron micrographs of the structures and second by analysing these in a Quantimet image analyser (Cambridge Instruments Ltd.). By using this instrument the radial distributions of separations from datum particles could be found, and this is precisely what is needed to test the recent theory of Hill and Coutts⁷². In addition, conductivity-temperature data are required over a far wider range of temperature than has previously been used if the various power laws are to be observed. (This problem is currently¹⁸ under investigation using an ultrahigh vacuum cryostat for deposition, electrical and optical measurements from 1 to 77 K.)

Inevitably, however, there will be room for ambiguity even after making such sophisticated structural measurements. Island shape is still in doubt, although a newly developed technique to extract topographical information from small islands offers hope in this region¹⁸. If the islands are actually prolate ellipsoids, then some information is required about contact angles. If the conduction path for tunnelling electrons is via the substrate (as seems to be very probable), then a simple estimate of the separations of the extremities of islands would be inadequate since this would suggest a shorter path length than would be expected. This raises the interesting possibility that, for contact angles that are greater than a certain value, the preferred path could still be through the vacuum between islands rather than through the substrate. Perhaps it will be possible to exploit the recent developments in scanning transmission electron microscopy to obtain some numerical information about contact angles. If adequate resolution can be obtained, it could also be interesting to examine the change of contact angle with temperature. This mechanism has been suggested as being responsible for post-deposition drift, and it is therefore relevant to any device application.

The change in contact angle on different substrates raises yet a further problem to structural analysis. The majority of films have been deposited on glass and then, for the purposes of microscopy, removed from this by overcoating with a carbon overlayer. This is generally floated off the substrate in water, and it is assumed that the film is removed with it. If indeed the entire film does leave the substrate, then presumably this must be because the adhesion is greater to the carbon than to the glass, in which case we would expect a change in island shape and in contact angle. Perhaps the way to overcome this would be to coat the substrate with carbon or silicon dioxide and to deposit the metal film directly onto this. No change of structure would then be expected when the carbon was removed from the substrate. Equally, to be of any relevance, the electrical measurements must also be made on the carbon support film.

The situation is further complicated by the need to transfer the film from the vacuum system to a water bath and then to an electron microscope. In view of the various changes in interfacial energies it would be surprising if the film structure remained stable, so that we are gradually directed towards what would be a very elegant solution to the problem of structural analysis and correlation with electrical

properties. There is no reason, in principle, why an entire deposition, measurement of conductivity and analysis of structure should not be carried out within an electron microscope. Similar experiments have been performed in studies of the nucleation of metals¹⁹⁰ and inert gases¹⁹¹, and the only additional complication would be that of providing electrical contacts. By using a combined cooling and heating stage it should be possible to make measurements of conductivity over a very wide range of temperature on stable (perhaps annealed) well-characterized structures and thereby to obtain very reliable data. This, together with subsequent information from a Quantimet analysis, would provide precisely what is required to test the percolation model of conduction of Hill and Coutts⁷³. Of course, great care would need to be taken to ensure that the electron beam did not influence the conductivity measurements, but this could perhaps be tested by varying the beam current and energy.

This experiment would also perhaps provide the opportunity to study field-induced coalescence both during and after growth. If, as has been suggested, charge transfer causes the distortion of islands, in the applied field, from prolate to oblate ellipsoids, then we have the interesting possibility of being able to determine preferential conduction paths and to observe the formation of new paths during heating or when increasing the applied field. It might also be interesting to attempt this using the voltage contrast mode of a scanning microscope, although great care would have to be taken to avoid substrate charging effects.

Throughout this review we have emphasized what we believe to be a comparatively neglected, but nevertheless important, aspect of this topic—that of the role of the contacts. The suggestion is that it is the charge injection process from the contacts into the body of the film which is activated and which limits the charge density. If this is indeed the case, then it should be interesting to perform localized heating in various areas of the film. This would be particularly informative if structures were available which were identical in the body of the film and near the electrodes, in which case it may be possible to separate out the activation energy corresponding to charge injection and to mobility (the latter arises from the additional energy barrier to tunnelling electrons due to the unrelaxed state of the lattice). Thus the problem reduces to that of ensuring a consistent structure along the length of the film, and it seems that this could be achieved by depositing the electrodes through a mask after deposition of the film. If the presence of the electrodes influences the growth process because of selective heating, then this could assist matters. Alternatively, non-uniform filament heating effects* may perhaps be eliminated by using sputter deposition (with its attendant problems of contamination and heating, albeit fairly uniformly) or ion beam deposition. If we wished to observe growth within an electron microscope, this modification would not be possible to implement and we would need to look for an alternative such as charging the vapour beam and deflecting this to the substrate (which would of course be out of direct line of the filament). The difficulty of implementing this would be very great and such an experiment is unlikely to be carried out.

* If the filament operates at 1500 K, then its black-body radiation curve will peak at a wavelength of about 2 μm . Since this is roughly on the absorption edge of glass we could expect that the centre of the substrate would get hotter than the areas near the electrodes because of the high reflectivity of the latter (at least if they are formed from gold or any other typical metal) in the near IR.

We also need quantitative information regarding the state of the substrate between islands since this is likely to have a radical influence on charge transfer. Tick and Fehlner⁵⁰ have speculated about the existence of an amorphous carpet between islands, and it is important to confirm or refute this. It is of interest to note the work of Pollard *et al.*¹⁶⁴ on Au-SiO₂ cermet thin films. By carrying out analyses of average sizes and separations of the metal particles from micrographs, they calculated the metallic volume fraction. This was compared with the value obtained using X-ray fluorescence and was found to be considerably less. The conclusion was that only a proportion of the gold existed in the form of islands greater than 15 Å in diameter (the minimum resolvable island size); the remainder comprised an amorphous material dispersed throughout the insulator. If the amorphous component exhibits the typical electrical characteristics of such materials, then its properties could dominate those of the island component. It is possible that an amorphous surface layer could be formed in two-dimensional discontinuous films by diffusion into the substrate, and again it is important to look for this by using one of the many surface analytical techniques developed over the last decade.

Holland¹⁹² has pointed out that, if glass is heated within a vacuum system, it continues to desorb water vapour for *many hundreds of hours*. If this is the case, it seems inevitable that there will be a thin surface layer of water vapour on the substrate, and we are left to speculate about its effect on the conduction process. Would it, for example, present a d.c. hopping path between islands? How would it behave if polarized? What would its effect be on the potential barrier? And so on. Quite clearly we have virtually no information on this topic, and it would therefore be invaluable to undertake surface analysis. In addition it would also be interesting to investigate the behaviour of the films by using thermal stimulation which could give information regarding trapping sites¹⁹³. On the same theme, has photo-conductivity ever been observed in discontinuous structures? If not, this could be useful, and an effect which might almost be expected.

The whole question of static charge transfer between islands and substrate still needs clarification, although Hill¹⁶ and Licznerski⁴⁶ have attempted to deal with the subject.

(i) Hill's suggestion that the transfer will lead to quantized shifts in the island Fermi level is rather dubious (since presumably Fermi level equalization is the driving mechanism) and was never adequately incorporated in his model.

(ii) If electron injection into the substrate creates interfacial Schottky barriers, then these must also be accommodated theoretically. In the text, however, considerable reliance has been placed upon Hill's assertion that platinum islands form depletion layers at the interface. Similar experimental results for platinum films and for films of other metals have consequently led to the conclusion that none of them form Schottky barriers.

(iii) The speed of the transfer may also be important if this is the mechanism causing post-deposition drift.

(iv) A further implication of the charge transfer process is that the island itself is statically charged before any conduction takes place. The ellipsoidal particle shape is well established, and the excess charge is expected to reside at points of maximum curvature. What effect will this have upon the localized Fermi level and the tunnelling probabilities?

The concept of two distinct types of tunnelling transfer (activated charge separation and non-activated charge transfer) along a single percolation path has some interesting implications. Let us consider a line of identical island-gap combinations. The activated process requires a higher gap voltage than the simple charge transfer. In the ideal structure the potential distribution along the line must be fluctuating, and the only reasonable model would be a dynamic one. (This idea has special implications for the origins of noise in the films.)

The switching effects discussed earlier are also worthy of further investigation, and it would be useful to determine whether they have a thermal or a field-induced origin. This could perhaps be achieved by using a pulsed voltage whose amplitude and duration can be varied.

At the fundamental level it would be interesting to evaluate m^* and $\bar{\phi}$ by using the curvature of conductivity-temperature plots²⁴ and, on the same film, by using the technique suggested by Gundlach^{19a} involving crossed electric and magnetic fields. Estimates of these quantities are certainly required before an attempt is made to compare the theoretical and the actual conductivities.

The above comments all represent suggestions for further practical work, but there are also a number of areas which require theoretical effort.

The work of Buffat¹⁰⁶ suggests that the melting point of metal islands decreases with their size, and very small particles may thus be not far removed from the liquid state. It would therefore be useful to determine the energy structure of a metal in such a finely divided state. For the reasons given earlier, it is unlikely that significant quantization of individual energy levels would occur, and certainly the quantum mismatch between the Fermi levels of adjacent islands cannot be responsible for the activation energy; however, if the islands are in a nearly liquid state, presumably the normal concepts of band structure break down. In this case quite radical modifications to the theory may be needed. For example, the dynamics of the shape of a charged liquid drop within a fluctuating potential field will certainly not simplify the analysis.

If we assume that we can neglect any liquid-like properties of the islands and if the conventional activated tunnelling theory can, at least broadly, be accepted, then it is essential that an attempt be made to develop the theory for the case of coplanar electrodes. Ideally this should account for variations in shape, size, contact angle, effective tunnelling area and inter-island separation. Realistically, such a theory is unlikely to be developed and it is probable that we shall have to be content with the "simple" coplanar theory.

The presence of Hall and magnetoresistive effects is surprising, and it would almost certainly be informative to look for the origins of these. The Hall effect can usually be used to infer a mobility so that, if some indication of its temperature dependence could be observed, then evidence related to the activated nature of the mobility could be anticipated. Equally, magnetoresistive effects are commonly used to infer details of band structure so that further efforts in this direction, and indeed all magnetoelectric and thermoelectric effects, are likely to be fruitful.

In conclusion therefore we can reasonably say that, even though it is commonly believed that the general features of charge transfer in discontinuous structures are understood, there are many important aspects which are not understood at all and which could, at worst, demand the development of a completely new theory. At

best, they could be expected to complement the existing theory and to bring about the quantitative agreement that has been sought after for so long. In any event the topic is likely to provide a continuing source of interest for some considerable time to come.

REFERENCES

- 1 W. F. G. Swann, *Philos. Mag.*, **28** (1914) 467.
- 2 M. R. Neumann and W. G. Sutton, *J. Vac. Sci. Technol.*, **6** (1969) 710.
- 3 Z. H. Meiksin, E. J. Stolinski, H. B. Kuo and K. J. Shah, *3rd Int. Conf. on Medical Physics, Chalmers University of Technology, Göteborg, Sweden, 1972*.
- 4 G. R. Witt, *Ph.D. Thesis*, Imperial College, London University, 1976.
- 5 E. J. Hsieh and R. Y. Shimoda, *Solid-State Electron.*, **12** (1969) 493.
- 6 L. L. Kazmerski and D. M. Racine, *J. Appl. Phys.*, **46** (1975) 791.
- 7 J. C. Mitchinson and R. D. Pringle, *Electron. Lett.*, **7** (1971) 38.
- 8 P. Borziak, I. Konovalov, Yu. Kulyupin and K. Pilipchak, *Thin Solid Films*, **35** (1976) 19.
- 9 C. R. Wronski, B. Abeles and A. Rose, *Phys. Rev. Lett.*, **27** (1975) 19.
- 10 E. Ahilea and A. A. Hirsch, *Phys. Lett A*, **28** (1969) 273.
- 11 V. I. Gol'danski, B. M. Levin, V. P. Shantarovich, D. A. Agienskii and M. V. Landau, *JETP Lett.*, **19** (1974) 241.
- 12 C. A. Neugebauer and R. H. Wilson, in R. Niedermayer and H. Mayer (eds.), *Basic Problems in Thin Film Physics*, Vandenhoeck and Ruprecht, Göttingen, 1966.
- 13 T. J. Coutts, *Electrical Conduction in Thin Films*, Elsevier, Amsterdam, 1974.
- 14 Z. H. Meiksin, *Phys. Thin Films*, **8** (1975).
- 15 L. I. Maissel, in L. I. Maissel and R. Glang (eds.), *Handbook of Thin Film Technology*, McGraw-Hill, New York, 1970.
- 16 R. M. Hill, *Proc. R. Soc. London, Ser. A*, **309** (1969) 375.
- 17 C. J. Gorter, *Physica*, **17** (1951) 777.
- 18 E. Darmois, *J. Phys. Radium*, **17** (1956) 210.
- 19 C. A. Neugebauer and M. B. Webb, *J. Appl. Phys.*, **33** (1962) 74.
- 20 B. Abeles, *RCA Rev.*, **36** (1975) 594.
- 21 J. E. Dryer, *Ph.D. Thesis*, Ohio State University, 1972.
- 22 J. G. Swanson, D. S. Campbell and J. C. Anderson, *Thin Solid Films*, **1** (1967-1968) 325.
- 23 K. D. Leaver, *J. Phys. C*, **10** (1977) 249.
- 24 R. M. Hill, *Proc. R. Soc. London, Ser. A*, **309** (1969) 397.
- 25 R. Kiernan and D. W. Stops, *Nature (London)*, **224** (1970) 907.
- 26 A. Barr and R. D. Finney, *Thin Solid Films*, **24** (1974) 511.
- 27 J. E. Morris, *J. Vac. Sci. Technol.*, **9** (1972) 437.
- 28 J. E. Morris, *Ph.D. Thesis*, University of Saskatchewan, 1971.
- 29 J. Dryer, D. Goer and R. Speiser, *J. Vac. Sci. Technol.*, **10** (1973) 316.
- 30 J. E. Morris, *Thin Solid Films*, **11** (1972) 259.
- 31 J. E. Morris, *Thin Solid Films*, **11** (1972) 81.
- 32 B. Abeles, P. Sheng, M. D. Coutts and Y. Aric, *Adv. Phys.*, **24** (1975) 407.
- 33 P. K. Dubey, *J. Phys. D*, **3** (1970) 145.
- 34 P. Schnupp, *Thin Solid Films*, **2** (1968) 177.
- 35 F. W. Schmidlin, *J. Appl. Phys.*, **37** (1966) 2823.
- 36 J. W. Gadzuk, *J. Appl. Phys.*, **41** (1970) 286.
- 37 R. M. Hill, *Nature (London)*, **204** (1964) 35.
- 38 C. A. Neugebauer, *Trans. 9th Natl. Vacuum Symp.*, American Vacuum Society, New York, 1962, p. 45.
- 39 P. J. Dobson and B. J. Hopkins, *J. Appl. Phys.*, **39** (1968) 3074.
- 40 F. G. Ball, E. P. Feldman and V. I. Sashara, *Sov. Phys.—Solid State*, **16** (1974) 570.
- 41 N. Nifontoff, *C. R. Acad. Sci.*, **236** (1953) 1538.
- 42 N. Nifontoff, *C. R. Acad. Sci.*, **237** (1953) 24.
- 43 C. Feldman, *J. Appl. Phys.*, **34** (1963) 1710.

- 44 N. Mostovetch and B. Vodar, in H. K. Henisch (ed.), *Semiconducting Materials*, Butterworths, London, 1951.
- 45 T. E. Hartman, *J. Appl. Phys.*, **34** (1963) 943.
- 46 B. W. Licznanski, *Thin Solid Films*, **36** (1976) 61.
- 47 J. Kadlec, *Solid-State Electron.*, **17** (1974) 496.
- 48 D. S. Herman and T. N. Rhodin, *J. Appl. Phys.*, **37** (1966) 1594.
- 49 L. Y. Wei, *J. Chem. Phys.*, **39** (1963) 2709.
- 50 P. A. Tick and R. P. Fehner, *J. Appl. Phys.*, **43** (1972) 362.
- 51 L. S. Palatnik, B. T. Boiko, A. N. Synelnikov and V. R. Kopach, *Sov. Phys.—Dokl.*, **14** (1970) 1211.
- 52 B. T. Boiko, L. S. Palatnik and A. N. Synelnikov, *Thin Solid Films*, **7** (1971) 305.
- 53 I. C. E. Pike and C. H. Seager, *Phys. Rev., Sect. B*, **10** (1974) 1421.
- 54 T. J. Coutts, *Thin Solid Films*, **38** (1976) 313.
- 55 H. L. Frisch, V. A. Vyssotskiy, S. B. Gordon and J. M. Hammersley, *Phys. Rev.*, **123** (1961) 1566.
- 56 M. F. Sykes and J. W. Essam, *Phys. Rev., Sect. A*, **133** (1964) 310.
- 57 J. W. Essam, C. M. Place and E. H. Sondheimer, *J. Phys. C*, **7** (1974) 169.
- 58 B. Abeles, H. L. Pinch and J. I. Gittleman, *Phys. Rev. Lett.*, **35** (1975) 247.
- 59 N. T. Liang, Yueh Shan and Shou-yin Wang, *Phys. Rev. Lett.*, **37** (1976) 526.
- 60 T. Andersson, *Ph.D. Thesis*, Chalmers University, Göttenburg, 1976.
- 61 H. R. Zeller and I. Giaever, *Phys. Rev.*, **181** (1969) 789.
- 62 H. R. Zeller, *Phys. Rev. Lett.*, **28** (1972) 1452.
- 63 N. F. Mott and E. A. Davis, *Philos. Mag.*, **17** (1968) 1269.
- 64 M. L. Knotek, M. Pollak, T. M. Donovan and H. Kurtzman, *Phys. Rev. Lett.*, **30** (1973) 853.
- 65 J. Bernasconi, *Phys. Rev., Sect. B*, **7** (1973) 2252.
- 66 C. G. Granqvist and R. A. Bührman, *Appl. Phys. Lett.*, **27** (1975) 693.
- 67 T. Andersson and C. G. Granqvist, *J. Appl. Phys.*, **48** (1977) 1673.
- 68 R. M. Hill and T. J. Coutts, *Thin Solid Films*, **35** (1976) L17.
- 69 R. M. Hill, *Phys. Status Solidi A*, **35** (1976) K29.
- 70 P. Sheng and B. Abeles, *Thin Solid Films*, **41** (1977) L39.
- 71 J. Heinrichs, A. A. Kumar and N. Kumar, *J. Phys. C*, **7** (1974) 3560.
- 72 V. Ambegaokar, B. I. Halperin and J. S. Langer, *Phys. Rev., Sect. B*, **4** (1971) 2612.
- 73 R. M. Hill and T. J. Coutts, *Thin Solid Films*, **42** (1977) 201.
- 74 T. Andersson, personal communication, 1976.
- 75 R. M. Hill, personal communication, 1975.
- 76 H. Fredriksson, B. Persson and L. Ystrom, *Phys. Scr.*, **3** (1971) 169.
- 77 B. Hök, R. Hyholm, O. Groth and P. A. Tove, *Thin Solid Films*, **17** (1973) 113.
- 78 R. M. Hill, *Nature (London)*, **210** (1966) 512.
- 79 J. E. Morris, *Metallography*, **5** (1972) 41.
- 80 Z. H. Meiksin and R. A. Mirchandi, *Thin Solid Films*, **22** (1974) 549.
- 81 M. J. Knight and K. N. Jha, *Thin Solid Films*, **2** (1968) 131.
- 82 G. Desrousseaux, *Thin Solid Films*, **32** (1976) 255.
- 83 A. N. Chester and B. B. Kosicki, *Rev. Sci. Instrum.*, **41** (1970) 1817.
- 84 J. C. Mitchinson and R. D. Pringle, *J. Phys. E*, **5** (1972) 742.
- 85 J. E. Morris, *Thin Solid Films*, **5** (1970) 339.
- 86 H. Schmeisser, *Thin Solid Films*, **22** (1974) 99.
- 87 M. H. Jacobs and D. W. Pashley, in S. S. Breese (ed.), *5th Int. Congr. for Electron Microscopy*, Vol. 1, Academic Press, New York, 1962.
- 88 J. C. Maxwell Garnett, *Philos. Trans. R. Soc. London, Ser. A*, **203** (1904) 385.
- 89 J. C. Maxwell Garnett, *Philos. Trans. R. Soc. London, Ser. A*, **205** (1905) 237.
- 90 R. H. Doremus, *J. Appl. Phys.*, **37** (1966) 2775.
- 91 E. David, *Z. Phys.*, **114** (1939) 389.
- 92 H. Schopper, *Z. Phys.*, **130** (1951) 565.
- 93 A. Donnadiu, *Thin Solid Films*, **6** (1970) 249.
- 94 D. N. Jarrett and L. Ward, *J. Phys. D*, **9** (1976) 1515.
- 95 C. Reale, *Infrared Phys.*, **10** (1970) 175.
- 96 M. B. Heritage and M. F. Tompsett, *J. Appl. Phys.*, **41** (1970) 407.

- 97 R. H. Wade and J. Silcox, *Phys. Status Solidi*, **19** (1967) 57.
- 98 J. J. Metois, M. Gauch, A. Masson and R. Kern, *Thin Solid Films*, **11** (1972) 205.
- 99 H. K. Chaurasia and W. A. G. Voss, *Nature (London)*, **249** (1974) 28.
- 100 D. Dobrev and L. Boicheva, *Thin Solid Films*, **32** (1976) 237.
- 101 E. Ahilea and A. A. Hirsch, *J. Appl. Phys.*, **42** (1971) 5601.
- 102 J. E. Morris, *Thin Solid Films*, **29** (1975) L9.
- 103 M. J. Stowell, T. J. Law and J. Smart, *Proc. R. Soc. London, Ser. A*, **318** (1970) 231.
- 104 J. R. Sambles, *Proc. R. Soc. London, Ser. A*, **324** (1971) 339.
- 105 S. J. Peppiatt and J. R. Sambles, *Proc. R. Soc. London, Ser. A*, **345** (1975) 387.
- 106 P. A. Buffat, *Thin Solid Films*, **32** (1976) 283.
- 107 R. W. Alexander and R. J. Bell, *J. Non-Cryst. Solids*, **19** (1975) 93.
- 108 J. E. Morris, *Thin Solid Films*, **36** (1976) 29.
- 109 K. Van Stoensel, *Philips Res. Rep.*, **22** (1967) 246.
- 110 T. Dingle, personal communication, 1975.
- 111 R. M. Hill, *J. Appl. Phys.*, **37** (1966) 4590.
- 112 S. Monchaud, S. Rolland and R. Granger, *Electrocomp. Sci. Technol.*, **1** (1974) 97.
- 113 R. Kiernan and D. W. Stops, *J. Vac. Sci. Technol.*, **9** (1972) 433.
- 114 J. E. Morris, *Thin Solid Films*, **11** (1972) 299.
- 115 B. T. Boiko, L. S. Palatnik, A. N. Synelnikov, Y. L. Friman and V. B. Arakelyan, *Sov. Phys.—Dokl.*, **17** (1972) 395.
- 116 J. E. Morris, *J. Appl. Phys.*, **39** (1968) 6107.
- 117 L. A. Weitzenkamp and N. M. Bashara, *Trans. Metall. Soc. AIME*, **236** (1966) 351.
- 118 A. A. Hirsch and S. Bazian, *Physica*, **30** (1964) 258.
- 119 A. Van Itterbeek, A. Dupre and P. Cept, *Proc. 10th Int. Congr. on Refrigeration, Copenhagen, 1959*, in *Prog. Refrig. Sci. Technol.*, **1** (July 1960).
- 120 G. Dittmer, *Thin Solid Films*, **9** (1972) 317.
- 121 J. E. Morris and M. D. Newberry, to be published.
- 122 A. A. Hirsch and N. Friedman, *Physica*, **30** (1964) 389.
- 123 A. A. Milgram and C. S. Lu, *J. Appl. Phys.*, **37** (1966) 4773.
- 124 A. V. Joglekar, R. N. Karekar and K. Sathianandan, *Phys. Lett. A*, **46** (1973) 69.
- 125 F. J. Kollaris and R. Speiser, *Thin Solid Films*, **35** (1976) 131.
- 126 D. Goer, J. Dryer and R. Speiser, *J. Vac. Sci. Technol.*, **11** (1974) 575.
- 127 R. M. Hill, *Thin Solid Films*, **12** (1972) 367.
- 128 E. Ahilea and A. A. Hirsch, *J. Vac. Sci. Technol.*, **6** (1969) 698.
- 129 C. J. Roussow, I. A. Kotze and C. A. O. Henning, *Thin Solid Films*, **29** (1975) 85.
- 130 P. Borziak, V. Diukov, A. Kostenko, Yu. Kulyupin and S. Nepijko, *Thin Solid Films*, **36** (1976) 21.
- 131 J. E. Morris, *Thin Solid Films*, **35** (1976) 165.
- 132 J. E. Morris, *Thin Solid Films*, **28** (1975) L21.
- 133 M. J. Knight, *J. Vac. Sci. Technol.*, **6** (1969) 706.
- 134 R. L. Parker and A. Krinsky, *J. Appl. Phys.*, **34** (1963) 2700.
- 135 Z. H. Meiksin and R. A. Hudzinski, *J. Appl. Phys.*, **38** (1967) 4490.
- 136 M. Nishiura, S. Yashida and A. Kinbara, *Thin Solid Films*, **15** (1973) 133.
- 137 G. R. Witt, *Thin Solid Films*, **22** (1974) 133.
- 138 B. T. Boiko, L. S. Palatnik, V. P. Kopach and S. B. Melent'ev, *Sov. Phys.—Solid State*, **14** (1972) 1090.
- 139 G. R. Witt and T. J. Coutts, *Thin Solid Films*, **7** (1971) R1.
- 140 B. S. Verma and H. J. Juretschke, *J. Appl. Phys.*, **41** (1970) 4732.
- 141 G. R. Witt, *Thin Solid Films*, **13** (1972) 109.
- 142 F. P. Fehner, *Nature (London)*, **210** (1966) 1035.
- 143 S. M. Deshpande, *Thin Solid Films*, **39** (1976) 125.
- 144 H. Dannielsson, B. Kasemo and I. Markland, *Thin Solid Films*, **9** (1971) 121.
- 145 J. L. Williams and I. L. Stone, *J. Phys. C*, **5** (1972) 2105.
- 146 J. E. Morris, *J. Vac. Sci. Technol.*, **9** (1972) 1039.
- 147 J. L. Williams and I. L. Stone, *Thin Solid Films*, **11** (1972) 329.
- 148 A. Barr, *Ph.D. Thesis*, London University, 1973.
- 149 D. Lazarov and St. Manev, *Thin Solid Films*, **10** (1972) 217.

- 150 F. P. Fehlner and S. M. Irving, *J. Appl. Phys.*, **37** (1966) 3313.
 151 J. L. Williams and R. K. Burdett, *J. Phys. C*, **2** (1969) 298.
 152 P. Mazzetti and A. Stepanescu, *Thin Solid Films*, **13** (1972) 67.
 153 A. Bobbio, A. Masoero, P. Mazzetti and A. Stepanescu, *Thin Solid Films*, **36** (1976) 33.
 154 S. Offret and M. B. Vodar, *J. Phys. Radium*, **17** (1955) 237.
 155 J. R. Harkay and A. D. Crowell, *J. Appl. Phys.*, **47** (1976) 4504.
 156 A. V. Joglekar, *J. Phys. D*, **7** (1974) 1270.
 157 K. Fuchs, *Proc. Cambridge Philos. Soc.*, **34** (1938) 100.
 158 A. F. Mayadas and M. Shatzkes, *Phys. Rev., Sect. B*, **1** (1970) 1382.
 159 J. R. Harkay, *Ph.D. Thesis*, University of Vermont, 1976.
 160 S. M. Deshpande and A. D. Crowell, *J. Vac. Sci. Technol.*, **9** (1972) 97.
 161 M. A. Jeppesen, *J. Appl. Phys.*, **37** (1966) 1940.
 162 G. Comsa, *Thin Solid Films*, **4** (1969) 1.
 163 N. C. Miller, B. Hardiman and G. A. Shirn, *J. Appl. Phys.*, **41** (1970) 1850.
 164 J. K. Pollard, R. L. Bell and G. G. Bloodworth, *J. Vac. Sci. Technol.*, **16** (1969) 702.
 165 B. T. Boiko, L. S. Palatnik, V. R. Kopach, A. N. Synelnikov and V. B. Arakelyan, *Sov. Phys.—Solid State*, **13** (1972) 2352.
 166 J. S. Helman and B. Abeles, *Phys. Rev. Lett.*, **37** (1976) 1429.
 167 J. I. Gittleman, Y. Goldstein and S. Bozowski, *Phys. Rev., Sect. B*, **5** (1972) 3609.
 168 S. K. Dey and G. D. Dick, *J. Vac. Sci. Technol.*, **11** (1974) 97.
 169 U.S. Patent 3,474,305 (1969), to R. E. Szupillo; U.K. Patent 1-211-669 (1970), to R. E. Szupillo
 170 P. Borziak, Yu. Kulyupin and S. Nepijko, *Thin Solid Films*, **36** (1976) 235.
 171 A. H. Madjid and W. F. Anderson, *J. Appl. Phys.*, **2** (1973) 133.
 172 P. G. Borziak, O. G. Sarbei and R. D. Fedorovich, *Phys. Status Solidi*, **8** (1965) 55.
 173 Yu. A. Kulyupin and K. N. Pilipchak, *Phys. Status Solidi A*, **11** (1972) K15.
 174 M. Bischoff and H. Pagnia, *Thin Solid Films*, **29** (1975) 303.
 175 Yu. A. Kulyupin and S. A. Nepijko, *Sov. Phys.—Solid State*, **17** (1976) 1822.
 176 P. Borziak, Yu. Kulyupin and P. Tomchuk, *Thin Solid Films*, **30** (1975) 47.
 177 A. A. Milgram and C. S. Lu, *J. Appl. Phys.*, **39** (1968) 1624.
 178 T. Andersson, *Thin Solid Films*, **29** (1975) L21.
 179 T. Andersson, *J. Phys. D*, **9** (1976) 76.
 180 J. Le Bas, *C. R. Acad. Sci., Ser. B*, **273** (1971) 1044.
 181 J. E. Morris, *Vacuum*, **22** (1972) 153.
 182 M. Nishiura and A. Kinbara, *Thin Solid Films*, **24** (1974) 79.
 183 R. Elsdon and F. R. G. Mitchell, *J. Phys. D*, **9** (1976) 1445.
 184 T. Andersson, *J. Appl. Phys.*, **47** (1976) 1752.
 185 W. B. Phillips, E. A. Desloge and J. G. Skofronick, *J. Appl. Phys.*, **39** (1968) 3210.
 186 J. G. Skofronick and W. B. Phillips, *J. Appl. Phys.*, **38** (1967) 4791.
 187 G. G. Paulson and A. L. Friedberg, *Thin Solid Films*, **5** (1970) 47.
 188 S. Norman and T. Andersson, *Vacuum* (1977) to be published.
 189 M. J. Yacaman and T. Ocana, *Appl. Phys. Lett.*, **30** (1977) 359.
 190 H. Poppa, *Thin Solid Films*, **34** (1976) 94.
 191 J. Venables, *Thin Solid Films*, **32** (1976) 135.
 192 L. Holland, *Properties of Glass Surfaces*, Chapman and Hall, London, 1962.
 193 H. M. Gupta and R. J. van Overstraeten, *J. Phys. C*, **7** (1974) 3560.
 194 K. H. Gundlach, *Thin Solid Films*, **3** (1969) 251.

APPENDIX A

Derivation of the activated tunnelling equation

(a) The number of electrons leaving the metal island surface in the x direction (N_L per unit area per unit time) is given by

$$N_L = \int_0^{E_m} v_x n(v_x) D(E_x) dv_x$$

ELECTRICAL CONDUCTION IN DISCONTINUOUS METAL FILMS

$$= m^{-1} \int_0^{E_m} n(v_x) D(E_x) dE_x$$

where $n(v_x) dx$ is the volume density of electrons with x-directed velocities in the range v_x to $v_x + dv_x$ (up to a maximum v_m) and $D(E_x)$ is the emission probability of an electron with x-directed energy $E_x = \frac{1}{2}mv_x^2$. We now require $n(v_x)$ and $D(E_x)$.

(b) For an isotropic velocity distribution (i.e. a zero field within the island) the density of available energy states $S(E) dE$ between $E = p^2/2m$ and $E + dE$ is

$$S(E) dE = \frac{\pi}{2} \left(\frac{8m}{h^2} \right)^{3/2} E^{1/2} dE$$

i.e.

$$S(p) dp = \frac{8\pi}{h^3} p^2 dp$$

in momentum space. The fraction of states in $dp_x dp_y dp_z$ is $dp_x dp_y dp_z / 4\pi p^2 dp$ and hence

$$S(p) dp_x dp_y dp_z = \frac{2}{h^3} dp_x dp_y dp_z$$

$$S(v) dv_x dv_y dv_z = \frac{2m^3}{h^3} dv_x dv_y dv_z$$

Since $n(v) = S(v)f(E)$, then

$$n(v) dv_x dv_y dv_z = 2 \left(\frac{m}{h} \right)^3 f(E) dv_x dv_y dv_z$$

and

$$n(v_x) = 2 \left(\frac{m}{h} \right)^3 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(E) dv_y dv_z$$

Putting

$$v_1^2 = v_x^2 + v_z^2 \quad v_y = v_1 \cos \theta \quad v_z = v_1 \sin \theta \quad E_1 = \frac{1}{2}mv_1^2$$

we can show that

$$n(v_x) = \frac{4\pi m^2}{h^3} \int_0^{\infty} f(E) dE_1$$

Hence

$$N_L = \frac{4\pi m}{h^3} \int_0^{E_m} D(E_x) \left\{ \int_0^{\infty} f(E) dE_1 \right\} dE_x$$

(c) Tunnelling requires both an electron with appropriate energy in the supply island and a vacant energy level in the recipient. The tunnelling current from island A to island B is therefore

$$J_{A \rightarrow B} = \frac{4\pi me}{h^3} \int_0^{E_2} D(E_x) \left[\int_0^\infty f_A(E) \{1 - f_B(E)\} dE \right] dE_x$$

(d) In the present case the "emission" probability $D(E_x)$ is clearly the tunnelling transmission probability given by the standard WKB approximation

$$D(E_x) = \exp\left(-\frac{4\pi}{h} \int_{s_1}^{s_2} [2m\{V(x) - E_x\}]^{1/2} dx\right)$$

for tunnelling through a barrier $V(x)$ of width $s_2 - s_1$. Simmons¹ has further reduced this to the form

$$D(E_x) = \exp\left[\frac{-4\pi(2m)^{1/2} \Delta s \{V(x) - E_x\}^{1/2}}{h}\right]$$

provided that

$$\int_{s_1}^{s_2} \{f(x) - \bar{f}(x)\}^2 dx \ll 8 \bar{f}(x)^2 \Delta s$$

(i.e. provided that fluctuations in the barrier shape are negligible) where $\bar{f}(x)$ is the average value of $f(x) = V(x) - E_x$ and $\Delta s = s_2 - s_1$. Hence

$$D(E_x) = \exp(-A\bar{\phi}^{1/2}) \exp\{-B(E_F - E_x)\}$$

where

$$A = \frac{4\pi\Delta s(2m)^{1/2}}{h} \quad B = \frac{A}{2\bar{\phi}^{1/2}} \quad \bar{\phi} = \Delta s^{-1} \int_{s_1}^{s_2} \phi(x) dx$$

and

$$\phi(x) = V(x) - E_F$$

The current density from the island surface is therefore

$$J_{A \rightarrow B} = \frac{4\pi me}{h^3} \exp(-A\bar{\phi}^{1/2}) \int_{-\infty}^{\infty} \exp\{-B(E_F - E_x)\} \times \left[\int_0^\infty f_A(E) \{1 - f_B(E)\} dE \right] dE_x$$

where the integral over E_x has been taken to ∞ as a maximum and has been extended below zero to $-\infty$ as a mathematical convenience with negligible error.

(e) Figure A1 enables the Fermi functions to be evaluated. For the determination of $J_{A \rightarrow B}$,

$$f_A(E) = \left\{ 1 + \exp\left(\frac{E - E_F + \delta E}{kT}\right) \right\}^{-1}$$

$$f_B(E) = \left\{ 1 + \exp\left(\frac{E - E_F + eV}{kT}\right) \right\}^{-1}$$

while for $J_{B \rightarrow A}$

$$f_A(E) = \left\{ 1 + \exp\left(\frac{E - E_F}{kT}\right) \right\}^{-1}$$

$$f_B(E) = \left\{ 1 + \exp\left(\frac{E - E_F + \delta E + eV}{kT}\right) \right\}^{-1}$$

Evaluation of the integrals requires the standard results

$$\int_0^\infty \{1 + \exp(at + b)\} dt = a^{-1} \ln\{1 + \exp(-b)\}$$

$$\int_{-\infty}^\infty \exp(at) \ln\{1 + \exp(b - t)\} dt = a^{-1} \exp(ab) \frac{\pi}{\sin(a\pi)}$$

and gives eqn. (3) for $J = J_{A \rightarrow B} - J_{B \rightarrow A}$.

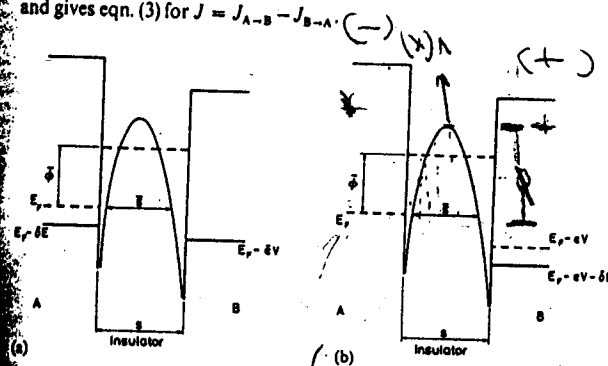


Fig. A1. The energy diagram for electron tunnelling between metal islands A and B at a potential difference V : (a) electron transfer from A to B; (b) electron transfer from B to A.

Reference for Appendix A

1. J. G. Simmons, *J. Appl. Phys.*, 34 (1963) 1793.